

Simuloidun tehtävän vaikutus  
kyselyn muodostukseen:  
INEX 2004-hankkeen  
vuorovaikutteisen tiedonhaun tutkimuslinja

Mervi Nurmela  
Tampereen yliopisto  
Informaatiotutkimuksen laitos  
Pro gradu –tutkielma  
Marraskuu 2006

## TIIVISTELMÄ

### TAMPEREEN YLIOPISTO

Informaatiotutkimuksen laitos

NURMELA, MERVI

Simuloidun tehtävän vaikutus kyselyn muodostukseen: INEX 2004-hankkeen vuorovaikutteisen tiedonhaun tutkimuslinja

Pro gradu –tutkielma, 65 s., 5 liites.

Informaatiotutkimus

Marraskuu 2006

---

Tutkielmassa tarkastellaan kyselyn muodostukseen liittyvää hakukäyttäytymistä vuorovaikutteisessa tiedonhakutilanteessa. Vuorovaikutteisen tiedonhaun tutkimus tarkastelee tiedonhakijan hakukäyttäytymistä tutkimustilanteessa, jossa otetaan huomioon todellisten käyttäjien vuorovaikutus, tiedontarpeen käyttäjäkohtaisuus sekä hakutulosten arvioinnin dynaamisuus ja moniulotteisuus. Tavoitteena on käyttäjän ja järjestelmän vuorovaikutteista tiedonhakutilannetta vastaavan realistisen, mutta suhteellisen kontrolloidun tutkimusasetelman saavuttaminen. Tällaista tutkimusasetelmaa tiedonhaun tutkimuksen kahden perinteisen eli järjestelmäkeskeisen ja käyttäjäkeskeisen suuntausten mukaiset, rajoittuneeseen näkökulmaan keskittyneet tutkimusmallit eivät ole mahdollistaneet.

Tutkielman aineiston muodostaa otos INEX 2004-hankkeen vuorovaikutteisen tiedonhaun tutkimuslinjasta. Siinä hakijoille luotiin mahdollisuus henkilökohtaiseen tulkintaan tiedontarpeesta käytämällä testitilanteessa simuloitua tehtävää. Sen tarkoituksena on kokonaisvaltaisemman tehtäväkuvauksen avulla motivoida ja auttaa hakijaa eläytymään tiedonhakutilanteeseen siten, että hakukäyttäytyminen olisi mahdollisimman realistista. Tutkielmassa selvitettiin, miten hakijat muodostavat kyselyjään tällaisessa tutkimustilanteessa. Tarkemmin perehdyttiin siihen, millaisia hakusanoja hakijat valitsevat suhteessa annettuun tehtäväkuvaukseen ja miten he muokkaavat kyselyään uudelleen. Hakijoiden käyttämiä hakusanoja verrattiin tehtäväkuvauksen sisältöön laskemalla näiden välisiä yhdenmukaisuuksia ja tarkastelemalla käsitteiden välisiä suhteita. Kyselyn uudelleen muotoilua tarkasteltiin muokkaustaktiikoiden luokittelun avulla. Lisäksi analysoitiin valittujen taustamuuttujien mahdollista vaikutusta kyselyn muodostukseen ja tehtiin vertailua kahden erityyppisen tehtäväkuvauksen kesken.

Tulosten perusteella huomattiin, että hakijat olivat tiedonhaussa hyvin aktiivisia kyselyjen ja hakusanojen määrän sekä muokkaustaktiikoiden soveltamisen perusteella. He olivat kuitenkin pitäytyneet varsin läheisesti tehtäväkuvauksissa, sillä hakusanoista keskimäärin noin 80–90 prosenttia oli suoraan tehtäväkuvauksista löytyviä. Myös kyselyjen kymmenen suosituinta käsitettä olivat peräisin tehtäväkuvauksista. Vaikka tehtäväkuvauksien ulkopuolisiksi luokiteltavia hakusanoja oli suhteellisen vähän, omia hakusanoja ideoitiin varsinkin taustatietoon liittyvän tehtäväkuvauksen kohdalla. Käsitesuhteita koskien selkeästi eniten käytettiin rinnakkaiskäsitteitä tehtäväkuvauksien käsitteistä. Vaikka erityyppisten tehtäväkuvauksien välillä havaittiin joitain eroja, ei taustamuuttujilla todettu olevan kovin suurta vaikutusta kyselyn muodostuksessa.

Asiasanat: INEX 2004, kyselyn muodostus, simuloitu tehtävä, vuorovaikutteinen tiedonhaku

## SISÄLLYSLUETTELO

<b>1 JOHDANTO.....</b>	<b>1</b>
1.1 Tutkielman rakenne .....	2
1.2 Keskeisiä käsitteitä .....	3
<b>2 TIEDONHAUN EVALUOINNIN MUUTTUMINEN .....</b>	<b>7</b>
2.1 Hakukäyttötymisen käyttäjäkeskeisiä tutkimuksia.....	8
2.2 Vuorovaikutteinen tiedonhaku .....	9
2.2.1 Simuloitu tehtävä.....	11
2.2.2 Relevanssiarviointi .....	13
<b>3 VUOROVAIKUTTEISEN TIEDONHAUN TUTKIMUKSIA .....</b>	<b>15</b>
3.1 TREC .....	16
3.2 INEX.....	18
3.2.1 XML-kieli ja XML-dokumentin rakenne .....	19
3.2.2 XML-tiedonhaun erityispiirteitä.....	21
<b>4 ONGELMANASETTELU JA TUTKIMUSKYSYMYKSET .....</b>	<b>23</b>
<b>5 INEX 2004 AINEISTO .....</b>	<b>25</b>
5.1 Tutkimusasetelma.....	25
5.1.1 Dokumenttikokoelma.....	25
5.1.2 Hakutehtävät.....	26
5.1.3 Relevanssiarviointi .....	28
5.2 Aineistonkeruu .....	29
5.3 Tutkimusaineiston valinta ja käsittely .....	30
<b>6 TUTKIMUSAINEISTON ANALYYSIMENETELMÄT.....</b>	<b>32</b>
6.1 Tasojen periaatteet ja hakusanayhdenmukaisuus.....	34
6.1.1 Esiintymätaso .....	34
6.1.2 Ilmaisutaso .....	35
6.1.3 Käsitetaso .....	35
6.2 Käsitesuhteet ja kyselyn uudelleen muotoilu.....	36
6.3 Korrelaatioanalyysi .....	38
<b>7 TUTKIMUSTULOKSET .....</b>	<b>39</b>
7.1 Yleisiä tuloksia.....	39
7.2 Hakusanayhdenmukaisuus.....	40
7.2.1 Kaikki kyselyt.....	40
7.2.2 Ensimmäinen ja viimeinen kysely.....	42
7.3 Suosituimmat käsitteet.....	43
7.4 Hakusanat tehtäväkuvauksesta ja sen ulkopuolelta.....	45
7.5 Tehtäväkuvauksen ulkopuolisten hakusanojen suhde tehtäväkuvaukseen.....	48

7.6 Kyselyn uudelleen muotoilu .....	50
7.7 Taustamuuttujien vaikutus kyselyn muodostukseen .....	51
<b>8 YHTEENVETO .....</b>	<b>56</b>
<b>LÄHTEET .....</b>	<b>61</b>
<b>LIITTEET .....</b>	<b>66</b>
LIITE 1: Tehtäväkuvauksen käsiteanalyysi .....	66
LIITE 2: Tehtäväkuvauksen ulkopuolelta olevien hakusanojen edustamien käsitteiden suhde tehtäväkuvauksen käsitteisiin .....	68

# 1 JOHDANTO

Tutkielman aiheena on tiedonhakijan kyselyn muodostukseen liittyvä hakukäyttäytyminen vuorovaikutteisessa tiedonhakutilanteessa. Viime aikoina on korostettu tiedonhakutilanteen dynaamisuutta ja todettu, että tiedonhakijan vuorovaikutus on tärkeä osa tiedonhakuprosessia. Perinteinen järjestelmäkeskeinen tiedonhaun arviointimalli ei enää ole riittävä pyrittäessä kohti realistisempia tiedonhakutilanteiden arviointia. Järjestelmäkeskeinen arviointi on yleensä jättänyt lopullisen käyttäjän tarkastelun ulkopuolelle. Tällaisessa testitilanteessa toimineet tiedonhakijat eivät välttämättä ole olleet esimerkkejä todellisista käyttäjistä, eivätkä testiä varten suunnitellut hakuaiheet ole vastanneet realistisia tiedontarpeita. Toisaalta toisen koulukunnan eli käyttäjäkeskeisen arvioinnin kohteena on puolestaan ollut käyttäjän ongelmanratkaisuprosessi siten, että järjestelmäominaisuudet on jätetty huomioimatta. Käyttäjiin keskittyvä mutta järjestelmästä erillään toteutettu arviointi ei sekään vastaa käyttäjän ja järjestelmän dynaamista, vuorovaikutteista tiedonhakutilannetta.

Vuorovaikutteisen tiedonhaun tutkimus tarkastelee, miten tiedonhakijat käyttäytyvät tutkimustilanteessa, jossa pyritään ottamaan huomioon todellisten käyttäjien vuorovaikutus, tiedontarpeen käyttäjäkohtaisuus sekä hakutulosten arvioinnin dynaamisuus ja moniulotteisuus. Järjestelmä- ja käyttäjäkeskeisen koulukunnan lähentymisessä ja vuorovaikutteisen tutkimustilanteen kehittämisessä haasteena on ollut kontrolloitujen tutkimusolosuhteiden säilyttäminen, kun pyritään kohti realistisempia tiedonhakutilanteita.

Tutkielman empiirinen aineisto kuuluu INEX 2004-yhteistyöverkoston vuorovaikutteisen tiedonhaun (*IIR, Interactive Information Retrieval*) tutkimuslinjaan. Siinä hakijoille on luotu vuorovaikutusmahdollisuuksia hakutehtävän ja relevanssiarvioinnin kohdalla. Tutkimuslinjan ajankohtaisuutta ja käytännönläheisyyttä lisää kokoelman XML-muotoinen esitys, joka mahdollistaa dokumenttien rakenteisuuden huomioimisen. Tässä tutkielmassa keskitytään tarkastelemaan tutkimustilannetta varten luodun simuloidun tehtävän vaikutusta hakijan kyselyn muodostukseen. INEX 2004-hankkeen vuorovaikutteisen tiedonhaun tutkimuslinjassa hakutehtävänä on käytetty simuloitua tehtävää, jonka tarkoituksena on luoda testihenkilöille mahdollisuus omakohtaiseen tulkintaan annetus-

ta hakutilanteesta, jolloin hakijan tulkinta ja sitä seuraava hakukäyttäytyminen vastaisi paremmin realistista tiedonhakutilannetta.

Nyt tavoitteena on tutkia, millainen on hakijoiden kyselyn muodostukseen liittyvä hakukäyttäytyminen realismiin pyrkivässä, mutta kontrolloidussa tutkimustilanteessa. Onko simuloidun tehtävän tehtäväkuvauksella keskeinen rooli käyttäjien hakusanavalinnoissa vai innoittaako tehtäväkuvauksen omakohtainen tulkinta käyttäjiä ideoimaan hakusanoja monipuolisesti tehtäväkuvauksen ulkopuolelta? Tutkielmassa tehdään myös vertailua kahden erityyppiseen tehtäväkuvaukseen liittyvän hakukäyttäytymisen kesken. Tehtäväkuvauksen vaikutusta hakusanojen valintaan tutkitaan vertaamalla käytettyjä hakusanoja tehtäväkuvauksen sisältöön. Lisäksi analysoidaan kyselyiden ja tehtäväkuvauksien käsitesuhteita, kyselyn muokkaustaktiikoita sekä tutkitaan taustamuuttujien vaikutusta kyselyn muodostukseen.

## **1.1 Tutkielman rakenne**

Tutkielma koostuu seuraavista osista: Tutkielman keskeisten käsitteiden määrittelyn jälkeen luvuissa 2 ja 3 tarkastellaan tiedonhaun evaluoinnin muuttumista kohti vuorovaikutteisuutta ja sitä, mihin vuorovaikutteinen tiedonhaku sijoittuu tiedonhaun tutkimuskentässä. Ennen kuin katsotaan vuorovaikutteiseen tiedonhakuun liittyviä uusia piirteitä ja esitellään vuorovaikutteisen tiedonhaun tutkimuksia, käydään läpi tuloksia realistisista käyttäjätutkimuksista koskien kyselyn muodostusta.

Luku 4 koskee tutkimusongelmien asettamista. Luvussa 5 esitellään yleisesti INEX 2004-hanketta sekä tutustutaan tarkemmin sen vuorovaikutteisen tiedonhaun tutkimuslinjaan, josta otettu otos toimii tämän tutkielman tutkimusaineistona. Tutkimusaineiston analyysimenetelmät käydään läpi luvussa 6, jota seuraa tutkimustulosten esittäminen. Tutkimustulosten yhteenveto esitetään tutkielman lopussa.

## 1.2 Keskeisiä käsitteitä

### Tiedonhaku

Tiedonhaku (*IR, Information Retrieval*) kuuluu osa-alueena tiedonhankinnan laajempaan kokonaisuuteen. Tiedonhaun tavoitteena on tiedonhakijan tiedontarpeiden tyydyttäminen. Siinä pyritään löytämään tiedontarpeiden tyydyttämistä mahdollisimman hyvin palveleva dokumentti tai dokumenttijoukko. (Järvelin 1995, 26.) Baeza-Yates & Ribeiro-Neto (1999, 1) määrittävät tiedonhaun laajassa merkityksessä kattavan tiedon esittämisen, tallentamisen, järjestämisen ja hankinnan. Usein tiedonhaku tapahtuu hakujärjestelmän avulla tietokannasta syöttämällä kysely, johon on valittu tiedontarvitsijan tiedontarvetta kuvaavat hakusanat. Hakujärjestelmä palauttaa kyselyä vastaavan aineiston (dokumentin tai sen osan) tulosjoukkoon.

### Hakukäyttäytyminen

Lyhyesti ilmaistuna hakukäyttäytymisellä tarkoitetaan sitä, mitä hakijat tekevät ja/tai ajattelevat tiedonhakuprosessin aikana (Bates 1979).

### Simuloitu tehtävä

Kirjallisuudessa simuloitun hakutehtävän, työtehtävän ja tehtävän käsitteet ovat epäselvät. Borlund (2000, ks. luku 2.1.1) esittää simuloitun työtehtävän käsitteen. Hänen esittämänsä tehtävät liittyvät kuitenkin ennemminkin vapaa-aikaan kuin työelämään. Joskus tällainen simuloitu työtehtävä saattaa sisältää erillisenä hakutehtävän (esim. White, Ruthven & Jose 2002). Kaiken kaikkiaan kysymys on siitä, että hakijalle pyritään luomaan simulaation kautta hakutilanne, johon hän pystyy eläytymään. Tässä tutkielmassa käsitteestä käytetään nimeä simuloitu tehtävä, ja se koostuu kirjallisesta kuvauksesta, joka sisältää sekä hakutilanteen että -tehtävän. Kirjallista kuvausta kutsutaan tässä työssä tehtäväkuvaukseksi.

### Kysely

Hakujärjestelmään syötetyt kyselyt voivat olla rakenteisia tai rakenteettomia. Rakenteisissa kyselyissä hakusanojen väliset suhteet ilmaistaan operaattoreilla, joista tunnetuimpia lienevät Boolean logiikan perusoperaattorit and, or ja not. Tämän tutkimusaineiston kyselyt eivät sisällä perusoperaattoreita, mutta hakijoilla oli mahdollisuus käyttää + ja - -merkkejä hakusanojen edessä painoarvon nostamiseksi tai laskemiseksi. Merkkien avulla hakija saattoi ilmaista, minkä hakusanan hän

katsoo välttämättömäksi tai mitä hakusanaa hän ei tahdo dokumentissa esiintyvän. Hakijoilla oli myös mahdollisuus muodostaa hakusanoista fraaseja lainausmerkkien avulla. Tiedonhaun kirjallisuudessa kysely-nimitystä vastaten esiintyvät usein myös nimitykset hakukysely ja hakulauseke.

### **Hakusana**

Hakuavain on yleisnimitys kyselyissä esiintyville luonnollisen kielen sanoille, yleiskieleen sisällyville lyhenteille ja koodeille sekä hakutermeille (Järvelin 1995, 176). Tässä tutkielmassa käytetään hakusanaa vastaavanlaisena yleisnimityksenä, sillä aineiston kyselyt muodostuvat muutamaa lyhennettä lukuun ottamatta luonnollisista hakusanoista. Koodeja tai erityisiä hakutermejä ei esiinny kyselyissä lainkaan. Kyselyissä esiintyviä fraaseja käsitellään joko fraaseina tai yksittäisinä sanoina. Siten tilanteen mukaan hakusanalla voidaan viitata sekä itse fraasiin että fraasiin sisältyviin yksittäisiin sanoihin.

### **Kyselyn muodostus ja uudelleen muotoilu**

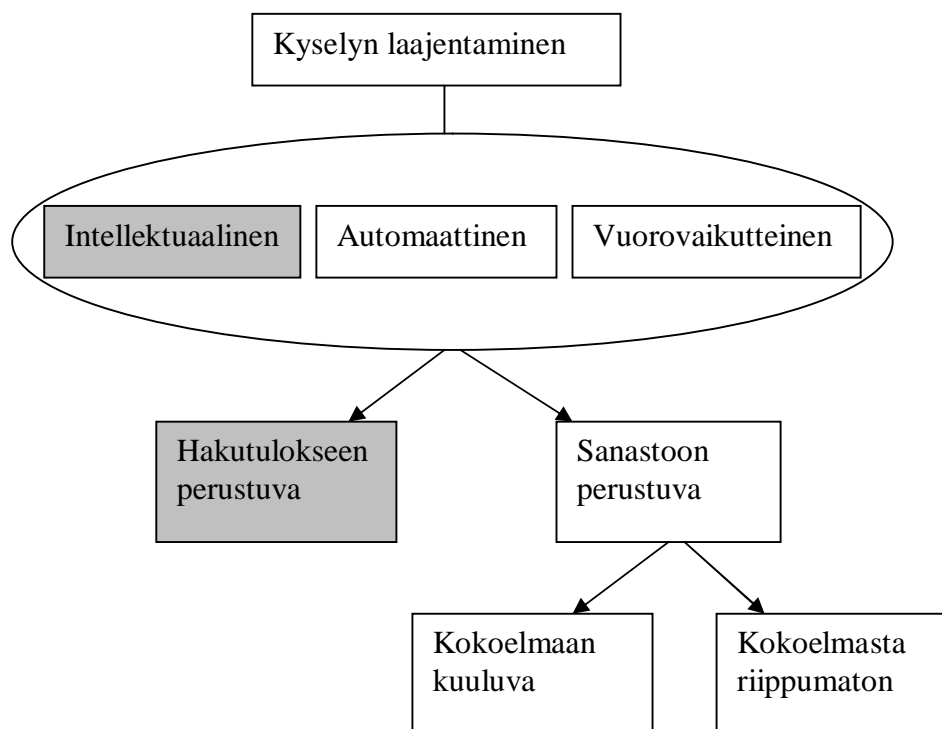
Kyselyn muodostuksessa muotoillaan haluttu kysely valitsemalla hakustrategia ja hakusanat. Haun suorittamisen jälkeen saatuja tuloksia voidaan tarkastella ja kyselyä voidaan muotoilla uudelleen (*query reformulation*) sopivan tuloksen saavuttamiseksi. Kyselyn laajentaminen tai kaventaminen voi tapahtua esimerkiksi syöttämällä uusia hakusanoja tai poistamalla jo kokeiltuja hakusanoja kyselystä. (Järvelin 1995, 220-224.)

Kyselyn laajennus (*query expansion*) on kyselyn uudelleen muotoilua, joka Efthimiadiksen (1996, 122) mukaan kattaa laajempaa käsitteenä sekä hakusanojen lisäämisen että poistamisen. Vaikka kyselyn laajennus yleensä käsittelee hakutulosten lisäksi jonkin muun lähteen käyttöä paremman hakutuloksen saavuttamiseksi, on kyselyn laajennus syytä käsitellä myös tässä tutkielmassa, jossa kyselyn uudelleen muotoilun lähteenä on tulosjoukko ja dokumenttien (tai niiden osien) tarkastelumahdollisuus.

Kyselyn laajennuksen menetelmät (intellektuaalinen, automaattinen ja vuorovaikutteinen) ja lähteet (hakutulos ja sanasto) on kuvattu kuviossa 1. Tätä tutkielmaa koskevat vaihtoehdot on erotettu tummennetulla taustalla. Kyseessä on intellektuaalinen kyselyn laajentaminen, jossa hakija itse muotoilee kyselyään lisäämällä, poistamalla tai vaihtamalla hakusanoja paremman hakutuloksen



saavuttamiseksi. Aiemmin saatu hakutulokset voi toimia kyselyn laajennuksen ja hakusanojen valinnan lähteenä.



Kuvio 1. Kyselyn laajentamisen menetelmät ja lähteet (Efthimiadis 1996, 124).

### Hakutaktiikka

Bates (1979) määrittää hakutaktiikan siirroiksi (*moves*), joita hakija tekee edistääkseen tiedonhakuun. Hakustrategia on laaja, koko tiedonhakutilannetta koskeva suunnitelma, joka voi koostua useista erilaisista lyhyempää ajanjaksoa koskevista hakutaktiikoista. Tiedonhakutilanteessa hakijalla on valittavanaan useita erilaisia hakutaktiikoita. Ne voivat olla hakustrategiaan liittyen ennalta päätettyjä; jotkut hakutaktiikat voidaan valita vasta hakuprosessin kuluessa. Samoin hakuprosessin aikana hakutaktiikoita voidaan muokata, hylätä tai vaihtaa. Hakutaktiikat voidaan jakaa ryhmiin esimerkiksi sen mukaan, millaista toimintoa ne edustavat tai mihin hakuprosessin vaiheeseen ne kyseisellä hetkellä kuuluvat. (Efthimiadis 1996, 131-132.)

Bates (1979) nimeää 29 hakutaktiikkaa ja jakaa ne neljään kategoriaan, joista kaksi ensimmäistä liittyvät tiedonhakutilanteen etenemiseen ja tietolähteen valintaan. Kahden jälkimmäisen kategorian taktiikat liittyvät kyselyn suunnittelu- ja muokkausprosessiin sekä hakusanojen valintaan ja tarkistukseen. Kyselyn muokkaustaktiikoita ovat muun muassa tarkennus, laajennus, supistaminen tai

poissulkeminen (*specify, exhaust, reduce* ja *block*). Hakusanoja koskeviksi taktiikoiksi Bates esittää esimerkiksi synonyymien, vastakohtien tai hierarkiassa eritasoisten hakusanojen valinnan sekä hakusanojen kirjoitusasun muokkaamisen taktiikat.

## 2 TIEDONHAUN EVALUOINNIN MUUTTUMINEN

Tiedonhaun järjestelmäkeskeisessä evaluoinnissa on perinteisesti keskitytty tekstin ja kyselyjen esittämismuotoihin sekä näiden esitysten vertailuun. Myöhemmin huomiota on kiinnitetty yhä enemmän myös tiedonhaun vuorovaikutteiseen luonteeseen. Vuorovaikutteisen tiedonhaun evaluoinnissa pyritään järjestämään mahdollisimman realistisesti todelliseen tiedonhakuun perustuva koetilanne kuitenkin siten, että suhteellisen kontrolloidut tutkimusolosuhteet säilytetään. Vuorovaikutteisten piirteiden lisääntyminen luo tarvetta lisäksi vaihtoehtoisille järjestelmän tulosten mittausmenetelmille. (Belkin, Cool, Stein, & Thiel 1994; Borlund 2003.)

Samalla kun tiedonhaun evaluointi on muuttunut yksiselitteisestä ja tarkkaan rajatusta laboratoriomallista eri lähestymistapojen huomioon ottamiseen, on lähestymistapojen mukaisen tiedonhakupäristön rajojen määrittäminen vaikeutunut. Laboratoriomallin vuorovaikutteinen järjestelmä on vaikea toteuttaa niin, että se jäljittäisi realistisesti käyttäjien toimintaa ja reaktioita. Tiedonhaun evaluoinnissa käytettävää relevanssin määrittelyä varten tulisi myös tietää todelliset käyttäjät ja heidän tiedontarpeensa. (Robertson & Hancock-Beaulieu 1992, 457-460.)

Ingwersen (1992, 78-84; 87) huomauttaa, että järjestelmäkeskeinen ja myöhemmin enemmän käyttäjien roolin huomioon ottanut käyttäjäkeskeinen malli ovat molemmat olleet oikeastaan erillisiä suuntauksia ja että kumpikin suuntaus on omalta osaltaan muodostanut rajoittuneen näkemyksen tiedonhausta. Suuntausten väliset erot ilmenevät tutkijoiden asenteissa informaation käsitteeseen, tiedontarpeen luonteeseen, tutkimusympäristöön, tiedonvälittäjän, käyttäjän ja tiedonhakupäristelmän rooliin sekä tiedonhakupäristelmän ja automaattisen välittäjän toimintarajoihin. Perinteisessä järjestelmäkeskeisessä mallissa testikokoelma on luotu huomioimatta varsinaisia käyttäjiä. Käyttäjäkeskeisen mallin kohteena puolestaan on ollut käyttäjän ongelmanratkaisuprosessi tiedonhaun aikana olosuhteissa, joissa käytettävä järjestelmä on kuitenkin oletettu vakioksi ja jätetty niin ikään huomioimatta. Ingwersen toteaa, että vasta mallien lähentyminen on ollut askel kohti vuorovaikutteista ja kognitiivista tiedonhakua.

## 2.1 Hakukäyttäjien käyttäjäkeskeisiä tutkimuksia

Tiedonhakijan hakukäyttäjyyttä kyselyn muodostuksen näkökulmasta on tutkittu useissa käyttäjäkeskeisen mallin mukaisissa tiedonhaun tutkimuksissa. Seuraavaksi tarkastellaan muutamia tällaisia tiedonhaun tutkimuksia, joissa tutkimuksen kohteena ovat todelliset käyttäjät kognitiivisine toimintoineen. Lisäksi kahdessa jälkimmäisessä tutkimuksessa käyttäjät hakivat tietoa oman, todellisen tiedontarpeensa perusteella.

Sutcliffe, Ennis ja Watkinsonin (2000) empiirisessä tutkimuksessa tarkkailtiin lääketieteen opiskelijoiden hakukäyttäjyyttä heidän hakiessaan tietoa annetuista hakuaiheista Medline-tietokannasta. Tutkimuksessa kokeneet käyttäjät käyttivät keskimäärin 9,1 hakusanaa ja noviisit käyttäjät keskimäärin 6,6 hakusanaa kyselyä kohden. Lääketieteen ja hakujärjestelmän ammattilaiset laativat yhteistyössä hakuaiheita vastaavat parhaimmat hakutulokset saavuttavat standardikyselyt (*gold standard query*), joihin käyttäjien muotoilemia kyselyjä hakusanoineen myös vertailtiin. Kokeneilla käyttäjillä noin 78 % ja noviiseilla noin 75 % hakusanoista oli samoja näihin standardikyselyihin määriteltyjen hakusanojen kanssa. Moni standardikyselyihin kuuluvista hakusanoista jäi kuitenkin hakijoiden kyselyissä käyttämättä, ja kyselyt olivat, varsinkin noviisikäyttäjillä, rakenteeltaan melko yksinkertaisia.

Jansen, Spink, Bateman ja Saracevic (1998) tutkivat käyttäjien tiedonhakua web-ympäristössä. Excite-hakupalvelun lokitiedostojen tarkastelu osoitti, että käyttäjien kyselyt olivat lyhyitä sisältäen keskimäärin vain 2,2 hakusanaa. Lisäksi suurin osa käyttäjistä suoritti vain yhden kyselyn eikä muokannut kyselyä uudelleen. Saman hakupalvelun jatkotutkimuksessa (Spink, Wolfram, Jansen & Saracevic 2001) selvisi, että käyttäjien muotoilemat kyselyt olivat edelleen yksinkertaisia sekä lyhyitä ja sisälsivät vähän hakusanoja.

Vakkarin (2000) pitkittäistutkimuksessa tutkittiin opiskelijoiden hakukäyttäjyyttä tiedonhakuprosessin eri vaiheissa. Tutkimuksen kohteena olivat todelliset käyttäjät hakuaiheineen eli kysymyksessä oli ei-kokeellinen, pienehkön opiskelijaryhmän käyttäjätutkimus. Opiskelijat tekivät seminaariinsa kuuluvaan tutkimussuunnitelmaan liittyen kolme tiedonhakua. Tutkimuksessa tarkasteltiin opiskelijoiden hakusanojen ja -taktiikoiden käytön muuttumista suhteessa heidän kasvavaan aihetuntemukseensa hakuprosessin edetessä. Tulokset osoittivat, että aihetuntemuksen lisääntyminen

nosti sekä yleisesti hakusanojen että tarkempien hakusanojen käyttöä ja lisäsi hakuoperaattoreiden ja –taktiikoiden käyttöä. Opiskelijat käyttivät ensimmäisessä tiedonhakuosiossa keskimäärin 3,0 hakusanaa, toisessa 4,2 ja kolmannessa 5,5 hakusanaa hakijaa kohti. Hakuosioiden edetessä ja aiheutuneen kasvaessa opiskelijat muokkasivat kyselyitään pääasiassa valitsemalla synonyymejä, suppeampia hakusanoja ja rinnakkaisia hakusanoja. Hakutaktiikoiden valintaan vaikuttavaksi tekijäksi ilmeni hakuprosessin vaiheiden ja aihetietämyksen ohella myös tulosjoukon koko.

## 2.2 Vuorovaikutteinen tiedonhaku

Tiedonhakujärjestelmät ovat muuttuneet vuorovaikutteisimmiksi eikä niitä siten voi arvioida niin, että vuorovaikutteinen tiedonhakuprosessi jätettäisiin tutkimuksen ulkopuolelle. Vuorovaikutteiset tiedonhakujärjestelmät kuvaillaan järjestelmiksi, joissa käyttäjä suorittaa hakutehtäviä dynaamisesti ja reagoi järjestelmän antamiin vastauksiin hakuprosessin aikana. Siten vuorovaikutteisen tiedonhaun arvioinnissa tutkimuksen kohteena voivat olla käyttäjän vuorovaikutteiset toiminnot, koskien sekä prosessin aikana tapahtuvia haku- ja palautetoimintoja että varsinaista hakutulosta. (Borlund 2000, 47.)

Robertson ja Hancock-Beaulieu (1992, 458-459) esittävät kolme vallankumousta, jotka ovat olleet tärkeimpiä vaikuttajia tiedonhakujärjestelmien ja vuorovaikutteisen tiedonhaun uusien lähestymistapojen tarpeen taustalla. *Relevanssimuutos* viittaa muutokseen relevanssikäsitteen tulkinnassa. *Kognitiivinen muutos* koskee tiedontarvetta ja sen muodostumista. Tiedonhakuprosessissa käyttäjä on kognitiivisine toimintoineen vuorovaikutuksessa hakujärjestelmän kanssa, ja käyttäjän kognitiivisen tilan muutos vastaa laboratoriomallin kuvaamaa rajattua kysely-tulos-prosessia. Samoin *vuorovaikutuksen* merkitys kasvoi relevanssin ja kognitiivisen muutoksen rinnalla. Tiedonhaun järjestelmät saivat yhä enemmän vuorovaikutteisia ominaisuuksia, eikä niitä voitu enää ajatella erillään tarkkaan rajattuna kokonaisuutena vaan osana koko tiedonhaun prosessia.

Borlundin (2000, 2-4) mukaan edellä mainittuihin muutoksiin ei voitu vastata tiedonhaun järjestelmäkeskeisessä, Cranfieldin malliin perustuvassa lähestymistavassa. Cranfieldin malli on peräisin 1960-luvulla toteutetusta Cranfield II –tutkimuksesta. Cranfieldin ja muiden laboratoriomallin mukaisten testien peruskoeasetelmaan kuuluu dokumenttikokoelma, hakuaiheet ja relevanssiarvioinnit. Monissa järjestelmäkeskeisissä tutkimusasetelmissa hakujärjestelmien arviointi on perustunut jär-

jestelmän kykyyn palauttaa kyselyjä tai hakusanoja vastaavia dokumentteja tarkkaan kontrolloidussa laboratorio-olosuhteissa. Hakutulosten arvioimiseksi tarvitaan arviointikriteerejä, joista tavallisimmat ja laajimmin käytetyt hakutulosta koskevat arviointikriteerit ovat saanti ja tarkkuus. Tarkkuudella tarkoitetaan relevanttien dokumenttien määrää tulosjoukkoon kuuluvista dokumenteista. Saannilla mitataan, miten suuri osa testikokoelman kaikista mahdollisista relevanteista dokumenteista saadaan tulosjoukkoon. Järjestelmän kehittämisen tavoitteena on usein ollut suoraan tarkkuuden ja saannin kasvattaminen ja näitä mittareita on käytetty järjestelmän tehokkuuden mittaamiseen. (Ellis 1996, 19-20; Harter & Hert 1997, 8; Järvelin & Sormunen 1999, 119.)

Cranfieldin tutkimukseen perustuvaa järjestelmäkeskeistä tiedonhakumallia käytettiin myöhemmin useissa tutkimuksissa ilman, että mallin kognitiivisia ja hakijan käyttäytymistä koskevia oletuksia olisi kyseenalaistettu (Hert 1997, 6-8). Malli ei ota huomioon tiedontarpeen dynaamisuutta vaan olettaa sen olevan staattinen käsite. Lisäksi käyttäjän vuorovaikutus sekä relevanssin moniasteinen ja dynaaminen luonne jää huomioimatta. Kokoelmaan sisältyvät hakutehtävät ovat yleensä tarkkaan muotoiltuja ja muokattu testiä varten tietoisesti aihe-suuntautuneiksi. Siten ne eivät vastaa realistisia tiedontarpeita, jotka voivat olla hyvinkin epämääräisiä. (Borlund 2000, 4-5; Ingwersen & Järvelin 2005, 123-130.)

Kognitiivisen muutoksen ja relevanssimuutoksen mukaan vuorovaikutteisen tiedonhaun arvioinnissa tiedontarvetta tulisi tarkastella käyttäjäkohtaisena ja mahdollisesti dynaamisena käsitteenä. Samoin relevanssin moniulotteisuus ja dynaaminen luonne olisi otettava huomioon. Henkilön, jolla on tiedontarve tietyssä tilanteessa, tulisi tehdä relevanssiarvointi tätä tilannetta eikä staattista kyselyä vastaan arvioituna. (Borlund 2000, 78.)

Nämä lähtökohdat huomioon ottaen Borlund (2000) esittää vaihtoehtoisen mallin vuorovaikutteisen tiedonhaun arvioinnille (*Evaluation package*). Mallin tarkoituksena on taata toimiva, perusteltu ja realistinen tutkimusasetelma. Relevanssimuutoksen, kognitiivisen muutoksen ja vuorovaikutteisen muutoksen mukaisten vaatimusten vastaamiseksi mallissa esitetään kolme komponenttia, jotka antavat määrityksiä tutkimuskehikseen kuuluville hakijoille, hakutehtäville ja relevanssiarvioinneille. Mallin mukaan vuorovaikutteisen arviointiympäristön saavuttamiseksi olisi käytettävä:

- 1) Tiedonhakijoita, jotka voisivat olla todellisia käyttäjiä
- 2) Tiedontarvetta, joka mahdollistaa henkilökohtaisen ja dynaamisen tulkinnan tiedontarpeesta
- 3) Moniulotteista ja dynaamista relevanssiarviointia (Borlund 2000, 77-78.)

Tiedontarpeen yksilöllinen ja dynaaminen luonne otetaan huomioon sallimalla potentiaalisten käyttäjien toimiminen testihenkilöinä siten, että heillä on mahdollisuus henkilökohtaiseen tulkintaan tiedontarpeesta. Tämän saavuttamiseksi mallissa esitetään simuloitun tehtävän käyttö vastaamaan realistista tiedontarvetta. Lisäksi sallitaan käyttäjän yksilöllisen tiedontarvetulkinnan muuttuminen ja kehittyminen hakuprosessin aikana. Tulkintaan liittyvä muutos taas on yhteydessä käyttäjän suorittamaan relevanssiarvioinnin prosessiin. Relevanssiarviointi suoritetaan verrattuna tiedontarvetulkintaan ja siten kyseiseen tilanteeseen, joka käyttäjällä on. Tilannerelevanssi arvioidaan sen dynaamisen ja moniulotteisen luonteen mukaisesti monitasoisesti eikä binäärisesti. (Borlund 2000, 78).

INEX-hankkeen vuorovaikuttaisen tiedonhaun tutkimuslinjassa on kiinnitetty erityistä huomiota tiedontarpeen ja relevanssiarvioinnin soveltamiseen mallin mukaisesti. Näin tutkimukseen osallistuneille käyttäjille luotiin mahdollisuus henkilökohtaiseen tulkintaan tiedontarpeesta. Seuraavaksi tarkastellaan simuloitun tehtävän ja relevanssin käsitettä vuorovaikuttaisen tiedonhaun kannalta.

### **2.2.1 Simuloitu tehtävä**

Borlundin (2000) esittämä simuloitu tehtävä on väline, joka saa aikaan simuloitun tiedontarpeen. Borlund itse käyttää käsitteestä nimitystä simuloitu työtehtävä. Hänen esimerkkinsä simuloitusta työtehtävästä liittyvät työelämän sijaan kuitenkin enemmän vapaa-aikaan, joten tässä tutkielmassa käsitteeseen viitataan nimityksellä simuloitu tehtävä.

Simuloitu tehtävä on taustakertomus, jossa kuvaillaan jokin tiedontarvetilanne, joka saattaa johtaa tiedonhankintaan ja tiedonhakuun. Semanttisesti avoimen kuvauksen tarkoituksena on luoda lukijalle (käyttäjä) kuva tiedontarpeesta, ilman että hänelle annetaan tarkkaan määrätty hakutehtävä. Kuvailtuun ongelmatilanteeseen perustuen käyttäjä muodostaa kyselyn ja syöttää sen järjestelmään. Simuloitu tehtävä vahvistaa käyttäjän vapautta, sillä sen tarkoituksena on tarjota konteksti, jonka

avulla käyttäjä kohtaa annetun tilanteen ja tulkitsee sitä yksilöllisesti. Simuloitu tehtävä täytyisi nähdä kognitiivisena tilana, joka luo käyttäjälle tiedontarpeen. Hakija reagoi siihen kohtaamalla, käsittelemällä ja ratkaisemalla sen. Simuloitu tehtävä toimii lisäksi perustana, jota vastaan relevanssiarviointi voidaan suorittaa. (Borlund 2000, 21-22; 79-80.)

Esimerkki simuloidusta tehtävästä:

Some friends of yours are about to visit you and as a surprise you are planning a trip for all of you to the Isle of Arran. You have heard rumours that some of the ferries to Arran are less safe than others. In addition to this you have recently seen the movie Titanic. You would therefore like to retrieve information about the causes of safety problems on ferries as well as some information about how to prevent accidents. (Borlund 2000, 215.)

Borlundin raportoimassa tutkimuksessa tutkittiin testihenkilöiden tiedonhakupäätöstyyppeistä erityyppisiä tiedontarpeita sovellettaessa. Testitilanteessa käytetyt tiedontarpeet olivat testiä varten luotuja erilaisia simuloituja tehtäviä sekä aitoja, hakijoiden henkilökohtaisia tiedontarpeita. Kullakin hakijalla oli haettavanaan yksi henkilökohtainen tiedontarve sekä viisi simuloitua tehtävää. Hakijoita pyydettiin hakemaan heidän tulkitsemaansa tiedontarvetta tyydyttävää tietoa sekä arvioidaan tulostokumenttien relevanssia 3-portaisella asteikolla. Tutkimusaineiston analysoinnissa käytettiin määrällisiä ja laadullisia menetelmiä. Tilastollisissa testeissä t-testien ja khiin neliö-testien perusteella todettiin, että simuloitun tehtävän ja hakijan henkilökohtaista tiedontarvetta koskevan hakukäyttäytymisen välillä ei ollut merkittävää eroa. Hakukäyttäytymisen vertailussa tutkittiin tehtyjen kyselyjen, relevanssiarvioiden ja käytettyjen hakusanojen määriä sekä tiedonhakuun käytettyä aikaa. Simuloitun tehtävän todettiin siten vastaavan hyvin hakijan todellista tiedontarvetta ja olevan käytettävissä vuorovaikutteisen tiedonhaun arvioinnissa realistisuuden saavuttamiseksi. (Borlund 2000, 101-121; 139-140.) Simuloitun tehtävän mukaista menetelmää on käytetty myöhemmin useissa vuorovaikutteisen tiedonhaun tutkimuksissa (ks. esim. Ruthven, Lalmas & van Rijsbergen 2003; Suomela & Kekäläinen 2005; Tombros, Ruthven & Jose 2005).



## 2.2.2 Relevanssiarviointi

Tiedonhaussa tavoitteena on löytää hakua vastaavaa relevanttia tietoa (Saracevic 1996, 201). Siten relevanssi on tiedonhaun keskeisimpiä käsitteitä. Relevanssin yksiselitteistä määrittelyä vaikeuttaa kuitenkin sen moniulotteinen ja dynaaminen luonne.

Relevanssi ilmaisee suhdetta, ja erilaisiin suhteisiin perustuen Saracevic (1996, 212) jakaa relevanssin viiteen ilmentymään, jotka ovat dynaamisessa vuorovaikutuksessa keskenään tiedonhaun konseptissa. Ilmentymät ovat:

- Järjestelmä/algoritminen relevanssi: kyselyn ja dokumentin välinen suhde kyseisessä hakujärjestelmässä käytettävän menetelmän (algoritmin) mukaisesti
- Aiherelevanssi: kyselyssä ja dokumentissa ilmenevien aiheiden välinen suhde käyttäjän odotustensa mukaisesti tulkitsemana
- Kognitiivinen/osuvuusrelevanssi: käyttäjän tietämyksen tai tiedontarpeen ja dokumentin välinen suhde
- Tilannerelevanssi/hyödyllisyys: käyttäjällä olevan kyseisen tilanteen, hakutehtävän tai ongelman ja dokumentin välinen suhde
- Motivaatio/affektiivinen relevanssi: käyttäjän aikomusten, tavoitteiden tai motivaation ja dokumentin välinen suhde.

Relevanssia on käsitelty tiedonhaun tutkimuksen kirjallisuudessa eri ilmentymien näkökulmasta. Algoritminen relevanssiarviointi on tyypillisesti perustunut binääriseen eli kaksitasoiseen arviointiin, jolloin dokumentit arvioidaan joko relevanteiksi tai ei-relevanteiksi. Algoritminen relevanssi onkin tyypiltään objektiivinen relevanssin ilmentymä neljän muun ilmentymän ollessa subjektiivisia (käyttäjä mukana). Aiherelevanssin kohdalla, jossa käyttäjä tulkitsee kyselyssä ja dokumentissa ilmenevien aiheiden välistä suhdetta, arviointi on toteutettu tyypillisesti tarkkuuden ja saannin mittareilla.

Myöhemmin Borlund (2000) sekä Cosijn ja Ingwersen (2000) huomauttavat, ettei Saracevicin luokituksen mukaista affektiivista, käyttäjän aikomuksiin, tavoitteisiin ja motivaatioon pohjautuvaa relevanssia voida pitää itsenäisenä ja spesifinä relevanssin lajina. Heidän mukaansa käyttäjän aiheet, tavoitteet ja motivaatio toimivat tekijänä, joka saa käyttäjän hakemaan tietoa, toteuttamaan

tiedonhaun ja arvioimaan saatujen tulostensa relevanssia. Siten affektiivisuus on luonnollisesti sidoksissa kaikkiin subjektiivisen relevanssin ilmentymiin.

Relevanssimuutoksen mukainen käsitys siitä, että käyttäjän esittämä kysely ei täysin vastaa hänellä olevaa tiedontarvetta johti ajatteluun, että relevanssia tulisi verrata alkuperäiseen tiedontarpeeseen eikä sen pohjalta esitettyyn kyselyyn (Robertson & Hancock-Beaulieu 1992, 458-459). INEX-hankkeen (käsitellään luvussa 5) vuorovaikutteisen tiedonhaun tutkimuslinjassa simuloitun tehtävän soveltaminen mahdollisti hakijoiden henkilökohtaisen tulkinnan tiedontarpeesta. Siten simuloitu tehtävä ja siitä tehty tulkinta toimi perustana, jota vastaan hakijat arvioivat tulostensa relevanssia.

Schamberin, Eisenbergin ja Nilanin (1990) mukaan relevanssi on käsitteenä:

- Moniulotteinen ja kognitiivinen. Sen merkitys riippuu käyttäjän tietämyksestä ja tiedontarpeen tilanteista.
- Dynaaminen. Käyttäjän tekemä arviointi tietyn hetken tiedontarpeen ja sen tyydyttämiseksi hankitun tiedon välisestä suhteesta vaikuttaa relevanssin merkitykseen.
- Monitahoinen mutta systemaattinen ja mahdollista mitata, jos käyttäjänäkökulma otetaan huomioon käsitteellisellä ja toiminnallisella tasolla. (Schamber, Eisenberg & Nilan 1990, 774.)

Näitä relevanssin määrittämiä on myöhemmin käytetty määrittämään relevanssia vuorovaikutteisen tiedonhaun arvioinnissa, jossa relevanssin moniulotteisuus, monitahoisuus ja dynaamisuus korostuvat. Samoin käyttäjän ja tiedontarpeen merkitystä painotetaan. Relevanssia on alettu arvioimaan binääritason arvioinnin sijaan yhä useammin monitahoisilla arviointiasteikoilla, joissa asteikkoa on laajennettu esimerkiksi melko relevantin ja marginaalisesti relevantin käsitteen avulla. INEX-hankkeen vuorovaikutteisen tiedonhaun tutkimuslinjan relevanssiarviointi suoritettiin monitahoisella 10-portaisella asteikolla, jossa arvioitiin sekä käyttökelpoisuutta että spesifisyyttä, luku 5.1.3. INEX-hankkeessa testikokoelmiin kuuluvien, etukäteen tehtyjen relevanssiarviointien lisäksi tai sijaan käyttäjille on suotu mahdollisuus itsenäiseen relevanssiarviointiin tutkimustilanteessa. Relevanssiarvioinnin tapojen ja asteikkojen muuttuminen on edellyttänyt myös vaihtoehtoisten arviointimenetelmien kehittämistä esimerkkinä relatiivinen relevanssi ja Ranked Half-Life (Borlund & Ingwersen 1998, 324-331) sekä kumulatiiviseen hyötyyn perustuvat mittarit (Järvelin & Kekäläinen 2002, 422-446).

### 3 VUOROVAIKUTTEISEN TIEDONHAUN TUTKIMUKSIA

Järjestelmäkeskeisessä tiedonhaussa usein toteutetun laboratoriokokeen etuna on muuttujien hyvä kontrolloitavuus ja havainnointi, kokeen luotettavuus ja toiston mahdollisuus. Toisaalta laboratoriomallissa vuorovaikutteinen järjestelmä on vaikea toteuttaa niin, että se vastaisi realistisesti käyttäjien toimintaa ja reaktioita. Lisäksi relevanssin määrittelyä varten tulisi tuntea todelliset käyttäjät ja heidän tiedontarpeensa. (Robertson & Hancock-Beaulieu 1992, 460.) Vuorovaikutteisen tiedonhaun tutkimuksessa keskeiseksi kysymykseksi nousikin se, miten laboratoriomallia on mahdollista laajentaa ottaen todelliset käyttäjät huomioon säilyttäen samalla muuttujien ja tutkimustilanteen kontrolloitavuus (Ingwersen & Järvelin 2005, 251).

OKAPI on esimerkki käyttäjäsuuntautuneesta tutkimushankkeesta, jossa laboratorio-olosuhteiden ja todellisten käyttäjien osallistumisen yhteensovittamisessa onnistuttiin jossain määrin. Vuonna 1985 alkaneen hankkeen lähtökohtana oli suunnitella ja kehittää Online-järjestelmän prototyyppejä, joiden erilaisia ominaisuuksia sekä hankkeen yhteydessä kehitettyjä ideoita ja teorioita testattiin todellisessa käyttäjäympäristössä. OKAPI-projekteissa keskityttiin noviisikäyttäjien hakukäyttäytymisen ymmärtämiseen; testeissä kerättiin tietoa ja havainnoitiin käyttäjiä monin tavoin. Arviointimenetelmien kehittyessä OKAPI-tutkimuksissa kyettiin yhä paremmin myös yhdistämään laboratoriomalli ja käyttäjille suunnattu kenttätesti. Kolmannessa OKAPI-testissä käyttäjän vuorovaikutus järjestelmän kanssa otettiin huomioon osittain jo laboratoriovaiheessa (aiheervalinta, kyselyn laajenusominaisuus, relevanssiarviointi) ennen järjestelmän ominaisuuksien reaalityilanteessa suoritettavaa testausta. OKAPI:n haasteeksi jäi edelleen tutkimuksen kontrolloitavuus esimerkiksi sovellettaessa käyttäjien omia tiedontarpeita ja relevanssiarvioita vaihtoehtona valmiiksi annetuille hakuaiheille. Henkilökohtaisten tiedontarpeiden käyttö vähentää tutkijoiden kontrollin mahdollisuutta tutkimustilanteessa. Tutkimusta varten luotujen hakuaiheiden avulla kontrolli säilytetään, mutta realiteetti heikkenee. (Robertson & Hancock-Beaulieu 1992, 461-462; Robertson 1997, 5-6; Ingwersen & Järvelin 2005, 251.)

TREC ja INEX ovat tiedonhaun tutkimuksen yhteistyöverkostoja. Molemmat yhteisöt ovat luoneet vuosittain arviointiympäristöjä, joissa on keskitytty erilaisiin tiedonhaun osa-alueisiin. Yhteistyöverkoston kuuluvat ryhmät ovat osallistuessaan sitoutuneet tutkimussuunnitelmaan ja -aikatauluun.

Sekä TREC- että INEX-hankkeiden yhdeksi osa-alueeksi on otettu vuorovaikutteisen tiedonhaun arviointi järjestelmäkeskeisen tutkimusasetelman mukaisessa kontrolloidussa ympäristössä. Seuraavaksi tutustutaan näihin kahteen hankkeeseen. Koska tutkielman tutkimusaineisto on otos INEX-hankkeen vuorovaikutteisen tiedonhaun tutkimuslinjan aineistosta, käsitellään se myöhemmin yksityiskohtaisemmin luvussa 5.

### 3.1 TREC

Vuodesta 1992 lähtien järjestetty järjestelmäkeskeiseen malliin pohjautuva TREC (*Text REtrieval Conference*) on ollut merkittävä hanke tiedonhaun tutkimuksessa. Hankkeen taustalla on toiminut NIST (*National Institute of Standards and Technology*), DARPA (*Defense Advance Research Projects Agency*) sekä ARDA (*Advanced Research and Development Activity*) Yhdysvaltain puolustusministeriöstä. TREC:in päätavoitteena oli tukea tekstitiedonhaun tutkimusta tarjoamalla tutkimuksen käyttöön laajoja kokotekstikokoelmia. Lisäksi tavoitteena oli lisätä yhteistyötä eri alojen tutkijoiden kesken, nopeuttaa teknologisten keksintöjen tuotteistamista ja luoda uusia arviointimenetelmiä. Jokaista TREC-konferenssia varten on luotu koeasetelmia, jotka koostuvat tutkimuslinjoista (*tracks*). TREC käsittää testikokoelman, valmiit hakuaiheet ja aiheita vastaavien dokumenttien relevanssiarviot. Testikokoelman laajuus vaihtelee tutkimuslinjoittain. Konferensseihin osallistuneet ryhmät ovat suorittaneet koeasetelman mukaisen tiedonhaun omilla hakujärjestelmillään. TREC:in testikokoelman dokumenttien, hakuaiheiden sekä osallistujaryhmien ja sitä kautta käytettyjen hakujärjestelmien määrä on kasvanut vuosittain. (Text REtrieval Conference 2000.)

TREC-kokoelman relevanssiarvioinnissa käytettiin ”poolaus” –menetelmää. Menetelmässä pooliin kuuluvat mahdollisesti relevantit kokoelman dokumentit kerätään eri hakujärjestelmien palauttamista ranking-listoista. Pooliin kerätään yleensä sata ensimmäistä dokumenttia jokaisesta listasta. Tämän jälkeen arvioijat tekevät relevanssiarvioinnin jokaiselle poolin dokumentille. Menetelmän oletuksena on, että enemmistö mahdollisista relevanteista dokumenteista sisältyy luotuun pooliin ja toisaalta dokumentit, jotka jäävät poolin ulkopuolelle, eivät ole relevantteja. TREC:in relevanssiarvioinnissa käytettiin binääristä arviointiasteikkoa. Binääriarvostelu korostaa staattista aiheenmukaisuutta, eikä siinä ole mukana arvosteluprosessiin vaikuttavaa toiston ominaisuutta (Harter & Hert 1997, 25; Baeza-Yates & Ribeiro-Neto 1999, 89; Ingwersen & Järvelin 2005, 123-130.)

Ensimmäinen vuorovaikutteisen tiedonhaun tutkimuslinja oli TREC:issä mukana vuonna 1997. Käyttäjänäkökulma ja tiedonhakuprosessin vuorovaikutteisuus pyrittiin ottamaan huomioon laboratoriomallin mukaisessa arvioinnissa. Tavoitteena oli tutkia vuorovaikutusta koko tiedonhakuprosessin sekä järjestelmän tuloksen näkökulmasta. (Harter & Hert 1997, 25-26; Text REtrieval Conference 2000.) TREC:in muissa tutkimuslinjoissa käytetty arvioinnin järjestelmäkeskeinen lähestymistapa toimi pohjana suurelta osin myös vuorovaikutteisessa linjassa (Borlund 2000, 55).

Vaikka vuorovaikutteisia kokeita järjestettiin jo TREC-3:ssa, hakuaiheet niissä olivat rajoittuneita ja epärealistisia, eivätkä ne siten toimineet todellisia tiedontarpeita vastaavina. Hakijat suorittivat haun annetun aiheen mukaisesti eikä heitä kannustettu tiedontarpeen omakohtaiseen tulkintaan. Siten myös hakijoiden tiedontarpeesta muodostaman tulkinnan mahdollinen vahvistuminen ja kehittyminen hakuprosessin aikana rajattiin tutkimuksessa ulkopuolelle. Hakijoilla ei myöskään ollut mahdollisuutta tehdä relevanssiarviointia, mutta heillä oli pääsy etukäteen tehtyihin relevanssiarviointeihin. Edellä esitettyjen rajoitusten vuoksi hakijat toimivat kyseisessä tilanteessa pikemminkin tiedonvälittäjinä kuin lopullisina käyttäjinä, joita heidän oli tarkoitus vastata. Hakuaiheiden ja relevanssiarviointien määrittelyllä ja käytöllä tutkimuksessa voidaan vaikuttaa siihen, mikä on hakijan rooli tutkimuksessa. TREC-4:sta lähtien hakijat ovat voineet tehdä relevanssiarviointia tuloksista, mutta taustalla on kuitenkin ollut tarkka ohjeistus siitä, mitä hakutehtävässä tulisi hakea. Hakijoiden tekemien relevanssiarvioiden perusteella on todettu, että relevanssiarviointeja tehdään eri tarkoituksiin ja eri kriteereihin pohjautuen. (Borlund 2000, 55-56.)

TREC:in vuorovaikutteisesta linjasta voidaan todeta, että käyttäjä oli osittain mukana tutkimuksessa ja vuorovaikutuksessa järjestelmän kanssa, mutta järjestelmäarviointi tehtiin kuitenkin sen perusteella, miten käyttäjä löytää ennalta arvioituja dokumentteja. TREC:in testikokoelmaan kuuluvista dokumenteista suurin osa oli uutiskokoelmista. Dokumentit oli merkitty SGML:llä, mutta ne olivat rakenteeltaan ja muodoltaan hyvin yksinkertaisia. Ammattilaisten muotoilemissa hakuaiheissa oli aiheotsikon ja lyhyen aihekuvauksen lisäksi ohjeistus, mitä aihetta vastaavien relevanttien dokumenttien tulee sisältää. Kaikille hakuaiheille oli ominaista, että ne olivat muuttumattomia ja tarkkaan muotoiltuja eli ne eivät oikeastaan vastanneet todellisia tiedontarpeita. (Baeza-Yates & Ribeiro-Neto 1999, 88-90; Ingwersen & Järvelin, 2005, 123-130.)

TREC-8:n jälkeen pidetyssä SIGIR (*Special Interest Group on Information Retrieval*) seminaarissa 19 osallistujaa pohtivat TREC:in vuorovaikutteisen tiedonhaun tutkimuslinjan ongelmia, rajoitteita ja kehittämistä. Osallistujien esityksissä ja keskusteluissa tuli esiin yleisiä vuorovaikutteisen koetilanteen ongelmia varsinkin verrattuna laboratoriomalliin, esimerkiksi ajalliset ja taloudelliset panokset, käyttäjien tekemien mahdollisten virheiden ja järjestelmässä esiintyvien virheiden vaikutus. Osallistajat miettivät, miten TREC:in kontrolloitua ja siten rajattua tutkimuskehystä voitaisiin parhaiten hyödyntää vuorovaikutteisten koetilanteiden järjestämiseen ja arviointiin. Seminaarissa päätettiin seuraaviin suosituksiin TREC:in vuorovaikutteisen linjan kehittämiseksi. Hakijoiksi täytyisi saada monentyylisiä, todellisia käyttäjiä edustavia hakijoita. Hakuaiheissa tulisi olla enemmän vaihtelua ja niiden tulisi muistuttaa mahdollisimman hyvin todellisessa elämässä esiintyviä tiedonhakuja, joille on ominaista muun muassa tiedon päällekkäisyys sekä uutuusarvon ja tietolähteen arvovallan merkitys. Nämä seikat tulevat esille varsinkin Internet-tiedonhaussa, joka on merkittävä osa reaali maailman tiedonhakua. Seminaarissa esitettiin tutkimuksen perusteella Internet-tiedonhaussa yleisimmin esiintyvät hakuaiheet, jotka käsittelevät henkilökohtaista terveydenhoitoa, ostotapahtumaa, matkasuunnittelua ja taustatiedonkeruuta projektia varten. Osallistajat näkivät olennaisena osana Internet-tiedonhaun, web-tiedon heterogeenisten ja dynaamisten ominaisuuksien sekä Internet-tiedonhaussa usein esiintyvien hakuaiheiden hyväksikäytön. Lisäksi havainnollisten tutkimusten tulisi tukea kokeelliseen tutkimukseen perustuvaa järjestelmävertailua. (Hersh & Over 2000, 1-4.)

### **3.2 INEX**

INEX (*INitiative for the Evaluation of XML Retrieval*) tutkii ja arvioi nimensä mukaisesti tiedonhakua XML-ympäristössä. INEX-tutkimukset aloitettiin vuonna 2002. Kahden ensimmäisen vuoden aikana hankkeeseen osallistui 41 ryhmää. Vuosittaisissa tutkimuksissa tiedonhakua on tutkittu eri näkökulmista hankkeeseen luotujen tutkimuslinjojen avulla. Vuodesta 2004 lähtien hankkeeseen on kuulunut INEX *Interactive Track*, jossa on keskitytty vuorovaikutteisen tiedonhaun erityispiirteiden tutkimiseen. Vuorovaikutteisen tiedonhaun tutkimuslinjan päämääränä on tutkia hakijoiden käyttäytymisistä heidän hakiessaan tietoa XML-dokumenteista. Lisäksi tavoitteena on tutkia ja kehittää menetelmiä XML-tiedonhakua varten. (INEX 2004; Tombros, Larsen & Malik 2004.)

Tombros, Larsen ja Malik (2004) ovat tehneet alustavan yhteenvedon INEX:in vuorovaikutteisen tutkimuslinjan tuloksista hakijoiden käyttäytymisestä koskien. He tutkivat lokitietojen perusteella,

miten hakijat olivat tarkastelleet tuloksiin saamiaan komponentteja: suoraan listasta, dokumenttien rakennetta hyödyntäen vai *next/prev*-painikkeiden avulla. He raportoivat lisäksi, miten hakijoiden tekemä relevanssiarviointi jakautui käytetyllä relevanssiasteikoilla. Pharo ja Nordlie (2005) tutkivat tarkemmin hakijoiden tekemiä relevanssiarvioita dokumenttien eri rakennetasoilla ja sitä, miten rakenteisuus vaikuttaa hakijoiden relevanssiarviointiin ja heidän kykyynsä arvioida rakenteisia dokumentteja.

Merkittävä ero on TREC:in tutkimuskehikseen verrattuna INEX-tutkimuskokoelman dokumenttien tallennus XML-muodossa, mikä mahdollistaa dokumenttien loogisen rakenteen hyödyntämisen tiedonhaussa. Toinen merkittävä edistys vuorovaikutteisuuden huomioon ottamiseksi on se, että realistista tiedonhakua ja todellisen käyttäjän tiedontarvetilannetta on pyritty jäljittelemään luomalla käyttäjälle testiaineiston kohdalla valinta- ja vaikutusmahdollisuuksia perinteisen laboratoriomallin puitteissa. INEX-hankkeen vuorovaikutteisessa tutkimuslinjassa hakijoille suotiin mahdollisuus tiedontarpeen omakohtaiseen tulkintaan. Samalla hakijoiden tekemä relevanssiarviointi perustui heidän omakohtaiseen tulkintaansa toisin kuin TREC:issä, jossa annettuihin aihekuvauksiin kuului ohjeistus, mitä relevanttien dokumenttien tulee sisältää.

INEX 2004 –aineiston esittelyn yhteydessä luvussa 5 käsitellään näitä vuorovaikutteisen tutkimuslinjan erityispiirteitä tarkemmin. Koska dokumenttien XML-muoto ja XML-tiedonhaun menetelmien kehittäminen ovat keskeisessä roolissa INEX:in vuorovaikutteisen tiedonhaun tutkimuslinjassa, käsitellään seuraavissa alaluvuissa myös XML-ympäristöä ja XML-muotoisten dokumenttien merkitystä tiedonhaussa.

### **3.2.1 XML-kieli ja XML-dokumentin rakenne**

XML-kielen käyttö on yleistynyt varsinkin tieteellisen tiedon säilyttämisessä, digitaalisissa kirjoissa ja yleensä Internetissä. XML-ympäristöön suunniteltujen tiedonhakujärjestelmien tavoitteena on hyödyntää dokumenttien loogista rakennetta, jonka XML-kielen käyttö mahdollistaa. (INEX 2004).

Elektroniset dokumentit tai niitä kuvaavat viitteet ovat tietueita, joilla on sekä fyysinen että looginen rakenne. Fyysinen rakenne koskee tietueen esittämistä tietokoneissa bittijonoina ja näiden jäsentämistä eritasoisiksi kokonaisuuksiksi tietojenkäsittelyä varten. Looginen rakenne jäsentää tietue-

een määrämuotoiseksi tietoalkioiden kokonaisuudeksi, joka kuvaa jotakin kohdetta, asiantilaa tai tapahtumaa. Dokumenttien rakenteiden kuvauskielet jäsentävät dokumenttien sisältöä ja kuvaavat sen rakenteena. Tällöin dokumentti nähdään rakenteena, jonka osia ovat otsikko, tekijämerkintä ja joukko lukuja, joilla voi olla alalukuja. Alaluvut voivat koostua kappaleista, taulukoista ja kuvista. (Järvelin 1995, 85; 91).

Merkkauskieltä (*Markup*) käytetään esimerkiksi kuvailemaan muotoiluja, rakennetietoja, tekstin semanttisia tietoja tai ominaisuuksia. SGML (*Standard Generalized Markup Language*) on standardi metakieli, joka määrittää merkkauškieliä. SGML:lla tehtyjen määrittelyjen avulla on mahdollista esittää yhtenäisesti minkä tahansa dokumentin looginen rakenne. SGML ei ota kantaa dokumentin esitysmuodon määrittämiseen. Eli jokainen SGML-dokumenttia käsittelevä ohjelma voi esittää saman dokumentin juuri kyseiselle ohjelmalle tyypillisellä tavalla. Siten dokumentit voidaan tallettaa laitteisto- ja ohjelmistoriippumattomalla tavalla, riippumatta myös niiden esitysmuodoista. (Baeza-Yates & Ribeiro-Neto 1999, 149; Jaakohuhta 2003, 489.)

XML (*eXtensible Markup Language*) on SGML-kielestä johdettu, yksinkertaistettu metakieli. Se ei ole varsinainen merkkauškieli, mutta se voi sisältää merkkauškieliä SGML:n tapaan. XML on erityisesti webiä varten suunniteltu metakieli. Sen kehittäminen on lähtöisin yleisimmin käytetyn webin merkkauškielen, HTML:n, asettamista rajoituksista. XML muistuttaaakin HTML-kieltä, mutta se ei ole tarkoitettu sivunkuvauskieleksi, vaan sillä kuvataan tiedon rakenne ilman ennalta määrättyjä koodeja. Siten XML on yleinen metakieli dokumenttien rakenteelliselle kuvaamiselle, mikä mahdollistaa sen käytön hyvinkin erilaisiin tarkoituksiin. (Baeza-Yates & Ribeiro-Neto 1999, 154.)

Dokumentin rakennetta koskien XML-dokumentin tulee täyttää kansainvälisen W3-konsortion (W3C) määrittämät vaatimukset ollakseen rakenteellisesti oikein ja hyvin muotoiltu. Vaatimukset vaikuttavat XML-dokumentin loogiseen ja fyysiseen rakenteeseen. Lisäksi XML:ssa on tarkkaan määrätty kielioppi, muun muassa pienillä ja isoilla kirjaimilla on eri merkitys. Vaatimusten merkitystä katsotaan tarkemmin alla, ja havainnollisuuden vuoksi esitetään yksinkertainen esimerkki oikein muotoillusta XML-dokumentista.

XML-dokumentit koostuvat *elementeistä*. Dokumentissa täytyy olla yksi (ja ainoastaan yksi) juurielementti. Jokaisen elementin alku ja loppu täytyy merkitä tunnisteilla eli tageilla. Elementin alkutunniste voi olla esimerkiksi muotoa <artikkeli> ja lopputunniste </artikkeli>. Elementit voivat si-



sältää tekstiä ja niillä voi olla arvoja sisältäviä ominaisuuksia, jotka on merkitty tunnisteiden sisään lainausmerkeillä, esimerkiksi <ryhmä nimi="Uutisartikkeli">. Ilman sisältöä oleva elementti voidaan lyhentää muotoon <alaotsikko/>. Lisäksi elementit voivat sisältää toisia elementtejä. Sisäkkäin olevat elementit muodostavat yläelementti-alaelementti –parin, joiden alku- ja lopputunnisteiden järjestys on säilytettävä eli yläelementtiin kuuluvan alaelementin alku- ja lopputunnisteiden on oltava yläelementin tunnisteiden sisällä, niin että elementit eivät mene ristiin keskenään. (Downing, Covington & Covington 2000, 541; Klein 2001, 26, W3C. 2004. Extensible Markup Language.) Kuviossa 2 esimerkin ryhmä-alaelementti on alku- ja lopputunnisteineen artikkeli-yläelementin sisällä.

```
<?xml version="1.0"?>
<artikkeli>
  <ryhmä nimi="Uutisartikkeli">
    <otsikko>Korot nousussa</otsikko>
    <alaotsikko/>
  </ryhmä>
</artikkeli>
```

Kuvio 2. Esimerkki XML-dokumentista.

### 3.2.2 XML-tiedonhaun erityispiirteitä

XML-perustaisessa tiedonhaussa järjestelmän on mahdollista palauttaa hakijalle tuloslistaan dokumenttien elementtejä, joiden voidaan olettaa olevan hakijan kannalta kaikkein relevanteimpia. Tämä tarkoittaa sitä, että tuloslistassa olevien dokumenttien rakennetasot voivat vaihdella. Kun perinteisesti hakujärjestelmä on palauttanut tulosjoukkoon listan kokonaisista dokumenteista, on XML-rakenteita hyödyntävän järjestelmän mahdollista palauttaa vain se osa dokumentista, jonka katsotaan olevan hakijan kannalta olennaisin tyydyttämään hänen tiedontarpeensa. (Pharo & Nordlie 2005, 1.) Toisin sanoen kun laaja, kokonainen dokumentti on saattanut sisältää paljon hakijalle epärelevanttia tietoa, mahdollistaa dokumentin jakaminen rakenteen mukaisiin elementteihin epärelevantin tiedon karsimisen tulosjoukosta.

Rakenteisten dokumenttien tiedonhaku voidaan jakaa rakenne- ja sisältökeskeisiin lähestymistapoihin. Rakenteikeskeisessä (*structural*) lähestymistavassa hakua pyritään rikastamaan dokumenttirakenteen käyttämisellä hakuehtoina: esimerkiksi hakusanan pitää esiintyä tietyissä dokumenttiraken-

teen elementissä tai jonkin ehdon täytyminen dokumenttirakenteen elementissä edellyttää elementille asetetun aiemman ehdon täyttymistä. Sisältökeskeisessä (*content-based*) tiedonhaussa tähdätään dokumenttien kyselyyn nähden relevanteimpien elementtien palauttamiseen. (Fuhr & Großjohann 2001, 172.)

## 4 ONGELMANASETTELU JA TUTKIMUSKYSYMYKSET

Tutkielmassa tarkastellaan hakijan hakukäyttäytymistä vuorovaikutteisessa tiedonhakutilanteessa. Tarkemmin tutkimuksen kohteena on hakijoiden kyselyn muodostukseen ja uudelleen muotoiluun liittyvä hakukäyttäytyminen. Eritasoisten elementtien selailu sekä hakijoiden suorittama relevanssiarviointi rajataan tarkastelun ulkopuolelle.

Tässä tutkielmassa tutkitaan, millaisia kyselyjä hakijat muodostavat suhteessa annettuun tehtäväkuvaukseen, kun tutkimustilanteessa on sovellettu reaali maailmaa simuloivaa tehtävää. Simuloidun tehtävän käytössä oletuksena on, että hakija pystyy orientoitumaan hakutilanteeseen ja kokee annetun tehtävän ikään kuin omaksi hakuaiheekseen. Siten hakija suorittaisi testitilanteen haun samalla tavalla kuin toimisi todellisessa tiedonhakutilanteessa.

Simuloidussa tehtävässä ei anneta yksityiskohtaisia ohjeita haun suorittamiseen, vaan luodaan kokonaisvaltaisempi tehtäväkuvaus hakutilanteesta. Tällä pyritään siihen, että tehtäväkuvaus ei vaikuttaisi liikaa hakijoiden toimintaan, vaan jättäisi tilaa hakijoiden omille tulkinnoille ja esimerkiksi omille hakusanavalinnoille. Toisaalta voidaan olettaa, että hakua suorittaessaan hakija käyttäisi alussa paljon hakusanoja suoraan tehtäväkuvauksesta, varsinkin jos aihealue ei ole hakijalle ennestään kovin tuttu. Haun edetessä hakijan taas voisi olettaa käyttävän enemmän tehtäväkuvauksesta löytymättömiä hakusanoja, esimerkiksi selailtuaan saamaansa tulosjoukkoa ja valitsemalla sieltä uusia hakusanoja. Joka tapauksessa simuloidun tehtävän käytön tavoitteena on vuorovaikutteinen tiedonhakutilanne, jossa kontrolloituun testiympäristöön luodaan tehtäväkuvauksen osalta realistiseen tiedonhakutilanteeseen verrattavissa olevat olosuhteet.

### **Tutkimuskysymykset**

Tutkimuksen pääkysymys voidaan muotoilla seuraavasti:

Millaisia kyselyjä hakijat muodostavat suhteessa tehtäväkuvaukseen ja millainen on heidän kyselyn uudelleen muotoiluun liittyvä hakukäyttäytymisensä vuorovaikutteisuuteen pyrkivässä tutkimusasetelmassa?

Tarkennetut alakysymykset:

- Kuinka suuri osa hakijoiden käyttämistä hakusanoista löytyy tehtäväkuvauksesta ja kuinka suuri osa tehtäväkuvauksen ulkopuolelta?
- Mitä tehtäväkuvauksen käsitteitä hyödynnetään eniten ja millaisia suosituimmat käsitteet ovat?
- Missä käsitesuhteessa tehtäväkuvauksen ulkopuoliset hakusanat ovat tehtäväkuvauksen käsitteisiin?
- Millaista kyselyn muokkaustaktiikkaa (tässä POISTO, LISÄYS, MUU MUUTOS ja UUSI) hakijat käyttävät?
- Onko kyselyn muodostuksessa eroja erityyppisten tehtäväkuvauksien välillä?
- Vaikuttaako tehtäväkuvaukseen liittyvän haun suorittamisen koettu helppous tai aiheutuneisuus kyselyn muodostukseen?

## 5 INEX 2004 AINEISTO

Tutkielman empiirinen aineisto on osa INEX 2004-hankkeen vuorovaikutteisen tiedonhaun tutkimuslinjan (*Interactive Track*) aineistoa. INEX-tutkimuslinjan tutkimusasetelma ja aineistonkeruumenetelmät esitellään seuraavaksi. Sitten katsotaan, mistä tutkielmaan valittu tutkimusaineisto koostuu.

### 5.1 Tutkimusasetelma

Vuorovaikutteisen tiedonhaun tutkimuslinjaan osallistuneet yhteistyötahot olivat yliopistoja ja korkeakouluja Aasiasta, Australiasta ja Euroopasta. Osallistujaryhmiä oli yhteensä kymmenen ja jokaisesta ryhmästä tutkimukseen osallistui vähintään kahdeksan henkilöä.

Tutkimuslinjassa käytettiin web-pohjaista käyttöliittymää sekä tutkimusta varten kehitettyä HyREX (*Hyper-media Retrieval Engine for XML*) -tiedonhakujärjestelmää. Peruskäyttöliittymässä (*Baseline*) hakutulokset näytettiin selainikkunassa yksinkertaisessa järjestetyssä listamuodossa. Toista, graafista käyttöliittymää, käyttivät vain yhden tahon osallistujat. Tämän tutkielman aineiston kaikki haut on suoritettu peruskäyttöliittymällä. Aikaa hakujen suorittamiseen ja samalla relevanssiarvioiden tekemiseen oli annettu korkeintaan 30 minuuttia. Ennen varsinaista tiedonhaun aloitusta hakijoita opastettiin esimerkkihaun avulla HyREX-järjestelmän käytössä. Tutkimuslinjan testiaineisto koostuu laboratoriomallille tyypillisesti dokumenttikokoelmasta, hakutehtävistä ja relevanssiarvioista.

#### 5.1.1 Dokumenttikokoelma

INEX-testiaineiston kokotekstinen dokumenttikokoelma on kerätty IEEE Computer Sciencen tiedellisistä julkaisuista ja aikakauslehtijulkaisuista vuosilta 1995-2002. Yhteensä 12 107 artikkelista koostuva kokoelma on kooltaan 494 MB. Dokumentit on tallennettu kokoelmaan XML-muodossa. XML-ympäristöön suunniteltujen tiedonhakujärjestelmien tavoitteena on hyödyntää dokumenttien loogista rakennetta. Tällöin järjestelmän on mahdollista palauttaa tulosjoukkoon kokonaisten do-

kumenttien sijaan komponentteja eli tiettyjä elementtejä dokumenteista, jotka vastaavat käyttäjän tekemää kyselyä. (INEX 2004; Kazai & Lalmas & Vries 2004, 72.)

### 5.1.2 Hakutehtävät

INEX-hankkeen vuorovaikutteisen tiedonhaun tutkimuslinjassa käyttäjille luotiin mahdollisuus henkilökohtaiseen tulkintaan tiedontarpeesta soveltamalla hakutehtävän osalta simuloitun tehtävän mallia. Mallin tarkoituksena on motivoida hakijaa ja auttaa häntä eläytymään tutkimusta varten järjestettyyn tiedonhakutilanteeseen niin, että kyselyn muodostus ja uudelleen muotoilu vuorovaikutuksessa järjestelmän kanssa olisi mahdollisimman realistista. (Tombros ym. 2004.) Simuloitun tehtävän kirjallista esitystä kutsutaan tässä tehtäväkuvaukseksi. Tehtäväkuvaus sisältää pelkän hakutehtävän lisäksi myös laajemman kertomuksen hakutilanteesta.

Vuorovaikutteiseen linjaan valittiin INEX-testiaineistosta neljä tehtäväkuvausta. Tehtäväkuvaukset jaettiin erilaista tiedonhakua edustavaan kahteen tyyppiin, taustatietoon ja vertailevaan tietoon. Taustatiedossa tavoitteena on löytää mahdollisimman paljon tehtäväkuvauksen aiheeseen liittyvää yleistä tietoa. Tehtäväkuvaus voidaan esittää yleisesti muodossa: ”Haluaisin löytää tietoa aiheesta X”. Vertailevassa tiedonhaussa halutaan löytää eroavaisuuksia eri käsitteiden välillä ja tehtäväkuvaus voitaisiin esittää seuraavasti: ”Etsi X:n ja Y:n välisiä eroja”. Aiemmin hakutehtäviä koskevissa ja hakutehtävätyypejä vertailevissa tutkimuksissa on osoitettu, että eri hakutehtävätyypit luovat hakijalle erilaisia kriteerejä heidän arvioidessaan hakutulosten relevanssia. Hakutehtävätyyppien voidaan olettaa vaikuttavan myös hakijan muuhun tiedonhakukäyttäytymiseen kuten selailuun ja navigointiin. (Tombros ym. 2004.)

Vuorovaikutteiseen tutkimuslinjaan osallistuneet hakijat valitsivat kummastakin tyypestä (taustatieto ja vertaileva tieto) haluamansa tehtäväkuvauksen eli he suorittivat yhteensä kaksi hakua. Tehtäväkuvausvalinnalla oli tarkoitus suoda hakijalle mahdollisuus valita häntä enemmän kiinnostava tehtäväkuvaus. Tehtäväkuvauksien esittämisessä käytettiin kiertävää järjestystä. B- ja C-tehtäväkuvauksen, jotka edustivat eri tyyppiä, esittämisjärjestystä vaihdeltiin siten, ettei kumpikaan tehtäväkuvauksista ollut koko ajan ensimmäinen tai jälkimmäinen hakijoille esitetyistä tehtäväku-

vauksista. (Tombros ym. 2004.) Kiertävän järjestyksen käytön tarkoituksena on estää tehtäväkuvausten järjestyksestä johtuva tutkimustulosten mahdollinen vinoutuminen.

Hakijat syöttivät kyselyn selainikkunassa olevaan kyselykenttään. Kysely sai sisältää yksittäisiä sanoja (hakusanoja) tai fraaseja lainausmerkeillä erotettuna. Erillisiä hakuoperaattoreita ei ollut mahdollista käyttää, mutta hakijat saattoivat käyttää + ja - merkkejä hakusanojen edessä painoarvon nostamiseksi tai laskemiseksi. Haun tulokset avautuivat uuteen selailuikkunaan. Tuloslistaan tuli sata parasta tulosta, kymmenen tulosta kullakin sivulla. Tuloslistassa tulokset olivat sijoitusjärjestyksessä ja jokaisen elementin otsikko, tekijätiedot sekä linkki elementin katselua varten olivat näkyvissä (Kuva 1).



dbdk\_training in Baseline System

Search

query was: text classification naive bayes  
Results 1 - 10 of 100.  
Result pages: 1 2 3 4 5 6 7 8 9 10 next

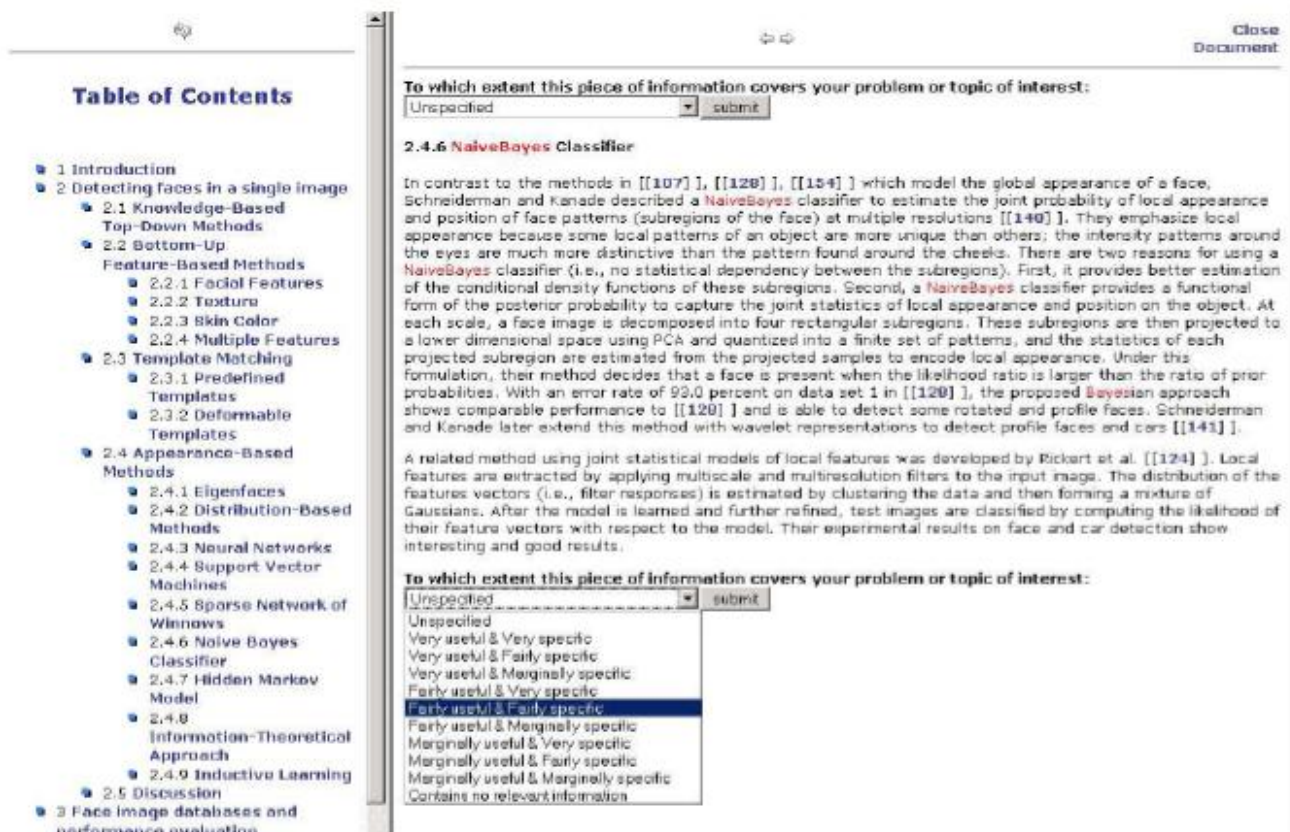
**HyREX**

**Search Result**

- 1: (0.247) **Scalable Feature Mining for Sequential Data**  
Neel Lesh Mitsubishi Electric Research Lab Mohammed J. Zaki Rensselaer Polytechnic Institute Mitsunori Ogihara University of Rochester  
Result path: /article[1]/body[4]/sec[5]
- 2: (0.204) **Probability and Agents**  
Marco G. Valtorta University of South Carolina mgv@coe.sc.edu Michael N. Huhns University of South Carolina huhns@sc.edu  
Result path: /article[1]/body[4]/sec[3]
- 3: (0.176) **Combining Image Compression and Classification Using Vector Quantization**  
Karen L. Oehler Member IEEE Robert M. Gray Fellow IEEE  
Result path: /article[1]/body[4]/sec[4]/ss1[2]/ss2[4]
- 4: (0.175) **Text-Learning and Related Intelligent Agents: A Survey**  
Dunja Mladenic J. Stefan Institute  
Result path: /article[1]/body[5]/app[4]/sec[5]
- 5: (0.175) **Detecting Faces in Images: A Survey**  
Ming-Hsuan Yang Member IEEE David J. Kriegman Senior Member IEEE Narendra Ahuja Fellow IEEE  
Result path: /article[1]/body[4]/sec[2]/ss1[9]/ss2[10]

Kuva 1. Tuloslista peruskäyttöliittymässä.

Hakijan valitessa tuloslistasta tarkasteltavan elementin avautuu jälleen uusi ikkuna (Kuva 2), jonka vasemmassa reunassa näkyy sisällysluettelo dokumentista, johon elementti kuuluu. Kyseinen elementti on erotettu sisällysluettelossa keltaisella korostuksella, tekstistä löytyvät hakusanat punaisella.



Kuva 2. Elementin selailuikkuna ja relevanssiarviointi.

### 5.1.3 Relevanssiarviointi

Hakijat arvostelivat tulosedokumenttien eri elementtien relevanssia käyttökelpoisuuden ja spesifisyyden mukaan. Relevanssiarviointi tehtiin valitsemalla alavetovalikosta haluttu arvio 10-portaiselta asteikolta. Käyttökelpoisuudella viitataan siihen, miten tyhjentävästi dokumentti tai sen elementti käsittelee tehtäväkuvausta. Erittäin käyttökelpoinen dokumentti auttaa useimpien tehtäväkuvaukseen liittyvien näkökulmien ratkaisemisessa, mutta melko käyttökelpoinen auttaa vain osittain. Spesifisyyden on tarkoitus määrittää sitä, missä määrin dokumentti on keskittynyt kyseessä olevaan tehtäväkuvaukseen eli onko tehtäväkuvauksen aihe elementissä ainoa, keskeinen vai vähäinen tema. Sekä käyttökelpoisuudelle että spesifisyydelle luotiin kolme arviointitasoa: erittäin, melko ja marginaalinen. (Pharo & Nordlie 2005, 2-3.) Näiden yhdistelmästä muodostui Relevanssimäärittelmän mukainen asteikko.



### **Relevanssimääritelmä**

- A Erittäin käyttökelpoinen & erittäin spesifi
- B Erittäin käyttökelpoinen & melko spesifi
- C Erittäin käyttökelpoinen & jonkin verran spesifi
- D Melko käyttökelpoinen & erittäin spesifi
- E Melko käyttökelpoinen & melko spesifi
- F Melko käyttökelpoinen & jonkin verran spesifi
- G Jonkin verran käyttökelpoinen & erittäin spesifi
- H Jonkin verran käyttökelpoinen & melko spesifi
- I Jonkin verran käyttökelpoinen & jonkin verran spesifi
- J Epärelevantti
- U Määrittelemätön (Tombros ym. 2004.)

Asteikossa kirjain J kertoo, että dokumentti ei sisällä tehtäväkuvauksen kannalta relevanttia tietoa lainkaan. Kirjain U tallentui lokitiedostoon, jos käyttäjä oli katsonut dokumentin, mutta ei ollut tehnyt sille relevanssiarviointia.

## **5.2 Aineistonkeruu**

INEX-tutkimuksessa käytettiin useita aineistonkeruumenetelmiä. Ennen koetilannetta hakijat täyttivät kyselylomakkeen, jossa saatiin tietoa hakijoiden aiemmasta hakukokemuksesta. Koetilanteen jälkeisellä kyselylomakkeella kerättiin palautetta hakutilanteesta. Samoin hakijakohtaista tietoa, muun muassa tehtäväkuvaukseen liittyvästä aiemmasta aiheutuksesta ja haun suorittamisen koetusta vaikeudesta, kerättiin kyselylomakkeella ennen ja jälkeen kummankin suoritetun haun. Hakuprosessin eri vaiheissa osallistujien hakukäyttäytymistä havainnoitiin ja prosessin lopussa osallistujia haastateltiin. (Pharo & Nordlie 2005, 240.)

Lokitiedostoihin tallentui tapahtumatietoja hakuprosesseista. Petersin (1993) määritelmän mukaan lokitiedosto on sähköisesti tallennettuja tietoja tiedonhakijan ja tiedonhakujärjestelmän välisestä vuorovaikutuksesta henkilön hakiessa tietoa kyseisestä järjestelmästä (Peters 1993, tässä Griffiths, Hartley & Willson 2002, 2). INEX:in lokitiedostoista nähdään tapahtuma-ajan ja tunnistetietojen lisäksi hakijan suorittamat kyselyt hakusanoineen, tuloslistan elementit, mitä elementtejä hakija on katsonut ja arvioinut sekä millaisia hakijan antamat relevanssiarviot ovat olleet.

### 5.3 Tutkimusaineiston valinta ja käsittely

INEX:in vuorovaikutteisen tiedonhaun tutkimusaineistosta valittiin tutkielman tutkimusaineistoon kaikki hakijat, jotka olivat tutkimuksessa valinneet taustatietoa edustavasta tyypistä tehtäväkuvauksen B1 tai vertailevan tiedon tyypistä tehtäväkuvauksen C2. Valinta tehtiin sen perusteella, että hakijat olivat tutkimustilanteessa valinneet selvästi useimmin tehtäväkuvaukset B1 ja C2. Näin saatiin käsittelyyn tarpeeksi laaja aineisto, jossa tehtäväkuvaukset edustavat erityyppisiä tehtäväkuvauksia. Hakijoista 54 oli valinnut tehtäväkuvauksen B1 ja 67 tehtäväkuvauksen C2. Hakijoille esitetyt tehtäväkuvaukset B1 ja C2 näyttivät seuraavanlaisilta.

#### Tehtäväkuvaus B1:

You are writing a large article discussing virtual reality (VR) applications and you need to discuss their negative side effects.

What you want to know is the symptoms associated with cybersickness, the amount of users who get them, and the VR situations where they occur. You are not interested in the use of VR in therapeutic treatments unless they discuss VR side effects.

#### Tehtäväkuvaus C2:

You are working on a project to develop a next generation version of a software system. You are trying to decide on the benefits and problems of implementation in a number of programming languages, but particularly Java and Python.

You would like a good comparison of these for application development. You would like to see comparisons of Python and Java for developing large applications. You want to see articles, or parts of articles, that discuss the positive and negative aspects of the languages. Things that discuss either language with respect to application development may be also partially useful to you.

Ideally, you would be looking for items that are discussing both efficiency of development and efficiency of execution time for applications. (Tombros ym. 2004.)

Kyselyjen ja hakusanojen käytön tarkastelua varten lokitiedostoista tallennettiin tiedot hakijoista ja kyselyistä hakusanoineen Excel-taulukkoihin. Lokitiedostoissa peräkkäin olevista täsmälleen samanlaisista kyselyistä tallennettiin vain ensimmäinen, sillä jälkimmäinen (jälkimmäiset) kertoi hakusession aikana tapahtuneesta selailusta eri elementtien välillä, jolloin paluu alkuperäiselle kyselysivulle tallentui lokitiedostoon uudelleen kyseisenä kyselynä. On mahdollista mutta ei kovin todennäköistä, että hakija olisi kirjoittanut täysin samanlaisen kyselyn toistamiseen, eikä tällaisen kyselyn poisjättämisellä ole siten suurta merkitystä tulosten analysoinnissa. Tehtäväkuvauksesta B1 hakijat olivat tehneet yhteensä 292 kyselyä ja tehtäväkuvauksesta C2 yhteensä 460 kyselyä.

Korrelaatioanalyysin avulla haluttiin lisäksi tutkia, onko taustamuuttujilla vaikutusta kyselyn muodostukseen. Koetilanteessa hakijoilta kerättiin vastauksia kyselylomakkeella sekä ennen haun suo-

rittamista että sen jälkeen. Korrelaatioanalyysiin valittiin haun suorittamisen jälkeen täytetystä kyselylomakkeesta kaksi kysymystä, joiden mahdollisen vaikutuksen kyselyn muodostukseen katsottiin olevan merkityksellistä tämän tutkimuksen kannalta. Kysymykset olivat alkuperäismuodossaan:

2. Was it easy to do the search on the given task?
7. Did you know a lot about the topic of the task in advance?

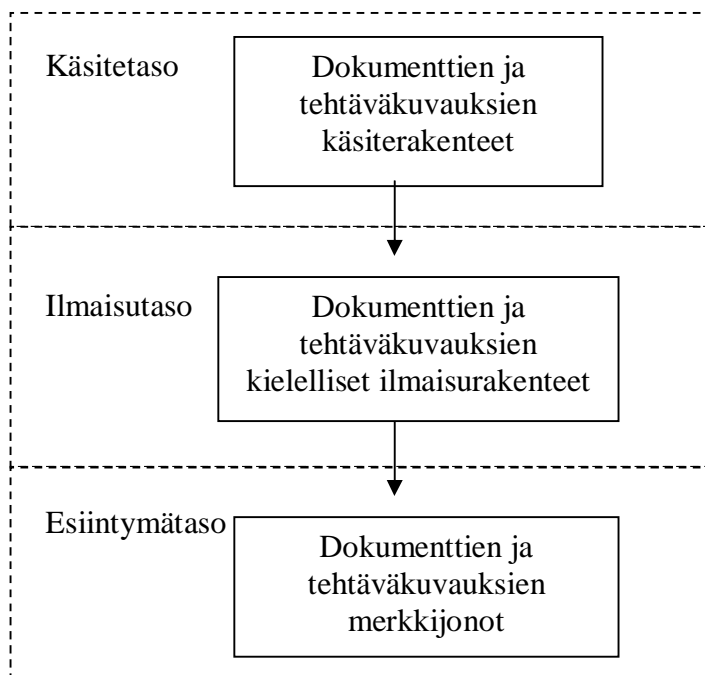
Valituissa taustamuuttujissa kysymyksessä numero kaksi hakijat siis arvioivat valitsemaansa tehtäväkuvaukseen liittyvän haun suorittamisen helppoutta ja kysymyksessä numero seitsemän tehtäväkuvaukseen liittyvää aiempaa aihetuntemustaan. Molemmissa kysymyksissä vastaukset annettiin asteikolla yhdestä viiteen, jossa arvo yksi merkitsi ”ei yhtään” ja arvo viisi ”erittäin”.

Kyselylomakkeen kysymykset ennen haun suorittamista koskien kyseiseen tehtäväkuvaukseen liittyvän haun suorittamisen helppoutta ja tehtäväkuvaukseen liittyvää aihetuntemusta jätettiin pois, koska haluttiin tutkia nimenomaan haun toteuttamisen jälkeisiä vastauksia eikä hakijoiden arvioin- teja ennen haun suoritusta. Samoin muut haun suorittamisen jälkeisen kyselylomakkeen kysymyk- sistä jouduttiin hylkäämään. Kysymykset olivat liian samankaltaisia (Oliko haun aloittaminen help- poa?), niiden ei todettu liittyvän tämän tutkielman aihepiiriin (Miten hakujärjestelmä tuki tiedonha- kuasi?) tai niissä ei todettu olevan tarpeeksi variaatiota käytetyllä viisiportaisella asteikolla, jotta niiden yhteyden tutkiminen kyselyn muodostukseen olisi ollut mielekästä.

## 6 TUTKIMUSAINEISTON ANALYYSIMENETELMÄT

Molemmille tehtäväkuvauksille tehtiin käsiteanalyysi, jossa niistä tunnistettiin käsitteet ja niiden väliset suhteet (Liite 1). Hakijoiden käyttämiä hakusanoja verrattiin tehtäväkuvauksen sisältöön kyselykohtaisesti kolmella eri tasolla: esiintymä-, ilmaisu- ja käsitetasolla. Kolme tasoa ovat tiedon tallennuksen ja haun yleisen tasoperiaatteen mukaisia. Tasoperiaate esitetään kuviossa 3.

Tasot ovat periaatteessa aina mukana, kun ihmiset tarkastelevat dokumentteja ja kyselyjä tiedontarpeiden esityksinä. Tekniseltä kannalta dokumentit ja kyselyt ovat merkkijonojen joukkoja (Esiintymätaso). Kirjoituskielellä esitetään luonnollisen kielen ilmaisuja (Ilmaisutaso), jotka puolestaan edustavat dokumentin ja kyselyn käsitteellistä sisältöä (Käsitetaso). (Järvelin & Sormunen 1999, 124.)



Kuvio 3. Tasoperiaate. Muunnelma tiedon tallennuksen ja haun tasoperiaatteen kuviosta (Järvelin 1995, 69).

Tasojen tarkastelussa selvitettiin, miten suuri osa kyselyjen hakusanoista on peräisin tehtäväkuvauksesta kunkin tason periaatteiden mukaisesti. Esiintymätasolla hakusanojen ja tehtäväkuvauksen sanojen vertailu suoritettiin tarkkaan merkkijonotasolla, jolloin kyselyissä esiintyvät hakusanat ja tehtäväkuvauksen sanat luokiteltiin samoiksi, mikäli niiden merkkijonot vastasivat täysin toisiaan

(*development - development*). Ilmaisutason vertailussa tarkasteltiin vaihtoehtoisia ilmaisumuotoja tehtäväkuvauksen sanoista (*traeatment - treatments*). Käsitetason tarkasteluun lisättiin edellisten tasojen lisäksi kuuluvaksi johdokset ja synonyymit, jotka edustavat tehtäväkuvauksen käsitteitä.

Esiintymätasolla pitäisi periaatteessa puhua ainoastaan merkkijonoista eikä hakusanoista, sillä hakujärjestelmät käsittelevät vain merkkijonoja (Järvelin 1993, 127). Tässä tutkielmassa esiintymä- ja ilmaisutason nimityksien osalta tehdään kuitenkin yksinkertaistaminen: kyselyjen merkkijonoista ja ilmaisuista puhuttaessa käytetään molemmista nimitystä hakusana. Vastaavasti puhuttaessa näillä tasoilla tehtäväkuvauksen merkkijonoista ja ilmaisuista käytetään nimitystä sana. Käsitetasolla käytetään nimityksiä käsite tai hakusanan edustama käsite.

Seuraavaksi aineistosta käsiteltiin hakijoiden käyttämiä, tehtäväkuvauksen ulkopuolelta olevia hakusanoja ja tutkittiin niiden edustamien käsitteiden suhdetta tehtäväkuvauksen käsitteisiin. Hakusana määriteltiin tehtäväkuvauksen ulkopuoliseksi, jos se ei ollut ilmaisutason kriteerien mukaan yhdenmukainen tehtäväkuvauksen sanojen kanssa. Tehtäväkuvauksen ulkopuolisten hakusanojen edustamat käsitteet jaettiin edelleen viiteen eri luokkaan. Tarkastelu tehtiin synonymian, hierarkian (ylä- ja alakäsite) ja assosiaatiosuhteen luokituksella. Viides luokka oli hakusanoille, jotka eivät olleet käsitteellisessä suhteessa tehtäväkuvauksen käsitteisiin. Tehtäväkuvauksen ulkopuolisten hakusanojen edustamien käsitteiden luokitus löytyy liitteestä 2. Hakijoiden valitsemien hakusanojen ja tehtäväkuvauksien sanojen merkitysten ja käsitteiden keskinäisten suhteiden selvittämiseksi apuna käytettiin Kielikoneen NetMot-sanakirjaa, Oxford English –sanakirjaa, Oxford American Thesaurus of Current English –tesaurusta ja Yleistä suomalaista asiasanastoa (YSA).

Kyselyn uudelleen muotoilun tarkastelussa hakijoiden suorittamat jatkokyselyt luokiteltiin neljään eri muokkaustaktiikkaan. Jatkokyselyt ovat hakijoiden muokkaamia, ensimmäisten kyselyjen jälkeisiä seuraavia mahdollisia kyselyjä. Kyselyn muokkaustaktiikat olivat POISTO, LISÄYS, MUU MUUTOS ja UUSI.

Esiintymä-, ilmaisu- ja käsitetason periaatteet ja laskemisen menetelmät käydään tarkemmin läpi luvussa 6.1. Sitten katsotaan käsitesuhteiden ja kyselyn uudelleen muotoilun luokitteluja luvussa 6.2. Viimeisessä luvussa käsitellään korrelaatioanalyysimenetelmä, jonka avulla tutkitaan taustamuuttujien mahdollista vaikutusta kyselyn muodostukseen.

## 6.1 Tasojen periaatteet ja hakusanayhdenmukaisuus

Hakijoiden käyttämien hakusanojen ja niiden edustamien käsitteiden sekä tehtäväkuvauksen sanojen/ käsitteiden vertailussa sovellettiin Iivosen (1995, 72) väitöskirjassaan käyttämää kyselyjen hakusanojen yhdenmukaisuuden menetelmää. Iivonen laski menetelmällä sekä hakijoiden välisiä että hakijoiden sisäisiä hakusanojen ja käsitteiden yhdenmukaisuuksia Online-viitehaussa. Hakijan sisäisessä yhdenmukaisuudessa yhden ja saman hakijan käyttämien hakusanojen ja käsitteiden yhdenmukaisuutta tarkasteltiin eri hakusessioissa. Hakijoiden välisessä yhdenmukaisuudessa eri hakijoiden yhdenmukaisuuksia verrattiin keskenään. Samaa yhdenmukaisuuden laskemistapaa ovat Iivosen lisäksi käyttäneet muun muassa Saracevic (1984) ja Fidel (1985).

Tässä tutkielmassa yhdenmukaisuus laskettiin tehtäväkuvauksen ja kaikkien siitä tehtyjen kyselyjen välille hakusanojen ja käsitteiden osalta. Näin saatiin selville, miten paljon hakijat ovat käyttäneet tehtäväkuvauksesta löytyviä sanoja ja käsitteitä hakusanoinaan kutakin kyselyään kohden.

Olkoon  $K_x$  tehtäväkuvauksen  $X$  perusteella laadittu käsitteiden/hakusanojen joukko ja  $T_x$  tehtäväkuvauksen  $X$  sanojen joukko. Siten kaikilla kolmella tasolla yhdenmukaisuusluku kyselyittäin laskettiin oheisella kaavalla. Yhdenmukaisuusluku sijoittuu välille  $[0,1]$ , jolloin arvo 1 ilmaisee kyselyn hakusanojen/niiden edustamien käsitteiden ja tehtäväkuvauksen sanojen/käsitteiden täydellistä yhdenmukaisuutta.

$$\frac{| K_x \cap T_x |}{| K_x |}$$

### 6.1.1 Esiintymätaso

Esiintymätasolla yhdenmukaisuutta vertaillessa kyselyn hakusanoja ja tehtäväkuvauksen sanoja verrataan toisiinsa merkkijonon tarkkuudella. Tällöin hakusanat ja tehtäväkuvauksen sanat luokitellaan yhdenmukaisiksi, mikäli merkkijonot vastaavat merkki merkiltä täysin toisiaan (*symptoms – symptoms*). Kaavan mukainen yhdenmukaisuusluku kertoo sen, kuinka suuri osuus hakijan yhdessä kyselyssään käyttämistä hakusanoista löytyy tehtäväkuvauksesta sellaisenaan.

### 6.1.2 Ilmaisutaso

Toisella tasolla hakusanoja vertailtiin väljemmin ilmaisutasolla, jossa myös tehtäväkuvauksen sanoja niin sanotusti lähellä olevat hakusanat otettiin huomioon yhdenmukaisuudessa. Tällöin kyselyn hakusana luokitellaan samaksi tehtäväkuvauksen sanan kanssa, jos se kuuluu edellisen, esiintymätason mukaisen yhdenmukaisuuden lisäksi johonkin alla esitettyyn luokkaan.

- 1) yksikkö-monikko -variaatio tehtäväkuvauksen sanasta (*treatment - treatments*)
- 2) aikamuotovariaatio tehtäväkuvauksen sanasta (*develop – developed*)
- 3) virhe sanassa tai sanayhdistelmässä, jotka voidaan liittää tehtäväkuvauksen sanoihin joko esiintymätasolla tai ilmaisutason kahden edellisen kohdan 1) tai 2) mukaan (*lanuage – language, sideeffect – side effect*)

Kielitieteen määritelmän mukaan perussanalla voi olla useita toteutumia eli taivutusmuotoja (Karls-son 2004, 83). Tehtäväkuvauksien sanoja käsitellään tässä perussanoina, joiden taivutuksia, muunnelmia ja tulkintoja esiintyy hakijoiden käyttämissä hakusanoissa. Koska englannissa on taivutusmuotoja varsin vähän, koskevat perussanojen erilaiset toteumat molempien tehtäväkuvauksien osalta enimmäkseen yksikkö-monikko-variaatiota (esimerkiksi *symptom-symptoms*). Ilmaisutasolla käsitellään tehtäväkuvauksen sanoista käytettyjä taivutus- ja aikamuotoja kirjoitusvirheineen. Ilmaisutasolla kysymys on hakusanojen muodon normittamisesta ja yhdenmukaisuuslukujen oletetaan nousvan esiintymätason tiukempaan määrittelyyn verrattuna, edellyttäen tietenkin että ilmaisutasoon luokiteltavia hakusanoja on kyselyissä käytetty.

### 6.1.3 Käsitetaso

Käsitetason yhdenmukaisuuden tarkasteluun kuuluvat ilmaisutasoon kuuluvien hakusanojen lisäksi johdokset ja synonyymit, jotka edustavat tehtäväkuvauksen käsitteitä oheisen luokituksen mukaisesti.

- 1) saman tai eri sanaluokan johdos tehtäväkuvauksen sanasta (*therapy – therapeutic*)
- 2) synonyymi tehtäväkuvauksen sanalle (*advantage – benefit*)

Johdokset eli johdannaiset ovat perussanoista johtimilla muodostettuja sanoja, jotka ovat myös sanakirjoissa itsenäisiä hakusanoja toisin kuin toteumat. Englannin kielessä on useita sanan vartalon eteen liitettäviä johtimia kuten un-, pre- ja dis-. (Karlsson 2004, 84; 196.) Esiintymä- ja ilmaisutasolla hakusanoja tarkasteltiin yksittäin, jolloin lainausmerkeillä ilmaistujen fraasien sisältämät hakusanat eroteltiin ja käsiteltiin yksittäisinä hakusanoina (*virtual reality* -> [*virtual*], [*reality*]). Käsitetasolla fraasit tunnistettiin, jolloin yhden ja mahdollisesti useimmankin hakusanan sisältävä fraasi käsiteltiin yhtenä hakusanana (*virtual reality* -> [*virtual reality*]). Fraasien ilmaisemat käsitteet luokiteltiin samoiksi tehtäväkuvauksen käsitteiden kanssa, jos fraasin sisältämät hakusanat olivat jo esiintymä- tai ilmaisutason määritelmien perusteella luokiteltu yhdenmukaisiksi tai ne täyttivät jonkun käsitetason kolmesta määritelmästä.

## 6.2 Käsitesuhteet ja kyselyn uudelleen muotoilu

Tehtäväkuvauksen ulkopuolelta olevien hakusanojen edustamien käsitteiden suhdetta tehtäväkuvauksen käsitteisiin tarkasteltiin hakijoiden suorittamien ensimmäisten kyselyjen (aloituskyselyt) ja sitä mahdollisesti seuraavien, hakijoiden uudelleen muotoilemien kyselyjen (jatkokyselyt) osalta. Aloituskyselyissä esiintyvien hakusanojen edustamia käsitteitä verrattiin tehtäväkuvauksen käsitteisiin. Jatkokyselyiden käsitesuhteiden tarkastelua varten hakijoiden uudelleen muotoilemat kyselyt luokiteltiin ensin eri hakutaktiikoihin vertaamalla kyseistä jatkokyselyä sitä edeltäneeseen kyselyyn ja siinä käytettyihin hakusanoihin. Hakutaktiikoista valittiin kaksi, joissa esiintyvät, tehtäväkuvauksen ulkopuolelta olevat hakusanat luokiteltiin sen mukaan, millaisessa suhteessa niiden edustamat käsitteet olivat tehtäväkuvauksen käsitteisiin nähden.

Erilaisissa tiedonhakuympäristöissä ja -tilanteissa mahdollisesti ilmeneviä, yksityiskohtaisesti jaoteltuja hakutaktiikoita ja siirtoja (ks. esim. Bates 1979; Sormunen 1989, Fidel 1991) yhdistäen sovelletaan tässä tutkielmassa neljää yksinkertaistettua kyselyn muokkaustaktiikkaa: POISTO, LISÄYS, MUU MUUTOS ja UUSI. POISTO, LISÄYS JA MUU MUUTOS –taktiikat voivat esiintyä samassa kyselyssä, mutta UUSI voi esiintyä vain yksinään. Näiden luokkien lisäksi kyselyissä esiintyvät kirjoitusvirheet kirjattiin ylös erilliseen luokkaan. Kyselyjen hakusanat otetaan sellaisina kuin hakijat ovat ne antaneet. Toisin sanoen fraasi tunnistetaan, jos se on merkitty lainausmerkein,



muuten hakusanat käsitellään yksittäisinä. Hakusanan katsottiin olevan tehtäväkuvauksesta, mikäli se täytti ilmaisutason määritelmän ehdot.

- POISTO – kyselystä poistetaan 1 tai useampia hakusanoja siten, että siihen jää ainakin yksi edellisen kyselyn sana
  - 1) Tehtäväkuvauksesta (Ilmaisutaso)
  - 2) Tehtäväkuvauksen ulkopuolelta
- LISÄYS – kyselyyn lisätään 1 tai useampia hakusanoja
  - 1) Tehtäväkuvauksesta (Ilmaisutaso)
  - 2) Tehtäväkuvauksen ulkopuolelta
- MUU MUUTOS – kyselyssä jo olevia hakusanoja koskevat muokkaukset
  - 1) Fraasin muodostaminen
  - 2) Fraasin purkaminen
  - 3) Kirjoitusvirheen korjaus
- UUSI - kaikki edellisen kyselyn hakusanat on poistettu ja uudet hakusanat ovat suhteessa tehtäväkuvaukseen
  - 1) Tehtäväkuvauksesta (Ilmaisutaso)
  - 2) Tehtäväkuvauksen ulkopuolelta

Aloituskyselyissä sekä jatkokyselyiden LISÄYS- ja UUSI- taktiikoissa käytetyt tehtäväkuvauksen ulkopuoliset hakusanat järjestettiin luokkiin, joista neljä ensimmäistä on peräisin Batesin (1979, 210) määrittämistä kyselyn muokkaustaktiikoista: *neighbour* (synonyymit), *sub* (hierarkiassa suppeammat käsitteet), *super* (hierarkiassa laajemmat käsitteet) ja *relate* (hierarkiassa rinnakkaiset käsitteet). Viides luokka käsitti hakijoiden ideoimat omat hakusanat, jotka eivät olleet käsitteellisessä suhteessa tehtäväkuvauksen käsitteisiin. Viisikohtainen luokitus muodostui seuraavanlaiseksi:

- 1) synonyymit (*advantages – benefits*)
- 2) suppeampi käsite, alakäsite (*nausea – symptoms*)
- 3) laajempi käsite, yläkäsite (*sickness – cybersickness*)
- 4) rinnakkaiskäsite, assosiaatiosuhde (*cyberspace – virtual reality*)
- 5) oma hakusana, joka ei kuulu edellä mainittuihin luokkiin (*witness*)

Sitä mitä tehtäväkuvauksen sisällöstä hyödynnetään eniten, on mielekkäintä tarkastella käsitetason määritysten mukaan, jolloin saadaan mukaan myös hakijoiden laajempia tulkintoja tehtäväkuvauksen kuuluvista käsitteistä. Samaan tehtäväkuvauksen käsitteeseen kuuluvat hakusanat lasketaan kaikki yhteen, esimerkiksi hakusanojen *progaming*, *progaming* ja *program* käytön lukumäärä lasketaan yhteen *programming* käytön lukumäärän kanssa. Näin saadaan selville, miten paljon tehtä-

väkuvauksen käsitettä ja siihen käsitetasolla liittyvää hakusanaa käytettiin kyselyissä kaiken kaikkiaan. Tehtäväkuvauksista tunnistettujen kaikkien käsitteiden kattavuutta kyselyissä ei ole järkevää tutkia, koska tehtäväkuvauksessa pikemminkin tilannetta taustoittavat käsitteet (esim. *You would like to see ---*), eivät olleet haun suorittamisen kannalta olennaisia hakusanoja.

### 6.3 Korrelaatioanalyysi

Taustamuuttujien vaikutusta kyselyn muodostukseen tutkittiin korrelaatioanalyysin avulla. Kyselyn muodostus koskee suoritettujen kyselyjen ja valittujen hakusanojen määrää sekä kyselyjen hakusanojen ja tehtäväkuvauksen sanojen välisiä yhdenmukaisuuslukuja.

Korrelatiivisessa tutkimuksessa tutkitaan ilmiöiden välisiä yhteyksiä ja pyritään saamaan selville, miten jokin tietty asia tai ominaisuus voisi vaikuttaa jonkin toisen asian tai ominaisuuden ilmenemiseen. Tutkimuksessa selvitetään muuttujien oletettuja vaikutuksia. Kovin voimakkaita päätelmiä ilmiöiden välisistä syy-seuraussuhteista sen avulla ei voida tehdä, koska mitattavaan ominaisuuteen vaikuttavaa muuttujaa ei ole suoranaisesti mahdollista säädellä. (Nummenmaa 2004, 28.)

Korrelaatioanalyysia varten tehtiin Excel-taulukko, johon laskettiin hakijakohtaiset keskiarvot kyselyissä käytettyjen hakusanojen sekä esiintymä-, ilmaisu- ja käsitetasojen yhdenmukaisuuslukujen osalta. Lisäksi merkittiin suoritettujen kyselyiden määrät per hakija. Sitten poimittiin hakijoiden vastaukset koskien kyselylomakkeen kysymyksiä 2 (helppous) ja 7 (aihetuntemus), jotta voitiin tutkia niiden yhteyttä hakukäyttäytymiseen.

## 7 TUTKIMUSTULOKSET

Tutkimusaineiston yleisten tulosten jälkeen tutkimustuloksissa käsitellään kyselyissä esiintyvien hakusanojen ja tehtäväkuvauksen sanojen yhdenmukaisuutta kolmen eri tasoperiaatteen mukaan määriteltynä. Sen jälkeen tarkastellaan, mitä tehtäväkuvauksen käsitteistä hyödynnettiin eniten ja millaisia nämä käsitteet ovat. Sitten katsotaan tarkemmin eriteltynä, miten tehtäväkuvauksesta ja sen ulkopuolelta olevien hakusanojen käyttö jakaantui. Lisäksi tutkitaan hakijoiden käyttämien, tehtäväkuvauksen ulkopuolisten hakusanojen edustamien käsitteiden suhdetta tehtäväkuvauksen käsitteisiin sekä perehdytään kyselyn uudelleen muotoilussa käytettyihin muokkaustaktiikoihin. Viimeisenä käydään läpi korrelaatioanalyysin tuloksia taustamuuttujien mahdollisesta vaikutuksesta kyselyn muodostukseen.

### 7.1 Yleisiä tuloksia

B1-tehtäväkuvauksen oli valinnut 54 hakijaa, joiden suorittamien kyselyiden määrä oli yhteensä 292. Hakijakohtainen kyselyjen keskiarvo oli 5,4 kyselyä. Kyselyissä käytettiin keskimäärin 3,2 hakusanaa ja kokonaishakusanamäärä oli 933. Hakijoista 67 oli valinnut tehtäväkuvauksen C2. He tekivät tätä tehtäväkuvausta kohden yhteensä 460 kyselyä, jolloin hakijakohtaiseksi keskiarvoksi saatiin 6,9 kyselyä. Kyselyjä tehtiin C2-tehtäväkuvauksen osalta jonkin verran enemmän kuin tehtäväkuvauksen B1, mutta myös siinä käytettiin keskimäärin noin 3 hakusanaa kyselyä kohden. Kokonaishakusanamäärä C2-tehtäväkuvauksessa oli 1538 (Taulukko 1).

Tehtäväkuvaus	Hakijat	Kyselyt	Kokonaishakusanamäärä	Hakusanat fraaseineen	Uniikit <sup>1</sup> hakusanat	Kyselyt/hakija	Hakusanat/kysely	Kirjoitusvirheet hakusanoissa	Virheet/kysely
B1	54	292	933	851	112	5,4	3,2	15	5,1 %
C2	67	460	1538	1438	181	6,9	3,3	44	9,6 %
Yhteensä	121	752	2471	2289	293	6,2	3,3	59	7,8 %

Hakusanojen kokonaismäärään on laskettu jokainen kyselyissä esiintynyt hakusana, myös fraaseista yksittäin eroteltuna. Kun fraaseiksi merkityt hakusanat yhdistetään yhdeksi hakusanaksi, saadaan

<sup>1</sup> Jokainen hakusana on laskettu vain kerran

hakusanamääräksi 851 (B1-tehtäväkuvaus) ja 1438 (C2-tehtäväkuvaus). Uniikkeja hakusanoja oli B1-tehtäväkuvausta koskevissa kyselyissä 112 ja C2-tehtäväkuvausta koskevissa kyselyissä 181. Hakijoiden muotoilemissa hakusanoissa esiintyi jonkin verran erilaisia kirjoitusvirheitä. Kyselyiden määrään suhteutettuna B1-tehtäväkuvauksessa 5 %:ssa kyselyistä ja C2-tehtäväkuvauksessa lähes 10 %:ssa kyselyistä esiintyi kirjoitusvirheitä (Taulukko 1).

## 7.2 Hakusanayhdenmukaisuus

Hakusanojen ja tehtäväkuvauksen sanojen väliset yhdenmukaisuusluvut kertovat sen, miten paljon hakijat ovat hyödyntäneet tehtäväkuvausta kyselyjä muodostaessaan. Samalla yhdenmukaisuusluku- jen käänteisarvot ilmaisevat, miten paljon hakijat ovat keskimäärin käyttäneet tehtäväkuvauksen ulkopuolisia hakusanoja. Ensin käsitellään yhdenmukaisuusluvut kaikkien kyselyjen osalta, sitten katsotaan lukuja ensimmäisten ja viimeisten kyselyjen osalta.

### 7.2.1 Kaikki kyselyt

Molempien tehtäväkuvauksien kaikki kyselyt huomioon ottaen keskimäärin noin 80-90 % hakijoi- den käyttämistä hakusanoista oli suoraan tehtäväkuvauksista löytyviä hakusanoja. Yhteensä vajaa 80 % hakusanoista oli esiintymätason eli tiukimpien kriteerien mukaan suoraan tehtäväkuvauksista. Ilmaisutasolla tehtäväkuvauksista löytyviksi luokiteltavien hakusanojen osuus kasvoi viisi prosent- tiyksikköä ja käsitetasolla tästä edelleen kolme prosenttiyksikköä. Tehtäväkuvauksien yksittäiset ja yhteenlasketut eri tasojen yhdenmukaisuusluvut kaikkien kyselyjen osalta löytyvät taulukosta 2.

Taulukko 2. Kyselyjen ja tehtäväkuvauksen keskinäinen yhdenmukaisuus esiintymä-, ilmaisu- ja käsitetasolla				
Tehtäväkuvaus	# Kyselyt	Esiintymä- taso	Ilmaisu- taso	Käsite- taso
B1	292	0,75	0,80	0,81
C2	460	0,82	0,86	0,90
<b>Yhteensä</b>	<b>752</b>	<b>0,79</b>	<b>0,84</b>	<b>0,87</b>

Tehtäväkuvauksia erikseen tarkastellessa todetaan, että B1-tehtäväkuvauksen osalta hakusanojen ja tehtäväkuvauksen sanojen vertailussa esiintymätason yhdenmukaisuusluvun keskiarvo oli 0,75. Lu- ku kertoo sen, että tiukimmankin eli merkkijonotason vertailun mukaan keskimäärin 75 % hakijoi-

den valitsemista hakusanoista oli täysin yhdenmukaisia tehtäväkuvauksesta löytyvien sanojen kanssa. C2-tehtäväkuvauksen kohdalla luku oli vieläkin suurempi. Siinä keskimäärin 82 % hakijoiden käyttämistä hakusanoista oli sellaisenaan tehtäväkuvauksissa esiintyviä sanoja. Ilmaisutasolla, jossa yhdenmukaisiksi tehtäväkuvauksen sanojen kanssa laskettiin myös joitain ilmaisullisia eroja sisältävät hakusanat, yhdenmukaisuuslukujen keskiarvot nousivat molemmissa tehtäväkuvauksissa melkein saman verran. B1-tehtäväkuvauksen yhdenmukaisuuskeskiarvo oli 0,80 ja C2-tehtäväkuvauksen 0,86.

Käsitetasolla, jossa hakijan merkitsemät fraasit otettiin huomioon, yhdenmukaisuuskeskiarvot olivat B1-tehtäväkuvauksessa 0,81 ja C2-tehtäväkuvauksessa 0,90. Toisin sanoen B1-tehtäväkuvausta koskien 81 % ja C2-tehtäväkuvausta koskien 90 % kyselyjen hakusanojen edustamista käsitteistä oli laajimpien eli enemmän tulkintavaihtoehtoja sallivien vertailukriteerien mukaan peräisin tehtäväkuvauksesta. Molempien tehtäväkuvauksien ilmaisutason ja käsitetason keskiarvojen välisiä eroja tutkimalla voidaan yleisesti todeta, että B1-tehtäväkuvauksen hakijat olivat ideoineet hakusanoja varsin vähän käsitetasoon kuuluvilla tavoilla kuten synonyymein tai johdoksien, koska ilmaisu- ja käsitetasojen ero oli varsin pieni. C2-tehtäväkuvauksen kohdalla ideointia oli tapahtunut enemmän ilmaisu- ja käsitetasojen eron ollessa kuusi prosenttiyksikköä. Hakusanojen ideointia ja niiden suhdetta tehtäväkuvauksen käsitteisiin tarkastellaan tarkemmin luvussa 7.5.

B1- ja C2-tehtäväkuvauksien yhdenmukaisuuslukujen eri tasojen välisessä vertailussa huomataan, että C2-tehtäväkuvauksen yhdenmukaisuuslukujen keskiarvot ovat 5-9 prosenttiyksikköä suuremmat kuin B1-tehtäväkuvauksen vastaavat luvut. C2-tehtäväkuvauksen hakijat ovat kaikilla kolmella tasolla kyselyjä muodostaessaan hyödyntäneet kyseistä tehtäväkuvausta enemmän kuin mitä B1-tehtäväkuvauksen hakijat ovat B1-tehtäväkuvausta hyödyntäneet. Tehtäväkuvauksien laajuus ja tehtäväkuvauksista tunnistettujen käsitteiden määrä on huomionarvoinen tekijä yhdenmukaisuuslukuja ja tehtäväkuvauksien hyödyntämistä tutkittaessa. Vertailevan tiedon tyyppiä edustava C2-tehtäväkuvaus oli pituudeltaan noin kaksinkertainen B1-tehtäväkuvaukseen verrattuna (ks. luku 5.3), samoin siitä tunnistettujen käsitteiden määrä oli huomattavasti suurempi (ks. Liite 1). Laajempi ja käsitteiltään kattavampi C2-tehtäväkuvaus on suonut hakijoille monipuolisemman lähtökohdan kyselyn muodostukseen.

## 7.2.2 Ensimmäinen ja viimeinen kysely

Kaikkien kyselyiden osalta lasketussa vertailussa useamman kyselyn tehneet hakijat vaikuttavat vertailutuloksiin enemmän kuin vähemmän kyselyitä tehneet. Hakijoiden painotuksen tasoittamiseksi laskettiin lisäksi yhdenmukaisuuskeskiarvot, joissa kyselyistä otettiin mukaan ainoastaan hakijan suorittama ensimmäinen (Taulukko 3) ja viimeinen (Taulukko 4) kysely. B1-tehtäväkuvauksen kohdalla kolme ja C2-tehtäväkuvauksen kohdalla kaksi hakijaa olivat suorittaneet vain yhden kyselyn tehtäväkuvausta kohden. Näiden viiden hakijan osalta tämä yksi kysely otettiin huomioon kahteen kertaan eli sekä ensimmäisen kyselyn että viimeisen kyselyn yhdenmukaisuuskeskiarvojen laskennassa.

Tehtäväkuvaus	# Kyselyt	Esiintymätaso	Ilmaisu-taso	Käsitetaso
B1	54	0,88	0,93	0,93
C2	67	0,88	0,92	0,95
<b>Yhteensä</b>	121	0,88	0,92	0,94

Taulukosta 3 nähdään, että hakijat ovat tehtäväkuvauksesta riippumatta turvautuneet ensimmäisissä kyselyissään lähes samankaltaisesti tehtäväkuvaukseen; tehtäväkuvauksien väliset yhdenmukaisuuskeskiarvot eri tasoilla eivät juuri eroa toisistaan. Yhdenmukaisuusluvut kertovat sen, että tehtäväkuvauksen sisältämät sanat ovat olennainen lähde ensimmäisen kyselyn hakusanavalinnoissa. Käsitetaso määritelmien mukaan käytettyjen hakusanojen ja tehtäväkuvauksen sanojen välinen yhdenmukaisuus oli molemmissa tehtäväkuvauksissa keskimäärin peräti noin 95 %. Esiintymätasolakin yhdenmukaisuusluvut olivat korkeita. Esiintymätason määritelmien mukaan eli merkki merkiltä tehtäväkuvauksen sanoja vastaavia hakusanoja oli keskimäärin 88 % ensimmäisissä kyselyissä käytetyistä hakusanoista.

Viimeisten kyselyjen kohdalla (Taulukko 4) molempien tehtäväkuvauksien yhteenlasketut yhdenmukaisuuskeskiarvot käytettyjen hakusanojen ja tehtäväkuvauksen sanojen välillä jäävät kaikilla kolmella tasolla selvästi alhaisemmiksi kuin ensimmäisten kyselyjen kohdalla. Erot ensimmäisten ja viimeisten kyselyjen yhdenmukaisuusluvuissa tasojen välillä ovat 12-14 prosenttiyksikköä. Viimeisten kyselyjen keskimääräiset yhdenmukaisuudet vaihtelevat esiintymätason 0,75:stä käsitetaso-

0,82:een. Viimeisissä kyselyissä hakijat ovat käyttäneet huomattavasti enemmän tehtäväkuvauksen ulkopuolisiksi luokiteltavia hakusanoja kuin ensimmäisissä kyselyissään.

Taulukko 4. Viimeisten kyselyjen ja tehtäväkuvauksen keskinäinen yhdenmukaisuus esiintymä-, ilmaisu- ja käsitetasolla

Tehtäväkuvaus	# Kyselyt	Esiintymä-taso	Ilmaisu-taso	Käsite-taso
B1	54	0,67	0,71	0,76
C2	67	0,82	0,83	0,87
<b>Yhteensä</b>	121	0,75	0,78	0,82

Hakijat ovat B1-tehtäväkuvauksen kohdalla toimineet viimeisessä kyselyssä hyvinkin innovatiivisesti. B1-tehtäväkuvauksessa esiintymätasolla yhdenmukaisuus käytettyjen hakusanojen ja tehtäväkuvauksen välillä on enää 67 %. Saman tehtäväkuvauksen käsitetason eli enemmän hakusanavariaatioita sallivamman tason yhdenmukaisuusluku (76 %) on jopa vähemmän kuin C2-tehtäväkuvauksen tiukimman eli esiintymätason yhdenmukaisuusluku (82 %).

B1-tehtäväkuvauksen osalta viimeisten kyselyjen yhdenmukaisuusluvut eri tasoilla ovat 5-9 prosenttiyksikköä pienemmät myös kaikkien kyselyiden keskimääräisiin yhdenmukaisuuslukuihin verrattuna. Tämä tarkoittaa, että hakijat ovat hakuprosessin lopussa, viimeisissä kyselyissään, ideoineet vielä lisää omia hakusanoja. C2-tehtäväkuvauksessa hakijat ovat taas ideoineet hakusanoja tasaisesti kaikkien kyselyjen ja viimeisten kyselyjen kesken, lukuerojen ollessa vain 0-3 prosenttiyksikköä, kuten nähdään vertailemalla taulukoita 2 ja 4.

### 7.3 Suosituimmat käsitteet

Taulukoissa 5 ja 6 on kumpaakin tehtäväkuvausta koskevien kyselyjen kymmenen suosituinta tehtäväkuvauksen käsitettä. Määrät laskettiin käsitetason kriteerien mukaan siten, että kyselyissä esiintyviä fraaseja tarkasteltiin erikseen. Hakijoilla oli mahdollisuus käyttää kyselyissään lainausmerkkejä fraasien ilmaisemiseen. Kummankin tehtäväkuvauksen kymmenen suosituinta käsitettä kattoivat vähän yli 75 % kyseistä tehtäväkuvausta koskevistä tehtäväkuvauksen käsitteistä edustavista hakusanoista.

Taulukko 5. B1-tehtäväkuvauksen 10 suosituinta käsitettä			
n=851			
	Käsite	Käyttömäärä	%
1	cybersickness	115	13,5 %
2	virtual	102	12,0 %
3	reality	98	11,5 %
4	vr	76	8,9 %
5	effects	61	7,2 %
6	side	46	5,4 %
7	symptoms	43	5,1 %
8	"virtual reality"	42	4,9 %
9	negative	31	3,6 %
10	"side effects"	25	2,9 %
	Yhteensä	639	75,1 %

Taulukko 6. C2-tehtäväkuvauksen 10 suosituinta käsitettä			
n=1438			
	Käsite	Käyttömäärä	%
1	python	308	21,4 %
2	java	281	19,5 %
3	development	103	7,2 %
4	application	96	6,7 %
5	comparison	90	6,3 %
6	efficiency	62	4,3 %
7	language	43	3,0 %
8	large	40	2,8 %
9	programming	28	1,9 %
10	negative	17	1,2 %
10	positive	17	1,2 %
	Yhteensä	1085	75,5 %

Tehtäväkuvauksien suosituimpia käsitteitä tutkimalla huomataan, että B1-tehtäväkuvauksen suosituimpien käsitteiden käyttömäärät ja -prosentit ovat jakautuneet jonkin verran tasaisemmin kuin C2-tehtäväkuvauksen. B1-tehtäväkuvauksessa hyödynnetty käsite oli *cybersickness*, jonka osuus oli 13,5 % B1-tehtäväkuvauksen käsitteistä. Myös käsitteet *virtual* ja *reality* ylsivät yli 10 % käyttömääriin. Näiden kolmen suosituimman käsitteen ja samoin aina edellistä seuraavan käsitteen käyttömäärien erot olivat tasaisesti enintään 1,8 prosenttiyksikköä. C2-tehtäväkuvauksen osalta hyödynnetyimmät käsitteet olivat erisnimet *Python* ja *Java*. Näiden molempien käyttömäärä (noin 20 %) kaikista käsitteistä oli selvästi enemmän kuin tehtäväkuvauksen kolmanneksi eniten hyödynnetyn käsitteen, *development* (7,2 %).

C2-tehtäväkuvauksen kymmenen hyödynnetyimmän käsitteen joukkoon ei mahtunut yhtään tehtäväkuvauksen käsitteeseen fraasimuodossa viitattua käsitettä. B1-tehtäväkuvauksessa *virtual reality* -käsitettä on fraasimuodossa ilmaistuna käytetty 42 kertaa. Yksittäin, ilman fraasi-ilmaisua on käsitettä *virtual* käytetty 102 kertaa ja käsitettä *reality* 98 kertaa. Nämä käsitteet ovat tosin silloin muutamaa kyselyä lukuun ottamatta esiintyneet yhdessä, samassa kyselyssä. Samalla tavalla *side effects*-käsite esiintyi fraasimuodossa vain 25 kertaa, mutta käsitteet *effects* sekä sitä määrittävät *side* ja *negative* vastaavasti 58, 46 ja 31 kertaa.



## 7.4 Hakusanat tehtäväkuvauksesta ja sen ulkopuolelta

Luvussa 7.2 tutkittiin kolmen tasomäärityksen perusteella, miten paljon hakijat käyttivät keskimäärin kyselyissään tehtäväkuvauksesta löytyviä hakusanoja. Tehtäväkuvauksesta peräisin olevien ja tehtäväkuvauksen ulkopuolisten hakusanojen käyttöä tarkastellaan seuraavaksi tarkemmin hakijoiden suorittamien ensimmäisten kyselyiden (aloituskyselyt) ja kyselyn uudelleen muotoilun (jatkokyselyt) osalta. Kun aiemmin tasomäärityksen perusteella laskettiin erikseen jokaisen kyselyn sisältämien hakusanojen yhdenmukaisuus tehtäväkuvauksen sanoihin, jatkokyselyiden tarkastelussa kyseessä olevaa kyselyä edeltäneessä kyselyssä jo mahdollisesti esiintyviä hakusanoja ei luokitella kahteen kertaan. Tehtäväkuvauksesta ja sen ulkopuolelta oleva luokitus tehtiin ilmaisutason määrittelyjen perusteella. Hakusanan katsottiin olevan tehtäväkuvauksesta, jos se täytti ilmaisutason määrittelymäärän ehdot. Muuten hakusana luokiteltiin tehtäväkuvauksen ulkopuolelta olevaksi.

Tässä luvussa jatkokyselyihin viitattaessa tarkoitetaan vain LISÄYS- ja UUSI-taktiikoita sisältäviä kyselyjä. POISTO-taktiikkaa koskevia kyselyjä tarkastellaan erillään, sillä niissä oli hakusanojen lisäämisen sijaan kyse hakusanojen poistamisesta. MUU MUUTOS –taktiikka käsitteli muunlaisia hakusanoille tehtyjä muutoksia, joten siinä ei ollut mahdollista tehdä kyselyissä käytettyjen hakusanojen jakoa tehtäväkuvauksesta ja tehtäväkuvauksen ulkopuolelta oleviin.

Yhteensä molempien tehtäväkuvauksien aloitus- ja jatkokyselyissä käytetyistä tai edellisiin kyselyihin lisätyistä 1364 hakusanasta noin 81 % oli tehtäväkuvauksesta peräisin olevia ja noin 19 % tehtäväkuvauksen ulkopuolelta (Taulukko 7). Hakijat olivat poistaneet kyselyistään lähes 600 hakusanaa, joista noin 80 % oli tehtäväkuvauksesta ja noin 20 % sen ulkopuolelta eli tehtäväkuvauksesta peräisin olevia hakusanoja oli lisätty lähes samassa suhteessa kuin niitä oli poistettu.

Taulukko 7. Tehtäväkuvauksesta ja sen ulkopuolelta olevien hakusanojen käyttö								
Tehtäväkuvaus		n	Hakusanat tehtäväkuvauksesta		Hakusanat tehtäväkuvauksen ulkopuolelta		Hakusanat yhteensä	
				%		%		%
<b>B1</b>	Aloituskyselyt	54	<b>154</b>	93,9 %	<b>10</b>	6,1 %	<b>164</b>	100 %
	Jatkokyselyt							
	<b>LISÄYS</b>	144	<b>161</b>	71,2 %	<b>65</b>	28,8 %	<b>226</b>	100 %
	<b>UUSI</b>	73	<b>117</b>	72,7 %	<b>44</b>	27,3 %	<b>161</b>	100 %
	<b>Yhteensä</b>	217	<b>278</b>	71,8 %	<b>109</b>	28,2 %	<b>387</b>	100 %
	<b>Yhteensä</b>	271	<b>432</b>	78,4 %	<b>119</b>	21,6 %	<b>551</b>	100 %
	<b>POISTO</b>	114	<b>162</b>	78,6 %	<b>44</b>	21,4 %	<b>206</b>	100 %
<b>C2</b>	Aloituskyselyt	67	<b>210</b>	90,9 %	<b>21</b>	9,1 %	<b>231</b>	100 %
	Jatkokyselyt							
	<b>LISÄYS</b>	259	<b>311</b>	77,8 %	<b>89</b>	22,3 %	<b>400</b>	100 %
	<b>UUSI</b>	80	<b>151</b>	83,0 %	<b>31</b>	17,0 %	<b>182</b>	100 %
	<b>Yhteensä</b>	339	<b>462</b>	79,4 %	<b>120</b>	20,6 %	<b>582</b>	100 %
	<b>Yhteensä</b>	406	<b>672</b>	82,7 %	<b>141</b>	17,3 %	<b>813</b>	100 %
	<b>POISTO</b>	230	<b>317</b>	80,9 %	<b>75</b>	19,1 %	<b>392</b>	100 %
<b>Yhteensä</b>	Aloituskyselyt	121	<b>364</b>	92,2 %	<b>31</b>	7,9 %	<b>395</b>	100 %
	Jatkokyselyt							
	<b>LISÄYS</b>	403	<b>472</b>	75,4 %	<b>154</b>	24,6 %	<b>626</b>	100 %
	<b>UUSI</b>	153	<b>268</b>	78,1 %	<b>75</b>	21,9 %	<b>343</b>	100 %
	<b>Yhteensä</b>	556	<b>740</b>	76,4 %	<b>229</b>	23,6 %	<b>969</b>	100 %
	<b>Yhteensä</b>	677	<b>1104</b>	80,9 %	<b>260</b>	19,1 %	<b>1364</b>	100 %
	<b>POISTO</b>	344	<b>479</b>	80,1 %	<b>119</b>	19,9 %	<b>598</b>	100 %

Tehtäväkuvauksesta B1 tehtiin yhteensä 54 aloituskyselyä (hakijoita 54). Jatkokyselyjä, joissa lisättiin hakutermejä (LISÄYS) tai muodostettiin täysin uusi kysely (UUSI), oli 217. Yhteensä aloitus- ja jatkokyselyissä käytetyistä tai edellisiin kyselyihin lisätyistä hakusanoista 78,4 % oli tehtäväkuvauksesta peräisin olevia ja 21,6 % tehtäväkuvauksen ulkopuolelta. Aloituskyselyihin kuuluvista 164 hakusanasta suurin osa eli 93,9 % oli suoraan tehtäväkuvauksesta olevia hakusanoja. Tehtäväkuvauksen ulkopuolisia oli ainoastaan 6,1 %. Jatkokyselyjä tarkastelemalla huomataan, miten hakijat ovat kyselyjä uudelleen muotoillessaan käyttäneet selvästi enemmän tehtäväkuvauksen ulkopuolisia hakusanoja kuin aloituskyselyissä. Jatkokyselyissä lisätyistä hakusanoista 71,8 % oli suoraan tehtäväkuvauksesta ja 28,2 % tehtäväkuvauksen ulkopuolelta.

Tehtäväkuvauksesta C2 tehtiin 67 aloituskyselyä ja 339 jatkokyselyä. Yhteensä aloitus- ja jatkokyselyissä käytetyistä tai edellisiin kyselyihin lisätyistä hakusanoista lähes 83 % oli tehtäväkuvauk-

sesta ja melkein 18 % tehtäväkuvauksen ulkopuolelta. Aloituskyselyissä hakijat ovat tässäkin tehtäväkuvauksessa pitäytyneet varsin tarkkaan tehtäväkuvauksessa, sillä tehtäväkuvauksesta olevia hakusanoja oli yli 90 % ja tehtäväkuvauksen ulkopuolisia hakusanoja vain noin 9 %. Jatkokyselyissä lisätyistä hakusanoista tai uusien kyselyiden hakusanoista tehtäväkuvauksesta olevien osuus pieneni ollen kuitenkin edelleen lähes 80 %. Tehtäväkuvauksen ulkopuolisiksi luokiteltavia hakusanoja jatkokyselyissä oli hieman yli 20 %.

Tehtäväkuvauksesta ja tehtäväkuvauksen ulkopuolelta olevien hakusanojen jakaumaa voidaan katsoa vielä kyselyn muokkauksen mukaan, jossa muokkaustaktiikat ovat LISÄYS, UUSI ja POISTO. Muokatessaan edellisiä kyselyjä täysin uudelleen eri hakusanoin, hakijat olivat molemmissa tehtäväkuvauksissa taipuvaisia käyttämään enemmän tehtäväkuvauksesta peräisin olevia hakusanoja kuin LISÄYS-taktiikan kohdalla, mutta C2-tehtäväkuvauksen kohdalla ero korostui selvemmin. C2-tehtäväkuvauksessa lisätyistä hakusanoista noin 78 % oli tehtäväkuvauksesta, mutta täysin uusien kyselyiden hakusanoista 83 % oli tehtäväkuvauksesta. Sen sijaan B1-tehtäväkuvauksessa LISÄYS- tai UUSI-taktiikan välillä oli vain reilun prosenttiyksikön ero tehtäväkuvauksesta ja sen ulkopuolelta olevien hakusanojen käyttömäärissä. Lisätyistä hakusanoista 71,2 % oli tehtäväkuvauksesta; täysin uusien kyselyiden hakusanoista vastaava osuus oli 72,7 %.

Eri tehtäväkuvauksien kyselyistä poistettujen hakusanojen suhde tehtäväkuvauksiin oli suunnilleen sama. Hakijat olivat poistaneet B1-tehtäväkuvausta koskevissa kyselyissä yhteensä 206 hakusanaa, joista melkein 80 % oli tehtäväkuvauksesta olevia hakusanoja ja tehtäväkuvauksen ulkopuolisia hieman yli 20 %. C2-tehtäväkuvauksen kohdalla poistettuja hakusanoja oli lähes 400, joista 81 % oli tehtäväkuvauksesta ja 19 % sen ulkopuolelta.

B1-tehtäväkuvauksen kyselyjen uudelleen muotoilussa tehtiin enemmän tehtäväkuvauksesta olevien hakusanojen poistoja kuin mitä tehtiin näiden hakusanojen lisäyksiä. Kokonaisuudessaan kuitenkin sen aloitus- ja jatkokyselyiden yhteenlaskettujen hakusanojen jakauma tehtäväkuvauksesta (78,4 %) ja sen ulkopuolelta (21,6 %) oleviin on lähes sama kuin hakusanojen poistoja koskeva vastaava jakauma (78,6 % ja 21,4 %). C2-tehtäväkuvauksessa hakijat turvautuivat hakusanavalinnoissaan tiiviimmin tehtäväkuvaukseen, ja sen aloitus- ja jatkokyselyiden yhteenlasketuista hakusanoista tehtäväkuvauksesta olevien osuus oli suurempi (82,7 %) kuin poistettujen hakusanojen tehtäväkuvauksesta oleva osuus (80,9 %).

Hakusanojen valintaan ja siten niiden jakautumiseen tehtäväkuvauksesta ja tehtäväkuvauksen ulkopuolelta oleviin saattanut vaikuttaa eri tehtäväkuvaustyyppiä edustavien tehtäväkuvauksen pituus ja käsitteiden määrä, kuten jo luvussa 7.2.1 todettiin. Pidemmän ja käsitteiltään kattavamman C2-tehtäväkuvauksen kohdalla hakijoilla on ollut tarjolla laajempi valikoima tehtäväkuvauksesta olevia hakusanoja, joita he ovat voineet hyödyntää kyselyssään. Hakijat olivatkin C2-tehtäväkuvauksessa turvautuneet jatkokyselyjen hakusanoissaan selvästi enemmän tehtäväkuvaukseen. Jatkokyselyissä C2-tehtäväkuvauksen kohdalla lisätyistä hakusanoista tai täysin uusista hakusanoista tehtäväkuvauksesta peräisin olevia hakusanoja oli 79,4 %, kun taas B1-tehtäväkuvauksen kohdalla vastaava luku oli 71,8 %.

Jatkokyselyissä poistettujen hakusanojen jakautumisessa tehtäväkuvauksesta ja tehtäväkuvauksen ulkopuolelta oleviin ei kuitenkaan ollut suurta eroa tehtäväkuvauksien välillä, sillä molemmissa tehtäväkuvauksissa on poistettu lähes yhtä suuressa suhteessa tehtäväkuvauksesta olevia hakusanoja kuin tehtäväkuvauksen ulkopuolisia hakusanoja. Lisäksi mielenkiintoinen huomio jatkokyselyihin verrattuna on, että aloituskyselyissä C2-tehtäväkuvauksen hakusanoista tehtäväkuvauksesta olevia hakusanoja (90,9 %) oli hiukan vähemmän kuin mitä B1-tehtäväkuvauksen hakusanoista oli tehtäväkuvauksesta olevia (93,9 %).

## **7.5 Tehtäväkuvauksen ulkopuolisten hakusanojen suhde tehtäväkuvaukseen**

Seuraavaksi tarkastellaan aloitus- ja jatkokyselyissä esiintyvien tehtäväkuvauksen ulkopuolisiksi määriteltyjen hakusanojen edustamien käsitteiden suhdetta tehtäväkuvauksen käsitteisiin (Taulukko 8). Kuten edellä, jatkokyselyistä otettiin huomioon sellaiset kyselyt, joissa kyselyä oli uudelleen muotoiltu lisäämällä edelliseen kyselyyn hakusana/hakusanoja (LISÄYS) tai kysely muodostui täysin eri hakusanoista kuin edellinen kysely (UUSI). Hakusanojen edustamien käsitteiden suhteeseen viittaavat lyhenteet taulukossa ovat Syn=Synonyymi, SK=Suppeampi käsite, LK=Laajempi käsite, RK=Rinnakkaiskäsite ja Oma=Oma käsite.

Aloituskyselyissä käytetyistä hakusanoista, jatkokyselyihin lisätyistä hakusanoista ja täysin uudelleen muotoiltujen uusien kyselyjen hakusanoista tehtäväkuvauksen ulkopuolisia oli molemmat teh-

täväkuvaukset yhteenlaskettuna kaiken kaikkiaan 260 hakusanaa. Näiden hakusanojen edustamien käsitteiden suhde tehtäväkuvaukseen jakautui niin, että noin 46 % oli rinnakkaissuhteessa, lähes 26 % omia käsitteitä, noin 13 % laajempia käsitteitä, hiukan vajaa 10 % synonyymejä ja noin 6 % suppeampia käsitteitä. Rinnakkaissuhde tehtäväkuvauksen käsitteisiin oli siten selvästi suurin suhde-tyyppi, johon tehtäväkuvauksien ulkopuolisten hakusanojen edustamat käsitteet kuuluivat.

Tehtäväkuvaus	Suhde Kysely							Hakusanoja					
		Syn	%	SK	%	LK	%	RK	%	Oma	%	Yhteensä	%
B1	Aloituskyselyt	0	0,0 %	0	0,0 %	6	5,0 %	3	2,5 %	1	0,8 %	10	8,4 %
	Jatkokyselyt												
	LISÄYS	2	1,7 %	13	10,9 %	9	7,6 %	28	23,5 %	13	10,9 %	65	54,6 %
	UUSI	1	0,8 %	2	1,7 %	17	14,3 %	16	13,4 %	8	6,7 %	44	37,0 %
	Yhteensä	3	2,5 %	15	12,6 %	26	21,8 %	44	37,0 %	21	17,6 %	109	91,6 %
	Yhteensä	3	2,5 %	15	12,6 %	32	26,9 %	47	39,5 %	22	18,5 %	119	100,0 %
C2	Aloituskyselyt	6	4,3 %	0	0,0 %	0	0,0 %	13	9,2 %	2	1,4 %	21	14,9 %
	Jatkokyselyt												
	LISÄYS	12	8,5 %	0	0,0 %	1	0,7 %	41	29,1 %	35	24,8 %	89	63,1 %
	UUSI	3	2,1 %	1	0,7 %	0	0,0 %	19	13,5 %	8	5,7 %	31	22,0 %
	Yhteensä	15	10,6 %	1	0,7 %	1	0,7 %	60	42,6 %	43	30,5 %	120	85,1 %
	Yhteensä	21	14,9 %	1	0,7 %	1	0,7 %	73	51,8 %	45	31,9 %	141	100,0 %
Yhteensä	Aloituskyselyt	6	2,3 %	0	0,0 %	6	2,3 %	16	6,2 %	3	1,2 %	31	11,9 %
	Jatkokyselyt												
	LISÄYS	14	5,4 %	13	5,0 %	10	3,8 %	69	26,5 %	48	18,5 %	154	59,2 %
	UUSI	4	1,5 %	3	1,2 %	17	6,5 %	35	13,5 %	16	6,2 %	75	28,8 %
	Yhteensä	18	6,9 %	16	6,2 %	27	10,4 %	104	40,0 %	64	24,6 %	229	88,1 %
	Yhteensä	24	9,2 %	16	6,2 %	33	12,7 %	120	46,2 %	67	25,8 %	260	100,0 %

Tehtäväkuvauksia erikseen tarkastellessa havaitaan, että B1-tehtäväkuvauksen ulkopuolisesta 119 hakusanasta tehtäväkuvaukseen rinnakkaissuhteessa olevia käsitteitä oli eniten, lähes 40 %. Käsitehierarkiassa tehtäväkuvauksen käsitteiden yläkäsitteitä eli laajempia käsitteitä edustavia hakusanoja oli noin 27 % ja alakäsitteitä eli suppeampia käsitteitä edustavia hakusanoja melkein 13 %. Alakäsitteiden käyttöä enemmän hakijat ideoivat omia hakusanoja, joita oli vajaa 19 %. Tehtäväkuvaukseen synonyymisuhteessa olevia oli varsin vähän, vain 2,5 %.

C2-tehtäväkuvauksen kyselyissä käytetyt 141 tehtäväkuvauksen ulkopuolista hakusanaa jakaantuvat siten, että suuren enemmistön muodostivat joko tehtäväkuvaukseen rinnakkaissuhteessa olevat

hakusanat (51,8 %) tai hakijoiden ideoimat omat hakusanat (31,9 %). Synonyymeja oli 14,9 % tehtäväkuvauksen ulkopuolisista hakusanoista. Ala- ja yläkäsitteiden käyttö jäi todella vähäiseksi, sillä kumpaankin luokkaa edusti vain 1 hakusana.

Kuten luvussa 7.4 todettiin, aloituskyselyissä käytetyistä hakusanoista vain hyvin pieni osa oli tehtäväkuvauksen ulkopuolelta, B1-tehtäväkuvauksessa ainoastaan noin 6 % ja C2-tehtäväkuvauksessakin vain noin 9 %. B1-tehtäväkuvauksessa nämä olivat suhteessa tehtäväkuvaukseen laajempia hakusanoja (6 kpl), rinnakkaishakusanoja (3 kpl) tai hakijan oma hakusana (1 kpl). Synonyymeja tai suppeampia hakusanoja tehtäväkuvauksen sanastosta ei käytetty lainkaan aloituskyselyissä. C2-tehtäväkuvauksessa 13 hakusanaa oli tehtäväkuvaukseen assosiaatiosuhteessa olevia ja 6 synonyymeja. Omia hakusanoja esiintyi aloituskyselyissä vain 2 kappaletta. Ala- ja yläkäsitteitä ei ollut lainkaan.

B1-tehtäväkuvauksen kohdalla tehtäväkuvauksen ulkopuolelle luokitelluista hakusanoista jatkokyselyiden osalta 37 % oli tehtäväkuvaukseen assosiaatiosuhteessa olevia, 21,8 % laajempia hakusanoja ja 17,6 % hakijoiden itse ideoimia hakusanoja. Samoin kuin B1-tehtäväkuvauksen aloituskyselyissä, sen jatkokyselyissä käytettiin tehtäväkuvauksen ulkopuolisissa hakusanoista vähän tehtäväkuvauksen käsitteistä suppeampia hakusanoja (12,6 %) tai synonyymejä (2,5 %). C2-tehtäväkuvauksen kohdalla jatkokyselyiden hakusanojen edustamista käsitteistä suuri enemmistö oli joko rinnakkaissuhteessa tehtäväkuvauksen käsitteisiin (42,6 %) tai omiksi hakusanoiksi luokiteltavia (30,5 %). Myös synonyymisuhde oli aika yleistä, sillä 10,6 % käsitteistä oli synonyymeja. C2-tehtäväkuvauksen aloituskyselyissä ala- ja yläkäsitteitä ei ollut lainkaan ja jatkokyselyissäkin niiden käyttö oli jäänyt erittäin vähäiseksi.

## **7.6 Kyselyn uudelleen muotoilu**

Kyselyn uudelleen muotoilun tarkastelu koskee jatkokyselyjä, joissa hakijat poistivat hakusanoja, lisäsivät hakusanoja tai tekivät muita hakusanojen esitystapaan liittyviä muutoksia. Muokkaustaktiikat, joita käsitellään, ovat siten POISTO, LISÄYS, MUU MUUTOS ja UUSI (Taulukko 9). Hakija saattoi samassa kyselyssään sekä poistaa että lisätä hakusanoja tai muuttaa niiden esitystapaa (joten muokkaustaktiikoiden yhteenlaskettu summa tehtäväkuvausta kohden on suurempi kuin jatkokyselyjen määrä ja prosentiosuuksien yhteenlasketusta summasta tulee enemmän kuin sata prosent-

tia). Melkein kaikki hakijat jatkoivat tiedonhakua ensimmäisen kyselyn jälkeen ja muotoilivat kyselyjään jollain tavalla uudelleen. Tutkimuksessa oli yhteensä viisi sellaista hakijaa, jotka tekivät vain yhden kyselyn (aloituskysely) kyseistä tehtäväkuvausta kohden.

Taulukko 9. Kyselyn muokkaustaktiikat							
Tehtäväkuvaus	Hakijat	Jatkokyselyjä	POISTO %	LISÄYS %	MUU MUUTOS %	UUSI %	
<b>B1</b>	51	238	114 47,9 %	144 60,5 %	15 6,3 %	73 30,7 %	
<b>C2</b>	65	393	230 58,5 %	259 65,9 %	29 7,4 %	80 20,4 %	
<b>Yhteensä</b>	116	631	344 54,5 %	403 63,9 %	44 7,0 %	153 24,2 %	

Tehtäväkuvauksien yhteisessä tarkastelussa hakijat olivat käyttäneet POISTO- ja/tai LISÄYS-taktiikkaa yli puolessa tekemistään jatkokyselyistä. Reilussa 60 %:ssa jatkokyselyistä oli lisätty hakusanoja edelliseen kyselyyn. Lähes 55 %:ssa kyselyistä oli poistettu hakusana tai hakusanoja. Täysin uusia kyselyjä, joissa edellinen kysely oli muutettu kokonaan, oli noin neljäsosa jatkokyselyistä. MUU MUUTOS-taktiikkaa, eli fraasien muutoksia tai kirjoitusvirheiden korjauksia, oli käytetty 7 %:ssa jatkokyselyistä.

B1-tehtäväkuvauksessa hakijat tekivät ensimmäisen kyselynsä jälkeen yhteensä 238 jatkokyselyä, joissa edellistä kyselyä oli jollain tavalla uudelleen muotoiltu. C2-tehtäväkuvauksessa jatkokyselyjä oli 393. B1- ja C2 tehtäväkuvausta verratessa hakijat sovelsivat eri muokkaustaktiikoita lähes samankaltaisesti jatkokyselymäärän suhteutettuna. Eroja oli lähinnä POISTO- ja UUSI-taktiikoiden käytössä. B1-tehtäväkuvauksessa hakijat olivat käyttäneet POISTO-taktiikkaa vähemmän kuin C2-tehtäväkuvauksessa oli käytetty, mutta UUSI-taktiikan kohdalla tilanne oli päinvastainen. Kummankin taktiikan käytön määrän väliset erot olivat noin 10 %. Sen sijaan LISÄYS-taktiikka koski molemmissa tehtäväkuvauksissa reilua 60 % ja MUU MUUTOS-taktiikka noin 7 % jatkokyselyistä.

## 7.7 Taustamuuttujien vaikutus kyselyn muodostukseen

Tehtäväkuvaukseen liittyvän haun suorittamisen helppouden ja aiemman aihetuntemuksen vaikutusta kyselyn muodostukseen tutkittiin korrelaation avulla. Hakijat olivat haun suoritettuaan vastanneet kyselylomakkeen kysymyksiin ja antaneet vastauksensa viisiportaisella asteikolla, jossa 1=ei yh-

tään, 3=jonkin verran ja 5=erittäin. Kyselylomakkeen kysymys 2 koski tehtäväkuvaukseen liittyvän haun toteuttamisen helppouden arvioimista ja kysymys 7 tehtäväkuvaukseen liittyvää aihetuntemusta.

Korrelaation avulla tutkittiin tehtäväkuvaukseen liittyvän haun suorittamisen helppouden ja aihetuntemuksen yhteyttä tehtyjen kyselyjen ja käytettyjen hakusanojen määrään sekä eri tasojen yhdenmukaisuuslukuihin. Muuttujien keskinäisessä vertailussa korrelaatiokerroin voi saada arvot välillä  $[1, -1]$ . Suuri positiivinen kerroin ilmaisee muuttujien välistä positiivista yhteyttä eli toisen muuttujan arvon kasvaessa myös toinen muuttujan arvo kasvaa. Negatiivinen kerroin lähellä  $-1$ :tä taas kertoo negatiivisesta yhteydestä eli toisen muuttujan arvon kasvaessa toinen vastaavasti pienenee. Kun korrelaatiokerroin on lähellä arvoa 0, muuttujien välillä ei ole lineaarista yhteyttä.

Tehtäväkuvauksen B1 valinneet hakijat pitivät haun suorittamista melko helppona, mutta tehtäväkuvauksen aihepiiriä suhteellisen vieraana. B1-tehtäväkuvauksen kysymysten 2 ja 7 vastausten keskiarvot olivat 3,43 ja 1,89. C2-tehtäväkuvausta koskevien kummankin vastauksen keskiarvot olivat hyvin lähellä arvoa kolme eli haun suorittaminen oli hakijoille melko helppoa ja tehtäväkuvauksen aihe entuudestaan tuttu. Vaikka C2-tehtäväkuvauksen aihe oli hakijoille jonkin verran tutumpi (keskiarvo 2,79), he pitivät kyseisen tehtäväkuvauksen mukaista hakua keskiarvon (3,18) mukaan hieman vaikeampana suorittaa, kuin B1-tehtäväkuvauksen valinneet olivat pitäneet B1-tehtäväkuvauseen liittyvää hakua.

Taulukon 10 korrelaatiomatriisista nähdään, että kysymysten 2 (tehtäväkuvaukseen liittyvän haun suorittamisen koettu helppous) ja 7 (tehtäväkuvaukseen liittyvä aihetuntemus) vaikutus kyselyn muodostukseen on melko vähäistä tarkasteltaessa tehtäväkuvauksia sekä erikseen että yhteenlasketuna korrelaatiokertoimien jäädessä lähelle arvoa 0. Koska korrelaatiokertoimet ja niiden erot ovat erittäin pieniä, on muistettava, että nämä tulkinnat ovat pääosin huomioita muuttujien välisestä keskinäisestä tai heikosta yhteydestä, eikä tarkoituksena ole tehdä selkeitä osoituksia yhteyden olemassaolosta.

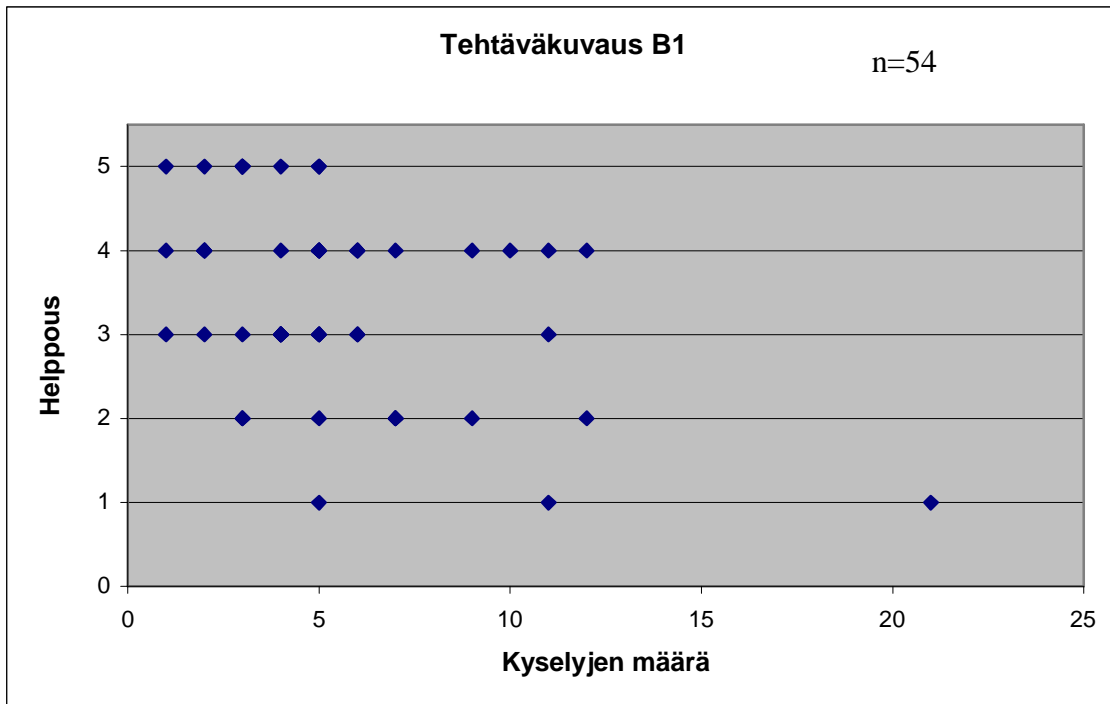


Taulukko 10. Korrelaatiomatriisi

Tehtäväkuvaus	Kysymys	Kyselyjen määrä	Hakusanojen määrä	Esiintymätaso	Ilmaisutaso	Käsitetaso	Kysymys 2 (koettu helppous)	Kysymys 7 (aihetuntemus)
B1	2	-0,42	-0,05	0,04	0,05	0,21	1	
	7	-0,14	0,19	0,06	0,12	0,15	0,14	1
C2	2	-0,21	-0,10	-0,04	-0,09	-0,12	1	
	7	-0,08	-0,04	-0,05	-0,20	-0,07	0,41	1
Yhteensä	2	-0,33	-0,09	-0,01	-0,03	0,02	1	
	7	-0,01	0,07	0,06	0,00	0,11	0,21	1

Korrelaatiokertoimien ja muuttujien yhteyden matemaattisessa arvioinnissa esimerkiksi arvon  $\pm 0.9$  on esitetty osoittavan voimakasta ja arvon  $\pm 0.5$  keskinkertaista lineaarista yhteyttä. Kuitenkin monien käyttäytymistieteissä tutkittavien muuttujien väliset yhteydet ovat paljon heikompia ja esimerkiksi  $\pm 0.5$ :n suuruisia korrelaatiokertoimia voidaan pitää osoituksena voimakkaasta muuttujien välisestä yhteydestä. (Nummenmaa 2004, 278.)

Siten B1-tehtäväkuvauksen kohdalla haun suorittamisen helppouden (kysymys 2) ja kyselyjen määrän yhteyttä voitaneen pitää keskinkertaisena. Korrelaatiokertoimen  $-0,42$  mukainen keskinkertainen negatiivinen yhteys on havaittavissa kuviossa 4. Sirontakuviokuva haun suorittamisen helppouden yhteyden voimakkuutta kyselyjen määrään B1-tehtäväkuvauksen osalta. Sirontakuviossa jokainen piste kuvaa yksittäisen muuttujan tulosta x- ja y-akseleilla. Kuviossa 4, samoin kuin myöhemmin esitettyssä kuviossa 5, osa pisteistä kuvaa useampaa havaintoa näiden muuttujien arvojen ollessa täysin samoja. Kuvion 4 pisteiden sijoittuminen melko hajalleen koordinaatistoon osoittaa, että selvää lineaarista yhteyttä muuttujien välillä ei ole. Kuten todettiin, korrelaatiokertoimen osoittama jonkinasteinen negatiivinen yhteys on kuitenkin havaittavissa kuviosta: mitä helpommaksi hakijat kokivat tehtäväkuvaukseen liittyvän haun toteuttamisen (kysymyksen 2 vastauksen arvo kasvaa), sitä vähemmän he tekivät kyselyjä tehtäväkuvausta kohden.

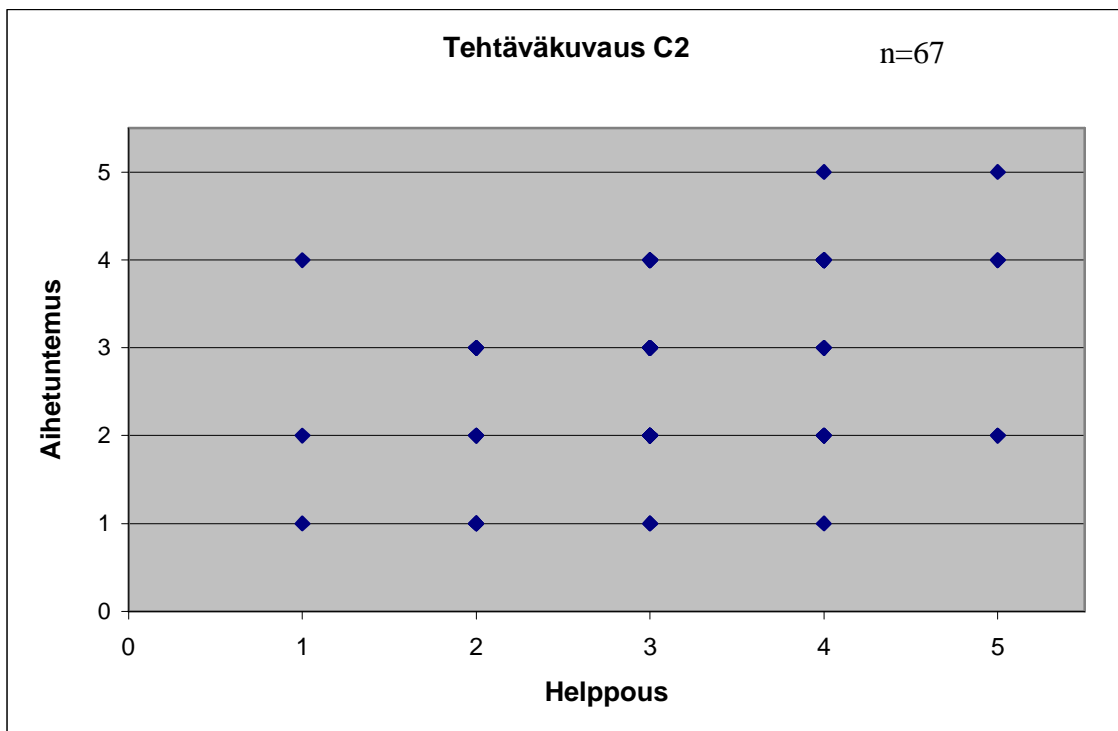


Kuvio 4. Haun suorittamisen helppouden ja kyselyjen määrän yhteys B1-tehtäväkuvauksen osalta.

B1- ja C2-tehtäväkuvausta tarkasteltaessa (Taulukko 10) voidaan molempien kysymysten todeta saavan noin  $\pm 0.2$  suuruisen korrelaatiokertoimen arvon muutaman muuttujan välisessä vertailussa. B1-tehtäväkuvauksen koettu helppous näyttäisi vaikuttavan kohtalaisesti yhdenmukaisuuteen käsitetasolla (0,21). Yhdenmukaisuus oli sitä suurempi mitä helpompina hakijat tehtäväkuvauksen koivat. B1-tehtäväkuvauksen aihepiirin tunnettuus oli myös heikossa positiivisessa yhteydessä hakusanojen määrään (0,19). C2-tehtäväkuvausta koskeva aihetuntemus näytti olevan heikossa negatiivisessa yhteydessä ilmaisutason yhdenmukaisuuteen (-0,20). Sen mukaan hakijoiden aihetuntemuksen kasvaessa ilmaisutason yhdenmukaisuus laski.

Tehtäväkuvauksien välisessä korrelaatiokertoimien vertailussa huomataan, että B1-tehtäväkuvauksen kertoimet ovat enimmäkseen positiivisia ja C2-tehtäväkuvauksen kertoimet yhtä lukuun ottamatta negatiivisia. Esimerkiksi B1-tehtäväkuvauksen helppous ja suurempi aihetuntemus vaikuttivat yhdenmukaisuuslukujen kasvuun, kun taas C2-tehtäväkuvauksen osalta koetun helppouden ja aihetuntemuksen kasvaessa yhdenmukaisuus kyselyjen hakusanojen ja niiden edustamien käsitteiden ja tehtäväkuvauksen sanojen ja käsitteiden välillä väheni.

Tehtäväkuvauksen koetun helppouden ja aihetuntemuksen keskinäisessä korrelaatiokertoimen vertailussa C2-tehtäväkuvauksen kerroin 0,41 ilmaisee keskinkertaista positiivista yhteyttä. Sirontaku-  
vio (Kuvio 5) osoittaa C2-tehtäväkuvauksen koetun helppouden ja aihetuntemuksen yhteyden: mitä helpompina hakijat pitivät C2-tehtäväkuvausta, sitä enemmän heillä oli siitä aikaisempaa tietoutta. B1-tehtäväkuvauksen osalta vastaavaa lineaarista yhteyttä ei ollut havaittavissa.



Kuvio 5. Aihetuntemuksen ja haun suorittamisen helppouden yhteys C2-tehtäväkuvauksen osalta.

## 8 YHTEENVETO

Tämän tutkimuksen tavoitteena oli selvittää, millaisia kyselyjä hakijat muodostavat suhteessa annettuun tehtäväkuvaukseen ja millainen on heidän kyselyn uudelleen muotoiluun liittyvä hakukäyttämisenä vuorovaikutteisessa tiedonhakutilanteessa. Innovatiivinen hakukäyttämisen eli omien hakusanojen käyttäminen kyselyissä viittaisi enemmän todelliseen tiedonhakutilanteeseen, kun taas annettuun tehtäväkuvaukseen pohjaava kyselyjen mekaaninen muodostus vastaisi pikemminkin perinteistä, laboratorio-olosuhteiden mukaista testikäyttämistä. Tutkimustilanteen realistisuutta saattaa heikentää se, että hakijoille oli annettu aika, jonka kuluessa tiedonhaku tuli suorittaa. Oikeassa elämässä tiedonhaku on usein kiireistä ja tietyllä tavalla ajallisesti rajattua, mutta testitilanteen kaltaista aikarajaa todelliset käyttäjät harvoin kohtaavat.

Todelliset käyttäjät näyttävät web-tiedonhaku koskevien tutkimusten perusteella toimivan hyvin eri tavalla kuin tiedonhakijat perinteisten tiedonhakujärjestelmien tutkimuksissa. Tutkimustulokset osoittavat, että hakijat käyttävät vain vähän aikaa ja vaivaa hakiessaan tietoa yksittäisellä hakukoneella web-ympäristössä; he käyttävät vähän hakusanoja, tutkivat tulosjoukkoa suppeasti ja muotoilevat kyselyjä harvoin uudelleen. Toisaalta web-ympäristö myös tarjoaa hyvin erilaista ja eritasoista tietoa toisin kuin tieteelliset tietokannat, ja pelkästään lokitiedostoja analysoimalla on vaikea arvioida, millainen tiedontarve käyttäjällä on tilanteen taustalla ollut. Tällaisten tutkimusten suurin rajoite onkin niiden suppea näkökulma käyttäjien vuorovaikutuksesta web-ympäristössä. Hakija itse, hänen aikomuksensa, strategiansa ja motivaationsa jäävät yleensä selvittämättä. Lisäksi tutkimusten tulosten välinen vertailu on vaikeaa, jos hakusessioiden ja yksittäisten hakijoiden määrittäminen lokitiedostoista ei ole yhtenäisen käytännön mukaista. (Ingwersen & Järvelin 2005, 225-227.)

Realistisiin tiedonhakutilanteisiin, esimerkiksi web-tiedonhakuun, keskittyneiden käyttäjätutkimusten keskinäinen vertailu on siis melko hankalaa. Samalla tavoin vuorovaikutteisten tiedonhakututkimusten, jossa on pyritty sekä säilyttämään kontrolloitavuus että saamaan mukaan todellinen käyttäjä, vertailua tällaisiin käyttäjäsuuntautuneisiin tutkimuksiin on pidettävä vain suuntaa-antavana.

Borlundin (2003) esittämän vuorovaikutteisen tiedonhaun evaluointimallin mukaan toimiva, perusteltu ja realistinen tutkimusasetelma saavutetaan, kun todellisia käyttäjiä edustaville tiedonhakijoille

tarjotaan mahdollisuus henkilökohtaiseen tulkintaan tiedontarpeesta sekä käytetään moniulotteista ja dynaamista relevanssiarviointia. Tutkimusasetelman saavuttamiseen liittyy malliin kuuluva simuloitu tehtävä, jota sovellettiin tämän tutkielman tutkimusaineiston muodostavassa INEX-hankkeen vuorovaikutteisen tiedonhaun tutkimuslinjassa. Simuloidun tehtävän avulla hakijalle pyritään luomaan kokonaisvaltaisempi kuva hakutilanteesta ja siten suomaan mahdollisuus suorittaa tiedonhaku realistisen tiedonhakutilanteen mukaisesti.

Web-tiedonhaun tutkimukset ovat osoittaneet, että todellisessa tiedonhakutilanteessa web-kyselyt ovat rakenteeltaan melko yksinkertaisia ja ylipäänsä lyhyitä. Jansenin ym. (1998) raportoimassa tutkimuksessa hakusanoja käytettiin keskimäärin vain 2,21 per kysely. Luku on huomattavasti pienempi kuin Jansenin ym. esittämä joidenkin perinteisen tiedonhakujärjestelmien käyttötutkimusten mukainen keskimääräinen hakusanamäärä. Niissä kyselykohtaiset hakusanamäärät vaihtelivat 7-15 hakusanan välillä. (Jansen ym. 1998, 5-17.)

Tässä tutkielmassa keskimääräinen hakusanamäärä oli 3,3 hakusanaa kyselyä kohden, mikä on hyvin lähellä edellä mainittua web-tiedonhaun keskimääräistä hakusanamäärää. Toisaalta kyselyn uudelleen muotoilun osalta hakukäyttäytyminen oli päinvastaista. Jansenin ym. (1998) tutkimuksen mukaan kyselyn uudelleen muotoilu oli vähäistä ja suurin osa hakijoista suoritti vain yhden kyselyn. INEX-tutkimuslinjaan osallistuneista käyttäjistä (yhteensä 121) lähes kaikki muokkasivat kyselyään uudelleen, ja yhden kyselyn suorittaneita hakijoita oli vain viisi. POISTO- ja LISÄYS-taktiikat olivat käytetyimmät muokkaustaktiikat, sillä molempia taktiikoita oli sovellettu selvästi yli puolessa jatkokyselyistä. Tulosten perusteella voidaan olettaa, että hakijat ovat yleensä poistaneet kyselystä yhden tai useamman hakusanan korvaten sen yhdellä tai useammalla hakusanalla.

Hakijoiden käyttämien hakusanojen ja tehtäväkuvauksen sanojen keskimääräisiä yhdenmukaisuuksia tarkastelemalla huomattiin, että käyttäjät olivat pitäytyneet varsin läheisesti tehtäväkuvauksessa. Hakusanoista suhteellisen vähän oli tehtäväkuvauksen ulkopuolisiksi luokiteltavia. Molemmissa tehtäväkuvauksissa kaikkien kyselyjen osalta kyselyissä käytettyjen hakusanojen ja tehtäväkuvauksen sanojen yhdenmukaisuus oli vähintään 80 % ilmaisutason määritelmien mukaan. Käsitetasolla hakusanojen edustamien käsitteiden yhdenmukaisuus tehtäväkuvauksen käsitteisiin oli melkein 90 %. Käyttäjät olivat odotetusti turvautuneet tehtäväkuvaukseen enemmän hakusessiodien alussa ja lopussa ryhtyneet käyttämään vaihtoehtoisia, tehtäväkuvauksen ulkopuolisia hakusanoja. Ensimmäisten kyselyjen hakusanoista keskimäärin 88-94 % hakusanoista löytyi

mäisten kyselyjen hakusanoista keskimäärin 88-94 % hakusanoista löytyi tehtäväkuvauksesta, kun viimeisten kyselyjen hakusanojen kohdalla suhde tehtäväkuvauksen sanoihin oli 75-82 %, mikä on kuitenkin edelleen varsin korkea.

Myös kyselyissä käytettyjen suosituimpien käsitteiden tarkastelu osoitti, että tehtäväkuvauksen sisältö vaikutti vahvasti hakijoiden käsitteiden valintaan. Suosituimpien kymmenen käsitteen ilmaiset, jotka olivat molemmissa tehtäväkuvauksissa peräisin tehtäväkuvauksesta, kattoivat kaiken kaikkiaan kummassakin tehtäväkuvauksessa käytetyistä hakusanoista yli 75 %.

Yhdenmukaisuuslukujen tehtäväkuvauksien välisessä vertailussa havaittiin, että hakijat pitäytyivät annetun tehtäväkuvauksen sanastossa enemmän C2-tehtäväkuvauksessa kuin B1-tehtäväkuvauksessa. C2-tehtäväkuvauksessa keskimäärin vain 10 % hakusanoista oli täysin hakijoiden itsensä ideoimia hakusanoja. Nämä hakusanat eivät olleet minkään kolmen vertailutason perusteella tehtäväkuvauksesta löytyviä. B1-tehtäväkuvauksessa vastaava osuus oli 19 %. B1-tehtäväkuvauksessa, joka oli tyypiltään taustatietoa koskeva tehtäväkuvaus, hakusanoja siis ideoitiin enemmän kuin vertailevaa tietoa koskevassa C2-tehtäväkuvauksessa.

Tehtäväkuvauksien pituudella on saattanut olla vaikutusta tehtäväkuvauksesta olevien hakusanojen käyttöön ja tehtäväkuvauksen ulkopuolisten hakusanojen ideointiin. C2-tehtäväkuvaus oli pituudeltaan noin kaksinkertainen B1-tehtäväkuvaukseen verrattuna. Samoin C2-tehtäväkuvauksesta tunnistettiin enemmän yksittäisiä käsitteitä (yli 50 käsitettä) kuin B1-tehtäväkuvauksesta (vajaan 40 käsitettä). Siten C2-tehtäväkuvauksen suorittaneilla hakijoilla on ollut hakusanojen valinnassa tarjolla laajempi valikoima ja enemmän tulkinnan mahdollisuuksia. Mielenkiintoista on lisäksi se, että molempien tehtäväkuvauksien hakijat arvioivat valitsemiinsa tehtäväkuvauksiin liittyvien hakujen olevan suunnilleen yhtä helppoja, mutta C2-tehtäväkuvauksen valinneet hakijat olivat kokeneet kyseisen aiheen itselleen tutummaksi. Toisaalta taustamuuttujien vaikutusta tutkittaessa ei koetulla helppoudella tai aihetuntemuksella havaittu olevan kovin suurta vaikutusta eri tasojen hakusanayhdenmukaisuuksiin; ainoastaan muutama heikko yhteys muuttujien välillä havaittiin. C2-tehtäväkuvauksen osalta todettiin negatiivinen yhteys aihetuntemuksen ja ilmaisutason yhdenmukaisuuden välillä eli aihetuntemuksen kasvaessa ilmaisutason yhdenmukaisuus laski. B1-tehtäväkuvauksessa taas yhdenmukaisuus käsitetasolla oli sitä suurempi mitä helpommaksi hakijat tehtäväkuvaukseen liittyvän haun kokivat.

Tehtäväkuvauksen ulkopuolisiksi luokiteltuja hakusanoja oli käytetty kaikkiaan siis suhteellisen vähän. Kuitenkin kun hakijat olivat tehtäväkuvausta tulkitessaan käyttäneet tehtäväkuvauksen ulkopuolisiksi luokiteltuja hakusanoja, he olivat toiseksi eniten ideoineet täysin omia hakusanoja. Omat hakusanat olivat niitä, joiden edustamien käsitteiden ei katsottu olevan synonyymi-, hierarkia- tai rinnakkaisuhteessa tehtäväkuvauksen käsitteisiin. Tällaisiksi luokiteltiin reilu 25 % tehtäväkuvauksen ulkopuolisista hakusanoista. Omia hakusanoja enemmän hakijat olivat käyttäneet tehtäväkuvauksen käsitteisiin rinnakkaisuhteessa olevia hakusanoja (46 %). Synonyymejä sekä ylä- ja alakäsitteitä tehtäväkuvauksen käsitteistä oli käytetty kaikkein vähiten.

Tehtäväkuvauksen ulkopuolisiksi luokitelluiden hakusanojen käytössä oli eroja tehtäväkuvauksien kesken. B1-tehtäväkuvauksessa hakijat olivat käyttäneet rinnakkaiskäsitteiden jälkeen toiseksi eniten laajempia käsitteitä, joita ei C2-tehtäväkuvausta koskien käytetty juuri lainkaan. Vakkarin (2000) tutkimus osoitti hakuprosessin edetessä lisääntyneen aihetuntemuksen lisäävän erityisesti synonyymien, suppeampien käsitteiden ja rinnakkaiskäsitteiden käyttöä. Laajempien käsitteiden käyttö kyselyissä väheni selvästi aihetietämyksen kasvaessa. Tutkimustulokset ovat samansuuntaisia tämän tutkielman tulosten kanssa, sillä B1-tehtäväkuvauksen hakijat, jotka käyttivät enemmän laajempia käsitteitä, vastasivat tehtäväkuvauksen aihepiirin olevan heille vieraampi kuin mitä C2-tehtäväkuvauksen hakijat olivat vastanneet. B1-tehtäväkuvaukseen liittyvän haun suorittamisen helppouden ja tehtäväkuvauksesta laadittujen kyselyjen määrän välillä havaittiin lisäksi keskinertainen käänteinen yhteys. Hakijat tekivät sitä vähemmän kyselyjä mitä helpommaksi he kokivat haun suorittamisen.

Hakusanojen ja tehtäväkuvauksen sanojen yhdenmukaisuutta vertaillen on muistettava, että yhdenmukaisuusluvut eivät kerro mitään haun onnistumisesta. Sitä, onko kyselyllä onnistuttu löytämään relevanttia tietoa tai miten viimeistä kyselyä edeltäneet kyselyt ovat onnistuneet, täytyisi tutkia erikseen. Jos hakija on muissa kyselyissään hyödyntänyt tehtäväkuvauksen sanastoa, mutta ei ole näitä hakusanoja käyttämällä löytänyt mielestään relevanttia tietoa, on hän voinut suorittaa vielä yhden kyselyn täysin erilaisten, itse ideoimien hakusanojensa pohjalta. Mielenkiintoista olisi tutkia, ovatko käyttäjät tehtäväkuvaukseen turvautuessaan löytäneet mielestään tiedontarvetta (tulkinta tehtäväkuvauksesta) vastaavaa tietoa ja minkä tasoisesti taas vähemmän tehtäväkuvauksen kanssa yhdenmukaisten kyselyjen tulosten relevanssia on arvioitu. Kyselyn muodostuksen analysointia voisi

myös jatkaa tutkimalla tarkemmin jatkokyselyjen muodostusta ja sitä, miten hakijoiden tarkastelemat dokumentit ovat vaikuttaneet hakusanojen valintaan.



## LÄHTEET

Baeza-Yates, R. & Ribeiro-Neto, B. 1999. Modern information retrieval. New York: ACM Press.

Bates, M.J. 1979. Information search tactics. *Journal of the American Society for Information Science* 30(4), 205-214.

Belkin, N.J., Cool, C., Stein, A. & Thiel, U. 1994. Cases, scripts, and information-seeking strategies: on the design of interactive information retrieval. Luettu 26.10.2005. Saatavilla: <http://www.scils.rutgers.edu/~belkin/articles/eswa.pdf>

Borlund, P. 2000. Evaluation of interactive information retrieval systems. Åbo : Åbo Akademi University Press.

Borlund, P. 2003. The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research* 3 (8). Luettu 15.10.2005. Saatavilla: <http://informationr.net/ir/8-3/paper152.html>

Borlund, P. & Ingwersen, P. 1998. Measures of relative relevance and ranked half-life: performance indicators for interactive IR. Teoksessa: Croft, B.W., Moffat, A., van Rijsbergen, C.J., Wilkinson, R. & Zobel, J. Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval. ACM Press/York Press, 324-331.

Cosijn, E. & Ingwersen, P. 2000. Dimensions of relevance. *Information Processing and Management* 36(4), 533-550.

Downing, D.A., Covington, M.A. & Covington, M.M. 2000. Dictionary of computer and internet terms. New York: Barron's Educational Series.

Efthimiadis, E.N. 1996. Query expansion. *Annual Review of Information Systems and Technology* 31, 121-187.

Ellis, D. 1996. Progress and problems in information retrieval. London: Library Association.

Fidel, R. 1985. Individual variability in online searching behaviour. Teoksessa: ASIS '85: Proceedings of the American Society for Information Science 48<sup>th</sup> annual meeting (22). Toim. Parkhurst, C.A. 69-72.

Fidel, R. 1991. Searchers' selection of search keys: III. Searching styles. Journal of the American Society for Information Science 42(7), 515-527.

Fuhr, N. & Großjohann, K. 2001. XIRQL: A Query language for information retrieval in XML documents. Proceedings of SIGIR '01 September 9-12, 2001, New Orleans, Louisiana, USA. New York, NY: ACM, 172-180.

Griffiths, J. R., Hartley, R. J. & Willson, J. P. 2002. An improved method of studying user-system interaction by combining transaction log analysis and protocol analysis. Information Research, 7(4).  
Luettu 18.3.2006. Saatavilla: <http://informationr.net/ir/7-4/paper139.html#pet93>

Harter, S. & Hert, C. 1997. Evaluation of information retrieval systems: approaches, issues, and methods. Annual Review of information Science and Technology 32, 3-94.

Hersh, W. & Over, P. 2000. SIGIR Workshop on Interactive Retrieval at TREC and Beyond. SIGIR Forum 34(1). Luettu 7.12.2005. Saatavilla:  
[http://www.sigir.org/forum/S2000/Interactive\\_report.pdf](http://www.sigir.org/forum/S2000/Interactive_report.pdf)

Hert, C. 1997. Understanding information retrieval interactions: theoretical and practical implications. London: Ablex.Publishing Corporation.

Iivonen, M. 1995. Hakulausekkeiden muotoilun yhdenmukaisuus onlineviitehaussa. Acta Universitatis Tamperensis ser A vol. 443. Tampere: Tampereen yliopisto.

INEX 2004. Initiative for the evaluation of XML Retrieval. Luettu 24.10.2005. Saatavilla:  
<http://INEX.is.informatik.uni-duisburg.de:2004/>

Ingwersen, P. 1992. Information retrieval interaction. London: Taylor Graham.

Ingwersen, P. & Järvelin, K. 2005. The turn: integration of information seeking and retrieval in context. Dordrecht: Springer.

Jaakohuhta, H. 2003. IT-ensyklopedia. Helsinki: IT Press.

Jansen, B.J., Spink, A., Bateman J. & Saracevic, T. 1998. Real life information retrieval: A study of user queries on the web. SIGIR Forum 32(1), 5-17.

Järvelin, K. 1993. Merkkijonot, sanat, termit ja käsitteet informaation haussa. Kirjastotiede ja informatiikka 12 (4): 199-128.

Järvelin, K. 1995. Tekstitiedonhaku tietokannoista: johdatus periaatteisiin ja menetelmiin. Espoo: Suomen atk-kustannus Oy.

Järvelin, K. & Kekäläinen, J. 2002. Cumulated gain-based evaluation of IR techniques. ACM Transactions on Information Systems 20(4), 422-446.

Järvelin, K. & Sormunen, E. 1999. Dokumentit kateissa? Tiedon tallennus ja haku avuksi. Teoksessa Ilkka Mäkinen (toim.), Tiedon tie: johdatus informaatiotutkimukseen. Helsinki: BTJ Kirjastopalvelu Oy, 110-143.

Karlsson, F. 2004. Yleinen kielitiede. Helsinki: Yliopistopaino.

Kazai, G., Lalmas, M. & Vries, A. 2004. The Overlap Problem in Content-Oriented XML Retrieval Evaluation. 72-79.

Klein, M. 2001. XML, RDF and Relatives. IEEE Intelligent systems 16(2), 26-28.

Nummenmaa, L. 2004. Käyttäytymistieteiden tilastolliset menetelmät. Helsinki: Tammi.

Pharo, N. & Nordlie, R. 2005. Context matters: an analysis of assessments of XML documents. Proceedings of CoLIS5, 238-248.

Robertson, S.E. 1997. Overview of the Okapi projects. Journal of Documentation 53(1), 3-7.

Robertson, S.E. & Hancock-Beaulieu, M.M. 1992. On the evaluation of IR systems. Information Processing and Management 28(4), 457-466.

Ruthven, I., Lalmas, M. & van Rijsbergen, C. J. 2003. Incorporating user search behaviour into relevance feedback. Journal of the American Society for Information Science and Technology, 54 (6), 528-548.

Saracevic, T. 1984. Measuring the degree of agreement between searchers. Teoksessa: ASIS '84: Proceedings of the American Society for Information Science 47<sup>th</sup> annual meeting (21), 227-230.

Saracevic, T. 1996. Relevance reconsidered. '96. Teoksessa P. Ingwersen & N. Pors (toim), Information Science : Intergration in perspective. Copenhagen: Royal School of Librarianship, 201-218.

Schamber, L., Eisenberg, M.B. & Nilan, M.S. 1990. A re-examination of relevance: toward a dynamic, situational definition. Information Processing and Management 26(6), 755-776.

Sormunen, E. 1989. An analysis of online searching knowledge for intermediary systems. VTT, tutkimuksia 630. Espoo: VTT.

Spink, A., Wolfram, D., Jansen, B.J. & Saracevic, T. 2001. Searching the web: The public and their queries. Journal of the American Society for Information Science and Technology 52(3), 226-234.

Suomela, S. & Kekäläinen, J. 2005. Ontology as a search-tool: A study of real users' query. Formulation with and without conceptual support. Proceedings of ECIR 2005, 315-329.

Sutcliffe, A.G., Ennis, M., Watkinson, S.J. 2000. Empirical studies of end-user information searching. *Journal of the American Society for Information Science* 51(13), 1211-1231.

Tombros, A., Larsen, B. & Malik, S. 2004. The interactive track at INEX 2004. Luettu 24.10.2005. Saatavilla:  
<http://INEX.is.informatik.uni-duisburg.de:2004/pdf/INEX2004PreProceedings.pdf>

Tombros, A., Ruthven, I. & Jose, J.M. 2005. How users assess web pages for information seeking. *Journal of the American Society for Information Science and Technology* 56(4), 327-344.

Text REtrieval Conference (TREC). 2000. Luettu 8.11.2005. Saatavilla: <http://TREC.nist.gov/>

Vakkari, P. 2000. Cognition and changes of search terms and tactics during task performance: A longitudinal case study. *Proceedings of the RIAO 2000 conference*, 894-907. Paris: C.I.D.

W3C. 2004. Extensible Markup Language (XML) 1.0 (Third Edition). Luettu 5.5.2006.  
<http://www.w3.org/TR/2004/REC-xml-20040204/>

White, R.W., Ruthven, I. & Jose, J.M. 2002. Finding relevant documents using top ranking sentences: an evaluation of two alternative schemes. *Proceedings of the 25th Annual international ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '02 August 11-15, 2002, Tampere, Finland. New York, NY: ACM Press, 57-64.*

# LIITTEET

## LIITE 1: Tehtäväkuvauksen käsiteanalyysi

### 1.1 Tehtäväkuvaus B1

amount	to be
amount of user / user	to discuss
application	to get
article / large article	to know
cybersickness	to need
effect	to occur
interested	to use
large	to want
negative	to write
negative side effect	treatments
<b>not interested</b>	unless
reality	virtual
side effect	virtual reality / vr
side	virtual reality application / vr application
situation	vr situation
symptoms / symptoms of	what
therapeutic	where
therapeutic treatments	who
they	you
to associate	

#### **Hierarkkinen suhde:**

- application
  - virtual reality application / vr application
- side effect
  - negative side effect
- situation
  - vr situation
- treatments
  - therapeutic treatments
- effect
  - side effect
  - negative effect

#### **Assosiatiosuhde:**

- cybersickness – symptoms
- cybersickness – virtual reality / vr
- negative side effect – cybersickness
- symptoms -- treatments
- virtual reality / vr – virtual reality application / vr application
- virtual reality / vr – vr situation
- vr situation – cybersickness

## 1.2 Tehtäväkuvaus C2

application / large application  
application development  
article  
aspects  
benefit  
comparison / good comparison  
comparisons of python and java  
development  
efficiency  
efficiency of development  
efficiency of execution time  
execution  
execution time  
generation  
good  
ideal  
implementation  
item  
java  
language  
large  
negative  
negative aspects  
next  
next generation  
next generation version  
number  
part

partial  
particular  
positive  
positive aspects  
problem  
programming  
programming language / number of programming languages  
project  
python  
respect  
software  
software system  
system  
thing  
time  
to be  
to decide  
to develop  
to discuss  
to like  
to look  
to see  
to try  
to work  
useful  
version  
you

### Hierarkinen suhde:

article

- part of article

aspects

- positive and/or negative aspects

development

- application development

efficiency

- efficiency of development
- efficiency of execution time

language

- programming language
  - java
  - python

system

- software system

### Assosiatiohde:

application – application development  
benefit – useful  
benefits – positive aspects  
development – project  
implementation – application  
java – programming  
java – python  
problem – negative aspects  
programming – programming language  
project – to work  
python – programming  
software – programming  
ideal – good – positive  
benefits of implementation – problems of implementation

## LIITE 2: Tehtäväkuvauksen ulkopuolelta olevien hakusanojen edustami- en käsitteiden suhde tehtäväkuvauksen käsitteisiin

### 2.1. Tehtäväkuvaus B1

addiction	LK negative side effects
affecting	RK effect
analysis	Oma
blur	LK symptoms
cause	Oma
causes	Oma
computer	RK applications
cyber	RK cybersickness
cybersex	RK cybersickness
cybersick	RK cybersickness
cyberspace	RK virtual reality
disorder	SK cybersickness
disorders	SK cybersickness
disorientation	LK symptoms
dizzy	LK symptoms
environment	RK situations
Eugenia	Oma
experiencing	Oma
"eye strain"	LK symptoms
flashbacks	LK symptoms
games	Oma
harm	RK negative side effects
health	Oma
human	RK users
illness	SK cybersickness
immersion	LK negative side effects
it	Oma
kennedy	Oma
Kolasinski	Oma
Ltd.	Oma
maladies	SK cybersickness

medical	RK symptoms
motion	Oma
"motion sick"	RK cybersickness
"motion sickness"	RK cybersickness
nausea	LK symptoms
number	Syn amount
percent	RK amount
percentage	RK amount
phobias	LK negative side effects
prevalence	RK amount
problem	RK negative side effects
problems	RK negative side effects
profile	Oma
pshychological	Oma
quasy	Syn virtual
queasiness	LK symptoms
related	Syn associated
residual	RK amount
sick	RK cybersickness
sickness	SK cybersickness
simulation	RK virtual reality
Simulator	RK virtual reality
"simulator sick- ness"	RK cybersickness
statistics	Oma
strain	LK symptoms
systems	Oma
therapy	RK therapeutic
treating	RK treatments
virtuality	RK virtual
vision	Oma
witness	Oma



## 2.2. Tehtäväkuvaus C2

"Alan Brown"	Oma
"best practise"	Oma
"development cost"	Oma
"execution speed"	RK execution time
"Exploiting Python re-sources"	Oma
"java versus python"	RK comparisons of python and java
"java vs python"	RK comparisons of python and java
"java vs. python"	RK comparisons of python and java
"lage scale develop-ment"	RK developing large applica-tions
"large scale"	RK large applications
"man hours"	RK efficiency of development
"man months"	RK efficiency of development
"object oriented progam-ing languages"	LK programming languages
"object oriented pro-gramming languages"	LK programming languages
"rapid application devel-opment"	RK efficiency of development
advantage	Syn benefits
advantages	Syn benefits
Alan	Oma
api	RK application development
appropriate	Oma
as	Oma
bad	RK negative
benchmark	RK comparison
best	RK ideal / good / positive
big	Syn large
book	Oma
Brown	Oma
by	Oma
combining	Oma
coming	RK next
compare	RK comparison
compared	RK comparison
comparing	RK comparison
compiler	RK programming language
components	Oma
computing	RK applications / programming
cons	Syn negative aspects
conversion	Oma
costs	Oma
cots	Oma
design	RK develop

difference	RK comparison
disadvantage	Syn negative aspects
disadvantages	Syn negative aspects
distibuted	Oma
distributed	Oma
drawback	Syn negative aspects
drawbacks	Syn negative aspects
efficient	RK efficiency
elctronic	Oma
electronic	Oma
enterprise	Oma
experience	Oma
faqs	Oma
faster	Oma
feature	Syn aspects
features	Syn aspects
flexible	Oma
for	Oma
fortran	LK programming language
framework	Oma
future	RK next
implementing	Syn implementation
interactive	Oma
library	Oma
memory	Oma
net	Oma
new	Oma
object	Oma
open	Oma
performance	RK implementation
pitfall	RK negative aspects
program	RK programming
pros	Syn benefits
reengineering	RK develop
sax	Oma
scalability	RK application development
scale	RK application development
scaleability	RK application development
script	RK programming
scripting	RK programming
slower	RK execution time
source	RK programming
suitable	Oma
suited	Oma
tutorial	Oma
using	Oma
Watters	Oma
weakness	RK negative aspects

versus	RK comparison
vs	RK comparison
vs.	RK comparison
xml	Oma