

Svar på DIGG:s remiss “Principer för tillgängliggörande av information för vidareutnyttjande” (ärendenummer 2019-139)

Om Wikimedia Sverige

Wikimedia Sverige (WMSE) är en ideell förening som arbetar för att alla människor ska ha fri tillgång till mänsklighetens samlade kunskap.

Vi är en lokal, juridiskt obunden stödförening till [Wikimedia Foundation](#), som driver några av världens största fria kunskapskällor, inklusive [Wikipedia](#), [Wikidata](#) och [Wikimedia Commons](#). Målet med vårt arbete är att verka för att sprida upphovsrättsligt fri kunskap till så många människor som möjligt. [Våra medlemmar](#) styr vårt arbete.

Vi arbetar framför allt tillsammans med andra, till exempel myndigheter, kulturarvsinstitutioner, andra ideella föreningar och volontärer. Vår organisation förändrar samhället och att experimentera är centralt för oss då det inte finns en färdig mall för vad vi gör. Vi utvecklar bland annat en [talsyntes för Wikipedia](#) så att det ska gå att lyssna på artiklarna. Vidare bygger vi även olika verktyg och gör förbättringar för att volontärer och partnerorganisationer enklare ska kunna bidra med sina data inom t.ex. kampanjer som [Wiki Loves Earth](#) och projekt som [FindingGLAMs](#).

Övergripande synpunkter

- De föreslagna principerna är i huvudsak mycket bra och WMSE bedömer att de, om de tillämpas av svensk offentlig sektor, skulle resultera i **mycket stort värde** för Wikimediarelsen i dess arbete att förbättra Wikipedia, Wikidata och andra tjänster.
- I den inledande texten anser vi att en del *bör*-formuleringar istället borde vara *ska*-formuleringar. Detta gäller i synnerhet formuleringen “Verksamheter bör sträva efter öppenhet som standard inom de juridiska gränserna för att öka möjligheten till användning och värdet av informationen.” som vi anser bör lyda “Verksamheter **ska** sträva efter öppenhet som standard inom de juridiska gränserna för att öka möjligheten till användning och värdet av informationen.”

Utan en sådan förändring kommer öppenhet i praktiken att vara ett undantag som måste motiveras internt i organisationen – till skillnad från “stängdhet” som sällan behöver motiveras då den uppfattas som “naturlig”. Det omvända, att öppenhet är regeln och inte undantaget, är det som bör eftersträvas ([jämför t.ex. med Nederländerna](#)).

- För att principerna ska leverera värde måste de tillämpas i praktiken. Vi vill därför *starkt rekommendera* att principerna kompletteras med råd, vägledningar och fallstudier om hur de kan och bör tillämpas i praktiken. Principerna riskerar annars att vara alltför abstrakta för att vara omsättningsbara i praktiken, i synnerhet för de offentliga verksamheter som är i början av sin digitala verksamhetsutveckling.
- DIGG bör en tid efter principernas publicering regelbundet utvärdera om och i vilken grad de följs i praktiken. Utan sådan uppföljning finns det risk för att principerna blir icke-förpliktiga principer som alla bekänner sig till, men få efterlever.
- DIGG bör verka för att regeringen gör det tydligt för alla svenska offentligt finansierade organisationer att principerna ska efterföljas.
- Vi anser att även källkod bör betraktas som en form av information och att det bör tydliggöras i principerna. Principerna bör då rekommendera källkod som produceras eller upphandlas av en offentligt finansierad organisation, som regel ska vara fri att vidareutnyttja. Källkod bör licensieras i enlighet med [The Open Source Definition](#).
- Principerna skulle kunna göras tydligare och mer konkreta genom att de struktureras upp och tillförs underrubriker. Detta skulle t.ex. kunna göra formuleringarna mer konsekventa än de är för närvarande, där en del av principerna innehåller motiveringar till principen (t.ex. Princip 9) och en del inte.

- Ett exempel på principer som formulerats strukturerat är de som [Digisam](#) (PDF) formulerat för sina medlemmars tillgängliggörande av information.
- För standarder generellt, bör [W3C](#)-standarder och andra öppna standarder framhållas som lämpliga att efterfölja vid tillämpning. Dessa är kostnadsfria, icke-proprietära, och webb- och internetvänliga. The Open Source Initiative har [goda principer](#) för vilka standarder som bör anammas vid utveckling av öppen källkod.

Synpunkter per princip

Princip 1: Använd öppna standardiserade licenser

“För användare är det viktigt att veta villkoren för vidareutnyttjande av information. Det är speciellt viktigt i de fall användare vill bygga en verksamhet på information. Information som erbjuds ska därför ha tydliga villkor för hur den får användas. Detta sker med fördel genom licensiering med öppna och standardiserade licenser. Villkoren ska öppna för så många användningsområden som möjligt och ska omfatta både information och tekniker för tillgängliggörande, till exempel API:er (applikationsprogrammeringsgränssnitt). Om information som görs tillgänglig skyddas av upphovsrätt som innehas av verksamheten bör verksamheten avstå från sina egna exklusiva rättigheter om möjligt.”

Våra synpunkter:

- Principen kan med fördel döpas om till “Använd öppna standardiserade licenser och rättighetsmarkeringar”. Detta då t.ex. CC0 och Public Domain Mark ej är licenser.
- Blanda inte juridisk och teknisk interoperabilitet och öppenhet i samma princip. Dessa är oavhängiga varandra. Bryt därför ut rekommendationen om tekniskt tillgängliggörande (via API, LOD, datadumpar, etc.) till en egen princip.
 - Alternativt lämna alla tekniska detaljer till mer detaljerade vägledningar om teknisk tillämpning av dessa principer (se också övergripande synpunkter ovan om behovet av sådana).
- Principen bör även rekommendera att licenser och rättighetsmarkeringar ska vara maskinläsbara. Det är också i linje med Princip 7.
- [Creative Commons](#) bör, i principerna eller tillhörande fördjupningsmaterial, pekas ut som den nuvarande de facto-standarderna för maskinläsbara licenser och rättighetsmarkeringar.
- För data och innehåll bör öppenhet definieras i enlighet med [Open Definition-kriterierna](#).
 - För digitala mångfaldiganden av objekt med verkshöjd, men där upphovsrätten i objektet är utgången, bör Public Domain Mark användas för att tydliggöra att den digitala representationen inte har någon egen verkshöjd (detta är också i linje med Artikel 14 i Digital Single Market-direktivet).
 - För databaser och metadata som innehåller eller representerar faktauppgifter, bevarandemetadata, samt teknisk metadata, rekommenderar

vi att principerna tydliggör att dessa inte kan anses hålla verkshöjd och att en CC0-markering är bästa sättet att tydliggöra det på. Allt fler dataplattformar har också detta som krav för att data ska kunna laddas upp dem - ett exempel är Wikidata.

- Tills dess att Artikel 14 går i kraft bör fotografiska bilder, verk och andra verk producerade eller upphandlade inom ramen för offentlig verksamhet licensieras CC BY eller CC BY-SA.
- Källkod, som också är att betrakta som en form av information, bör licensieras i enlighet med [The Open Source Definition](#) och med val av licens som inte sätter begränsningar på hur återanvändaren använder källkoden. Exempel på sådana licenser är BSD-licenser och derivat av dessa, t.ex. MIT-licensen.

Princip 2: Erbjud information avgiftsfritt

“Avgifter för information är ett stort hinder för användning, framförallt för små och medelstora företag och privatpersoner. Huvudregeln är därför att information ska vara avgiftsfri då det ökar den potentiella användningen.

Det finns rättigheter och skyldigheter i avgiftsförordningen och annan författning som ger anledning för vissa verksamheter att ta betalt för information. Avgifter får bara överskrida marginalkostnaden om verksamheten har en skyldighet att ta ut avgifter. Om avgifter tas ut ska beräkningsgrunden publiceras elektroniskt.”

Våra synpunkter:

- Avgifter är inte bara ett ofta för högt hinder för företag och privatpersoner utan likaledes för många andra organisationsformer och verksamheter såsom ideella initiativ och föreningar. Studenter bör kanske även speciellt omnämnas då de i många fall behöver data för t.ex. sina uppsatsarbeten, men inte kan inte i alla fall få den betald av sitt universitet.

Princip 3: Erbjud information utan krav på registrering

“Att registrera sig för att få åtkomst till information är ett hinder för användning. Information ska huvudsakligen göras tillgänglig utan att användaren behöver ansöka om tillstånd eller registrera sig. Samtidigt bör användare ges möjlighet att tillhandahålla kontaktinformation så att utgivaren kan meddela användare, om till exempel ändringar, driftstörningar eller avbrott.

I de fall det finns användarvillkor som begränsar användningen kan registrering vara nödvändig. Då bör exempeldata för utvärdering och övrig kringinformation som dokumentation och annat stöd vara tillgängligt utan registrering eller avgifter.”

Våra synpunkter:

- Utgångspunkten, regeln, bör vara att användningsvillkor inte ska kräva registrering. Undantag till detta bör motiveras. Detta för att etablera normen att maximal öppenhet är regeln, inte undantaget.

Princip 4: Beskriv och synliggör information

“För att användare ska kunna hitta och använda information måste de veta vilken information verksamheten har. Verksamheten ska därför upprätthålla och publicera en översikt över all information som vanligtvis tillhandahålls elektroniskt på begäran på Sveriges dataportal.

Även information som vanligtvis inte tillhandahålls för vidareutnyttjande kan synliggöras på Sveriges dataportal. På det viset underlättar verksamheten för användarna när de söker information oavsett om de kan ta del av den eller inte.”

Våra synpunkter:

- Principen bör uppdateras med en rekommendation att information med fördel bör tillgängliggöras på flera webbtjänster än den egna och Sveriges dataportal. Detta för att nå (fler) användare där de redan är, snarare än att helt förlita sig på den svåra utmaningen att locka användare till sin egen webbplats.
 - Exakt var informationen bör publiceras beror på dess natur och tilltänkt målgrupp.
- Vid tillgängliggörande av information på flera webbplatser blir Princip 6 extra viktig så att det är tydligt för användarna vad informationens aktualitet är. Vidare bör informationsutgivaren sträva efter att hålla information de publicerat på externa plattformar uppdaterad gentemot originaldata.

Princip 5: Tillgängliggör information efter användarnas behov

“Vid val av vilken information som ska tillgängliggöras för vidareutnyttjande ska användarnas behov vara en viktig parameter. Verksamheten bör sträva efter att anpassa innehåll, format och stödmaterial utifrån användarnas behov för att öka möjligheterna till att informationen används.”

Våra synpunkter:

- Det är otydligt från hur principen är formulerad om den handlar om informationens slutanvändarbehov eller indirekta användares behov (~ User Experience), återanvändarnas behov (~ Developer Experience), eller bådadera. Vi har i detta svar antagit att det handlar om återanvändarnas behov.
- Rekommendationen att utgå från (åter)användarbehov vid utformningen av tjänster för öppna data bör förstärkas. Kanske med en formulering om att organisationer bör bedriva "proaktiv, regelbunden och systematisk kartläggning av återanvändarnas behov" eller dylikt. En god användarupplevelse för utvecklare är ofta avgörande för om en informationsmängd återanvänds eller inte.
- Om det är slutanvändarbehov som avses bör principen innehålla en formuleringen om att sådana i princip är omöjliga att förutse på ett heltäckande vis, utan enbart användningsfall till användningsfall.

Princip 6: Erbjud aktuell, komplett och uppdaterad information

“Information som tillgängliggörs ska vara aktuell, komplett och omfatta så mycket som möjligt till exempel när det gäller detaljeringsgrad och omfång. Verksamheten bör sträva efter att tillgängliggöra information med en servicenivå anpassad efter användarnas behov.”

Våra synpunkter:

- Att data skall hålla god kvalitet är rätt. Däremot visar erfarenhet att många organisationer undviker att publicera sina data för att den inte är bra nog. Data är aldrig helt perfekt så det kan vara bra att uppmuntra en förståelse av att det är bättre att publicera dataset som är adekvata men inte felfria än att inte publicera data alls (vilket är resultatet när man väntar tills den är perfekt). Att publicera data kan för övrigt vara ett väldigt gott första steg till att få hjälp med att förbättra den.
- Det vore bra om principen lägger extra emfas på god versionshantering av publicerad data.
- Principens andra mening är egentligen överflödigt givet Princip 5. Generellt kanske det bör tydliggöras att principerna även “gäller för varandra”.

Princip 7: Använd maskinläsbara, öppna och standardiserade format, stabila adresser och unika identifierare

“Information bör vara tillgänglig i öppna, standardiserade och maskinläsbara format. Det ökar möjligheterna att kombinera information (interoperabilitet) och lägger inte onödiga begränsningar på vad informationen kan användas till. Exempel på maskinläsbara och standardiserade format är CSV, XML, JSON och RDF.

Information bör ha en unik och permanent adress på internet så att det är möjligt att länka ihop den med annan information. Information får mervärde genom att hänvisa till andras information om till exempel samma sak, samma plats eller samma händelse.”

Våra synpunkter:

- Principen kan och bör vara tydligare med att rekommendera användning av kontrollerade vokabulärer/termlistor/auktoriteter för t.ex. platser, ämnen, datum, språk, namn (både på personer och organisationer), etc. Sådana vokabulärer bör i sig också efterfölja princip 7, d.v.s. vara maskinläsbara öppna, standardiserade och sammanlänkade.
- För konsekvensens skull bör specifika filformat och datamodeller inte omnämnas alls i principerna, utan istället i den tillhörande vägledning för tillämpning vi rekommenderar.
 - Alternativt bör liknande mer specifika rekommendationer ingå i fler principer. T.ex. i Princip 1 där när API nämns exemplifieras dessa inte med t.ex. REST, GraphQL eller SPARQL.
- Teknisk not: RDF är ett ramverk enligt vilket data kan serialiseras i t.ex. både JSON och XML. Det är alltså lite “äpplen och päron” som blandas i exemplen just nu.

Princip 8: Erbjud ändamålsenlig dokumentation

“Verksamheten bör dokumentera informationsmängder så att de är lätta att använda för både människor och maskiner. Med dokumentation menar vi beskrivningar som gör att andra kan upptäcka, förstå och använda informationen.

Dokumentationen kan innehålla kodexempel, beskrivning av hur informationen skapats, vad den används till idag, exempel på användning och vad den inte kan användas till. Kvaliteten på uppgifterna bör också dokumenteras och kända felaktigheter och utmaningar bör uttryckligen nämnas i beskrivningarna.”

Våra synpunkter:

- Det bör tydliggöras att även dokumentationen bör vara öppet licensierad (d.v.s. också självt följa de föreslagna principerna).
- Överväg att rekommendera att det bör finnas mekanismer för användare att ge feedback och förbättringsförslag på rekommendationen (se också Princip 9).

Princip 9: Uppmuntra till användning och återkoppling

“Att göra information öppet tillgänglig räcker inte för att information ska kunna användas. Verksamheten bör interagera med användare och aktivt uppmuntra till användning. Att kunna ställa frågor om informationen och få hjälp i tolkning och användning är viktigt för många användare och användarstöd bör erbjudas från verksamheten. Att föra dialog och skapa en gemenskap kring dataområden är ofta bra och bör övervägas.

Genom att erbjuda användare möjlighet att ge feedback kommer verksamheten att få bättre förutsättningar att förstå användarnas behov. Det kommer att skapa möjligheter till att förbättra kvaliteten på publicerad information och skapa förtroende mellan utgivaren och användarna.”

Våra synpunkter:

- Rekommendationen att föra dialog med användarna bör förstärkas. Kombinationen av “bör” och “övervägas” är väldigt mjuk. Eftersom att en övervägning kan ge ett negativt svar kan formuleringen vara “ska övervägas”.
- Vi anser att när informationsutgivare öppnar upp för återkoppling bör man göra det på ett öppet, transparent och spårbar sätt, gärna genom ett helpdesksystem. Det bygger förtroende och minskar även risken att man får samma frågor och återkopplingar om och om igen.