

Vad är digital historia?

JACOB ORRJE*

Uppsala universitet

De senaste decennierna har datorer blivit en integrerad del av forskning och undervisning, en utveckling som bara har accelererat under pandemin med distansarbete och Zoom-undervisning. Denna digitaliseringsprocess har förändrat hur vi söker efter källor, hur vi besöker fysiska arkiv (där digital fotografering allt mer ersatt en intensiv läsning under själva besöket), hur vi skriver (där ordbehandlare och referenshanteringssystem genomsyrar vårt författande) och hur vi publicerar oss (nu ofta som Open Access online, och i tidskrifter med digitala publiceringssystem).¹ Dessa förändringar av historievetenskapens forskningspraktiker har även fått stora konsekvenser för historia som kunskapsfält.² Exempelvis möjliggör sökningar i Google Books och andra databaser fyllda med digitiserade historiska källor undersökningar som ofta även hade varit möjliga för några decennier sedan, men som hade tagit mångdubbelt längre tid att genomföra. Parallellt med dessa förändringar av historia i stort har även en rad historiker börjat presentera sin forskning som "digital historia". Med detta begrepp följer ofta anspråk på att de på mer genomgripande och medvetna sätt integrerar digitala metoder i undersökningarna. Återkommande svenska konferenser i digital historia har arrangerats de senaste åren och även kurser ges på temat. Digital historia knuten till centra för digitala humaniora drivs vid många universitet, och

* Forskare i idé- och lärdoms historia

1. Om digitala publiceringssystem som OJS, se redaktörens inledning i *Historisk tidskrift* nr 2 2021. Se också "From the Editor's Desk: The Times They Are a-Changin'", *The American Historical Review* 126:1 (2021) s. xv–xix.

2. För ett kritiskt perspektiv på digitaliseringen av universiteten och en diskussion kring vad vi förlorar genom denna process, se Anne Berg, "Det postpandemiska zoomuniversitetet – en pessimistisk prognos", *Historisk tidskrift* 140:4 (2020) s. 687–695.

nyligen har ett dedikerat digitalhistoriskt centrum etablerats vid Lunds universitet.

I ett internationellt perspektiv har framväxten av detta digitalhistoriska fält pågått under relativt lång tid.³ Redan 1997–1998 grundades Virginia Center for Digital History som var bland de första att använda begreppet. År 2008 publicerade *Journal of American History* en diskussion med titeln "The promise of digital history".⁴ I ett svenskt sammanhang introducerade Anna Nilsson Hammar begreppet digital historia i en artikel i *Scandia* år 2015.⁵ Fältet är nu någonstans i 20-årsåldern och börjar bli allt mer etablerat och vedertaget. Att presentera digital historia som något nytt eller framväxande, vilket har varit ett typiskt sätt att beskriva fältet de senaste decennierna, börjar därför kanske nå vägs ände.

Samtidigt är digital historia ett relativt marginaliserat fält i Sverige, i relation till mer traditionell historisk metodologi. I jämförelse med Sverige har vårt grannland Finland en mer utvecklad digitalhistorisk forskning. Ett exempel på detta är två engelskspråkiga böcker som nyligen författats av finska forskare. Hannu Salmis monografi *What is Digital History* (Cambridge: Polity 2021) ger en översikt av fältet. Mats Fridlund, Mila Oiva och Petri Pajus antologi *Digital Histories: Emergent Approaches within the new Digital History* (Helsingfors: Helsinki University Press 2020) samlar ett stort antal (18) kapitel av huvudsakligen finska forskare och ger en intressant inblick i ett brokigt fält fullt av metodologiska reflektioner och experiment. Med anledning av dessa böcker kan det finnas skäl även för historiker i Sverige att stanna upp och fråga sig vad fältet digital historia är idag. Hur kan självidentifierande digitalhistoriker relateras till dagens övriga historievetenskapliga forskning? Skulle historisk forskning i stort gynnas av mer avancerad digital analys av historiska källor?

3. En aktuell översikt av det internationella forskningsläget kan hittas exempelvis i C. Anemiek Romein m.fl., "State of the Field: Digital History", *History* 105:365 (2020) s. 291–312.

4. "Interchange: The Promise of Digital History", *Journal of American History* 95:2 (2008) s. 452–491.

5. Anna Nilsson Hammar, "Scandia introducerar: Digital history", *Scandia: Tidskrift för historisk forskning* 81:2 (2015). Värt att notera är att digitala metoder såklart användes och diskuterades inom svensk historisk forskning även tidigare. Ett exempel är Tomas Nygren, "Digitala material och verktyg: Möjligheter och problem utifrån exemplet spatial history", *Historisk tidskrift* 133:3 (2013) s. 474–482.

Digital historia som fält och forskningspraktik

Salmi introducerar läsaren för centrala frågor inom digital historia. I ett första kapitel diskuteras de problem som kan uppstå när vi gör digitala representationer av kulturarv.⁶ I nästa kapitel är fokus på relationen mellan närläsning och datorunderstödd textanalys med hjälp av statistisk metod. I de följande kapitlen berörs kartor och spatialitet, interdisciplinärhet, och digitala sätt för historiker att presentera det förflutna. Sammantaget redogör Salmi för de mest centrala metoderna inom digital historia, både för analyser av historisk text och rumslighet. Dessutom visar han hur dessa digitala metoder förlitar sig på tvärvetenskapliga kopplingar inte bara till litteraturvetenskap, konstvetenskap och mediestudier, utan även till mer tekniska fält så som datavetenskap och informationsteknologi.⁷

Där Salmi ger en strukturerad förståelse för fältet, ger *Digital Histories* en bra bild av den mångfald av metoder, studieobjekt och empiri som finns inom digital historia idag. I några av kapitlen anlägges ett mer kritiskt perspektiv där författarna försöker ringa in vad digital historia är. Petri Paju historiserar exempelvis datoranvändandet inom historisk forskning i Finland från 1960-talets kvantitativa undersökningar fram till dagens digitala historia – och frågar sig hur denna äldre digitala historia förhåller sig till den som idag presenteras som ”ny”. Mats Fridlund närmar sig frågan från en mer teoretisk utgångspunkt. Han ser den typ av historia som alla historiker ägnar sig åt (som genomsyras av allt fler digitala forskningsverktyg och -praktiker) som en typ av digitalhistorisk normalvetenskap i Kuhnsk mening (Fridlund kallar det digital historia 1.0). Denna forskning ställer han mot mer radikala kvantitativa digitala historikers forskning, som han identifierar som ett nytt paradigm (digital historia 2.0). Mitt emellan dessa två paradigm placerar Fridlund en hybridvariant – en ”digital historia 1.5” – som visserligen systematiskt använder sig av digitala verktyg, men som anpassar dessa verktyg så att

6. Inom digitala humaniora och digital historia skiljer man ofta mellan begreppen ”digitalisering” (som är den mer konkreta praktiken att göra analogt material digitalt) och ”digitalisering” som snarare beskriver mer övergripande förändringar av verksamheter, organisationer och processer (så som exempelvis digitalisering av historisk forskning).

7. För en diskussion om digital historia och interdisciplinärhet, se Tim van der Heijden & Andreas Fickers, ”Inside the Trading Zone: Thinkering in a Digital History Lab”, *Digital Humanities Quarterly* 14:3 (2020).

de kan besvara frågor som är relevanta för mer traditionell historisk forskning.⁸

I ljuset av dessa två böcker utkristalliserar sig två problemområden digitala historiker huvudsakligen förhåller sig till. För det första vilken roll digitala representationer av historiska källor har kommit att få i historisk forskning. Hur bör vi representera historiska samlingar i digital form? Hur behöver vår källkritik utvecklas när vi arbetar med digitala representationer av källor, eller med källor som aldrig varit något annat än digitala? Dessa frågeställningar öppnar i sin tur till större frågor: Hur ser våra digitala infrastrukturer ut? Vad är politiken kring urval av källor som görs digitalt tillgängliga och hur bör relationen mellan historisk forskning och minnesinstitutioner se ut? Vilken typ av beskrivande data, så kallad metadata, ska finnas tillsammans med digitala representationer av kulturarv? För det andra är digital historia ett fält för metodologiska experiment, där datavetenskapliga tekniker använts för att analysera och strukturera stora mängder historiskt material. På så sätt vill dessa forskare överkomma envisa historiografiska problem inom dagens historia – och främst klyftan mellan mer översiktliga diakrona studier och detaljerad, men begränsad, mikrohistoria.

Digital historia, digitala metoder och digitala humaniora

Genomgående för kapitlet i *Digital Histories* är att de behandlar olika metoder som huvudsakligen utvecklats inom digitala humaniora. Främst handlar det om så kallad *distant reading* (på svenska ofta översatt som fjärrläsning), lanserat av litteraturvetare så som Franco Moretti som ett alternativ, eller snarare komplement, till traditionell närläsning. Moretti har presenterat fjärrläsning som ett sätt att bryta sig loss från en smal litterär kanon, för att i stället se sådant som både är mycket mindre och mycket större än enskilda texter.⁹ I praktiken innebär denna läsart att forskare analyserar digital text genom mer eller mindre avancerade statistiska metoder eller genom det som inom datavetenskapen kallas maskininlärning. På så sätt går det att hitta olika former av mönster i stora mängder text: exempelvis hur frekvenser av specifika ord ändras

8. Samma poäng kring behovet av att integrera digital historia och mer traditionella historievetenskapliga metoder görs i Gerben Zaagsma, "On Digital History", *BMGN: Low Countries Historical Review* 128:4 (2013) s. 3–29.

9. Franco Moretti, "Conjectures on World Literature", *New Left Review* 1:1 (2000) s. 54–68; Franco Moretti, *Distant reading* (London 2013).

över tid, vilka ord som förekommer tillsammans (d.v.s. samvarierar), eller vilka källor som kan organiseras i tematiska kluster på grund av likheter i vokabulär och syntax.

Salmis kapitel "Reading and Textuality in Digital History" ger en utmärkt introduktion till dessa analysmetoder. Han pekar också på hur dessa ursprungligen litteraturvetenskapliga metoder blir något annat när de plockas upp av historiker. När närläsning lanserades inom 1900-talets litterära nykritik bestod metoden i att bortse från kontext för att i stället undersöka textens interna struktur. Inom historia har begreppet närläsning dock snarare kommit att involvera just tolkningar av strukturella aspekter av historiska källor i relation till vidare kulturella och sociala sammanhang.¹⁰ På samma sätt presenterade Moretti inledningsvis fjärläsning som ett sätt att vidga perspektiven bortom västerländsk kanon. Inom digital historia ses dock också ännu tydligare en rörelse bort från detaljerade mikrohistorier mot långa diakrona analyser med hjälp av materialmängder som är så stora att vi genom traditionella läsningar omöjligt kan få en systematisk överblick över dem (så kallad "Big Data"). I sitt avslutande kapitel i *Digital Histories* diskuterar Jo Guldi just denna möjlighet att skriva en digitalhistorisk *longue durée*, och hur sådana studier förhåller sig till arbeten som genomförts efter 1980- och 1990-talens kulturella vändning.¹¹

Flera av studierna i *Digital Histories* är dessutom goda exempel på hur digitala metoder möjliggör sådana längre historiska narrativ. Exempelvis används dessa metoder för att kartlägga den nordiska allemansrättens historia genom digitiserat källmaterial från nordiska parlament (Matti La Mela). På liknande sätt undersöks också olika konceptualiseringar av begreppet internationalism i Storbritanniens parlament (Pasi Ihalainen och Aleksi Sahala), och olika sätt att skriva om Venuspassagen i *Philosophical Transactions* 1753–1777 (Reetta Sippola). Samtidigt understryks i många av kapitlen i *Digital Histories*, precis som flertalet digitalhistoriker i stort gör, en medvetenhet om behovet att kombinera historisk fjärläsning med traditionella närläsningar som i högre grad bibehåller historisk kontext. Exempelvis diskuterar Johan Jarlbrink i sitt kapitel på

10. Salmi (2021) s. 29–56.

11. Se även Guldi & David Armitage, *The History Manifesto* (Cambridge 2014); Jo Guldi, "The Measures of Modernity: The New Quantitative Metrics of Historical Change Over Time and Their Critical Interpretation", *International Journal for History, Culture and Modernity* 7:1 (2019) s. 899–939.

ett skarpsynt sätt allt det praktiska och manuella arbete som gör digital analys möjlig.

Det finns en vana inom digitalhistoriska studier att, på gott och ont, fokusera på de metoder som utgör grunden för analysen. Detta gör just att de kan fungera inspirerande för potentiellt intresserade som vill experimentera med liknande verktyg. Samtidigt gränsar studierna ibland mot det lite för tekniska, och historiska narrativ och argument hamnar stundtals i skymundan. Den amerikanske historikern Cameron Blevins har diskuterat just det som han ser som en "widespread absence of academic argumentation in digital history",¹² det vill säga att i många digitalhistoriska studier diskuteras vad som framöver skulle gå att genomföra med hjälp av nya metoder snarare än att fokus sätts på att faktiskt utveckla rådande forskningslägen genom nya metodingångar. De kapitel i *Digital Histories* som i högre grad lyckas använda digitala metoder för att intervensera i rådande forskningslägen är också några av de som i mina ögon lyser starkast. Allra tydligast ses det i kapitlet av Reima Välimäki, Aleks Vesanto, Anni Hella, Adam Poznański och Filip Ginter där fjärläsning för att fastställa författaren av anonyma senmedeltida kätterska manuskript används.

Som även Salmi påpekar, är digitala metoder oftast mycket textcentrerade, vilket kan ses som ett steg tillbaka efter de senaste decenniernas visuella och materiella vändningar.¹³ Ytterligare en kritik som kan riktas mot digital historia är att historiska aktörers agens tenderar att drunkna i dessa stora textmassor. Fjärläsning ser ut att passa bättre för diakrona begreppshistoriska analyser, än för att synliggöra enskilda människors levda erfarenhet i historiska skeenden. Melanie Conroy och Kimmo Elos kapitel i *Digital Histories* visar dock att digitala metoder, använda på ett kreativt sätt, även kan generera en tät beskrivning av specifika sociala sammanhang. I sin studie av den östtyska oppositionsrörelsen 1975–1990 använder de sig av metadata från en databas över fotografier, och gör en historisk nätverksanalys över vilka personer som befinner sig på samma fotografier. På detta sätt skapar de diagram över dessa nätverk,

12. Cameron Blevins, "Digital History's Perpetual Future Tense", i Matthew K. Gold & Lauren F. Klein (red.), *Debates in the Digital Humanities 2016* (Minneapolis 2016) s. 315. Samma observation har även gjorts av Lincoln Mullen, "A Braided Narrative for Digital History", i Matthew K. Gold & Lauren F. Klein (red.), *Debates in the Digital Humanities 2019* (Minneapolis 2019) s. 382–388.

13. Salmi 2021 s. 42–43.

som både synliggör kopplingar (vilka syns på samma fotografier) och vilka aktörer som hade störst tyngd i nätverken (de som förekommer på många fotografier). Genom analyser av var fotografierna tagits, visar författarna även vilka platser som var viktigast för oppositionsrörelsen, samt hur betydelsen av olika platser förändrades över tid. Det är lätt att se att liknande metoder skulle kunna ge nya perspektiv på en rad historiska frågor och sammanhang.

Data och källor

Grunden för digitalhistoriska analyser är att historiska källor har omvandlats till data. Flertalet av metoderna ställer också krav på att denna digitala data är strukturerad på rätt sätt. Även om textkällor idag ofta finns tillgängliga som digitala bilder, kräver digitala analysmetoder att de har hög upplösning och att texten tolkats och gjorts sökbar via OCR (Optical Character Recognition). Urvalskriterier vid arkiv, bibliotek och finansierare av storskalig digitisering påverkar alltså vilka källor som det är möjligt att analysera digitalt.¹⁴ Vad gäller äldre handskrivet material är möjligheterna till automatiserad transkribering mer begränsad. Även om tekniker för handskriftsigenkänning (HTR, Handwritten Text Recognition) de senaste åren har utvecklats i snabb takt, med plattformen Transkribus som främsta exempel, så gör varierande handstilar och icke-standardiserat språkbruk denna typ av material mer svårarbetat.¹⁵ Å ena sidan är risken alltså att fjärrläsning skapar en ny kanon – bestående av de källor som är möjliga att bearbeta med digitala metoder, och som huvudsakligen består av tryckt text från 1900-talet eller så kallad *born-digital data* från 2000-talet (det vill säga data som aldrig existerat i annat än digital form).¹⁶ Å andra sidan har teknologier så som HTR, som kan omvandla äldre historisk handskrift till sökbar text, potentialen att kunna öppna upp äldre svårlästa materialkategorier för exempelvis studenter.

En annan fråga är hur begreppet data – eller mer specifika begrepp som historisk data och forskningsdata – formar de sätt vi konceptualise-

14. Nanna Bonde Thylstrup, *The Politics of Mass Digitization* (Cambridge, Mass. 2018) s. 52.

15. Guenter Muehlberger m.fl., "Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study", *Journal of Documentation* 75:5 (2019) s. 954–976.

16. Att urval för digitisering premierar vissa historiska aktörer och aspekter av historiska aktörers liv har påpekats av Lara Putnam, "The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast", *The American Historical Review* 121:2 (2016) s. 377–378.

rar historievetenskaplig verksamhet och vår relation till det förflutna. I sitt kapitel i *Digital Histories* relaterar Risto Turunen data till källor med hjälp av DIKW-pyramiden, en datavetenskaplig modell som slår fast en hierarki mellan (d)ata, (i)nformation, (k)unskap och (v)isdom.¹⁷ Turunen argumenterar för att pyramiden bör revideras för att bli historievetenskapligt relevant, genom att källor läggs till under data (källor är inte läsbara för maskiner, men data är det). Givet en viss typ av historisk analys är den hierarki Turunen föreslår begriplig – det vill säga analyser som är kvantitativa och ligger nära samhällsvetenskapliga undersökningar. Samtidigt förutsätter hierarkin ett visst förhållningssätt till historiska källor som inte är given för mer kvalitativt inriktade historiker. För flertalet forskare, som interagerar med digitala källor genom fulltextsökning eller mer traditionell närläsning av digitalt material, skulle nog en modell där data ses som mer abstrakt än historiska källor bättre avspegla deras arbete. Det vill säga, för dessa forskare är det fortfarande källor som möjliggör historisk analys – och för dem är det därmed fortsatt viktigt att data presenteras på ett samlat sätt, med ett visst gränssnitt och med metadata, som skapar en digital representation (av exempelvis en bok eller ett arkiv) som eftersträvar att bibehålla källans historiska form samtidigt som den öppnar upp nya möjligheter för digital analys. Skillnaden i förhållningssätt blir möjligtvis tydligare genom en parallell till hur de flesta idag förhåller sig till digitala fotografier i vardagen: även om bilderna idag tekniskt betraktat består av ettor och nollor, så är det sällan vi förhåller oss till dem som data. I stället tittar vi på dem, och redigerar dem, via gränssnitt som förbigår datalagret och presenterar digitala fotografier i form av en fysisk bild (samtidigt som gränssnitten såklart samtidigt förändrar hur vi tittar på, redigerar och delar med oss av fotografiskt material). Just hur historiker ser på relationen mellan begreppen "data" och "källor" verkar alltså till stor del vara avhängigt vilka metoder som används, vilken typ av historiskt narrativ vi vill skriva och hur vi relaterar digitala arbetssätt till mer traditionell analys.

Till stor del beror vårt förhållningssätt troligen också på om vi utbildats i en mer kvalitativ eller kvantitativ historievetenskaplig tradition. Det faktum att mycket digital historia utgår från data snarare än källor

17. Kritik av denna modell för information och kunskap har exempelvis framförts i Martin Frické, "The Knowledge Pyramid: A Critique of the DIKW Hierarchy", *Journal of Information Science* 35:2 (2009) s. 131–142.

kan i många fall bidra till att den känns främmande för en bredare grupp historiker (och kanske främst då de som är mer kvalitativt inriktade). Hur hemma forskare känner sig i detta digitala fält hänger såklart även ihop med tidigare erfarenheter inte bara av digitala teknologier utan också av mer matematiska och statistiska analysmetoder. Till stor del ser digital historia också ut att brottas med samma typ av skev könsfördelning som den datavetenskap från vilket fältet lånar många av sina metoder. Heidi Kurvinen diskuterar dessa problem i ett utmärkt självreflexivt kapitel i *Digital Histories*. Hon beskriver sin metodologiska resa, som feministisk historiker, in i digitala humaniora, och pekar på hur informationsteknologiskt kunnande och genus, samt samvariationen av de två, är faktorer som påverkar forskares möjlighet att delta i detta manligt dominerade fält. Samtidigt skildrar Kurvinen erfarenheter av att övervinna dessa hinder, och visar på hur digital analys öppnar upp nya möjligheter för hennes studie av Finlands tidiga rösträttsrörelse (genom fjärläsning av texter i tidningar från sekelskiftet 1900).

En digital vändning på vilket sätt?

Vad är då digital historia? Flera av kapitlen i *Digital Histories* talar om en digital vändning inom historia. Av fallstudierna i boken att döma, är det i så fall en något annorlunda vändning än den uppsjö av andra vändningar som lanserats de senaste decennierna. Den så kallade kulturella vändningen gjorde kulturteori central för stora delar av humanistisk forskning, och möjliggjorde därmed nya perspektiv på dåtiden. Andra vändningar, så som exempelvis den visuella och materiella, har huvudsakligen handlar om att belysa nya materialkategorier, och därmed också nya historiska sammanhang. Om digital historia är en vändning, så innebär den i så fall nytt fokus på två områden: för det första på infrastrukturella frågor kring hur vi bevarar dåtiden (samt tillhörande källkritiska frågor kring hur vi kan arbeta med digitala historiska representationer), för det andra på nya typer av metoder för att besvara huvudsakligen redan existerande forskningsproblem. På samma sätt pekar exempelvis Fridlunds beskrivning av en "digital historia 1.5" på styrkan i en digital historia som på nya sätt undersöker och besvarar frågor som är relevanta för bredare historievetenskapliga fält. För att detta ska vara möjligt behöver digitala metoder integreras inom historisk normalvetenskap, något som poängterades av den nederländske digitalhistorikern Gerben

Zaagsma redan i en artikel från 2013.¹⁸ Möjligen behöver användningen av mer avancerade digitala metoder avdramatiseras och ses som en möjlig metodingång bland andra. Fjärrläsning skulle mycket väl kunna vara användbar för historiker i stort, som en mer kvalificerad form av fritextsökning. Det är möjligt att de metodologiska insikter som behövs för att omsätta dessa metoder i konkret forskning på sikt blir mer spridda bland historiker i allmänhet, och att det då blir möjligt att skriva digital historia som fokuserar mindre på metoderna som sådana och mer på de historiska sammanhang som de kan synliggöra och de frågor de kan besvara. De senaste årens digitala historia visar tydliga tendenser att utvecklas åt det hållet. På så sätt mognar fältet även långsamt från något med potentiella användningar i framtiden, till något som är relevant och viktigt för historisk forskning här och nu.

18. Zaagsma 2013.