

# 1. Datavaalit -pilotin tausta ja tavoitteet



Datavaalit on sitoutumaton, itseorganisoituva ja kaikille avoin yhteisövetoinen hanke, joka kokoaa yhteen vaali- ja päätöksentekodatan parissa toimivia tahoja ja edistää resurssien jakamista ja laajempaa hyödyntämistä yhteisen tietokannan ja tapaamisten myötä.

Datavaalit-pilotissa Kansanmuisti ry. on vastuullinen toteuttajaorganisaatio ja lisäksi sillä on teknisen alustan kehittämisen päävastuu. Kansanmuisti ry:llä on asettanut hankkeelle ohjausryhmän.

Uusi Demokratia-foorumi (2011-2012) kuuluu Suomen itsenäisyyden juhlarahasto Sitran Elinvoima-kehitysohjelmaan, joka on osa Sitran yhteiskunnallista koulutustoimintaa. Uusi Demokratia foorumin puitteissa oli keväällä 2012 avoin ideakuulutus, jossa etsittiin sekä uusia ideoita että jo käynnissä olevia hankkeita, jotka vievät eteenpäin suomalaista demokratiaa. Tällä pyrittiin tunnistamaan demokratian kehittäjät ja heidän ajatuksensa. Datavaalit -idea lähetettiin ideahakuun ja se valittiin karsinnan jälkeen yhdeksi kokeiluksi, jonka käynnistymistä Sitra tukee Uusi Demokratia -kiihdytysrahoituksella (12 000€).

Datavaalit -pilotissa Sitra on rahoittajan roolissa. Uusi demokratia-kiihdytysrahoituksen lisäksi Datavaalit -pilotilla on omarahoitusta, joka tulee pääasiassa datajournalismin koulutuksen myynnistä.

## 2. Lyhyt historia




Datavaalit [hanke julkistettiin perjantaina 21.9.2012](#) kansainvälisen Open Knowledge -festivaalin yhteydessä, kuvassa Datavaalit -hankkeen aktiiveja Leo Lahti vas. ja Joona Lehtomäki oik. (Kuva: [Daniel Schildt](#))

Hankkeen ideointi käynnistyi erilaisissa kokoontumisissa kevättalvella 2012. Datavaalit lähetettiin Sitran järjestämään uusi demokratia ideahakuun ja se valittiin karsinnan jälkeen yhdeksi kokeiluksi, jonka käynnistymistä Sitra tuki Uusi Demokratia -kiihdytysrahoituksella. Aihetta lämmiteltiin toukokuun HSopenissa (Helsingin Sanomien järjestämä datajournalismin hackathon-tapahtuma), ja kesäkuussa 2012 pidettiin toiminnan suunnittelutapaaminen YLE:llä.

Datavaalit-sivusto julkistettiin syksyn 2012 kunnallisvaalien alla Helsingissä kansainvälisen Open Knowledge-festivaaleilla. Samalla hankkeen aktiivit kokosivat vaalidataa hyödyntävää verkostoa (aktiivinen Facebook-ryhmä), laativat koulutusmateriaaleja ja julkaisivat ehdokkaiden vaalirahoitusta ja sosiaalisen median aktiivisuutta valottavia sovelluksia, joista esimerkkeinä Datavaalit-sivustolla tarjottava kunnallisvaaliehdokkaiden sosiaalisen median aggregaatti (<http://www.datavaalit.fi/candidates/social/>) ja Louhos-blogissa julkaistut analyysit (<https://louhos.wordpress.com/2012/10/26/kunnallisvaaliehdokkaiden-aktiivisuus-sosiaalisessa-mediassa-puolueiden-vertailua/>).

Joulukuussa 2012 Datavaalit vei tuplavoiton (yleisö- ja tuomariäännet) Apps4Finland-kilpailun Yksityinen datan avaus-sarjassa. Alkuvuodesta 2013 Datavaalit-hankkeessa kehitettyjä ideoita jalostettiin edelleen Helsingin Sanomain säätiön järjestämässä Uutisraivaaja-kilpailussa. Datavaalien pohjalta tehty Protomedia -ehdotus menestyi Uutisraivaajakilpailussa varsin hyvin päästen sadan kilpailuehdotuksen joukosta 15 parhaan joukkoon. Vaikka Protomedia ei loppujen lopuksi päässyt Uutisraivaajan finaaliin oli kilpailuosallistuminen tärkeää ja se kiteytti Datavaalien jatkokehityksen suuntia.

### 3. Pilotin tapahtumat ja toimenpiteet



**Henri Ahti** @HenriAhti  
viikonloppuna Suomi datahub täydentyi mukavasti vaalidatalla  
[fi.thedatahub.org/group/datavaalit](http://fi.thedatahub.org/group/datavaalit) Hyvä @apoikola ja #datavaalit  
#avoindata  
Expand

8 Oct

Datavaalit -hanke hyödynsi ja omalta osaltaan täydensi yhteisöllisesti ylläpidettyä suomen datakatalogia  
<http://fi.thedatahub.org/group/datavaalit>

Datavaalit sivusto [julkistettiin perjantaina 21.9.2012](#) kansainvälisen Open Knowledge -festivaaleilla. Tietokanta sisältää nyt ehdokas- ja äänestysaktiivisuustiedot 2012 kunnallisvaaleissa, luettelon kunnista ja kuntien rajatiedot (geospaatialisessa tietokannassa), sekä Helsingin kaupunginvaltuutetut (nimet, puolueet). Edellämainitut ovat saatavilla rajapinnan kautta ja tietokannalle on toteutettu rajapintaa hyödyntävät R-kieliset hakurutiinit, jotka mahdollistavat aineistojen sujuvan jatkoanalyysin.

Jatkotavoitteena on joukkoistaa kuntiin liittyvien vaaliaineistojen koostamista muun muassa tiedonkaapija (screen scraper) -ohjelmien hajautetun kehittämisen ja tietopyyntöjen avulla. Esimerkkitoiteutuksena tehtiin hyvin dokumentoitu [Jyväskylän kaupunginvaltuuston jäsenien skreippi](#) ja tarjottiin yhteisölle mahdollisuutta vastaavien skreippereiden hostaamiseen. Esimerkin mukaisia kuntaskreippereitä ei syntynyt, mutta yhteisö tuotti sen sijaan ehdokkaiden sosiaalisen median päivityksiä automaattisesti tietokantaan tallentavan ohjelmiston, vaalirahoitusilmoitusten ajantasaisen tilanteen tallentavan ohjelmakoodin, ja avoimen lähdekoodin laskenta-algoritmeja aineiston analysointiin ja visualisointiin (Ks. 5. Datan käyttö ja sovellukset).

Pilotin aikana järjestettiin ja osallistuttiin seuraaviin tapaamisiin ja tapahtumiin: [Datavaalit -tempaus tapaaminen YLE:llä 12.6.2012](#), [Hacks&Hackers Helsinki \(elokuu\)](#), [Pariisin Open Legislative Data-konferenssi](#), [OKfestival ja sivuston julkistamistilaisuus 21.9.2012](#), [Häkseröintisunnuntai 7.10.2013 \(Helsinki ja Jyväskylä\)](#), [Esiintyminen Hacks & Hackers Helsinki -tilaisuudessa 16.10.2012](#), [Lehdistötiedote vaalirahoitussovelluksesta 26.10.2012](#), [Datavaalit syystapaaminen 19.12.2012](#) ja [Open Knowledge Finland Convention 8-9.2.2013](#)

Yhteisön muodostaminen on onnistunut hyvin (Ks. 4. Sivustojen kävijäävät ja yhteisön aktiivisuus) ja hanke on saanut näkyvyyttä printtimediassa, verkossa ja lyhyesti jopa televisiossa (Ks. 6. Media- ja muu näkyvyys). Joukkoistamisprosessit eivät vielä pilotin aikana lähteneet toimimaan odotetulla tavalla, yhtenä syynä saattaa olla hankkeen käynnistymisen viivästyminen. Mikäli yhteisö olisi ollut nykyisen vahvuinen jo joitain kuukausia ennen vaaleja olisi pohja joukkoistamiselle ollut vahvempi. Vuoden 2014 Europarlamenttivaalien alla Datavaalit-yhteisöllä onkin paremmat pohjavalmiudet toimintaan.

Pilotin puitteissa on kehitetty kaikille avointa datajournalismin opintomateriaalia ja koulutuspäiviä, sekä ylläpidetään Datajournalismi.fi -sivustoa, jossa oppimateriaalit julkaistaan. Tällä hetkellä sivustolla on yhteensä 11 datajournalismin tutoriaalia, jotka ovat keränneet positiivista palautetta käyttäjiltä. Yhteistyökumppanina koulutusten toteutuksessa on ollut Otavanopisto ja datajournalismin tutoriaalien kehitystä tuetaan osaltaan Avoimuudesta Voimaa Oppimisverkostoihin projektin kautta.

Jatkosuunnitelmissa on hyödyntää nyt koostettu data ja ideat ja jalostaa niistä myös vaalien väliseen aikaan soveltuvia uusia avoimen demokratian palveluita, edistää ja tukea hankkeen tarjoamien resurssien laajempaa hyödyntämistä, sekä valmistautua vuoden 2014 europarlamenttivaaleihin yhdessä muiden EU-maiden vastaavien demokratian ja läpinäkyvyyden edistäjäorganisaatioiden kanssa.

## 4. Sivustojen kävijämäärät ja yhteisön aktiivisuus



**Antti Jogi Poikola**  
Vuoden 2008 kunnallisvaaleissa Humppi Virpi / KESK / Karstula on saanut valintatiedokseen 4 ... ainoa laatuaan, mitähän se meinaa?

1 = valittu, 2 = vara, ja 3 = ei valittu  
Tykkää · Kommentoi · Älä seuraa julkaisua · 9. lokakuuta kello 2:26

👍 Kari A. Hintikka, Juuso Parkkinen ja 2 muuta Tykkäävät tästä. ✓ Tämän on nähnyt 84

💬 Näytä kaikki 10 kommenttia

 **Jaana Huhta** Antti, kyseessä oli todellakin virhe. Kaksi samannimistä henkilöä oli summautunut äänestysnumeroitaan myöden (!) ja toisen tiedot jääneet nolliksi. Kyseinen henkilö ei tullut valituksi. Tiedot tullaan korjaamaan näiltä osin viimeistään perjantaihin mennessä. Kiitos tarkkanäköisyydestä ja ilmoituksesta 😊  
10. lokakuuta kello 14:15 · Tykkää · 1

Yksi esimerkki Datavaalit -yhteisön toiminnasta: Facebook-ryhmän keskusteluun osallistuneet ihmiset löysivät ja selvittivät Tilastokeskuksen vuoden 2008 vaalitulosdatassa olleen virheen, jossa kahden saman nimisen ehdokkaan äänimäärät ja muut tiedot olivat vahingossa summautuneet yhteen.

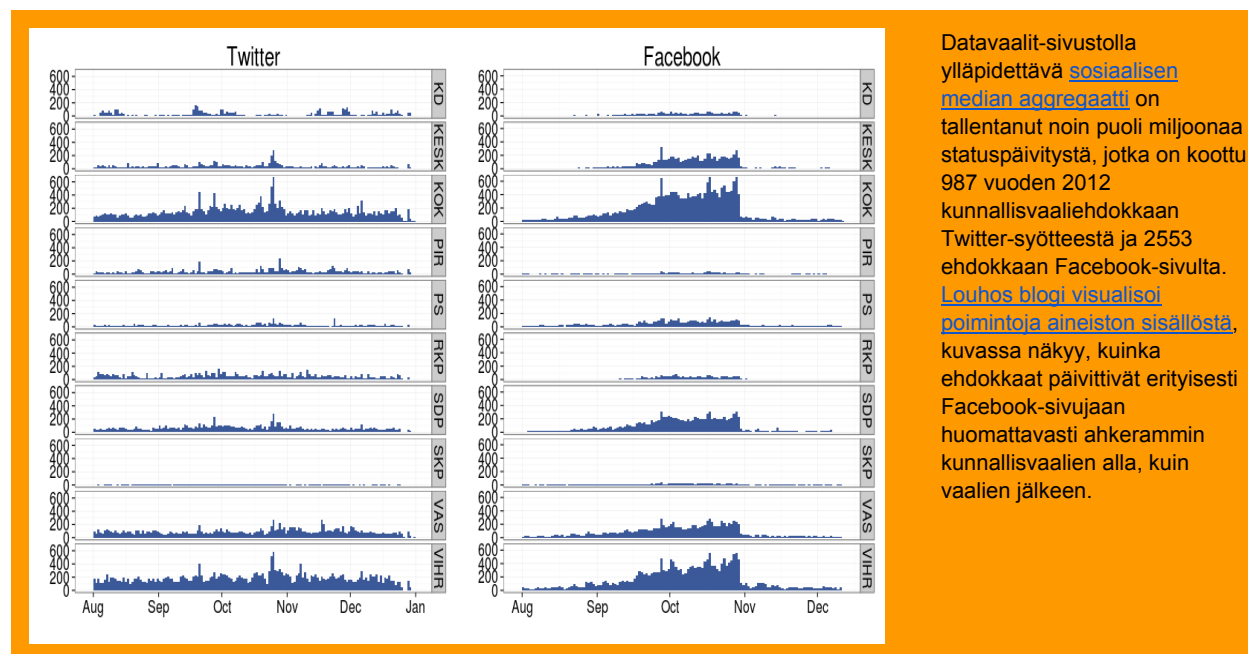
Syyskuussa 2012 avatulla varsinaisella Datavaalit (<http://www.datavaalit.fi/>) -sivustolla eniten kävijöitä vaalien aikaan lokakuussa (9000 kävijää lokakuun aikana). Suurimpia hittituotteita olivat ehdokkaiden [somefiidiagregaatti](#) ja [vaalirahoitusgraafi](#), jotka tehtiin proof-of-concept toteutuksina näyttämään, mihin dataa voidaan käyttää. Syyskuussa, sekä vaalien jälkeen sivuston kävijämäärä on ollut merkittävästi pienempi (n. 100-200 kävijää kuukaudessa).

Pilottihankkeen puitteissa toteutettiin myös Datajournalismi -sivusto (<http://datajournalismi.fi>), jonka kautta on julkaistu muun muassa vaalidataa hyödyntäviä datajournalismin oppimateriaaleja (<http://datajournalismi.fi/tutoriaaleja>). Sivustolla on ollut tasaisesti noin 200-300 kävijää kuukaudessa.

Datavaalit -hankkeen Facebookryhmässä (<http://www.facebook.com/groups/datavaalit>) on käyty aktiivista keskustelua ja se on toiminut datan hyödyntäjien vertaistukikanavana. 25.3.2013 mennessä ryhmässä on 195 jäsentä, joista 142 on osallistunut aktiivisesti keskusteluihin.

Githubin avoindata-tilin kautta (<https://github.com/avoindata>) on jaettu Datavaalit-hankkeen keskeisiä aineistoja helppokäyttöisessä CSV-muodossa yhteisön toimesta. Lataustilastoja ei ole saatavilla.

## 5. Datan käyttö ja sovellukset



Datavaalit-yhteisö tuotti sosiaalisen median päivityksiä automaattisesti tietokantaan tallentavan ohjelmiston ja vaalirahoitusilmoitusten ajantasaisen tilanteen tallentavan ohjelmakoodin. Nämä ovat keskeisiä palveluita tulevaisuutta silmällä pitäen, sillä sosiaalisen median syötteiden saatavuus jälkikäteen on ongelmallista, ja vaalirahoitusilmoituksetkin ovat vaalien jälkeen saatavilla Oikeusministeriön sivuilta vain läpi menneiden ehdokkaiden osalta. Datavaalit-tietokantaan taltioidaan katoavia yhteiskunnalliseen päätöksentekoon kytkettyjä tietoaineistoja.

Datavaalit-hankkeen kokoamia aineistoja käytettiin aktiivisesti, ja analyysyjä julkaistiin sosiaalisessa mediassa. Louhos-blogissa on raportoitu hankkeen tuloksista (<http://louhos.wordpress.com>). Blogissa julkaistiin mm. Tilastokeskuksen aineistoihin pohjautuva yhteenvedoikasarja 2004-2012 kunnallisvaaleista ja analyysi ehdokkaiden sosiaalisen median aktiivisuudesta kunnallisvaalien aikana. Blogilla oli 5587 käyntiä Datavaalit-hankkeen aktiivisimmassa vaiheessa elo-marraskuussa 2012.

Datawikiin ja avoimen datan lähdekoodikirjastoon sorviin ([louhos.github.com/datawiki](http://louhos.github.com/datawiki)) on koottu hyvin dokumentoituja avoimen lähdekoodin välineitä Datavaalit-aineistojen jatkoanalyysiin ja visualisointiin.

Suurin osa yhteisön resursseista kului kunnallisvaaleissa 2012 tarpeellisen tietokantainfrastruktuurin kokoamiseen. Tavoitteenamme on seuraavien vaalien (Europarlamenttivaalit 2014) yhteydessä vahvistaa vaaliaineistojen jatkokäyttöä.



## 6. Media- ja muu näkyvyys



Datavaalit-sivustolta löytyy tuhansien ehdokkaiden Facebook- ja Twitter-päivitykset. Kuva: Yle (YLE 17.10.2012)

24.05.2012 / Vapa -media YouTube

[Vapa tapasi Datavaalit 2012 -hankkeen Antti Poikolan, jonka mukaan datajournalismin suurin haaste on tiedon muokkaaminen käytettävään muotoon](#)

17.10.2012 / YLE

[Yle avasi vaalikonedatan - syntyi työkalu Facebook-seurantaan](#)

21.9.2012 / Maaseudun tulevaisuus

[Datavaalit kokoaa pysyvää tietovarastoa kuntavaaleista](#)

7.11.2012 / Jyväskylän ylioppilaslehti

[Datavaaleista tuli totta](#)

12.11.2012 / Sateen verjolla -blogi (Joona Lehtomäki & Leo Lahti)

[Avoin data, avoimet välineet ja avoimet haasteet](#)

4.12.2012/ Datavaalit-videoita Apps4Finland-gaalasta:

1) [Hankkeen esittely kohdasta 1:13:00 eteenpäin](#)

2) [Datavaalit yleisemmin avoimen datan kontekstissa kohdasta 01:47:05 alkaen](#)

8.1.2013 / YLE TV2 Ajankohtainen kakkonen

[Jakso 1/396 Naapurivahdit poliisia paikkaamaan. Median murros: jyrääkö digikumous painokoneet? Natsit marssilla Unkarissa](#)

Ohjelmassa oli Antti Poikolan lyhyt haastattelu datajournalismiin liittyen.

## 7. Yhteenveto

Datavaalit-hankkeessa onnistuttiin suunnitellusti luomaan avoimen vaalidatan hyödyntämisestä kiinnostuneiden toimijoiden verkosto, ja rakentamaan tietokantainfrastruktuuri helppokäyttöiseen muotoon siistityn vaalidatan jakeluun. Datavaalit-yhteisö on saanut jo verrattain lyhyen toimintakautensa aikana positiivista mediahuomiota ja luonut vahvan pohjan toiminnalle tulevien vaalien yhteydessä. Suurin osa yhteisön resursseista kului kunnallisvaaleissa 2012 tarpeellisen tietokantainfrastruktuurin kokoamiseen. Tavoitteenamme on seuraavien vaalien (Europarlamenttivaalit 2014) yhteydessä vahvistaa vaaliaineistojen jatkokäyttöä ja yhteyksiä vastaavien ulkomaisten hankkeiden kanssa.