
INTERNETÖVERVAKNING SOM METOD

Vårt digitala samhälle genererar mängder av data, men drar dagens samhällsforskare full nytta av all ny data? Den franske sociologen **Gabriel de Tarde** hade gärna hade dykt ner i den nya datafloden, kan dagens forskare lära sig något av hans ansats? **Christopher Kullenberg** undersöker digital övervakning inom samhällsforskning.

Text: Christopher Kullenberg

Varje gång du går och handlar en liter mjölk registreras, arkiveras och kalkyleras stora mängder data. I fickan avläser din mobiltelefon din position med gps samtidigt som den sparar en lista på vem du ringer på väg till butiken. Dessa platsdata sparas även av din teleoperatör tack vare att en finurligt kamouflerad mobiltelefonmast hela tiden lagrar uppgifterna i en gigantisk databas. När du sedan betalar mjölken registrerar din bank och din kreditkortsleverantör beloppet och tidpunkten. Har du ett medlemskort sparas även vilka produkter du köpte. Samtidigt sänder kassaapparaten en signal till en dator som beräknar hur många liter mjölk butiken kommer att behöva köpa in för att möta ett beräknat behov i morgon.

Något så enkelt som en liter mjölk blir – utan att vi tänker på det – föremål för en hel kedja av informationsbearbetning, som i huvudsak används av företagen själva, eller, i vissa fall, av polismakten om ett brott har begåtts. Men kan denna information även vara värdefull för samhällsforskare?

Det senaste decenniet har internet genomgått två stora förändringar. För det första har ett fåtal plattformar för sociala medier, utvecklade framförallt av Facebook, Google och Apple, kommit att dominera hur vi navigerar, interagerar och organiserar våra digitaliserade liv. För det andra har de så kallade smarttelefonerna kommit att förändra hur, när och på vilket sätt som vi ansluter till internet och vilken typ av data vi delar med andra. Konsekvenserna av dessa två förändringar är svåra att överblicka, därför tänker jag i denna artikel koncentrera mig på två specifika aspekter, nämligen vilken typ av *data* som våra uppkopplade liv genererar och hur dessa enorma datamängder kan användas i samhällsvetenskaplig forskning.

Datafloden

Om vi tänker oss det inledande exemplet, och datamängden som automatiskt genererades bara av att man gick och köpte en liter mjölk, går det att föreställa sig de enorma kvantiteterna av data som skapas varje sekund i en globaliserad och uppkopplad värld. Denna oerhörda informationsmassa kallas ibland *the data deluge*, eller datafloden, och betecknar ett tillstånd där så mycket information genereras att det blir oöverskådligt för en enskild institution eller ett enskilt system att skapa meningsfulla resultat och slutsatser ur den aldrig sinande strömmen av ettor och nollor. Varje liten transaktion, varje samtal, varje klick på internet och varje gång vi åker buss eller spårvagn bidrar

”Dessa ändlösa detaljer betydde inget som isolerade enheter. Det är först, menade Tarde, när man ser dem som imitativa beteendemönster som de blir intressanta.”

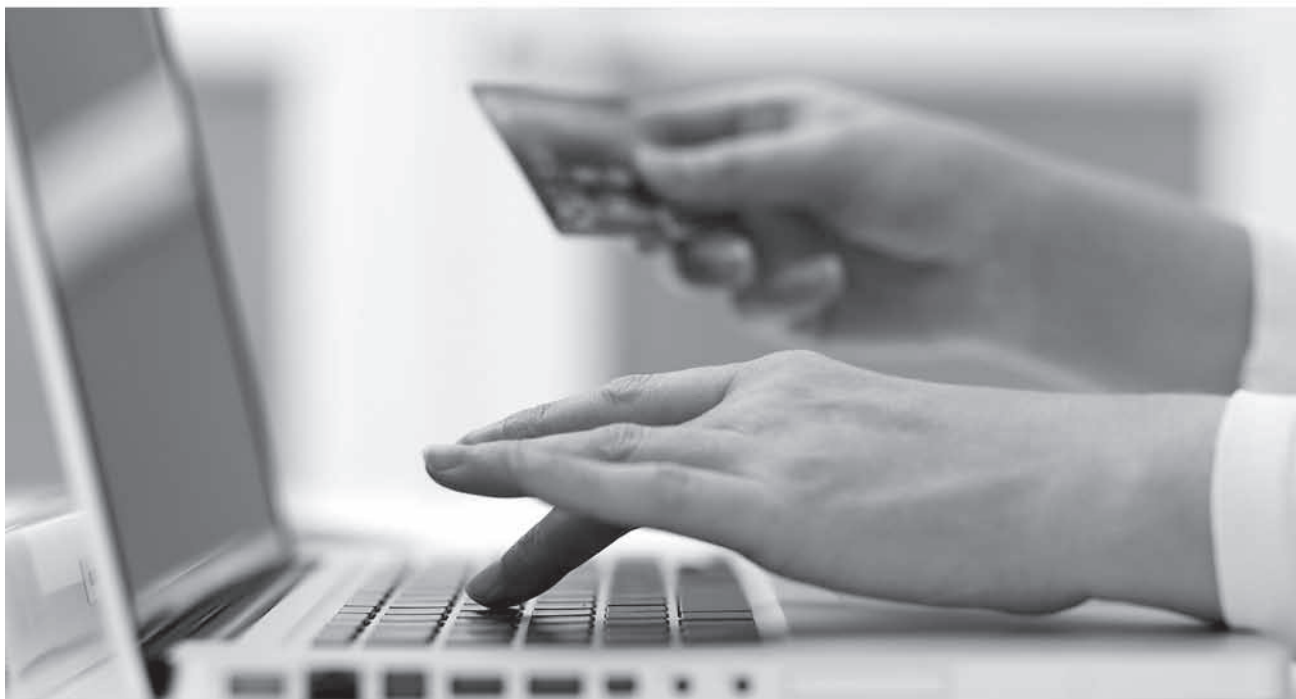
vi till datafloden, som förvisso inte driver omkring på ett öppet hav, utan har sin fysiska form i serverhallar runt om i världen. För de stora aktörerna måste dessa ständigt byggas ut, gärna i länder med kallt klimat och billig elektricitet, exempelvis Facebooks stora serverhall i Luleå.

Datafloden medför två saker. För det första skapas massvis med data som aldrig kommer att användas. Endast någon promille av mobilmasternas positionsdata kommer att

användas i en polisutredning, mängder av uppdateringar på Facebook och Twitter förfaller sakta mot en långsam radering och många bloggar och hemsidor försvinner efter några år av inaktivitet. Att bevara data kostar i slutänden pengar och på så sätt kommer datafloden att torka ut samtidigt som den fylls på.

För det andra tvingar datafloden fram ett nytt sätt att tänka om data och stora informationsmängder. Eftersom mängderna av information inte kan begripas i sin helhet, varken

dag för dag, med exakta tidsstämplar och positioner. Arkeologerna skulle med ett sådant fynd i sina händer kunna arbeta nästan som sociologer. Databasen innehåller inte bara en mängd beskrivande data om varje individ. Framförallt innehåller den en massa interaktionsdata, det vill säga information om vem som pratar med vem, vilka som är vänner, vilka som har gått på samma event, vilka som gillar samma klädmärke eller musik, vilka som har befunnit sig på samma platser samtidigt. Framtidens arkeologer kommer inte att behöva dechifrera inskriptioner på kruk-



av människa eller av maskin, måste de göras vetbara med hjälp av algoritmer och visuella behandlingar. Med en algoritm kan man definiera regler för hur data ska sorteras, behandlas och förändras. När väl inställningarna är gjorda, kan algoritmerna arbeta inuti datorprogram som aldrig vilar, och på så sätt kan man skapa meningsfulla mönster som blir lättare att tolka. Den kanske mest kända algoritmen heter *Page-Rank* och utgör grunden i det system som Google använder sig av i sin sökmotor för att sortera fram de viktigaste sökresultaten ur den ständigt växande världsväven.

Övervakning som samhällsvetenskap

Vad vore det mest intressanta arkeologiska fyndet från 00-talet, om låt säga hundra eller tusen år? *Burj Khalifa*, *Ground Zero* eller en nedgrävd 40-tums platt-TV? Inget av dessa fynd säger så mycket om hur människor levde under det första decenniet av 2000-talet som en kopia av Facebooks databas. I denna gigantiska fil, som om hundra år troligtvis skulle rymmas i våra fickor, har en femtedel av jordens befolkning registrerat vad de gör och tänker på –

skärvor för att få fram ett fåtal ledtrådar om det förflutna. Snarare kommer problemet att vara omvänt; den sederade datafloden kommer att innehålla en miljard människors digitala fotavtryck och det primära arbetet kommer att bestå av att hitta det generella i det partikulära, det relevanta i alla upprepningar av detsamma.

För ett sekel sedan spekulerade den franske sociologen **Gabriel de Tarde** i verket *Les Lois de l'imitation (Imitationens lagar, 1895)* vilken form av statistik som behövdes för att man på allvar skulle förstå sociala beteenden. Han var mycket missnöjd med den förenklade statistik över födslar, brott, och dödsorsaker som fanns tillgänglig vid denna tid. Istället föreställde sig Tarde att människor, ur samhällsvetenskapernas synvinkel, optimalt sett levde i "glashus" så att sociologerna kunde se rakt in i varje hem, ned i varje byrålåda och in i varje garderob. På så sätt kunde man föra statistik över varje liten detalj av människors liv och kartlägga vilka färger på strumpor som var populära, vilken sorts mjöl som fanns i påsarna i skafferiet och, inte minst,

vilka böcker som stod i bokhyllorna. Dessa ändlösa detaljer betydde inget som isolerade enheter. Det är först, menade Tarde, när man ser dem som imitativa beteendemönster som de blir intressanta. Hur kommer det sig att människor helt plötsligt börjar läsa *Fifty Shades of Grey*, köper selfiepinnar till sina mobiltelefoner och börjar baka surdegsbröd enligt gamla recept? Tarde tänkte inte att det fanns någon övergripande (eller undermedveten) struktur som stakade ut riktningen för dessa beteenden. Massbeteenden strålar istället genom samhällskroppen, som samtidigt konsti-

Inget säger så mycket om hur människor levde under det första decenniet av 2000-talet som en kopia av Facebooks databas.

tueras av deras multiplicerande. Varje gång vi imiterar ett beteende skapar vi en liten byggsten, och när tillräckligt många imiterar samma ursprungliga innovationer, uppstår en samhällelig kontinuitet.

Tarde fick aldrig sitt glashus. Det skulle vara oändligt dyrt att inventera varje franskt hem kring sekelskiftet 1800/1900 i syfte att föra statistik över alla livets detaljer. Men om vi återvänder till samtiden och datafloden kan man undra om det inte är just ett sådant glashus vi har byggt på Internet. Facebook och Google tjänar pengar på precis denna transparens; att de vet vilken färg på strumporna du vill ha, om du har googlat efter selfiepinnar eller läst bloggar som innehåller recept på bröd. Alla dessa detaljer vill de veta för att kunna optimera sin marknadsföring och sälja informationen vidare. Datafloden ter sig som en aldrig sinande källa till värde, just på grund av att vi fotograferar vår frukost, taggar våra vänner i bilderna vi laddar upp och ser på Youtube-filmer om hur man renoverar sitt kök. Varje gång vi utför dessa beteenden skapas en (digital) imitativ

stråle som fångas upp och därmed utgör en potentiell profit för nätjättarna. Transparensen är gränslös, samtidigt som den bara framträder fragmentariskt för oss vanliga användare. De stora mönstren räknas fram av algoritmer i de stora datorhallarna nära polcirkeln, nära de norrländska älvorna som med sina vattenkraftverk levererar grön el samtidigt som lokalpolitikerna som levererar skattelättnader. Ingen vill hamna efter i digitaliseringen för den för med sig ett löfte om en ny ultrasnabb ekonomi och en ljus framtid.

Samhällsforskningen ligger tyvärr långt efter nätjättarna. På samma sätt som survey-metoderna utvecklades av privata företag, exempelvis Gallup, under första halvan av nittonhundratalet, sker dagens metodutveckling i Silicon Valley snarare än i Oxford eller Cambridge. Trots att datafloden fortsätter att tillta i styrka är det väldigt få samhällsvetare som har börjat mejsla fram intressanta analyser ur detta material. Detta beror troligtvis på att samhällsvetare ännu inte fullt ut har sett värdet av att samarbeta långt över disciplingränserna, för att på så sätt nå fram till den nödvändiga kompetens som programmerare har och gemensamt tillverka mjukvara och algoritmer som gör datafloden mätbar och vetbar. Men det beror även på att vi ännu inte har funnit de teoretiska och metodologiska redskap som krävs för att kvantifiera det där tillfället när vi går och köper en liter mjölk eller söker information. Men det är nog bara en tidsfråga innan samhällsvetenskaperna har hunnit ifatt. Lockelsen att se rakt igenom glashusets väggar – in i människors liv – är alldeles för stor för att vi ska kunna fortsätta blunda.

Vidare läsning

Gabriel de Tardes *Les Lois de l'imitation* (1895) finns översatt till engelska som *The Laws of Imitation* (1903) och kan laddas ned fritt från archive.org



Christopher Kullenberg är doktor och forskare i vetenskapsteori vid Göteborgs universitet.