

Gunilla Widén-Wulff

Tiedonhankinta- ja tiedonhaku- tutkimuksen käännekohta

Peter Ingwersen & Kalervo Järvelin: *The Turn – Integration of Information Seeking and Retrieval in Context*. Dordrecht: Springer, 2005. – XIV, 448 s. ISBN 1-4020-3850-X

Informaatiotutkimuksen keskeiset tutkimusalueet, tiedonhankinta ja tiedonhaku, ovat kehittyneet yllättävän pitkään erillään. Suhteellisen pienelle tieteenalalle tämä ei ole paras mahdollinen suuntaus. Kirjassaan *The Turn* Peter Ingwersen ja Kalervo Järvelin puuttuvat tähän erillisyyden ongelmaan tarkastelemalla käännekohta, jossa nämä aiemmin pitkälti erilliset tutkimusalueet voisivat lähentyä toisiaan. Toisin sanoen, kirjan päämääränä on löytää uusi suunta tiedonhankinnan ja -haun tutkimukseen yhdentämällä nämä kaksi tutkimusalaa, niin että konteksti olisi suuntauksen yhteinen perusta. Toinen keskeinen perusta yhdentyneelle suuntaukselle on kognitiivinen näkemys. Tiedonhaun tutkimus kohdistuu usein tiedonhakujärjestelmiin ja teknologiaan kun taas tiedonhankintatutkimus pohjautuu yhteiskuntatieteellisiin ongelmanasetteluihin. Alueiden eroavaisuuksista huolimatta yhteisiä kiinnostuksen kohteita ovat esimerkiksi relevanssin tarkastelun kysymykset sekä internet-ympäristö tutkimuskohteena. Tästä syystä on hyvä tuoda yhä enemmän esille miten tiedonhankinnan ja tiedonhaun tutkimus linkittyvät ja hyötyvät toisistaan.

Kirjassa on yhdeksän informatiivista lukua, joissa käsitellään tiedonhankintatutkimusta, tiedonhakututkimusta sekä perusteluja uudelle yhtenäiselle tutkimussuuntaukselle. Kirjan painopiste on luvuissa kuusi ja seitsemän, joissa esitellään yhtenäistettyä suuntaa, sekä uuden suuntauksen vaikutuksia tutkimuskysymyksiin ja -menetelmiin. Seuraavassa esitetään tarkempi katsaus kirjan eri lukuihin.

Ensimmäinen luku on johdanto kirjoittajien näkemykseen siitä, miten kolme eri tutkimusaluetta voisivat hyötyä konkreettisemmasta integroitumisesta. Ensinnäkin on tärkeätä katsoa

miten tiedonhankintatutkimus on kehittynyt. Tämän lisäksi on tärkeätä katsoa tiedonhaun kahta pääaluetta, järjestelmäperusteista sekä käyttäjäperusteista tiedonhakua. Päämääränä on osoittaa mitä lisäarvoa nämä alueet voisivat antaa toisilleen ja miten alueita voisi jopa yhtenäistää.

Toisessa luvussa, *The Cognitive Framework For Information*, Ingwersen ja Järvelin keskustelevat kognitiivisesta viitekehuksesta suhteessa tietoon. Tarkastelu luo perustaa kirjan päämäärälle. Luku on mielenkiintoinen ja valaiseva, koska tiedon määrittäminen on koko tieteenalan jatkuva ongelma. Kognitiivinen näkökanta on kehittynyt 1970-luvulta lähtien ja sitä on monta kertaa kritisoitu siitä, että inhimillisiä tekijöitä ei ole otettu huomioon ja koska yksilöä on tarkasteltu kontekstien ulkopuolella. Kirjoittajat haluavat huomioida tämän arvostelun ja kehittää kognitiivista näkökantaa sisällyttämällä teknologisia, inhimillisiä ja kontekstuaalisia tekijöitä kokonaisvaltaisella tavalla viitekehukseensä.

Luvuissa kolme, *The Development Of Information Seeking Research*; neljä, *System-Oriented Information Retrieval* ja viisi, *Cognitive And User-Oriented Information Retrieval* kirjoittajat esittävät perusteellisen katsauksen kolmen tutkimusalueen (tiedonhankinta, järjestelmäperustainen tiedonhaku sekä käyttäjäperustainen tiedonhaku) kehityksestä vuosina 1960-2005. Vaikka kyseessä on laaja kuvaus, kirjoittajat korostavat, että kyseessä ei ole ollut kirjoittajien kaikenkattavaa katsausta ja he viittaavat esimerkiksi ARIST-julkaisun katsausartikkeleihin, jos lukija haluaa syventyä tutkimuksen eri osa-alueisiin perusteellisemmin.

Luvussa kolme tarkastellaan tiedonhankintatutkimuksen kehitystä. Tyypillisiä tutkimuskysymyksiä

olivat tarkasteluajankohdan alkuvaiheessa tiedonhankinnan suhde tietojärjestelmiin, mutta tutkimus on sittemmin kehittynyt tarkastelemaan yksilöitä, eli miten käyttäjä käyttäytyy, millainen on tiedonhankintaprosessi sekä miten relevanssia arvoidaan ja käsitteellistetään. Ajanjakson 1986-2000 aikana kehittyi teoreettinen näkemys jossa integroitiin ymmärtämystä kontekstista, työtehtävistä, tiedonhankinnasta sekä järjestelmistä. Tämä kehitys johti uusiin tiedonhankintamalleihin (mm Dervinin kontekstiorientoitunut sense-making malli, Kuhlthauin prosessimalli, Byström & Järvelinin tehtävä-perustainen malli). Kirjoittajat toteavat, että teorioiden ja käsitteiden kehitys tiedonhankintatutkimuksessa on ollut hidasta, koska monia käsitteitä on käytetty määrittelemättöminä eri muodoissa ja tilanteissa (kuten tiedontarve, tiedonkäyttö). Tutkimukselta on puuttunut myös yhteinen malli tai metateoreettinen viitekehys.

Neljännessä luvussa tarkastetaan miten järjestelmäperustainen tiedonhaku on kehittynyt. Tämä 80 sivuinen tarkastelu on pisin yksittäinen luku kirjassa mikä osoittaa tutkimusalueen laajuutta ja tutkimusaktiivisuutta. Tutkimus on kehittynyt yksittäisistä kokeellisista tutkimuksista järjestelmä- ja käyttäjäperusteiseen sekä kognitiiviseen tutkimukseen. Text retrieval, TREC-konferenssit 1990-luvusta lähtien ovat merkinneet paljon kehitettäessä tiedonhakutekniikkaa ja tutkimusta. Tutkimus on myös laajentunut kattamaan uusia alueita kuten tietämyksen muodostaminen tekstiaineistoista (text mining), kieltenvälinen tiedonhaku (cross-language information retrieval) sekä eri medioiden, kuten ääni-, musiikki-, kuva- ja videotiedonhakua.

Tiedonhaun tutkimuksen hallitsevin malli on ns. laboratoriomalli, jonka tarkastelun avulla paljastetaan tiedonhakuun liittyvät haasteet kuten relevanssiongelmat ja epätäydelliset hakutulokset. Laboratoriomalli on tutkimusalueen vahvuus, koska mallia on käytetty paljon eri tutkimuskohteissa ja se on testattu moneen otteeseen. Mallia kehittävät lukuisat tutkijat ja malli perustuu siten vankkaan empiiriseen tutkimukseen. Kirjoittajat huomauttavat kuitenkin, että tämä tilanne saattaa pysähdyttää teoreettista kehitystä monen käyttäessä samaa mallia.

Luvussa viisi tarkastetaan kognitiivista eli käyttäjäperusteista tiedonhakututkimusta. Kognitiivinen lähestymistapa tarkastelee käyttäjän roolia tiedonhakuprosessissa. Tutkimuksen alkuvaiheessa keskityttiin käyttäjään ja käyttö-

liittymään. Viime vuosikymmenten saatossa kognitiivinen lähestymistapa on alkanut integroida myös vuorovaikutteista tiedonhaun ja -hankinnan tutkimusta. Nykyinen tutkimus perustuu usein tiedonhankintaan internet-ympäristössä. Ensimmäiset mallit olivat prosessimallit joissa korostettiin hakuprosessin eri vaiheita. Monet näistä malleista kehitettiin 1990-luvulla kuten esimerkiksi tiedonhankinnan ja tiedollisen käyttäytymisen mallit (Wilson), Kuhlthauin prosessimalli, mallit hakuprosessin vaiheista (Vakkari, Wang-Soergel), sekä käsitteelliset mallit (Ingwersen, Saracevic). 1990-luvulla relevanssikysymykset, jotka ovat hyvin vaikeita, ovat saaneet yhä enemmän jalansijaa. Relevanssitarkastelun ongelmana on relevanssikäsitteen määrittely sekä käytännössä käyttäjien kyky arvioida esimerkiksi mitä määritelmä osittainen relevanssi tarkoittaa. Kognitiivisen tiedonhaun tutkimuksen ongelmana on usein ollut kontekstuaalisten näkökulmien puute. Nytemmin ollaan kuitenkin rakentamassa siltaa kognitiivisen ja järjestelmälähtöisen tutkimuksen välille esimerkiksi TREC konferensseissa.

Kuudennessa luvussa, The Integrated IS&R Research Framework, kirjoittajat rakentavat käsitteellisen viitekehysten ja mallin integroidulle tutkimussuuntaukselle, jossa tiedonhankinta- ja tiedonhakututkimus yhdentyvät. Ingwersenin malli vuodelta 1992 on tämän yhtenäistetyn viitekehysten perusta. Mallissa tarkastellaan tiedonhaku- ja hankintaprosessia rakenteellisista (IT, indeksit, tesauurukset), yksilöllisistä ja kontekstuaalisista näkökulmista. Mallissa on seitsemän eri toimijaa, erilaisilla kognitiivisilla taustoilla, jotka kommunikoivat interaktiivisessa tiedonhaussa (hankinnassa). Malli auttaa keskittämään tarkastelun koskemaan tiedonhakututkimuksessa tärkeiksi osoittautuneita keskeisiä elementtejä, eli järjestelmää, toimijaa ja kontekstia. Malli antaa myös työvälineitä analysoida kognitiivisten ja sosiaalisten elementtien suhdetta. Tiedonhankinta on monitahoinen ilmiö ja sen vuoksi on tärkeää että koko prosessia tutkitaan monesta eri näkökulmasta ja yhdistetään tietämystä eri tutkimusalueilta.

Luvussa seitsemän, Implications Of The Cognitive Framework For IS&R, Ingwersen & Järvelin tarkastelevat esittämänsä uuden viitekehysten vaikutuksia tutkimukselle. Tutkimukselle asetetaan uusia vaatimuksia ja kirjoittajat tuovat esille erityisesti viisi tutkimusulottuvuutta, jotka yhtenäistävät tiedonhankinta- ja tiedonhakututkimusta.

Kyseessä on organisaatio – työtehtävä, toimija (hakija, kirjoittaja, indeksoija jne.), dokumentti, algoritmit sekä saatavuus ja interaktiivisuus. Näillä osa-alueilla on tiedonhankinnassa ja -haussa ollut eri painopisteitä, ja osa-alueet pyritään liittämään yhteiseen viitekehykseen. Näitä mahdollisia tutkimusohjelmia esitetään luvussa kahdeksan, Towards a research program, ja viimeisessä, yhdeksännessä luvussa esitetään kirjan yhteenveto.

Kirjalla on kunnianhimoinen päämäärä, jonka kirjoittajat Peter Ingwersen ja Kalervo Järvelin pystyvät hyvin täyttämään. He tarjoavat tutkijoille analyyttisiä työvälineitä, joiden avulla tiedonhankinta- ja tiedonhauntutkimusta integroivaa viitekehystä pystytään käsittelemään yhtenäistetyssä suuntauksessa. Kirjassa keskus-

tellaan keskeisistä käsitteistä kuten tiedosta ja tiedon kognitiivisesta näkökulmasta. Kirjan katsaukset kolmesta eri tutkimusalueesta ovat hyviä ja antavat paljon tietoa tutkimusalueiden vahvuuksista ja heikkouksista. Kirjoittajat osoittavat mitä tutkimuksen alueita pitäisi kehittää edelleen, jotta esimerkiksi tiedonhankintatutkimus saavuttaisi analyyttisemmän ja käsitteellisemmän tason. Vaikka kirjoittajat eivät pyri kaikenkattavaan kuvaukseen olisi tähän yhteyteen voinut lisätä esimerkiksi arkielämän tiedonhankinnan malleja. Kirja *The Turn – Integration Of Information Seeking An Retrieval In Context* sopii moneen eri käyttöön. Se on tärkeä tutkimuksen kehitystä varten, mutta sopii myös hyvin peruskirjaksi alan koulutukseen.