



Vägledning för implementering av kriterier för FAIR forskningsdata

Vägledning för implementering av kriterier för FAIR forskningsdata

Dnr 5.2-2018-01016

Swedish Research Council
Vetenskapsrådet
Box 1035
SE-101 38 Stockholm, Sweden

Innehållsförteckning

| | |
|---|-----------|
| Inledning | 4 |
| FAIR principer och kriterier | 6 |
| F1. (Meta)data are assigned a globally unique and persistent identifier..... | 6 |
| F2. Data are described with rich metadata..... | 7 |
| F3. Metadata clearly and explicitly include the identifier of the data it describes | 8 |
| F4. (Meta)data are registered or indexed in a searchable resource..... | 9 |
| A1. (Meta)data are retrievable by their identifier using a standardized communications protocol..... | 10 |
| A1.1. The protocol is open, free, and universally implementable | 11 |
| A1.2. The protocol allows for an authentication and authorization procedure, where necessary | 12 |
| A2. Metadata are accessible, even when the data are no longer available | 13 |
| I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation | 14 |
| I2. (Meta)data use vocabularies that follow FAIR principles..... | 15 |
| I3. (Meta)data include qualified references to other (meta)data..... | 16 |
| R1. (Meta)data are richly described with a plurality of accurate and relevant attributes | 17 |
| R1.1. (Meta)data are released with clear and accessible data usage license... .. | 18 |
| R1.2. (Meta)data are associated with detailed provenance | 19 |
| R1.3. (Meta)data meet domain-relevant community standards | 20 |
| Begreppslista | 21 |

Inledning

För att forskningsdata ska anses vara FAIR ("FAIR forskningsdata"), det vill säga hanterade enligt de så kallade FAIR-principerna, har ett antal kriterier tagits fram som ska bidra till att forskningsdata uppfyller krav på att vara sökbara (Findable), tillgängliga (Accessible), interoperabla (Interoperable), och återanvändningsbara (Reusable).

För att FAIR datahantering ska kunna implementeras i forskningsprocessen har Vetenskapsrådet tagit fram en mer detaljerad beskrivning i form av en vägledning till dessa kriterier.

Vägledningen är skriven i första hand för att bidra till ökad tydlighet i vad som krävs att implementera FAIR datahantering samt för att utgöra ett stöd för framförallt de stödjande funktionerna för datahantering vid lärosäten och infrastrukturer, där anpassning till FAIR-datahantering och praktisk implementering av FAIR-principerna ofta sker. Vägledningen innefattar både internationella och nationella aspekter som kan vara relevanta vid implementeringen. Avsikten är att formuleringen av kriterierna ska vara allmängiltig, teknikoberoende och så beständig över tid som möjligt. Kriterierna har därför formulerats på konceptuell nivå, som en tolkning av var och en av de femton FAIR principerna. I varje kriteriums vägledning har vägledande frågor formulerats som kan besvaras med ja eller nej. Ett kriterium kan uppfyllas på olika sätt och i olika grad. Om tillämpliga vägledande frågor besvaras jakande kan kriteriet betraktas vara uppfyllt. Observera att det i slutet på vägledningen finns en begreppslista, där olika begrepp förklaras närmare och exempel ges.

För att tillämpa FAIR kriterierna krävs ofta att ett antal olika stödjande tjänster finns på plats som kan användas i ekosystemet för forskningsdatahantering. Vissa av dessa är generella för alla typer av forskningsdatahantering och andra är domänspecifika.

Exempel på tillämpning kan vara användning av tjänster som tillhandahåller identifierare som möjliggör unik identifiering av de olika digitala objekten, allt från dataset till begrepp. Behovet av unika och beständiga identifierare som kan användas för olika typer av digitala objekt och på olika nivåer är således grundläggande inom alla områden. Dessa är också en av förutsättningarna för att data ska vara sökbara av både människor och maskiner och ökar möjligheter till att hitta och koppla samman relevanta data och kunna samarbeta över disciplins- och landsgränser.

För att uppfylla kriterierna för interoperabilitet behöver de olika digitala objektens betydelse definieras. Då betydelser naturligt skiljer sig åt mellan olika domäner, t ex medicinska begrepp jämfört med arkeologiska begrepp, innebär det att en del termer som används vid beskrivning av data är domänspecifika.

Genom användning av FAIR-kriterier ökar möjligheterna att förtydliga vad som avses med olika begrepp inom och mellan olika discipliner.

Även digitala objekt, t ex data, som inte är skapade av forskarna själva används ofta som källa till olika typer av forskning, exempel på detta är data från register. I sådana fall bör FAIR kriterierna appliceras på dessa digitala objekt redan när de skapas. Detta då unika identifierare, tydliga maskinläsbara definitioner, insamling av information om digitala informationens proveniens m.m. behöver tillämpas genom hela datahanteringsprocessen för att uppnå god vetenskaplig kvalitet, inte bara under den del som kontrolleras av forskaren.

Data behöver hanteras på ett sätt som skapar förutsättningar för reproducerbarhet, tillförlighet och användbarhet, både i dag och i framtiden. Att möta FAIR-kriterierna är ett steg i den processen.

När det gäller användning av stödjande digitala tjänster och funktioner för implementering av FAIR omfattas dessa av särskilda tekniska beskrivningar. Ett exempel på sådan teknisk beskrivning är referensarkitekturen för EOSC Interoperability Framework, som är framtagen med datahantering enligt FAIR-principerna som utgångspunkt.

Referensarkitekturen utgör en mer teknisk beskrivning av EOSC Interoperability Framework och beskriver en struktur för att vägleda arbetet med att nå god interoperabilitet mellan datasystem och tjänster inom EOSC. Fokus ligger på de ramverk och tjänster som krävs för att möjliggöra ett "Minimal Viable EOSC", dvs för att implementera grundläggande FAIR datahantering och interoperabilitet. Arkitekturen är planerad att utvidgas och utvecklas baserat på återkoppling från olika communitys och vetenskapliga domäner.

Med förhoppningen att kriterierna, vägledningen för implementering samt de vägledande frågorna nedan ska bidra till att skapa en harmoniserad tolkning av FAIR principerna och underlätta ert arbete!

FAIR principer och kriterier

F1. (Meta)data are assigned a globally unique and persistent identifier

Principen lyfter fram behovet av ett system som är beständigt över tid och som avser att unikt identifiera ett digitalt objekt eller samling av objekt.

Kriterium

Data och metadata har en unik och beständig identifierare.

Vägledning för implementering

Det digitala objektet har en tilldelad uppslagingsbar globalt unik beständig identifierare (eng Persistent Identifier (PID)).

Vägledande frågor

- Har metadata (beskrivande information som hänvisar till data) en tilldelad globalt unik beständig identifierare?
- Har data en tilldelad globalt unik beständig identifierare?
- Är identifieraren uppslagingsbar?
- Finns det någon organisation som garanterar och kontrollerar kontinuerligt att identifieraren alltid länkar till metadata som definierar det digitala objektet?
- Administreras identifieraren av en organisation som garanterar dess beständighet över tid?

F2. Data are described with rich metadata

För att kunna hitta ett digitalt objekt utan förkunskaper behöver denna vara beskriven med rikt metadata så att den tjänst som används för sökningen inte bara kan använda objektets namn, eller motsvarande, i sin sökning utan också andra typer av relaterade beskrivningar. Exempel på sådana är termer och begrepp som beskriver det digitala objektet, den person som publicerat objektet, tidsperioden som den är insamlad, ämnesområdet som avses osv.

Kriterium

Data beskrivs med utförliga maskinläsbara metadata på ett sätt som möjliggör att sökningar kan processas maskinellt.

Vägledning för implementering

Det digitala objektet är beskrivet med rikt metadata i ett maskinläsbart format.

Vägledande frågor

- Finns metadata som beskriver det digitala objektet?
- Innehåller den framtagna metadatat även beskrivningar som exempelvis innehåll, den aktör som publicerat det digitala objektet, tidsperiod för insamling, ämnesområde, och så vidare?
- Struktureras metadata på ett sätt som gör det tekniskt effektivt att söka och hitta resursen?

F3. Metadata clearly and explicitly include the identifier of the data it describes

Metadata som är till för att göra data sökbara finns ofta separerade från data. För att säkerställa att metadata länkar till de data som de beskriver ska metadata inkludera datasetets/innehållets uppslagningsbara globalt unika beständiga identifierare (PID). Via uppslagningstjänsten kan då data nås med hjälp av metadata i de fall användaren har rätt behörighet.

Kriterium

Metadata innehåller identifieraren till det data som beskrivs.

Vägledning för implementering

Metadata som beskriver det digitala objektet, eller samlingen av objekt, innehåller en uppslagningsbar globalt unik beständig identifierare.

Vägledande frågor

- Innehåller metadata en identifierare till det digitala objektets data?
- Är det en beständig identifierare, dvs. finns det någon organisation som garanterar dess beständighet?
- Är denna identifierare uppslagningsbar så att det går att nå det data den beskriver?

F4. (Meta)data are registered or indexed in a searchable resource

De rika metadata beskriver digitala objekt samt de identifierare som kan användas för att slå upp vart data finns tillgängligt eller vilka regler och process som gäller för att få åtkomst till data. De rika metadata behöver användas för att synliggöra digitala objekt ([se webbplats go-fair.org](http://se.webbplats.go-fair.org)) samt samlingar av objekt och göra att de går att hitta på internet. Indexering i generella sökmotorer är en del av detta, tillgängliggörande via repositorier/metadatakataloger som använder de rika metadata som beskriver digitala objekt för att göra forskningsdata sökbara är en annan viktig del.

Kriterium

Metadata, och om tillämpligt data, går att hitta genom en sökbar webbtjänst.

Vägledning för implementering

Data och/eller metadata som beskriver digitala objekt är möjliga att hitta i minst en sökbar webbtjänst via minst en uppslagingsbar globalt unik beständig identifierare.

Vägledande frågor

- Är det möjligt att hitta det digitala objektet, eller samlingen av objekt, via internet, exempelvis genom en domänspecifik, nationell eller internationell webbportal, katalog eller repositorium?
- Är det möjligt att hitta det digitala objektet, eller samlingen av objekt, på en nationell domänövergripande webbportal, katalog eller repositorium för olika typer av forskningsdata?
- Går det att hitta denna både genom att söka på identifieraren och på metadata som beskriver digitala objekt, eller samlingen av digitala objekt?
- Länkar en eventuell generell webbportal, katalog eller repositorium vidare till metadata som beskriver innehållet mer i detalj?
- Går det att hitta en beskrivning av metadata för ett digitalt objekt genom att söka i en generell sökmotor?

A1. (Meta)data are retrievable by their identifier using a standardized communications protocol

Metadata och data kommuniceras mellan sändare och mottagare via regelverk som styr hur utbytet ska gå till, vad olika saker betyder, hur fel hanteras mm. Ofta används generella kommunikationsprotokoll som grund varefter lager av mer specialiserade protokoll läggs till för att möta olika behov.

Då tillgängligheten gynnas av att standardiserade och välkända kommunikationsprotokoll används eftersträvas detta ur FAIR-perspektiv.

Kriterium

Metadata, och om tillämpligt data, kan nås via sina identifierare via ett standardiserat kommunikationsprotokoll.

Vägledning för implementering

Data och/eller metadata går att hämta via ett standardiserat kommunikationsprotokoll.

Vägledande frågor

- Finns det möjlighet att hämta data och/eller metadata, med information om hur man får åtkomst till data, via ett standardiserat kommunikationsprotokoll.

A1.1. The protocol is open, free, and universally implementable

Även om kommunikationsprotokollet är standardiserat enligt principen ovan kan tillgänglighet till den digitala resursen påverkas negativt i de fall protokollet inte är öppet och fritt tillgängligt, och/eller om användning eller implementation är förknippat med kostnader. Se webbplats [Github.com Protokoll öppet tillgängligt](#)

Kriterium

Digitala objekt kan nås och läsas av på ett öppet, kostnadsfritt och generellt implementerbart sätt.

Vägledning för implementering

Kommunikationsprotokollet är öppet tillgängligt utan kostnad och/eller möjligt att implementera utan kostnad.

Vägledande frågor

- Är beskrivning av protokollet, som är nödvändig vid eventuell implementering, åtkomligt via en länk på internet?
Se webbplats [Github.com Protokoll öppet tillgängligt](#)
- Är användning av protokollet eller egen implementering av detsamma kostnadsfri?

A1.2. The protocol allows for an authentication and authorization procedure, where necessary

För att kunna möta detta behov behöver använt protokoll stödja verifiering av att en användare är den som den utger sig för att vara, dvs. autentisering, samt kunna avgränsa tillgång till det forskningsdata som användaren har rättighet att få åtkomst till, dvs. accesskontroll.

Kriterium

Det är möjligt att skapa olika användarroller och mekanismer för verifiering av användare och accesskontroll för tillgång till digitala objekt, när så är nödvändigt.

Vägledning för implementering

Vid tillgängliggörandet av ett digitalt objekt i en begränsad omfattning behöver en policy och ett protokoll för auktorisering och åtkomstkontroll finnas.

Vägledande frågor

- Om det digitala objektet, eller samlingen av objekt inte är öppet tillgängligt, finns det en beskrivning av de regelverk, processer och policyer som gäller för att få tillgång till det skyddade materialet?
- Finns det ett sätt att säkerställa att de personer som ska få tillgång till resursen är de som de utger sig för att vara?
- Finns det ett sätt att säkerställa att enbart de som ska få tillgång till det digitala objektet får det?

A2. Metadata are accessible, even when the data are no longer available

Att logiskt separera beskrivande metadata som är till för att data ska gå att hitta från de data de beskriver förbättrar sökbarhet och förståelse av innehållet utan att exponera data som inte ska eller kan tillgängliggöras öppet.

Underhåll och lagring av metadata är generellt mindre resurskrävande än lagring av data samt mindre känsligt att spara över tid. Detta möjliggör fortsatt tillgång till beskrivningar för att förstå innehållet i ett digitalt objekt, eller en samling av objekt, och återanvända denna information även om de data som beskrivs inte finns kvar.

Kriterium

Metadata finns tillgängliga även när data inte längre finns tillgängliga.

Vägledning för implementering

Lagring och underhåll av det digitala objektets metadata sker logiskt separerat från det data de beskriver, på ett sådant sätt så att metadata fortsätter att vara tillgängliga även om data inte längre finns att tillgå.

Vägledande frågor

- Lagras och underhålls metadata logiskt separerade från data?
- Lagras metadata i ett maskinläsbart format?
- Är metadata som beskriver den digitala resursen och dess innehåll möjliga att hitta i en sökbar webbtjänst även om data inte längre är tillgängligt?
- Finns det en dokumenterad plan för förvaltning och tillgängliggörande av dessa metadata över tid?

Exempel:

Se webbplatser:

- registerforskning.se Register Utiliser Tool (RUT),
- snd.gu.se Forskningskatalog Svensk Nationell Datatjänst (SND),
- Etsin.fairdata.fi Fair Datamängder

I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation

För att kunna tolka beskrivningar av digitala objektets innehåll på ett effektivt sätt behöver betydelsen av dessa uttryckas på ett sätt som gör att de som ska använda informationen kan tolka beskrivningarna och den kontext som de finns i. Det kan innebära både maskinläsbar representation (t ex RDF, SKOS) och semantiska resurser (som terminologier/ontologier/kontrollerade vokabulärer).

För att programvara ska kunna tolka betydelse på ett effektivt sätt behöver både begrepp och relationer mellan dem kunna uttryckas formaliserat. På så sätt kan semantiska strukturen visualiseras som en graf för att det ska vara enklare att förstå relationer.

Kriterium

Metadata och data redovisas enligt semantiska beskrivningar som är standardiserade/vedertagna, dokumenterade och tillgängliggjorda.

Vägledning för implementering

Metadata6 som beskriver betydelsen av det digitala objektet, uttrycks med en formell, vedertagen samt maskinläsbar representation av begrepp och termer samt semantiska resurser tydliggör deras inbördes relationer.

Vägledande frågor

- Möjliggörs referenser/mappning från använda termer och begrepp till standardiserade terminologier/ontologier/kontrollerade vokabulär?
- Finns det en beskrivning av hur använda begrepp relaterar till varandra i sammanhanget, antingen i en graf eller hur olika terminologier mappas till varandra?
- Används en vedertagen representation av de termer, begrepp och relationer som uttrycker betydelse, exempelvis genom ett språk som kan avbildas som en graf?

Exempel:

Se webbplatser

- [w3.org Ramverk för resursbeskrivning \(RDF\)](http://www.w3.org/Ramverk_för_resursbeskrivning_(RDF)) ,
- [w3.org Webb ontologispråk \(OWL\)](http://www.w3.org/Webb_ontologispråk_(OWL))

I2. (Meta)data use vocabularies that follow FAIR principles

De kontrollerade vokabulärer, terminologier och/eller ontologier som används för att uttrycka betydelsen med metadata behöver själva uppfylla FAIR-principerna. Genom att de tilldelas en unik och uppslagingsbar beständig identifierare och beskrivs med metadata och proveniens går de att hitta, förstås och återanvändas, och kan användas för mappning och harmonisering för förbättrad interoperabilitet.

Kriterium

Vokabulärer, terminologier och/eller ontologier som används är vedertagna och kontrollerade och beskrivningar av dessa finns tillgängliga.

Vägledning för implementering

De digitala objektens data och/eller metadata använder kontrollerade vokabulärer, terminologier och/eller ontologier som uppfyller FAIR-principerna.

Vägledande frågor

- Finns använda vokabulärer, terminologier, ontologier väl dokumenterade med rikt metadata på en sökbar webbplats och i maskinläsbart format?
- Har dessa en tilldelad uppslagingsbar beständig identifierare?
- Används ett vedertaget sätt att beskriva de termer, begrepp och relationer som uttrycker betydelse?
- Finns det dokumenterat vilka aktörer som deltagit i arbetet med att ta fram använda vokabulär, terminologier, ontologier och hur processen såg ut?
- Finns det tydliga licensregler för hur de får användas?

Exempel:

Se webbplatser

- [browser.ihtsdotools.org SnomedCT](http://browser.ihtsdotools.org/SnomedCT),
- [mesh.kib.ki.se MeSH](http://mesh.kib.ki.se/MeSH)

I3. (Meta)data include qualified references to other (meta)data

För att få bättre förståelse för innehållet i en samling digitala objekt, och/eller dess relation till andra digitala objekt/samlingar av digitala objekt, beskrivs inte bara enskilda delar av innehållet för sig utan också i det sammanhang, den kontext, de verkar.

Sammanhanget beskrivs genom att uttrycka olika typer av relationer. Det kan vara relationer mellan olika delar av aktuellt metadata, mellan objekttyper, olika begrepp, olika terminologier eller ontologier, men det kan också vara relationer till andra typer av digitala objekt eller samlingar av objekt.

Relationerna behöver därför tillföra semantik för att utöka förståelsen för det sammanhang som beskrivs, de behöver uttrycka innebörden av relationen. En HTML-länk mellan exempelvis två populationsbegrepp uttrycker inte innebörden av relationen mellan dem, en relation i en graf som uttrycker att den ena populationen är en "del av" den andra tillför däremot innebörd.

Kriterium

Relationer mellan olika data och/eller metadata beskrivs på ett sätt som möjliggör att sammanhang kan förstås.

Vägledning för implementering

Inbördes relationer mellan de olika typerna av metadata som används för att beskriva de digitala objekten och deras innehåll beskrivs. Inbördes relationer mellan olika delar av de digitala objekts data samt relationer till eventuella externa data beskrivs.

Vägledande frågor

- Uttrycks eventuella inbördes relationer mellan olika typer av metadata inom de digitala objekten?
- Uttrycks relationer som finns mellan av objekten använda termer och begrepp för att beskriva sammanhanget inom vilka de är framtagna?
- Uttrycks eventuella relationer som finns mellan inom objektet inkluderat data och externa digitala objekt eller samlingar av objekt?
- Uttrycks innebörden av relationerna i ett maskinläsbart format?

Exempel:

RDF + Vokabulär, t.ex. SKOS se webbplats:

[wikipedia.org Simple Knowledge Organization System \(SKOS\)](http://wikipedia.org/Simple%20Knowledge%20Organization%20System%20(SKOS))

R1. (Meta)data are richly described with a plurality of accurate and relevant attributes

För att kunna utvärdera om, och på vilket sätt, de digitala objekten är användbara behöver det finnas metadata som beskriver objektens innehåll och betydelse samt den kontext inom vilken data är skapat/insamlat.

Genom att beskriva innehållet med så utförliga metadata som möjligt ges så goda förutsättningar som möjligt för att innehållet ska vara sökbart, kunna bevaras över tid, användas och återanvändas och att egenskaper som påverkar återanvändbarheten upptäcks och kan hanteras så tidigt som möjligt i processen.

Kriterium

Digitala objekt innehåller olika typer av kontextuella beskrivningar som möjliggör förståelse och att avgöra om data passar för syftet med återanvändningen.

Vägledning för implementering

De digitala objekten, deras proveniens samt dess innehåll beskrivs med rikt metadata.

Vägledande frågor

- Är samlingen av digitala objekt beskrivna med metadata och semantik, t.ex. identifierare, skapare, namn, datum, ämnesområde, version osv.?
- Är de digitala objekten beskrivna med metadata och semantik, som är relevant inom domänen. Exempel på detta kan vara variabler, versioner och datum, begrepp, begreppssystem, kodlistor, kategorier osv.?
- Är alla värden (t ex även valfria) som används i metadata beskrivna?
- Är kontexten beskriven på ett sätt som möter kriteriet under I3?

R1.1. (Meta)data are released with clear and accessible data usage license

Om det inte är tydligt på vilket sätt digitala objekt får användas och under vilka förutsättningar, kan detta påverka om resursen kommer att kunna återanvändas då det blir svårt att bedöma vilka regler som ska följas. Det behöver därför framgå vilka licenser eller regler som gäller för alla delar i det digitala objektet, både avseende data och metadata. Se webbplats

[Github.com FAIRMetrics/Metrics](https://github.com/FAIRMetrics/Metrics)

Kriterium

Villkor för hur metadata och/eller data kan användas är angivna.

Vägledning för implementering

De digitala objekten inkluderar tydliga licenser för användning och återanvändning. Detta gäller för både data och metadata.

Vägledande frågor

- Finns tydliga licenser/regler för vad som gäller vid användning och återanvändning av metadata och data?
- Finns referens till beskrivningar av dessa licenser?
- Är licenserna maskinläsbara?

Kommentar: Om licenserna/rättighetsmärkningen tillåter öppen tillgång främjas den digitala resursens återanvändbarhet.

Exempel:

Se webbplats

creativecommons.se [Creative Commons](https://creativecommons.org) , [opensource.org](https://opensource.org/licenses) [MIT licenser](https://opensource.org/licenses)

R.1.2. (Meta)data are associated with detailed provenance

För att kunna bedöma om, och på vilket sätt, de digitala objekten går att återanvända behöver det finnas metadata som beskriver objektets proveniens. Metadata som beskriver processen kring hur de digitala objekten tagits fram, i vilket syfte, vilka aktiviteter som genomförts, av vem och med hjälp av vilka resurser (källor, metoder osv.) bör tillhandahållas i maskinläsbar form. Se webbplats go-fair.org [Proveniens](#)

Kriterium

Datas och metadata proveniens beskrivs i detalj.

Vägledning för implementering

De digitala objektens proveniens finns beskriven, helst i maskinläsbar form och innehåller information om entiteter, aktörer som bidragit till processen för att skapa forskningsdata, aktiviteter som genomförts samt i vilken ordning och när respektive aktivitet är utförd.

Vägledande frågor

- Finns det information om personen, organisationen eller infrastrukturen som har skapat och datum (t ex för skapande, tillgängliggörande, m m) angivna i metadata?
- Finns övriga aktörer som bidragit via publikationer, metoder, aktiviteter, finansiering etc. beskrivna i metadata?
- Finns syftet med skapande av de digitala objekten och forskningsfrågor som använts under processen beskrivna?
- Finns använda metoder, källor, publikationer, källkod, variabler, ontologier, terminologier beskrivna i metadata?
- Finns processen beskriven med vilka aktiviteter som ägt rum när samt vilka aktörer som genomfört dem och vilka entiteter som använts, skapats förändrats?

R1.3. (Meta)data meet domain-relevant community standards

Genom att använda ett inom domänen vedertagen metadatabeskrivning eller metadatastandard tillsammans med vanligt förekommande terminologier, ontologier, vokabulär och värdemängder för att beskriva de digitala objekten och deras innehåll så förbättras återanvändbarheten.

Användning av en vedertagen metadatabeskrivning eller standard ger ett gemensamt språk. Detta gör att missförstånd undviks när en beskrivning av innehållet kommuniceras, dels då användaren av de digitala objekten vet vilket innehåll denne kan förvänta sig och dels på grund av förbättrad interoperabilitet mellan olika digitala objekt. Det kan ge en ökad grad av harmonisering då samma metadataelement används.

Kriterium

Metadata och data är strukturerade och dokumenterade enligt för ändamålet tillämpliga standarder och vedertagna format.

Vägledning för implementering

Data och/eller metadata är organiserade och dokumenterade enligt inom domänen vanligt förekommande standarder.

Vägledande frågor

- Används en gemensam inom området vanligt förekommande metadatabeskrivning eller metadatastandard?
- Används standardiserade eller inom domänen vanligt förekommande värdemängder, t.ex. klassifikationer, kodlistor etc.

Begreppslista

Efter varje begrepp finns en länk som hänvisar till annan webbplats med fördjupad information.

Accesskontroll/Åtkomstkontroll: Metod för att ge användare tillgång till de digitala objekt de är auktoriserade till.

[ICANN.org Accesskontroll](#)

Aktiviteter: Händelser som sker över tid och förändrar och/eller använder entiteter. Beräkning, transformation, skapande av entiteter, kommunikation.

[W3.org PROV](#)

[W3.org PROV - Ontologi Aktiviteter](#)

Aktörer: En person, organisation eller mjukvara som ansvarar för en aktivitet eller resurs. Exempel: Forskare, data steward, infrastruktur, lärosäte, program.

[W3.org PROV](#)

[W3.org PROV - Ontologi Aktörer](#)

Auktorisering: Specificerar vad en användare, eller grupp av användare, har rätt att få tillgång till.

[ICANN.org Auktorisering](#)

Autentisering: Verifiering av en användares identitet. Exempel SWAMID.

[ICANN.org Autentisering](#)

[sunet.se SWAMID](#)

Beständig Identifierare: Identifieraren är beständigt länkad till det metadata som definierar det objekt den identifierar. Beständigheten garanteras av tjänsteleverantören som tilldelat densamma.

[Cessda.eu Beständig Identifierare](#)

Digitalt objekt: Ett digitalt objekt representeras av en binär sekvens, refereras till av en beständig identifierare och har egenskaper beskrivna av metadata. Begreppet används här även för att referera till samlingar av digitala objekt. Exempel: Digital museisamling, Digital bild, Dataset, Datamängd, Klassifikation, Ontologi.

[RD-alliance Digitalt objekt](#)

Domän: Verksamhets och aktivitetsområde som forskningsfält eller organisatoriskt område.

Domänspecifikt metadata: Metadata som beskriver de digitala objekten inom en domän på ett inom fältet vedertaget sätt. Exempel: SDTM, GSIM, DDI, FHIR, DICOM.

Cdisc.se SDTM

statswiki.unece.org GSIM

ddialliance.org DDI

hl7.org FHIR

ddicomstandard.org DICOM

Domänövergripande metadata: Metadata där ett gemensamt urval beskrivande element som är oberoende av domän används för beskrivning av det digitala objektet. Exempel: DataCite, DCAT-AP.

zenodo.org Domänövergripande metadata

schema.datacite.org Exempel DataCite

Entiteter: Entiteter använda av aktiviteter under processen. Exempel: Metoder, Källkod, Standarder.

W3.org PROV

W3.org PROV - Ontologi Entiteter

Generell sökmotor: Sökmotor som indexerar innehåll på webben. Exempel: Google, Bing.

Globalt unik Identifierare: Globalt unik identifierare avser globalt unik referenskod. Referenskoden kan enbart referera till ett digitalt objekt. Exempel: URI, DOI.

go-fair.org Globalt unik identifierare

Kommunikationsprotokoll: Formellt regelverk som specificerar hur utbyte av data mellan två system ska gå till och vilket format som används. Exempel http, TCP/IP, FTP.

ietf.org http

cisco.com TCP/IP

tools.ietf.org FTP

Kommunikationsprotokollet öppet tillgängligt: Kommunikationsprotokollet och dess beskrivning distribueras fritt över internet utan kostnad.

Kontext: Sammanhang inom vilket metadata och data till det digitala objektet skapas, beskrivs, samlas in och används.

Lagras och underhålls logiskt separerat: Metadata lagras inte som en del av det data det beskriver utan exempelvis i en separat fil eller i en metadata katalog/repositorium. Exempel RUT, SND, Etsin.

registerforskning.se RUT

snd.gu.se SND

etsin.fairdata.fi Etsin

Licens: "...Ordet används för att beskriva ett särskilt tillstånd för licensinnehavaren att göra sådant som normalt är förbjudet. Genom ett licensavtal kan du ge utvalda personer eller företag rätt att utnyttja din uppfinning, design eller ditt varumärke mot betalning. Du kan upplåta rätt att utnyttja din rättighet helt eller delvis. Licensen kan gälla ett visst land eller bara delar av ett visst geografiskt område." Exempel: Creative Commons.

prv.se Citat licens

creativecommons.se Creative Commons

opensource.org MIT

Maskinläsbar form: "Med maskinläsbart format avses ett filformat som är strukturerat på ett sådant sätt att det enkelt kan läsas av ett datorprogram." Exempel: QR-kod, XML-fil.

[Innovation genom information SOU 2020:55 QR-kod XML-fil](#)

Metadata som beskriver proveniens: Metadata som beskriver datats proveniens, dvs de entiteter, aktiviteter och personer involverade i att producera data eller objekt och kan användas för att utvärdera dess kvalitet och trovärdighet. Exempel: PROV.

w3.org Metadata som beskriver proveniens

Metadata: Data om data. Dessa kan innehålla beskrivande, kontextuell, och proveniens-relaterad information om det digitala objektet och dess egenskaper. Exempel: Titel, namn på forskare, ämnesområde, datatyp.

rd-alliance.org Metadata

Metadata beskrivning/Metadata standard: Metadata standard på konceptuell nivå, inkluderar semantik och struktur som används för att beskriva hur metadata

ska beskrivas i en metadatastandard. Inkluderar ej vilka metadataelement som beskrivs utan fokuserar på hur det beskrivs. Exempel ISO11179, GSIM.

iso.org [ISO11179](#)

statswiki.unece.org [GSIM](#)

Metadataelement: De delar med information som beskriver egenskaper inom en metadatastandard. Har ibland hierarkiska relationer, dvs. ett metadataelement består av andra metadataelement.

Metadatastandard: ”Metadatastandarden anger vilka element som ingår i den specifika standarden och en förklaring till vilken information som ska skivas under respektive element.” Exempel: HL7, FHIR, DDI, SDTM.

snd.gu.se [Metadatastandard](#)

Ontologi: Formaliserad modell för att uttrycka kunskap inom en specifik domän. En ontologi beskriver begrepp/klasser som finns, relationerna mellan dem samt hur de kan användas tillsammans. Exempel: MEDDRA.

w3.org [Ontologi](#)

bioportal.bioontology.org [MEDDRA](#)

Protokoll: Regelverk och process som gäller för tillgängliggörande.

Representation: Information nödvändig för att förstå och återge digitalt material och metadata.

opendatahandbook.org [Representation](#)

Rikt metadata: Det ska vara tillräckligt med metadata så att den som söker efter data kan finna, bedöma huruvida data kan användas, samt förstå vad det handlar om utan att ha tillgång till själva data. Detta kan exempelvis innebära att det finns grundelement i metadata som tydliggör datans benämning, vem (t ex person, organisation), som har skapat metadata, vem som har publicerat metadata/data och när, en sammanfattning och begrepp som underlättar sökning.

zenodo.org [Rikt metadata](#)

Semantiska resurser: Semantik beskriver betydelse. Semantisk beskrivning görs exempelvis genom att terminologier och ontologier definierar begrepp och begreppssystem för att tydliggöra betydelse för mjukvara och/eller människa. Exempel: SnomedCT, Mesh, LOINC.

snomed.org [SnomedCT](#)

mesh.kib.ki.se [Mesh](#)

loinc.org [LOINC](#)

Standardiserat: Av standardiserings-organisation fastställt eller inom domänen överenskommen praxis.

Terminologi: En uppsättning termer som används inom ett fackområde och de begrepp som definierar dem. Exempel: SnomedCT

[browser.ihtsdo.org SnomedCT](http://browser.ihtsdo.org/SnomedCT)

Tilldelad: Utgiven av tjänsteleverantör vilken administrerar globalt unika beständiga identifierare och garanterar tillgänglighet och integritet.

Tillgängliggöra: ”Ett forskningsmaterial som görs tillgängligt blir sökbart och synligt både nationellt och internationellt samt möjligt för andra att återanvända och citera.”

[snd.gu.se Citat Tillgängliggöra](http://snd.gu.se/Citat_Tillgängliggöra)

Uppslagningsbar identifierare: Identifieraren länkar till den resurs den identifierar.

Vokabulär: En specificerad lista med utvalda termer som tilldelas en beständig betydelse. Exempel: Mesh.

[mesh.kib.ki.se Vokabulär](http://mesh.kib.ki.se/Vokabulär)

[mesh.kib.ki.se Mesh](http://mesh.kib.ki.se/Mesh)

Värdeområde: Tillåtna värden för att beskriva egenskaper för en variabel. Värdena kan vara i form av en lista eller beskrivna i ett uttryck.

[statswiki.unece.org Värdeområde](http://statswiki.unece.org/Värdeområde)

Webbportal, metadata katalog eller repositorium: Sökbar webbaserad (meta)data katalog/repositorium eller webbportal. Exempel: European Genome-phenome Archive Zenodo.

[snd.gu.se Sökbar webbaserad \(meta\)data katalog](http://snd.gu.se/Sökbar_webbaserad_(meta)data_katalog)

[ebi.ac.uk European Genome-phenome Archive](http://ebi.ac.uk/European_Genome-phenome_Archive)

[zenodo.org Zenodo](http://zenodo.org/Zenodo)

Vetenskapsrådet
Västra Järnvägsgatan 3
Box 1035, 101 38 Stockholm, Sweden
Tel +46 (0)8-546 44 000
vetenskapsradet@vr.se
Vetenskapsrådet.se