

Erityisryhmille soveltuvan suomenkielisen tiedonhakupalvelun toteuttaminen -projekti

4.5.2010

Kimmo Kyyhkynen

IT-suunnittelija

Papunet-verkkopalveluyksikkö

Kehitysvammaliitto ry

Tiivistelmä

Kehitysvammaliiton Papunet-verkkopalveluysikössä toteutettiin vuosien 2008-2009 aikana ”Eri-tyisryhmille soveltuvan suomenkielisen tiedonhakupalvelun toteuttaminen” -projekti. Projektissa kartoitettiin selkokielen käyttäjien tarpeita ja mahdollisia ongelmakohtia internetin hakukoneiden käytössä. Kartoituksen perusteella suunniteltiin näiden käyttäjien tarpeita vastaava hakukoneen käyttöliittymäprototyyppi.

Projektin päämääränä oli parantaa erityisryhmien tietoyhteiskuntavalmiuksia ja -taitoja kehittämällä tehokas tiedonhaun väline. Pääasialliset tavoitteet olivat tiedon kerääminen erityisryhmien hakukäyttäytymisestä ja kerätyn tiedon soveltaminen helppokäyttöisen hakukäyttöliittymän toteuttamisessa.

Projektin kohderyhmänä olivat lukutaitoiset puhe- ja kehitysvammaiset ihmiset, joilla on riittävät tietotekniset taidot. Projektin tuloksia voi helposti hyödyntää tulevaisuudessa myös monet muut tiedonhakijat, mm. erilaiset oppijat, ikäihmiset ja maahanmuuttajat. Näissä ryhmissä on selkokielen käyttäjiä tai he hyötyvät selkokielestä. Tästä johtuu projektin työnimi ”Selkohaku”.

Ensimmäisen projektivuoden aikana kartoitettiin taustatietoa aikaisempaa tutkimusmateriaalia vertailemalla ja kerättiin käyttäjätietoa projektin kohderyhmien tiedonhaun tavoista seuraamalla kohderyhmien edustajien toimintaa olemassa olevissa hakupalveluissa, esimerkiksi Googlessa.

Toisen projektivuoden aikana kerättyä tutkimusaineistoa analysoitiin ja sen perusteella tehtiin erilaisia käyttöliittymäprototyyppiejä. Prototyyppien käytettävyyttä arvioitiin asiantuntija-arvioilla ja käyttäjäryhmien edustajien toimesta suoritettuihin käyttäjätesteihin.

Projektissa kehitettiin helppokäyttöinen tiedonhaun käyttöliittymä. Projektin tuotoksia ja havaintoja on hyödynnetty www.selko.fi-sivustolla suomenkielisen internetin selkosivustojen sisällöt kattavassa haussa.

Projekti toteutettiin yhteistyössä HIITin (Helsinki Institute for Information Technology), Helsingin kauppakorkeakoulun CKIR:n (Center for Knowledge and Innovation Research), BitLips oy:n, Espoon Eestinmetsän koulun, Ikäihmisten tietotekniikkayhdistys Enter ry:n, Helsingin Diakonissalaitoksen säätiön Diakoniaopiston kanssa, toimintakeskus Cäpsän ja Vallilan työkeskus Valtterin kanssa. Projektin rahoitti Raha-automaattiyhdistys.

Sisällys

Tästä dokumentista.....	5
Projektin lähtökohdat.....	6
Projektin toteutus.....	7
Projektin tulokset.....	12
Pohdintaa.....	14

LIITE 1: Käyttöliittymäongelmien analyysi
LIITE 2: Luettelo käytettävyydestien osallistujista
LIITE 3: Luettelo testitehtävistä

Tästä dokumentista

Tämä on "Erityisryhmille soveltuvan suomenkielisen tiedonhakupalvelun toteuttaminen" -projektin loppuraportti, jossa käydään läpi projektin tavoitteet, projektin kulku sekä sen keskeisimmät tulokset ja muut suoritteet. Raportin selkeyttämiseksi olemassa olevista hakukoneista löydettyjen käyttöliittymäongelmien analyysi (Liite 1) on erotettu omaksi dokumentikseen. Lisäksi liitteenä on luettelo kaikista käytettävyydesteihin osallistuneista henkilöistä (Liite 2) sekä luettelo käytettävyydestien testitehtävistä (Liite 3).

Dokumentissa käytettyjä termejä

Papunet	Papunet-termillä viitataan tässä dokumentissa Kehitysvammaliiton Papunet-verkkopalveluysikköön, jonka päätehtävä on kehittää ja toteuttaa esteettömiä verkkopalveluita puhe- ja kehitysvammaisille ihmisille, selkokielen käyttäjille, heidän lähi-ihmisiilleen sekä alan ammattilaisille ja opiskelijoille.
Selkokieli	Selkokieli on sisällöltään, sanastoltaan ja rakenteeltaan yleiskieltä luettavammaksi ja ymmärrettävämmäksi mukautettua kieltä niitä ihmisiä varten, joilla on vaikeuksia lukemisessa ja ymmärtämisessä (tai molemmissa). Suomessa selkokielen kehittämisestä ja kirjoittajien koulutuksesta vastaa Kehitysvammaliiton Selkokeskus.
Hakukone	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, joita kehitetään jatkuvasti paremmiksi ja toimivammiksi.
Käyttöliittymä	Käyttöliittymä on se laitteen, ohjelmiston tai minkä tahansa muun tuotteen osa, jonka kautta käyttäjä käyttää tuotetta. Esimerkiksi tietokoneohjelmassa käyttöliittymä tarkoittaa sitä ohjelman osaa, jonka käyttäjä näkee tietokoneen näytöllä, ja sitä tapaa (hiiri, näppäimistö), jolla hän käyttää ohjelmaa.
Käytettävyys	Käytettävyydellä tarkoitetaan tuotteen sopivuutta aiottuihin tehtäviin aiotuilla käyttäjillä: laitteen käyttäjäryhmän tulisi pystyä käyttämään tuotetta tehokkaasti, tuottavasti ja miel lyttävästi tiettyssä käyttöympäristössä ja tilanteessa ¹ . Toinen usein käytetty määritelmä ² määrittelee käytettävyyden viiden mitattavan ominaisuuden avulla: opittavuus, muistettavuus, tehokkuus, virhetilanteiden vähyys ja käyttäjän tyytyväisyys. Käytettävyyden arviointimenetelmillä pyritään löytämään jonkinlainen arvio näille ominaisuuksille.

¹ ISO 9241-11: Guidance on usability. (1998)

² Nielsen, J. (1993) Usability Engineering. Academic Press. USA

Projektin lähtökohdat

Projektin taustaa

Tiedonhaku on keskeinen tietoverkkojen käyttötapana; tieto on arvotonta, ellei se ole helposti saatavilla ja esitettävissä käyttäjän ymmärtämässä muodossa. Nykyiset tiedonhakumenetelmät on suunniteltu käyttäjille, joilla on tavanomaiset motoriset ja kognitiiviset taidot, ja joiden hahmottamis- ja ymmärtämiskyky on normaali. Tietomassan hallinta on hankalaa käyttäjille, joilla ei ole vastaavia tietoja ja taitoja. Tietoyhteiskuntavalmiudet vaikuttavat olennaisesti mahdollisuuksiin löytää ja käyttää yhteiskunnan tarjoamia palveluja sekä mahdollisuuksiin saada koulutusta ja työtä. Perinteisiä hakukoneita hyödyntävä tiedonhaku on entistä vaikeampaa erityisryhmille, joiden taidot eivät ole riittävät hakurajausten tekemiseen. Hakutulosten esittämistapa ei myöskään tue oppimista. Tämän vuoksi onkin tärkeää kehittää tiedonhaun mahdollisuuksia, jotka huomioivat erilaisten käyttäjien kyvyt kehittyä tiedon hakemisessa ja tulosten hyödyntämisessä.

Projektin tavoitteet

Projektin valmisteluvaiheessa tavoitteeksi asetettiin rakentaa helppokäyttöinen tiedonhakupalvelu, joka perustuisi kotimaiseen hakuteknologian perustutkimukseen. HIITissä kehitetty teknologia eroaa perusperiaatteiltaan nykyisistä avainsanapohjaisista hakumenetelmistä (esim. Google). Projektissa suunniteltaisiin helppokäyttöinen käyttöliittymä, jota käytettäisiin esimerkiksi selkokielen hakupalvelun, Kehitysvammaliiton verkkosivustot kattavan haun sekä Wikipedian ja YLEn aineistot kattavan haun toteutuksessa.

Projektille määriteltiin lisäksi seuraavat osatavoitteet:

1. Käyttäjätiedon keruu ja käyttöympäristöjen määrittely. Projektissa kerätään tietoa kohde-ryhmän tiedonhakutavoista käyttäjätutkimuksilla haastatteluin ja havainnoinein. Tieto julkaistaan Papunet-verkkopalvelussa.
2. Tiedonhakupalvelun käyttöliittymän käyttäjälähtöinen suunnittelu kerätyn käyttäjätiedon perusteella käytettävyytutkimusten avulla.
3. Suomessa kehitetyn avoimen lähdekoodin hakuteknologian mukauttaminen suunnitellun käyttöliittymän tarpeisiin.
4. Kehitysvammaliiton verkkosivustojen (Vernerin, Papunet, Kehitysvammaliitto), suomenkielisen Wikipedian ja YLEn aineistoja käyttävän tiedonhakupalvelun toteutus yhteistyössä loppukäyttäjien kanssa.
5. Uuden tiedonhakupalvelun lanseeraus ja ylläpito.

Projektin toteutus

Projekti toteutettiin kahden rinnakkaisen osaprojektin kautta. Nämä olivat käytettävyystiedon kerääminen ja sen soveltaminen käyttöliittymän kehittämisessä sekä hakukoneteknologian kehittäminen. Osaprojektien vastualueet jaettiin niin, että HIIT oli vastuussa kehittämänsä hakukoneteknologian jatkokehittämisestä projektin tarpeisiin vastaavaksi ja Papunet keskittyi käytettävyystiedon keräämiseen ja käyttöliittymäsuunnitteluun. Projektin loppuvaiheessa osaprojektit yhdistettiin valmiin hakukoneen muodossa.

Projektin organisaatio ja henkilöstö

Projektin päävastuu oli Kehitysvammaliiton Papunet-verkkopalveluyksiköllä, jossa projektin johtajana toimi yksikön johtaja Marianna Ohtonen ja projektisuunnittelijana IT-suunnittelija Kimmo Kyyhkynen.

Projektissa pääyhteistyökumppanina toimi Helsingin yliopiston tietotekniikan tutkimuslaitos HIIT³in CoSCo⁴-tutkimusryhmä. Lisäksi valmistelutyössä oli mukana Bitlips oy.

Käyttäjätesteissä yhteistyötä testikäyttäjryhmien hankkimisessa ja testien toteuttamisessa tekivät Espoon Eestinmetsän koulu, Ikäihmisten tietotekniikkayhdistys Enter ry, Helsingin Diakonissalaitoksen Diakoniaopisto sekä Helsingin kaupungin toimintakeskus Cäpsä ja Vallilan työkeskus Valtteri.

Ohjausryhmä

Projektiä koordinoi ja seurasi yhteistyöorganisaatioista ja asiantuntijoista koottu ohjausryhmä, joka kokoontui projektin aikana kolme kertaa. Ohjausryhmän muodostivat Kehitysvammaliiton toiminnanjohtaja Veli-Pekka Sinervuo, yksikön johtaja Marianna Ohtonen sekä IT-suunnittelijat Kimmo Kyyhkynen ja Sami Älli Kehitysvammaliiton Papunet-verkkopalveluyksiköstä, Selkokeskuksen johtaja Hannu Virtanen, tutkija Hannu Vesala Kehitysvammaliiton tutkimus- ja kehittämisyksikkö Tutkesta, professori Petri Myllymäki ja tutkija Antti Tuominen HIITistä, toimitusjohtaja Jarmo Välikangas Bitlips oy:stä sekä tiiminjohtaja Ilpo Kojo Helsingin kauppakorkeakoulun CKIR⁵-yksiköstä.

Käytettävyystiedon keruu

Jo projektia valmisteltaessa oli selvää, että aikaisempaa, vapaasti saatavilla olevaa tutkimustietoa erityisryhmien hakukonekäyttäytymisestä tulisi löytymään hyvin vähän. Ylipäänsä puhuttaessa muiden selkokielen kohderyhmien kuin ikäihmisten verkkosivujen käytettävyyttä koskevista käytettävyystudkimuksista voidaan todeta, että Suomessa niitä on suoritettu pääasiassa Papunetin toimesta^{6,7}.

³ Helsinki Institute for Information Technology.

⁴ Complex Systems Computation Group.

⁵ Center for Knowledge and Innovation Research.

⁶ Kuusela, M. (2004) Selkokielisten Internet-sivujen käytettävyyks kehitysvammaisilla käyttäjillä. Diplomityö, Teknillinen korkeakoulu

⁷ Kyyhkynen, K. (2005) Selkokielisten verkkosivuston käytettävyys. Diplomityö, Teknillinen korkeakoulu.

Hakukoneiden käytettävyyttä ikäihmisillä sen sijaan on tutkittu niin muualla maailmassa kuin Suomeksi⁸ jonkin verran, ja näissä tutkimuksissa tehtyjä havaintoja voitaneen soveltaa myös muihin selkokielen kohderyhmiin.

Luonnollisesti alan suuret toimijat (Google, Microsoft, Yahoo) ovat käytettävyydestään omia hakukoneitaan runsaasti myös ikäihmisillä ja todennäköisesti myös muilla selkokielen kohderyhmillä, mutta ymmärrettävästi näiden testien tulokset eivät ole julkisesti saatavilla olevaa tietoa.

Käytettävyytustutkimukset

Koska valmista tietoa kohderyhmien hakukäyttäytymisestä ei esiselvityksissä löytynyt riittävästi, tieto oli hankittava projektin toimesta käytettävyytustutkimuksen keinoin. Tässä kontekstissa tämä tarkoitti sitä, että projektin kohderyhmään kuuluvien käyttäjien kanssa suoritettiin käytettävyystestejä olemassa olevilla hakukoneilla.

Käytettävyydestä on tyypillisimmillään tilanne, jossa määritelty käyttäjä testaa käytettävää tuotetta tilanteessa, joka muistuttaa tuotteen aitoa käyttötilannetta. Käyttäjä käyttää tuotetta itsenäisesti, joko oman mielensä mukaan (nk. vapaa läpikäynti) tai suorittaen samalla jotain ennalta määrättyä tehtävää.

Projektin yhteydessä toteutettiin käytettävyystestejä kahdessa eri yhteydessä; projektin alussa suoritettiin testejä olemassa olevilla hakukoneilla, jotta saatiin kerättyä tietoa niiden mahdollisista käytettävyysongelmista kohderyhmän käyttäjillä. Projektin lopussa suoritettiin testejä toteutetuilla käyttöliittymäprototyypeilla, jotta saatiin kerättyä tietoa niissä käytettyjen ratkaisuiden toimivuudesta kohderyhmän käyttäjillä.

Suurimmaksi haasteeksi käytettävyystesteissä osoittautui sopivien testiryhmien löytäminen; selkokielen kohderyhmiin kuuluvia käyttäjiä kyllä löytyi, mutta kun lisäksi vaatimuksena oli tietokoneenkäyttötaito, sopivat ryhmät vähenivät merkittävästi. Testeihin osallistuneilla aikuisilla kehitysvammaisilla ei ollut ollut koulutuksensa osana järjestelmällistä tietotekniikkaopetusta, kun taas testeihin osallistuneella nuorten kehitysvammaisten ryhmällä opiskeluun kuului säännöllisesti Tikas-opetusta⁹, mikä näkyi suoraan heidän tietokoneenkäyttötaidoissa. Tikas-opetukseen kuuluu osana myös hakukoneiden käytön harjoittelu.

Testeissä käytettäväksi olemassa oleviksi hakukoneiksi valikoitiin markkinajohtaja¹⁰ Googlen suomenkielinen hakukone osoitteessa www.google.fi ja täysin suomalainen Ihmemaan Haku, joka on sittemmin lopettanut toimintansa.

Käytettävyystestejä olemassa olevilla hakukoneilla suoritettiin kolmella eri käyttäjäryhmällä; kuuden kehitysvammaisen nuoren ryhmän, kolmen ikäihmisen ryhmän ja seitsemän kehitysvammaisen aikuisen ryhmän kanssa. Käyttöliittymäprototyyppejä testasi kaksi aikuisista kehitysvammaisista

⁸ Aula, A. (2005) Older adults' use of web and search engines. *Universal Access in the Information Society*, 4(1-2), 67-81. Springer-Verlag

⁹ Tikas on erityistä tukea tarvitsevien oppijoiden tieto- ja viestintätekniikkataitojen (TVT) opettamiseen kehitetty pedagoginen koulutusmalli, jonka tavoitteena on tarjota helpompi ja yksilöllisempi tapa oppia TVT-taitoja.

¹⁰ Net Applications Search Engine Market Share Q4/2009

muodostettua ryhmää. Näissä ryhmissä oli yhteensä 13 henkilöä. Kaikkien testihenkilöiden ikä ja sukupuoli on lueteltu erillisessä liitteessä (Liite 2).

Kaikille testihenkilöille lähetettiin ennakoon esitietokysely, jolla kartoitettiin henkilön tietoteknisiä taitoja ja henkilökohtaisia kiinnostuksen kohteita. Kiinnostuksen kohteita kysyttiin, jotta kunkin testitehtävät voitaisiin laatia niin, että tehtävien aiheet olisivat testihenkilöitä kiinnostavia ja tätä kautta enemmän motivoivia kuin satunnaisesti valitut tehtävät. Tyypillisesti käytettävyydestien testitehtävät sidotaan jonkinlaiseen kehystarinaan, joka antaa motivaation tehtävien suorittamiseen. Tällä kertaa motivaatiota yritettiin hakea kuitenkin toista kautta, koska hakukoneilla suoritettaviin hakuihin oli vaikea sovittaa kehystarinaa, ja Papunetin aikaisemmissa selkokielen käyttäjillä suoritetuissa käytettävyydesteissä kehystarinan on todettu hankaloittavan tehtävän ymmärtämistä.

Kukin testihenkilö sai suoritettavakseen kymmenen testitehtävää, jotka oli laadittu niin, että niistä 60 % oli "informatiivisia" tehtäviä. Näissä haettiin jotain spesifiä tietoa annetusta aiheesta. Loput olivat tehtäviä, joissa haettiin tietoa yleisemmin, ilman varsinaista "oikeaa" vastausta. Jaottelu perustuu tutkimukseen, jossa selvitettiin minkä tyyppisiin hakuihin käyttäjien tekemät haut yleisesti jakautuvat¹¹. Testitehtävät on lueteltu erillisessä liitteessä (Liite 3).

Testitilanteessa käyttäjät haastateltiin ennen testitehtävien suorittamista ja testitehtävien jälkeen. Lisäksi testitilanteesta saadut kokemukset purettiin lyhyesti heidän kanssa. Testitilanteet tallennettiin joko videoimalla tai nauhoittamalla käyttäjän puheääni ja näytön tapahtumat, jotta niitä voitiin analysoida uudelleen jälkepäin.

Käyttäjät saivat suorittaa testitehtävät omaan rauhalliseen tahtiinsa ilman, että testin valvoja puuttui tilanteeseen. Mikäli käyttäjä pyysi apua, sitä annettiin. Jos käyttäjä koki jonkun tehtävän niin vaikeaksi, ettei halunnut suorittaa sitä loppuun, sai tämä siirtyä vapaasti seuraavaan tehtävään.

Käyttöliittymä

Käytettävyydesteissä olemassa olevilla hakukoneilla ilmeni useita haasteita, joista osa liittyi suoraan käyttöliittymään ja osa tiedonhaun käsitteeseen yleisesti. Testeissä ilmeni, että mikäli käyttäjä ei ollut aikaisemmin juurikaan käyttänyt hakukoneita, ei niiden toimintaperiaate ollut kovin intuitiivinen, mikä ilmeni vaikeuksina keksiä sopivaa hakusanaa. Erityisen vaikeaksi havaittiin hakusanojen keksiminen silloin, jos ensimmäisellä hakutermillä ei löydetty haluttua tietoa, ja hakutermiin olisi pitänyt keksiä lisää sanoja tarkentamaan hakua.

Tähän ongelmaan ei ole olemassa mitään ilmeistä teknistä ratkaisua; käyttäjälle voisi toki tarjota yläkäsitteitä tai satunnaisia vaihtoehtoja, joista tämä voisi valita sen, mikä on lähinnä haettua tietoa, mutta riittävän laajan käsittehierarkian rakentaminen ja sen soveltaminen hakukoneessa olisi niin valtavasti resursseja vaativa tehtävä, ettei se ole käytännössä mahdollista toteuttaa kuin alan suurempien toimijoiden toimesta. Projektin puitteissa keskityttiin näin ollen ratkaisemaan ainoastaan ongelmat, jotka on mahdollista ratkaista käyttöliittymätasolla.

¹¹ Rose, D. E., Levinson, D. (2004) Understanding user goals in Web search. Proceedings of 13th International World Wide Web Conference - WWW 2004.

Käyttöliittymäongelmien ratkaisuvaihtoehtoja pohdittaessa vertailtiin myös suurimpien kaupallisten hakukoneiden (Google, Yahoo!, Microsoft Live Search) käyttöliittymiä keskenään ja yritettiin löytää paras käytössä oleva malli toteuttaa tietyt asiat.

Tarkempi kuvaus havaituista käyttöliittymäongelmista ja niihin kehitetyistä ratkaisuista löytyy erillisestä dokumentista (Liite 1).

Hakukoneteknologia

Projektin aikana HIIT jatkoi hakukoneteknologiansa kehittämistä edelleen valmiimpaan suuntaan mm. lisäämällä hakumoottoriin rajapinnan, jonka avulla sitä voitiin kokeilla yhdessä projektissa kehitetyn käyttöliittymän kanssa.

Lisäksi HIIT kehitti joitakin kokeellisia demokäyttöliittymiä, joissa tiedonhakua lähestyttiin perinteisestä eroavilla tavoilla. Näitä käyttöliittymiä ei kuitenkaan projektin puitteissa kehitetty sen pidemmälle, eikä niitä testattu oikeilla käyttäjillä.

Tarkempi kuvaus HIITin osuudesta hakuteknologian kehittämisessä kerrotaan seuraavassa, professori Petri Myllymäen kirjoittamassa osuudessa.

Tiedonhaku on osoittautunut yhdeksi keskeiseksi tietoverkkojen käyttötavaksi; tieto sinänsä on arvotonta, ellei se ole helposti saatavilla ja esitettävissä käyttäjän ymmärtävässä muodossa. Tiedonhaun tulokset voivat siis auttaa hakijaa kommunikoimaan maailman kanssa ja ymmärtämään sitä paremmin. Tiedonhakua varten tieto voidaan järjestellä semanttisesti loogisiin hierarkkisesti selailtaviin hakemistoihin tai yksittäisiin tietoalkioihin (dokumentteihin) voidaan liittää haussa käytettäviä avaimia (tags). Tietomäärän kasvaessa inhimilliset resurssit eivät kuitenkaan riitä tiedon manuaaliseen luokitteluun tai rikastamiseen. Tämä on johtanut uusien automaattisten keinojen etsintään tiedon hallitsemiseksi ja hakukoneiden (esim. Google) syntyyn; internetin dokumenttimassa käydään automaattisesti läpi, ja siitä muodostetaan ns. indeksi, jonka avulla voidaan tehdä avainsanahakuja dokumenteissa esiintyviin sanoihin perustuen.

Nykyisten hakukoneiden toimivuus perustuu laajalti siihen, että käyttäjä pystyy "arvaamaan" haettavassa dokumentissa esiintyviä sanoja ja kykenee muodostaman pienen joukon aihepiiriä hyvin kuvaavia avainsanoja. Tehtävä on käyttäjälle kuitenkin haastava; mikäli avainsanoja on liian monta, hakutuloksia on nolla kappaletta (yhdessäkään dokumentissa ei esiinny kaikki annettuja avainsanoja), ja mikäli sanoja on liian vähän, voi kuvaus olla liian väljä ja hakutuloksia liian paljon.

Ongelmaa voidaan kiertää esimerkiksi loogisten operaattorien ("or", "not", jne.) tai erillisten muiden lisäominaisuuksien ("Tarkennettu haku") avulla, mutta on selvää, että monille käyttäjille tämä on aivan ylivoimainen tehtävä. Tällaiset tiedonhakumenetelmät onkin selvästi suunniteltu käyttäjille, joilla on normaalit motoriset, kognitiiviset ja kielelliset taidot sekä keskivertoinen tietämys maailmasta, ja kasvavan tietomassan hallinta tuottaa hankaluuksia käyttäjille, joilla ei ole vastaavia tietoja eikä taitoja. Tämän vuoksi on tärkeää kehittää sellaisia tiedonhaun mahdollisuuksia, jotka ottavat huomioon erilaisten käyttäjien tarpeet ja kyvyt kehittyä tiedon hakemisessa sekä tulosten hyödyntämisessä.

Tietotekniikan tutkimuslaitos HIITin tutkimusryhmä CoSCo on kehittänyt perinteisille avainsanalogiikkaan perustuville hakukoneille vaihtoehtoisia lähestymistapoja, uusia teknologioita, jotka tarjo-

avat käyttäjille luonnollisempia keinoja kuvailla haun kohteena olevaa käsitettä semanttisesti. Kehitetyt menetelmät perustuvat tekstin tilastolliseen analyysiin ja yksinkertaiseen "sanasäkki" (bag of words) -lähestymistapaan, jossa tekstiä tarkastellaan joukkona sanoja, ja tekstihakuja tehdään päästelemällä tekstin semanttinen sisältö siinä esiintyvien sanojen yhteisesiintymistodennäköisyyksien perusteella. Tällaisessa lähestymistavassa ei kieliopilla ole analyysin kannalta merkitystä, ja kieliopilliset roolit tai odotukset eivät siis vaikuta tulokseen, joten tilastollisesta lähestymistavasta on etua erityisesti tilanteissa, joissa analysoitava teksti ei ole kieliopillisesti oikeassa muodossa. Niinpä tilastolliseen lähestymistapaan perustuvilla tiedonhakutekniikoilla on tärkeä rooli tarjottaessa hakupalveluja käyttäjille, joille perinteinen avainsanahaku tuottaa vaikeuksia.

Selkohaku-projektissa toteutettiin tekstidokumenttien sisällön semanttiseen tilastolliseen analyysiin perustuva hakukone, jonka tietovarastona toimi suomenkielinen Wikipedia (<http://www.wikipedia.fi>) yhdessä selkokielisten verkkosivustojen sisällön kanssa. Itse hakukone toimii HIITin palvelimella, mutta hakukoneeseen luotiin rajapinta, jonka kautta hakukoneen käyttöliittymä oli mahdollista toteuttaa osana Kehitysvammaliiton järjestelmiä, ja hakukonetta on testattu paikallisesti siellä. Hakukoneen normaalisti integroitu rajapinta toteutettiin verkon yli käytettävänä, jotta prototyyppikäyttöliittymien kehittäminen helpottuisi, mutta toisaalta säilytettäisiin käyttöliittymän ja hakukoneen integroinnin helpous.

Yksi ongelma verkosta ladattujen sivustojen indeksoinnissa on sivun varsinaisen sisällön erottaminen navigaatio- ja muihin tarkoituksiin sivulle lisätyistä osioista, joiden tekstisisältö johtaa hakutuloksia harhaan. Tähän tarkoitukseen projektissa kehitettiin tekstin osittamiseen ja sisältöpohjaiseen analyysiin perustuvia automaattisia menetelmiä epäoleellisten osioiden tunnistamiseksi ja poistamiseksi.

Hakusanojen syöttämiseen pohjautuva käyttöliittymä on suurelle yleisölle tutuin, sillä tätä periaatetta kaikki suuret hakukoneet noudattavat. Paitsi itse hakujen laatiminen, myös hakutulosten hyödyntäminen otsikoiden ja tekstikatkelmien listana esitettynä vaatii käyttäjältään kuitenkin sellaista hahmotuskykyä, jota kaikilla tiedonhakijoilla ei ole. Tämän ongelman kiertämiseksi projektissa kehitettiin hakutulosten listaesityksen ohittava käyttöliittymä, jossa vastauksena tarjotaan suoraan parhaiten hakua vastaava verkkosivu. Kertomalla oliko tarjottu sivu halutun kaltainen vai ei, käyttäjä voi helposti jalostaa hakua eteenpäin antamalla sisältöpohjaiselle hakukoneelle positiivisia ja negatiivisia sisältöesimerkkejä. Demo käyttöliittymästä on osoitteessa <http://cosco-demo.hiit.fi/fikipedia/simppeli.html>.

Projektissa toteutettiin myös täysin uudentyyppinen graafinen hakukoneen käyttöliittymä, joka perustuu kuvien käyttöön hakuavaimena tekstin sijaan. Perusideana on se, että yksittäisellä Wikipedian sivulla esiintyvä kuva voi edustaa koko sivun tekstimassaa, ja valitsemalla kuvan käyttäjä voi ikään kuin valita sivun koko tekstimassan semanttiseksi syötteeksi, vaikka käyttäjän ei itse asiassa tarvitse tekstiä edes lukea; mikäli sivulla on esimerkiksi jalkapalloa esittävä kuva, on todennäköistä, että teksti liittyy jotenkin jalkapalloon, jolloin siitä on hyötyä jalkapalloa koskevissa hauissa. Kuviin perustuvaa hakua voidaan laajentaa tai rajata semanttisesti valitsemalla useita kuvia (dokumentteja); valitsemalla esimerkiksi kuva sekä jalkapallosta että jääkiekosta voidaan osoittaa, että käyttäjän semanttinen kohden on laajempi ("urheilu"), tai hakukohdetta voidaan rajata valitsemalla esimerkiksi jalkapallokuva ja kuva, joka edustaa tiettyä maata tai vaikkapa aikakautta. Koska kehitetty ratkaisu mahdollistaa toimimisen pelkästään kuvien perusteella, on ratkaisu tietenkin täysin kieliriippumaton ja sopii käyttäjille, jotka eivät osaa kieltä lainkaan, tai joilla on vaikeuksia kielen kanssa. Toimiva pilottidemo tästä täysin uudentyyppisestä hakukoneesta löytyy osoitteesta <http://cosco-demo.hiit.fi/onkuvia/onkuviademo.html>.

Tutkimussuunnitelmassa mainittu rajapinta YLEn materiaaleihin päätettiin jättää toteuttamatta, ja HIITin osuutta projektista leikattiin vastaavasti.

Projektin tulokset

Yhteenvetona projektin tuloksista voidaan todeta, että osa projektin valmisteluvaiheessa asetetuista tavoitteista ylitettiin merkittävästi ja osa saavutettiin osittain. Tavoitteet ylitettiin erityisesti kerätyn käyttäjätiedon osalta. Osa saavuttamattomista tavoitteista johtui hakuteknologian kehittämisen haasteista.

Tässä kappaleessa projektin tulokset on jaoteltu projektille asetettujen osatavoitteiden mukaan ja niitä tarkastellaan tavoitteiden toteutumisen näkökulmasta.

Käyttäjätiedon keruu ja käyttöympäristöjen määrittely

Tiedonhakupalvelun kehittämisessä tarvittavan käyttäjätiedon keruu oli alkuperäistä suunniteltua laajempaa sekä määrällisesti että laadullisesti, koska vastaavaa tietoa ei ole ollut saatavilla kansallisesti tai kansainvälisesti. Projektia valmisteltaessa oli tiedossa, että aikaisempaa tietoa olisi saatavilla niukasti, mutta aikaisempia tutkimuksia kartoitettaessa selvisi, että nimenomaan erityisryhmien näkökulmasta aihetta ei ollut juurikaan aikaisemmin julkisesti tutkittu.

Selkokielen käyttäjäryhmien käyttöliittymähaasteista ja hakukäyttäytymisestä saatiin runsaasti arvokasta tietoa, jota voidaan varmasti hyödyntää jatkossa muutenkin kuin tämän projektin puitteissa. Kerätty tutkimusdata ja sen perusteella tehdyt havainnot julkaistaan Papunet-verkkopalvelussa kaikkien saataville.

Tiedonhakupalvelun käyttöliittymän käyttäjälähtöinen suunnittelu kerätyn käyttäjätiedon perusteella käytettävyystudkimusten avulla

Käytettävyystesteissä havaitut haasteet pyrittiin ratkaisemaan käyttöliittymäprototyypeissä, joita edelleen testattiin kohderyhmän käyttäjillä. Osa ongelmista oli sen luonteisia, että niihin ei ollut mahdollista löytää projektin puitteissa kokeiltavaa ratkaisua.

Useimpiin havaituista ongelmista löydettiin kuitenkin toimiva ratkaisu, jotka pystyttiin konkretisoimaan myös käyttöliittymätasolla. Käyttöliittymäprototyyppien perusteella luotiin yksi käyttöliittymä, joka sisältää parhaiksi havaitut ratkaisumallit, ja jota käytetään pohjana mm. selkokielisille internet-sivuille hakuja suorittavassa hakupalvelussa osoitteessa www.selko.fi.

Suomessa kehitettyjen avoimen lähdekoodin hakuteknologian mukauttaminen suunnitellun käyttöliittymän tarpeisiin

HIIT jatkoi perinteisistä ratkaisuista poikkeavan hakukoneteknologiansa kehittämistä eteenpäin. Projektin aikana tuotettiin mm. rajapinta, jonka avulla hakukoneteknologiaa pystyttiin testaamaan yhdessä kehitettyjen käyttöliittymäprototyyppien kanssa. Hakukoneteknologiaa ei kuitenkaan saatu projektin puitteissa niin valmiiksi, että sitä olisi voitu testata selkokielen käyttäjillä tai ottaa tuotan-

tokäyttöön jossain julkisesti saatavilla olevassa hakupalvelussa. Lisäksi HIIT kehitti tekstin osittamiseen ja sisältöpohjaiseen analyysiin perustuvia automaattisia menetelmiä epäoleellisten osioiden tunnistamiseksi ja poistamiseksi, joita hyödynnetään selkokielisille internet-sivuille hakuja suorittavassa hakupalvelussa osoitteessa www.selko.fi.

Kehitysvammaliiton verkkosivustojen, suomenkielisen Wikipedian ja YLEn aineistojen käyttävän tiedonhakupalvelun toteutus

Projektin kuluessa tuli selväksi, ettei Wikipedian tai YLEn aineistoihin kohdistuvan ulkopuolisen haun toteuttaminen tarjoaisi loppukäyttäjille mitään merkittävää lisäarvoa näiden palveluiden itse tarjoamiin hakuun verrattuna. Suomenkielisen Wikipedian ja Kehitysvammaliiton verkkopalveluiden sisältöä käytettiin projektin kuluessa testiaineistona HIITin hakukoneteknologiaalle.

Myös Kehitysvammaliiton ylläpitämien eri sivustojen käyttäjät todettiin tarpeiltaan niin heterogeeniseksi ryhmäksi, ettei nämä sivustot kattavalle yhteiselle haulle nähty todellista tarvetta.

Sen sijaan projektin kohderyhmä huomioiden todettiin, että internetin selkokielliset sivustot ovat niin toisistaan erillisiä ja hajallaan, että niiden yhdistäminen saman haun alle voisi olla selkokielen käyttäjille tervetullut ja hyödyllinen ominaisuus. Projektin tuotoksia päätettiin näin ollen hyödyntää hakukoneessa, jolla voi suorittaa hakuja kaikkeen suomenkielisessä internetissä julkaistuun selkokielliseen materiaaliin. Suomenkielisen internetin selkosivustot on kerätty osoitteessa www.selko.fi sijaitsevalle sivustolle, jonne myös kyseiset sivustot kattava haku sijoitettiin. Haku kattaa mm. Papunetin selkokiellisten sivujen sisällön, Vernerin selkokiellisten sivujen sisällön sekä YLEn Selkouutisten sisällön.

Projektissa kerättyä tietoa ja toteutettua käyttöliittymää käytetään hyväksi myös Kehitysvammaliiton verkkopalveluiden kehityksessä.

Uuden tiedonhakupalvelun lanseeraus ja ylläpito

Projektin tuotoksia hyödynnetään www.selko.fi-sivustolle lisätyssä, suomenkielisen internetin selkosivustojen sisällöt kattavassa haussa. Sivusto on ollut entuudestaan olemassa ja sen sisällöllisestä ylläpidosta vastaa Selkokeskus. Sivuston teknisestä ylläpidosta, kuten myös avatun hakutoiminnallisuuden ylläpidosta vastaa Papunet.

Muut suoritteet

Projektin tuloksia on esitelty kansainvälisessä apuvälinealan AAATE 2009 -seminaarissa Firenzessä ja kansallisessa puhetta tukevan ja korvaavan kommunikoinnin Sillalla 2010 -seminaarissa Helsingissä. Projektia on esitelty EASPD:n (European Association of Service Providers for Persons with Disabilities) eurooppalaisille vammaisalan palveluntuottajien järjestön eri jaostoille yhteisten projektien ja seminaarien yhteydessä vuosina 2008 ja 2009.

Toisen projektivuoden aikana projektia esiteltiin laajasti opetustoimen henkilökunnalle suunnatussa Kehitysvammaisen oppilas perusopetuksessa -koulutuksessa (Opetushallituksen tukemaa täy-

dennyskoulutusta). Lisäksi projektia on esitelty Kehitysvammaliiton liittovaltuustolle ja -hallitukselle sekä Papunetin johtoryhmälle. Papunetissa vuosina 2008-2009 vierailleet alan ammattilaiset ja opiskelijat ovat saaneet tietoa tiedonhaku-hankkeesta.

Projektista on myös kirjoitettu artikkeleita Tietotekniikka- ja kommunikaatiokeskus Tikoteekin Tikon tiedotteeseen, Kehitysvammaliiton jäsentiedotteeseen ja Ketju-lehteen. Projektista on tiedotettu kohdennetusti Papunetin sähköpostitiedotteen kautta. Projektin päätiedotuskanavana on toiminut Papunet-sivusto.

Pohdintaa

Projekti oli aihepiiriltään ja kohderyhmältään haastava. Toisaalla on hakukoneteknologia, jonka kaupalliset sovellukset perustuvat maailman suurimpien pörssiyritysten rahoittamiin yksityisiin tutkimuksiin, jotka on toteutettu käytännössä rajattomien taloudellisten ja työvoimallisten resurssien turvin. Toisaalla ovat selkokielen käyttäjät, joka on hyvin heterogeeninen ryhmä, jonka tarpeista käyttöliittymien suhteen on hyvin vähän tutkimustietoa. Ylipäättänsä selkokielistä materiaalia on internetissä saatavilla hyvin vähän. Silti yhteiskunta muuttuu kovaa vauhtia siihen suuntaan, että peruspalvelut siirtyvät verkkoon, missä niiden tulisi olla samalla tavalla kaikkien saavutettavissa, ikään tai vammaan katsomatta.

Hakukoneet, kuten internetin luonteeseen kuuluu, ovat jatkuvassa muutoksessa elävä toimiala. Kentässä on tapahtunut suuria muutoksia tämänkin projektin kuluessa. Esimerkiksi suomalainen Ihmemaahan Haku lakkasi olemasta, markkinoiden kolmanneksi suurin hakukone Microsoft Live Search muuttui niin teknisiltä ratkaisuiltaan kuin nimeltäänkin ja tunnetaan nykyään nimeä Microsoft Bing ja markkinoiden toiseksi suurinta toimijaa Yahoo!:ta on myyty vuorotellen Googlelle, Microsoftille ja mediajätti News Corp:lle. Kesällä 2009 Yahoo! ja Microsoft tekivät yhteistyösopimuksen, jonka seurauksena markkinoiden toiseksi ja kolmanneksi suurimmat toimijat käyttävät ainakin seuraavat kymmenen vuotta yhteistä teknologia-alustaa. Google on hakukoneiden ylivoimainen markkinajohtaja; vuoden 2009 viimeisellä neljänneksellä sen markkinaosuus oli lähes 85 %¹².

Ylivoimaisesta johtoasemastaan huolimatta Googlenkin on jatkuvasti kehityttävä, minkä johdosta sen hakukoneen käyttöliittymä muuttuu useita kertoja vuodessa. Jotkut muutokset huonontavat hakukoneen käytettävyyttä selkokielen käyttäjien näkökulmasta, jotkut puolestaan parantavat sitä; syyskuussa 2009 Google kasvatti tekstin kokoa hakukoneensa käyttöliittymässä - liian pieni teksti oli yksi tässäkin tutkimuksessa havaituista ongelmista. Googlella on myös "kokeellinen" heikkonäköisille suunnattu esteetön haku¹³, mutta selkokielen käyttäjille tai hiirtä ohjaimena käyttäville se ei tuo käytännössä muuta parannusta kuin koneellisesti "esteettömäksi" todetut hakutulokset.

HIIT:in kehittämä, perinteisestä poikkeava lähestymistapa tiedon hakuun mahdollistaa uudenlaiset käyttöliittymäratkaisut. Projektin puitteissa teknologia ei kuitenkaan kehittynyt siihen pisteeseen asti, että siihen perustuvia käyttöliittymiä olisi voitu testata selkokielen käyttäjillä. Toisaalta perinteisen kaltaisen käyttöliittymän käyttämistä puoltaa se, että kaupalliset hakukoneet tuskin muuttuvat toimintaperiaatteeltaan radikaalisti, ja niitä luultavasti kaikkien ikään tai vammaan katsomatta tulee

¹² Net Applications Search Engine Market Share Q4/2009

¹³ Google Accessible Search, <http://labs.google.com/accessible>

tulevaisuudessa osata jollain tasolla käyttää pystyäkseen toimimaan täysipainoisesti yhteiskunnan jäsenenä. Hakukoneiden käyttöliittymät ovat lähtökohtaisesti melko teknisiä, ja niiden käyttöä on käytettävyydestien perusteella opeteltava, ennen kuin käyttäjä pystyy käyttämään niitä tehokkaasti ja turhautumatta. Projektin käytettävyydesteissä henkilöillä, jotka olivat saaneet jonkinlaista opastusta hakukoneen käyttöön esimerkiksi Tikas-koulutuksen muodossa, oli luonnollisesti vähemmän ongelmia testitehtävien suorittamisessa.

LIITE 1: Käyttöliittymäongelmien analyysi

Yleistä

Tässä dokumentissa kuvataan selkohaku-projektissa tuotetuissa kokeellisissa käyttöliittymissä käytetyt ratkaisut sekä niihin johtaneet seikat.

Dokumentissa viitataan Papunetin käyttäjätesteihin, jotka suoritettiin Papunet-verkkopalveluyksikön toimesta vuoden 2008 aikana. Käyttäjätestien osanottajat olivat eri selkokielen käyttäjäryhmistä ja testeissä tutkittiin osanottajien toimintaa olemassa olevissa hakupalveluissa.

Dokumentin viittaukset muihin hakukoneisiin ja kuvankaappaukset niistä on tehty vuoden 2009 ensimmäisellä neljänneksellä. Hakukone-alan jatkuvan kehityksen johdosta osa tiedoista oli jo kirjoitushetkellä vanhentunut.

Termillä selkohaku viitataan yhteisesti kaikkiin niihin käyttöliittymävaihtoehtoihin, joita tässä dokumentissa esiteltujen ongelmien ja niiden ratkaisuiden perusteella tuotetaan jatkotestattavaksi. Jatkotesteissä vertaillaan käyttäjien hakutehokkuutta ja tehtyjen virheiden määrää eri käyttöliittymävaihtoehtojen kesken sekä selkohaun ja Googlen välillä.

Käyttöliittymä

Hakupalvelun käyttöliittymä voidaan jakaa loogisesti kahteen osaan; hakusanan syöttämiseen ja hakutulosten esittämiseen. Yleinen konventio on, että hakusanan voi syöttää myös suoraan hakutulosten yhteydessä ilman, että käyttäjän tarvitsee palata erikseen taaksepäin tai käyttää "uusi haku"-tyyppistä haun alkuun vievää toimintoa.

Hakusanan syöttäminen - ongelmat ja havainnot

Tavassa, jolla hakusanan syöttäminen on yleisesti kaikissa hakukoneissa toteutettu (tekstikenttä, johon hakusanat syötetään, ja painike, josta haku aloitetaan), ei havaittu Papunetin käyttäjätesteissä ongelmia. Mahdollisimman yksinkertainen näkymä auttaa käyttäjää pääsemään alkuun hakutehtävän suorittamisessa, kun huomiosta ei kilpaile muita, mahdollisesti hakutehtävään liittymättömiä toimintoja tai visuaalisia elementtejä. Lisäksi tämänkaltaisen malli on ollut käytössä hakukoneissa käytännössä aina, joten se on muodostunut myös konventioksi, joka on toiminnaltaan käyttäjille entuudestaan tuttu.

Ylimääräiset toiminnot kilpailevat käyttäjän huomiosta varsinaisen hakulausekkeen syöttämisen kanssa. Esimerkiksi Googlen hakulomakkeessa olevia haun tarkentamiseen ja hakuasetuksiin liittyviä toimintoja ei joko käytetty ollenkaan, tai sitten niitä käytettiin tarpeettomasti, esimerkiksi osa käyttäjistä valitsi säännönmukaisesti erikseen toiminnon "Haku suomenkielisiltä sivuilta", vaikka syötetty hakulauseke oli suomenkielinen eikä näin ollen tuottanut muun kuin suomenkielisiä tuloksia.



Kuva 1. Googlen hakulomakkeessa on useita "ylimääräisiä" lisätoimintoja.

Anne Aulan ikäihmisten hakukoneiden käyttöä selvittäneen tutkimuksen¹⁴ käyttäjätesteissä oli havaittu, että hakusanan syöttölomakkeen toteuttamisessa on joitakin potentiaalisia ongelmakohtia etenkin heikkonäköisillä käyttäjillä. Hakusanakentän koon on luonnollisesti oltava riittävän suuri sekä tekstin koon, että siihen mahtuvan merkkimäärän osalta. Lisäksi käyttäjille ei välttämättä ollut selvää, milloin selaimen fokus on hakusanakentässä niin, että siihen voi kirjoittaa. Useimmissa selaimissa oletuksena visuaalisena vihjeenä käytetty vilkkuva kursori erottuu liian heikosti. Erityisesti ikäihmisillä, jotka eivät olleet entuudestaan tottuneita näppäimistön käyttäjiä, tämä oli ongelma, sillä käyttäjien tarvitsi suunnata katseensa näppäimistöön kirjoitettaessa, eivätkä he näin ollen havainneet, että hakusanakenttään ei kirjoitettaessa ilmestynyt mitään.

Sekä Papunetin että Aulan testeissä käyttäjillä oli oikeinkirjoitukseen liittyviä ongelmia. Kirjoitusvirheet ovat hakukoneissa kaikilla käyttäjillä huomattavan yleisiä; jopa 10–15 % hakulausekkeista sisältää kirjoitusvirheitä¹⁵. Googlen oikeinkirjoitusehdotus-toimintoa osattiin hyödyntää melkein kaikissa oikeinkirjoitusongelmissa, mutta osittain sitä hyödynnettiin "väärin" niin, että oikeinkirjoitusehdotus kopioitiin käsin kirjoittamalla hakusanakenttään sen sijaan, että oikeinkirjoitusehdotusta olisi suoraan klikattu.

Tekstikenttään kirjoitetun tekstin muokkaaminen havaittiin vaikeaksi; mikäli käyttäjä halusi korjata kirjoittamaansa hakulauseketta, tämä usein pyyhki ensin aikaisemman hakulausekkeen joko kokonaan tai osittain ja sitten kirjoitti korjatun version hakulausekkeesta sen sijaan, että olisi käyttänyt korjaamiseen nuolinäppäimiä tai korjattavan kohdan maalaamista ja poistamista. Tämä on kuitenkin tekstikentän toimintalogiikkaan liittyvä ongelma ja on selitettävissä testihenkilöiden yleisillä tietokoneenkäyttötaidoilla sen sijaan, että asialle olisi käyttöliittymän osalta tehtävissä hirveän paljoa.

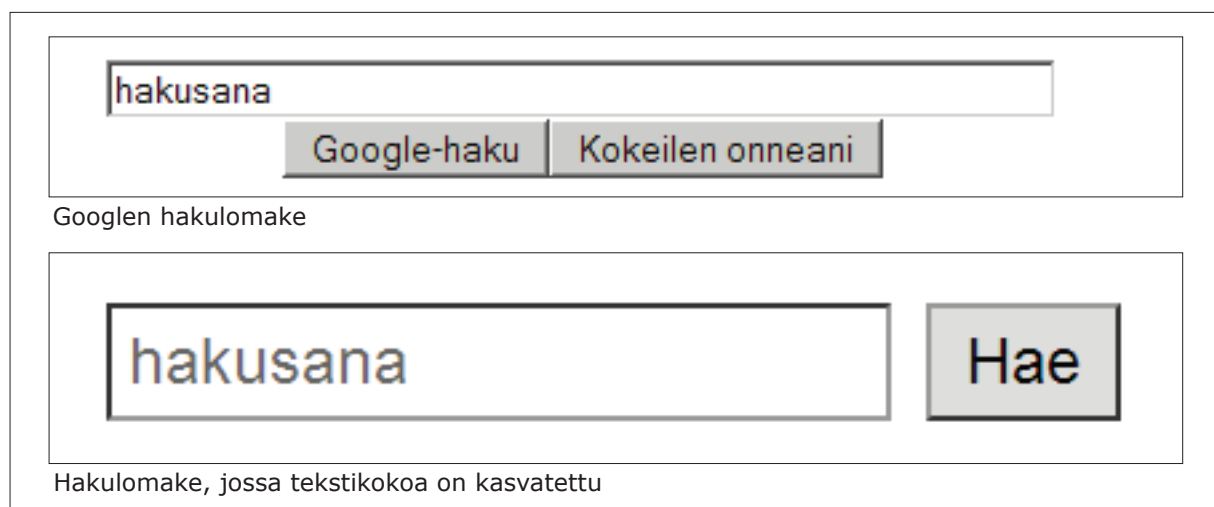
Papunetin testeissä joillakin käyttäjillä oli vaikeuksia päästä alkuun hakutehtävän suorittamisessa, koska sopivan hakutermin keksiminen oli vaikeaa. Tähän on kuitenkin lähes mahdotonta tarjota apua niin, että se oikeasti aina auttaisi käyttäjää. Ainoa mahdollinen ratkaisu olisi ehdottaa käyttäjälle rajattua joukkoa hakusanoja, joilla käyttäjä pääsisi hakutehtävässään alkuun, mutta tarjottavia hakusanoja on mahdoton valita niin, että ne kattaisivat edes jollain tasolla kaikki mahdolliset hakutarpeet. Käyttöliittymässä lähdetään siis siitä perusoletuksesta, että käyttäjä osaa jollain tavoin nimetä hakutarpeensa.

¹⁴ Aula, A (2005) User study on older adults' use of the Web and search engines. Universal Access in the Information Society. Volume 4, Number 1. Springer Berlin / Heidelberg, s. 67-81

¹⁵ Cucerzan, S., Brill, E, (2004) Spelling correction as an iterative process that exploits the collective knowledge of web users. <http://acl.ldc.upenn.edu/acl2004/emnlp/pdf/Cucerzan.pdf>

Hakusanan syöttäminen - ratkaisut

Hakusanan syöttäminen on yleisesti käytössä olevissa ratkaisuissa toteutettu visuaalisesti hyvin pelkistetyksi ja selkeästi, joten perusajatus ei voi enää juurikaan yksinkertaistaa. Hakusanakentän käytettävyyttä heikkonäköisillä käyttäjillä voi kuitenkin helposti parantaa yksinkertaisesti suurentamalla hakulomakkeen kontrolleja ja jättämällä ylimääräiset toiminnot pois. Hakusanakenttää ei kuitenkaan ole syytä suurentaa siihen mahtuvan merkkimäärän osalta, koska hakulausekkeet ovat kaikilla käyttäjäryhmillä verrattain lyhyitä, yleensä alle kolmisanaisia. Esimerkiksi Googlen hakulomakkeen hakusanakenttään mahtuu kerrallaan 50 merkkiä, Yahooon hakusanakenttään 52 tai 45 merkkiä, riippuen siitä ollaanko pää- vai hakutulossivulla, ja Microsoftin Live Searchissa 45 merkkiä.

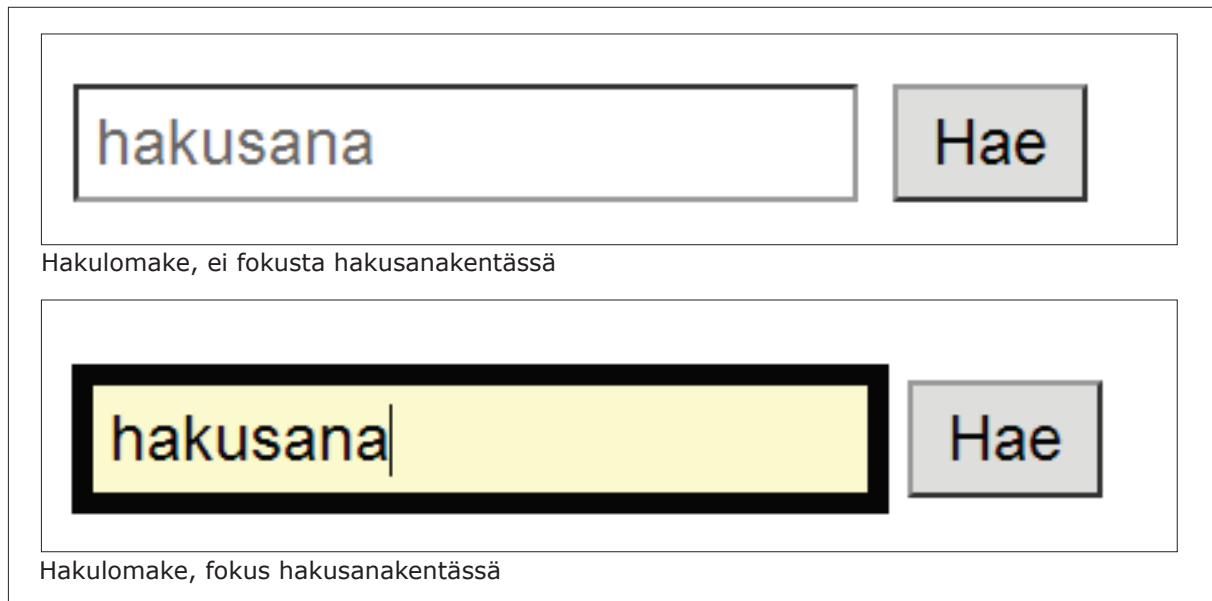


Kuva 2. Hakulomakkeen visuaalista selkeyttä parannetaan suurentamalla tekstin kokoa

Hakulomakkeessa ei teknisesti tarvittaisi muita kontrolleja kuin hakusanakenttä, jolloin haku suoritettaisiin rivinvaihtonäppäintä painettaessa. Suuri osa testikäyttäjistä suoritti kuitenkin haut klikkaamalla hiirellä hakulomakkeen lähetys-painiketta, joten käytettävyyden näkökulmasta hakulomakkeessa kannattaa säilyttää myös hakupainike. Kaikki muut toiminnot ovat selkokäyttäjän kannalta epäoleellisia, joten ne voidaan jättää pois. Esimerkiksi Googlen hakulomakkeessa on useita haun tarkentamiseen liittyviä toimintoja (vrt. Kuva 1), jotka kilpailevat käyttäjän huomiosta, ja joita testikäyttäjät eivät joko käyttäneet ollenkaan, tai käyttivät tarpeettomasti. Ainoa jotenkin perusteltu lisäkontrolli voisi olla hakusanakentän tyhjennyspainike, joka joillakin käyttäjillä saattaisi helpottaa tai nopeuttaa hakusanan muuttamista tai uuden haun tekemistä, mutta koska yleensä tyhjennyspainikkeiden käyttö lomakkeissa ei ole suositeltavaa¹⁶, ei sellaista selkohaunkaan tapauksessa käytetä.

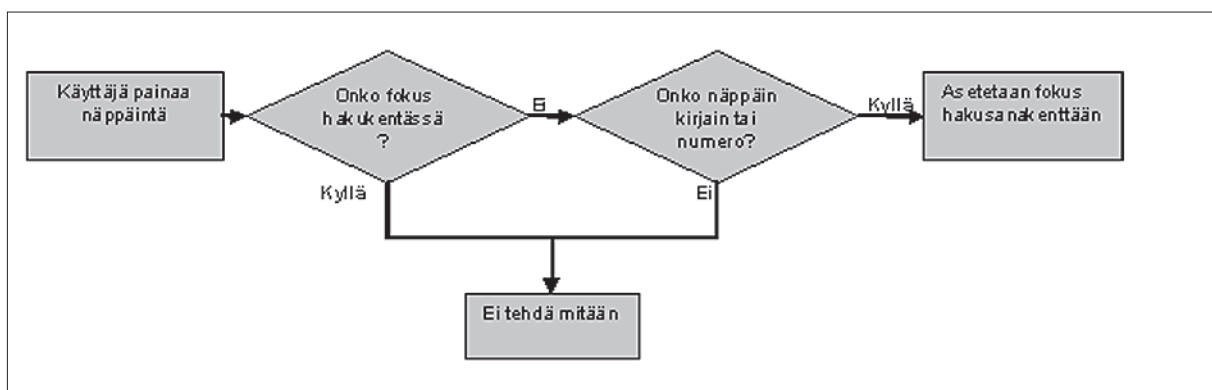
Hakusanakentän fokuksen visuaalista ilmentämistä parannetaan käyttämällä huomattavan paksua reunaviivaa ja erottuvampaa taustaväriä silloin, kun fokus on hakusanakentässä. Vilkkuvan cursorin muuttaminen erottuvammaksi olisi myös ilmeinen parannus, mutta tämän toteuttamiseen ei ollut keväällä 2010 helposti ja saavutettavasti toteutettavaa teknistä keinoa.

¹⁶ Nielsen. J. (2000) Reset and Cancel Buttons, <http://www.useit.com/alertbox/20000416.html>



Kuva 3. Hakusanakentän fokuksen visuaalinen ilmentäminen

Tilannetta, jossa fokus ei ole hakusanakentässä, kun käyttäjä alkaa kirjoittamaan voidaan lähestyä kahdella tavalla; joko keskeytetään käyttäjä ja ohjataan tämän huomio siihen, että fokus ei ole hakusanakentässä, tai oletetaan, että käyttäjä haluaa kirjoittaa hakusanaa ja siirretään fokus automaattisesti hakusanakenttään käyttäjää häiritsemättä. Näistä ensimmäinen vaihtoehto on todennäköisesti ongelmallisempi; käytännössä ainoa teknisesti toteutettavissa oleva menetelmä käyttäjän huomion saamiseen on jonkinlaisen äänimerkin käyttö, jolloin sopivan äänimerkin valitseminen niin, ettei se ole liian häiritsevä, mutta kuitenkin niin, että käyttäjä kiinnittää siihen huomionsa, voi olla haasteellista. Toinen vaihtoehto puolestaan voisi aiheuttaa ongelmia, mikäli samalla sivulla hakusanakentän kanssa olisi muita tekstisyötettä vastaanottavia kontroleja, tai aiheuttaa konflikteja selainohjelman näppäinkomentojen kanssa. Selkohaun testikäyttöliittymässä tällaiset potentiaaliset ongelmat on kuitenkin helppo rajata pois.



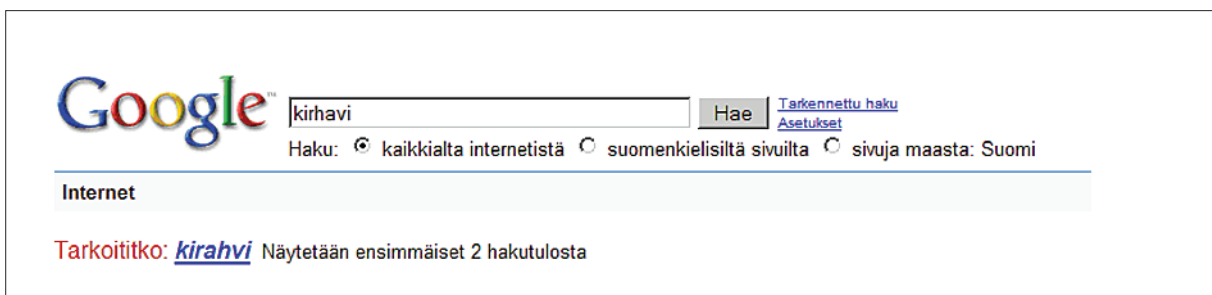
Kuva 4. Kaavio: Miten menetellään, jos fokus ei ole hakukentässä kun käyttäjä alkaa kirjoittaa

Oikeinkirjoitusta ja hakusanan valintaa olisi mahdollista edesauttaa jo hakusanan syöttövaiheessa esimerkiksi ehdottamalla käyttäjälle mahdollisia hakusanoja tämän kirjoittaessa (vrt. Google Suggest). Koska kohderyhmien käyttäjien huomio oli testeissä kuitenkin niin keskittynyt itse hakusanan kirjoittamiseen, ettei tämänkaltaisen lähestymistavan katsottu toimivan.



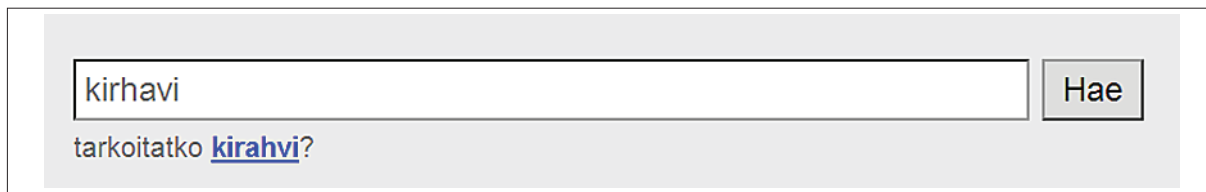
Kuva 5. Google Suggest ehdottaa käyttäjälle mahdollisia hakusanoja.

Sen sijaan esimerkiksi Googlen oikoluku-ominaisuus, jossa käyttäjälle ehdotetaan mahdollista oikoluettua hakulauseketta vasta hakutulosten yhteydessä, oli testikäyttäjillä hyvin usein käytetty toiminto, joten oikeinkirjoitukseen päätettiin puuttua selkohaussakin hakutulos-sivulla; monesti hakulauseke voi olla täysin oikea, vaikka sille ehdotetaan vaihtoehtoisia kirjoitusmuotoja, joten loogisinta on näyttää mahdolliset hakutulokset oikeinkirjoitus-ehdotusten kanssa samassa yhteydessä.



Kuva 6. Googlen oikoluku ehdottaa oikeinkirjoitettua sanaa hakutulossivulla

Selkohaun hakulomakkeessa oikeinkirjoituksen tukena kokeillaan varovaista lähestymistapaa; näytetään hakusanakentän yhteydessä mahdollinen oikeinkirjoitusehdotus, kun käyttäjä on ollut riittävästi kauan kirjoittamatta. Sama oikeinkirjoitusehdotus näytetään myös hakutulossivulla.



Kuva 7. Oikeinkirjoitusehdotus hakusanakentän yhteydessä

Myös oikeinkirjoitusehdotuksen sanavalintaan on kiinnitettävä huomiota; Papunetin testeissä osa käyttäjistä ei osannut klikata oikeinkirjoitusehdotusta suoraan, vaan kopioi sen sijaan oikeinkirjoitusehdotuksen käsin hakusanakenttään. Kysymysmuoto ei välttämättä ohjaa käyttäjää tarpeeksi vahvasti klikkaamaan linkkiä. Toisaalta liian pitkä teksti voi heikentää käytettävyyttä ja oikeinkirjoitusehdotuksen käytön turha korostaminen saattaa häiritä niitä käyttäjiä, joille sen toiminta on tuttua. Hakulomake näytetään yleensä myös hakutulossivulla, jotta käyttäjä voi halutessaan tarkentaa haun tai tehdä uuden haun palaamatta edelliselle sivulle. Esimerkiksi Googlessa ja Yahoossa hakulomakkeen ulkoasu on hakutulossivulla erilainen kuin haun "aloitussivulla". Lisäksi hakulomakkeen sijainti sivupohjassa muuttuu näiden sivujen välillä. Microsoftin Live Searchissa hakulomakkeen ulkoasu pysyy samana, mutta sijainti muuttuu. Käyttöliittymäprototyypissä hakulomake pyritään pitämään samannäköisenä ja samassa paikassa koko ajan.



Kuva 8. Googlen hakulomake näyttää erilaiselta pää- ja hakutulossivuilla

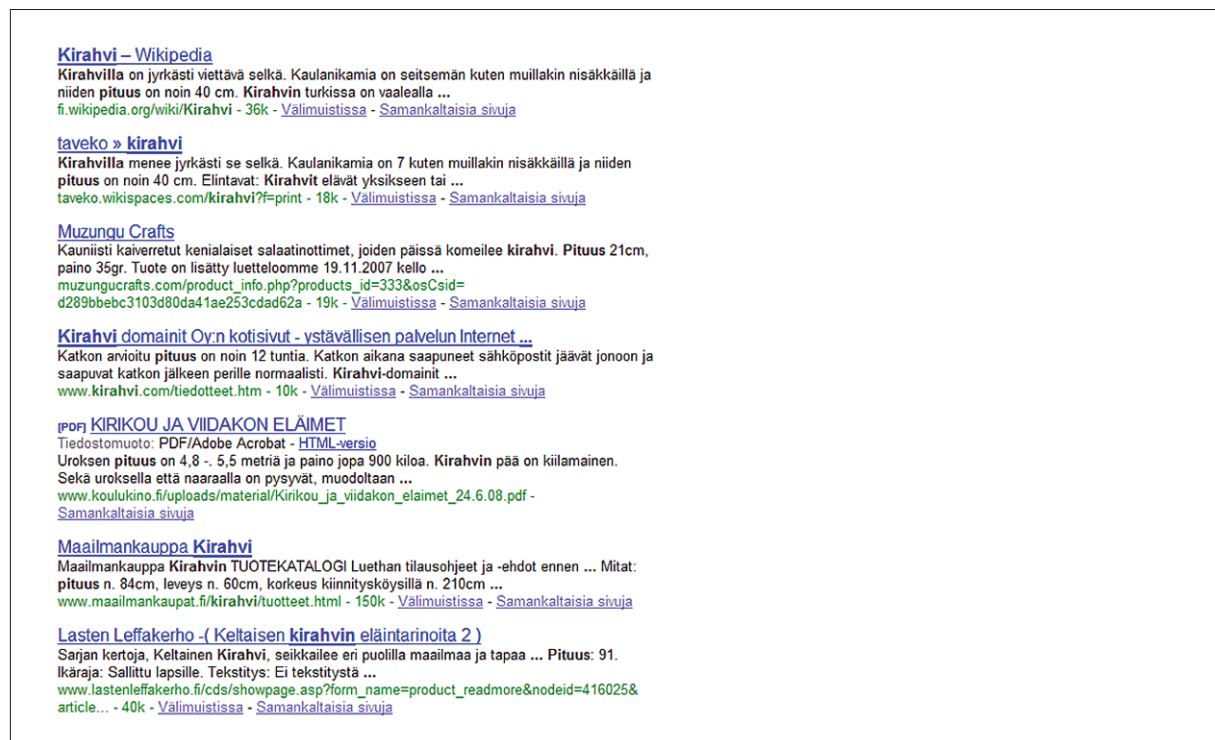
Hakutulosten esittäminen - ongelmat ja havainnot

Hakutulosten esittämisessä suurin ongelma on luonnollisesti oikeiden hakutulosten löytäminen. Tähän emme voi vaikuttaa muuten kuin auttamalla käyttäjää oikeaan suuntaan hakulausekkeensa keksimisessä ja tarkentamisessa, mutta hakutulosten laadukkuus on silti aina subjektiivista. Hakutulossivulla voidaan ajatella olevan kaksi funktiota; esittää löydetty hakutulokset käyttäjälle niin, että tämä pystyy hyödyntämään niitä mahdollisimman hyvin, sekä tarjota apua haun tarkentamiseen, mikäli annetut hakutulokset eivät vastanneet käyttäjän tarpeita.

Hakutulokset koostuvat useimmiten otsikosta, ”abstraktista” ja kohdesivun osoitteesta. Lisäksi hakupalvelusta riippuen näytetään joitakin kohdesivuun liittyviä tietoja ja toimintoja. Abstrakti on yleensä jonkin algoritmin avulla valittu ote, tai useampi ote, kohdesivun tekstisisällöstä. Abstraktiin viitataan jatkossa termillä tekstiote.

Suurin ongelma Papunetin käytettävyydesteissä oli ylipääntensä hakutulosten tulkitseminen; käyttäjät eivät osanneet valita annetuista hakutuloksista oikeata, vaikka juuri heidän tarpeisiinsa vastannut hakutulos olisikin ollut tarjolla. Todennäköinen syy tähän on hakutulosten esittämistapa; esimerkiksi Googlen hakutuloksissa tekstiote on harvoin selkeässä lausemuodossa, ja hakutuloksessa on kohdesivun otsikon ja lyhyen tekstiotteen lisäksi runsaahkosti ylimääräistä tietoa ja toimintoja, jotka heikentävät hakutulosten luettavuutta. Käyttäjät kyllä lukivat hakutulosten tekstiotteita, mutta myös niiden laatu riippuu paljolti kohdesivusta; valitettavan usein tekstiotteessa on jotain kohdesivun navigaatioon liittyvää tekstiä.

Googlen hakutulosten tekstiotteet on esitetty otteena sivun sisällöstä, joka on valittu niin, että hakutermit esiintyvät otteessa. Hakutermit on myös lihavoitu otteen tekstissä. Tutkimusten¹⁷ perusteella termien lihavointi parantaa relevanttien hakutulosten valintaa, mutta esimerkiksi osittain lihavoitu teksti saattaa huonontaa tekstin luettavuutta. Esimerkiksi yleisesti linkkien merkitsemiseen käytetävän alleviivaamisen on todettu heikentävän tekstin luettavuutta merkittävästi¹⁸.



Kuva 9. Googlen hakutuloksia hakulausekkeella ”kirahvin pituus”

¹⁷ Kirkmeier, M.D. and Albert. D. (2003) The effects of scanability on information search: An online experiment. <http://wundt.uni-graz.at/publications/file1097836870.pdf>

¹⁸ Obendorf, H and Weinreich. H. (2003) Comparing link marker visualization techniques - Changes in reading behaviour. <http://vsis-www.informatik.uni-hamburg.de/getDoc.php/publications/119/www2003.pdf>

Yksi mietittävä osa-alue on myös tarjottujen hakutulosten määrä. Papunetin testeissä vain murto-osa käyttäjistä kävi toisella hakutulossivulla, eikä suurin osa käyttäjistä vierittänyt edes ensimmäistä hakutulossivua alas asti. Tämä korostaa entisestään mahdollisimman hyvien hakutulosten merkitystä.

Mikäli käyttäjä ei löydä tarpeensa täyttävää hakutulosta, tulisi tälle tarjota keino tarkentaa hakuun. Käytännössä ainoa mahdollinen ratkaisu on ehdottaa käyttäjälle rajattua joukkoa mahdollisia hakutuloksia, jotka on valittu joko tilastollisesti tai jotenkin algoritmisesti. Google ehdottaa joillakin (riittävän yleisillä) hakulausekkeilla jatkohakuja sekä sivun ylä- että alareunassa. Jostakin selvittämättömästä syystä ehdotukset ovat keskenään erilaisia riippuen siitä missä ne sijaitsevat.

Aiheeseen liittyvät haut: [hotellimailma](#) [scandic](#) [holiday inn](#)

Googlen hakuehdotukset hakusanalle "hotelli" sivun yläreunassa

Seuraavaan aiheeseen liittyvät haut: **hotelli**

[hotelli tampere](#)

[hotellimailma](#)

[hotelli turku](#)

[hotelli jyväskylä](#)

[hotelli lahti](#)

[scandic](#)

[kylpylä](#)

[holiday inn](#)

Googlen hakuehdotukset hakusanalle "hotelli" sivun alareunassa

Kuva 10. Googlen hakuehdotukset eroavat toisistaan riippuen siitä, ovatko ne hakutulossivun ylä- vai alareunassa.

Yahoon hakutuloksissa esitetään samoin ehdotuksia haun tarkentamiseksi sekä sivun ylä- että alareunassa. Lisäksi hakutulossivun hakukenttään kirjoitettaessa on käytössä nk. Search Assist, joka Googlen Suggest -ominaisuuden tapaan ehdottaa reaaliajassa vaihtuvia hakuja hakulausetta kirjoitettaessa. Normaalien hakutarkennusten lisäksi Search Assist ehdottaa "käsitehakuja", joissa hakutermin on liitetty muita hakutermejä, jotka esiintyvät usein käytetyn hakutermin yhteydessä.

The screenshot shows the Yahoo search interface with the search bar containing the word "hotelli". Below the search bar, there are two columns of suggestions. The left column lists related terms: "omina hotelli", "hotelli helsinki", "hotelli cumulus", "hotelli torni", and "hotelli helka". The right column, under the heading "Explore concepts:", lists: "hotelli 'Helsinki, Finland'", "hotelli 'Hotel Reviews'", "hotelli 'hotel rates'", and "hotelli TripAdvisor".

Kuva 11. Yahoos Search Assist ehdottaa hakutarkennuksia sekä "käsitehakuja".

Hakutulosten esittäminen - ratkaisut

Hakutuloksen valitsemisen vaikeus johtune osaltaan hakutulosten esitystavasta, osaltaan hakutulossivun taustahälystä, joka häiritsi hakutulosten tutkimista. Hakutulosten esitystapaa voi kuitenkin yksinkertaistaa vain jonkin verran; loppukädessä esitystapa riippuu suurelta osin varsinaisen hakutuloksen viittaaman sivun sisällöstä ja siitä, miten se pystytään esittämään irrallisen tekstiotteenä mahdollisimman hyvin. Selkokielen käyttäjillä hakutulosten esitystapa korostuu entisestään; irralli-

nen lauseenpätkä kohdesivulta ei välttämättä riitä kertomaan, mistä sivulla oikeasti puhutaan. Puhumattakaan siitä, että tekstiote on vain lista navigaatiolinkkien tekstejä.

Viime kädessä tekstiotteen luominen riippuu käytettävästä hakukoneteknologiasta, koska useimmiten se palautetaan valmiina hakutuloksen yhteydessä, eikä sen sisältöön pysty käyttöliittymätasolla enää vaikuttamaan.

Anne Aula on vertaillut¹⁹ erilaisia esitystapoja hakutulosten tekstiotteiden esittämiseen, ja saatujen tulosten perusteella tehokkain esittämistapa on esittää tekstiotteet listamuodossa. Tämänkaltaista esitystapaa verrattiin myös selkohaussa normaaliin tekstiotteen esitystapaan selkokielen käyttäjillä. Hakutuloksen tekstisisällön tueksi on selkohaussa lisätty ääni. Joillakin selkokielen käyttäjillä tekstin kuunteleminen ääneen luettuna auttaa tekstin lukemista ja luetun ymmärtämistä. Äänituessa on samat ongelmat kuin tekstiotteessa ylipäättänsä; mikäli teksti ei ole selkeää kieltä ja kokonaisia lauseita, ei sen ääneen luettu versiokaan auta tekstin tulkittamista. Äänituki tuotetaan puhesyntetisaattorin avulla, eikä sen olemassaoloa mainosteta testikäyttäjille.

Selkohaun hakutulosten esityksessä pyritään luettavuuden ohella mahdollisimman suureen visuaaliseen selkeyteen suurentamalla tekstin kokoa ja riviväliä. Ylimääräinen informaatio on tärkeää pitää tässäkin tapauksessa minimissään. Hakutuloksen otsikon ja abstraktin lisäksi ei periaatteessa tarvittaisi muuta informaatiota, mutta yleinen käytäntö on lisäksi ilmoittaa kohdesivun url-osoite, mikä joissakin tapauksissa voi olla selkokäyttäjienkin näkökulmasta hyvä asia. Entuudestaan käyttäjälle tuttu osoite kertoo kohdesivun luotettavuudesta ja antaa osviittaa siitä, mitä linkin takaa löytyy.

Googlen hakutulokset

¹⁹ Aula, A. (2004) Enhancing the readability of search result summaries.
http://anne.aula.googlepages.com/aula_summary.pdf

Lorem ipsum hakusana

Lorem ipsum dolor sit amet, consectetur adipiscing elit. **Hakusana**. Mauris vehicula nisi in turpis. Duis lobortis interdum mauris. Nullam eget massa vel libero eleifend aliquet. Vivamus elit. Praesent sodales commodo arcu. Sed purus. In id elit a ipsum tincidunt mollis.

(sivulta osoitteessa www.lipsum.com)

Integer hakusana tristique

Maecenas lorem. Vivamus sagittis condimentum dui. **Hakusana**. Vivamus volutpat, eros et dignissim dignissim, nisl erat convallis lacus, quis pharetra ipsum nulla sit amet nunc. Duis tortor. Praesent auctor turpis at massa. Praesent leo lacus, vestibulum eget, ornare et, convallis quis, eros.

(sivulta osoitteessa www.lipsum.com)

Aenean sollicitudin tempus sapien

In hac habitasse platea dictumst. Sed congue. Proin pretium rutrum nibh. **Hakusana**. Donec pellentesque nunc quis urna. Integer cursus dui vel magna. Fusce commodo porttitor nunc. Nam sit amet nunc. Maecenas ligula. Phasellus dignissim congue nisl. Quisque massa.

(sivulta osoitteessa www.lipsum.com)

Selkohaun hakutulokset

Kuva 12. Selkohakun hakutuloksissa tekstin kokoa ja riviväliä kasvatetaan. Kuvat samassa suhteessa.

Käyttäjätestien perusteella hakutulosten määrää voi helposti vähentää ja lisä-hakutulosten hakemista yksinkertaistaa. Selkohaussa kokeillaan kerralla näytettävien hakutulosten määrän vähentämistä viiteen ja lisä-hakutulosten selaamista pelkillä eteen- ja taaksepäin-linkkeillä yleisesti käytetyn sivunumeroinnin sijaan. Jatkotesteissä seurataan, lisääkö tämä käyttäjien siirtymistä toiselle hakutulossivulle, vai kokeilevatko käyttäjät edelleen hakusanan tarkentamista, mikäli ensimmäisten hakutulosten joukosta ei löydy halutunlaista tulosta.

Jatkohakujen tekemiseen ja hakutermien tarkentamiseen ei selkohaun prototyyppien suunnittelu-
hetkellä ollut tarjolla teknisiä ratkaisuja, joiden avulla niitä pystyisi testaamaan. Eli esimerkiksi selkohaun prototyypeissä käytetystä Googlen ohjelmointirajapinnasta ei saa ulos annettuun hakulauseeseen liittyviä muita hakulausekkeita.

LIITE 2: Luettelo käytettävyydestien osallistujista

Ikä	Sukupuoli	Internetin käyttö, kuinka usein
13	Mies	Useita kertoja päivässä
15	Nainen	Päivittäin
16	Nainen	Useita kertoja päivässä
16	Nainen	Useita kertoja päivässä
16	Mies	Useita kertoja päivässä
17	Mies	Päivittäin
18	Mies	Päivittäin
18	Mies	Useita kertoja päivässä
20	Mies	Kuukausittain
22	Mies	Viikottain
24	Mies	Harvemmin kuin kerran kuukaudessa
24	Mies	Kuukausittain
26	Nainen	Useita kertoja päivässä
26	Nainen	Kuukausittain
27	Nainen	Kuukausittain
28	Mies	Päivittäin
35	Mies	Kuukausittain
36	Mies	Viikoittain
42	Nainen	Viikoittain
43	Nainen	Useita kertoja päivässä
44	Nainen	Viikoittain
47	Mies	Harvemmin kuin kerran kuukaudessa
48	Nainen	Päivittäin
49	Mies	Viikoittain
49	Mies	Harvemmin kuin kerran kuukaudessa
50	Mies	Harvemmin kuin kerran kuukaudessa
56	Nainen	Harvemmin kuin kerran kuukaudessa
67	Nainen	Harvemmin kuin kerran kuukaudessa
71	Mies	Päivittäin
75	Nainen	Päivittäin

Testihenkilöiden keski-ikä oli 35 vuotta ja heistä 57 % oli miehiä. 67 % kertoi käyttävänsä internetiä vähintään kerran viikossa.

LIITE 3: Luettelo testitehtävistä

Tässä on listattuna kaikki testitehtävät aakkosjärjestyksessä. Tehtävistä on muokattu pois maininnat siitä, millä hakukoneella tehtävä tulee suorittaa sekä mahdollinen pohjustus.

- Etsi [käyttäjän suosikki-joululaulu]n sanat.
- Etsi [käyttäjän suosikki-tv-ohjelma]-aiheinen keskustelupalsta.
- Etsi [käyttäjän suosikki-tv-ohjelma]n kotisivut.
- Etsi Digitoday:n uutinen Enter ry:stä.
- Etsi Eestinmetsän koulun henkilökunnan yhteystiedot.
- Etsi Eestinmetsän koulun osoite.
- Etsi glögivertailu.
- Etsi Helsingin kasvitieteellisen puutarhan pääsymaksut.
- Etsi Helsingin kaupungin sivuilta (www.hel.fi) tietoa työväenopiston tietotekniikkakoulutuksista.
- Etsi Helsingin kaupungin sivuilta kaupungin viestintäpäällikön puhelinnumero.
- Etsi Jääkiekon MM-kilpailuiden ohjelma.
- Etsi kinkunpaisto-ohjeet.
- Etsi Kuningaskuluttaja-ohjelman sivut.
- Etsi linkki Euroviisujen sivuille.
- Etsi linkki Papunetin palapeleihin.
- Etsi lista huvila-teltan esiintyjistä Helsingin juhlaviikoilla.
- Etsi lista jääkiekon MM-kilpailujen kaikista mitalisteista kautta aikojen.
- Etsi lista Ruotsin kuninkaista kautta aikojen.
- Etsi lista Suomen Euroviisuedustajista.
- Etsi luettelo [käyttäjän suosikki-tv-ohjelma]n henkilöistä.
- Etsi luettelo tämän vuoden Euroviisu-ehdokkaista.
- Etsi osoitteessa Kaisaniemenkatu 6 sijaitsevan rakennuksen rakennusvuosi.
- Etsi pajatso-peli.
- Etsi Papunet-verkkopalveluyksikön katuosoite.
- Etsi Pekingin olympialaisten avajaispäivämäärä.
- Etsi piparkakkuresepti.
- Etsi Salattujen Elämien sivut.
- Etsi Teräsbetonin perustamisvuosi.
- Etsi tieto kuinka pitkä kirahvi voi olla.
- Etsi tieto paljonko norsu voi painaa.
- Etsi tieto siitä, kuinka pitkä Matti Vanhanen on.
- Etsi tieto siitä, kuinka vanha Paavo Lipponen on.
- Etsi tietoa juhannuskokosta.
- Etsi tietoa siitepöly-allergioista.
- Etsi tämän vuoden joulukalenteri.
- Etsi tämän vuoden tangomarkkinoiden ajankohta.
- Etsi YLE:n sivuilta uutinen siitä, kun Helsinki valittiin vuoden 2005 yleisurheilun maailmanmestaruuskilpailujen järjestäjäksi.
- Koska Linnanmäki aukeaa?
- Kuinka vanha Suomen Euroviisuedustaja Waldo on?
- Kuinka vanhoina lapset oppivat kävelemään?

- Kuka oli Suomen ensimmäinen presidentti?
- Mikä on Suomen yleisin lintu?
- Selvitä mikä on lämpötila Norsunluurannikolla huhtikuussa.
- Selvitä missä katuosoitteessa Papunetin porukka majoilee.

