

Datavaalit – hankesuunnitelma

1. Hankkeen kuvaus

”Moni uuden demokratian innostavimmista hankkeista perustuu avoimeen dataan. Usein tarvittava tieto on jo nyt olemassa, mutta se on hajallaan erilaisissa järjestelmissä, hankalasti saatavilla ja käyttökelvottomassa muodossa. Uuden demokratian luominen edellyttää siis demokratiasovelluksiin sopivaa tietopohjaa, jonka luomiseksi kiinnostava ajatus on Datavaalit 2012 -hanke, jossa erilaisilta julkishallinnon verkkosivuilta etsitään ja koneluetaan aineistoa demokratiahankkeissa hyödynnettäväksi.”

Sitra Uusi Demokratia -yhteenveto

Syksyn ja kevään mittaan on useiden eri tahojen kanssa kehitelty Datavaalit 2012 -idea. Hyvissä ajoin ennen kunnallisvaaleja (Vaalipäivä 28.10.2012) aletaan keräämään ja julkaisemaan yhteen paikkaan keskitetysti helppokäyttöisessä muodossa vaaleihin liittyvää avointa dataa. Ensisijaisena kohderyhmänä olisivat datajournalistit ja avoimen demokratian sovellusten kehittäjät.

Tässä ehdotuksessa kuvataan, miten kunnallisvaaleihin tähtäävästä ensimmäisestä puristuksesta muodostetaan samalla jatkuva toimintamalli. Datavaalit olisi pysyvä ja ajan myötä karttuva vaaleihin ja vaalien väliseen keskusteluun liittyvien avoimien tietovarantojen kokoelma. Lisäksi sivusto linkittäisi avoimen demokratian sovelluskehittäjä ja datajournalisteja yhteen verkostoksi.

Verkosto

Datavaalit on yhteinen nimi ja brändi yhteiskunnallisen datan ja erityisesti vaaleihin liittyvien tietoaaineistojen parissa toimiville hankkeille ja tahoille. Datavaalit-sivusto linkittää kaikkia datan journalistisia ja demokratiaa edistäviä sovelluksia ja niiden tekijöitä löyhästi linkittyneeksi verkostoksi, joka tarjoaa mahdollisuuksia asioiden opetteluun yhdessä ja tiedon jakamiseen. Datavaalit-verkosto tuo yksittäisille projekteille näkyvyyttä ja vahvistaa toimijoiden välisiä suhteita. Tämä motivoi tekijöitä ja auttaa keskittämään resursseja esimerkiksi tietoaaineistojen hakuun ja esikäsittelyyn liittyvien teknisten välivaiheiden osalta. Toimijaverkosto auttaa myös tunnistamaan datan saatavuuteen ja käyttöön liittyviä keskeisimpiä pullonkauloja ja tuottaa lisätietoa yhteiskunnallisten tietoaaineistojen kehitystarpeista.

Jatkuvuus

Datavaalit-sivustolle kertyvä tieto ja sivuston ympärille muodostuvat yhteistyöverkostot edistävät osallistumista, demokratiaa ja läpinäkyvyyttä pitkällä tähtäimellä. Datavaalit-sivusto ja toimintamalli perustetaan ennen vuoden 2012 kunnallisvaaleja. Pääasiallinen tavoite ei kuitenkaan ole vain näissä vaaleissa ja näkyvyyspurskauksessa, vaan samalla luodaan avoimen datan infrastruktuuria, joka palvelee yhteiskunnallista keskustelua, datajournalismia ja muitakin tarkoituksia myös tulevaisuuden vaaleissa ja vaalien välillä.

Sekä toimintamalli, että tekninen infrastruktuuri suunnitellaan ja toteutetaan sellaiseksi, että tietovarantojen kartuttaminen on mahdollista ja helppoa laajan toisistaan riippumattoman verkoston toimesta. Ennen kunnallisvaaleja avataan keskeisiä tietovarantoja liittyen kuntiin, ehdokkaisiin, aiempiin vaaleihin jne. Perustettu infrastruktuuri ja alkuun kerätyt ja julkaistut tietovarannot toimivat luurankona, jonka ympärille verkoston piirissä toimivat tahot voivat tuottaa lisää täydentäviä tietovarantoja ja näin Datavaaleista muodostuu karttuva yhteiskunnallisen datan keskus.

Sivuston ylläpidon ja kehityksen jatkuvuuden turvaamiseksi tehdään sopimuksia mediapartnereiden kanssa. Alkuun saattamisen jälkeen Datavaalit ei kuitenkaan edellytä jatkuvia henkilöresursseja, vaan kehityspanostukset voidaan ajoittaa esimerkiksi aina tulevien vaalien yhteyteen erillisiksi kehitysprojekteiksi, joilla tavoitellaan parempaa näkyvyyttä, kattavampaa tietopohjaa tai kehittyneempää infrastruktuuria sivustolle.

2. Mihin hankkeella halutaan vaikuttaa

Tarkoituksena on nostaa kaikkien tulevien vaalien yhteydessä ja niiden välilläkin keskusteluun tärkeitä yhteiskunnallisia teemoja ja ohjata keskustelua nojautumaan entistä enemmän yhteiseen tietopohjaan.

Esimerkiksi keskustelu kunnallisveroasteesta on erilaista silloin, jos ihmiset voivat suhteuttaa veroäyriin sekä omiin tuloihinsa, että kunnan talouteen. Jos veroäyri nousee yhdellä, niin paljonko lisää minä joudun maksamaan veroja ja kuinka paljon lisää kunta silloin kerää veroja yhteensä? Tämä on vain yksi esimerkki lukuisista mahdollisuuksista, joita helposti käytettävä data datajournalistien ja avoimen demokratian sovellusten avustuksella voi tuoda vaalikeskusteluun.

Yksittäisellä kansalaisella tai poliitikollakaan ei ole mahdollista etsiä faktoja siinä määrin, mitä talkoilla saadaan aikaiseksi. Sama pätee myös journalisteihin ja avoimen demokratian sovellusten kehittäjiin, joilla on vain rajalliset resurssit datan keräämiseen ja jalostamiseen. Datavaaleilla pyritään madaltamaan kynnystä sellaisen journalismin ja sovellusten syntymiselle, joka ohjaisi yhteiskunnallista keskustelua yhteiseen tietoon pohjautuvaksi.

Yhteinen tietopohja

Vaikka julkishallinnon avoimesta datasta puhutaankin Suomessa paljon aina hallitusohjelmaa myöten, niin etenkin kuntatasolla demokratian ja osallistumisen kannalta keskeisten tietovarantojen avautuminen laajasti ei ole vielä näköpiirissä lähiaikoina. Nimenomaan helppokäyttöisessä muodossa olevan datan saatavuus on nykyisin suurimpia esteitä muun muassa yhteiskunnallisesti vaikuttavan datajournalismin, sekä demokratiaa ja läpinäkyvyyttä edistävien verkkopalveluiden yleistymiseen. Datavaalit ei itsessään tuota sovelluksia, mutta se luo tätä yhteistä tietopohjaa.

Tiedon löydettävyys

Vuonna 2011 Googlen avulla tehtiin Suomessa noin 20 miljoonaa hakua päivässä.

Suosituimmat hakutermit olivat – Gmail ja Facebook! Mistä tämä kertoo? Hakukenttää käytetään ensisijaisena tapana etsiä toimintoja silloinkin kun taatusti tiedetään url-osoite. Googlen dominointi tiedonhaussa tekee siitä normin. Tietoa ei käytännössä ole olemassa jollei Google sitä löydä. Valitettavasti suuri osa erilaisissa tietokannoissa olevasta julkisesta datasta on hakukoneiden ulottumattomissa. Teknisenä alustana Datavaalit tuo julkisen datan myös hakukoneiden indeksoitavaksi.

Datan käytettävyys

Esimerkiksi datajournalistisen jutun tekemisessä 80% ajasta kuluu datan hankkimiseen ja siivoamiseen ennen analyysivaihetta saatika itse jutun kirjoittamista. Datavaalit tarjoaa runsaasti valmiiksi kerättyä ja käsiteltyä vaalii aiheista dataa, jolloin jutunteossa pääsee alkuun huomattavasti nopeammin. Vastaavalla tavalla avoimen demokratian sovelluksissa tarvitaan usein dataa monista eri lähteistä ja sovellusten kehittäminen olisi huomattavasti helpompaa, mikäli monipuolista dataa olisi saatavilla standardien rajapintojen kautta koneluettavassa muodossa – juuri tätä Datavaalit tarjoaa.

Mitä medialle?

Datajournalismi saattaa tuoda pysyviä muutoksia poliittiseen journalismiin. Esimerkiksi vaalijournalismi on muuttunut vaalikoneiden myötä, nyt datajournalismi voi muuttaa vaalikoneita vaalikauden mittaisiksi seurantapalveluiksi. Siksi datajournalismi kiinnostaakin monia mediataloja, mutta siihen ollaan vasta varovasti tutustumassa.

Datavaalit tarjoaa kevyen mahdollisuuden olla mukana suomalaisen datajournalismin eturintamassa ilman, ettei tarvitse tehdä mittavia investointeja datan keräämiseen. Datavaalit kerää hajautetusti yhteiselle alustalle sellaisia datavarantoja, joiden kerääminen yksittäiselle toimitukselle olisi työlästä tai lähes mahdotonta (esimerkiksi kaikki tällä hetkellä luottamustoimissa eri kuntien lautakunnissa, valtuustoissa yms. toimivat henkilöt).

Uutiskilpailu siirtyy kerralla porrasta ylemmäs, kun kaikilla on sama pohjatietovaranto käytössään, niin mehukkaimmat skuupit saa se, joka keksii parhaan tavan hyödyntää ja rikastaa dataa, aikaa ei tarvitse käyttää perustietojen etsimiseen ja siivoamiseen. Bonuksena päälle mediatalojen omia IT-yksiköitä ei tarvitse häiritä lainkaan.

3. Tavoitteet ja työsuunnitelma

Tavoitteena on luoda tarvittava tekninen infrastruktuuri ja vahva tunnettu brändi, joiden avulla yhteisen vaaleihin liittyvän avoimen tietopohjan kartuttaminen voi tapahtua hajautetusti huomattavasti nopeammin, mitä se tapahtuisi pelkästään julkisen sektorin toimesta.

a. Rahoitus

a.i. Tavoitteet, tuotokset ja mahdolliset mittarit

Infrastruktuuri (suluissa mahdolliset mittarit)

- a. Datavaalit.fi-sivusto (sivuston kävijämäärä, sivuston hakukonenäkyvyys)
- b. Tietovarannot ja datakatalogi (avattujen ja harmonisoitujen datasettien määrä, latausmäärät)
- c. Avoimet rajapinnat (rajapintojen käyttäjämäärät)
- d. Ruudunraapijoiden (screen scrapers) hostaus (raavittujen datasettien määrä)

Vahva tunnettu brändi

- a. Datavaalit-badge, joka linkittää vaaleihin liittyvistä datajournalistisista jutuista ja demokratiasovelluksista takaisin Datavaalit-sivustolle (sovellusten ja datajournalististen juttujen määrä ja laatu)
- b. Jatkuvuuden turvaavat pitkäaikaiset yhteistyösopimukset (sopimusten taloudellinen ja muu arvo)
- c. Medianäkyvyys (maininnat mediassa)

a.ii. Tekeminen ja aikataulu: ks. viimeinen sivu

a.iii. Budjetti

Tässä on esitetty hankkeen budjetti siten, että aktiivinen vaihe kestää vuoden 2012 kunnallisvaalien yli, jäljelle jäävillä varoilla katetaan ylläpitokustannukset ja jätetään siemenrahaa sen verran, että seuraavien vaalien alla voidaan palkata uudelleen osa-aikainen koordinaattori, joka käynnistää uuden aktiivisen vaiheen. Kokonaisbudjetiksi esitetään 22 000€, josta 10 000€ katetaan omarahoitussuutena ja Uusi demokratia kiihdytysrahaa haetaan 12 000€. Lisäksi hankkeeseen kuuluu noin kahden henkilötyökuukauden verran vapaaehtoistyötä.

Omarahoitusosuus

- Mediapartnerit 10 000€ (toistaiseksi varmistettu 2000€)
- Vapaaehtoistyö 2 henkilötyökuukautta

Haettava kiihdytysrahoitus

- 12 000€ (+ ALV 23%)

Kulut

- Koordinaattorin palkka ja muut henkilöstökulut aktiivisessa vaiheessa 7000€
- Infrastruktuurin toteutus 10 000€
- Markkinointi ja muut kustannukset 1000€
- Tapahtumat ja tilaisuudet 1000€
- Ylläpitokulut ja siemenraha seuraavan aktiivisen vaiheen käynnistämiseen 3000€

a.iii.1. Henkilöresurssit

Koordinaattori (2,5 henkilötyökuukautta, 7000€)

Noin viiden kuukauden ajalle käynnistymisvaiheessa kesäkuusta 2012 kunnallisvaalien yli lokakuun 2012 loppuun saakka palkataan osa-aikainen koordinaattori, jonka toimenkuvaan kuuluu:

- a. Varainhankinta
- b. Verkoston koordinointi

- c. Datan avaamiseen liittyvät neuvottelut
- d. Markkinointi

Vapaaehtoistyö (2 henkilötyökuukautta, 0 €)

- a. Infrastruktuurin ja toimintamallien suunnittelu
- b. Datan avaaminen ja siivoaminen
- c. Markkinointi

a.iii.2. Muut kustannukset

Infrastruktuurin toteutus (10 000€)

Perusinfrastruktuuri pyritään toteuttamaan 30.6. mennessä, jolloin Datavaalit-sivusto julkaistaan (sivuston esittely 6.-7.7. Pariisissa kansainvälisessä konferenssissa). Julkaisun jälkeen teknistä alustaa kehitetään vielä palautteen perusteella.

- a. Arkkitehtuurin suunnittelu
- b. Ohjelmointityö
- c. Graafinen suunnittelu
- d. Ylläpito

Muut kustannukset (5000€)

- a. Markkinointi ja muut kustannukset 1000€
- b. Tapahtumat ja tilaisuudet 1000€
- c. Ylläpitokulut ja siemenraha seuraavan aktiivisen vaiheen käynnistämiseen 3000€

a.iv. Yhteiskunnallisen vaikuttavuuden perustelut

Tarkoituksena on nostaa kaikkien tulevien vaalien yhteydessä ja niiden välilläkin keskusteluun tärkeitä yhteiskunnallisia teemoja ja ohjata keskustelua nojautumaan entistä enemmän yhteiseen tietopohjaan. Yksittäisellä kansalaisella tai poliitikollakaan ei ole mahdollista etsiä faktoja siinä määrin, mitä talkoilla saadaan aikaiseksi. Sama pätee myös journalisteihin ja avoimen demokratian sovellusten kehittäjiin, joilla on vain rajalliset resurssit datan keräämiseen ja jalostamiseen.

Datavaaleilla pyritään madaltamaan kynnystä sellaisen journalismin ja sovellusten syntymiselle, joka ohjaisi yhteiskunnallista keskustelua yhteiseen tietoon pohjautuvaksi. Vaikka julkishallinnon avoimesta datasta puhutaankin Suomessa paljon aina hallitusohjelmaa myöten, niin etenkin kuntatasolla demokratian ja osallistumisen kannalta keskeisten tietovarantojen avautuminen laajasti ei ole vielä näköpiirissä lähiaikoina. Nimenomaan helppokäyttöisessä muodossa olevan datan saatavuus on nykyisin suurimpia esteitä muun muassa yhteiskunnallisesti vaikuttavan datajournalismin, sekä demokratiaa ja läpinäkyvyyttä edistävien verkkopalveluiden yleistymiseen.

Datavaalit ei itsessään tuota sovelluksia, mutta se luo tätä yhteistä tietopohjaa. Teknisenä alustana Datavaalit tuo julkisen datan myös hakukoneiden indeksoitavaksi.

Datajournalismi saattaa tuoda pysyviä muutoksia poliittiseen journalismiin. Esimerkiksi

vaalijournalismi on muuttunut vaalikoneiden myötä, nyt datajournalismi voi muuttaa vaalikoneita vaalikauden mittaisiksi seurantapalveluiksi. Siksi datajournalismi kiinnostaakin monia mediataloja, mutta siihen ollaan vasta varovasti tutustumassa. Datavaalit tarjoaa kevyen mahdollisuuden olla mukana suomalaisen datajournalismin eturintamassa ilman, että tarvitsee tehdä mittavia investointeja datan keräämiseen.

Datavaalit kerää hajautetusti yhteiselle alustalle sellaisia datavarantoja, joiden kerääminen yksittäiselle toimitukselle olisi työlästä tai lähes mahdotonta (esimerkiksi kaikki tällä hetkellä luottamustoimissa eri kuntien lautakunnissa, valtuustoissa yms. toimivat henkilöt). Uutiskilpailu siirtyy kerralla porrasta ylemmäs, kun kaikilla on sama pohjatietovaranto käytössään, niin mehukkaimmat skuupit saa se, joka keksii parhaan tavan hyödyntää ja rikastaa dataa.

Tekeminen ja aikataulu

Aloitus 11.6.2012

Kesäkuu

Kunnallisvaaleihin liittyvien verkkosivujen avaaminen

Kunnallisvaaleihin liittyvien tietovarantojen kokoaminen julkiselle verkkosivustolle datakatalogin muodossa

Hankkeesta tiedottaminen toimittajille, avoimen datan aktiiveille ja muille kohderyhmille

Heinäkuu

Datakatalogin täydentäminen

Esittely Pariisin Open Legislative Data -konferenssissa

Elokuu

Kuntien talous- ja päätöksentekodatan ruudunraavinnan ja tietopyyntöjen joukkoistamisen organisointi

Aineistojen siistiminen ja koneluettavien rajapintojen avaaminen

Aineistojen täydentäminen

Syyskuu

Hankkeesta tiedottaminen toimittajille ja muille kohderyhmille

Esittely Open Knowledge Festivalilla Helsingissä (<http://okfestival.org/>)

Tapaustutkimukset datakatalogiin kerätyn vaali- ja päätöksentekodatan käytöstä

Lokakuu

Hankkeesta tiedottaminen toimittajille ja muille kohderyhmille

Kunnallisvaalit

Marras-joulukuu

Hankkeen onnistumisen arviointi

Seuraaviin vaaleihin valmistautuminen

Toimitukset

3.9.2012

- Datavaalit.fi-sivusto sisältöineen. Sivusto sisältää:
 - Kuntiin liittyviä avoimen datan tietoaaineistoja
 - Avoimet rajapinnat em. aineistoihin
- Tiedonkaapijaohjelmistojen (skreippereiden) hostaus
- Hyvin dokumentoituja esimerkkiskreippereitä kehittäjäyhteisölle
- Joukkoistamisprosessit laadittu
- Mediapartnereiden haku aloitettu
- Hankkeesta tiedotettu toimittajille ja muille avoimen datan kohderyhmille

7.1.2013

- Analyysi hankkeen toiminnasta kunnallisvaaleihin liittyen
- Avaamamme datan käyttöön liittyvät käyttötapaustutkimukset
- Suunnitelmat jatkosta
- Joukkoistamista ja tiedottamista jatkettu