



VALTIOVARAINMINISTERIÖ

# **NettiDogi loppuraportti**

Turvallisuus- ja kemikaalivirasto

VM/2310/02.02.03.09/2018

**Versio 0.2**

**24.9.2019**

## Sisällys

Sisällys .....	2
Dokumentin versiohistoria .....	2
<b>1. Yhteenveto .....</b>	<b>3</b>
<b>2. Kokeilun toteutuminen .....</b>	<b>3</b>
2.1. Kokeilun tiedot .....	3
2.2. Kokeilun rahoitus, kustannukset ja henkilötyöpäivät .....	5
2.3. Hankintakäytännöt .....	6
2.4. Riskienhallinta .....	6
2.5. Kokeilun tavoitellut hyödyt ja niiden toteutuminen .....	7
<b>3. Kokeilun päättäminen .....</b>	<b>8</b>
3.1. Kokeilun opit .....	8
3.2. Kokeilun kokemusten jakaminen .....	8
3.3. Kokeilun hyödyntäminen .....	8

### Dokumentin versiohistoria

Versio	Päiväys	Laatija	Muutoksen kuvaus
0.1	16.9.2019	Tarja Valvisto	
0.2	24.9.2019	Tarja Valvisto	

## 1. Yhteenveto

Tämä dokumentti on uuden toimintamallin tai teknologiaratkaisun todentamiseen tähtäävän kokeilun ”NettiDog” loppuraportti. Kokeilun tarkoituksena oli kehittää yksinkertainen datalouhintaan ja -analysointiin perustuva sovellus uudenlaisten, laajaan levitykseen tulevien tuotteiden tunnistamiseen verkkokauppa-alustoilta, mahdollisesti jo ennen kuin ne saapuvat Euroopan/Suomen markkinoille.

Kokeilun tuloksena saatiin selaimessa käyttäjätunnuksilla toimiva palvelu, joka kerää ja analysoi tietoa verkkokauppa-alustan tuotekategoriassa tapahtuvista muutoksista. Palvelu kerää, analysoi ja raportoi tietoa esiintyvien sanojen lukumääristä ja lukumäärien kehityksestä. Kokeilun aikana ei vielä onnistuttu tunnistamaan uudenlaista laajaan levitykseen tulevaa tuotetta.

Kokeilun jälkeen suunnitelmassa on laajentaa palvelua toiseen tuotekategoriaan ja kehittää kerätyn tiedon analysointia.

## 2. Kokeilun toteutuminen

### 2.1. Kokeilun tiedot

Tarkoituksena oli kehittää yksinkertainen datalouhintaan ja -analysointiin perustuva sovellus uudenlaisten, laajaan levitykseen tulevien tuotteiden tunnistamiseen verkkokauppa-alustoilta, mahdollisesti jo ennen kuin ne saapuvat Euroopan/Suomen markkinoille.

Sovelluksen oli suunniteltu toimivan siten, että se louhisi usean kansainvälisen verkkokauppa-alustan tuoteryhmätiedot määrävälein (päivittäin/viikoittain), ja seuraisi muutoksia tuoteryhmissä. Mikäli sovellus havaitsisi jonkin uuden tuoteryhmän syntyvän useaan verkkokauppaan (tai useita samalla tavalla nimettyjä uusia tuotteita vanhan tuoteryhmän sisällä), lähettäisi sovellus hälytyksen markkinavalvonta-asiantuntijalle, joka voisi ryhtyä arvioimaan tuotteeseen/tuoteryhmään mahdollisesti liittyviä vaaroja ja tekemään tarvittavia ennakoivia toimenpiteitä.

Sovellusta oli myös tarkoitus myöhemmin jatkokehittää hakemaan tietoa nettikeskusteluista ja siten palautetta tuotteesta saaduista kokemuksista kuten esimerkiksi tuotteeseen liittyvistä ongelmista ja vaaroista. Ohjelma ikään kuin piirittäisi joka louhinnalla taulun verkkokauppa-alustojen tuotetarjonnasta ja vertaisi kutakin piirrettyä taulua aikaisempiin tauluihin. Mikäli muutos aikaisempaan on tarpeeksi suuri, lähtisi hälytys käyttäjälle, jossa näkyisi mikä muutos on syntynyt. Idean perustana toimii verkkokauppa-alustojen toimintamalli, jossa samalla alustalla toimii lukuisia kauppia. Oletus on, että mikäli uudentyyppinen tuote saapuu markkinoille (ja sille vaikuttaa olevan laajasti kysyntää) alkaa usea kauppias markkinoimaan samaa tuotetta alustoilla. Sovellusta voitaisiin myös jatkokehittää tunnistamaan potentiaalisesti vaaraa aiheuttavia tuotteita, kun tietyt parametrit täyttyvät.

Alla on kokeilun suunniteltu aikataulu ja työpaketit:

Valmis Q1/2019

TP1: Vaatimusten määrittely valituille tuotealueille (sähkötuotteet, lelut, ilotulitteet)

---

TP2: Olemassa olevan ohjelmiston modaaminen TP1 määritellyille tuotealueiden verkkovalvonnalle

TP3: Pilotointi ja korjaukset

TP4: Käyttöönotto

Ohjelmiston kehittämisessä oli tarkoitus keskittyä joko yhteen tai muutamaa myyjään. Alla on suunniteltu eteneminen:

- 1) Rakennettaisiin yksinkertainen tietyille verkkosivulle räätälöity hakurobottisovellus (web crawler) (mieluiten hyödyntäen olemassa olevaa, avointa lähdekoodia)
- 2) Arvioitaisiin kuinka helppoa/vaikeaa on "mönkiä" ison yrityksen verkkosivut (jotka olisivat tässä lähtökohtaisesti kiinnostavimpia esim. uusien tuotteiden näkökulmasta)
- 3) Tehtäisiin tiettyjen tuoteryhmien "koeajoja"
- 4) Rakennettaisiin analysointityökalu
- 5) Tehtäisiin uudelleen vastaava ajo/mönkiminen kuin aiemmin
- 6) Analysoitaisiin data
- 7) Arvioitaisiin toimiiko työkalu/toteuttamistapa ja lähdetäänkö rakentamaan lopullinen työkalu laajemmassa mittakaavassa

Idean testaamiseksi tehtävän proof of concept-sovelluksen (POC) tekijäksi valittiin Haltu Oy. Kokeilun kick off-palaveri oli 20.3.2019 ja tilauksessa suunniteltu valmistumisaika 31.5.2019.

Kokeilussa tehty työ oli luonteeltaan iteratiivista, jonka johdosta kokeilun sisältö, etenemissuunnitelma, työtehtävien prioriteetti sekä projektin aikataulu täsmentyivät kokeilun edetessä. Suurimmat suunnitelman täsmennykset liittyivät projektin aikatauluun eli aikataulun venymiseen.

Hakemuksessa mainittiin, että datan keräämistä koneellisesti oli tarkoitus tehdä usealle sivustolle kohteena ilotulitteet ja lelut. Kokeilun edistyessä päätettiin kuitenkin keskittyä yhteen suureen b2b verkkokauppalustaan ja tuotteista leluihin, koska näin pystyttiin keskittämään kehitysresurssit paremmin tässä POC-vaiheessa.

Kokeilun aikana selvitettiin myös mahdollisuutta kerätä tietoa verkkopalvelun julkisten rajapintojen kautta. Tämä ei onnistunut rajapintojen puutteellisen dokumentaation takia. Jatkossa kehitettiin verkkosivujen "harvointityökalua" (web scraping- työkalua), joka kävi läpi verkkokaupan laajaa lelu- ja harrastustuotekategoriaa. Kokeilussa selvitettiin, kokeeko verkkokauppa screippaamisen haitalliseksi ja pyrkiikö se estämään toiminnan. Screippaamisessa käytettiin apuna ilmaisia välityspalvelimia ja testattiin niiden käytettävyyttä. Toiminta rajoitettiin vastaamaan tavallista verkkosivuston käyttöä. Kokeilussa testattiin myös eri laajuista tuotekategorioiden screippaamista. Datasetit screipattiin viikon välein. Kokeilussa selvitettiin myös eri kerättyjen tuotetietojen hyödynnettävyyttä trendien tunnistamisessa.

Käytetty screipperi ohjelmoitiin Python-ohjelmointikielellä samoin kuin web-käyttöliittymä. Kun palvelu ja screipperi todettiin toimivaksi, perustettiin postgresSQL tietokantapalvelin, johon kerätty data sijoitettiin.

## 2.2. Kokeilun rahoitus, kustannukset ja henkilötyöpäivät

Kokeilun suunnitellut (käyttö- ja kirjausoikeuspäätöksen mukaiset) ja toteutuneet kustannukset euroina eriteltynä omaan ja ostettuun työhön sekä muihin kustannuksiin ovat seuraavat:

Kustannus	Suunniteltu €	Toteutunut €
Oma työ (nykyresursseilla tehtävä työ)	22 500,00	6 723,00
Oma työ (kokeiluun erikseen palkattavien resurssien työ)	0	0
Palvelujen ostot	70 000,00	34 149,60
Muut kustannukset	57 500,00	0
<b>Kokonaiskustannus</b>	<b>150 000,00</b>	<b>40 872,60</b>

Kustannukset eriteltynä rahoituslähteittäin euroina ovat seuraavat:

Rahoituslähde	Suunniteltu €	Toteutunut €
28.70.22 Hallinnon palveluiden digitalisoinnin tuki	50 000,00	34 149,60
32.40.05	100 000,00	6 723,00
<b>Kokonaiskustannus</b>	<b>150 000,00</b>	<b>40 872,60</b>

Kokeilun kustannusten alitus johtuu pääosin seuraavista tekijöistä:

- Kokeilussa jouduttiin etenemään suunniteltua hitaammin iteroivan työtapaa, jolloin sekä omia että ostettuja työtunteja kertyi käytettävissä olevassa aikataulussa suunniteltua hitaammin. Kokeilun POC -vaihe oli suunniteltu saatavan valmiiksi toukokuun loppuun mennessä, mutta valmistui vasta kokeilun loppuun mennessä (5.8).
- Kokeilussa ”haettiin tuntematonta” eli selvitettiin sopivaa tapaa ja työkalua kerätä tietoa verkkokauppa-alustalta ja tunnistaa kerättävästä datasta tietoja mitkä soveltuisivat trendin tunnistamiseen. Kokeilun aikana yritettiin myös saada tunnistettua jokin valittuihin tuoteryhmiin liittyvä trendi tai huomattava muutos, joka auttaisi kehittämään työkalua edelleen

Oman, kokeiluun erikseen palkatun henkilöstön toteutunut kustannus euroina ja henkilötyöpäivinä:

€	htp
0	0

## 2.3. Hankintakäytännöt

Hankinta käynnistyi lähettämällä tammi – helmikuussa viiteen yritykseen hinta-tiedustelut. Saatuja tietoja tarkennettiin lisätietopyynnöillä, tapaamisilla ja skype-kokouksilla. Yhdeltä yrityksistä ei saatu hinta-arvioita ja yksi hinta-arvio ylitti budjetin. Kolmelta muulta yritykseltä saatuja tietoja verrattiin keskenään. Vertailtuja tietoja olivat yrityksen koko, kotimaa, käytössä olevat ohjelmistot ym., referenssit, tarjottu palvelu, palvelun/mahdollisen palvelun kuvaus, projektiehdotus, aikatauluarvio ja hinta-arvo. Koska jäljellä olevien yritysten hinta-arvio jäi alle 60 000 €:n tehtiin päätös suorahankinnasta. Toimittajaksi valittiin saatujen tietojen ja keskustelujen perusteella Haltu Oy. Tehdyllä hankinnalla oli tarkoitus testata konseptin toimivuus, minkä jälkeen tehtäisiin päätös varsinaisen työkalun kehittämisestä, jonka toimittaja valittaisiin kilpailuttamalla.

## 2.4. Riskienhallinta

Rahoitushakemuksen arviointikehikossa kokeilun riskejä on tarkasteltu siinä tapauksessa, että kokeilu toteutettaisiin jatkokehittämällä aikaisemmin kehitettyä Nordcrawler -ohjelmaa. Tämä toteuttamisvaihtoehto ei kuitenkaan toteutunut, koska ko. sovelluksen toimittaja ei arviointimme jälkeen valikoitunut kokeilun toimittajaksi ja koska em. vaihtoehto olisi voinut asettaa rajoituksia kokeilun suunnittelussa ja toteutuksessa.

Rahoitushakemuksen toisessa liitteessä (kuvauspohja ohjelmistorobotiikkaa tai tekoälyä hyödyntäville uuden toimintamallin tai teknologian todentamiseen tähtäville hankkeille) riskejä on tarkasteltu myös siitä näkökulmasta, että kokeilu toteutettaisiin yhteistyössä tanskalaisten, virolaisten ja hollantilaisten viranomaisten kanssa, mikä ei myöskään toteutunut, vaan kokeilu päätettiin toteuttaa omin voimin. Riskeiksi kuvauspohjassa oli tunnistettu:

1. säädösvaatimukset osataan mallintaa riittävän tarkasti ohjelmistoon, jotta NettiDogi löytää mahdollisimman tehokkaasti ja oikeaan osuvasti vaarallisia tuotteita
2. kansainvälinen yhteistyö tuo aikatauluhaasteita, kansallinen rahoitus vauhdittaa Tukesin edellytyksiä edetä tässä kehittämisetäivässä
3. toimittajan kyvykkyys ja hankkeeseen varaamat resurssit

Kokeilun tavoitteena oli tunnistaa verkkokauppa-alustalta uusia laajaan levi-tykseen tulevia tuotteita, joiden mahdollista vaarallisuutta asiantuntijat voisivat arvioida ennen kuin ne tulisivat Suomen markkinoille. Ensimmäinen tunnis-tettu riski liittyy kokeilulla saadun työkalun jatkokehittämiseen. Toinen hake-muksessa tunnistettu riski ei toteutunut, koska kokeilu päätettiin toteuttaa omin voimin. Kolmas tunnistettu riski toteutui osittain toimittajalla tapahtuneiden henkilöstövaihdosten takia sekä kokeilun ajoittumisesta kesälomakau-delle.

Kokeilun riskien tilanne kokeilun päättyessä:

Riski	Lopullinen tila	Toimenpiteet	Toimenpiteiden vaikutus
1	ei toteutunut		
2	ei toteutunut		
3	toteutui osittain	henkilöresurssien uudelleen järjestelyjä	aikataulun venyminen

## 2.5. Kokeilun tavoitellut hyödyt ja niiden toteutuminen

Kuvaa alla olevaan taulukkoon kehitettävän prosessin vaikuttavuus- ja asiakas-hyötypotentialiaali hakemuksen mukaan ja arvioi sen toteutumista kokeilun jäl-keen:

Rahoitushakemuksen liitteessä (kuvauspohja ohjelmistorobotiikkaa tai tekoälyä hyödyntäville uuden toimintamallin tai teknologian todentamiseen tähtääville hankkeille) NettiDogi -sovelluksen on arvioitu parantavan merkittävästi verkko-markkinoilla olevien tuotteiden turvallisuuden valvontaa sekä edesauttavan uu- denlaisten tuotteiden tunnistamista. NettiDogin on arveltu parantavan myös vi- ranomaisen tilannekuvaa markkinoiden kehittymisen volyyymistä ja nopeudesta sekä vaarallisten tuotteiden määristä ja saatavuudesta EU:n ja kolmansien maiden verkkokaupoissa, joka edesauttaisi jäsenmaiden yhteistyötä ja lainsää- dännön kehittämistä. Verkkokauppojen tehokkaampi valvonta varmistaisi myös tasapuolisia kilpailuolosuhteita kivijalkakauppojen ja verkkokauppojen välillä, erityisesti EU:n ulkopuolella toimiviin verkkokauppoihin nähden.

Arvio kehitettävän prosessin vaikuttavuus- ja asiakashyötypotentialista		
Tavoiteltava yhteiskunnallinen vaikuttavuus	Hyötyjen realisoituminen ha- kemuksen mukaan	Arvio hyötyjen realisoitumisen toteutumisesta, jos kokeilussa rakennettu muutos otetaan tuotantoon
nykyistä tehokkaampi verkko- kaupoissa myytävien tuotteiden turvallisuuden valvonta	Tukesilla mahdollisuus tunnis- taa digitaalisesti verkkokau- poissa myytäviä uusia laajalle leviäviä tuotteita ja selvittää nykyistä aikaisemmin näiden mahdollinen vaarallisuus	kokeiluna kehitetty työkalu on vielä trendien tunnistamisen osalta kesken onnistuessaan tuo uusia edel- leen kehitettäviä mahdolli- suuksia verkkokaupoissa myy- tävien tuotteiden turvallisuus- den valvontaan
nykyistä tasapuolisemmat kil- pailuolosuhteet kivijalkakaup- pojen ja verkkokauppojen vä- lillä	Tukesilla mahdollisuus seu- rata ja valvoa verkkokau- poissa myytäviä tuotteita yhtä tehokkaasti kuin kivijalkakau- poissa myytäviä tuotteita	verkkokauppojen myynnin kasvaessa, verkkokaupoissa ja kivijalkakaupoissa myytä- vien tuotteiden yhtä tehokas valvonta ylläpitää tasapuolisia kilpailuolosuhteita

Kuvaa alla olevaan taulukkoon kehitettävän prosessin vaikuttavuus- ja asiakas- hyötypotentialiaali hakemuksen mukaan ja arvioi sen toteutumista kokeilun jäl- keen:

Nykyinen verkkokauppojen valvonnan toimintamalli on täysin manuaalinen eli Tukesin asiantuntija etsii eri tavoin (esim. hakukoneilla, kuten Google, yms.) tietoa uudenlaisista ja mahdollisista vaarallisista tuotteista. Työn tekeminen on aikaa vievää ja kohdistuu satunnaisesti valittuihin kauppapaikkoihin, ja tällä ta- valla ei ole mahdollista pysyä verkkomarkkinoiden perässä. NettiDogin avulla tietojen keruu tehostuu ja systematisoituu, työaika säästöt ovat vähintään noin 6

htkk/v (45 000 €/v). Verkkokauppojen määrien lisääntyessä suhteellinen säästö kasvaa merkittävästi. Suoria kustannuksia merkittävämpää on uudenlaisten tuotteiden aiempi havaitseminen sekä valvonnan osuvuuden parantuminen vaarallisiin, erityisesti henkeä uhkaaviin tuotteisiin ja sitä kautta mahdollisuudet niiden poistamiseen markkinoilta paranevat.

Arvio kehitettävän prosessin tuottavuuspotentialista		
Taloudelliset hyödyt	Hyötyjen realisoituminen	Arvio hyötyjen realisoitumisen toteutumisesta, jos kokeilussa rakennettu muutos otetaan tuotantoon
työajan säästö	Tukesin työaikasäästöt ovat vähintään noin 6 htkk/v	säästöä henkilöstökuluissa 45 000 €/v
vaarallisten tuotteiden aiheuttamien vahinkojen väheneminen ja sitä kautta vahingoista aiheutuneiden kustannusten pieneneminen tuotteiden ostajille		arvio riippuu mahdollisuudesta tunnistaa uusia markkinoille tulevia vaarallisia tuotteita ja mahdollisuudesta poistaa ne markkinoilta

### 3. Kokeilun päättäminen

#### 3.1. Kokeilun opit

Kokeilun avulla saatiin tietoa ja kokemusta käyttötarkoitukseen sopivan screiperin ohjelmoinnista ja sopivuudesta tiedon keräämiseen verkkokauppa-alustalta. Kokeilussa kerättiin myös tietoa ja kokemusta kerätyn tiedon erilaisista analysointitavoista.

#### 3.2. Kokeilun kokemusten jakaminen

Palvelu on käytettävissä käyttäjätunnuksilla selaimessa ja palvelua on tarkoitus esitellä markkinavalvonta- ja muille viranomaisille (kotimaisille ja EU:n jäsenmaiden viranomaisille) yhteisissä kokouksissa.

#### 3.3. Kokeilun hyödyntäminen

Kokeilua on tarkoitus laajentaa kattamaan sähkölaitteet. Kokeilun testausaika on vuoden loppuun, minkä jälkeen päätetään kokeilun jatkosta.