

*Kim Holmberg\**

## Webometrisk nätverksanalys

“Oh, what a tangled web we weave...” skrev Sir Walter Scott år 1808. Citatet är idag kanske mer passande än någonsin tidigare. World wide web eller webben har vuxit till ett enormt nätverk av webbsajter och hyperlänkar som kopplar ihop dem. Genom att skapa nya webbsajter och nya hyperlänkar skapar man också nya strukturer och kopplingar på webben. Det finns ingen kvalitetskontroll eller några regler för hur man borde eller hur man får skapa hyperlänkar mellan olika webbsajter, utan vem som helst kan skapa vilka hyperlänkar som helst. På grund av detta skulle man kunna tro att det råder främst ett kaos på webben. Webometri är ett forskningsområde som försöker hitta mönster och ny kunskap i detta kaos av hyperlänkar och innehåll på webben.

Webometrin har sina rötter i informations- och biblioteksvetenskapen, och närmare bestämt i bibliometrin, vilket är det forskningsområde som analyserar vissa kvantitativa egenskaper av böcker och andra publikationer. Bibliometriska metoder har bland annat använts för att kartlägga samarbetsmönster mellan forskare som publicerat tillsammans och i citeringsanalys för att kartlägga hur olika vetenskapsområden är relaterade och kopplade till varandra. När man använder bibliometriska metoder i webbkontext ser man på hyperlänkarna tekniskt som analogier för citeringar och referenser i vetenskapliga artiklar. Hyperlänkarna skapar kopplingar mellan olika webbdokument på samma sätt som citeringar och referenser kopplar ihop vetenskapliga artiklar. När dessa länkar har blivit skapade på grund av några så kallade ”offline” orsaker så kan man med webometrisk metod analysera data på webben för att studera ”offline” fenomen. Webometri har tidigare använts framgångsrikt till att bland annat mäta kvalitet hos sökmaskiner på

webben, kartlägga organisationers synlighet på webben, förhållanden mellan universitet, kopplingar mellan olika länder, och konkurrens mellan företag. I och med den avhandling som här idag skall granskas, har webometrin även använts för att kartlägga samarbete och geografisk likhet mellan kommunerna.

Det finns också ett annat sätt att närma sig saken. I hyperlänk-nätverksanalys använder man metoder från nätverksanalys för att studera webben och de nätverk som länkarna formar. Nätverksanalys är egentligen en benämning på en samling metoder som kan användas för att analysera olika egenskaper hos nätverk och de komponenter som bildar nätverken. Egenskaper som man vanligtvis analyserar är bland annat olika noders positioner i nätverk och deras förhållanden till varandra. Centralitet och klusteranalys är också vanliga metoder i nätverksanalys. Metoder för nätverksanalys används i flera olika forskningsområden och beroende på vad man vill undersöka så kan det handla om att kartlägga och mäta förhållanden och egenskaper hos nätverk av t.ex. människor, organisationer, webbsajter, datorer eller neuroner, som är sammanlänkade och skapar ett nätverk där till exempel information, trafik eller kommunikation överförs.

Hyperlänk-nätverksanalys har vanligtvis studerat hur existerande teorier och metoder från social nätverksanalys kan användas i forskning på webben och man har oftast fokuserat på nätverksstrukturer och enskilda noder i nätverk. Det informationsvetenskapliga sättet att studera webben, webometrin, har däremot fokuserat främst på att utveckla metoder för både datainsamling och tolkning av data, samt att utveckla metoder för att validera data genom innehållsanalys och klassificering av länkar för att skapa en djupare förståelse för länkar och orsaker för att skapa länkar. Både webometrin och hyperlänk-nätverksanalysen utgår ifrån att länkarna tyder på vissa handlingar hos dem som skapat länkarna eller förhållanden mellan de människor eller organisationer som länkarna kopplar till varandra. Hyperlänkarna är vanligtvis skapade för att uppfylla en uppgift och därför kan

---

\*Kim Holmbergs doktorsavhandling *Webometric Network Analysis – Mapping Cooperation and Geopolitical Connections between Local Government Administration on the Web* granskades vid Åbo Akademi den 4.12.2009.

de ibland avspegla ”offline” fenomen. Under de senaste åren har dessa båda forskningsområden lånat metoder av varandra och de har närmat sig varandra både praktiskt och vetenskapligt.

Den avhandling som här idag skall granskas har tagit det bästa av de båda världarna och kombinerat metoder från både webometri och hyperlänk-nätverksanalys samt utvecklat nya metoder som kan användas i webometriska studier. Med detta har denna avhandling fört de två disciplinerna närmare varandra och visat hur de kan dra nytta av varandra. Denna avhandling är även ett led i en lång tradition av webometrisk forskning som utvecklat metoder för att samla och analysera data från webben. Med detta har avhandlingen dels bidragit med ökad kunskap om hur kommunerna använder webblänkar och hur de länkas till från andra webbsajter, och dels med ökad förståelse för orsakerna till varför hyperlänkar skapas till och från kommunernas webbsajter.

Hyperlänkar är inte en perfekt informationskälla, eftersom en stor del av länkarna kan vara skapade för att fungera som navigeringslänkar, och resten kan vara skapade av positiva, negativa eller neutrala orsaker. För att kunna säga något om de upptäckta mönstren och tendenserna som man kan upptäcka i webometrisk forskning, måste man veta något om varför länkarna har skapats. I den aktuella forskningen upptäcktes att länkarna mellan kommunerna var skapade p.g.a. så kallade ”officiella orsaker” och att de reflekterade kommunernas uppgifter och plikter. Tack vare detta kunde man säga att kommunernas länkar kunde användas för att undersöka förhållanden mellan kommunerna i verkligheten.

Webben är en enorm källa av lätt tillgänglig information och detta är också den största potentialen med webometrisk forskning som använder data insamlat från webben. Data från webben, som t.ex. hyperlänkar, kan insamlas automatiskt. Hela forskningsmaterialet finns gratis tillgängligt på webben. Och material finns t.o.m. i överskott.

Webben växer, utvecklas och förändras hela tiden och därför bör även forskning om webben utvecklas och förändras. Webben spelar en allt viktigare roll i vår vardag. Den är en viktig informationskälla och en plats för både arbete och fritid. Vi använder webben för kommunikation och för att umgås med andra människor. Mängden av material som vi publicerar på webben är enorm. Vi laddar upp 750 bilder till Facebook varje sekund, vilket gör ca 65 miljoner bilder varje

dag. Vi laddar upp nästan 15 000 timmar video till YouTube varje dag och tjänsten Technorati hittar 10 nya blogginlägg varje sekund eller ungefär en miljon inlägg varje dag. På Twitter skickar vi ca 315 meddelanden varje sekund eller ca 27 miljoner meddelanden varje dag. För att hitta något relevant i denna enorma mängd av innehåll använder vi sökmaskiner och vi gör faktiskt ca 4000 sökningar varje sekund eller 340 miljoner sökningar varje dag i Google. Och det här är endast några av de tjänster vi använder på webben.

Det vi gör på webben avspeglar ofta våra intressen utanför webben. Våra aktiviteter på webben berättar om vem vi är och vad vi gör och därför utgör webben ett intressant och viktigt forskningsobjekt. Det vi gör i verkligheten påverkar webben och det vi gör på webben påverkar våra liv på ett eller annat sätt, och eftersom länkarna ibland är skapade av ”offline” orsaker, kan man använda data på webben för att analysera ”offline” fenomen .

I den aktuella forskningen kom jag fram till att förhållanden mellan kommunerna, såsom samarbete, avspeglas på webben och man kan använda data insamlat från webben för att analysera ”offline” fenomen . Allt vi gör på webben lämnar spår och dessa spår kan användas för att kartlägga våra intressen och vårt beteende på webben. Sökmaskinerna på webben vet exakt vad vi är intresserade av och kan t.o.m. förutspå kommande Idols- eller Big Brother-vinnare. Våra blogginlägg kan användas för att analysera trender och våra åsikter om t.ex. stamcells forskning eller strålning från mobiltelefoner. Hyperlänkar som vi skapar mellan olika sajter kan berätta till vilka andra människor eller organisationer vi har någon koppling och till hurdana nätverk vi eller organisationen vi representerar hör. Det vi kollektivt gör på webben kan användas för att analysera trender, åsikter, nätverk och intressen.

Det är också märkligt att vi idag öppet själva på webben berättar om sådana saker om våra privatliv som vi för tio år sedan inte skulle ha berättat ens för våra grannar. Vad vi gör på webben ger också upphov till vissa hotbilder, som risk för identitetsstöld och ryktesspridning. Därför är det också viktigt att förstå varför vi gör som vi gör på webben och hurdana konsekvenser det kan ha.

Under de senaste åren har webometrisk forskning även undersökt de s.k. sociala medierna på webben och människors beteende i dem. En rätt så aktuell definition av webometri inkluderar faktiskt möjligheten att

med webometriska metoder analysera sociala medier. Webmetrin har nyligen definierats som ”forskningsområdet som analyserar innehåll på webben med huvudsakligen kvantitativa metoder för mål i sociala vetenskaper genom att använda tekniker och metoder som inte hör till endast en specifik disciplin”. Enligt denna definition handlar webometri om att inom sociala vetenskaper studera webben med de metoder som är lämpliga

just för den aktuella forskningen och som inte nödvändigtvis behöver begränsa sig till t.ex. informationsvetenskaper eller nätverksanalys. Detta öppnar också nya möjligheter för webometrisk forskning och tack vare den enorma mängd av innehåll som vi varje dag skapar på webben ser jag också att webmetrin allt mera kommer att användas för forskning om mänskors beteende på webben, på den sociala webben.