



Foto: © tilla eulenspiegel / photocase.com

HUMANISTEN I VÄRLDSVÄVEN

Sedan 1990-talet har humanisternas arbetsmiljöer och metoder förändrats mycket. Även de mest konservativa forskarna och skribenterna använder sig i dag, på några undantag när, åtminstone av textbehandlingsprogram och e-post i sitt arbete. För många är datorn och internet helt integrerade delar av arbetsprocesserna. De är så självklara att man sällan reflekterar över hur dessa ogenomskinliga system påverkar forskningen.

Text: Jessica Parland-von Essen

Ju mer lättanvända och skenbart enkla dessa redskap är, desto mer förrådiska är de. Att den för användaren osynliga tekniken i dag är så komplicerad att majoriteten av människorna inte alls kan ta den till sig, gör ändå inte att humanister kan låta bli att förhålla sig analytiskt till it-tekniken. Forskningens kvalitet kan bli mycket dålig om forskarna själva frånsäger sig ansvaret. Också om man själv vägrar använda dator överhuvudtaget, används den till exempel arkivpersonalen eller bibliotekarien som hjälper en.

Alla människor och all forskning är beroende av informationsteknologi vare sig man vill det eller inte. Nyligen bytte man till exempel söksystem vid Lunds Universitet, med katastrofala följder. I det fallet gjorde ett systembyte att forskare blev medvetna om kvalitetsskillnaderna mellan olika tekniska lösningar. Under andra omständigheter kunde de dåliga sökresultaten ha förblivit oupptäckta, eller föremål för diffusa och spridda klagomål, då enstaka kunder inte hittat material de vetat om att borde finnas. I själva verket presterar väldigt många söksystem mycket dåliga resultat. I och med att informationsmängderna ökar så dramatiskt blir relevans och precision allt viktigare. I Lund handlade det endast om referensinformation, om att man kanske helt missar annan forskning eller källor på grund av teknik. Men för humanister, som också arbetar med hela textkorporusar och stora fulltextmaterial, såsom Det historiska tidningsbiblioteket, blir det allt viktigare hur informationen är strukturerad. Framför allt måste forskaren veta hur informationen är strukturerad och hur det påverkar hela forskningsprocessen. I annat fall kan metoden inte granskas.

Varje beslut om hur informationen ska struktureras och presenteras är en tolkning. Mer semantiskt uppbyggda system är sannolikt den enda vägen att gå för att uppnå bättre relevans. I semantiska system skriver man in relationer och

en begrepps nivå kring ord i textmassan (se faktarutan). Samtidigt som detta radikalt förbättrar användbarheten, minskar överskådligheten, och källkritiken försvåras ytterligare för den enskilda forskaren inom humaniora. Det innebär nämligen att informationssystemet på sätt och vis kommer att "dra allt fler egna slutsatser" om hur saker och ting i världen förhåller sig till varandra, alltså i praktiken utsträcks datorns logiska kalkyler över en större del av forskningsprocessen då språket inkluderas. Det är nämligen mycket viktigt att komma i håg att datorn bara opererar med representanter (*signifiant*), och att de innebörder (*signifié*) som alltid är närvarande då människor behandlar betydelser, alltså tänker och kommunicerar sina tankar, inte alls finns med i processen. Det vore därför av största vikt att forskare inom humaniora skulle engagera sig i utvecklandet av digitala material, både för att de själva ska förstå dem bättre och för att materialen ska kunna utvecklas så att de befördrar forskningen i stället för att försvaga dess kvalitet.

Texten och språket är humanistens primära arbetsred-

I själva verket presterar väldigt många söksystem mycket dåliga resultat.

I och med att informationsmängderna ökar så dramatiskt blir relevans och precision allt viktigare.

skap. De utgör kärnan både i forskningen och i presentationen av forskningens resultat. Den förändring som skett i textens kvalitet i och med webbens genomslag är på sätt och vis mer genomgripande för humaniora än man vid ett första ögonkast kunde tänka sig. Det hänger samman med hur betydelser uppstår både på en semiotisk och på en social nivå. Den sociala nivån påverkar också forskaren inom humaniora på genomgripande sätt, både befriande och utmanande.

Den befriande aspekten gäller tillbakagången av normaliseringen eller standardiseringen både av språk och texter. Den snabbt växande digitala textmassan och den upplösta kontrollen över offentligt publicerad text, har gjort att språk och stilar faller sönder i otaliga varianter och genrer, som förändras mycket snabbt och lever helt utom språkgranskningens eller språkvårdens kontroll. Den skrivna texten har till sin karaktär närmast sig talspråket på många sätt.

Den snabbt växande digitala textmassan och den upplösta kontrollen över offentligt publicerad text, har gjort att språk och stilar faller sönder i otaliga varianter och genrer, som förändras mycket snabbt och lever helt utom språkgranskningens eller språkvårdens kontroll.

Eftersom hypertext, som tidigare tvingades in i fotnoter och marginaler, liksom olika varianter av texter, kan integreras smidigt i den elektroniska texten, är kritiska utgåvor i dag i regel elektroniska. Idéhistoriskt kunde man se att den s.k. gutenbergska parentesens tvingade fram, inte bara fenomen som normalisering av text och språk, utan även prioritering mellan textversioner och varianter. Texten, och skrivandet och läsandet som kommunikation, tvingades in i en modell där den tryckta texten blev statisk och auktoritativ. Men texten och den skriftliga kommunikationen är givetvis mer än så. Det är en kommunikativ process lika väl som den muntliga kommunikationen, en process fylld av tolkning, vaghet och ömsesidighet. Genom att webben förvandlat en stor del av det som tidigare varit muntlig kommunikation till skriftlig har detta blivit påtagligt och omöjligt att ignorera. Också den skrivna texten sönderfaller i allt fler olika tonfall och genrer. Skillnaden mellan original och kopia är utradad, texterna är flytande, föränderliga, klipps och klistras, länkas, försvinner och dyker upp igen i nya kontexter.

Här uppstår en för forskaren mycket utmanande situation. Där det tidigare fanns auktoritativa texter, fanns det också auktoriteter. Så länge publikationskanalerna var bristvara kunde de kontrolleras. Så är det inte längre. Expertens inom humaniora har ett arbetsredskap och en kompetens som långt består i text och språk. Var och en har tillgång till dem. Allt fler har också tillgång till våra källmaterial. I dag är publiceringsmonopolet och den självskrivna auktoriteten borta. Vad gör vi då? Hur kan man få genomslag, tyngd och respekt i detta överflöd av information av olika kvalitet?

Svaret är att gå in i dialog och målmedvetet bygga sig ett

anseende på webben. Att modigt diskutera och dela med sig av också mindre färdiga tankar och resultat. Fördelarna med att rapportera öppet om sitt arbete är många. Genom att inte vänta i årtal tills en monografi står färdig kan man både reservera ett ämne, och – vilket är ännu bättre – få respons av andra forskare av andra än den egna närmaste kretsen eller dem som råkar ta del av ens presentationer vid någon konferens i ett tidigt skede. Dessutom ger man andra möjlighet att

hänvisa till pågående forskning. Här måste man lita på kollegernas självkritik och kompetens att kunna tillmäta olika typer av så kallade gråa publikationer rätt betydelse.

Eftersom den traditionella publiceringen varit så långsam är fördelarna med större öppenhet under forskningsprocessen betydande för både de enskilda forskarna och vetenskapen som helhet. Mindre grundligt belagda påståenden kommer sannolikt ohjälpligt att krypa in i forskningen, men de kan också korrigeras av en större publik. Framtidens forskare måste idka en annorlunda och större kritik mot andra forskares texter, en sak som inte nödvändigtvis är bara dålig. En forskares eller en forskningsinstitutions anseende kommer att få allt större betydelse.

I den digitala världen sliter sig tecken och symboler från sina ursprungliga sammanhang och kan segla i väg i samtiden och i framtiden. Deras betydelse förändras på ett sätt som är tydligare än vid verbal kommunikation. Upphovsmannen har ingen möjlighet att kontrollera alla tolkningar som görs, nu och i framtiden. Humanisternas mångordiga kontextualiseringar, som kunde florera under den gutenbergska erans rätt statiska texter, måste

anpassa sig och omformas i den digitala världen, där texter styckas och återanvänds i ett hisnande tempo och där läsandet ofta är ytligt. Samtidigt ger det digitala mediet möjligheter till nya typer av kontextskapande i form av flexibla användning av metatext och länkning.

Digitala material kan kopieras och därvid förvrängas av tekniska orsaker eller av mänsklig hand. Också gamla texter, skapade före den digitala erans enkla och snabba kommunikation förflyttas i ökande takt in i den digitala världen. Endast professionella forskare kan ha riktigt goda insikter i vil-

FULLTEXTSÖKNING OCH ANNOTERING AV TEXT

För att text skall kunna vara fullt sökbar måste den vara maskinläsbar. Den kan antingen vara en text som är digitaliserad (skriven på dator) eller en tryckt text som skannats in och sedan lästs in med visuell igenkänning av bokstäver (OCR). Handskrivna källor måste tillsvidare skrivas in för hand.

Om man har sparat material som text kan man annotera det. Man märker då ut ord i texten, antingen gällande utseende (t ex "blyertspenna") eller innehåll (t ex "personnamn"). TEI-koderna är allmänt använda vid kritiska utgåvor. Själva texten blir då en xml-fil. Den som någon gång arbetat med enkel html vet idén i stort, principen är den samma. Ett exempel på en modern kritisk utgåva med annoterad text är den digitala utgåvan av Zacharias Topelius skrifter.

Språkforskare använder också ofta annotering för att kunna hitta i och analysera textkorpusar.

ken kontextuell information som kan gå förlorad och måste beaktas vid digitalisering och skapande av digitala resurser. I den digitala världen finns nästan obegränsade möjligheter till återanvändning och kreativitet.

Den humanistiska experten har kunskap av det slag som behövs för att förhindra att information går förlorad då allt mer inom forskning och kultur sker i den digitala sfären. Expertens ansvar består i att se till att korrekt information också finns att få, att tekniska lösningar erbjuder vad de ser

SEMANTISK KODNING OCH SEMANTISKA SÖKNINGAR

För många forskningsändamål är berikandet av text en stor hjälp och den gör det möjligt att få fram strukturer och samband man annars kanske aldrig kunde ha fått syn på. Allra bäst är då fulltextsökning i annoterad text. Då kan användaren söka inne i texterna och inte bara i metadata, alltså kataloger och register, som ibland kan dölja information snarare än hjälpa en rätt.

Semantiska sökningar innebär att man kodat dokument eller annoterat själva texten med också en begreppslig nivå. Till exempel att man faktiskt identifierat enskilda personer i en text så att söksystemet förstår att "Urho Kekkonen" på ett ställe är samma sak som "Presidenten" på ett annat ställe. I praktiken kodar man alltså in roller eller hur saker förhåller sig till varandra. Då mängden information hela tiden ökar är det allt viktigare att kunna skilja på texter skrivna av eller om en viss person, för att ta ett exempel.

I semantiskt kodade system använder man sig ofta av ontologier, som i detta sammanhang har en specifik betydelse. Det är en datauppsättning där begrepp och deras relationer finns sparade; en modell av hur begrepp förhåller sig till varandra. Genom att använda dylika lösningar kan man få en dator att förstå till exempel att då man söker saker som hör ihop med Åboland, avser man alla orter i landskapet och alla de olika historiska kommunnamn som kan finnas med i materialet, utan att man själv behöver komma ihåg eller ens känna till alla dessa namn.

Se exemplet Kulttuurisampo.

ut att göra och att man inte tappar bort information av ren och skär okunskap. Vi måste försöka förhindra att man gör dålig digitalisering som är dyr, eftersom den inte fyller de krav forskningen ställer. Därför måste vi engagera oss i den digitala kulturen och outtröttligt kommunicera om dessa frågor med kulturarvsinstitutioner, med varandra och med allmänheten. Vi måste försöka främja att framgångsrika förfaringssätt sprider sig och att man undviker onödigt arbete genom otillräcklig planering.



Jessica Parland-von Essen är doktor i historia och arkivchef vid Brages pressarkiv