

**Mikko Pennanen**  
**Pertti Vakkari**

# Ongelman jäsentymisen yhteys tiedonhaun muutokseen tehtäväprosessin aikana

Mikko Pennanen & Pertti Vakkari, Ongelman jäsentymisen yhteys tiedonhaun muutokseen tehtäväprosessin aikana. *Informaatiotutkimus* 19 (1), : 3-10. [The effects of problem structure on information searching during task performance process]

The aim of this study was to analyse how changes in user's cognition and problem stages during task performance process are related to changes in search tactics and term choices. Eleven students majoring in information studies were investigated during a four-month seminar. The students were expected to complete a research proposal at the end of the seminar. The students were asked to make an IR-search in LISA-database at the beginning, middle and end of the seminar. Data for describing their understanding of the task, search goals and tactics as well term choices were collected during the search sessions. A pre- and post-search interview was conducted during each session. The students were asked to think aloud during the sessions. The students were asked to think aloud during the sessions. The transaction logs were captured and think alouds were recorded.

Students mainly proceeded in their work according to Kuhlthau's model of the information seeking process. The result show that the stages of task performance have an effect on term choices. Information searches became longer and they includes more facets and terms by the end of the seminar. Especially the number of narrowing and related terms increased. Students also used more operators at the end of the seminar (especially the use of OR-operator increased).

*Address: University of Tampere, Department of Information Studies, FIN-33014 University of Tampere, Email mikko.pennanen@uta.fi, pertti.vakkari@uta.fi*

## Johdanto

Informaatiotutkimus jakaantuu perinteisesti kahteen eri tutkimusalueeseen: tiedonhaun ja tiedonhankinnan tutkimukseen. Tutkimussuuntauksien välillä ei ole juurikaan ollut yhteistyötä, vaikka monet tutkijat ovat korostaneet yhteistyön ja tutkimustulosten integroinnin tarvetta (ks. esim. Bates 1989; Belkin 1993; Ingwersen 1992; Kuhlthau 1993; Marchionini 1995). Tiedonhakua ei ole yhdistetty siihen tehtävään, joka on saanut aikaan henkilön tiedonhankintakäyttäytymisen. Tiedonhakua ei ole myöskään

tarkasteltu prosessinäkökulmasta. Perinteisesti tiedonhaun tutkimuksessa on tarkasteltu vain yhtä yksittäistä hakua. Usein kuitenkin tiedonhakua tehdään useaan kertaan tehtävän suorittamisen aikana (vrt. Spink et al. 1998). Yksittäisen haun tarkastelu ei siis anna kokonaiskuvaa tiedonhakijan toiminnasta ja hakijan tavoitteiden muuttumisesta prosessin aikana (vrt. Hert 1996). Tiedonhankintaprosessia koskevan kokonaiskuvan saaminen edellyttää tämän vuoksi tiedonhan-

kinnan ja tiedonhaun tutkimussuuntauksien integroimista. Yksittäisen hakuistunnon sijaan tiedonhaku tulee tarkastella osana koko tehtäväprosessia. Tehtäväprosessin tutkiminen auttaa paremmin ymmärtämään tiedonhakuun liittyviä ilmiöitä ja niissä tapahtuvia muutoksia. (Vakkari 1999, 35-37.)

Tämä tutkimus yhdistää tiedonhankinnan ja haun tutkimusta tarkastelemalla kompleksin tehtävän suorittamisen aikana tapahtuvaa ongelman jäsentymistä ja sen yhteyttä tiedonhakatietokoneissa tapahtuviin muutoksiin. Tutkimus on yksi ensimmäisistä, jossa tiedonhakuja tarkasteltiin osana jonkin tehtävän suorittamista.

## Teoreettinen viitekehys

Tutkimuksen teoreettisena viitekehystenä hyödynnettiin Vakkarin (1999) esittämää mallia ongelman jäsentymisen ja tietotoimintojen välisestä yhteydestä. Vakkarin mallissa erilaiset tietotoiminnot (tiedonhakustrategiat, tietotyypit ja relevanssiarviot) kytketään tehtävän kompleksisuuteen ja ongelman jäsentymiseen.

Tehtävän kompleksisuudella tarkoitetaan sen ennalta määräytyvyyden astetta. Tehtävän ennalta määräytyvyys koskee tehtävän tietovaatimuksia, suoritusprosessia ja tulosta. Tässä tutkimuksessa tehtävän kompleksisuus vakioitiin. Tutkittavien henkilöiden tehtävänä oli laatia pro gradu-työn tutkimussuunnitelma annetun ajan kuluessa, mitä voidaan pitää vaativana työtehtävänä.

Ongelman jäsentyneisyyden asteella tarkoitetaan sitä, missä määrin henkilö on tietoinen ongelmaan liittyvistä erilaisista muuttujista ja näiden välisistä suhteista. Heikosti jäsentyneille ongelmille on tyypillistä se, että ongelmanratkaisija ei ole kyennyt tunnistamaan ongelman keskeisiä elementtejä ja niiden välisiä suhteita. (Partridge & Hussain 1995, 82.) Jäsentyneillä ongelmilla on puolestaan fokus eli henkilö on muotoillut oman henkilökohtaisen näkemyksen tarkastelun kohteena olevasta asiasta. Fokuksen hahmotus on keskeinen osa ongelman jäsentämistä. Fokuksen muotoiluun jälkeen henkilö tuntee ongelman keskeiset elementit ja niiden väliset suhteet. Ongelman jäsentyneisyyden asteen kasvaessa henkilön kyky määrittellä se, minkälaista tietoa hän tarvitsee ongelman ratkaisemiseksi, minkälaisia tiedonhakustrategioita hänen tulee soveltaa tarvittavan tiedon hankkimiseksi ja mitä prosesseja tarvitaan ongelman ratkaisemiseksi lisääntyy. (Vakkari 1999).

Ongelman jäsentyneisyyden astetta arvioitiin Kuhlthau (1993) esittämän tiedonhankinnan prosessimallin avulla. Kuhlthau (1993, 41-52) jakaa tiedonhankintaprosessin kuuteen eri vaiheeseen: aloitusvaiheeseen, valintavaiheeseen, tunnusteluvaiheeseen, muotoilu-

vaiheeseen, keruuvaiheeseen sekä esitysvaiheeseen. Malli liittyy yhteen hakijan kokemat tunnetilat, toiminnot sekä ajatukset. Tiedonhankintaa pidetään «järkeistämisen prosessina», jossa erilaiset ongelman jäsentyneisyyden asteet ja tunnetilat vaihtelevat. Tiedonhankinta lähtee liikkeelle hakijan tuntemasta epävarmuuden tilasta ja päättyy parhaimmillaan tapauksessa ongelmallisten asioiden jäsentymiseen ja ymmärtämiseen tai vaihtoehtoisesti pahimmillaan entistä vaikeampaan epävarmuuden ja ahdistuksen tunteeseen. Tiedonhankinta on siis matka kohti jäsentyneempää kuvaa ongelmasta. Käytännössä tiedonhankintaprosessi ei yleensä etene täysin suora- viivaisesti mallin mukaan, vaan henkilö saattaa olla useassa eri vaiheessa samanaikaisesti tai vaiheiden järjestys voi vaihdella. Usein tiedonhankinta etenee lähinnä syklistä (vrt. esim. Swain 1996).

## Tutkimustehtävä ja tutkimuksen suorittaminen

Tutkimuksen tehtävänä oli tarkastella kompleksin tehtävän suorittamisen aikana tapahtuvaa ongelman jäsentymistä sekä sen vaikutusta tiedonhakijoiden hakuun sisällyttämisen fasettiin, hakuterminien ja operaattoreiden valintaan. Tutkimukseen osallistui yhteensä 11 informaatiotutkimuksen pääaineopiskelijaa, jotka osallistuivat kevään 1999 aikana pidettyihin tutkimusseminaareihin. Opiskelijoiden tehtävänä oli seminaarin aikana laatia ja esittää oman pro gradu-työn tutkimussuunnitelma.

## Tutkimuksen aineistonkeruu

Opiskelijoilta kerättiin tietoa kolmeen eri kertaan tehtäväprosessin aikana. Aineistonkeruu tehtiin seminaarin alussa, keskivaiheilla ja sen päättyessä. Näin pyrittiin saamaan tietoa opiskelijoiden tiedonhausta ennen tehtävän fokuksen muotoutumista, sen muotoutuessa ja muotoutumisen jälkeen. Tärkeimpiä aineistonkeruumenetelmiä olivat haastattelu ja tutkittavien haastattelutilanteessa oman tutkimuksensa aiheesta tekemä tiedonhaku. Tiedontarpeet olivat näin todellisia ja ajankohtaisia.

Haastattelun ja tiedonhaun lisäksi opiskelijat pitivät tutkimusseminaarin aikaisista toimistaan tiedonhaku- ja tutkimuspäiväkirjaa. Tutkimuspäiväkirjassa tutkittavat saivat vapaasti kertoa tutkimuksen tekemiseen liittyvistä ajatuksista, toiminnoista ja tunnetiloista. Tiedonhankintapäiväkirjassa tutkittavia pyydettiin merkitsemään ylös heidän löytämänsä tärkeimmät lähteet sekä se, mistä he olivat saaneet tietää lähteistä. Haastattelutilanteessa tutkittavat

täyttivät myös ongelman jäsenyneyssyden arviointiin käytettävän Kuhlthaun Process Survey lomakkeen. Viimeisellä haastattelukerralla tutkittavat piirsivät lisäksi aikajanana koko tutkimusprosessin kulusta ja antoivat oman arvionsa ongelman jäsenyneyssydestä. Koko kevään aikana tehtiin yhteensä 33 haastattelua, joissa tehtiin 32 tiedonhakua. Yksi haastateltava ei tehnyt hakua viimeisellä haastattelukerralla.

Haastattelumenetelmänä sovellettiin teemahaastattelua. Haastattelun aluksi tutkittava täytti ongelman jäsenyneyssyden arviointiin käytetyn Kuhlthaun (1993) Process Survey lomakkeen. Tämän jälkeen tutkittavalta kysyttiin hänen tutkimuksen tekoa koskevia tuntemuksiaan. Seuraavaksi haastateltava teki tiedonhaun informaatiotutkimuksen alan keskeisestä tietokannasta, LISA:sta. Haun kulku tallennettiin kokonaisuudessaan lokitiedostoon. Hakijaa pyydettiin ajattelemaan ääneen haun aikana ja puhe nauhoitettiin. Tutkijat tekivät samalla myös havaintoja hakijoiden toiminnasta. Hakijoilla ei ollut haun aikana käytössään tesaurusta, koska tämän uskottiin vaikeuttavan päätelmien tekoa siitä, johtuivatko haun muutokset ongelman jäsenymisestä vaiko tesauruksen tarjoamasta valmiista asiahierarkiasta. Haun kohteena ollut tietokanta pidettiin myös vakiona. Muutoin haun suoritustapa oli vapaa.

## Aineiston analyysi

Ongelman jäsenyneyssyden tarkasteluun käytettiin lähinnä Process Survey-lomakkeita sekä haastattelua. Henkilön tyypittäminen johonkin tiettyyn vaiheeseen oli käytännössä vaikeaa, koska monilla henkilöillä vaiheet menivät osittain päällekkäin erityisesti kolmannella haastattelukerralla. Kuhlthaun mallin mukaisten kuuden vaiheen sijaan aineistosta eriteltiin kolme vaihetta: pre-fokus-, fokuksen hahmotus- ja post-fokus-vaiheet.

Tiedonhaussa käytettävien termien, fasettien ja operaattoreiden analyysissä käytettiin tiedonhakutallennetta ja haastattelua. Tiedonhaun tuottama hakutallenne tulostettiin paperille jokaisen haun jälkeen. Nauhoitusten purkamisen jälkeen tiedonhakutallenne ja purettu haastattelu yhdistettiin analyysin helpottamiseksi työpaperiksi, jossa hakijan ääneen ajattelu ja vuorovaikutus hakujärjestelmän kanssa esitettiin yhdessä. Jokaisen kerran haut analysoitiin erikseen henkilöittäin. Työpapereista laskettiin kunkin hakijan käyttämien hakulausekkeiden, fasettien sekä hakutermin ja operaattoreiden määrät. Varsinainen tulosten kokoaminen ja tarkastelu tehtiin koko ryhmän tasolla.

Fasetit määriteltiin tarkastelemalla koko haussa esiintyneitä termejä ja niiden välisiä yhteyksiä (ts. mitä näkökulmaa termit edustavat) (vrt. Järvelin 1995). Fasetit käytiin läpi henkilöittäin erikseen jokaisen hakukerran osalta.

Termit jaettiin synonyymeihin sekä hakua laajentaviin, hakua kaventaviin ja rinnakkaisiin termeihin (vrt. Wang 1997). Termien tyypittely tehtiin vertailemalla niitä edellisellä hakukerralla esiintyneisiin termeihin. Tarvittaessa termien välisistä suhteista laadittiin käsittekartta.

Operaattoreista laskettiin jokaisen operaattorityypin (AND, OR, NOT) kokonaismäärät haussa. Laskussa otettiin huomioon jokainen hakuun uutena tuotu operaattori. Läheisyysoperaattoreita ei tässä tutkimuksessa laskettu lainkaan.

## Tulokset

### Ongelman jäsenyminen tiedonhankintaprosessin aikana

Hakijoiden ongelma jäsenyi yleisesti tutkimusajanjakson kuluessa. Ongelman jäsenymisessä esiintyi tosin jonkin verran vaihtelua eri henkilöillä. Kaikki kuitenkin etenivät työssään kevään aikana ja muodostivat tutkimukselleen fokuksen, mikä on edellytyksenä koko tutkielman teolle.

Ensimmäisellä hakukerralla tutkittavat sijoituivat Kuhlthaun (1993) tiedonhankinnan prosessimallin kolmeen ensimmäiseen vaiheeseen. Henkilöt mainitsivat I hakukerralla tärkeimmiksi tehtävikseen aihepiiriin yleisen tutkimisen, tiedonlähteiden selailun sekä alustavien tiedonhakujen tekemisen. Tunnetilojaan tutkittavat kuvasivat optimistisiksi ja luottavaisiksi, mutta osin myös hieman hämmentyneiksi ja epävarmoiksi. Tärkeimpinä tehtävinään hakijat pitivät aihepiiriin yleistä tutkimista, tiedonlähteiden selailua sekä alustavien tiedonhakujen tekoa.

Toinen hakukerta sijoittui maaliskuun alkuun. Tällöin kaikki tutkittavat olivat edenneet oman ongelman jäsentämisessä. Henkilöt olivat pääsääntöisesti hahmotamassa oman ongelmansa fokuksia tai olivat jo sen tehneet. Tutkittavat pitivät tärkeimpinä tehtävinään työn tarkan aiheen muotoilua sekä aiheeseen liittyvän yleisen tiedon tutkimista. Henkilöt kuvailivat tunteitaan lähinnä optimistisiksi ja luottavaisiksi, mutta myös hieman hämmentyneiksi ja epävarmoiksi.

Kolmas hakukerta ajoittui huhti-toukokuuhun. Tuolloin henkilöiden sijoittaminen Kuhlthaun eri vaiheisiin oli I ja II kertaa vaikeampaa. Kuhlthaun mallin periaatteiden mukaan kaikkien hakijoiden olisi pitänyt olla viimeisessä vaiheessa (esitysvaihe). Osalla hakijoista työn fokuksen hahmotus oli kuitenkin vielä osittain menossa ja osa oli selvästi vielä tiedon-keruuvaiheessa. Osalla henkilöistä työn laatimisen ja hakujen perspektiivi oli osittain muuttunut: tutkittavat miettivät jo koko pro gradu-tutkielman tekoa, eivät enää

pelkän seminaariesitelmän kirjoittamista. Tutkittavat pitivät III hakukerralla tärkeimpinä tehtävinään tarkkan aiheen muotoilemista koko tutkielmaa varten, tarkkaan aiheeseen liittyvän tiedon keruuta sekä esitelmän viimeistelyä. Tunnetilojaan haastateltavat kuvasivat luottavaisiksi ja helpottuneiksi. Osa sanoi kuitenkin tuntevensa myös pientä epävarmuuden tunnetta koko pro gradu-tutkielman teon suhteen.

## Ongelman jäsentymisen vaikutus tiedonhakuun

### Hakujen kokonaiskuva

Tutkittavat tekivät kevään 1999 aikana tiedonhakuja kolmella eri haastattelukerralla. Hakuja tehtiin koko kevään aikana yhteensä 32. Koko kevään aikana hakuistunnoissa käytettiin yhteensä 191 hakulauseketta, haun käsitteellistä sisältöä kuvattiin 92 fasetilla, 129 hakutermillä ja 126

operaattorilla.

Hakijoiden aloitusvaiheessa tekemät tiedonhaut olivat varsin lyhyitä (ks. taulukko 1.). Hakujen lyhyyttä ensimmäisellä kerralla selittää luultavasti osaltaan se, että osa hakijoista oli tottumattomia tutkimuksen kohteena olleen tietokannan käyttöön.

Hakijoiden ongelman jäsentymisen lisäksi sekä hakujen pituutta että haun käsitteellisen sisällön rikkautta. Hakujen pituus kasvoi eniten II ja III hakukerran välillä hakijoiden ollessa tiedonkeruu- ja esitysvaiheissa. Haun käsitteellisessä rikastumisessa oli pieniä eroja eri hakukertojen välillä. I ja II hakukerran välillä haun käsitteellisen sisällön rikastuminen tuli esille sekä fasettien että termien määrän kasvuna. Opiskelijat pyrkivät aluksi jäsentämään aiheitaan laajoiksi kategorioksi sisällyttämällä hakuun mahdollisimman paljon aiheita koskevia näkökulmia. Hakijat eivät vielä pystyneet kuvaamaan aiheitaan eriytyneesti. Laajojen kategorioiden hahmottaminen luultavasti auttoi oman tutkimuksen fokuksen muotoilua. II ja III hakukerran välillä opiskelijoiden oman tutkimuksen fokuksen muotoilun jälkeen puolestaan

Taulukko 1. Tiedonhakujen kokonaiskuva eri hakukerroilla

	I HAKUKERTA		II HAKUKERTA		III HAKUKERTA	
	N	KA	N	KA	N	KA
Hakulausekkeet	51	4,6	61	5,6	79	7,9
Fasetit	23	2,1	33	3,0	36	3,6
Termit	32	2,9	44	4,0	53	5,3
Operaattorit	28	2,6	46	4,2	52	5,2

Taulukko 2. Fasettien muutos kevään aikana

	I hakukerta		II hakukerta		III hakukerta		
	Kaikki	Kaikki	Uudet	I kerralta pudotetut	Kaikki	Uudet	II kerralta pudotetut
N	23	33	25	15	36	23	16
KA	2,1	3,0	2,3	1,4	3,6	2,3	1,6
%	100 %	100 %	76 %	65 %	100 %	64 %	55 %

Taulukossa pudotettujen fasettien suhteellinen osuus on laskettu edellisen hakukerran fasettien määrästä. III hakukerralla hakijoita oli muita kertoja yksi vähemmän, mikä on otettu huomioon pudotettujen fasettien suhteellista osuutta laskettaessa vähentämällä II hakukerran fasettien kokonaismäärästä poisjääneen hakijan neljä fasettia.

tiedonhaussa käytettyjen termien määrä lisääntyi fasettien määrää enemmän. Lisääntynyt tietämys siis auttoi hakijoita kuvaamaan aihettaan runsaammin termein.

### Fasetit

Haussa mukana olevat fasetit kuvaavat niitä näkökulmia, joita tiedonhakijat ovat ottaneet hakuun mukaan (vrt. Järvelin 1995). Hakuun mukaan otettujen fasettien määrä kasvoi ongelman jäsenyessä, mutta muutokset eivät yleisesti olleet suuria (ks. taulukko 2.). Hakijoilla oli loppukeväällä tiedonkeruu- ja esitysvaiheessa haussa mukana keskimäärin 1,5 fasettia enemmän kuin heidän aloittaessa tehtävän suoritusta. Sen sijaan faseteissa tapahtui varsin suurta vaihtumista eri kertojen väleillä. II hakukerralla hakijat olivat pudottaneet lähes kaksi kolmasosaa edellisessä haussa mukana olleista faseteista. III hakukerralla (tiedonkeruu- ja esitysvaihe) samoina pysyneiden fasettien suhteellinen osuus lisääntyi hieman II hakukertaan verrattuna. Tosin myös III kerralla valtaosa faseteista oli uusia. II hakukerran faseteista hakijat olivat pudottaneet hieman yli puolet. Ongelman jäsenytyminen siis lisäsi hakujen monipuolisuutta, mikä kertoi hakijoiden aihetta koskevan erottelukyvyn kasvusta

sekä ongelman luonnetta koskevan käsityksen vakiintumisesta.

### Hakutermit

Hakijoiden hakutermin määrä lisääntyi tutkimusprosessin edetessä ja ongelman jäsenyessä. (ks. taulukko 3.). Termien määrän kasvu oli eri kertojen välillä varsin tasaista. Sen sijaan termien vaihtuminen oli II ja III hakukerralla fasettien vaihtumisen tavoin varsin suurta. Hakijat pudottivat sekä II että III hakukerralla 59 prosenttia edellisen hakukerran termeistä. Hakuun uutena tuotujen termien osuus oli II hakukerralla 70 prosenttia ja III kerralla 74 prosenttia. Valtaosa uusista hakutermeistä oli sekä II että III hakukerralla rinnakkaisia termejä (ks. taulukko 4.). Hakijat eivät juurikaan tuoneet hakuun uusia laajentavia termejä. Synonyymien ja rinnakkaisten termien suhteellinen osuus pysyi lähes samana sekä II että III hakukerralla. Kaventavien termien osuus sen sijaan kasvoi hieman III hakukerralla. Rinnakkaisten termien suuri määrä III hakukerralla selittyi luultavasti osittain sillä, että osa hakijoista oli jo ehtinyt esittää oman työnsä ennen III kerran hakuja ja he ajattelivat hakutilanteessa koko tutkielman laatimista.

Tulokset osoittavat, että hakutermin määrä lisääntyi

Taulukko 3. Termien muutos kevään aikana

	I hakukerta		II hakukerta			III hakukerta		
	Kaikki		Kaikki	Uudet	I kerralta pudotetut	Kaikki	Uudet	II kerralta pudotetut
N	32		44	31	19	53	39	20
KA	2,9		4,1	2,9	1,7	5,3	3,9	2,0
%	100 %		100 %	70 %	59 %	100 %	74 %	59 %

Taulukossa pudotettujen termien suhteellinen osuus on laskettu edellisen hakukerran termien kokonaismäärästä. III hakukerralla hakijoita oli muita kertoja yksi vähemmän, mikä on otettu huomioon pudotettujen termien suhteellista osuutta laskettaessa vähentämällä II hakukerran termien kokonaismäärästä poisjääneen hakijan kymmenen hakutermiä.

Taulukko 4. Uusien termien jakautuma kevään aikana

	II hakukerta					III hakukerta				
	Uudet yht.	ST	NT	BT	RT	Uudet yht.	ST	NT	BT	RT
N	31	6	8	3	14	39	7	11	3	18
KA	2,8	0,5	0,7	0,3	1,3	3,9	0,7	1,1	0,3	1,8
%	100 %	19 %	26 %	10 %	45 %	100 %	18 %	28 %	8 %	46 %

Taulukon lyhenteet: ST=synonyymit, BT=laajemmat termit, NT=kapeammat termit, RT=rinnakkaiset termit.

ongelman jäsenyydessä ja hakujen monipuolistuessa ja erityydessä. Monipuolisuus ja hakujen eriytyminen tuli esille rinnakkaisten termien ja kaventavien termien määrän kasvuna. Synonyymien osuus hakuun uutena tuoduista termeistä pysyi lähes samana II ja III hakukerralla. Ongelman fokuksen muotoilun jälkeen tietoa haettiin siis lähinnä joko samoilla tai kapeammilla käsitteillä taikka rinnakkaisilla hakutermeillä.

## Termit/fasetti

Fasettiin kuuluvien termien määrän voidaan katsoa kuvaavan ongelman jäsentymistä: mitä enemmän fasetissa on termejä, sitä tietoisempi henkilö on ongelman osa-alueista. Alkuperäisenä hypoteesinä oli, että termien määrä fasettia kohden lisääntyy prosessin kuluessa. Hypoteesi ei kuitenkaan saanut tässä tutkimuksessa tukea. Fasettiin kuuluvien termien määrässä ei juuri tapahtunut muutosta I ja II hakukerralla. Termien määrä kasvoi hieman hakijoiden tullessa tiedonkeruu- ja esitysvaiheeseen (III-hakukerta), mutta kasvu oli niin pientä ettei sillä ollut juurikaan merkitystä (ks. taulukko 5.). Ongelman jäsentymisellä ei tässä tutkimuksessa siis ollut juurikaan vaikutusta faseteissa olevien termien määrään.

## Operaattorit

Ongelman jäsentyminen tuli esille fasettien ja niihin kuuluvien termien lisäksi myös termejä ja fasetteja yhdistävissä operaattoreissa. Operaattoreista AND ja NOT ovat hakua kaventavia operaattoreita. Kaventavien operaattorien käytöllä tiedonhakija sulkee hausta pois sellaisia käsitteitä, jotka eivät kuulu ongelman piiriin. Haun kaventaminen ja ei-toivottujen käsitteiden poistaminen hausta kertoo siitä, että henkilö hahmottaa ongelmaa paremmin. AND-operaattoria käytetään tiedonhaussa yleensä fasettien välillä niiden yhdistämiseen toisiinsa. OR-operaattori on taas hakua laajentava operaattori. OR-operaattori osoittaa ongelman jäsentyneisyyttä silloin, kun sitä käytetään fasetin sisällä. Hakijan laajentaessa fasettiin kuuluvaa aihetta uusilla termeillä hän on tietoinen ongelmaan liittyvistä erilaisista muuttujista. OR-operaattorin käyttö fasettien välillä puolestaan kertoo siitä, että hakija ei ole hahmottanut ongelmaa vielä hyvin. OR-operaattoria käytetään yleensä fasetin sisällä. OR-operaattorin käyttö fasettien välillä on harvinaista. (vrt. Harter 1986, 173.)

Ensimmäisellä hakukerralla 90 prosenttia operaattoreista oli hakua kaventavia AND-operaattoreita (ks. taulukko 6.). OR-operaattorin käyttö oli vielä I hakukerralla vähäistä. Tämä kertoo siitä, että hakijat eivät olleet vielä

Taulukko 5. Termien määrä fasettia kohti eri hakukerroilla

	Termit	Fasetit	Termit/fasetti
I hakukerta	32,0	23,0	1,39
II hakukerta	44,0	33,0	1,33
III hakukerta	53,0	36,0	1,47

Taulukko 6. Operaattorien jakautuma eri hakukerroilla

	I hakukerta				II hakukerta				III hakukerta			
	AND	OR	NOT	Yht.	AND	OR	NOT	Yht.	AND	OR	NOT	Yht.
N	26	3	0	29	35	9	2	46	36	14	2	52
KA	2,4	0,2	0,0	2,6	3,2	0,8	0,2	4,2	3,6	1,4	0,2	5,2
%	90 %	10 %	0 %	100 %	76 %	20 %	4 %	100 %	67 %	26 %	7 %	100 %

kunnolla hahmottaneet ongelmaa koskevia kaikkia osa-alueita.

Operaattorien käyttö lisääntyi II hakukerralla hakijoiden lähestyessä oman työnsä fokuksen hahmottamista. Valtaosa operaattoreista oli myös II hakukerralla hakua kaventavia AND-operaattoreita, mutta OR-operaattorien käyttö lisääntyi. Tämä kertoo ongelman jäsentymisestä siten, että hakijat osasivat II hakukerralla tunnistaa ongelman osa-alueita I kertaa paremmin ja että he halusivat hauissa varmistua kaikkien ongelmaa kuvaavien alueiden mukaantulosta hakuun.

Operaattorien määrän kasvu jatkui III hakukerralla hakijoiden siirryessä tiedonkeruu- ja esitysvaiheeseen. Valtaosa operaattoreista oli myös III hakukerralla hakua kaventavia AND-operaattoreita, mutta niiden osuus kaikista operaattoreista oli laskenut kahteen kolmasosaan. OR-operaattorien käyttö kasvoi III hakukerralla hieman verrattuna edelliseen hakukertaan. OR-operaattorien käytön lisääntyminen voi kertoa toisaalta ongelman jäsentymisestä tai siitä, että hakijat halusivat III hakukerralla laajentaa hakua ajatellen koko tutkielman tekemistä.

Ongelman jäsentyminen lisäsi myös operaattorien käytön monipuolisuutta (ks. taulukko 7).

I ja II kerralla valtaosa hakijoista käytti ainoastaan yhtä operaattorityyppiä (lähinnä AND-operaattoria). III kerralla, hakijoiden ollessa tiedonkeruu- ja esitys-vaiheessa, operaattorien käyttö oli hieman aikaisempia hakukertoja monipuolisempaa. Tämä selittyy OR-operaattorien lisääntyneellä käytöllä III hakukerralla. Ongelman jäsentyessä hakijoiden kyky kuvata aiheitaan useammin termein siis kasvoi.

## Johtopäätökset

Opiskelijoiden ongelman jäsentyminen eteni pääasiassa Kuhlthaun (1993) tiedonhankinnan prosessimallin mukaisesti. Tutkittavat olivat tehtävää aloittaessaan epävarmoja tulevan tehtävän suunnasta. Prosessin kuluessa he siirtyivät kohti varmempaa ja selkeämpää kuvaa omasta tutkimusaiheestaan. Kaikki tutkittavat hahmottivat työilleen fokuksen ja kaikki esittivät oman seminaariesitelmänsä. Loppukeväästä heidät voitiin sijoittaa post-fokus-vaiheeseen.

Tutkimuksen hypoteesi ongelman jäsentymisen vaikutuksesta hakijan tekemiin tiedonhakuihin sai tukea. Tutkittavien ongelman jäsentyessä tiedonhakujen pituus kasvoi ja niiden käsitteellinen ala rikastui. Hakujen piteneminen tuli esille hakulausekkeiden määrän lisääntymisenä. Käsitteellisen sisällön rikastuminen näkyi puolestaan fasettien ja hakutermien määrän kasvuna sekä operaattorien käytön monipuolistumisena. Hakutermeistä erityisesti rinnakkaisten ja kaventavien termien määrä kasvoi.

Tulokset vastaavat Wangin tutkimuksen tuloksia. Myös Wangin tutkimuksessa kaventavien ja rinnakkaisten termien käyttö lisääntyi prosessin kuluessa (vrt. Wang 1997). Rinnakkaisten termien käytön lisääntyminen sekä kapeampien termien tuominen hakuun kertoo hakujen monipuolistumisesta ja eriytymisestä. Ongelman jäsentyminen auttoi siis tiedonhakijaa kuvaamaan hakuaiheitaan entistä täsmällisemmin. Kaventavilla termeillä hakijat lisäsivät haun tarkkuutta, rinnakkaisilla termeillä

Taulukko 7. Eri operaattorityyppiä käyttäneiden henkilöiden määrät eri hakukerroilla

	I hakukerta	II hakukerta	III hakukerta
Ei operaattoreita	1	0	0
Yksi operaattorityyppi	8	8	3
Kaksi operaattorityyppiä	2	1	6
Kolme operaattorityyppiä	0	2	1
YHTEENSÄ	11	11	10

puolestaan haun kattavuutta eli hakijat pyrkivät varmistumaan siitä, että kaikki ongelmaa koskevat osat alueet tulivat mukaan hakuun. Ongelman jäsentymässä myös hakijoiden operaattorien käytön monipuolisuus lisääntyi.

Tämä tutkimus lisää tietoa tehtäväprosessin aikana tehtävästä tiedonhausta. Tiedonhaun kehittymistä prosessin edetessä ei ole juurikaan tutkittu. Tulosten yleistäminen ei ole kuitenkaan mahdollista, koska tarkastellun ryhmän koko oli varsin pieni: 11 opiskelijaa. Kyse on siis tapaustutkimuksesta.

Ongelman jäsentymisen ja tiedonhaussa tapahtuvien muutoksien tutkimuksessa on kyse teoreettisesta perustutkimuksesta, jonka tuloksia ei voida suoraan soveltaa käytäntöön. Tiedon määrän kasvaessa tarvitaan lisää perustutkimusta siitä, kuinka henkilöt hankkivat ja käyttävät tietoa erilaisissa tehtävissä, tilanteissa ja ympäristöissä. Tiedonhaun ja tiedonhankinnan tutkimuksessa tulisi kiinnittää huomiota erityisesti tiedonhankinnan prosessiluonteeseen ja tiedonhankinnan asemaan osana muuta henkilön toimintaa kuten oppimista. Parempien tietopalveluiden ja tietojärjestelmien kehittäminen ei ole mahdollista ilman niiden kysyntää ja käyttöä koskevan perustan tuntemusta.

Hyväksytty julkaistavaksi 10.4.2000

## KIRJALLISUUS

Bates, M. (1989) *The Design of Browsing and Berrypicking Techniques for the Online Search Interface*. *Online Review* 13(5), s. 407–424.

Belkin, N. (1993) *Interaction with Texts: Information Retrieval as Information-Seeking Behavior*. Teoksessa: *Information retrieval '93: von der Modellierung zur Anwendung*. Konstanz: Universitätsverlag Konstanz, 55–66.

Harter, S. P. (1986) *Online Information Retrieval*. Orlando: Academic Press.

Hert, C. (1997) *Understanding Information Retrieval Interactions: Theoretical and Practical Implications*. London: Ablex.

Ingwersen, P. (1992) *Information Retrieval Interaction*. London: Taylor Graham.

Järvelin, K. (1995) *Tekstitiedonhaku tietokannoista*. *Asiantuntija-sarja*. Espoo: Suomen ATK-kustannus.

Kuhlthau, C. (1993), *Seeking Meaning. A Process Approach to Library and Information Services*. Norwood, NJ: Ablex.

Marchionini, G. (1995) *Information Seeking in Electronic Environments*. Cambridge University Press.

Partridge, D. & Hussain, K. (1995) *Knowledge Based Information Systems*. London: McGraw-Hill.

Pennanen, M. (1999) *Ongelman jäsentymisen yhteys tiedonhakutaktiikoiden muutokseen*. Tutkimus opiskelijoiden tiedonhausta tutkimussuunnitelman laadinnan aikana. Informaatiotutkimuksen pro gradu-tutkielma. Tampereen yliopisto, informaatiotutkimuksen laitos. Tutkielma on luettavissa kokotekstinä Internetissä osoitteessa

<http://www.uta.fi/~mp58522/infogradu.html>.

Spink, A. & Wilson, T. & Ellis, D. & Ford, N. (1998), *Modeling Users Successive searches in Digital Environments*. A National Science Foundation / British Library Funded Study. *D-Lib Magazine*, March 1998. hdl:cnri.dlib/april98-spink.

Swain, D. (1996) *Information Search Process Model: How Freshmen Begin Research*. *ASIS-96. Proceedings of the 59th ASIS Annual Meeting*, vol. 33, s. 95–99.

Vakkari, P. (1999) *Task Complexity, Information Types, Search Strategies and Relevance: Integrating Studies on Information Seeking and Retrieval*. *Information Processing & Management* 35 (6): 819–838.

Wang, P. (1997) *User's Information Needs at Different Stages of a Research Project*. A Cognitive view. Teoksessa: Vakkari, P. & Savolainen, R. & Dervein, B. (toim.) *Information Seeking in Context*. London & Los Angeles: Taylor Graham.