



Next Media
Hyperlocal-tuloskooste
2013



Hyperlocal

Sisällys

Lukijat mukaan tehokkaaseen sisällöntuotantoon	3
Median tuottamia tunne-elämyksiä mittaamassa	3
Tekstitys paljastaa tunteet	5
Datajournalismi on tärkeä osa mediakentän muutosta	5
Miten hyötyä avoimesta datasta?	6
Lisätty todellisuus tuo uuden ulottuvuuden kaupunkioppaaseen	8
Lukijat sanomalehden tekijöinä.....	10
Hyperlokaalista "Glokaaliksi"	13
Automaattinen uutisena haastelee sosiaalista mediaa	15
Menetelmiä kuvien sisällön ja laadun automaattiseen tunnistamiseen	18

Tämä PDF on optimoitu tablettitietokoneelle. Saat halutessasi sen näkymään esim. iPadin iBooks ohjelmassa avaamalla tiedoston iTunesiin joko käyttämällä "Arkisto > Lisää kirjastoon" -komentoa tai raahaamalla PDF:n suoraan avoimen iTunesin Kirjat-kirjastoikkunaan.

Kun iPadin seuraavan kerran synkronisoi, kirja siirtyy iPadin iBooks ohjelmaan selattavaksi (mikäli asetukseksi mahdollisavat automaattisen synkronoinnin. Jos ei, niin tulee ruksia tämä kirja iTunesista synkronoitavaksi). Huomaa, että kirja löytyy iBooksin erillisestä PDF-hyllystä.

Next Media on yksi tieto- ja viestintäteknikan strategisen huippuosaamisen keskittymän (ICT-SHOK) tutkimusohjelmista, joita koordinoi Tivit Oy. Next Media -ohjelmaa rahoittaa Teknologian ja innovaatioiden kehittämiskeskus Tekes.

Toimitus: Olli Kuusisto, VTT, Olli.Kuusisto@vtt.fi

Ulkoasu: Harri Heikkilä, Aalto ARTS

Kansi: Sari Halme, VTT

Taitto: Tapio Leppänen, VTT

Next Media-tutkimusohjelma: Eskoensio Pipatti, ohjelmajohtaja

Jakelu: Next Media -tutkimusohjelmaan osallistuvat organisaatiot.

Tilaa uutiskooste sähköpostiisi: Olli.Kuusisto@vtt.fi

Uutiskooste verkossa: www.nextmedia.fi

Next Media Tuloskooste kertoo Next Media -ohjelmassa syntyneistä tutkimustuloksista Next Media -tutkimusohjelmaan osallistuvien organisaatioiden edustajille.

Lukijat mukaan tehokkaaseen sisällöntuotantoon

Hyperlocal-tutkimuksen taustalla on tarve löytää uusia tapoja julkaista sisältöä ja luoda uudenlaisia mobiilipalveluja. Lukijoita halutaan enemmän mukaan tehokkaaseen sisällöntuotantoon, koska se rikastaa sisältöä ja sitouttaa lukijat entistä paremmin paikalliseen uutissivustoon. Sitouttaminen taas johtaa tiheämpiin ja pitempiin käynteihin uutissivustolla, mikä vuorostaan mahdollistaa mainosmyynnin kehittämisen.

Helmikuun tulosseminaarissa Santtu Parkkonen Sanoma Newsiltä esitteli Hyperlocal-työpakettin tuloksia vuodelta 2012. Vuonna 2012 lukijatoimittajien osuus sisällöntuotannosta Metro-lehdessä oli huikea. Lehdellä oli avustajina yli 30 000 lukijaa, jotka lähettivät yhteensä yli 35 000 valokuvaa. Metro-lehti teki 4 000 uutisjuttua, joista suuri osa perustui lukijoiden lähettämiin kuviin. Itse asiassa sadasta suosituimmasta jutusta 90 perustui lukijan kuviin. Hyperlocal-ohjelman saavutuksiin Sanoma News listaa myös Helsingin Sanomat -dataryhmän perustamisen muutaman kuukauden pilottiprojektin jälkeen.

Vuoden tärkeimpinä saavutuksina Parkkonen nosti esiin seuraavat tulokset: ohjelmassa kerättiin suomalaista slangisanastoa crowdsourcing-menetelmällä, löydettiin kolme avoimeen dataan perustuvaa liiketoimintamallia mediayhtiöiden tarpeisiin, tehtiin lisätyn todellisuuden (Augmented Reality, AR) teknologiaan perustuva mobiili kaupunkiopas ja testattiin sitä, luotiin lukijatoimittajille työkaluja (teejuttu.fi) ja testattiin niitä sekä koulutettiin lukijareporttereita valokuvauksen ja kirjoittamisen saralla. Näihin tuloksiin ja osahankkeisiin tutustutaan tarkemmin seuraavissa kappaleissa.

→ *Santtu Parkkosen esitys*

→ *SanttuParkkosen kalvot*

MEDIAN TUOTTAMIA TUNNE-ELÄMYKSIÄ MITTAAMASSA

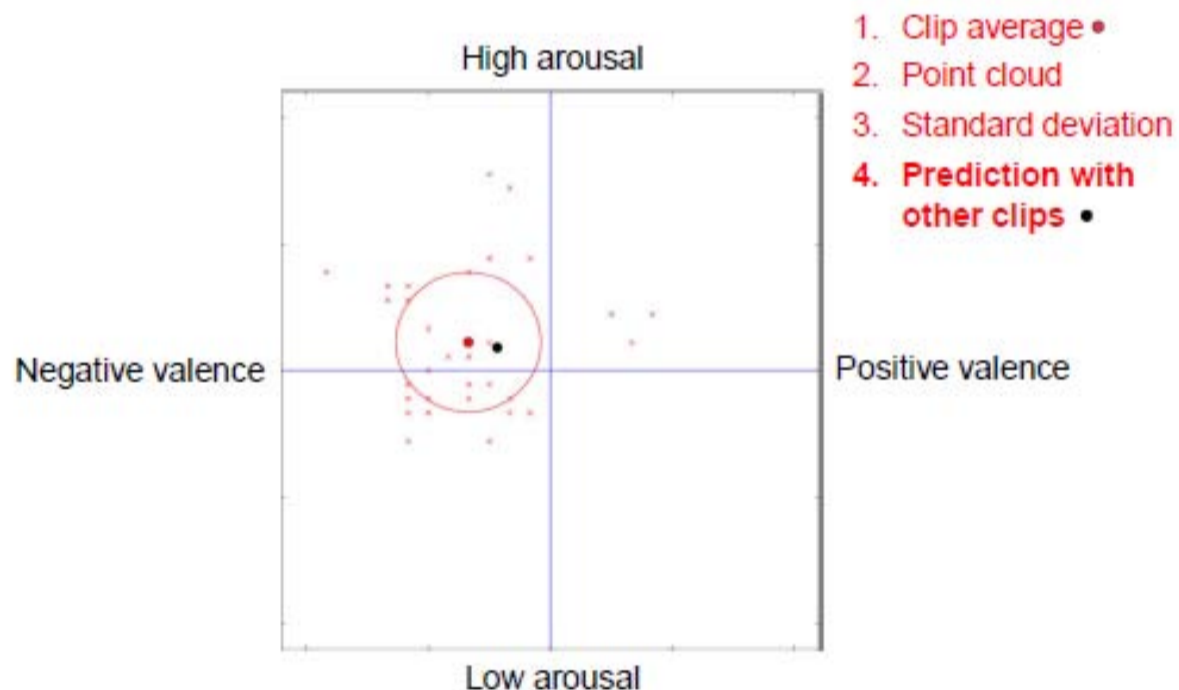
Affective Facets of Multimedia Content -hankkeessa Aalto-yliopiston Perustieteiden korkeakoulun tutkijat halusivat selvittää, millaisen tunnevaikutuksen elokuva herättää katsojissa. Erityisenä tutkimuskohteena oli äänen ja visuaalisen videon yhteisvaikutus. Hankkeen tärkeimpänä tavoitteena oli parantaa audiovisuaalisen aineiston hakua, jota voitaisiin hyödyntää erilaisissa videomateriaalia tarjoavissa sisältöpalveluissa.

Yhteensä 74 käyttäjälle näytettiin lyhyitä pätkiä 14 elokuvasta, jonka jälkeen heitä pyydettiin arvioimaan tunnetilojaan kirjallisesti. Testidata sijoitettiin koordinaatistoon, jonka akselit olivat **valence** (epämukava – mukava) sekä **arousal** (nukuttava – korkea virittyneisyys).

Käyttäjien omien arvioiden mukaiset tunneilmaisut korreloitiin videoista mitattuihin ja laskettuihin piirteisiin ja löydettiin yhteyksiä elokuvan tyyllisten teki-
jöiden ja tunnetilojen välillä. Menetelmän avulla voidaan esimerkiksi luokitella ja kuvailla elokuvia automaattisesti ja sisällyttää tieto erilaisiin elokuvia tarjoaviin palveluihin luokittelukeinoina.

Tuloksista voitiin huomata, että esimerkiksi keskimääräinen väri kuvaili hyvin sitä, miten rento tunnelma kohtauksessa on, ja dominoiva väri taas kuvasi elokuvan masentuneisuusastetta. Laskennallisella menetelmällä päästiin kohtalaiseen tarkkuuteen, kun sitä verrattiin käyttäjien antamiin arvioihin. Jatkosuunnitelmissa on laskennallisten menetelmien kehittäminen edelleen sekä aineiston laajentaminen.

Mood prediction



Kuva 1. Esimerkki mielialan ennustamisen onnistumisesta muiden videopätkien perusteella (verrokkina käyttäjien arviot).

→ *Jussi Tarvaisen esitys*

→ *Jussi Tarvaisen kalvot*

- Raportti multimediasisältöjen tunnevaikutuksista *Affective approaches to multimedia content*
- Raportti testisarjasta (dataset) multimediasisältöjen tunnevaikutusten tutkimuksiin *Test collection for affective video studies*
- Tieteellinen artikkeli multimediasisältöjen tunnevaikutuksista *Content-based Prediction of Movie Style, Aesthetics and Affect*

TEKSTITYYS PALJASTAA TUNTEET

Oulun yliopistossa analysoitiin videotekstityksiä tunneanalyysin perusteella, eli pyrittiin laskennallisesti selvittämään kirjoittajan asennetta tai tekstiin sisältyvää mielipidettä. Analyysin eräänä tavoitteena on tunnistaa tekstin sisältämä asennepolariteetti, eli se, onko kirjoitus positiivinen, negatiivinen vai neutraali.

Tutkimuksessa käytettiin mielipideanalyysiin kehitettyä SentiWordNet 3.0 tunnesanastoa, joka sisältää lähes 118 000 sanakuvausta ja 150 000 englanninkielistä sanaa. Tarkoituksena oli testata, voisiko sanastoa käyttää elokuvien ja tv-ohjelmien tekstitysdatan louhinnassa, jolloin videoista voitaisiin löytää automaattisesti sellaiset kohdat, joissa voimakkaimmin ilmaistaan mielialoja, asenteita ja mielipiteitä. Suomenkieliset tekstitykset käännettiin Google Translatella englanniksi, koska SentiWordNet on englanninkielinen. Tekstitysegmentit analysoitiin yksittäin ja niille laskettiin kokonaissentimenttiarvot yksittäisten sanojen arvojen perusteella. Positiivisten ja negatiivisten arvojen erotus ja summa laskettiin kullekin segmentille. Menetelmällä voidaan korostaa tekstityksestä löytyviä asenteellisia kohtia, joissa käytetään voimakasta ilmaisua. Nämä kohdat ovat yleensä myös ohjelman kerronnan kannalta mielenkiintoisia kohtauksia. Menetelmä toimi varsin hyvin kohtausten löytämisessä. Tavoitteena oli helpottaa multimediadatan semanttista prosessointia ja organisointia. Tätä voi testata esimerkiksi Oulun yliopiston lähetettyjen ohjelmien live seuranta palvelu Mediaseinän (<http://kuukkelitv.fi/mediaseina>) uutena ominaisuutena.

DATAJOURNALISMI ON TÄRKEÄ OSA MEDIAKENTÄN MUUTOSTA

Tampereen yliopisto tutki suomalaista datajournalismia ja sen tekijöiden näkemyksiä. Tutkimuksen toteuttamiseksi Auli Harju haastatteli HS Open sovellustapahtumaan osallistuneita ja teki nettikyselyn heille. HS Openissa journalistista, ohjelmoijasta ja graafikosta koostuvien kolmen hengen ryhmien on tarkoi-

tus saada aikaan kiinnostava sovellus Helsingin Sanomien tarjoamasta datasta. Tilaisuuteen osallistuttiin esimerkiksi työkalujen oppimisen, datajournalismin ymmärtämisen tai verkostoitumisen takia. Toisaalta motiivina saattoi olla myös työnantajan määräys tai oma harrastustoiminta.

HS Open -osallistajat toivoivat lisää avointa dataa sekä julkiselta että yksityiseltä puolelta. Datajournalismia pidettiin tärkeänä osana mediakentän muutosta, eikä pelkästään Helsingin Sanomien kaltaisten isojen tekijöiden mahdollisuutena. Osallistujien keskuudessa pidettiin tärkeänä, että myös pienemmät media-alan toimijat hyödyntäisivät avointa dataa aktiivisesti.

→ Raportti datajournalismi-yhteistyöstä toimituksen ja kehittäjien kesken
Experiences from Collaboration Between Newsroom and Developers

MITEN HYÖTYÄ AVOIMESTA DATASTA?

Avoin data (open data) on nimensä mukaisesti avointa, eli ilmaista ja lisenssi-vapaata joko julkishallinnon tai yrityksen julkaisemaa dataa. Datan avoin saata-vuus ja sen käsittely muuttavat mediakenttää merkittävästi.

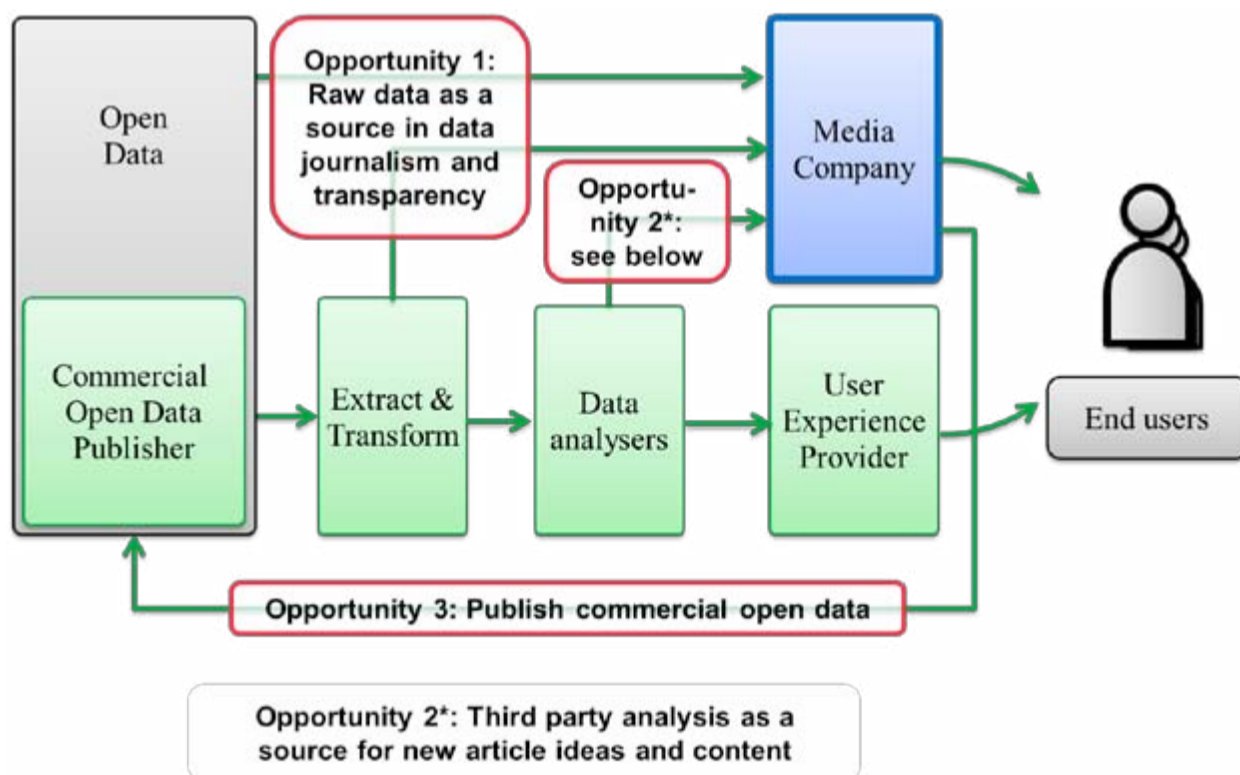
Avoimen datan käyttö Suomessa on kasvussa. Tämä näkyy muun muassa avoi-men datan innovaatiokilpailu Apps4Finlandin osallistujamäärissä: kun vuonna 2010 kilpailutöitä jätettiin 30 kappaletta, niin lukumäärä oli jo 140 vuonna 2011. Esimerkiksi Maanmittauslaitoksen karttadatan avautumista toukokuussa 2012 voidaan pitää merkittävänä tekona avoimen datan saralla.

Aalto-yliopiston Kauppakorkeakoulussa tutkittiin sitä, millaisilla liiketoimin-tamalleilla yritykset voivat toimia avoimen datan parissa, yritysten kohtaamia haasteita ja yhteisiä menestystekijöitä. Tutkimusta tehtiin haastattelemalla yri-tyksiä, jotka hyödyntävät avointa dataa toiminnassaan. Yhteensä haastatteluja tehtiin 14 kappaletta keväällä 2012.

Tutkimuksen tavoitteena oli tarkoitus tunnistaa mediakentän muutoksia media-yhtiön näkökulmasta, kun käytössä on päivittäin myös avointa dataa. Tässä tilanteessa on tärkeä tunnistaa, miten avointa dataa voi hyödyntää taloudelli-sesti ja miten voisi olla mahdollista luoda sen ympärille tuottavaa liiketoimintaa. Pitemmän aikavälin tavoitteena on kehittää datajournalismin tuotantomalleja sekä uudenlaisia avoimeen dataan perustuvia palveluja. Tämän tavoitteen täyttä-miseksi on olennaisen tärkeää ymmärtää kehittäjäyhteisön motivaatioita ja tar-peita.

Avointa dataa käytetään yrityksissä erittäin monipuolisesti ja yhdessä kaupallisten lähteiden kanssa. Data sinänsä ei kuitenkaan ole arvokasta ilman sovellusta, joka jalostaa ja hyödyntää sitä. Liiketoiminta on toistaiseksi pienimuotoista ja useat yritykset etsivät vielä parasta liiketoimintamallia. Kasvua avoimen datan parissa pidetään hyvin todennäköisenä. Liiketoimintamalleista yrityksillä oli käytössä muun muassa freemium (palvelun käyttö on maksutonta, ja tarkoituksena on sitouttaa käyttäjä - tarjolla on lisäksi mahdollisesti maksullisia lisäpalveluita), bannerimainokset tai vuosimaksu. Kaikkien sovellusten tekijöiden ei ollut tarkoituskaan luoda liiketoimintaa, vaan sovelluskehitystä tehtiin puhtaasti harrastuspohjalta.

Yritysten haasteina on ollut muun muassa julkisen hallinnon tuntemuksen puute, ohjelmoinnissa käytettyjen rajapintojen yllättävät muutokset sekä kannustavien tukirahoitusmuotojen puute. Avoimen datan toimijakenttä hahmoteltiin jakamalla toimijat kolmeen luokkaan: **Data analyzers** (rakentavat uutta tietoa kokoamalla avointa dataa eri lähteistä ja havainnollistamalla sitä uudella tavalla), **Extract & transform** (hakevat ja käsittelevät raakadataa ja muuntavat sen muille toimijoille helpommin jatkokäsiteltävään muotoon) ja **Mobile & Web UI** (tarjoavat uuden ja helpomman käyttöliittymän raakadatan tarkasteluun ja käyttöön, esim. yhdistelemällä eri lähteiden dataa ja visualisoimalla sitä karttapohjan päällä).



Kuva 2. Kolme vaihtoehtoista liiketoimintamahdollisuutta.

Avoimen datan liiketoimintamalleista löytyi kolme mahdollisuutta mediataloille. Ensimmäisenä vaihtoehtona on juttujen kirjoittaminen raakadatasta. Vaikka tämä vaatiikin osaamista ja resursseja, on lopputuloksena mahdollisesti erittäin kiinnostavia juttuja. Kun lähteenä on avoin raakadata, sen avaaminen luo luotamusta ja läpinäkyvyyttä itse juttuun. Toisena vaihtoehtona on kolmannen osapuolen laatimien analyysien käyttö juttulähteenä. Tässä vaihtoehdossa kuluu vähemmän resursseja, ja datan käsittelystä aiheutuvien virheiden riski laskee. Toisaalta data-analyysille ei Suomessa ole vielä kehittyntä markkinaa, joten saatavuus on ongelma. Kolmas ja kenties kiinnostavin vaihtoehto on kaupallisen datan avaaminen, jota voitaisiin testata tulevaisuudessa esimerkiksi Sanomien tapahtumatiedoilla.

- *Matti Rossin ja Tomi Kinnarin esitys*
- *Matti Rossin ja Tomi Kinnarin kalvot*
- Raportti avoimen datan yritysten uusista arvoverkkoprofiileista *New Value Network Profiles for Open Data*
- Raportti yksityiskohtaisemmista suomalaisten avoimen datan yritysten toimintaprofiileista avoimen datan ekosysteemissä *Detailed Sample Profiles of Finnish Open Data Companies*

LISÄTTY TODELLISUUS TUO UUDEN ULOTTUVUUDEN KAUPUNKIOPPAASEEN

Mediakentän muuttuessa mainontamuodot joutuvat muuttumaan sen mukana. Uudenlaisia mainoskanavia ja -vaihtoehtoja aukeaa teknologian kehittymisen myötä. Eräs uudehko mahdollisuus on lisätyn todellisuuden (Augmented Reality, AR) rajapinta, jonka avulla voidaan luoda linkki digitaalisen ja oikean maailman välille esittämällä digitaalista sisältöä tosielämän kohteiden yhteydessä, esimerkiksi kuvan päällä.

Augmented Reality -tutkimuksessa tehtiin ensin selvitys jo yleisessä käytössä olevista AR-sovelluksista ja -alustoista. Selvityksen perusteella lisätyn todellisuuden käytön uskotaan olevan jo ihan tuttua Suomessa muutaman vuoden kuluttua.

VTT:llä tehtiin lisätyn todellisuuteen pohjautuva iOS-pilottisovellus, Kaupunkiopas. Sen tarkoitus on auttaa käyttäjiä löytämään kännykän kameranäky-

mässä lisätyn todellisuuden sovelluksen avulla tapahtumia, palveluita ja tarjouksia. Sovelluksen avulla voi myös paikantaa uutistapahtumia. Sisältö muodostuu Omakaupunki.fi-palvelun uutisista ja tapahtumista, mutta myös käyttäjien itsensä lisäämistä viesteistä ja kuvista.



Kaupunkioppaan karttanäkymä. Tapahtumat, uutiset ja käyttäjien kommentit näkyvät erivärisinä merkintöinä kartalla.



Kaupunkioppaan AR-näkymä, jossa lisätyn todellisuuden elementit näkyvät sinisinä laatikoina.

Kuva 3. Kaupunkioppas-sovelluksen erilaiset näkymät.

Kaupunkioppaan käyttäjätesteillä selvitettiin yleisesti asenteita AR-sovelluksia kohtaan ja kerättiin yksityiskohtaista palautetta sovelluksen ideasta ja toiminnoista. Loppuvuodesta sovellukselle tehtiin käytettävyystestit, joihin osallistui 11 käyttäjää. Käyttäjien tuli suorittaa kolme erilaista tehtävää, jotka saattoivat liittyä esimerkiksi tapahtuman löytämiseen palvelusta, viestin lisäämiseen tai uutisen kommentointiin. Käytettävyystestin jälkeen käyttäjät haastateltiin.

Karttapohjaisen sovelluksen käytössä tutkijat tunnistivat neljä käyttökokemustavoitetta, joita olivat hallinnan, onnistumisen ja yhteenkuuluvuuden tunteet sekä viihtyminen. Näihin vaikuttavia tekijöitä ovat muun muassa käytön ja käyttöönoton helppous, luotettava ja tarkka data, sovelluksen hyödyllisyys, mahdollisuus

sosiaaliseen verkottumiseen, mahdollisuus tyhjen hetkien täyttämiseen ja positiiviset yllätykset. Käyttöliittymän miellyttävyys vaikuttaa myös viihtymiseen ja kokemukseen käytön miellyttävyydestä.

Sovelluksesta havaittiin käyttäjätesteissä myös kehityskohteita. Koska näytettävää dataa oli paljon, oli sitä välillä hankalaa mahduttaa selkeässä muodossa puhelimen pienelle näytölle. Käyttäjät kaipasivat lisäksi hakutoimintoa, navigointia, kohteen lukitsemista ja tietoa kohteen etäisyydestä. Viestien käytössä oli hieman epäselvyyttä siitä, näkyykö viestin kohdalla lähettäjän nimi ja mihin viesti sijoittuu kartalla.

Tavoitteena oli sovelluskehityksen jälkeen päästä analysoimaan pilotin ja kolmannen osapuolen sovellusten ympärille rakentuvia palveluekosysteemejä ja liiketoimintamalleja. Analyysin perusteella voidaan kerätä tietoa siitä, millaisia mahdollisuuksia lisätty ja yhdistetty todellisuus tuovat media-alalla sekä sisältötuotannon, mobiilimainonnan että palvelujen saralla.

- *Tatu Harviaisen ja Maiju Aikalan esitys*
- *Tatu Harviaisen ja Maiju Aikalan kalvot*
- Raportti mobiilin kaupunkioppaan käytettävyydestä tuloksista *Requirement specifications for the pilot Mixed Reality application*
- Kalvosarja hyperlokaalin sisällön yhdistetyn todellisuuden prototyypistä *Pilot Mixed Reality application with hyperlocal content*
- Raportti mobiilin kaupunkioppaan käytettävyydestä tuloksista *Analysis of usability testing of Mobile Augmented Reality City Guide*

LUKIJAT SANOMALEHDEN TEKIJÖINÄ

Sanomalehtien levikit ovat olleet laskussa ympäri maailmaa, kun kuluttajien mediaeuroista ja ajasta on kilpailemassa yhä kasvava määrä erilaisia ajanvietteitä. Tutkimusten mukaan lukijat kokevat usein, että perinteinen talous- ja politiikkauutisten esitystapa on vaikeaselkoinen ja etäällä arkielämästä. Tästä syystä sanomalehdet ovat halunneet keskittää huomiotaan enemmän siihen, mitä sanomalehteen painetut tai verkossa julkaistut uutiset todella merkitsevät ihmisille. Verkkokeskustelut ja lukijoiden tuottama sisältö ovat saamassa yhä suurempaa roolia. Tämän tutkimuksen tavoitteena oli lisätä käyttäjien osallistumista tehokkaaseen sisältötuotantoon, koska sen ajatellaan sitouttavan käyttäjiä paremmin uutissivustoon. Sitouttaminen vuorostaan palvelee mainosmyynnin tarpeita.

Aalto-yliopiston Taiteiden ja suunnittelun korkeakoulun tutkijat tekivät kokeilun Omakaupunki.fi:n lukijoiden keskuudessa marraskuun 2011 ja helmikuun 2012 välillä. Kokeilussa lukijoita pyydettiin osallistumaan uutisten tekemiseen kertomalla omia kokemuksiaan erilaisista aiheista. Ensimmäisessä vaiheessa lukijoille esitettiin 12 päivänä kysymys jostain lukijakeskustelusta nousseesta aiheesta. Tämän jakson jälkeenkin toimitus jatkoi lukijoiden osallistamista pari kertaa viikossa kahden kuukauden ajan. Aiheet poikivat useita artikkeleita Vartti-lehteen.

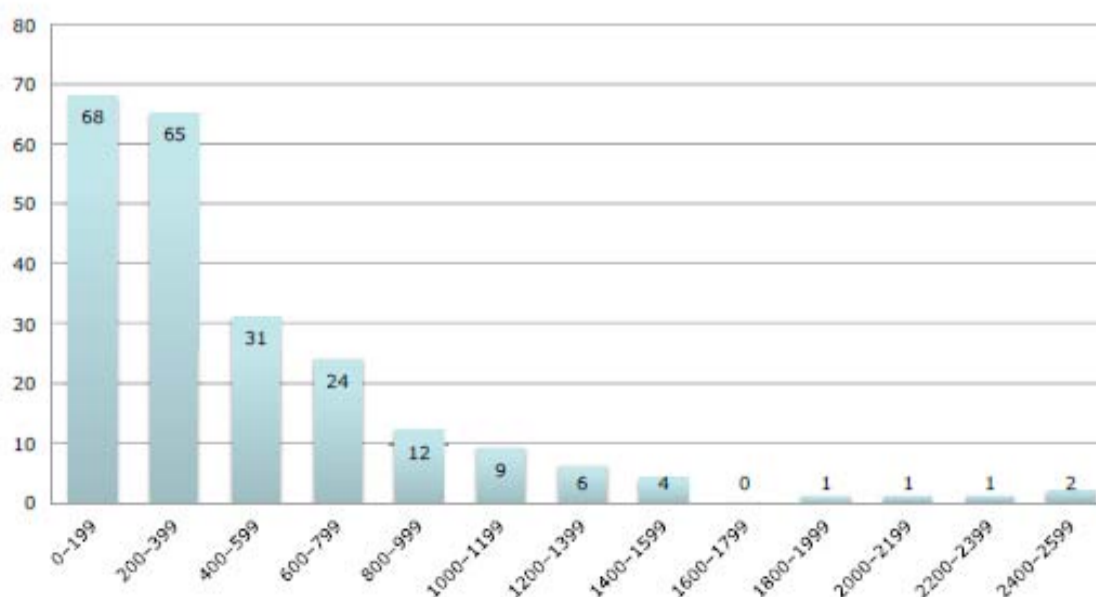
Tutkimuksessa tarkkailtiin myös Kaupunkilehtien toimituksen työskentelyä ja toimintatapojen muutosta. Kun Vartti-lehti uudistui toukokuussa, oli tavoitteena hyödyntää sisällöntuotannossa aiempaakin enemmän lukijoiden osallistamista. Kajsia Hytönen haastatteli toimituksen kunnallistoimittajia ja uutispäälliköitä. Uudistuneen lehden koettiin tuoneen enemmän työtä toimitukselle muun muassa lukijamateriaalin läpikäynnin muodossa.

Lisäksi hankkeessa tarkasteltiin syyskuussa käynnistyneeseen teejuttu.fi-palveluun tulleen materiaalin määrää ja laatua. Etenkin lukijoiden kolumnien taso oli yllättänyt toimituksen positiivisesti. Ylipäätään palvelun kautta saadun materiaalin määrä oli todella suuri, vaikka sitä ei ollut markkinoitu vielä juuri lainkaan. Suurin osa lehden verkkoversion klikatuimmista jutuista oli saanut alkunsa lukijan kuvasta tai vihjeestä. Eniten juttuja tuli tapahtumiin ja vapaa-aikaan liittyen, usein juuri viikonlopun jälkeen. Suurin osa lukijoiden kirjoittamista jutuista oli hyvin lyhyitä ja niissä oli mukana yksi kuva. Tärkeimpänä motivoivana tekijänä tuntui olevan usein jokin havainto, joka haluttiin jakaa muiden kanssa. Myös erilaisiin epäkohtiin puuttuminen johti usein lukijajutun tekoon.

Length of the articles

(marks without spaces)

(N=224)

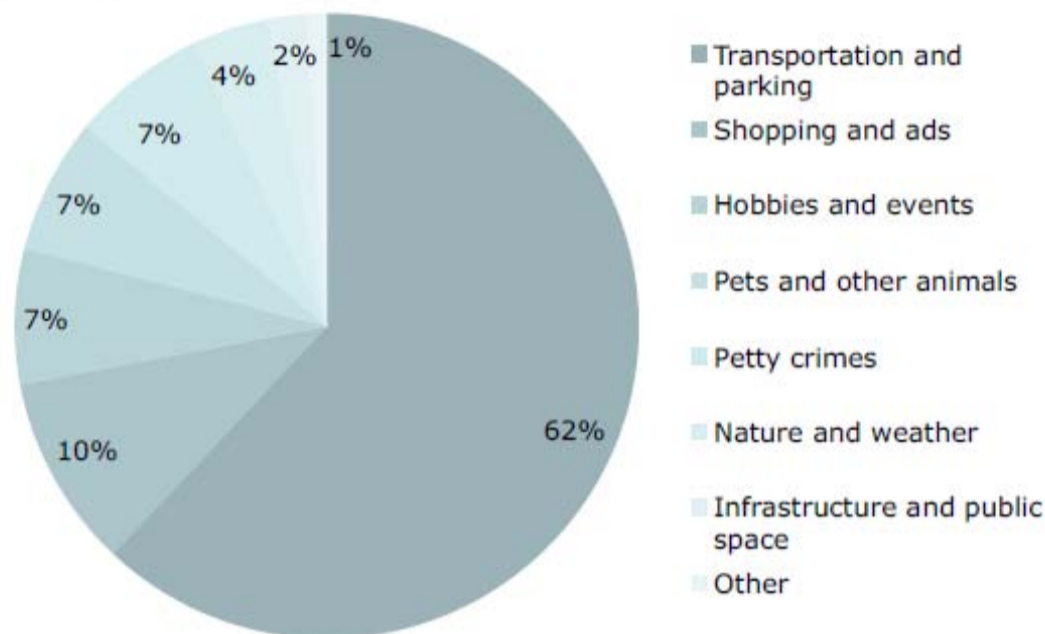


Kuva 4. Lukijajuttujen pituus.

Osana analyysiä tutkijat kävivät myös läpi sata eniten klikattua ja sata eniten kommentoitua kuvaa. Suuri osa kuvista oli huonolaatuisia, ja ne oli otettu todella kaukaa tai ikkunan läpi. Kuvan laatu ei kuitenkaan vaikuttanut klikkausten määrään, vaan sisältö oli paljon ratkaisevampi tekijä. Klikatuimmat ja kommentoiduimmat jutut koskivat liikennettä ja parkkeeraamista.

Most clicked photos

(Jan. – Oct. 2012)
(N=99)



Kuva 5. Klikatuimmat lukijakuvat kategorioittain.

Tutkijat olivat vertailleet vuoden 2012 lehteä vuoden 2010 Varttiin. Lukijakuvien käyttö oli lisääntynyt huomattavasti etenkin uutisjuttujen yhteydessä, vaikka pääkuvina käytettiin yhä paljon ammattilaisten ottamia valokuvia. Samaan aikaan kuvista maksettavat palkkiot olivat laskeneet ja kuvien koko oli pienentynyt. Lukijan kuvat olivat tuoneet aiheet lähemmäs lukijoiden elämää. Nämä jutut, esimerkiksi väärinpysäköintiin liittyvät kuvat, olivatkin kommentoiduimpien joukossa.

Tampereen Teknillisessä Yliopistossa tehdyn selvityksen mukaan maailmalla lukijoiden osallistamiseen käytetään muun muassa Twitterin ja joidenkin puhe- tai tablettisovellusten kautta tuotettua sisältöä. Lukijamateriaalin laadun parantamiseen soveltuviksi todettuja keinoja olivat muun muassa moderointi, tutoriaalit ja ohjeet, palkitseminen ja ranking-systeemit.

Lukijareporttereiden parissa tehtiin käyttäjätesti, jossa lukijareportterit jaettiin kahteen ryhmään. Toisessa ryhmässä oli aktiivisia kuvaajia (60 henkilöä) ja toisessa ihmisiä, joilla ei ole erityistä valokuvaustausta (104 henkilöä). Ensimmäiselle ryhmälle lähetettiin tehtäviä tekstiviestillä ja jälkimmäiselle Scoopshot-pal-

velun avulla. Tavoitteena oli selvittää esimerkiksi sitä, mikä motivoi lukijoita osallistumaan sekä löytää uusia osallistumisen muotoja. Scoopshot-ryhmän jäsenten mielestä kokeilu oli mielenkiintoinen ja kuvaustehtäviä toivottiin useammin kuin kerran viikossa. Osallistujat ilmoittivat olevansa valmiita käyttämään aikaa ja vaivaa tehtävien tekemiseen, mikä ei toisaalta näkynyt aktiivisuutena kokeilun aikana. Toisaalta taas Aalto ARTS:n haastattelututkimuksessa huomattiin, ettei toimitus ollut järin tyytyväinen Scoopshot-kokeiluun, sillä kuvia tuli liian vähän, niitä ei voinut arkistoida ja Scoopshot oli kallis käyttää. Ainoana positiivisena puolena toimituksen kannalta pidettiin sen helppokäyttöistä maksujärjestelmää.

Samalla tehtiin tutkimus, jossa pyrittiin selvittämään sitä, miten lukijamateriaalin laatua voitaisiin parantaa. Tutkimus tehtiin haastattelemalla 20 henkilöä. Laadun parantamisen kannalta tärkeänä pidettiin ja arvostettiin toimituksen erityisesti sanallista palautetta ja asiallista kritiikkiä. Lisäksi toimitukselta toivottiin lisätietoa siitä, miten lukijoiden tuottamaa materiaalia käytetään. Sen sijaan kilpailullisuutta lukijoiden kesken ei arvostettu tai kaivattu.

- *Merja Helteen ja Maria Ruuskan esitys*
- *Merja Helteen ja Maria Ruuskan kalvot*
- *Heli Väättäjän kalvot*
- Raportti lukijoiden tuottaman sisällön hallinnasta *Facilitating Community Management*
- Raportti uutisisältöjen joukkoistamisen kenttätestauksesta *Mobile Crowdsourcing of News Content – Findings from a Field Trial with Scoopshot*
- Raportti käyttäjälähtöisen sisällön laadun kehittämistä online-palautteen avulla *Improving the Quality of User-Generated Content by Online Feedback*

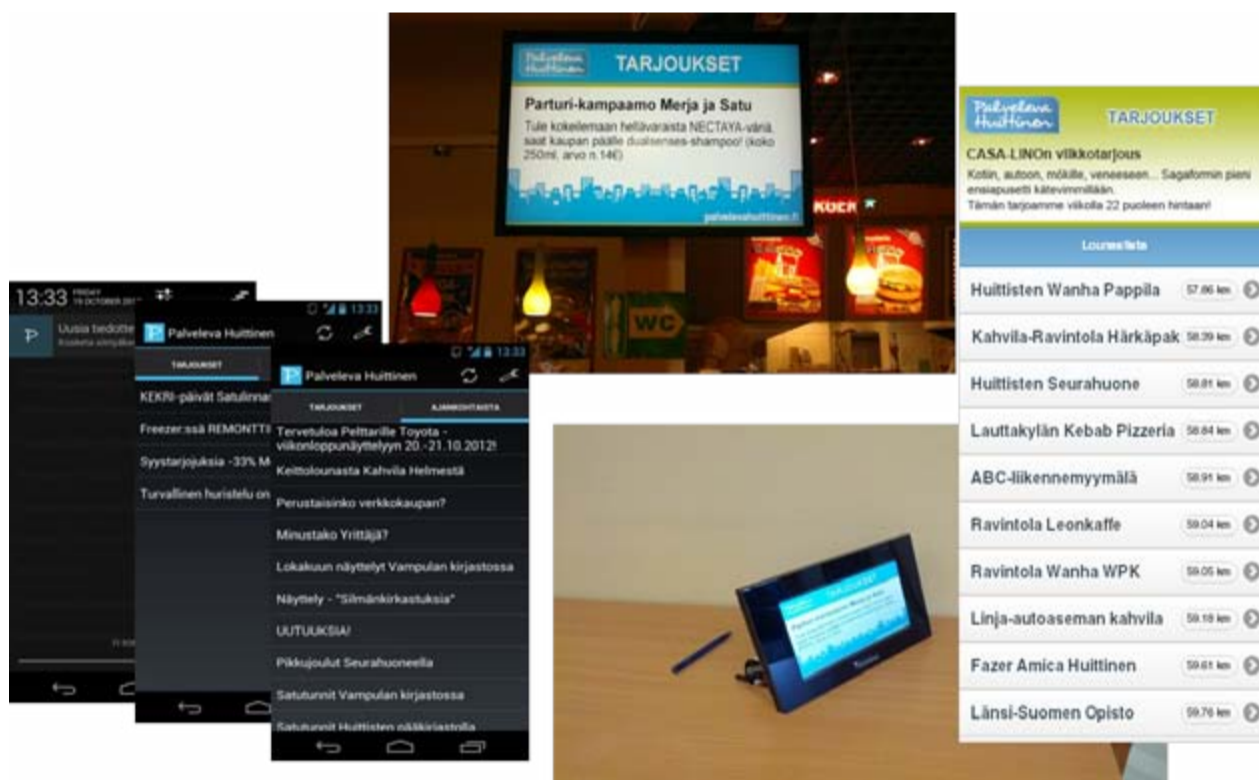
Hyperlokaalista "Glokaaliksi"

Upscaling Hyperlocal to Glocal -hankkeessa tutkimuksen tavoitteena on kehittää menetelmiä, teknologiaa ja liiketoimintamalleja, joiden avulla voitaisiin laajentaa hyperlokaalin sisällön käyttöä suuremmalle maantieteelliselle alueelle muodostamalla palvelu- ja yhteistyöverkkoja paikallisten toimijoiden tuottamien sisältöjen välille. Hyperlokaalien palveluiden takana on tarve aktivoida ja luoda yhteishenkeä paikallisen yhteisön keskuudessa. Samalla saadaan lisänä-

kyvyttä paikallisille toimijoille, esimerkiksi mainostajille. Tutkimus tähtää sekä hyperlokaalien palveluiden toimittamiseen soveltuvien teknologioiden kehittämiseen sekä käyttäjiä osallistavien ja motivoivien tekijöiden tunnistamiseen ja käyttämiseen.

Hankkeessa jatkettiin Palveleva Huittinen -portaalin konseptin kehittämistä selvittämällä mitä uusia elementtejä siihen olisi tarvetta lisätä. Huittisissa oli aloitettu keskusteluja yhteistyön parantamiseksi erilaisten paikallisten medioiden kanssa. Palvelusta oli julkaistu mobiilisovellus, joka sisälsi muun muassa ajankohtaisia ilmoituksia sekä mukana olevat yritykset jaoteltuna kategorioihin. Pilottiversio oli saanut paljon positiivista palautetta, vaikka sekä yritys- että käyttäjäpuolella toivottiin lisää aktiivisuutta. Portaalin mainosmyynti oli nousut 20 % tammi-helmikuussa 2012. Sovellusta jatkokehitettiin edelleen ja siihen lisättiin muun muassa toiminnallisuus, joka ilmoittaa automaattisesti lähistöllä olevista tarjouksista ja tapahtumista.

Yrityspartnereiden keskuudessa tehtiin kysely, jossa pyydettiin palautetta Palveleva Huittinen palvelun kehittämistä varten. Kyselyssä selvisi, että paikalliset mediatilat etsivät jatkuvasti uusia mahdollisuuksia esimerkiksi paikkatietoa hyödyntävistä mobiilisovelluksista.



Kuva 6. Palveleva Huittinen -palvelun käyttöesimerkkejä.

Myös Kaleva työskenteli hankkeessa paikallisen mobiiliproton parissa. Palvelu, nimeltään Plink, sisältää paikallisia uutisia, tapahtumatietoja ja mainoksia.

Tiedot tapahtumista ja uutiset saadaan Kaleva.fi:stä. Käyttäjät voivat niin ikään lisätä sovellukseen kuvia ja kommentteja. Sovellus tunnistaa käyttäjän sijainnin ja näyttää sen perusteella sisältöä käyttäjän lähiympäristöstä.



Kuva 7. Plink-sovellus käytössä.

Sekä Palveleva Huittinen- että Plink-prototyyppiä testattiin käyttäjillä. Plinkin käyttäjätesteihin osallistui 21 ihmistä, joiden mukaan sovelluksen parhaita puolia olivat sen helppokäyttöisyys, paikkatiedon hyödyntäminen sekä mainokset ja tarjoukset. Parannuskohteita löydettiin etenkin käyttöliittymästä, ja latausajat koettiin liian pitkiksi. Osa käyttäjistä oli myös haluttomia lisäämään sovellukseen omia viestejä eli plinkkejä. Sosiaalista aspektia, esimerkiksi omien kavereiden sijaintitietoa, toivottiin jo olemassa olevien ominaisuuksien rinnalle. Lisäsisältöä kaivattiin myös uutisista ja liikenteestä. Mainoksilta odotettiin ajan-kohtaisuutta, esimerkiksi kaivattiin houkuttelevia päivän tarjouksia.

- *Upscaling hyperlocal to glocal -esitys*
- *Upsaling hyperlocal to glocal -kalvot*
- Glocal-mallin vaatimusmäärittely *Description and requirement of Glocal Model*
- Raportti yhteistoiminnallisten hyperlokaalien palvelukonseptien evaluoinnista *Evaluation report for co-operative hyperlocal service concepts*

AUTOMAATTINEN UUTISNENÄ HAISTELEE SOSIAALISTA MEDIAA

Sosiaalinen media ja muu avoimesti verkossa liikkuva data on tuonut mittavia määriä uutta tietoa kaikkien saataville. Toimittajat tarvitsevat tehokkaita auto-

maattisia työkaluja uudenlaisten uutislähteiden sisällön löytämiseen ja analysoimiseen. Olennaista on tiedon huolellinen tarkastelu ja laadunvarmistus, sillä Internetistä löytyy paljon turhaa, virheellistä ja harhaanjohtavaa tietoa. Tässä tutkimuksessa yhdisteltiin useita kehittyneitä laskennallisia menetelmiä, joiden avulla pyrittiin vastaamaan tietotulvien asettamiin haasteisiin. Tutkimuksen tavoitteena on kehittää menetelmiä, joita voitaisiin hyödyntää automaattisen uutisneen kehittämässä. Työvaiheisiin kuuluu uutistiedonlouhinnan menetelmien ja algoritmien kehittäminen sekä analyysi-, arviointi ja luokittelumenetelmien parantaminen. Toisena haasteena on kehittää työkaluja ja menetelmiä, jotta ensimmäisessä vaiheessa kehitettyjä työkaluja voitaisiin käyttää suomalaisessa sosiaalisessa mediassa esiintyvien murre- ja puhekielisten tekstien louhintaan. Työkaluille on tarkoitus kehittää myös soveltuvat graafiset käyttöliittymät.

Helsingin yliopiston osiossa oli edistytty puhekielisen ilmaisun ymmärtämisessä. Nettikielen sanastoa koottiin sosiaalisen median käyttäjiltä crowdsourcing-menetelmällä. Slangisanasofta julkaistiin Omakaupunki.fi-palvelussa ja löytyy myös osoitteesta <http://blackhole.astro.helsinki.fi/swnr/>

Mikä ihmeen snagari!

Keskustelupalstat, Facebook ja Twitter viisevät sanoja, joita on aivan turha sanakirjasta etsiä. Useimmat ihmiset ymmärtävät ne, mutta kone ei. Tekesin rahoittamassa Next Media - tutkimushankkeessa etsimme näille sanoille selityksiä. Opetamme koneelle, että esimerkiksi Stadi ja Hki tarkoittavat samaa asiaa kuin Helsinki.

Auta meitä luomaan slang- ja puhekielen sanakirja. Lue esimerkkilause ja kerro mitä siitä eroteltu sana tarkoittaa.

SDPsälä SDP,sä Jutta, Junger, Guzeniina ja erityisesti "mungopaviaani" romahdutti koko puolueen kannatuksen olemattomiin! :(Lukekaa Jussi Parviaisen blogista, mitä hän sanoo Lipposesta

VALITTU SANA:
SDPsä

PUHEKIELI	YLEISKIELI
Alkuperäinen: SDPsä Esim. Stadin	Suora käännös: Esim. Helsingin
Perusmuoto: Esim. Stadi	Perusmuotoinen käännös: Esim. Helsinki

Sanaluokka:
 Ei tietoa Substantiivi Verbi Adjektiivi Adverbi

Nimetty luokka:
 Ei tietoa Henkilo Organisaatio Paikka -roskasana

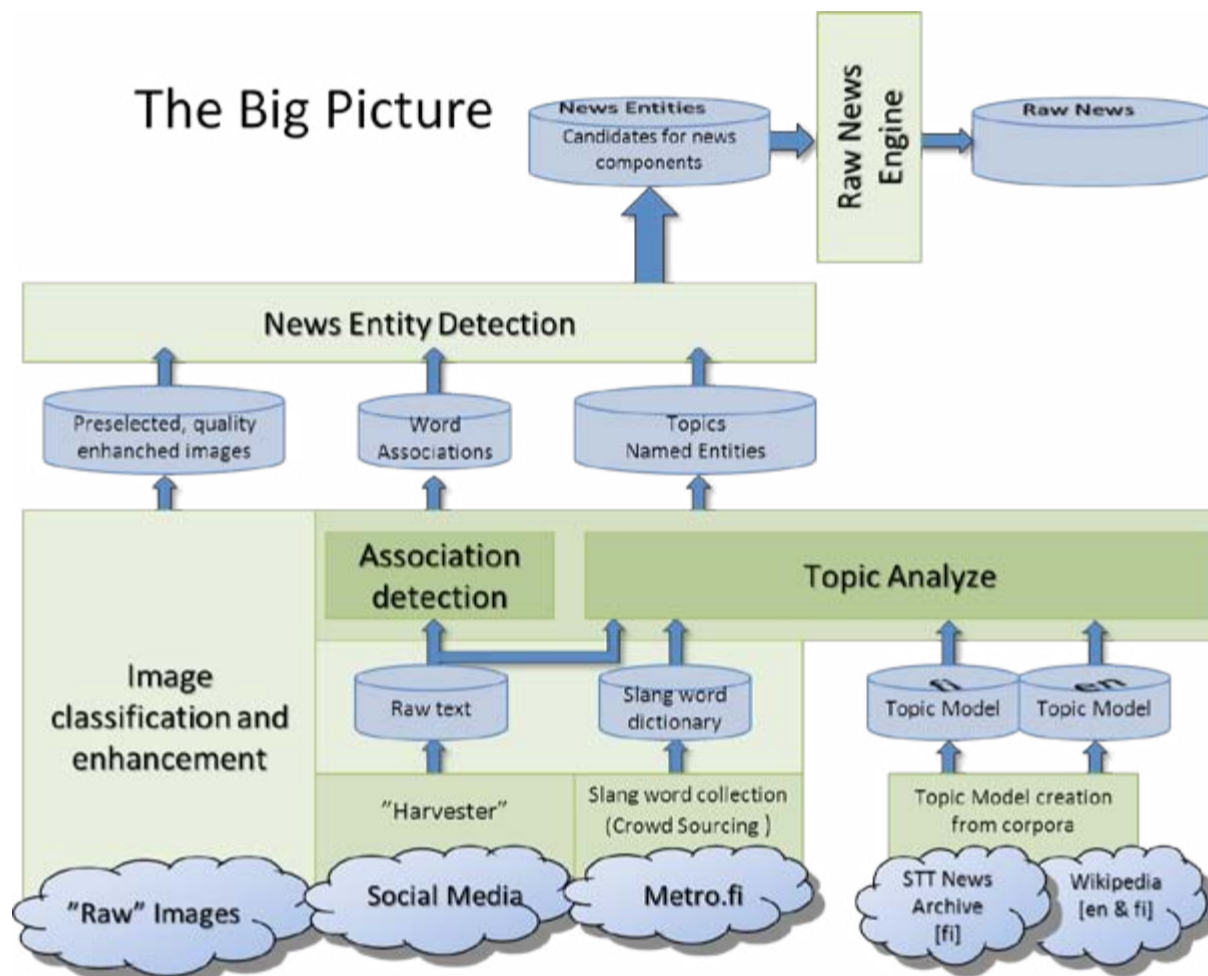
"Feeling":
 Positiivinen Neutraali Negatiivinen

Kuva 8. Slangityökalu kerää käännöksiä.

Jokainen slangisana on käännetty noin seitsemän ihmisen toimesta, jotta saataisiin parhaiten vastaava käännös.

Uusia tiedonkeruulähteitä lisättiin Software Newsroomiin; mukaan tulivat Ilta-Sanomien ja Helsingin Sanomien keskustelut sekä muroBBS-keskustelupalsta.

Työkaluun lisättiin 16 uutta kategoriaa; esimerkiksi viihde, urheilu ja talous. Loppuvuodesta lähteiksi lisättiin myös RSS-syötteitä, kun niitä varten oli kehitetty geneerinen parseri.



Kuva 9. SWNR-työkalun toiminta.

Software Newsroomin keräämää dataa käytettiin testaamaan aiheiden assosiointiin liittyvää työkalua, jonka avulla saattoi löytää sosiaalisessa mediassa tiettyinä aikana puhuttaneita aiheita. Työkalulla oli mahdollista yhdistää sanoja – esimerkiksi organisaatioita, paikkoja ja nimiä – jotka liittyivät toisiinsa. Sitä on tarkoitus myös tutkia ja testata luonnollisen kielen prosessointiin tarkoitettua OMorFi (Open Source Finnish Morphology) -työkalun kanssa. STT-Lehtikuvan uutisarkesta hyödynnetään siihen, että saadaan kerättyä suuri massa suomenkielistä tekstiä korpusta varten.

Suomen kieli on suurena haasteena, kun yritetään saada konetta ymmärtämään tekstiä. Kun englannin kielessä on 370 000 NER (named entity recognition, esimerkiksi ihmisten, paikkojen ja organisaatioiden nimet) -tagattua sanaa, on niitä suomen kielessä vain 2 500. Tätä määrää on tarkoitus päästä kasvattamaan tulevaisuudessa. POS (Part of Speech, esimerkiksi verbit) tagattuja sanoja on englannin kielessä 133 000 ja suomessa 51 000.

- *Tero Oittisen esitys*
- *Tero Oittisen kalvot*
- *Kuvaus (raportti) raakauutistustyökalupakin prototyypistä*

MENETELMIÄ KUVIEN SISÄLLÖN JA LAADUN AUTOMAATTISEEN TUNNISTAMISEEN

Aalto-yliopiston Perustieteiden korkeakoulussa työskenneltiin kuvien tunnistamisen parissa. Lukijareporttereiden kuvat ovat tärkeä osa uutta hyperlokaalia sisällöntuotantoa. Kuvatiedon automaattiseen käsittelyyn tarvitaan uudenlaisia työkaluja, jotta kuvat saadaan luokiteltua helposti ja nopeasti. Lisähaasteen tehtävään tuo lukijoiden kuvien hyvinkin vaihteleva laatu, joka vaikeuttaa nykyisten tunnistustyökalujen toimintaa. Tutkimuksen tavoitteena oli selvittää, millälaisia elementtejä kuvista on ylipäänsä mahdollista tunnistaa. Automatisoitu kuvan visuaalisen laadun tunnistaminen oli niin ikään tärkeä tekijä, jotta useista kuvista voitaisiin valita automaattisesti paras. Kuvan korkea visuaalinen laatu ei silti takaa sitä, että kuva sinänsä on arvokas uutislähde, joten tutkimuksessa selvitettiin erikseen vielä uutisarvoa.

Kuvan uutisarvon määrittelemiseksi Aalto-yliopiston tutkijat tekivät käyttäjäkokeen seitsemän henkilön koeryhmälle. Käyttäjiä pyydettiin järjestämään yhdeksän kuvaa ja juttuaihetta siten, että kullekin uutisotsikolle etsittiin sille parhaiten sopiva kuva. Kuvan uutisarvoon vaikutti kaksi päätekijää: sen laatu ja sisältö. Oli tärkeää, että kuvan sisältö vastasi uutisotsikkoa.

Kuvien automaattista järjestelemistä algoritmien avulla tutkittiin Omakaupunki.fi-palvelun kautta. Omakaupunki.fi:n kuvista luotiin 11 erilaista kategoriaa. Tavoitteena oli tutkia, millä tarkkuudella on mahdollista automaattisesti löytää eri kategorioihin kuuluvat kuvat. Paras tulos saavutettiin ”näyttötaulut”-kategoriassa (94 %), ja myös ”luonto” ja ”ihmismassat” löytyivät hyvin. Sen sijaan heikoimmin tunnistettiin kuvat kategoriassa ”tuli/savu”, joiden kohdalla tarkkuus oli vain 60 %. Tutkimusta jatkettiin STT-Lehtikuvan kuvilla, joissa on runsaasti metadataa.

- *Henri Kivisen esitys*
- *Henri Kivisen kalvot*
- *Raportti kuvan sisältö- ja laatuanalyysin metodeista ja teknologioista **Methods and technologies for image content and quality analysis***