

Fra 200.000 tweets til klassisk humanistisk nærlæsning

□ **Max Odsbjerg Pedersen**, informations-specialist, AU Library
Helle Strandgaard Jensen, Lektor, Historie og Klassiske Studier, Aarhus Universitet.

I 2020 kombinerede vi vores fælles interesse for digital humaniora i projektet *Sesame Street at 50: The Shapes of Digital Memory*, der skulle anvende computerassisterede metoder på data fra Twitter (nu X, *red.*). Med udgangspunkt i Helle Strandgaard Jensens daværende bogprojekt om det mere end 50 år gamle amerikanske børne-tv program *Sesame Streets'* transnationale historie ønskede vi at undersøge den erindringskultur, som knytter sig til programmet på det sociale medie.

De computerassisterede metoder, vi benyttede, var relativt simple, hvilket var et bevidst valg. Vi ønskede at demonstrere, hvordan meget simple statistiske metoder kan hjælpe humanister med at fokusere deres kvalitative nærlæsning, når de forsker i omfattende strukturerede data. Denne tilgang kaldte vi 'skalbar læsning', og projektet resulterede i udgivelsen af en artikel i tidsskriftet *Memory Studies* og en tutorial på hjemmesiden programminghistorian.com.

De to historiske begivenheder, vi valgte at trække data ud på baggrund af, var da Sesame Street skiftede fra public service-kanalen PBS til det kommercielle HBO i 2015, og da programmet fyldte 50 år i USA i 2019. For

hver af de to begivenheder trak vi alle tweets ud 14 dage før og efter, der enten indeholdt "#sesamestreet" eller "sesame street" (uden hensyn til store og små bogstaver). Udtrækene blev udført gennem Twitters såkaldte API (Application Programming Interface).

Da vi startede projektet, havde vi adgang via en betalings-API, der tillod os at trække en mængde data per måned. Senere blev denne adgang gjort gratis for forskere, for så igen at blive en betalingstjeneste. I samme periode ændrede Twitter navn til X. Valget af Twitters API-løsning var afgørende, da de fleste sociale medieplatforme ikke tillader direkte adgang til deres tjenester som ren data. Derfor valgte vi Twitter, hvor vi kunne få adgang til data lovligt. Data fra API'et blev leveret i JSON-format, hvilket krævede en betydelig mængde tid at transformere til et format, der kunne arbejdes med i statistiksproget R.

En digital bro til erindringsstudier

Studier, der har beskæftiget sig med Twitter-data inden for erindringsstudier, har primært indsamlet data gennem hashtag-søgninger på Twitters hjemmeside, hvorefter udvalgte tweets er blevet analyseret. Vi ønskede at imødegå denne praksis ved at udtrække tweets, der indeholdt "#sesamestreet" og/eller "sesame street", for at undersøge i hvilken grad man gik glip af materiale ved kun at fokusere på hashtag-søgninger – den mest

udbredte metode blandt erindringsforskere. Undersøgelsen viste, at kun 9,9 procent af tweetsene i datasættet fra 2015 og 7,9 procent fra datasættet fra 2019 indeholdt hashtagget "#sesamestreet". Dette indikerer, at ved at fokusere kun på hashtags risikerer man at gå glip af en betydelig mængde materiale.

Nogle af de mange metadatafelter fra Twitters API er antallet af *likes*, svar og *retweets* et enkelt tweet har. Disse tal indikerer, hvor meget interaktion der har været med hvert tweet. På baggrund heraf kunne vi relativt nemt i R udtrække de mest populære tweets målt på disse parametre, hvilket gav os en kvantitativ begrundelse for, hvilke tweets der skulle udvælges til nærlæsning. Samlet har vi kaldt kombinationen af fjernlæsning af de 200.000 tweets og den nære læsning af de mest populære tweets 'skalerbar læsning.'

Vores metode til udvælgelse af de mest populære tweets omkring de to begivenheder resulterede i en selektion af 249 tweets, som havde høj interaktion. 129 af disse stammer fra datasættet fra 2015, og af disse omhandlede 87 skiftet fra PBS til HBO. Langt størstedelen af disse 87 tweets gjorde grin med skiftet på baggrund af HBO's store katalog af voldelige serier. F.eks. med spørgsmål om, hvorvidt Sesame Streets karakter skulle dræbes i samme hastighed, som det var tilfældet i *Game of Thrones*.

Nogle tweets var kritiske, især dem fra ikke-verificerede brugere, men langt de fleste af de 249 udvalgte tweets på tværs af de to begivenheder kom fra verificerede brugere. Verifikationen af brugere på Twitter var dengang en garanti for, at kontoen var autentisk og blev typisk tildelt sportsstjerner, politikere, kulturpersonligheder osv. Projektet kunne bidrage til diskussioner om, hvem der er gatekeepers i erindringspraksisser på sociale medier, men understregede samtidig behovet for yderligere analyser af, hvordan Sesame Street spiller en stor rolle i amerikansk erindringskultur.

Nettets foranderlighed

Twitter, som vi beskæftigede os med i 2020, har ændret sig betydeligt siden da. Salget til Elon Musk medførte, at mediet skiftede navn til X, og mange brugere forlod mediet. Derudover ændrede verifikationsprocessen sig også, så man nu skulle være betalende bruger for at opnå eller opretholde sin verificerede status. Dette giver anledning til refleksioner over, hvordan man fremover griber lignende projekter an. For eksempel vil man skulle tage højde for platformens historie samtidig med, at man forholder sig til problemstillingen. Med andre ord, hvilken generation af Twitter foregår din analyse i, og hvilken konsekvens har

det? I vores projekts tilfælde ville en analyse af nutidens gatekeepers i erindringspraksisser på Twitter/X skulle forholde sig til to vidt forskellige typer af verificerede brugere.

I projektet opnåede vi betydelige fremskridt ved at kombinere kompetencer og samarbejde på tværs. Domæneviden om Sesame Streets' historie og erindringsstudier skabte grundlaget for at udtrække et datasæt, der kunne bruges til at stille relevante spørgsmål. Vores fællesinteresse og forståelse af problemstillinger inden for digital humaniora skabte et fælles sprog, der gjorde samarbejdet frugtbar, uanset om man var ekspert i Sesame Streets' transnationale historie og erindringsstudier eller havde ekspertise i at trække data ud fra API'er og transformere det.

Selvom projektets resultater kun er baseret på en enkelt case og understreger behovet for yderligere lignende undersøgelser, har det bidraget med viden om, hvordan Sesame Street konkret blev husket. Derudover beskæftiger det sig med gatekeepernes rolle inden for erindringspraksis på sociale medier, som er en af de centrale diskussioner inden for erindringsstudier.

Læs mere

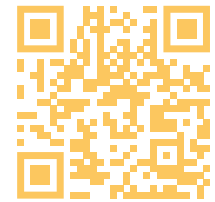


Læs om projektets resultater i tidsskriftet

Memory Studies



Læs om 'skalerbar læsning' som metode på programminghistorian.com.



i *Sesame Street*

Sesame Street er en amerikansk tv-serie for børn med dukker, såkaldte muppets, i hovedrollerne. Den skal ikke forveksles med *The Muppet Show* – dog er figuren Kermit med i begge serier. Sesame Street havde premiere i 1969 i USA og blev sendt på public service netværket PBS, inden den skiftede kanal til det kommercielle HBO i 2015. I Danmark tog man også serien til sig, der havde premiere som Sesamgade på TV2 i 2009.

Projektgruppen til Sesame Street at 50: The Shapes of Digital Memory bestod af lektor i Historie og Klassiske studier Helle Strandgaard Jensen og Max Odsbjerg Pedersen, informations-specialist AU Library. Derudover Josephine Møller Jensen og Victor Harbo Johnston, begge studerende på AU, samt ph.d-studerende på AU Alexander Ulrich Thygesen. Projektet er finansieret af interne midler på Aarhus Universitet samt med ressourcer fra AU Library.

“ I projektet opnåede vi betydelige fremskridt ved at kombinere kompetencer og samarbejde på tværs. Domæneviden om Sesame Streets’ historie og erindringsstudier skabte grundlaget for at udtrække et datasæt, der kunne bruges til at stille relevante spørgsmål.

↓ Foto: Karsten Paulick, Pixabay. Sesame Street er en amerikansk tv-serie for børn, som blev sendt første gang i 1969

