

Voiko verkkotietoon luottaa?

Jussi Nuorteva

Verkossa olevan tiedon määrä on viimeisten kymmenen vuoden aikana kasvanut räjähdysmäisesti. Samaan aikaan hakukoneiden ominaisuudet ovat parantuneet niin, että tiedon tulvan käsitteleminen ja luotettavan tiedon seulominen valtaisasta ja luonteeltaan erittäin kirjavasta tietomassasta on muodostunut todelliseksi ongelmaksi.

Hakusana "talvisota" tuotti suomenkielisenä hakuna tietokoneeni vakioselaimella 3272 osumaa. Englanninkielinen "Winter War", lisämäärittelyksellä "Finland", tuotti jo yli 45 000 osumaa. Hakusana "feminismi" pääsi, yllättävää kyllä, suomenkielisenä vain vähän yli 800 osumaan. Hakusanalla "feminism" osumia kertyikin sitten jo runsaat 330 000. Suomenkielinen "geeniteknologia" ylsi vain 400 osumaan, mutta englanninkielinen "genetic technology" saavutti peräti 650 000 löydön massan.

Harvinaisillakin sanoilla etsittäessä tietokoneen vakio-ohjelmistoon kuuluva hakukone pystyy löytämään verkosta kymmeniä osumia. Tavallisimmista aiheista englanninkielisten sivujen määrä liikkuu sadoissa tuhansissa. Mutta kuinka käyttökelpoista tämä hakukoneen suoltama tieto sitten on – ja voidaanko sitä luotettavalla tavalla käyttää muodostettaessa käsitystä yhteiskunnasta, historiasta tai vaikkapa tämän päivän ympäristöstä ja terveyttä käsittelevistä ilmiöistä?

Verkosta löytyy tietoa, mutta minkälaista?

Esimerkkeinä mainitsemani, täysin vapaasti mieleen nousseet hakusanat antavat varsin hyvän kuvan verkossa olevan tiedon ongelmista. Esimerkiksi talvisotaan liittyvää tietoa sisältävien Internet-sivujen järjestys oli hakukoneen vastausruudussa täysin sattumanvarainen. Suomenkielisessä haussa luettelossa oli ensimmäisenä puolustusvoimien ylläpitämä sivu <http://www.mil.fi/tiedotus/talvisota/>, joka si-

sälsi valikoituja sähköisiä asiakirjoja sodan jokaiselta päivältä. Aivan asiallinen ja hyvä sivusto, siis.

Toisen hakukoneen löytämän internetsivun ylläpitäjä oli WSOY:n Koulukanava <http://www.koulukanava.fi/historia/ww2/talvisot/>. Se sisälsi hyvät perustiedot talvisodasta, mutta kuvineen ja karttoineenkin sen laajuus jäi alle kymmenen sivun. Sivut olivat osa Koulukanavan Suomi toisessa maailmansodassa -kokonaisuutta. Mukavaa palvelua sinänsä, mutta tietosisällöltään vielä kaukana esimerkiksi siitä tasosta, jota talvisodasta on kirjoina tarjolla.

Kolmannen hakukoneen löytämän sivun talvisodasta olivat tuottaneet Äkäsloppolon ala-asteen yläluokkalaiset <http://www.akaslompolo.net/aksa/talvisota.htm> vuonna 1999, kun talvisodan syttymisestä oli tullut kuluneeksi 60 vuotta. Neljäs osuma oli talvisotaa koskeva osa Mannerheim-internetprojektista http://www.mannerheim.fi/10_ylip/s_talvis.htm. Tekstin oli käsikirjoittanut professori Ohto Manninen, sotahistorian todellinen asiantuntija. Viidenneksi nousi hakukoneen valinnassa Matti Yrjölän laatima sivu http://hkkk.fi/~yrjola/war/finland_wwii.html/, jonka tekijästä saattoi sivulla olevien merkintöjen perusteella päätellä ainoastaan, että hän oli Helsingin kauppakorkeakoulun opettaja tai opiskelija. Viimeinen päivitys sivulle oli tehty vuonna 1998.

Viiden edellä mainitun sivuston jälkeen luettelossa seurasi kirjava joukko sivustoja, joiden tiedot olivat luotettavuudeltaan varsin vaihtelevia. Talvisodasta verkon avulla tietoa etsivä henkilö olisi varmasti joutunut käyttämään niiden parissa tuntikausia ilman, että hän olisi välttämättä saanut juurikaan lisättyä todellista tietomääräänsä.

Minkälainen käsitys talvisodasta siis muodostuu vaikkapa koululaiselle, joka Internetin käyttöön tottuneena turvautuu pelkkään verkkotietoon ja kokoaa sen pohjalta koululaisten keskuudessa jo varsin vakiintuneella "leikkaa ja liimaa" -menetelmällä vaikkapa kouluesitelmän, jota hän sitten muokkaa hieman yhte-

näisemmäksi parilla sidelauseella ja joillakin omilla kommentteillaan. Lopputulos on huonossa tapauksessa hyvinkin sattumanvarainen, etenkin jos koululaisen perustiedot käsiteltävästä asiasta ovat ennestään vähäiset.

Entä sitten toinen esimerkkihakukohdalla tiedontarjonnan taso ei poikennut talvisodasta. Ensimmäinen teemaa käsittelevä osuma oli Carol Ehrlichin käännöksenä julkaistu artikkeli "Socialismi, anarkismi ja feminismi" <http://www.dlc.fi/~ravelre/feminismi/ehrich.htm>. Artikkelin oli kotisivun alussa olevan ilmoituksen mukaan julkaistu alunperin kirjoittaja- ja julkaisukollektiivi Research Group Onen raporttina n:o 26 tammikuussa 1977. Artikkelin julkaisija oli ilmoitettujen tietojen mukaan kuopiolainen Punamusta Kustannus ja suomennoksen olivat tehneet Ulla Vehaluoto ja Hannu Toivanen.

Sivu antaa siis varsin paljon taustatietoa, mutta kuinka moni kuitenkin tietää, mikä on sivulla mainittu kirjoittajakollektiivi tai kuopiolainen Punamusta Kustannus? Pääsivulta <http://www.dlc.fi/~ravelre/> löytyy pienen selaamisen jälkeen sivujen ylläpitäjä, Virtuaalinen itsehallinnollinen keskus, joka sekin oli ainakin minulle yhtä tuntematon. Kotisivun viitteet Kansan kuokkavierasjuhliin, mielenosoituksen kapitalistista globalisaatiota vastaan sekä linkit Anarkia-tekstiarkistoon, Kapinatyöläisen kotisivulle, Anarkofeminismiin ja Eläinten vapautusliikkeeseen antoivat kuitenkin tuntumaa sivun ylläpitäjän aateperustaan.

Mistä johtuu, että Internetin hakukone tarjoaa satunnaiselle tiedonhakijalle ensimmäisenä sivustoa, jonka sisältö on vahvasti ideologisesti väritynyt, ollaanpa sivuston tietosisällöstä sitten mitä mieltä hyvänsä. Millä periaatteella yhden verkkotiedon tuottajan sivu määrätty ensimmäiseksi tai sijoittuu haussa muuten korkealle, se on kysymys, johon hyvin harva Internetin käyttäjä ylipäättään lähtee etsimään vastausta.

Ensi näkemältä asiallinenkaan sivu ei välttämättä tyydytä riittävällä tavoin tietoa etsivän tarpeita. Esimerkiksi feminismiä käsittelevässä listauksessa sijoille 2–4 sijoittui kolme kotisivua, jotka kaikki olivat syntyneet yliopistoissa. Niistä kaksi oli yksittäisten opiskelijoiden määrätystä temasta valmistelemia alustuksia tutkijaseminaareissa. Ne oli selvästi tarkoitettu suppealle ryhmälle, ilmeisesti lähinnä muille seminaariin osallistujille. Kolmannella sivulla, jonka tuottaja oli Tampereen yliopiston täydennyskoulutuskeskus [http://www.uta.fi/tyt/avoin/verkko-](http://www.uta.fi/tyt/avoin/verkko-opinnot/sospsyka/tnfemin.htm#essentia)

esiteltiin sosiaalipsykologiaan liittyvää luentosarjan osaa, joka käsitteli tutkimuksen ns. toisia näkökulmia, kuten postmodernismia, sosiaalista konstruktionismia, Michel Foucault'n teorioita sekä feminismiä naistutkimuksena ja feminisminä. Mutta luentosarjan sisällysluettelosta ei löytynyt tällaista kokonaisuutta. Oliko senkin laatinut "Anu Pietikäinen sosiologian ja sosiaalipsykologian laitoksen ohjauksessa", kuten muun sosiaalipsykologian peruskurssin aineiston? Vai mikä oli sen suhde kokonaisuuteen – se ei käynyt ilmi täydennyskoulutuskeskuksen kotisivulta. Sinänsä valitettavan tavallinen puute, mutta tiedon käyttäjän kannalta kiusallinen. Kehen voi viitata, jos tietolähteen haluaa ilmoittaa, kuten asiaan kuuluu.

Internetin asema tiedonhankinnassa voimistuu

Esittelemäni ongelmat ovat arkisia kysymyksiä Internetiä tietolähteenään käyttävälle. Verkossa olevan tiedon määrä on viimeisten kymmenen vuoden aikana kasvanut räjähdysmäisesti. Samaan aikaan hakukoneiden ominaisuudet ovat parantuneet niin, että tiedon tulvan käsitteleminen ja luotettavan tiedon seulominen valtaisasta ja luonteeltaan erittäin kirjavasta tietomassasta on muodostunut todelliseksi ongelmaksi.

Joulukuussa 2002 julkaistu yhteispohjoinen informaatioteknologiatilasto, *Nordic Information Technology Statistics 2002*, osoittaa, että vuonna 2002 jo 60 prosenttia 16–74 vuotiaista Pohjoismaiden asukkaista saattoi käyttää Internetiä kotoaan. Islannissa ja Ruotsissa kotikäyttö ylsi peräti 70 prosenttiin, Suomi piti pohjoismaisessa tilastossa perää 53 prosentilla, mikä sekin on hyvää länsieurooppalaista tasoa [1].

Viikoittain Internetiä käytti Pohjoismaissa yli puolet väestöstä. Käyttö on koko ajan voimakkaasti lisääntynyt, eikä mikään näytä viittaavan siihen, että kehitys voisi tulevina vuosina saada toisen suunnan. Suomessa koko koulutusjärjestelmämme on peruskouluista yliopistoihin saakka liitetty Internet-verkkoon ja koululaisen ja opiskelijoiden mahdollisuudet käyttää verkkotietoa ovat muutamassa vuodessa oleellisesti parantuneet. Yhä useammin keskeinen käyttötieto haetaan painetun tietolähteen sijasta verkosta, olipa sitten kyse yksityisestä käytöstä tai opetuksen tai tutkimuksen tarpeista. Myös tiedot yhteiskunnan tarjoamista palveluista ovat jo erittäin hyvin tarjolla sähköisessä muodossa

ja niiden käyttö on merkittävästi lisääntynyt kaikkialla läntisessä Euroopassa.

Periaatteessa verkkotieto on demokraattista. Sen selaaminen on yhtä lailla mahdollista Helsingin Kruununhaassa kuin vaikka Äkäslompolon ala-asteella. Tekniset edellytykset käyttää tietoverkkoa tehokkaasti vaihtelevat silti paljon Suomen sisälläkin. Ruotsissa valtio on asettanut tavoitteeksi kaikkien kansalaisten tiedonsaannin turvaamisen mittavalla ohjelmalla, jonka tavoitteena on liittää kaikki taloudet nopeaan kaapeliverkkoon. Tanskassa puolestaan opetuksen tietohallinnosta vastaava UNI-C solmi joulukuussa 2002 Tanskan Sun Microsystemsin kanssa sopimuksen, jonka perusteella kaikki opettajat ja koululaiset voivat ladata palvelimesta SunOffice-ohjelman myös kotikoneilleen [2]. Suomessa kehitys on jätetty markkinavoimien tehtäväksi, mutta yhteiskunnan vähäisemmästä suorasta panostuksesta huolimatta tekninen kehitys on myös täällä ollut erittäin nopeaa.

Sähköisen tiedonhaun mahdollisuuksia tulevat lisäämään tietokoneiden ohella myös muiden viestimien, kuten puhelimen ja televisiovastaanottimien tekninen kehitys, voimakas panostus digitalisointiin, tietoverkkojen nopeutuminen ja langattoman viestinnän ennustettavissa oleva erittäin voimakas kehittyminen hieman pitämällä, ehkä kymmenen tai viidentoista vuoden aikavälillä. Nämä tekijät tulevat epäilemättä muuttamaan myös käyttötottumuksiamme voimakkaasti.

Vuonna 1988, jolloin juhlittiin ensimmäisen Suomea varten painetun kirjan ilmestymistä vuonna 1488, ennustettiin painetun kirjan häviämistä. Sen jälkeen on yhteisesti voitu todistaa, että painettu kirja ja painetut viestimet ovat ennustuksista huolimatta säilyttäneet vahvan asemansa nopeasti kehittyvän sähköisen viestinnän rinnalla.

Olen vanhan kirjan ystävänä iloinen siitä, ettei painettu kirja ole menettänyt merkitystään. Siitä huolimatta uskon vakaasti, että sähköiset tietoverkot tulevat jo aivan lähivuosina ohittamaan painetut tiedonlähteet tiedonhankinnan keskeisimpänä välineenä. Suurten tietomassojen hallinnassa sähköiset välineet ovat painettuihin tietolähteisiin verrattuna monessakin suhteessa yliverkaisia. Sen sijaan kaunokirjallisuuden kohdalla ei ole nähtävissä, että sähköiset välineet kykenisivät lähivuosina tosissaan haastamaan perinteisen kirjan.

Verkossa oleva aineisto on sisällöltään erilaista kuin painettu aineisto. Tästä johtuen kysymys verkkotiedon sisällöstä ja tiedon haku-

minaisuuksista tulee olemaan aivan keskeinen. Verkkotiedosta käytävässä keskustelussa tärkeitä ovat ainakin seuraavat kysymykset:

- 1) Verkossa olevan tiedon pysyvyys ja arkistointi
- 2) Verkossa olevan tiedon riittävyys tutkimukselle
- 3) Verkkoaineistoon kohdistuvan tiedonhaun kehittäminen ja tiedon laatuluokittelu

Verkossa olevan tiedon pysyvyys ja arkistointi

Jokainen verkkokäyttäjä on usein törmännyt tilanteeseen, jossa yhteyttä etsittyyn sivuun ei enää synny, vaan ruutuun ilmestyy tieto: "Sivua ei löydy. Etsimäsi sivu on voitu poistaa, sen nimi on voinut muuttua tai se on tilapäisesti pois käytöstä". Mitä siis eteen, jos sivu on sisältänyt korvaamatonta tietoa tai sitä on muuten tottunut käyttämään.

Teoriassa poistuneen sivun löytäminen on mahdollista, mutta käytännössä erittäin työlästä. Suomessa Internet-sivujen arkistoinnista vastaa Helsingin yliopiston kirjasto, Suomen Kansalliskirjasto. Kirjaston verkkoarkistoon tallennetaan suomalaisia Internet-sivuja automaattisella haravointiohjelmistolla, jonka CSC Tieteellinen laskenta Oy on kehittänyt eurooppalaisessa Nedlib-projektissa. Vuodesta 2001 lähtien vapaasti käytössä olevia tiedostoja on haravoitu fi-domainista siten, että kesäkuuhun 2002 mennessä oli arkistoitu jo 11 miljoonaa erillistä tiedostoa 40 miljoonasta verkko-osoitteesta [3].

Tulos on tietysti hieno, mutta tällä hetkellä aineistoa voi käyttää vain Helsingin yliopiston kirjaston ja muiden vapaakappalekirjastojen tiloissa verkkoarkistoa varten varatuilta työasemilta. Luulenpa, että käyttäjien määrä on pieni, sillä verkkotiedon käyttö perustuu yleensä juuri siihen, että tiedot ovat tarjolla omalta koneelta. Vain harvassa tapauksessa käyttäjä vaivautuu jäljittämään tietoa niistä lähteistä, jotka edellyttävät menemistä paikan päälle tietoja kaivamaan. Ennen vanhaan, kun kaikki vanhempi tieto oli saatavissa vain arkistoista, arkistoissa työskentely oli itsestään selvyyttä.

Internetissä olevan tiedon yhtenä etuna on pidetty sitä, että tietokantoja voidaan helposti lisätä, korjata ja päivittää. Tämäkin kuuluu asioihin, jolla on hyvät ja huonot puolensa. Jatkuva päivittäminen takaa sen, että aineisto pysyy ajan-tasaisena. Tutkimuksen kannalta ongelmallista on kuitenkin se, että lähteenä käytettyyn, jatkuvasti korjattavaan ja päivitettävään tietokantaan

on vaikea viitata. Helposti käy niin, että tietoa, johon on aikanaan on viitattu, ei muutaman vuoden kuluttua enää löydykään tai se löytyy aivan toisenlaisessa muodossa. Jos tutkimuksessa on tehty johtopäätöksiä, jotka ovat perustuneet vaiheeseen, jonka jälkeen tiedot ovat oleellisesti muuttuneet, aiheuttaa verkkotietoon viittaaminen suuria vaikeuksia arvioitaessa tutkimustiedon lähteiden käyttöä. On mahdotonta ajatella, että jälkikäteen voitaisiin enää rekonstruoida sähköisiin lähteisiin perustuvat tiedot sellaisina kuin ne olivat tutkimuksen valmistuessa ja löytää viitteet samassa muodossa kuin ne tuolloin olivat.

Painettua aineistoa käytettäessä kumuloituvan ja täydentyvän tiedon ongelmaa ei ole samassa muodossa kuin sähköisiä aineistoja käytettäessä, sillä kirjastolaitos säilyttää kaikkialla maailmassa erinomaisella tavalla kaikki painotuotteet ja niiden eri versiot. Ne voidaan vaivatta tunnistaa ja erottaa toisistaan bibliografisten tietojen perusteella. Näin tieteelliselle tutkimukselle välttämätön johtopäätösketju on kiistattomasti osoitettavissa ja johtopäätösten perustana oleva tieteellinen näyttö voidaan arvioida. Verkkotiedon kohdalla tilanne ei ole yhtä helppo.

Vaikeaksi on osoittautunut myös kysymys sähköisestä arkistoinnista. Toki digitaalisessa muodossa olevaa aineistoa voidaan periaatteessa muuttaa tallennusmuodosta toiseen, mutta tärkeiden aineistojen säilyttämistä pelkästään sähköisessä muodossa ei pidetä suotavana, ainakaan pysyvässä arkistoinnissa. Erilaisia sähköisen asiakirjahallinnon järjestelmiä on toki kaupallisestikin runsaasti tarjolla. Muistiorganisaatioille – kirjastoille, arkistoille ja museoille – sähköiset arkistointijärjestelmät eivät kuitenkaan vielä tarjoa sellaista mahdollisuutta, joka riittävän hyvin tarjoaisi mahdollisuuden aineistojen pysyvään säilyttämiseen. Yhtä lailla ongelmallista on silti sekin, että suuri osa sähköisestä aineistosta hävietään arkistoista. Sen vuoksi sähköisen asioinnin lisääntyessä on vaara, että asioiden pysyvän dokumentoinnin taso heikkenee, vaikka tuntuukin siltä, että paperimuotoisia tulosteita tuotetaan entistä enemmän.

Riittääkö verkkotieto tutkimuksen aineistoksi?

Järjestelmällisen digitoinnin ansiosta tutkimuksen käytössä on jo varsin huomattavia aineistomassoja [4]. Esimerkiksi Suomalaisen

Kirjallisuuden Seura on yhteistyössä Tarton Kirjallisuusmuseon kanssa digitoinut perinteen tutkimuksen keskeisimmän aineistokokoelman, 36-osaisen mammutisarjan Suomen Kansan Vanhat Runot, sekä lisäksi kymmeniä tuhansia samaan kokonaisuuteen liittyviä täydentäviä kortteja ja hakemistoja. Ylivertaisten hakuominaisuuksien ansiosta tietyn sanan tai teeman etsiminen massiivisesta aineistosta voidaan tehdä alle minuutissa. Silloin on koneen ruudulla valmis luettelo kaikista niistä kohdista, joissa hakusana esiintyy runoaineistossa. Aikaisemmin kattava tiedonhaku kyseisestä aineistosta olisi edellyttänyt viikkojen tai kuukausien puuduttavan selailemisen eikä tulos todennäköisesti olisi ollut edes yhtä luotettava ja kattava kuin koneen hakuominaisuuksia käyttäen voidaan saavuttaa.

Suurten tietomassojen saatavuus digitaalisessa muodossa on merkittävä apuväline tutkimukselle. Verkkokäyttö on parantanut yliopistojen ja tutkijoiden välistä tasavertaisuutta myös siten, että keskeisiin aineistoihin pääsee käsiksi vaikkapa kotikoneelta joutumatta työskentelemään pitkiä aikoja etäällä sijaitsevista kirjastoista ja arkistoista.

Verkkoaineistojen käytön helppoudessa piilee myös vaaransa, etenkin digitoitujen aineistojen määrän lisääntyessä. Hyvät käyttöominaisuudet houkuttelevat valitsemaan esitelmien tai opinäytetöiden aiheet siten, että työ voi perustua digitaalisessa muodossa oleviin materiaaleihin. Kuitenkin digitoiminen on valinnan tulos. Vain pieni osa kaikesta ihmisen tuottamasta on digitoitu tai voidaan ylipäätään digitoida. Aineistoja valitsevat henkilöt voivat sen vuoksi suunnata tieteen tai tietämättään tutkimusta, mikä tutkimuksessa on vakavasti muistettava ottaa huomioon tutkimuksen tietolähteitä valittaessa.

Historiantutkijalle pelkästään digitaalisiin aineistoihin turvautuminen tuntuu koko lailla mahdottomalta [5]. Etenkin vanhan ajan tutkija on oppinut, että alkuperäiseen lähdeaineistoon tuntuman säilyttäminen on oleellista myös lähdekritiikin kannalta. Vanhan kirjan sidos, paperi, reunamerkinät, täydennykset, poistot, korjaukset, ex-librikset sekä lukemattomat muut yksityiskohdat ovat oleellinen osa lähteen yksilöllistä historiaa. On täysin mahdotonta ajatella, että esimerkiksi 1700-luvun pitäjänkuvauksia tutkittaisiin pelkästään digitoitujen standardiversioiden avulla. Ne voivat toki monilta osin tukea tutkimusta, mutta ne eivät korvaa historiantutkijan ammattitaitoon välttämättä kuuluvaa perehtymistä autenttiseen aineistoon.

Verkkosensuuri – toimiiko se?

Yksi tapa estää ei-toivottua informaatiota on niin sanottu sisältösuodatus, jonka avulla pyritään estämään pääsy ennen muuta laitonta tai moraalitonta aineistoa sisältäville Internet-sivuille. Yhdysvalloissa on 8. helmikuuta 1996 hyväksytyn televiestintälain viidennen luvun, niin sanotun viestinnän siveellisyylain (Communications Decency Act, CDA) [6] avulla määrätty julkisten kirjastojen, koulujen ja virastojen sivut noudattamaan sisältösuodatusta, jolla ei-toivotun aineiston käyttö pyritään ehkäisemään. Tällä hetkellä Yhdysvalloissa on käynnissä suoranaisten sensuurisota, jossa lyömäaseiksi on otettu toisaalta vanhempien vastuu lasten moraalisen kasvatukselta toisaalta kansalaisten oikeus julkaista ja käyttää julkaistua aineistoa ilman ennakkosensuuria.

Sisältösuodatusta tukee voimakkaimmin vuonna 1996 Yhdysvalloissa toimintansa aloittanut PICS – Platform for Internet Content Selection <http://www.w3.org/PICS/>, jonka tavoitteena on kehittää menetelmiä, joilla Internetin käyttöä voidaan kontrolloida siten, että se tukee haluttujen sisältöjen löytämistä ja estää ei-toivotun aineiston käytön. Kyse on siis pikemminkin aktiivisesta valinnasta kuin suoranaisestä sensuurista, vaikka tarkkaa eroa näiden tavoitteiden välille onkin käytännössä hyvin vaikea vetää. PICS-järjestön aloitteesta ohjelmistotuottajat ovat kehittäneet ohjelmia, jotka tukevat mahdollisuuksia suorittaa valintaa ja tehdä rajoituksia. Internetin sisältöaineiston rajoitettua valvontaa ovat tukeneet sekä Euroopan Unioni että monien maiden kansallinen lainsäädäntö. Suomessa esimerkiksi Mannerheimin lastensuojeluliitto on koonnut Lapset.net –sivuille ohjeet lasten ja nuorten Internetin käytön ohjaamiseksi <http://www.mll.fi/>.

Toisaalta Internetin ”sananvapautta” on asettunut puolustamaan joukko järjestöjä, jotka katsovat verkon vapauden kuuluvan samanlaisiin kansalaisyhteisöihin kuin sananvapauden ja painovapauden. Vahvimmin vapaan Internetin puolesta on asettunut puhumaan yhdysvaltalainen Electronic Frontier Foundation, EFF <http://www.eff.org/>, joka on käynyt voimakasta taistelua niin PICS –järjestöä kuin viestinnän siveellisyylakiakin vastaan [7].

Kysymys Internetin sisältösuodatuksista ja ei-toivotun sisällön rajoituksesta on keino välttää sivuja, joita ei halua käyttää tai joita pitää hyödyttöminä. Se ei kuitenkaan paljontakaan auta sellaista verkkokäyttäjää, joka pyrkii nimenomaan

etsimään tarvitsemiaan tietolähteitä. Sensuurin jälkeenkin verkosta löytyy niin suuri määrä sivuja, että yleisessä haussa osuimen määrä nousee helposti tuhansiin, vaikka suodattamisessa käyttäisi hyvinkin tiheää kampa. Tiedon etsijälle rajoitukset ja sensuuri eivät siis tuo merkittävää helpotusta.

Verkoaineistoon kohdistuvan tiedonhaun kehittäminen ja tiedon laatuluokittelu

Internetissä olevan tiedon epätasaisuus havaittiin ongelmaksi jo 1990-luvun alkupuoliskolla. Silloin käynnistyi myös keskustelu niistä menetelmistä, joilla verkossa olevan tiedon käyttöä voitaisiin ohjata siten, että käyttäjät voisivat helposti arvioida käyttämiensä tiedonlähteiden luotettavuuden ja pitäytyä niissä lähteissä, joiden tarjoama tieto täyttää määrättyä laatuvaatimukset joiden vahvistaminen voi tapahtua erilaisin tavoin, joita edempänä käsitellen. Voimakkaimmin keskustelua käytiin lääketiedettä ja terveyttä koskevien sivujen sisällöstä, sillä epäluotettavat tiedot saattoivat olla suorastaan hengenvaarallisia käyttäjälle, joka ei kyennyt arvioimaan niiden turvallisuutta ja luotettavuutta.

Yksi varhaisimmista ja vakiintuneimmista terveystiedon laatuajusteista on kansainvälisesti perustavaa laatua ollut sveitsiläisen The Health On the Net säätiön vuonna 1996 esittelemä HONcode – HON käyttäytymissääntö <http://www.hon.ch/HONcode/>. Säätiön taustalla olivat Geneven kantonin, Geneven yliopistollinen sairaala, Swiss Institute of Bioinformatics – Sveitsin bioinformatiikan keskus – ja Sun Microsystems yhtiö. Nykyisin voimassa oleva HON käyttäytymissääntö hyväksyttiin huhtikuussa 1997 ja se on sittemmin käännetty 23 kielelle. Suomi on ollut vahvasti mukana sääntöä kehittämisessä ja kansainvälisessä tunnetuksi tekemisessä, sillä Turun yliopiston lääketieteen tietotekniikkaan (<http://www.utu.fi/research/mircit/forsstro.html>) erikoistunut dosentti Jari Forsström on yksi alan johtavista kansainvälisistä asiantuntijoista. Sääntöä avulla pyritään tukemaan ennen muuta maallikoille ja lääke- ja terveystieteellistä tutkimusta muuten vähän tunteville suunnattavaa tiedontuottamista. Nykyinen HONcode – eli HON käyttäytymissääntö edellyttää seuraavien periaatteiden noudattamista tarjottaessa lääke- ja terveystieteellistä tietoa verkossa <http://www.hon.ch/HONcode/Finnish/>:

1. Mikä tahansa lääketieteeseen tai terveydenhuoltoon liittyvä neuvo, joka palvelimella esiintyy, on lääketieteen tai terveydenhuollon ammattihenkilön antama ellei erikseen mainita, että tieto on muun kuin lääketieteen tai terveydenhuollon ammattihenkilön antama.
2. Tällä palvelimella tarjottu tieto on tarkoitettu tukemaan, ei korvaamaan, hoitosuhdetta joka on olemassa potilaan/tiedon lukijan ja hänen lääkärinsä välillä.
3. Palvelimen ylläpitäjä pitää luottamuksellisina tietoja niistä potilaista ja muista vierailijoista, jotka käyvät lääketieteellistä tietoa sisältävillä sivuilla. Palvelimen omistaja täyttää tai ylittää ne lain asettamat vaatimukset lääketieteen tietojen luottamuksellisuudesta, jotka ovat voimassa siinä maassa, jossa palvelin tai sen peilauspalvelin sijaitsevat.
4. Jos vain mahdollista, sivulla esitetyn tiedon tueksi pitäisi esittää viitteet tiedon lähteeseen ja jos mahdollista myös suora HTML-linkki ko. tietolähteeseen. Päivä, jolloin lääketieteellistä sivua on viimeksi muutettu, tulee olla selvästi näkyvissä (esim. sivun alalaidassa).
5. Kaikkien väitteiden tietyn hoidon, kaupallisen tuotteen tai palvelun hyödyistä/tehosta tulee tukeutua asianmukaiseen puolueettomaan näyttöön, kuten edellä Periaatteessa 4. on määritelty.
6. WWW-sivun tekijät pyrkivät esittämään tiedon selkeimmällä mahdollisella tavalla ja antamaan sivuilla kävijöille kontaktitiedot, joiden avulla heidän on mahdollista kysyä lisätietoa tai tukea. Sivun ylläpitäjä laittaa oman sähköpostiosoitteensa selkeästi esille kaikille sivuilleen.
7. Sivulla on selkeästi ilmoitettava mahdollinen palvelimen saama ulkopuolinen tuki, kuten kaupallisten ja ei-kaupallisten organisaatioiden tuki rahoituksen, palvelun tai materiaalin muodossa.
8. Jos rahoituslähteenä käytetään mainostuloja, se on selkeästi ilmoitettava. Tässä tapauksessa palvelimen omistaja kirjaa sivulleen lyhyen kuvauksen mainonnan suhteen noudatettavasta politiikasta. Mainokset ja muu myynninedistämiseen tähtäävä materiaali on esitettävä sillä tavalla ja sellaisessa yhteydessä, että se on mahdollisimman hyvin erotettavissa palvelimen ylläpitäjän toimittamasta alkuperäisestä aineistosta.

HON käyttäytymissääntöön liittyy laatumerkki, joka takaa sen, että sitä käyttävät tiedontuottajat ovat sitoutuneet toimimaan HON käyttäytymissäännön periaatteiden mukaisesti. Health On the Net -säätiö tukee luotettavien tietojen tarjoamista myös muodostamansa terveystietovestojen portaalin avulla. Säätiön MedHunt <http://www.hon.ch/MedHunt/> hakukone

tarjoaa pääsyn suureen joukkoon luotettavana pidettävää tietoa, joista osa on sitoutunut HON käyttäytymissääntöön, mutta joukossa on monia muitakin sivustoja, joiden tiedot perustuvat luotettavana pidettävään tieteelliseen arviointiin. Portaalin etuna tässä tapauksessa on se, että se sulkee automaattisesti hakujen ulkopuolelle yleisten hakukoneiden suoltamat sadat tuhannet osumat, joiden tietosisältö on tavanomaisen kirjavaa ja luotettavuudeltaan paikoin hyvinkin kyseenalaista.

HON käyttäytymissääntö perustuu vapaaehtoisuuteen, haluun sitoutua yleisesti hyväksyttäviin periaatteisiin. Laatumerkin käyttö on luvanvaraista ja edellyttää tehdyn hakemuksen hyväksymistä. Käyttäytymissääntöä myös valvotaan. Jos sivuilla epäillään väärinkäytöksiä, epäluotettavuutta tai vilppiä, voi käyttösääntöä valvova säätiö varottaa tiedontuottajaa ja toisessa vaiheessa, jos huomautus ei ole johtanut toivottuun tulokseen, evätä laatumerkin käyttöoikeuden. Jos tämäkään ei johda tulokseen katkaistaan yhteydet HON säätiön sivuille, jolloin laatumerkki – jos sellainen on vielä omavaltaisesti pidetty sivulla – muuttuu passiiviseksi. Passiivinen merkki, josta ei ole yhteyttä säätiön sivuille on osoitus siitä, ettei luokitus enää ole pätevä. Käytännössä varsinaisia väärinkäytöksiä on ollut hyvin vähän, useimmin ongelmana on tietolähteen jääminen päivittämättömäksi, yhteystietojen muutos tai sivun poistuminen käytöstä.

Kansainvälisten laatuluokittelujen käyttö näyttää olevan se suunta, johon verkkotiedon hallinnassa ollaan suuntautumassa. Euroopan komissio julkaisi aivan hiljan, marraskuun 29. päivänä 2002, eEurope-hankkeeseen liittyvän kommunikation, jonka aiheena oli *Quality Criteria for Health related Websites* – terveystietoa sisältävien verkkosivujen laatukriteerit [8].

Komission tarkoituksena ei ole ollut antaa tarkkoja ohjeita laatuluokittelusta eikä luoda EU standardeja, joiden valvonnan se ottaisi vastuulleen. Lähtökohtana on ollut tukea jäsenvaltioiden ja niissä toimivien organisaatioiden omaehtoista pyrkimystä tuottaa tietoa tavalla, joka noudattaa voimassa olevaa lainsäädäntöä ja täyttää yhteisesti hyväksytyt laatuvaatimukset.

Keskeisiä laatukriteerejä ovat tiedon tuottamisen avoimuus, johon liittyy tiedontuottamisen tavoitteiden ja kohderyhmien selkeä määrittely ja rahoituslähteiden esitleminen. Tieto ei myöskään saa olla nimetöntä, vaan sekä tietoa tuottaneiden henkilöiden että heidän edustamiensa laitosten tai organisaatioiden tiedot tulee

ilmoittaa. Sivunkäyttäjien ja kohdehenkilöiden tietoturva on toteutettava komission tietosuojalainsäädännön edellytysten mukaisesti. Tietoja tulee päivittää säännöllisesti ja päivityspäivämäärät tulee ilmaista selkeästi tietojen ajantasaisuuden varmistamiseksi. Käyttäjien tulee voida antaa palautetta sivuista ja palautteen käsittelystä tulee olla selkeä ohjeistus. Sivujen toimituspolitiikan tulee olla käyttäjän tiedossa, samoin sivuihin sisältyvien käyttömahdollisuuksien ja mahdollisten käyttörajoitusten.

Lääke- ja terveystieteellisten sivustojen ylläpitäjät ovat olleet edelläkävijöitä kehitettäessä tietosisältöön perustuvia laatu- ja luokitusjärjestelmiä. Vastaavia luokitusjärjestelmiä tarvittaisiin myös muille aloille. Menettelytapoja voi olla useita, niistä komission 29.11.2002 antama kommunikointiluetteleeseen muutamia. Monilla aloilla asiantuntijat ovat keskenään ryhtyneet tekemään työtä kehittääkseen yhteisiä käyttäytymissääntöjä, joihin sitoutuminen on vapaaehtoista. Seuraava aste on laatumerkin (Quality Label) käyttö, mikä edellyttää jo pitemmälle menevää sitoutumista ja valvontamekanismien kehittämistä. Valvontaa voi suorittaa esimerkiksi tunnustettu kansainvälinen järjestö, samaan tapaan kuin HON käyttösäännön valvonnan osalta on toteutettu. Erittäin tärkeää on joka tapauksessa se, että käyttäjille ilmoitetaan kaikissa tapauksissa selkeästi käyttäytymissäännöt ja laadunvalvonnan toteuttamistapa.

Ongelmatonta laatujärjestelmien kehittämisen ei suinkaan ole. Ei ole hyvä, jos eri aloille syntyy eri organisaatioiden määrittelemiä laatuluokitteluja, jotka ovat periaatteiltaan vaihtelevia. Moninaisuus on ongelma käyttäjän kannalta, sillä hän ei välttämättä kykene arvioimaan laatumerkin todellista pätevyyttä, jos niitä ryhtyy myöntämään organisaatio, jonka menettelytavat eivät vastaakaan tieteellisen tiedontuottamisen menetelmiä. Voidaan kysyä hyvin aiheellisesti, miten suhtauduttaisiin esimerkiksi luontaislääketiedettä edustavien organisaatioiden laatumerkkiin – epäilemättä tällaiselle voitaisiin helpostikin tehdä luotettavilta vaikuttavat periaatteet – jos kyseessä ei kuitenkaan olisi tieto, jonka luotettavuudesta tiedeyhteisö on samaa mieltä. Vastaavia esimerkkejä olisi helppo nostaa esiin muiltakin aloilta, samoin instituutteja, jopa yliopistoiksi itseään nimittäviä laitoksia, joiden opetus perustuu ennalta määriteltyyn ideologiseen tai uskonnolliseen oppiin tai ahtaaksi rajoitettuun arvojärjestelmään, ei vapaaseen tieteelliseen tutkimukseen. On tärkeää, että laatuluokittelua

voidaan kehittää siten, että laatumerkin voivat antaa vain yhteisesti tunnustetut, keskeiset kansainväliset organisaatiot, joilla on mahdollisuus ottaa monipuolisesti huomioon eri tieteenalojen menetelmät ja tiedontuottajien laadukkuus. Esimerkkejä toimivista kansainvälisistä standardeista on jo olemassa, esimerkiksi SIS merkki tai vaikkapa pohjoismainen ympäristömerkki. Niitä ei toistaiseksi vain ole merkittävässä laajuudessa sovellettu verkkotiedon tuottamisen alueelle.

Internetistä löytyy runsaasti tietosisällöltään korkealaatuista aineistoa, kysymys on vain siitä, miten ne parhaalla tavalla voidaan löytää. Yleiset hakukoneet eivät ainakaan toistaiseksi tunnista tiedon erilaisia laatuluokkia, minkä vuoksi tietoa on osattava etsiä verkosta toisia menetelmiä käyttämällä. Yleisimmin käytettyjä ovat portaalityyppiset ratkaisut, joissa määrätyltä sivulta löytyy linkkejä keskeisiin tiettyä aihepiiriä käsitteleviin sivustoihin. Portaaleja on sekä kansainvälisiä että kansallisia. Euroopan komissio ylläpitää Euroopan kulttuuriportaalia http://europa.eu.int/comm/culture/index_en.htm, Suomessa toimivat muun muassa kansallinen kulttuuriportaali <http://www.kulttuuri.net/> ja opetusministeriön ylläpitämä Research.fi –tietopalvelu <http://www.research.fi/>

Euroopassa ja Suomessa kirjastot ja arkistot tarjoavat pääsyn laajoihin tietokantoihin ja aineistoihin, muun muassa keskeisiin sähköisessä muodossa julkaistuihin tieteellisiin aikakauskirjoihin. Monissa tapauksissa esimerkiksi yliopistojen kirjaston kautta tietokantoihin pääsy on rajoitettu yliopiston piirissä toimiviin tutkijoihin, opettajiin ja opiskelijoihin. Näihin ryhmiin kuulumattomien on vaikeampi tavoittaa samantasoiset tietopalvelut, mikä on vakava puute ja suuntaa osaltaan tavallisia tiedontarvitsijoita hakemaan tietoa Internetissä tarjolla olevista maksuttomista lähteistä, jotka kuitenkin ovat luotettavuudeltaan ongelmallisia.

Yhdysvalloissa etenkin koulujen tiedontarpeeseen on kiinnitetty voimakasta huomiota. Koulujen tiedekasvatuksen tarpeisiin aineistoa tuottavat monetkin organisaatiot. Kongressin kirjasto tarjoaakin kotisivuillaan tiedeviitepalvelun, Science Reference Services <http://lcweb.loc.gov/rr/scitech/>. Osana sitä on koulujen tiedeopetuksen tarpeisiin koottu valikoima Internet sivustoja, jotka tarjoavat runsaasti tietoa aineistoa vapaasti käytettäväksi. Yhdysvaltain tiukan lainsäädännön vuoksi kirjasto sanoutuu kuitenkin irti vastuusta kotisivuillaan olevien sivustojen tietosisällöstä. Laatumerkin sijasta sivustoilla onkin tätä osoittava merkki,

Standard Disclaimer for External Links <http://www.loc.gov/global/disclaim.html>, jossa kirjasto toteaa, että mainitut sivut on liitetty kotisivuille käyttäjien palvelemiseksi, mutta kirjasto ei valvo niiden sisältöä eikä vastaa niillä esitetystä tiedoista, niiden laillisuudesta tai ajanmukaisuudesta tai mielipiteistä, joka on yksinomaan sivustojen tuottajien vastuulla.

Rajoituksista huolimatta palvelu on hyödyllinen ja paljon käytetty ja tarjoaa erinomaisia esimerkkejä hyödyllisistä tietopalveluista myös meille. Onkin todettava, että Suomessa koulujen tiedekasvatuksessa on vielä huomattavan paljon tehtävää ja opittavaa niin Yhdysvaltain kuin johdettavien Euroopan maiden, kuten Ison-Britannian ja Ranskan tai Italian käytännöistä. Kongressin kirjaston sivustoilta löytyvät viitteet lukuisille sivuille, joilla on tarjolla tieteseen ja tekniikkaan liittyvää tietoa <http://www.loc.gov/rr/scitech/teachers.html>. Esimerkiksi ERIC Clearinghouse for Science, Mathematics and Environmental Education tarjoaa kasvattajille, vanhemmille ja lapsille sopivaa materiaalia luonnon- ja ympäristötieteistä ja matematiikasta. Tiedeopettajien valtakunnallinen järjestö, National Science Teachers Association on koonnut runsaan tietopaketin, joka sisältää linkkejä aikakauslehtiin, museoihin, organisaatioihin ja opetusjärjestöjen tuottamiin aineistoihin Yhdysvalloissa ja Kanadassa.

Miten sitten esittämäni pohjalta vastaan kirjoitukseni kysymykseen, voiko verkkotietoon luottaa? Vastaan, että kyllä verkkotietoon voi luottaa, mutta vai silloin, kun tietoa on osattu etsiä luotettavista lähteistä, joita kyllä verkosta löytyy runsaasti. Nopeasti kehittynyt laatuluokittelu, huolimatta siitä, että paljon on vielä tehtävää silläkin alueella, on oleellisesti helpottanut luotettavan aineiston erottamista epäluotettavasta.

Tänä päivänä suurimmat haasteet liittyvät kansalaisten verkkolukutaitojen kehittämiseen. Vain harvat tuntevat jo nyt tarjolla olevia mahdollisuuksia, monet niistä ovat lisäksi englanninkielisiä. Kansallisella tasolla Suomessa on vielä paljon työtä tehtävänä. Erityinen vastuu verkkolukutaitojen kehittämisestä on opetuslaitoksella ja kirjastolaitoksella. Kansalaisyhteiskunnassa on kiinnitettävä huomiota myös siihen, että kaikilla kansalaisilla on sekä pääsy tiedon lähteille että kyky käyttää verkossa olevaa tietoa.

Kehitettäessä Suomea tietoyhteiskuntana – mikä on jo pitkään ollut keskeinen kansallinen kehittämisstrategiamme – on tärkeää kyetä tarjoamaan käyttäjille korkeatasoista ja luottavaa, viime kädessä tieteelliseen tutkimukseen perus-

tuva tietoa. Vain sillä tavoin voimme taata sen, että sekä kansalaisten tietous että yhteiskunnallinen päätöksenteko eri tasoilla perustuvat parhaaseen saatavilla olevaan tietoon, olipa se sitten tarjolla painettuna tai sähköisessä muodossa. Suomessa perusoikeuksiin on luettava myös se, että maailmankuvaa voidaan rakentaa omalla äidinkielellä ja että siihen on saatavilla riittävästi luotettavaa aineistoa. Tämä ei voi toteutua ilman yhteiskunnan määrätietoista panostusta eri käyttäjäryhmille soveltuvaan, tutkimukseen perustuvaan tietoon.

VIITTEET

- [1] Nordic Information Society Statistics 2002, Helsinki 2002, 153 s (http://stat.fi/tk/yr/tietoyhteiskunta/nordic_iss_02.pdf).
- [2] Karim Pedersen, Sun forærer StarOffice til danske skoleelever, Com.On 20.12.2002, (<http://www.comon.dk/index.php?page=news:show,id=12816>).
- [3] Kaisa Kaunonen, Apua, hävisikö se verkosta! Onks se nyt arkistossa? Kansalliskirjasto verkkoharavoi tiedostoja. Helsingin yliopiston kirjaston tiedotuslehti 3/2002, s. 19-22.
- [4] Ks. Tieteen tietokantojen toimenpideohjelma 1998–2001, CSC – Tieteen tietotekniikan keskus, 1998, (<http://www.csc.fi/raportit/tietokannat.html>).
- [5] Ks. esim. *Tapio Onnela*, Mihin historian-tutkimus ja opetus tarvitsevat digitaalisia aineistoja? – Helsingin yliopiston kirjaston tiedotuslehti 3/2002, s. 19-22.
- [6] Hannele Ihonen, Sananvapaus kansainvälisissä sopimuksissa ja Internet, luku 3.11. Yhdysvallat, (<http://www.helsinki.fi/~mscheini/ihonen.txt>).
- [7] Elias Aarnio, EFF hyökkää Yhdysvaltain Internet-sensuurilakeja vastaan. Digitoday. fi, 2.4.2001 (http://digitoday.fi/digi98fi.nsf/print/dd20010402110913_ea_97633783).
- [8] Commission of the European Communities, Brussels, 29.11.2002 COM (2002) 667 final. Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions. EEurope 2002: Quality Criteria for Health related Websites.

Kirjoittaja on dosentti ja Tiedonjulkistamisen neuvottelukunnan puheenjohtaja sekä Suomalaisen Kirjallisuuden Seuran pääsihteeri. Kirjoitus perustuu esitelmään Tieteen päivillä 8.1. sessiossa "Tiede verkossa – bulkkia vai laatua?"