

Problemen med att kräva attribuering vid användning av data

Att välja licensiering på sin data kan upplevas som knepigt. Många aktörer funderar initialt på möjligheten att använda sig av krav på attribuering även för data, vilket ofta är det krav de ställer vid användning av texter eller mediafiler. Tanken är ofta att det ska leda till ökad synlighet för den egna organisationen samt en tro att krav på källangivelse inte skulle ha någon direkt skadliga påverkan. Ibland övervägs även att göra undantag för specifika utvalda aktörer. För att göra data användbart vid vidareanvändning är det dock viktigt att använda så fria licenser som möjligt.

E-delegationen rekommenderar att myndigheter ”som huvudregel gör information tillgänglig på nätet och tillåter vidareutnyttjande fritt eller enligt standardiserade och generösa villkor. Därigenom bidrar myndigheten till att förverkliga de politiska målen om ökat vidareutnyttjande av offentlig information. Samtidigt blir det lättare för myndigheten att följa PSI-lagens krav.”¹

Ett krav på attribuering gör det dock svårt att inkludera dataseten i många produkter. I olika tjänster kombineras ofta stora mängder av data och beräknas på olika sätt vilket gör att enormt långa listor skulle krävas med källor för varje enskild datapunkt.

Attribueringskrav gör det även svårt att veta när man gjort tillräckligt för att uppnå licensen. All typ av osäkerhet gör det svårare att arbeta med materialet och gör att många kommer att välja bort dataset med otydliga eller licenser som kräver attribuering eller andra saker. Att göra fel innebär att du kanske i ett senare läge måste bygga om hela din tjänst. Denna osäkerhet drabbar framförallt mindre aktörer.

För dataägaren innebär en minskad användning paradoxalt ofta **mindre synlighet** och att värdet av det egna arbetet därigenom minskar då informationen skulle ha inkluderats i fler produkter.

Ett exempel: Din tjänst visar befolkningens mängd på en karta med en s.k. heatmap där färger representerar befolkningstätheten. Du kan zooma upp och ner i kartan. Tjänsten börjar med att visa kartan utzoomad och visar hela världen (med data från

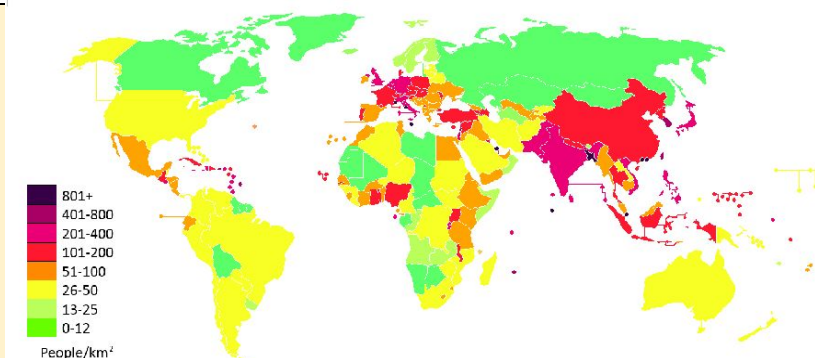


Illustration: <https://commons.wikimedia.org/w/index.php?curid=670091> Licens: CC BY-SA 3.0

¹ E-delegationen. 2. *Sammanfattning och checklista för att komma igång.*

<http://feedback.edelegationen.se/vpsi/vidareutnyttjande-av-offentlig-information/2-sammanfattning-och-checklista-fo%CC%88r-att-komma-iga%CC%8Ang/>

hundratals olika datainsamlingar och kompletterande geodata från tusentals andra källor). Hur ska då attribuering ske? Ska alla som kräver attribuering vara tydligt angivna redan på världskartan? Eller ska källan endast visas när man zoomar ned några nivåer i kartan? När ska attribueringen då dyka upp? Hur ska attribueringen se ut? När anser den som skapat datasetet att attribueringen är fullgod? Vad händer när olika dataägare ställer olika krav?

Alla dessa frågor måste tjänstens skapare brottas med när tjänsten skall realiserars. Att ställa krav på attribuering låter enklare än det faktiskt är att implementera det.

Wikimedia-rörelsen lever som vi lär och all data på Wikidata är licensierad med CC0. Ett dataset som har en licens som har attribuering som krav (och därmed sätter begränsningar på användandet) kan därför **inte** inkluderas på Wikidata, vilket är Wikimedia Foundations projekt för strukturerad data. Wikidata kommer att bli en central nod och förse alla de 293 språkversionerna av Wikipedia, och dess 500 miljoner månatliga läsare, med uppdaterad och korrekt data.

Vad ska man då välja?

CC0 är en licens som är specifikt framtagen för data i databaser. Den rekommenderas bland annat av Europeiska kommissionen.²



Målet med licensen är att underlätta

återanvändningen av intressant data så mycket som möjligt och därigenom stimulera framtagandet av nya produkter, underlätta för små och medelstora företag m.m.

CC0 är en licens för de organisationer som vill att den information de samlar in skall användas av så många som möjligt och som vill hålla allmänheten informerad med ny och uppdaterad information.

Licensen innebär att det inte finns något krav på att beskriva varifrån informationen kommer,³ vilket kan ses som något negativt för dataägaren. Frågan som dataägaren måste ställa sig vad som är viktigast: användningen och samhällsnyttan kontra organisationens synlighet? Är det viktigt att skattefinansierat material praktiskt och enkelt kan användas av olika aktörer i samhället?

Viktigt att påpeka är även att användningen av CC0 inte förhindrar användarna från att ändå attribuera. I många fall kommer attribuering att göras trots att licensen CC0 har valts. Detta då

² “[T]he CC0 public domain dedication is of particular interest. As a legal tool that allows waiving copyright and database rights on PSI, it ensures full flexibility for re-users and reduces the complications associated with handling numerous licences, with possibly conflicting provisions.”, citat från Timothy Vollmer, Creative Commons. *European Commission endorses CC licenses as best practice for public sector content and data*. 17 juli 2014.

<https://blog.creativecommons.org/2014/07/17/european-commission-endorses-cc-licenses-as-best-practice-for-public-sector-content-and-data/>

³ Läs mer här: <https://creativecommons.org/publicdomain/zero/1.0/deed.sv>

många användare ser ett värde på att trycka på att materialet kommer från en trovärdig källa (ex. på Wikipedia kommer källor att anges även när det inte är ett krav från dataägaren).

<https://www.wikidata.org/wiki/Q34>

folkmängd	9 747 355	redigera
bestämningsmetod	Folkbokföring i Sverige	
tidpunkt	31 december 2014	
▼ 1 referens		
referens-url	http://scb.se/sv/_Hitta-statistik/Statistik-efter-amne/Befolkning/Befolkningens-sammansattning/Befolkningsstatistik/25788/25795/Kvartals--och-halvarsstatistik---Kommun-lan-och-riket/244147/	
utgivare	Statistiska centralbyrån	
titel	Folkmängd i riket, län och kommuner 30 juni 2014 och befolkningsförändringar 1 januari - 30 juni 2014 (svenska)	
		+ lägg till referens

Så här kan de se ut i Wikidata när en källa läggs till för ett påstående. Målet är att alla påståenden skall ha en källa, även om all data där är under CC0.

Andra myndigheter som använder sig av CC0

CC0 används av allt fler myndigheter i Sverige och i andra länder. En lista med exempel på organisationer från olika delar världen som använder sig av CC0 finns här:

https://wiki.creativecommons.org/wiki/CC0_use_for_data

Några svenska exempel är:

- **Styrelsen för internationellt utvecklingssamarbete (SIDA)** – För SIDA stod valet endast mellan att licensiera sin data med Public Domain Mark eller CC0. SIDA valde i slutändan att licensiera all data på Openaid.se med Public Domain Mark (PDM) snarare än CC0 "eftersom vi som myndighet inte har någon upphovsrätt till vår öppna data".⁴
- **Kungliga biblioteket (KB)** – All data i Libris använder sig av CC0 och KB har även ingått avtal med andra parter om att CC0 skall användas.⁵ KB arbetar aktivt inom området.
- **Riksantikvarieämbetet (RAÄ)** – All metadata i K-samsök är licensierat med CC0.⁶ Det gäller inte bara data från RAÄ utan alla andra myndigheter och organisationer som vill dela sin metadata där måste enligt avtalet licensiera den under CC0.⁷ RAÄ

⁴ Joachim Beijmo, Sida. *Licenser för data på Openaid.se*. 19 mars 2015.

<http://openaid.se/sv/blog/licensing-for-the-openaid-se-data/>

⁵ Maria Kadesjö, KB. *KB tecknar avtal med OCLC om CC0-licens*. 13 februari 2014.

<http://librisbloggen.kb.se/2014/02/13/kb-tecknar-avtal-med-oclc-om-cc0-licens/>

⁶ Johan Carlström, RAÄ. *K-samsök*. 21 september 2015. <http://www.raa.se/hitta-information/k-samsok/>

⁷ Johan Carlström, RAÄ. *Samverkansavtal/överenskommelse för leveranstill K-samsök*. 20 mars 2013. <http://www.ksamsok.se/wp-content/uploads/2013/04/K-sams%C3%B6k-avtal.pdf> sid. 4.

argumenterar att det är viktigt att markera upp data då “[ä]ven om metadata generellt inte är skyddat av upphovsrätt eftersom det saknar verkshöjd kan databasen i sig vara skyddad; CC0-märkning är därför ett sätt att tydliggöra att det i grunden fria materialet faktiskt är fritt att vidareanvända.”⁸

- **Livrstkammaren, Skoklosters slott och Hallwylska museet (LSH)** – Är ett av flera centralmuseer⁹ i Sverige som har licensierat sina samlingsdatabaser med CC0.¹⁰
- **Örebro kommun** – Har bl.a. valt att för all kartdata och geodata från kommunen använda sig av CC0.¹¹

⁸ Riksantikvarieämbetet. *Riktlinjer för digitalisering av kulturarvsmaterial på Riksantikvarieämbetet*. September 2013. http://www.raa.se/app/uploads/2013/10/digitaliseringsriktlinjer_riksantikvarieambetet.pdf, sid. 17.

⁹ Haxpett *et al.*, Wikipedia. *Centralmuseum*. 8 januari 2016.

<https://sv.wikipedia.org/w/index.php?title=Centralmuseum&oldid=32360788>

¹⁰ LSH.se. *Öppet bildarkiv*. http://lsh.se/sites/lsh.se/files/pdf/oppet_bildarkiv_0.pdf

¹¹ Stefan Peterson, Örebro kommun. *Kartdata och geodata*. 10 november 2014. www.orebro.se/36470.html