Miten uusi teknologia mahdollistaa reaaliaikaisen riskien arvioinnin

Professori, laitosjohtaja Sasu Tarkoma
Tietojenkäsittelytieteen laitos
Helsingin yliopisto

Esineiden Internet

Esineet satoja miljardeja

Ihmiset 7 miljardia

Paikat miljardi

Globaalit yhteydet

1875 1900 1925 1950 1975 2000 2025

Esineiden Internet

Esineiden Internetin kolme keskeistä osaa ovat toimintaympäristön havainnointi ja siihen liittyvä telemetria, kerätyn tiedon analyysi ja sen perusteella tapahtuva järjestelmän toiminnan säätäminen ja tehostaminen.

Digitaalisuuden kaksi keskeistä yhteiskunnan eri toimialoja koskettavaa veturia ovat esineiden Internet (tai teollinen Internet) sekä skaalautuvat data-analytiikkapalvelut (Big Data). Nämä kaksi keskeistä teemaa tukevat tosiaan ja muodostavat pohjan **datavetoiselle reaaliaikaiselle digitaaliselle infrastruktuurille**.

Kohti esineiden Internetin reaaliaikaista analytiikkaa

hadoop Spark

Reunalla tapahtuva analytiikka

Big Data -kehikot

Reaali-aikainen käsittely

Reaali-aikainen käsittely

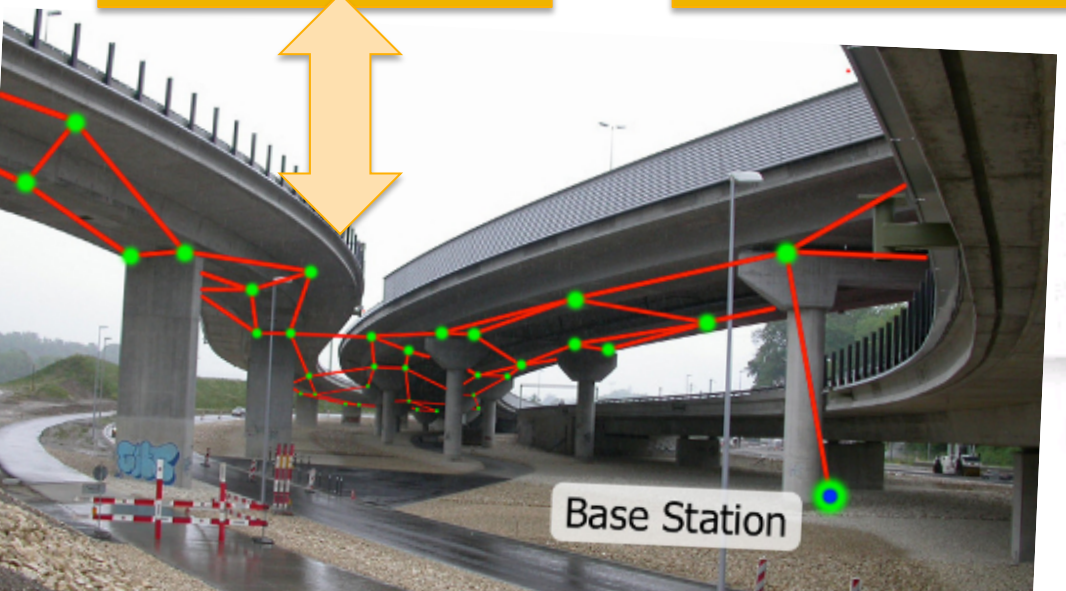
Reaali-aikainen käsittely

Eräajo-käsittely

Datan kerääminen ja prosessointi

Datan välitys ja prosessointi verkossa (4G ja 5G)

Datan prosessointi pilvialustan ja Big Data –kehikon avulla hajautetussa ympäristössä



Esineiden Internet ja digitalisaatio

Teknologia mahdollistaa esineiden ja asioiden reaali-aikaisen seurannan ja säätämisen

Tilat, liikenne, teollisuus, ketjut ja verkostot

Kehittyneet data-analytiikkaratkaisut mahdollistavat uudenlaisen lisäarvon löytämisen datasta

Kone-oppiminen ja tekoäly tulevat muuttamaan toimialoja ja luomaan uusia. ETLA ennustaa, että 36% työnimikkeistä katoaa Suomessa tämän muutoksen seurauksena.

Esineiden Internet visiot

Tämä muutos on jo käynnissä:

- Teollinen Internet / Industrial Internet (GE),
- Kaiken Internet / Internet of Everything (Cisco)
- Teollisuus 4.0 (Saksa)
- Esineiden Internet ja Teollinen Internet (Suomi)

Teknologia mahdollistaa tiedon keräämisen, yhdistämisen ja jalostamisen, jotka muuttavat ja uudistavat rakenteita ja toimintatapoja

Esimerkiksi älyliikenne ja Uber

Esineiden Internet, biotalous ja elintarvikkeiden välitysketju

Esineiden Internet ja biotalous

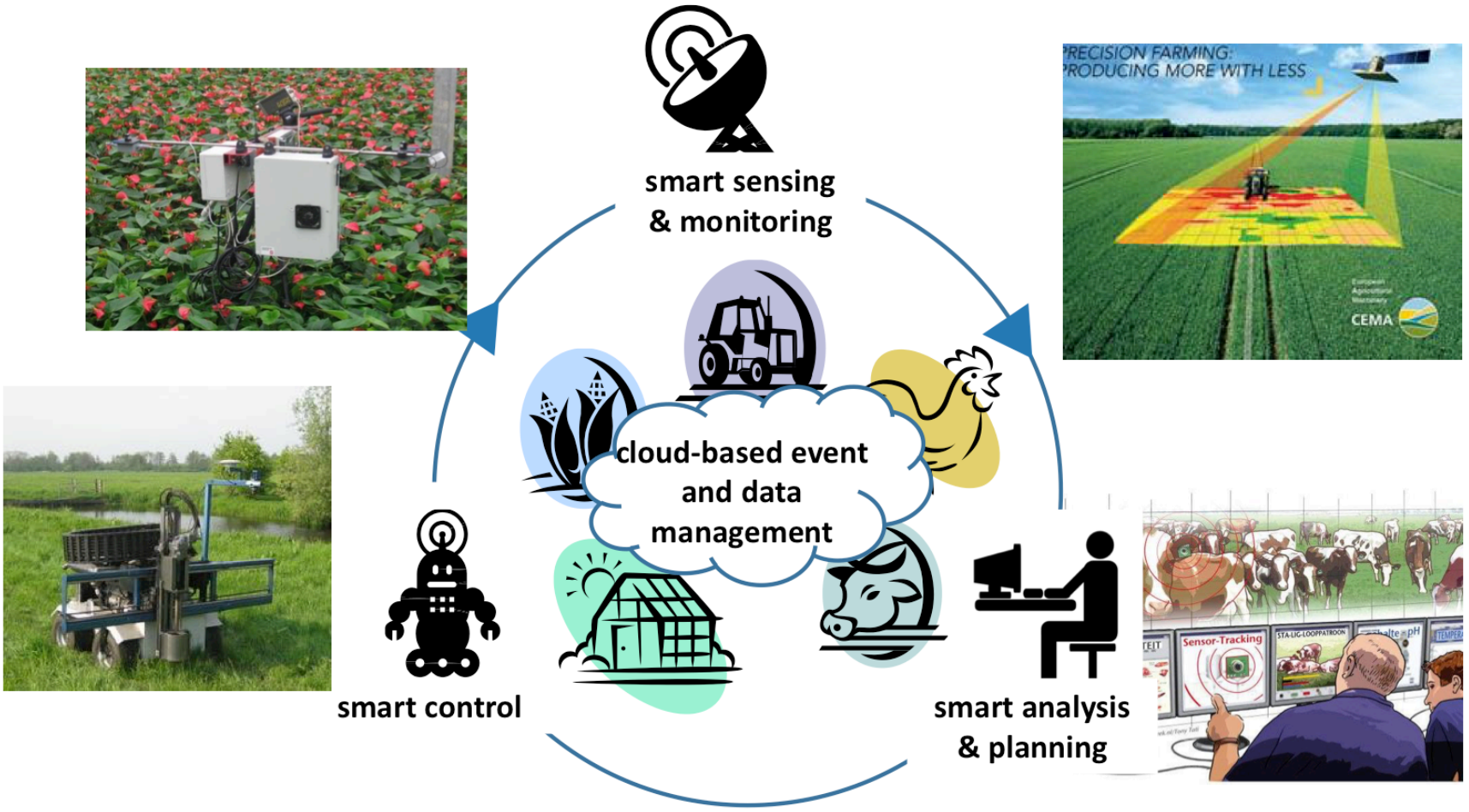
Teollinen Internet ja data-analytiikka mahdollistavat tuotantoprosessien tarkan seurannan ja ohjauksen kehittyneiden dataan ja tilannekuvaan nojaavien algoritmien avulla.

Tuotantoketjun tehokkuutta ja toimintaa voidaan seurata ja optimoida ottaen huomioon ketjun osien ominaisuudet, sähkön hinta ja tuotanto-olosuhteet.

Data-analytiikka löytää mahdollisia pullonkauloja ja voi säätää esimerkiksi maatilan toimintaa ja antaa neuvoja työntekijöille miten tehokkuus paranee ja eläimet voivat paremmin.

Esimerkiksi pellon vedenkäyttöä voidaan seurata tarkasti ja mahdolliset viat putkissa ja laitteissa voidaan tunnistaa ennakoivalla analytiikalla.

Esineiden Internet ja biotalous



Yhdysvalloissa maanviljelijät ovat raportoineet 5% parannuksesta satoon kahden vuoden aikana uuden teknologian vaikutuksesta.

Kuva: Krijn Poppe, LEI Wageningen UR. February 2016 OECD Workshop.

Elintarvikkeiden välitysketju

Elintarvikkeiden tuotantoprosessia ja välitysketjua seurataan erilaisin sähköisin tunnistein ja näin voidaan taata tuore ja turvallinen ruoka kuluttajille.

Kuluttajan lopputuotteen tuotantoketjut voivat olla monimutkaisia ja käsittää kymmeniä toimijoita ja maita.

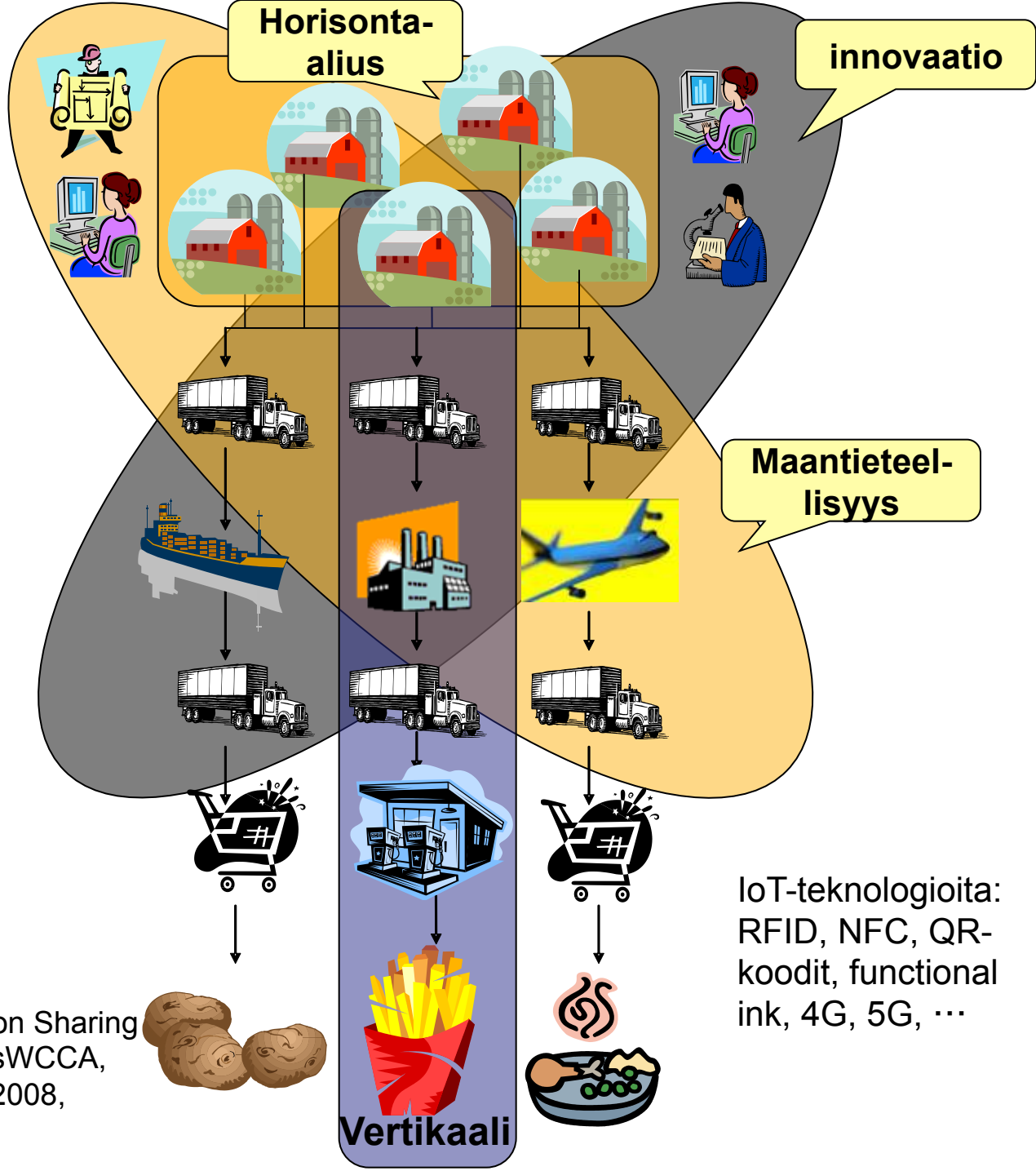
Esineiden Internet mahdollistaa tuotteen ja sen osien tarkan seurannan. Tämä mahdollistaa paitsi tuotannon optimoimisen, myös sen ympäristövaikutusten ja turvallisuuden mallintamisen.

Ennakoiva analytiikka

Moniulotteiset välitysketjut

IoT ja Big Data:
tiedon keräys ja analyysi ketjun läpi ja dataan perustuva tilannekuva ja riskianalyysi

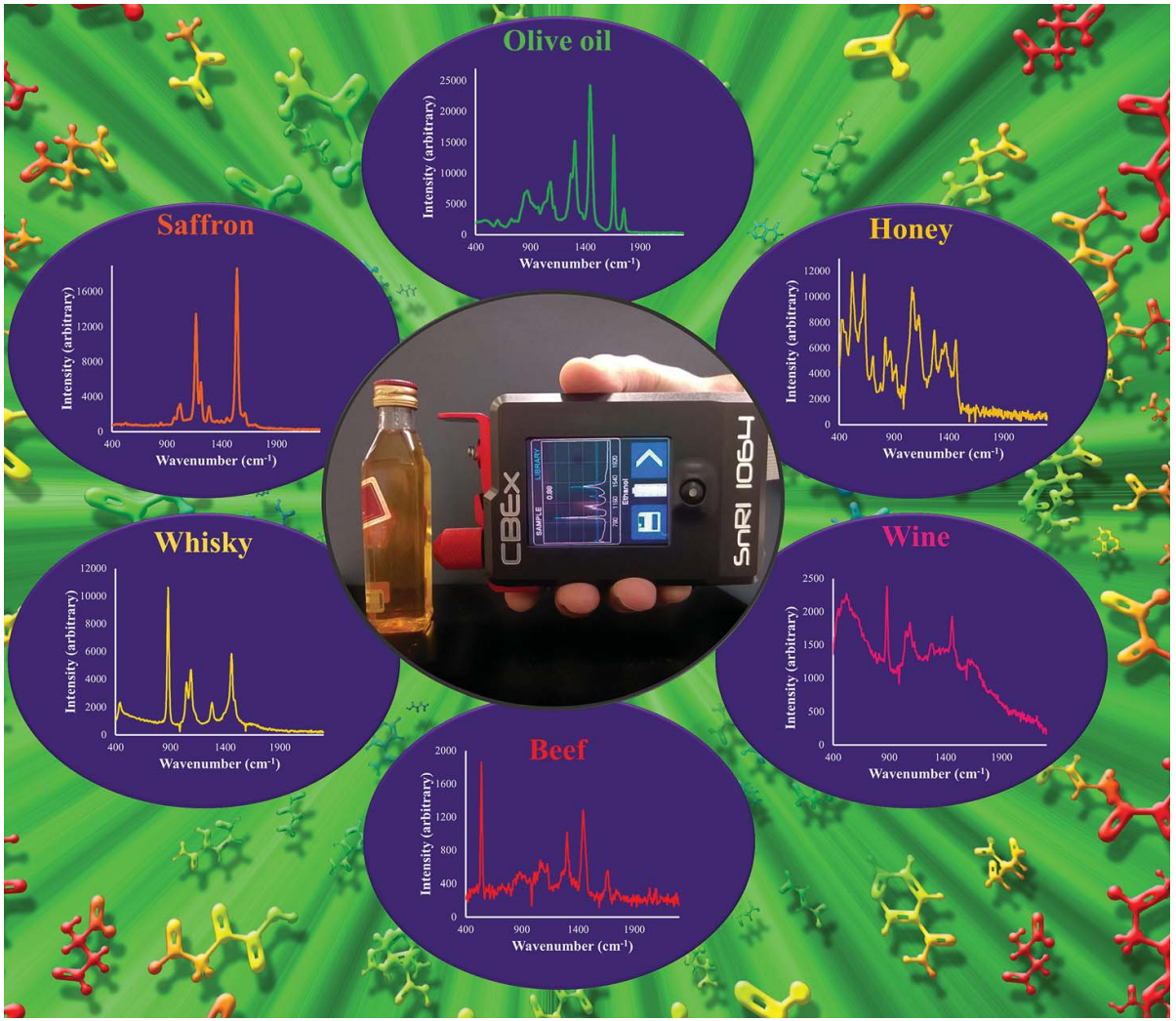
Läpinäkyvyys ja jäljitettävyys



IoT-teknologioita: RFID, NFC, QR-koodit, functional ink, 4G, 5G, …

Kuva: Dr.Ir. Sjaak Wolfert. Information Sharing in Agri-Food Supply Chain NetworksWCCA, workshop ICT adoption, 24 August 2008, Tokyo, Japan

Esimerkki: kannettava spektrometri



Lähde: Point-and-shoot: rapid quantitative detection methods for on-site food fraud analysis – moving out of the laboratory and into the food supply chain. David I. Ellis, Howbeer Muhamadali, Simon A. Haughey,Christopher T. Elliott and Royston Goodacre. Analytical Methods, 2015, 7, 9401.

Esimerkki: kannettava DNA sekvensseri

Kannettava DNA sekvensseri mahdollistaa joustavan näytteiden tutkimisen

Tulevaisuudessa on mahdollista kerätä näytteitä reaali-ajassa

Pilvipohjainen genomidatan analyysi

Kuva: http://www.malaghan.org.nz/news/speedy-sequencing-the-international-verdict-is-out/

Haasteita

Laitteiden ja ratkaisuiden yhteentoimivuus ja toimintavarmuus

- Yhteiset standardit, esitysmuodot, protokollat
- Mistä alustat tulevat?
- Missä data ja palvelut sijaitsevat?

Tietoturva ja tietosuoja

- Kuka omistaa laitteet, verkon ja datan?
- Kuka päivittää ja suojaa laitteet ja palvelut?
- Kuka vastaa monimutkaisissa verkoissa tietojärjestelmien yhteensovittamisesta?

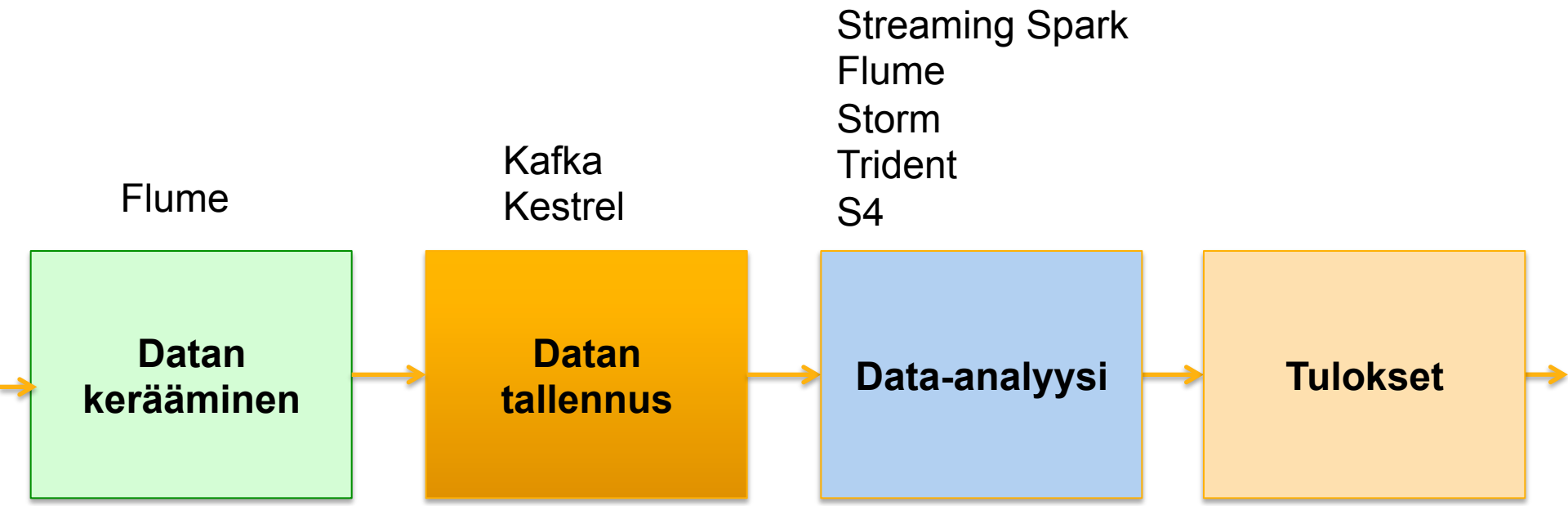
Valtava tiedon ja yhteyksien määrä

Big Data -ratkaisut

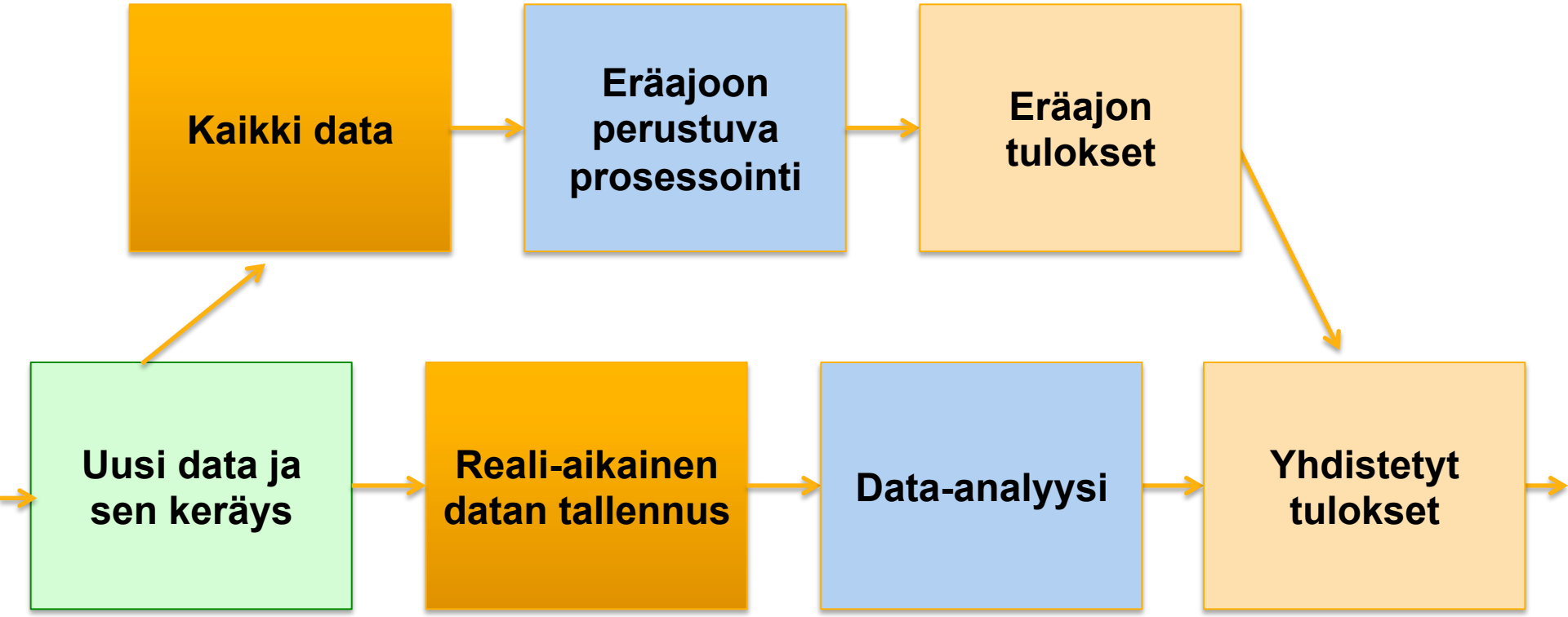
Carat tunnistaa poikkeamat

IoT -ympäristön suojaaminen

Esimerkkejä Big Data -ratkaisuista

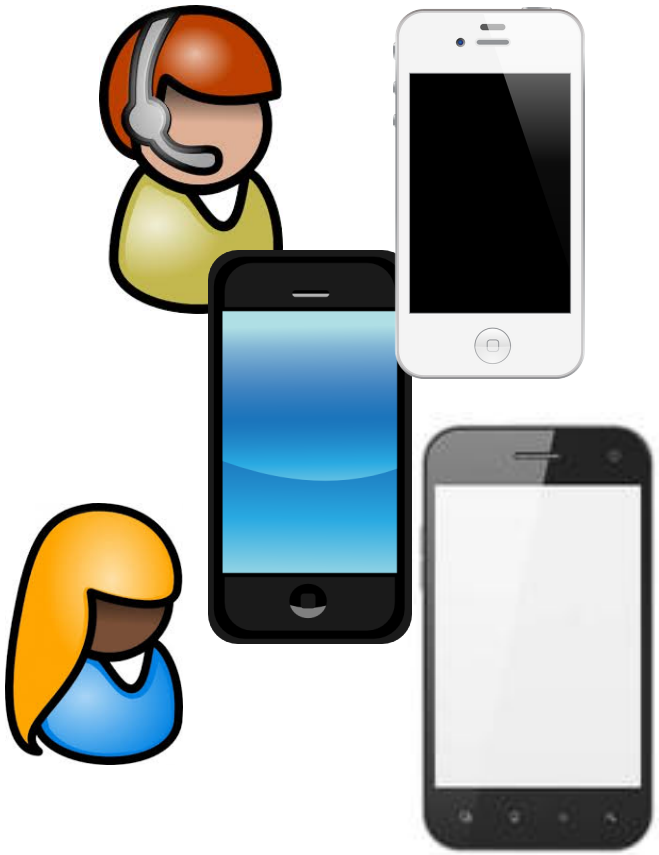


Lambda-arkkitehtuuri



Integroituna esimerkiksi Apache Sparkissa

Carat tutkimus: Johdanto



Akun kesto?

Tietoturva?

Paljon erilaisia laitteita ja käyttäjiä
Paljon erilaisia käyttötapoja

Mikä on normaalia ja tyypillistä käyttöä?

Carat -järjestelmä

Carat on ensimmäinen järjestelmä, joka käyttää mobiililaiteyhteisöä energiaongelmien löytämiseen sekä korjaamiseen

Kehitetty uusi menetelmä diagnosoi energiapoikkeamia yhteisöstä mallinnetun energiaspesifikaation avulla. Anomalia tunnistetaan keskivertokäyttäytymismallin avulla.

Yhteisöllinen datan keruu

Jokainen laite kerää ja lähettää tietoa toiminnasta

- Akun tilanne, aikaleima, sovellukset, asetukset

Data yhdistetään, analysoidaan ja loppukäyttäjä saa raportin energiankulutuksestaan

Yhteisöllisyys: yhteisön avulla voidaan määritellä mikä on tyypillistä energiakäyttäytymistä

Menetelmä on yleinen ja sitä voidaan soveltaa myös tukiasemiin, taloihin, palvelimiin, kannettaviin…

Carat

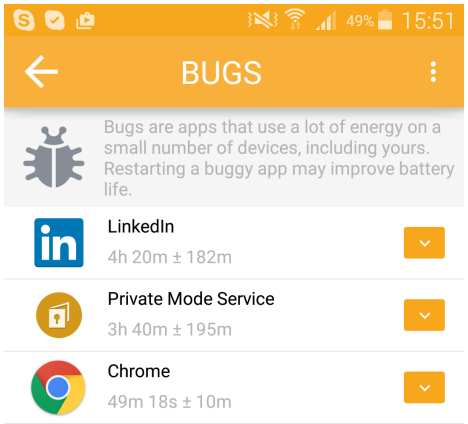
- Yhteistyöprojekti UC Berkeleyn ja Helsingin yliopiston välillä
- Ilmainen mobiilisovellus Androidille sekä iOSille
- Yli 850 000 käyttäjää
- Yli 4 vuotta dataa
  - >2,5 TB dataa, > 250 miljoonaa näytettä
- Yli 450 000 eri sovellusta
- Tutkimusprojektilla on monta suuntaa
- http://carat.cs.helsinki.fi

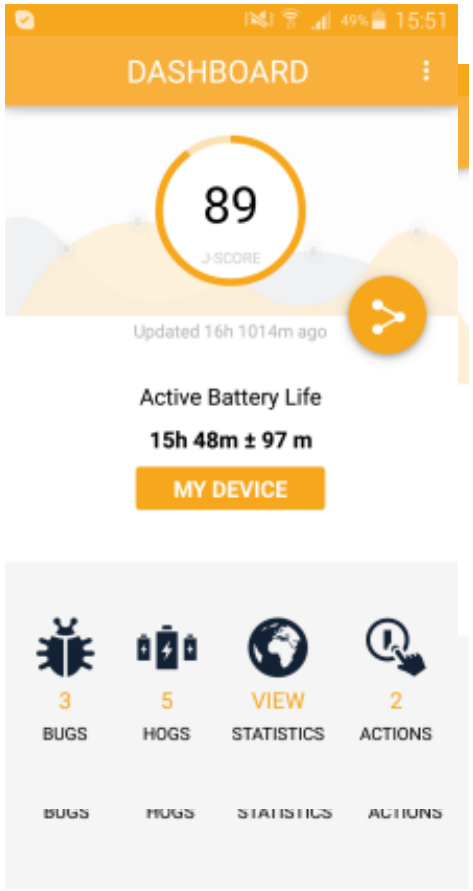
UNIVERSITY OF HELSINKI

amplab UC BERKELEY

Sovellus käyttäjän silmin

- Carat näyttää, mitkä sovellukset syövät enemmän akkua kuin muilla käyttäjillä
  - Kaikki saman sovelluksen käyttäjät vertailussa mukana
- Käyttäjä voi sulkea sovelluksen Caratin kautta
- ja säästää akkua
- Carat kertoo myös J-Scoren, joka mittaa akunkestoa verrattuna muihin

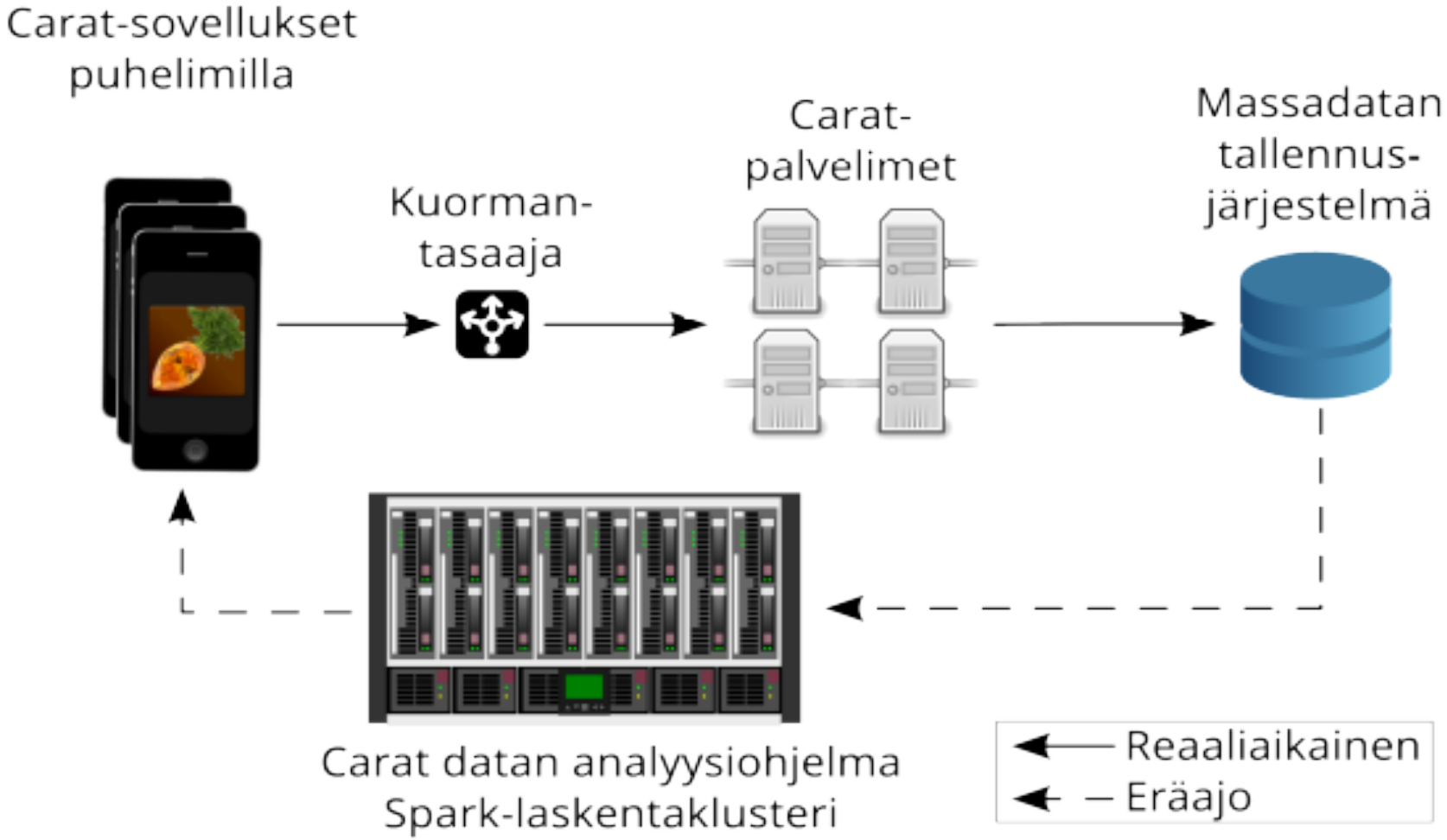




Caratin toiminta

- Käyttäjät näkevät energiasyöpöt ja -bugiset ohjelmat
- Bugiset ohjelmat käyttävät enemmän energiaa tietyllä laitteella kuin muilla laitteilla
- Käyttäjät näkevät mahdolliset energiaongelmat ja niiden korjaamisen aiheuttamat hyödyt
- Ohjelmistokehittäjät ja laitevalmistajat saavat tietoa ohjelmien ja laitteiden energiakäyttäytymisestä

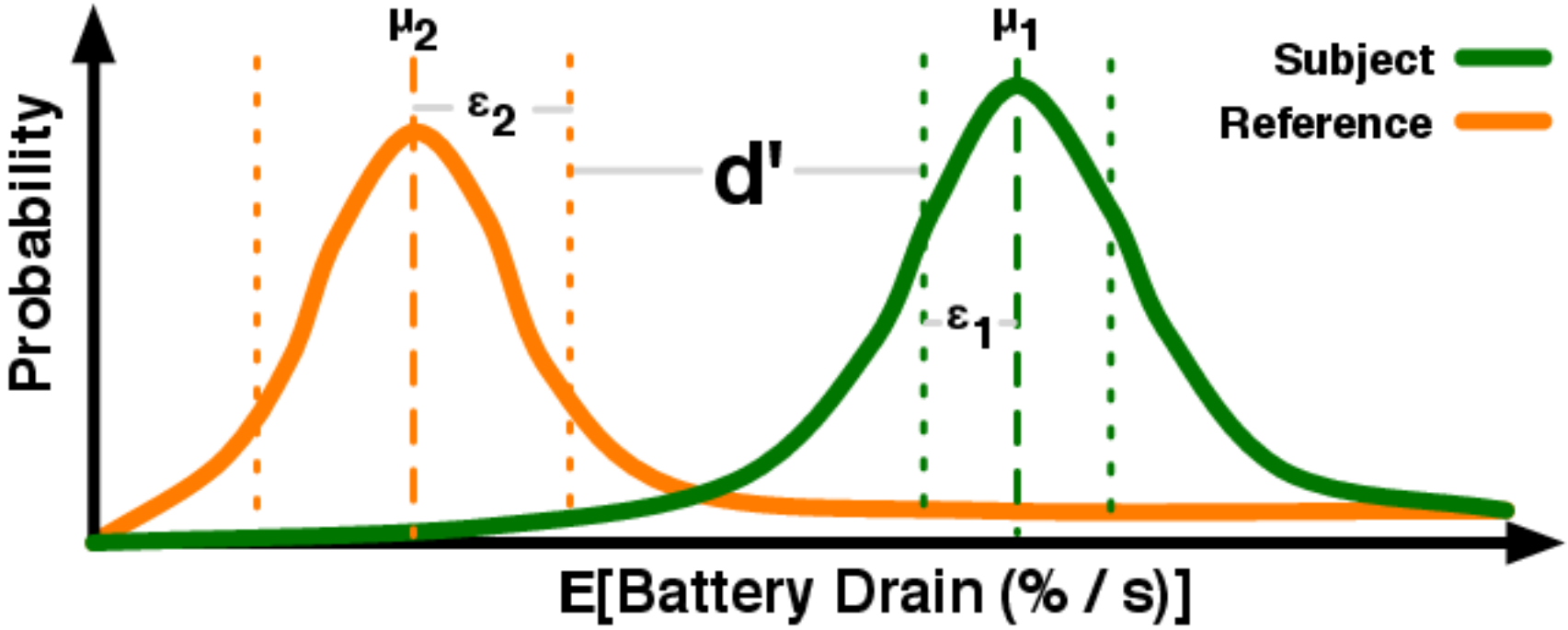
Carat –järjestelmän malli



Energiapoikkeamien (bugien) tunnistaminen

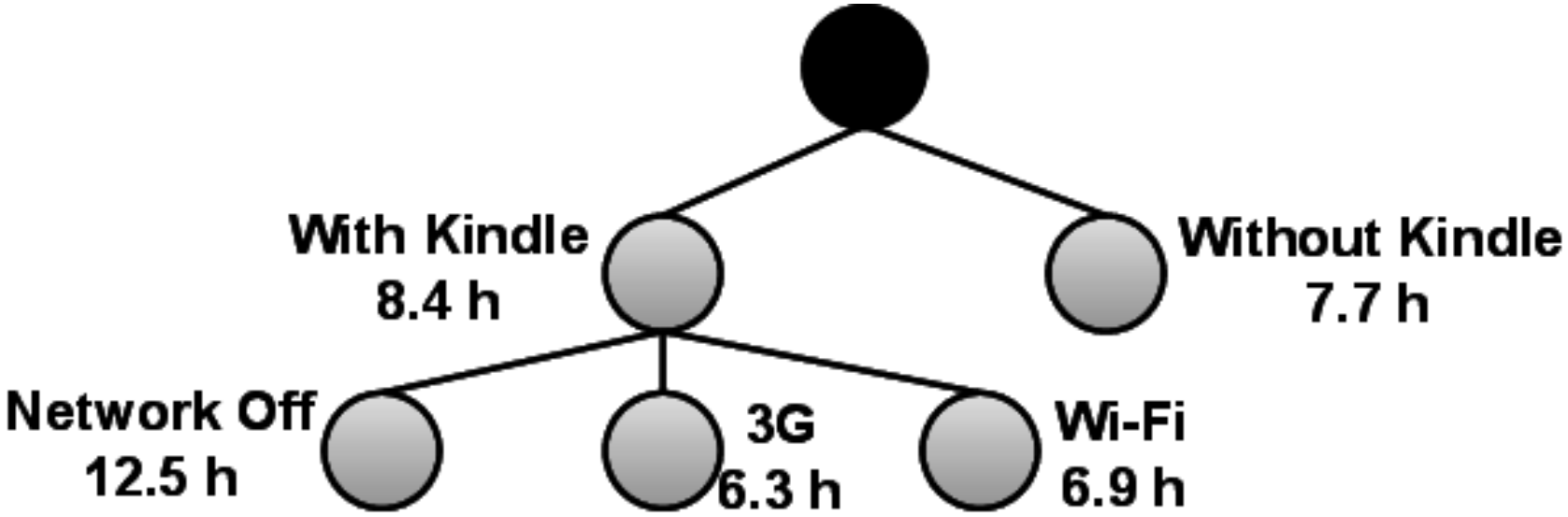
Sovellus X muilla laitteilla

Sovellus X minun laitteellani



…

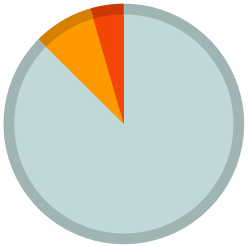
Esimerkki: Kindlen WhisperSync energiapoikkeama



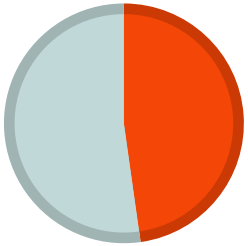
Päätöspuu mahdollistaa laitteiden ja sovellusten pidemmälle menevän diagnosoinnin sekä automaattiset neuvot loppukäyttäjille

Tyypillisiä energiasyöppöjä kaikilla laitteilla

Tilastoja (lokakuu 2016)



471 645 Android ja iOS sovellusta
10% energiasyöppöjä, 4% energiapoikkeamia

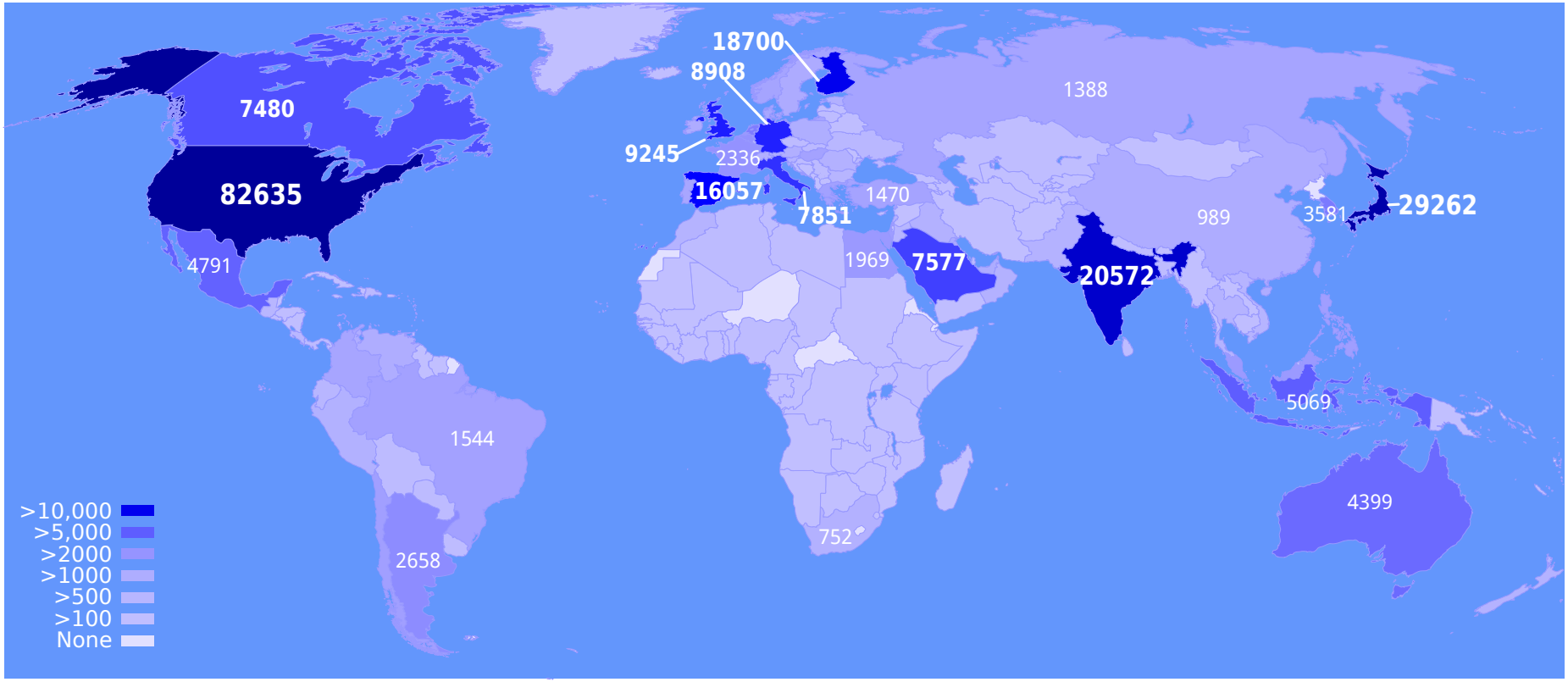


50% kaikista laitteista sisältää vähintään yhden energiapoikkeaman

Androidilla pitkä häntä erilaisia laitteita. Tämä kertoo alustan laitteiston fragmentoitumisesta.

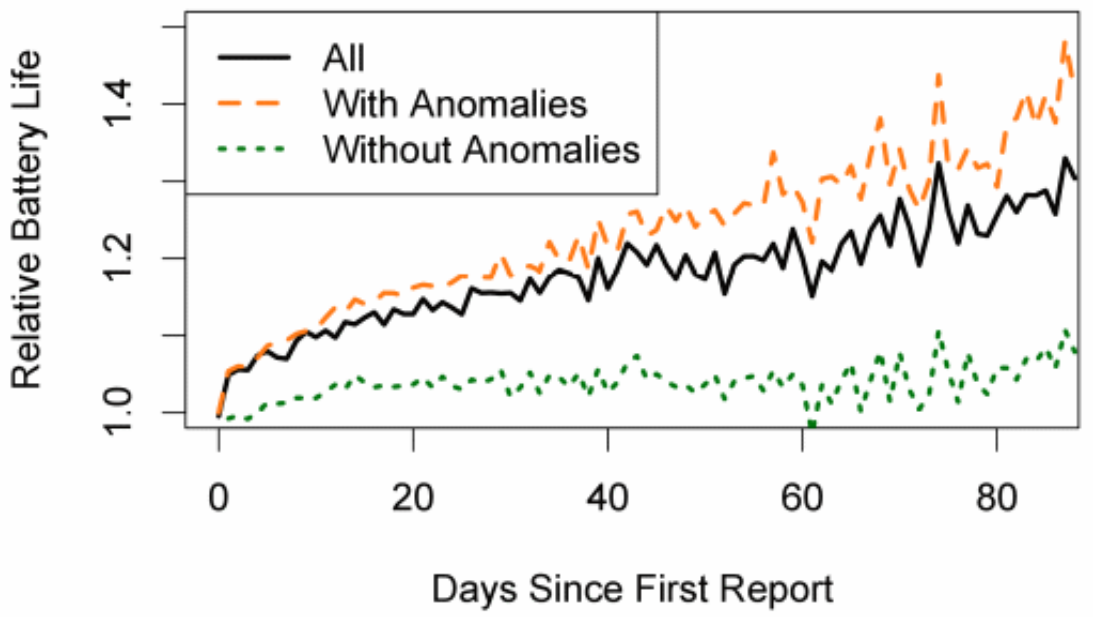
carat.cs.helsinki.fi/statistics

Carat -projektin käyttäjät (Android, 2015)



Caratin hyöty loppukäyttäjälle

- Caratin käyttäjät saavat keskimäärin 20% lisää irti akusta 3kk käytön jälkeen
- Ne joilla on ongelmallisia sovelluksia saavat jopa 41% lisää akunkestoa



Kuinka yleisiä ovat mobiilihaittaohjelmat?

domains. We make several important observations. The mobile malware found by the research community thus far appears in a minuscule number of devices in the network: 3,492 out of over 380 million (less
during the course of our analysis

Charles Lever
Georgia Institute of Technology
chazlever@gatech.edu

Manos Antonakakis
Damballa
manos@damballa.com

Brad Reaves
Georgia Institute of Technology
brad.reaves@gatech.edu

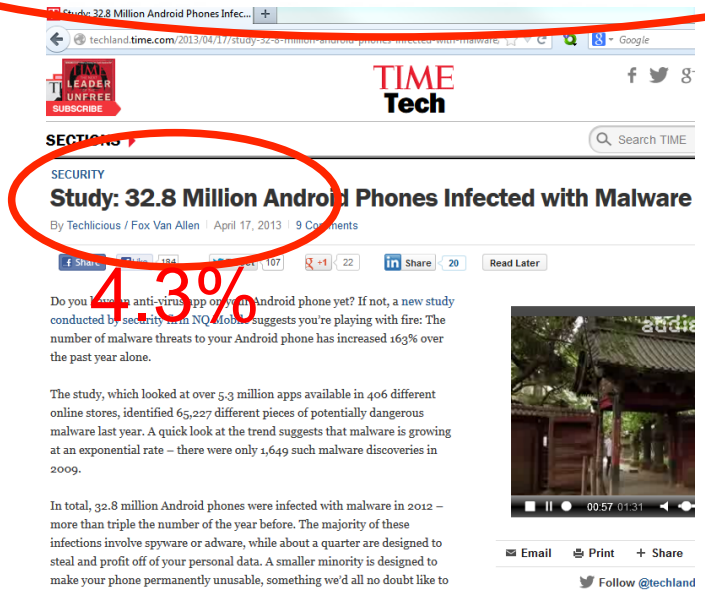
Patrick Traynor
Georgia Institute of Technology
traynor@cc.gatech.edu

Wenke Lee
Georgia Institute of Technology
wenke@cc.gatech.edu

NDSS 2013

?

Lookout



Haittaohjelmien tunnistaminen

Carat instrumentoitiin keräämään sovellusten allekirjoituksia (julkisia allekirjoitusavaimia)

55 000 Android-laitetta

Vertasimme kerättyjä tietoja virustorjuntayritysten

Haittaohjelmatietokantoihin

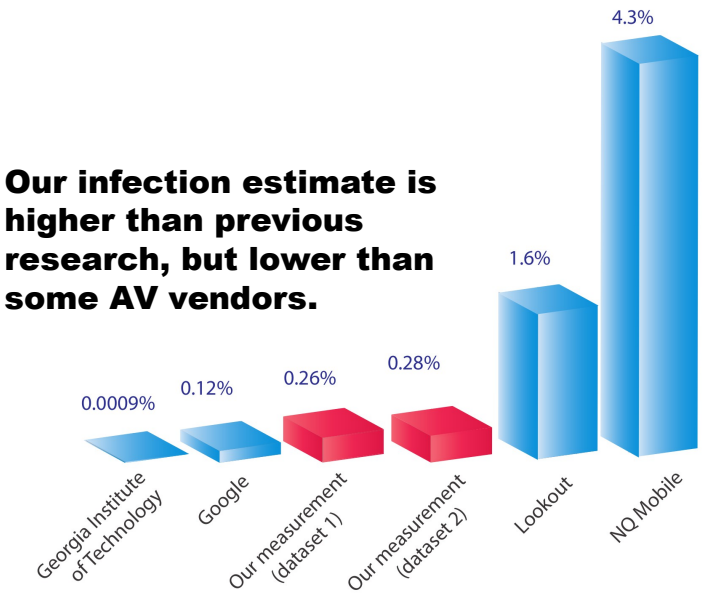
McAfee, Mobile Sandbox, MalGenome, ...

Haittaohjelmien yleisyys

Haittaohjelmia on enemmän kuin konservatiivisimmat arviot olettavat (0,26%)

Google raportoi, että 0,12% manuaalisesti asennetuista ohjelmista on haittaohjelmia

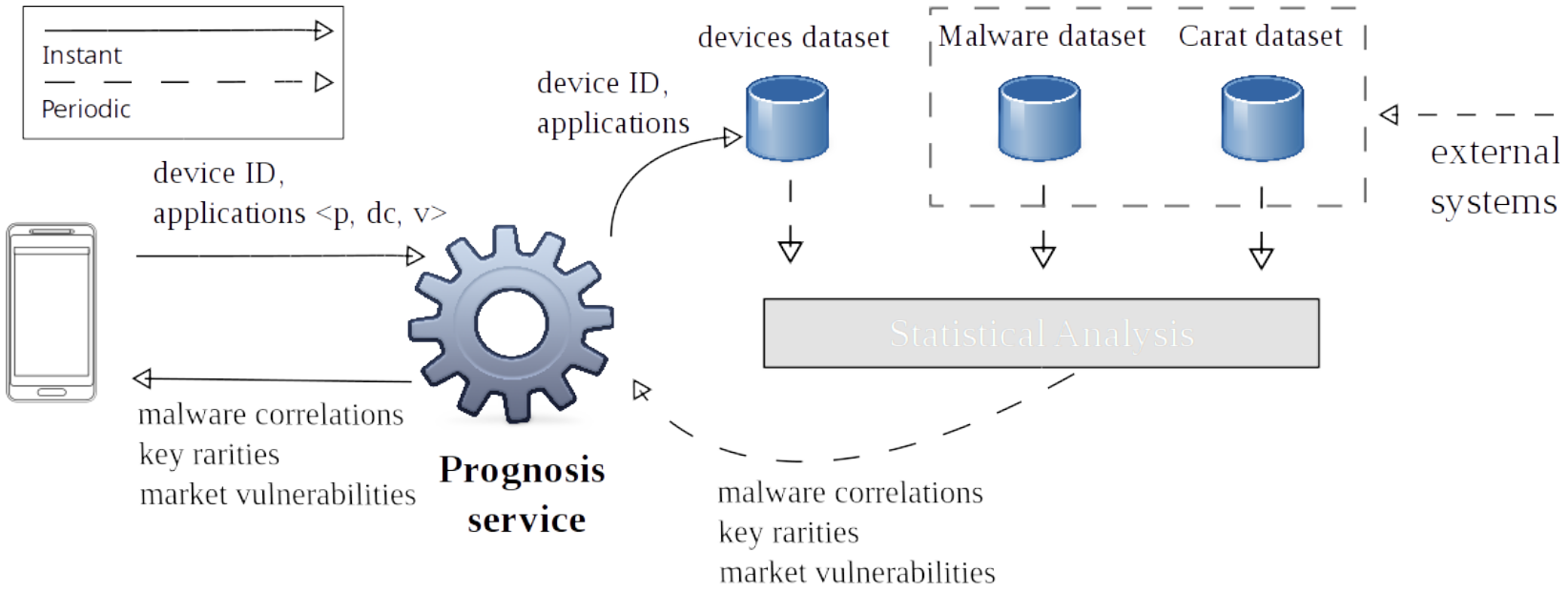
Lookout Antivirus ennustaa >1%



Varhainen varoitusjärjestelmä

Kehitimme varhaisen varoitusjärjestelmän, joka tunnistaa riskipitoiset laitteet

Käytettyjen sovelluksien ja haittaohjelmien välistä korrelaatiota käyttämällä voidaan ennustaa laitteen saastuminen 5 kertaa paremmin kuin laitteiden satunnainen valikoiminen testausta varten

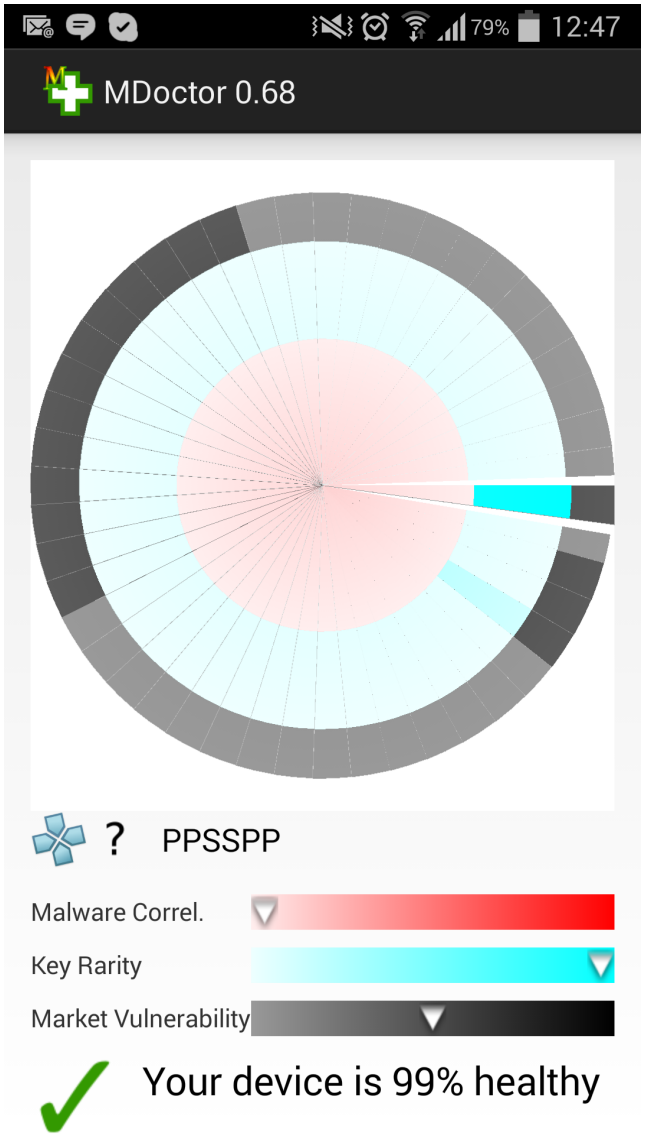


MDoctor: Mobiililaitteen haittaohjelma-tilannekuva

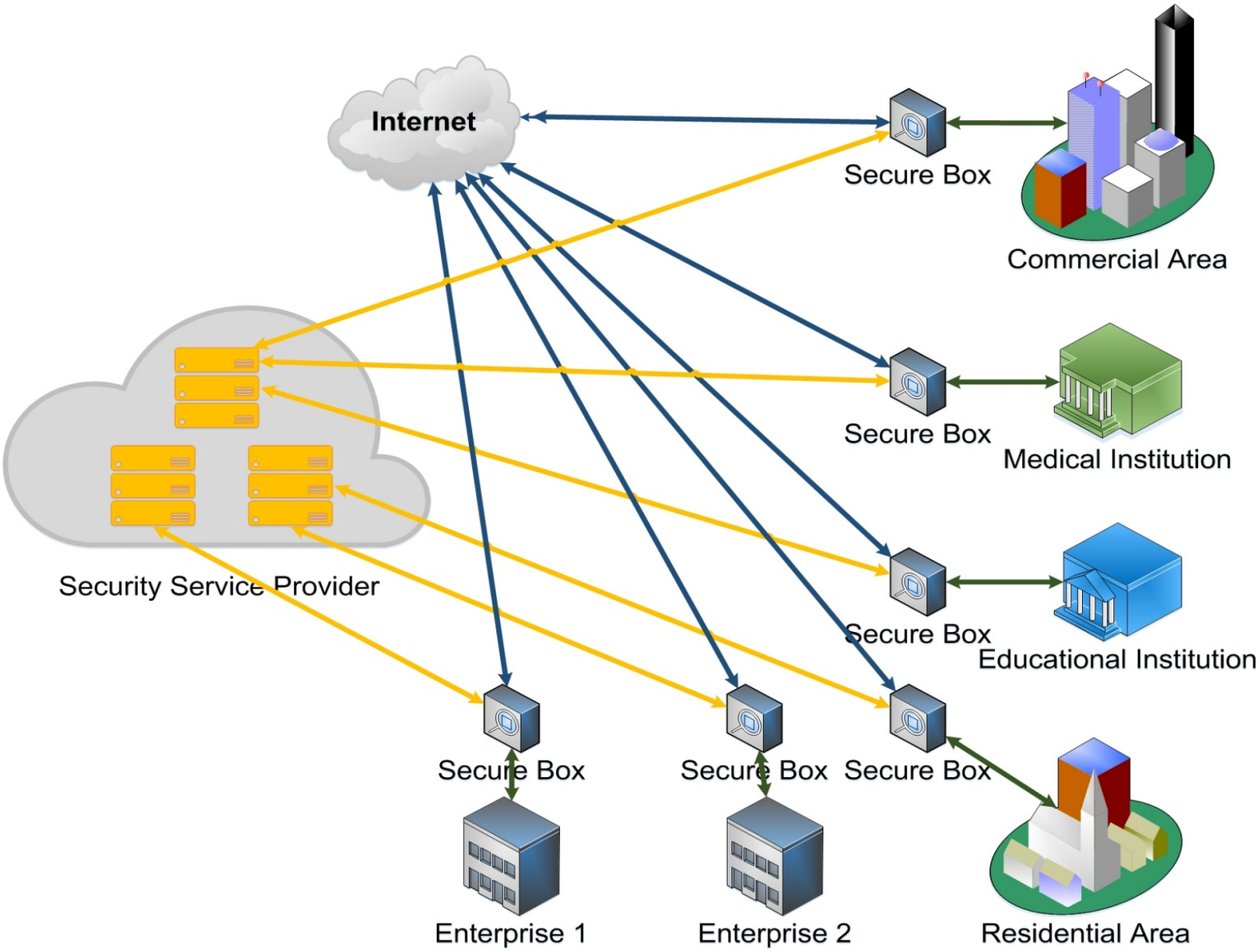
MDoctor sovellus näyttää tilannekuvan mobiililaitteen haittaohjelmariskeistä

Ohjelma laskee tunnusluvun riskille saada haittaohjelmia ja näyttää kunkin sovelluksen vaikutuksen kokonaisriskiin

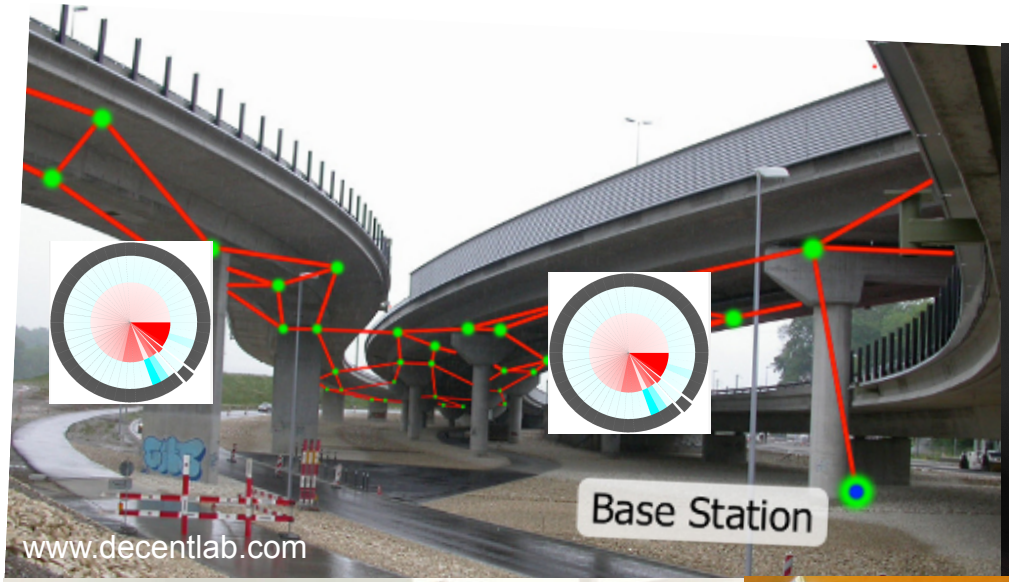
Ohjelma käyttää kolmea keskeistä metriikkaa:
haittaohjelmakorrelaatio
allekirjoitusavainten analyysi
sovellusmarkkina-analyysi

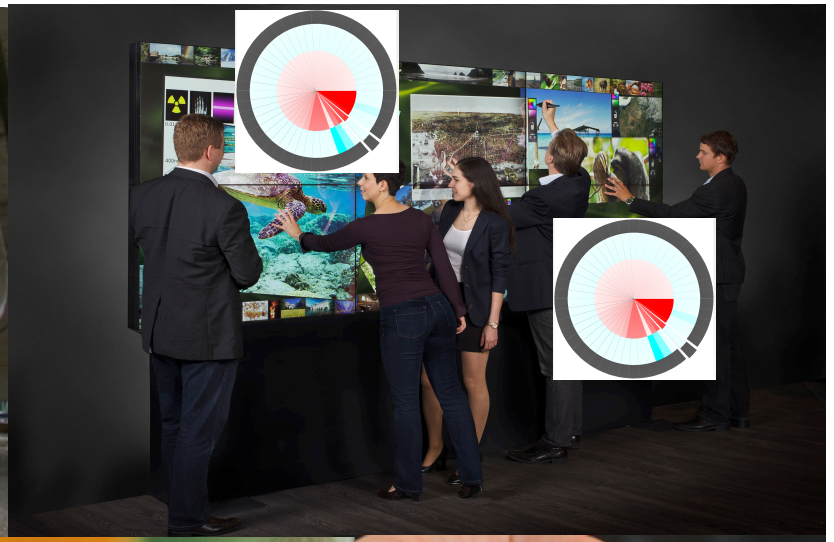


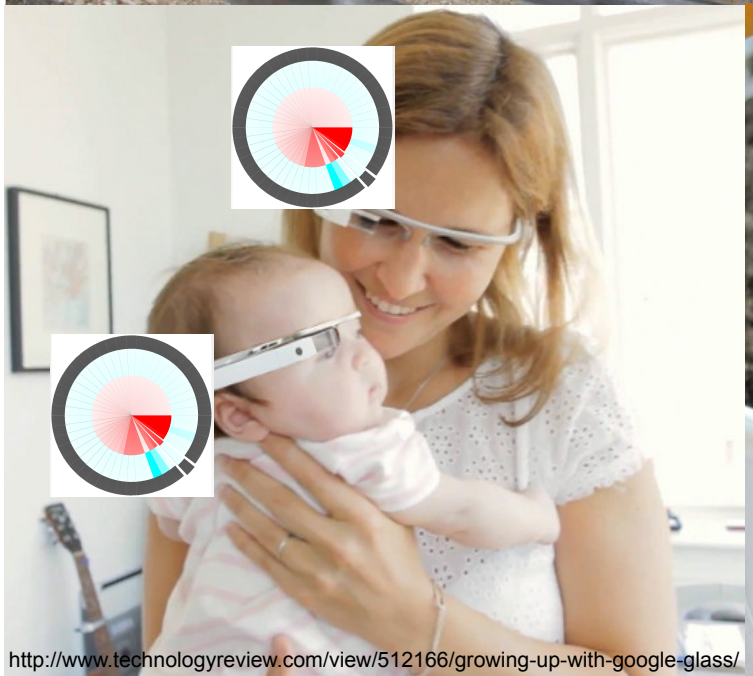
SecureBox -arkkitehtuuri

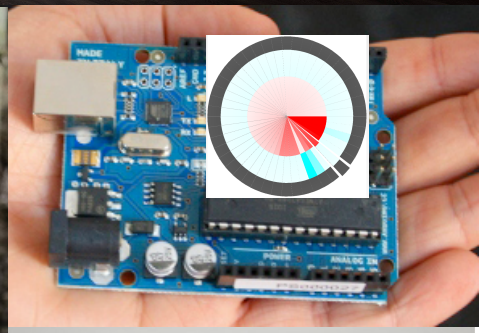


Kohti esineiden Internetin reaaliaika-analytiikkaa









Yhteenveto

- Esineiden Internet ja Big Data tukevat tosiaan ja muodostavat pohjan **datavetoiselle reaaliaikaiselle digitaaliselle infrastruktuurille**.
- Esineiden Internet mahdollistaa tuotteen ja sen osien tarkan seurannan. Tämä mahdollistaa paitsi tuotannon optimoimisen, myös sen ympäristövaikutusten ja turvallisuuden seurannan.
- Nykyiset ratkaisut mahdollistavat reaali-aikaisen data-analyytikan, poikkeamien tunnistamisen sekä ennakoivan analytiikan.
- Haasteita: miten tieto kerätään ajantasaisesti ja aikaansaadaan reaali-aikainen tilannekuva?

Kiitokset