

Web 2.0 muuttaa tutkimuksen näkökulmia – NORLIS-tutkijaverkoston seminaari Tallinnassa

Päivikki Karhula

Osallistuin ensi kertaa pohjoismaisen NORLIS-tutkijaverkoston seminaariin. NORLIS on 15 pohjoismaisen ja Baltian alueen kirjasto- ja informaatioalan oppilaitoksen verkosto, joissa on mahdollista suorittaa alan jatko-opintoja. Tutkijoiden verkosto kokoontuu vuosittain jossakin jäsenmaista. Tämänkertaiseen seminaariin, joka pidettiin Tallinnassa oli erinomaisella tavalla tuotu esimerkkejä verkkomaailman haasteista tutkimukselle Web 2.0 ympäristössä, mutta myös monia silmiä avaavia havaintoja aiemmista kehitysvaiheista.

Seminaarissa syvennyttiin Internetin mukanaan tuomiin tutkimuskohteisiin ja -menetelmiin. Puhujat tulivat Tanskasta, Norjasta ja Englannista. Ensimmäinen seminaaripäivänä oli avoin myös niille, jotka eivät itse aktiivisesti tutkimusta. Seuraavana päivänä tapahtuma jatkui tutkimusta tekevien omien töiden esittelyinä.

Web 2.0 kutoo verkon syvemmälle

Tanskalaista kirjasto- ja informaatioalan seminaarissa tutkimusta edusti Lennart Björneborn, jonka esitys perehdytti Internetin viimeaikaiseen kehitykseen ja sen haasteisiin alan tutkimukselle. Web 2.0 oli tiivistetysti esityksen teemoille yhteinen nimittäjä.

Hän kuvaa Internetiä eräänlaisena kolmiulotteisena hämähäkin verkkona, joka on jatkuvassa liikkeessä ja uudelleen muotoutumisen tilassa. Käyttäjät kutovat verkkoa ("weaving the Web") hetki hetkeltä tuottaen tietoa ja luoden yhteyksiä linkkien tai viittausten kautta. Ajatus on tuttu jo Tim Berners-Leen kirjoituksista ja Internetin alkua ajoilta.

Tämän Webin luonteeseen pohjautuvan dynamiikan Björneborn näkee vaikuttavan myös tiedon organisointiin. Moniulotteinen hämähäkin verkko on niiltä osin ontuva symboli, ettei verkko ole tasainen.

Vaikka tietoa tuotetaan massoittain, eivät kaikki palvelut ole koskaan yhtä vetovoimaisia. Palveluja, joita käytetään hyvin paljon, on vähän ja sivustoja, joilla on hyvin vähän kävijöitä, löytyy valtava määrä. Koska verkkoon muodostuu eniten käytettyjen palvelujen keskittymiä sekä erilaisten käyttäjäryhmien runsaasti käyttämiä ryppäitä tai klustereita, hänen kiinnostuksensa kohdistuu näihin keskittymiin.

Björneborn esitteli erilaisia visuaalisen analyysin malleja käyttäjien ryppäiden erittelyyn. Käyttäjien ryhmittelyä samankaltaisten käyttötapojen mukaan ("patterns") hän kuvaa esimerkiksi tähti-taivaana, jossa eri teemoista kiinnostuneet käyttäjät erotetaan erivärisinä tähtinä. Toinen esitysmuoto oli käyttää teemakarttoja, joissa suosituimmat teemat erottuvat fontin koon mukaan "teemapilviksi" ("tag clouds"), mikä on käytössä esimerkiksi Technoratin verkkopalvelussa.

Tagit ovat Technoration kaltaisissa uusissa verkkopalveluissa käytössä eräänlaisina käyttäjien tuottamina asiasanastoina, joiden avulla käyttäjät voivat merkitä kiinnostuksensa kohteet teemoittain, linkittää niihin sopivat tiedonlähteet tai käyttää jo muiden valitsemia tageja (esim. <http://www.technorati.com/tag/>).

Sosiaalinen suodatin

Björneborn korosti sosiaalisten ja informaatioverkostojen merkitystä, joka saa erilaisia muotoja Web 2.0 ympäristössä. Sosiaalinen verkosto, samoin kuin informaatioverkosto muotoutuu hetki hetkeltä uudelleen. Web 2.0 ympäristön käännekohtana on tämän verkoston laajentuminen ja syventyminen.

Web 2.0 ympäristö tuo monia uusia välineitä tiedonhakuun ja tuottamiseen, jakamiseen ja yhteyksien luomiseen, joille leimallista on, että ne ovat jakamisen ja sosiaalisen suodattamisen välineitä. Verkostomaisuutta, yhteyksien luomista ja kommentointia suosivat uudet välineet ovat muuttamassa tiedon haun ja organisoinnin ympäristöä. Tietoa tuotetaan, kommentoidaan ja kokemuksia jaetaan erilaisissa ryhmissä. Puhutaan myös tiedon sekoittamisesta (mix) ja uudelleen sekoittamisesta (remix).

Sosiaalisten palvelujen ympäristöä edustavat jo monet laajalti käyttöön levinneet ja käyttäjien suosimat verkkopalvelut, kuten wikipedia, folksomiat, blogit ja sekä erilaiset palvelumallit, jotka perustuvat resurssien jakamiseen. Wikipedia eli käyttäjien tuottama verkkotietosanakirja on laajalti jo levinnyt käyttöön erikielisinä versioina. Folksomiat ovat käyttäjien luomia teemapohjaisia palveluja. Resurssien jakamisesta hyvä esimerkki on Flickr, joka toimii kuvien tallennuspalveluna. Samalla käyttäjä voi kuitenkin luoda palveluun oman profiilinsa kuvauksen, sekä luoda linkkejä muihin arvostamiinsa kuvaajiin.

Blogien tähtitaivas

Blogit eli verkkopäiväkirjat edustavat erästä verkossa räjähdysmäisesti levinnyttä tiedon tuottamisen ja verkostoitumisen tapaa. Ensimmäiset blogit julkaistiin jo ennen vuosituhannen vaihdetta, kun nykyisellään blogien määrä arvioidaan yli 30 miljoonaksi.

Vaikka blogieja tuottavat eniten 13-19 -vuotiaat nuoret, niiden käyttö on levinnyt myös eri alojen asiantuntijoiden ja erityisesti medioiden käyttöön.

Se, mikä blogien luonteessa on muuttunut, on niiden laajentuminen ja monimuotoisuus tietovarastona. Blogeihin on pakattu tiiviisti paitsi profiili tekijästä ja arkisto hänen kirjoituksistaan, myös tiivis kuvaus hänen yhteyksistään tai mielenkiinnon kohteistaan. Blogit eivät ole enää vain päiväkirjoja, vaan ne edustavat yksittäisten ihmisten maailmankuvia ja sosiaalisia verkostoja kytkentöineen.

Blogit ovat monimuotoisia tietovarastoja tekijälle ja lukijalle. Teksti ei välttämättä hallitse blogieja, vaan puhutaan myös mobiiliblogeista, video- ja kuvablogeista. Kirjoituksista muodostuu automaattisesti arkisto, mikä antaa kirjoituksille ajallista perspektiiviä. Tekijän oman tuotannon lisäksi blogieissa tyypillisesti viitataan muihin bloggaajiin ja teemoitetusti esimerkiksi kirjallisuuteen, dokumentteihin, kuviin tai musiikkiin. Björneborn piti blogieja viittausanalyysistä kiinnostuneille kirjasto- ja informaatioalan tutkijoille aarreaittana.

Verkkopäiväkirja ei ole enää ajanmukainen kuvaus blogille, koska sen muoto ei ole yksinpuhele, vaan se voi olla myös monitahoista kommunikaatiota muiden bloggaajien ja kommentoijien kanssa. Blogieista on myös muodostunut sosiaalisten verkostojen solmukohtia. Bloggaajat viittaavat muihin blogeihin (esim. blogit, joita seuraan), kommentoivat toistensa tekstejä ja perustavat yhteisöjä, joissa tavataan myös livenä. Päiväkirjamainen kerronta on vain osa pakettia.



Onko kirjastoihin keskusteluvirtaa?

Suuremman kysymysmerkin muodostaa se, miten näitä välineitä voisi kirjastojen ja tietopalvelujen ympäristössä hyödyntää. Saako käyttäjät mielekkäällä ja toimivalla tavalla osallisiksi verkkopalvelujen tuottamiseen?

Miten tämä kytkeytyy kirjaston palveluihin tai miten sitä voitaisiin hyödyntää? Björneborn kertoo paikallisen kirjaston ottaneen käyttöön esimerkiksi suositushyllyn, jonne käyttäjät voivat valita aineistoa. Tämän rinnalla hän pohdiskelee esimerkiksi mahdollisuuksia kytkeä käyttäjiä mukaan esimerkiksi tietueiden kuvailuun tai sanastojen tuottamiseen

Onko kirjastoon sitten keskusteluvirtaa? Kun olen joitakin kirjastojen tarjoamia blogeja Library Weblogs -hakemistosta seurailut, ei käyttäjien kommentteja näistä juuri löydy.

Muuttuvan tiedonhakuympäristön vaikutukset

Professori Nils Pharo Oslon yliopistosta esitteli tiedonhakuun liittyvää tutkimustaan, jonka asetelma oli peräisin Web-ajan ensimmäisestä aallosta 1990-luvun lopulta. Silloin alussa ollut Web-ympäristö haastoi tutkijat pohtimaan, millaisia vaikutuksia uudella ympäristöllä näyttää olevan tiedonhakuun.

Pharon tutkimuksessa pyrittiin tunnistamaan tiedonhakukäyttäytymisen malleja ("identifying patterns in people's behaviour") yhdistelemällä erilaisia lähestymistapoja ja tutkimusmenetelmiä. Tutkimusasetelmassa hakeuduttiin tiedonhakukysymyksiä ja -strategioita syvemmälle tiedonhakutilanteen kokonaisuuteen: Miten ja mihin tiedonhakija etsii vastauksia?

Tutkimuksen pohjana käytettiin opiskelijoiden todellisia tiedonhakutehtäviä, joissa he jäljittivät aineistoa lopputyöhönsä. Hakutilanteita tutkittiin erilaisilla toisiaan täydentävillä menetelmillä. Hakijat vastasivat kyselyyn ja hakutilannetta havainnoitiin samalla, kun hakija kertoi hakunsa etenemisestä ääneen ("talking aloud"). Tämän lisäksi haut videoitiin.

Havainnointien ja videointien aineistot avattiin hakuun suoranaisesti liittyvien tapahtumien osalta. Tämän jälkeen yhteenvedoista rakennettiin hakukohtaiset kertomukset, joiden pohjalta vedettiin lopulliset johtopäätökset.

Hakujärjestelmien mallit siirtyvät

Tutkimuksen tulokset antoivat odotetusti monia yleisempiä viitteitä hakijoiden käyttäytymisestä tiedonhakutilanteessa. Pharon mukaan työtehtävä ja tilanne suuntasivat tiedonhakuja eri tavoin.

Tiedonlähteiden jäljittämiseen käytettiin hakutilanteissa verrattain paljon aikaa ja haun aikana saaduista palautteista, vastauksista ja virheistä otettiin oppia. Ongelmallista näytti silti olevan haluttomuus muotoilla uudelleen hakustrategiaa, vaikka haku ei olisikaan näyttänyt tuottavan tulosta.

Pharon tutkimustulokset kertoivat joissakin kohdin myös verkkomaailman ja tiedonhakujärjestelmien silloisesta vaiheesta ja näiden välisestä suhteesta. Esimerkiksi hakukomentojen käytössä ilmeni, että käyttäjien hakutapoihin siirtyi malleja kirjastojärjestelmien käyttötavoista.

Koska tiedonhaku verkon hakukoneilla ja kirjastojärjestelmissä toimii monilta osin erilaisin periaattein, näiden hakukäytäntöjen toisiinsa sekoittaminen näytti heikentävän haun tuloksia. Nykyisellään samankaltainen ongelma näkyy käytännössä ehkä toisensuuntaisena eli hakuko-



neiden hakutapojen soveltamisena muilla periaatteilla toimiviin hakujärjestelmiin.

Hakijoiden todettiin myös haun aikana viettävän huomattavasti aikaa (42%) ns. siirtymävaiheissa, kuten linkkilistalta toiselle siirtymisessä ja näihin perehtymisessä.

Tässä on todennäköisesti huomioitava, että 1990-luvun lopussa hakukoneet olivat toisella tasolla ja erilaiset aihehakemistot olivat merkittävämmässä asemassa tiedonlähteinä kuin nykyisin.

Blogianalyysin uudet välineet

Mike Thelwall Wolverhamptonin yliopistosta palautti menetelmakeskustelun miltei hämmäntävillä tavoilla reaaliaikaan. Thelwallin tutkimuskohteena on tieteen kehitykseen ja uusiin innovaatioihin liittyvän yleisen mielipiteen jäljittämisen blogeista.

Analyysien eräänä tavoitteena on jäljittää ja ennakoita yleiseen keskusteluun tulevien ”kuumien aiheiden” aaltoja. Analyysi ei siis niinkään pyri selittämään jo tapahtunutta kuin ennakoimaan, mistä keskustellaan tulevaisuudessa. Tietoa pyritään hyödyntämään päätöksenteossa.

Blogien analysoinnissa hän on koonnut aineistoa verkon hakukoneilla ja analysoinut niitä Intelliseekin analyysiohjelmalla sekä työryhmänsä tuottamalla Mozden RSS –monitorilla. Thelwallin esittelemillä menetelmillä käsitellään häkellyttävän nopeasti suuria määriä ajankohtaista aineistoa, organisoidaan se teemakohtaisesti ja analysoidaan esimerkiksi aiheen mukaan ja aikasarjoina.

Aikasarjat luovat kuvaa esimerkiksi tietyn keskustelunaiheen kehityksestä. Thelwall tuo esimerkiksi keskusteluun aiheen ”sarjakuvat” ja ”Tanska”? Aikasarjaksi jalostettu blogikeskustelujen analyysi paljastaa, etteivät piirroksen julkaisuajankohta ja keskustelun ajankohta osu yhteen, vaan keskusteluyöry lähti liikkeelle huomattavasti myöhemmin.

Monia muutkin hänen esimerkeistään osoittivat, millaisia yllätyksiä julkisen keskustelun aaltoiin liittyy.

Thelwallin analyysit ovat kohdentuneet myös erilaisiin suhdeverkostoihin. Kiinnostuksen kohteena on ollut selvittää, mitä Internet- sivustojen välisten linkkien määrä kertoo esimerkiksi yhteyksistä eri maissa sijaitsevien tutkimusyksiköiden välillä? Kun hakukoneella haettiin tiettyjen osoitteiden väliset yhteydet, tietojen pohjalta saatiin esille nopeasti päätettyä suuntaa vahvoista, vähäisemmistä ja melko lailla puuttuvista yhteyksistä korkeakoulujen ja tutkimusyksikköjen välillä.

Nopeuden vaikutukset

Thelwallin analyysin välineistö ja tutkimuskohdekin vievät ajatuksia pidemmälle tutkimuksen suuntautumiseen. Se, miten blogit edustavat yleistä mielipidettä on oma kysymyksensä. Kirjoittajajoukko sijoittuu käytännössä tiettyihin ikäryhmiin sekä niihin, joilla on tiivis verkkoyhteyden käyttömahdollisuus. Kuinka hyvin niistä voi päätellä yleistä mielipideilmastoa, on kysymys, jonka hän itse myös otti esille.

Hänen mallinsa ja välineensä tutkia ajan mielipideilmastoa, on kuitenkin häkellyttävän nopea ja reaaliaikainen. Harva blogin kirjoittaja ymmärtänee, millaiset välineet hänen kirjoitushistoriansa, laajemman kirjoittajajoukon tekstien analysointiin sekä sosiaalisten verkostojen haravointiin ovat käytettävissä.

Erilaisia mielipidemittauksia on toki tehty kymmeniä vuosia, mutta mitä tapahtunee, kun tietojen käsittely alkaa olla lähes reaaliaikaista, samalla, kun sen ulottuvuudet tietojen määrässä, tiedon luonteessa, niiden ajallisessa katteessa ja tiedon käsittelyn nopeudessa kasvavat. Mitkä ovat asian vaikutukset ja miten analyysien tietoja tullaan vielä käyttämään? 