

Humanioras utfordringer i det digitale landskapet

AKTUELT: Er digitale verktøy bare pynt på søknader til forskningsprosjekter i academia eller nyttige hjelpemidler innen humanistisk forskning? Norske akademikere bør se til digitala humaniora, for å finne noe av svaret.

Av Ane Vestby Øiesvold

Dette feltet anvender nettopp digitale verktøy, metoder og databaser i humanistisk forskning. Men har det en reell nytteverdi? Jeg dro til den internasjonale konferansen Digital humanities at Oxford summer school 3. - 7. juli for å undersøke dette nærmere.

Ulike problematikker knyttet til det digitale har hatt gjenklang blant de store internasjonale forskningsmiljøene de to siste tiårene og resultert i diverse sommerskoler, her representert ved Oxford summer school. I norsk offentlighet har debatten rundt dette engasjert og i 2014 sparket Fritt ord for alvor i gang diskusjonen rundt humanioras utvikling og fremtid. Da publiserte de rapporten «Hva skal vi med humaniora?» om tilstanden til humaniora i Norge i dag. Den ble overrakt kunnskapsminister Torbjørn Røe Isaksen som et innspill til stortingsmeldingen, den såkalte «humaniorameldingen», som skulle komme på et senere tidspunkt. I mars 2017 ble stortingsmeldingen «Humaniora i Norge» lagt frem av regjeringen Solberg.

I stortingsmeldingen finner man et lengere avsnitt om digitalisering som peker på en del uløste utfordringer på feltet. To viktige elementer påpekes: For det første anser myndighetene «kobling mellom e-vitenskap (nye modeller, metoder, algoritmer) og mellom humanistiske disipliner»¹ som et behov i kunnskapssamfunnet. «Humanistiske disipliner» sikter i denne sammenheng til arkiv-, bibliotek- og informasjonsvitenskap. For det andre påpekes det at det ligger et stort potensial for forskere i det å ta i bruk informasjonsteknologi som et «effektivt og kraftfullt verktøy»².

i sin forskning. Utvilsomt gode ambisjoner, men er de realistiske? Arkiv- og bibliotekvitenskap er helt klart to fag som kan ha store fordeler av å benytte digitale modeller, algoritmer og lignende, men det finnes en rekke andre fag som ikke nødvendigvis har bruk for slike digitale verktøy for å prosessere store mengder informasjon. Få forskere innenfor for eksempel idéhistorie eller filosofi jobber med å lage, eller behøver, digitale modeller, grafer og algoritmer. Hvis det aktuelle kildematerialet er så omfangsrikt at slike verktøy faktisk har en funksjon, er det riktignok flott å ha slike dem tilgjengelig. Forskningsprosjekter på denne skalaen er derimot sjeldne fordi idéhistorie og filosofi stort sett driver med nærlesning, altså benytter en annen metode i forskningsprosessen, og i tillegg krever de store ressurser.

LES OGSÅ: [Å samtale med maskiner](#) • [Når bøkene forsvinner](#)

Et tredje punkt som kommer frem i regjeringens gjennomgang av humaniora i Norge er viktigheten

av tverrfaglighet. Digital humaniora er et felt som ansees for å muliggjøre samarbeid på tvers av ulike disipliner. Det heter i rapporten at: «Sammen med nye arbeidsmåter og metodeutvikling utvider det samarbeidsflaten mot andre fagområder.»³ «Nye arbeidsmåter og metodeutvikling» sikter her til digitale metoder og verktøy. Gjennom disse kan nytteflaten utvides betraktelig og nye spennende perspektiver komme til syne. Ved å flette ulike felt sammen i et større perspektiv og ta i bruk ressurser fra ulike forskningsmiljø, skapes det et potensial for å spre kunnskap til offentligheten på en god måte.

Et internasjonalt fellesskap

Her hjemme er det Nasjonalbiblioteket som har ledet an i utviklingen med å digitalisere samlingen sin og gi offentligheten fri tilgang til den, slik også kjente forskningsbiblioteker i USA har gjort. Nasjonalbiblioteket har fått skryt fra flere hold for den store andelen av bøkene i den totale samlingen deres som er scannet og digitalisert. Utviklingen har kommet langt, og det er (kanskje) i slike sammenhenger at nytten av det digitale blir mest synlig. Forskere, så vel som privatpersoner, kan nå sitte hjemme og bla seg side for side gjennom bøker i Nasjonalbibliotekets digitale samling, Bokhylla.no, på pc'en hjemmefra. Det er av verdi innenfor, så vel som utenfor, academia, fordi dette digitale verktøyet gir direkte tilgang til innholdet i millioner av bøker. Det er altså snakk om en helt annen grad av tilgjengelighet og spredning av kunnskap til offentligheten enn tidligere.

Spørsmålet er heller om teknisk tunge prosjekter, som TEI-prosjektet, har noen reell verdi for andre enn kodere som arbeider med det. TEI, Text Encoding Initiative, er et prosjekt som tidlig utviklet en standard for presentasjon av tekst i digital form og konkrete retningslinjer for koding av dette. Siden 1994 har disse retningslinjene blitt brukt av forskere, museer og biblioteker verden over og er svært kjent i miljøet. Digitale verktøy i humanistisk forskning bør brukes kritisk for å kunne ha en nytteverdi. OpenRefine, er for eksempel et frittstående skrivebordsverktøy som renser og konverterer data til ulike format – det ligner en database for å strukturere datamengder. For en forsker kan et godt digitalt verktøy, som OpenRefine eller TEI-standardene, ta seg av mye av grovarbeidet og spare tid i forskningsprosessen. Tolkningen av funnene slike digitale verktøy løfter frem, er fremdeles nødvendig for at settet med data skal gi mening i en større sammenheng.

Dette er noe av det den internasjonale sommerskolen i Oxford har forsøkt å gjøre. Utviklingen av feltet ved blant annet universitet i Oxford, som har pågått siden 90-tallet, er gjennom en rekke prosjekter både et godt eksempel på hva som er mulig, og et forbilde for det den norske regjeringen etterlyser i forskning her hjemme. Det som er blitt Digital Humanities Oxford Summer school begynte i det små som TEI summer school i 1997. Sommerskolen besto da kun av en workshop som konsentrerte seg om retningslinjene fra et av de første prosjektene innen digital humaniora: TEI-prosjektet omtalt ovenfor. I takt med den teknologiske utviklingen har sommerskolen vokst seg større siden den startet for tjue år siden. Nå tilbys det workshops på ulike nivåer, over 80 foredrag, og i 2016 huset den 166 deltagere fra hele verden.

I humaniorameldingen er de tre hovedpunktene når det gjelder det digitale: koblingen mellom e-

vitenskapen og humanistiske disipliner, potensialet i informasjonsverktøy for forskere, og utvidelsen av samarbeidsflater til andre fag. Kan vi se det potensialet som regjeringen etterlyser i eksempler fra seminarrekkene i Oxford?

Tre ulike innganger til digitala humaniora

Det tverrfaglige feltet digital musikkvitenskap kombinerer musikkvitenskap og digitale verktøy. Digital musicology ser på hvordan digitale teknikker, analyseverktøy, programmer og andre elementer fra IT-verdenen kan anvendes i musikkvitenskap. Mange av teknikkene som anvendes her har sin opprinnelse i helt andre fagfelt, som datavitenskap, bibliotek- og informasjonsvitenskap. Det tverrfaglige samarbeidet som regjeringen etterlyser, er tydelig i programmet til Digital musicology. Et eksempel er prosjektet, Music Encoding Initiative, som ligner på Text Encoding Initiative. I fellesskap prøver forskere å utvikle et system for koding av musikkdokumenter, som i de fleste tilfeller betyr noteark. Med deres egne ord:

The Music Encoding Initiative (MEI) is an open-source effort to define a system for encoding musical documents in a machine-readable structure. MEI brings together specialists from various music research communities, including technologists, librarians, historians, and theorists in a common effort to define best practices for representing a broad range of musical documents and structures.⁴

Teknologer, bibliotekarer, historikere og programmerere bidrar altså til dette åpne prosjektet for å finne en god standard som en rekke disipliner kan anvende på sitt materiale gjennom et samarbeid mellom e-vitenskapen og kodere. Music Encoding Initiative viser at den digitale delen av musikkvitenskapen ikke kun benytter programvare for å skape selve musikken, men også for å analysere den i større forskningsprosjekter, altså anvendes teknologien hensiktsmessig på sitt eget og helt andre felt.

Andre interessante ting som skjer på dette feltet, blant annet i Oxford, er analyser av opptak ved hjelp av datamaskiner, bruk av dataanalyse for å katalogisere og identifisere opptak, samt se på hvordan forskere har trent opp datamaskiner for automatisk å gjenkjenne mønstre i musikk, også når det gjelder eldre noteark.⁵

Music Encoding Initiative – prosjektet som bruker disse digitale analyseverktøyene for musikalske verker, representerer en utvidelse av samarbeidsflatene og en god utnyttelse av det digitale i et tverrfaglig perspektiv. Historikere, musikkforskere, IT-avdelinger og bibliotekarer kan da jobbe sammen om å kode nye og gamle musikkdokumenter. Av den grunn kan prosjektet betraktes som et vellykket eksempel innen digitala humaniora, til etterfølgelse av norske akademikere. Informasjon om prosjektene som har brukt denne standarden for koding, selve grunnkodingen (kilden) og video med retningslinjer spres til offentligheten gjennom prinsippet om frivillig gitt, gratis tilgang til kunnskapen (open-source).

Et fugleperspektiv

Seminarrekken «Introduction to digital humanities» ga en god innføring i et felt med en mengde navn og databaser. Her er noen av de største og mest vellykkede prosjektene representert, hvor databasene er helt åpne. Disse prosjektene viser igjen mulighetene som ligger i slike digitale verktøy når de brukes som midler for å fremme god humanistisk forskning. Forskere kan konstruere kart, vise dokumenter digitalt og få ut mye god kvantitativ informasjon.

Early English books online (EEBO) er en samling av engelske, tidligmoderne verker, og databasen inneholder nå 125 000 titler.⁶ Tematisk hører verkene til historie, filosofi, politisk tenkning, lingvistikk, teologi, musikk og matematikk. Slike databaser er en ressurs for både studenter og forskere, men ofte må man abonnere for å få tilgang. Heldigvis eksisterer det åpne databaser og samlinger som er et resultat av gjennomførte forskningsprosjekter som har brukt digitale verktøy og modeller. To eksempler på det er Shakespeare quartos archive, Quill-prosjektet og Mapping the Republic of letters ved Stanford universitet. Shakespeare quartos archive er et digitalt arkiv over de skjøreste kvarter⁷ ulike biblioteker har av hans skuespill. Gjennom dette samarbeidet er skuespillene scannet digitalt fra flere forskjellige forskningsbiblioteker og er nå samlet i databasen. Et slikt prosjekt viser hvor verdifulle digitale verktøy kan være for å bevare dokumenter som er så skjøre at de ikke kan bli behandlet manuelt lenger, og derfor må bevares digitalt, dog på bekostning av materialiteten; det vil si kunnskap om type papir, blekk, innbinding, format og sanseopplevelsen, som er knyttet til å ta og føle på den. Når noe fremstilles digitalt, mister man denne dimensjonen ved et verk og forståelsen av det.

LES OGSÅ: [Forskningsrådet og humaniora](#) • [Humanioras fremtid](#)

Quill-prosjektet bruker datamodeller i større grad enn Shakespeare quartos archive for å kartlegge og vise forhandlingsmønstre. I databasen viser digitale programmer hvordan politiske forhandlinger har foregått, med fokus på konstitusjoner, lovverk og traktater. Blant handler Quill-prosjektet om grunnlovsforhandlingene i 1787, hvor den amerikanske grunnloven ble tegnet opp. Ved hjelp av ulike datamodeller kan man se hvordan forhandlingsprosessene mellom flere aktører foregikk. Slike prosjekter er morsomme, men hvor mye nyttig kvalitativ informasjon man får ut av modellene er mer tvilsomt. På en mer generell basis kan det absolutt være nyttig. For eksempel for å se at i sesjon no. 88 av forhandlingene, var delstaten Rhode Island representert med én, mens Connecticut hadde fem representanter.

Mapping the Republic of letters er et stort prosjekt ved Stanford universitet. Republic of letters var et intellektuelt nettverk som korresponderte i brevform over hele Europa på 1600- og 1700-tallet. Medlemmene utvekslet kunnskap og oppdagelser. Kjente personer som Voltaire, John Locke og Jean le Rond D'Alembert var alle med i nettverket, som Stanford-prosjektet bokstavelig talt har kartlagt. Brevenes destinasjon, antall og mottakere er tegnet opp på kart for hver tenker. Brevskrivning var essensielt for dette fellesskapet av opplysningstidens intellektuelle, fordi det var slik de distribuerte og overførte kunnskap. Nettopp derfor er det digitale verktøyet, nemlig programmet som genererer kartet, samt plotter inn bevegelsesmønstrene til brevene, svært nyttig. Forskere kan se nye mønstre som gir viktig informasjon om kontakt mellom ulike filosofer og

vitenskapsmenn på denne tiden, for eksempel hvem Voltaire eller Galileo brevvekslet med. En forsker kan deretter bruke informasjonen til å si noe om hvilken betydning denne korrespondansen hadde idéhistorisk og kunnskapshistorisk sett. Som tidligere nevnt, etterlyste regjeringens melding om digitala humaniora bruk av teknologirenskaper for å spre kunnskapen fra akademikernes side. Alle disse prosjektene er eksempler på forskere som har anvendt kraftige digitale verktøy i arbeidet med klassisk humanistisk kildemateriale, med stor suksess og i stor skala. Materialet står fremdeles i sentrum, men hjulpet av det digitale. Både EEBO og Mapping of the Republic of letters sier noe om potensialet som ligger i slike verktøy. I slike store prosjekter med store ressurser kan digitale verktøy og datamodeller komme til sin rett som informasjonsverktøy.

En verden av koder

Koding av tekst også mulig å tilegne seg i Oxford; enten gjennom Data science for the humanities, From text to tech seminarerne eller An introduction to the guidelines of the TEI. Disse seminarrekke dreier seg om koding, men på to forskjellige nivåer: koding av tekst og maskinlæring. Datascience for the humanities er den mest avanserte av de tre og dreier seg om maskinlæring. Maskinlæring er en gren av kunstig intelligens som fokuserer på å utvikle og designe algoritmer som datamaskiner lærer for å utvikle adferd basert på empiriske data. Dette tverrfaglig fagområdet har bidrag fra bla. informatikk, matematikk og statistikk. Et hovedfokus innenfor feltet er at datamaskiner skal gjenkjenne komplekse mønstre for dermed å fatte intelligente beslutninger basert på tilgjengelige data. Med andre ord er dette avansert og særdeles teknisk felt for de fleste humanister. På en annen side er det samtidig svært nyttig for en forsker å kode sitt eget dataprogram eller verktøy, som gir langt større kontroll i forhold til materialet, mulighet til å sjekke sitt eget datasett og mulighet for å oppdage feilkilder.⁸

Seminarrekken som tar for seg Tei-retningslinjene er en lettere versjon av en lignende problematikk. Til forskjell fra Datascience for the humanities dreier denne seminarrekken seg om å kode tekst, gjerne i litt større mengder gjennom bruk av Text Encoding Initiative standarden for koding. Det innebærer å bruke digitale kommandoer for å orientere seg i en tekst og hente ut kvantitativ informasjon. Dets tilgjengelighet gjør at disse retningslinjene er tatt i bruk av klassiske humanistiske disipliner som bibliotekvitenskap og museumsvitenskap, og viser dermed godt koblingen mellom e-vitenskap og humanister.

Det beste fra to verdener

Humanities data: A hands-on approach er seminarrekken som kombinerer den(mer) datatekniske siden med typisk humanistisk forskningsdata. Digitale verktøy gir oss ofte kvantitativ informasjon, uten noe videre nytte.

Potensialet i slike verktøy ble i midlertidig tydelig gjennom arbeidet med et ganske standard humanistisk datasett. Det var en bunke brev fra Thomas «Mummy» Pettigrew (1791 - 1865). Han

var en lege, antikvar og egyptolog som har etterlatt seg omfangsrik korrespondanse. Vår samling var et utvalg på rundt 150 brev, og oppgaven gikk ut på og rydde disse rotete dataene (brevene) og kategorisere dem. I kombinasjon med foredrag om hvordan man som forsker kan håndtere og forholde seg til slike rotete datasett, jobbet vi med tre programmer for å gjøre nettopp dette: OpenRefine, Python og SQL. Alle tre fungerer slik at man importerer datasett i et valgt format. Deretter bruker man enkle kommandoer eller koding for å sortere ut materialet som man trenger. Da kurset var ferdig var hele gruppen av humanister, deriblant mange kunsthistorikere, begeistrede tilhengere av OpenRefine. Når disse programmene brukes som informasjonsteknologi, kan også bibliotekarer, kunsthistorikere og idéhistorikere bruke digitale verktøy i sitt forskningsarbeid på en funksjonell og hensiktsmessig måte. Informasjonsverktøyene kan da brukes, slik vi gjorde, til å gjøre mye av grovarbeidet for humanister med mye rotete data, og samtidig styrke koblingen mellom e-vitenskapen og humaniora, som regjeringen etterlyser i humaniorameldingen. OpenRefine brukt på Pettigrews brev er et eksempel på hvordan digitale verktøy fungerer best for humanister: når den kritiske sansen til det digitale blir beholdt og man anvender det på humanioras egne premisser. Nysgjerrigheten på ny kunnskap og evnen til å systematisk jobbe seg gjennom ofte tungt, ukjent materiale er et akademisk fortrinn. Denne kompetansen kommer man langt med i møte med programmer og rotete data.

Veien videre

Denne artikkelen åpnet med spørsmålet om digitale verktøy bare er pynt på søknader til forskningsprosjekter eller faktiske nyttig hjelpemidler innen humanistisk forskning. Nyansene i dette feltet har jeg forsøkt å vise gjennom å løfte frem noen eksempler fra de åtte ulike seminarrekkene som løp parallelt i Oxford på sommerskolen. Digitale informasjonsverktøy som dataanalyseprogrammer for ulike materiale, kart, grafer og lignende er svært nyttige for å spre kunnskapen prosjektene har funnet fram til. Det som står i fare for kun å bli pynt, og for teknisk for den store gruppen av akademikere, er for eksempel maskinlæring som har en begrenset nytteverdi for den typen materiale som humanister jobber med i sin forskning.

Sommerskolen i Oxford speiler noen elementer fra humanioramelding fra 2017. Under de tre inngangene til feltet digital humaniora slik de er skissert opp her, finner vi eksempler som oppfyller én eller flere av regjeringens oppfordringer. De ulike seminarrekkene viser enten en direkte kobling mellom e-vitenskap og humanistiske disipliner, gjennom koding og programmering, potensialet i informasjonsverktøy for forskere i store prosjekter som Early English books online og Shakespeare quartos archive, og utvidelse av samarbeidsflater til andre fag, som digital musicology.

Så, hvor er digital humaniora på vei? Svaret er ikke enkelt, grunnet hvor utrolig mye nytt som foregår på feltet. En ting ble midlertidig klart i løpet av uken i Oxford: Alle seminarrekkene var tydelig tverrfaglige. Vellykkede prosjekter innen digital humaniora, er vellykkede nettopp fordi så ulike fagretninger anvender digitale verktøy i et samarbeid for å samle inn data og spre god humanistisk forskning. Feltet har i fremtiden en gyllen mulighet til å kommunisere og løfte frem humanistisk forskning ved hjelp av digitale verktøy. De digitale verktøyene bør derfor brukes som

nyttige hjelpemidler for å nå nettopp dette målet. Noe politikerne bør være klar, over er at avanserte digitale verktøy er et dårlig mål i seg selv. Fagmiljøene i humaniora, må på sin side ha en løpende diskusjon om hvordan man forholder seg til kvalitative data digitalt og hvordan bruk av digitale verktøy påvirker materialet. Samtidig må norsk academia være åpen for at digitale verktøy kan være nyttige for innsamling, bearbeiding og spredning av humanistisk forskning. Da kan vi gå den digitale verden i møte med et kritisk blikk.

Noter

¹ St. Ml. 25, *Humaniora i Norge* (2016-2017) 18

² Ibid.,

³ Ibid.,

⁴ <http://music-encoding.org/> .

⁵ Denne delen av musikkvitenskap forsker også på forholdet mellom matematikk og musikk ved hjelp av programvare som kan analysere musikalske verker i større detalj.

⁶ <https://eebo.chadwyck.com/marketing/about.htm>.

⁷ Kvart er et eldre trykkeformat for bøker og skuespill som ble brukt på denne tiden. Det er brettet to ganger og ligner på dagens A4 format.

⁸ Forskerens egen data og datasett sjekkes her ved hjelp av programmeringsspråket kjent som Wolfram language.