

Strategisk inkludering av biblioteksdata på Wikidata

Slutrapport

Introduktion

Denna rapport redogör för aktiviteterna genomförda inom ramarna för projektet Strategisk inkludering av biblioteksdata på Wikidata 2018. Projektet genomfördes av Wikimedia Sverige i samarbete med Kungliga biblioteket mellan den 1 maj 2018 och den 15 augusti 2019.

Rapporten innehåller en kortfattad sammanfattning av projektet och länkar till fördjupande material. För ökad transparens och för att erbjuda ett mervärde för den svenskspråkiga Wikimediagemenskapen har vi lagt upp materialet i en portal på föreningens wiki, där det kan läsas av alla. Detta material kan nås på adressen <http://bit.ly/biblioteksdata2018>.

I rapporten hänvisar vi till de fem effektmålen som sattes upp i projektansökningen, det vill säga:

1. Öka användningen av svensk biblioteksrelaterad data.
2. En grund finns för inkludering av verksdata inför att denna skapas i Libris XL.
3. Sporra bibliotekens tekniska personal som jobbar med LOD m.m. genom att ge konkreta exempel på det värde deras arbete skapar samt tillsammans identifiera framtida utvecklingsmöjligheter för fri kunskap.
4. Visa på bibliotekens värde för allmänheten genom att tydliggöra insatserna och det värdefulla material de förvaltar.
5. Öka kunskapen om hur Wikimediaprojekten kan användas av biblioteken för att prioritera sin digitalisering.

Förberedande arbeten

Innan vi satte igång med uppladdningarna, genomförde vi en undersökning av Wikidata som en plattform för länkade öppna bibliografiska data, samt hur den svenskspråkiga Wikipediagemenskapen i nuläget hanterar bibliografiska data. Syftet med denna aktivitet var att kartlägga vilka möjligheter vi hade och möjliga riktningar för fortsatt arbete.

Bakgrundsinformation för gemenskapen

Vi skrev en kortare introduktion till projektet på engelska för att kunna hänvisa den internationella gemenskapen i samband med t.ex. konferenser. Denna återfinns på adressen https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_p%C3%A5_Wikidata_2018/Background.

Uppfyller effektmål 1.

Biblioteksdata på Wikidata – en översikt

Vi skrev en kortare introduktion till den bibliografiska infrakonstruktursystemen på Wikidata, alltså det ekosystem i vilket vi skulle arbeta. Detta återfinns på adressen https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_p%C3%A5_Wikidata_2018/Biblioteksdata_p%C3%A5_Wikidata.

Uppfyller effektmål 1.

Hur man kan undersöka kvalitén på Wikidata med hjälp av Shape Expressions

Vi skrev en kortare introduktion till Shape Expressions vilket är ett sätt att undersöka RDF-strukturers beskaffenhet och har nyligen implementerats i Wikidata. Detta återfinns på adressen https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_p%C3%A5_Wikidata_2018/Rapportering/Kvalite

Uppfyller effektmål 1.

Källor på Wikipedia – vilka är mest populära?

Vi undersökte hur olika sorters källor används på svenskspråkiga Wikipedia. Målet var att identifiera de vanligaste källorna som underlag till en diskussion om hur samarbetet med Kungliga biblioteket kan hjälpa wikipedianerna i deras arbete. Rapporten och slutsatserna återfinns på adressen https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_p%C3%A5_Wikidata_2018/K%C3%A4llor_p%C3%A5_Wikipedia.

Uppfyller effektmål 1 och 5.

En inventering av Libris deldatabaser

Vi genomförde en inventering av de deldatabaser som ingår i Libris för att kartlägga vilka sorts bibliografiska data det finns. Inom det här projektet fokuserade vi på Svenska nationalbibliografen på grund av dess CC0-licens, vilken är kompatibel med Wikidata, men flera andra dataset identifierades som potentiellt intressanta ur Wikipediasynpunkt. Kartläggningen återfinns på adressen

https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_p%C3%A5_Wikidata_2018/Libris_deldatabaser.

Uppfyller effektmål 4.

Uppladdningar av data

För att öka användningen av svensk biblioteksrelaterad data importerade vi ett antal poster från Libris: ca. 500 bokposter från Svenska nationalbibliografin och ca. 60 000 auktoritetsposter. En detaljerad rapport från detta arbete återfinns under adressen

https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_p%C3%A5_Wikidata_2018/Rapportering/Arbete.

Vidare arbetade vi med bibliografiska data från [Projekt Runeberg](#), ett digitalt bibliotek med äldre nordiska verk. Webbplatsen är välkänd inom den svenska öppen kunskap-rörelsen, och flera av verken hör till de mest citerade på svenskspråkiga Wikipedia. Genom att ladda upp data om nästan 5 000 verk bidrog vi till projektets synlighet bland en målgrupp som aktivt intresserar sig för svenska kunskapsresurser. En mer detaljerad beskrivning återfinns under adressen

https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_p%C3%A5_Wikidata_2018/Rapportering/Projekt_Runeberg.

Uppfyller effektmål 1 och 2.

Djupdykning i biblioteksdata

För att skapa en mer nyanserad bild av biblioteksdata, nu och i framtiden, för både Wikimediagemenskapen och bibliotekens personal skrev vi en rad rapporter med specifika teman inom området.

Att uppdatera Librisdata på Wikidata

Vi undersökte hur Librisdata på Wikidata kan hållas uppdaterad, med syfte att säkerställa att informationen på Wikidata fortsätter att utvecklas, fel kan rättas till och nya poster kan läggas till automatiskt. Rapporten återfinns under adressen

https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_p%C3%A5_Wikidata_2018/Rapportering/Verktyg_f%C3%B6r_att_h%C3%A5lla_Libris-information_p%C3%A5_Wikidata_uppdaterad.

Uppfyller effektmål 3.

Inklusionskriterier på Wikidata

För att klarlägga vilka Librisposter som kan importeras till Wikidata sammanställde vi en rapport om Wikidatas inklusionskriterier, vilken återfinns under adressen

https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_p%C3%A5_Wikidata_2018/Rapportering/Inklusionsguide.

Uppfyller effektmål 3.

Wikimedias roll i bibliotekets digitaliseringsarbete

Vi undersökte hur Wikimediarelsen kan bistå biblioteket i dess digitaliseringsarbete, inte minst gällande vilka prioriteringar som bör göras för att nå ut med materialet. Rapporten återfinns under adressen

https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_på_Wikidata_2018/Rapportering/Prioritera_digitalisering.

Uppfyller effektmål 5.

Evenemang

Vi deltog i flera evenemang, där vi spred kunskap om biblioteksdata och om vårt samarbete med Kungliga biblioteket. Dessutom arrangerade vi ett hackathon med fokus på Libris och Wikidata för Kungliga bibliotekets personal.

Evenemang där vi deltog

Hack for Heritage, Visby (5–7 oktober 2018)

Evenemangets målgrupp var personer från olika kulturarvsinstitutioner intresserade av att aktivt arbeta med kulturarvsdata från olika källor. Vi höll en kort presentation (*Wikimedias data och data via Wikimedia*) av våra dataset för samtliga närvarande. Vi höll även en presentation/Q&A/workshop för att ge deltagarna ingångar till hur de kan konsumera Wikimedias olika datamängder (*Konsumera Wikimedias data*). Dessutom svarade vi på frågor och hjälpte på andra sätt till när olika deltagare var intresserade av hur de kunde använda Wikimedia-material för att sprida kunskap om kulturarvet.

Uppfyller effektmål 1 och 4.

WikiCite 2018, USA (27–29 november 2018)

Konferensens målgrupp var personer intresserade av och engagerade i utvecklingen av bibliografiska data på Wikidata. Vi höll en presentation (*Strategic inclusion of library data on Wikidata*) vilken handlade om vårt projekt, om dess plats i Wikimedia-landskapet samt om våra framtidsplaner. Dessutom inhämtade vi fördjupade kunskaper om Wikidata som en ständigt utvecklad plattform för bibliografiska data, samt om liknande projekt på andra håll i världen.

Uppfyller effektmål 1 och 4.

Libris användardag, Stockholm (5 december 2018)

Konferensens målgrupp var bibliotekspersonal samt andra intresserade av utvecklingen Libris och av bibliografiska data och kataloger i synnerhet. Vi höll ett föredrag på 15 minuter om vårt samarbete med Kungliga biblioteket, i vilken vi presenterade bakgrundsinformation om Wikidata som en fri kunskapsdatabas som lämpar sig bra för bibliografiska data.

Uppfyller effektmål 3.

Digikult 2019, Göteborg (25 april 2019)

Konferensen riktade sig till kulturarvsinstitutionspersonal med intresse för digitalt kulturarv, av vilka många kände till Wikidata sedan innan. Vi höll ett anförande på 20 minuter med tema öppna biblioteksdata på Wikidata. I presentationen berättade vi om vårt samarbete med Kungliga biblioteket samt i mer breda termer om varför Wikidata är en bra plattform för att jobba med kulturarvsdata. I synnerhet betonades värdet av att länkade data berikar varandra, samt av Wikidatas öppenhet och att alla kan redigera.

Uppfyller effektmål 3 och 4.

Libris utvecklingsråd, Stockholm (22 maj 2019)

Vi blev inbjudna att hålla en presentation på Libris utvecklingsråds möte. Syftet är dels att berätta om projektet och dels att placera bibliotekets länkade öppna data i ett större perspektiv, med exempel på vilka möjligheter de skapar utanför institutionen.

I samband med utvecklingsrådet blev vi även inbjudna till en intervju om den pågående utvecklingen av verksnivån i Libris XL, där vi kan dela av oss med våra erfarenheter av att arbeta med Librisdata.

Uppfyller effektmål 3

Genomförda evenemang

Biblioteksdatahackathonet, Stockholm (14 januari 2019)

Tio personal på Kungliga biblioteket deltog i hackathonet med fokus på Wikidata, Libris och länkade öppna data. En fördjupad rapport från evenemanget återfinns på adressen https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_på_Wikidata_2018/Hackathon/Resultat

Uppfyller effektmål 3 och 5.

Extern kommunikation

Vi kommunicerade om projektet till allmänheten genom två blogginlägg:

1. [Mer biblioteksdata på Wikidata](#)
2. [Biblioteksdatahackathonet – en dag av lärdomar och erfarenhetsutbyte](#)
3. [Kungliga Biblioteket och länkade öppna data – vart nu?](#)
4. [Om gamla böcker och nya plattformar](#)

För att kommunicera om projektet med den internationella gemenskapen samt bjuda in till kommentarer och återkoppling publicerade vi en rapport från vårt arbete på engelska: https://se.wikimedia.org/wiki/Projekt:Strategisk_inkludering_av_biblioteksdata_på_Wikidata_2018/Strategic_Inclusion_of_Library_Data_in_Wikidata. Vi såg det som särskilt värdefullt då

det finns ett aktivt initiativ för att förbättra hanteringen och täckningen av bibliografiska data på Wikidata ([WikiCite](#)), som i samband med vår närvaro på evenemang m.m. uttryckt stort intresse för vårt samarbete med Kungliga biblioteket.

Uppfyller effektmål 4.

Ändringar och avvikelser från projektplanen

Inledningsvis planerade vi för att genomföra två hackathons. I samförstånd med KB reviderades den siffran till ett hackathon.

Inledningsvis planerade vi för att undersöka möjligheterna att inkludera utvalda delar av andra KB-dataset på Wikidata, exempelvis Biblioteksdatatabasen eller andra Libris-deldatabaser förutom Nationalbibliografin. Biblioteksdatatabasen valde vi att istället fokusera på inom projektet FindingGLAMs, som är inriktat på att samla in data om världens kulturarvsinstitutioner och därmed mer lämpligt för denna sorts dataset. I skrivande stund genomför vi en crowdsourcingkampanj riktad till svensk bibliotekspersonal där vi uppmuntrar dem att berika den grundläggande informationen om svenska bibliotek på Wikidata (https://meta.wikimedia.org/wiki/FindingGLAMs_Campaign).

Andra Libris-deldatabaser visade sig vara okompatibla med Wikidata, trots den stora mängd intressant information de innehåller, eftersom de förvaltas av andra aktörer än KB och omfattas därmed inte av samma öppna licens som Svenska nationalbibliografin.

Inledningsvis planerade vi att importera data om såväl verk som deras utgåvor från Libris till Wikidata. Detta då denna struktur stämmer överens med visionen av hur bibliografiska data på Wikidata bör modelleras, och till fullo utnyttjar möjligheterna med länkade öppna data; den gör det möjligt att t.ex. identifiera olika utgåvor av samma verk, inklusive översättningar. Detta kunde dock ej genomföras på grund av att arbetet med verk på Libris inte var komplett under projektets gång.

Vi laddade upp en bråkdel av de flera hundra tusen bokposter som finns i Svensk nationalbibliografi. Man hade lätt kunnat tänka sig att ladda upp många tusen, eller till och med hela databasen. Vi ser på denna uppladdning som ett pilotprojekt, som gav oss tillfälle till att undersöka Librisdata, dess struktur och lämplighet för Wikidata. Även ur ett gemenskapsperspektiv var det lämpligt att begränsa mängden uppladdade data, då Wikidata är ett gemenskapsdrivet projekt och det pågår fortfarande diskussioner om vilken riktning det skulle ta när det gäller bibliografiska data: ska man ladda upp "allt som finns i hela världen" eller en begränsad, kurerad delmängd? Detta är något som diskuteras livligt inom WikiCite, där vi fick göra ett uppskattat bidrag med vårt projekt.

En annan begränsning vi fick göra var att endast lägga till auktoritetsidentifierare till befintliga poster. Vi skapade inga nya Wikidataobjekt för författare. Detta på grund av att det ofta är komplicerat att matcha mellan Libris och Wikidata med hundra procent säkerhet, och vi ville

undvika att duplicera existerande objekt i onödan. Vi blev ändå nöjda med resultatet, 60 000 tillagda identifierare.

Slutsatser och framtidsmöjligheter

Sammanfattningsvis bedömer vi projektet som lyckat. Vi fick göra en satsning på biblioteksdata på Wikidata vid en läglig tidpunkt, då länkade bibliografiska data är ett högaktuellt ämne både på bibliotek världen över och bland den internationella Wikimediagemenskapen. Att vi fick samarbeta med Kungliga biblioteket samtidigt som Libris XL var under pågående utveckling var väldigt tacksamt, då den involverade personalen var nyfikna på möjligheterna med länkade öppna data.

Samtidigt innebar våra tidsramar vissa begränsningar, då vårt projekt var till viss del beroende av utvecklingen av Libris XL. Till exempel kunde vi inte implementera den tvålagrade verk/utgåvestrukturen, som ännu inte är klar i Libris. Vi saknar även ett smidigt sätt att följa uppdateringarna i Libris, medelst RSS-flöden eller dylikt, vilka skulle kunna göra det möjligt att skapa verktyg för automatisk uppdatering av Wikidata. Vi fick begränsa oss till att beskriva hur ett sådant verktyg skulle kunna se ut.

Det finns fortfarande mycket att göra inom området biblioteksdata på Wikidata, varför vi har lämnat in en ansökan om medel till ett fortsättningsprojekt. I detta vill vi fokusera delvis på att inkludera ännu mer bibliografiska data, och delvis för att öka dess synlighet och användning, t.ex. genom att bygga verktyg för källhänvisningar på svenskspråkiga Wikipedia.