

Alexi Hietaniemi

Kuinka automatisoitu data-analytiikka voidaan integroida hakukonemarkkinoinnin toteutukseen?

Opinnäytetyö

Kuinka automatisoitu data-analytiikka voidaan integroida hakukonemarkkinoinnin toteutukseen?

Aleksi Hietaniemi
Opinnäytetyö
Kevät 2021
Liiketalous
Oulun ammattikorkeakoulu

TIIVISTELMÄ

Oulun ammattikorkeakoulu
Liiketalous, Myynti ja markkinointi

Tekijä: Aleksi Hietaniemi

Opinnäytetyön nimi: Kuinka automatisoitu data-analytiikka voidaan integroida hakukonemarkkinoinnin toteutukseen?

Työn ohjaaja: Miikka Blomster

Työn valmistuslukukausi ja -vuosi: Kevät 2022

Sivumäärä: 58

Opinnäytetyön tavoitteena oli koota opas, jota yritykset voivat hyödyntää, kun he ottavat askeleen kohti älykkäämpää, dataan pohjautuvaa hakukonemarkkinointia. Oppaassa käsitellään hakukoneita, hakukonemainontaa sekä hakukoneoptimointia. Opas tarjoaa myös tietoa siitä, kuinka dataa voidaan hyödyntää helpottamaan markkinoijien työtehtäviä sekä opas auttaa yrityksiä tarjoamaan asiakkaille kohdennettuja www-sivustoja ja mainoksia.

Hakukoneet ja niiden tarjoamat kohdennetut hakutulokset toimivat nykyään yhä suuremmassa roolissa imagoansa rakentavien yritysten ja kohdeasiakkaiden välillä. Hakukonemarkkinointiin käytettävät menot kasvavat jatkuvasti ja ne tulevat tulevaisuudessa täyttämään leijonanosan markkinoinnin budjetista.

Teoria osuudessa käsitellään alan kirjallisuuteen pohjaten mahdollisia ratkaisuja sekä yleisiä toimintamalleja, kuinka automatisoitu datan analysointi voitaisiin ottaa käyttöön yrityksessä ja millaisia toimenpiteitä sen käyttöönotto vaatii. Oppaan avulla lukija saa ymmärryksen hakukonemarkkinoinnin automatisaation kannalta tärkeisiin osa-alueisiin ja täten toimii ohjenuorana automaattisen data-analyysin integroimisessa hakukonemarkkinoinnin toteutukseen. Opinnäytetyö ei ota kantaa siihen, miten yritysten tulisi toteuttaa työssä esille nostettavia toimenpiteitä.

Asiasanat: Hakukoneoptimointi, Automatisoitu data-analyysi, SEO, Markkinointi, Automatisoitu hakukonemarkkinointi, Hakukonemarkkinointi, Tekoäly, Koneoppiminen

ABSTRACT

Oulu University of Applied Sciences
Degree Programme in Bachelor of Business Administration, Option of marketing

Author: Aleksi Hietaniemi

Title of thesis: How automated data analysis can be integrated into the implementation of search engine marketing?

Supervisor: Miikka Blomster

Term and year when the thesis was submitted: Spring 2022

Number of pages: 58

The purpose of this thesis was to construct a guidebook that businesses could use as they progressed toward smarter, data-driven search engine marketing. Search engines, search engine advertising, and search engine optimization are all covered in the guide. The guide explains how the data may be used to help marketers do their jobs and give customers with targeted webpages and adverts.

Search engines and the customized search results they deliver are now playing an increasingly important part in the relationship between organizations that are building their image and their target customers. Search engine marketing spending is steadily rising and will soon account for the lion's share of marketing budgets.

The theoretical section discusses alternative solutions based on the literature, as well as general operational models, how automated data analysis could be implemented in a corporation, and what actions are required for its execution. The guide's understanding of machine learning models important for search engine marketing automation, as well as how machine learning and artificial intelligence can be used in automated data analysis and how that data can be integrated into search engine marketing implementation, provides guidance to its reader. The study does not take a stand on how corporations should put the measures suggested in the paper into action.

Keywords: Search Engine Optimization, Automated Data Analysis, SEO, Marketing, Automated Search Engine Marketing, Search Engine Marketing, Artificial intelligence, Machine Learning

KÄYTETYT LYHENTEET JA SANASTO

AdWords Googlen maksullinen hakusanamainonta.

Alt-teksti Teksti, joka sisältää tietoa kuvasta tai kuviosta.

Big data Erittäin suuri tietojoukko, jota voidaan analysoida laskennallisesti kuvioiden, trendien ja assosiaatioiden paljastamiseksi, erityisesti asioissa, jotka liittyvät ihmisten käyttäytymiseen ja vuorovaikutukseen.

Indeksointi Indeksointi on prosessi, jolla hakukoneet järjestävät tiedot ennen hakua mahdollistaakseen erittäin nopeat vastaukset kyselyihin. Avainsanojen ja aiheiden etsiminen yksittäisiltä sivuilta olisi erittäin hidas prosessi hakukoneiden tunnistaa asiaankuuluvat tiedot.

Indeksioija Googlen verkkoindeksointirobotti etsii uusia ja päivitettyjä sivuja Googlen hakemistoon lisäämistä varten. Hakukonerobotti on algoritmi, joka automaattisesti indeksoi sivustoja ja sivuja.

Konversio Hetki, jossa markkinointiviestin vastaanottaja suorittaa halutun toimenpiteen.

KPI-mittari Määrällisesti mitattava suorituskyvyn mittari tietyn tavoitteen aikana.

Metakuvaus Kuvaus, joilla kuvaillaan ja luokitellaan sisältöä; esimerkiksi metatunnisteiden, avainsanojen tai kuvaustietojen avulla.

Napsautus (valinta) Käyttäjän napsauttaessa linkkiä tai mainosta haku ohjautuu mainostajan tai hakukoneen määrittelemälle verkkosivulle.

Off-page SEO Verkkosivuston ulkopuolella tehdyt toimet, jotka vaikuttavat sijoitukseen hakukoneiden tulossivuilla.

On-page SEO Yksittäisten web-sivujen optimointi ylemmän sijoituksen ja osuvamman liikenteen ansaitsemiseksi hakukoneissa.

Otsikot (H1, H2, H3) H1 = Tärkeimmät avainsanat ja aihe, mistä postauksessa on kyse. H2 = Osiot sisällön hajottamiseen käyttämällä H1-tunnisteen kaltaisia avainsanoja. H3 = Alakategoriat sisällön hajottamiseksi edelleen, mikä tekee siitä helposti ymmärrettävän indeksoijille.

Page Rank Algoritmi, jolla Google määrittää web-sivun arvon ja sijainnin hakutuloksissa.

Robots.txt-tiedosto Rajoittaa Internetiä indeksoivien hakurobottien pääsyä sivustolle. Robots.txt-tiedostoilla ei kuitenkaan ole lainvoimaa, joten kaikki robotit eivät kunnioita pyyntöä. Luottamukselliset tiedot kannattaa aina laittaa salasanan taakse, etteivät roskapostittajat ja muut häiriköt mene sivustolle.

SEO Hakukoneoptimointi.

SEO-Auditointi Prosessi, jossa analysoidaan, kuinka hyvin verkkonäkyvyys liittyy parhaisiin käytäntöihin.

SEO-skeema Tunnisteiden semanttinen sanasto, jonka voi lisätä HTML-koodiin parantaakseen tapaa, jolla hakukoneet lukevat ja näyttävät sivuston hakukoneen tulossivulla.

SERP Hakukoneentulossivu, jonka hakukone näyttää, kun käyttäjä on lähettänyt hakukyselyn.

Sisällönkuvauskenttä Verkkovastaavat voivat toimittaa hakukoneille tietoja sivustoistaan kätevästi käyttämällä metatietoja. Hakukoneet tunnistavat ne, kun sisällönkuvauskentät lisätään HTML-sivun-osioon.

Title-attribuutti Määrittää lisätiedot elementistä.

URL Internet-osoite. Määrittää myös tiedonsiirtoprotokollan, isäntäkoneen nimen ja resurssien hakemistopolun.

XML-sitemap Tiedosto, joka luettelee verkkosivuston tärkeät sivut ja varmistaa, että Google voi löytää ja indeksoida ne.

SISÄLLYS

1	JOHDANTO	8
1.1	Työn tavoitteet ja rajaukset	9
1.2	Työn rakenne, luonne ja menetelmät.....	9
2	HAKUKONEET.....	11
2.1	Hakukonemarkkinoinnin rakenne.....	13
2.2	Hakukoneiden toiminnan ymmärtäminen	14
2.3	Hakukonemainonta.....	17
2.4	Hakukoneoptimointi.....	21
3	DATA JA DATAN ANALYSOINTI DIGITAALISESSA MARKKINOINNISSA.....	24
3.1	Hakukonemarkkinoinnin kannalta oleelliset datatyypit	25
3.2	Suuret datajoukot, Big Data	27
3.3	Datan analysointi hakukonemarkkinoinnissa.....	31
4	KONEOPPIMINEN JA TEKOÄLY	33
5	KUINKA TOTEUTTAA AUTOMAATTISEN DATA-ANALYYSIN INTEGROINTI HAKUKONEMARKKINOINNIN TOTEUTUKSEEN	37
5.1	Automatisoidun tiedonlouhinnan toteuttaminen.....	39
5.2	Automatisoidun data-analyysin integrointi hakukonemarkkinointiin ...	43
5.3	Hakukoneoptimoinnin automatisointi	45
6	JOHTOPÄÄTÖKSET.....	48
7	POHDINTA.....	51
	LÄHTEET	53

1 JOHDANTO

Tiedonhaku verkossa on kaikkialla läsnä oleva ja erittäin tärkeä osa sähköistä kaupankäyntiä. Hakukoneet ovat merkittävässä asemassa online-maailmassa; yli puolet verkkosivujen vierailijoista saapuu verkkosivulle hakukoneen kautta, eivätkä suoran linkin kautta esimerkiksi toiselta verkkosivulta (Introna & Nissenbaum 2000). Hakujen kasvavan merkityksen myötä hakukoneilla on myös yhä suurempi rooli kriittisinä linkkeinä imagonsa rakentamiseen internetiä käyttävien yritysten ja heidän kohdeasiakkaidensa välillä. Yritysten hakukonemarkkinointiin käyttämät menot kasvavat nopeammin, kuin menot muihin verkkomainonnan keinoihin ja analyytikot arvioivat, että hakukonemarkkinoinnin menot täyttävät pian suurimman osan verkkomainonnan kuluista. Verkkomainonta on sähköisen kaupankäynnin tärkein kasvuala. (Garside 2007.)

Digitaalisen markkinoinnin näkökulmasta hakukoneiden käyttö tehokkaina sähköisen kaupankäynnin työkaluina voi olla laajaa PR-toimintojen laajentamista, joka auttaa yrityksiä myymään maailmanlaajuiselle yleisölle, mainostamaan pieniä tai paikallisia myymälöitä edullisesti ja rakentamaan brändejä (Linh 2008, 39-40). Yksi yleisimmistä indikaattoreista, minkä avulla digitaalisen markkinoinnin tehokkuus arvioidaan, on näkyvyys hakukoneissa, sosiaalisissa verkostoissa sekä muilla www-sivuilla. Digitaalisen markkinoinnin tulevaisuus tulee muokkautumaan tekoälyn (Artificial Intelligence, AI) myötä huomattavasti. Tekoälyn avulla hakukonetta käyttävät henkilöt tulevat saamaan kohdistetumpia ja henkilökohtaisempia mainoksia, hakutulokset tulevat olemaan tarkempia ja hakijan mielenkiinnonkohteita vastaavia. Tekoäly vastaa myös yritysten kysymyksiin koskien brändiään ja mitä siitä ajatellaan. (Narayanan 2019.)

Tutkimuksessa keskitytään, kuinka automatisoitu datan analysointi voi auttaa hakukonemarkkinointia, kuinka sen avulla voidaan vaikuttaa hakutulosten näyttöjärjestykseen sekä siihen, voivatko yritykset, tunnetut tai suhteellisen tuntemattomat, erottaa bränditarjouksensa optimoimalla verkkosivustojensa

näyttöjärjestyksen haussa tai saada lisää näkyvyyttä display-mainonalla. Tutkimuksessa käsitellään myös suurten datajoukkojen (Big Data), tekoälyn sekä koneoppimisen (Machine Learning, ML) hyödyntämistä tulevaisuuden digitaalisessa markkinoinnissa. Oppaan tarkoitus on antaa suuntaviivat ja ohjenuorat lukijalleen siitä, miten ja miksi hakukoneita käytetään, kuinka hakukoneissa näkyminen optimoidaan ja kuinka tätä työmäärää voidaan vähentää, kun hyödynnetään tekoälyä ja koneoppimista.

1.1 Työn tavoitteet ja rajaukset

Opinnäytetyön tarkoituksena on avata lukijalleen, mitä on hakukonemarkkinointi, mistä se koostuu ja miksi sitä tulisi tehdä. Tutkimuksessa syvennyttään datan analysointiin ja kuinka se voidaan automatisoida helpottamaan hakukonemarkkinoinnin toteuttamista. Lisäksi tutkimuksessa selvitetään tekoälyn sekä koneoppimisen hyödyntämisen mahdollisuuksista ja tavoista hyödyntää sitä hakukonemarkkinoinnin tukemiseksi. Teoriaosuudessa käsitellään alan kirjallisuuteen pohjaten mahdollisia ratkaisuja sekä yleisiä toimintamalleja, kuinka automatisoitu datan analysointi voitaisiin ottaa käyttöön yrityksessä ja millaisia toimenpiteitä sen käyttöön saattaminen vaatisi. Lukijalle muodostuu tutkimuksen myötä ymmärrys hakukonemarkkinoinnin automatisaation kannalta tärkeisiin koneoppimismalleihin sekä kuinka koneoppimista ja tekoälyä voidaan hyödyntää datan automaattisessa analysoinnissa ja kuinka se data voidaan integroida hakukonemarkkinoinnin toteuttamiseen. Työ ei ota kantaa siihen, miten yritysten tulisi toteuttaa työssä esille nostettavia toimenpiteitä, vaan työn tarkoituksena on toimia ohjenuorana.

1.2 Työn rakenne, luonne ja menetelmät

Opinnäytetyö koostuu kattavasta teoriaosuudesta. Luonteeltaan työ on kirjallisuuskatsaus, joka muodostaa oppaan ja se toteutettiin laadullisena tutkimuksena. Laadullinen tutkimus on luonteeltaan aineistovetoinen. Aineiston tuottamisella ja sen analysoinnilla on keskeinen rooli tutkimuksen

muodostumisessa (Günther & Hasanen 2021). Opinnäytetyön teoriapohjana on laadukas kirjallisuus, tutkimukset sekä blogikirjoitukset. Opinnäytetyössä on edetty Günther ja Hasanen (2021) ohjeiden mukaan tutkimusideasta tutkimusaiheen perehtymiseen. Näiden vaiheiden jälkeen aineisto kerättiin, analysoitiin sekä raportoitiin tutkimukseen. (Günther & Hasanen 2021.)

Teoriaosuus käsittelee kokonaisvaltaisesti digitaalisen markkinoinnin osa-alueita koskien hakukonemarkkinointia, kuten myös dataa, datan analysointia sekä koneoppimista ja tekoälyä. Aktiivisin tiedonhaku tapahtui helmi-, maaliskuu- ja huhtikuussa 2022, jolloin työtä edistettiin laadukkaan kirjallisuuden etsimisellä sekä työn kirjoittamisella. Hakukielenä käytettiin suomea ja englantia. Pääsääntöisenä hakukielenä kuitenkin toimi englanti, sillä tutkimus käsittelee aihetta, josta ei toista vastaavaa tutkimusta löytynyt ja tutkimusongelma on verrattain uusi. Siksi tieto tuli koota ja jäsenellä useista poikkitieteellisistäkin tutkimuksista. Pääsääntöisenä tiedonhakukantana toimi: Google Scholar, Oamk Finna sekä satunnaisesti Googlen hakukone ja sitä kautta blogikirjoitukset. Vastaan tulleet tutkimukset valittiin luettavaksi niiden otsikon sekä tiivistelmän pohjalta, mikäli ne vaikuttivat vastaavan opinnäytetyön tutkimusongelmaa. Hakusanojen käyttö vaihteli, mutta hakusanat sisälsivät taulukossa 1 olevia hakutermejä.

TAULUKKO 1. Hakutermit.

Hakutermit				
SEM	SEO	Big data	AI	ML
Search engine	Automated data analysis	Implementation of automated data analysis	Artificial intelligence	Machine learning
Data analysis	Digital marketing	Digital marketing automation		

2 HAKUKONEET

Maailman suosituin hakukone on Google, jota ihmiset käyttävät pääasiallisesti. Googlen markkinaosuus Globalstats.com-verkkosivuston mukaan on 91.4 % kaikista hakukoneilla suoritettavista hauista. Seuraavaksi suurin markkinaosuus on Bing-nimisellä hakukoneella, jonka markkinaosuus kattaa 3.4 % hakukoneilla suoritetuista hauista. Kolmanneksi suurin markkinaosuus on Baidu-nimisellä hakukoneella, joka kattaa markkinaosuudesta 1.76 % (Statcounter 2022). Googlen suuren markkinaosuuden vuoksi yrityksen tulisi tehdä hakukonemarkkinoinnin toimenpiteet vastaamaan Googlen algoritmeja.

Uuden tuotteen tai palvelun ostoprosessi aloitetaan lähes aina hyödyntämällä hakukonetta. Hakukoneiden tehtävänä on tyydyttää tiedon tarve tai ratkaista hakijan ongelma. Tästä syystä hakukonenäkyvyys ja hakukoneissa korkealle sijoittuminen on erittäin tärkeää tuotetta tai palvelua tarjoavalle yritykselle. Näkyvyyden saaminen hakukoneissa ei kuitenkaan ole helppoa, sillä se on erittäin kilpailtu osa-alue digitaalisessa markkinoinnissa. (Blomster ym. 2020, 119-128.)

Useimmilla hakukonemarkkinointikampanjoilla on kaksi päätavoitetta: brändin kehittäminen sekä suora ostoärsyke. Sopivan markkinointikanavan valinta riippuu viime kädessä siitä, mitkä strategiat tarjoavat suurimman sijoitetun pääoman tuottoprosentin (ROAS, Return On Ad Spent). Yritykset, jotka tarjoavat tuotteita tai palveluita verkon kautta, hyötyvät selvästi hakukonemarkkinoinnista, koska niiden potentiaaliset asiakkaat ovat jo verkossa. Muut kuin verkkopohjaiset yritykset voivat valita hakukonemarkkinoinnin lisätäkseen näkyvyyttä ja mainostaakseen brändiä. (Jacob 2022.)

Hakukonemarkkinointi kuuluu yleensä osaksi markkinointi- tai brändäysstrategiaa. Yksi hakukonemarkkinoinnin (SEM, Search Engine Marketing) ensisijaisista tavoitteista on nostaa sivustolle tulevien verkkovierailijoiden määrää. Sivustolle siirtyvien vierailijoiden määrää voidaan

parantaa hakukonemainontaa lisäämällä tai hakukoneoptimoinnilla. Tulosprosentti on myös erityisen tärkeä hakukonemarkkinoinnille, koska se määrittää viime kädessä mainonnan kustannusten ja siten myös sijoitetun pääoman tuottoprosentin. (Wolf 2007, 32.)

Google Ads-hakusanamainonta on maksu-klikkauksesta-periaatteella (Pay-Per-Click, PPC) toimivaa maksettua mainontaa. Tämä tarkoittaa sitä, että mainoksen näkyminen Googlen hakukoneessa ei maksa mitään. Veloitus tapahtuu vasta, kun Googlen näyttämää mainosta klikataan ja verkkosivuvierailija siirtyy mainoksen asettajan www-sivulle. Hinta PPC mainonnassa määräytyy Google Ads-palvelussa olevan huutokaupan mukaan. Hakusanamainonta on yrityksen kannalta oleellinen mainonnan tapa. Se kuitenkin vaatii, että Google Ads-tiliä ja huutokaupan hinnastoa tarkkaillaan (Google 2022a). Hakusanamainonnassa kriittistä on oikeiden hakusanojen käyttö.

Hakukoneoptimointi on joukko teknisiä optimointistrategioita, joita hyödynnetään esimerkiksi verkkokaupoissa tai www-sivustoilla. Näiden optimointistrategioiden tarkoitus on parantaa näkyvyyttä hakukoneissa orgaanisesti. Hakukoneet käsittelevät miljoonia hakuja joka ikinen sekunti (Mousinho 2022). Lähes kaikki hakukoneoptimointistrategiat suunnataan vastaamaan Googlen hakukonealgoritmeja Googlen markkinaosuudesta johtuen.

Viime vuosien tutkimusten mukaan Googlen hakukone on paljon tehokkaampi ja onnistuneempi ymmärtämään verkkosivustoja kuin muut hakukoneet. Hakukoneoptimoinnin toimintalogiikan kannalta on kriittistä päästä alkuun sivuille semanttisilla avainsanoilla, jopa esimerkkiplogisivuilla. Googlen hakukoneiden luonnollisten kielten käsittely on jo todella edistynyt. Googlen algoritmi osaa korjata esimerkiksi hakijan väärin kirjoittaman sanan lauseen sisällä, jotta hakutulokset ovat sitä, mitä hakija hakee, huolimatta siitä, että alkuperäinen hakukäsky ei sen mukainen suoranaisesti ollut. (Tekin 2021.)

2.1 Hakukonemarkkinoinnin rakenne

Oppaassa käytettävä termi hakukonemarkkinointi toimii sateenvarjoterminä kaikelle toiminnalle, jonka tavoitteena on lisätä verkkovierailujen määrää www-sivustolla. Hakukonemarkkinointi koostuu kahdesta osa-alueesta; hakukoneoptimoinnista (Search Engine Optimization, SEO) sekä hakukonemainonnasta, joka sisältää hakusanamainonnan, display- eli näyttömainonnan ja maksetut listaukset. Hakukoneoptimointi ja hakukonemainonta muodostavat kokonaisuuden: hakukonemarkkinointi. (Blomster ym. 2020, 119-128.)

Hakukonemarkkinointia tulee lähestyä kuten muitakin markkinointikampanjoita. Deckerin (2019) mukaan markkinointi on tapa, jolla yritys lisää tietoisuutta brändistään ja vakuuttaa asiakkaat ostamaan, kun taas mainonta on prosessi, jolla luodaan vakuuttavia viestejä näiden laajojen tavoitteiden ympärille. Markkinointikampanjoiden tulee täten olla strategisia ja organisoituja toimenpiteitä kohti markkinointia tekevän yrityksen tavoitteita. Kuten perinteisessäkin markkinoinnissa, hakukonemarkkinoinnille täytyy asettaa strategia ja tavoitteet, jotta toimet ovat yhdenmukaisia ja tavoiteltava päämäärä selkeä (Decker 2019). Jotta hakukonemarkkinoinnin onnistumista voidaan mitata, tulee sille asettaa esimerkiksi numeraaliset tavoitteet koskien konversiota, verkkovierailujen määrää taikka impressiota ja tavoitteet tulisi saavuttaa ennalta päätetyssä aikamääreessä.

Hakukonemarkkinoinnin rakenne voi vaihdella tavoitteista riippuen. Esimerkiksi verkkovierailijoiden hankintataktiikoita on useita, joista hakukonemarkkinointia tekevä yritys valitsee yritystä parhaiten palvelevat taktiikat. Taktiikkana voi olla esimerkiksi, bannerimainokset, jotka ovat teksti tai www-sivuston mediapohjaisia mainoksia, joissa on yleensä linkki mainostajan www-sivustolle. Toinen yleisesti käytössä oleva taktiikka on hakukoneoptimointi, jolla pyritään vaikuttamaan hakukoneiden tulossivulla (SERP-sivu, Search Engine Result Page) esiintymiseen orgaanisesti, eli yritys sijoittuu korkealle hakutuloksissa. Kolmas hankintataktiikka on maksettu mainonta (Boughton 2005, 29-32).

Hakukonemarkkinointi ja siihen liittyvät toteutukset tulee tehdä harkiten ja yrityksen kohdeyleisöä ajatellen.

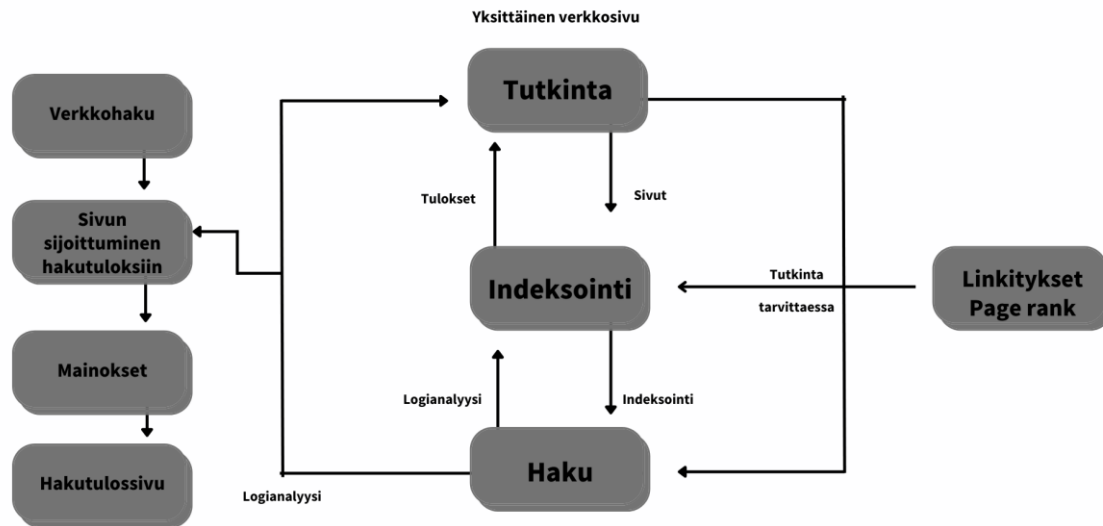
2.2 Hakukoneiden toiminnan ymmärtäminen

Hakukonemarkkinointia tehtäessä on tärkeää ymmärtää hakukoneiden toiminta sekä kuinka kilpailu ja muiden yritysten toimet vaikuttavat hakutuloksiin. Hakukoneissa kilpailu tapahtuu kilpailemalla hakutulossivun sijoituksista (Blomster ym. 2020, 119-128). Glick ym. (2014) tekemässä tutkimuksessa havaittiin, että SERP-sivun yläosassa esiintymisellä on suuri syy-seuraussuhde käyttäjän sivulle siirtymisen todennäköisyyteen. Pelkästään verkkosivuston siirtäminen sijoituksesta yksi (1) sijoitukselle kaksi (2) pienensi sivulle siirtymisten määrää 1/3 tai 2/3 riippuen hausta. Heidän tuloksensa viittaavat siihen, että algoritmissa korkealle sijoittumisella on todella suuri merkitys (Glick ym. 2014, 101-105). Hakukoneiden tulossivuilla ensimmäinen sivu on kaikista merkityksekkäin.

Googlen hakukone tarjoaa käyttäjälleen kohdennettuja hakutuloksia edellisen puolen vuoden hakuihin pohjautuen, tästä johtuen hakutulokset voivat erota suurestikin markkinointia tekevien henkilöiden sekä heidän kohdeasiakkaidensa välillä (Blomster ym. 2020, 119-128). Hakutulosten pohjautuminen edellisen puolen vuoden hakuihin aiheuttaa myös vaikeuksia tulosten validoinnin suhteen, mikäli tästä ei ole tietoinen. Yritys saattaa näyttäytyä omissa hauissaan korkeammalla sijoituksella, kuin he oikeasti ovat, joka antaa väärän kuvan yrityksen suoriutumisesta hakukoneissa.

Hakukoneet toimivat kolmen pääasiallisen funktion avulla. *Tietojen etsiminen indeksoimalla (Crawling)*: Selaa sisältöä internetistä ja tutkii jokaisen löytämänsä URL-osoitteen sisällön. *Tietojen järjestäminen indeksoimalla*: Tallentaa ja järjestää tietojen etsimisen aikana löydetyn sisällön. Kun sivu on indeksoitu, se on hakemistossa, jotta se näytetään parhaiten osuvien kyselyiden tuloksena. *Googlen Page Rank* (Pages rank=Sijoitus hakukonesivulla) tarjoaa hakukyselyä parhaiten vastaavan sisällön hakijan hakusivulle. Tämä tarkoittaa sitä, että

parhaimman sijoituksen saanut sivusto vastaa algoritmin tulkin mukaan hakijan tarpeita (Muller ym. 2021). Hakutulokset määräytyvät hakusanainosten sekä yksittäisten sivujen indeksoinnin perusteella. Toteuttaakseen tehokkaasti hakukoneoptimointia, tulee sitä tekevän yrityksen ymmärtää, mitä hakukone sivuilta indeksoi, ja vastaako se sitä sisältöä, jonka he haluavat hakijalleen tarjota. (Blomster ym. 2020, 119-128.)



KUVIO 1. Hakukoneiden toiminta. (Mukaillen Blomster ym. 2020, 119-128.)

Yllä olevassa kuviossa kuvataan, kuinka hakutulokset määräytyvät hakukoneissa. Hakukonesijoituksissa ylöspäin nousu vaatii aikaa, sillä se sisältää kohdeyleisön hakukäyttäytymisen ymmärtämistä, sivurakenteen käsittelyä ja testausta, sisällöntuotantoa sekä linkitysten rakentamista. Toimenpiteiden indeksoituminen vaatii aikaa, ja tuloksia pääsee usein tarkastelemaan noin kolmen-neljän (3-4) kuukauden päästä. Yllä mainittujen hakukoneoptimointiin liittyvien tekojen lisäksi hakukonemainonnan tehostamisella voidaan saavuttaa hakukoneissa näkyvyyttä. Hakukonemainonnassa on kuitenkin tärkeää, että yritys kykenee käyttämään kohdeasiakkaidensa käyttämiä hakusanoja, jolloin mainokset kohdistuvat oikeaan asiakassegmenttiin. Mikäli hakusanat eivät vastaa ostavien asiakkaiden hakutuloksia, voivat ne ohjata www-sivustolle epätoivottua liikennettä (Blomster ym. 2020, 119-128). Epätoivotulla liikenteellä tarkoitetaan tässä tapauksessa

sitä, että verkkovierailija, jolla ei ole potentiaalia konversioon, klikkaa PPC-mainosta ja täten aiheuttaa veloituksen yrityksen Google Ads-tilille.

Tietojen indeksointi on etsintäprosessi, jossa hakukoneet lähettävät robottiryhmän (tunnetaan nimellä indeksoijat tai hämähäkit) etsimään uutta tai päivitettyä sisältöä. Sisältö voi vaihdella – se voi olla verkkosivu, kuva, video tai PDF-tiedosto, mutta muodosta riippumatta sisältö löydetään aina linkkien avulla. Internetiä voi verrata jatkuvasti kasvavaan kirjastoon, jossa on useita miljardeja kirjoja ilman loogista lajittelujärjestelmää. Tästä syystä indeksointirobotit ovat käytössä; löytääkseen yleisesti käytössä olevat www-sivustot. Indeksointirobotit tutkivat www-sivuja ja seuraavat sivustolle sijoitettuja linkkejä ikään kuin verkkovierailija, joka olisi vierailemassa www-sivustolla. Indeksointirobotit siirtyvät linkistä toiseen ja koostavat raportin Googlen palvelimille. Kun indeksointirobotit ovat etsineet tiedot, on vuorossa tietojen järjestäminen. Googlen järjestelmä tekee huomioita www-sivulla olevista tärkeistä asioista, kuten: avainsanoista ja sivuston iästä ja lisää tiedot Caffeine-hakemistoonsa (Google 2022d), valtavaan löydettyjen URL-osoitteiden tietokantaan, josta se voidaan myöhemmin hakea ja ohjata hakuparametreja parhaiten vastaavalle verkkosivulle. (Muller ym. 2021.)

Indeksointi on siis prosessi, jolla hakukoneet järjestävät tiedot ennen hakua mahdollistaakseen erittäin nopeat vastaukset kyselyihin. Määritelmän mukaan, jos verkkosivusto ei ole hakemistossa, sitä ei voida eikä näytetä missään hakutuloksissa ja se menettää arvonsa. On tärkeää ymmärtää, kuinka Google ja muut hakukoneet indeksoivat www-sivustoja ja tehdä verkkotunnuksesta teknisesti saavutettavissa oleva, joka tarkoittaa, että indeksointirobotti voi tarkastella kaikkia sivuja ja niiden sisältöä vaikeuksitta ja, että kaikki verkkotunnuksen URL-osoitteet voidaan sisällyttää asiaankuuluvaan hakemistoon (XML.sitemap) (Marsden 2021). Halutessaan yritys tai www-sivuston ylläpitäjä voi myös piilottaa haluamansa alisivut indeksoinnilta. Indeksoinnin ymmärtäminen ja oikeat toimet auttavat sivuston näkyvyyttä hakukoneiden tulossivuilla.

Avainsanojen ja aiheiden etsiminen yksittäisiltä sivuilta olisi erittäin hidas prosessi indeksointiroboteille. Sen sijaan hakukoneet käyttävät käänteistä indeksiä. Käänteinen indeksi on hakemistotietorakenne, joka tallentaa kartoituksen sisällöstä, kuten sanoista tai numeroista, sen sijaintiin asiakirjadokumenttijoukossa. Yksinkertaisesti sanottuna se on hajautustaulun kaltainen tietorakenne, joka ohjaa indeksointirobotin sanasta asiakirjaan tai verkkosivulle. (Marsden 2021.)

Alla olevassa taulukossa 2 esitetään, mitä dataa Googlen indeksoijat indeksoivat sivustolta. (Google 2022e.)

TAULUKKO 1. *Indeksoitavat tiedot.*

Indeksoitavat tiedot	Sisältö
Tekstisisältö	Kaikki tekstimuotoinen sisältö
<title>-elementit	Määrittää asiakirjan otsikon, joka näkyy selaimen otsikkorivillä tai sivun välilehdessä.
Tärkeimmät sisältötunnisteet	Määrittää asiakirjan pääsisällön.
Kuvat	Kuvallinen sisältö
Videot	Videomuodossa oleva sisältö
Taulukot	Taulukkomuodossa oleva sisältö

2.3 Hakukonemainonta

Nykypäivänä hakukoneet ovat erittäin suuressa osassa asiakkaan ostoprosessia. Tuotteiden tai palveluiden ostoprosessi alkaa lähes aina hakukonetta hyödyntämällä. Hakukoneita käytetään tarpeen tyydytykseen tai ongelman ratkaisuun. Hakusanamainonta on yksi digitaalisen markkinoinnin

keskeisimmistä toimenpiteistä, sillä sen avulla yritys näkyy kohdennetusti hakukoneiden mainoksissa vastaten hakukonetta käyttävän henkilön hakua. (Blomster ym. 2020, 119-128.)

Hakukonemainonta ja maksetut listaukset ovat nopea tapa saada www-sivuston näkyvyys ja esiintyvyys hyvälle tasolle. Hakukoneoptimoinnilla tuloksista voidaan tehdä pysyviä ja pitkäaikaisia. Hakukonemainonta on kokonaisvaltainen yhdistelmä teknisiä- sekä maksettuja toimenpiteitä, joiden tarkoituksena on saavuttaa lisää verkkovierailuja. Hakukonemainonnan avulla yritykset voivat kohdistaa mainoksia kuluttajiin sijoittamalla mainoksia hakukoneisiin ja se on osoittautunut tehokkaaksi yleisön hankintastrategiaksi. Toisin kuin perinteisessä verkkomainonnassa, mainostajat maksavat vain, kun kuluttajaa klikkaa mainosta ja päätyy yrityksen asettamalle laskeutumissivulle. Onnistuneesti toteutettu hakukonemainonta voi tuottaa tasaisen liikenteen ja valtavan sijoitetun pääoman tuoton (Return On Investment, ROI). (Google 2022a.)

Näyttömainontaa (Display-mainonta) näkyy kolmansien osapuolien verkkosivustoilla ja käyttää video-, kuva- tai tekstielementtejä tuotteiden tai palveluiden mainostamiseen. Näyttömainontaa on monenlaisia, esimerkiksi bannerimainokset ovat näyttömainontaa. Useimmat mainokset ovat muodoltaan suorakaiteen tai neliön muotoisia, ja niiden sisältämä sisältö on yleensä suunniteltu vastaamaan isäntäsivuston sisältöä ja valittuja kohdeyleisöasetuksia. Näyttömainoskampanjoita voidaan ajaa mainosverkostojen, kuten Facebook-mainonnan tai Google-mainosten kautta, jotka tarjoavat tehokkaita yleisökohdistusominaisuuksia sekä mainosmuotoja ja ne voidaan yhdistää hakusanamainosten kanssa (Mailchimp 2022). Näyttömainontaa voidaan tehdä myös uudelleenmarkkinointikampanjoina, jolloin mainokset näkyvät vain henkilöille, jotka ovat jo vierailleet mainosta näyttävän yrityksen verkkosivuilla tai muulla sivustolla.

Hakusanamainonnassa yritykset tekevät hintatarjouksia haluamastaan avainsanasta ja se perustuu huutokauppaan. Yritykset maksavat mainoksen näytöstä tai klikkauksesta riippuen heidän sopimuksestaan. Tätä kutsutaan PPC-taikka CPC-mainonnaksi. Jotta hakusanamainonnan mainos näkyy korkeilla

sijoituksilla SERP:issä, tulee sivuston olla Googlen mielestä laadukas, avainsanojen vastata hakuparametrejä sekä Google Ads-tilillä olla positiivinen historia. (Google 2022b.)

Hakusanamainonta on digitaalisen markkinoinnin kilpailluimpia osa-alueita. Näkyvyyden saavuttamisen haasteellisuus hakusanamarkkinoinnin avulla on riippuvainen kilpailusta kyseistä hakusanaa kohtaan. Kilpailluimmilla hakusanoilla ylhäälle listautuminen on haasteellista ja hintavaa. Suomessa hakukonemarkkinoita hallitsee Google, joten yritykselle useimmiten riittää Google Ads-palvelun hallitseminen. Globaalisti hakusanamainonnassa ei Google Ads-palvelun hallinta välttämättä riitä, sillä esimerkiksi Kiinassa tai Etelä-koreassa Google ei toimi (Blomster ym. 2020, 119-128). Hakusanamainonnassa hakusanatutkimus ja asiakkaiden hakujen ymmärtäminen voi johtaa hyviin tuloksiin kohtuullisilla rahallisilla panostuksilla.

Hakukonemainonnan yleistyessä ja täten kilpailun kasvaessa tarjoushinnat nousevat nykytasoa korkeammalle. Pienemmät toimijat markkinoilla ovat huomanneet, että heillä ei ole enää varaa tehdä hintatarjouksia haetuimmista sanoista. Sen sijaan heidän on laajennettava kampanjaansa sisältämään useita hakukoneita ja kymmeniä tuhansia avainsanoja (Fleischner 2010). Mainostajat kohtaavat myös menetettyjen napsautusten ongelman (ero sen välillä, mistä he maksavat ja mitä heidän www-sivustonsa todellisuudessa vastaanottaa), jotka ovat usein seurausta napsautuspetoksista. Näistä esteistä huolimatta hakusanamarkkinointi voi tarjota paremman tuoton, kuin muut markkinointikanavat ja sen tulisi olla osa kaikkia kehittyneitä monikanavaisia markkinointikampanjoita. (Google 2022c.)

Google Ads-palvelussa on mahdollista käyttää useita eri mainonnan muotoja, kuten napautuskohtaiseen hintaan perustuvaa mainontaa, joka tunnetaan usein nimellä Cost Per Click, CPC. Napautuskohtaiseen hintaan perustuva hakukonemainonta on yleisin tapa saavuttaa asetettu konversio näyttämällä mainoksia verkkosivustolla tai hakukoneen tulossivulla. Mainostaja maksaa sovitetun maksun aina, kun joku napsauttaa mainosta sivustolla tai hakukoneen tulossivulla, ja mainosta näyttävän sivuston omistaja saa prosenttiosuuden tästä

maksusta. Mainokset, jotka sisältävät kontekstuaalista materiaalia mainospaikassa, eli ovat kohdistettuja mainoksia, toimivat tehokkaasti ja liittyvät läheisesti CPC-mainontaan. Kohdennettu CPC-mainonta on tehokkaampaa ja tuo sivuston omistajalle enemmän rahaa kuin ei-kohdistetut mainokset. CPC-mainonnan lähtökohtana on, että sivuston vierailijat napsauttavat mainoksia. Koska vierailijoiden kiinnostuksen kohteet ovat samankaltaisia, mainokset menestyvät parhaiten, kun ne näkyvät yhdistetyissä sivustoissa. (Davis 2006.)

Hakukonemainontaa voidaan käyttää myös mainospaineen luomiseen näyttökampanjoiden sivun näyttökertojen avulla. Pääosin mainoksia käyttävät verkkokaupat, koska hakukonemainonnan tarkoituksena on saada käyttäjä ostamaan. Yksi päätavoitteista on kasvattaa konversioita myynnin muodossa. (Pan 2015.)

Hakukonemainonta antaa yrityksille mahdollisuuden seurata sijoitetun pääoman tuotto prosenttia tarkasti yleisöhankinnan näkökulmasta. 1990-luvun lopun dotcom-buumin aikana yritykset käyttivät miljoonia dollareita mainoksiin, eikä tuloksista ollut selkeää käsitystä. Nykyisin yritykset voivat tehdä yksityiskohtaisia kustannus-hyötyanalyyskejä. Tämä mahdollistaa realistiset liiketoimintamallit. (Suskova 2021, 19.)

Vuonna 2004 kaksi tunnettua markkinointialan yritystä julkaisivat yhteistyössä työkalun, jonka tarkoituksena on mitata hakusanojen hintavaihtelua markkinoilla. Kyseisen työkalu tunnetaan nykyään nimellä avainsanojen hintaindeksiksi (Key Performance Indicator, KPI). KPI-mittari on tarkoitettu käytettäväksi perustana, kun vertaillaan hakukoneen suhteellista kustannustehokkuutta hakukonemarkkinointikampanjoissa. Se seuraa napautuskohtaisen hinnan vaihtelua. (Fadner 2004.)

KPI-mittarin avulla on voitu todeta napautuskohtaisten hintojen vaihtelevan suurestikin toimialojen välillä, joillakin toimialoilla keskiarvoakin korkeammat napautuskohtaiset hinnat voivat olla toimivia ratkaisuja, johtuen korkeasta ROI:sta. Esimerkiksi rahoitus- ja asuntoluottoalalla oli painotettu keskimääräinen napautuskohtainen hinta 3,17 dollaria, joka kuulostaa kohtuulliselta, kun otetaan

huomioon mahdolliset tulot asuntolainan tai sijoitustilin myynnistä. Vähittäiskaupan painotettu keskiarvo sen sijaan oli 0,32 dollaria, koska mahdolliset voitot tai konversiot eivät ole riittävän merkittäviä oikeuttaakseen korkeampaa tarjoushintaa. Tämä indeksi antaa markkinoinnin ammattilaisille työkalun ymmärtää paremmin tarjosten hintavaihtelua toimialojen välillä sekä mitata mahdollisimman tehokasta markkinointibudjetin käyttöä. (Fadner 2004.)

Yrityksellä voi olla jopa tuotekohtaisesti kannattavaa olla joustava napautuskohtaisen hinnan arvioinnissa. Napautuskohtaista hintaa ja sen kipurajaa pohdittaessa tulee kuitenkin ottaa huomioon mahdollisen konversion arvo.

2.4 Hakukoneoptimointi

Hakukoneoptimoinnin tarkoituksena on parantaa www-sivuston orgaanista näkyvyyttä hakukoneissa ja SERP-sivulla. Hakukoneoptimointia tekevän yrityksen sijoitus paranee hakukoneissa tiettyjä hakusanoja käyttämällä, silloin kun hakukoneoptimointi on tehty oikein. Hakukoneoptimointi on tärkeää löydettävyyden kannalta, ja oikein tehtynä se ohjaa oikeaa asiakaskuntaa yrityksen www-sivustolle. Hakukoneoptimointi koostuu: Hakukoneoptimoinnin nykytilan selvittämisestä, teknisestä hakukoneoptimoinnista, sivuston sisäisestä hakukoneoptimoinnista (On-page SEO), ja sivuston ulkopuolisesta hakukoneoptimoinnista (Off-page SEO). (ks kuvio 2 alla). (Blomster ym. 2020, 119-128.)



Kuvio 2. Hakukoneoptimoinnin osa-alueet. (Mukaillen Blomster ym. 2020, 119-128.)

Teknisen hakukoneoptimoinnin tarkoituksena on tehdä www-sivuston koodissa olevia tiedostoja hakukoneille helpommin ymmärrettäväksi. Www-sivustolla sisällöksi hakukoneet lukevat: tekstin, kuvat, videot, taulukot sekä muut elementit. Sivustolla olevien hakukoneoptimointitoimien elementtejä ovat esimerkiksi title-attribuutti, metakuvaus, XML-tiedostot ja otsikot (H1, H2, H3). (Halsey 2021.)

Hakukoneisiin indeksointia parantaakseen www-sivustolle tulisi rakentaa XML-tiedosto, joka toimii ikään kuin sivuston pohjapiirustuksena (XML.sitemap). Www-sivuston looginen rakenne ja se, että sivusto vastaa standardeja auttaa hakukoneita näkemään sivuston hyvin rakennettuna kokonaisuutena ja täten se vaikuttaa positiivisesti näkyvyyteen hakukonetulossivulla (Backlinko 2021). Sivuston nopeutta pidetään myös useasti yhtenä tärkeimmistä hakukoneoptimoinnin tekijöistä, tämä juontaa siitä, että Google pyrkii jatkuvasti kehittämään sivustojen käyttäjäkokemusta. Myös mobiiliystävällisyys vaikuttaa huomattavasti hakukoneoptimoinnin tuloksiin sillä suurin osa verkkovierailuista tapahtuu mobiililaitteilla (Blomster ym. 2020, 119-128). Sivuston hakukoneoptimointia tehdessä yrityksellä tulee pitää loppukäyttäjää ja käyttökokemus mielessä.

Sivustolla olevien hakukoneoptimointitoimien tarkoituksena on optimoida tieto ja tiedon löytäminen, niin hakukoneiden, kuin verkkovierailijoiden kohdalla (Halsey, 2021). Sisäinen hakukoneoptimointi tulee aloittaa siitä, että yrityksen URL-osoitteet ovat suunniteltu niin hyvin, että hakukone ja verkkovierailija ymmärtävät, mitä kohdesivu pitää sisällään. (Blomster ym. 2020, 119-128.)

Sivuston metaotsikko (Metatitle) kertoo hakukoneille, mitä sivustolla on tarjota. Otsikkoon tulee sisällyttää tärkeitä avainsanoja. Sivuston metakuvaus (Meta description) on tärkeää asiakkaalle, sillä se kertoo, mitä sivusto pitää sisällään, mutta se ei vaikuta sijoittumiseen hakukoneissa. Sivuston sisällön optimointi on myös tärkeää. Sisällöissä on tärkeää optimoida myös otsikot (Headings, H1, H2, H3...) vastaamaan sisältöä ja sisältämään haluttuja avainsanoja (Blomster ym. 2020, 119-128). Sivuston sisältö ja avainsanojen sisällyttäminen sivustolle on tärkeää, jotta hakukoneet tarjoavat sivustoa sen sisällöstä kiinnostuneille verkkovierailijoille.

Otsikkojen avulla Google ymmärtää sivuston tärkeät osa-alueet ja niitä tehdessä kannattaakin muotoilla ne samaan muotoon, mitä hakukonetta käyttävä henkilö voisi käyttää. Googlen hakukoneet arvostavat ainutlaatuista ja ajankohtaista sisältöä, tästä syystä avainsanojen ja – lauseiden käyttö ja niiden vastaaminen hakusanoja on tärkeää. (Blomster ym. 2020, 119-128.)

Kuvat ovat myös tärkeä osa sisäistä hakukoneoptimointia. Kuvissa olevat alt-tekstit vaikuttavat hakukoneiden havaitsemiseen koskien sivustoa. Myös Googlen kuvahaun kautta on mahdollista saada verkkovierailijoita, joten kuvat tulisi optimoida nopeasti avautuviksi ja alt-tekstit vastamaan kuvaa. Sisäiset linkit ovat tärkeitä niin indeksointirobotin, kuin verkkovierailijan kannalta. Tämä helpottaa verkkovierailijoita liikkumaan sivuston sisällä sivulta toiselle (Blomster ym. 2020, 119-128). Sivuston ulkopuolisilla hakukoneoptimointitoimilla pyritään nostamaan www-sivuston luotettavuuspisteitä Googlen silmissä. Tämä onnistuu kerryttämällä www-sivustolle tulevia linkkejä jo valmiiksi Googlen turvalliseksi määrittelemiltä sivustoilta. Sivuston ulkopuolista hakukoneoptimointia on myös sosiaalisen median markkinointi, vieraskirjoitukset blogeissa, linkatut ja linkittömät maininnat brändistä sekä vaikuttajamarkkinointi. (Muller ym. 2021.)

3 DATA JA DATAN ANALYSOINTI DIGITAALISESSA MARKKINOINNISSA

Digitaalisesti hankitun tiedon lisääntynyt saatavuus ja edistynyt teknologia sen analysoimiseksi ajaa monia toimialoja digitaaliseen muutokseen. Monet globaalit yritykset puhuvat suurten datajoukkojen potentiaalista ja he uskovat, että suurten datajoukkojen analysointi voi auttaa yrityksiä saamaan kilpailukykyisiä näkemyksiä sekä muokkaamaan organisaatioiden markkinointistrategisiapäätöksiä. Suurten datajoukkojen vaikutus vaihtelee toimialakohtaisesti. Toimialat, jotka tarjoavat enemmän virtuaalisia, kuin fyysisiä tuotteita ovat herkempiä digitaalisille muutoksille. Vuorovaikutus digitaalisen teknologian sekä organisaatioiden välillä on kuitenkin monimutkainen ja sen integroinnin edessä on monia esteitä. Teknologian tuomat muutokset haastavat niin tutkijoita, kuin alan toimijoita. (Grishikashvili ym. 2014, 1-12.)

Korkeakoulujen sekä yritysten välistä yhteistyötä on tarpeellista korostaa ja lisätä, jotta globaalit liiketoiminnat ja digitaalisen markkinoinnin tekijät voivat hyötyä enemmän akateemisista tutkimuksista. On olemassa akateemisia käyttöteorioita, joita ei kuitenkaan voida täysin soveltaa käytännön toimiin. Alustavat tulokset osoittavat, että ne toimialat, jotka yrittävät kaventaa kuilua käyttöteorioiden ja käytännön välillä, sekä ottaa käyttöön tarvittavat mekanismit suurten datajoukkojen hyödyntämiseksi markkinoinnissa, ovat muita toimialoja valmiimpia markkinoilla tapahtuviin muutoksiin. (Grishikashvili ym. 2014, 1-12.)

2000-luvun alusta saakka digitaalisen markkinoinnin käyttö ja kannattavuus on kehittynyt huomattavasti. Tämä on synnyttänyt digitaalisen ekosysteemin, joka yhdistää käyttäjänsä siihen jatkuvasti, ja se myös muokkaa käyttäjien tapoja sekä käyttäytymistä. Lukuisat tutkimukset ovat osoittaneet, että keskeinen tapa lisätä digitaalisen markkinointistrategian tehokkuutta on datatieteiden ja datan hyödyntäminen. Saura (2021) mainitsee tutkimuksessaan Kelleher ja Tierney (2018) työn, jossa he nostavat esille käytännön toimia, joiden avulla digitaalista markkinointia ja sen tehokkuutta voidaan parantaa. Se onnistuu yritysten käyttäjiltä kerätyn tiedon hallinnalla, tiedon tyyppin ja tietolähteen hyödyntämisellä

yrittäjien tietojoukoissa sekä uusien tietoanalyysien innovatiivisella soveltamisella. (Kelleher & Tierney 2018, 18-32.)

3.1 Hakukonemarkkinoinnin kannalta oleelliset datatyypit

Jotta hakukonemarkkinoinnissa voidaan hyödyntää datatyyppejä optimaalisella tavalla, tulee ymmärtää datatyypin erot sekä niiden käyttökohteet. Erityisesti siirtymävaiheessa, kun yritys on siirtymässä kohti dataohjattua hakukonemarkkinointia, on tärkeää ymmärtää, millaiset tiedot ovat tärkeitä hakukonemarkkinointia ajatellen. Tärkeää on myös ymmärtää, kuinka halutut datatyypit voidaan hankkia ja kuinka ne tallennetaan. Prosessin aloittaminen laadukkaasti ja tarpeita pohtien toimii kivijalkana tulevaisuuden tuloksille. Luotettavan ja vankan datainfrastruktuurin luominen helpottaa tulevaisuutta.

Datatieteissä tietokannat koostuvat erilaisista muuttujista tai indikaattoreista. Näitä tietokantoja kutsutaan "tietojoukoiksi". Jokainen tietojoukkojen sisältämä muuttuja ilmaisee tietyn ominaisuuden. Tietojoukot voivat sisältää jäseneltyä eli strukturoitua dataa tai jäsentämätöntä eli strukturoimatonta dataa. Strukturoitua dataa voidaan tallentaa taulukoihin, joissa jokaisessa taulukossa on sama rakenne tai attribuutti. Sen sijaan jäsentämättömällä datalla on oma sisäinen rakenne ja siten attribuutit voivat järjestyä eri tavoin kussakin taulukossa. Attribuutit tulevat joskus raakatiedoista. Raakadata sisältää erityyppisiä tietoja, jotka voivat tarjota oivalluksia ongelman ratkaisemiseen. (Saura 2021.)

Strukturoitu data on tietoa, joka on järjestetty laadukkaasti ja se on datan standardisoitu muoto, se on yksinkertaista ja selkeää dataa. Strukturoidun datan avulla sivusto tarjoaa tietoja sivusta sekä luokittelee sivuston sisällön. Esimerkiksi tuotetiedot kirjoitetussa muodossaan ovat strukturoitua dataa, koska niiden avulla tuotteen koko, valmistaja, hinta sekä muut tarpeelliset tiedot saadaan luokiteltua. (Google 2022d.)

Puolistrukturoitu data on strukturoidun datan muoto, joka ei noudata tietotaulukoihin liittyvien tietomallien taulukkorakennetta, mutta sisältää silti

tunnisteita. Tunnisteiden avulla tietueet ja kenttien hierarkiat voidaan tunnistaa. Puolistrukturoitu data tunnetaan itsekuvaavana rakenteena. Puolistrukturoitua dataan saadaan esimerkiksi kuvista, sähköposteista, XML sekä muista ohjelmointikielistä, ZIP-tiedostoista ja verkkosivuilta (Wikipedia 2022). Algoritmi, joka poimii dataa, ei välttämättä ymmärrä mitä esimerkiksi kuva sisältää ja mitä se esittää, mutta algoritmi kykenee ymmärtämään kuvalle asetetun määreen taikka muulle datajoukolle asetetun tunnisteiden. Hakukoneoptimoinnin kannalta kuvat ja kuvien määreet voivat olla hyvinkin merkityksellisiä.

Strukturoimaton data on tietoa, jolla joko ei ole ennalta määritettyä tietomallia tai jota ei ole järjestetty ennalta määrättyllä tavalla. Strukturoimaton data koostuu kaikista tiedoista, jotka on tallennettu jäsentämättömässä muodossa. Strukturoimattomalla datalla ei ole käsitteellistä määritelmää eikä tietotyypin määritelmää, toisin, kuin esimerkiksi tekstidokumenteissa sana on yksinkertaisesti sana. Jotkin nykyiset tekniikat, joita käytetään sisällönhakuun strukturoimattomista tiedoista, edellyttävät entiteettien, kuten nimien merkitsemistä tai avainsanoja ja sisällönkuvauskenttiä. Siksi tarvitaan ihmisen validitaatiota, jotta jäsentämätön tieto saadaan koneellisesti luettavaksi. Strukturoimaton data koostuu kahdesta perusluokasta: Bittikarttaobjektit, jotka ovat luonnostaan ei-kieli-pohjaisia, kuten video-, kuva- tai äänitiedostoja sekä tekstistä, joka on esimerkiksi Microsoft Word- tai Microsoft Excel-tiedosto. Molemmat objektityypit voidaan luokitella dataksi, mutta tekniikka ja menetelmät niiden hyödyntämiseen ovat vielä alkeellisia. (Weglarz & Geoffrey 2004.)

Tietojoukon analysoinnissa datatieteet luottavat koneoppimiseen perustuviin malleihin. Nykyaikaisen datatieteen ydin, koneoppiminen, tarjoaa algoritmeja suurten tietojoukkojen automaattiseen analysointiin. Tutkijat tai markkinoijat voivat kouluttaa näitä malleja saadakseen hyödyllisiä oivalluksia tai tunnistamaan esimerkiksi käyttäytymismalleja. Saatavilla on laaja valikoima algoritmeja. Näitä algoritmeja voidaan käyttää ja kouluttaa muodostamalla yhteys yritysten tai tutkijoiden tietokantoihin. (Saura 2021.)

3.2 Suuret datajoukot, Big Data

Suurten datajoukkojen aikakaudella hakukoneoptimointi käsittää tietojoukkojen kapseloimista, jotka liittyvät www-sivuston suorituskykyyn, arkkitehtuuriin, sisällön kuratointiin sekä käyttäjien käyttäytymiseen www-sivustolla. Tarkoituksena on muuntaa ne käyttökelpoisiksi analyyseiksi ja näiden analyysien avulla parantaa näkyvyyttä sekä löydettävyyttä verkossa. Tässä suhteessa suurten datajoukkojen analyysi laajentaa mahdollisuuksia kehittää uusia metodologisia puitteita, jotka koostuvat pätevistä, luotettavasta ja johdonmukaisesta analytiikasta, jotka ovat käytännössä hyödyllisiä kehitettäessä tietoisia strategioita organisaation liikenteen optimointiin.

Suurten datajoukkojen analysointi ja integrointi ei automaattisesti johda parempiin markkinointitoimiin, mutta potentiaalia on. Suuria datajoukkoja voidaan käyttää raaka-aineena, välttämättöminä elementteinä askeleessa kohti kohdennetumpaa markkinointia. Suurten datajoukkojen tiedot eivät itsessään ole niin suuressa osassa. Pikemminkin se on suurten datajoukkojen avulla saadut oivallukset, päätökset ja toimet, jotka tekevät muutoksen. Suuret datajoukot tarjoavat kattavamman ymmärryksen markkinoinnin tarpeesta.

Koska suuret datajoukot tarjoavat uusia ominaisuuksia, sen tiedon laatu kohtaa myös monia haasteita. Suurten datajoukkojen ominaisuudet ovat Cai & Zhu (2015) mukaan 4V:tä: Laajuus (Volume), Nopeus (Velocity), Monimuotoisuus (Variety) ja Arvo (Value) kuten Katal, Wazid ja Goudar (2013) julkaistussa tutkimuksessaan toteavat. *Laajuus* viittaa valtavaan datamäärään. Yleensä datamäärät ovat Terabittejä (TB) tai suurempia tämän datamäärän mittaamiseen. *Nopeus* tarkoittaa, että dataa muodostuu ennennäkemättömällä nopeudella ja se on käsiteltävä ajoissa. (Cai & Zhu 2015.)

Monimuotoisuus osoittaa, että suurissa datajoukoissa on kaikenlaisia tietotyyppisiä, ja tämä monimuotoisuus jakaa datan strukturoituun dataan ja strukturoitumattomaan dataan. Nämä monityyppiset tiedot tarvitsevat korkeampia tietojenkäsittelyominaisuuksia. *Arvo* edustaa pieniarvoista tiheyttä. Arvotiheys on kääntäen verrannollinen koko datan kokoon, mitä suurempi

suurten datajoukkojen skaala on, sitä vähemmän arvokasta dataa on. (Cai & Zhu 2015.)

Gandomi ja Haider (2015) tutkimuksessaan toivat aiemman 4V:n lisäksi kolme datajoukkojen ominaisuutta. Gandomin ja Haiderin mukaan tarvittavat lisät datajoukkojen ominaisuuksiin ovat: todenmukaisuus, vaihtuvuus, monimutkaisuus sekä arvo. *Todenmukaisuus (Veracity)* edustaa joidenkin tietolähteiden luontaista epäluotettavuutta. Esimerkiksi asiakkaiden tunteet sosiaalisessa mediassa ovat luonteeltaan epävarmoja, koska ne sisältävät ihmisen harkintaa, tästä huolimatta ne silti sisältävät arvokasta tietoa. Näin ollen tarve käsitellä epätarkkoja ja epävarmoja tietoja on toinen ison datan puoli, jota käsitellään epävarman datan hallintaan ja louhintaan kehitetyillä työkaluilla ja analytiikan avulla. (Gandomi & Haider 2015, 1-8.)

Gandomi ja Haider (2015) mukaan viides ja kuudes V ovat: *Vaihtuvuus (Variability)* ja *monimutkaisuus (complexity)*. *Vaihtelu* viittaa tiedon virtausnopeuksien vaihteluun. Usein suurten datajoukkojen nopeus ei ole johdonmukainen, ja sillä on ajoittain huippuja ja ajoittain nopeus on huomattavasti matalampaa. *Monimutkaisuus* viittaa siihen, että suuria datajoukkoja hankitaan lukemattomista lähteistä. Tämä asettaa kriittisen haasteen: tarpeen yhdistää, sovittaa, puhdistaa ja muuntaa eri lähteistä saatua dataa. *Arvon (Value)* voi nähdä suuret datajoukot määrittävänä attribuuttina. Suurille datajoukoille on usein ominaista suhteellisen matala arvotiheys. Tämä tarkoittaa sitä, että alkuperäisessä muodossaan olevalla tiedolla on yleensä pieni arvo suhteessa datan määrään. Analysoimalla tätä dataa voidaan sille kuitenkin saada suurikin arvo. (Gandomi & Haider 2015, 1-8.)

Koska suurilla datajoukoilla on V-ominaisuudet, yritysten käyttäessä ja prosessoidessa suuria datajoukkoja, korkealaatuisen ja todellisen datan poimiminen massiivisista, vaihtelevista ja monimutkaisista tietojoukoista on haastavaa. Tällä hetkellä suurten datajoukkojen laatua kohtaavat seuraavat haasteet: *Tietolähteiden monimuotoisuus* tuo mukanaan runsaasti tietotyyppejä ja monimutkaisia tietorakenteita ja se vaikeuttaa tietojen integrointia. *Tietojen määrä* on valtava, ja tietojen laatua on vaikea arvioida kohtuullisessa ajassa.

Data muuttuu erittäin nopeasti ja tiedon "ajantasaisuus" on hyvin lyhyt, mikä edellyttää korkeampaa vaatimustasoa datan käsittelyteknologialle ja ajalle, jossa se käsitellään. (Cai & Zhu 2015.)

Koska suuret datajoukot ovat uusi käsite, eikä tiedemaailma ole tehnyt yhtenäistä määritelmää datan laadusta ja laatukriteereistä, aiheuttaa se käsitteenä vaikeuksia. Kirjallisuudessa on eroja tiedon laadun määritelmästä, mutta yhdestä asiasta on konsensus: tiedon laatu ei riipu pelkästään sen ominaisuuksista, vaan myös dataa käyttävästä liiketoimintaympäristöstä, mukaan lukien liiketoimintaprosessit ja yrityskäyttäjät. Vain ne tiedot, jotka vastaavat käyttötarkoituksia ja täyttävät vaatimukset, voidaan katsoa kelpuutetuiksi tai laadukkaiksi. Yleensä tiedon laatustandardeja kehitetään tiedontuottajien näkökulmasta. Aiemmin tiedonkuluttajat olivat joko suoria tai epäsuoria tiedontuottajia, mikä varmisti tiedon laadun. Suurten datajoukkojen aikakaudella datalähteiden moninaisuuden vuoksi datan käyttäjät eivät kuitenkaan välttämättä ole tiedon tuottajia. Siksi tiedon laatua on erittäin vaikea mitata. (Cai & Zhu 2015.)

Cai & Zhu (2015) ehdottavat tutkimuksessaan seuraavaa mallia tiedon laadun arvioimiseksi: *Saavutettavuus* tarkoittaa vaikeustasoa, jolla käyttäjät voivat saada tietoja. Saavutettavuus liittyy läheisesti tiedon avoimuuteen, mitä korkeampi tiedon avoimuusaste on, sitä enemmän tietotyyppisiä saadaan ja sitä korkeampi saavutettavuusaste on. *Ajantasaisuus* määritellään aikaviiveeksi tiedon tuottamisesta ja hankinnasta käyttöön. Tietojen pitäisi olla saatavilla tämän viiveen kuluessa hyvän analyysin mahdollistamiseksi. Suurten datajoukkojen aikakaudella datan sisältö muuttuu nopeasti, joten ajantasaisuus on erittäin tärkeää. *Valtuutus* viittaa siihen, onko henkilöllä tai organisaatiolla oikeus käyttää tietoja. (Cai & Zhu 2015.)

Uskottavuutta käytetään ei-numeerisen tiedon arvioimiseen. Se viittaa lähteen tai viestin uskottavuuden objektiivisiin ja subjektiivisiin komponentteihin. Tiedon uskottavuudella on kolme keskeistä tekijää: tietolähteiden luotettavuus, tiedon normalisointi ja aika, jolloin data tuotetaan. Asiakirja koostuu tietomääryyksistä, jotka sisältävät tiedon nimen, määritelmän, kelvollisten arvojen alueet, standardimuodot, liiketoimintasäännöt. *Metadata*, tiedontuottajien on toimitettava

metadataa, joka kuvaa tietojoukkojen eri puolia vähentääkseen väärinkäsitysten tai epäjohdonmukaisuuksien aiheuttamia ongelmia. *Data-arvon tarkkuuden* varmistamiseksi sitä verrataan tunnettuun viitearvoon. Joissakin tilanteissa tarkkuus voidaan mitata helposti, kuten sukupuoli, jolla on vain kaksi varmaa arvoa: mies ja nainen. Mutta muissa tapauksissa ei ole tunnettua viitearvoa, mikä vaikeuttaa tarkkuuden mittaamista. Koska tarkkuus korreloi jossain määrin kontekstin kanssa, tietojen tarkkuus tulisi päättää sovellustilanteen mukaan. (Cai & Zhu 2015.)

Tietojen johdonmukaisuus viittaa siihen, onko korreloitujen tietojen välinen looginen suhde oikea ja täydellinen. *Eheällä datalla* on täydellinen rakenne. Tietoarvot standardoidaan tietomallin ja/tai tietotyypin mukaan. Tietojen kaikkien ominaisuuksien on oltava oikein, jotta dataa voidaan kutsua eheäksi. *Kompetenssi* tarkoittaa, että yhden peruspisteen kaikkien komponenttien arvot ovat voimassa. Esimerkiksi kuvan värien osalta RGB:tä voidaan käyttää kuvaamaan punaista, vihreää ja sinistä, ja RGB edustaa kaikkia väritietojen osia. Jos tietyn komponentin väriarvo puuttuu, kuva ei voi näyttää todellista väriä ja sen täydellisyys tuhoutuu. (Cai & Zhu 2015.)

Tarkastettavuus tarkoittaa, että tarkastajat voivat arvioida tietojen tarkkuutta ja eheyttä rationaalisissa aika- ja työvoimarajoissa tiedon käyttövaiheen aikana. *Sopivuudella* on kaksitasoiset vaatimukset: käyttäjien käyttämän datan määrä sekä se, missä määrin tuotettu data vastaa käyttäjien tarpeita indikaattoreiden-, elementtien, luokituksen määrittelyssä. Luettavuus määrittellään tietosisällön kyvyksi selittää oikein tunnettujen tai hyvin määriteltyjen termien, attribuuttien, yksiköiden, koodien, lyhenteiden tai muun tiedon mukaisesti. *Rakenne*, yli 80 % kaikesta tiedosta on strukturoimatonta, joten rakenne viittaa vaikeusasteeseen puolistrukturoidun tai strukturoimattoman datan muuntamisessa strukturoiduksi dataksi teknologian avulla. (Cai & Zhu 2015.)

Kun suurten datajoukkojen analysointi integroidaan osaksi hakukonemarkkinoinnin strategiaa, pystyy tätä tekevä yritys saavuttamaan merkittäviäkin muutoksia kolmella osa-alueella; Asiakkaan sitouttaminen, Asiakkuuden säilyttäminen ja lojaalisuus sekä markkinoinnin optimointi tai

suorituskyky. *Asiakkaan sitouttaminen*: Suurten datajoukkojen avulla yritys saa tietoonsa asiakkaansa, sen missä he ovat ja mitä viestintäkanavaa ja viestintätapaa he suosivat ja milloin. *Asiakkuuden säilyttäminen ja lojaalisuus*: Suurten datajoukkojen avulla voidaan selvittää, mikä tekee asiakkaan tyytyväiseksi palveluun ja mikä saa heidät palaamaan tuotteen tai palvelun ääreen. *Markkinoinnin optimointi tai suorituskyky*: Suuria datajoukkoja hyödyntäen markkinointikanavat ja markkinoinninkulut pystytään määrittämään ja niitä pystytään seuraamaan todella tarkasti. (Rani & Rani 2017.)

3.3 Datan analysointi hakukonemarkkinoinnissa

Suurten datajoukkojen aikakaudella hakukoneoptimoinnilla on ratkaiseva rooli laatua heijastavan personoidun sisällön levittämisessä. Tämä laatu liittyy sisällön kuratointiin ja oikeaan käytettävyyteen web-pohjaisissa järjestelmissä käyttäjien tietotarpeiden kattamiseksi. Hakukoneoptimoinnin päätarkoitus liittyy strategisten vaiheiden tarjoamiseen web-kehittäjille ja sisällöntuottajille, jotta he voivat optimoida verkkosivustoja korkeammalle sijoitukselle hakutuloksissa ja siten lisäämään hakukoneilta tulevaa orgaanista liikennettä. Huolimatta verkkosivuston suorituskyvystä tai käyttäjien käyttäytymiseen liittyvien tietojoukkojen ja analyysien suuresta määrästä, tutkimuskenttä suurten datajoukkojen analytiikan hyödyntämiseksi strategisissa SEO-skeemoissa on vielä alkutekijöissään. Tähän on useita syitä, kuten tietojen ylikuormitus ja täten vaikeudet tulosten esikäsittelyssä, analysoinnissa, visualisoinnissa sekä tulkinnessa. (Drivas ym. 2020.)

Riippumatta organisaation tyypistä, mitä suurempi se on, sitä paremmat mahdollisuudet yrityksellä on hyödyntää suurten datajoukkojen analytiikkastrategioita mahdollisen verkkonäkyvyyden arvioimiseksi ja parantamiseksi. Suurten datajoukkojen analytiikka tuo kuitenkin mukanaan huomioon otettavia asioita. On huomioitava, että kooltaan suuren verkkosivuston olemassaolo ja siten suuri datatuotanto ei välttämättä tarkoita ehdotusten ja ratkaisujen lisääntymistä strategiseen SEO-suunnitteluun. Mitä suurempi sisällön määrä on, sitä vaikeampaa sen manipulointikin on. Myös teknisten ja

käyttäytymiseen liittyvien tekijöiden tunnistaminen vaikeutuu. Teknisten ja käyttäytymiseen liittyvien tekijöiden tunnistaminen ja optimointi vaikuttavat verkkosivuston suorituskykyyn sen näkyvyyden ja löydettävyyden kannalta hakukoneiden tulossivuilla. Käytännössä on hallittavampaa optimoida jopa 100 alasivua sisältävä verkkosivusto sen sijaan, että se sisältäisi tuhansia alasivuja, kuvia ja siten monimutkaisia linkkien arkkitehtuuria. (Drivas ym. 2020.)

Hakukoneoptimointia ja -markkinointia ajatellen olennaisimpia datamuotoja ovat: Teksti-, numeerinen- ja visuaalinen-data. Näiden datamuotojen analysoinnin avulla voidaan saavuttaa kokonaisvaltaisempi ymmärrys käyttökohteista ja -tarkoituksista. Tekstidatan tyyppejä ovat esimerkiksi: Avainsanat, sanalliset kuvaukset, asiakastiedot ja asiakkaiden vuorovaikutukset. Näitä yritys voi kerätä itse verkkolomakkeiden, dokumenttien, myyntitietojen ja verkkosivujen avulla. (Blomster & Koivumäki 2021, 18-24.)

Numeerista dataa saadaan verkkoanalytiikan sekä myynti- ja markkinointikanavien avulla. Numeerista dataa voidaan hyödyntää analytiikassa ja datan esittämisessä liiketoiminnan kehittämistä koskien. Visuaalista dataa ovat kuvat, videot ja grafiikat. Visuaalista dataa voidaan hyödyntää myynnin ja markkinoinnin tukena, esimerkiksi sisältömarkkinoinnissa, markkinoinnin automaatiassa ja hakukoneoptimoinnissa. (Blomster & Koivumäki 2021, 18-24.)

4 KONEOPPIMINEN JA TEKOÄLY

Tekoälyn tarkoituksena on kehittää ja hyödyntää tietokonejärjestelmiä siten, että tietokonejärjestelmät voivat simuloida ihmisälyä annettujen tehtävien suorittamiseen. Koneoppiminen on osa tekoälyä. Koneoppimisen tehtävä on tärkeä data-analytiikan kannalta ja se sisältää algoritmeja, jotka voivat oppia itsenäisesti. Koneoppimista hyödyntämällä tietojärjestelmä voidaan opettaa ottamaan vastaan tietoa ja analysoimaan niitä tulosten ennustamista varten, ilman ennalta määriteltyä johtopäätöstä.

Tekoälystä on tullut osa useiden toimialojen päivittäistä arkea. Tekoäly on ollut valokeilassa jo jonkin aikaa. Se on kuitenkin monimutkainen ja laaja aihe. Tekoälyn määritelmä voi olla luonteensa vuoksi haastava tehtävä, mutta McCarthy (1998) luonnehtii sitä ihmisen tekemäksi tietojenkäsittelytieteen ohjelmistoksi ihmisten työn jäljittelymiseen. Tekoälyn alaisuuteen sisältyy useita alalajeja kuten: Koneoppiminen, Neuroverkot, Syvä oppiminen, Kognitiivinen tietojenkäsittely, Konenäkö sekä Luonnollisen kielen käsittely (NLP, Natural language processing). (SAS 2019.)

Viime aikoina tekoäly on noussut erittäin suosituksi aiheeksi johtamis- ja markkinointitieteiden alueella, vaikkakin paradoksaalisesti sen kehittämistyöt muilla tieteenaloilla ovat jatkuneet yli puoli vuosisataa. Tekoäly on vuosien varrella noussut valokeilaan ja kadonnut valokeilasta riippuen sen mahdollisen sovellettavuuden lisääntymisestä. Laajan keskustelun myötä kiinnostus tekoälyä kohtaan on toiminut omanaan vauhdittamaan ensimmäisten laajamittaisten kaupallisten sovellusten markkinoille tuontia, jotka ovat osoittaneet tämän teknologian potentiaalin ja kyvyt myös markkinoinnin alalla. Tekoälyn nopea kehitys viime vuosina on ollut mahdollista tekoälyn kognitiivisten mekanismien ja koneiden oppimiskyvyn kehittymisen ansiosta saatujen tietojen perusteella (Lieto ym. 2017). Tekoälyn voima piilee erilaisten datamuotojen käsittelyssä, numeerisen datan lisäksi tekoäly käsittelee tekstejä, kuvia, ja ääniä, tarjoten niille merkityksen ja relevanssia lisäanalyysien kannalta (Dhar 2016). Tekoäly

mahdollistaa yrityksille resurssien vapauttamisen vaikeampiin työtehtäviin, mutta myös parantaa analyysien laatua ja objektiivisuutta. Tekoäly pystyy analysoimaan tulokset ja datan objektiivisesti, tämä on tärkeää, kun tuloksista halutaan todenmukaisia.

Markkinoinnin ammattilaiset hyödyntävät asiakasnäkemyksiä sekä kilpailutietoja tutkiessa dataa suurissa määrin, esimerkiksi kuluttajien tarpeita, markkina-analyysejä, eri viestintä- tai jakelukanavien toimintojen harjoittamiseen tai jo käytössä olevien strategioiden tulosten sekä vaikutusten mittaamisessa. Markkinointi hyötyy suuresti tietotekniikan kehittymisestä. Tärkeää on tulevaisuudessa korostaa tekoälyn ja saatavilla olevien tekoälypohjaisten työkalujen potentiaalia ja keskustella tekoälyn kaupallisista sovelluksista markkinoinnin alalla.

Hakukoneoptimoinnissa on useita erilaisia tekijöitä, jotka tulee ottaa huomioon ja avainsanat ovat yksi näistä tekijöistä. Avainsanat voivat olla avainlauseita tai avainsanoja, joita hakukoneen käyttäjä kirjoittaa löytääkseen etsimänsä tiedon hakukoneista. Avainsanat ovat olleet hakukoneoptimoinnin kulmakivi jo pitkään ja markkinoijat ovat jatkuvasti yrittäneet löytää tarkkoja avainsanoja, jotka sopivat yrityksen mainostamiseen, tämä auttaa myös yrityksen verkkosivustoa saamaan paremman aseman hakutuloksissa verkossa, mikä antaa yritykselle enemmän näkyvyyttä. (Dumitriu & Popescu 2020.)

Tekoäly muuttaa tapaa, jolla avainsanatutkimus tehdään. Tekoäly tulee yleistymään kaikkialla, joten myös hakukoneoptimoinnin asiantuntijoiden tulee ymmärtää automatisoidusta oppimisesta ja automaatiosta. Digitaalisen markkinoinnin maailmassa hakukoneoptimointi erottuu olennaisena osana yrityksen strategiaa (Dumitriu & Popescu 2020). Hakukoneoptimoinnin ammattilaisille tekoäly ja sen tuomat työkalut ovat eduksi työn laadun ja ajankäytön kannalta. Tekoälyn avulla pystytään tekemään esimerkiksi kilpailijan verkkosivujen avainsana-analyysi muutamassa minuutissa, ihmisen tuntiin, tai jopa pidempään aikaan verrattuna. Tämä vapauttaa resursseja vaativampiin prosesseihin ja toimenpiteisiin.

Luonnollisesti avaintekijänä on houkutella tuotteen tai palvelun ostajia. Tekoälyn pitäisi pystyä oppimaan ja parantamaan itseään pysyvästi. Koneoppimisen kulmakivi on tarjota ohjelmistoille kyky oppia. Tämä tavoite voidaan saavuttaa käyttämällä algoritmeja, jotka tunnistavat kuvioita vastaanotetuista tiedoista ja niistä riippuen ohjelmistojärjestelmä tekee päätöksiä tai ennusteita. Älykkäiden algoritmien käyttöönoton jälkeen markkinoinnin alalle on kasvanut alatoimiala nimeltä automatisoitu markkinointi. Automatisoitu markkinointi liittyy uusimpiin teknologiatrendeihin. (Dumitriu & Popescu 2020.)

Automatisoitua markkinointia on monenlaista. Automatisoitu markkinointi voidaan hyödyntää esimerkiksi verkkokaupoissa asiakkuuden ylläpitämiseksi. Automatisoituun markkinointiin voidaan asettaa esimerkiksi tavoitelaskutus, joka tulisi täyttyä puolen vuoden aikana. Mikäli tämä tavoite ei täyty, asiakas saa automatisoidusti markkinointiviestejä taikka alennuskoodeja. Automatisoitua markkinointia voidaan hyödyntää todella laajasti ja sitä hyödynnetään niin uusien asiakkaiden hankintaan, kuin olemassa olevien asiakkuuksien ylläpitämiseen.

Monet markkinoijat suosivat tekoälyä muuntaakseen dataa arvokkaiksi asiakasnäkemyksiksi. Koneoppiminen on erityisen tärkeää digitaalisessa markkinoinnissa, sillä se auttaa luomaan yhtenäisen tukijärjestelmän. Kun tekoäly otettiin käyttöön digitaalisessa markkinoinnissa, tuli siitä nopeasti yksi markkinointiprosessin viimeistelevistä askeleista ja tehokas tapa saattaa markkinointikampanja loppuun (Arasu ym. 2020). Tekoälytyökalut käyttävät dataa ja asiakasprofileja oppiakseen kommunikoimaan asiakkaiden kanssa parhaiten ja täten tarjoamaan heille räätälöityjä viestejä oikeaan aikaan ilman markkinointitiimin jäsenten puuttumista asiaan, mikä varmistaa maksimaalisen tehokkuuden.

Luvun johdannossa mainitut tekoälyn alaisuuteen kuuluvat alalajit: Koneoppiminen, Neuroverkot, Syvä oppiminen, Kognitiivinen tietojenkäsittely, Konenäkö sekä Luonnollisen kielen käsittely (NLP, Natural language processing) on hyvä ymmärtää, ennen niiden käyttöönottoa. *Koneoppiminen* automatisoi teoreettisten rakenteiden kehittämistä. (SAS 2019.)

Neuroverkot ovat koneoppimisen tyyppi, joka koostuu toisiinsa yhdistetyistä yksiköistä (kuten neuroneista), jotka käsittelevät tietoa vastaamalla ulkoisiin palautteisiin ja välittävät tietoa kunkin yksikön välillä. Prosessi vaatii useita läpikulkujia tiedoilla, jotta yhteyksiä löydetään ja saadaan merkitys määrittelemättömälle tiedolle (SAS 2019). *Syväoppiminen* on koneoppimisen osajoukko, joka on pohjimmiltaan hermoverkko, jossa on kolme tai useampia kerroksia. Nämä hermoverkot yrittävät simuloida ihmisaivojen käyttäytymistä (IBM 2022). *Kognitiivinen tietojenkäsittely* on tekoälyn alakenttä, joka yrittää normaalia, ihmisen kaltaista viestintää. (SAS 2019.)

Konenäkö perustuu kuvioiden tunnistamiseen ja syvälliseen oppimiseen tunnistaaakseen, mitä kuvassa tai videossa on. Kun koneet voivat käsitellä, analysoida ja ymmärtää kuvia tai videoita, voivat koneet ottaa kuvia tai videoita reaaliajassa ja tulkita ympäristöään (SAS 2019). *Luonnollisen kielen käsittely* tavoitteena on se, että tietokoneet pystyisivät käsittelemään suuria määriä luonnollisen (eli ihmisten) kielen dataa (Aaltonen 2019). Luonnollisten kielten käsittelyn avulla tietokone voi myös analysoida ja tuottaa ihmiskieltä, kirjoitettuna tai puhuttuna.

Monet koneoppimisenmenetelmät ovat poikkitieteellisessä käytössä. Esimerkiksi lineaarista- ja logistista regressiota käytetään laajasti niin koneoppimisessa, kuin markkinoinnissa. Vaikka toimintaperiaate ja näkökulma on sama, painotettavat asiat vaihtelevat alan mukaan. Muita koneoppimisenmenetelmiä käytetään kuitenkin harvemmin markkinoinnissa ja markkinointitutkimuksissa. (Ma & Sun 2020, 4-13.)

Valvottua oppimista voidaan hyödyntää markkinoinnin alalla. Valvotuissa oppimistehtävissä järjestelmälle tarjotaan harjoitustietojoukko. Valvottu oppiminen pyrkii oppimaan harjoitustietojoukosta funktion $Y = f(X)$ ja tämän avulla ennustamaan tuloksia. Mikäli Y on numeerinen/kategorinen muuttuja, tehtävää kutsutaan regressioksi/luokitukseksi. Ennustaminen on valvotun oppimisen pääpaino. Ennustava tarkkuus on arvioitava käyttämällä erilaista testausaineistoa. (Ma & Sun 2020, 4-13.)

5 KUINKA TOTEUTTAA AUTOMAATTISEN DATA-ANALYYSIN INTEGROINTI HAKUKONEMARKKINOINNIN TOTEUTUKSEEN

Viime vuosien teknologinen kehitys on sysännyt markkinoinnin kehityksen liikkeelle. Markkinointi on saavuttanut kehityksessään pisteen, jossa digitaalisiin trendeihin sopeutuminen on välttämätöntä. Vaikka sen kuvittelisi olevan lisätyötä markkinoijille, itse asiassa kaikki automatisoidut sovellukset ja tekoälyyn perustuvat järjestelmät vähentävät normaalia kohdentamisen ja räätälöinnin monimutkaisia prosesseja. Monissa tilanteissa online-mainontaan käytetyt alustat sisältävät algoritmeja parhaiden yhdistelmien tunnistamiseksi, muissa tilanteissa yritykset ovat aloitteellisia kehittämään ja toteuttamaan omia räätälöityjä järjestelmiään (Dumitriu & Popescu 2019). Luvun tavoitteena on tarjota malleja, jotka käyttävät tekoälyä markkinointiratkaisuihin, joiden avulla yritys voi parantaa verkkosivuston näkyvyyttä.

Internet, sosiaalinen media ja mobiililaitteet ovat dramaattisesti lisänneet yritysten ja kuluttajien välistä vuorovaikutusta. Multimediamuodot, kuten teksti, kuvat ja videot sisältävät koodattua tietoa. Yritysten on ymmärrettävä kuluttajien käsitykset ja mieltymykset sekä hankittava oivalluksia brändin sijoittelusta tämän perusteella. Informatiivinen suunnittelu ja houkutteleva sisältö tietoisuuden, havainnoinnin ja hyväksynnän lisäämiseksi on tärkeä kilpailualue (Ma & Sun 2020, 4-13). Tekoälytyökalujen kapasiteetin ja prosessointitehokkuuden vuoksi näitä tuleekin hyödyntää kohdennettua vuorovaikutusta tehdessä.

Tekoälyn käytön yleistyessä kaikkialla, hakukoneoptimoinnin asiantuntijoiden tulee ymmärtää koneoppimisen ja automaation merkitys. Tulevaisuudessa yhä useammat ihmiset alkavat tehdä hakuja vokaalihaun avulla, joten hakukoneoptimoinnin ammattilaisten tulee sopeutua tähän suhteellisen uuteen hakuun. Vokaalihaun vaikeus syntyy siitä, että puhepyynnöt eroavat tavallisista kyselyistä ja sitten avainsanatutkimusprosessi todennäköisesti muuttuu. Äänihaku ei sisällä pelkästään enemmän sanoja, vaan on nykyisellään monin

tavoin vaikeammin ymmärrettävää koneille. Tekoälylle on vaikeaa ymmärtää puhetta, esimerkiksi murteen, puhenopeuden tai äänen vuoksi. Avainsanat ovat muiden komponenttien ohella yksi sopivan SEO-strategian tärkeimmistä elementeistä, olipa kyseessä online-yritys tai esittelysivusto. (Dumitriu & Popescu 2019.)

Jotta yritys voi vastata tulevaisuudessa vokaalihakua käyttävien hakukoneen käyttäjien hakuihin, tulee heidän käyttää myös todella alueellisia avainsanoja tai -lauseita esimerkiksi murteiden vuoksi. Vokaalihaun optimointi vaatii suurta manuaalista työmäärää.

Avainsanat ovat lauseita, pidempiä lauseita tai sanoja, joita käyttäjä käyttää löytääkseen sivuston hakukoneissa. Periaatteessa optimoidun sivuston on puhuttava ihmisten puhekieltä käyttämällä samoja avainsanoja kuin käyttäjät. Tätä kieltä tulisi käyttää sisällössä, metatiedoissa, kuvien vaihtoehtoisessa tekstissä. Tässä muuttuvassa ympäristössä yksi harvoista asioista, jotka ovat pysyneet muuttumattomina, on avainsanat, -sopivimpien ja osuvimpien sanojen löytäminen tietyn yrityksen toimipaikan tai tuotteen/palveluiden tunnistamiseksi. (Dumitriu & Popescu 2019.)

Tarve löytää avainsanat ovat pysyneet samoina, mutta tapa, jolla se tehdään, on kehittynyt. Tutkimalla tietyillä markkinoilla käytettyjä avainsanoja ei ole mahdollista vain määrittää termejä ja lauseita, jotka tulisi sisällyttää mukaan SEO-strategiaan (Dumitriu & Popescu 2019). Avainsanoja ja -lauseita tutkiessa ja otettaessa käyttöön tulee pohtia, käytetäänkö niissä asiakkaan käyttämää kieltä.

Tulevaisuudessa tulisi määritellä enemmän potentiaalisista asiakkaista ja heidän toiveista, tarpeista, odotuksista sekä kiinnostuksen kohteista. Tärkeintä on houkutella kiinnostuneita vierailijoita, joilla on suuret mahdollisuudet konversioon. Avainsanatuutkimus antaa mahdollisuuden ennakoita markkinoiden muutoksia, sopeutua näihin muutoksiin ja luoda tuotteita, palveluita ja tarjouksia, jotka ovat käyttäjät etsivät jatkuvasti. (Dumitriu & Popescu 2019.)

Laadukkaiden verkkovierailijoiden saaminen verkkosivustolle vaikuttaa myös positiivisesti sivuston laatuun Googlen algoritmeissa. Mikäli verkkovierailija viettää sivulla aikaa ja siirtyy alasivujen välillä, viestii tämä algoritmille, että sivusto vastasi hakukyselyä. Tästä syystä asiakkaiden odotusten ja tarpeiden ymmärtäminen on tärkeää.

5.1 Automatisoidun tiedonlouhinnan toteuttaminen

Tietojen validoinnille on olemassa erilaisia määritelmiä, mutta yleensä tietojen validointi on prosessi, jossa toimitetaan puhdasta ja tarkkaa tietoa tietyille ohjelmille, sovelluksille tai palveluille. Esimerkiksi koneoppimisessa tietojen validointi tarkoittaa lähteen laadun ja tarkkuuden tarkistamista. Erilaisia validointityyppejä voidaan suorittaa riippuen tietojoukon tavoitteista ja rajoituksista. Tietojen validointi tapahtuu yleensä tiedonpoiminnan, siirron- ja latausprosessin (ETL, Extract, Transfer and Load) muunnosvaiheessa, jossa tiedot ensin poimitaan tietolähteestä. Vaihe 1; kun data on validoitu, puhdistettu, muotoiltu ja liitetty muihin tietojoukkoihin. Vaihe 2; valmis ladattavaksi ja vaihe 3; valmis prosessoitavaksi annetun käyttötapauksen mukaisesti. (Ahmand & Dias 2021.)

Tietojen validointi mahdollistaa sen, että yritys voi luottaa tietoihinsa. Yritykset luottavat korkealaatuiseen dataan tehdessään kriittisiä päätöksiä organisaationsa kannalta. Jos tiedot eivät ole oikeita ja tarkkoja, loppukäyttäjät eivät välttämättä saa vastaansa heille kohdennettuja mainoksia tai www-sivustoja.

Toimialasta riippuen on monia syitä, miksi tietojen validointi on kriittinen osa tietopohjaisissa projekteissa. Kaksi tärkeintä syytä, jotka usein otetaan liian kevyesti, ovat: *virheiden varhainen havaitseminen* ja *ajan säästö*. Nämä kaksi kulkevat käsi kädessä, mutta on tärkeää korostaa niitä erikseen. Päittäjillä ja data-asiantuntijoilla on taipumus sulkea pois tietojen validointi, koska heillä ei ole tietoa näiden kahden monista seurauksista (Ahmand & Dias 2021). Tietojen

validoinnilla on käytännön hyötynä se, että siinä varmistetaan tiedon vastaavan yrityksen haluamaa dataa.

Tietojen validoinnin suorittaminen: tietojen validoinnissa käytetään sääntöjä, jotka varmistavat tietojen eheyden, esimerkiksi oikean muodon sähköpostiosoitteen syöttämiselle tai salasanan vähimmäispituuden. Nämä perustietojen validointisäännöt auttavat ylläpitämään standardeja, jotka tehostavat tehokkaasti tietojen käsittelyä. Tietojen validointiprosessin aikana on tärkeää, että myös tietomallin standardit ja rakenne ymmärretään hyvin. Strukturoidut tiedot ovat tietoa, jotka on muotoiltu hyvin määritettyyn tietomalliin, kun taas jäsentämätön data on dataa raakamuodossa. On myös puolistrukturoitua dataa, joka sijoittuu strukturoidun ja strukturoimattoman tiedon väliin. (Ahmand & Dias 2021.)

Tällaisten tietojen välisen eron ymmärtäminen auttaa ylläpitämään yhteensopivuutta sovellusten ja muiden tietojoukkojen kanssa, joiden kanssa dataa integroidaan ja tallennetaan. Tietojen validointiin on saatavilla monia menetelmiä. Yleisimmät menetelmät ovat koodien ja ohjelmistojen käyttö. Koodin kirjoittaminen voi olla vaihtoehto niille, jotka hallitsevat sujuvasti koodauskieliä, kuten Python ja R. Koodi mahdollistaa tietojen paremman tulkinnan, kuten datan arvojen ja rakenteen vertaamisen ennalta määritettyihin sääntöihin ja kaikkien tarvittavien tietojen tarkistamisen vaaditun rajoissa. Laatuparametrit, tämä menetelmä voi kuitenkin olla hyvin aikaa vievä tietojoukon monimutkaisuudesta ja koosta riippuen. Nykyään on olemassa monia kaupallisia ohjelmistoja, joita voidaan käyttää tietojen validointiin. Useimmille ihmisille tämä on suositeltavin menetelmä, koska ohjelma on kehitetty ymmärtämään yleisiä sääntöjä ja tiedostorakenteita, joita käytetään tietojen validoinnissa alalla. Käyttäjältä ei vaadita syvällistä ymmärtämistä taustalla olevasta formaatista. (Ahmand & Dias 2021.)

Automaattisesta tiedonlouhinnasta puhuttaessa tarkoitetaan prosessia, jossa tietoa poimitaan automaattisesti ohjelmiston avulla. Automaattisella tiedonkeruulla voidaan toteuttaa esimerkiksi: hintaseurantaa, hintatietojen vertailua, dynaamista hinnoittelua taikka arvosteluja. Tekoälyn ja koneoppimisen

avulla tiedon automaattisesta keräämisestä voidaan luoda uusia kasvumahdollisuuksia sekä toimintatapoja. (Benson 2021.)

Kaikkien yritysten, jotka haluavat kerätä näkemyksiä, uudistaa toimintaansa tai luoda uusia kasvunäkymiä, tulee pyrkiä liiketoiminnan automatisointiin. Massiivisista tiedonkeruutoiminnoista niinkin yksinkertaisiin kuin useiden sosiaalisen median tilien käyttöön, oikeat työkalut ovat välttämättömiä, jotta organisaation sisällä saavutetaan ymmärrys ja tieto, kuinka dataa hyödyntää. Mitä enemmän tietoja on käytettävissä analytiikkaa varten, sitä parempi on kyky tehdä perusteltuja arvioita (Benson 2021). Automaattista tiedonkeruuta hyödyntämällä yritys voi saavuttaa kattavamman näkemyksen haluamastaan asiasta, mikäli. Datan määrä ja laatu vaikuttavat kuitenkin siihen, voidaanko tiedonkeruun avulla hankittuun tietoon luottaa.

Automaattisen tiedonkeruun etuja ovat aikaa säästävät ja kustannustehokkaat lähestymistavat tiedonkeräämiseen ja -käsittelyyn. Automaattinen tiedonkeruu ei vaadi juurikaan ihmisten osallistumista ja sen avulla voidaan hakea tietoa nopeammin, ennustaa riskejä sekä sen tarjoamien analyysien perusteella parantaa kuluttajakokemusta. Tulee kuitenkin muistaa, että automaattisen tiedonkeruun prosessit eivät ole virheettömiä, mutta virheprosentti laskee jatkuvasti teknologian kehittyessä.

Tiedonlouhintaa yritykset käyttävät saadakseen tehokkuutta tiedonhakuprosesseihinsa. Yksinkertaisesti sanottuna tiedonlouhinta voidaan jakaa kahteen erilliseen luokkaan, joita kutsutaan kuvaaviksi ja ennustaviksi luokiksi. Yhdistämällä erilaisia koneoppimisen algoritmeja tiedonlouhinnan järjestelmään voidaan rakentaa useita erilaisia käyttöliittymiä, joilla kaikilla on eri tarkoitus. Tekstilouhinnalla on merkittävä rooli. Internetissä vain 20 % tiedosta on strukturoitua, ja loput 80 % on strukturoimatonta tietoa. (Arasu ym. 2020.)

Yleensä tekstinlouhinnalla saatava tieto on arvokasta. Tavoitteena on hankkia dataa hakemalla tietoa, jota voidaan käyttää nykyisen strategian parantamiseen. Tämä koostuu tiedostojen tutkimisesta sekä olennaisten datamallien löytämisestä. Kuluttajat kirjoittavat verkossa erilaisista asioista, kuten palaute

yrittäjien tuotteista tai palveluista. Tekstin louhintaa voidaan siis käyttää järjestämään se hienostuneemmaksi malliksi, jotta se on helpompi analysoida ja lopulta käyttää markkinointistrategian viimeistelyyn taikka tuotteen/palvelun parantamiseen (Arasu ym. 2020). Tiedonlouhintaan löytyy markkinoilta jo useita, luotettavia työkaluja niitä tarvitsevien yritysten tai henkilöiden käyttöön. Esimerkiksi Rapid Miner, Oracle Data Mining tai Weka.

Tiedonlouhintavaiheet tiedonhakuprosessissa ovat seuraavat: *Tietojen puhdistus*: epäjohdonmukaisten tietojen poistaminen. *Tietojen integrointi*: useiden tietolähteiden yhdistelmä. *Tietojen valinta*: analyysin kannalta merkitykselliset tiedot haetaan tietokannasta. *Tietojen muuntaminen*: tietojen yhdistäminen ja muuntaminen louhintaan sopivaan muotoon. *Tiedon louhinta*: älykkäiden menetelmien käyttö kuvien poimimiseen tiedoista. *Kuvion arviointi*: mielenkiintoisten kuvien tunnistaminen. *Tiedon esittäminen*: esittelyssä käytetään visualisointi- ja tiedonesitystekniikoita, kun poimittua tai louhittua tietoa esitellään loppukäyttäjälle. (Jagtap & Kodge 2013, 1-6.)

Blomster ja Koivumäki (2020) mukaan markkinointiorganisaatioiden koneoppimisen kehitysprojektit noudattavat usein 5-vaiheista 5-porttimallia ja täten koneoppimisen kehitysprojektit tulisi jakaa viiteen (5) vaiheeseen. Vaihe 1: Projektien arviointi, vaihe 2: data-analyysi, vaihe 3: koneoppimisen algoritmien tutkiminen ja kehitys, vaihe 4: koneoppimismallien kehittäminen, vaihe 5: iterointi ja markkinoille tuominen. Jokainen vaihe sisältää viisi tunnistettavaa porttia: toteutettavuus, koulutussarja, algoritmiraikaisu, koneoppimismalli sekä julkaisun jälkeinen tarkastelu ja kehitys. (Blomster & Koivumäki 2021, 18-24.)

Yrityksen koneoppimisen kehitysprojekteissa koneoppimista kehittävä henkilöstön tulee täten muodostaa ymmärrys koneoppimismallin kilpailutarpeesta ja markkinointijärjestelmästä, joihin se on sisällytetty, jotta oikean datan saaminen varmistetaan. Hakukonemarkkinoinnista vastaavan henkilöstön tulee kyetä tarjoamaan näkemys siitä, kuinka hyvin koneoppimismallin tarjoamat tiedot vastaavat tarpeita. Tästä syystä yrityksen tulee omata taikka hankkia asiantuntevia henkilöitä vastaamaan hakukonemarkkinoinnin rooleista, sillä se on erityisen tärkeää, jotta kehitysprojektista saadaan kattava kuva.

5.2 Automatisoidun data-analyysin integrointi hakukonemarkkinointiin

Drivas ym. (2020) mukaan ensimmäisen hypoteesin mukaan verkkosivuston koko johtaa suurempaan tai pienempään orgaaniseen hakukoneliikenteeseen. Tämä liittyy hakukoneoptimointitekniisiin yhteensopivuustekijöihin, kuten: Sivuston suojaus, sivuston nopeus, onko sivustolla sama sisältö kahdesti ja kuinka strukturoitu data on merkattu verkkosivuille ja siihen, seuraavatko sivustot niitä ja missä määrin. On olemassa aikaisempia hakukoneoptimoinnillisia lähestymistapoja, jotka osoittavat, että Google käyttää luokitusalgoritmissaan yli 250 tarkasti varjeltua yhteensopivuustekijää. Kuitenkin monet tutkimukset ovat tutkineet vain osaa niistä. Drivas ym. (2020) mukaan tämä liittyy kyvyttömyyteen hakea laajamittaisia tietojoukkoja koskien teknisiä ja käyttäytymiseen liittyviä tekijöitä, jotka vaikuttavat hakukoneiden orgaanisen liikenteen prosenttiosuuteen (Drivas ym. 2020). Vaikkakin Googlen varjelemat yhteensopivuustekijät ovat tarkoin varjeltuja, voi hakukoneoptimointitekniisiä yhteensopivuustekijöitä kuitenkin tarkastella ja parantaa, joko valmiin ohjelmiston avulla taikka manuaalista työtä hyödyntäen.

Tätä taustaa vasten suurten datajoukkojen analytiikka tarjoaa uusia mahdollisuuksia strategiseen hakukoneoptimoinnin suunnitteluun ja käyttöönottoon. Sovelluskäyttöliittymän/ohjelmointirajapinnan (Application Programming Interface, API) integroinnin ja esikäsittelyn avulla kerätyt tiedot johtavat useisiin teknisiin ja käyttäytymiseen liittyviin SEO-muuttujiin, jotka todennäköisesti vaikuttavat verkkosivustojen orgaanisen hakukoneliikenteen prosentuaaliseen vaihteluun. Pääkomponenttianalyysin perusteella Drivas ym. (2020) jakoivat nämä tekniset ja käyttäytymiseen liittyvät muuttujat neljään lisätekijään: hakukoneoptimoinnin indeksointi, verkkosivuston latausnopeus, verkkosivuston suojaustila ja käyttäjien käyttäytyminen. (Drivas ym. 2020.)

Drivas ym. (2020) tutkimuksessaan toteuttivat suurten datajoukkojen analytiikan haun integroimalla ja parametriomalla kolmea toisistaan poikkeavaa API: a. He käyttivät Googlen Search Consolen URL-sovellusliittymää hakemaan tietoja verkkosivustojen koosta. Tutkimuksessaan he tutkivat 171 URL-osoitetta ja niistä kuvien sekä ulkoisten ja sisäisten linkkien määrää. Kaikilta verkkosivustoilta

haettiin yhteensä 141 851 verkkosivua, 86 788 kuvapolkutiedostoa ja 254 373 sisäistä ja 97 626 ulkoista linkkiä ja ne tallennettiin Googlen Arts & Culture tietoaarkistoon. (Drivas ym. 2020.)

Checkbot-sovellusliittymää käytettiin teknisten tekijöiden, kuten hakukoneoptimoinnin indeksoinnin, verkkosivuston latausnopeuden ja verkkosivuston suojaustilan, yhteensopivuuden kiteyttämiseen. API:ssa sovellettiin lisäparametreja HTML-, JavaScript- ja CSS-koodin teknisten mallien tutkimiseksi ja löytämiseksi. Kullekin haetulle muuttujalle otettiin käyttöön asteikko 0–100 hyödyllisenä johtamismenetelmänä verkkosivuston suorituskyvyn mittaamisen monimutkaisuuden ratkaisemiseksi. Yhteensä 9405 yksilöllistä esiintymää haettiin 55 tutkitulle muuttujalle, vaihteluvälit 0 % - 100 %. (Drivas ym. 2020.)

Drivas ym. (2020) tutkimuksen tulosten perusteella kaikilla tuotetuilla malleilla on selkeä tilastollinen merkitsevyys p-arvoilla alle 0,05. Siitä huolimatta verkkosivuston latausnopeudessa on marginaalinen poikkeus, p-arvo 0,061. Tulokset kuvaavat yhdenmukaisuutta aikaisempien tutkimusten kanssa, joiden mukaan ranking-algoritmit sisältävät valtavan määrän salaisia muuttujia SEO-kontekstissa. Www-sivustojen koko selitti jopa 30.06 % vastaustietojen vaihtelusta sen keskiarvon ympärillä; SEO indeksointi, 17.7 %; Verkkosivuston latausnopeus, 10.01 %; Verkkosivuston suojaustila, 17.9 %; ja käyttäjien käyttäytyminen, 29.5 %. (Drivas ym. 2020.)

API:n tehokkuus lisääntyy huomattavasti, kun testeissä käytetään realistista dataa, joka edustaa todellisia olosuhteita. Google Page Speed Insights-sovellusliittymän avulla verkkosivun kehittäjät voivat mitata verkkosivun suorituskykyä ja saada ehdotuksia sivun suorituskyvyn, käytettävyyden ja hakukoneoptimoinnin parantamiseksi. Tämä API palauttaa tiedot Chromen käyttökokemusraportista ja laboratoriotiedot: Lighthouse:sta (Googlen avoimen lähdekoodin automaattinen työkalu verkkosivujen laadun parantamiseen). Esimerkiksi Semrush ja Supermetrics tarjoavat laadukkaita ja luotettavia API-työkaluja.

5.3 Hakukoneoptimoinnin automatisointi

Tekoäly on tärkeä kestävä kehityksen kannalta. Tekoälyn kilpailuetuna on se, että se on aina reaaliajassa, jonka perusteella markkinoijat voivat jatkuvasti toteuttaa tekoälyn räätälöityjä, oivalluksiin perustuvia vuorovaikutuksia asiakkaiden kanssa yksilöllisesti. Yritykset, jotka ovat ymmärtäneet tekoälyn merkityksen ja ovat onnistuneet hankkimaan toimivan tekoälyn perustuvan järjestelmän onnistuvat luomaan kilpailuedun itselleen.

Hakukoneoptimoinnin automaatio on prosessi hakukoneoptimoinnin tehtävien systematisoimiseksi ohjelmistojen tai verkkosovellusten avulla. Ajatuksena on tehostaa markkinointikampanjaa. Hakukoneoptimoinnin automaation käyttö vähentää markkinointikampanjassa tarvittavan manuaalisen työn määrää. Tämä nopeuttaa aikaa vievää toimintaa. Vaikka kaikkia hakukoneoptimointitoimintoja ei voida automatisoida, jotkin tehokkuutta lisäävät toimet voivat olla hyödyksi yrityksille.

Hakukoneoptimoinnin automaation kysyntä kasvaa, koska hakukoneoptimoinnin alalla on edessään suuri ongelma: resurssien ja käytettävissä olevien kykyjen puute. Tämä tulee haastavammaksi hakukoneoptimoinnissa, jossa tietojoukot ovat laajoja, tiheitä ja monimutkaisia. Hakukoneoptimoinnin automatisoinnin avulla digitaaliset markkinoijat voivat lyhentää datasta oivallusten saamiseen kuluvaa aikaa. Tämä vapauttaa resursseja toteuttaa strategioita nopeammin. (Srivastava 2021.)

Hakukoneoptimoinnin automatisoinnissa tulee yrityksen hyödyntää dataa esimerkiksi verkkosivuvierailijoista, verkkosivuvierailuista, siitä kuinka verkkosivuille päädytään (mistä ja miten) sekä mitkä alasivut ovat sivun sisällä suosittuja. Tätä dataa hyödyntäessä yritys voi paikantaa ”pullon kaulat” verkkosivuiltaan ja automaatiotyökalujen avulla parantaa näitä. Esimerkiksi ScreamingFrog ja Google Analytics-työkalut voivat toimia yrittäjän apuna datan hakemisessa ja analysoinnissa.

Tekninen SEO: Sivuston sijoitukseen vaikuttavien teknisten osien tarkastelu keskittyy verkkosivuston nopeuteen, käyttäjäkokemukseen, Verkon perustietoihin, mobiilivastauksiin ja strukturoituun dataan. *Automaatio:* Tässä tapauksessa automaatio toimii jo hyvin useiden SEO-ohjelmistojen, kuten MOZ, SEMRUSH ja WooRank, verkkosivustojen indeksointiohjelmistojen, kuten ScreamingFrog, Sitebulb avulla. Kyseessä olevat ohjelmistot kuitenkin vaativat käyttäjältään alan ymmärrystä ja käsityksen siitä, kuinka dataa tulee analysoida ja hyödyntää. (Volpini 2022.)

Otsikkotunnisteet, sisällön kuvaukset ja otsikot hakukoneoptimointia varten voidaan ratkaista myös automaatiolla; tekoäly ja syväoppiminen on hyödyllistä tässä tapauksessa. SaaS-palvelulle voidaan kouluttaa kielimalleja, jotka on räätälöity mihin tahansa sisällön optimointivaatimukseen, esimerkiksi metakuvausten luomiseen tai otsikkotunnisteen optimointiin. Myös luonnollisen kielen käsittelyä voidaan käyttää. (Volpini 2022.)

Sivun ulkopuolinen SEO: Käänteisten linkkien luominen ja parantaminen on sivun ulkopuolisen hakukoneoptimoinnin yleisin osa-alue. Käänteisten linkkinen tarkistus on suositeltavin vaihtoehto sivun ulkopuoliseen SEO-tarkasteluun. *SEO-strategia* sisältää liikennemallitutkimuksen, A/B-testauksen ja tulevaisuuden ennustamisen. Automaationa tässä tulee hyödyntää aikasarjaennusteita, joissa voi soveltaa myös koneoppimista. Koneoppimismalleja käyttämällä voi ennustaa tulevaisuuden trendejä ja tunnistaa aiheita, joissa verkkosivusto pärjää jo hyvin. (Volpini 2022.)

Hakukoneoptimointi ilman automaatiota sisältää useita, manuaalisesti alusta asti suoritettavia tehtäviä. Hakukoneoptimointiprosessin aloittaminen manuaalisesti alkaa verkkosivuston näkyvyyden tarkistamisella. Tarkistusvaiheessa selvitetään, kuinka paljon sivustolla on verkkovierailijoita, millä avainsanoilla se sijoittuu hakukoneiden tulossivuilla ja mille sijoille. Sivustot tulee myös tarkistaa teknisten virheiden varalta. Kun alustavat tehtävät on tehty, tulee selvittää, onko toimilla ollut kannattava vaikutus. Ovatko toimet saaneet lisää maksavia asiakkaita sivustolle ja ovatko verkkovierailijat löytäneet sivuilta etsimänsä.

Hakukoneoptimoinnin automatisoinnilla voidaan helpottaa yllä mainittuja prosesseja. Hakukoneoptimoinnin automatisoinnilla voidaan seurata verkkosivuston sijoittumista hakukoneen tulossivulla, verkkosivuston indeksoitumista, mainintoja brändistä ja koostaa automaattisesti raportit niihin tarkoitetuilla työkaluilla.

Automatisointi helpottaa huomattavasti hakukoneoptimointia harjoittavien ammattilaisten työmäärää. Sen avulla voidaan myös analysoida yrityksen linkkiprofiilia sekä linkkien laatua. Tekoäly ja sen automaation avulla voidaan jo tänä päivänä tuottaa sisältöä automaattisesti, avainsanatutkimukseen pohjautuen. Tekoälyn tuottamaa sisältöön ei voida kuitenkaan vielä tällä hetkellä luottaa täysin, varsinkaan muulla, kuin englannin kielellä.

6 JOHTOPÄÄTÖKSET

Opinnäytetyössä tutkittiin, kuinka automatisoitu datan analysointi voidaan integroida osaksi hakukonemarkkinoinnin toteuttamista. Lisäksi tavoitteena oli tehdä opinnäytetyöstä opastuinen teos, jonka avulla lukija voi itse aloittaa hakukonemarkkinoinnin automatisoinnin. Opinnäytetyöhön pohjaten, tekoälyn vaikutus näkyy jo nyt digitaalisessa markkinoinnissa. Tulevaisuudessa lisääntyvät automaatiot tulevat auttamaan markkinoinnin ammattilaisia hakukonemarkkinoinnissa suuresti. Hakukonemarkkinointia tekevien ammattilaisten työ tulee olemaan luovempaa ja heillä on aikaa kohdata enemmän aikaa vieviä haasteita, sillä tekoäly hoitaa yksitoikkoiset ja helpot tehtävät. Tekoälyn integroiminen hakukonemarkkinoinnin toteutukseen on silti vielä lapsenkengissä ja tulee vaatimaan poikkitieteellistä yhteistyötä, jotta tekoäly ja koneoppiminen saadaan vietyä yleiseen käyttöön tavalla, joka helpottaa ammattilaisten arkea. Uskon, että tekoälyn myötä pienet, inhimilliset virheet esimerkiksi oikolukemisessa taikka muissa helposti ohimennen tarkastettavissa asioissa tulevat vähenemään huomattavasti.

Suurten datajoukkojen avulla yritykset voivat hankkia tarkemman, sekä laajemman käsityksen asiakaskunnastaan, asiakkaidensa tarpeista ja mielenkiinnon kohteista. Tämä auttaa yrityksiä tarjoamaan asiakkailleen kohdennetumpia ja tarkempia mainoksia, koskien tuotteita tai palveluita. Suurten datajoukkojen automaattisen data-analyysin integroinnissa on kuitenkin vielä haasteita. Esimerkiksi henkilö, joka tilaa vaatetuotteita verkkokaupasta muille kuin itselleen, tulee jatkossa saamaan kohdennettuja mainoksia edellisten tilausten pohjalta, mikä ei toimi omanaan herättämään asiakkaan omaa ostoimpulssia.

Yritys, joka aloittaa hakukonemarkkinoinnin automatisaation automaattiseen data-analyysiin pohjaten tulee omata taikka hankkia ammattitaitoisia henkilöitä tämän toteuttamiseen. Koska ala on itsessään niin laaja, on hyvin todennäköistä, että riittävää osaamista ei löydy valmiiksi yrityksen sisältä, eikä välttämättä yksi konsultti taikka osaava henkilö riitä rakentamaan luotettavaa ja kaikki huomioon

ottavaa järjestelmää. On kuitenkin olemassa jo valmiiksi suunniteltuja ja rakennettuja ohjelmistoja, joita voi hankkia. Yleensä nämä palvelut toimivat kuukausimaksu-periaatteella ja ovatkin tästä syystä pienemmille yrityksille kustannustehokkaampia ja parempia vaihtoehtoja.

Koneoppimisen integrointia hakukonemarkkinointiin tulee lähestyä yrityksen kyvyistä tarjota riittävät teknologiset- sekä tietoresurssit. Koneoppimisen integroinnista hakukonemarkkinointiin suunnittelevan yrityksen tulee kyetä rekrytoimaan, säilyttämään ja kehittämään henkilöstönsä markkinoinnin ja koneoppimisen osaamista (Blomster & Koivumäki 2021, 18-24). Jotta yritys onnistuu toteuttamaan yllä mainitun lähestymistavan, tulee rekrytoinnista vastaavien henkilöiden, sekä johtajien ymmärtää alaa ja alan tarpeita riittävän syvällisesti.

Myös tiedonhallinta tulee nähdä tärkeänä ja olennaisena kykynä hakukonemarkkinoinnin automaatiosta kiinnostuneissa yrityksissä henkilöstön osaamisen rinnalla. Jotta verkkosivustoille ja mainoskanaville, kuten Google Ads voidaan tarjota toteuttamiskelpoisia ratkaisuja, tulee organisaation sisällä olla osaamista markkinointitietojen ymmärtämisestä. Ymmärrys siitä, että Luonnollisten kielten käsittelyn algoritmeilla hakukoneoptimoinnin ja hakukonemarkkinoinnin tekstidata on hallittava oikein, jotta koneoppimisen ja digitaalisen markkinoinnin tarkoituksiin käytettävässä kielessä on yhtenäinen rakenne. (Blomster & Koivumäki 2021, 18-24.)

Yritykset voivat myös luottaa ulkoisiin palveluntarjoajiin, tarjoamaan heille automaattisen data-analyysin tuloksia hakukonemarkkinointiin integroimiseksi. Ulkoisten palveluntarjoajien hyödyntäminen voi olla perusteltua riippuen hakukonemarkkinoinnin automaatiosta kiinnostuneen yrityksen koosta ja resursseista. Mikäli yritys haluaa rakentaa oman tukijärjestelmän ja kouluttaa oman koneoppimismallin, tulee yrityksessä olla: Hakukonemarkkinoinnin johtamisen osaamista, teknologisen ympäristön ymmärtämistä, analytiikkaa, asiakasnäkemyksiä ja järjestelmien tarjoamaa dataa. Tämän kaiken yhteen saattaminen kuitenkin vaatii kokonaisvaltaista ymmärrystä koneoppimisesta,

automaattisesta datan analysoinnista ja sen integroinnista hakukonemarkkinoinnin toteutukseen.

Tulevaisuuden tutkimuksissa aihetta voisi rajata kapeammaksi ja keskittää tutkimus automaattisen data-analyysi työkalun rakentamisen vaiheisiin, konkreettiseen rakentamiseen hyödyntäen tekoälyä ja koneoppimista sekä valmiin palvelun testaamiseen ja markkinoille vientiin.

7 POHDINTA

Opinnäytetyön eteneminen sujui vaiheittain. Opinnäytetyöprosessin aloitus tapahtui elokuussa 2021, yhteydenotolla aiheen antaneeseen yritykseen, jonka kanssa keskustelu koski digitaalisen markkinoinnin tulevaisuutta ja sitä millainen työ toisi lisäarvoa toimialalle, mutta myös minulle arvokasta oppia koskien tulevaisuuden digitaalista markkinointia. Työn oli tarkoitus toimia empiirisenä tutkimuksena, jossa teoriaosuudessa käsitellään datan automaatiota ja hakukoneita sekä hakukonemarkkinointia ja tutkimusosiossa käytettäisiin toimeksiantajan SaaS-palvelua ja täten saataisiin tietoa palvelun hyödyistä ja ominaisuuksista. Työn edetessä kävi kuitenkin selväksi, että opastyylinen opinnäytetyö palvelee tarkoituksia paremmin ja antaa enemmän mahdollisuuksia pureutua aiheeseen kokonaisvaltaisesti. Opinnäytetyöhön oli mielekästä, mutta todella hankalaa etsiä tietoa ja tutkimuksia opinnäytetyön poikkitieteellisyyden ja aiheen uutuuden vuoksi. Opinnäytetyössä lähteinä käytettiin suomen- sekä englanninkielisiä kirjoja, verkkolähteitä sekä erilaisia aiheeseen liittyviä tutkimuksia ja blogikirjoituksia. Opinnäytetyöhöni tehdyt valinnat, siinä käsitellyt väittämät ja käsitteet on johdettu alan kirjallisuudesta ja aikaisemmasta tutkimustiedosta.

Opinnäytetyöprosessia kun katson taaksepäin, tekisin muutamia asioita toisin. Rajaisin työn aiheen huolellisesti, ennen työn aloittamista ja suunnittelisin työn rakenteen alusta asti huolellisesti ohjaavan opettajan kanssa. Opinnäytetyöprosessin edetessä jouduin tekemään suuriakin muutoksia työn rakenteeseen ja rajauksia aiheiden suhteen. Aiheenvaihtelu tuotti vaikeuksia työn kirjoittamiseen, sillä opinnäytetyöprosessin alkaessa ymmärrys tietotekniikasta ja tekoälystä ei ollut riittävällä tasolla. Työn poikkitieteellisyyden näen kuitenkin toimineen omanaan kehittämään minua kohtia asiantuntijuutta.

Kokonaisuuteen olen varsin tyytyväinen. Koen opinnäytetyöprosessin tarjonneen minulle tilaisuuden haastaa itseni oppimaan uutta ja antaneen minulle mahdollisuuden kehittyä. Prosessin aikana kehittyneen tietotaidon ja asiantuntijuuden näen suurena käytännön etuna ajatellen tulevaisuudessa

enenevissä määrin käytettävää automaattiseen data-analyysiin pohjautuvaa hakukonemarkkinointia.

LÄHTEET

Aaltonen, Toni 2019. Mitä NLP tarkoittaa?. KPMG.

Hakupäivä 10.05.2022. <https://home.kpmg/fi/fi/home/Pinnalla/2019/09/digiajan-dynaaminen-tilintarkastus-nlp.html>

Ahmand, Norita & Dias, Kevin 2021. Validating “Data validation”. Computer.org 54 (10), 129-132. <https://doi.ieeecomputersociety.org/10.1109/MC.2021.3099045>

Arasu, Senthil; Jonathan, Seelan & Thamaraiselvan, N 2020. A machine learning-based approach to enhancing social media marketing. Computers & Electrical Engineering 86. <https://doi.org/10.1016/j.compeleceng.2020.106723>

Backlinko, 2021. Sitemaps. Backlinko. Hakupäivä 03.12.2021.

<https://backlinko.com/hub/seo/sitemaps>

Benson, Hannah 2021. How To Automatically Pull Data From A Website With Web Scraping. Scrapingrobot. Hakupäivä 22.2.2022.

<https://scrapingrobot.com/blog/how-to-automatically-pull-data-from-a-website/>

Blomster, Miikka & Koivumäki, Timo 2021. Exploring the resources, competencies, and capabilities needed for successful machine learning projects in digital marketing. Information Systems and E-Business Management 20, 123–169. 18-24. <https://doi.org/10.1007/s10257-021-00547-y>

Blomster, Miikka; Kurtti, Jonna-Riikka; Määttä, Mika & Sinisalo, Jaakko 2020. Oulun ammattikorkeakoulu. Digitaalisen markkinoinnin kasikirja mikro- ja pk-yritykselle. 119-128. <https://urn.fi/URN:ISBN:978-951-597-197-5>

Boughton, Bartlett 2005. Search Engine Marketing. St Edward’s university. 29-32. Hakupäivä 03.02.2022.

<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.4419&rep=rep1&type=pdf>

Cai, Li & Zhu, Yangyong 2015. The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. Data Science Journal, 14 (2).
<http://doi.org/10.5334/dsj-2015-002>

Davis, Harold 2006, Google: kuinka varmistat verkkonäkyvyytesi. Readme.fi Oy. Hyvinkää.

Decker, Allie 2019. Marketing campaigns. HubSpot. Hakupäivä 26.2.2022.
<https://blog.hubspot.com/marketing/marketing-campaigns>

Dhar, Vasant 2016. The Future of Artificial Intelligence. New York: Mary Ann Liebert Inc, 4 (1). <https://doi.org/10.1089/big.2016.29004.vda>

Drivas Ioannis; Sakas, Danianos; Giannakopoulos, Gergios & Kyriaki-Manesi, Daphne 2020. Big Data Analytics for Search Engine Optimization. Big data and cognitive computing, 4 (2). <https://doi.org/10.3390/bdcc4020005>

Dumitriu, Dan & Popescu, Mirona 2020. Artificial Intelligence Solutions for Digital Marketing. Procedia Manufacturing, 46. 25-26.
<https://doi.org/10.1016/j.promfg.2020.03.090>

Fadner, Ross 2004. New tracking data reveals marked disparities in paid search costs, yield. Media Post 12.10.2004. Hakupäivä 10.12.2021.
<https://www.mediapost.com/publications/article/838/new-tracking-data-reveals-marked-disparities-in-pa.html>

Fleischner, Michael 2010. PPC Made Simple: Pay Per Click Strategies for Dominating Google Adwords. Createspace Independent Publishing Platform. Lonto.

Gandomi Amir & Haider Murtaza 2015. Beyond the hype: big data concepts, methods, and analytics. International Journal of Information Management, 35 (2). 1-8. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>

Garside, Juliette 2007. Google Phobia: A Rational Fear of a Search Engine Seeking to Dominate Internet Advertising. The Sunday Telegraph, 12.04.2007.

Glick, Mark; Richards, Greg; Sapozhnikov, Margarita & Seabright, Paul 2014. How Does Ranking Affect User Choice in Online Search?. Review of Industrial Organization 45 (2), 99-119. 101-105. <https://doi.org/10.1007/s11151-014-9435-y>

Google 2022a. Benefits of online advertising and Google Ads. Google. Hakupäivä 22.3.2022. <https://support.google.com/google-ads/answer/6123875?hl=en>

Google 2022b. About ad position and Ad Rank. Google. Hakupäivä 19.4.2022. <https://support.google.com/google-ads/answer/1722122?hl=en>

Google 2022c. Traffic quality. Google. Hakupäivä 03.4.2022. <https://www.google.com/ads/adtrafficquality/>

Google 2022d. Miten Haku järjestää tietoa. Google. Hakupäivä 12.4.2022. <https://www.google.com/intl/fi/search/howsearchworks/crawling-indexing/>

Google 2022e. Google. How Search works for site owners. Hakupäivä 09.05.2022. <https://developers.google.com/search/docs/advanced/guidelines/how-search-works>

Grishikashvili, Ketty; Dibb, S & Meadows, M 2014. Investigation into Big Data Impact on Digital Marketing. Online Journal of Communication and Media Technologies, Erikoisnumero 10/2014. 1-12. <https://www.ojcmt.net/download/investigation-into-big-data-impact-on-digital-marketing-5702.pdf>

Günther, Kirsi & Hasanen, Kirsi 2022. Laadullisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. Hakupäivä 10.05.2022.

<https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/laadullisen-tutkimuksen-prosessi/tutkimuksen-kulku/>

Halsey, Lindsay 2021. SEO 101 – 4 Pillars of SEO. PathfinderSEO. Hakupäivä 24.1.2022. <https://pathfinderseo.com/blog/4-pillars-seo/>

IBM 2022. Deep Learning. IBM. Hakupäivä 10.05.2022. <https://www.ibm.com/cloud/learn/deep-learning>

Introna, Lucas, D & Nissenbaum, Helen 2000. Shaping the Web: Why the Politics of Search Engines Matters. The Information Society, 16 (3), 169-185. <https://doi.org/10.1080/01972240050133634> . Vaatii käyttöoikeuden.

Jagtap, Sudhir & Kodge, B, G 2013. Census Data Mining and Data Analysis using WEKA. 1-6. 28-29. <https://doi.org/10.48550/arXiv.1310.4647>

Kelleher, John & Tierney, Brendan 2018. Data Science. 18-32. MIT Press. Lontoo.

Lieto, Antonio; Bhatt, Mehul; Oltamari, Alessandro & Vernon, David 2017. The role of cognitive architectures in general artificial intelligence. Cognitive Systems Research 48. <https://doi.org/10.1016/j.cogsys.2017.08.003>

Linh, Nguyen 2008. Internet marketing acquisition process. Lahden ammattikorkeakoulu. Kansainvälisen liiketalouden tutkinto-ohjelma. Opinnäytetyö. 39-40. <https://urn.fi/URN:NBN:fi:amk-200812104343>

Ma, Liye & Sun, Baohong 2020. Machine learning and AI in marketing: connecting computing power to human insights. International Journal of Research in Marketing 37 (3), 481-504. 4-13. <https://doi.org/10.1016/j.ijresmar.2020.04.005>

Mailchimp 2022. Display Ads. Mailchimp. Hakupäivä 09.05.2022. <https://mailchimp.com/marketing-glossary/display-ads/>

Marsden, Sam 2021. Search Engine Indexing. Deepcrawl.com. Hakupäivä 22.4.2022. <https://www.deepcrawl.com/knowledge/technical-seo-library/search-engine-indexing/>

McCarthy, J., 1998. What is artificial intelligence. The Search for Terrestrial Intelligence, 335-360. https://doi.org/10.1142/9789812816818_0009

Mousinho, Anre 2022. What is SEO (Search Engine Optimization): The Complete Guide to Reach Top Google Results. Rock Content. Hakupäivä 25.3.2022. <https://rockcontent.com/blog/what-is-seo/>

Muller, Britney 2021. How search engines work: Crawling, indexing and ranking. Moz. Hakupäivä 03.02.2022. <https://moz.com/beginners-guide-to-seo/how-search-engines-operate>

Narayanan, Ramya 2019. Future of digital marketing in 2020. Konferenssi: Digital Marketing: Scope and opportunities. 10/2019.

Pan, Bing 2015. The power of search engine ranking for tourist destinations. Tourism Management 47, 79–87. <https://doi.org/10.1016/j.tourman.2014.08.015>

Rani, K & Rani, Sophia 2017. Impact Of Big Data Analysis in Digital Marketing. International Journal of Innovative Research and Advanced Studies 4 (1). 140-142. http://www.ijiras.com/2017/Vol_4-Issue_1/paper_29.pdf

SAS, 2019. Artificial Intelligence, What it is and why it matters. SAS Institute Inc 2019. Hakupäivä 17.3.2022. https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html

Saura, Jose 2021. Using Data Sciences in Digital Marketing: Framework, methods, and performance metrics. Journal of Innovation & Knowledge 6 (2), 92-102. <https://doi.org/10.1016/j.jik.2020.08.001>

Sherice, Jacob 2022. Which Marketing Channels Deliver the Biggest Impact on ROI?. Neil Patel. Hakupäivä 13.03.2022. <https://neilpatel.com/blog/marketing-channels-roi/>

Srivastava, Prashant 2021. SEO Automation: Understand Its Key Stages. Jivaso. Hakupäivä 4.4.2022. <https://jivaso.com/blog/seo-automation-understand-their-key-stages/>

Statcounter 2022. Search engine market share. Statcounter. Hakupäivä 1.4.2022. <https://gs.statcounter.com/search-engine-market-share>

Suskova, Darja 2021. A Case Study for improving a Travel Company's Digital Marketing activities. Arcada ammattikorkeakoulu. Kansainvälisen liiketalouden tutkinto-ohjelma. Opinnäytetyö. 19. <https://urn.fi/URN:NBN:fi:amk-2021122290468>

Tekin, Buase 2021. Natural Language in Search Engine Optimization (SEO)—How, What, When, And Why. Towards AI. Hakupäivä 10.05.2022. <https://towardsai.net/p/nlp/natural-language-in-search-engine-optimization-seo-how-what-when-and-why-b390364b5d3d>

Volpini, Andrea 2022. SEO Automation in 2022. Wordlift. Hakupäivä 28.2.2022. <https://wordlift.io/blog/en/seo-automation/>

Weglarz, Geoffrey 2004. Two Worlds of Data Unstructured and Structured. Digital Marketing Review 14 (9), 19. <https://www.proquest.com/scholarly-journals/two-worlds-data-unstructured-structured/docview/214691423/se-2?accountid=191112>

Wikipedia 2022. Semi-structured data. Wikipedia. Hakupäivä 20.3.2022.

https://en.wikipedia.org/wiki/Semi-structured_data

Wolf, Sylvia 2007. Developing a Search Engine Marketing Strategy for Casual Games for International Markets – Case Studies on European Markets.

Hochschule Furtwangen university. Tutkielma. 32. <https://webuser.hs-furtwangen.de/~heindl/ThesisSylviaWolf.pdf>