



**LAUREA**  
AMMATTIKORKEAKOULU

*Uuden edellä*

# Internet-hakukoneiden tulevaisuus

---

Ojala, Lauri

2012 Kerava

Laurea-ammattikorkeakoulu  
Laurea Kerava

## Internet-hakukoneiden tulevaisuus

Lauri Ojala  
Tietojenkäsittelyn koulutusohjelma  
Opinnäytetyö  
Toukokuu, 2012

## Sisällys

|   |  |    |
|---|--|----|
| 1 | Johdanto .....   | 6  |
| 2 | Hakukoneen määritelmä ja toiminnallisuus .....                       | 6  |
| 3 | Hakukoneiden menneisyys ja nykyhetki.....                            | 8  |
|   | 3.1 Alkuvuodet .....   | 8  |
|   | 3.2 Nykytilanne.....   | 9  |
| 4 | Hakukoneiden tulevaisuuden kehityslinjoja.....                       | 10 |
|   | 4.1 Uudenlaisia teknologioita ja hakukonetyyppejä.....               | 10 |
|   | 4.2 Visioita .....   | 11 |
| 5 | Hakukoneiden tulevaisuus-kyselyn tulokset ja niiden analysointi..... | 14 |
| 6 | Oma näkemys .....  | 20 |
|   | 6.1 Yksityisyys .....  | 21 |
|   | 6.2 Käyttäjäkokemus .....  | 21 |
| 7 | Yhteenveto ja johtopäätelmät .....                                   | 21 |
|   | Lähteet .....  | 23 |
|   | Liitteet.....  | 25 |

Ojala, Lauri

### Internet-hakukoneiden tulevaisuus

Vuosi 2012 Sivumäärä 16 + 4

---

Tämä opinnäytetyö on tutkimustyö, jossa tarkastellaan millaiseksi Internet-hakukoneiden tulevaisuus on mahdollisesti muodostumassa. Tavoitteena oli erityisesti kartoittaa asiantuntijoiden mielipiteitä kyselyn avulla. Opinnäytetyö perustuu pääasiassa laadullisiin tutkimusmenetelmiin.

Aineistona käytettiin lähdekirjallisuutta, Internet-lähteitä sekä tutkimuksessa tehdyn kyselyn vastausmateriaalia. Kysely suoritettiin web-pohjaisena, ja se lähetettiin informaatioteknologia-alan tutkijoille Jyväskylän, Oulun ja Helsingin yliopistoihin.

Opinnäytetyöprosessin aikana saavutettiin tavoitteet asiantuntijamielipiteiden kartoittamisesta ja näin muodostui myös oma mielikuva hakukoneiden tulevaisuudesta.

Prosessin loppupuolella koottiin yhteen kyselyn tulokset. Tämän jälkeen ne analysoitiin ja näiden pohjalta muodostettiin oma näkemys ja loppupäätelmät.

Kyselyn tulosten perusteella voidaan sanoa, että perinteinen hakukone ei ole luopumassa valtaistuimestaan mobiililaitteiden hakusovelluksille, eikä Googlen markkinajohtajuus ole siirtymässä kilpailijoille. Lisäksi semantiikan eli tiedon merkitysoopin asema tulevaisuuden hakukoneissa tulee korostumaan.

Ojala, Lauri

**The future of web search engines**

|      |      |       |        |
|------|------|-------|--------|
| Year | 2012 | Pages | 16 + 4 |
|------|------|-------|--------|

---

This thesis is a research, which aims to answer what possibly could be the future of web search engines. The main purpose is to map out what the opinion of experts of the field is about the subject. The research in this thesis is based on the qualitative method of research.

The sources used herein were the books about the subject, internet sources and the results from the study that was conducted for the research. The study was carried out as a web survey, which was sent to the information technology research staff at the University of Jyväskylä, Oulu and Helsinki.

The thesis process can be considered a success at meeting its targets. The experts' opinions were received and the results were analyzed. Based on the results the views and ideas could be formed about what might be the way search engines are heading in the future. At the end of the process, the experts' opinions were summarized and personal visions and conclusions were drawn from this.

Based on the results of the survey can be deducted that the traditional web search engine is not going to sway from its leading position for the search applications in mobile devices, and Google's leading market share isn't threatened by the competition. The meaning of semantics in search engines is also going to strengthen in the future.

Keywords: Google, search engines, Internet

## 1 Johdanto

Hakukoneita käyttää todennäköisesti jokainen, jolla on mahdollisuus päästä Internetiin. Maailmassa tehdään joka sekunti yksinomaan Googlen kautta 34000 hakua, Yahoo!:n kautta 3200 ja Bingin kautta 927 hakua jokaista sekuntia kohden (Search Engine Land 2010). Tänä päivänä hakukoneissa liikkuu paljon rahaa, erityisesti hakukonemarkkinoinnissa ja optimoinnissa.

Alun perin oli tarkoitukseni tehdä opinnäytetyö Puolustusvoimille selvityksenä semanttisista hakukoneista. Tämän kuitenkin jäädessä toteutumatta päätin jatkaa hakukoneteeman saralla, ja aiheeksi tuli lopulta internet-hakukoneiden tulevaisuus.

Opinnäytetyön tärkeimpänä tavoitteena oli saada selville asiantuntijoiden mielipide hakukoneiden tulevaisuudesta ja muodostaa tämän pohjalta oma kokonaiskuva siitä, millaiselta internet-hakukoneiden tulevaisuus voisi näyttää.

Tutkimustyössäni käytin pääasiassa laadullisia menetelmiä, määrällisiä menetelmiä myös, lähinnä kyselytulosten analysoinnissa. Alkuasetelmana selvitin paljon hakukoneiden toimintatapa ja niiden historiaa sekä tutustuin erilaisiin uudentyyppisiin hakukoneisiin. Tämän jälkeen lähetin internetissä toteutetun kyselyn yliopistojen asiantuntijoille sekä muutamalle kotimaiselle hakukoneasiantuntijalle.

Tutkimusongelmanani halusin erityisesti selvittää, onko hakukoneiden tulevaisuus nimenomaan kannettavien laitteiden, kuten kännyköiden ja tablet-tietokoneille räätälöidyissä aihepiireihin erikoistuneissa hakusovelluksissa.

## 2 Hakukoneen määritelmä ja toiminnallisuus

Internet-hakukone määritellään seuraavasti:

“Hakukone on internet-pohjainen ohjelma, joka etsii jatkuvasti internetistä uusia sivuja eritellen ja liittäen ne hakemistoonsa erityisten hakusanojen mukaan. Näitä hyväksi käyttäen hakukone tulostaa ruudulle käyttäjän syöttämiä hakusanoja lähimpänä olevat sivut. Analysointi tapahtuu käytännössä eri hakukoneissa erilaisilla menetelmillä. Näitä menetelmiä kehitetään jatkuvasti paremmiksi ja toimivammiksi, ja niiden täsmällinen toiminta on joka hakukoneessa suuri salaisuus, jota ei paljasteta.” (Wikipedia 2011.)

Hakukone terminä voi kuitenkin tarkoittaa mitä tahansa ohjelmaa, jolla haetaan tietoa tietokannasta tai vastaavasta hakusanoilla ja ehdoilla. Tällaisia ovat esimerkiksi kirjastojen aineistohakukoneet.

Hakukone yhdistää hakusanan tietokantaan, jonka se on luonut web-sivustoista ja hakee listan URL:ja sekä tiivistelmiä sisällöstä, joiden se katsoo liittyvän parhaiten hakuun. Tällä tavalla kaikki tämän hetken käytetyimmät hakukoneet toimivat, vaikka on olemassa myös muunlaisia hakukoneiden toimintatapoja. (Battelle 2005, 19-25; Wikipedia 2011.)

Hakukone koostuu crawlerista, indeksistä ja ajonaikaisesta järjestelmästä tai kyselyprosessorista, joka on samalla käyttöliittymä ja yhdistävä tekijä käyttäjän sekä indeksin välillä. Ajonaikainen järjestelmä myös vastaa relevanssin ja sijoituksen (kuten PageRankissa) prosessoinnista. (Battelle 2005, 19-25; Wikipedia 2011.)

Crawlerin vastuulla on selata internetiä automatisoidusti seuraten jokaista sivustolla X sijaitsevaa linkkiä. Jokaisen sivun sisältö analysoidaan indeksointia varten, esimerkiksi otsikon sanat ja metatietokentät. Tämä tieto tallennetaan indeksiin hakukyselyitä odottamaan. (Battelle 2005, 19-25; Wikipedia 2011.)

Indeksin tarkoituksena on nopeuttaa tiedon löytämistä. Edellä mainittu analysoitu tieto on tallennettu "välimuistiin". Kyseessä on sivuston karsittu versio, joka sisältää kaiken teksti- ja metatiedon, jota crawlerin analysoimalla sivustolla on. Tätä versiota sivustosta, joka on indeksiin tallennettuna, käytetään hakukyselyssä vertailtaessa hakusanojen relaatiota sivustoon. (Battelle 2005, 19-25; Wikipedia 2011.)

Käyttäjä kirjoittaa hakukyselyn hakukoneen hakukenttään. Hakukone vertailee kyselyn sisältöä indeksinsä sisältöön ja antaa listan sivustoista, joiden se katsoo parhaiten sopivan hakukyselyn motiiveihin. Listalla on sivustoista otsikko sekä lyhyt tekstitiivistelmä. Googlen tapauksessa voi halutessaan myös katsella sivuston välimuistiversiota. Tämä on hyödyllistä, jos vaikkapa alkuperäisen sivuston palvelin on alhaalla. (Battelle 2005, 19-25; Wikipedia 2011.)

### 3 Hakukoneiden menneisyys ja nykyhetki

#### 3.1 Alkuvuodet

Useiden lähteiden mukaan ensimmäinen hakukone internetissä oli Archie.

Tämä hakukone syntyi yliopistojen FTP-arkistojen selaamisen helpottamiseksi. Lopulta se kuitenkin päätyi internet-hakukoneeksi. Archie ei kuitenkaan ollut yksittäisessä osoitteessa, vaan eri tahoilla oli omat versionsa Archiestä ja omat hakutietokantansa. (Battelle 2005, 39-40; Wikipedia 2011.)

Alkuaikoina tietoa hakiessa tapana oli yhdistää telnet-asiakasohjelmalla Archie-serverille tai lähettää kyselyt sähköpostilla. Myöhemmin saatiinkin jo WWW-käyttöliittymä. Archie oli kuitenkin hakukoneena varsin rajoittunut, sillä se ei hakenut ja indeksoinut muuta kuin palvelimella olevien tiedostojen nimet ja tämän vuoksi tiedon etsijän tuli tietää tarkkaan, minkänimistä tiedostoa oli hakemassa. (Battelle 2005, 39-40; Wikipedia 2011.)

Veronica ja Jughead tulivat seuraavaksi. Veronica lisäsi aiempaan toiminnallisuuteen asiasanahaun, ja Jughead oli erikoistunut hakemaan ainoastaan yhdeltä nimenomaiselta serveriltä. Ensimmäinen varsinainen koko internetin skaalassa toimiva hakukone oli sarja perl-skriptejä, jotka loi Oscar Nierstrasz Geneven yliopistossa. Skriptit kokosivat muiden listat internet-resursseista yhteen luetteloon, joka tultiin myöhemmin tuntemaan nimellä W3Catalog. (Battelle 2005, 40; Wikipedia 2011.)

Vuodesta 1993 vuoteen 1996 internet kasvoi 130 sivusta yli kuuteensataantuhanteen sivuun. Asioiden etsiminen ja löytäminen netistä oli edelleen ongelmallista ja vaikeahkoa, kunnes Matthew Gray kehitti ensimmäisen crawler-robotin. Wanderer oli robotti, joka automaattisesti loi indeksin sivustoista, ja Gray loi tämän indeksin päälle hakukäyttöliittymän, joka antoi käyttäjien hakea tietoa sieltä. Grayn luomuksella oli kuitenkin yllättävä sivuvaikutus: alkujaan se söi kaiken kaistanleveyden, mitä indeksoitavan sivuston palvelimella oli tarjota. Pie-nellä robotin toimintalogiikan hienosäädöllä kuitenkin ongelma saatiin hallintaan. (Battelle 2005, 40-41; Stross 2008, 24; Wikipedia 2011.)

Wanderer jäi pian tehokkaampien hakukoneiden varjoon. Tällainen oli esimerkiksi WebCrawler. Toisinkuin edeltäjiensä avulla, tällä hakukoneessa saattoi etsiä mitä tahansa sanaa, joka löytyi web-sivulta. Tämä onkin nykyään hakukoneiden vakio-ominaisuus. Se oli ensimmäinen, jonka löysivät myös IT-alan ei-ammattilaiset. Samana vuonna 1994 kuvioihin ilmestyi myös Lycos, josta tuli merkittävä tekijä hakukonepalveluiden tarjoajana. Vähitellen hakukoneita alkoi syntyä kiivaaseen tahtiin, esimerkiksi Magellan, Excite, Infoseek, Inktomi, Northern Light



ja Altavista ilmestyivät hetkeä myöhemmin hakukoneiden maailmankartalle. (Battelle 2005, 41-42; Wikipedia 2011.)

Yahoo! oli näihin aikoihin suosituin tapa etsiä tietoa internetistä. Se tosin ei vielä sisältänyt varsinaista hakutoiminnallisuutta, vaan oli portaalisivusto, josta käyttäjät etsivät haluamaansa tietoa sisältävää sivustoa hakemistoihin lajiteltuina aihepiirin tms. mukaan. (Battelle 2005, 57-63; Wikipedia 2011.)

Vuonna 1996 selainyritys Netscape halusi tehdä sopimuksen jonkin hakukoneyrityksen kanssa yksinoikeudesta olla aloitussivustonsa hakukone. Lopulta sopimukset tehtiin viiden eri yrityksen (Yahoo!, Magellan, Lycos, Excite ja Infoseek) kanssa johtuen kovasta kiinnostuksesta tarjoukseen. Hakukoneet olivat myös 2000-luvun vaihteessa suosittu sijoituskohde ennen IT-kuplan puhkeamista. (Battelle 2005, 39-63; Stross 2008, 21-29; Wikipedia, 2011.)

### 3.2 Nykytilanne

Vuoden 2000 tienoilla Googlen hakukone nousi internetin suosituimmaksi tavaksi etsiä tietoa. Googlen suurimpana innovaationa oli yrityksen oma PageRank-algoritmi hakutulosten järjestämiseksi. (Battelle 2005, 65-93, 123-152; Stross 2008, 65-70; Wikipedia, 2011.)

Samaan aikaan Yahoo! siirtyi käyttämään Inktomin hakukonetoiminnallisuutta, nielaisten yrityksen itseensä vuonna 2002. Seuraava vuonna se kuitenkin siirtyi käyttämään Googlen palveluita, kunnes 2004 yritys kehitti oman hakukoneensa. (Battelle 2005, 65-93, 123-152; Stross 2008, 65-70; Wikipedia, 2011.)

Microsoftin panos hakukonemarkkinoille oli MSN Search. Tämä käytti myös aluksi muiden kehitteitä teknologioita, kunnes samoihin aikoihin kuin Yahoo! se siirtyi Microsoftin sisäisen kehitystyön tuottamaan hakukoneteknologiaan. Hakukone uudelleenbrändättiin Bingiksi vuonna 2009. Samana vuonna Yahoo! siirtyi Bingin teknologia-asiakkaaksi. (Battelle 2005, 65-93, 123-152; Stross 2008, 65-70; Wikipedia, 2011.)

Tällä hetkellä merkittävimmät hakukonepalvelut internetissä ovat Google, Yahoo! ja Bing (Battelle 2005, 30). Lisäksi on myös paljon muita, mutta suurin osa internetissä tehdyistä hauista suoritetaan edellä mainituilla hakukoneilla. Googlen hakukone osaa rajallisesti semantiikkaa tällä hetkellä. Esimerkiksi siltä löytyy resepteihin erikoistunut hakumahdollisuus sekä hakukoneen hakulomakkeessa on jo pidemmän aikaa kyennyt tekemään yksinkertaisia laskutoimituksia ja mittakäännöksiä, joihin on saanut vastauksen.

## 4 Hakukoneiden tulevaisuuden kehityslinjoja

### 4.1 Uudenlaisia teknologioita ja hakukonetyyppejä

Semantiikka eli merkitysoppi on löytämässä tiensä hakukoneisiin. Semanttisen hakukoneen päämääränä on parantaa haun tarkkuutta selvittämällä haun suorittajan perimmäinen tarkoitus haun suhteen, hakutermien merkitys kontekstin sisällä ja täten antaa myös paremmin haettuun asiaan liittyviä hakutuloksia. Semanttinen haku tulee todennäköisesti tehostamaan perinteistä hakukonetta, ei korvaamaan sitä. Suuri osa hakukyselyistä on internetissä navigoimiseen liittyviä, ja näihin semanttinen haku soveltuu huonosti.

Esimerkkinä semanttisen haun toiminnasta voidaan pitää seuraavaa käyttäjäskenaariota: olet hakemassa tietoa vaaleista. Tällöin hakukone saattaisi hakea dokumentteja, jotka sisältävät sanat äänestää, kampanjointi tai äänestys, vaikka dokumentti ei sisältäisikään suoranaisesti sanaa "vaalit". Tärkeänä osana tätä prosessia on sanojen merkitysopin tunteminen soveltaen sekä hakukyselyihin että internetin sisältöön. Hakukoneen tulee tietää, haetko isoa kissaa vai autoa hakiessasi hakusanalla "Jaguar". Semanttisia hakukoneita ovat mm. Hakia, SenseBot, Powerset, DeepDyve ja Cognition. (Pandia Search & Social 2009.)

Kesän 2011 MikroPC-lehdessä toimittaja Jukka K. Korpela on etsinyt ratatouillen reseptiä, mutta on saanut enimmäkseen samannimisen elokuvan tuloksia. Apuna ongelmaan hän on havainnut Googlen hakukoneen käyttöliittymän työkalupalkissa olevat erikoistuneet haut, kuten tässä tapauksessa reseptihaut. (Korpela 2011.)

Moni entinen Googlen työntekijä on perustanut oman hakukonebisneksen. Kaksi näistä, Aardvark ja FriendFeed, erikoistuvat sosiaaliseen hakuun. Aardvark osti Google vuonna 2010 kuitenkin itselleen takaisin (palvelun tarjonta lakkasi yrityskaupan myötä) ja FriendFeedin taas osti Facebook. FriendFeed on sosiaalinen hakukone, joka on erikoistunut tiedon hakemiseen sosiaalisista medioista. Tietoa haetaan siis ihmisten luomasta sisällöstä eri sosiaalisista yhteisöpalveluista. (Wikipedia 2011.)

Aardvarkin toimintatapa oli hieman erilainen edelliseen nähden. Käyttäjän pikaviestinpalveluun lisättiin kaveriksi Aardvark-robotti, joka kysyi väliajoin käyttäjältä, haluaisiko tämä vastata kysymykseen, jonka joku toinen palvelun käyttäjä oli hakurobotilta kysynyt. Vastatessa myöntävästi robotti välitti käyttäjälle kysymyksen ja kysyjän nimen. Käyttäjä saattoi tällöin vastata kysymykseen tai välittää kysyjälle tahon, joka osaisi vastata kysymykseen. Robotille pystyi samalla tavalla keskustelussa välittämään kysymyksen, joka edellämämainitulla tavalla välitettiin jollekulle muulle vastattavaksi. (Search Engine Journal 2010; Wikipedia 2011.)

Tämäntapainen sosiaalinen hakukone olisi erinomainen tapa hakea tietynlaista tietoa, koska ihminen tietää sanojen merkitysoopin paremmin kuin tietokone ja antaa parempia hakutuloksia. Toisaalta hakutuloksiin voivat vaikuttaa ihmisen omat näkökannat ja mielipiteet, esimerkiksi kysymys jääkiekkoseurojen paremmuudesta voi antaa puolueellisia hakutuloksia riippuen vastaajasta. (Search Engine Journal 2010; Search Engine Land 2010; Wikipedia 2012.)

Qwiki on hakukone, joka tekee hakemastasi aiheesta pienen multimediaesityksen, jossa kerrotaan tärkeimmät faktat aiheesta. Qwikin aihe sivujen parantamiseen voivat käyttäjät myös wiki-periaatteen mukaan osallistua itse. (Qwiki 2012; Unknown Pixels 3.0, 2011)

Shazam on mobiililaitteissa toimiva hakukone, joka hakee puhelimen tai muun mobiililaitteen mikrofoniasta ottamansa ääninäytteen perusteella tietokannastaan tiedon, mitä musiikkikappaletta käyttäjä tällä hetkellä kuuntelee. Kyseisellä sovelluksella voi tutustua uuteen musiikkiin ja hakea tietoa siitä.

Applen Siri on puheentunnistukseen perustuva ohjelma, joka antaa iPhone 4S:n käyttäjille mahdollisuuden tehdä puhelimessaan toimintoja puhumalla. Siri on kuitenkin myös hakukone, joka toimii puhumalla. Ohjelma luottaa toiminnoissaan isoon partnerien muodostamaan ekosysteemiin. (Wikipedia 2012.)

Tähän joukkoon kuuluu toimijoita kuten Yelp, CitySearch, Eventful, MovieTickets, RottenTomatoes, New York Times, Bing Answers, Wolfram Alpha, Bing, Yahoo! ja Google. Näiden avulla Siri hakee mm. ravintoloita, yrityksiä, tapahtumia, konsertteja, elokuva-tietoja sekä arvioita, faktakysymyksien vastauksia ja tavallisista web-hauista tietoa. (Wikipedia 2012.)

On huomion arvoista, että Googlen toimitusjohtaja Eric Schmidt näkee Sirin vakavana uhkana bisnekselleen. (Gizmodo 2011; TUAW 2011; Wikipedia 2012.)

#### 4.2 Visioita

John Battelle, Wired-lehden perustajajäsen ja The Industry Standardin perustaja, on pohtinut sekä kirjassaan *The Search* (2005, 229-251) muutamassa kappaleessa että internetissä sivustolla Search Engine Land otsikon ”John Battelle On The Future Of Search” alla hakukoneiden, joista lähinnä Googlen, tulevaisuutta.

Battelle on jälkimmäisessä kirjoituksessaan pitkälti teemoihin erikoistuneiden hakusovellusten kannalla, mitä tulevaisuuteen tulee. ”Ehkäpä tulevaisuus on ennemminkin hakusovellus kuin hakukone.” Battelle käyttää Expediaa esimerkkinä: ”Teemme sofistikoituneen haun, jonka

avulla päätämme kolmessa minuutissa, minkä lennon otamme. Kaikki data on strukturoitua sovelluksessa. Expedia on hakusovellus ja päätöksenteon tukisovellus. Tämän jälkeen avaa-  
me Googlen ja etsimme tietoa jääkaapeista tai autoista. Kirjoitat “1967 Mustang” ja toivot,  
että saat haluamasi vastauksen hakukoneelta, kun sinun tarvitsisi tehdä vain 2-3 minuuttia  
töitä strukturoidun haun parissa sovelluksella geneerisen hakukoneen sijasta.”

Samassa artikkelissa pohditaan ansaintalogiikan muutoksia. “Koska hakukone on kaikkialla,  
Google tahkoo miljardeja dollareita jokaisena neljänneksenä. Haku sovelluksena, jossa en-  
simmäinen hakusi ei ole haku itsessään, vaan oikean sovelluksen hakemista, on erilainen käyt-  
täjäskenaario. Markkinoiden vaikutus ja määräävyys on pirstoutunut yhdestä pelaajasta kym-  
meniin tuhansiin. Sinä tai minä istuessamme toimistossa viikonlopun yli voisimme saada aikai-  
seksi absoluuttisesti parhaan strukturoidun hakusovelluksen parhaan puutarhurin päättelemi-  
seen, joka siistisi ruusupensaasi. Ja se on uhka Googlen paikalliselle haulle. Jos paras sovellus  
putkimiehen löytämiseen on putkimies-sovellus iPhonella, lataat sen ja se hakee automaatti-  
sesti tulokset Yahoosta, Bingistä ja Googlestä ja sitten hakee kaikki arvostelut Yelpistä ja An-  
gie’s Lististä, ristiin vertailee ne Better Business Bureauhun ja Diamond Certifiediin tehtyjen  
valitusten kanssa. Jos käytät siis sovellusta, missä on Googlen osuus? Ellei Google ohjelmoi  
tuollaista sovellusta. Ja siinä on yksi syy, miksi Steve Jobs ja Eric Schmidt eivät pidä toisistaan  
paljoa tällä hetkellä. Meillä on interface-sota käynnissä ja minusta se on kiehtovaa.”

Googlen työntekijöiden virallisessa blogissa on kirjoitettu näkemys hakukoneiden tulevaisuu-  
desta. Kirjoittajan mielestä haussa on vielä tilaa innovaatioille, muutoksille ja kehitykselle  
kohti parempaa käyttäjän tarpeisiin vastaamista. Haku ei ole tarpeeksi käyttäjäystävällinen,  
ja haun tulisi olla enemmän mobiili, eli helposti tai helpommin käytettävissä kännyköissä,  
tableteissa ja autoissa sekä mahdollisissa tulevaisuuden laitteissa. (Google Official Blog  
2008.)

Esimerkkinä blogaaja käyttää kaveriporukassaan syntynyttä keskustelua: Mikä on vanhin elävä  
olento? Kukaan ei halunnut ottaa kömpelöä, hidasta tietoteknistä laitettaan esille tehdäkseen  
haun asiaa koskien. Hänen mielestään olisi kätevämpää, jos olisi laite, jossa olisi hyvät yhtey-  
det ja joka osaisi tehdä hakuja ilman keskeytyksiä. (Google Official Blog 2008.)

Blogissa esitetään ajatus puettavasta laitteesta, joka tekee hakuja taustalla sanoista, joita se  
poimii keskustelun seasta ja esittää aiheeseen liittyvät faktat. Tai miksi haun yleensä pitäisi  
olla kirjoittamista, sen sijasta, että tekisi haun vaikkapa kameralla otetusta kuvasta linnuista,  
jotka sattuivat lentämään ohitse tai pätkästä ääntä (tällainen hakukone on jo nykyään osittain  
olemassa, tosin se hakee vain kuulemansa musiikkikappaleen tiedot), jonka hakukone ana-  
lysoisi ja esittäisi siihen liittyvää tietoa. (Google Official Blog 2008.)

Seuraavan kymmenen vuoden aikana blogissa pohditaan tapahtuvan mullistavia muutoksia hakutyypeissä internetiin kykenevien mobiililaitteiden lisääntyessä ja yleistyessä. Selvyytenä pidetään kuitenkin, että avainsanapohjainen haku on äärettömän tehokas, mutta samalla rajoittava tekijä. (Google Official Blog 2008.)

Toukokuussa 2007 Googlen julkistamaa ”Universal Search”-hakukonetta pidetään blogissa ensimmäisenä ja tärkeänä askeleena kohti tätä päämäärää. Universal Search kattaa kartta-, blogi-, kuva-, video- sekä kirjahaut. Presentaatiotapaa eli linkkilistaa pidetään kuitenkin turhan lineaarisena. (Google Official Blog 2008.)

Googlen vastaus tulevaisuuden haasteisiin on blogin kirjoittajan mielestä selkeä: tulevaisuuden hakukone on parempi, koska se tietää enemmän sinusta. Blogaaja kiirehtii kuitenkin mainitsemaan, että yksittäinen käyttäjä hallitsee kuitenkin, mitä tietoja hänestä säilytetään ja kerätään. Tosin nykyisellään on huomattava, että esimerkiksi oman hakuhistorian poistaminen Googlen syövereistä on tehty melkoisen hankalaksi valikkoviidakossa seikkailemiseksi. Silti jopa aivan perustason tiedoilla käyttäjistä saadaan blogin mukaan hakutuloksia enemmän relevanteiksi ja yksilöllisiksi. (Google Official Blog 2008.)

Yksi tärkeä tällainen yksilöivä tieto hakukoneen kannalta on sijaintisi. Halutessasi tietää vaikkapa, milloin lähin Pizzeria X aukeaa, tulisi hakukoneen tietää nykyinen sijaintisi. Toinen yksilöivä tieto on sosiaalinen konteksti. Tarkemmin sanoen ystäväsi ja suhteesi heihin. Tälläkin alueella Google on jo blogin kirjoittamisen jälkeen ehtinyt tehdä oman avauksen aihealueeseen, Googlen plus-palvelulla, joka on heidän oma sosiaalinen yhteisöpalvelunsa. Hakiessasi tietoa kyseisestä pizzeriasta hakukone voisi tietää, että ystäväsi työskentelee kyseisessä pizzeriassa ja haku ”Pizzeria X” tarkoittaa juuri tätä pizzeriaa hakukontekstissä. (Google Official Blog 2008.)

Lisäksi on hakuja, joissa ystävät voivat auttaa. Blogaaja kertoo hakeneensa tietoa särkylääkkeistä. Hän halusi tietoa aspiriinista ja cumitiniinista, mutta kirjoitti jälkimmäisen väärällä tavalla. Lähetettyään myöhemmin sähköpostia lääkäriystävälleen hän tiesi oikean kirjoitusasun ja lisäksi sai ystävänsä mielipiteen lääkkeiden eroista. (Google Official Blog 2008.)

Tietoa voi myös olla saatavilla, mutta väärällä kielellä. Google investoi paljon koneelliseen kääntämiseen blogin mukaan. Googlella on tietysti Translate-palvelunsa, jolla voi kääntää sivuja kielelle, jota ymmärtää paremmin, mutta blogissa annetaan vihjeitä, että tämä on vain pieni osa tulemassa olevista käännöspalveluista. (Google Official Blog 2008.)

## 5 Hakukoneiden tulevaisuus-kyselyn tulokset ja niiden analysointi

Kyselyssä oli kymmenen väittämää, joihin hakukone- ja internet-alan ammattilaiset vastasivat asteikolla yhdestä viiteen, miten samaa mieltä he väittämästä olivat. Numero kolmen valitsemista tai valitsematta jättämistä käytettiin tässä kyselyssä, jos ei osannut vastata kysymyksen. Kaikki kyselyyn vastanneet olivat anonymoijia, joten heidän vastauksiaan ei voitu yksilöidä johonkin tiettyyn henkilöön. Kyselyn lopussa oli mahdollisuus ilmaista oma mielipiteensä tekstikentässä. (Kuva 1.)

The screenshot shows a web browser window with the URL [https://elomake3.laurea.fi/lomake/lomake.php?lomake\\_id=5127&nakyma=lomake&kategoria=tekstit](https://elomake3.laurea.fi/lomake/lomake.php?lomake_id=5127&nakyma=lomake&kategoria=tekstit). The page header features the logo 'E-lomake' and navigation buttons for 'lomakkeet' and 'uusi lomake'. Below the header, there are links for 'Lomakkeen muokkaus', 'Esikatselu', and 'Julkaistu'. The main content area is titled 'Hakukoneiden tulevaisuus - Kysely' and includes instructions: 'Luekaa väittämät ja vastatkaa 1-5 asteikolla miten paljon olet samaa mieltä väittämän kanssa (1 = eri mieltä, 3 = ei eri eikä samaa mieltä ja 5 = samaa mieltä). Jos ette halua vastata tai ette ymmärrä jotakin kysymystä, voitte jättää vastaamatta ao. kysymyksen kohdan. Jos haluatte, kommentoikaa väittämän aihetta väittämien jäljessä olevaan tekstikenttään. Kun olette valmis, painakaa tallenna-nappia lomakkeen alaosassa. Strukturoitu haku, esim: <http://searchengineland.com/structure-search-coming-to-yahoo-12774>'. Below the instructions is a section labeled 'Valinnat' with a dropdown menu. The survey questions are listed in a table with a 5-point rating scale (1-5) for each. At the bottom, there is a text input field for 'Omat kommentit aiheeseen liittyen:'.

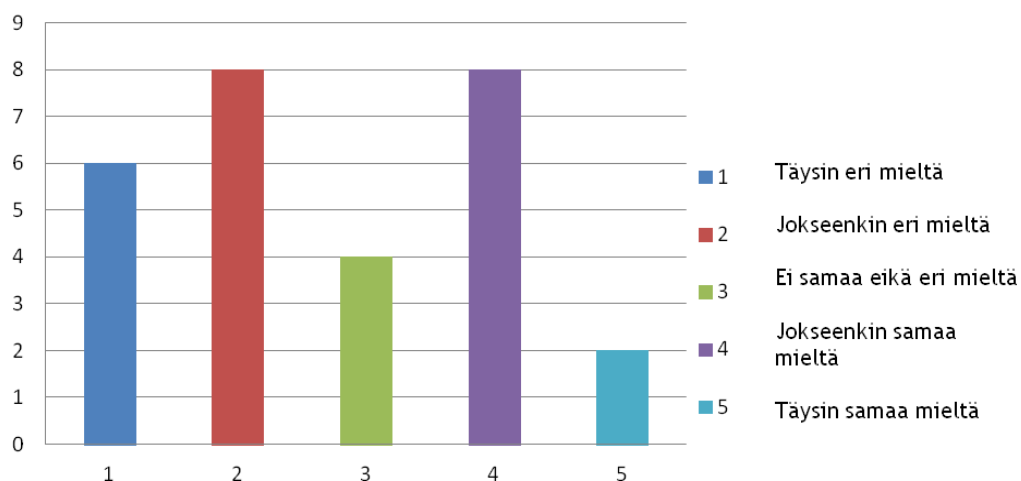
| Väittämiä:   | 1                     | 2                     | 3                     | 4                     | 5                     |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Google jää tulevaisuudessa yleishakuteokseksi ensyklopediakirjojen tapaan                                      | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Paikannustiedot ovat ongelma yksityisyyden suhteen mobiileissa hauissa.  | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Ainut hakukone internetissä, jolla on merkitystä, on Google.   | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Hakukoneet osaavat tällä hetkellä soveltaa semantiikkaa hakuhiin hyvin vähän.                                  | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Semanttinen hakukone ei tule syrjäyttämään perinteistä hakukonetta.  | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sosiaalinen hakukone on paras tapa hakea tietoa, sillä ihminen on osallisena hakutuloksiin.                    | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Hakukoneiden tulevaisuus on mobiililaitteissa toimiva johonkin tiettyyn aihepiiriin erikoistunut hakusovellus. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Mobiililaitteen hakusovellukset säästävät käyttäjänsä aikaa.   | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Tulevaisuudessa hakukoneita ei enää käytetä näppäimistöyöteellä lainkaan, vaan esimerkiksi puhumalla.          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Strukturoitu haku säästää aikaa verrattuna umpimähkäiseen tiedon etsimiseen Googlella.                         | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Omat kommentit aiheeseen liittyen:

Kuva 1. Hakukoneiden tulevaisuus-kyselyn lomake.

Väittämät oli laadittu hieman provosoiviksi, jotta saatiin vaihtelevia ja erilaisia mielipiteitä aikaiseksi. Kyselyyn vastasi 28 henkilöä. Kymmenen heistä kirjoitti omia näkemyksiään itse kyselyvastausten oheen.

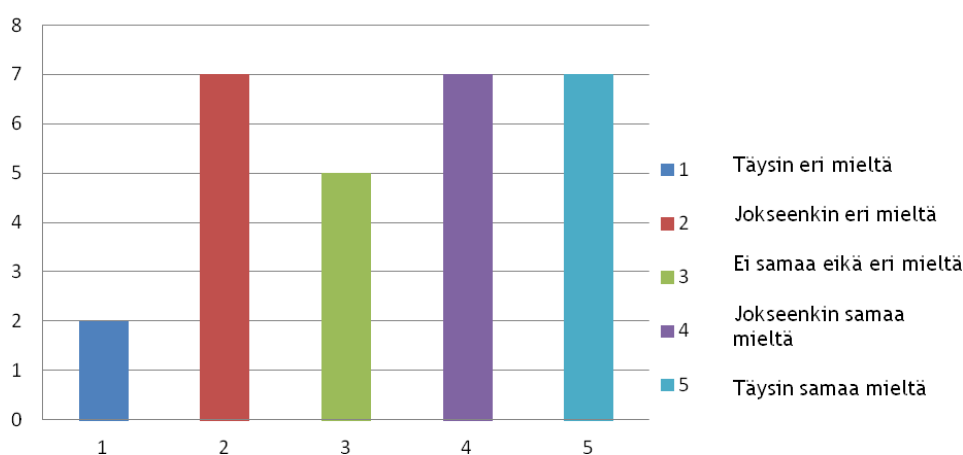
Ensimmäinen väittämä oli, että Google jää tulevaisuudessa yleishakuteokseksi ensyklopediakirjojen tapaan. Enemmistö vastaajista eli noin puolet oli väittämän ajatusta vastaan (Kuvio 1).



Kuvio 1. Google jää tulevaisuudessa yleishakuteokseksi ensyklopediakirjojen tapaan.

Onkin hyvin todennäköistä, että Google tulee yrityksenä pitämään itsensä hakukoneiden ykkössijalla tavalla tai toisella, esimerkiksi ostamalla nuoria ja lupaavia uudenlaisia hakukone- tai sosiaalisen median alan yrityksiä. Eräs vastaajista kommentoi tätä väittämää seuraavanlaisesti: ”Yleishakukoneella on vankka käyttöpohja ja tämä kehityspolku näyttää jatkuvan. Muiden hakukoneiden on vaikea syrjäyttää tätä perinteistä.” Toisen vastaajan mielestä taas Google ei ole ”yleishakuteos” eikä pyri sellaiseksi.

Toisessa väittämässä väitettiin paikannustietojen olevan ongelma yksityisyyden suhteen mobiileissa (mobiililaitteilla tehdyissä) hauissa. Noin puolet vastaajista oli väittämän aiheesta samaa mieltä (Kuvio 2).



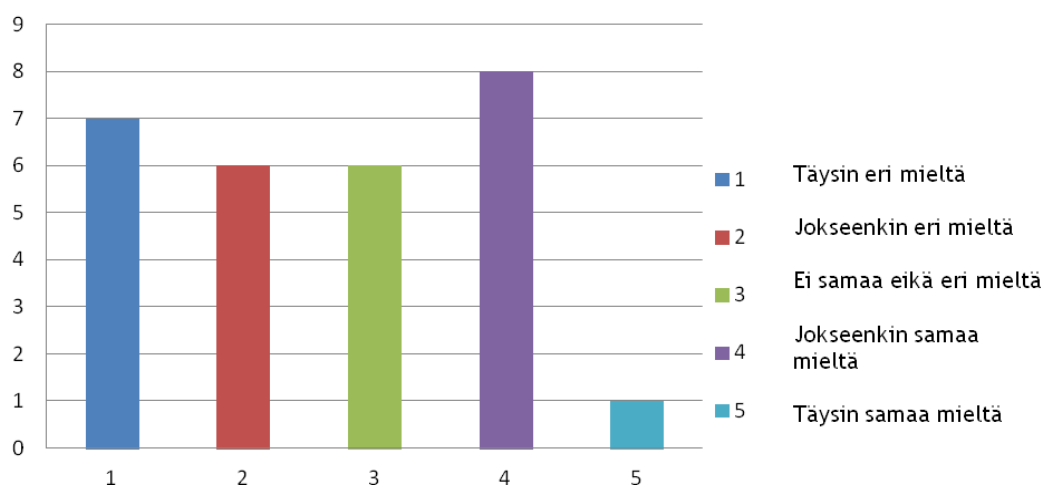
Kuvio 2. Paikannustiedot ovat ongelma yksityisyyden suhteen mobiileissa hauissa.

Paikannustietojen lisääminen hakuun helpottaa tietysti mobiileja hakuja huomattavasti, koska oletettavasti suurin osa näillä laitteilla tehdyistä hauista etsii jotakin lähiympäristöstä,

esimerkiksi erikoisliikkeitä. Lisäksi tämä tuo Googlelle ja muille hakukoneyrityksille jälleen yhden tavan tienata rahaa hauilla sponsoroitujen linkkien muodossa.

Ongelmana tässä on loppukäyttäjän yksityisyyden ja tietoturvan kannalta se, että joku tietää aina missä olet. Lisäksi mobiilien datayhteyksien suojaus ja salaaminen ei ole aivan samaa luokkaa kuin edes kehnoimman WLAN-yhteyden vastaavien. (Engadget, 2010.)

Kolmannessa väittämässä pidettiin Googlea ainoana hakukoneena internetissä, jolla on merkitystä. Tätä näkemystä tuki niukka enemmistö vastaajista (Kuvio 3).

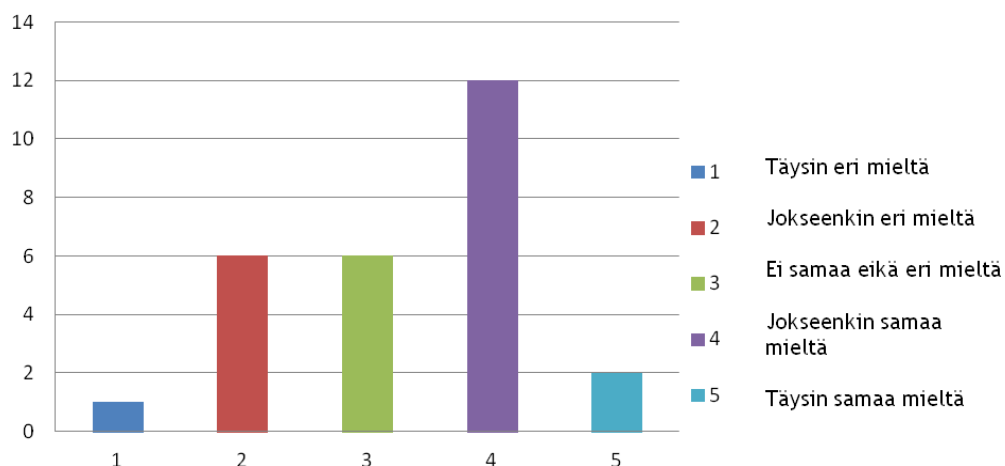


Kuvio 3. Ainut hakukone internetissä, jolla on merkitystä, on Google.

Näin tällä hetkellä Google on ainakin markkinaosuutensa (Google 66 %, Bing 15.1 % ja Yahoo 14.5 % sekä muilla loput) perusteella ylivoimaisesti merkittävin hakukone (Forbes 2012).

Neljäs väittämä: ”Hakukoneet osaavat tällä hetkellä soveltaa semantiikkaa hakuihin hyvin vähän.” Hieman yli puolet vastaajista olivat väittämän kanssa samaa mieltä (Kuvio 4).

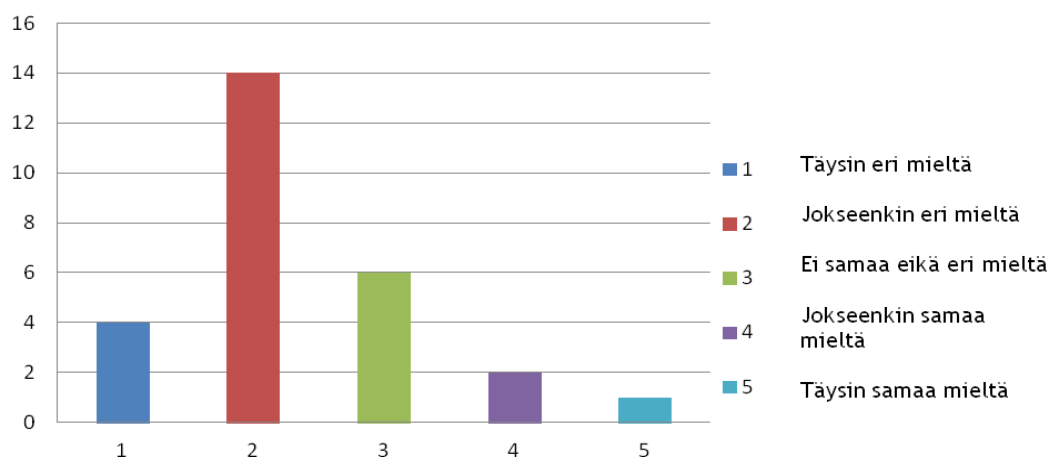




Kuvio 4. Hakukoneet osaavat tällä hetkellä soveltaa semantiikkaa hakuihin hyvin vähän.

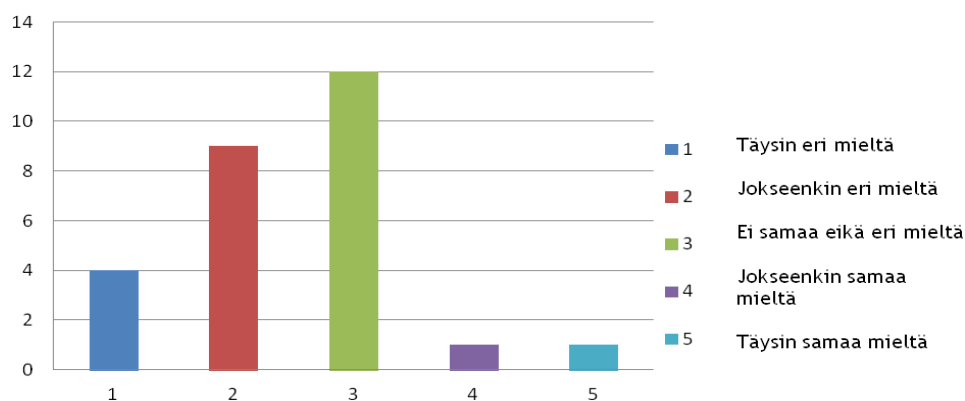
Muuan vastaajista kiteytti tilanteen seuraavanlaisesti: ”Semanttiset hakukoneet ovat olleet tulemaisillaan viimeiset 20 vuotta. Tilanne on todennäköisesti sama vielä 20 vuoden päästäkin”. Semantiikan toteuttamisessa suurin ongelma tällä hetkellä on, kuten eräs vastaajista oli todennut, on valtava määrä sisältöä, joka täytyisi käydä läpi ja ’korvamerkitä’ eli lisätä sisällön kuvailu tiedostoihin.

Viidennes väittämä kuului: ”Semanttinen hakukone ei tule syrjäyttämään perinteistä hakukonetta.” Suurin osa vastaajista oli toista mieltä (Kuvio 5). Kuitenkin edellisessä väittämässä kuvatut ongelmat tulevat tekemään perinteisen hakukoneen syrjäyttämistä hyvin hankalaa toistaiseksi.



Kuvio 5. Semanttinen hakukone ei tule syrjäyttämään perinteistä hakukonetta.

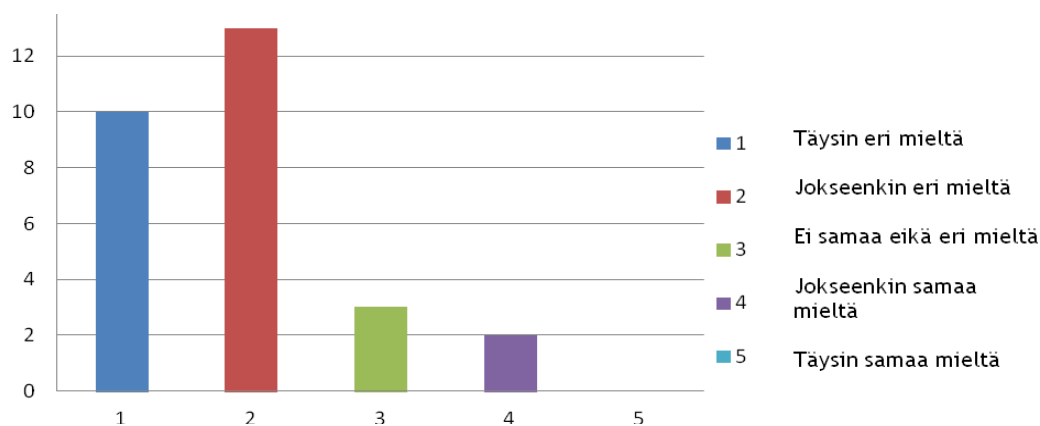
Kuudennessa väittämässä sosiaalista hakukonetta pidettiin parhaana tapana tiedon hakemiseen. Tämä väittämä jakoi hieman mielipiteitä (Kuvio 6).



Kuvio 6. Sosiaalinen hakukone on paras tapa hakea tietoa, sillä ihminen on osallisena hakutuloksiin.

Eräs asiantuntijavastaajista yltyi jopa sanomaan sosiaalista hakukonetta huonoimmaksi asiakasi, mitä haulle on tapahtunut. Sosiaalisella hakukoneella on kuitenkin paikkansa tietynlaista tietouden muotoa haettaessa, esimerkiksi ihmisten mielipiteitä ja kokemuksia jostakin asiasta.

Seuraava väittämä: Hakukoneiden tulevaisuus on mobiililaitteissa toimiva johonkin tiettyyn aihepiiriin erikoistunut hakusovellus. Rungas enemmistö vastaajista oli väittämän kanssa eri mieltä (Kuvio 7).



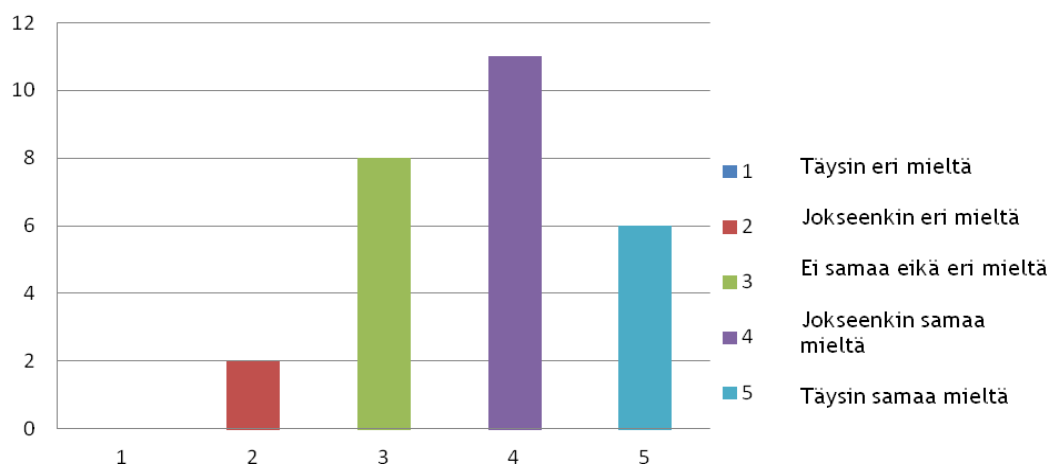
Kuvio 7. Hakukoneiden tulevaisuus on mobiililaitteissa toimiva johonkin tiettyyn aihepiiriin erikoistunut hakusovellus.

Tuloksen perusteella erikoishakusovellukset eivät tule syrjäyttämään perinteistä hakukonetta, vaan jatkossakin tietoa haetaan yleishakukoneella, kuten Googlessa.

Kahdeksannessa väittämässä kysyttiin asiantuntijoiden mielipidettä kysymykseen siitä, säästävätkö mobiililaitteiden hakusovellukset käyttäjänsä aikaa. Asiantuntijat olivat paria

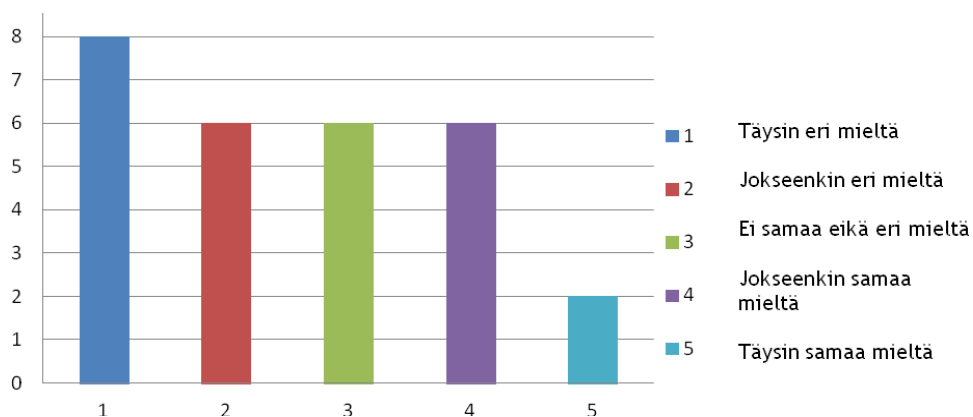
poikkeusta lukuunottamatta lähes yksimielisiä siitä, että yllämainitut sovellukset todellakin säästävät käyttäjiensä aikaa (Kuvio 8).

Sovelluksista on hyvänä esimerkkinä älypuhelimien Junat-sovellus, jolla voi hakea vaikka Helsingistä Purolaan kulkevien junien aikataulut. Tai vaikkapa erään asiantuntijan mainitsema hintatason tarkistaminen kaupassa asioidessa.



Kuvio 8. Mobiililaitteen hakusovellukset säästävät käyttäjänsä aikaa.

Yhdeksäs väittämä jakoi asiantuntijoiden mielipiteet. Niukan enemmistön mielestä hakukoneita käytetään jatkossakin myös näppäimistöyötteellä (Kuvio 9).

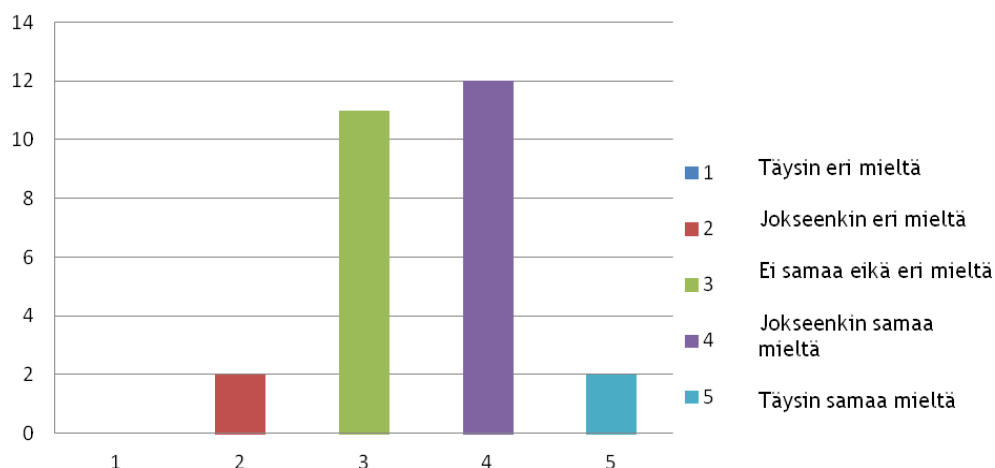


Kuvio 9. Tulevaisuudessa hakukoneita ei enää käytetä näppäimistöyötteellä lainkaan, vaan esimerkiksi puhumalla.

Kuitenkin jatkossa tulemme todennäköisesti näkemään erilaisia hakukäyttöliittymiä, joista yhtenä esimerkkinä Applen puhelimissa oleva Siri. Lisäksi myös kuva- tai ääninäytteellä tehtävät haut laajemmassa mittakaavassa ovat vain ajan kysymys. Tällä hetkellä näitä hakuja voi suorittaa vain rajatussa kontekstissa, esim. Shazamilla radiossa soivan kappaleen tunnistaminen. Muuan asiantuntijoista kirjoitti tähän väittämään lisätietoa vastaukseensa

seuraavanlaisesti: ”Niin kauan kuin aikaa menee enemmän väärinymmärrysten korjaukseen puheen tulkinassa, pysytään kirjoituksessa.”

Viimeiseksi väitettiin strukturoidun haun säästävän aikaa verrattuna Googlella haluamansa tiedon umpimähkäiseen löytämiseen. Enemmistö asiantuntijavastaajista oli väittämän kanssa jokseenkin samaa mieltä (Kuvio 10).



Kuvio 10. Strukturoitu haku säästää aikaa verrattuna umpimähkäiseen tiedon etsimiseen Googlella.

Kuitenkin usein koettiin strukturoidun haun soveltuvan parhaiten vain tietynlaisiin sovellutuksiin, kuten yliopistojen tutkimuksiin. Suurin osa hauista tehdään yhä periaatteella, jossa etsitään tiettyä sivua, jolla ovat haussa käytetyt sanat. Tässä tapauksessa ei rakenteista juuri olisi hyötyä.

## 6 Oma näkemys

Oman näkemykseni mukaan Google tulee tulevaisuudessa jäämään yhä enemmän yleishakuteoksen rooliin, kuten aikanaan tapahtui kirjastoille ja ensyklopediateoksille Googlen vallatessa alaa. Tietoa haetaan jatkossa yhä enemmän kännykällä, kuten junamatkoista kyseiseen tarkoitukseen räätälöidyllä hakusovelluksella. Kyse on siis mobiililaitteisiin sovelletuista erikoishakukoneista. Todennäköisesti Google onnistuu kuitenkin pitämään itsensä ajan tasalla yritysostoilla.

Hakuja tehdään yhä vähemmän hakusanoilla. Sen sijaan hakuprosessissa saatetaan käyttää apuna kuvia, videoita ja ääninäytteitä, joita saadaan helposti jaettua nykyaikaisista kannettavista älylaitteista. Vaikka lintubongarin linnun tunnistus helpottuisi huomattavasti, jos tuntemattomasta linnusta voisi vain näpätä kuvan ja lähettää sen hakupalveluun, josta saisi takaisin linnun lajinimen ja muuta olennaista faktatietoutta kyseisestä linnusta.

## 6.1 Yksityisyys

Ongelmia yksityisyyden suhteen muodostaa mielestäni paikannustietojen käyttö paikallisiin hakuihin. Mikäli haun tarjoaja tai hakuohjelman valmistaja kerää itselleen statistiikkaa paikannustietoihin liittyen, tulisi sen tapahtua hyvän tavan mukaisesti eli paikannustiedot tulisi salata ja yksilöivät tiedot tulisi jättää keräämättä. Varsin usein tällaiset tiedot tulitaisiin lähettämään mobiiliverkon ylitse, jolloin tietoliikenteen salausta olisi vaikea toteuttaa, ellei jopa mahdotonta.

Kaikkien yksilöivien tietojen keräämättä jättäminen kuitenkin huonontaa hakutulosten relevanssia käyttäjälle. Tulisi siis miettiä, mitä tietoa on tarpeen kerätä hakukokemusta parantamaan ja mitä ei.

## 6.2 Käyttäjäkokemus

Mielestäni tulevaisuuden hakukonetta käyttäjä käyttäisi mobiililaitteellaan, joko tablet-tietokoneella tai älypuhelimella. Sovellus ladataan kyseisen puhelinmallin käyttöjärjestelmän valmistajan kaupasta. Ensiksi halutaan esimerkiksi hakea tieto, millä bussilinjalla pääsee Helsingin rautatieasemalle Jätkäsaaresta. Käyttäjä avaa Kyyti-sovelluksen ja saa eteensä karttanäkymän, josta näkyvät mahdolliset pysäkit, niillä kulkevat linja-autot ja raitiotievaunut.

Päästyään rautatieasemalle hän avaa Junat-sovelluksen, josta hakee itselleen sopivan junan, jolla pääsee Järvenpäähän ja ostaa lipun vaivattomasti internetin kautta kännykkäänsä. Rautatieasemalla aikaa viettäessään käyttäjä kuulee radiosta tulevan mielenkiintoisen kappaaleen. Hän ei kuitenkaan tunnista kappaletta, joten avaa Shazam-sovelluksen kännykästään ja lähettää ääninäytteen verrattavaksi heidän tietokantoihinsa. Pian hän saakin vastauksena kappaaleen nimen ja esittäjän. Päällimmäisenä käyttäjälle jää mieleen hakuprosessin vaivattomuus ja aikaasäästävyys verrattuna perinteiseen mediaan (bussiaikataulut sekä juna-aikataulut).

## 7 Yhteenveto ja johtopäätelmät

Kyselyn lopputuloksista ja muusta kerätystä aineistosta päätellen Google tuskin lienee tulevaisuudessakaan syrjäytymässä Internet-hakukoneiden valtaistuimelta. Johtava markkina-asema, luovuutta kahlitsemaan yrityskulttuuri sekä lähes rajattomat resurssit mahdollistavat Googlen pysymisen hakukoneiden kehityksen aallonharjalla.

Lisäksi voidaan todeta, että perinteinen hakukone ei ole katoamassa mihinkään aivan lähitulevaisuudessa, ainakaan niin kauan kuin muut tavat kertoa hakukoneelle haettava tieto ovat parhaimmillaankin epäluotettavia. Nämä tavat syöttää hakukoneelle haettava tieto, kuten

vaikkapa ääninäyte ja kuvat, ovat parhaimmillaan rajatuissa konteksteissa, kun hakukoneella ei ole suuria mahdollisuuksia ymmärtää hakijan tarkoitusta väärin.

Sen sijaan semantiikan merkitys tulee todennäköisesti tulevaisuudessa hakukoneissa korostumaan, kuten viime aikoina olemme jo Googlen osalta nähneetkin. Tämä vaatii tosin vielä paljon käsien tehtävää metatiedon syöttämistä ja asiayhteyksien luomista kunnes tämä voidaan luotettavasti automatisoida.

Mobiililaitteiden hakusovellukset tulevat myös todennäköisesti pysymään omissa rajatuissa aihepiireissään eivätkä ole syrjäyttämässä vanhaa hakukonemallia. Sen sijaan Googlen varsinaista hakua tullaan käyttämään yhä enemmän, taulutietokoneilla ja älypuhelimilla varsinaisten pöytäkoneiden ja kannettavien tietokoneiden sijasta.

## Lähteet

Battelle, J. 2005. The Search - How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture 3. painos. Lontoo : Nicholas Brealey Publishing.

Engadget 2010. 3G GSM encryption cracked in less than two hours, Viitattu 3.12.2011.  
<http://www.engadget.com/2010/01/15/3g-gsm-encryption-cracked-in-less-than-two-hours/>

Forbes 2012. Bing's market share squeaks higher, Google still gorilla of search, Viitattu 23.5.2012. <http://www.forbes.com/sites/greatspeculations/2012/01/17/bings-market-share-squeaks-higher-google-still-gorilla-of-search/>

Gizmodo 2011. Google's Eric Schmidt is scared of Siri, Viitattu 3.12.2011.  
<http://gizmodo.com/5856834/googles-eric-schmidt-is-scared-of-siri>

Google Official Blog 2008. The future of search, Viitattu 10.1.2012.  
<http://googleblog.blogspot.com/2008/09/future-of-search.html>

Korpela, J. 2011. Web alkaa ymmärtää merkityksiä. MikroPC 6-7/2011.

Pandia Search & Social, Top 5 semantic search engines, Viitattu 3.3.2012.  
<http://www.pandia.com/sew/1262-top-5-semantic-search-engines.html>

Qwiki 2012. Qwiki.com, Viitattu 15.1.2012. <http://www.qwiki.com>

Search Engine Journal 2010. Google buys social search engine Aardvark, Viitattu 3.3.2012.  
<http://www.searchenginejournal.com/google-buys-social-search-engine-aardvark/17637/>

Search Engine Land 2010. By the numbers: Twitter vs. Facebook vs. Google Buzz, Viitattu 23.5.2012. <http://searchengineland.com/by-the-numbers-twitter-vs-facebook-vs-google-buzz-36709>

Search Engine Land 2010. John Battelle on the future of search, Viitattu 18.9.2011.  
<http://searchengineland.com/john-battelle-on-the-future-of-search-38382>

Search Engine Land 2010. John Battelle on the future of search part two, Viitattu 18.9.2011.  
<http://searchengineland.com/john-battelle-on-the-future-of-search-part-two-40094>

Stross, R. 2008 Planet Google. New York : Free Press.

TUAW 2011. Eric Schmidt says Siri poses competitive threat to google, Viitattu 3.12.2011.  
<http://www.tuaw.com/2011/11/07/eric-schmidt-says-siri-poses-competitive-threat-to-google/>

Unknown Pixels 3.0 2011. Qwiki - täysin uudenlainen hakukone, Viitattu 15.1.2012.  
<http://unknownpixels.com/blogi.php?kategoria=sekalaiset&id=94>

Wikipedia 2012. Aardvark (search engine), Viitattu 23.4.2012.  
[http://en.wikipedia.org/wiki/Aardvark\\_\(search\\_engine\)](http://en.wikipedia.org/wiki/Aardvark_(search_engine))

Wikipedia 2011. Hakukone. Viitattu 10.11.2011. <http://fi.wikipedia.org/wiki/Hakukone>

Wikipedia 2011. Search engine, Viitattu 10.11.2011.  
[http://en.wikipedia.org/wiki/Search\\_engine](http://en.wikipedia.org/wiki/Search_engine)

Wikipedia 2012. Siri (software), Viitattu 3.4.2012.  
[http://en.wikipedia.org/wiki/Siri\\_\(software\)](http://en.wikipedia.org/wiki/Siri_(software))

Wikipedia 2011. Web search engine, Viitattu 10.11.2011.  
[http://en.wikipedia.org/wiki/Web\\_search\\_engine](http://en.wikipedia.org/wiki/Web_search_engine)



## Liitteet

|   |    |
|---|----|
| Liite 1 Internet-hakukoneiden tulevaisuus-kyselyn vastaukset excel-muodossa. .... | 26 |
| Liite 2 Kyselyyn vastanneiden kommentteja väittämiin. ....                        | 27 |

Internet-hakukoneiden tulevaisuus-kyselyn vastaukset excel-muodossa.

| Tall.id | Tallennus   | Google ja | Paikannus | Ainut haku | Hakukone | Semanttin | Sosiaalinen | Hakukone | Mobiililaitte | Tulevaisuus | Strukturoit |
|---------|-------------|-----------|-----------|------------|----------|-----------|-------------|----------|---------------|-------------|-------------|
| 1       | 5.3.12.11.1 | 1         | 2         | 1          | 4        | 2         | 2           | 2        | 3             | 2           | 4           |
| 2       | 13.3.12.12  | 4         | 5         | 3          | 3        | 1         | 2           | 1        | 4             | 2           | 4           |
| 3       | 13.3.12.12  | 4         | 4         | 4          | 3        | 2         | 3           | 2        | 4             | 2           | 4           |
| 4       | 13.3.12.12  | 1         | 5         | 1          | 4        | 2         | 2           | 1        | 2             | 1           | 2           |
| 5       | 13.3.12.12  | 1         | 1         | 1          | 1        | 2         | 3           | 2        | 5             | 2           | 4           |
| 6       | 13.3.12.12  | 3         | 4         | 4          | 2        | 1         | 3           | 2        | 3             | 2           | 2           |
| 7       | 13.3.12.13  | 4         | 3         | 1          | 2        | 2         | 3           | 1        | 5             | 5           | 4           |
| 8       | 13.3.12.13  | 4         | 2         | 2          | 3        | 1         | 5           | 1        | 4             | 4           | 5           |
| 9       | 13.3.12.13  | 2         | 2         | 3          | 5        | 2         | 1           | 2        | 2             | 1           | 3           |
| 10      | 13.3.12.13  | 4         | 3         | 3          | 4        | 4         | 1           | 2        | 3             | 1           | 3           |
| 11      | 13.3.12.13  | 1         | 5         | 3          | 4        | 3         | 3           | 1        | 3             | 1           | 3           |
| 12      | 13.3.12.13  | 2         | 3         | 2          | 2        | 2         | 2           | 2        | 3             | 4           | 4           |
| 13      | 13.3.12.13  | 1         | 1         | 2          | 2        | 2         | 2           | 1        | 3             | 1           | 3           |
| 14      | 13.3.12.13  | 2         | 3         | 4          | 5        | 2         | 3           | 2        | 5             | 3           | 3           |
| 15      | 13.3.12.13  | 4         | 4         | 3          | 4        | 3         | 4           | 2        | 4             | 2           | 3           |
| 16      | 13.3.12.14  | 4         | 2         | 4          | 3        | 2         | 1           | 2        | 3             | 4           | 4           |
| 17      | 13.3.12.14  | 2         | 5         | 1          | 4        | 4         | 3           | 3        | 5             | 3           | 3           |
| 18      | 13.3.12.14  | 1         | 4         | 3          | 4        | 2         | 3           | 1        | 5             | 4           | 4           |
| 19      | 13.3.12.15  | 2         | 4         | 2          |          |           |             | 2        |               | 3           | 3           |
| 20      | 13.3.12.17  | 2         | 4         | 2          | 4        | 3         | 2           | 2        | 4             | 3           | 3           |
| 21      | 13.3.12.21  | 3         | 3         | 1          | 2        | 3         | 3           | 3        | 4             | 1           | 3           |
| 22      | 14.3.12.10  | 5         | 5         | 4          | 4        | 2         | 3           | 4        | 4             | 5           | 5           |
| 23      | 15.3.12.09  | 3         | 5         | 1          | 3        | 2         | 2           | 1        | 4             | 4           | 3           |
| 24      | 15.3.12.14  | 4         | 4         | 2          | 4        | 3         | 2           | 2        | 4             | 1           | 4           |
| 25      | 16.3.12.08  | 2         | 2         | 4          | 2        | 2         | 2           | 4        | 5             | 4           | 4           |
| 26      | 19.3.12.09  | 3         | 5         | 5          | 4        | 5         | 1           | 1        | 4             | 3           | 4           |
| 27      | 12.2.12.13  | 2         | 2         | 4          | 4        | 1         | 3           | 1        | 4             | 1           |             |
| 28      | 12.2.12.16  | 5         | 2         | 4          | 3        | 3         | 3           | 3        | 3             | 3           | 4           |

Kyselyyn vastanneiden kommentteja väittämiin.

Tulevaisuudessa hakukoneissa tulee olemaan enemmän visualisointeja, joita seuraamalla voidaan pudottaa tai valita suuntauksia. Käyttöliittymät tulevat muuttumaan merkittävästi nykyisesti, ihminen on hyvä tunnistamaan asioita kuvista ja sitä tullaan hyödyntämään enemmän.

Lisäksi haluaisin sanoa, että "sosiaalinen" hakukone on huonointa, mitä hakukoneille on tapahtunut/tapahtumassa. Sosiaalisten sivustojen vaikutus pilaa hakutulosten objektiivisuuden. Miksi ihmeessä minun näkemiäni hakutuloksia pitäisi manipuloida sillä perusteella, että joku kaveri on käyttänyt/jakanut tiettyä linkkiä sosiaalisessa mediassa? Järkyttävää väärinkäyttöä. Semanttisesta hausta sanoisin sen verran, että idea on hyvä, mutta ei tule toimimaan, mikäli sivustojen tekijöiden täytyy käyttää valtava työmäärä sisällön "tagittamiseen" manuaalisesti.

Mikä on strukturoitu haku? Entä umpimähkäinen? Miten määrittelet nämä? Mikä katsotaan semantiikan soveltamiseksi? Pelkästään sosiaalinen hakukone - vai myös muita tekniikoita sovellettu? Jne. Hieman hankala vastata, kun ei tiedä, mitä aidosti tarkoitetaan, kun kaikki ymmärtävät/määrittelevät asioita hieman eri lailla. Mutta. Sosiaalisuus (social graph etc.) tulee ilman muuta enenevissä määrin osaksi hakupalveluita - esim. Facebookhan on jo merkittävä pelaaja alalla. Mutta tuskin yksinään, vaan osana kokonaisuutta (ts. "Sosiaalinen hakukone" ei taida kuitenkaan olla pelkästään ja vain ja ainoastaan sosiaalisuuten pohjautuva...) Mutta jo mainonnallisistakin syistä - kun raha hommaa kuitenkin ajaa - sosiaaliset aspektit varmasti halutaan mukaan ja talteen. Yksityisyys? Buy-buy privacy (tarkoituksellinen väärintavaus...): Paremmaksi ainakin väitetyllä palvelulla tullaan viemaan vähäinkin yksityisyys ja tekemään kaikesta kaupallista dataa, jolla meitä luodataan ja meille myydään. Mistä hakukoneet saavat rahansa? Eivät hakujen tekijöiltä. Joten niiden on palveltava tahoja, joista raha tulee - sen ohella tietenkin, että niiden on tuotettava hyötyä käyttäjilleen, että käyttäjät eivät mene toiseen palveluun (joka sitten puolestaan myy ko. dataa). Follow the money trail - vielä ei taida olla yhtään kunnollista hakukonetta, jonka toiminta perustuisi hyvän tekemiseen eikä voiton tavoitteluun. Suosittelujärjestelmät - joista tässä ei kysytä ollenkaan!?! - tulevat integroimaan sosiaalisia aspekteja enenevissä määrin ja ne tulevat integroitumaan yhä enemmän hakupalveluihin. Näinhän on jo nyt ja tämä kehitys tulee vaan jatkumaan. Aiheesta hieman vanha mutta yhä loistava artikkeli: <http://www.mobilegroove.com/guest-column-what-is-the-recommender-industry/> - hakemisella on osuutensa/paikkansa, mutta discovery/serendipity... Now that's fun.

Yleishakukoneella on vankka käyttäjäpohja ja tämä kehityspolku näyttää jatkuvan. Muiden hakukoneiden on vaikea syrjäyttää perinteistä, ehkä tietyissä käyttäjäryhmissä tai yliopistoissa, joissa tutkimuksessa strukturoidulla haulilla olisi paikkansa.

Uskon, että Google on valmis kehittymään jatkossakin eri suuntiin, niin kuin se on tälläkin hetkellä tehnyt. Sillä on resursseja tehdä tutkimusta ja laajentaa toiminta-alueitaan myös hakukentällä lisäämällä aina uusia ja uusia tapoja hakea informaatio. Sen vuoksi en usko, että Google jää kovin marginaaliseksi toimijaksi, vaikka kilpailijoitakin tulee varmasti jatkossa olemaan ainakin erikoistuneemmissa hakukoneratkaisuissa. Yksityisyydensuoja on aina ongelma hakukoneiden käytössä. Käyttäjien tietojen kerääminen on hyvä tapa tehdä rahaa mainostarkoitukseen, mutta hakukoneita käyttämällä ihmiset paljastavat myös paljon itsestään hakupalvelun tarjoajalle. Tällä hetkellä yleisessä tekstihaussa Google on erittäin hyvä. Kuitenkin tietoa voi hakea myös muulla tavalla, kuten ääni- tai kuvanäytteillä. Mobiililaitteiden tallennusominaisuuksien vain kehittyessä, tämä tulee olemaan varsin todennäköinen kehityssuunta myös tulevaisuudessa. Tälläkin hetkellä esimerkiksi Tineye-palvelu tarjoaa Internetissä mahdollisuuden hakea samankaltaisia kuvia esimerkkikuvilla. Tulevaisuudessa todennäköisesti tarjolla on myös palveluja hakea erilaista informaatiota kuvien kohteista. Tällä hetkellä osa sosiaalisista järjestelmistä tarjoavat mahdollisuuden annotoida kuvia lisäinformaatiolla esimerkiksi löydettyjen kasvojen identiteeteistä. Tätä tekniikkaa on mahdollista laajentaa myös muihin kohteisiin. Esimerkiksi hintavertailu tai muuta informaatiota mistä tahansa kiinnostavasta tuotteesta, kappaleesta tai kohteesta. Tällaisen datan automaattinen keruu ja assosiaatioiden luominen eri asioiden välille on vain hankalaa. Lisäksi skaalautuvuus ja erilaisten luokittelujen oikeellisuus muodostaa haasteita tällaisten palveluiden rakentamiselle. Syöte hakukoneille voi siis olla jotain muutakin kuin tekstiä. Jos tekstisisällön saa riittävän luotettavasti ja nopeammin puhuttuna, puheeseen tullaan varmaankin siirtymään. Niin kauan kuin aikaa menee enemmän väärinymmärrysten korjaukseen puheen tulkinassa, pysytään kirjoituksessa.

Tulevaisuudessa sanan hakukone merkitys hämärtyy. Keinotekoinen, biologisten aivojen ulkopuolella tapahtuva päättely ja muistaminen lisääntyy, ja ihmisten ja koneiden kollektiivinen toiminta saa aivan uusia ulottuvuuksia. Tietoa ei pelkästään haeta, vaan sitä myös tallennetaan ja muokataan.

Käsittäakseni Google on yrittänyt parantaa hakutuloksia viime aikoina sosiaalisen verkoston perusteella (?). Olen kokenut, että hakutulokseni ovat huonontuneet viime aikoina. Ensimmäistä kertaa pitkään aikaan olen harkinnut hakukoneen vaihtoa.

Yksi kehityssuunta diskreeteistä hauista olisi myös eräänlainen jatkuva haku mobiililaitteen kautta, jossa laite olisi tietoinen päivän agendasta esim. kalenterin kautta, ja tarjoaisi hakutuloksia paikkatietoon ja kalenteriin liittyen, ilman erillistä hakupyynnöä. Tässä voisi hyödyn-

tää esim. augmented reality -tekniikoita. Tosin tähän suuntaan liittyy myös isoja haasteita erityisesti käyttöiittymäsuunnittelun kannalta.

Kommentteja väittämiin 1) Ei Google ole mikään "yleishakuteos" eikä pyri sellaiseksi, siksi tähän väittämään ei oikeastaan voi vastata. 2) Totta. Mutta kenelle? 3) Mitä tällä tarkoitetaan? "Olla merkitystä" == toimii parhaiten? 4) Semanttiset hakukoneet ovat olleet tulemaisillaan viimeiset 20 vuotta. Tilanne on todennäköisestä sama vielä 20 vuoden päästäkin. 5) No ei, katso edellinen kommentti. 6) "Sosiaalinen hakukone" == "kilauta kaverille"? 7) Ei, kyllä se Google sielläkin jyrää. 8) Kyllä, voi kaupassa googlata oikean hintatason. 9) Onhan tämä mahdollista, mutta ei kovin todennäköistä. 10) Riippuu hakutehtävästä. Ylivoimaisesti tavallisin tarve on "näytä ne sivut, joissa esiintyy nämä sanat". Siinä ei rakenteista ole kovin paljon iloa.