



LUND UNIVERSITY

Digitalisering av kulturarvssamlingar

En inledning

Näslund, Anna; Snickars, Pelle

Published in:
Digitalisering av kulturarvssamlingar

2025

Document Version:
Förlagets slutgiltiga version

[Link to publication](#)

Citation for published version (APA):
Näslund, A., & Snickars, P. (2025). Digitalisering av kulturarvssamlingar: En inledning. I A. Näslund, & P. Snickars (Red.), *Digitalisering av kulturarvssamlingar: Om forskningsprogrammet Digarv* (s. 9-44). Makadam förlag.

Total number of authors:
2

Creative Commons License:
CC BY-NC-ND

General rights

Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

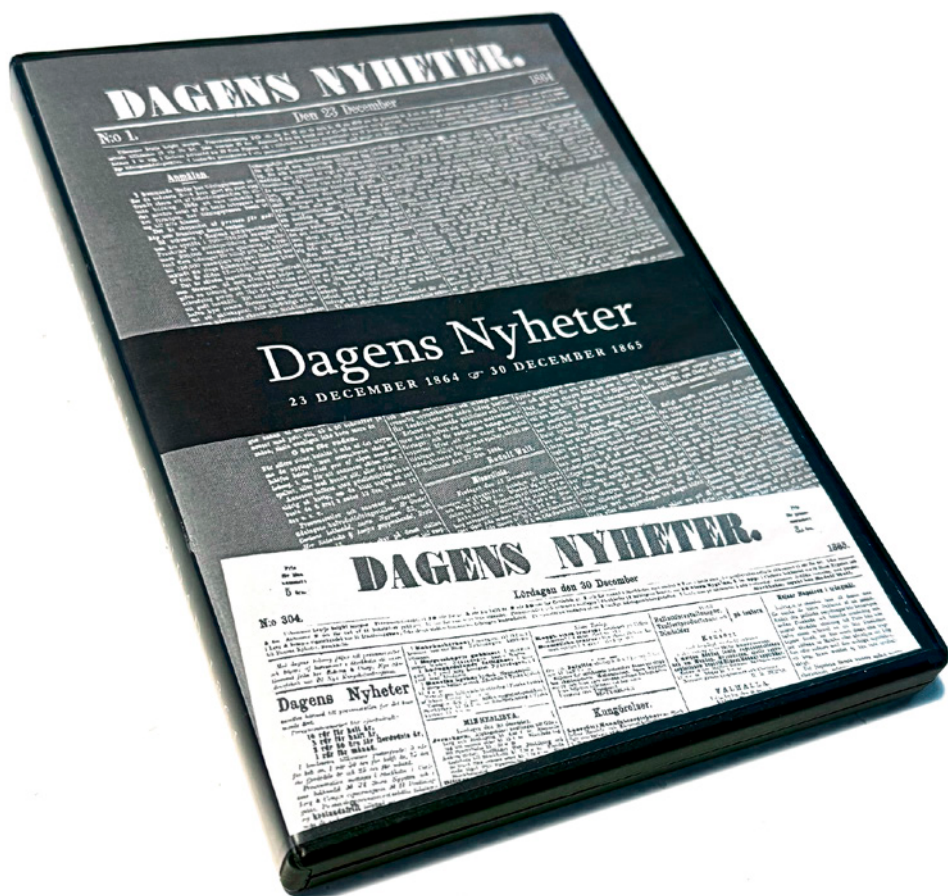
Digitalisering av kulturarvssamlingar

EN INLEDNING

Anna Näslund & Pelle Snickars

I slutet av 00-talet cirkulerade en CD-skiva som innehöll en digital version av samtliga sidor av *Dagens Nyheters* första årgång bland svenska humanistiska forskare. Upphovspersonen, Jonas Harvard, var själv humanist och verksam som forskare i historia vid Mittuniversitetet. Digitaliseringen var gjord med enkla medel. En digital kamera hade ställts upp på stativ framför en mikrofilmläsare. Fotoformatet och sidformatet överensstämde inte, så varje tidnings-sida krävde två exponeringar med kameran. Fotografierna lades ihop till en pdf och med hjälp av en betalversion av ett program för OCR-läsning (ABBYY FineReader) gjordes varje ord i tidningen sökbart. Resultatet blev en digital och fullständigt sökbar faksimil av *Dagens Nyheter* 1865.

Kopior av CD-skivan delades bland annat ut på en mediehistorisk workshop på Linköpings universitet år 2009, lägligt nog i samband med arbete på en forskningsantologi om *1800-talets mediesystem*.¹ Uppskattningen var närmast översvallande; Harvards initiativ och dess betydelse för historiskt orienterade forskare kan i efterhand inte överskattas. Framför allt innebar det en enorm tidsbesparing. Alla som har gått igenom dagspress från 1800-talet på mikrofilm vet hur tidskrävande det är. Typerna är tätt packade, rubrikerna små och illustrationerna få. Att läsa ord för ord i tidningen är möjligt i teorin – men knappast i praktiken. Att ens hitta artikelrubriker på det tema man intresserar sig för kräver enormt tålamod och mycket tid. Tidningars rubriksättning ger som bekant inte heller alltid information om ämnet. Det gäller särskilt noveller och kåserier men även vanliga nyhetsnotiser. Digitaliseringen av



Dagens Nyheter – om än bara under ett enda utgivningsår – gav därför helt nya möjligheter att hitta information. Att en enskild forskare stod bakom detta uppslag säger något om den svenska ABM-sektorns saktfärdighet när det gällde massdigitalisering.

Jonas Harvard fortsatte sitt digitaliseringsarbete de följande åren med svenska tidskrifter och riksdagsmaterial som bearbetades och gjordes sökbara med olika programvaror. Bland annat experimenterade han med mjukvara för så kallad *fuzzy search* och indexering. Det förra innebär att man kunde söka bredare genom att också få träffar på associerade ord, det senare att man kunde skapa ett register med sökord och därigenom få en överblick över vilka ämnen som präglade en publikation under ett visst år. Den här femton år gamla historien har flera poänger i relation till ämnet för den här boken. Den visar att humanistiska forskare relativt

[illegible]

tidigt insåg att digitalisering av historiskt material hade en avgörande betydelse, och agerade på denna insikt. Men än viktigare att poängtera är att detta hobbyprojekt drivet av en enskild forskare överträffar befintliga digitala infrastrukturer för forskning på svenskt dagpressmaterial när det gällde öppenhet och funktionalitet. Det var kostnadsfritt, vilket gjorde att även doktorander och andra forskare som inte hade stora forskningsanslag i ryggen kunde nyttja det. Dessutom kunde användarna fritt bearbeta informationen, till exempel använda valfri programvara för att få ut de data man önskade – vilket än idag inte alltid är möjligt när det exempelvis gäller Kungliga bibliotekets databas Svenska tidningar. I kvantitet kan Harvard's digitaliseringsprojekt givetvis inte mäta sig med dagens digitala resurser inom kulturarvsinstitutionerna, men i termer av forskaranvändbarhet står det sig förvånansvärt väl.

Forskningsprogrammet Digarv startade med en utlysning från Vetenskapsrådet 2018 och avslutades 2025. Vetenskapsrådet (VR) har varit huvudfinansiär, men Riksbankens Jubileumsfond (RJ) och

Kungl. Vitterhetsakademien har också delfinansierat programmets fjorton forskningsprojekt. Projektbidrag utlystes – i två omgångar 2018 och 2020 – för digitalisering och tillgängliggörande av kulturarvssamlingar; det har med andra ord handlat om projekt som (i huvudsak) både ägnat sig åt forskning och att digitalisera olika typer av samlingar. Några av forskningsprojekten har med humanistiska mått mått varit omfattande och inkluderat många medarbetare, en del med en budget på över tjugo miljoner kronor. Räknar man in alla som medverkat i Digarv-programmet rör det sig om fler än etthundrafemtio personer. Vi som är redaktörer för den här boken har både varit projektledare för varsitt forskningsprojekt och haft uppdrag som koordinators för forskningsprogrammet. I den sistnämnda rollen har vi fungerat som en sorts kontaktyta mellan forskningsprojekten, men också som dialogpartner för berörda kulturarvsinstitutioner och finansiärer.

Denna bok är tänkt att summera de sinsemellan skiftande forsknings- och digitaliseringsverksamheterna inom Digarv-programmet. Det kommer dock inte bara att handla om de forskningsresultat som uppnåtts. I lika hög grad står vunna erfarenheter när det gäller praktiska tillvägagångssätt och samarbeten mellan akademi och kulturarvssektor i fokus. Grundtanken har varit att i fjorton kapitel – ett för varje projekt – beskriva vad projekten ägnat sig åt, samt att resonera om de erfarenheter och utmaningar som de givit upphov till, såväl forsknings- som digitaliseringsmässigt. Antologin innehåller därmed en palett av *best practices*, men indikerar också blindskär att hålla utkik efter i tvärvetenskapliga digitaliseringsprojekt. Vår inledning kommer förstås i sedvanlig stil att (mot slutet) presentera vart och ett av bokens kapitel, dispositionsmässigt grupperade i fem avsnitt. Men vi vill också ta tillfället i akt och resonera om några av de forsknings- och digitaliseringskontexter som programmet tangerat under de år som det pågått. Ett sådant tematiskt fält rör AI, kulturarv och digitalisering; för utvecklingen från Harvards hobbyprojekt till dagens AI-modeller, ställvis baserade på äldre digitaliserade samlingar, har varit rasande snabb. Visserligen har några av projekten inom Digarv sysslat med artificiell intelligens. Parallellt har dock de senaste årens intensiva AI-utveckling satt tydliga spår inom arkiv, bibliotek och museer (ABM-sektorn) – rentav omdefinierat själva sektorn. Vid sidan av AI vill vi också uppmärksamma de forskningspolitiska sammanhang som Digarv-programmet varit en del av. I ljuset av kulturarvets digitali-

sering önskar vi bland annat lyfta frågan om vilken typ av forskningsinfrastruktur som landets kulturarvsinstitutioner egentligen utgör, samt svårigheten för humanistiska infrastrukturer att konkurrera om medel när utlysningar inte har en tydlig ABM-profiler. Eftersom vi redaktörer bägge har tidigare erfarenhet av att arbeta som forskare inom ABM-sektorn innehåller inledningen också ett avsnitt om Digarv-programmets bakgrund och upprinnelse, där vi framhåller vikten av att svenska arkiv, bibliotek och museer fortsatt anställer disputerade medarbetare. Dessutom vill vi förstås i denna inledning lyfta fram de många forskningsaktiviteter som vi i egenskap av koordinators arrangerat inom ramen för detta forskningsprogram.

Digarv och artificiell intelligens

Kontrasten mellan dammiga läderband och artificiell intelligens kunde inte vara större, lät *Svenska Dagbladet* Näringsliv meddela sina läsare på sin förstasida i maj 2024 efter ett besök på Kungliga bibliotekets nya datalabb. För att lära upp AI-modeller krävs enorma mängder data, påpekade tidningen. Det är något som nationalbiblioteket har i sina arkiv; KB står modell för ny AI, var den fyndiga rubriken i *SVD*. Eftersom nationalbiblioteket har så stora mängder digitaliserat svenskt tal ”kan vi leverera de bästa svenska språk- och tal-till-text-modellerna”, påpekade Leonora Vesterbacka, då anställd sedan något år tillbaka på KB-labb som *senior data scientist*. Inom kulturarvssektorn var det en titel som knappt någon tidigare hade stoltserat med. Men som disputerad partikelfysiker skolad på forskningscentret CERN i Schweiz hade Vesterbacka inte heller en bakgrund som de flesta andra bibliotekarier på KB.² Artikeln i *SVD* följdes upp nio månader senare med en snarlik text; nu hade KB nämligen tagit fram en ny tal-till-text-modell, KB-Whisper, en vidareutveckling av det amerikanska bolaget Open AI:s omskrivna Whisper-modell. ”Det är otroligt värdefullt att vår modell har tränats på en bred variation av svenska”, påpekade Vesterbacka. Dels hade modellen använt KB:s omfattande audiovisuella samlingar, dels samlingar vid Institutet för språk och folkminnen, en myndighet med ett stort ljudarkiv med olika svenska dialekter. Träningsdata för KB-Whisper uppgick till mer än femtiotusen timmar tal. Alla AI-modeller som KB-labb utvecklat var öppet tillgängliga, framhöll *SVD* – de var mycket populära och hade redan

då laddats ned över två miljoner gånger. Vesterbacka var också övertygad om att KB-Whisper skulle kunna vidareutvecklas. Ett område var sjukvården, påpekade hon, där det fanns behov av att transkribera läkares diktafoninspelningar och automatiskt göra om tal till anteckningar i patientjournaler. Men då behöver modellen först fintränas mer på medicinska termer som läkare använder, framhöll hon. ”Det finns startups som jobbar med sånt, som skulle kunna utgå från vår modell och träna den vidare, och då skulle även en skånsk läkare bli förstådd, jämfört med om man utgick från en amerikansk modell.”³

AI, startups, riskkapital, hälsovård och appar baserade på tal-till-text-modeller som automatiskt omvandlar läkares diktafonmeddelanden till patientanteckningar är inte något som vi vanligtvis associerar med kulturarvsinstitutioner, eller för den delen digitalisering av äldre samlingar.⁴ Men som inom många andra samhällsdomäner har maskininlärning och AI snabbt stöpt om ABM-sektorn. Vår poäng med att accentuera detta är att denna utveckling i närtid ägt rum exakt parallellt med Digarv-programmet. För vad vi numera menar med digitalt kulturarv – och framför allt hur det kan användas – är idag inte längre detsamma som för bara några år sedan. På kulturarvsinstitutioner arbetar förstås sedan länge data- och IT-kunnig personal, men datavetare har det funnits färre av. Förändringarna gäller också de AI-uppgifter som dessa medarbetare ägnar sig åt, där digitaliserade samlingar på kort tid fått helt nya funktioner, inte sällan med betydande samhällsnytta, vilka för bara några år sedan varit svåra att föreställa sig.

Som forskningsprogram hade Digarv sin bakgrund i (den socialdemokratiska) regeringens forskningsproposition från hösten 2016, *Kunskap i samverkan*, enligt vilken staten uttryckligen ville ”främja datadriven forskning, särskilt inom humaniora och samhällsvetenskap”.⁵ Propositionen innehöll faktiskt bara en enda rad om artificiell intelligens. I efterhand kan det te sig underligt. Men som de flesta av oss kommer ihåg har de allra senaste åren präglats av en rekordsnabb utveckling på det området, med introduktion och tillämpning av stora språkmodeller och generativ AI, en utveckling som satte högsta fart just under det sena 2010-talet: en *AI-boom*, som den kallats. Att den skulle påverka arbetet med att digitalisera kulturarv var det emellertid få som trodde när 2016 års forskningsproposition publicerades. Knappt någon av dem som arbetade inom eller samarbetade med kulturarvssektorn



- I mediearkivet. Statens ljud- och bildarkiv innehöll några år efter millennieskiftet videokassetter (i olika format) med en sammanlagd speltid på mer än en miljon speltimmar. Att överföra denna enorma mängd rörliga bilder till digitala format var endast möjligt genom automatiserad migrering; en bandrobot utförde arbetet. Numera är det sådana audiovisuella samlingar av hundratusentals filer som Kungliga biblioteket använt som träningsdata för att bland annat utveckla AI-modellen KB-Whisper, som kan transkribera svenskt tal till text. Metadata till nationalbibliotekets tv- och radiosamlingar har alltid varit tämligen rudimentära. Det märkliga i sammanhanget är att KB-Whisper förmodligen kommer att göra det möjligt att på automatisk väg – och långt mer utförligt än tidigare – metadata-sätta de träningsdata (äldre tv- och radioprogram) som AI-modellen byggs upp av.

– inklusive forskare som vi själva – hade för bara några år sedan någon aning om hur snabbt artificiell intelligens skulle förändra delar av ABM-sektorn, och då framför allt de verksamheter som ägnade sig åt digitalisering av företrädesvis äldre samlingar. Faktum är att ingen av Digarv-ansökningarna som beviljades medel 2018 nämnde AI, och endast två i den andra ansökningsomgången 2020. Från att ha varit ett ämne som mest diskuterades inom data- och systemvetenskap fick AI sitt verkligt breda genombrott i början av 2020-talet, inom såväl forskarvärlden som kulturarvssektorn. Ur ett akademiskt perspektiv rör det sig om mycket kort tid. För som SVD framhöll: det är en stor kontrast mellan gamla läderband och AI. Fast egentligen haltade den liknelsen, för KB-Whisper var inte tränad på text – utan på tiotusentals timmar tal, vilka främst hämtats från de stora samlingar av audiovisuellt innehåll som digitaliserades redan under 00-talet av KB:s föregångare, Statens ljud- och bildarkiv, och sedermera införlivades i nationalbiblioteket.

Flera av forskningsprojekten inom Digarv-programmet som presenteras i den här boken har arbetat med olika former av maskininlärning och AI – inte sällan i samarbete med några av landets större kulturarvsinstitutioner. I ett projekt har forskare tillsammans med Riksarkivet ägnat sig åt automatiserad handskriftstolkning, så kallad HTR (*handwritten text recognition*), i flera andra har AI, mönsterigenkänning och objektidentifiering använts för att automatiskt klassificera bildmotiv i fotografier (Stadsmuseet i Stockholm och KB) eller maskininlärning utnyttjats för stora textmassor för att korrekt annotera riksdagsprotokoll (KB). Om forskningspropositionen 2016 inte alls ägnade spaltutrymme åt AI, så har flera projekt inom Digarv-programmet alltså gjort det.

Som redaktörer (och koordinatörer) vill vi emellertid poängtera att flera av landets större kulturarvsinstitutioner parallellt med Digarv-programmet kommit att lägga alltmer fokus på de möjligheter som artificiell intelligens skapat för hantering och bearbetning av digitalt kulturarv. På ett sätt har Digarv-programmet bidragit till denna nyorientering, på ett annat har de forskningsprojekt som den här boken behandlar fått konkurrens av AI om uppmärksamhet och resurser. När en av oss (Snickars) hösten 2018 publicerade en förstudie om att etablera ett datalabb på Kungliga biblioteket, ja då nämndes artificiell intelligens nästan inte alls.⁶ Men när KB-labb sedan lanserades kom verksamheten snabbt att

främst ägna sig åt utveckling av AI-modeller. Nationalbibliotekets datalabb har därefter återkommande figurerat i den mediala offentligheten, som i SVD Näringsliv – och KB-labb är på flera sätt en utmärkt verksamhet, men det framstår inte alltid som glasklart varför ett nationalbibliotek egentligen ska lägga resurser på att ta fram språkmodeller. Även på Riksantikvarieämbetet och Riksarkivet har AI på kort tid blivit en del av den ordinarie verksamheten. Strax före jul 2024 fick till exempel Riksantikvarieämbetet i uppdrag av (den moderata) kulturministern Parisa Liljestrand att kartlägga hur AI används inom kulturmiljöarbetet, liksom inom den egna verksamheten – potentiell rationalisering hägrade. ”Riksantikvarieämbetet påverkas på olika sätt av den pågående AI-utvecklingen. Jag anser därför att det är viktigt att följa [myndighetens] arbete genom [...] återsrapporteringskrav och uppdrag eftersom de har potential att effektivisera sin verksamhet med hjälp av AI”, påpekade kulturministern i en kommentar.⁷ Samma typ av potentiell effektivisering, till exempel när det gäller automatiserade metadata – med radikalt förbättrade sökmöjligheter för forskare – återkommer också i de sätt på vilka KB rättfärdigar sin AI-satsning.⁸

Även Riksarkivet fick 2023 ett särskilt uppdrag från regeringen att redogöra för hur AI kunde användas i verksamheten. När arbetet redovisades i maj 2024 menade riksarkivarie Karin Åström Iko i ett pressmeddelande att utvecklingen och användning av artificiell intelligens rentav inneburit ”ett paradigmskifte för arkivsektorn, eftersom AI möjliggör tillgängliggörande av information på ett helt nytt sätt”.⁹ Fast myndighetens arbete med AI hade faktiskt påbörjats redan 2019; då sjuöskattes nämligen projektet AIRA – Artificiell intelligens inom Riksarkivet – med fokus på interna arbetsflöden för storskalig indexering av arkivhandlingar.¹⁰ För att stärka detta arbete med AI-baserad innovation och utveckling inrättade Riksarkivet också ett AI-labb, med FD Olof Karsvall som föreståndare, en verksamhet som snabbt vuxit till att omfatta fyra heltidsanställda *data scientists*. De har under de senaste åren framför allt ägnat sig åt att utveckla språkmodeller för automatiserad handskriftstolkning: ”Nu blir det möjlighet att maskintolka och tillgängliggöra digitaliserade arkiv, flera miljoner ovärderliga handskrivna dokument kan [om]tolkas” till sökbar text – ungefär så har det låtit i en strid ström av pressmeddelanden.¹¹ På portalen Huginface finns bland annat Riksarkivets AI-applikation Swedish



- En skärmdump från Riksarkivets interna gränssnitt visar möjligheterna med automatiserad handskriftstolkning. En av arkivets modeller som tränats på handskrifter mellan åren 1600 och 1900 har i skrivande stund en genomsnittlig noggrannhet på teckennivå om cirka 95 procent.

Lion, en språkmodell som tränats på cirka femton miljoner handskrivna ord på svenska från tre århundraden (1600–1900), där träningsdata aggregerats i samarbete med Stockholms stadsarkiv och finska Riksarkivet (*Kansallisarkisto*) i Helsingfors.

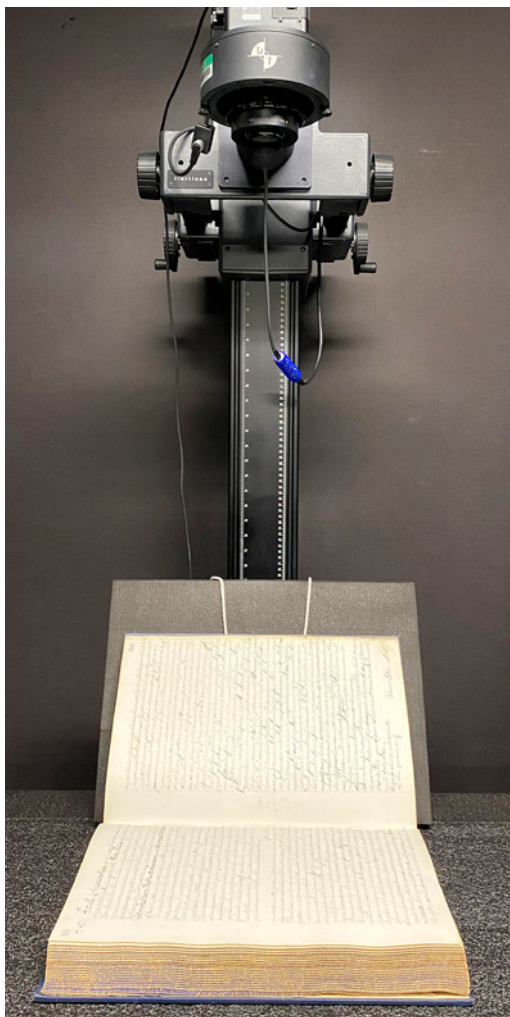
Att ABM-sektorns AI-entusiasm varit (och fortsatt är) betydande kan ingen förneka, inte heller att stora framsteg gjorts – inte minst eftersom de språkmodeller som Riksarkivet och KB lanserat är öppettillgängliga. Noterbart är också att bägge dessa datalabb ledes av disputerade medarbetare. Den samhällsnytta som ABM-sektorn alltid vinnlagt sig om har nu alltså fått ytterligare en dimension. AI har, i korthet, inneburit en möjligen oväntad renässans för (vissa) av landets större kulturarvsinstitutioner. Gamla pappershögar och dokumenttravar som under sekler samlats i magasin och bevarats – ibland trots ringa forskningsintresse – har nu för olika typer av *text mining* visat sig vara en guldgruva. Redan 2021 korades föreståndaren för KB-labb, FD Love Börjesson, till

finalist i tävlingen årets AI-svensk, och i den senaste forskningspropositionen har regeringen uppmärksammat KB:s arbete med AI-modeller, samt inte minst hur denna verksamhet ”kan bidra till att möta behov av AI från [andra] relevanta verksamheter i offentlig sektor”.¹² Även i den medialt upphäussade AI-kommissionens slutbetänkande från våren 2025 figurerar nationalbibliotekets datalabb, eftersom KB ”har tillgång till ett i princip världsunikt material [som] är en stor fördel i arbetet med att skapa en svensk språkmodell med god kvalitet”.¹³ Samtidigt förefaller samma kommission ha missat att en förutsättning för att utveckla sådana språkmodeller är att arkivens innehåll återfinns i digitaliserad form, samt att alla data därpå strukturerats – ett minst sagt omfattande arbete. ”Om data inte finns uppordnad eller tillgänglig digitalt begränsas den potentiella samhällsnyttan, eftersom informationen blir oåtkomlig för modellen”, som företrädare för KB och Riksarkivet påpekade i en debattartikel i SVD i mars 2025.¹⁴

Utifrån vårt perspektiv som koordinators för Digarv-programmet kan vi konstatera att den samtida AI-utvecklingen på mycket kort tid påverkat den nationella ABM-sektorn. I den här boken finns det tydliga spår av denna förändring, även om de enskilda kapitlen sällan resonerar om mer övergripande strukturella förskjutningar. Vi vill därför ta tillfället i akt och påpeka att en tech-optimistisk syn på AI och ABM ger vid handen att kulturarvsinstitutionernas samlingar numera utgör en sorts strategisk resurs och konkurrensfördel, i synnerhet visavi allehanda internationella aktörer och stora teknikföretag. Men på sikt kommer det att förutsätta att staten tillför medel till massdigitalisering av dessa samlingar, samt resurser för att strukturera data, ungefär på det sätt som den förra forskningspropositionen argumenterade. Därtill behöver konfliktytan mellan techbolag och kreativa branscher – vilka anser att det icke tillståndsgivna bruket av deras verk för att träna AI-modeller utgör ett brott mot upphovsrätten – adresseras, en juridisk träta där ABM-sektorn bör agera både kraftfullt och varsamt.¹⁵ *AI-boomen* innebär, i korthet, en betydande monetär möjlighet för en ofta anemisk ABM-sektor. Något nytt Digarv-program är inte i sikte, men KB-labb har som antytts tilldelats en tämligen framträdande roll i den senaste forskningspropositionen. Att kulturministern skickar ut AI-relaterade dekret och uppdrag till myndigheter som Riksantikvarieämbetet och Riksarkivet tyder också på ett lovande intresse från politiskt håll.



● Digitalisering av kulturarv kräver alltid ett *dispositif technique* (en term från Michel Foucault), som i regel påverkar det material som digitaliseras: ibland mer, ibland mindre. Automatiserad boksanning och efterföljande OCR-behandling tillhör den tidigare kategorin, enskild fotografering av sidor med handskrift den senare, förutsatt att automatiserad handskriftstolkning inte används, vilken som namnet antyder är en tolkande verksamhet – även för mjukvara.



Samtidigt är det uppenbart att AI berör (delar av) ABM-sektorn också på mindre smickrande sätt. Under de senaste tvåhundra åren har nya medietekniker förvisso kontinuerligt inverkat på hur kulturarv samlats in, dokumenterats och representerats.¹⁶ AI utgör på flera sätt en ny fas i kulturarvets mediehistoria; riksarkivarien har i så måtto rätt i att information ur arkiven nu potentiellt kan tillgängliggöras på helt nya sätt (inte minst beträffande handskrivna dokument). Men synen på vad digitala samlingar egentligen är har också ställvis förändrats; att de numera i många sammanhang betraktas som *träningsdata* bär syn för sägen. Dagens AI-modellering

innebär helt enkelt en sorts epistemologisk reduktion av de mödosamt uppbyggda samlingar som finns inom svenska arkiv, bibliotek och museer. De kommer nu till samhällsnytta – men inte riktigt på de sätt historieorienterade forskare som vi, eller personal som arbetar på kulturarvsinstitutioner, nog föreställt sig. Satsningarna på data-labb har även mött kritik; merparten av de forskare som använder Riksarkivet hade kanske hellre sett att öppettiderna utökades, eller kanhända att prissättningen för att digitalisera dokument reducerades. Upphovsrätt har också gjort att forskare som vill använda KB-labb alltid måste vara på plats i Stockholm: ”vad är det för forskarservice om datan bara finns på ett enda ställe i hela Sverige”, har någon undrat.¹⁷ Att det därtill innebär en anse- nlig kostnad för forskare att få tillgång till de dataset som finns innanför KB-labbs väggar gör också att dessa knappast är öppet tillgängliga.¹⁸ KB-labb fungerar utmärkt som forskningsinfrastruktur, men främst för den privilegierade grupp av forskare som beviljats större forsknings- anslag vilka kan bekosta användaravgiften. Andra aktörer som AI Sweden ”har nekats tillgång till Kungliga bibliotekets digitaliserade samlingar”.¹⁹ Ytterligare en annan kritik – som återkommer i denna bok – är att KB-labb blivit till en aktör med egen agenda, där natio- nalbiblioteket givit verksamheten ett slags *carte blanche* och gjort avkall på sin egen roll som öppen forskningsinfrastruktur för fors- kare vid lärosäten.

Digarv-programmet som forskningspolitik

Utifrån ett forskningspolitiskt perspektiv brottas den svenska ABM-sektorn med problemet att dess uppdrag är tudelat. Natio- nella institutioner av större format – såsom KB, Riksarkivet, Riks- antikvarieämbetet eller Statens historiska museer – ska fungera som en sorts forskningsinfrastrukturer. Men de har också ett lika viktigt ansvar för att bevara och tillgängliggöra kulturarv för landets med- borgare (nu och i framtiden). Som arkivmyndighet för den statliga förvaltningen framhåller exempelvis Riksarkivet på webben: ”Ut- gångspunkten för vårt uppdrag finns i Tryckfrihetsförordningen, som ger alla rätt att ta del av allmänna handlingar.” Riksarkivet bevarar, kort och gott, samhällets ”arkivinformation för framtida generationer och ser till att den kan användas”.²⁰ Riksarkivet är en statlig förvaltningsmyndighet under Kulturdepartementet; även Riksantikvarieämbetet lyder under samma departement, liksom

Statens historiska museer. KB sorterar visserligen under Utbildningsdepartementet, men får även uppdrag från Kulturdepartementet via Kulturrådet.

Denna insortering av kulturarvsinstitutioner inom svensk statsförvaltning har periodvis varit föremål för vissa omstruktureringar. Men grundackordet har sedan decennier tillbaka varit detsamma: att bevara och tillgängliggöra kulturarv (i olika modaliteter) för sam- och framtida generationer, samt att understödja forskning och ny kunskapsproduktion. Vi vill här inte på något sätt ifrågasätta detta tudelade uppdrag. Men däremot uppmärksamma – och göra det forskningspolitiska påpekandet – att större institutioner inom den svenska ABM-sektorn inte alltid betraktas som forskningsinfrastrukturer i akademiska sammanhang, eller av större forskningsfinansiärer för den delen. Det beror på att de inte främst (eller endast) används av forskare, och att de ofta saknar ett tydligt vetenskapligt uppdrag på det sätt som det formuleras (och utövas) inom universitet och forskningsinstitut. En konsekvens är att ABM-sektorns digitaliseringsarbete ofta inte är anpassat efter forskares behov, åtminstone inte i första hand, något som flera av de projekt som presenteras i den här antologin vittnar om. Forskare idag efterfrågar inte bara ökade öppettider till de fysiska arkiven, de önskar emellanåt en annan sorts digitalisering än vad de förvaltande ABM-institutionerna behöver för att kunna vårda, ordna och förmedla sina samlingar till allmänheten. Forskare kan till exempel vara i behov av högupplösta fotografier framför lågupplösta, rena identifikationsbilder, eller önska mera eller andra sorters metadata. Många gånger är de också intresserade av smalare typer av material, till exempel facktidsskrifter framför dagspress. Inom kulturarvsinstitutionerna är samlingen styrande när det gäller digitalisering, vilket inte sällan står i kontrast till den enskilda forskarens *on demand*-behov av partikulära handlingar. Från institutionernas sida vill man digitalisera material av samma format, det som kallas digitaliseringslina, gärna från samma samling (eller placering). Här finns alltid fysiska hanteringsaspekter som inte är kända eller relevanta för forskare. Det är i en sådan forskningskontext som Digarv-programmet haft viss betydelse, inte minst då det finansiellt administrerats av Vetenskapsrådet. Utan att vara ett tydligt uttalat infrastrukturellt forskningsprogram har Digarv haft en klar vetenskaplig slagsida; flera projekt har utgjorts av samarbeten mellan forskare och kulturarvsinstitutioner.

Som vi påpekat tidigare argumenterade forskningspropositionen 2016 för stöd till datadriven forskning inom humaniora och samhällsvetenskap (HS). Det fanns då (som nu)

ett behov av att digitalisera material [...] vid kulturarvsinstitutionernas arkiv, samlingar och bibliotek. Sådant material utgör en resurs vars forskningspotential till stor del är outnyttjad då bara en bråkdel är digitaliserat. En digitalisering av materialet vid kulturarvsinstitutionernas arkiv, samlingar och bibliotek ökar tillgängligheten och skapar nya möjligheter för forskning [...].²¹

Det var dessa formuleringar som bildade upptakten till Digarv-programmet, och noterbart är att ABM-sektorn då var tydligt framskriven som en forskningsinfrastruktur för HS-området – något som senare års AI-utveckling delvis ställt på huvudet. I VR-utlysningarna för medel inom Digarv 2018 och 2020 välkomnades därför ansökningar där samarbete skulle bedrivas mellan forskare och kulturarvsinstitutioner, med betoning på metodutveckling. Utlysningen syftade till att bygga nya forskningsmiljöer med inriktning mot digitalisering och tillgängliggörande av kulturarv, samt att stödja samverkan mellan forskare och kulturarvssektorn. Ånyo: det handlade inte om att etablera nya forskningsinfrastrukturer, men väl om forskningsprojekt med en infrastrukturell komponent.

Digarv-programmet har tveklöst varit betydelsefullt för att organisera och koordinera forskningsverksamheten på några av landets större kulturarvsinstitutioner. Till programmet knöts (som sagt) vi koordinators, liksom en styrgrupp med representanter från forskningsfinansiärerna, kulturarvsinstitutioner och Digisam, det nationella sekretariatet för digitalisering, digitalt tillgängliggörande och digitalt bevarande av kulturarvet. På ett praktiskt plan vill vi framhålla att programmet både uppmärksammat och accentuerat att ABM-sektorn kan vara en tydlig forskningsaktör; projektet *Stadens ansikten* har till exempel framgångsrikt drivits av Stadsmuseet i Stockholm, och projektet *Samisk audiovisuell samling* har varit fast förankrat vid Kungliga biblioteket. Bägge dessa institutioner har sedan tidigare disputerade forskare fast anställda i den ordinarie verksamheten. Annars är bristen på disputerad personal en anledning till att ABM-sektorn inte alltid ses som forskningsinfrastruktur, exempelvis av större forskningsfinansiärer. Frånvaro av disputerade leder till avsaknad av vetenskaplig förankring, vilket om inte annat våra egna forskningsbiografier vittnar om. Det är

nämligen ofta komplicerat att vara anställd som forskare på en kulturarvsinstitution. När en av oss (Näslund), efter att ha haft en femårig så kallad ABM-postdok och *tenure track*-anställning vid Nordiska museet med uppnådd docentkompetens, blev erbjuden av museet att arbeta med barn- och ungdomsvisningar av arkivet var valet enkelt. Hon gick vidare till en forskarassistenttjänst på Stockholms universitet för att kunna fortsätta bedriva forskning. Den andra av oss (Snickars) fick för mer än ett decennium sedan sluta på Kungliga biblioteket eftersom dåvarande riksbibliotekarie bryskt avvecklade nationalbibliotekets forskningsavdelning för vilken han var chef – en verksamhet som för övrigt numera återuppstått som Enheten för FOU och KB-labb.

Vi inser naturligtvis att tiden är förbi då sextiofem procent av Riksarkivets personal, för att ta ett exempel, var forskarutbildad.²² Men Digarv-programmet – liksom flera projekt som presenteras i denna bok – har uppmärksammat att en minskad disputationssfrekvens inom ABM-sektorn får konsekvenser: dels i synen på kulturarvsinstitutioner som forskningsinfrastrukturer, dels i kontakter med forskarsamhället och dels i relationen till etablerade forskningsfinansiärer, där man som bekant behöver vara disputerad för att söka medel. Ovan nämnda Digisam är ett illustrativt exempel på en ABM-verksamhet med svag akademisk förankring. I VR:s regleringsbrev 2018 framgick att Digarv-programmet skulle samordnas med Digisam, ett samordningssekretariat för digitalisering av kulturarv som då var placerat vid Riksantikvarieämbetet.²³ Som kulturarvsvetaren Jonathan Westin påpekat i en forskningsöversikt var Digisams uppdrag att verka för en mer ”resurseffektiv digitalisering och utgöra ett komplement till regeringens nationella strategi för digitalt kulturarv”.²⁴ Sekretariatet etablerades 2011 och placerades till en början på Riksarkivet, men flyttades (något förvånande) 2017 till Riksantikvarieämbetet – där det (ännu mera snöpligt) helt avvecklades 2023. Inom Digarv-programmet var kopplingen till Digisam närmast obefintlig, trots att samordningen avsåg digitalt kulturarv som en infrastrukturell resurs. I den utredning om svensk forskningsinfrastruktur som den tidigare högskole- och forskningsministern Tobias Krantz stod bakom, *Stärkt fokus på framtidens forskningsinfrastruktur* (SOU 2021:65), var faktiskt Digisam (och KB-labb) de enda kulturarvsanknutna exemplen vilka lyftes fram som humanistiska infrastrukturer. Men beskrivningen av Digisam var närmast förklenande:

”en plattform där 22 kulturarvsaktörer arbetar tillsammans för ett digitalt kulturarv”.²⁵

Vår avsikt är här inte att vara raljanta. Men om kulturarvsinstitutioner i akademiska sammanhang ska uppfattas som den forskningsinfrastruktur ABM-sektorn gärna vill kalla sig, ja då gäller det att flytta fram positionerna och öka de forskningspolitiska ambitionerna. För forskningsinfrastruktur är som bekant ett område som inte sällan präglas av politiska direktiv och utspel. Här går det emellertid trögt; ordet kulturarv figurerade inte alls i regeringens senaste forskningsproposition. Och i den som publicerades 2020 fanns termen bara nämnd i en minimal bisats till landets dyraste forskningsinfrastrukturer: ”de potentiella vetenskapliga genombröten som ESS och MAX IV kan bidra till [...] gäller avancerade och nya material inom många olika områden såsom läkemedel, kemi, livsmedel, förpackning, energi, elektronik, *kulturarv och fordon*” (vår kursiv).²⁶ Att VR administrerat Digarv-programmet har varit ett steg i rätt riktning; professor Stefan Svallfors, rådets (tidigare) huvudsekreterare inom HS, samt handläggare FD Eva Bergström och (tidigare) FD Kim von Hackwitz har gjort ett utmärkt arbete. Samtidigt har Digarv varit ett humanistiskt forskningsprogram på en statlig myndighet där HS-området länge kämpat i motvind. Trots inriktningen mot kulturarvsinstitutioner har Digarv-programmet på VR till exempel inte haft något att göra med (den relativt omfattande) Avdelningen för forskningsinfrastruktur. Frågan är rentav om personalen vid avdelningen ens känner till vad programmet överhuvudtaget ägnat sig åt.

En av oss (Snickars) är sedan några år tillbaka ledamot i RFI, Rådet för infrastruktur, på VR. Det är RFI som har hand om finansieringen av (stora delar av) landets forskningsinfrastruktur; årsbudgeten är nästan tre miljarder kronor, där merparten av pengarna går till att finansiera de ovan nämnda anläggningarna ESS och MAX IV vid Lunds universitet. Arbetet inom RFI är tämligen omfattande, det handlar om mellan fem och sju tvådagarsmöten per år. Men under de tjugotalet RFI-möten som Snickars deltagit i som ledamot har ABM-sektorn faktiskt inte nämnts vid ett enda tillfälle. Ingen inom RFI vet vad Digarv-programmet är. Och om KB, Riksarkivet eller Riksantikvarieämbetet ser sig själva som forskningsinfrastrukturer – ja, även statsförvaltningen – då är det inte riktigt en syn som delas av VR. Det är svårt att göra exakta beräkningar kring vilka forskningsområden som tilldelas medel för

att bygga forskningsinfrastrukturer, men mellan tummen och pek-fingret överstiger anslaget till humaniora inte ens fem procent av RFI:s totala budget. Kort sagt: Vetenskapsrådet satsar nästan inga pengar alls på humanistisk forskningsinfrastruktur.²⁷

Till dels har svenska humanister sig själva att skylla för denna skrala tilldelning; initiativen och de tydliga artikuleringarna av forskningsbehov har varit få (eller vetenskapligt undermåliga) – dock med arkeologi som betydande undantag. Samtidigt går inte allt i moll. I den guide till svensk forskningsinfrastruktur som VR publicerade 2023 framgår under rubriken ”Framtida utmaningar” att landets kulturarvsinstitutioner har betydande samlingar vilka ”i stor utsträckning ännu ej digitaliserats, och [...] därför [finns] ett fortsatt stort behov av kulturarvsdigitalisering på bred och sam-lad front”. Med referens till Digarv-programmet framhålls vidare att ”forskningsfinansiärer tagit initiativ för att tillgängliggöra och harmonisera redan digitaliserat kulturarv”. Sådana ansatser har visserligen varit lovvärda, enligt VR, men de betonar att det ”fort-farande [är] av största vikt att kulturarvsdata samordnas i relation till forskningsfrågor och problemområden, till exempel genom att digitaliseringsinitiativ svarar mot forskningsbehovsformulerade områden för infrastrukturella satsningar”. Även om det är en minst sagt knölig formulering, så stämmer beskrivningen tämligen väl in på de fjorton olika projekt som utgjort Digarv-programmet. För guiden påpekar mycket riktigt att digitalisering av kulturarvsmate-rial ”ofta genomförts på initiativ av och inom ramen för enskilda forskningsprojekt eller av enskilda forskargrupper [vilket] inne-burit att den humanistiska infrastrukturen som helhet betraktad är delvis fragmenterad, samt att användningen är ojämnt fördelad geografiskt och ämnesmässigt”.²⁸ Därtill brister tillgängligheten. Men med en forskningsmässig koppling till en ABM-institution blir samlingarna öppet tillgängliga på ett helt annat sätt. Detta tillsammans med att samlingsinstitutionerna har en betydligt lägre prislapp för administration (*overhead*) än lärosätena bör i ökad omfattning göra dem till attraktiva att förlägga forskningsmedel på inom materialnära discipliner som till exempel konst- och bild-vetenskap, historia och etnologi.

Det ligger dessvärre i sakens natur att forskardriven digitalise-ring av kulturarvet riskerar att balkanisera och splittra upp det-samma. Att landets humanistiska digitala infrastruktur delvis är fragmenterad kan vi som redaktörer för denna bok därför inte



● Svensk humaniora behöver tänka större när det gäller forskningsinfrastruktur – inte minst som ett sätt att konkurrera med andra akademiska discipliner som av tradition tilldelats oändligt mycket mer resurser för att bygga infrastrukturer. Det svenska trycket, ett samarbetsprojekt mellan fem universitetsbibliotek och Kungliga biblioteket, har varit ett steg i en sådan riktning. Tillsammans med ett Digarv-projekt ligger det till grund för en större infrastrukturell satsning (SwePrint), som med medel från Vetenskapsrådet kommer att bygga en nationell digital infrastruktur av allt svenskt tryck. På bilden skannas den illustrerade tidskriften *Idun* på Uppsala universitetsbibliotek.

annat än instämma i. Även om Digarv-programmet omfattat ett par större forskningsprojekt har det inte skilt sig från tidigare satsningar där mindre digitaliserings- och forskningsprojekt dominerat. Men kanske är något på väg att hända. Riksbankens Jubileumsfond har våren 2025 exempelvis gjort om sitt stöd för forskningsinfrastruktur så att det nu ska vara av mer omfattande art, och endast gälla ”infrastrukturer av nationell eller internationell signifikans”. De aviserade tekniska lösningarna ska redan vara utprovade och väl fungerande; därutöver ska forskarsamhällets

behov och efterfrågan av infrastrukturen ”vara väl dokumenterad vid ansökningstillfället” och infrastrukturen ska ”redan ha ett vidare antal användare”.²⁹ I efterföljden av två Digarv-projekt – inriktade på handskriftstolkning och digitalisering av svensk skönlitteratur – har Riksarkivet och KB därtill via två ansökningar till VR och RFI kvalificerat sig för att ansöka om bidrag till forskningsinfrastrukturer av nationellt intresse. Även i detta sammanhang handlar det om att etablera humanistiska forskningsinfrastrukturer av större mått, där AI i bägge fall spelar en betydande roll. I den ena ansökan, NAD+, ligger fokus på att med hjälp av AI-modeller göra tryckta arkivhandlingar på Riksarkivet både sök- och analyserbara för olika typer av datadriven forskning. Den andra, SwePrint, handlar om att på Kungliga biblioteket bygga upp en nationell digital infrastruktur som ger tillgång både till allt svenskt tryck och kurerade datamängder av detsamma.³⁰ Både NAD+ och SwePrint beviljades medel av Vetenskapsrådet hösten 2025.

Om Digarv: bakgrund, forskningsaktiviteter – och denna bok

Sedan millennieskiftet har flera olika initiativ tagits för att öka antalet disputerade forskare på svenska kulturarvsinstitutioner och indirekt därigenom forskning på deras samlingar. 2006 publicerade Sven Rentzhog den lilla skriften *Forskning & museer: En debattskrift om museernas behov av kunskap och forskningens behov av museer*. Han hade då ett förflutet som landsarkivarie och museichef, och talade sig varm för fler forskarutbildade inom museerna, samt hur en gradvis ökande kompetensbrist skulle kunna hävas.³¹ Med hänvisning till en enkät genomförd av DIK-förbundet 2004 konstaterar Rentzhog bland annat att de flesta museianställda saknade vetenskaplig utbildning; bland den lilla skara disputerade som då fanns var nästan hälften födda före 1950. Slutsatsen var att flera forskarutbildade behövde anställas inom museisektorn, framför allt på de kulturhistoriska museerna där andelen var som allra lägst – långt färre än vid konstmuseerna och de naturvetenskapliga museerna.³²

Forskningsfinansiärer som RJ och Vitterhetsakademien har under lång tid spelat en avgörande roll för forskning med nära anknytning till ABM-institutioner i Sverige, inklusive Digarv-programmet. Möjligen har inspiration kommit från Danmark, där andelen forskarutbildade inom arkiv, bibliotek och museer är

betydligt högre än i Sverige. Under alla förhållanden var dessa forskningsfinansiärer för mer än tjugo år sedan med om att lansera Nordiska museets forskarskola för museianställda. Den startade hösten 2002, och genom forskarskolan fick elva personer med anställning vid museer och Riksantikvarieämbetet möjlighet att avlägga doktorsexamen vid något av landets universitet. Målet var att "återupprätta kompetensen inom stora och viktiga kärnområden på våra museer och öka antalet anställda med vetenskaplig skolning".³³ Deltagarna fick under fem år bedriva doktorandstudier samtidigt som de förblev anställda på sina ordinarie arbetsplatser.

Några år senare (2005) kom ytterligare ett initiativ för att öka antalet disputerade forskare inom ABM-sektorn. Genom ett samarbete mellan RJ och Vitterhetsakademien utlystes sjutton så kallade ABM-postdoktjänster.³⁴ Syftet var både att stärka den vetenskapliga kompetensen på landets arkiv, bibliotek och museer och att svenska samlingar skulle beforskas i högre grad. Tanken var också att innehavarna av tjänsterna skulle fungera som ett slags broar mellan akademi och ABM-sektor. Tjänsterna var femåriga och bestod av sjuttiofem procent forskning och tjugofem procent tjänstgöring på respektive institution. De var så kallade *tenure-track*-anställningar, med målet att uppnå docentkompetens för att därefter få möjlighet att fortsätta som tillsvidareanställda forskare eller specialister inom respektive kulturarvsinstitutionen. En av oss (Näslund) hade en sådan postdok mellan 2006 och 2011 på Nordiska museet, och det var just då för tjugo år sedan som vi gemensamt började arbeta med frågor som rörde forskning och forskares villkor inom den svenska ABM-sektorn. Snickars var då verksam som forskningschef på Statens ljud- och bildarkiv, och tillsammans med FD Ulrika Torell, även hon ABM-postdok vid Nordiska museet, och FD Jan Garnert, då forskningssamordnare vid Tekniska museet, initierade vi ett nätverk för forskare verksamma på ABM-institutioner. Till det första mötet sommaren 2007 bjöds ett tiotal disputerade personer in, somliga med anställning som ABM-postdok, andra som forskare eller chefer på Riksarkivet, KB, Nordiska museet, Nationalmuseum och Nobelmuseum. Två år senare, 2009, hade nätverket knappt trettioalet medlemmar och hösten samma år arrangerade vi "Prognoser och profetior: Konferens om ABM-forskning" på landets nationalbibliotek. Konferensen innehöll fyra tematiska paneldebatter: om pendlare (forskare som varvat ABM och akademien), chefer (inom ABM), ABM-

satsningen (postdoks) och prognoser (forskare och ABM-företrädare). På KB diskuterades till exempel skillnader, för- och nackdelar med att verka som disputerad forskare inom ABM-institutioner och inom lärosäten, vilka de stora framtidsfrågorna var för kulturarvsforskning och vad som behövdes för att locka fler forskare till ABM-sektorn.

Vi nämner dessa gamla nätverksaktiviteter för att uppmärksamma att de frågor som turnerats inom Digarv-programmet varit på agendan under en längre tid. Digitalisering var visserligen under 00-talet ännu inte en avgörande fråga, men hur samarbetet mellan forskare och kulturarvsinstitutioner borde organiseras har varit föremål för en hel del tankeverksamhet. Det löper såtillvida en röd tråd mellan Nordiska museets forskarskola, satsningen på ABM-postdok och Digarv-programmet. De forskningsverksamheter som vi ägnat oss åt inom Digarv har därför utgjort ett slags fortsättning på det nätverk som vi en gång initierade – även om de delvis haft helt andra accenter.

Under den tid som forskningsprogrammet rullade, mellan 2019 och 2025, har vi årligen genomfört både programmöten och öppna seminarier. Programmötena inom Digarv riktade sig enbart till forskare och andra medarbetare i de fjorton olika projekten. Till de öppna seminarierna däremot var alla välkomna; de var avgiftsfria och genom åren deltog både forskare och anställda från lärosäten och ABM-institutioner. Totalt genomfördes tio seminarier med inbjudna talare från både Europa och USA på olika aktuella teman med relation till programmets forskningsprojekt. Här presenterades till exempel plattformar som European Time Machine (Amsterdams universitet) och @Artlas och Visual Contagion (Genèves universitet). Vår idé var från början att placera Digarv-programmet i en internationell forskningskontext; medarbetare vid forskningscentra som Alan Turing Institute (British Library), Visual Geometry Group (Oxford University) och Getty Institute var därför inbjudna.³⁵ Naturligtvis innebar pandemin problem – vilket även återkommer i beskrivningarna av flera projekt i den här boken. Fyra av de öppna seminarierna genomfördes digitalt, från våren 2020 till och med våren 2022. Den uppenbara nackdelen var att de därmed inte blev till mötesplatser för internationellt utbyte med utrymme för informell och spontan kunskapsöverföring. Samtidigt innebar de digitala seminarierna att Digarv-programmet blev bekant internationellt; flera av dem hade en mycket stor och inter-

nationell publik (i mycket högre grad än de seminarier som genomfördes på plats). Seminariet ”Digital Humanities Now!” i januari 2021 hade till exempel fler än trehundra registrerade deltagare, vilket kan jämföras med seminarierna på plats som ofta lockade femtiotalet åhörare. Noteras bör att vi före pandemin strävade efter att genomföra de fysiska, öppna seminarierna vid olika lärosäten i landet – Stockholm, Lund och Göteborg – liksom på Stadsmuseet i Stockholm, där koordineringen av Digarv-programmet var administrativt baserad från 2022 och framåt.

Som vi tidigare framhållit handlar den här boken inte så mycket om de forskningsresultat som uppnåtts inom Digarv-programmets fjorton projekt. Snarare utgör kapitlen (liksom denna inledning) en sorts sammanfattning och dokumentation över programmets direkta och indirekta resultat. Visserligen presenteras i varje kapitel de övergripande resultaten av respektive projekt. Men framför allt lyfter författarna fram mer generiska erfarenheter, metoder och slutsatser när det gäller digitalisering av kulturarvsmaterial och samarbeten mellan forskarvärlden och ABM-sektorn. På så sätt ger bokens kapitel inblick i en rad olika utmaningar – men också i en palett med frågeställningar, upplägg, arbetsmetoder och verktyg som visat sig vara framgångsrika (eller inte). Boken är indelad i fem avsnitt: I. Digitalisering och etik, II. Digitalisering och materialitet, III. Digitalisering och textualitet, IV. Digitalisering och visuellitet och V. Digitalisering och data. Dispositionen är löst sammanhållen och ska främst ses som ett uttryck för ascender i de olika projekten. Alla projekt inom Digarv-programmet har givetvis handlat om data, materialitet, etik, text och bild, men dessa teman är i olika grad karaktäristiska för projekten som helhet. De fjorton kapitlen följs av två längre förteckningar. Den första är det vi kallar programbibliografi, där samtliga publikationer och dataset som kommit ut inom programmets forskningsprojekt listas, projekt för projekt. Den andra är en förteckning över de öppna seminarier som har genomförts inom ramen för programmet, med korta beskrivningar av de deltagande forskarna och deras presentationer.

Boken första avsnitt behandlar projekt som i större utsträckning än andra ägnat sig åt **DIGITALISERING OCH ETIK**. I tre forskningsprojekt har etiska överväganden – när det gäller vilket material som bör digitaliseras och göras tillgängligt och på vilket sätt – varit särskilt centrala. Först ut är kapitlet ”Svensk karibisk kolonialism och slaveri”, författat av Fredrik Thomasson (projektledare), Ale

Svensk karibisk kolonialism och slaveri, s. 47–63

Sápmi på film och tv,
s. 65–94

Digital repatriering av
kulturarv i Amazonas,
s. 97–129

De gåtfulla
hällristningarna i
Bohuslän,
s. 133–157

Pålsson och Victor Wilson. Projektet har genomförts vid Historiska institutionen på Uppsala universitet 2019–2023, i samarbete med Riksarkivet och Fonds suédois de Saint Barthélemy i de franska kolonialarkiven i Aix-en-Provence. Det övergripande syftet har varit att genom digitalisering foga samman dokument som bevaras på vitt skilda fysiska arkiv, i akt och mening att på nya sätt kunna beforska den svenska slavekonomin och juridiken på den karibiska ön Saint Barthélemy, som var en svensk koloni mellan 1784 och 1878. Det följande kapitlet, ”Sápmi på film och tv”, är skrivet av Kajsa Kuoljok, Ulf Mörkenstam (projektledare), Ragnhild Nilsson och Mats Rohdin. Projektet pågick 2019–2022 och syftade till att på ett inkluderande sätt metadatasätta audiovisuella medier som framställer samer i KB:s mediedatabas (SMDb). Det digitaliserade materialet metadatasattes bland annat på fyra olika samiska språk, syd-, ume-, lule- och nordsamiska, i nära samarbete med samiska representanter från Sametinget och Ájttmuseet. Även projektet som presenteras i kapitlet ”Digital repatriering av kulturarv i Amazonas”, av Per Stenborg, Lilian Rebellato (projektledare), Silvia Sánchez, Adriana Muñoz och Cinthya Lana, har präglats av genomtänkta etiska överväganden. Projektet pågick 2021–2025 och dess huvudsyfte var att både utveckla och pröva möjligheter att med digitala metoder – till exempel 3D-modeller av föremål och texttolkning och OCR av handskrivna dokument – göra kulturarv tillgängligt för de befolkningar från vilka de etnografiska samlingarna på Statens museer för världskultur egentligen härstammar. Utöver själva digitaliseringen, som inte bara gör text sökbar utan också översättningsbar till andra språk, visar kapitlet på vikten av att involvera och samarbeta med representanter för den kultur vars objekt man studerar.

I bokens andra avsnitt ligger fokus på **DIGITALISERING OCH MATERIALITET**, där särskilt två projekt inkluderade metodutveckling beträffande digital avbildning av tredimensionella föremål. Båda rörde forntida bildskapande på sten, där digitala metoder på olika sätt skapat nya möjligheter att kartlägga föremålens materialitet. I kapitlet ”De gåtfulla hällristningarna i Bohuslän”, av Cecilia Lindhé, Johan Ling (projektledare), Carina Liebl, Christian Horn, Ashely Green, Mark Peternell, Julián Moyano och Ellen Meijer, beskrivs nya metoder för att analysera, identifiera och klassificera hällristningar. Inom projektet (2021–2026) användes olika tekniker såsom objektidentifiering samt 3D- och datavisualiseringar.

Tanken var inte bara att studera de synliga dragen på stenens yta – sprickor, textur och mönster – utan också undersöka hur den var uppbyggd på djupet. Kapitlet betraktar hållristningar som uttrycksfulla bilder, men studerar också deras tillkomstshistoria, det vill säga de olika tekniker som har använts för att framställa dem. Även kapitlet "Forntida bilder 2.0" av Sigmund Oehrl (projektledare), Michaela Helmbrecht, Per Widerström och Michael Ferguson handlar om hur avancerad digital teknik skapar nya möjligheter att upptäcka och se ristningar i sten. Inom ramen för projektet, som pågick 2019–2025, har omkring sjuhundra bildstenar på Gotland inventerats och avbildats med en så kallad RTI-kamera (*reflectance transformation imaging*). Genom denna avbildningsteknik – samt med hjälp av fotogrammetri – har tidigare oupptäckta bildmotiv, osynliga för blotta ögat, kunnat spåras.

Forntida bilder 2.0,
s. 159–179

I bokens tredje avsnitt, **DIGITALISERING OCH TEXTUALITET**, handlar det om projekt som på olika sätt ägnat sig åt digitalisering och bearbetning av stora textmängder. Projektledaren Astrid von Rosen skriver i kapitlet "Älska monstret" om den databas hon varit med om att bygga över scenkonstens utövare i Göteborg under perioden 1965 till 2000. Databasen – det är den som är monstret – konstruerades på basis av digitaliserad dagspress från Göteborgsområdet som tidigare endast funnits bevarad i analog form på KB. Inom von Rosens projekt, som pågick 2019–2023, digitaliserades en rad Göteborgstidningar, och en direkt effekt var därför att databasen Svenska tidningar utökades – till gagn för alla typer av forskning som använder sig av tidningsmaterial. En viktig fråga som ställs i kapitlet är hur databaser av detta slag ska bevaras långsiktigt, samt med vilka resurser och huvudmän. I kapitlet "Att tilltala överheten", författat av Jezzica Israelsson, Jonas Lindström, Joakim Nivre, Linda Oja och Maria Ågren (projektledare), handlar det istället om handskrifter som textuella data. Projektet har bedrivits mellan 2019 och 2025 vid Historiska institutionen på Uppsala universitet i samarbete med Riksarkivet, där tiotusentals så kallade suppliker (böneskrifter) från 1700-talet digitaliserats. Kapitlet redogör på ett övergripande plan för denna källtyp – hur supplikväsendet fungerade och vad dessa böneskrifter berättar om människors försörjningsarbete – men också för arbetet med att göra handskrifterna användbara för datorlingvistisk bildanalys, samt för metodutveckling för att automatiskt läsa handskriven text. Även i projektledaren Pelle Snickars kapitel, "Välfärdsstaten (o)analyserad", handlar det

Älska monstret,
s. 183–203

Att tilltala överheten,
s. 205–225

Välfärdsstaten
(o)analyserad,
s. 227–266

om metodutveckling och textanalys, men nu med basis i svensk efterkrigstid, utifrån textmodellering av svensk politik, media och kultur mellan 1945 och 1989. Fokus i kapitlet ligger på samspelet mellan forskare och kulturarvssektor, mellan akademiker och utvecklare, samt inte minst mellan vetenskapliga ideal och krass forskningspraktik. *Välfärdsstaten analyserad*, som pågick 2019–2024, kan bäst beskrivas som ett forskningsprojekt i skärningspunkten mellan mediehistoria och digital humaniora – med ett tredelat syfte: att digitalisera *äldre* textmaterial, att kurera sådan text och omforma den till strukturerade dataset, samt att bedriva forskning på dessa dataset med hjälp av olika slags digitala metoder.

Bokens fjärde avsnitt innehåller fyra kapitel som alla behandlar **DIGITALISERING OCH VISUALITET**. Bild som kunskapsobjekt har varit centralt i flera Digarv-projekt. Några av de redan nämnda projekten har skapat digitala visualiseringar av olika slag – 3D-visualiseringar, karteringar, ordmoln och visualiseringar av statistik – men i några projekt har det primära undersökningsmaterialet varit bilder. Projektet som beskrivs i kapitlet ”Kartläggning av religion i vardagen” pågick 2019–2025 och har haft Sara Ellis Nilsson som projektledare. Digitalisering av Ikonografiska registret, ett register över motiv ur den svenska medeltida kyrkokonsten, var en av projektets uppgifter. Kapitlet är samskrivet tillsammans med Terese Zachrisson, Anders Fröjmark, Lena Liepe och Johan Åhlfeldt, och handlar om arbetet med att skapa en digital karta över olika slags spår som vördnaden av helgonen avsatte i den medeltida svenska kyrkoprovinsen. Kartan var tänkt som en sorts forskningsresurs, ett hjälpmedel för var och en som vill studera helgonkulten och dess betydelse i det medeltida Sverige. Härvidlag handlar kapitlet också om hur forskningsdriven digitalisering tillämpats, samt om projektets metodik för digitalisering, med nyttjande av etablerade digitala standarder och vokabulärer. Även i kapitlet ”Stadens ansikten” är bilder i fokus. Det är skrivet av Rebecka Lennartsson (projektledare), Anders Hast och Mats Hayen, och baserar sig på ett forskningsprojekt med samma namn, förlagt vid Stadsmuseet i Stockholm 2021–2025. I projektet kopplas den textdatabas som hör till det historiska Rotemansarkivet samman med urval av den stora samling porträttfotografi som finns i museet. Kapitlet resonerar om hur ansiktigenkänning gör det möjligt att spåra individer genom bildsamlingen – och på så vis koppla samman klass, civilstånd och bostadsadress med fotografiska porträtt. Inom ramen för

Kartläggning av
religion i vardagen,
s. 271–288

Stadens ansikten,
s. 291–326

projektet har även digitala visualiseringar av Stockholms invånare på basis av socioekonomisk status gjorts. Om länkar mellan metadata och bilder varit centrala i projektet *Stadens ansikten*, gäller detsamma i kapitlet ”Deltagande i det visuella kulturarvet” av Anna Näslund (projektledare), Karin Hansson och Amanda Wasielewski. Detta projekt, som pågick 2019–2023, har emellertid till skillnad från alla övriga inom Digarv-programmet inte inkluderat digitalisering av material vid en kulturarvsinstitution. Projektets syfte har snarare varit att utveckla metoder för att kvalificera metadata-sättningen för bildsamlingar så att de bättre stämmer överens med de intressen som finns bland olika typer av forskare. Förutom att undersöka de behov av metadata som förekommer i olika grupper av användare visar kapitlet hur metadata på avgörande sätt styr både sökbarhet och tillgänglighet av bildsamlingar. Kapitlet resonerar också om hur olika digitala gränssnitt, AI-modeller och bilddataset påverkar vad som blir betonat, synligt eller mer eller mindre dolt i det visuella kulturarvet. Snarlika frågeställningar återkommer även i kapitlet ”Att digitalisera barns kulturarv”, författat av Anna Sparrman (projektledare), Pål Aarsand, Yelyzaveta Hrehaniuk, Johanna Sjöberg, Ylva Ågren och Alex Ormalm. Jämfört med andra projekt i Digarv-programmet har detta forskningsprojekt, som pågick 2021–2025, på ett genomgripande sätt påverkat arbetet på den kulturarvsinstitution, Svenskt barnbildarkiv, som varit en central samarbetspartner. En av totalt två anställda arkivarier har varit involverad i projektet, som är det första digitaliseringsarbete som genomförts vid detta bildarkiv. Kapitlet aktualiserar dels relationen mellan digitaliserade respektive digitalt födda bilder, dels de upphovsrättsliga och etiska överväganden som digitaliseringen medför – detta eftersom det främst rört sig om bilder skapade av minderåriga.

Slutligen innehåller boken ett femte avsnitt som fokuserar på **DIGITALISERING OCH DATA**. Även om i stort sett alla Digarv-projekt inbegripit digitalisering och byggande av nya databaser och plattformar för tillgängliggörande, så sticker två projekt ut när det gäller omfång. Projektledaren Per Pettersson Lidbom beskriver i kapitlet ”Att bygga en databas med en miljard observationer” det omfattande och mödosamma arbetet med att sammanställa data från flera olika typer av arkiv, samlingar och datakällor som vanligen inte är sammankopplade. Värt att notera är att Digarv-programmet bara har varit en av flera finansiärer av detta projekt

Deltagande i det visuella kulturarvet, s. 329–347

Att digitalisera barns kulturarv, s. 349–373

Att bygga en databas med en miljard observationer, s. 377–402

(2019–2026). Förutom att databasbyggande av denna storlek kräver flera stora anslag lyfter Pettersson Lidbom utmaningen med långsiktigt bevarande. Eftersom den information som samlats i databasen kommer från flera olika kulturarvsinstitutioner och myndigheter finns det ingen given huvudman som tar ett långsiktigt ansvar för bevarande och underhåll av databasen. Till sist så kapitlet ”Tvärvetenskap i arbete”, skrivet av Jens Edlund (projektledare) och Mats Fridlund, i vilket förekomsten av begrepp som terror och terrorism undersöks i den svenska riksdagen mellan 1968 och 2018. Projektet (2021–2026) baserades på språkteknologisk analys av ljudinspelningarna från riksdagens debatter och tydliggör bland annat genom sitt tvärdisciplinära angreppssätt skillnaden mellan digitisering och digitalisering. Riksdagens debatter fanns redan digitiserade och öppet tillgängliga men i kapitlet argumenteras det för att debatterna ännu inte var digitalt forskningsbara, det vill säga digitaliserade. Genom att utveckla existerande språkteknologiska och talteknologiska verktyg visar Edlund och Fridlund på komplexiteten i det som i allmänt tal kallas för digitalisering.

Kulturarv, digitalisering, framtid

Vetenskapsrådets Digarv-program har med humanistiska mått mätt varit ett stort forskningsprogram som involverat runt hundra-femtio forskare, digitaliserings- och samlingsexperter från svenska lärosäten och ABM-institutioner under en tidsperiod av sju år. Generellt har ett mycket omfattande arbete genomförts, men samtidigt är vi bara i början när det gäller utvecklingen av digitalt tillgängliggörande och digitala metoder och verktyg i relation till forskning. Avslutningsvis vill vi som koordinators för programmet därför säga några ord om vägar framåt.

För det första är det tydligt att de mest lyckade forskningsprojekten inom Digarv-programmet haft en gemensam komponent: ömsesidig förståelse och kommunikation mellan akademiska forskare, de som använder eller utvecklar digitala verktyg och de som förvaltar och råder över kulturarvssamlingar. Här finns inte en enda kungsväg, utan snarare en rad olika framgångsfaktorer. I något fall har det handlat om att involverade forskare haft en anställning på en kulturarvsinstitution, med såväl kunskap om samlingarna som direkt tillgång till dem och deras förvaltare. I

andra fall har administrativa organisationer och infrastrukturer aktivt fungerat som dynamiska kontaktpunkter eller noder mellan humaniora och teknikvetenskap, såsom till exempel GRIDH, Göteborgs universitets forskningsinfrastruktur för digital humaniora. Genom öppna kommunikationskanaler har sådana noder skapat bred förståelse för vilken typ av data och digitala verktyg som forskare behöver. Omvänt har de givit forskare inblick i vilka möjligheter – och begränsningar – som återfinns i befintlig teknik, i samlingar och i samlingarnas systemiska strukturer, såväl analogt som digitalt. I de Digarv-projekt där kontakt mellan de olika typerna av kompetenser varit ett uttalat mål, har inte bara själva forskningsuppgiften blivit enklare och snabbare att realisera. Projekten har även fungerat som en grund för utveckling, både när det gäller digitala verktyg och samlingsförvaltning. En kontinuerlig dialog har, i korthet, skapat ett ömsesidigt lärande.

För det andra visar flera Digarv-projekt på att öppen tillgång till data fortsatt ligger högt på forskarnas önskelista. Visserligen fungerar Jonas Harvards CD-skiva med *Dagens Nyheter* 1865 inte längre på de flesta datorer; lagringsmediet CD har blivit obsolet. Men grundtanken att genom *open access* distribuera kulturarvsdata har förblivit central. Öppen betyder i sammanhanget tre saker: att data kan nås oberoende av var i landet man är verksam som forskare, att data är kostnadsfria och att de kan bearbetas fritt. På samma sätt som forskare idag behöver fri tillgång till databaser för publikationer vill de ha fri tillgång till forskningsmaterial i digital form. Den enskilda forskare (Jonas Harvard) som för femton år sedan digitaliserade dagspress hade förstås inte alls de ekonomiska och personella resurser som kulturarvsinstitutioner besitter. För att skala upp behövs en betydande budget, men också uthållig finansiering och ett långsiktigt förvaltningsansvar för att säkra öppen tillgång. För mer modernt kulturarvsmaterial figurerar givetvis en rad upphovsrättsliga problematiker. Men Digarv-programmet har med all önskvärd tydlighet visat att de inläsningstendenser av digitaliserade samlingar som fortsatt återfinns inom många kulturarvsinstitutioner behöver försvinna.

För det tredje har programmet accentuerat att kulturarvets digitalisering är en framtidsfråga – på fler än ett sätt. En av oss (Näslund) är numera verksam på Institutet för framtidsstudier; i en sådan forskningskontext är det tydligt att även bevarande och tillgängliggörande av historiskt kulturarv är en framtidsfråga. Redan digita-

- AI-modellen Perplexity genererar (sommaren 2025) denna bild utifrån prompten "AI supporting digital preservation" som ska illustrera hur artificiell intelligens kan hjälpa till med att bevara digitalt kulturarv – måhända inte helt övertygande.



liserat kulturarv är också något som ABM-sektorn framgent behöver adressera i ökad utsträckning, om så via migrering av äldre digitala format, ny digitalisering eller med hjälp av artificiell intelligens. När vi i skrivande stund (sommaren 2025) ställer frågan hur AI kan hjälpa till med digitalt bevarande, så svarar AI-modellen Perplexity bland annat följande:

AI kan automatisera sortering, klassificering och gallring av stora datamängder, vilket är särskilt värdefullt för arkiv och museer som hanterar mycket digitalt material. Detta minskar risken för mänskliga fel och frigör tid för mer kvalificerade arbetsuppgifter [...] Text- och taligenkänning används för att göra dokument och inspelningar sökbara och lättillgängliga – AI kan exempelvis skapa text från gamla ljudupptagningar eller handskrivna dokument [...] AI kan [också] bistå med att kontrollera och säkerställa autenticitet och dataintegritet, så att bevarad digital information verkligen är oförvanskad och det går att spåra ursprung och förändringar över tid. Detta är avgörande för att kunna lita på bevarat material.³⁶

Perplexity har uppenbart gott självförtroende och obefintligt självtvivel, när den säger sig kunna arbeta felfritt och garantera autenticitet. I praktiken gör förstås AI "fel" ibland. Ironiskt nog är

det kanske just självkritiken som är människans största tillgång i sammanhanget. De kultur- och samhällskritiska perspektiv som är basen för humanistisk och samhällsvetenskaplig forskning sätter fenomen, aktörer och händelser i ett större historiskt och samhällsreligt perspektiv. Forskning är ett ständigt prövande och omprövande. Genom forskning kan akademi och ABM-institutioner – när de är som bäst – vara en motvikt till samtidens (över)tro på den egna tidens unicitet och nya teknikers revolutionerande karaktär. Forskningen kan även relativisera samtidens naturaliserade föreställningar, det vi tar för sanning och fakta, och visa att det är historiskt, kulturellt och socialt bestämda uppfattningar och konventioner. Forskare inom humaniora och samhällsvetenskap kan därmed sätta samtidens digitalisering i perspektiv till dåtiden – men även till framtiden.

Kulturarvsinstitutionernas uppgift och avgörande samhälleliga roll är att samla, bevara och förmedla vårt gemensamma kulturarv.³⁷ Digitalisering för bevarande och tillgängliggörande av kulturarv är därför inte bara en aktuell fråga för samtida kulturpolitik och forskning. För att upprätthålla vår tids demokratiska system och förstå de historiska dimensionerna i nuet krävs medvetna överväganden när det gäller bevarande av kulturarvet. Både *vad* som bevaras och *hur* det görs tillgängligt är en demokratisk fråga. Medan arkivhandlingar traditionellt har betraktats som unika objekt kan de idag, genom digitaliseringen, uppträda i flera exemplar. Numera både skapas och förvaltas dokument och bilder digitalt, och en utmaning inför framtiden är därför att avgöra vad som är original och vad som är reviderade versioner.³⁸ Den samtida digitaliseringen väcker därmed frågor om hur vi i framtiden ska kunna garantera och bedöma äkthet och autenticitet i historiska dokument, både de som är digitalt födda och de som digitaliserats från analoga format. En annan framtidsfråga handlar om i vilken utsträckning ökad digital tillgänglighet av historiska dokument kommer att påverka utvecklingen av generativ AI. Inom bild- och ljudområdet förs idag som bekant en intensiv diskussion om *deepfakes* och deras potentiella samhälleliga, demokratiska och legala effekter.³⁹ När det gäller bild har debatten främst handlat om samtida fenomen – påhittade spektakulära eller komprometterande bilder som används för politiska syften eller som ren underhållning. Samtidens ideologiska landskap kan enkelt modifieras med hjälp av AI, men ännu så länge inte historiska bilder, i och med att modellerna

främst är tränade på samtida fotografier från webben – eftersom det är dessa som varit tillgängliga online, och eftersom de är så oändligt många fler än de som återfinns i äldre digitaliserade bildsamlingar.⁴⁰ Möjligen är det dock bara en tidsfråga innan modeller tränade på stora mängder historiska fotografier och bilder generellt kommer att lanseras – forskningsprojekt i mindre skala finns redan i Europa⁴¹ – och det skulle radikalt förändra förutsättningarna för att även skapa ett slags förflutenhetens *deepfake*.

Slutligen – och för att återknytta till den här inledningens början – är även frågan om tillgång till data en framtidsfråga. På kommersiella plattformar som Meta och Open AI, vilka inledningsvis var kostnadsfria att använda, har konsumenten idag två val. Antingen får man betala månadsavgift, eller betala med sina användardata, vilket i regel innebär att dessa data lagras, används, säljs vidare till tredje part eller disponeras fritt av plattformen. Inom flera projekt i Digarv-programmet har behovet av och utmaningarna med öppen tillgång till kulturarvssamlingarnas digitaliserade material aktualiserats. Med den kommersiella webbens terminologi finns på vissa håll fortfarande betalväggar. Lika central som frågan om användarnas öppna tillgång till kulturarvsdata är därför frågan om användarnas integritet när det gäller användardata. En viktig demokratifråga för framtiden är att försäkra oss om att vårt gemensamma kulturarv görs tillgängligt i system som *inte* registrerar *vem* som söker *vad*. Den demokratiska friheten att ta del av det gemensamma kulturarvet behöver både skyddas och utvecklas.

NOTER

1. Den mediehistoriska workshopen ägde rum på Linköpings universitet och var föranledd av den kommande boken *1800-talets mediesystem* (red.) Patrik Lundell och Jonas Harvard (Stockholm: Mediehistoriskt arkiv, 2010).

2. Sara L. Bränström, "Kungliga biblioteket står modell för ny AI", *Svenska Dagbladet* 24/5 2024.

3. Sara L. Bränström, "Då kan även en skånsk läkare bli förstörd", *Svenska Dagbladet* 22/2 2025.

4. Termen digitalisering (av kulturarv) har på senare tid inom ABM-sektorn (i dess interna språkbruk) alltmer ersatts av begreppet digitisering, efter engelskans *digitization*. Vi är medvetna om denna begreppsförskjutning, men eftersom Digarv-programmet – och den forskningsproposition (Prop. 2016/17:50) som låg till grund för programmet – uttryckligen handlade om digitalisering av material ur kulturarvsinstitutionernas arkiv, så har vi valt att i den här boken fortsatt använda termen digitalisering. Det är vi nu inte ensamma om; även en stor myndighet som Riksantikvarieämbetet använder fortsatt bägge termerna.

5. *Kunskap i samverkan: För samhällets utmaningar och stärkt konkurrenskraft* (Prop. 2016/17:50), 95.
6. Pelle Snickars, "Datalabb på KB: En förstudie", <https://urn.kb.se/resolve?urn=urn:nbn:se:kb:publ-339> (senast kontrollerad 31/8 2025).
7. "Uppdrag om AI till myndigheter inom kulturområdet", pressmeddelande Kulturdepartementet 20/12 2024, <https://www.regeringen.se/pressmeddelanden/2024/12/uppdrag-om-ai-till-myndigheter-inom-kulturomradet/> (senast kontrollerad 31/8 2025).
8. <https://www.kb.se/forskning-pa-kb/kb-labb---ai-och-digital-forskning.html> (senast kontrollerad 14/9 2025).
9. "AI ett paradigmskifte för arkiven", pressmeddelande Riksarkivet 31/5 2024, <https://riksarkivet.se/inlagg/ai-ett-paradigmskifte-for-arkiven>. Se även Riksarkivet, "Redovisning av uppdrag om AI i verksamheten", skrivelse 31/---- 2024, <https://riksarkivet.se/files/2025/03/rapport-aiiverksamheten-riksarkivet-240531.pdf> (senast kontrollerade 31/8 2025).
10. Årsredovisning 2024 för Riksarkivet, https://riksarkivet.se/files/2024/09/riksarkivet_arsredovisning_2024.pdf (senast kontrollerad 31/8 2025).
11. "Ny banbrytande AI-modell för svenska historiska texter", pressmeddelande Riksarkivet 7/2 2024, <https://riksarkivet.se/inlagg/ny-banbrytande-ai-modell-for-svenska-historiska-texter> (senast kontrollerad 31/8 2025).
12. *Forskning och innovation för framtid, nyfikenhet och nytta* (Prop. 2024/25:60), 67. För en diskussion om utvecklingen av KB-labb, se Love Börjesson et al., "Transfiguring the Library as Digital Research Infrastructure: Making KBLab at the National Library of Sweden", *College, Research & Libraries* nr 4, 2024.
13. *AI-kommissionens färdplan för Sverige* (SOU 2025:12), 122.
14. Faton Rekathati och Erik Lenas, "Risk för svenskt självsmål inom AI", *Svenska Dagbladet* 24/3 2025.
15. Se t.ex. Henning Eklund, "Strid om AI och böcker: 'Massiv piratkopiering'", *Svenska Dagbladet* 30/6 2025.
16. För en historisk diskussion om olika mediepraktiker och kulturarv, se Pelle Snickars, *Kulturarvets mediehistoria: Dokumentation och representation 1750–1950* (Lund: Mediehistoriskt arkiv, 2020).
17. Martin Röshammar, "KB-labb: Infrastruktur för AI-utveckling med förhinder", *Biblioteksbladet* 16/6 2023.
18. I ett mejl från KB-labb till en av oss (Näslund) 16/4 2025 framgår att kostnaden för en arbetsstation i labbet uppgår till 100 000 kronor under det första projektåret. Priset inkluderar kostnader för uppstart och introduktion (men inte tekniskt stöd), därefter är prisbilden 50 000 kronor per år oberoende av hur många forskare som nyttjar platsen.
19. Per Gudmundson, "Svensk AI får läsa på Flashback istället för KB", *Svenska Dagbladet* 10/12 2022.
20. Om Riksarkivet, <https://riksarkivet.se/om-riksarkivet> (senast kontrollerad 31/8 2025).
21. Prop. 2016/17:50, 95.
22. *Forskarutbildningens meritvärde* (SOU 1981:30), 48.
23. Vetenskapsrådets regleringsbrev för år 2018 återfinns på Ekonomistyrningsverkets webbplats, <https://www.esv.se/regleringsbrev/18765/pdf?Version=HelaBrevet> (senast kontrollerad 31/8 2025).
24. Jonathan Westin, "Kulturarvets digitalisering", *Digitalisering inom*

- kultursektorn: Förhållanden i nordiska länder (2025), <https://pub.norden.org/kulturanalysnorden2025-2/nord2025-009.pdf> (senast kontrollerad 31/8 2025).
25. *Stärkt fokus på framtidens forskningsinfrastruktur* (SOU 2021:65), 217.
26. *Forskning, frihet, framtid: Kunskap och innovation för Sverige* (Prop. 2020/21:60), 139.
27. De forskningsinfrastrukturer som i skrivande stund (sommaren 2025) finansieras av VR där kulturarvsinstitutioner medverkar är tre till antalet: 1) Swedpop (Svenska befolkningsdatabaser för forskning) där Riksarkivet och Stockholms stadsarkiv medverkar, 2) SweDigArk (Swedish National Infrastructure for Digital Archaeology) med Riksantikvarieämbetet som part, samt 3) Huminfra där Riksarkivet och KB-labb ingår.
28. "Vetenskapsrådets guide till forskningsinfrastrukturen 2023", <https://www.vr.se/download/18.72136e7218779250836274ec/1683110770640/Vetenskapsrådets%20guide%20till%20forskningsinfrastrukturen%20VR%202023.pdf> (senast kontrollerad 31/8 2025).
29. RJ Infrastruktur för forskning 2025, <https://www.rj.se/utlysningar/rj-infrastruktur-for-forskning-2025/> (senast kontrollerad 31/8 2025).
30. Ansökningarna om bidrag till forskningsinfrastruktur av nationellt intresse skickades in till VR i februari 2025 – NAD+ av Olof Karsvall, Riksarkivet och SwePrint av Pelle Snickars, Lunds universitet och KB.
31. Sten Rentzhog, *Forskning & museer: En debattskrift om museernas behov av kunskap och forskningens behov av museer* (Stockholm: Nordiska museets förlag, 2006).
32. "Tillbaka till gå? Enkätundersökning om forskning och forskarkompetens på museer 1/3 2004". DIK och Museimannaförbundet, 2004.
33. "Nordiska museets forskarskola utökas", osignerad, *Nättidningen Svensk Historia* 29/8 2004, <https://svenskhistoria.se/nordiska-museets-forskarskola-utokas/> (senast kontrollerad 31/8 2025).
34. Jussi Nuorteva och Staffan Wahlén, *ABM postdok-program: Utvärderingsrapport 2012* (Stockholm: Kungl. Vitterhets historie och antikvitets akademien och Stiftelsen Riksbankens Jubileumsfond, 2012).
35. I Digarvs seminarier deltog ett flertal forskare från institutioner och centra för digitala humaniora utanför Sverige, t.ex. företrädare för King's College London, Oxford Internet Institute, Lancaster University, University of Malaga, Helsingfors universitet, Centre for the Study of the Networked Image / London South Bank University och University of California, Santa Barbara. Flera internationella presentationer gjordes även av personer verksamma inom ABM-institutioner med tydlig digital profil, t.ex. amerikanska Getty Institute och Stanford University Libraries, danska Det Kongelige bibliotek och British Library. Ett stort antal forskare inom fältet digitala metoder, digitala humaniora och digitalisering har gjort presentationer inom Digarv-programmet, såsom Oliver Grau (Danube University), Stuart Dunn (King's College London), Anna Bentkowska-Kafel och Béatrice Joyeux-Prunel (University of Geneva), Johanna Drucker (University of California, Los Angeles), Thomas Smits (University of Amsterdam), Annet Dekker (University of Amsterdam) och Fabian Offert (University of California, Santa Barbara). Se vidare bokens appendix, "Digarv-programmets öppna forskningsaktiviteter 2019–2025", på s. 461.
36. Hur kan AI hjälpa till med digitalt bevarande? Sökfråga ställd till AI-modellen Perplexity 15/8 2025.

37. Se t.ex. UNESCO-rapporten från 2015, *Recommendation concerning the protection and promotion of museums and collections, their diversity and their role in society*, <https://unesdoc.unesco.org/ark:/48223/pf0000246331> (senast kontrollerad 31/8 2025). Se även den svenska arkivutredningen, *Härifrån till evigheten: En långsiktig arkivpolitik för förvaltning och kulturarv* (SOU 2019:58).

38. SOU 2019:58, 138. Se även Staffan Smedberg, "Sverige", *Det globala minnet: Nedslag i den internationella arkivhistorien* (red.) Lars Jörwall et al. (Stockholm: Riksarkivet, 2012).

39. Se t.ex. Sami Alanazi och Seemal Asif, "Exploring deepfake technology: Creation, consequences and countermeasures", *Human-Intelligent System Integration* nr 6, 2024, liksom Laishram Hemanta Singh, Panem Charanarur och Naveen Kumar Chaudhary, "Advancements in detecting deepfakes: AI algorithms and future prospects – a review", *Discover Internet Things* nr 5, 2025.

40. Sanjay Chawla et al., "Ten years after ImageNet: A 360° perspective on artificial intelligence", *Royal Society Open Science* nr 10, 2023. "ImageNet is an image database organized according to the WordNet hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images", <https://www.image-net.org> (senast kontrollerad 31/8 2025).

41. Se t.ex. projektet *Digital Curator* som inkluderar knappt 200 000 äldre målningar från museer i Österrike, Tyskland, Tjeckien och Slovakien, <https://digitalcurator.art> (senast kontrollerad 31/8 2025). Stora historiska bilddataset finns redan online, som t.ex. Visual Contagion med 3,6 miljoner bilder från 1791 ur europeiska illustrerade tidskrifter och tidningar från samlingarna i Bibliothèque Nationale de France, Universitätsbibliothek Heidelberg, Princeton Blue Mountain Project m.fl., <https://visualcontagions.unige.ch/explore/corpus/corpus/95/> (senast kontrollerad 18/8 2025).