

Johan Jarlbrink

Inledning – digitala metoder

I en tid när det mesta medieinnehåll föds digitalt och när historiska samlingar samtidigt digitaliseras skapas möjligheter att utforska medier med hjälp av digitala metoder. En metodologisk uppdatering kan innebära såväl förändrade frågor och upptäckten av nya studieobjekt, som digitala redskap att använda i forskningen (Jensen 2011). Datorernas förmåga att beräkna och sortera gör det möjligt att få fram frekvenser och samförekomster i stora materialsamlingar, hur texter och medieanvändare relaterar till varandra, och hur det enskilda förhåller sig till den större massan av data. Samtidigt behöver digitala analysstrategier inte nödvändigtvis vara kvantitativa. Som Rogers (2013) betonat kan digitala metoder även innebära att följa och utforska hur exempelvis en sökmotor organiserar data och villkorar tillgången till den. Internationellt finns ett växande intresse för digitala forskningspraktiker och ambitionen med detta temanummer är att presentera aktuella metoddiskussioner inom fältet.

Helt nya är metoderna inte – särskilt inte inom medieforskningen. Redan under 1960-talet användes datorer för att hitta kvantitativa mönster i exempelvis tidningstexter (Jarlbrink 2015). Och inom datavetenskap, informatik och språkvetenskap har det under flera decennier utvecklats metoder för att analysera texter och nätverk digitalt. Humanistiskt och samhällsvetenskapligt orienterade medie- och kommunikationsforskare har goda möjligheter att bygga vidare på denna forskning och använda de verktyg som redan finns. Samtidigt finns det förutsättningar att urskilja och

ställa andra typer av frågor än dem som ofta intresserat föregångarna inom andra discipliner (Burdick m.fl. 2012). Hur kan man exempelvis förstå digitala nätverk ur ett maktperspektiv? Och hur kan man förklara diskursiva förändringar i tiotusentals texter över tid? I detta ligger också en utmaning: Med så många forskningsmöjligheter till hands är det ibland svårt att komma på vilka frågor som är angelägna att ställa.

Att anamma en digital forskningspraktik får emellertid inte innebära att forskare bortser från de problem som digitala metoder kan föra med sig. Med tillgången till och möjligheten att undersöka stora datamängder kommer ofta föreställningen om att storskaligheten representerar en högre form av vetande (Boyd & Crawford, 2012). Som påpekats av bland andra Lagoze (2014) och Gitelman & Jackson (2013) bygger även stora datamängder på urval som kan snedvräta resultaten. Data är aldrig 'rå' utan har alltid skapats och samlats med specifika syften. Tillgången både till samtida flöden i sociala medier och till digitaliserade historiska samlingar är dessutom ofta reglerad och beroende av urvalsmekanismer som enskilda forskare har svårt att kontrollera och få insyn i. Det är heller inte säkert att innehållet i datamängderna är vad det utges för att vara – automatiskt genererade samlingar innehåller ofta 'rutten' data såsom felkategoriserade texter, spam och avbrutna länkar (Boellstorff 2013). Samtidigt som digitala metoder gör stora datamängder lätt hanterbara riskerar de att urskilja mönster som inte existerar. Av forskare krävs ett kritiskt förhållningssätt både till

hur data har tillkommit och samlats in och till de digitala verktygens sätt att bearbeta dem. Här utöver bör det finnas beredskap att kombinera kvantitativa metoder med kvalitativa.

Men icke desto mindre: En medievetenskap som vill utforska medielandskapet i all dess mångfald behöver behärska en bred metodarsenal – och här blir digitala metoder en allt viktigare del. Alla metoder bär spår av de medier som de utvecklats för att studera. Att överföra analoga metoder till digitala sammanhang kan därmed innebära att forskningen förbiser det digitalt mediespecifika (Hayles 2004). Kommersiella aktörer använder redan digitala verktyg för att få syn på trender, nätverk och beteenden. Medieforskningen behöver hitta sina egna angreppssätt och verktyg, men för att en sådan utveckling ska ta fart krävs att fler forskare lär sig hantera och modifiera de digitala metoder som redan finns.

Artiklarna som samlats i detta nummer behandlar och utvecklar flera av de teman som antytts ovan. Några presenterar fallstudier där de digitala metodernas möjligheter och problem illustreras, andra diskuterar digitala metoder mer principiellt. Två av studierna beskriver hur digitala metoder för textanalys kan användas för att undersöka digitaliserat historiskt material, medan två studier berör metodfrågor i anslutning till flöden i sociala medier.

Fredrik Norén undersöker hur begreppen 'information' och 'sambandsinformation' blir en del av den svenska efterkrigstidens politiska diskussioner. Utgångspunkten är ett stort antal utredningar, utgivna i serien Statens Offentliga Utredningar (SOU, 1922-), som nyligen gjorts digitalt tillgängliga. Norén inleder sin studie med att ta reda på hur frekvenserna av ordet 'information' förändras från 1920-talet till 1990-talet, och undersöker därefter hur ordet användes och vilka politikområden det blev en del av, från 1960-tal till 1980-tal. Med hjälp av verktyget MALLETT visar Norén hur samhällsproblem under den undersökta perioden börjar beskrivas som informationsproblem, samtidigt som sambandsinformationen blir viktig inom flera av statens verksamhetsområden.

Johan Jarlbrink, Pelle Snickars och Cristian Colliander beskriver hur digitala verktyg för textanalys kan användas i studier av äldre tidningspress. I artikeln tecknar de den mediehistoriska bakgrunden till bibliotekens pågående digitaliseringsprojekt och beskriver den omvandling som historiska tidningar genomgår när de digitaliseras. När det pappersbaserade kulturarvet har OCR-tolkats finns nya möjligheter att hitta mönster i de texter som tidigare har varit svåra att söka i och utforska. Digitaliseringen av äldre tidningspress är emellertid förbunden med flera problem: kvaliteten på OCR-läsningen är ofta bristfällig, samtidigt som stora delar av materialet är fortsatt otillgängligt på grund av upphovsrättsskäl. För att illustrera de möjligheter som ändå finns presenterar författarna en fallstudie om hur det nya mediet Internet presenterades för 1990-talets tidningsläsare.

Jonas Andersson Schwarz och Johan Hammarlund behandlar i sin artikel de mediala och sociala kontexter som digitala analyser av twitterdata både kan framhäva och riskerar att förvränga och osynliggöra. Utifrån ett antal fallstudier av twitterdebatter diskuterar de hur digitala metoder kan användas för att analysera användare och nätverken de bildar, både inom forskningen och inom omvärldsbevakningsbranschen. Kvalitativa analyser riskerar här ofta att framhäva enskilda exempel som inte nödvändigtvis är representativa för den större datamängden. Kvantitativa analyser riskerar å andra sidan att föra samman olika användare på ett sätt som skapar sammanhang som egentligen inte finns. För att hantera kontexter på ett rimligt sätt krävs att forskare kombinerar kvantitativa metoder med en kvalitativ lyhördhet för materialets egenart.

Pernilla Seversons artikel behandlar hur användargenererat material på internet kan förstås och undersökas utifrån tre forskningsperspektiv och referenspunkter: framhävandet av publikens njutning inom kulturstudieforskningen, designforskningens persona-begrepp, samt medieinnehållet betraktat som data, så som det framträder i ett datavetenskapligt perspektiv. Med publikens njutning i åtanke framstår det som intressant att studera medie-

användarnas egna drivkrafter, samtidigt som det med designforskningens betoning av medieplattformarnas gränssnitt och struktur är viktigt att undersöka hur användargenererat material villkoras av de tekniker som samtidigt utgör dess förutsättning. Med medieinnehåll i digitalt format blir det möjligt att samla in och analysera det med hjälp av digitala verktyg, men Severson betonar samtidigt att detta kan

föra med sig såväl etiska som analytiska problem. Kvantitativa metoder bör därför kombineras med kvalitativa.

Härutöver innehåller temanumret presentationer av digitala forskningsmiljöer, en recension av Richard Rogers bok *Digital Methods*, samt en översikt över nyutkommen litteratur inom området.

Referenser

- Boellstorff, Tom (2013). Making Big Data, in Theory. *First Monday*, 18(10): 1-17.
- boyd, danah & Crawford, Kate (2012). Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information, Communication & Society*, 15(5): 662-679.
- Burdick, Anne, Drucker, Johanna, Lunenfeld, Peter, Presner, Todd & Schnapp, Jeffrey (2012). *Digital Humanities*. Cambridge & London: MIT-Press.
- Gitelman, Lisa & Jackson, Virginia (2013). Introduction, pp. 1-14 in Gitelman, Lisa (ed.) *"Raw Data" Is an Oxymoron*. Cambridge & London: MIT-Press.
- Hayles, Katherine N. (2004). Print Is Flat, Code Is Deep: The Importance of Media-Specific Analysis. *Poetics Today*, 25(1): 67-90.
- Jarlbrink, Johan (2015). Historievetenskapens mediehantering, pp. 225-247 in Hyvönen, Mats; Snickars, Pelle & Vesterlund, Per (eds.) *Massmedieproblem: Mediestudiets formering*. Lund: Mediehistoriskt arkiv.
- Jensen, Klaus Bruhn (2011). New Media, Old Methods: Internet Methodologies and the Online/Offline Divide, pp. 43-58 in Consalvo, Mia & Ess, Charles (eds.) *The Handbook of Internet Studies Edited*. Chichester: Blackwell Publishing.
- Lagoze, Carl (2014). Big Data, Data Integrity, and the Fracturing of the Control Zone. *Big Data & Society* 1(2): 1-11.
- Rogers, Richard (2013). *Digital Methods*. Cambridge & London, MIT-Press.

Johan Jarlbrink är docent i mediehistoria och lektor i medie- och kommunikationsvetenskap vid Umeå universitet. Han forskar om 1800-talets tidningspress i alla dess skepnader: tidningen som informationsbärare i arkiv och klippböcker, som avfall och omslagspapper, som forskningsdata i digitala databaser.