

Att digitalisera arkivmaterial och bibliotekssamlingar – en nödvändig förutsättning för framtida forskning och innovation

Digitaliseringen av det svenska samhället har fört med sig fundamentalt nya förutsättningar att beforska material. Med hjälp av program och verktyg som främjar datadriven analys kan digitala dokument idag bearbetas utifrån både innehåll och metadata. Nya typer av frågor kan ställas till materialen på nya sätt, och på kvantiteter av material som tidigare varit arbeteekonomiskt otänkbart. Den snabba utvecklingen av AI-tekniker för automatiserade analyser driver detta område framåt ytterligare.

Digitaliseringen innebär också att kulturarvssamlingar används av fler discipliner än de humanistiska. Forskarna som använder Riksarkivets och Kungliga bibliotekets material idag kommer från en allt bredare skara av discipliner, som till exempel ekonomi, medicin, demografi, lingvistik, statsvetenskap och sociologi.

Dessa forskare efterfrågar ofta stora mängder data. Omfattande, enhetliga serier av data analyseras på digital väg, snarare än den manuella analys av enskilda dokument som präglar traditionella arbetsmetoder inom samhällsvetenskap och humaniora.

För att Sverige ska kunna möta de nya krav som dagens forskare ställer måste vi öka takten på digitaliseringen av arkiv- och biblioteksbestånden. Insatserna behöver fokusera på breda satsningar av infrastrukturell karaktär. Detta för att skapa bredast möjligast nytta för framtida forskning och innovation, för utveckling av nya offentliga och kommersiella tjänster samt för demokratisk insyn och samhällelig informationsförsörjning.

Kulturarvsdata är samhällsviktig information

I Riksarkivets bestånd finns spåren av hur den svenska staten blev till från senmedeltiden till idag – statliga myndigheters arkiv, samlade, förtecknade och bevarade på ett ställe. Där finns, bland mycket annat, ett rikt material om byar, gårdar och skatter i Sverige från 1500-talet och framåt, detaljerad geografisk information från hela Sveriges tillblivelse, och enskilda arkiv från tusentals inflytelserika personer och organisationer. Från mitten av 1700-talet tills idag finns även information om alla människor som levat i Sverige – var de bott och vad de arbetat med, ägt, betalat i skatt, när de fötts och dött, vem de gift sig med och vilka barn de fått. Sådan information har av staten vid flera tillfällen definierats som *grunddata*.¹ År 2018 gav den svenska regeringen Myndigheten för digital förvaltning (DIGG) ett uppdrag att standardisera infrastrukturen för informationsutbyte i den offentliga sektorn vad gäller

¹ Se t.ex. *Grunddata – i samhällets tjänst*, SOU 1997:146

grunddata i Sverige, som tillhandahålls av andra myndigheter för samhällsservice.² De *historiska grunddata* som Riksarkivet förvaltar ingick dock inte i uppdraget. Den potential för forskning, innovation och kunskapsproduktion som detta material skulle kunna utgöra om det blev digitalt tillgängligt för analys, och därtill kodat på samma sätt som grunddata hos andra myndigheter, går dock knappast att överskatta. Liknande potential finns även i andra materialgrupper.

Kungliga bibliotekets samlingsbyggande baseras på pliktlagstiftningen. Genom sammangåendet med Statens Ljud och Bildarkiv 2009 utvidgades uppdraget till att utöver den svenska tryckproduktionen också omfatta den audiovisuella publiceringen. 2015 vidgades uppdraget ytterligare genom att e-pliktlagstiftningen trädde i full kraft. Detta inflöde kompletteras genom insamling och donationer av både historiskt och samtida material (t.ex. handskrifter och författararkiv). Som ett resultat av detta förvaltar Kungliga biblioteket idag omfattande och växande samlingar av fysiska och digitala medier. Ett uttalat mål är att öka tillgängligheten och utnyttjandegraden hos dessa samlingar. Enskilda objekt ur samlingarna är värdefulla i sig, men det digitala formatet möjliggör ett fokus på beroenden och samband mellan olika element såväl på objekt- som samlingsnivå – kontextualiserade och föränderliga över tid får dessa samlingar på så sätt ännu större värde som informationsbärare. Länkning till data från källor utanför biblioteket stärker ytterligare dessa samlingars relevans för forskning.

Tillsammans utgör Riksarkivets och Kungliga bibliotekets bestånd en ovärderlig källa till det svenska samhällets tillblivelse och politiska, sociala, ideologiska och kulturella utveckling. Som del av vårt gemensamma kulturarv har denna information ett oersättligt värde i sig. Men mycket av denna samlade data utgör även *samhällsviktig information*. Som sådan är den särdeles viktig att tillgängliggöra för forskning, innovation och demokratisk insyn.

Riksarkivets och Kungliga bibliotekets digitalt tillgängliga bestånd anno 2019

Sverige har under de senaste decennierna utvecklats från ett industrisamhälle till ett tjänstesamhälle. I ett sådant har kunskapsindustrin en central roll för ekonomi och tillväxt.³ En förutsättning för att informationen ska kunna användas för forskning, innovation och tillväxt är dock att den finns digitalt tillgänglig. Det har även betydande effekter för demokratins välmående. Därför har Riksarkivet och Kungliga biblioteket de senaste tjugo åren satsat mycket kraft på att digitalisera sina samlingar.

² I uppdraget ingick grunddata från Bolagsverket, Domstolsverket, E-hälsomyndigheten, Försäkringskassan, Lantmäteriet och Skatteverket. *Uppdrag om säker och effektiv grunddata*, DIGG Dnr: 2018-31.

³ Detta har även uppmärksammats i Kulturarvspropositionen 2016/17: 116, sid 53 där regeringen konstaterar att digitaliseringen av samhället bidragit till att kulturarvsinformation har börjat utnyttjas ”för ny ekonomisk verksamhet”, samt att kulturarvet i allt högre utsträckning ”får en dynamisk roll i en expanderande postindustriell ekonomi mer inriktad på tjänsteproduktion”.

Kungliga biblioteket har sedan 2014 löpande digitaliserat den pliktlevererade svenska dagspressen. Med hjälp av externa medel (forskningsfinansiärer och fonder) genomförs insatser för att säkra digital tillgång även till dagspressen före 2014. Ett sådant exempel är den pågående digitaliseringen av allt dagstidningsmaterial fram t.o.m. 1906 – ett projekt finansierat av den brittiska forskningsstiftelsen Arcadia. Kungliga biblioteket har idag drygt 23 miljoner tidningssidor tillgängliga via söktjänsten *Svenska dagstidningar*.⁴ Kungliga biblioteket genomför också en systematisk digitalisering av det audiovisuella material som levererats i analogt format. Kungliga bibliotekets audiovisuella samlingar är tillgängliga via *Svensk mediedatabas*.⁵

Av Riksarkivets totala bestånd om 80 hyllmil fysiskt material är idag runt cirka procent digitaliserat. Det motsvarar runt 200 miljoner digitalt avbildade arkivhandlingar. Digitaliseringsarbetet har finansierats av särskilda anslag (det s.k. FOSAM-projektet), arbetsmarknadspolitiska insatser, av aktörer inom släktforskningen och med medel från externa forskningsprojekt.

En stor del av digitaliseringen av Sveriges kulturarvsmaterial utförs idag vid *Enheten för digitisering* (DIT) inom Riksarkivet i Fränsta och Ramsele. DIT är norra Europas främsta digitaliseringscenter för storskalig digitalisering av kulturarvsmaterial. Här finns en samlad kompetens inom området med över tjugo års erfarenhet. I enlighet med Riksarkivets förordning tar DIT emot finansierade digitaliseringsuppdrag gällande dokument med text- och bildbaserat informationsinnehåll från ett flertal svenska offentliga institutioner, och fungerar som kompetenscentrum för digitaliseringsfrågor för hela det svenska kulturarvsområdet.⁶ Det finns planer på att på liknande sätt erbjuda Kungliga bibliotekets produktionskapacitet för digitalisering av audiovisuellt material som en resurs på nationell nivå.⁷

Vid Riksarkivet finns med andra ord tekniken, systemen och kunskapen för att genomföra storskalig digitalisering av sina egna och andra myndigheters data. Både Riksarkivet och Kungliga biblioteket medverkar också, genom deltagandet i Digisam, till utvecklingen av gemensamma tekniska lösningar och standarder kulturarvsmyndigheter emellan.

Vad både Riksarkivet och Kungliga biblioteket däremot saknar är finansiering för det löpande arbetet vad gäller digitaliseringen av de egna bestånden. De bestånd ur Riksarkivets och Kungliga bibliotekets samlingar som digitaliseras för forskningens behov idag finansieras huvudsakligen av externa medel. Framför allt handlar det om enskilda framgångsrika forskares projekt i vilket digitalisering av ett specifikt kulturarvsmaterial ingår. Större projekt av mer infrastrukturell karaktär, som fokuserar på omfattande enhetliga serier likt den Arcadia-finansierade tidningsdigitaliseringen, hör till undantagen. Digitaliseringstakten av bestånden hos Riksarkivet och Kungliga biblioteket är därför betydligt lägre än hos våra

⁴ <https://tidningar.kb.se/>

⁵ <https://smdb.kb.se/>

⁶ Förordning 2017: 1347, §§9a och 19a

⁷ Konstenius, G. (2018). *Digitalisering av audiovisuella samlingar – förslag på hur ökad samverkan kan effektivisera digitaliseringen av audiovisuellt material inom ABM-området*. Kungliga biblioteket & Svenska Filminstitutet. Dnr 1.1.5-2018-120

grannländer. I den reviderade norska statsbudgeten för 2019 har till exempel den norska regeringen nyligen skapat 70 nya tjänster med inriktning digitalisering av både tryckt och audiovisuellt material vid Nasjonalbibliotekets avdelning i Mo i Rana. Nasjonalbiblioteket har också tilldelats NOK 10 miljoner för att inleda ett arbete med att digitalisera audiovisuellt material för arkiv, bibliotek och museer. Det norska Arkivverket å sin sida har fått 140 miljoner kronor extra för att utveckla nya lösningar för lagring och publicering av historiska data. Detta sammantaget kommer med största sannolikhet att innebära ett stort uppsving för forskning och innovation i Norge.

Riksarkivets ska enligt sin förordning ”tillgängliggöra arkiven för forskning och annat nyttjande, bl.a. genom digitalisering, tjänster för elektronisk tillgång och information om arkiven (§8).”⁸ I Kungliga bibliotekets förordning står att myndigheten ska ”främja den svenska forskningens kvalitet genom att tillhandahålla en effektiv forskningsinfrastruktur” (§1) och ”...verka för utveckling och samordning av digitala tjänster och system inom biblioteksväsendet” (§3).⁹ Anno 2019 betyder detta att omfattande satsningar behövs på digital infrastruktur och digitalt tillgängliggörande vid båda institutionerna.

Urvalen – vem bör bestämma?

Skillnaden som uppstår vad gäller faktorer som *tillgänglighet*, *användbarhet* och *interoperabilitet* mellan digitalt och analogt material är enorm. Det material som är digitalt tillgängligt på webben – och dessutom maskinläsbart – är tillgängligt för alla, alltid, och kan analyseras digitalt med en rad olika metoder, medan det som bara finns i analog form endast är tillgängligt på särskilda platser, och under särskilda tider. Därför väljer ofta forskare bort material som inte finns digitalt tillgängligt. Det får givetvis allvarliga konsekvenser för svensk kunskapsproduktion.

Skillnaden för dock även med sig utmaningar av existentiellt slag. Förutsatt att man inte kan digitalisera allt – enligt vilka kriterier bör man göra urvalet? Vems informationsbehov är viktigast att tillgodose?

I Forskningspropositionen 2016/17:50 konstaterades att kulturarvsinstitutionernas material utgör en outnyttjad forskningsresurs, eftersom merparten av detta material inte är digitaliserat.¹⁰ Därför gavs Vetenskapsrådet i sitt regleringsbrev för år 2017 i uppdrag att fördela särskilda medel för digitalisering av kulturarvssamlingar.¹¹ Uppdraget resulterade 2018 i projektbidraget *Digitalisering och tillgängliggörande av kulturarvssamlingar*, en satsning i samarbete mellan Vetenskapsrådet och Riksbankens Jubileumsfond. För projektperioden 2019-2023 avsattes drygt 110 miljoner kronor till forskningsprojekt som, förutom att besvara en traditionell forskningsfråga, även skulle bidra till digitalisering av

⁸ Förordning 2009: 1593, samt Förordning 2014:265 med instruktion för Riksarkivet

⁹ Förordning 2008: 1421 med instruktion för Kungliga biblioteket

¹⁰ Proposition 2016/17: 50

¹¹ <https://www.esv.se/statsliggaren/regleringsbrev/?RBID=18424>

kulturarvsmaterial. De beviljade projekten pågår nu som bäst, och spänner över skilda områden som till exempel samiska audiovisuella samlingar, gotländska runstenar, utominstitutionell scenkonst i Göteborg, medeltida helgonkulturer, svensk karibisk kolonialism på 1700- och 1800-talet, och svensk välfärdspolitik.¹² En ny projektperiod om 2021-2025 kommer att utlysas under våren 2020.

Satsningen på att öka digitaliseringstakten av det svenska kulturarvet är mycket välkommen. Förhoppningsvis kommer de material som nu digitaliseras som ett resultat av de beviljade forskningsprojekten att komma flera forskare till nytta. Faktum kvarstår dock, att när urvalet på detta sätt lämnas i händerna på enskilda forskare, blir nyttan av satsningen mindre än om man på ett mer övergripande sätt skulle finansiera storskalig digitalisering av källmaterial riktade till forskarsamhället i stort. Urvalet riskerar att bli för begränsat när det styrs av de forskningsfrågor som enskilda forskargrupper ställer upp. Den hantering som krävs inför och under digitaliseringen blir inte optimerad. Relaterade och liknande material som borde digitaliseras samtidigt förbises. Ansvaret för den långsiktiga förvaltningen och bevarande riskerar också att bli otydligt när digitaliseringen sker genom beställningar från enskilda forskare.

Man kan givetvis inte avkräva enskilda forskare, som har sina egna forskningsprojekt i fokus, ett helhetsansvar för samhällets informationsförsörjning och forskningsinfrastrukturer. Det ansvaret vilar ytterst på regering och riksdag. Huvudsyftet för framtida satsningar av detta slag bör därför vara digitaliseringen av kulturarvsmaterialen och dess nyttoskapande för bredast möjliga forskningsområden, snarare än den specifika forskningsfrågans besvarande. *Återanvändningsbarheten* – både med hänsyn till informationen och de digitaliserade materialens kvalitet – bör alltså få komma i första hand. Med bredd och helhetsperspektiv kan man skapa förutsättningar för betydligt omfångsrikare framtida tvärvetenskapliga forskningsfält – och för forskningsfrågor som kanske ännu inte hunnits föreställas – än om enbart enskilda forskares dagsaktuella forskningsintressen tillåts styra urvalet.

Behovet av att digitalisera kulturarvssamlingar för att möta forskningens och allmänhetens behov och förväntningar är alltså fortsättningsvis mycket stort. Kungl. Vitterhetsakademiens skrivelse "Om betydelsen av offentlig finansiering för digitaliseringen av Sveriges skriftliga kulturarv", [2019-04-15](#), avseende förslag till Nationell biblioteksstrategi visar både på allvaret i situationen och på det behov som finns.¹³ Riksarkivet och Kungliga biblioteket anser därför att staten bör investera i digitalt tillgängliggörande av kulturarvsdata på bred front för att främja högkvalitativ forskning, kunskapsproduktion, demokrati och innovation. De materialförvarande myndigheterna själva bör också ges inflytande över urvalet. Detta eftersom det inom dessa myndigheter finns överblick och helhetsansvar över bestånden. Därtill finns där unik kompetens kring samlingarna, och god förståelse för vilken nytta dessa samlingar skulle kunna skapa för forskning och innovation om de tillgängliggjordes för

¹² <https://www.vr.se/utlysningar-och-beslut/beslut-om-bidrag/beslut/2018-09-06-humaniora-och-samhallsvetenskap.html>

¹³ http://www.vitterhetsakad.se/ckeditor_assets/attachments/656/till_regeringen_om_betydelsen_av_offentlig_finansiering_for_digitaliseringen_av_sveriges_skriftliga_kulturarv.pdf

storskalig digital användning. Dessa myndigheter har också täta kontakter både med varandra och med forskarsamhället, och har därför goda insikter i vilka forskningsfält som är på kommande.

För att möta forskningens och den växande kunskapsindustrins behov föreslår därför Riksarkivet och Kungliga biblioteket att regeringen satsar särskilda medel för att digitalisera det svenska kulturarvet – på bred front, och med så omfattande forskningsintressen som möjligt i fokus. En sådan satsning skulle även gå i linje med OECD:s slutsatser i rapporten *Digital Government Review of Sweden*, i vilken OECD nyligen rekommenderade Sverige att med större kraft och tydligare politiskt ledarskap driva på för att tillgängliggöra myndigheters data i syfte att främja datadriven innovation.¹⁴

¹⁴ *Digital Government Review of Sweden. Towards a Data-driven public sector*. OECD 2018 (<http://www.oecd.org/gov/digital-government/digital-government-review-of-sweden-2018.pdf>)