DLM tutustui Norjaan ja E-Arkin tuloksiin

Anssi Jääskeläinen
TKI-asiantuntija
Digitalia-tutkimuskeskus
Mikkelin ammattikorkeakoulu

Tällä kertaa DLM-Forumin pitopaikaksi oli valikoitunut Oslo, lähinnä siksi että DLM-Forumilla ei ole jäseniä EU:n tämän hetkisessä puheenjohtajamaassa Slovakiassa. Lisäksi Norjan kansallisarkisto on ollut *Hans Fredrik Bergin* voimin jo pitkään aktiivisesti mukana DLM-Forumin toiminnassa. Kansallisarkisto, jossa tapaaminen pidettiin, sijaitsee 20min metromatkan päässä keskustasta. Rakennus on tehty osittain rinteeseen ja osittain maan alle. Kierroksella kävi ilmi että maanpäällisiä kerroksia on kolme ja maanalaisia kerroksia neljä.

Bergin alkupuheenvuoron jälkeen *Zoltan Szatucsek* avasi DLM:n puheenjohtajan roolissa tämänkertaisen kokouksen. Aluksi hän totesi ettei Oslossa ole ollenkaan niin kylmä kuin hän oli olettanut. Sitten hän hieman sarkastisesti viittasi tiettyyn EU:n jäsenmaahan joka ei kuitenkaan tuo mitään DLM-Forumiin, käyttäen tätä aasinsiltana Norjan arkistolaitokseen joka on DLM-Forumin jäsen vaikka Norja ei EU:n jäsen olekaan. Norjan arkistolaitoksen johtajan *Inga Bolstadin* avauspuheenvuoro päätti alustuksen, puheessaan hän korosti kansainvälisten tiedonjako- ja levityskanavien tärkeyttä. Näistä hyvänä esimerkkinä hän mainitsi DLM-Forumin.

Varsinaiset puheenvuorot aloitti *Herbjørn Andresen* (Oslo and Akershus University College). Puheessaan hän keskittyi siihen, onko RM (Records Management – asiakirjahallinta) toiminta hyötyjen saavuttamista varten vai riskien välttämistä varten. Omaan kysymykseensä hän vastasi että molempia. Etujen korostaminen on kuitenkin hänen mielestään jäämässä taka-alalle, mutta näin ei kuitenkaan saisi olla, koska koko RM-toiminnan olemassaolo ei voi perustua vain riskien hallintaan. Riskienhallinnassakin tulisi olla myös muita näkökulmia kuin liiketoimintariski, jota kuitenkin ylikorostetaan.

Pia Virmalainen Jøsendal (Agency for eGovernment) kysyi puheenvuorossaan voiko julkinen sektori olla hyödyntämättä hallinnassaan olevaa tietoa strategisena välineenä. Alalla hänen mukaansa paljon erilaisia termejä, joita pitäisi tunnistaa ja pystyä hyödyntämään. Kaiken tiedon tulisi olla digitaalista, avointa ja rajat ylittävää jo oletuksena.

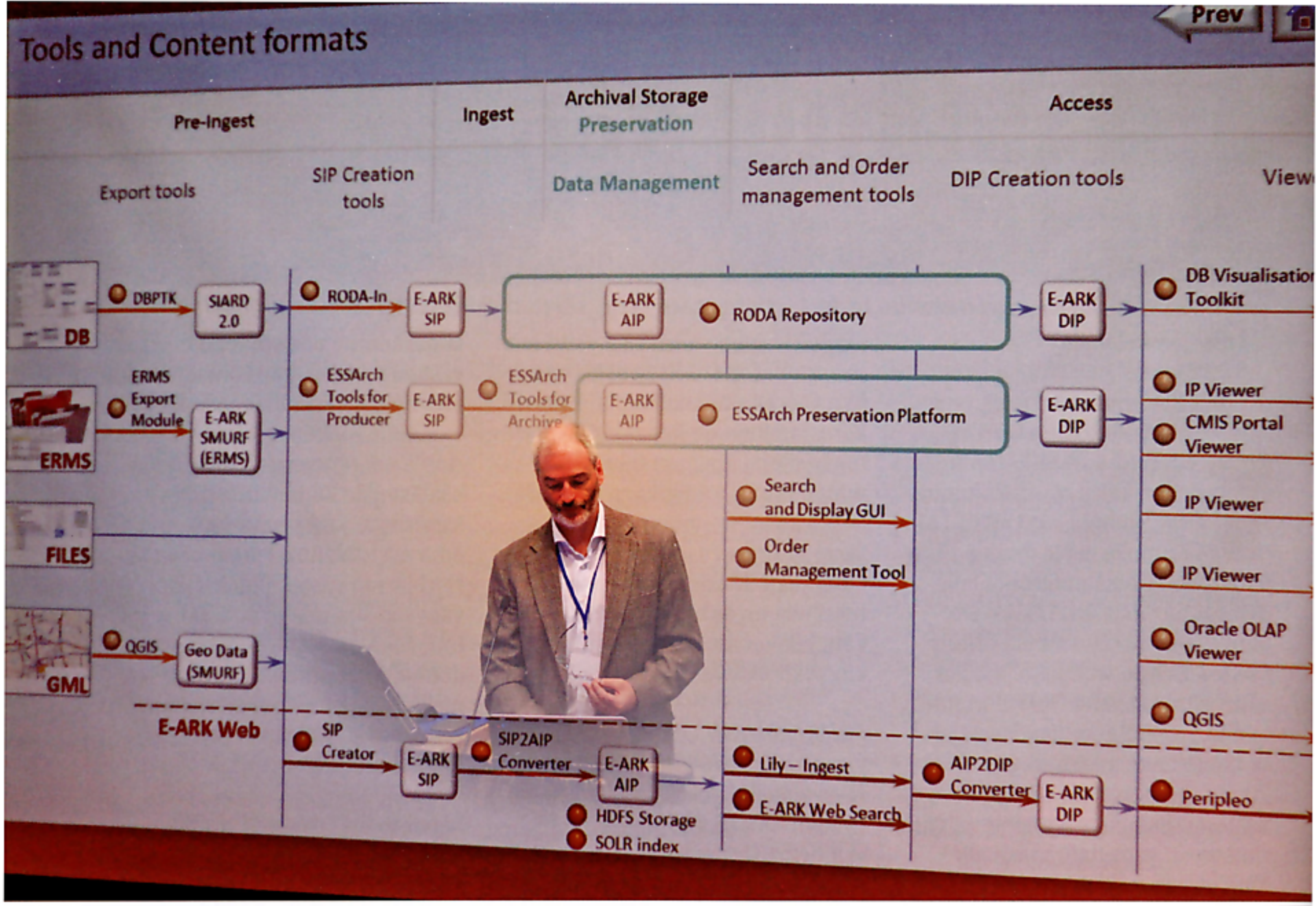
Norjassa on käytössä "once only"-periaate, joka edellyttää, että kansalaisen kerran syöttämä tieto on suoraan kaikkien julkisten toimijoiden käytettävissä. Hän esitti myös näkemyksiään siitä millaisia asioita alalla toimijoiden olisi syytä ottaa huomioon, tässä muutamia poimintoja: Nuorempi sukupolvi näkee asiakirjat kuolleina; arkistointi on muinaismuisto, tietoturva on ensiarvoisen tärkeää tiedonvaihtamisessa: jos odottaa viisi vuotta kuullakseen saako jotakin tehdä, on jo pudonnut kelkasta. Viimeinen asiakohta on mielestäni TKI-toiminnan ydin, vaikka nykyinen lainsäädäntö olisikin jotakin asiaa vastaan se ei saa tarkoittaa sitä, ettei-kö asiaa voisi tutkia ja kokeilla.

Ann Keen (Preservica Ltd) kertoi osallistujille luonnollisesti Preservicasta, jonka palveluita mm. oma arkistolaitoksemme käyttää. Herbjørnin vastakohdaksi Ann oli vahvasti sitä mieltä että RM ja IM (Information Management – tiedonhallinta) profiilia ja arvoa voi nostaa nimenomaan painottamalla riskejä ja sitä mitä voi seurata jos tiedot katoavat.

Sanoisin että tämä oli tyypillinen kaupallisen toimijan näkökulma, vaikkakin hyvin toimiva. Siitäkin huolimatta että, kyseessä on kaupallinen toimija, heidän verkkosivuillaan www.preservica.com/resource on tarjolla monia raportteja sekä työkaluja ilmaiseksi.

Esityksen jälkeen käyneen eloisan keskustelun keskeytti käytävältä kaikunut kuorolaulanta ja yllättäen koko "lallattava porukka" pelmahti suoraan auditorioon ja jatkoi esiintymistään. Kyseessä oli Norjan

Kirjoittaja esitteli Mikkelissä kehitettyä sähköpostien arkistointimallia.



István Alföldi, Poliphon Ltd:stä kertoi E-Arkin yleismallista.

kansallisarkiston kuoro. Kuoro sai ansaitusti koko tapaamisen suurimmat suosionosoitukset ja tuskin siksi että lopettivat laulamisensa.

Iltapäiväkahvin jälkeen oli vuorossa vielä kolme esitystä ennen kierrosta arkistolla. Ensimmäisenä oli *Aleksandra Mrdavsic* Sloveniaan kansallisarkistosta. Hän kertoi seulontakäytännöstä Sloveniassa ja siitä miten he olivat haastatelleet asiakirjojen luojia saadakseen paremmat lähtökohdat seulonnan toteuttamiseen.

Seuraavana kerroin itse päivitykset sähköpostien arkistointiratkaisusta, josta voitte lukea lisää tämän Failin toisessa jutussa. Yksi taustaltaan tuntemattomaksi jäänyt britti kyseenalaisti PDF/A formaatin käyttämisen ja ehdotti html5 muotoa sähköpostien arkistoinnissa. Sinänsä varteenotettava vaihtoehto etenkin kansalaisarkiston käyttöön, mutta toistaiseksi seuraamme arkistolaitosten näkemystä PDF/A formaatista. Puheenvuoroni jälkeen kuulin myös että Mamkissa käytetty OSA-lyhenne on ollut käytössä jo reilu 10 vuotta Blinken Open Society Archives nimisellä toimijalla Unkarissa (http://www.osaarchivum.org/). Heidän edustajansa oli kuitenkin sitä mieltä että saman lyhenteen käyttäminen ei aiheuta ongelmia.

Päivän viimeisenä esityksenä oli Norjan kansallisarkiston *Erik Gunnar Aabergin* puheenvuoro Arkade-nimisestä yleiskäyttöisestä sovelluksesta, jonka tavoite lyhykäisyydessään on "Kuka tahansa missä vain voi testata minkä tahansa datasetin" Sovelluksen pitäisi olla valmis huhtikuussa 2017 ja sen rakentaa 13 työntekijän firma nimeltä Arkitektum AS.

Uskallan karrikoidusti todeta että vastaavan julkishallinnon käyttöön tulevan sovelluksen rakentamisurakka olisi Suomessa mitä luultavammin mennyt raskaan kilpailutuksen jälkeen jollekin suurelle toimijalle ja se olisi valmis aikaisintaan 2025.

E-Arkissa kehitettyjä työkaluja ovat myös pilotoineet Euroopan ulkopuoliset kuten USA:n NARA sekä Uuden-Seelannin kansallisarkisto.

Arkistokierros vei osallistujat aina arkiston alla kulkevaan tunneliin asti. Luonnollisesti myös arkistomakasiinit kierrettiin nopeasti läpi mutta mihinkään erityiseen aineistoon emme kierroksella tutustuneet. Fyysistä tilaa arkistossa on esittelijän kertoman mukaan noin kolmen jalkapallokentän verran, hyllykilometreistä ei ollut muuta

mainintaa kuin että heiltä alkaa olla tila lopussa.

Keskiviikon ohjelma oli sisällöllisesti Digitalian kannalta vähemmän kiinnostava. Osittain koska kolme puheenvuoroa liittyivät E-Arkkiin, jotka käsiteltiin tarkemmalla tasolla torstain E-Ark päivässä. Toisekseen siksi, että luvassa oli monia maakohtaisia esityksiä, esim. *Karin Bredenbergin* esitys Ruotsin yleismäärittelystä (common specification) sekä päivän ensimmäinen puheenvuoro, jossa keskityttiin Norjan yhteisen arkistointiratkaisun vatvomiseen.

Mainittakoon kuitenkin jo nyt, että E-Ark projekti on päätösvaiheessa ja uutta rahoitusinstrumenttia etsitään kuumeisesti. Tämänhetkisistä H2020, CEF ja ISA rahoituslähteissä ei ole E-Ark 2 projektiin soveltuvia hakuja auki. Kaikkiin on oltu yhteydessä ja mahdolliseksi rahoituslähteeksi on todennäköisesti valikoitumassa H2020 joulukuussa 2016 aukeava BigData haku.

E-Arkissa kehitettyjä työkaluja ovat myös pilotoineet Euroopan ulkopuoliset kuten USA:n NARA sekä Uuden-Seelannin kansallisarkisto. Tulevaa rahoitusta varten nykyiset toimijat haluat laajentaa osaajajoukkoaan mm. arkistoilla (tämä vinkkinä meidän Arkistolaitokselle), e-hallinto asiantuntijoilla, sovelluskehittäjillä sekä yliopistoilla ja tutkimuslaitoksilla. Mamk, tai siis tuleva Xamk ja Digitalia-tutkimuskeskus ilmaisivat halukkuuteensa olla mukana projektissa. Tähän koordinaattori vastasi lyhyesti ja ytimekkäästi "ehdottomasti kyllä".

Lisäksi kuulimme myös E-arkin yleismallin (general model) lopullisesta versiosta, joka on myös E-arkin verkkosivuilla. Päivän lopuksi käytiin ryhmien vetäjien johdolla läpi aamun ryhmätöiden tulokset. Näissä ei mielestäni mitään mainitsemisen arvoista tai normaalilla maalaisjärjellä pääteltävissä olevaa.

Varsinaisen DLM-Forumin päätteeksi suomeakin hienosti puhuva *Leonard Callus* Maltan kansallisarkistosta toivotti kaikki tervetulleeksi seuraavaan DLM-Forumin jäsenkokoukseen 11-12.5.2017 Maltalle. Jälleen ainakin kirjoittaja pyrkii olemaan paikalla koska toukokuussa kelit Maltalla ovat jo hyvin suotuisat.

"hän totesi kaikkien nähtävien demojen perustuvan täysin avoimen lähdekoodin ratkaisuihin ja kaikkiaan pyrkimyksenä on saada aikaiseksi täysi digitaalisen arkistoinnin ekosysteemi."

Viron kansallisarkiston *Kuldar Aas* avasi E-Ark päivän. Avauspuheessaan hän totesi kaikkien nähtävien demojen perustuvan täysin avoimen lähdekoodin ratkaisuihin ja kaikkiaan pyrkimyksenä on saada aikaiseksi täysi digitaalisen arkistoinnin ekosysteemi. Projektin aikana samaan käyttötarkoitukseen on tehty useita työkaluja, joka mahdollistaa suoraan niiden tulosten ristiinvertailun.

Päivän ensimmäisen varsinaisen puheenvuoron piti "toveri" *Tarvo Kärberg* myöskin Viron kansallisarkistosta. Hän kertoi modulista, jolla ERMS (Electronic Records Management System) järjestelmästä voidaan viedä tiedot ulos E-Arkin smurf-formaatissa. Teknisestä toteutuksesta ei noussut kysymyksiä, mutta ERMS-järjestelmien epämääräisten kirjoitusvirheitä sisältävien tietojen suora vienti arkistoon nostatti tietyillä henkilöillä karvat pystyy. Kuldar totesi että etenkin arkistonhoitajien asenteiden on muututtava.

Seuraavaksi *Luis Faria* Keep Solutionilta esitteli sekä SIARD2 modulin sekä RODA-in työkalut. Yksinkertaistettuna SIARD lukee tiedot yhdestä tietokannasta ja kirjoittaa ne yleisluontoiseen formaattiin joka voidaan viedä takaisin joko samaan tai toiseen tietokantaympäristöön. Siis työkalu, joka helpottaa tietokannan siirtämistä järjestelmästä toiseen. RODA-in puolestaan tekee E-Ark tai muuta formaattia noudattavia SIP-paketteja, jotka voidaan viedä suoraan esim. RODA arkistoon. Mottona RODA-in työkalulla on "create thousands of sips in just few clicks" ja sanottava on, että selvästi vähemmillä klikkauksilla se onnistui kuin ES Solutionin kehittämällä vastaavalla SIP-työkalulla. Sitten esiteltiin E-Ark web -työkalu, joka pyrkii olemaan kaikki E-arkin toiminnot yhteen nivova opetusympäristö, joka pyörii AIT:n (Austrian Institute of Technology) valvonnassa kohtalaisen järeän klusterointiratkaisun päällä.

E-Arkin työkalujen esittely herätti ainakin Digitalian henkilökunnassa ajatuksia esimerkiksi tietokantavisualisaattorin (db-visualization-toolkit) hyödyntämisestä. Yksinkertaistettuna se siis osaa näyttää visuaalisesti tietokantojen sisältöä ilman tietokantasovellusta. Kaikkiaan tämänkertainen DLM-Forumin jäsenkokous oli antoisa kokonaisuus ja sisällöllisesti hyvällä tasolla muutamia poikkeuksia lukuun ottamatta.