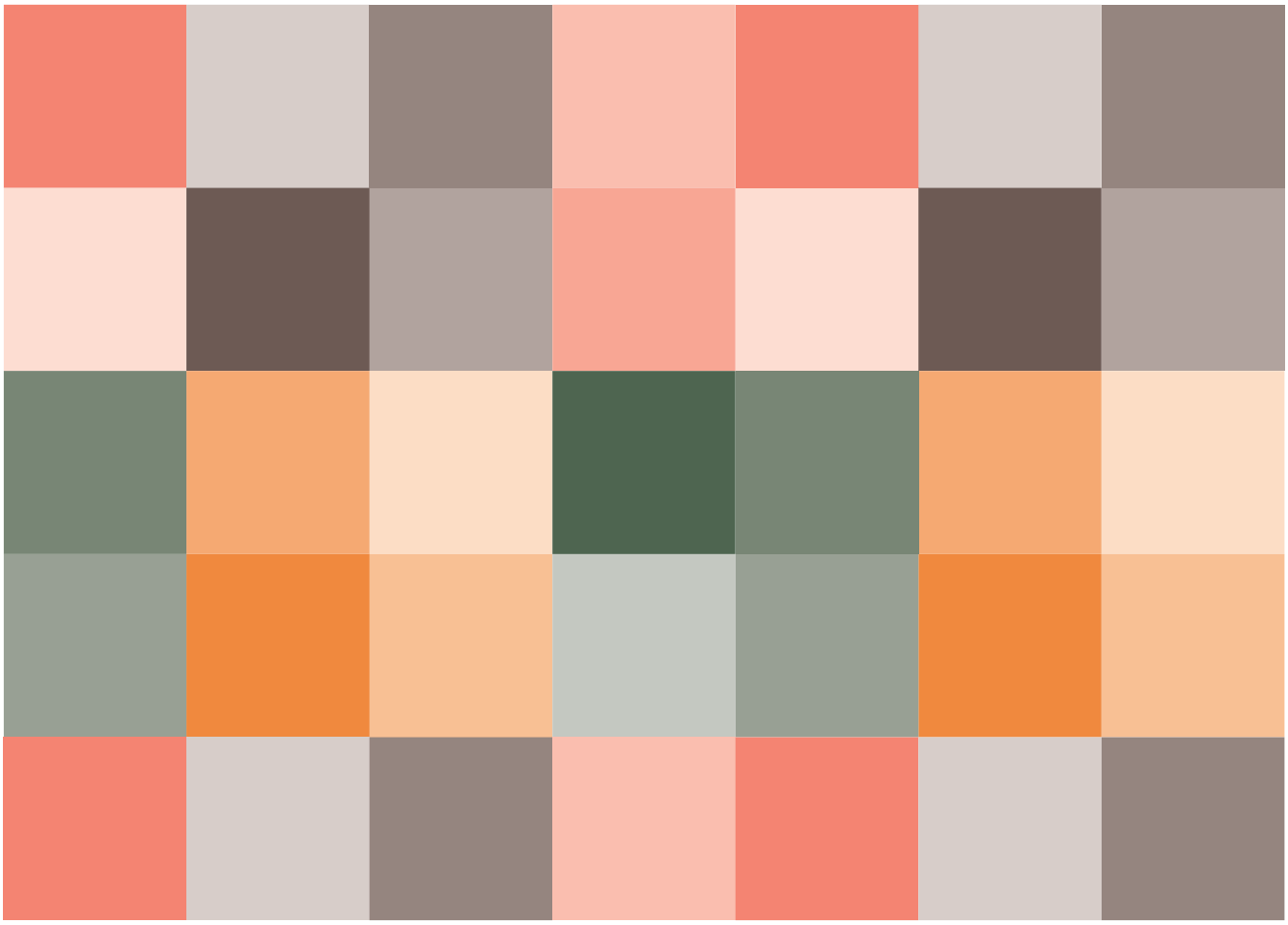


Metadata för AI-resurser på Sveriges dataportal

Bilaga till Slutredovisning av uppdrag till Myndigheten för digital förvaltning att föreslå en samverkansstruktur för utveckling av tjänster för data och artificiell intelligens

Diarienummer: Fi2026/00137

MYNDIGHETEN FÖR DIGITAL FÖRVALTNING
Agency for Digital Government

Datum: 2026-05-29

1. Inledning

Föreliggande rapport har tagits fram på uppdrag av Digg (Myndigheten för digital förvaltning) och syftar till att analysera hur AI-resurser kan beskrivas på Sveriges dataportal. Målet är att möjliggöra sökbarhet, spårbarhet och regulatorisk efterlevnad i det svenska AI-ekosystemet. Analysen avgränsar AI-resurserna till maskininlärningsmodeller, datamängder, AI-tjänster och associerad mjukvara, behandlingsmiljöer och riskbeskrivningar. Rapporten har särskilt beaktat de unika infrastrukturella förutsättningarna inom svensk offentlig förvaltning, vilket inkluderar hantering av säkra behandlingsmiljöer, krav- och tillitsmodeller samt integritetsbevarande metoder. En inventering av befintliga standarder, såsom DCAT-AP-SE, MLDCAT-AP, CPSV-AP och CCCEV, visar att ingen enskild standard täcker samtliga identifierade nationella behov. Baserat på denna analys föreslås därför etableringen av en nationell applikationsprofil för AI. Denna profil syftar till att integrera och harmonisera dessa internationella byggblock med specifika svenska krav, vilket säkerställer teknisk interoperabilitet och efterlevnad av nya regelverk såsom AI-förordningen.

Syftet med rapporten är att ge en strategisk inriktning för en nationell applikationsprofil för AI, snarare än att beskriva varje detalj. Arbetet inleds med en analys av svenska behov och hur dessa kan passas in på Sveriges dataportal genom att återanvända europeiska och internationella standarder. Vidare lämnas ett förslag på hur profilen för AI-resurser kan tas fram och förvaltas.

2. Analys

2.1 Identifiering och avgränsning av AI-resurser

Här analyseras vad som faktiskt utgör kärnan i AI-infrastrukturen och hur behoven ser ut.

Den centrala komponenten i den AI-infrastruktur som uppdraget pekar ut utgörs av **maskininlärningsmodellen**. I föreliggande analys avgränsas behovet av metadata till de aspekter som direkt bidrar till förståelsen av modellens tillämpning och användningsområde. Detta innebär ett medvetet val att exkludera djupare beroenden relaterade till modellens framtagande, reproducerbarhet eller bakomliggande vetenskapliga forskning. Maskininlärningsmodellen är i sin tur nära kopplad till olika typer av **datamängder**, vilket inkluderar data för träning, validering och testning, samt de indata som modellen exekveras på och resulterande utdata.

I de flesta fall sker användningen inte genom lokal driftsättning av användaren, utan via **AI-tjänster** där maskininlärningsmodeller tolkar och genererar data i ett tjänstebaserat gränssnitt. För att möjliggöra effektiv kommunikation mellan dessa modeller och externa system, mjukvara eller datakällor används ofta **Model Context Protocol** (MCP). MCP utgörs av mjukvarukomponenter som kan driftsättas antingen lokalt eller som publika tjänster. Utöver dessa AI-specifika resurser bör även traditionella datatjänster för distribution av modeller, som ett alternativ till direktnedladdning, beaktas. Vidare identifieras

riskbeskrivningar som en kritisk resurs; dessa dokumenterar potentiella risker vid modellanvändning samt tillhörande riskminimerande åtgärder.

Mot bakgrund av denna genomgång framstår det som ändamålsenligt att i den fortsatta analysen kategorisera de nödvändiga resurserna i maskininlärningsmodeller, datamängder och AI-tjänster, kompletterat med stödjande resurser för mjukvara och riskhantering.

2.2 Infrastrukturella förutsättningar i en svensk kontext

Här analyseras de specifika krav som den svenska offentliga förvaltningen ställer på miljön samt behovet av en enhetlig modell för tillit och anslutning.

I en svensk kontext, och i synnerhet inom offentlig förvaltning, tillkommer specifika begrepp och förutsättningar som kräver närmare definition. En central sådan resurs är behandlingsmiljöer, varav vissa av dessa är att betrakta som **säkra behandlingsmiljöer** (*Secure Processing Environment*, SPE), vilken utgörs av en kontrollerad fysisk eller virtuell plattform för analys av känsliga data, exempelvis hälsodata. Ett exempel på en behandlingsmiljö som finns idag är AI-verkstaden. En grundläggande princip för SPE är att data förblir inom miljön; användaren ges tillgång till analytiska verktyg lokalt, men kan endast exportera granskade och avidentifierade resultat.

För att realisera dessa analyser krävs en adekvat **beräkningsmiljö**, vilket omfattar den tekniska infrastruktur i form av hård- och mjukvara som möjliggör själva databehandlingen. Ett exempel på en beräkningsmiljö som finns idag är AI-fabriken. Vid resurskrävande AI-processer är det nödvändigt att beräkningskapaciteten tillgängliggörs i direkt anslutning till datans lagringsplats, särskilt inom ramen för en säker behandlingsmiljö. Det är därför av stor vikt att kunna definiera och beskriva beräkningsmiljöns egenskaper, oavsett om den verkar autonomt eller som en integrerad del av en skyddad miljö.

För att garantera efterlevnad av säkerhetskrav vid hantering av känslig information, tillämpas en tillitsmodell baserad på formella **krav** och **bevis** på att dessa uppfylls. Detta innebär att tillgång till en resurs styrs av fastställda **tillitsnivåer** (*Level of Assurance, LoA*) samt specifika kriterier rörande exempelvis organisationstillhörighet eller lagligt mandat. För att sänka trösklarna för användaren uppstår ett behov av att maskinellt kunna beskriva hur dessa krav verifieras, exempelvis via etablerade federationer som SIB (OpenID Federation) eller SAML-federationer. Detta fungerar som ett ramverk för att styra och beskriva behörighet till känsliga datamängder, tjänster och behandlingsmiljöer.

Vidare spelar **anslutningsmodellen** en avgörande roll i att förstå tillgängligheten hos en resurs. En infrastruktur i en svensk kontext måste kunna beskriva om en resurs nås direkt, via en teknisk brygga (proxy/adapter) eller genom en icke-federerad lösning. Även graden av automatisering i dessa processer – från manuell prövning till fullt automatiserad behörighetstilldelning – kan vara kritisk information för den som önskar nyttja resurserna.

Sammantaget innebär dessa nationella behov att beskrivningen av AI-resurser inte enbart kan fokusera på modellens tekniska prestanda, utan även måste omfatta de regulatoriska

och infrastrukturella ramverk som säkerställer en tillförlitlig, laglig och tekniskt tillgänglig hantering av data i en svensk kontext.

2.3 Mekanismer för integritetsskydd och innovation

Här analyseras hur juridik och teknik samverkar för att möjliggöra datadriven utveckling under strikt regelefterlevnad.

För att säkerställa ett robust skydd av personlig integritet vid avancerad dataanalys är det nödvändigt att etablera en enhetlig terminologi för **integritetsbevarande metoder**. En sådan terminologi bör kunna refereras till från såväl säkra behandlingsmiljöer och AI-tjänster som från specifika datamängder vilka genererats genom integritetshöjande processer. Genom att definiera dessa metoder som en sammanhållen struktur skapas en teknisk och juridisk flexibilitet; det möjliggör dels en successiv inkludering av framtida tekniska innovationer, dels en tillämpning inom kontexter som i dagsläget inte kan förutses.

Inom ramen för flöden i en säker behandlingsmiljö (SPE) spelar integritetsbevarande metoder som anonymisering, pseudonymisering, syntetiska data och *differential privacy* en avgörande roll. Dessa tekniker möjliggör exempelvis träning av maskininlärningsmodeller på individnära data, samtidigt som den enskildes integritet värnas. Implementeringen av dessa metoder är dock oskiljaktig från den krav- och tillitsmodell som omger miljön. För att en aktör ska ges tillgång till att applicera dessa metoder eller ta del av resultaten krävs en formell verifiering av identitet och behörighet. Här fyller standardiserade beskrivningar av krav och bevis en kritisk funktion. Genom dessa kan en SPE-ansvarig definiera exakt vilka kriterier (t.ex. "godkänd forskningsorganisation" eller "LoA3-identitet") som krävs för att få tillgång.

Genom att systematiskt beskriva både de tillämpade integritetsmetoderna och de kravställda tillitsnivåerna stärks tilliten till det samlade systemet. Det är denna obrutna kedja av transparens som det möjliggör datadriven innovation under strikt regelefterlevnad.

2.4 Befintlig metadatainfrastruktur som tekniskt fundament

Här analyseras hur den befintliga nationella infrastrukturen och etablerade metadatastandarder utgör fundamentet för beskrivningen av AI-resurser.

Sveriges dataportal, som förvaltas av Myndigheten för digital förvaltning (Digg), utgör den nationella infrastrukturen för att tillgängliggöra data som en strategisk samhällsresurs för innovation och transparens. Portalen fungerar som en metadatakatalog där information om datamängder och datatjänster beskrivs enligt DCAT-AP-SE. Utöver detta stödjer portalen även beskrivningar av applikationsprofiler samt begrepp och terminologier. Denna arkitektur, som vilar på principerna för länkade data och RDF, erbjuder robusta byggstenar som är väl lämpade att utvidgas för beskrivning av AI-resurser. Genom användningen av persistenta identifierare i form av URI:er säkerställs en entydig identitet för varje resurs; vid behov kan dessa kompletteras med ytterligare identifierare, såsom UUID:er, för att möjliggöra interoperabilitet med standarder utanför RDF-ekosystemet.

Genom att tillämpa dataportalens etablerade principer för informationskvalitet och tillgängliggörande kan AI-resurser integreras i det nationella ekosystemet för data på ett enhetligt och säkert sätt, där ansvaret för resursens korrekthet och aktualitet vilar kvar hos den tillhandahållande organisationen.

Analysen visar att även om DCAT-AP-SE tillhandahåller flera av de grundläggande komponenter som krävs, och portalens begreppsstöd möjliggör hantering av nödvändig terminologi, krävs ytterligare integration av väletablerade strukturer. Detta är nödvändigt för att fullt ut kunna beskriva de unika egenskaper som är förknippade med AI-specifika resurser och deras regulatoriska kontext.

I arbetet med att skapa en sammanhållen nationell infrastruktur för AI-resurser fungerar Diggs initiativ **INSPEC** (*Interoperable Specifications Profile*) som en metodologisk grund. Enligt INSPEC definieras en specifikation inte som ett enskilt statiskt dokument, utan som en abstrakt behållare av resurser – såsom informationsmodeller, terminologier och maskinläsbara scheman – vilka tillsammans preciserar hur data ska uttryckas och förstås.

Genom att tillämpa denna systematik i framtagandet av en svensk applikationsprofil för AI säkerställer vi att vi inte bara namnger metadatafält, utan skapar en profil som består av ett nätverk av formella definitioner. Arbetet bör därför fokusera på att återanvända existerande resurser från internationella initiativ likt MLDCAT-AP, och endast vid behov komplettera med nya definitioner som i sin tur utformas för att kunna återanvändas i andra sammanhang. Detta angreppssätt underlättar avsevärt för en framtida utbyggnad av applikationsprofilen; i en snabbt förändrad AI-miljö kan nya moduler och resurser adderas till behållaren utan att den befintliga strukturen bryts ner, vilket ger en nödvändig teknisk agilitet.

Individuella AI-resurser kan i detta ekosystem göras mer tillgängliga och validerbara genom att tydliggöra sin regelefterlevnad via egenskapen `dcterms:conformsTo`. Genom att den specifikation som resursen pekar på i sin tur är beskriven enligt INSPEC, skapas en obruten kedja av tillit och teknisk interoperabilitet. Detta tillvägagångssätt möjliggör en maskinell verifiering av att relationerna mellan data, modeller och tjänster följer de avsedda instruktionerna, vilket är en förutsättning för att AI-arbetsflöden ska kunna samverka säkert och effektivt över organisationsgränser.

2.5 Inventering av etablerade specifikationer och ramverk

Här analyseras internationella och europeiska standardiseringsinitiativ för att identifiera lämpliga byggblock för att realisera den nationella infrastrukturen.

Analysen fokuserar inledningsvis på metadata för maskininlärningsmodeller och AI-tjänster. För att skapa en praktiskt fungerande infrastruktur krävs dock även metadata för de stödjande komponenterna i ekosystemet; vilket omfattar de säkra behandlingsmiljöer där data bearbetas, och formella krav som styr åtkomsten till miljöer, tjänster och data.

2.5.1 Metadata för AI-modeller och datastrukturer

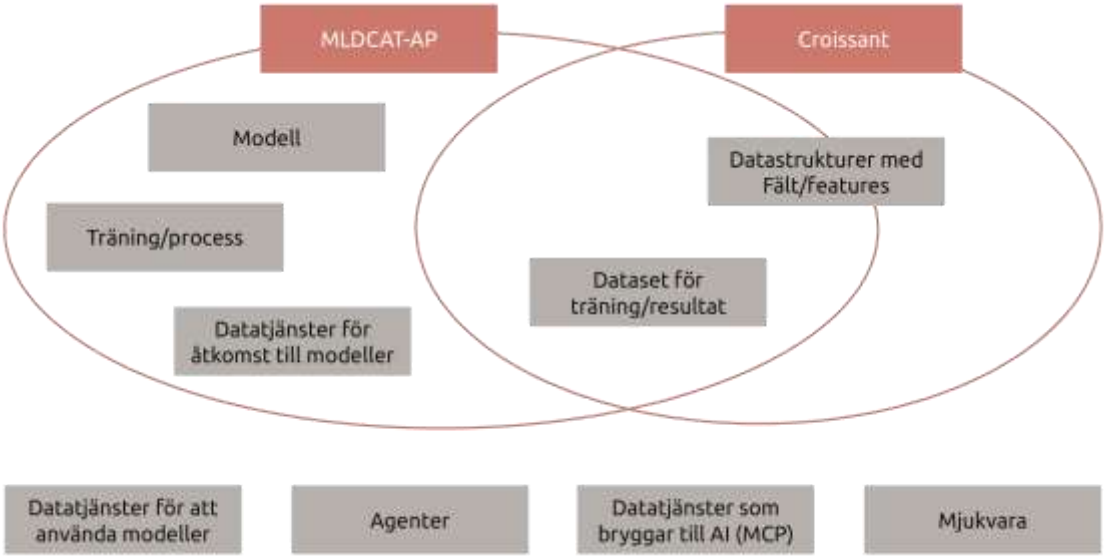
För att realisera de behov som identifierats krävs tillämpning av specifikationer som kan hantera de unika metadataattributen för maskininlärningsmodeller och deras datastrukturer. Två dominerande initiativ har identifierats: **MLDCAT-AP** och **Croissant**. Dessa kompletterar den befintliga infrastrukturen på olika sätt genom att adressera olika delar av AI-livscykeln.

MLDCAT-AP är ett initiativ inom ramen för EU:s SEMIC (Semantic Interoperability Community). Som en utvidgning av DCAT-AP är den särskilt relevant i en svensk och europeisk kontext, då den sömlöst kan integreras med den arkitektur som Sveriges dataportal redan vilar på. Specifikationen är omfattande och erbjuder vokabulär för att beskriva maskininlärningsmodeller, deras koppling till vetenskaplig forskning samt identifierade risker (*Harm Risk*), definierat som kombinationen av sannolikhet och allvarlighetsgrad för en skada.

Vidare innehåller MLDCAT-AP strukturer för att dokumentera **Uppgifter** (*Tasks*), **Flöden** (*Flows*) och **Körningar** (*Runs*). Dessa komponenter möjliggör en hög grad av reproducerbarhet genom att inkludera exakta mjukvaruberoenden, hyperparametrar och detaljerade utvärderingsresultat, vilket även underlättar analys av exempelvis modellens klimatpåverkan. Det bör dock noteras att dessa delar av specifikationen till stor del faller utanför den avgränsning som gjordes i avsnitt 2.1, där fokus ligger på metadata som bidrar till förståelse för modellens praktiska tillämpning och användningsområde snarare än dess tekniska återskapande. Det bör även noteras att specifikationen har kommit att anpassas för att fungera som en teknisk brygga för de dokumentationskrav som ställs i **AI-förordningen**, särskilt för modeller med allmänna ändamål (*General Purpose AI – GPAI*) samt den närliggande **AI CoP** (*The General-Purpose AI Code of Practice*).

Croissant (utvecklat av MLCommons) fokuserar istället primärt på att bädda in AI-relevant metadata direkt i datamängdsbeskrivningar för att göra dem "maskinläsbara" för AI-verktyg. Genom att bygga på schema.org möjliggör Croissant en djupare beskrivning av datastrukturer, såsom fält och särdrag (*features*), vilket underlättar upptäckt och användning i miljöer som Hugging Face, OpenML och Kaggle.

Dessa två standarder överlappar i sin beskrivning av datamängder för träning och resultat, men täcker i övrigt olika behov i ekosystemet. Nedanstående illustration visar hur MLDCAT-AP och Croissant förhåller sig till varandra samt till de angränsande resurserna för AI-tillämpning.



Figur 1. Överlapp och komplementaritet mellan MLDCAT-AP och Croissant.

2.5.2 Modell för AI-tjänster och dynamiska resurser

Vidare har analysen undersökt tillgången på specifikationer för att beskriva mer dynamiska resurser såsom AI-tjänster och Model Context Protocol (MCP). Här konstateras en avsaknad av bredare, internationellt etablerade ramverk specifikt framtagna för dessa resurstyper. Inom ramen för befintliga standarder bedöms dock klassen `dcat:DataService` i DCAT-AP kunna fungera som en teknisk grund för AI-tjänster. Genom att betrakta en AI-tjänst som en specialiserad datatjänst kan man dra nytta av existerande egenskaper, samtidigt som det finns ett behov av att utöka beskrivningen för att fånga unika beteenden såsom exekverad modell och aktiva skyddsmekanismer.

Denna brist på specifika standarder blir särskilt tydlig vid en granskning av MCP, som utgör en framväxande metod för att ge AI-agenter kontextuell tillgång till resurser. Trots stöd från flera stora aktörer tillhandahåller befintliga register i regel endast rudimentär information såsom titel, beskrivning och parametrar för anslutning. Ur ett förvaltnings- och katalogiseringsperspektiv innebär MCP-begreppet i praktiken en sammanblandning av två olika resurstyper:

- **Körande tjänster:** Likställs med en `dcat:DataService` med fokus på en aktiv slutpunkt för interaktion.
- **Källkod och bibliotek:** Saknar direkt motsvarighet i DCAT-standarden, vilket skapar ett glapp i hur dessa resurser ska kategoriseras och beskrivas.

2.5.3 Vokabulärer för krav, tillit och verifiering

För att reglera åtkomst till känsliga resurser, såsom datamängder i en säker behandlingsmiljö eller specifika AI-tjänster, krävs en metod för att formellt uttrycka

behörighetskrav och verifieringsmetoder. Här identifieras **CCCEV** (*Core Criterion and Core Evidence Vocabulary*) som ett centralt verktyg för att realisera den krav- och tillitsmodell som beskrivs i avsnitt 2.3.

CCCEV är medvetet utformat för att vara generellt, vilket innebär att dess kärnklasser kan appliceras brett på alla typer av resurser i AI-ekosystemet:

- **Requirement (Krav):** Genom denna klass kan en resursägare definiera exakta kriterier som måste uppfyllas för åtkomst. I en svensk kontext utgår dessa ofta från det nationella tillitsramverket (baserat på ISO/IEC 29115) och de etablerade tillitsnivåerna LoA2, LoA3 och LoA4. Kraven kan dock även omfatta organisatoriska kriterier, såsom att användaren måste tillhöra en "godkänd forskningsorganisation" eller inneha ett specifikt lagligt mandat.
- **EvidenceType (Bevistyp):** Denna klass används för att beskriva vilka typer av bevis eller intyg som accepteras för att styrka att ett krav är uppfyllt. Det kan röra sig om tekniska bevis utfärdade via etablerade federationer som SIB (OpenID Federation) eller SAML-federationer, men även fristående intyg.

Genom att använda dessa klasser skapas en maskinläsbar brygga mellan resursen och tillitsinfrastrukturen vilken gör det tydligt vad som krävs av den aktör som vill nyttja resursen. För säkra behandlingsmiljöer (se 2.5.4) kan dessa klasser kopplas direkt, för en datamängd eller en AI-tjänst kan kopplingen göras till tillitsnivåer via egenskapen `dcterms:accessRights` (åtkomsträttigheter) i DCAT-AP.

Analysen visar att de svenska tillitsnivåerna med fördel kan representeras genom de URI:er som tillhandahålls av Sweden Connect (t.ex. `http://id.elegnamnden.se/loa/1.0/loa2`). En initial analys visar att även om dessa identifierare ursprungligen är framtagna för e-legitimering, utgör de den mest stabila nationella referenspunkten för att uttrycka krav på säkerhet vid autentisering. Denna slutsats bör verifieras i ett nästa skede.

Det är dock viktigt att notera den begreppsmässiga skillnaden gentemot EU:s eIDAS-förordning, som definierar tre egna tillitsnivåer: *låg*, *väsentlig* och *hög*. För att säkerställa interoperabilitet i en europeisk kontext bör metadata-arkitekturen ha kapacitet att hantera båda dessa referenssystem, särskilt vid gränsöverskridande datadelning inom ramen för AI-förordningen.

2.5.4 Beskrivning av säkra behandlingsmiljöer, beräkningsmiljöer och anslutningsvägar

När det gäller beskrivningen av **säkra behandlingsmiljöer (SPE)** har analysen identifierat fyra möjliga spår baserade på etablerade internationella vokabulär. Valet av uttryck styr huruvida miljön betraktas som en teknisk resurs, ett bibliotek eller en samhällstjänst:

- **Infrastrukturspåret (MLDCAT-AP):** Genom att kombinera klasser för datorinfrastruktur (*ComputingInfrastructure*), mjukvarubibliotek (*Library*) och fysiska komponenter (*Hardware*) kan en SPE beskrivas utifrån sin tekniska kapacitet och de resurser som krävs för drift.

- **Förvaltningsspåret (ADMS):** Genom *AssetRepository* kan en SPE ses som ett system för lagring och underhåll av tillgångar. Detta fokuserar på miljöns roll som förvaltare av beskrivningar och åtkomstpunkter.
- **Applikationsspåret (Schema.org):** Här beskrivs miljön som en mjukvaruapplikation (*SoftwareApplication*) med specifika ingångar (*EntryPoint*) och plattformar för exekvering (*actionPlatform*). Detta spår är starkt kopplat till webbaserade protokoll.
- **Tjänstespåret (CPSV-AP):** Detta spår beskriver en SPE som en offentlig tjänst (*PublicService*) som tillhandahålls av en organisation (*PublicOrganisation*). Det möjliggör beskrivning av kostnader (*Cost*) och de kanaler (*Channel*) genom vilka användaren interagerar med miljön.

Bland dessa framstår **CPSV-AP** (Common Public Service Vocabulary) som det mest intressanta alternativet för den svenska kontexten. Eftersom en SPE ofta utgörs av en sammansatt tjänst snarare än en enskild teknisk komponent, medger CPSV-AP en beskrivning som täcker både de legala och organisatoriska aspekterna. Genom att utgå från en tjänsteorienterad modell kan metadata om miljön inkludera villkor för användning och organisatoriskt ansvarstagande, vilket är centralt för att bygga tillit. Genom användning av kanaler (`cpsv:Channel`) blir det möjligt att i metadata precisera hur användaren praktiskt når tjänsten vilket kan omfatta viktiga parametrar såsom anslutningstyp och grad av automatisering av behörighetskontroll. De integritetsbevarande metoder som används i den säkra behandlingsmiljön beskrivs via regler (`cpsv:Rule`), med kopplingar till en etablerad terminologi för dessa.

Genom att kombinera tjänstebeskrivningen med dessa kanalspecifika attribut ges informationskonsumenten en tydlig bild av den tekniska och administrativa vägen fram till data.

Analysen pekar på möjligheten att betrakta **beräkningsmiljön** som en fristående entitet i de fall där ett sådant behov identifieras, exempelvis när en specifik beräkningsresurs görs tillgänglig för flera olika säkra behandlingsmiljöer (SPE). För att uppnå en enhetlig beskrivningsmodell kan även beräkningsmiljön betraktas som en public service (`cpsv:PublicService`), med fördelen att den vid behov kan dubbeltypas som teknisk infrastruktur (`it6:ComputingInfrastructure` från MLDCAT-AP). En sådan hantering medger en specifik beskrivning av de krav som ställs på de anslutande parterna, samtidigt som de tekniska egenskaperna hos resursen bibehålls.

Denna ansats innebär att en organisation kan definiera en beräkningsmiljö som en separat resurs när det är lämpligt, för att sedan beskriva hur den tillgängliggörs inom ramen för olika säkra behandlingsmiljöer med varierande legala och administrativa förutsättningar. Som en avslutande reflektion möjliggör detta synsätt att tjänster kan aggregeras; genom att utnyttja att publika tjänster kan ha inbördes beroenden kan en säker behandlingsmiljö beskrivas som beroende av en separat beräkningsmiljö. Detta skapar en modulär arkitektur där komplexa beroendekedjor mellan miljöer, beräkningsresurser och behörighetskrav kan uttryckas utan att tvinga fram en onödig fragmentering i enklare användarfall.

Sammanfattande analys: Behovet av en integrerad profil

Analysen av tillgängliga specifikationer visar sammanfattningsvis att stora delar av de relevanta AI-resurserna kan beskrivas med etablerade metoder, men att ingen enskild standard täcker hela det nationella behovet. Nedanstående gap-analys sammanställer hur de inventerade byggblocken svarar mot de identifierade behoven:

Behovsområde	**Tillgänglig standard (Byggblock)**	**Status**
Grundläggande metadata	DCAT-AP	Täcks väl av DCAT-AP-SE
Modellspecifik metadata	MLDCAT-AP	Täcks väl
AI-tjänster och MCP	DCAT-AP	Täcks delvis genom en utökning av DCAT-AP-SE
Formella behörighetskrav	CCCEV	Täcks delvis, kräver nationell profilering
Säkra behandlingsmiljöer	CPSV-AP	Täcks delvis, kräver profilering som tydliggör kopplingen till övriga AI-resurser
Beräkningsmiljöer	MLDCAT-AP	Täcks väl, kan behöva komplettering från CPSV-AP

Genom att kombinera grundstandarden DCAT-AP med specialiserade ramverk som MLDCAT-AP med generella vokabulärer som CCCEV för kravställning och CPSV-AP för tjänsteinteraktion, skapas en teknisk förmåga att beskriva de relevanta delarna av AI-ekosystemet.

Samtidigt visar analysen att det finns behov av en nationell sammanhållen metadataprofil som formellt sammanför dessa delar, adresserar identifierade glapp och säkerställer en enhetlig beskrivning av AI-ekosystemets samtliga komponenter.

2.6 AI förordningen och AI CoP

Här analyseras hur den föreslagna metadatastrukturen förhåller sig till kraven i AI-förordningen samt hur harmonisering sker med framväxande uppförandekoder.

I utvecklingen av MLDCAT-AP 3.0.0 har SEMIC integrerat explicita referenser till AI-förordningen (AI Act), särskilt avseende kraven på teknisk dokumentation för AI-modeller med allmänna ändamål (GPAI). Specifikationen är utformad för att fungera som en teknisk bärare av den information som krävs enligt förordningens bilagor XI och XII. En viktig aspekt i detta arbete är även anpassningen till AI Code of Practice (AI CoP), där MLDCAT-

AP strävar efter att vara kompatibel med framtagna mallar såsom *Model Documentation Form*. Detta innebär att den information en leverantör sammanställer för att efterleva uppförandekoden direkt kan speglas i metadata, vilket minskar dubbelarbete och säkerställer att dokumentationen är maskinläsbar och sökbar i en nationell kontext.

Stora delar av de referenser som finns i MLDCAT-AP täcker de behov som identifierats som kärnresurser i denna analys, inklusive beskrivning av träningsdata, avsedd användning och riskhantering. Genom att i föreliggande analys fokusera på en tillämpningsorienterad profil har dock vissa tekniska klasser i specifikationen medvetet exkluderats. Analysen av dessa bortval visar att vi genom att utelämna klasser som `it6:Run`, `it6:Flow` och `it6:Task` förlorar den direkta kopplingen till AI-förordningens krav på detaljerad teknisk spårbarhet och reproducerbarhet i metadata. På samma sätt innebär exkluderingen av bibliografiska referenser (`lpwcc:Paper` och `biro:BibliographicReference`) att den explicita länken till den vetenskapliga validering som betonas för högrisk-AI saknas i den förenklade profilen.

Det bör dock understrykas att dessa exkluderingar inte innebär ett brott mot specifikationens krav, då MLDCAT-AP är modulärt uppbyggd och de berörda klasserna är definierade som valfria. En profil fokuserad på maskininlärningsmodeller, datamängder och tjänster är därmed fullt valid enligt standarden. Vidare säkerställer användningen av persistenta identifierare (URI:er) att den nationella profilen förblir öppen och länkningsbar. Det skapas därmed en maskinläsbar brygga mellan den tillämpningsorienterade beskrivningen på Sveriges dataportal och mer omfattande dokumentation fokuserad på reproducerbarhet som kan publiceras på andra plattformar.

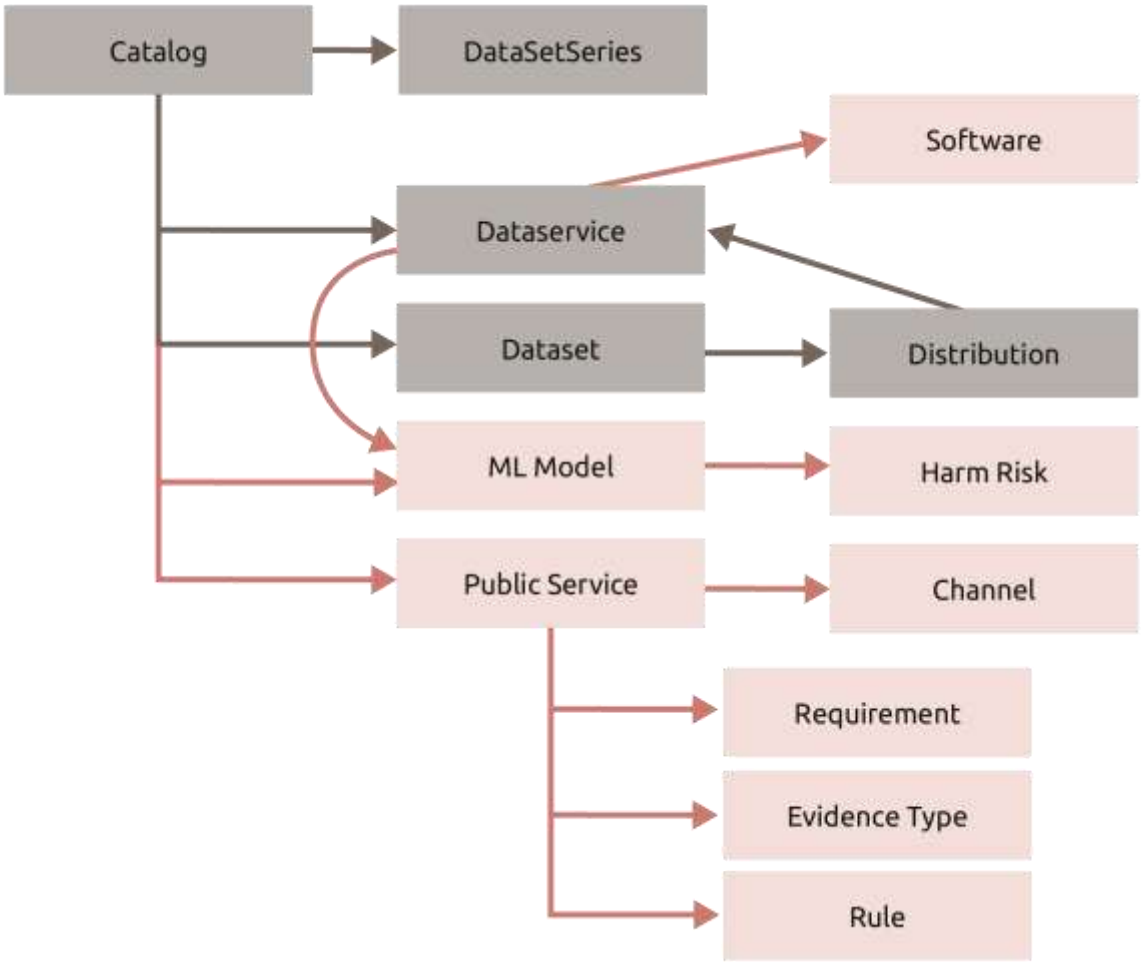
Då den föreslagna profilen förblir tekniskt kompatibel med MLDCAT-AP finns inget hinder för att en informationsleverantör väljer att tillhandahålla båda informationsmängderna tillsammans. Detta underlättar för leverantörer som redan har en fullständig dokumentation i enlighet med AI-förordningens mest omfattande krav, samtidigt som det möjliggör för den nationella infrastrukturen att bibehålla fokus på tillämpbarhet och sökbarhet för slutanvändaren.

Det bör tydliggöras att vissa delar inte är lösta ännu, t.ex. hur tillämplighet, roller, riskkategori och dokumentation ska uttryckas i metadata.

3. Förslag till nationell profil och förvaltningsmodell

Baserat på analysen i kapitel 2 föreslås etablerandet av en nationell applikationsprofil för AI-resurser. Förslaget syftar till att skapa en teknisk och organisatorisk brygga som möjliggör sökbarhet, spårbarhet och regulatorisk efterlevnad i det svenska AI-ekosystemet.

3.1 Specifikationsarbete: DCAT-AP-SE med AI-utvidgning



Figur 2. Översiktsbild för applikationsprofilen, delar i grått finns redan i DCAT-AP-SE, övriga delar införs som del av AI-utvidgningen.

Det centrala förslaget är att etablera en applikationsprofil vilken integrerar nödvändiga delar från MLDCAT-AP och CPSV-AP för att möjliggöra beskrivning av AI-resurser. För det praktiska införandet den nationella infrastrukturen identifieras två huvudsakliga vägar:

- **Alternativ A: Utvidgning av DCAT-AP-SE.** I detta scenario integreras AI-resurserna direkt i nästa version av den nationella grundstandarden. Detta skapar en sammanhållen specifikation för alla typer av digitala tillgångar och säkerställer att kopplingarna mellan träningsdata, modeller och tjänster uttrycks inom ett och samma valideringsbara ramverk. DCAT-AP-SE har redan en etablerad referensgrupp och uppdateringsprocess.
- **Alternativ B: Etablering av separat profil.** Här skapas en separat men strikt kompatibel profil som fungerar som en utvidgningsmodul till DCAT-AP-SE. Detta möjliggör en snabbare iterationstakt för AI-specifika behov, särskilt under initial utveckling, samtidigt som full interoperabilitet med existerande katalogtjänster

bibehålls. Samtidigt förblir grundprofilen enklare för de som inte har behov av att beskriva AI-resurser.

Oavsett vald väg ska specifikationen utformas för att:

- **Kombinera DCAT-AP-SE med MLDCAT-AP:** Fokus ligger på att integrera kärnklasser för maskininlärningsmodeller (`ml:MachineLearningModel`) med etablerade klasser för datamängder (`dcat:Dataset`) och distributioner (`dcat:Distribution`). Detta skapar en obruten kedja från rådata till färdig modell och AI-förädlad data.[1]
- **Integrera säkra behandlingsmiljöer (SPE) och beräkningsmiljöer:** Genom att använda **CPSV-AP** inkluderas SPE som en sökbar resurs (`cpsv:PublicService`). Genom användning av kanaler (`cpsv:Channel`) går det att beskriva de tekniska och praktiska steg som krävs för att en användare ska få tillgång till miljön. Förslaget medger en enhetlig beskrivningsmodell där även beräkningsmiljöer betraktas som tjänster, med möjligheten att vid behov dubbeltypa dessa som teknisk infrastruktur (`it6:ComputingInfrastructure`). Detta möjliggör förvaltning av beräkningsmiljöer som fristående entiteter vilka kan göras tillgängliga inom ramen för olika legala och tekniska kontexter genom aggregering av tjänster, där en säker behandlingsmiljö kan uttrycka ett beroende till en separat beräkningsmiljö.
- **Harmonisera tillträdeskrav och verifiering via CCCEV:** För att beskriva åtkomstrestriktioner används klasserna krav (`cv:Requirement`) och bevistyp (`cv:EvidenceType`). Detta möjliggör ett standardiserat sätt att ställa krav på exempelvis tillitsnivå (LoA) och organisationstillhörighet, samt att definiera vilka bevis som accepteras. Dessa har stöd direkt i CPSV-AP för den säkra behandlingsmiljön eller via åtkomsträttigheter (`dcterms:accessRights`) för datamängder och AI-tjänster.
- **Standardisera AI-tjänster via DCAT:** Genom en förfining (*refinement*) av kraven på `dcat:DataService` skapar en enhetlig metod för att beskriva AI-tjänster och MCP-tjänster, vilket förenklar maskinell upptäckt utan att öka modellens komplexitet.
- **Integritetsbevarande metoder:** En gemensam terminologi för integritetsbevarande metoder (Privacy Enhancing Technologies, PET) etableras, vilket är kritiskt för delning av känsliga datamängder och resultat inom ramen för en SPE. Denna knyts an till public service via regler (`cpsv:Rule`) och det finns även möjlighet att knyta an till dess från beskrivningen för hur en datamängd har genererats
- **Koppling mellan MCP och mjukvara:** En lämplig stödjande klass (exempelvis via *Schema.org* eller *DOAP*) identifieras för att hantera kopplingen mellan en MCP-

[1] MLDCAT-AP har valts som primär utgångspunkt framför alternativ som Croissant, då den förstnämnda är utformad för att harmonisera med DCAT-standarden och de specifika transparenskrav som ställs i AI-förordningen. Medan Croissant fokuserar på den interna datastrukturen och maskinläsbarhet för träningsbibliotek, erbjuder MLDCAT-AP den administrativa och regulatoriska kontext (såsom riskbeskrivningar och organisatoriskt ansvar) som krävs för en nationell dataportals infrastruktur.

tjänst och den specifika mjukvara som exekveras. Detta möjliggör även beskrivning av MCP som ej tillhandahålls som en tjänst.

3.2 Harmoniseringsarbete och bevakning på EU-nivå

Eftersom MLDCAT-AP 3.0 för närvarande är ett utkast (draft), är en aktiv bevakning och påverkan en central del av förslaget.

- **Aktivt deltagande:** Sverige bör delta i de europeiska arbetsprocesserna, vilka leds av SEMIC, för att säkerställa att nationella behov – såsom specifika behov kring säkra behandlingsmiljöer – förblir kompatibla med den europeiska standarden.
- **Kontrollerade vokabulärer:** Särskilt fokus bör läggas på rekommendationer kring riskbeskrivningar och transparensrapportering med direkt koppling till AI-förordningen.
- **Identifierarstrategi:** En strategi för persistenta och länkbara identifierare (PID) måste etableras. Detta är avgörande för att kunna referera till modeller och resurser som skapas utanför den egna organisationen, men som tillgängliggörs via den nationella infrastrukturen.

3.3 Integrering och vägledning

För att förslaget ska få praktiskt genomslag krävs åtgärder på både system- och användarnivå:

- **Systemanpassning:** Sveriges dataportal behöver utveckla förmågan att skörda, validera och presentera de nya resurstyperna och deras inbördes relationer.
- **Praktisk vägledning:** Pedagogiska guider för informationsleverantörer ska tas fram. Dessa ska steg för steg visa hur en myndighet dokumenterar allt från en säker behandlingsmiljö till den specifika användningen av en maskininlärningsmodell, inklusive kopplingar till legala krav.

Bilaga A: Ordlista och förkortningar

Förkortning/Term	Definition
AI CoP	*The General-Purpose AI Code of Practice*. Uppförandekod för AI med allmänna ändamål.
AI-förordningen (AI Act)	Regelverk från EU som ställer krav på bland annat teknisk dokumentation för AI-modeller.
AI-tjänster	Tjänstebaserade gränssnitt där maskininlärningsmodeller tolkar och genererar data.
Applikationsprofil	Tillägg eller anpassning av en specifikation för mer specifika användningsfall. Oftast för att möta domänspecifika eller nationella behov.
Beräkningsmiljö	Den tekniska infrastruktur (hård- och mjukvara) som möjliggör själva databehandlingen, t.ex. AI-fabriken.
Bevis	Dokumentation eller intyg som används i tillitsmodellen för att styrka att formella krav uppfylls.
CCCEV	*Core Criterion and Core Evidence Vocabulary*. Vokabulär för att formellt uttrycka behörighetskrav och verifieringsmetoder.
Croissant	Standard (utvecklad av MLCommons) som syftar till att bädda in AI-relevant metadata direkt i datamängdsbeskrivningar.
CPSV-AP	*Common Public Service Vocabulary*. Används för att beskriva resurser som offentliga tjänster (*PublicService*), vilket föreslås för säkra behandlingsmiljöer.
Datamängd	En samling data, publicerad eller förvaltad av en enskild aktör, tillgänglig för åtkomst eller nedladdning i en eller flera representationer. Ur ett AI-perspektiv är datamängden relevant för träning, validering och testning, samt indata och resulterande utdata från en maskininlärningsmodell.

DCAT	*Data Catalog Vocabulary*. En rekommendation från W3C om hur man uttrycker dataktaloger. Rekommendationen tillhandahåller en mängd olika uttryck som kan specialiseras för olika användarfall genom så kallade applikationsprofiler
DCAT-AP	En applikationsprofil av DCAT framtagen av EU:s SEMIC med målsättningen att förenkla utbytet av datakataloger inom EU. (Borde egentligen heta DCAT-EU.)
DCAT-AP-SE	Sveriges nationella applikationsprofil för metadatakatalogisering (baserad på DCAT-AP).
Digg	Myndigheten för digital förvaltning.
GPAI	*General Purpose AI*. AI-modeller med allmänna ändamål. Omfattas av särskild reglering i AI-förordningen till skillnad från mer specialiserad AI, som exempelvis översättning mellan olika naturliga språk.
Harm Risk	Identifierad risk definierad som kombinationen av sannolikhet och allvarlighetsgrad för en skada.
INSPEC	*Interoperable Specifications Profile*. En metodologisk grund där en specifikation definieras som en container av resurser (informationsmodeller, terminologier, scheman).
Integritetsbevarande metoder (PET)	*Privacy Enhancing Technologies*. Tekniska metoder som anonymisering, pseudonymisering, syntetiska data och *differential privacy* för att skydda personlig integritet.
Krav	Kriterier som måste uppfyllas för åtkomst till en resurs, tjänst eller miljö, t.ex. LoA-nivå eller organisationstillhörighet.
LoA	*Level of Assurance*. Se **Tillitsnivåer**.
Maskininlärningsmodell	Den centrala komponenten i AI-infrastrukturen.
MCP	*Model Context Protocol*. Mjukvarukomponenter som möjliggör effektiv kommunikation (kontextuell tillgång) mellan AI-agenter och externa resurser.

MLDCAT-AP	Metadata Standard (inom EU:s SEMIC) som är en utvidgning av DCAT-AP för att beskriva maskininlärningsmodeller och deras koppling till forskning och risker.
PET	*Privacy Enhancing Technologies*. Se **Integritetsbevarande metoder**.
RDF	*Resource Description Framework*. Arkitekturen för länkade data som Sveriges dataportal vilar på.
Riskbeskrivning	Dokumentation av potentiella risker vid modellanvändning samt tillhörande riskminimerande åtgärder.
SAML	*Security Assertion Markup Language*. En öppen standard för utbyte av autentiserings- och behörighetsuppgifter mellan parter.
SIB	*Samordnad identitet och behörighet*. Nationell infrastruktur för inloggning och behörighetskontroll, från Digg. Skalbar och kompatibel med moderna standarder, inklusive OpenID Federation.
SPE	*Secure Processing Environment*. Se **Säkra behandlingsmiljöer**.
Säker behandlingsmiljö (SPE)	*Secure Processing Environment*. Kontrollerad fysisk eller virtuell miljö för behandling av känsliga data (t.ex. hälsodata), där data förblir inom miljön.
Tillitsnivå (LoA)	*Level of Assurance*. Fastställda nivår (t.ex. LoA2, LoA3, LoA4) för säkerhet vid autentisering och behörighetskontroll.
URI	*Uniform Resource Identifier*. Persistent identifierare som säkerställer en entydig identitet för varje resurs.
UUID	*Universally Unique Identifier*. Ytterligare unik identifierare som kan komplettera URI:er.

Bilaga B: Referenser

Standard/Specifikation	Typ	Länk
AI CoP (*The General-Purpose AI Code of Practice*)	Uppförandekod	https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai
AI-förordningen (*AI Act*)	EU-förordning (Regelverk)	https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
CCCEV (*Core Criterion and Core Evidence Vocabulary*)	Europeisk vokabulär (SEMIC)	https://semiceu.github.io/CCCEV/releases/2.1.0
CPSV-AP (*Common Public Service Vocabulary Application Profile*)	Europeisk vokabulär (SEMIC)	https://semiceu.github.io/CPSV-AP/releases/3.2.0
Croissant	Global datastandard (MLCommons)	http://mlcommons.org/croissant/1.0
DCAT-AP (*Data Catalog Vocabulary Application Profile*)	Europeisk metadatastandard (SEMIC)	https://semiceu.github.io/DCAT-AP/releases/3.0.1/
DCAT-AP-SE (*Sveriges nationella applikationsprofil*)	Nationell metadatastandard	https://docs.dataportal.se/dcat/3.0.1/sv/
INSPEC (*Interoperable Specifications Profile*)	Nationell metodologisk grund (Digg)	http://w3id.org/inspec/specification
MLDCAT-AP (*Machine Learning Data Catalog Vocabulary Application Profile*)	Europeisk metadatastandard (SEMIC)	https://semiceu.github.io/MLDCAT-AP/releases/3.0.0