

LAPPEENRANNAN TEKNILLINEN YLIOPISTO
TIETOTEKNIIKAN OSASTO

HAKUKONEOPTIMOINTI INTERNETMARKKINOINNIN TUKENA

Diplomityön aihe on hyväksytty Lappeenrannan teknillisen yliopiston tietotekniikan osaston osastoneuvoston kokouksessa 14.8.2006.

Työn tarkastajina toimivat Prof. Heikki Kälviäinen ja TkT Pekka Jäppinen ja ohjaajana tekn.yo Kalle Pyörälä.

Lappeenrannassa 23.10.2006

Juha Paananen
Teknologiapuistonkatu 4 b 11
53850 Lappeenranta
050-5448298
juha.paananen@lut.fi

TIIVISTELMÄ

Lappeenrannan teknillinen yliopisto
Tietotekniikan osasto

Juha Paananen

Hakukoneoptimointi internetmarkkinoinnin tukena

Diplomityö

2006

84 sivua, 17 kuvaa, 12 taulukkoa

Tarkastajat: Prof. Heikki Kälviäinen, TkT Pekka Jäppinen

Hakusanat: internetmarkkinointi, hakukonemarkkinointi, hakukoneoptimointi, kävijäseuranta, web-metriikka

Keywords: internet marketing, search engine marketing, search engine optimization, clickstream analysis, web metrics

Yhä useampi etsii nykyään tietoa tuotteista ja palveluista internetin kautta. Vastapainoisesti lähes jokainen yritys käyttää internetsivujaan markkinointikanavana. Mietittäessä markkinoinnin peruskysymyksiä kuten kohdesegmentin saavuttamista tai kampanjan tuottoastetta ei vastausta usein osaa internetsivujen osalta antaa niin markkinointiosasto kuin IT-osastokaan. Hakukoneoptimointi on yksi hakukonemarkkinoinnin muoto, jonka avulla internetsivujen saavutettavuutta voidaan parantaa. Kehityksen toteamiseksi on oltava mittareita, joina internetsivuilla voidaan käyttää internetsivuille tarkoitettuja kävijäseurantaohjelmistoja.

Tässä työssä käsitellään hakukoneoptimointia ja sen mahdollisuuksia parantaa sivustojen näkyvyyttä internetin hakukoneissa. Hakukoneoptimoinnilla tarkoitetaan sivustojen teknisen toteutuksen muokkaamista hakukoneystävälliseksi ja sisällön muokkaamista niin, että sivustot sijoittuvat halutuun hakusanoin hakutulosten kärkipäähän. Onnistumisen mittaamiseksi työssä perehdytään kävijäseurannan mahdollisuuksiin ja toteutukseen. Työn tavoitteena oli tuoda Primesoft Oy:lle riittävä tietotaito hakukoneoptimoinnista, toteuttaa hakukoneoptimointipalvelu ja muokata yrityksen ohjelmistot hakukoneoptimointia tukeviksi.

Työn tavoitteet saavutettiin pääosin ja tutustuminen hakukoneoptimointiin avasi portin koko internetmarkkinoinnin maailmaan. Palvelun toimivuutta testattiin Primesoftin omilla sivuilla ja tulokset osoittautuivat varsin rohkaiseviksi. Jatkossa hakukoneoptimointia voidaan tarjota palveluna asiakkaille.

ABSTRACT

Lappeenranta University of Technology
Department of Information Technology

Juha Paananen

Search Engine Optimization as a Part of Internet Marketing

Master's Thesis

2006

84 pages, 17 pictures, 12 tables

Supervisors: Prof. Heikki Kälviäinen, Dr.Tech. Pekka Jäppinen

Keywords: internet marketing, search engine marketing, search engine optimization, clickstream analysis, web metrics

A growing number of people use internet to search information about products and services. Counterbalance that, most of the companies use their websites as their marketing channel. Still, neither the marketing nor IT department is usually able to answer the fundamental questions of marketing, like is the target group able to find the website or what the return on investment of the website is. Search engine optimization is a form of search engine marketing that makes the websites to be more easily reachable to the target audiences. As every marketing act, also, search engine optimization has to be measured to see the development. The web metrics software available can be used for this purpose.

This thesis concerns search engine optimization and its possibilities to enhance website visibility in search engines. Search engine optimization means improving the technical implementation and content of a website, so that the website will position to the top of the search engine results. To measure the success, some research on the web metrics is, also, done. The objective of this thesis was to bring the required knowledge of search engine optimization to Primesoft Ltd., to develop a search engine optimization service and to rework the Primesoft's marketing support software to be search engine optimization friendly.

The objectives were mostly achieved and search engine optimization widened the company's view to internet marketing on the whole. Functioning of the search engine optimization service was tested on the company web pages. The results were encouraging and in the future the service can be offered to the customers.

Alkusanat

Tämä diplomityö on tehty Primesoft Oy:ssä Lappeenrannan teknillisen yliopiston Tietotekniikan osastolle. Työ toteutettiin toukokuun ja lokakuun välisenä aikana vuonna 2006. Etsiessäni diplomityöpaikkaa ja -aihetta Primesoft tarjosi mahdollisuuden tehdä työn hakukoneoptimoinnista. Otin työn vastaan innolla, koska aiheesta on Suomessa tehty hyvin vähän tutkimusta, vaikka internetmarkkinointi kokonaisuutena on maailmanlaajuisesti nopeimmin kasvava markkinoinnin osa-alue. Suurimmat kiitokset tästä työstä kuuluvat työni ohjaajalle, mahdollistajalle ja Primesoftin toimitusjohtajalle Kalle Pyörälälle. Aihetta koskevien tieteellisten julkaisujen tarjonnan ollessa niukkaa tahdon osoittaa myös kiitokset aiheesta vapaa-aikanaan artikkeleita julkaisseille alan ammattilaisille ja erityisesti Seomozin Rand Fishkinille Search Engine Ranking Factors -kartoituksesta. Edellä mainittujen lisäksi tahdon kiittää työni virallisia tarkastajia Heikki Kälviäistä ja Pekka Jäppistä rakentavasta palautteesta ja mielenkiinnosta työtäni kohtaan. Lopuksi tahdon antaa vielä erityiskiitokset työni kolmannelle tarkastajalle Anna Väistöille, joka helpotti työtaakkaani tarkastamalla ja lukemalla työtäni useita kertoja.

23.10.2006

Juha Paananen

SISÄLLYSLUETTELO

1	JOHDANTO	6
1.1	Tausta	6
1.2	Työn tavoite ja rajoitteet	6
1.3	Työn rakenne	7
2	HAKUKONEOPTIMOINTI	9
2.1	Toiminta lyhyesti	9
2.2	Miksi hakukoneoptimoida?	10
2.3	Hakukoneoptimoinnin asema internetmarkkinoinnissa	11
3	HAKUKONEET	13
3.1	Hakukoneiden toiminta	13
3.1.1	Indeksointi	13
3.1.2	Hakujen käsittely	16
3.2	Hakukäyttäytyminen	17
3.2.1	Kuluttajakäyttäytyminen	17
3.2.2	Kuluttajakäyttäytyminen internetissä	19
4	OPTIMOINTIPROSESSI	22
4.1	Avainsana-analyysi	22
4.1.1	Avainsanojen kehittäminen	23
4.1.2	Kehittelyperiaatteita	24
4.2	Sivuston optimointi	25
4.2.1	Tekninen optimointi	26
4.2.2	Sisällön optimointi	27
4.3	Linkkianalyysi	28
4.3.1	Verkkoteoria	28
4.3.2	Linkitys ja hakukoneet	30
4.3.3	Linkkien kerääminen	31
4.4	Indeksointi	33
4.4.1	Lisääminen hakukoneisiin	33
4.4.2	Miksi sivut eivät indeksoidu?	34
4.4.3	Indeksoinnin seuranta	36
4.5	Seuranta	37
4.5.1	Mitä?	37
4.5.2	Miten?	39
4.5.3	Mitä hyödytään?	40

5	PRIMESOFTIN TARPEET HAKUKONEOPTIMOINNIN SUHTEEN	43
5.1	Primesoft Oy	43
5.1.1	Palvelut ja tuotteet	43
5.1.2	Toimintaperiaate	44
5.2	Yrityksen internetpalvelut ja tavoitteet	45
5.2.1	Tunnettavuuden ja näkyvyyden lisääminen	46
5.2.2	Kävijäseuranta	47
6	HAKUKONEOPTIMOINNIN TOTEUTUS PRIMESOFTILLA	49
6.1	PrimeContent ja hakukoneoptimointi	49
6.1.1	PrimeContent-julkaisujärjestelmä	49
6.1.2	Korjaustoimenpiteet	51
6.2	Kävijäseuranta	52
6.2.1	Google Analytics -kävijäseurantaohjelmisto	53
6.2.2	Palvelinlokien analysointiohjelma Webalizer	54
6.2.3	Snoobi-kävijäseurantaohjelmisto	56
6.3	Kävijäseurannan valinta	56
6.4	Kävijäseurantaohjelmisto PrimeTrack	58
6.4.1	Sivulatausten käsittely	59
6.4.2	Tietokanta	61
6.4.3	Tiedon analysointi	62
7	YRITYKSEN SIVUJEN HAKUKONEOPTIMOINTI	64
7.1	Alkutilanne	64
7.2	Toimenpiteet	65
7.3	Tulokset	66
7.4	Kokemuksia toiminnasta	68
8	YHTEENVETO	70
	LÄHTEET	72

LYHENTEET

ASCII	American Standard Code for Information Interchange
CLF	Common Log Format
CPA	Cost Per Acquisition
CPC	Cost Per Click
CSS	Cascading Style Sheets
DMOZ	The Open Directory Project
ECLF	Extended Common Log Format
HTML	HyperText Markup Language
HTTP	HyperText Transfer Protocol
IP	Internet Protocol
PDF	Portable Document Format
PPC	Pay-Per-Click
ROI	Result on Investment
SEM	Search Engine Marketing
SEO	Search Engine Optimization
URL	Unified Resource Locator
W3C	World Wide Web Consortium
WWW	World Wide Web

TERMISTÖ

Avainsanaohjelma	Ohjelma, jota käytetään avuksi sivustojen avainsana-analyysissä. Ohjelma käyttää avainsanojen etsintään apuna hakukoneiden hakusanatilastoja.
Bow-tie -teoria	Bow-tie -teoria on teoria, joka selittää sivustojen linkittymistä toisiinsa. Sivut ovat joko irrallisia sivuja, sivuja, jotka linkittävät toisiin sivustoihin, ydinsivuja, joihin linkitetään ja jotka linkittävät, tai kohdesivuja, jotka eivät linkitä muihin sivuihin, mutta muut linkittävät niihin.
Julkaisujärjestelmä	Sivustojen tuotanto-ohjelma, jossa ulkoasu ja sisältö ovat erotettu toisistaan. Näin myös muutkin kuin sivujen tekijät voivat tuottaa sivuille ulkoasultaan yhdenmukaista sisältöä.
Eväste	Eväste (cookie) on alunperin verkkokaupoille kehitetty tekstitiedosto, jonka avulla voidaan tallettaa käyttäjän koneelle istuntokohtaista tietoa.
Jäljite	Jäljite (web beacon) on sivulle lisättävä elementti, joka raportoi palvelimelle sivulatauksen.
Kehykset	Kehykset (frames) ovat sivurakenteen toteutustapa, jonka käyttö ei ole suositeltavaa.
Klikkihinta	Hakusanamarkkinoinnissa hinnoittelu perustuu käyttäjien suorittamiin mainoksiin kohdistuneisiin klikkeihin. Yhden klikin hintaa kutsutaan klikkihinnaksi (Cost Per Click).
Linkittäjä	Linkittäjä on sivu tai sivusto, jolta on linkki omalle sivustolle.
Maksulliset hakemistot	Yleisiä tai yrityshakemistoja, joita ylläpitävät tulosta tavoittelevat organisaatiot.
Metatiedot	WWW-sivujen tunnisteessa olevat sivun sisältöä koskevat tiedot. Hakukoneoptimoinnin kannalta mielenkiintoiset metatiedot koostuvat kuvauksesta (description) ja avainsanoista (keywords).
Ohjelmistorajapinta	Rajapinta, jonka kautta ohjelma voi keskustella toisen ohjelman kanssa.

Selainlaajennus	Selainlaajennus (plug-in) on WWW-selaimen yhteydessä käytettävä komponentti, joka mahdollistaa erilaisten multimediatiedostojen katselun.
Sivu	Sivulla tarkoitetaan yksittäistä WWW-sivua.
Sivusto	Sivusto koostuu yhtenäisistä samaan kokonaisuuteen kuuluvista sivuista.
Sponsoroidut linkit	Hakukoneissa esiintyvät hakutulokset, jotka ovat erillään normaaleista hakutuloksista. Ne sisältävät lähinnä mainoksia ja näkyvät mainostajan haluamin hakusanoin.
Viittaaja	Viittaaja (referer) on selaimen palvelimelle antama tieto, joka kertoo viimeksi ladatun sivun. Tiedon avulla voidaan päätellä, mistä kävijä on tullu juuri ladatulle sivulle.

1 JOHDANTO

1.1 Tausta

Yhä useampi etsii nykyään tietoa tuotteista ja palveluista internetin kautta. Vastapainoisesti lähes jokaisella yrityksellä on internetsivut, monet yrityksen ovat tehneet suuria sijoituksia noudattaakseen alan viimeisimpiä standardeja ja ollakseen mahdollisimman käyttäjäystävällisiä. Kun kysytään, miksi yrityksellä on kotisivut, perusteet ovat usein markkinoinnillisia. Sivuilla pyritään myymään yrityksen tuotteita ja palveluita sekä tuomaan ne potentiaalisten asiakkaiden tietouteen. Vastoin perinteisen markkinoinnin perussääntöjä tulee internetmarkkinoijan mieleen aniharvoin kysymyksiä, kuten saavutettiin kohdesegmentit tai mikä oli kampanjan tulos. Näihin kysymyksiin antaa vastauksen internetmarkkinoinnin tukitoiminto hakukoneoptimointi. Sivustojen hakukoneoptimoinnin avulla parannetaan sivustojen näkyvyyttä hakukoneissa ja luodaan yhteys asiakkaan tarpeen ja yrityksen tuotteiden ja palveluiden välille. Hakukoneoptimointi on jatkuva markkinointiprosessi. Jotta onnistumista voidaan mitata ja jatkaa kehitystä, on hakukoneoptimoinnille asetettava mittarit. Kävijäseurantaohjelmistot ovat ohjelmistoja, jotka keräävät tietoa ja tilastoivat sivustojen kävijämääriä ja ihmisten toimintaa internetsivuilla. Ne siis kertovat markkinoijalle, löytävätkö asiakkaat sivuston, ja tuloksista voidaan myös päätellä sivuston tuottoasteita muiden markkinointitoimenpiteiden rinnalla.

1.2 Työn tavoite ja rajoitteet

Primesoft Oy on markkinointipalveluja ja markkinoinnin tukiohjelmia myyvä yritys. Perinteisten markkinointimateriaalien ohella Primesoft on erikoistunut internetmarkkinointiin. Internetmarkkinointi on nopeimmin kasvava markkinointimuoto ja erityisesti hakukonemarkkinoinnin kasvu on ollut viime aikoina voimakasta. Yritys ei ole aiemmin tarjonnut markkinointia hakukonemarkkinoinnin muodossa, joten vastatakseen kehityksen haasteisiin ja asiakkaiden tarpeisiin oli yrityksen laajennettava myös tälle alueelle. Tämän diplomityön tavoitteena oli selvittää yhden hakukonemarkkinoinnin muodon eli hakukoneoptimoinnin toteutus osana yrityksen muita palveluita. Työssä tutkittiin, mistä hakukoneoptimoinnissa on kysymys, miten se voidaan toteuttaa palveluna, mitä lisäohjelmistoja tarvitaan ja millaisia muutoksia vanhoihin ohjelmistoihin joudutaan tekemään. Tässä työssä keskityttiin sivustojen sijoituksen parantamiseen eniten käytetyissä ja kaupallista arvoa omaavissa hakukoneissa. Sivustojen muokkauksen lisäksi tutustuttiin sivustojen kävijäseurantapalveluihin ja toteutettiin yritykselle sellainen. Perehtymisen jälkeen

yrittäjien omille sivuille toteutettiin hakukoneoptimointi ja arvioitiin saatuja tuloksia.

1.3 Työn rakenne

Työ aloitetaan tutustumalla lyhyesti hakukoneoptimointiin, jotta tämän jopa hieman harhaanjohtavan termin tarkoitus selviää ja voidaan ymmärtää, mihin hakukoneoptimoinnilla pyritään. Toiminnan pääpiirteisen esittelyn jälkeen tutkitaan tarkemmin hakukoneita ja niiden käyttäjiä. Hakukoneoptimointi on jo nimensäkin perusteella sidottu voimakkaasti hakukoneisiin. Tunteamatta omia työvälineitään työskentely on vaikeaa eikä työntekijään anna itsestään kovin vakuuttavaa kuvaa. Kun hakukoneiden toimintaperiaatteet ovat tulleet selväksi, tutustutaan hakukoneiden käyttäjiin. Liikkeelle lähdetään normaalia kuluttajakäyttäytymisestä, joka avaa näkökulmia muun muassa hakukoneoptimoinnin avainsana-analyysiin ja helpottaa markkinointihenkilöstön kanssa kommunikointia. Ennen kuin siirrytään itse hakukoneoptimointiprosessiin, laajennetaan kuluttajakäyttäytymisen käsitettä internetissä ilmeneviin kuluttajakäyttäytymisen erityispiirteisiin.

Neljäs luku käsittelee optimointiprosessia ja sen eri vaiheita. Prosessi aloitetaan suorittamalla avainsana-analyysi, jossa kehitellään sivuston avainsanat, joita uskotaan hakijoiden käyttävän tarjottuja palveluita etsiessä. Seuraavaksi suoritetaan sivuston tekninen ja sisällöllinen optimointi, jossa sivustosta tehdään hakukoneystävällinen ja muokataan sisältöä niin, että sen sijoitus hakukonelistauksissa kehittyy halutuun avainsanoin. Merkittävä tekijä sivuston luokituksessa on sivustolle osoittavat linkit. Linkkianalyysissä selvitetään linkkien merkitys ja miten sivustolle hankitaan ulkopuolisia linkkejä. Linkityksen jälkeen lisätään sivusto hakukoneiden indekseihin ja pohditaan syitä, miksi sivusto ei välttämättä näy tuloksissa. Prosessin loppuvaiheessa perehdytään optimointiprosessin onnistumisen mittaukseen ja erilaisiin mittausvaihtoehtoihin.

Teorian jälkeen tarkennetaan yrityksen asettamia vaatimuksia hakukoneoptimoinnille ja kävijäseurannalle. Vaatimusten asetusten jälkeen tutustutaan hakukoneoptimointipalvelun toteutukseen Primesoftissa. Toteutuksessa keskitytään lähinnä teknisiin muutoksiin, jotka koskevat yrityksen ohjelmistoja, sekä kävijäseurannan toteutukseen. Kävijäseurannan vaihtoehtoihin pureudutaan kirjallisen tutkimuksen kautta. Esille tulleiden vaihtoehtojen jälkeen valitaan paras ja toteutetaan kävijäseuranta.

Luvussa 6 toteutetaan hakukoneoptimointi yrityksen omille sivuille. Aluksi määritellään avainsanat, jonka jälkeen suoritetaan itse optimointi ja asetetaan sivustolle kävijäseuranta. Jotta hakukoneoptimoinnin toiminnasta saatiin viitteitä, seurattiin tilanteen kehitystä tou-

kokousta 2006 syyskuuhun 2006 asti. Luvun loppuksi esitellään saadut tulokset ja esitetään myös omia mielipiteitä hakukoneoptimoinnin toiminnasta.

2 HAKUKONEOPTIMOINTI

Lyhykäisyydessään hakukoneoptimoinnilla tarkoitetaan internetsivuston muokkaamista sellaiseen muotoon, että hakukoneet luokittelevat sivuston mahdollisimman hyvin. Luokituksella tarkoitetaan tässä yhteydessä arvoa, jonka mukaan hakukoneet järjestävät hakutuloksissa näytettävät tulokset. Luonnollisestihan jokainen sivun ylläpitäjä sekä yritystään internetissä mainostava yrittäjä toivoisi olevansa hakutulosten kärkipäässä. Näin on kuitenkin hyvin harvoin ja tällöin hakukoneoptimoinnin avulla voidaan parantaa tuloksia. Jo lyhyellä aiheen artikkelien lukemisella voidaan tuloksia parantaa, mutta hyviin tuloksiin päästäkseen on tunnettava niin ihmisten hakukäyttäytyminen, hakukoneiden toiminta, verkkosivujen toteutustekniikat kuin kohdealan termistökin. Tässä luvussa käsitellään hakukoneoptimoinnin periaatteet pääpiirteittäin, miksi hakukoneoptimointi on tärkeää ja mikä on hakukoneoptimoinnin asema internetmarkkinoinnissa.

2.1 Toiminta lyhyesti

Nopein reitti hakukoneoptimointiin kiinnipääsyy on tutustua hakukoneiden itsensä toimintaan. Hakukoneet kuten Google keräävät automaattisesti tiedot sivuilta. Tällaiset hakukoneet koostuvat kolmesta eri osasta, hakurobotista, sivukannasta sekä tulostennäyttösovelluksesta. Hakurobottien tehtävä on selata läpi internetissä olevia sivuja lisätä ne sivutietokantaan eli indeksiin. Käyttäjien suorittaessa hakuja kolmasosa etsii indeksistä vastaavat sivut ja näyttää tulokset luokituksen mukaisessa järjestyksessä. Hakukoneoptimoinnin kannalta mielenkiintoisia asioita ovat robotit sekä tuloksista vastaava osa. [1]

Etenemistapoja on useita, mutta hyvä tapa on aloittaa tarkastamalla vierailevatko robotit sivustolla ja löytävätkö ne kaikki sivut. Tarkastus voidaan suorittaa etsimällä hakurobottien jälkiä esimerkiksi palvelinten lokitiedostoista tai etsimällä omia sivuja suoraan hakukoneista. Mikäli robotti ei ole käynyt sivuilla, voidaan se ilmoittaa useimpiin hakukoneisiin ja robotti tulee tarkistamaan sivun muutaman päivän kuluessa. Useimmat sivustot tulevat automaattisesti indeksoiduksi ja onkin syytä tarkistaa, ettei esteenä ole esimerkiksi epästandardit toteutustavat tai esimerkiksi jollekin selainlaajennuksella toteutettu sisältö. Robotit näkevät ainoastaan sisällön, joka on nähtävissä myös tekstiselaimin. [1]

Kun sivusto on saatu indeksoitua, voidaan siirtyä luokituksen parantamiseen. Väärinkäytösten välttämiseksi hakukoneiden luokitusalgoritmit ovat hyvin salattuja. Varmaa kuitenkin on, että esiintyäkseen jollain sanalla on sana myös löydyttävä sivustolta. Viimeis-

tään tässä vaiheessa hakukoneoptimoijan on kehitettävä lista sanoista, joilla hän haluaa sivustonsa löytyvän. Sanalistan kehitykseen on useita tapoja. Lähteenä voidaan käyttää esimerkiksi omia tuoteryhmiä, palvelunimikkeitä ja internetistä löytyviä hakusanoja tilastoivista palveluista. Kun sanat on kehitetty, lisätään ne sisältöön. Löytymisen lisäksi luokitukseen vaikuttavat sanojen sijainti ja sanamäärät. Leipätekstin ohelle sanoja voidaan sijoittaa muun muassa linkkiteksteihin ja otsikoihin. [1]

Jos pelkkä sivuston tekninen ja sisällöllinen optimointi riittäisi, olisivat hakukoneiden tulokset lähes merkityksettömiä väärinkäytösten vuoksi. Sivuston lisäksi hakukoneet käyttävät luokitukseen myös ulkoisia tekijöitä. Yksi tärkeimmistä tekijöistä on linkitys ulkopuolisilta sivuilta. Sivuston ulkopuolisia linkkejä on tarjolla niin ilmaisista kuin maksullisistakin internetin hakemistopalveluista. Lisäksi apuna voidaan käyttää eri sidosryhmiä pyytämällä heitä asettamaan linkki sivustolle heidän omalta sivustoltaan. Linkit paljon käytetyiltä sivuilta ovat yksi tärkeimmistä avaimista hyvään hakukonenäkyvyyteen. Näiden toimenpiteiden jälkeen voidaan siirtyä odottamaan nousua tuloksissa ja kävijämäärissä. Luonnollisesti hyvä ja pysyvä näkyvyys vaatii jatkuvaa seuranta omien ja kilpailijoiden sijoitusten ajatellen. Optimoinnin onnistumisessa on useita kriittisiä tekijöitä. Esimerkiksi huonosti suoritettulla avainsana-analyysillä saatetaan saada hyvät sijoitukset, mutta asiakkaita ei tavoiteta. Luvussa 4 käsitellään eri vaiheita ja niiden onnistumiseen vaikuttavia tekijöitä tarkemmin. [1]

2.2 Miksi hakukoneoptimoida?

Mansikanmyyjä tietää, että kauppa käy paremmin valtatievarressa kuin hiljaisemman sivutien varressa. Markkinoidakseen tehokkaasti tuotteitaan ja palveluitaan yrityksen on mentävä asiakkaiden luokse eikä päinvastoin. 22. elokuuta 2006 Ylen uutiset kertoivat, että suomalainen käyttää keskimäärin enemmän internetiä päivässä kuin lukee lehteä. Tutkimusten mukaan internetin kautta on tavoitettavissa jopa 70 prosenttia suomalaisista [2, 3]. Oli siis kyseessä lähes toimiala kuin toimiala, potentiaalisia asiakkaita on tavoitettavissa verkon kautta.

Google-hakukoneessa tehdään päivittäin miljoonia hakuja pelkästään suomalaisten toimesta ja yli puolet ihmisistä käyttää hakukoneita aina käyttäessään internetiä [2]. Kuten luvusta 2.1 selvisi, hakukoneoptimoinnilla tehostetaan näkyvyyttä yksittäisillä avainsanoilla. Tällöin siis saavutetaan myös juuri yrityksen ratkaisemiin ongelmiin ratkaisuja etsivät. Tarkkaan rajatulla aiheesta kiinnostuneella kohderyhmällä on suuri ostopotentiaali eikä mainontaa tunneta kiusalliseksi, kuten esimerkiksi on usein puhelinmainonnan

yhteydessä [2].

Tuotteiden ja palveluiden esilletuonnin lisäksi yritysten on vastattava myös kilpailijoiden liikkeisiin. Vuoden 2006 Tammi-heinäkuun välillä verkkomainonta on kasvanut Suomessa 28 prosenttia ja Vuonna 2005 Search Engine Marketing Professional Organizationin (SEMPO) mukaan Yhdysvalloissa 44 prosenttia. Yhä useampi yritys on siis löytänyt internetmarkkinoinnin. Asiakkaiden huomattavassa saman internet tulee kasvattamaan entisestään merkittävyyttään markkinoinnin työkaluna. [3, 4]

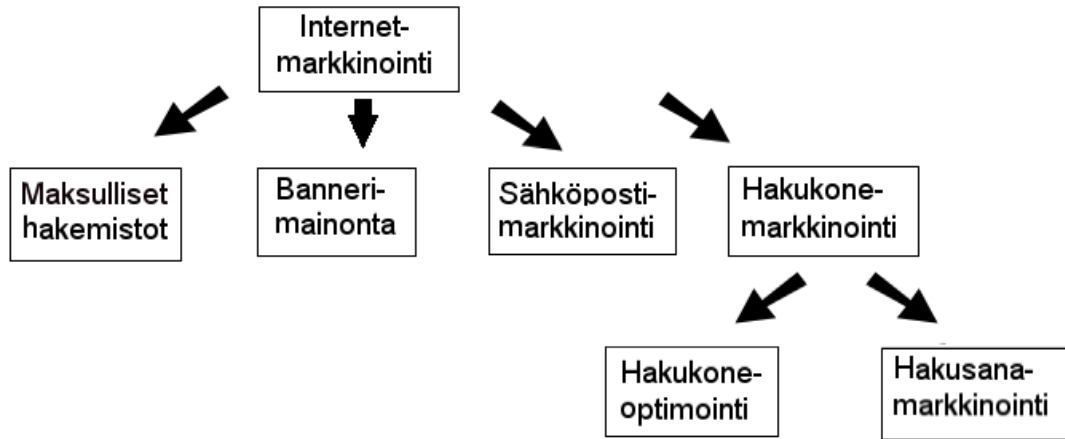
2.3 Hakukoneoptimoinnin asema internetmarkkinoinnissa

Internetmarkkinoinnissa käytetään internetiä markkinoinnin apuvälineenä eli operatiivisessa mielessä hankitaan uusia asiakkaita internetin kautta sekä ylläpidetään ja aktivoidaan vanhoja. Johtuen suhteellisen nuoresta iästään perustavanlaatuisia teoksia ei ole paljon tarjolla ja myös eri osa-alueiden jakaantumisen suhteen on erilaisia käsityksiä. Internetmarkkinointi sisältää mainostuksen maksullisissa hakemistoissa, bannerimainonnan (banner ads), sähköpostimarkkinoinnin sekä hakukonemarkkinoinnin. [5]

Internetmarkkinoinnin juuret juontavat 90-luvun alkupuolelle, jolloin toiminta rajoittui lähinnä tekstipohjaisiin tuotelistauksiin sekä tuote-esittelyihin. Myöhemmin siirryttiin graafisten mainosten kautta asiakaspalveluun, myyntiin ja sidosryhmien hallintaan asti. Internetistä on tullut merkittävä kaupankäyntipaikka perinteisten ohelle. Syitä kasvulle ovat muun muassa jatkuva tiedonsaanti, laskeneet kustannukset esimerkiksi laskevien myyntihenkilöstötarpeiden vuoksi ja helppo laajentuminen kansallisilta markkinoilta kansainvälisille. Toisaalta kansainvälistyminen on myös asettanut paineita yrityksille ja perinteisen markkinoinnin lisäksi yritysten on laajennettava näkyvyyttä myös internetiin. [5, 6]

Kuvassa 1 on esitetty internetmarkkinointi ja sen osa-alueet puuna. Hakukoneoptimointi on siis osa hakukonemarkkinointia. Molempien tähän ryhmään kuuluvien toimenpiteiden päämäärä on sama eli tuoda näkyvyyttä hakukoneissa. Hakukoneoptimoinnilla tämä tapahtuu nostamalla sivuston luokitusta ja tuomalla sivusto esille niin sanotuissa oikeissa hakutuloksissa. Hakusanamarkkinoinnissa näkyvyyttä haetaan maksullisten ja sponsoroitujen linkkien avulla. Nämä näkyvät yleensä erillään oikeista hakutuloksista ja sijoitukset jaetaan hakusanakohtaisten tarjousten mukaan. Molemmat tukevat siis toisiaan muun muassa hakusananalyysin kautta. Hakusanamarkkinoinnin lisäksi hakukoneoptimointi on sidoksissa maksullisiin hakemistoihin. Maksulliset hakemistot ovat linkkikokoelmia, joihin yritykset voivat lisätä itsensä saadakseen yhteystietonsa näkymään internetis-

sä. Näihin hakemistoihin suhtaudutaan usein vakavasti ja ne ovat yritysmaailmassa arvostettuja. Tämän vuoksi niiden luokitus on myös korkea ja linkit hakemistoista korottavat myös oman sivuston luokitusta merkittävästi. [7]



Kuva 1: Hakukoneoptimoinnin asema internetmarkkinoinnissa.

3 HAKUKONEET

Hakukoneoptimoinnissa on kyse sivuston näkyvyydestä hakukoneissa. Vuonna 2004 arvioiden mukaan vain noin kolmannes internetissä olevista sivuista oli indeksoitu hakukoneiden toimesta [8]. Sivujen löytyminen Googlasta ei siis ole itsestäänselvyys ja jo nimensäkin perusteella hakukoneita voidaan pitää hakukoneoptimoijan tärkeimpinä työkaluina. Indeksoinnin ohella on myös hyvä tietää esimerkiksi, miten hakukoneet asettavat sivut arvojärjestykseen ja miten järjestykseen voi vaikuttaa. 19. huhtikuuta 2006 muun muassa Tietokone-lehti uutisoi satojen suomalaisten yritysten poistamisesta Googlen indeksistä. Syynä tähän oli suomalaisen Webfinder.fi -palvelun ostaminen, joten on myös hyvä tietää, miten ei pidä menetellä [9]. Tässä luvussa käsitellään hakukoneiden toiminnan perusteita sekä näkyvyyteen vaikuttavia asioita.

Hakukoneoptimointi olisi helppoa, jos itse hakujen suorittajia ei tarvitsisi ottaa huomioon. Sivustoa laatiessa ja optimointia suoritettaessa on otettava huomioon, että hakijat eivät useinkaan ole kohdealan saatikka hakukoneiden käytön ammattilaisia. Hakijan etsiessä tietoa yrityksen tai tuotteen nimellä hakija tietää jo ennalta, mitä haluaa. Tällöin asiakas on jo löytänyt ratkaisun tarpeeseensa, mutta olennaista olisi tunnistaa aiheet, joihin ratkaisu on vasta haussa. Luvussa 3.2 käsitellään tiedonhakijaa kuluttajakäyttäytymisen näkökulmasta siltä osin, kuin se hakukoneoptimointiprosessissa on tarvittavaa. [10]

3.1 Hakukoneiden toiminta

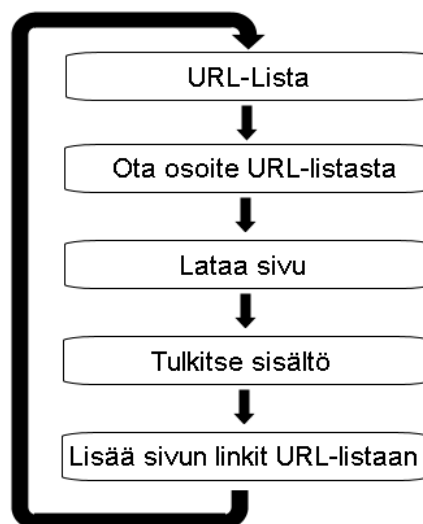
Hakukoneet ovat verkkosivustoja, joiden avulla käyttäjät voivat hakea tietoa erilaisin hakusanoin. Niiden avulla hakukoneet muodostavat listoja sivustoista, jotka sisältävät tietoa käytettyihin hakusanoihin liittyen. Yksinkertaisen toiminnan taustalta löytyy runsaasti kapasiteettia vaativa ja monimutkainen järjestelmä. Hakukoneet voidaan jakaa pääpiirteittäin kolmeen olennaiseen osaan, tiedonkeruuseen, tietokantaan ja käyttäjien hakujen käsittelyyn [11]. Seuraavissa luvuissa käsitellään tarkemmin optimoinnin kannalta tärkeitä osia, tiedonkeruuta eli indeksointia ja käyttäjähakujen käsittelyä.

3.1.1 Indeksointi

Ennen indeksoinnin tarkempaa käsittelyä on hyvä tehdä selväksi hakemistojen ja todellisten hakukoneiden ero. Esimerkiksi Yahoo ja Google mielletään usein samanlaisiksi ha-

kukoneiksi, mutta niiden tiedonkeruu poikkeaa toisistaan. Yahoo on toiminnaltaan niin sanottu hakemisto, johon asiakas voi usein maksua vastaan ilmoittaa sivustonsa. Tämän jälkeen Yahoohon asiakaspalvelu tarkistaa sivuston kelvollisuuden ja lisää sen hakemistosta löytyvään sopivaan alihakemistoon. Google taas käyttää apunaan hakurobotteja, jotka indeksoivat sivustoja siirtyen linkistä linkkiin. Hakukoneoptimointia ajatellen hakemistot tuovat luokitusta nostavia linkkejä, mutta merkittävämmässä roolissa ovat robotteja käyttävät hakukoneet. Niiden vuoksi on ymmärrettävä robottien toiminta ja niiden sivuston sisällölle asettamat vaatimukset. [12]

Hakurobotit ovat siis ohjelmia, jotka käyvät internetin graafimaista rakennetta läpi sivu sivulta. Hakurobottien tavoitteena on käydä läpi niin suuri osa internetsivuista kuin mahdollista ja syöttää kopiot tietokantamoottorille, joka luo sivusta optimoidun paikallisen kopion hakuja varten. Yksinkertaisimmillaan robotille annetaan alkusivu, jonka lukemisen jälkeen se siirtyy sivulta löytyviin linkkeihin ja niin edelleen. Ongelmia prosessiin aiheuttavat kuitenkin muun muassa sivukoodin tulkinta, tiedon ajantasaisuus, palvelinten kuormituksen käsittely ja robotteja vastaan kohdistettu ilkivalta. Hakurobotit ovat merkittävässä osassa mitattaessa hakukoneiden tehokkuutta. Tarkan kilpailijoilta salaamisen vuoksi tarkkaa tietoa robottien toiminnasta on siis vaikeaa saada. Kuvassa 2 on esitetty yksinkertaistettu esimerkki hakurobotin toiminnasta. Robotilla on lista läpikäytävistä osoitteista, joista löytyvät sivut se lataa ja tulkitsee niiden sisällön. Sisältö lähetetään hakukoneen tietokannalle ja sivulta löytyneet linkit lisätään URL-listaan. [13]



Kuva 2: Hakurobotin toiminnan eri vaiheet.

Robotin ensimmäinen osa ylläpitää listaa verkko-osoitteista URL (Unified Resource Locator) -muodossa (kuva 2: URL-lista). Sen tehtävänä on listata osoitteet, joihin hakuko-

ne siirtyy seuraavaksi. Sen tehtävänä on myös huolehtia, missä osoitteissa on jo käyty, milloin viimeksi ja onko tarvetta käydä uudestaan [13]. Koko internetiä on käytännössä mahdotonta indeksoida sen koon ja rikkonaisuudenkin vuoksi. Tämän vuoksi oikealla indeksointistrategialla on laajuuden ohella vaikutusta hakutulosten laatuun ja luotettavuuteen. Tutkimusten mukaan jo perinteisellä leveyshaulla saadaan luotettavasti keskeisimmät sivut indeksoitua. Leveyshaussa robotti käy sille annetun sivuskaalan läpi tasoittain aloittaen indeksoimalla kaikkien eri verkko-osoitteiden pääsivut siirtyen sen jälkeen seuraavalle tasolle. Apuna voidaan kuitenkin käyttää myös muun muassa sivujen luokitus-historiaa eli sivujen aiempaa tilastoitua luokitusta ja luokituksen kehitystä sekä sivustojen laajuutta. Indeksoimalla ensin ja päivittämällä useammin paljon muuttuvia, korkeasti luokiteltuja ja suuria sivustoja saadaan hakukoneen tiedot ajantasaisemmiksi ja kattavam-maksi eniten käytettyjä sivuja ajatellen. Edellisten kriteerien avulla saadaan indeksointia myös nopeutettua, jolloin aikaa jää myös alempien tasojen indeksointiin. [14]

Seuraavassa vaiheessa robotti ottaa yksittäisen sivun URL-listasta ja lataa sen (kuva 2). Vartenotettavia seikkoja ladatessa sivuja ovat robottien poissulkuprotokollan noudatta-minen, kuormituspiikkien välttäminen sekä nimipalveluhakujen optimoiminen. 90-luvun puolivälissä internetin käytön ja hakukoneiden yleistyessä hakurobotit eivät olleet terve-tulleita useille palvelimille. Ne aiheuttivat runsaasti kuormitusta, sotkivat äänestyksiä ja hakivat toistuvasti samoja tiedostoja. Nykyiset robotit ovat pääosin älykkäämpiä, mut-ta silti kaikki sivustot eivät halua robotteja sivuilleen. Ongelman ratkaisuksi on kehi-tetty robottien poissulkuprotokolla, joka on toteutettu palvelimen juureen asennettavan robots.txt-tiedoston avulla. Standardimuotoisesta tiedostosta robotti voi katsoa, mitä se saa indeksoida ja mitä ei. [15]

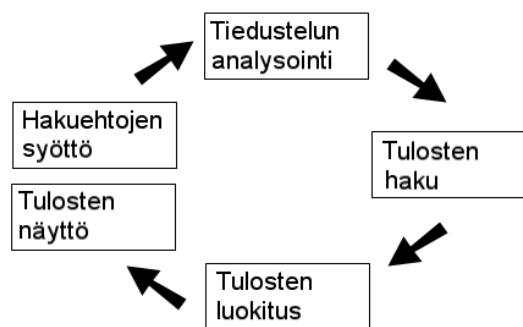
Suurin pullonkaula hakurobottien toiminnalle ovat erilaisten vastauspyyntöjen odotusajat. Tämän vuoksi ne hyödyntävät rinnakkaislaskentaa, jolloin prosessointiaikaa ei kulu huk-kaan. Ongelmana kuitenkin on, että rinnakkaiset prosessit saattavat pyynnöillään kuor-mittaa voimakkaasti vastapalvelinta ja pyynnöt voidaan tulkita esimerkiksi palvelunesto-hyökkäykseksi. Robottien on siis pyrittävä olemaan kohteliaita vastapalvelimia kohtaan ja ladattava sivuja riittävän pitkin väliajoin. Noudattamalla “if-modified-since” -tunnisteita ja lataamalla vain tietty pätkä tiedostoista eivät muut palvelimet kärsi niin paljon ja in-deksointi nopeutuu. [16]

Tunnistamalla peilipalvelimet ja rinnakkaissivustot vähenee kuormitus huomattavasti. On-gelmallisia ovat kuitenkin runsaat verkkotunnuskyselyt, jotka vievät runsaasti aikaa. Tut-kimusten mukaan tehokkaan robotin standardiosa onkin oma nimipalvelujärjestelmä. [16]

Viimeisenä operaationa prosessissa suoritetaan ladatun sivun tulkinta. Tulkintavaiheessa robotti jäsentää sivun, kerää käytetyt sanat, luokittelee ne ja ilmoittaa tiedot tietokannalle. Johtuen vaaditusta prosessointitehosta hakurobotit tulkitsevat sivua vain tietyn sanamäärän verran ja ovat kriittisiä esimerkiksi HTML (HyperText Markup Language)-standardien suhteen. Hakukoneiden kehittyessä ja niiden välisen kilpailun kiristyessä tulkintaa on kuitenkin laajennettu ja nykyaikaiset robotit osaavat käsitellä useanlaisia virheitä kuten viallisia HTML-elementtejä ja ei ASCII (American Standard Code for Information Interchange) -merkkejä. Myös tiedostovalikoimaa on laajennettu HTML-tiedostoista esimerkiksi PDF (Portable Document Format) ja Microsoft Word -dokumentteihin, joten kaikkea sisältöä ei ole välttämättä muuttoa HTML-formaattiin. [2]

3.1.2 Hakujen käsittely

Hakukoneoptimoinnin kannalta toinen tärkeä hakukoneiden osa on niiden indeksoidun tiedon käsittely. Hakujen käsittely voidaan jakaa kolmeen päävaiheeseen. Ensimmäisessä vaiheessa hakukone analysoi käytetyt hakusanat ja etsii hakusanoja vastaavat sivut. Sivujen löydyttyä niille lasketaan luokitus, jonka mukaan ne viimeisessä vaiheessa järjestetään. Kuvassa 3 on esitetty kaaviona yksittäisen haun tapahtuminen. [10]



Kuva 3: Tiedonkulku suorittaessa kysely hakukoneelta.

Käytetystä “hakurivistä” käytetään tässä yhteydessä nimeä tiedustelu. Jos käyttäjä kirjoittaa esimerkiksi Googlen rivin “markkinointi ja viestintä”, on rivi “markkinointi ja viestintä” käytetty tiedustelu (kuva 3, haku-ehdojen syöttö). Käsiteltäessä hakua tiedustelu puretaan yksittäisiksi sanoiksi, haku-ehdoiksi (kuva 3, tiedustelun analysointi). Purkamisen jälkeen haku-ehdot käsitellään yksittäin, jolloin hakukoneet etsivät ehdoille läheisiä vastaavia sanoja, oikeinkirjoitettua muotoja sekä muuttavat ehdot pienellä kirjoitetuiksi. Ohessa poistetaan myös yleiset merkityksettömät sanat sekä käsitellään loogiset operaattorit. Merkityksettömillä yleisillä sanoilla tarkoitetaan sanoja kuten esimerkiksi on, siis

ja tai. Edellä mainituin ohjein hakukone siis hakisi todennäköisesti tietoa tiedusteluin “markkinointi viestintä”, “markkinointiviestintä” ja “viestintä markkinointi”. [10]

Tulosten haun jälkeen hakukone suorittaa yhden eniten keskustelua aiheuttavista operaatioista eli tulosten luokituksen (kuva 3, tulosten luokitus). Tulosten luotettavuuden ja manipuloinnin vaikeuttamisen vuoksi tarkkaa tietoa luokituksesta on mahdotonta saada. Esimerkiksi Googlen epäillään käyttävän jopa 150 erilaista muuttujaa luokituksen laskemiseen. Yleisesti tiedossa olevia luokitukseen vaikuttavia asioita ovat muun muassa sivukoodin standardius, hakuehtojen määrä sivulla sekä sivulle osoittavien linkkien määrä. Luvussa optimointiprosessi käsitellään tarkemmin näitä kriteereitä ja niiden vaikutusta tuloksiin. Hakukoneoptimoinnin teknisessä toteutuksessa on kuitenkin pohjimmiltaan kyse luokituksen nostosta mahdollisimman korkeaksi halutuilla tiedusteluilla. [8, 10]

Tulosten esittäminen on aiempiin vaiheisiin nähden suoraviivaisempaa. Hakukoneet esittävät tulokset sivun otsikon mukaan lisäten niihin lyhyen kuvauksen sivun sisällöstä. Muutamit hakukoneet käyttävät vielä kuvaukseen tarkoitettua sivulta löytyvää metatietoa kuvauksena. Pääosa on kuitenkin luopunut metatietojen käytöstä niiden helpon väärinkäytön vuoksi. Sivun kuvaus tulee usein esimerkiksi alaotsikoista, sivun ensimmäisestä tekstikappaleesta tai sivustolle käytetystä DMOZ (The Open Directory Project) -kuvauksesta. [10]

3.2 Hakukäyttäytyminen

Hakukoneoptimoinnin ollessa markkinoinnin tukitoiminto on oltava perustietämys myös markkinoinnista itsestään. Useissa yrityksissä myös koko hakukonemarkkinointi on yksittäisten ihmisten vastuulla ja varsinaista markkinointiosastoa ei välttämättä löydy ollenkaan, jolloin alan tuntemisen tärkeys korostuu entisestään. Hakukäyttäytymiseen voidaan soveltaa normaaliin kuluttajakäyttäytymisen teorioita. Tässä luvussa käsitellään kuluttajakäyttäytymisen perusteita siltä osin, kuin ne ovat olennaisia, sekä hakukäyttäytymisen erityispiirteitä.

3.2.1 Kuluttajakäyttäytyminen

Kuluttajakäyttäytyminen tutkii prosesseja, jotka liittyvät yksilöiden tai ryhmien tapoihin täyttää tarpeensa ja tyydyttää halunsa. Prosessit sisältävät tehdyt valinnat, tuotteiden oston, käytön, käytöstä poiston sekä palveluiden hyödyntämisen. Kuluttajien skaala on laaja

ja se käsittää ihmiset lapsista suuriin yrityksiin. Kulutuksen kohteet vaihtelevat ihmisten päivittäisistä ruokatarpeista suuryritysten hankintoihin ja jopa henkilöpalvontaan. Tutkittaessa prosesseja on myös huomioitava tarpeiden monimuotoisuus perustarpeista statusarvoa kohottaviin kulutuksen kohteisiin. [17]

Kuluttajakäyttäytyminen nähdään usein yksittäisinä tapauksena eli ostona. Todellisuudessa se on kuitenkin jatkuva ketju, joka alkaa kuluttajan tarpeen havaitsemisesta ja päättyy joko tuotteen käytöstä poistoon tai jatkuu tuotteen uudelleenhankintana. Taulukossa 1 on esitetty esimerkki kulutusprosessista ja esille tulevista kysymyksistä kuluttajan sekä markkinoijan näkökulmasta. Taulukossa on esitetty kulutusprosessin kolme eri vaihetta. Prosessin eri vaiheissa esillä olevat asiat on esitetty kysymysmuodossa. Kysymyksistä voidaan päätellä, mitkä asiat kuluttaja näkee olennaisena ja millaisiin ongelmiin kuluttaja hakee ratkaisua eri vaiheissa. Vastaajan eli markkinoijan puolella on esitetty kysymyksiä, joihin vastaamalla markkinoija onnistuu vastaamaan kuluttajan kysymyksiin ja ongelmiin. [17]

Taulukko 1: Kulutusprosessi kuluttajan ja markkinoijan näkökulmasta.

Vaihe/Näkökulma	Kuluttaja	Markkinoija
Ennen ostoa	Mitkä tuotteet tyydyttävät tarpeen? Mistä saa tietoa vaihtoehtoista?	Miten kuluttaja muodostaa mielipiteensä tuotteesta? Mitkä seikat saavat asiakkaan valitsemaan tuotteen?
Oston aikana	Onko tuotteen hankkiminen helppoa? Millaisen kuvan tuotteen hankinta antaa ostajasta?	Miten erilaiset ostopäätökseen vaikuttavat ostopäätökseen? Miten tuoda esille myyjän ammattitaito?
Oston jälkeen	Täyttääkö tuote vaatimukset? Onko jatkossa tuotteesta eron pääseminen helppoa?	Miten selvittää onko asiakas tyytyväinen ja mitkä seikat siihen tyytyväisyyteen vaikuttavat? Millaista sanomaa ostaja levittää tuotteesta?

Kuten taulukon 1 prosessista voi päätellä, kuluttajan ymmärtäminen auttaa sekä kehittämään yrityksen toimintaa että luomaan kuluttajalle miellyttävän kokemuksen. Tarpeet voidaan täyttää ainoastaan tuntemalla palveluiden ja tuotteiden käyttäjät ja pärjätäkseen kilpailussa kohderyhmä on tunnettava paremmin kuin kilpailijat. Jotta käyttäjien tunteista ja markkinoinnin onnistumisesta voidaan mitata, on asetettava tarvittavat mittarit. Tärkein mittari kampanjan onnistumisessa myynnin ohella on asiakkailta saatu palaute,

ja mitä lähempänä on asiakas, sitä helpompi palautetta on kuulla ja kerätä. [17]

Muiden markkinointitoimenpiteiden ohella kuluttajan ymmärtämistä ja onnistumisen seuraamista voidaan helpottaa segmentoinnilla. Segmentoinnilla kuvataan asiakasryhmiä, jotka ovat samanlaisia yhden tai useamman piirteen perusteella ja sen vuoksi poikkeavat muista. Tuotteesta ja palvelusta riippuen rajaus voi olla hyvin tarkka tai massamarkkinoille suunnattaessa lähes olematon. Taulukossa 2 on esitetty perinteiset muuttujat, joita käytetään segmentoinnin pohjana. Muuttujia käytetään apuna kohdistettaessa markkinoita esimerkiksi niin, että, jos tuotteen käyttäjät ovat 50-60 vuotiaita (muuttuja: ikä) pohjoismaalaisia (muuttuja: maanosa) miehiä (muuttuja: sukupuoli), ei heille suunnattua tuotetta kannata markkinoida medioissa, jotka ovat suunnattu keskieurooppalaisille naisille. Ääripäässä kuitenkin käytettävien muuttujien rajana on vain mielikuvitus kunhan seuraavat ehdot säilyvät voimassa [17, 18]:

- Segmentin kuluttajia yhdistävät tekijät vastaa johonkin tuotteen täyttämään tarpeeseen.
- Merkittävät erot segmenttien välillä voidaan tunnistaa.
- Segmentti on niin laaja, että myynti on kannattavaa.
- Segmentti on saavutettavissa kohtuullisella markkinointisuunnitelmalla.
- Segmentin kuluttajat reagoivat markkinointisuunnitelmaan halutulla tavalla.

Taulukko 2: Segmentointiin perinteisesti käytettyjä muuttujia. [17]

Kategoria	Muuttuja
Demografiset tekijät	Ikä Sukupuoli Sosiaalinen luokka
Geografiset tekijät	Maanosa Maa
Käyttäytyminen	Brändiuskollisuus Statushakuisuus

3.2.2 Kuluttajakäyttäytyminen internetissä

Kuluttajakäyttäytyminen on tieteen haarana hyvin poikkitieteellinen sekä laaja. Alan tunteminen pelkästään aiheeseen vihkiytyneellekin on haastavaa. Tässä työssä aiheeseen ei

uppouduta syvällisemmin, vaan tarkoitus on luoda pohja hakukoneoptimoinnin eri sidosryhmien kanssakäymisen helpottamiseksi, hakukäyttämisen ymmärtämiseksi sekä lähtökohdiksi hakusana-analyysiin. Tutkimusten mukaan internet on kriittisessä roolissa ihmisten kulutus päätöksiensä teossa ja yli 90 prosenttia yrityksistä käyttää internetin haku palveluita tehdessään "business to business" -hankintapäätöksiä. Internetistä sekä hakukoneista on siis tullut viime vuosina tärkeä elementti ennen kauppaa tapahtuvassa päätöksenteossa ja se on myös tärkeä kuluttajakäyttämisen osa-alue. [19]

The Role of Search In Business Buying Decisions -tutkimus osoittaa, että internetistä on tullut perinteisten medioiden lisäksi merkittävä informaatiolähde tuotteiden ja palveluiden hakuun sekä ostoon [19]. Verrattuna perinteisiin mainoksiin, käyttöohjeisiin ja myyntimiesten muistiin ja asiantuntevuuteen internet tarjoaa rajattomasti tietoa ympäri vuorokauden. Poikkeuksellisen internetistä tekee myös sen interaktiivisuus, joka on huomiotava tehdessä tutkimusta kuluttajakäyttämisen näkökulmasta. Näiden ominaisuuksien vuoksi on tultu jopa tulokseen, että pallo on siirtynyt kaupankäynnissä myyjältä ostajalle [20]. Koska internet yhdistää niin kaupankäynnin kuin ajanvietonkin, on sivustoja suunniteltaessa otettava huomioon käyttämisen erot eri tilanteissa.

Menemättä tarkemmin käyttämiseen internetissä siirrytään suoraan hakukäyttämiseen ja hakulausekkeen muodostamiseen eri tilanteissa. Hakutilanteet voidaan karkeasti jakaa kolmeen pääryhmään: suunnistavat haut, informaationetsintä ja kauppahakuiset haut.

Suunnistavilla hauilla tarkoitetaan hakuja, joiden tarkoituksena on löytää joku tietty sivusto. Tällainen sivusto voi olla esimerkiksi jonkun yrityksen sivusto tai sivusto, jolla hakija on käynyt. Näihin hakijoihin hakukonemarkkinointi ei suuremmin vaikuta, koska he tietävät tarkalleen mitä haluavat. Hakukonemarkkinoinnin on kuitenkin otettava huomioon sivut niin, että tarkastaa sivujen löytymisen päähakukoneista ja että hakulistauksen kuvaukset ovat kunnossa. Vaikkei suunnistavissa hauissa olekaan potentiaalia konversioille, on ne huomiotava suuren osuutensa vuoksi.

Informaation hakijat etsivät informaatiota jostain tietystä aiheesta. Suunnistavista hauista nämä eroavat niin, että hakijat uskovat aiheesta löytyvän tietoa, mutta he eivät tiedä mistä. Haut tapahtuvat usein niin, että hakija aloittaa hyvin yleisellä termillä ja tarkentaa hakufraasiaan karsiakseen asiaan liittymättömiä tuloksia. Informaatiohaut ovat tärkeimpiä hakuja hakukonemarkkinoinnin kannalta. Tässä vaiheessa potentiaalinen asiakas on usein tunnistanut tarpeensa ja etsii ratkaisuja tarpeen täyttämiseen. Tarjoamalla informaatiopitoisia selkeitä sivuja oikein avainsanojin eli optimoimalla sivustot tavoitetaan nämä

hakijat.

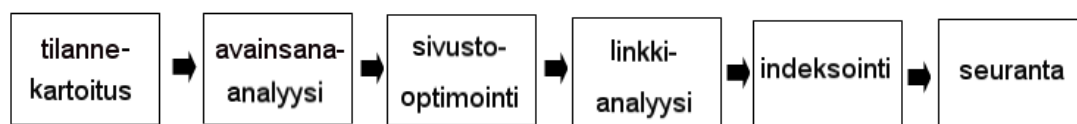
Viimeiseen ryhmään kuuluvat, jotka eivät enää etsi tietoa, vaan haluavat ostaa tuotteen, ladata tiedoston, liittyä postituslistalle tai tehdä jotain muuta vastaavaa. Hakukonemarkkinoinnin kannalta he ovat vaikeita asiakkaita. Hakijat, jotka tietävät, mitä haluavat, käyttävät hakufraaseja, jotka sisältävät tarkkoja tuotetietoja. Ongelmia tuottavat hyvien kuvausten hakukoneisiin saanti sekä hyvien hakukonesijoitusten saanti. Luomalla hyvät kuvaukset tuotelistauksiin sekä saattamalla erikoistarjoukset ja muut myyntiä kasvattavat seikat hakukoneiden kuvauksiin päästään hyviin tuloksiin. Taulukossa 3 on esitetty esimerkkeinä erilaisia hakuja, kerrottu mihin ryhmään ne kuuluvat ja miksi. [21]

Taulukko 3: Esimerkkejä erilaisista hakutilanteista. [10]

Hakutyyppi	Hakulauseke	Hakijan aikomus	Miksi hakija klikkaa
Suunnistava	iltalehti	Löytää Iltalehden kotisivut	Hakusanat otsikossa ja kuvauksessa
Informaatiohaku	digikamerat	Informaatiota useilta sivuilta	Hakusanat kuvauksessa ja sivusto luotettava
Kauppahakuinen	Canon A60	Haluaa ostaa kameran	Sama kuin yllä + tarjoukset

4 OPTIMOINTIPROSESSI

Hakukoneoptimoinnissa optimointiprosessilla tarkoitetaan prosessia, jolla sivusto saateetaan luokituksestaan mahdollisimman korkeaksi ja halutuilla hakusanoilla löytyväksi. Prosessissa lähdetään liikkeelle alkutilanteen kartoituksesta, jossa selvitetään nykytilanne ja parannusta vaativat osa-alueet. Kartoituksen jälkeen suoritetaan avainsana-analyysi, jossa pyritään selvittämään paljon käytetyt tuotteita ja palveluita vastaavat avainsanat. Avainsana-analyysi jälkeen suoritetaan itse sivuston optimointi. Tämä optimointi jakautuu tekniseen ja sisällölliseen osaan. Teknisessä osassa korjataan mahdolliset palvelinasetusten aiheuttamat näkyvyysongelmat, epästandardit toteutustavat ja muut lähinnä hakurobotteja häiritsevät asiat. Sisältöosassa muokataan sivut tarvittaessa ulkoasua myöten käyttäjiä ja hakukoneita ajatellen ystävälliseen asuun. Pääosin tässä vaiheessa keskitytään kuitenkin tekstisisällön muuttamiseen sellaiseksi, että se sisältää kehitellyt avainsanat. Kun sivustot ovat kunnossa, suoritetaan linkkianalyysi. Linkkianalyysissä etsitään toimialaan liittyviä maksullisia ja ilmaisia linkkihakemistoja ja lisätään sivusto niihin sekä julkisiin hakemistoihin. Linkityksen jälkeen sivusto on valmis lisättäväksi hakukoneiden indeksointilistoihin ja seurannan asetantaan. Internetmarkkinoinnin nopeasti muuttuvien tilanteiden vuoksi optimointiprosessi on iteratiivinen eli prosessissa edetään vaiheittain, jonka jälkeen työ aloitetaan uudelleen (kuva 4). Jatkuvan kehityksen avulla sivustoista muokkautuu tuloksista saatujen viitteiden avulla entistä optimaalisemmat ja vaatimuksiin vastaavat.



Kuva 4: Hakukoneoptimoinnin etenemisprosessi.

Seuraavissa luvuissa käsitellään kuvan 4 prosesseja, hakukoneoptimoinnin etiikkaa sekä pohditaan lähdemateriaalin avulla prosessin luonnetta.

4.1 Avainsana-analyysi

Avainsana-analyysissä selvitetään hakutermit ja hakulausekkeet, joilla sivuston halutaan näkyvän hakukoneiden hakutulostauksissa. Oikeiden avainsanojen tunnistaminen on merkittävä osa hakukoneoptimointia ja sillä on myös laajempaa vaikutusta internet- ja hakukonemarkkinointiin esimerkiksi sponsoroitujen linkkien muodossa. Avainsanojen ha-

kuun voidaan käyttää menetelmiä perinteisestä aivoriihestä maksullisiin hakusanatyökaluihin, jotka tilastoivat käytettyjä hakusanoja ja ehdottavat aihealueella paljon käytettyjä hakulausekkeitä. Avainsana-analyysi on usein aliarvioitu ja liian nopeasti läpikäyty prosessi. Usein ajatellaan, että tiedetään, mitä ihmiset hakevat etsiessään sivustoa tai että sivuston löytyminen esimerkiksi tuotenimellä on riittävää. Näissä tapauksissa asiakas kuitenkin jo yleensä tietää, mitä hakee. Potentiaalisia asiakkaita, jotka ovat tunnistaneet tarpeensa, mutta eivät tunne tuotteita tai palveluita, ei huomioda ja merkittävä osa asiakaskunnasta jää huomiotta. Toimialan tai hakukoneiden asiantuntijan on siis osattava astua tavallisen kuluttajan saappaisiin ja pohdittava asiaa heidän näkökulmastaan. [10]

4.1.1 Avainsanojen kehittäminen

Avainsanojen kerääminen voidaan aloittaa esimerkiksi aivoriihellä, jossa kerätään yhdessä mieleen tulleet potentiaaliset hakusanat. Aivoriihen aikana on hyvä miettiä seuraavia asioita:

- Etsitäänkö tuotetta/palvelua yleiskielellisin vai teknisin termein?
- Mitkä tuotteet/palvelut tuovat parhaat voitot ja mitkä tuotteet halutaan mieluiten löytyvän?
- Mitä avainsanoja kilpailijat käyttävät sivuillaan?
- Millaisin hakusanoin potentiaalinen asiakas etsii kilpailijan tuotteita?

Keksityistä sanoista muodostetaan muutamasta kymmeneen sanaan oleva ydinjoukko, joista muodostetaan hakulausekkeet. Aluksi ydinjoukon sanat voidaan taulukoida niin, että avaimeksi asetetaan yksittäinen ydinjoukon sana, jolle etsitään synonyymit, vastaavat tuoteryhmät ja tuotteet. Kun yksittäiset sanat on löydetty, kehitetään niiden ympärille fraaseja, joissa käytetään sanan kanssa usein käytettäviä adjektiiveja, mittayksiköitä ja muita määreitä. Taulukossa 4 on esitetty taulukkomalli ydinsanojen laajennukseen sekä fraasien luomiseen. [10]

Fraasien kehittelyn jälkeen listaa laajennetaan käyttämällä hakukoneiden tarjoamia avainsanojensuositusyökaluja. Avainsanojensuositusyökalut ovat tarkoitettu lähinnä sponsoitujen linkkien hakusanojen etsintään, mutta ne toimivat yhtä hyvin myös muuhun avainsanojen keräämiseen. Suositusyökaluille syötetään kehitellyt sanat ja fraasit ja työkalu ehdottaa syötteitä lähellä olevia hakusanoja. Hakusanasuositelijoiden toiminta perus-

Taulukko 4: Esimerkki hakusanojen laajennuksesta. [17]

Tuote	Määre	Ominaisuus	Kuvaaja	Toiminto
Kamera	halvin kamera	nopein digikamera	megapikselin kamera	Osta kamera
...

tuu hauista kerättyihin lausekkeisiin, joten saatu tieto on arvokasta, koska sitä kautta on mahdollista saada tietoa todellisista hakumääristä ja hakusanoista, joiden avulla on helppo päästä lähemmäs normaalikuluttajaa. Useimmat hakukoneiden tarjoamat avainsana-apuohjelmat ovat ilmaisia ja koskevat vain kyseistä hakukonetta. Tarkempiin tuloksiin päästään kuitenkin esimerkiksi maksullisella WordTracker-palvelulla. WordTracker on vastaavanlainen avainsanaohjelma, mutta se käyttää usean hakukoneen lausekkeitä, osaa kertoa sen hetkiset paljon käytetyt sanat ja kertoo myös tietoa hakusanasta kiinnostuneista kilpailijoista. [22]

4.1.2 Kehittelyperiaatteita

Tässä vaiheessa analyysiä ollaan jo pitkällä, mutta saavutettuihin tuloksiin voidaan vielä vaikuttaa pohtimalla, millaisia kävijöitä sivuille halutaan ja millaisin sanoin heidät saavutetaan. Toisille sivuille sopivat hyvin laajat termit, kun taas toisille sivuille ei haluta suuria kävijämääriä vaan halutaan tavoittaa palvelusta kiinnostunut maksava yleisö.

Jos hakufraasit ovat yksisanaisia ja käytetyt termit ovat laajoja, saadaan sivuille runsaasti kävijöitä. Ongelmana on kuitenkin, että pääosa kävijöistä todennäköisesti vain klikkaa sivua ja siirtyy takaisin hakusivulle todettuaan, että sisältö ei ollut se, mitä he etsivät. Laajojen termien ohella on varottava myös moni tarkoituksellisia sanoja ja mietittävä termejä luvussa 2 kerrottujen hakijan aikomusten pohjalta. Esimerkiksi termillä “Java” hakukoneet eivät osaa lukea hakijan ajatuksia ja ohjelmointikielen, teen ja saaren välille ei kontekstin puuttuessa tehdä eroa. Kunnostusyrityksen optimoidessa sivunsa sanalle “kerrostaloasunto” klikkausten määrä varmasti kasvaa, mutta pääosan aikomuksena kuitenkin on asunnon osto tai vuokraus eikä niinkään kunnostus. [23]

Liian laajojen hakufraasien valintaa ei kuitenkaan kannata pelätä liian paljon, koska liian tarkat ja pitkät fraasit rajaavat myös potentiaalisia asiakkaita. Tehottomiin ja vähän osu-
mia saaviin hakulausekkeisiin törmätään usein myös käännettäessä mainoskampanjoita eri kielille sekä olemalla huomioimatta esimerkiksi maiden nimien tai sanojen eri kirjoit-

tusasuja eri maissa. Jos kuluttajakäyttäymistä käsittelevät sivut optimoidaan pelkästään fraasille “consumer behaviour”, jää suuri osa potentiaalisista vierailijoista ulkopuolelle yleisemmän “behavior”-ulkoasun vuoksi. [10]

Luvussa 2 hakijat jaoteltiin heidän aikomustensa mukaan. Hakijoiden aikeet ja asiat, joita sivusto palvelee, vaikuttavat toimivien hakusanojen valintaan. Aikeet ovat tunnistettavissa hakufraaseista niiden tarkkuuden ja eri aikeisiin usein liitettävien sanojen perusteella. Tietokonekaupan ei kannata hankkiutua listan kärkeen fraasein “kannettavat tietokoneet arvostelut” eikä välttämättä myöskään “kannettavat tietokoneet”. Näissä tapauksissa käyttäjä usein etsii vain informaatiota tai on vasta suunnittelemassa ostoa. Tällaiset fraasit sopivat paremmin informaatiopitoisille sivuille, jotka rahoittavat itsensä esimerkiksi mainosin. Kun hakutermeiksi laitetaan “Acer X11 kannettava hinta”, on ostajalla jo todennäköisesti ostaminen mielessä tai ainakin hän suorittaa hintavertailua eri kauppojen välillä. Laitevalmistaja voi taas pitää palvelunsa tason korkealla hakeutumalla kärkipäähän esimerkiksi termillä “Acer ohjekirja”. Runsaasti käytetyn ja muutoin hyvänkin hakusanan kohdalla kannattaa siis miettiä, vastaako sivusto haun taustalla olevaan kysymykseen. Kapeilla markkinoilla ja vähän kilpailuissa maissa kuten Suomessa kannattaa kuitenkin harkita näiden “väärille poluille” johtavien avainsanojen käyttöä. Jotta oikeiden avainsanojen tehoa ei vähennetä, voidaan hakusanalista asettaa prioriteettijärjestykseen ja käyttää kaikkia sanoja, mutta vähemmissä määrin ja vähemmän merkitsevillä paikoilla. [10]

4.2 Sivuston optimointi

Hakukoneoptimointi käsitetään usein pelkkänä sivuston sisällön muokkailuna ja teknisenä kikkailuna, vaikka se sisältää paljon muutakin. Tässä luvussa käsitellään koko hakukoneoptimoinniksi miellettyä sivuston muokkausta sekä sisällöllisestä että teknisestä näkökulmasta. Optimoinnista mielenkiintoisen tekee se, että itse toteutus on yksinkertaista, mutta vaikuttavat tekijät ovat vaikeasti pääteltävissä sekä vaihtelevat jatkuvasti. Seattlilainen hakukoneoptimointiin erikoistunut yritys Seomoz teetti 12 alan ammattilaisella kyselyn, jossa he arvioivat eri asioiden vaikutusta sivun luokitukseen [24]. Liitteessä 1 on esitelty kyselyn tulokset tärkeimpien ja palvelinkohtaisten vaikuttajien osalta [24]. Ne antavat suuntaa eri toimintojen hyödyllisyydestä sekä toimivat hyvänä tarkistuslistana optimointia tehtäessä.

4.2.1 Tekninen optimointi

Teknisellä optimoinnilla tarkoitetaan sivun HTML-koodin muokkaamista hakukoneystävälliseksi, jotta sivun luokitus saadaan mahdollisimman hyväksi. Tässä luvussa käsitellään yleisimmät optimoinnin kohteet, joiden on todettu vaikuttavan merkittävästi luokitukseen.

Sivuston tekemistä suunniteltaessa yksi ensimmäisistä päätettävistä asioista on sivuston osoite. Suomalaisten yritysten tapauksessa fi-päätteen saamisen vaatimukset rajaavat melko paljon optimointia osoitteen suhteen [25]. Muun päätteen tullessa kyseeseen kannattaa hankkia myös vähintään rinnakkais- tai jopa pääosoitteeksi aiheeseen liittyvä osoite. Esimerkiksi kalastusvälineitä myyvän toiminimi Soikkelin kannattaa sijoittaa osoitteen *www.tmisoiikkeli.fi* lisäksi osoitteeseen *www.kalastus.com*. Haettaessa hakusanalla kalastus osoite *www.kalastus.com* nostaa jo luokitusta merkittävästi. Samaa taktiikkaa on myös suositeltavaa käyttää tiedostonimissä ja hakemistoissa. Sen lisäksi, että kuvaavat nimet ovat miellyttävämpiä käyttäjille, ne vaikuttavat luokitukseen.

Tutkittaessa sivun HTML-koodia lähes ensimmäisenä vastaan tulee sivun otsikkotiedot (header) ja niiden sisällä otsikkotunniste (title). Otsikkotunniste ilmenee käyttäjälle selaimen sivun otsikkona sekä hakutulosten otsikkoina. Hakukoneet siis käyttävät otsikkotunnistetta tuloslistauksissaan, joten se on merkittävässä roolissa vierailijoiden houkuttelessa. Tässä vaiheessa on hyvä tarkistaa, että otsikko on määritelty, mutta hyvän otsikon kehitystä käsitellään sisällön optimoinnissa. Kuten Seomozin kyselyn tuloksetkin (liite 1) osoittavat, otsikkotunnisteen vaikutus sivuston luokitukseen on suuri.

Seuraava optimoinnin kannalta mielenkiintoiset asiat koodissa ovat myös otsikkotiedoissa sijaitsevat metatiedot (meta tags). Useat sovellukset käyttävät hyödykseen erilaisia metatietoja, mutta 90-luvun puolivälissä Inktomin HotBot-hakukone toi mukanaan kaksi hakuihin liittyvää metatietoa, avainsanat (keywords) ja kuvaus (description). Avainsanojen ideana oli mahdollistaa sivustojen ylläpitäjille helppo tapa ilmoittaa hakukoneille sivuston aiheet ilman, että hakukoneen tarvitsee tutkia sisältöä. Kuten arvata saattaa, avainsanat katosivat lähes välittömästi hakukoneiden luokitusfunktioista väärinkäytön takia [26]. Kuitenkin muun muassa kotimainen Ihmemaan haku ilmoittaa käyttävänsä avainsanoja hyödykseen [27]. Avainsana-analyysissä laadittujen avainsanojen listausta kannattaa harvita, mutta on otettava myös huomioon, että kilpailija voi säästää huomattavasti aikaa omassa optimoinnissaan kopioimalla avainsanat metatiedosta. Jäljempi metatieto eli kuvaus on edelleen tuettu jossain määrin hakukoneiden toimesta. Alun perin kuvaus luotiin hakutuloksissa esiintyvää kuvausta varten. Väärinkäytösten vuoksi vain osa hakukoneis-

ta hyödyntää edelleen kuvausta ja myöskin Google hyödyntää osittain tietoa luodessaan kuvauksia. Optimoinnin kannalta kuvauksesta ei suurta hyötyä ole, mutta se auttaa hakukoneita luomaan houkuttelevan kuvauksen sivulle. [28]

Kun otsikkotiedot on käyty läpi, voidaan siirtyä sivun sisältöosiin eli runkoon (body). Tärkeintä rungossa on tarkistaa standardien mukainen toteutus sekä HTML:n muotoiluun tarkoitettujen elementtien käyttö. Useat sivustojen tekniset toteuttajat käyttävät mieluummin CSS:n (Cascading Style Sheets) luokkia muotoillakseen leipätekstin lisäksi otsikot ja tarvittavat korostukset ilman, että niissä hyödynnetään hakukoneiden paljon arvostamia HTML:n otsikkoelementtejä. Näin sivusto ei siis saavuta parasta mahdollista luokitusta. Asia voidaan korjata muuttamalla otsikot käyttämään HTML-standardin mukaisia otsikkoelementtejä. Otsikoiden lisäksi on kiinnitettävä huomiota muihinkin sivuelementteihin. Monet niistä kuten kuva- ja linkkielementit sisältävät käytettävyyttä helpottavia otsikko- ja kuvausmääreitä. Käytettävyyden paranemisen lisäksi määreisiin saadaan myös helposti toistoa avainsanoille. Listauksia elementeistä ja niiden oletetuista vaikutuksista löytyy myös liitteestä 1.

Standardien mukainen toteutus lisää sivuston käytettävyyttä sekä tekee sivujen indeksoinnin helpommaksi. Mikäli hakukoneoptimoiija ei itse myös tuota sivujen HTML-koodia, jää hänen vastuulleen lähinnä tarkastaa noudattaako koodi standardeja. Standardeista vastaava W3C (World Wide Web Consortium) tarjoaa kotisivuillaan sekä CSS-tyylimääritteiden että HTML-koodin validointityökaluja, joiden avulla voidaan nopeasti tarkastaa onko sivut toteutettu standardien mukaisesti. Kun sivuston tekninen toteutus on kunnossa, voidaan siirtyä optimoimaan sisältöä.

4.2.2 Sisällön optimointi

Hyvin tehdyn teknisen optimoinnin jälkeen vanha sisältö on helposti muokattavissa optimoituun muotoon. Sisällön optimoinnilla tarkoitetaan sivuston tekstisisällön tuottamista tai muokkaamista sellaiseksi, että se sisältää hakusana-analyyysissä esille tulleita sanoja ja niiden yhdistelmiä.

Luvussa tekninen optimointi muokattiin useita eri paikkoja valmiiksi sisällöntuotantoa varten. Sivusto käydään läpi eri osa-alueittain ja lisätään hakusanat niille sopiviin paikkoihin. Samalla saavutetaan myös usein näkyvyyden lisäksi muita etuja. Tutkimusten mukaan ihmiset eivät näe internetissä olevaa tietoa yhtä relevanttina kuin kirjoitettua, vaan teksti luetaan enemmän silmäillen Nielsenin F-teorian mukaisesti [29]. Kirjoitetun teks-

tin lukeminen on myös noin 30 prosenttia nopeampaa kuin ruudulla olevan. Lisäämällä hakusanoja sisältöön teksti usein tiivistyy ja mielenkiintoiset asiat löytyvät helpommin hakusanojen avulla.

Lisäämällä runsaasti avainsanoja ympäri sivua sijoitukset hakukoneissa voivat kohota merkittävästi. Järki on kuitenkin pidettävä kädessä ja sivut on suunniteltava ensisijaisesti kävijöitä varten. Mikäli hakusanoja on useita, voidaan ne jakaa esimerkiksi eri tuotteille ominaisiin sanoihin ja käyttää tuotekohtaisilla alisivuilla.

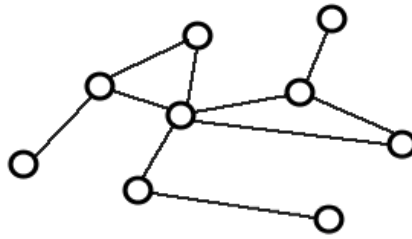
4.3 Linkkianalyysi

Linkkianalyysissä kartoitetaan sivustolle tällä hetkellä tulevat linkit, etsitään uusia potentiaalisia linkittäjiä sekä lisätään sivusto yleisiin sekä toimialakohtaisiin hakemistoihin. Linkkianalyysin tärkeyttä voidaan perustella seuraavassa esiteltävällä verkkoteoriallakin, mutta useat hakukoneet käyttävät sisääntulevien linkkien määrää luokitusperusteena ja kasvattavathan internetin keskeisiltä sivuilta tulevat linkit näkyvyyttäkin. Tässä luvussa käsitellään linkkien vaikutusta siihen, miksi tietyt sivut tulevat esiin liikuttaessa missä päin tahansa internetiä, miten hakukoneet suhtautuvat linkkeihin ja miten sivustolle saadaan sisäänpäintulevia linkejä.

4.3.1 Verkkoteoria

1700-luvun loppupuolella Leonhard Euler kehitti verkko- ja graafiteorian. Verkkoteoriasa asiat kuvataan kuvan 5 mukaisesti solmuina ja niiden välisinä reitteinä. Verkkoesityksen etuna on, että verkko voidaan irroittaa itse asiasisällöstä ja tätä kautta laajentaa asian ymmärtämistä takertumatta asiakohtaisiin seikkoihin. Useiden luonnonilmiöiden lisäksi esimerkiksi sosiaaliset suhteet ja internet noudattavat tätä teoriaa. Internet noudattaa niin sanottua skaalavapaata teoriaa, jossa verkko kasvaa yksittäisestä solmusta rajattomasti. [8, 30]

1900-luvulla tutkimukset osoittivat, että ominaista luonnollisille verkoille kuten internetillekin on reittien muodostamien linkkien epätasainen jakaantuminen. Esimerkiksi ihmisten ollessa kyseessä pääosa ihmisistä tuntee Yhdysvaltojen presidentin, mutta lappeenrantalaisen koodarin tuntee vain harva. Linkkien eksponentiaalista jakaumaa vahvistaa myös uusien solmujen mukaan tuleminen. Verkkoon tulevat uudet solmut menettelevät usein suosivasti valitessaan ulospäin lähtevien linkkiensä kohteita. Syntyvä lapsikin oppii tun-



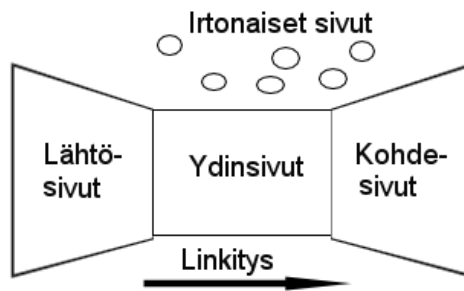
Kuva 5: Verkko.

temaan presidentin todennäköisemmin kuin jonkun paikallisen tekijän ja samoin myös tämän lapset. Linkkien määrä siis jo suosituilla kohteilla kasvaa eksponentiaalisesti, kun muilla kasvua ei välttämättä tapahdu ollenkaan. [30, 8]

Viime vuosikymmenen lopulla Altavista, Compaq ja IBM laajensivat teoriaa internetin osalta kehittämällä niin sanotun bow-tie -teorian. Sen mukaan internet koostuu neljänlaisista sivuista [10]:

- Ydinsivut: Internetin keskeiset sivut, joihin linkitetään runsaasti ja jotka myös linkittävät runsaasti ulkopuolisiin sivuihin.
- Lähtösivut: Sivustoja, jotka linkittävät ydinsivuihin, mutta ydinsivut eivät linkitä niihin. Tällaisia sivuja ovat laadultaan huonot sivut ja uudet sivut, jotka eivät ole vielä nousseet ydinsivuiksi.
- Kohdesivut: Korkeatasoisia sivuja, jotka sisältävät runsaasti informaatiota ja palveluita. Esimerkiksi ydinsivustojen keskitinsivustot kuten hakemistot ja linkkisivut linkittävät usein näille sivuille.
- Irtonaiset sivut: Sivuja ja sivustoja, jotka eivät linkitä eivätkä ole linkitetty muista ryhmistä. Ryhmään kuuluu sekä tarkoituksen mukaisesti ulkopuolella olevia sivustoja että sivustoja, joissa internetmarkkinointi on unohdettu.

Kuvassa 6 on esitetty yllämainittujen sivujen suhteet toisiinsa. Lähtösivut ovat linkkikoelmia ja yksittäisten ihmisten sivuja, joille ei juurikaan linkitetä, mutta ne tarjoavat linkejä yleisille sivuille. Ydinsivut muodostavat keskeiset sivut ja ne ovat kuvan mukaisesti molemminsuuntaisessa yhteydessä muiden sivujen kanssa. Kohdesivut ovat suurten yritysten sivuja ja paljon tietoa sisältäviä sivuja, joiden ei edes tarvitse linkittää muihin



Kuva 6: Internetsivujen bow-tie -teorian mukainen jakautuminen. [10]

ollakseen esillä. Tietoisesti tai tahtomattaan osat sivuista ovat irtonaisia sivuja, jotka eivät ole yhteydessä muihin sivuihin.

Hakukoneoptimoinnin kannalta verkkoteorian perusteiden tunteminen on tärkeää, koska linkkien määrä on, jos ei aivan suoraan verrannollinen niin hyvin vaikuttava tekijä kävijämääriin. Linkkejä kerätessä on huomioitavaa, että yksittäinen linkki ydinsivuilta on tärkeämpi kuin useat linkit irtonaisilta sivuilta. Tämän asian osaavat nykyään huomioida myös hakukoneet sivustojen luokitusta laskettaessa.

4.3.2 Linkitys ja hakukoneet

Hyvän sijoituksen edellytyksenä hakukoneiden tulostauksissa on oikean sisällön ja standardin koodin lisäksi korkea luokitus. Ajan mittaan hakukoneyritykset ovat huomanneet, että pelkästään sivulla olevan sisällön mukaan tehtävä luokitus on altis väärinkäytöksille ja hakutulokset ovat tätä kautta manipuloitavissa. Sen vuoksi nykyiset hakukoneet ja osa hakemistoistakin luokittelevat sivustot sivustolle osoittavien linkkien määrän ja laadun mukaan. [31]

90-luvun loppupuolella Brin ja Page esittelivät ensimmäistä kertaa Googlen julkaisussaan “The Anatomy of a Large-Scale Hypertextual Web Search Engine” [32]. Merkittävää julkaisussa oli tapa, jolla nykyäänkin hyödynnetään sivulle osoittavien linkkien määrää luokituksessa. Perusteina luokitustavalle oli sen riippumattomuus manipuloinnista ja sen kyky osoittaa käyttäjille tärkeät eli enemmistön hyödyllisiksi katsomat sivut. Akateemisessa maailmassa vastaavanlaista tapaa on käytetty mitattaessa julkaisun arvokkuutta laskeamalla julkaisuun viittaavat sitaatit. Internetissä etuna on, että sivustojen välisistä suhteista

voidaan muodostaa koneellisesti karttoja, joiden avulla voidaan laskea pitempiä yhteyksiä sivustojen välillä. Google-julkaisussa esitetty PageRank-luokittelualgoritmi ei siis pelkää laskea sivulle osoittavia linkkejä vaan normalisoidun arvon sivulle osoittavien sivujen luokituksista, mikä taas kertoo käsiteltävän sivun luokituksen. PageRank-arvo voidaan yksinkertaisimmillaan laskea seuraavasti [32]:

$$PR(A) = (1 - d) + d(PR(T1)/C(T1)) + \dots + PR(TN)/C(TN)). \quad (1)$$

Kaavassa 1 oletetaan, että sivulle A osoittaa sivut T1-Tn. Funktio C(X) kertoo sivulta X lähtevien linkkien määrän. Käsiteltävän sivun PageRank-arvoksi muodostuu siis sille osoittavien sivujen PageRank-arvo ja osoittavien linkkien osamäärien summa. Manipuloinnin estämiseksi kaavaan on lisätty satunnainen personointitekijä d, joka kuvaa ihmisten todennäköisyyttä siirtyä linkkiketjun kautta käsiteltävälle sivulle. [32]

Vaikka aiemmin esitelty luokitus on vaikeasti manipuloitavissa, nykyiset luokitusmenetelmät ovat huomattavasti monimutkaisempia. Ne sisältävät useita d:n tapaisia satunnaisista ja sivukohtaisista asioista riippuvia muuttuja. Jotta hakukoneet säilyttävät riippumattomuutensa ja tehokkuutensa kilpailijoihin verrattuna, ovat luokitusalgoritmit tarkkaan varjeltuja liikesalaisuuksia. Muutamia vaikuttavia tekijöitä niistä on kokemuksen kautta selvitetty. Useat hakukoneet osaavat yhdistellä sisältöä aiheisiin ja luokittaa linkin sisältöjen vastaavuuden mukaan. Myös sivulta lähtevillä linkeillä on huomattu olevan vaikutusta. Runsaat linkitykset alan korkeasti arvostettuihin sivuihin voivat nostaa sivun statuksen keskitinmäiseksi informaatiokeskukseksi, joiden on todettu olevan arvostettuja käyttäjien piirissä. Huomioista riippumatta parhaaksi toimintatavaksi on osoittautunut rehellinen mahdollisimman hyvien sivujen tuottaminen, joka on avain sisääntulevien linkkien houkutteluun ja kohti internetin ydinsivuryhmää siirtymiseen. [31]

4.3.3 Linkkien kerääminen

Paras ja luonnollisin tapa hankkia ja saada ulkopuolisia linkkejä omalle sivulle on luonnollisesti laadukkaan sisällön tuotanto. Kaikki sivustot eivät kuitenkaan luonteeltaan ole informaatiolähtöisiä, joten linkkejä on hankittava myös muilla tavoin, jotta kävijät löytävät sivuille ja sivusto näkyy korkealla hakukonelistauksissa. Yleisimpiä menetelmiä ulkopuolisten linkkien hankkimiseen ovat linkityksen pyytäminen, linkkien käyttö useissa paikoissa ja ilmoittaminen ilmaisiin ja maksullisiin hakemistoihin.

Arvokkaita linkkien lähteitä ovat erilaiset sidosryhmät. Harva asiakas, alihankkija tai muu

sidosryhmään kuuluva automaattisesti linkittää omia sidosryhmiään, mutta pyydettyä linkki yhteistyökumppaniin usein järjestyy. Samalla tavalla on mahdollista menetellä myös toimialan informaationsivustojen ja keskustelupalstojen kanssa. Jos sivustosi tarjoaa hyödyllistä tietoa aiheesta, ottavat ylläpitäjät mielellään sinut linkkikokoelmiinsa. Potentiaalisia linkittäjiä onkin helppo mieltä astumalla itse ulkopuolisen ylläpitäjän saappaisiin ja pohtien, mitä hyötyä ylläpitäjä saa linkin lisäyksestä. Sivuston ylläpitäjä tai ulkopuolinen hakukoneoptimoija voi myös itse lisätä linkkejä esimerkiksi liittämällä linkin keskustelupalstojen nimikirjoitukseensa. Tässä tapauksessa voi myös itse vaikuttaa näiden linkkien tuomiin kävijöihin laadukkain ja informaatiopitoisin kirjoituksin. [31]

Internetin hakemistot ovat hakukoneiden esiasteita. Hakemistoilla tarkoitetaan suuria linkkikokoelmia, joissa linkit on jaettu kategorioittain. Vaikka hakukoneet käsittelevät huomattavasti suurempia sivumääriä, ovat hakemistot säilyttäneet osittain asemansa niiden laadukkaiden linkkien ja helpon käytettävyyden vuoksi. On otettava kuitenkin huomioon, että hakemistojen käyttäjät ovat jakautuneet kuten hakukoneidenkin. Yagoon hakemistot keräävät runsaasti käyttäjiä, kun taas vähemmät tunnettujen käyttö on hyvin vähäistä. Hakemistoja löytyy moneen lähtöön. Osat ovat keskittyneet vain tiettyihin toimialoihin tai maanosiin ja osat ovat taas yleisiä hakemistoja. Jotkut hakemistoista ovat maksullisia ja jotkut ilmaisia. Eettisesti oikein toimivat ilmaiset linkkihakemistot ovat hyödyllisiä erityisesti uusille sivuille, vaikeivat ne asiakkaita toisikaan. Hakukoneet arvostavat internetin niin sanottuja keskitinsivustoja ja näin ollen niiltäkin linkitettyjen sivun luokitus kohoaa. [33]

Maksullisten hakemistojen kohdalla linkitys ei käy niin yksioikoisesti, vaan on tutkittava, onko paikka hakemistossa hintansa arvoinen. Kertamaksullisten hakemistojen ollessa kyseessä saatuja hyötyä on vaikea arvioida ennen kuin hakemistoon on itse lisännyt useamman sivuston. Taustatutkimusta voi kuitenkin tehdä etsimällä tietoja hakemistosta keskustelupalstoilta tai hakukoneista. Apuna voi myös käyttää esimerkiksi Alexa Web Searchin liikennemääräarvioita, joista voi arvioida hakemiston kuukausittaisia kävijämääriä [34]. Vuosittain maksettavissa hakemistolistauksissa päästään laskemaan jo tuottoasteitakin käyttämällä apuna seurannasta saatuja tietoja. Kun hakemistolistauksen taustalla on maksimaalisen luokituksen tavoittelu, myös listauksen epäsuoria vaikutuksia kannattaa tutkia. Suoraa luokituksen kohotusta on mahdotonta päätellä, mutta seuraavat asiat selvittämällä päästään suuntaa-antaviin tuloksiin [35]:

- Ovatko linkit suoria? Jos linkit kulkevat mainospalvelimen kautta tai on toteutettu jollain skriptikielellä, ei luokitus kasva.
- Käsitteleekö hakemistosivu aihealuetta ja saatko itse päättää linkkitekstin? Kuten

aiemmin todettiin, hakukoneet osaavat huomioda kontekstin.

- Montako linkkiä sivulla on? Aiemmin esitetyn PageRank-algoritmin mukaisesti linkkien määrän kasvaessa saatu luokitus laskee.
- Ketkä linkittävät hakemistoon ja mikä on hakemistosivun luokitus?

Kokemus on linkkien keräämisessäkin merkittävässä roolissa. Useita optimointeja tehneelle hakukoneoptimoijalle on kasautunut laaja lista eri alojen potentiaalisista linkittäjistä, hän tuntee julkiset hakemistot ja tietää, mitkä maksulliset palvelut ovat hintansa arvoisia. Tässä luvussa on painotettu runsaasti linkkien merkitystä luokituksen parannuksessa. Linkkejä kerätessä on kuitenkin muistettava, miksi linkit ovat pohjimmiltaan olemassa. Linkkien idea ei ole vain nostaa sivuston luokitusta vaan johdattaa kävijät halutun tiedon luo ja tätä kautta myös luonnollisesti nostaa kävijämääriä. Asiakasta ajatellen luotu sivu voittaa - olipa luokitus millainen tahansa.

4.4 Indeksointi

Hakukoneoptimoinnissa indeksoinnilla tarkoitetaan sivustojen lisäämistä hakukoneiden indekseihin tai sivustojen ilmoittamista hakukoneille. Kuten luvussa 3 tuli ilmi, hakurobotit eivät löydä läheskään kaikkia internetin sivustoja. Vaikka sivusto olisi optimoitu kuinka hyvin, koko työ menee hukkaan, mikäli sivustoa ei löydy hakukoneiden indekseistä. Tässä luvussa käsitellään sitä, miten sivut saadaan indeksoitua, miksi sivut eivät ole välttämättä indeksoitu ja miten seurata indeksointia.

4.4.1 Lisääminen hakukoneisiin

Kun uusi sivusto avataan internetiin, voi sen hakukoneista löytymiseen kulua pitkäkin aika. Indeksoitumista voidaan kuitenkin nopeuttaa ilmoittamalla hakukoneelle uudesta sivusta. Ennen ilmoittamista on kuitenkin hyvä tarkistaa löytyvätkö sivut hakukoneista. Tarkistaminen voidaan tehdä esimerkiksi suorittamalla hakuja yrityksen nimellä. Totuudesta riippuen voi kuitenkin olla niin, että sivustoa ei löydy yrityksen eikä tuotteiden nimelläkään ja indeksoituminen jää huomaamatta. Varmuus indeksoinnista saadaan paremmin käyttämällä yleisimpien hakukoneiden tukemaa "site"-toimintoa. Kun hakufraasiksi laitetaan esimerkiksi

site : www.yritys.fi,

hakukone etsii sisältöä pelkästään osoitteesta *www.yritys.fi*. Ilman muita ehtoja haku siis palauttaa kaikki hakukoneen kyseisestä osoitteesta indeksoimat sivut.

Jos sivustoa ei etsinnöistä huolimatta löydy, ei hakurobotti joko ole lisännyt sivustoa indekseihin tai muuten vain löytänyt sitä. Tällöin apuna voidaan hakukoneiden tarjoamia sivuston lisäys palveluita. Palvelut löytyvät hakukoneiden kotisivuilta. Taulukossa 5 on esitetty kolmen yleisimmän hakukoneen sivustojen lisäysosoitteet.

Taulukko 5: Osoitteet, joissa sivuston voi ilmoittaa eri hakukoneisiin.

Hakukone	osoite
Google	http://www.google.com/addurl
MSN	http://search.msn.com/docs/submit.aspx
Yahoo!	http://search.yahoo.com/info/submit.html

Lisäyspalvelut eivät lisää sivuja automaattisesti, vaan ne lisäävät sivut hakurobottien selausjonoihin. Näin robotit pääsevät niillekin sivustoille, joille ei ole ulkopuolista linkitystä. Lisäyksen jälkeen indeksoinnin edistymistä voidaan seurata esimerkiksi palvelinlokien avulla. Tutkimalla selainkenttää (user-agent) -kenttiä voi nimestä päätellä, kuka kävijöistä on hakurobotti. Taulukossa 6 on listattu yleisimmät sivuilla vierailevat robotit.

Taulukko 6: Yleisimpien hakukoneiden hakurobottien nimet.

Hakukone	Hakurobotin user-agent
Google	Googlebot
MSN	Msnbot
Yahoo!	Slurp!

Vaikka robotti käykin sivuilla, ei ole itsestään selvyys, että ne löytyvät indekseistä. Seuraavassa luvussa käsitellään syitä, miksi sivusto ei löydy hakukoneista robottien vierailuista ja lisäyspalvelujen käytöstä huolimatta.

4.4.2 Miksi sivut eivät indeksoitu?

Jos sivusto on ollut olemassa pitkään, on hyvin todennäköistä, että yleisimmät hakukoneet ovat indeksoineet sivuston. Jos näin ei kuitenkaan ole, syy voi olla muun muassa [10]:

- Hakukoneet ovat asettaneet sivuston pannaan (banned).
- Hakurobotti ei pääse liikkumaan sivuilla tai ei löydä sisältöä.

Kaikkein hankalin tilanteista on pannaan joutuminen. Tällöin sivustolla on aikaisemmin käytetty tai käytetään edelleenkin jotain epäeettistä optimointitekniikkaa. Sivustot eivät joudu helposti pannaan, vaan se vaatii toistuvia ja jatkuvia väärinkäytöksiä. Epäeettisten menetelmien käytöstä myös rangaistaan ennen kokonaista indeksistä poistoa. Osana hakukoneoptimoinnin seurantaan on myös hyvä seurata sivustojen indeksointia, jolloin mahdollisesti omista väärinkäytöksiksi tulkituista menetelmistä saadaan hyvissä ajoin tietoa. Koska indeksoinnin suorittavat koneet, on niiden huijaamista harrastettu hakukoneoptimoinnin alkuajoista asti. Näin myös tämän markkinoinnin maine on pilattu useissa piireissä eikä se myöskään ole saavuttanut ansaitsemaansa asemaa tieteen maailmassa. Eri-laisissa huijaustavoissa riittäisi materiaalia vaikka toiseksi diplomityöksi, joten niihin ei tässä yhteydessä sen tarkemmin mennä. Muutamia käytetyimpiä menetelmiä ovat [10]:

- Dynaamisen sisällön muokkaaminen selaintiedon ja verkkotunnuksen mukaan (cloaking).
- Piilotettu sisältö ja linkit.
- Roskapostitus blogeihin ja vieraskirjoihin.

Parempi lähestymistapa aiheeseen on toteuttaa sivut asiakasystävällisesti ja edetä rehellisin keinoin niin pitkälle, kuin sivustolla on yleensäkin rahkeita. Myös hakukoneet ovat ohjeistuksia, joita noudattamalla sivustot eivät joudu pannaan. Esimerkiksi Googlen ohjeistukset ovat saatavilla osoitteesta:

<http://www.google.com/webmasters/guidelines.html>

Osittain aiheelle on myös omistautunut Search Engine Marketing Professional Organization (SEMPO) ja hakukoneoptimoinnin etiikkaa käsittelevä foorumi SEO Code of Ethics. SEO Code of Ethics on myös julkaissut listan säännöistä, joita noudattamalla hakukoneoptimoiija saa sivulleen SEO Code of Ethicsin tunnusmerkin. [36]

Yleisin syy hakuindekseistä puuttumiseen kuitenkin on, että hakurobotit eivät pääse liikkumaan vapaasti sivuilla. Uutta sivustoa luotaessa tällaisia sivuja ei pääse tulemaan, jos aiemmin esiteltyjä sivusto-optimoinnin teknisiä seikkoja on noudatettu. Vanhemmilla sivuilla ongelmia voivat aiheuttaa esimerkiksi [10]:

- Robots.txt, joka estää robottien vierailun. Myös aiemmin palvelimella ollut robots.txt voi aiheuttaa hakurobotin indeksoinnin luovuttamisen.
- Robottiansat eli linkkien toteutus esimerkiksi JavaScript-koodin avulla.
- PopUp-ikkunat, kehykset (Frames), yms.
- Julkaisujärjestelmien aiheuttamat pitkät parametrijonot linkeissä.
- Epävalidi koodi.
- Evästeiden (cookies) “pakotus”.

On myös varmistettava, että pääsivulta pääsee jokaiselle alisivulle yksinkertaisia linkkejä seuraten. Sivukartat parantavat sekä indeksoitavuutta että käytettävyyttä.

4.4.3 Indeksoinnin seuranta

Pienten yritysten sivustojen indeksoinnin seuranta onnistuu helposti jo pelkän “site”-toiminnon avulla. Tuloksista on helppo käydä läpi sivut ja tarkastaa, että koko sivusto on indeksoitu. Suurempien yritysten sivustojen seurantaan tarvitaan jo järeämpiä menetelmiä, jotta voidaan jatkuvasti valvoa indeksoinnin tilannetta. Osana indeksoinnin seuranta saadaan helposti myös selville sivustolle viittaajat. Yleisimmät hakukoneet tukevat link-toimintoja ja esimerkiksi haku

link : www.yritys.fi

kertoo osoitteen *www.yritys.fi* viittaavat sivustot. Viittaajamäärät kertovat oman linkityksen onnistumisesta, sivuston informaatioarvosta sekä mahdollisista saaduista varoituksista. Kun indeksointi on onnistunut, seuranta on tärkeää potentiaalisten varoitusten vuoksi. Vaikka eettistä linjaa noudattaisikin, saatetaan jotkut toimenpiteistä tulkita väärinkäytöksi. Indeksointiseurannan avulla tämä selviää seuraavista asioista [10]:

- Viittaajien määrä romahtaa yhtäkkiä.
- Indeksoitujen sivujen määrä laskee.
- Sivustoa ei löydy hakukoneesta kuin haettaessa sivuston osoitteella.

Seuranta voi näyttää vähäpätöiselle toimenpiteelle, mutta erityisesti hakumarkkinointipalveluita tuottaville yrityksille seuranta on tärkeää jo pelkästään maineen mutta myös kattavien asiakasraporttien tuottamisen vuoksi.

4.5 Seuranta

Jotta minkä tahansa asian onnistumista ja toimintaa voidaan seurata, on asetettava oikeanlaiset mittarit. Hakukoneoptimoinnin seurannassa käytetään apuna sivustojen seurantaan käytettyjä kävijäseurantatyökaluja. Verkkosivujen seuranta voidaan toteuttaa usein eri tavoin. Tässä luvussa käsitellään erilaisten seurantapojen mahdollisuuksia, seurannan kannalta olennaisia tietoja ja tiedon analysointia. Analysoinnissa kiinnitetään myös huomiota markkinointiosastojen tarpeisiin, koska analyysit jätetään usein teknisen puolen toteutettavaksi ja tiedosta ei saada kaikkea olennaista irti.

4.5.1 Mitä?

Verkkosivujen kävijäseurantaan tarkoitetut järjestelmät ovat usein isovelvi valvoo -henkisiä järjestelmiä. Ne pyrkivät keräämään talteen kaiken saatavilla olevan tiedon. Olennaista saatavilla olevaa tietoa voidaan tutkia katsomalla W3C:n (World Wide Web Consortium) kehittämää CLF-lokitiedostoformaattia (Common Log Format) ja sen sisältämiä tietoja. ECLF (Extended Common Log Format) ja sen aiempi versio CLF olivat alun perin kehitetty palvelinkuormien seurantaan, mutta erityisesti ECLF:n laajennusten avulla pystytään suorittamaan kävijäseuranta. ECLF:n mukainen lokirivi sisältää taulukon 7 tiedot [37].

Käyttäjäseurannan kannalta useat ylläesiteltyt kentät ovat turhia. Käyttäjätunnuksilla ei intra- ja extranetseurantaan lukuun ottamatta ole suurempaa merkitystä. Statuksen avulla voidaan lähinnä seurata löytävätkö esimerkiksi hakurobotit sivuilta, rikkinäisiä linkkejä. Jos kävijäseuranta on eroteltu liikennemäärien seurannasta, kuten on suotavaa, kenttää bytes ei tarvita.

Seurannan kannalta olennainen mutta vaikea tehtävä on asiakkaan tunnistaminen ja erottelu. Ohjelmallisesti asiakkaan henkilöllisyyttä ei ilman kyselytutkimuksia tai kirjautumista voida selvittää, eivätkä ne niissäkään tapauksissa ole täysin varmoja. Yksittäisten asiakkaiden erotteluun sopii verkkotunnuksen lisäksi asiakkaan selain. Äkkiseltään kuulostaa, että konekohtaiset verkkotunnukset olisivat tehokas keino asiakkaiden tunnistamiseen. Yritysten ja palveluntarjoajien aliverkot ja sisällönsuodatuspalvelimet estävät yksi-

Taulukko 7: ECLF-lokirivin sisältämät tiedot.

Kentän nimi	Selitys
remotehost	Asiakkaan IP-osoite (Internet Protocol) tai verkkotunnus
rfc391	Asiakkaan käyttäjätunnus asiakkaan palvelimella
authuser	Käyttäjätunnus, jolla asiakas on kirjautunut verkkosivuille
date	Päivämäärä ja kellonaika
request	Käyttäjän lähettämä HTTP-protokollan (HyperText Transfer Protocol) mukainen käsky
status	HTTP-protokollan mukainen käskyn paluuarvo
bytes	Siirrettyjen tavujen määrä
referer	URL (Unified Resource Locator), jolla asiakas oli ennen tämän rivin osoittamaa ladattua sivua(viittaaaja)
user-agent	ohjelma, jota asiakas käyttää tai väittää käyttävänsä käskyn suoritukseen (selain)

löinnin ja organisaatiot näkyvät yksittäisinä kävijöinä, koska verkkotunnus sisältää näiden välipalvelinten osoitteet. Samanlaisia ongelmia aiheuttavat myös internetin toimintaa tehostavat välimuistipalvelimet. Ottamalla huomioon kentän selain saadaan asiakkaat eroteltua hieman paremmin. Jos tulosta halutaan tehostaa, on otettava käyttöön evästeet ja jäljitteet (web beacon), joita käsitellään tässä luvussa myöhemmin. [38]

Pelkkä tieto asiakkaan vierailusta sivustoilla ei usein riitä. On mielenkiintoista tietää myös, mitä asiakas on tehnyt sivuilla ja millä sivuilla hän on vieraillut. Request-kenttä tarjoaa mahdollisuuden tarkempaan tutkimukseen asiakkaan liikkeistä. Erityisesti tiedonhakuun käytetty GET-komento on kävijäseurannan kannalta mielenkiintoinen. Standardin mukaiset GET-pyyntöt näyttävät esimerkiksi seuraavalta:

```
GET /sivu.html HTTP/1.1
```

Komennot noudattavat standardin mukaisesti aina samaa muotoa ja tulkitessa lokitietoa saadaan helposti selville, millä sivuilla asiakas on käynyt. [39]

Hakukoneoptimoinnin ohella useat yritykset käyttävät internetmarkkinointiin maksullisia hakemistoja sekä hakusanamarkkinointia. Ainakin johtoa kiinnostavien tuottoasteiden selvittämiseen viittaajakenttä (referer) tarjoaa arvokasta tietoa. Mikäli käyttäjä on saapunut sivustolle jonkin ulkoisen sivun kautta eikä viittaajatiedonsaantia ole estetty, paljastaa viittaaja, mistä sivulle on saavuttu. Esimerkiksi viittaajan sisältäessä tiedon

<http://www.google.fi/search?hl=fi&q=click+tracking>

on asiakas saapunut sivustolle käyttäen googlea ja hakufraasia “click tracking”. Viittajätieto toimii myös siirryttäessä omalla sivustolla sivulta toiselle, jolloin analysoimalla istuntokohtaisia lokitietoja saadaan selville myös asiakkaan kulkureitit sivuilla. [39]

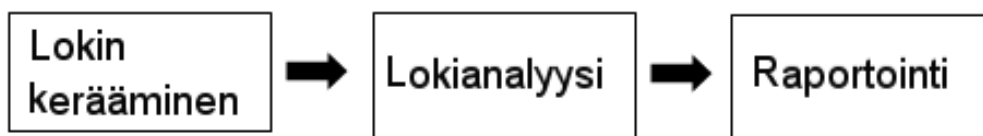
4.5.2 Miten?

Nykyisin kävijäseurantaan on tarjolla useita eri mahdollisuuksia. Korkeimmalla tasolla seurannan valinnassa voidaan päättää käytetäänkö ostettua omalle palvelimelle asennettavaa ohjelmaa, ostetaanko seuranta palveluna kolmannelta osapuolelta vai toteutetaanko järjestelmä itse. Seurantajärjestelmät voidaan jakaa seuraaviin pääryhmiin: lokianalysointiohjelmat, evästeisiin tai jäljitteisiin perustuvat ohjelmat ja verkkoliikenteenvalvontaohjelmat. Tässä luvussa käsitellään eri vaihtojen heikkouksia ja vahvuuksia.

Tehtäessä päätöstä järjestelmän toteutuksesta tärkeimpiä kriteerejä ovat järjestelmän joustavuus, toimivuus ja kustannukset. Mikäli yrityksellä ei ole osaamista eikä järjestelmää myydä eteenpäin, palvelun hankkiminen ulkopuolelta on vaihtoehtoista paras. Käyttämällä ulkoista palveluntarjoajaa säästetään kustannuksissa ja järjestelmä pysyy ajantasalla. Useat palveluntarjoajat antavat järjestelmänsä myös koekäyttöön, jolloin järjestelmää voidaan testata omassa ympäristössä ja tutustua raportointimahdollisuuksiin. Jos palvelua aiotaan myydä jatkossa myös omille asiakkaille, voi järjestelmän kehitys olla oikea vaihtoehto. Itse toteutettu seuranta on helposti integroitavissa muihin palveluihin ja muokkamalla raportointia saadaan selville juuri ne asiat, jotka koetaan tärkeinä. Suurista kehityksistä ja ylläpitokustannuksista huolimatta järjestelmä voi maksaa itsensä takaisin ja tarjota jopa kilpailuedun laajempia kokonaisuuksia käsitellessä.

Seurantajärjestelmien toteutustavoista perinteisin on WWW-palvelinten lokien analysointi. WWW-palvelimet keräävät edellisessä luvussa esitetyn lokimallin mukaisesti tietoa. Haluttaessa tietoa kävijöistä loki syötetään analysointiohjelmalle, joka koostaa tiedoista ennalta määrätyn raportin. Kuvassa 7 on esitetty lokianalysointiohjelman toimintaprosessi. Lokianalyysin heikkoina puolina voidaan pitää lokien vaikeaa tulkintaa, tietojen epätarkkuuksia, suurta herkkyyttä välityspalvelimille ja raportointi ei yleensä ole reaaliaikaista. [40]

Evästeet ovat kehitetty verkkokauppojen toimintaa varten, mutta niiden käyttö on ajamittaan laajentunut huomattavasti. Evästeet ovat käyttäjän koneella olevia tekstitiedosto-



Kuva 7: Palvelimen lokianalyysin eteneminen.

ja, joihin palvelin voi tallettaa tietoa. Käytettäessä istuntojen tai koneiden tunnistimina evästeitä evästeeseen voidaan tallettaa yksilöivä tunniste. Kävijän ladatessa ensimmäinen tiedosto palvelimelta talletetaan tunniste ja jatkossa joka tiedostonlatauksella asiakas voidaan yksilöidä tunnisteiden perusteella. Nyt tieto voidaan kerätä jokaisella sivulatauksella esimerkiksi PHP (Hypertext Preprocessor) -skriptin avulla tietokantaan. Etuna lokianalyysiin verrattuna evästepohjaiset järjestelmät yksilöivät kävijät huomattavasti tarkemmin, polkujen määrittäminen helpottuu, tiedon analysointi pystytään reaaliaikaistamaan ja herkkyys välityspalvelimille ei ole niin suuri. Pääosa käyttäjistä hyväksyy ensimmäisen tahon evästeet, mutta epätarkkuuksia voivat aiheuttaa kävijät, jotka eivät niitä hyväksy tai poistavat ne usein. [38]

Sekä sivustojen että seurantal palvelujentarjoajien perusteella suosituin seurantatapa ovat jäljitteet. Jäljitteiden toiminta perustuu sivulle sisällytettyyn jäljitteekoodiin. Jäljite on usein JavaScript-skriptikielellä toteuttu koodi, joka ottaa yhteyttä ulkopuoliseen palvelimeen. Palvelin kerää tiedot tietokantaan ja tarjoaa reaaliaikaisia raportteja kävijöistä. Aiemmin esitettyjen lokitietojen lisäksi jäljitteillä pystytään selvittämään kauan asiakas on ollut sivulla, onko hän täyttänyt lomakkeita ja miten hän on poistunut sivustolta. Menetelmä on myös täysin riippumaton välityspalvelimista, koska latauksen lisäys tapahtuu asiakkaan selaimesta eikä palvelimen saamasta latauspyynnöstä. Menetelmän ongelmana ovat kuitenkin asiakkaat, jotka estävät erityisesti kolmansien osapuolien palvelimilla olevat jäljitteet, sivulatausten hidastuminen ja kuorman aiheutus asiakkaan koneelle. [38]

4.5.3 Mitä hyödytään?

Tähän mennessä kävijäseurantaa on käsitelty runsaasti toteutukselliselta kannalta, mutta saatavaan tietoon ja tiedon lisäarvoon ei suuremmin ole kiinnitetty huomiota. Sama ongelma on vaivannut verkkosivujen kävijäseurantaa jo vuosia. Perinteisesti seurannan on toteuttanut ja tavoitteet sille on asettanut organisaation tekninen henkilöstö. Heidän intressinsä ovat toteutuksessa ja koneiden suorituskyvyn mittauksessa eikä niinkään ta-

loudellisissa tunnusluvuissa. Tässä luvussa pyritään selvittämään mitä kävijäseuranta voi tarjota yritysten markkinoinnista ja taloudesta vastaaville osastoille.

Ennen mittauksen ja koko sivustoprojektin aloittamista on määriteltävä sivuston tavoitteet ja asiakkaiden tavoitetoiminnot. Useat sivustot rahoittavat toimintansa mainoksin. Tällöin tavoitteet ovat lähinnä kävijämäärissä. Puhuttaessa yritysten sivuista tavoitteet ovat kuitenkin usein saada joko asiakas ostamaan jotain tai täyttämään yhteydenotto- tai tarjouspyyntölomake samalla tarjoten hyvän kontaktin. Ensimmäinen askel mittauksessa on muunnosasteen (Conversion Rate) laskeminen. Muunnosaste saadaan laskemalla asetetun tavoitteen toteuttaneiden vierailijoiden määrän suhde vierailijoiden kokonaismäärään kaavalla

$$\text{Muunnosaste} = \frac{\text{Tavoitteen toteuttaneet asiakkaat}}{\text{Yksittäiset kävijät}} \quad [41, 42]. \quad (2)$$

Pelkän muunnosasteen avulla on kuitenkin vaikeaa selvittää, miksi aste on kuinkakin hyvä tai huono. On selvitettävä, mitä asiakas tekee ennen kuin hän täyttää lomakkeen tai suorittaa ostoksen. Evästeiden avulla yksilöityjen kävijöiden reittejä voidaan tutkia polkuanalyysin avulla. Polkuanalyysissä käydään yksittäisten istuntojen avulla läpi reitit, joita asiakkaat ovat liikkuneet. Analyysi kertoo, mistä asiakkaat tulevat sivuille, mitä he tekevät sivuilla ja miltä sivulta he poistuvat sivustolta. Reittien avulla voidaan selvittää löytävätkö asiakkaat palvelut, onko sivuilla umpikujia tai selkeitä epämiellyttäviä sivuja, joilta asiakkaat poistuvat [43]. Tutkimalla poistumissivuja voidaan selvittää hylkäysaste (Abandonment rate). Hylkäysasteella tarkoitetaan verkkokaupan kassalle tai tilauslomakkeelle saapuneita asiakkaita, jotka jättävät kuitenkin viime silauksen tekemättä, suhteessa kokonaiskävijämääriin. Korkeat hylkäämisasteet kertovat muun muassa rajoittuneista maksutavoista ja ikävistä kysymyksistä lomakkeilla. [41]

Yhdistämällä verkkobudjetin muunnosasteen kanssa päästään käsiksi markkinointia ja taloutta kiinnostaviin tunnuslukuihin. Jakamalla verkkoon sijoitettu rahamäärä verkon kautta tehdyillä ostoilla saadaan seuraavan kaavan avulla selville sijoitetun pääoman tuottoaste ROI (result on investment) [18, 41]:

$$\text{ROI} = \frac{\text{Tuotot}}{\text{Sijoitettu pääoma}}. \quad (3)$$

Katelaskennassa hyödyllistä tietoa on yksittäisen verkko-oston aiheuttamat keskimääräiset kulut CPA (Cost Per Acquisition). [41]

Kuten aiemmissa luvuissa kerrottiin, kulutusprosessi ei sisällä pelkkää ostohetkeä. Seuraamalla palaavatko maksaneet asiakkaat sivustolle selviää heidän tyytyväisyytensä tuoteseen. Verkkomarkkinoinnin tunnuslukuja laskettaessa on otettava myös huomioon mainosarvo ja verkkomainosten perusteella muualta tehdyt ostokset. Niiden seuranta on kuitenkin hankalaa osoittaa ja verkkomarkkinoinnin korkeat tunnusluvut ovat usein todellisuudessa vielä korkeampi. [10]

Yksi muista poikkeava mittauksen kohde on klikkiväärennösten (Pay Per Click Click-fraud) seuraaminen. Johtuen internetin luonteesta ja vaikeasta valvonnasta hakukone-markkinointia voidaan helposti väärinkäyttää. Kilpailijat voivat käydä täyttämässä päivittäiset budjetit klikkailemalla mainoksia, jolloin omat mainokset pääsevät näkyviin. Valvonta tässä tapauksessa on luonnollisesti mainostoimiston eli hakukoneen päässä, mutta valvonta on vaikeaa, joten mainostajan itsensäkin on syytä olla valveilla. Seuraamalla samoista osoitteista usein tulevien ja vähän aikaa sivulla viihtyvien vierailuita voidaan väärinkäyttäjät saada kiinni mainostajankin toimesta. [44]

5 PRIMESOFTIN TARPEET HAKUKONEOPTIMOINNIN SUHTEEN

Primesoft Oy on erikoistunut markkinointipalveluihin ja markkinointia tukeviin ohjelmistoihin, joissa pääpaino on sähköisellä viestinnällä. Hakukonemarkkinointi on sekä Suomen mittakaavassa että kansainvälisestikin nopeimmin kasvava markkinointialue. Vastataksien asiakkaiden tarpeisiin ja laajentaakseen toimintaansa hakukoneoptimointipalveluiden kehittäminen oli Primesoftille luontainen valinta. Hakukoneoptimointiprojektin tavoitteena oli tuoda yrityksen omat sivut paremmin esille hakukoneissa ja toteuttaa hakukoneoptimointipalvelu osana muita markkinointitoimenpiteitä. Tässä luvussa selvitetään, mitä Primesoft tekee, mikä oli lähtötilanne ja millaisia vaatimuksia yrityksen toimintatavat ja ohjelmistot asettivat hakukoneoptimoinnin toteutukselle.

5.1 Primesoft Oy

Primesoft Oy on vuonna 2002 perustettu lappeenrantalainen markkinointi- ja markkinointiohjelmistopalveluihin erikoistunut yritys. Palvelut on suunnattu pääasiassa pk-yrityksille, joilla resurssit eivät riitä omaan markkinointitoimintaan. Palvelut kattavat koko markkinoinnin kirjon yrityksen ulkoasun suunnittelusta asiakashallintajärjestelmään. Primesoftin itsensäkin ollessa pk-yritys on se luonut alihankkijaverkoston perustuvan toimintamallin vastataksien asiakkaiden vaihteleviin ongelmiin. Tässä luvussa käsitellään lyhyesti Primesoftin palvelut sekä toimintamalli.

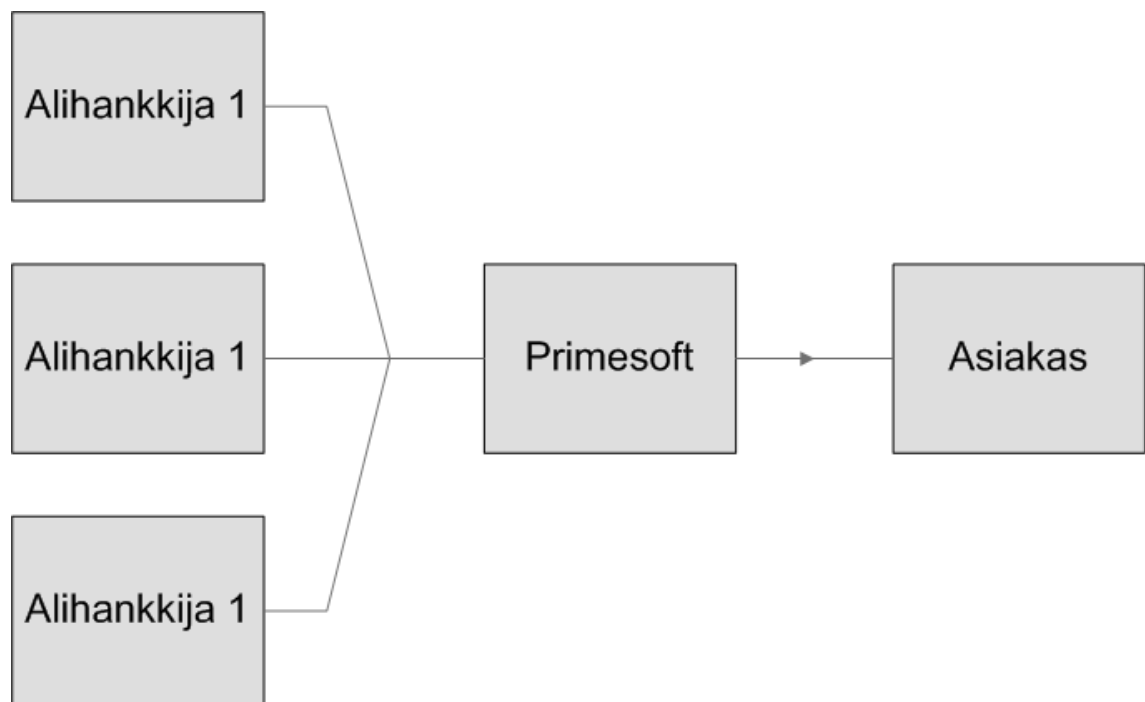
5.1.1 Palvelut ja tuotteet

Yrityksen tuotteet voidaan kahteen pääryhmään: perinteiseen markkinointiin sekä internetin vuoksi yleistyneeseen internet- ja sähköiseen markkinointiin. Perinteiset markkinointipalvelut kattavat yrityksen markkinoinnin suunnittelusta materiaalihankintaan. Suunnitteluvaiheessa yritykselle voidaan kehittää yhdenmukainen toimialaan sopiva ulkoasu sekä markkinointistrategia tarvittavine toimenpiteineen ja budjetiteineen. Konkreettiselle tasolle strategia viedään tekemällä yritykselle markkinointimateriaalit kuten käyntikortit, esitteet ja osastot messuja varten. Hakukoneoptimoinnin kannalta mielenkiintoisemman osuuden muodostavat internetpalvelut. Palvelut koostuvat myös suunnittelu- ja materiaali- palveluista. Ne sisältävät internetsivujen suunnittelun, sisällöntuotannon, teknisen toteu-

tuksen ja ylläpidon sekä materiaalin tuomisen sähköiseen muotoon, jolla tarkoitetaan vanhojen materiaalien kuten esitteiden ja logojen digitaaliseen muotoon vientiä. Sivujen toteutuksessa voidaan käyttää apuna muita markkinointipalveluja varten luotuja, vanhoja tai täysin uusia materiaaleja. Sivustot voidaan toteuttaa staattisina sivuina tai PrimeContent-julkaisujärjestelmän päälle toteutettuina sivuina.

5.1.2 Toimintaperiaate

Useiden muidenkin pk-yritysten tapaan Primesoft on panostanut ulkoistamiseen ja vahvoihin alihankkijaverkostoihin. Yritys hankkii ydinosaamisensa tukitoiminnot, fokusta haittaavat palvelut sekä voimakkaasti hintakilpaillut toiminnot tehtäviin orientoituneilta yhteistyöyrityksiltä. Kuvassa 8 on Primesoftin asema toimitusketjussa. Laajan toiminnan mahdollistamisen lisäksi strategialla saavutetaan hyötyjä yritykselle ja asiakkaille.



Kuva 8: Primesoftin asema toimitusketjussa.

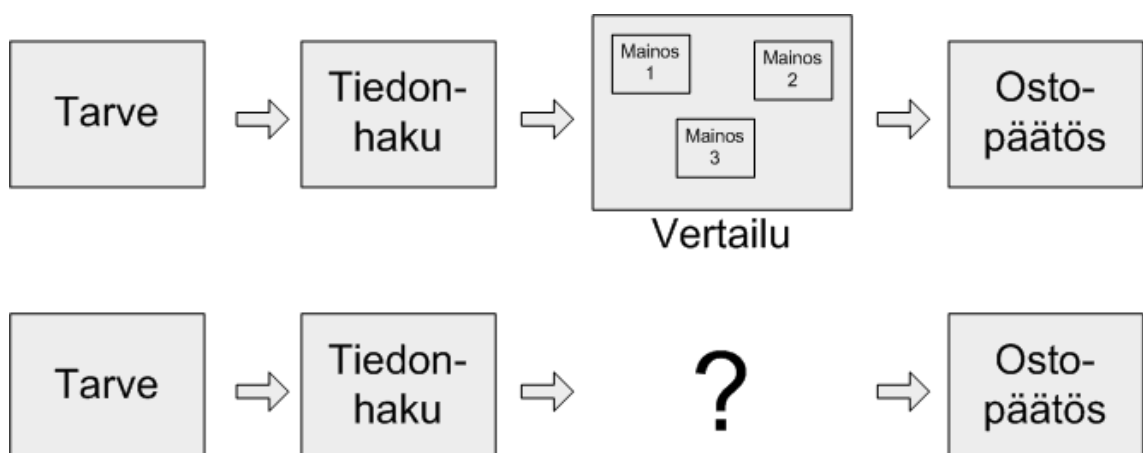
Riskejä aliarvioimatta Primesoft hyötyy toimintamallistaan niin strategisella kuin henkilöstötasollakin. Strategisella tasolla yritys pystyy vahvistamaan ydinosaamistaan, parantamaan yrityskuvaansa oikeilla kumppanivalinnoilla ja toimimaan joustavasti. Kehittämällä verkostoa yritys pystyy nopeasti saavuttamaan uusia segmenttejä, vastaamaan uusiin laatuvaatimuksiin ja kysyntämuutoksiin. Henkilöstöä ajatellen ydinosaamisen ulkopuolella olevien perustoimintojen ulkoistamisen vuoksi yrityksen sisäistä muutosvastarintaa

ei esiinny ja kaikkien keskittyessä ydintoimintoihin esteet urakehitykselle tulevat lähinnä työntekijästä itsestään. [45]

Asiakkaan saavuttamat edut ovat luonnollisesti sidoksissa yrityksen etuihin. Alihankkijaverkosto takaa asiakkaalle osaavan ja laadukkaan palvelun kaikilla osa-alueilla. Asiakasyritysten kehittyessä markkinoiden mukana Primesoft pystyy vastaamaan tehtäväorientoituneiden verkostotoimijoiden avulla asiakkaiden muuttuviin tarpeisiin. Keskitetty palvelu pienentää yrityksen yhteistyöverkostoa ja aikaa jää enemmän ydinosaiselle [45]. Markkinointipalveluissa merkittäväksi tekijäksi ikävä kyllä on osoittautunut hinta. Hintakilpailuilla aloilla erikoistuneet yritykset pystyvät tuottamaan palvelunsa kustannustehokkaasti ja alihankkijaverkostoja hyödyntävän yrityksen avulla päästään nauttimaan epäsuorasti myös massaeduista. [45]

5.2 Yrityksen internetpalvelut ja tavoitteet

Yrityksen edellisessä luvussa esitellyt palvelut vastasivat hyvin muutaman vuoden takaisia standardeja, mutta internetin merkitys ostopaikkana ja tuotteiden ja palveluiden etsintäpaikkana on kasvanut. Enää ei riitä, että yrityksellä on käyntikorttimaiset kotisivut, vaan niiden on tarjottava runsaasti ajankohtaista informaatiota tuotteista ja palveluista. Sen lisäksi yrityksen tuotteet täytyy olla löydettävissä etsittäessä ratkaisuja ongelmiin. Yrityksen tarjoamissa internetpalveluissa oli siis havaittavissa selkeä aukko asiakkaan tarpeen ja palvelun oston välillä. Kuvassa 9 on esitetty puute verrattuna perinteisten markkinointimateriaalien tarjoamaan ketjuun.



Kuva 9: Yrityksen internetmarkkinointipalveluissa havaittu puute verrattuna perinteisen markkinoinnin mainontaan.

Selkeä tavoite hakukoneoptimointiprojektille oli siis luoda yhteys asiakkaan tarpeiden ja palvelun välille. Täyden hyödyn saamiseksi projektista selvitettiin myös aiheeseen liittyviä ja ohessa helposti toteutettavia toimintaa kehittäviä tavoitteita. Alan artikkeleihin tutustumisen ja yrityksen johtoa sekä markkinointi- ja myyntivastaavaa haastatteleamalla päätavoitteiksi ilmenivät seuraavat asiat [46, 47]:

- Tunnettavuuden ja näkyvyyden lisääminen.
- Tehokas kontaktien kerääminen.
- Myyntityön tukeminen ja tehostaminen.
- Sähköisen markkinoinnin prosessiin liittyvien alihankkijoiden koulutus.
- Tehokas kävijäseuranta.

Tämä työ on rajattu käsittelemään näkyvyyden parantamista ja sivustojen kävijäseurantaa.

5.2.1 Tunnettavuuden ja näkyvyyden lisääminen

Hakukonenäkyvyys ja sen tuoma tunnettavuus on hakukoneoptimoinnin päätavoite. Tätä pidettiin myös tärkeänä Primesoftissa. Tavoitteeksi asetettiin:

- Hakukoneoptimointiin liittyvän tietotaidon tuonti yritykseen.
- Omien sivujen näkyvyyden lisääminen.
- Hakukoneoptimointipalvelun luominen asiakkaita varten.
- PrimeContent-julkaisujärjestelmän muokkaus hakukoneystävälliseksi ja hakukoneoptimointia tukevaksi.
- Hakukoneoptimoinnin operatiivisuus yrityksessä.

Hakukoneoptimointiin liittyvällä tietotaidolla tavoiteltiin osaamista ja osoitusta siitä, että osaamista löytyy. Tietotaitoa haluttiin kasvattaa laajalla aiheita koskevalla tutkimuksella ja tutkimuksen kautta tuotettu julkaisu diplomityön muodossa on konkreettinen osoitus asiaan perehtymisestä. Jatkossa osaaminen aiotaan myös osoittaa ottamalla osaa alan tutkimusprojekteihin ja markkinointihenkisiin tempauksiin sekä olla mukana luomassa itse vastaavanlaisia projekteja.

Hakukoneoptimoinnin toteuttajan valinnasta on kirjoitettu useita oppaita. Tärkeänä kriteerinä referenssien ohella pidetään optimointiyrityksen omaa hakukonenäkyvyyttä [48, 49]. Tämän vuoksi oma näkyvyys ja kärkisijat käytetyimmissä hakukonelistauksissa nähtiin tärkeinä. Oman sivuston näkyvyyden kasvattaminen toimii myös hyvänä kokemuksena ja osoittaa, mitä kilpailu kovimmillaan on, koska kärkisijat nähdään tärkeinä myös asiansa osaavissa kilpailijayrityksissä.

Hakukoneoptimoinnin operatiivisuudella tarkoitetaan hakukoneoptimointitoimenpiteiden automatisointia. Tavoitteena on siis saavuttaa tilanne, jossa kotisivujen tuotantoprosessiin osallistuvat tietävät, mistä on kyse, ja kotisivut toteutetaan näin ollen aina hakukoneoptimoinnin periaatteiden mukaisesti [10]. Yrityksen tavoitteena on myydä kokonaisia markkinointiratkaisuja, jolloin itse hakukoneoptimointi ei maksaisi asiakkaalle mitään. Ratkaisuissa asiakas maksaisi kuukausittaista maksua kokonaisesta palvelusta tai esimerkiksi kontaktipohjaisesti tietyn summan saatua kontaktia kohden.

Kun perustieto ja kokemus on hankittu, voidaan hakukoneoptimointia alkaa tarjota palvelun muodossa. Palvelua ajatellen tavoitteena oli selvittää toimitusmallit.

5.2.2 Kävijäseuranta

Kuten luvussa seuranta todettiin, ilman mittareita ja seuranta ei voida arvioida onnistumista eikä kehittää toimintaa. Kävijäseurannan tavoitteeksi oli asetettu vaatimusten kartoittaminen ja seurannan toteutus. Vaatimukset kartoitettiin pääosin jo kirjallisessa tutkimuksessa, mutta mukaan haluttiin myös vähemmän tekniikkaorientoituneiden ihmisten mielipiteitä haastattelujen kautta. Toteutusvaiheessa tehtäväksi jäi kartoittaa toteutusvaihtoehdot sekä valita toteutukseen parhaiten sopiva vaihtoehto.

Tavoitteenasetannassa sekä asiakkaiden että markkinoinnin ja myynnin mielipiteiden perusteella kävijäseuranta nähtiin tärkeänä. Haastattelujen perusteella tärkeimmiksi asioiksi osoittautuivat [46, 47]:

- Kävijämäärät ja niiden kehitys eri ajanjaksoina.
- Potentiaaliset asiakkaat osoitetietojen perusteella.
- Kävijöiden kulkemat reitit.
- Tietojen käyttö muissa markkinoinnin tukijärjestelmissä.

- A/B-testaus.

Kävijämäärillä eri ajanjaksoina tarkoitetaan vierailujen, eri kävijöiden määrän seuraamista ja vertailua esimerkiksi eri kuukausien välillä. Potentiaalisten asiakkaiden tiedoilla tarkoitetaan asiakkaiden verkko-osoitteita, jotka usein kertovat organisaation. Reitit eli järjestys, jossa vierailija on ladannut sivut, kertoo, mistä asiakkaat ovat kiinnostuneet ja eksyvätkö he sivustolla etsiessään tietoa. Tiedonkäytöllä muissa järjestelmissä tarkoitetaan ohjelmistorajapintojen kehitystä, jolla tieto voidaan viedä automaattisesti eri sovelluksiin. A/B-testaus on markkinointikampanjoiden tehokkuuden testausta, jossa vertaillaan mitkä kampanjat toimivat parhaiten.

Jatkossa kävijäseuranta tullaan tarjoamaan myös erillisenä palveluna. Tällöin järjestelmän on vastattava markkinointilähtöisten tietojen lisäksi kävijäseurantajärjestelmien perinteisesti tarjoamiin tietoihin [50]:

- Vierailut.
- Sivulataukset.
- Ladatuimmat sivut.
- Suosituimmat
 - saapumissivut
 - poistumissivut
 - sivulle viittaajat
 - vierailijamaat
 - vierailijaorganisaatiot.
- Hakusanat, joilla sivuille saavutaan.

Tietojen ohella seurannan on tarjottava tiedot listauksen lisäksi havainnollisemmissa muodoissa kuten kaavioina.

6 HAKUKONEOPTIMOINNIN TOTEUTUS PRIMESOFTILLA

Tässä luvussa keskitymme hakukoneoptimoinnin toteutukseen Primesoft Oy:ssä. Työn teknillisen tason säilyttämiseksi keskitytään tässä luvussa kotisivujen tuotto-ohjelmalle eli PrimeContent- julkaisujärjestelmälle asetettuihin sisällöllisen ja teknisen optimoinnin vaatimuksiin sekä kävijäseurannan toteutukseen. Tähänastiset PrimeContentin versiot on kehitetty sivuston ylläpitoa ja helppokäyttöisyyttä ajatellen. Hakukoneystävällisyyteen ja sivustojen indeksoitumiseen ei tähän mennessä ole kiinnitetty huomiota. Yhtenä markkinoinnin tärkeimmistä toiminnoista pidettiin tulosten seuranta. Tähän mennessä yritys ei tarjonnut internetmarkkinointiin seurantatapoja, joten kartoitus valmiista kävijäseurantamahdollisuuksista tai mahdollisen oman seurannan toteutus oli tarpeen. Diplomityöhön kuului osallistuminen PrimeContentin rakennemuutosten jälkeiseen uudelleenkirjoitukseen, PrimeContentin hakukoneoptimoinnillisten puutteiden kartoitus ja korjaus sekä kävijäseurannan toteutus.

6.1 PrimeContent ja hakukoneoptimointi

Tässä luvussa esitellään Primesoftin PrimeContent-julkaisujärjestelmä ja hakukoneoptimoinnin siihen vaatimat muutokset. Aluksi käydään läpi, mikä PrimeContent on ja miten se pääpiirteissään toimii. Järjestelmän esittelyn jälkeen esitetään hakukoneoptimointia vaikeuttavat ominaisuudet ja puutteet sekä tavat, joilla ne voidaan korjata.

6.1.1 PrimeContent-julkaisujärjestelmä

PrimeContent on Primesoftin tuottama internetsivujen julkaisujärjestelmä. Julkaisujärjestelmällä tarkoitetaan järjestelmää, jossa sivuston sisältö erotetaan esityskerroksesta. Näin sisällön kuten kuvien ja tekstin tuottaminen onnistuu helposti ilman, että sivujen muokkaajan tarvitsee ottaa kantaa niiden tekniseen toteutukseen. Vastapainoisesti taas pyritään jopa estämään käyttäjän mahdollisuuksia muokata graafista sekä teknistä toteutusta. Näin sivusto saadaan pysymään kauttaaltaan yhdennäköisenä ja eri teknisissä ympäristöissä toimivana.

PrimeContent koostuu hallintaosasta sekä julkaisu- ja sivuntuottamismoottorista. Hallintaosio koostuu sivustojen toteutukseen, sisällöntuotantoon ja oikeuksienmäärittelyyn vaa-

dittavista komponenteista. Hallinnan kautta luodaan sivuston rakenne, sisältö sekä määritellään esimerkiksi tietyille käyttäjäryhmille rajatut sisäverkot, joiden perusteella julkaisumoottori tuottaa sivut. Julkaisumoottori koostuu järjestelmän sisäisistä sivupohjista, sisällön ja tyylimääreiden jäsentimisestä sekä interaktiivisten elementtien käsittelijöistä. Julkaisumoottorin tehtävä on siis tuottaa sivuja sisäisten mallien ja hallintaosiossa tehtyjen määritteiden sekä sisällön mukaan.

Hallintaosan komponentit on jaettu julkaisujärjestelmässä rakennekohtaisiin ja sisältökohtaisiin komponentteihin. Rakenneosio sisältää sivuston rakennekohtaiset kuten sivuja ja sivuhierarkiaa koskevat toiminnot. Sisällön kautta pystytään muokkaamaan erilaisia sisältökomponentteja ja ulkoasun kautta sivuston tyylimääreitä.

Rakenteen hallinnassa voidaan määrittää sivuston sivut, sivuryhmät sekä käytettävät kielet. Yksittäisille sivuille voidaan määrittää omat tyylinsä, käyttäjäryhmät sekä ajastettu julkaisu. Sivuryhmien avulla voidaan määrittellä sivuista koostuvia ryhmiä, jotka näkyvät esimerkiksi tuoteryhmille ominaisin tyylein taikka vain tietyille käyttäjäryhmille. Jos sivut ovat monikieliset, ovat ne usein myös hyvin laajat. Kielimäärittely luo eri kielille omat juurisivut, jolloin laajojenkin sivujen sisällönmuokkaus ja hallinta pysyy helppona.

Sisällön muokkausosiossa järjestelmän peruskäyttäjät pystyvät muokkaamaan ja lisäämään sisältöä sivustolle. Tuotantovaihtoehdot sisältävät vapaamuotoisen tekstipalstan ja eri käyttötarkoituksiin sopivia työkaluja. Listatyökalu on kehitetty erilaisten listausten kuten uutisten, tuotekatalogien ja työntekijälistojen tekoon. Työkalun avulla käyttäjä voi helposti määrittellä kenttäkohtaiset otsikot, tekstit, infosivut sekä kuvat. Näin jatkuvan tiedon muokkaaminen on helppoa ja lista säilyttää alkuperäisen ulkoasunsa. Lomaketyökalu on tarkoitettu erilaisten kontaktointilomakkeiden tekoon. Käyttäjä voi näin määrittellä kentät, niiden pakollisuuden ja vastausten pätevöintitavat. Lomakkeilta tulleet yhteydenotot voidaan asettaa menemään suoraan sähköpostiosoitteisiin tai -listoihin tai julkaista eri muo-
maateissa olevina raportteina. Pääosa nykyisistä sivuista sisältää erilaisia valikkopalkkeja. Valikkotyökalun avulla voidaan määrittää sivustolle valikot ja kuinka monta alemmaa tasoa milläkin valikkotasolla näkyy.

Yksi julkaisujärjestelmien peruspiirteistä on esityksen ja sisällön erottamisen mukana tuleva ulkoasun muokkaamisen esto. PrimeContent sisältää työkalun myös tähän toimintoon. Sen käyttö on yleensä rajattu pelkästään sivuston ulkoasuntuottajiin. PrimeContentissa työkalua käytetään ulkoasun muokkauksen lisäksi sivuston tyylin syöttämiseen järjestelmälle ja sisältöpalstojen määrittelyyn. Sisältöä tuottaessa käyttäjä voi määrittää, mihin haluaa sisältöelementtinsä. Nämä elementit määritellään ulkoasun muokkauksen

avulla ja nimetään kuvaavasti, jotta käyttäjä voi helposti lisätä sisältöä haluamaansa kohtaan sivulle. Palstojen ohella ulkoasu määrittää kaikki eri elementtien sisältämät tyylit kuten otsikot ja leipätekstin.

6.1.2 Korjaustoimenpiteet

Yhtenä hakukoneoptimoinnin tavoitteista oli PrimeContentin muokkaaminen hakukone- ja hakukoneoptimointiystävälliseksi. Jo tutkimisen alkuvaiheessa oli selvää, että hakukoneoptimoinnin aiemmasta huomioimattomuudesta huolimatta PrimeContentilla tuotetut sivut olivat hakukoneystävällisiä. Vaikka muutamia puutteita järjestelmästä löydettiin, suuriin arkkitehtuurillisiin muutoksiin ei ollut tarvetta. Julkaisujärjestelmä ei käyttänyt alun alkujaankaan esimerkiksi kehyksiä (frame) tai taulukoita sivurakenteiden määrittämiseen vaan toteutus oli puhtaasti CSS-pohjainen.

Yksi kriittisimmistä ja myös yleisimmistä julkaisujärjestelmien ongelmista ovat dynaamisten sivujen URL:t ja niissä esiintyvät parametrit. Runsaiden parametrien johdosta hakukoneet eivät usein indeksoi sivuja [10]. PrimeContentissa erilliset sivut oli toteutettu jo alkuvaiheessa niin, että uusi sivu näkyy sivuja katsoessa alihakemistona. Eli esimerkiksi Primesoftin kotisivujen yritystietosivu näkyy käyttäjälle osoitteena

<http://www.primesoft.fi/yritys/>.

Käytön helpottamiseksi järjestelmässä oli kuitenkin yhdistetty sivun otsikko, hakemisto ja sivun valikkonimi. Otsikosta tai tässä tapauksessa sivun nimestä muodostettiin URL-kelpoinen hakemistonimi ja otsikkona käytettiin suoraan sivulle annettua nimeä. Hakukoneoptimoinnissa otsikko on tärkeässä asemassa ja kunnolliset otsikot muodostavat hyvin pitkiä hakemistonimiä eivätkä myöskään sovi kovin kätevästi valikkonimiksi. Esimerkiksi yrityksen hakukonemarkkinointisivulla otsikko on

Hakukonemarkkinointi - internetmarkkinointi, hakukoneoptimointi, hakusanamarkkinointi, googlemarkkinointi.

Tästä muodostuisi sivun URL:ksi

http://www.primesoft.fi/hakukonemarkkinointi-internemarkkinointi_hakukoneoptimointi_hakusanamarkkinointi_googlemarkkinointi/,

mikä ei taas ole kovin käytännöllinen sivun osoitteena kuten ei myöskään valikkonimenä

esimerkiksi alavetovalikossa. Ongelma korjattiin yksinkertaisesti lisäämällä sivun tietokantatauluun erilliset kentät hakemistolle, otsikolle ja valikkonimelle. Yhteensopivuuden ja käyttömukavuuden ylläpitämiseksi kentät voidaan jättää tyhjiksi, jolloin käytetään normaaliin tapaan sivun nimeä näissä kohdissa.

PrimeContent käyttää erilaisten komponenttien julkaisuun valmiita sivu- ja komponenttipohjia. Erilaisten komponenttien validoinnissa huomattiin, että niitä ei ollut tarkastettu toteutusvaiheessa ja virheitä löytyi. Komponentit käytiin yksitellen läpi ja testisivujen avulla korjattiin niistä löytyneet virheet.

Vaikka metatietokenttien kuvaus ja avainsanat tärkeydestä ollaan montaa mieltä, nähtiin ne selkeänä puutteena järjestelmässä. Kentät toteutettiin lisäämällä ne osaksi järjestelmän rakenneosion sivujen hallintaa. Jokaiselle sivulle lisättiin kentät kuvaus ja avainsanat, jotka käyttäjä voi halutessaan täyttää. Julkaisupuolella muutokset lisättiin sivun ylätunnistemalliin niin, että ne tulevat näkyviin, jos sivulle on määritetty kyseiset tunnisteet, mutta muussa tapauksessa niitä ei määritetä.

Jotta mahdollisimman paljon sivun sisällöstä tulee indeksoiduksi, on kaikki sisällön kannalta epäolennainen kuten tyylimääritteet oltava omassa tiedostossaan. Julkaisujärjestelmässä tyylimääritteet tuotettiin omana tiedostonaan, mutta JavaScript-lisäosat lisättiin suoraan sivun koodiin. Joissain tapauksissa nämä skriptit voivat venyä hyvinkin pitkiksi ja näin estää sivun sisällön indeksoinnin.

6.2 Kävijäseuranta

Kävijäseuranta ja web-analytiikka ovat olleet viime aikoina runsaasti esillä mediassa ja aiheesta on tehty tutkimuksia sekä kirjoitettu oppaita. Taustatietona vaihtoehtojen kartoitukseen käytettiin hakukoneita sekä Katriina Saransaaren vuonna 2005 Lappeenrannan teknilliseen yliopistoon tekemää diplomityötä “WWW-sivuston kävijäseurannan uudistaminen” [51]. Näiden pohjalta kerättiin potentiaalisimmat vaihtoehdot ja vertailtiin niitä.

Saransaaren diplomityössä tehty laaja ohjelmistojen ominaisuus- ja hintakartoitus antoivat hyvän pohjan ohjelmistovalinnalle. Koska vaatimukset vastasivat pääosin vaatimuk-siamme ja hintalaatusuhde oli ylivertainen, valitsimme diplomityön koeryhmästä parhaaksi valitun Google Analyticsin yhdeksi vaihtoehdoistamme. [51]

Hakukoneiden kautta löydettyjen oppaiden avulla vaihtoehtoja löytyi useita. Monet vaih-

toehdoista oli kuitenkin jo käsitelty Saransaaren diplomityössä, joten niihin ei perehdytty sen tarkemmin. Mielenkiintoisimmaksi vaihtoehdoksi jäi Snoobi-kävijäseuranta kotimaisuutensa ja mahdollisen kumppanuuden vuoksi. Ilmaisisista vaihtoehdoista otettiin mukaan vielä Webalizer-ohjelma, johon tutustumista helpotti myös se, että ohjelma tuli yhden palvelimemme Linux-jakelussa valmiiksi asennettuna.

6.2.1 Google Analytics -kävijäseurantaohjelmisto

Google Analytics on Googlen tarjoama ilmainen kävijäseurantaohjelmisto. Se tarjoaa luotettavasti tarkkoja ja helposti ymmärrettäviä raportteja kävijöistä. Analytics ei myöskään vaadi ohjelmistoasennuksia, koska Google tarjoaa sitä ASP (Application Service Provider) -palveluna eli ohjelmisto on palveluntarjoajan järjestelmässä, josta myös tulosten seuranta tapahtuu. Google Analyticsin avulla voidaan toteuttaa normaalien internetsivujen sekä yritysten sisäverkkojen seuranta. Analysoinnin kannalta se tarjoaa muun muassa seuraavia tietoja [52]:

- Kävijämäärät.
- Mistä kävijät tulevat.
- Miten kävijät käyttäytyvät sivulla.
- Mainoskampanjoiden tulokset ja vertailu (A/B-testaus).
- Kävijöiden sijainti.

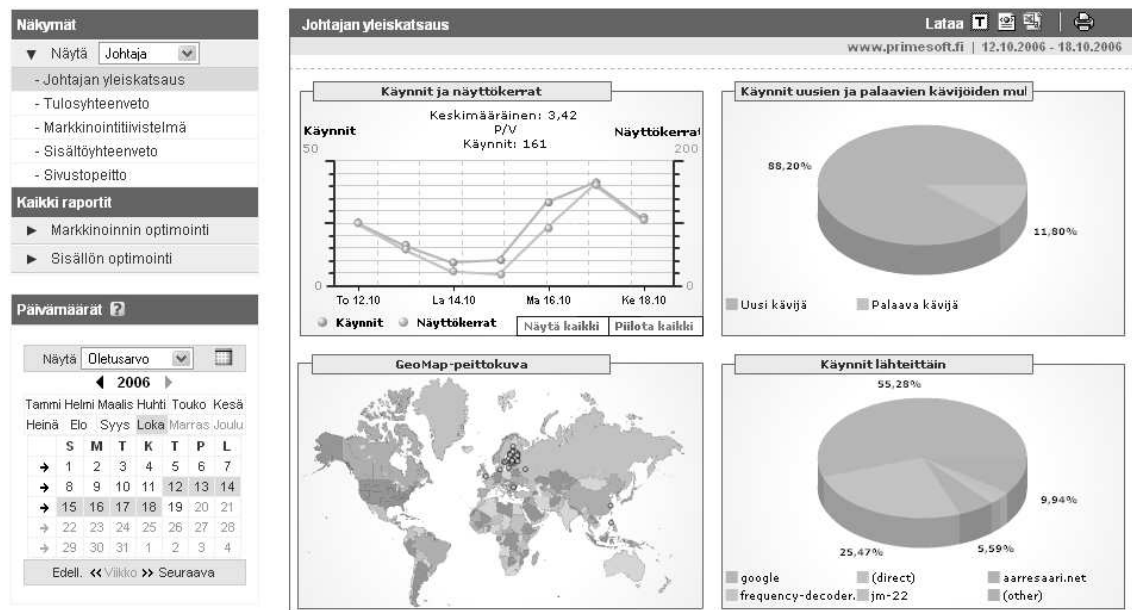
Google Analytics käyttää seurantaan jäljitteitä. Jäljitepohjaisissa seurantaohjelmistoissa seuranta suoritetaan lisäämällä yleensä JavaScript-skriptikielellä toteutettu seurantakoodi jokaiselle sivulle. Seurantakoodi tulee asettaa mahdollisimman lähelle sivun tai sivumalin loppua, jotta sivujen lataus ei hidastu. Analyticsissä kyseinen koodi näyttää seuraavalta [52]:

```
<script src="http://www.google-analytics.com/urchin.js"
type="text/javascript">
</script>
<script type="text/javascript">
_uacct="UA-xxxx-x";
urchinTracker();
```

</script>

Tiedon analysointiin Google Analytics tarjoaa web-käyttöliittymän (kuva 10), jossa tilastot on jaettu loogisiin kokonaisuuksiin. Tilastoihin voi tutustua erilaisin graafein sekä numeerisena tietona. Google ei ainakaan toistaiseksi tarjoa API:a (Application Program Interface) eli ohjelmistorajapintaa Analyticsiin. Tiedot on siis haettava manuaalisesti web-käyttöliittymän kautta, josta ne ovat saatavissa teksti-, Excel- ja XML-muodossa. [52]

Raportit: www.primesoft.fi



Kuva 10: Google Analyticsin web-käyttöliittymän tarjoama yleiskatsaus.

6.2.2 Palvelinlokien analysointiohjelma Webalizer

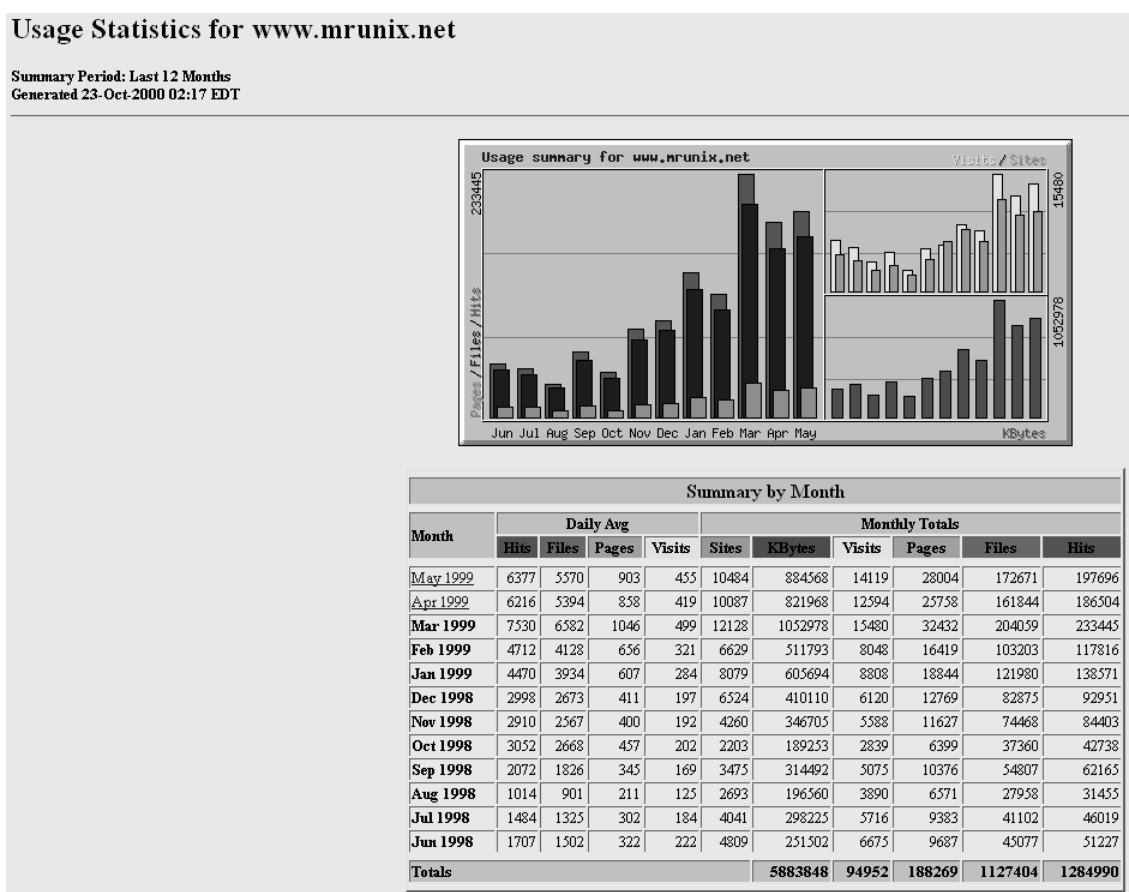
Webalizer on nopea ja ilmainen palvelinlokien analysointiohjelma. Webalizer tuottaa yksityiskohtaisia kävijäseurantareportteja HTML-formaatissa. Raportit tarjoavat useita eri ajanjaksoja kuten vuosi, kuukausi ja viikko. Raporteista ilmenevät muun muassa seuraavat tiedot [53]:

- Kävijämäärät.
- Ladatuimmat sivut.
- Suosituimmat saapumis- ja poistumissivut.

- Suosituimmat viittaajat ja hakusanat.
- Kävijöiden kotimaat.

Poiketen Google Analyticsista Webalizer suorittaa siis kävijäseurannan tutkimalla palvelinlokia. Tästä johtuen tieto ei ole reaaliaikaista vaan lokianalyysi suoritetaan esimerkiksi päivittäin palvelinkuorman vähentämiseksi. Webalizer tukee CLF-lokiformaattistandardia sekä esimerkiksi Apache-web-palvelimissa usein käytettyä Combined Log -formaattia, jonka avulla palvelimella sijaitsevat erilliset internetsivustot voidaan erotella toisistaan. [53]

Tiedon analysointiin Webalizer tarjoaa lokianalyysissä tuotetun HTML-raportin. Tilastot on jaettu aikaväleihin ja kävijämäärien kehitystä voidaan tutkia myös graafien avulla (kuva 11). Muuten tilastot koostuvat taulukoin toteutetuista top-listauksista. Webalizer ei tarjoa API-liittymää eikä tietoja ole mahdollista saada muissa formaateissa. [53]



Kuva 11: Webalizerin web-käyttöliittymän tarjoama yleiskatsaus.

6.2.3 Snoobi-kävijäseurantaohjelmisto

Toisena vaihtoehtona vertailuun otettiin suomalaisen Snoobi Oy:n Snoobi-kävijäseurantaohjelmisto. Snoobi on myös niin sanottu ASP-palvelu ja se tarjoaa web-pohjaisen käyttöliittymän tulosten seurantaan. Toteutustavasta johtuen Analyticsin tapaan Snoobi sopii myös julkisten sekä yritysten sisäisten sivujen seurantaan. Snoobin raportit tarjoavat muun muassa seuraavia tietoja [54]:

- Kävijäyhteenveto.
- Suosituimmat sivut ja hakusanat.
- Suosituimmat viittaajat eroteltuna hakukoneisiin, hakemistoihin ja linkkeihin.
- Kävijät maittain ja organisaatioittain.

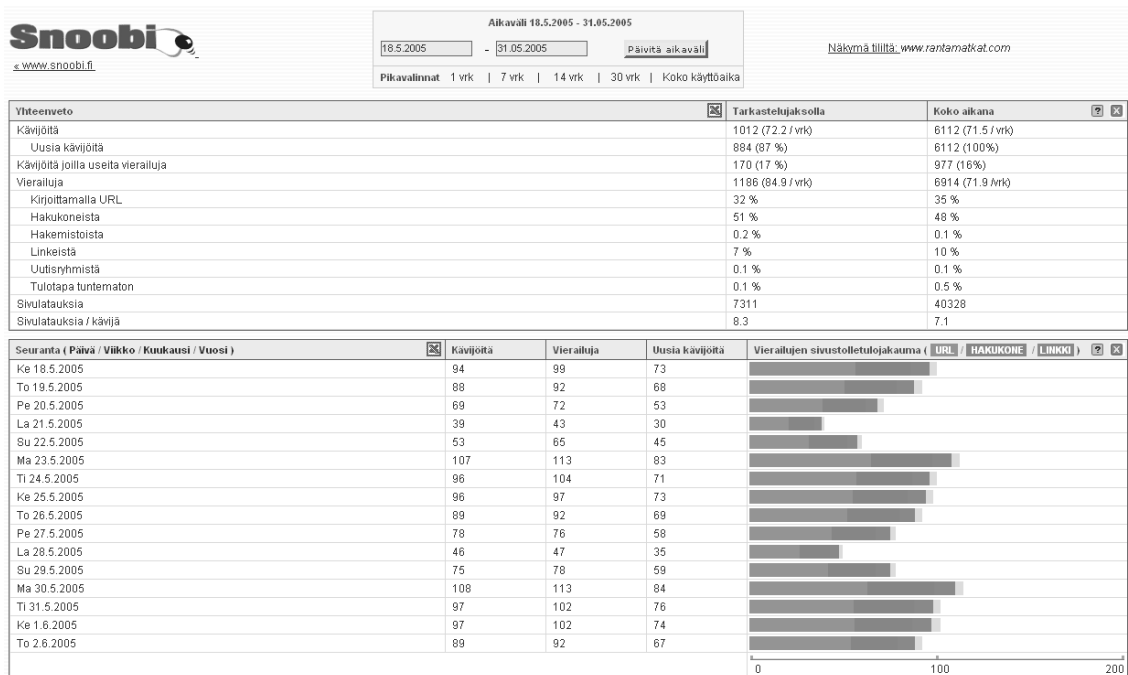
Snoobi käyttää seurannassaan apuna jäljitteitä ja evästeitä. Käyttöön otossa asiakkaan sivustolle asennetaan JavaScript-seurantakoodi, joka ottaa yhteyttä Snoobin palvelin keskuksen. Palvelin keskus rekisteröi sivulatauksen ja raportit ovat seurattavissa reaaliaikaisesti. [54]

Snoobin tuottamien tulosten seuranta tapahtuu web-käyttöliittymän kautta, joka tarjoaa HTML-sivun kävijätietoyhteenvetoiheen ja yksityiskohtaisine tuloksineen (kuva 12). Web-käyttöliittymän kautta tiedot voidaan tuoda HTML:n lisäksi Excel-tilukkona. API-liittymää Snoobikaan ei tarjoa, mutta Online Business Intelligence -konsultoinnin lisäksi Snoobi Oy tarjoaa teknistä tukea toiminnan tehostamiseksi muun muassa asiakaskohtaisesti räätälöityjen raporttien muodossa. [54]

6.3 Kävijäseurannan valinta

Kävijäseurannan valinta- ja toteutus päätös tehtiin pääosin ominaisuusvaatimusten perusteella. Ohjelmien toimivuutta olisi voitu testata esimerkiksi yrityksen omilla sivuilla, mutta kireän aikataulun ja resurssien puutteen vuoksi testausta ei suoritettu. Lähdemateriaalista kuten Saransaaren diplomityöstä katsottiin saatavan riittävät tulokset käytännön toiminnasta.

Vaihtoehtojen keskinäinen arviointi suoritettiin vertaamalla, miten eri vaihtoehdot vastaavat vaatimuksiin. Apuna käytettiin taulukon 8 matriisia, jossa ohjelmistot esiintyvät riveil-



Kuva 12: Snoobin web-käyttöliittymän tarjoamaa tietoa.

lä ja vaaditut ominaisuudet sarakkeilla. Sarakkeet jaettiin perustietoihin, A/B-testaukseen, kävijäreitteihin, ilmaisuuteen, API-liittymään sekä tiedontuontiin. Perustiedot tässä tapauksessa sisälsivät yleisesti ohjelmien tarjoamia ja niiltä perinteisesti odotettuja toimintoja kuten kävijämäärät, aikajaksotus ja verkko-osoitteet ja viittaajat. Eri kampanjoiden toimivuuden A/B-testaus ja kävijäreitit asetettiin omiksi sarakkeikseen, koska ne tuottavat tärkeää tietoa kampanjoiden ja sivuston tehokkuuden kehitystä ajatellen. Hinta ja ilmaisuus oli myös yksittäinen palveluntarjonnan kannalta olennainen kriteeri. Primesoftin tavoitteena on kehittää kokonaisvaltainen asiakas- ja markkinointitiedon hallintajärjestelmä, jolla saadaan esimerkiksi potentiaaliset asiakkaat suoraan asiakashallintajärjestelmään ja kampanjoiden aiheuttamat tuotot ja kulut tehokkuusmittareihin. Kävijäseuranta vastaa muun muassa edellisiin tarpeisiin sivustokävijöiden verkko-osoitteiden sekä esimerkiksi bannerikampanjoiden tuomien kävijöiden kautta. Tämän vuoksi API-liittymä ja tiedontuonti nähtiin niin tärkeinä, että ne asetettiin matriisissa omiksi sarakkeiksiin.

Taulukko 8: Eri kävijäseurantaehdokkaiden vertailumatriisi.

Ohjelmisto / ominaisuus	Perustiedot	A/B-testaus	Reitit	Ilmaisuus	API	Tiedontuonti
Google Analytics	X	X	X	X		X
Snoobi	X	X				X
Webalizer	X			X		

Kuten taulukosta 8 voidaan todeta, Google Analytics vastasi parhaiten asetettuihin vaatimuksiin. Webalizer oli oletetusti ikänsä ja kehityslähtökohtiensa vuoksi hyvin tekniikkaorientoitunut eikä vastannut markkinoinnin ja myynnin asettamiin vaatimuksiin. Snoobiin ei täysin pystynyt perusominaisuuksillaan vastaamaan ja hieman ongellisena nähtiin myös ASP (Application Service Provider) -mallin hinnoittelu palvelumyyntiä ajatellen.

Valmiista ohjelmista Google Analytics oli siis paras vaihtoehto. Valinta ei osoittautunut kuitenkaan niin yksinkertaiseksi, koska API:n painoarvo oli suuri ja eräitä toimituksellisia seikkoja ilmeni perehdyttäessä tarkemmin ohjelmistoon. Saransaaren diplomityössä yhtenä Analyticsin negatiivisena puolena nähtiin toimitusvarmuus. Palvelu vaikutti myös uutisten perusteella kärsineen toimitusongelmista ja ilman erityistä kutsua käyttäjäksi on päässyt vasta elokuusta 2006 lähtien [51, 55]. Esitetyistä kyselyistä huolimatta Google myös ilmoittaa, että API:a ei ole eikä myöskään viittauksia kehitykseen ole.

Johtuen toimitusepävarmuuksista ja ohjelmistorajapintojen puutteesta kävijäseuranta päätettiin toteuttaa itse osa diplomityötä. Korkeista kehitys- ja toteutuskuluista huolimatta itse tuotettu ohjelmisto saadaan helpoiten vastaamaan vaatimuksia. Jatkossa toteutus on myös helposti laajennettavissa, jolloin se pystyy nopeasti vastaamaan uusiin vaatimuksiin.

6.4 Kävijäseurantaohjelmisto PrimeTrack

Tuoteperheen nimien mukaisesti kävijäseurantaohjelmisto sai nimen PrimeTrack. PrimeTrack päätettiin toteuttaa osana PrimeContent-julkaisujärjestelmää. Tällä järjestelyllä etuina saavutetaan valmiit ohjelmistorajapinnat ja alusta. PrimeContentin rajapintojen avulla tietoa saadaan suoraan web-liikenteestä ja sitä voidaan tuottaa asiakashallintajärjestelmää varten. Samalla alustalla työskentely lisää toimintavarmuutta ja tietokantayhteydet ja muut perustoiminnot on valmiiksi alustettu. Vastapainoisesti verrattaessa irralliseen toteutukseen järjestelmä vaatii enemmän resursseja PrimeContentin ulkopuolisilla sivuilla.

PrimeTrack suunniteltiin rakennettavan kolmesta komponentista, sivulataustenkäsittelystä, tietokannasta ja tiedonanalysoinnista (kuva 13). Sivulataustenkäsittelyssä selvitetään käyttäjän identiteetti, ladatun sivun tiedot, ajankohta ja mahdolliset viittaajat ja muunnetaan ne käytettävään muotoon. Tiedonkäsittelyn jälkeen tiedot talletetaan tietokantaan. Tulosten luettavaan muotoon muutto ja muu visualisointi tapahtuu tiedonanalysointikomponentissa.

6.4.1 Sivulatausten käsittely

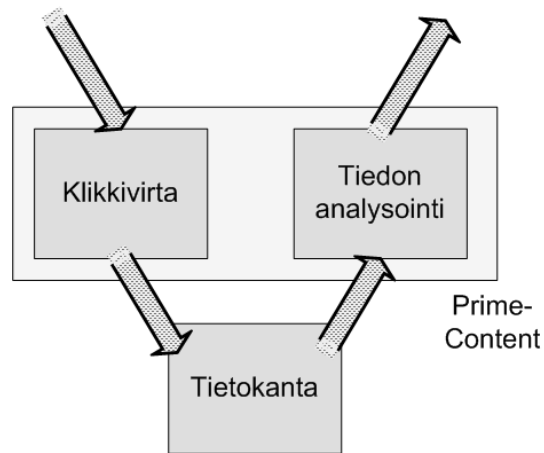
Tapahtumaketjussa ensimmäisenä komponenttina on sivulataustenkäsittelijä. Komponentti toteutettiin PHP-skriptikielisenä luokkarakenteena (kuva 14). Olion luonti ja tiedon prosessointi tapahtuu PrimeContentin sivun generoinnin yhteydessä, jolloin kävijästä saadaan tiedot parhaiten irti.

Prosessi lähtee käyntiin luomalla käyttäjä- tai hakurobottiluokka luokan *CClickStreamRecord* tehdasmetodia käyttäen. Tehdasmetodi (factory) sisältää tunnistimen, joka tunnistaa robotit hakurobottien selaimen listauksen avulla ja luo joko robotti- tai kävijäolion. Robotin ollessa kyseessä lähetetään tietokannalle robotin nimi ja vierailuaika. Kävijän ollessa kyseessä lähetetään tietokannalle tiedot kävijästä. Olion luonnin jälkeen kutsutaan yksinkertaisesti process-funktiota, joka huolehtii tiedonkäsittelystä ja eteenpäinviennistä.

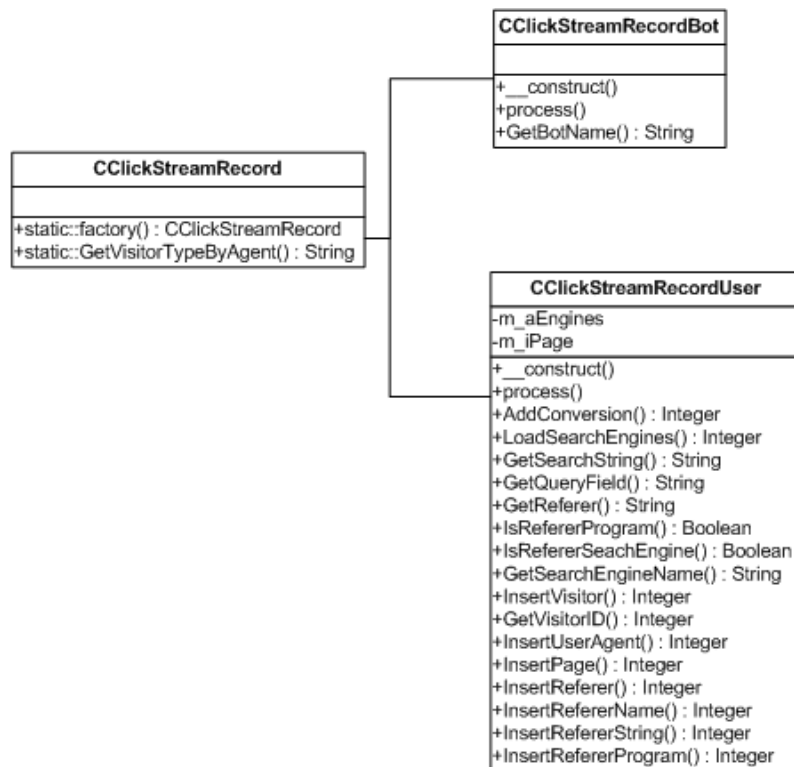
Kävijätietoihin käytetään apuna PHP:n esitäytettyjä assosiatiivisia tauluja *\$_SERVER* ja *\$_COOKIE*. *\$_SERVER*-taulun kautta saadaan tiedot kävijän alkuperästä indekseillä *REMOTE_HOST* ja *REMOTE_ADDRESS*, selain indeksillä *HTTP_USER_AGENT* ja edeltävän sivun viittaajatiedot indeksillä *HTTP_REFERER*. Indeksijä *REMOTE_ADDRESS*, *REMOTE_HOST* ja *HTTP_USER_AGENT* käytetään käyttäjän identifiointiin sekä käyttäjän organisaation tunnistamiseen. Aiemmin mainittujen lisäksi taulusta *\$_COOKIE* saadaan selville käyttäjän istuntoeväste *SESSION_ID* ja järjestelmän asettama pysyvä eväste *PERMANENT_ID*, jotka helpottavat yksittäisen käyttäjän tunnistamista. Markkinoinnin tehokkuusmittauksen ja käyttäjien reittien ja lähteiden kannalta olennainen tieto saadaan *\$_SERVER*-taulun indeksistä *HTTP_REFERER*. Käsittelijä tunnistaa viittaajan osoitteen perusteella, millainen viittaaja on kyseessä. Taulukossa 9 on esitetty tarvittavat tiedot ja niiden lähteet.

Taulukko 9: Sivulatauksesta tarvittavat tiedot ja niiden lähteet.

Tieto	Lähde
Sivun tunniste	Rakentajassa annettu parametri
Verkko-osoite	<i>\$_SERVER['REMOTE_HOST']</i>
Verkko-tunnus	<i>\$_SERVER['REMOTE_ADDR']</i>
Istuntoeväste	<i>\$_COOKIE['SESSION_ID']</i>
Pysyvä jäljite	<i>\$_COOKIE['PERMANENT_ID']</i>
UserAgent	<i>\$_SERVER['HTTP_USER_AGENT']</i>
Viittaajan tiedot	<i>\$_SERVER['HTTP_REFERER']</i>



Kuva 13: Tiedonkulku PrimeTrackissa ja ohjelma pääkomponentit.

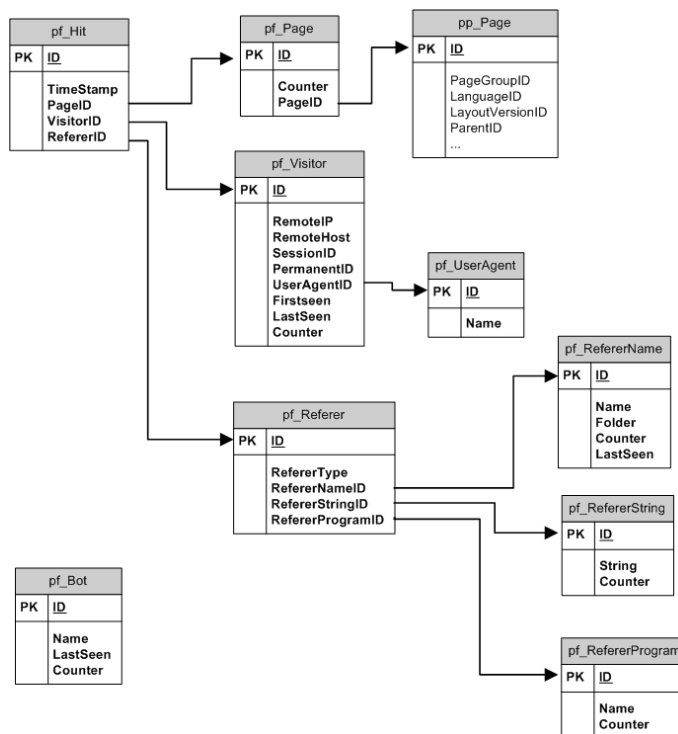


Kuva 14: PrimeTrackin sivulataustenkäsittelijän luokkarakenne.

6.4.2 Tietokanta

Käsittelyn jälkeen tiedot tallennetaan PrimeTrackiä varten suunniteltuun tietokantaan. Tietokanta oli suunniteltava huolella, koska pitkällä aikavälillä tiedonkäsittelyä ja -analysointia voidaan muuttaa, mutta tietokantarakennetta muutettaessa koko ohjelmisto voi tarvita muutoksia ja vääränlainen tietorakenne voi aiheuttaa ongelmia esteinä tai aika-kompleksisesti suurina tehtävinä tiedonanalysoinnissa. [38]

Kuvassa 15 tietokanta on esitetty relaatiomallina tauluineen ja kenttineen. Ensimmäisenä tehtävänä virrankäsittelijä loi joko robotti- tai käyttäjäolion. Tämä tunnistus vaikuttaa myös merkittävästi tietokantatalletukseen. Mikäli sivunlataaja tunnistetaan robotiksi, täytetään ainoastaan taulu *pf_Bot*. Kenttään *Name* tallennetaan robotin nimi, kenttään *Lastseen* tämän hetkinen aika ja laskuria kasvatetaan yhdellä, jotta voidaan tarvittaessa mitata robotin vierailufrekvenssiä. Kun sivulle saapuu oikea kävijä, täytetään vastaavasti sivulataustaulu *pf_Hit* sekä sen alitaulut omine alitauluineen. Sivulataustaulu *pf_Hit* koostuu latausajankohdasta sekä linkeistä sivu-, vierailija- ja viittaajatauluihin. Johtuen relaatioista taulujen täyttö on aloitettava alimmalta tasolta. Sivulataustiedot ja yhteys PrimeContentiin saadaan linkkitaulun *pf_Page* avulla. *CClickstream*-olio saa rakentajassaan parametrina ladattavan sivun tunnisteiden, jolla taulu *pf_Page* täytetään. Seuraavaksi täytetään vierailijan osoite- ja tunnistetiedot tauluun *pf_Visitor* käsittelijästä saatujen tietojen avulla. Taulu sisältää relaation selaintauluun *pf_UserAgent* tilansäästön vuoksi. Viimeisenä täytetään viittaustietotaulu *pf_Referer*. *RefererType*-kenttä on tyypiltään lueteltu tyyppi, joka sisältää vaihtoehdot hakemisto, markkinointiohjelma, hakukone tai oma sivu. Oman sivun ollessa kyseessä asetetaan vain *RefererType* ja muut jätetään tyhjiksi. Hakukoneen ollessa kyseessä täytetään kentät *Type*, *RefererNameID* ja *RefererStringID* sekä niiden alitaulut tarvittavilla tiedoilla. *pf_RefererName*-taulu sisältää viittaajan nimen tai lähdehakemiston sekä viittaussuurekset ja *pf_RefererString* hakukoneen tai hakusanamarkkinointiohjelman hakusanan. Viittaustyyppin ollessa hakemisto toimitaan samoin, mutta viittaajan nimi taulun nimen sijaan täytetään hakemistokenttä *Folder*. Viimeinen vaihtoehto markkinointiohjelma käsittää hakusanamarkkinointiohjelmat kuten Overture ja Google Adwords. Tunnistettaessa ohjelma täytetään myös kenttä *RefererProgramID* ja sen alitaulu. Alitaulu yksinkertaisesti sisältää hakusanamarkkinointiohjelmalle ennalta määritellyn nimen. Näin taulun *pf_Hit* alitaulut on täytetty ja relaatiotunnisteet on saatu kenttien täyttöä varten.



Kuva 15: Relatiokuvaus PrimeTrackin tietokannasta.

6.4.3 Tiedon analysointi

Tiedonanalysointikomponentissa yhdistettiin tulosten tuottaminen ja niiden käyttäjälle näyttäminen. Tulokset saatiin pääosin suoraan SQL (Structured Query Language) -kutsuin. Tiedonnäyttämistä varten tehtiin web-käyttöliittymä, joka toteutettiin osaksi PrimeContentin hallintajärjestelmää.

Liitteessä 2 on kuva PrimeTrackin käyttöliittymästä ja yhteenvetotilastoista. Tietokantatoetuksen vuoksi tilastotiedot on tuotettu melko yksinkertaisin kantahauin, kuten esimerkiksi sivulataukset ovat taulun *pf_Hit* rivien summa ja kävijät erillisten *pf_Hit*-taulun *VisitorID*-indeksien summa. Vierailijat ovat listaus vierailijoiden osoitteista, vierailukerroista ja ajoista. Tämäkin listaus on yksinkertaisesti muotoiltu tuloste annetun aikavälin riveistä taulussa *pf_Visitor*. Ohjelman osio polut sisältää tiedot vierailijan osoitteesta, mistä hän on tullut, mihin aikaan hän on sivuilla ollut ja millä sivuilla hän on vierailut. Tiedontuottaminen on toteutettu simuloimalla kävijän liikkumista sivuilla käymällä tietyn kävijän sivulataukset aikajärjestyksessä. Yksittäisen kävijän vierailu poluissa näyttää seuraavalta:
2006-10-12 16:23:27

Osoitteesta: hoas-fe17dd00-92.dhcp.inet.fi

Saapumistapa: Hakukoneesta Google Haku hakusanalla verkkosivut

Reitti: Etusivu -> Etusivu -> Mitä -> Etusivu -> Mitä -> Kenelle -> Miksi -> Yritys -> Yhteystiedot

Web-käyttöliittymä toteutettiin niin, että kirjautumalla sivustonhallintajärjestelmään pääsee katsomaan myös tilastoja. Apuna käytettiin PrimeContentin alustan tarjoamaa sivumallimootoria. Ensimmäisessä versiossa käyttöliittymä tarjoaa tilastot ainoastaan liitteen 2 mukaisina taulukoina ja valmiiksi määritellyjä aikarajoja ei toistaiseksi ole.

7 YRITYKSEN SIVUJEN HAKUKONEOPTIMOINTI

Tärkeä osa diplomityötä ja hakukoneoptimoinnin testausta oli sivuston optimointi. Ensimmäisenä pilottisivustona hakukoneoptimoinnille päätettiin käyttää yrityksen omia sivuja. Tavoitteeksi ei asetettu mitään konkreettisia paikkoja, vaan haluttiin lähinnä nähdä miten nopeasti ja millaisiin tuloksiin eri toimenpiteet johtavat. Koealueena hakukoneoptimointi ei ainakaan antaisi liian kaunista kuvaa tuloksista. Alaan liittyvät hakusanat ovat voimakkaasti kilpailtuja. Yritysten keskinäinen kilpailu on ammattitaidosta johtuen kovaa ja asiakkaille tehdyissä palveluntarjoajaoppaissaakin tärkeänä kriteerinä pidetään palveluntarjoajan omaa näkymistä [48, 49]. Sivusto optimoitiin myös perinteisiä markkinointipalveluita ja kotisivuja ajatellen, mutta tulosten vertailussa käytettiin hakusanoja hakukoneoptimointi, hakukonemarkkinointi, internetmarkkinointi, markkinointipalvelut ja hakusanamarkkinointi. Tässä luvussa esittelemme alkutilanteen, tehdyt toimenpiteet ja tuloksia optimoinnin aiheuttamasta kehityksestä.

7.1 Alkutilanne

Lähtötilanteessa toukokuussa 2006 sivuston näkyvyyteen hakukoneissa ei ollut kiinnitetty huomioita eikä kävijämääriä mitattu. Lähtökohtien selvittämiseksi tutkittiin sivustoa hakukoneoptimoinnin kannalta olennaisien elementtien kannalta. Vaikka kävijätietoa ei ollut aktiivisesti kerätty pyrittiin se selvittämään sivupalvelimen lokien avulla.

Sivuston alkutilanteen arviointi aloitettiin tutkimalla sivuston teknistä toteutusta. Primesoft mainostaa toteuttavansa sivustot W3C-standardin mukaisesti, joten yrityksen omatkin sivut läpäisivät validointitestit. Seuraavaksi tutkittiin sivusto tekstipohjaisen lynx-selaimen avulla, jotta saatiin selville näkevätkö hakurobotit sivuston sisällön. Myöskin tämän testin sivusto läpäisi. Tutkittaessa lähdekoodiin lisättäviä selitteitä, joiden avulla käytettävyyttä ja luokitusta voidaan korottaa, löytyi puutteita. Kuvien vaihtoehtoistekstejä, linkkien kuvauksia eikä metatietoja ollut käytetty hyväksi.

Teknisen toteutuksen jälkeen tarkistettiin sivuston sisältö. Sivujen otsikkokenttiä ei käytetty ollenkaan hyödyksi, vaan jokaisella sivulla oli sama teksti. Yhtäkään sanoista ei löytynyt sivuston otsikoista, kappaleista eikä linkkiteksteistä. Jo tässä vaiheessa sivuston esiintymisestä hakukoneissa halutuun sanoin voitiin vetää omat johtopäätökset.

Sivuston tutkimisen jälkeen siirryttiin ulkosivullisiin seikkoihin. Selvitys aloitettiin tar-

kistamalla, onko sivusto indeksoitunut. Johtuen standardista toteutuksesta, sivukartasta ja yksinkertaisesta rakenteesta hakukoneet olivat indeksoineet sivuston lähes täysin. Toisin esimerkiksi Google ei esittänyt listauksessaan kaikkia sivuja johtuen mahdollisesti kaikilla sivuilla olevasta samasta otsikkotiedosta. Ulkoisen linkityksen selvitykseen käytettiin apuna Googlea. Ainoastaan yrityshakemisto Menoksesta ja Primesoftin asiakkaan syntyneet.fi-palvelun sivuilta löytyivät tunnistettavat linkit.

Kävijämäärien arviointiin käytettiin apuna Webalizer-lokianalysointiohjelmaa ja palvelimen toukokuulta tallettamia tietoja. Analyysin perusteella sivuille oli tehty 350 vierailua. Vierailijoista 20 oli saapunut sivustolle hakukoneiden kautta ja viisi jonkun ulkoisen linkin kautta. Kävijämäärien lisäksi mitattiin sivuston sijainnit tärkeimpien hakukoneiden listauksissa annetuin hakusanoin. Sivusto ei esiintynyt kymmenellä ensimmäisellä hakutulossivulla yhdelläkään annetuista hakusanoista.

7.2 Toimenpiteet

Edellisen luvun perusteella lähtökohdat hakukoneoptimointiin olivat suhteellisen hyvät. Tekninen toteutus oli kunnossa, mutta sisältö ja ulkoinen näkyvyys olivat puutteellisia. Yrityksen sivuston tapauksessa avainsana-analyysiin ei ollut aihetta, koska avainsanat olivat jo tuttuja alan tuntemuksen ja aiempien markkina-analyysien kautta. Hakukoneoptimointia suoritettaessa yrityksen sivustoa uusittiin myös runsaasti, joten sisältöoptimointia tehtiin runsaasti jo uuden sisällön tuotantovaiheessa. Uusien sivujen julkaisun jälkeen sivustolle asetettiin PrimeTrack-kävijäseuranta ja kartoitettiin kotimaiset hakemistot ja ilmoitettiin sivusto niihin.

Uuden sisällön tuotantovaiheessa siis hakukonenäkyvyys otettiin jo huomioon. Muokkauksista jatkettiin kuitenkin aloittamalla eri sivujen otsikoista. Otsikoihin lisättiin yrityksen nimen lisäksi sivua koskevat tärkeimmät hakusanat. Esimerkiksi hakukonemarkkinointipalveluita esittelevälle sivulle otsikoksi laitettiin: Primesoft Oy - Hakukonemarkkinointi, hakusanamarkkinointi, hakukoneoptimointi Seuraavaksi aliotsikkoja ja sisältöjä muokattiin niin, että useat niistä sisälsivät avainsanoja perusmuodoissaan haittaamatta kuitenkaan luettavuutta. Käytettävyyden parantamiseksi kaikkiin kuviin lisättiin vaihtoehtoistekstit ja linkkeihin niiden kuvaukset.

Ulkoisen linkityksen lisäämiseksi kotimaiset hakemistot kartoitettiin hakukoneiden avulla. Sekä ilmaisia että maksullisia hakemistoja löytyi useita. Hakemistot hintatietoineen kerättiin tietopankkiin, jotta asiakkaat saadaan nopeasti lisättyä hakemistoihin ja hintatie-

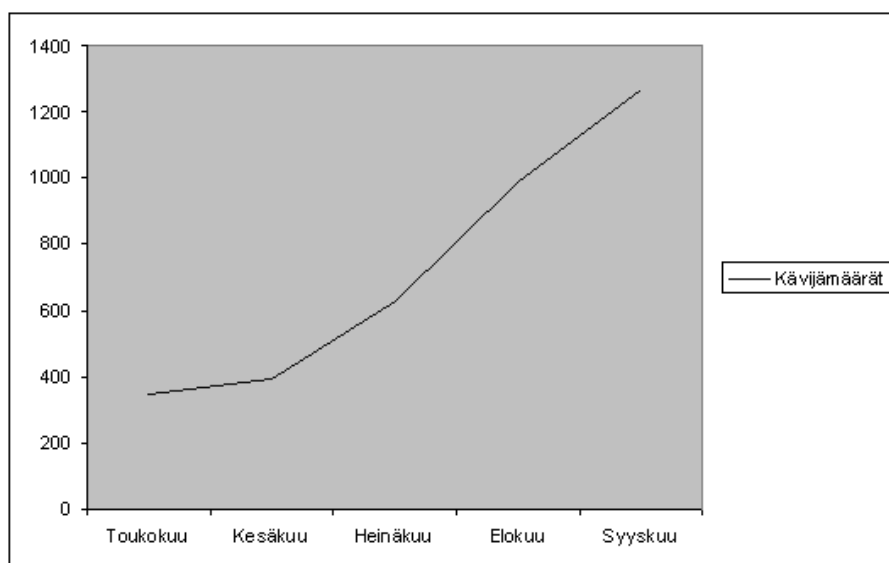
tojen perusteella voidaan tehdä kustannusarvioita näkyvyydestä eri hakemistoissa. Tässä vaiheessa ei vielä itse haluttu sijoittaa maksullisiin hakemistoihin, mutta sivusto ilmoitettiin seuraaviin ilmaisiin hakemistoihin:

- etusivu.info
- www.hakukone.net
- www.linkkeja.com
- dmoz.org

7.3 Tulokset

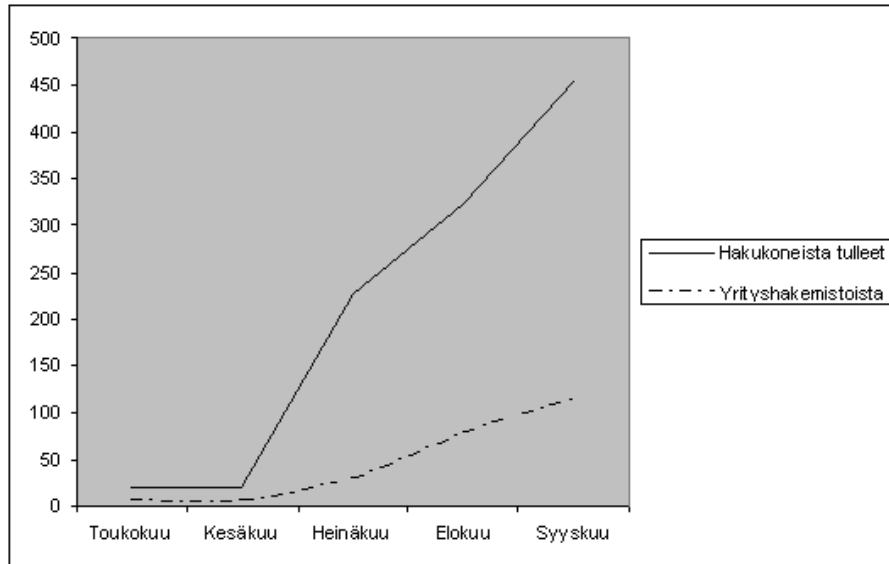
Hakukoneoptimoinnin tehokkuutta tutkittiin viiden kuukauden aikavälillä toukokuusta 2006 syyskuun 2006 loppuun. Mittareina käytettiin kävijämäärien muutoksia, sijoitusten kehitystä hakukonelistauksissa, viittaajien määriä ja hakukoneissa käytettyjä sanoja, joilla sivuille saavuttiin.

Tutkittaessa sivujen kävijämääriä lähtötilanne oli 350 kävijää kuukaudessa. Sivuston uusinnan ja paremman fokuksionnin vaikutusta kehitykseen on vaikea arvioida, mutta määrät kasvoivat runsaasti. Tutkimuksen viimeisessä kuussa kävijämäärät olivat nousseet noin yli 250 prosenttia 1268 kävijään. Kuvassa 16 on esitetty kävijämäärien kehitys.



Kuva 16: Yrityksen sivuston kävijämäärien kehitys seuranta-ajanjaksolla.

Linkkikampanjan ja optimoinnin ansiosta myös kävijämäärät hakemistoista ja hakukoneista kasvoivat. Optimoinnin alkuvaiheessa hakukoneista tuli 20 kävijää ja ulkoisista linkeistä viisi. Syyskuussa vastaavat lukemat olivat 455 ja 114. Kuvassa 17 on esitetty ulkoisista lähteistä tulevien määrät optimointiaikavälillä.



Kuva 17: Ulkoisista lähteistä tulevien kävijöiden kehitys seuranta-ajanjaksolla.

Myöskin käytetyissä hakusanoissa oli havaittavissa selvää kehitystä. Pienistä määristä johtuen yleisimpiä hakusanoja maaliskuussa on vaikea määrittää, mutta ne koostuivat pääosin asiakkaiden ja työntekijöiden nimistä sekä muista ei tuotteisiin kohdistuneista haustoista. Syyskuussa tilanne taas oli toinen. Epäolennaiset hakusanat puuttuivat listoilta ja haut olivat selvästi kohdistuneet yrityksen tuotteisiin. Taulukoissa 10 ja 11 on esitelty vertailun vuoksi viisi suosituinta hakusanaa maaliskuulta ja syyskuulta hakumäärineen.

Taulukko 10: Toukokuun käytetyimmät hakusanat ja niiden käyttömäärät.

1.	valofirma	8
2.	site:www.primesoft.fi	4
3.	cine light rental	1
4.	neliskulma oy	1
5.	ruutuvihko ohjelmisto	1

Taulukko 11: Syyskuun käytetyimmät hakusanat ja niiden käyttömäärät.

1. Kotisivut	160
2. nettisivut	25
3. primesoft	25
4. hakukoneoptimointi	12
5. kotisivu	11

Sijoitusten kehitystä selvitettiin etsimällä sivustoa hakukoneista mitattavin sanoin. Alkutilanteessa sivustoa ei löytynyt ollenkaan, mutta kehitystä oli havaittavissa jo muutamien päivien jälkeen. Osissa hakukoneista päästiin kärkisijoille hyvin lyhyellä aikavälillä ja osissa taas kehitystä ei suuremmin tapahtunut. Tutkimus antoi kuitenkin viitteitä, että hakukoneet käyttävät luokitukseen paljon samoja muuttujia mutta erilaisin painoarvoin. Suomalainen hakukone www.fi ilmoittaa, että se luokittelee sivut metatietojen perusteella [27]. Kärkisijoille päästiinkin nopeasti metatietojen lisäyksen jälkeen. Vastapainoisesti Yahoo:ssa ja Googlessa kehitys oli hitaampaa eikä selkeitä vaikutuksia pystytty päättämään. Taulukossa 12 on esitetty yrityksen sivuston sijainti käytetyin hakusanoin eri hakukoneissa.

Taulukko 12: Yrityksen kotisivujen sijoittuminen hakukoneissa syyskuun lopussa.

Hakusanat	Google	Yahoo	MSN	www.fi
Hakukoneoptimointi	Ei ensimmäisellä 10 sivulla	Ei ensimmäisellä 10 sivulla	sivulla 5	sivuilla 2-3
Hakukonemarkkinointi	sivulla 2	sijalla 5	sijalla 3	sijoilla 1-5
Internetmarkkinointi	sivulla 2	Ei ensimmäisellä 10 sivulla	sijalla 8	sijoilla 1-5
Markkinointipalvelut	sivulla 3	Ei ensimmäisellä 10 sivulla	sijalla 6	sivulla 3
Hakusanamarkkinointi	sijalla 3	Sijalla 1	sijalla 3	sijoilla 1-5

7.4 Kokemuksia toiminnasta

Suoritettaessa hakukoneoptimointia esille tuli muutamia yllättäviäkin seikkoja. Kehityksen vaikea seurattavuus, tulosten arviointi ja tulostensaannin kestäminen tekee hakukoneoptimoinnista vaikean artikkelin myydä.

Suomen mittakaavassa vaikuttaminen hakutuloksiin vaikutti suhteellisen helpolta. Tutkit-
taessa muita aloja vaikutti usein, että pelkkä otsikon muokkaaminen riittää kärkisijoihin.
Verrattuna omiin sivuihin pahimman vastuksen vaikuttavat kuitenkin antavan yritysten
avuksi tehdyt hakemistot. Hakemistot hyvine linkityksineen ja korkeine luokituksineen
ovat vaikeita ylittää ja tällöin yrityksen sivustot jäävät helposti taka-alalle.

Lähdemateriaalissa tärkeänä hakukoneoptimoinnin mittarina pidetään Google Pageran-
kia. Tulimme kuitenkin johtopäätökseen, että vähemmän kilpailluilla toimialoilla luoki-
tuksen kasvattaminen on vaikeaa ja sivustot ovat luokassa 0-2. Luokitus on selkeästi ke-
hitetty suurempia kävijä- ja linkitysmääriä ajatellen eikä skaalaudu pienemmille paikal-
lisille sivustoille. Oman sivustomme luokitus kehityksestä huolimatta vaihteli nollan ja
kahden välillä ollen viimeisessä kuussa etusivulla 1 ja muilla sivuilla nolla.

Tärkeänä tekijänä pidettiin myös ulkoisia linkittäjiä. Linkittäjien päivittyminen ei tapahdu
automaattisesti robottien vieraillessa sivuilla, vaan esimerkiksi Google tekee satunnaisin
ajoin linkityspäivityksiä, jolloin hankitut linkit tulevat huomioiduksi luokituksessa. Lisä-
simme yrityksemme hakemistoihin huhtikuussa ja mittauksen loppuessa syyskuussa uusia
linkkejä ei edelleenkään oltu huomioitu.

8 YHTEENVETO

Internetmarkkinointi on voimakkaimmin kasvava markkinoinnin ala. Tämän vuoksi jokaisen yrityksen olisi otettava huomioon internetmarkkinoinnin potentiaali sekä uhat. Säilyttääkseen pysyvän kehityksen on oltava siellä missä asiakkaatkin. Hakukoneoptimointi tarjoaa keinon parempaan näkymiseen internetissä. Toimintona sitä voidaan monessa mielessä pitää monen muun markkinointitoiminnon tapaan "listana hyväksi todettuja menettelytapoja", joka kehittyy kokemuksen kautta. Ymmärtämällä hakukoneoptimoinnin taustalla oleva filosofia voi sivustokehittäjä päästä pitkällekin. Peltikattoja myyvää yritystä on lähes mahdoton löytää, jos sivuilla ei mainita kertaakaan sanoja peltikatto tai peltikatot.

Tässä työssä tutustuttiin hakukoneoptimointiin kirjallisen tutkimuksen kautta ja toteutettiin Primesoftin vaatimukset aiheutta koskien. Kirjallinen tutkimus aloitettiin valaisemalla hieman internetmarkkinoinnin kenttää, jonka jälkeen siirryttiin hakukoneoptimointiin perehtymällä hakukoneiden maailmaan. Ymmärtääkseen hakukoneoptimoinnin toiminnan on hyvä myös tuntea hakukoneiden toiminnan perusteet ja erityisesti tuntea, miten potentiaaliset asiakkaat hakukoneita käyttävät. Taustaselvityksen jälkeen tutustuttiin itse hakukoneoptimointiprosessiin, sen eri vaiheisiin ja toiminnon mittaukseen kävijäseurannan kautta. Käytännön osio aloitettiin kartoittamalla Primesoftin asettamat vaatimukset ja tavoitteet projektin suhteen. Kartoituksesta siirryttiin toteutukseen, jossa käsiteltiin yrityksen ohjelmistojen hakukoneystävälliseksi muokkausta ja kävijäseurannan toteutusta. Lopuksi toteutetun järjestelmän toimintaa testattiin yrityksen sivuilla ja raportoitiin tulokset. Tieteellisten julkaisujen vähyydestä huolimatta kirjallisesta tutkimuksesta saatiin suhteellisen kattava ja se avasi yritykselle myös muun internetmarkkinoinnin maailman. Yrityksen asettamat tavoitteet saatiin pääosin tavoitettua. Siinä mielessä onnistuttiin jopa suhteellisen hyvin, että hakukoneoptimointi on vahvasti tietotaitoon perustuva toiminto eikä konkreettista toteutusta esimerkiksi työkalujen muodossa voida tehdä. Yrityksen omilla sivuilla tehdystä kokeilusta saadut tulokset olivat rohkaisevia ja hakukoneoptimointi todettiin toimivaksi.

Kirjallisen tutkimuksen aikana tuli esille myös muita jatkokehityskohteita, joista tärkeimpänä voidaan pitää toista hakukonemarkkinoinnin osa-aluetta, hakusanamarkkinointia. Hakusanamarkkinoinnilla tarkoitetaan hakukoneiden markkinointiohjelmien käyttöä. Tällaisia ohjelmia ovat muun muassa Googlen Google Adwords ja Yahaon Overture. Hakusanamarkkinoinnin toiminta on hyvin yksinkertaista. Markkinoija luo itselleen ohjelmaan tilin, määrittää hakusanat, joilla haluaa näkyä, ja tekee mainokset. Mainoksilla tarkoitetaan lyhyitä muutaman kymmenen merkin mittaisia mainostekstejä. Nämä mainokset näkyvät eroteltuna sponsoroiduissa tuloksissa ja sijoitukset riippuvat yleensä mainokselle määri-

tetystä klikkihintoista (Cost Per Click). Etuina hakukoneoptimointiin voidaan pitää välitöntä näkymistä, avainsanojen rajatonta määrittelyä ja mainonnan alueellista kohdistusta. Perinteiseen bannerimainontaan verrattuna etuina voidaan pitää hinnoittelua ja riskittömyyttä. Hakusanamarkkinoinnissa hinnoitteluperusteena käytetään sivustolle mainoksen kautta saapuneita asiakkaita. Ja mikä parasta, vain harva hakukoneiden käyttäjä osaa tehdä eron sponsoroidun ja oikean hakutuloksen välille.

LÄHTEET

- [1] Sullivan P. Intro to search engine optimization. 2002. Verkkosivu.
<http://searchenginewatch.com/showPage.html?page=2167921>. 30.8.2006.
- [2] Parantainen J., Poutiainen R. Google-markkinoijan opas. Total Management Oy & Noste Oy, 2005.
- [3] Summa S. Nettimainonta kiilaa radiomainonnan. Verkkouutinen 22.8.2006.
http://www.tietokone.fi/uutta/uutinen.asp?news_id=27774. 30.8.2006.
- [4] Jarboe G. Search engine marketers spent \$5.75 billion in 2005. Verkkouutinen 22.8.2006.
http://www.sempo.org/news/releases/Search_Engine_Marketers. 30.8.2006.
- [5] Eager B. Complete Idiot's Guide to Online Marketing. Alpha, 1999.
- [6] Search and Go. An industry build upon search. 2006. Verkkosivu.
<http://www.searchandgo.com/articles/search-engine-marketing.php>. 30.8.2006.
- [7] deGeyter S. Online marketing hierarchy and definitions. Technical report, Search Engine Guide, 2005. Verkkosivu.
<http://www.searchengineguide.com/degeyter/006428.html>. 30.8.2006.
- [8] Ramos A., Cota S. Insider's Guide to SEO. Jain Publishing, 2004.
- [9] Lehto T. Google rankaisee suomalaisyrityksiä. Verkkouutinen 19.4.2006.
http://www.tietokone.fi/uutta/uutinen.asp?news_id=26728. 30.8.2006.
- [10] Moran M., Hunt B. Search Engine Marketing, Inc. IBM Press, 2006.
- [11] Grossan B. Search engines: What they are, how they work, and practical suggestions for getting the most out of them. 1997. Verkkosivu.
<http://www.webreference.com/content/search/index.html>. 30.8.2006.
- [12] Sullivan D. How search engines work. 2002. Verkkosivu.
<http://searchenginewatch.com/showPage.html?page=2168031>. 30.8.2006.
- [13] Pant G., Srinivasan P., Menczer F. Crawling the web, 2004.
- [14] Baeza-Yates R., Castillo C., Marin M., Rodriguez A. Crawling a country: Better strategies than breadth-first for web page ordering. *Proceedings of the 14th international conference on World Wide Web / Industrial and Practical Experience Track*, pages 864–872, Chiba, Japan, 2005. ACM Press.

- [15] Koster M. A standard for robot exclusion. Verkkosivusto.
<http://www.robotstxt.org/wc/norobots.html>. 30.8.2006.
- [16] Baeza-Yates R., Castillo C. Balancing volume, quality and freshness in web crawling. *Hybrid Intelligent Systems*, 2002.
- [17] M. Solomon, Bamossy G., Askegaard S. Consumer Behaviour - A European Perspective. Prentice Hall Europe, 4th edition, 1999.
- [18] Kotler P. Marketing management: analysis, planning, implementation and control. Prentice Hall Europe, 11th edition, 2002.
- [19] Hotchkiss G, Jensen S, Jasra M, Wilson D. The role of search in business to business buying decisions. Technical report, Enquiro Search Solutions, 2004.
- [20] Armstrong L. Kicking tires on the web. *Business Week*, pages 120E1–120E2, 4 1999.
- [21] Muylle S., Moenaert R., Despontin M. A grounded theory of world wide web search behaviour. *Journal of Marketing Communications*, volume 5, pages 143–155. Routledge, part of the Taylor & Francis Group, 1999.
- [22] Beal A. Search engine optimization basics. Technical report, Search Engine Guide, 2003. Verkkosivu.
http://www.searchengineguide.com/beal/2003/0205_ab1.html. 30.8.2006.
- [23] Banks L. Search engine optimization basics, part 2: Seo keyword and infrastructure strategies. Technical report, IBM Press, 2006. Verkkosivu.
<http://www-128.ibm.com/developerworks/web/library/wa-seo2/>. 30.8.2006.
- [24] Fishkin R. Search engine ranking factors. Verkkosivusto.
<http://www.seomoz.org/articles/search-ranking-factors.php>. 31.10.2006.
- [25] Laki verkkotunnuslain muuttamisesta. 2006. 187/2006.
<http://www.finlex.fi/fi/laki/alkup/2006/20060187>. 1.11.2006.
- [26] Sullivan D. Death of a meta tag. 2002. Verkkosivu.
<http://searchenginewatch.com/showPage.html?page=2165061>. 30.8.2006.
- [27] WWW.FI. Hakuohjeet: näin pääset alkuun. verkkosivusto.
<http://www.fi/Help>. 30.8.2006.
- [28] Sullivan D. How to use html meta tags. 2002. Verkkosivu.
<http://searchenginewatch.com/showPage.html?page=2167931>. 30.8.2006.

- [29] Farber D. Eye tracking web usability. Verkkosivusto.
<http://blogs.zdnet.com/BTL/?p=2776>. 30.8.2006.
- [30] Savolainen V. Verkkoteorian perusteet ja algoritmit. Gaudeamus, Vaasa, 1978.
- [31] Brown R. emarketing - what is link popularity? 2003. Verkkosivu.
<http://www.seoachat.com/c/a/Link-Trading-Help/What-Is-Link-Popularity/>.
 30.8.2006.
- [32] Brin S., Page L. The anatomy of a large-scale hypertextual Web search engine.
Computer Networks and ISDN Systems, 30(1–7):107–117, 1998.
- [33] Boggs C. Choosing link partners. 2003. Verkkosivu.
http://www.sempo.org/learning_center/articles/link-partners. 30.8.2006.
- [34] Alexa Internet Inc. Alexa web search. Verkkosivusto.
<http://www.alexa.com/>. 31.10.2006.
- [35] Wall A. Determining the value of links from link renting. 2005. Verkkosivu.
<http://seotoday.com/browse.php/category/articles/id/484/index.php>. 30.8.2006.
- [36] Clay B. Seo code of ethics. Verkkosivusto.
http://www.bruceclay.com/web_ethics.htm. 30.8.2006.
- [37] Hallam-Baker P., Behlendorf B. W3c working draft: Extended log file format. verkkosivu.
<http://www.w3.org/TR/WD-logfile.html>. 30.8.2006.
- [38] Sweiger M., Madsen M., Langston J., Lombard H. Clickstream DataWarehousing.
 John Wiley & Sons Inc., 2002.
- [39] Connolly D. Rfc 2616: Hypertext transfer protocol – http/1.1. verkkosivu.
<http://www.w3.org/Protocols/rfc2616/rfc2616.html>. 30.8.2006.
- [40] Malacinski A., Dominick S., Hartrick T. Measuring web traffic. verkkosivu.
<http://www-128.ibm.com/developerworks/web/library/wa-mwt1/>. 30.8.2006.
- [41] Dainow B. Key performance indicators you should track. Verkkosivusto.
<http://www.webanalyticsassociation.org/en/art/?8>. 30.8.2006.
- [42] Jasra M. Analytics analysis for search marketers. Verkkosivusto.
<http://www.webanalyticsassociation.org/en/art/?72>. 30.8.2006.
- [43] Web Analytics Association. Onsite behaviour - path analysis. 2000.

- [44] Alexander J. Is someone draining your pay per click account? find out & stop them. 2005. Verkkosivu.
<http://www.searchengineworkshops.com/articles/clickfraud.html>. 30.8.2006.
- [45] Kiiskinen S., Linkoaho A, Santala R. Prosession johtaminen ja ulkoistaminen. WS Bookwell Oy, 2002.
- [46] Primesoft Oy:n toimitusjohtaja Kalle Pyörälä. Keskustelu 11.07.2006.
- [47] Primesoft Oy:n markkinointi ja myynti Simo Hämäläinen. Keskustelu 13.07.2006.
- [48] McGee M. Choosing a search engine marketing firm. 2005. Verkkosivu.
http://www.sempo.org/learning_center/articles/choosing_a_search_engine_marketing_firm/. 30.8.2006.
- [49] Poutiainen R. 10 kysymystä, jotka sinun kannattaa kysyä, kun valitset seo-firmaa. Infolehti, 2006.
- [50] Primesoft Oy:n toimitusjohtaja Kalle Pyörälä. Keskustelu 20.07.2006.
- [51] Saransaari K. WWW-sivuston kävijäseurannan uudistaminen. PhD thesis, Lappeenrannan teknillinen yliopisto, tietotekniikan osasto, 2006.
- [52] Google. Google analytics. Verkkosivusto.
<http://www.google.com/analytics/fi-FI/>. 30.8.2006.
- [53] Bradford B. The webalizer - what is your server doing today? Verkkosivusto.
<http://www.mrunix.net/webalizer/>. 30.8.2006.
- [54] Snoobi Oy. Snoobi-kävijäseuranta. Verkkosivusto.
<http://www.snoobi.fi/>. 30.8.2006.
- [55] Crosby B. The official google analytics blog: We're open! instant access now available. Verkkosivusto.
http://analytics.blogspot.com/2006/08/were-open-instant-access-now-available_15.html. 30.8.2006.

LIITE 1. SEOmoz: Search Engine Ranking Factors






Categories included

Most important factors






URL, Technical, Hosting & Server Side Factors:

Introduction to Ranking Factors

This article contains a large list of the factors that can influence a web document's rank at the major search engines (Yahoo!, MSN, Google & AskJeeves) for a particular term or phrase. Although it is impossible to say for certain which of these items affects which search engine or how important the factors are individually, I've created an estimated ranking importance scale as indicated by the following symbols:

	Exceptional Importance
	High Importance
	Moderate Importance
	Slight Importance
	Inconsequential Importance

For each factor, the 13 individuals (myself and 12 others in the field) voted on each factor's perceived importance. I then averaged these votes to come up with the scores you see listed next to the factors. In addition, the standard deviation (which measures how much fluctuation existed between votes) is also listed and a scale has been created to show if there was wide disagreement or general consensus on a factor's importance.

	Highly Disputed
	Somewhat Disputed
	Average Agreement
	Some Consensus
	High Consensus

List of Contributors

The following individuals made this project possible by submitting their opinions on each ranking factor, and adding comments where they felt it important. I have taken note of many of these comments in my description of the ranking factors and each score is representative of the group's overall opinions. Many thanks to:

Danny Sullivan, <http://www.searchenginewatch.com>

Dan Thies, <http://www.seoresearchlabs.com>

EGOL, Profile @ SEOmoz

Graywolf, <http://www.wolf-howl.com>

Jill Whalen, <http://www.highrankings.com>

Donna Fontenot, <http://www.seo-scoop.com>

Michael Martinez, <http://www.michael-martinez.com/>

Bill Slawski, <http://www.cre8asiteforums.com>

Ammon Johns, <http://www.webmarketingplus.co.uk/>

Scottie Claiborne, <http://www.successful-sites.com/>

(jatkuu)

Liite 1. Jatkoa

2K, <http://www.2kmediat.com/kkmediat/eng/>

Todd Malicoat, <http://www.stuntdubl.com>

Rand Fishkin, My votes are also included in the avgs







Most Important Factors

The following are the top 10 ranked factors across the 5 categories:

1. Title Tag - 4.57
2. Anchor Text of Links - 4.46
3. Keyword Use in Document Text - 4.38
4. Accessibility of Document - 4.3
5. Links to Document from Site-Internal Pages - 4.15
6. Primary Subject Matter of Site - 4.00
7. External Links to Linking Pages - 3.92
8. Link Popularity of Site in Topical Community - 3.77
9. Global Link Popularity of Site - 3.69
10. Keyword Spamming - 3.69







URL, Technical, Hosting & Server Side Factors:

These factors affect a document's rankings by virtue of their affect on technical, spidering/indexing or trust areas of the search engines' rankings.

	Factor Name	Description
	Accessibility of Document Avg 4.31 Std. Dev. 1.32	Inaccessible documents could be due to 404 errors, server mishaps, plug-in requirements or other technical issues.
	Session ID Variables Avg 3.62 Std. Dev. 1.08	Session variables can produce hazardous search rankings as duplicates of a URL's contents are indexed hundreds or thousands of times, diluting the incoming links and ranking power of the page as a stand-alone document.
	Noarchive/Noindex Use & Robots.txt Avg 3.31 Std. Dev. 1.54	The noarchive and noindex directives instruct search engines not to cache the page in their index or save its content. Robots.txt can also be used to issue these commands.
	Dynamic Parameters in URL Avg 3.23 Std. Dev. 1.19	Multiple parameters in a dynamic URL can often mean that a search engines spider will choose to ignore the document. SE reps have long urged site creators to limit dynamic parameters to 2, optimally to 1 so that spiders will easily crawl them.
	Use of Web Page Frames Avg 3.15 Std. Dev. 1.10	Frames can impact many SEO-related issues as the separation of URLs can devalue the links coming to the page(s) and dilute the ranking power of a single document over multiples.
	Hosting Uptime Avg 3.00 Std. Dev. 1.18	If a site's hosting is inconsistent and frequently not returning pages, the search engines may devalue the rankings of documents at the site, or worse, may not spider many pages.

(jatkuu)

Liite 1. Jatkoa

	Geographic Hosting Location Avg 2.54 Std. Dev. 1.39	Geography of hosting may play a small role in determining language or country specific targeting and relevancy at the search engines.
	IP-Based Content Delivery Avg 2.08 Std. Dev. 1.21	Although difficult to detect, search engines finding evidence of IP-based content delivery on a website are likely to penalize or devalue that site in the results.
	Domain Registration Information Avg 1.69 Std. Dev. 0.82	The domain registration information may be used by search engines to identify spammers or violators of guidelines and reduce or penalize the value of their sites in the SERPs.
	Domain Registration Length Avg 1.62 Std. Dev. 0.74	The length of time a domain has been registered for may be used to determine the legitimacy or trust given to a site. Longer registration periods may denote greater trust placed in the site.
	URL Length Avg 1.46 Std. Dev. 0.63	The overall length of a URL may be used as a component in trust or value - exceptionally long URLs might be viewed as unusable and more likely to be spam.
	Robots.txt Presence Avg 1.31 Std. Dev. 1.07	The existence of a robots.txt file indicating that crawling and caching by search engines is permitted may have influence at any of the major search engines.

Maintainer: Rand Fishkin (rand@seomoz.org), SEOmoz (www.seomoz.org)
Source: <http://www.seomoz.org/articles/search-ranking-factors.php>

LIITE 2. PrimeTrack esimerkkitulo

PrimeTrack - Tilastot

Tilasto Vierailijat Polut

Rajaa ajalle	<input type="text" value="1.10.2006"/>
-	<input type="text" value="15.10.2006"/>
Näytä	<input type="text" value="10"/> Ensimmäistä Top listoissa.
	<input type="button" value="Käytä"/>

Tilastot ajalle 1.10.2006-15.10.2006. Näytetään 10 ensimmäistä Top-Listoissa.

YHTEENVETO		
Kävijöitä		
Yhteensä	792	
Uusia	781	
Vierailuja	841	
Sivulataukset		
Yhteensä	1609	
Per käyttäjä	2.1	
Sivulle saapumistavat		
Kirjoittamalla URL/muutoin	423	
Hakukoneista	35	
Hakukonemarkkinointiohjelmista	249	
Linkit	44	
TOP - KATSOTUIMMAT SIVUT		
Sija	Sivu	Osumat
1.	Etusivu	698
2.	Kenelle	176
3.	Yhteystiedot	140
4.	Kotisivut	121
5.	Yritys	108
6.	Hakukonemarkkinointi	104
7.	Mitä	104
8.	Markkinointimateriaalit	74
9.	Miksi	69
10.	Starttipaketti	30
TOP - SAAPUMISSIVUT		
Sija	Sivu	Osumat
1.	Etusivu	497
2.	Kenelle	81
3.	Markkinointimateriaalit	54
4.	Hakukonemarkkinointi	52
5.	Kotisivut	51
6.	Yritys	19
7.	Yhteystiedot	18
8.	Mitä	17
9.	Miksi	17
10.	Starttipaketti	11

(jatkuu)

LIITE 2. Jatkoa

TOP - POISTUMISSIVUT		
Sija	Sivu	Osumat
1.	Etusivu	401
2.	Kenelle	88
3.	Kotisivut	68
4.	Markkinointimateriaalit	59
5.	Hakukonemarkkinointi	57
6.	Yritys	42
7.	Yhteystiedot	34
8.	Mitä	22
9.	Miksi	22
10.	Starttipaketti	19

TOP - HAKUSANAT		
Sija	Sana	Osumat
1.	kotisivut	43
2.	Kotisivu	16
3.	primesoft	12
4.	internet sivut	8
5.	markkinointi	8
6.	hakukoneoptimointi	7
7.	kotisivuja	5
8.	lisää kävijöitä kotisivuille	5
9.	mainos	5
10.	nettisivut	5

TOP - HAKUKONEET		
Sija	Hakukone	Osumat
1.	Google Haku	31
2.	Msn Haku	2
3.	Altavista Haku	2

TOP - HAKUMARKKINOINTIOHJELMAT		
Sija	Ohjelma	Osumat
1.	AdWords	249

TOP - LINKITTÄJÄT		
Sija	Linkki	Osumat
1.	http://www.alexa.com/site/ds/top_sites?cc=FI&ts_mode=country(=none	7
2.	http://www.alexa.com/site/ds/top_sites?ts_mode=lang(=fi	6
3.	http://www.lappeenrantaopas.com/tiedot/Lappeenranta/Primesoft_Oy.html	5
4.	http://www.whois.sc/jarjesto.net	3
5.	http://www.syntyneet.fi/	2
6.	http://www.whois.sc/televisiokauppa.com	2
7.	http://mike.cc.lut.fi/urapalvelut/cgi/lut/jobs.cgi?action=view&id=3975	2
8.	http://www.netcraft.com/survey/	2
9.	http://web.archive.org/web/20040927203932/http://www.primesoft.fi/	2
10.	http://www.whois.sc/xn--vrisuora-0za.com	2