

Intro: Tervetuloa kuuntelemaan CSC:n For A Better Future -podcastia!

Teemu Roos (jatkossa TR): Moi! Mun nimi on Teemu Roos ja mulla on täällä studiossa kaverina CSC:n Alekski Kallio. Me keskustellaan tänään it-vallankumouksesta.

Haluatko kertoa meidän kuulijoille ensinnäkin, että kuka olet?

Alekski Kallio (jatkossa AK): Joo eli olen Alekski Kallio CSC:ltä. Teen siellä töitä tekoäly- ja data-analytiikkaryhmän vetäjänä.

Taustaltani olen koodarinörtti ihan alunperin, todella pikkulapsesta asti. Ja sitten yliopisto-opintojen aikana vaihdoin tiedonlouhinnan piiriin, silloin oli vielä se sana mitä ei nykyään enää tunnetakaan, mutta kuitenkin data-analytiikan parissa oon hommia tehnyt siitä eteenpäin.

TR: Mä oon Helsingin yliopistolla edelleen, eli mä oon jäänyt sinne tielle, mutta hyvin pitkälti sama tausta. Tunnustaudun ylpeänä nörtiksi ja olen opiskellut itse asiassa samoissa paikoissa kuin Alekski. Ja nykyään oon proffana tietojenkäsittelytieteen osastolla ja vedän Suomen tekoälykeskuksessa FCAI:ssa opetusohjelmia.

Me ollaan Aleksin kanssa lukittauduttu tänne huoneeseen ja ajateltiin, että me turistaan tässä hetken aikaa tekoälystä, ohjelmistoista, mitä CSC tekee ja katsotaan mihin me päädytään eli tämmöinen hyvin chilli keskustelusessio tiedossa. Me juodaan tässä kahvia aamutuimaan ja pidetään teidät toivottavasti kiinnostuneina.

Hei, mä ajattelin, että jos me lähdetään liikkeelle ihan sana-assosiaatioilla, joilla ikään kuin rikotaan jättä. Jos mä kysyn sulta muutaman sanan ja sä sanot mitä sulle tulee ekana niistä mieleen.

AK: Joo, tehdään näin.

TR: Kasvisruoka?

AK: Vegaani.

TR: Tampere?

AK: Turku

TR: Yliopisto?

AK: Helsinki.

TR: Ensilumi?

AK: LUMI-supertietokone. Tämä oli maksettu mainos.

TR: Tiedätkö, mikä olisi ollut seuraava sana? Lumi. Mä arvasin, että sulle voi tulla tämä assosiaatio. Me avataan myöhemmin kuulijoillekin, että mistä tässä vähän sisäpiirijutussa on kyse, ettei se jää ihan näin arvoitukselliseksi. Mutta hyvä, selvästi meillä on ajatukset aika hyvin samalla taajuudella.

Mua kiinnostaisi kuulla mitä sä ajattelet tekoälyn, koneoppimisen ja ohjelmistojen yhdistämisestä? Joskus aikaisemmin sanottiin, että ohjelmistot syö hardwaren eli laitteistot. Eli tavallaan enemmän asioita pystytään tekemään ohjelmistojen avulla ja siitä tulee hardwaren sijaan merkittävä osa

liiketoimintaa. Vai onko tämäkin nyt vähän muuttunut, koska nykyään puhutaan taas GPUsta ja muista spesialisoituneemmasta hardwaresta?

Vai onko uusin trendi nyt se, että koneoppiminen tai tekoäly syö ohjelmistot? Onko niin, että tulevaisuudessa meidän ei tarvitse enää koodata? Koneet ikään kuin koodaisivat jollain tasolla itse itsensä?

AK: Ei se varmaan ihan näin mene, mutta kyllä mä uskon, että tässä on totuuden siemen. Tämä yhdistelmä: ohjelmointi, ohjelmistotekniikka ja tekoäly, se on hyvin tehokas liitto.

Tää voi johtua mun henkilökohtaisesta taustasta koska olen ohjelmoija, josta on tullut tekoälyihminen. On ikään kuin olisi kaksi vasaraa käsissä, alan hahmottamaan kaiken asioina, joita niillä kahdella vasaralla voi paukuttaa. Mutta mä olen aika vakuuttunut, että tässä on totuus taustalla, että ohjelmistot syö hardwaren. Eli ohjelmistot tulevat joka paikkaan esim. tietokone, sekin on oikeastaan lähinnä ohjelmisto nykyään.

TR: Tai auto.

AK: Tai auto. Auto on aika pitkälti ohjelmistotuote ja jatkossa vielä ehkä enemmän. Nyt ollaan sitten tilanteessa, että meillä on todella paljon ohjelmistoja ja joku ihminen kirjoittaa ja erityisesti ylläpitää sen kaiken itse. Se on valtava työ tehdä käsin. Mikä sitten on se tapa, jolla tätä voidaan tehostaa? Jos ajattelee ohjelmistokehityksen historiaa, niin kaikenlaista on yritetty – yleensä vähän huonolla menestyksellä. On kehitetty kaikenlaisia automaatteja, jotka tuottavat ohjelmistoja. Mutta eipä niistä nyt on jäänyt kovin paljon käteen viimeisen viidenkymmenen vuoden aikana.

Nyt sitten tekoälymenetelmät ja koneoppiminen muuttaa tämän tilanteen. Sitä voisi kysyä, että millä perusteella, jos 50 vuotta yritetty, niin miksi just nyt onnistuisi? Tietysti menetelmät on aika erilaiset ja mä itse ajattelen, että tähän asti ohjelmoijat ovat itse yrittäneet keksiä uutta tapaa ohjelmoida, ne on toimineet tietyn logiikan työkalupakilla, ja se mitä on löytynyt on löydetty.

Mutta nyt kun tulee tekoälytutkijat, joilla kaikki pohjautuu tilastotieteeseen, tilastolliseen ajatteluun ja rakentavat tätä samaa asiaa vähän eri logiikan työkalupakilla, niin on mahdollista, että paljon tapahtuu ja ohjelmistot tulee muuttumaan ja tekoäly sujahtaa atk:n sisään sujuvasti.

TR: Tuossa on monta mielenkiintoista kohtaa mistä sä voisit kertoa tarkemmin. Ensimmäiseksi kun sä sanoit, että ei ole tapahtunut erityisen paljon kehitystä siinä, että automatisoitaisiin ohjelmistojen tuotantoa, vaan ihmiset kirjoittaa sitä koodia. Mutta meillä on nykyään tosi korkean tason ohjelmointikieliä, siis siinä mielessä, että koodia täytyy kirjoittaa vain vähän, jotta tosi monimutkaisia asioita saadaan tapahtumaan. Alunperinhän koodattiin reikäkorteilla ja koodattiin tavallaan konekieltä, ja sitten tuli näitä jotain C-kieliä ja muita, jotka oli suhteellisen matalan tason ohjelmointikieliä ainakin nykykatsannossa. Nyt meillä on Pythonia ja sitten vielä Pythonin päälle on jotain syväoppimiskirjastoja, jotain Pytorchia ja TensorFlowta. Siinä kohtaa mennään sitten koneoppimiseen varmaankin, mutta vaikka ei puhuttaisi vielä koneoppimisesta, niin onhan Python ja jotkut korkean tason kirjastot, mitkä voi ladata valmiina, niin tavallaan osa sitä automaatiota.

Tuleeko semmoista "no code"-koodausta jossain vaiheessa, et sä vaan klikkailet jotain valintoja tai drag & dropilla laitat Scratch-ohjelmointikielessä, missä lapset voi opetella liikuttelemaan kissaa ruudulla?

AK: Varmaan sitäkin tulee. Tuo mitä sä sanoit, että ohjelmointikirjastot ja -kielet on korkeamman tason, niin on ollut just sitä oikeaa, aitoa kehitystä.

Kun nyt on näitä työkaluja, jotka tuottavat Java-koodia generoimalla valtavia määriä niin sanottua boilerplate-koodia, niin itsekin ajattelen, että eikö tässä nyt ole vaan valittu väärät kirjastot ja väärät työkalut, jos sitä pitää tämmöisellä typerällä skriptillä tuottaa? Miksei koodattaisi semmoisella kielellä, millä voi toimia oikealla tasolla. Kehitystä on kyllä tapahtunut, mutta että se ei ehkä ollut tässä koodin generoinnissa, vaan niissä kirjastoissa ja korkeamman tason kielissä ja virtuaalikoneessa ja mitä ohjelmistotekniikassa on tässä vuosien varrella tullutkaan.

TR: Miten sitten kun mennään tosiaan tähän kohtaan, että alkaa koneoppiminen tulla tähän kuvaan mukaan ja sä sanoit et se on oikeastaan meidän ajattelutyökaluista myös osittain kiinni. Haluatko sä avata vielä tuota vähän? Sanoit logiikkaan tai jollain tavalla loogisilla, vedenpitävillä päätelmillä. Ja sitten tilastollisen ajattelun avulla. Mitä sä tarkoittit sillä?

AK: Joo, mä voisin käyttää esimerkkinä Googlea, että jos ajattelee... no ennen Googlea oli Yahoo- ja AltaVista-hakukoneet. Olen jo niin vanha, että muistan tämmöisiäkin.

TR: Kyllä, sama täällä. Kerro kuinka vanha olet sanomatta, kuinka vanha olet. Tämä toimii esimerkkinä siihen.

AK: Kyllä, kyllä. Olen käyttänyt AltaVistaa, se oli aika kova juttu silloin kun se tuli. Se oli ehkä vähän sellaista perinteisempi, ohjelmistoinföörien tekemää hakukone. Vaikka AltaVistassa pystyikin tekemään aika monimutkaisia hakulauseita, kun oli massa niitä web-sivuja, ei oikein löytynyt se oikea. Ratkaisuna oli, että tehdään rikas hakukieli millä voi niin kuin naputella erilaisia ehtoja, poikkeuksia ja muita säättöjä. Sitten tuli suunnilleen se oikea hakutulos. Sitten taas Yahoo keräsi internetin hakemistoa, mikä on tietysti tänä päivänä aika hupsultakin kuulostava ajatus, mutta siellähän oli kategoriat, että nyt haluan katsoa esim. harrastukset ja jalkapalloaiheisia sivustoja.

Sitten tuli Google, joka oli nyt ehkä tämmöinen uuden ajan tekoälyfirma, joka sitten rupesi katsomaan dataa. Eli aluksi web-sivuja datana Page rank -algoritmilla, että kuka viittaa mihinkäkin ja että sinulla on tällainen sivu, johon olet laittanut hakusanoja, jotta nousisit ylös listauksissa ja sitten kun muut viittaa sinuun jollakin tietyllä termillä, niin se kertookin itse asiassa aika paljon enemmän siitä sinun sivusta. Ja kun ihmiset rupesivat tekemään hakuja ja käyttämään palvelua, niin siitäkin kertyy dataa. Sen jälkeen pystyikin sitten heittämään pois kaikki monimutkaiset hakukielet ja sen valtavan työn, minkä jonkun tällaisen hakemiston ylläpidossa oli ja sitten tuli vaan yksi laatikko, Googlen hakulaatikko. Kirjoitit siihen mitä tarvitset, ja sitten - paremmin kuin ne muut hakukoneet – se löysi sen mitä oikeasti olit etsimässä.

Niin tässä esimerkki siitä, että Googella oli ajattelua. Se ajatus tuli akateemisesta tutkimuksesta alun perin, että entäs jos katsotaankin dataa ja ajatellaankin vähän tilastollisesti ja datan kautta.

TR: Tosi mielenkiintoista... Mulle tuli mieleen, että mä oon kuullut tällaisesta käsitteestä kuin Computational Thinking 2.0. Tämä on ihan ”uusinta hottia” tietojenkäsittelytieteen opetuksen rintamalla. Itä-Suomen yliopistosta muun muassa on tullut tällaista tutkimusta, ja se liittyy siihen, että meidän pitäisi opettaa tätä jo peruskoulussa ja varmaankin jo esikoulussakin. Mä en muista mikä tämän suomenkielinen termi oli, varmaan joku laskennallinen tai ohjelmoinnillinen...

AK: Se on aika vaikea termi.

TR: Joo, mutta tiedät mitä tarkoitan. Alkuperäinen versio on jo aika hyvin integroitu meidän peruskoulun opetussuunnitelmaan, niin että lapset oppii peruskoulussa ymmärtämään, että tietokoneet suorittaa tällaisia askeleita algoritmista, että esimerkiksi... ota voileipä vasempaan käteen ja ota voiveitsi oikeaan käteen. Sitten ota voita voikupista ja sitten levitä sitä näin jne. Eli vähän niin kuin ruuanlaittoresepti. Eli proseduraalisesti ja askel askeleelta hyvin yksityiskohtaisesti määriteltynä etenevää on tietokoneen ohjelmointi.

Ja nyt seuraava havainto liittyy siihen, että tällainen paradigma, jossa kaikki on rikottu pieniin paloihin, vähän niin kuin se AltaVistan hakulauseke, että tämä sana saa esiintyä, mutta tuo sana ei ja sitten pitää jompikumpi näistä sanoista esiintyä, ja niin edes päin. Sitä nyt ollaan korvaamassa peruskoulun ohjelmoinnillisen ajattelun opetuksessa, et otetaan rinnalle sellainen ajattelu, että välillä ohjelmistoja voi tuottaa niin, että näyttää esimerkkejä, vaikka että tässä on hymyilevät kasvat ja tässä on irvistävät kasvat. Tämän ajattelun työkaluja haluttaisiin opettaa jatkossa myös kouluissa ja siinä on kyse osittain tästä samasta paradigmanmuutoksesta.

AK: Kyllä. Kuulostaa hyvin oikeasuuntaiselta.

Mietin tässä omaa lukioaikaani, josta on jo hetki kyllä kulunut, mitä silloin oppi vaikka matematiikassa. Opin derivointia ja integrointia, osaan monimutkaisten kappaleiden tilavuuksia ainakin periaatteessa, jos vielä muistaisi miten se menee... Mutta tilastotiedettä ei ollut ollenkaan. Oli todennäköisyyslaskennan kurssi ja siellä kerrottiin, että on olemassa normaalijakauma. Mutta ei siitä sen enempää. Se oli aika heikko johdatus aiheeseen. Sitten diskreettiä matematiikkaa, logiikkaa ynnä muuta oli hyvin vähän. Se on hienoa, että painotus siirtyy ja sillä osaamista kyllä tarvitaan. Ihan vaan tämän maailman ymmärtämiseen.

TR: Joo, samaa mieltä. Nyt on esimerkiksi ollut tapetilla vaikkapa lääketieteellisten testien väärät positiiviset ja väärät negatiiviset ja sensitiivisyys ja spesifisyys ja muut tällaiset termit, jotka ovat suht yksinkertaisia, mutta niitä on ehkä vaikea hahmottaa, kun niitä ei ole käyty läpi kouluissa. Se on käytännössä arkielämässäkin järkevä tietotaito. Sanoisinko, että ton tyyppistä tietämystä tarvitaan jopa yleisemmin kuin derivointia tai integrointia.

AK: Kyllä. Jos ottaa nyt hyvin ajankohtaisen ja mielipiteitä jakavan esimerkin kuin korona. Se, että meillä on nyt tällä hetkellä sairaaloissa noin fifty-fifty rokotettuja ja rokottamattomia. Sitä jotkut pitää vähän kuin näyttönä siitä, että ei sillä rokotteella ole mitään vaikutusta. Tämä esimerkiksi, että jos opetettaisi ihmisille tilastollista ajattelua, niin ei tarvitsisi kahdenkymmenen vuoden päästä painia sen ongelman kanssa, että ei hahmoteta mikä se perusjoukko tässä tapauksessa nyt olikaan.

TR: Nimenomaan. No miten sä näet, kun tällaisia ajattelutyökaluja ja koneoppimisen työkaluja integroidaan ohjelmistotuotantoon, niin mitä se tulee muuttamaan, jos nyt heittäydytään vähän spekulatiivisemmalle tuulelle. Sä vastasit jo siihen kysymykseen, että ei varmaankaan tule korvaamaan koodareita tekoälytyökaluilla. Kun tekoälystä ja työn murroksesta puhutaan, niin yleensä se fiksumpi kysymys on että miten koodareiden työ tulee muuttuman? Ja miten ohjelmistotuotannon kanssa tekemisissä olevien ihmisten ja organisaatioiden elämä tulee jatkossa muuttumaan?

AK: Mä ajattelen niin, että voi olla, että ohjelmoijia tarvitaan vähemmän jatkossa. Tavallisia ohjelmoijia. Voi käydä niin, että tavallaan se ihmisen tekemä ohjelmointi erikoistuu ja siitä tulee vaikeampaa. Mutta

sitten tulee data scientisteille eli datatieteilijöille tai hybridi datatieteilijä/ohjelmoija-tyyppisille ammattiteille enemmän kysyntää.

Nykyään hyvin yleinen tapa tulla it-alalle on, että aloitat perustason ohjelmoijana. Ei tarvitse olla kauhean hyvä siinä hommassa vielä aluksi, kun sinulla on joku asiakas, jolla on joku spesifi tarve, vaikka jonkinlaisen koneen osan varastokatalogin ylläpito. Jos sä oot ainoa ihminen maailmassa, joka sitä koodaa ja tuot hyötyä ja lisäarvoa ja saat siitä palkan kasaan. Tämmöiset tehtävät voi vähentyä ja tulevat älykkäät ohjelmistot voi tämän tarpeen hoitaa. Mutta sen sijaan tarvitaan joku, joka muotoilee sen datan, johon tämä ohjelmisto perustuu. Et sen sijaan, että kun nyt muotoillaan koodia, niin jatkossa ollaankin muotoillaan spesifiä dataa ja siihen analytiikka päälle. Joku tämmöinen siirtymä vois tapahtua.

TR: Nyt herää tällainen huoli, kun puhutaan koodaripulasta ja erilaisista ratkaisuista siihen. Ja toisaalta siitä, että juniorikoodareille jotka hakevat sitä ensimmäistä tai toista jobia, on yllättävän vähän työpaikkoja. Kysyntä painottuu kokeneiden seniorikoodareiden puolelle, jopa niin että pyydetään 15 vuotta kokemusta jostain ohjelmointityökalusta, joka on ollut olemassa 10 vuotta. Väheneekö nämä aloittelevien koodareiden työpaikat jatkossa entisestään?

AK: Ei välttämättä... sitten vaan tulee datavetoisempi reitti näille markkinoille. Ja ehkä moni seniorikoodari aloittaa datan kanssa työskentelyn ja sitten keskittyykin vaikka backend-ohjelmointiin jatkossa.

TR: Tarkoittaako se jatkossa sitä, että juniorikoodareita tarvitaan edelleen, mutta niiden työtehtävät vastaavat ehkä enemmän niiden seniorikoodareiden työtehtäviä, niin että ne pystyisi hahmottamaan tavallaan sitä kontekstia paremmin. Sehän on sellainen taito, joka kypsyy tosi hitaasti. Sitä on tosi vaikea opettaa. Mä oon itse tekemisissä datatieteilijöiden kouluttamisen kanssa yliopistolla, ja helppohan niille on opettaa koodaamista ja teknisiä taitoja. Mutta sitten tällaiset ei-tekniset taidot, miten linkitetään se ohjelmisto siihen kontekstiin, jossa sitä tullaan käyttämään käyttäjän tarpeisiin ja ylipäänsä liiketoiminnallisiin tarpeisiin. Se on tosi vaikeata. Pitäisikö meidän nyt yrittää jatkossa kouluttaa niin, että suoraan koulun penkiltä olisi jo paremmat valmiudet koodin ja datan ja bisneksen välimaastossa toimimiseen?

AK: No ainakin se olisi hienoa, jos näin voisi tehdä. Mutta taitaa olla tosi vaikeaa, niin kun sanoitkin, että katsoo koodinpätkää ja miettii, onko tämä tehokas ja paras mahdollinen. Mutta sellainen hahmotus voisi olla ainakin nykypäivän ajattelussa, että tässä on tämmöisiä käskyjä, ovatko ne ymmärrettäviä, jos tähän tulee joku toinen ohjelmoija? Onko tämä se ymmärrettävin mahdollinen tapa vai voisinko mä tehdä tämän jollakin toisella tavalla, joka on helpompi muiden tulkita.

Ja sitten myös se, että mites tämä suhtautuu muutoksiin? Meillä on semmoisia asiakkaita ja semmoinen kenttä missä me toimitaan, mutta että tämä tässä voi muuttua. Tämä yksi piirre, onko se helppo tässä kohdassa ottaa huomioon. Se on sitä kontekstin ymmärtämistä mistä puhuit.

Sitä on kyllä vaikea muuten saada kuin kokemuksen kautta. Mä ajattelisin näin, että aloittelevat ohjelmoijat tulevat enemmän datan kanssa työskentelemään ja data scientist -nimikkeelle töihin, koska data on yleensä aina enemmän kiinni siinä kontekstissa ja on helpompi johdatus tähän maailmaan. Että katsoppas nyt näitä meidän datalähteitä ja kun teet näitä tiettyjä operaatioita ja sitten rupeaa syntymään koodia ja ymmärrystä siitä kokonaisuudesta.

Mutta kyllähän se on hyvä kysymys, että jos tarvitaan tosi guruohjelmoijia, niin mistä ne oikein tupsahtaa, jos niitä ei ole jossakin kypsymässä kokoajan?

TR: No siinä jää mietittävää vielä monella viisaalle, jotka voi sitten hakata päitään tämän kanssa yhteen.

Toinen isompi aihe, mikä minua kiinnostaa, kun tässä alussa oli se sana-assosiaatio ja puhuttiin LUMI-supertietokoneesta.

CSC on sellainen organisaatio Suomessa ja jopa nyt Euroopan mittakaavassa, joka tästä asiasta jotain tietää. Ja sinä olet sellaisessa tiimissä, jossa vedetään hankkeita, jotka liittyvät myös superlaskentaan ja todella vaatimaan koneoppimiseen ja muuhun.

Mitä teillä on tällä hetkellä käynnissä? Nyt se LUMI-supertietokone on tässä kohtapuoliin. Joko se on painettu nappia, joko se hurisee siellä Kajaanissa?

AK: Joo, siitä puolet on olemassa ihan fyysisesti konesalissa ja se on käynnissä, mutta ei ole vielä asiakaskäytössä. Se on nyt niin sanotusti ylösajovaiheessa. Se iso kapasiteetti tulee tässä kevään aikana, eli se varsinainen raskain laskentatavara eli GPU-osiot asennetaan ensi kevään aikana.

TR: Kerro vähän enemmän vielä, mikä se LUMI on? Onko se joku tietokone? Onko se joku kokonainen farmi? Miltä se näyttää, miltä se kuulostaa tai kerro siitä jotain, kun siitä niin paljon puhutaan? Ainakin meillä puhutaan yliopistolla ja tekoälypiireissä puhutaan, että se tulee muuttamaan monia asioita. Mikä se oikeasti sitten on se LUMI?

AK: LUMI on eurooppalainen supertietokone. Sen omistaa tämmöinen EuroHPC-yhteistyöelin, jossa on monia maita mukana Suomen lisäksi. Tämä EuroHPC-yhteistyöelin järkesti kilpailutuksen tulevista esi-exascalen koneista. Esi-exascale viittaa siihen, että kuinka paljon laskentatehoa on, ja sitä lasketaan flopeissa. Exaflop on nyt raja, mitä seuraavaksi tavoitellaan ja nämä koneet on tavallaan se askel ennen seuraavaa varsinaisen exa-rajan ylitystä. Eli tämmöisiä isoja koneita tulee Eurooppaan 3 kappaletta ja kilpailutuksessa Suomi oli mukana osana LUMI-konsortiota eli me ollaan sen verran pieniä, ettei yksin voida tämmöiseen osallistua. Ja CSC kasasi tämmöisen konsortion eri maista, jotka lähtivät mukaan.

TR: Ketkä muut maat siinä on mukana?

AK: Siinä on 10 maata, että en nyt ihan ulkoa muista kaikkia, mutta pohjoismaita ja ainakin Belgia ja Sveitsi, tämän tyyppisiä... sekä Baltian maita... Pärjättiin hyvin kilpailutuksessa, saatiin se hankinta.

TR: Minkä takia pärjättiin hyvin? Onko Suomella jotain tiettyjä etuja siihen?

AK: Kyllähän se perustuu siihen, että toki meillä oli hyvä, huolella tehty hakemus ja hyvä infrastruktuuri. Ja sitten tietysti tämä meidän sijaintipaikka eli se tulee Kajaaniin vanhaan paperikonehalliin. Se oli hyvä lähtökohta, että siellä on valmis sähkönsyöttö, joka on tehty aivan eri mittaluokan juttuihin kuin tietokoneisiin. Eli se ottaa sujuvasti vastaan isonkin laitteiston. Kylmä ilmasto ja se että meillä menee paljon vähemmän sähköä hukkaan, koska meillä on mahdollisuus kierrättää se hukkalämpö Kajaanin kaukolämpöverkkoon. Operointikustannukset oli pienemmät, jolloin tietysti jää enemmän rahaa siihen varsinaiseen hankintaan, saadaan isompi, tehokkaampi laitteisto. Siinä oli varmaan ainakin nämä tärkeimmät avaimet siihen.

TR: Mitä se LUMI sitten tulee tarjoamaan? Varmaan paljon laskentatehoa. Onko jotain sellaista laadullista eroa tai onko jotain muuta, että ei tarvitse ostaa kapasiteettia sitten joltain kaupalliselta tarjoajalta, Amazonilta tai Googlelta... mutta onko jotain muuta mielenkiintoista mitä tapahtuu?

AK: LUMI tulee olemaan osa tätä jatkumoa, kun on koneita hankittu Suomeen, mutta tässä vaiheessa tapahtuu semmoinen tasapainotilan muutos, että LUMI:ssa suurin osa kapasiteettia on GPU-kapasiteettia eli alunperin grafiikkakäyttöön tarkoitettuja piirejä, jotka suurteholaskennassa alkaa olla se normaali tapa, että se teho mitä 1 €:lla saa on paljon suurempi GPU-puolella. LUMI on nyt pitkälti GPU-pohjainen laitteisto, mutta samaan aikaan siellä on myös pieni CPU-osio, joka on kuitenkin samankokoinen kuin esimerkiksi tämän hetken Suomen suurin kone. Ei sekään mikään pieni ole, mutta LUMI:n mittakaavassa pienempi. Sitten siellä on suuri muistin data-analytiikkaosio ja erilaisia tallennusjärjestelmiä – kolme erilaista tallennusjärjestelmää ainakin ja kaikki on kytketty yhteen.

Isoin idea on se että kaikenlaista kapasiteettia on, mutta se ei ole vaan varastohuone, minne on kerätty erilaisia tietokoneita, vaan ne kytketään tehokkaalla kytkentäverkolla toisiinsa niin, että se oikea kapasiteetti on käytettävissä siihen samaan laskentatyöhön tehokkaasti sellaisilla teknologioilla, mitä ei tuolta tavallisesta tietokonekaupasta tai verkkokaupasta saa. Että sehän se supertietokoneen lopullinen idea on, että kaikki kytketään yhteen tavalla, joka on tehokkaampi kuin mikään muu, mitä sillä hetkellä markkinoilla on ja siitä se tulee se “super” siihen eteen.

TR: Eli siellä on tallennuskapasiteettia. Siellä on laskentakapasiteettia. No, tarjoatteko te jotain siihen tueksi vai onko se niin, että syötät luottokorttitiedot tähän ja sen jälkeen voit ajaa siellä jotain ohjelmistoa? Onko CSC:llä jotain muuta tarjottavaa siihen, että ihmiset löytäisi ja pystyisi hyödyntämään sitä paremmin?

AK: Kyllä on. Merkittävä ero kun tätä tarjotaan tutkijoille ja yrityksille on se, että yleensä pilvipalvelun tarjoajat ei kovin herkästi vastaa puhelimeen, jos sinulla on ongelmia, ellet ole aika merkittävä asiakas. Mutta meillä CSC:llä on lähdetty siitä, että tarjotaan tukea ja myös erityisesti koitetaan ylipäänsä rakentaa yhteistyötä. Jos vaikka yrityksellä on tarpeita, niin me hyvin mielellään haetaan joku akateeminen osapuoli, joka tuntee sen asian, niin että syntyy semmoinen oikeasti hyvä yhteistyö laskentatehtävän ympärille. Ja toki LUMI:lla on oma User Support -tiimi, joka on koko Euroopan kattava. Ja LUMI-konsortion kattava tukitiimi, johon CSC:kin osallistuu. Ja tällä hetkellä meidän ryhmän yksi merkittävimpiä hommia on valmistautua siihen, että se GPU-osio tulee LUMI:in. Mitä se tarkoittaa, kun tulee mahdolliseksi tehdä todella isoja rinnakkaisia koneoppimistöitä ja onhan meillä valmiudet sitten auttaa ihmisiä ja omaa kokemusta riittävästi. Sitä tässä koko ajan rakennetaan.

TR: Niin mä uskon, että toi on yksi niistä merkittävimmistä eduista. Ei se välttämättä, että kuinka paljon kapasiteettia jossain tietokoneessa on. Totta kai sekin on tärkeää tai minkä hintaista sitä on käyttää, jos katsoo sitten maksajan näkökulmasta. Mutta se, että siihen tarjotaan asiantuntemusta ympärille, että miten sitä käytetään fiksusti eikä vaan olla tyytyväisiä, että siellä se hurisee ja tahkoaa euroja ja dollareita.

Sinä olet ilmeisesti sellaisessa tiimissä, jossa teillä on monenlaisia hankkeita. Eli myös tätä “hankehumpaa” mitä aina manaillaan meillä yliopistolla, mutta totta kai siinä on omat puolensa. Millaisia hankkeita teillä nyt sitten on ja miten ne kehittää tätä? Kehittääkö ne nyt sitten jotenkin datan ja ohjelmistojen kohtaamista ja auttaa asiakkaita löytämään niitä hyviä tapoja?

Itse oon törmännyt siihen ongelmaan, että vaikka meillä on uusia teknologioita eikä välttämättä niitten tarvitse olla uusiakaan teknologioita... koneoppimisessahan voidaan käyttää helposti kymmeniä vuosia vanhoja ideoita ja innovaatioita, jos niitä osataan vaan käyttää taitavasti. Mutta löytää just se kohta missä liiketoiminta, data ja algoritmi tai se menetelmäpuoli kohtaa, niin siinä on tavallaan se kohtaaminen, mikä täytyy tapahtua, sen käyttäjän tai sen organisaation, joka on kehittämässä jotain sovellusta ja sitten niiden teknisten ratkaisujen välillä.

Miten paljon te pystytte tulemaan vastaan ja onko teillä jotain tiettyjä toimialoja, joissa te operoitte ja joissa te pystytte miettimään, että miten supertietokonekapasiteettia esimerkiksi pystyisi hyödyntämään paremmin?

AK: Joo on kyllä kaksi eri toimialaa. Meillä on silleen helppo tilanne CSC:llä, kun me ollaan valtion erityistehtäväyhtiö, niin me voidaan laista lukea, että mitkä on meidän fokukset. Meidän suurin asiakaskunta on suomalaiset tutkijat korkeakouluissa ja valtion tutkimuslaitoksissa. Se on tärkein työllistäjä meidän porukalle. Toinen on valtionhallinto. Ollaan valtiokonsernin sisäinen toimija ja tehdään töitä ministeriöiden ja virastojen ja muiden vastaavien valtion alaisten laitosten kanssa. Tehdään tämmöistä tekoälykonsultointi-tyyppistä työtä, missä lähdetään aika laajalla skaalalla miettimään ratkaisua johonkin tiettyyn ongelmaan. Millä tavalla voisi tuoda koneoppimista ja data-analytiikkaa toimintaan mukaan? Ja mitkä on ne oleelliset datalähteet ja minkä tyyppisiä ongelmia voidaan ratkaista? Ja mitä pitää tehdä ennen kuin me voidaan edes tehdä koneoppimista.

Datan kanssa pitää ensin tehdä aika kauan työtä ennen kuin voi sitä ruveta käyttää sitten jonkun mallin kouluttamiseen. Mutta tärkein on tosiaan se tutkimusmaailma ja ne hankkeet – tämä niin sanottu hankehuppa mitä mekin tanssitaan... Siinä on aina ideana se, että mitä suomalainen tutkimus hyötyy siitä, jos me lähdetään mukaan? Se voi olla, että saadaan uusi palvelu tai palvelun osa rakennettua, joka sitten jatkaa elämäänsä sen jälkeen. Semmoisia tehdään aika paljon.

Sitten toinen on tietysti se, mihin me voidaan rakentaa osaamista... aika tärkeää on ollut suomen kielen käsittely, se on semmoinen aika erikoisalue, missä ei ihan hirveän monta toimijaa ole. Meillä on tuon Kansalliskirjaston kanssa ollut paljon yhteistyötä. Suomenkielisten dokumenttien luokittelumenetelmien läpikäynnissä ja kehittämässä, ja sitten he on rakentaneet semmoista Annif-työkalua, mihin mekin on osallistuttu ja on sitä aktiivisesti käytetty. Siinä on aika paljon taustalla CSC:llä tehtyä laskentaa, kun sitä varten on tutkittu erilaisia malleja ja niiden toimivuutta eri datoilla.

TR: Joo itsekin näin jotain demoja, kun kävin kirjastolla ja tutustuin tähän Annif:iin, joka on siis Finna-sana nurinpäin luettuna.

Tähän myöhemmin, mutta kun sä sanoit, että te teette korkeakoulujen ja Suomen valtionhallinnon kanssa yhteistyötä, niin onko se sun näkökulmasta niin että Suomi on tekoälyn edelläkävijämaa? Siinä mielessä, että meillä osataan hyödyntää valtion puolesta, julkisissa palveluissa digitalisaatiota, jos otetaan laajempi katsanto tähän. Onko se sun mielestä totta ja miten se näkyy meille? Onko se sellainen asia, joka on jossain hallinnon rattaiden sujuvammassa pyörimisessä vai näkykö se jotenkin kansalaisille? Onko teillä jotain sellaisia hankkeita, mitkä saisi tavallisille kansalaisille näkymään, että Suomi on digitalisaatiossa edelläkävijä?

AK: Tuo on erittäin hyvä kysymys. Mulla on valitettavasti aika vähän vertailukohtia, kun en ole muissa maissa tässä viime aikoina asunut, mutta minun käsitys on, että se voisi olla totta. Mutta toki ollaan vielä

kehitysvaiheessa, mutta ehkä tästä muutaman vuoden päästä käännytään tämmöiseen aika hyväänkin tilanteeseen.

TR: Mitä se voisi olla muutaman vuoden päästä?

AK: No tällä hetkellä on sellaisia merkittäviä hankkeita kuin Robo-hanke ja sitten Aurora AI -hanke, mitä valtionhallinnossa tehdään. Ne on semmoisia prosesseja, mitkä nyt sit kansalaisia koskee ja käydään läpi, että voisiko asioita tehdä jotenkin ihan eri tavalla tekoälymenetelmillä. Mennään aika syvällekin siihen miten ihmisen näkökulmasta palvelut oikeasti pitäisi mennä ja miten nämä palvelut voisi järjestellä uudelleen ja tietysti myös minkälainen työkalu tekoäly siinä voisi olla? Tämmöistä kehitystä on... tietysti se on aika kunnianhimoista, että on vielä vaikea ennustaa mikä on se lopputulema. Mutta ainakin kehitysasioissa ollaan edelläkävijöitä minun käsityksen mukaan.

TR: Mä oon tähän itse kiinnittänyt huomion, että ulkomailla kiinnostus Suomen tai julkisen hallinnon digitalisaatiohankkeita kohtaan on tosi valtava, sekä tähän Aurora AI:n että siihen robotiikkahankkeseen. Jostain syystä etenkin Aasian maissa sekä Japani että Korea esimerkiksi on ollut tosi kiinnostuneita... ehkä heillä on saman tyyppisiä väestön ikääntymiseen liittyviä haasteita tai siihen, että miten kansalaiset löytää digitaaliset palvelut tavallaan tasa-arvoisesti, että ei pelkästään palvelu niitä, jotka on tottunut käyttämään digitaalisia palveluita.

No miten sitten yksityisen puolen... mä oon ollut itse esimerkiksi mukana Työ- ja elinkeinoministeriön tämmöinen AI 4.0 -hanke, joka on tälle Suomen tekoälystrategialle tämän hallituskauden kärkihanke, ja siellä on kannettu paljon huolta siitä, että miten me tätä edelläkävijyyttä sitten levennetään? Että kun meillä on tämmöisiä keihäänkärkihankkeita, jotka voi olla kansallisia, mutta levennetään kärkeä niin, että otetaan pk-sektori ja pk-yritykset fokukseen myös, että saataisiin sitä tarjottua laajemmalle.

Miten sä näet että se parhaiten tapahtuu? Onko siinä jotain julkisia toimijoita? Onko meillä yliopistoilla tai teillä CSC:llä ja muilla julkisrahoitteisilla organisaatioilla oma roolinsa vai kannattaako se jättää yrityskentän omaksi murheeksi niin sanotusti?

AK: Ajattelisin, että lähtökohta on hyvä, koska Suomi on kuitenkin ohjelmoijien maa. Ja jos nyt uskotaan aiempiin puheenvuoroihin siitä, kuinka tekoäly ja ohjelmistokehitys kytkeytyy toisiinsa, niin se, että me ollaan tämmöinen ohjelmisto-orientoitunut maa, on hyvä lähtökohta myös tulla tekoälyorientoituneeksi maaksi.

Jos ajattelen julkisen puolen roolia, vaikka nyt ihan CSC:n roolia, niin mä itse näen, että varmaan se kaikista kovin juttu, mitä voi yritysmaailmalle tarjota, on data. Tätä on julkisella puolella paljon, mutta että sitten tullaankin näihin haastaviin kysymyksiin, että se on helppoa, kun Facebook voi vaan ottaa sinun datan ja joku EULA, jonka nyt puolihuolimattomasti hyväksyt, ja sitten kaikki on sisäisesti Facebookin käytettävissä. Mutta että nyt kun tehdäänkin tämmöinen demokraattinen versio siitä, niin se onkin aika paljon isompi työ. CSC:lläkin voisi olla merkittävä rooli neutraalina datan tallentajana siinä välissä. Jos näitä datavarantoja mitä meillä on, saataisiin hyötykäyttöön reiluilla pelisäännöillä, niin se olisi aika merkittävä kilpailukykyetä suomalaisille yrityksille. Ja myös muut edelläkävijäyritykset tulisi Suomeen tekemään tuotekehitystä, koska meillä olisi jotain semmoista tarjota, mitä ei muilla ole.

TR: Aivan eli tavallaan ajatellaan, että mitä julkinen puoli tekee, niin on tavallaan mahdollistaa ja tukea semmoista "government as a platform" -ajattelua, että sen päälle pystyy sitten kasvattamaan menestystä ja liiketoiminta.

No miten sä nyt ihan konkreettisesti sanoisit, että jos jossain on esimerkiksi pk-yritys tai sitten suuryritys, niin mitä tällä hetkellä kannattaisi tehdä, että pystyy hyödyntää näitä tulevaisuudessa tarjolla olevia palveluita? Kannattaako soittaa CSC:lle? Kannattaako ruveta laittamaan jotain omaa hankehumpaa tai koulutustua jonkun tietyn kyvykkyyden osalta?

AK: No ainakin niin sanotun hankehumpun osalta meillä CSC:llä on Business Finlandin kanssa hyvä yhteistyö eli Business Finland olisi varmaan se paras taho mihin soittaa. Sieltä löytyy business-alueen tuntemus ja he tuntevat CSC:n palvelut ja ylipäänsä tekoäly-asioita niin hyvin, että sieltä mä aloittaisin. Mutta totta kai sitten osaamisen kasvattaminen, nythän löytyy erinomaisia kursseja verkosta, mitä voi käydä tekemässä ja ylipäänsä se, että on hereillä ja seuraa kehitystä. Ja jos ei ole itse tekijä, joka pystyy näitä tekoälyratkaisuja rakentaa, niin etsii sitten yhteistyökumppaneita, semmoisia luottotoimijoita. Sillä on varmasti aika suuri arvo sitten jatkossa.

JATKA TÄSTÄ:

TR: CSC:hän täyttää 50 vuotta tänä vuonna eli on synttärjuhluvuosi. Tässä oli tietenkin vasta ensimmäiset 50 vuotta, ja nyt seuraavat 50 vuotta tulee tietenkin olemaan paljon mielenkiintoisempia ja muutostahti tulee nopeutumaan ja kaikkea muuta. Mulla ainakin itsellä on aina suuret estot, että mä rupeaisin ennustamaan tulevaisuutta, mutta että JOS me nyt kuitenkin ennustettais, että mitä tulee tapahtumaan seuraavan viidenkymmenen vuoden - puolen vuosisadan - aikana, niin millaisia terveisiä me voitaisiin lähettää tulevaisuuteen?

Tämäkin podcast tullaan laittamaan aikakapseliin, joka avataan 2071, silloin kun on CSC:n satavuotisjuhlat. Mitä terveisiä, mulla on tosi vaikea keksiä mitään... keksitsä Aleksi jotain terveisiä, mitä me voitaisiin lähettää nyt sinne 2071 vuoteen?

AK: Tää on aika vaikeeta. Varmaan terveiset on se, että tulevaisuuden ennustaminen on hirveän hankalaa. Mutta ehkä mä voisin ajatella näin, että yleensä kun ennustetaan tulevaa, niin otetaan ne sillä hetkellä päällä olevat kehityskulut ja niitä vaan viedään reippaasti eteenpäin ja unohdetaan kaikki muu.

Ehkä tällä hetkellä voisi olla, että kun näitä itseajavia autoja on kehitetty, että ne nyt tulisi tässä ihan just kohta ja se olisi jo ihan vanha juttu silloin tulevaisuudessa, ja sitten toisaalta esimerkiksi robotteja kehitetään nyt aktiivisesti. Boston Dynamicsilla on aika näppäriä robotteja, et ihmisen fyysisyys saavutetaan ihan tässä hetkenä minä hyvänsä... mä en usko, että nää nyt kovin äkkiä oikeasti ehkä sittenkään tapahtuu.

Mutta sen sijaan tapahtuu laajemmalla rintamalla oleva kehitys, että vähän arkisemmat ja tylsemät tekoälymenetelmät yleistyy monissa eri paikoissa. Ja sitten kun ollaan tilanteessa, missä ohjelmistot ja tekoäly on vähän niin kuin syönyt maailman, ja että älykkäitä komponentteja on joka puolella, niin siitä tuleekin sitten semmoinen aika iso muutosvoima. Ja ehkä sitten lopulta, vähän pidemmän ajan päästä tulee ne itsenäiset toimijat – robotit ja autot, jotka täysin sujuvasti pyörii ilman ihmisen apua. Mutta jotta se voi tapahtua, niin maailman pitää olla tekoälyvalmiimpi. Se ei välttämättä ihan nyt kävele se robotti vielä huoneeseen sisään ja ala esimerkiksi pitämään podcastia tässä sujuvasti. Vaan vaaditaan aika paljon taustatyötä ja laaja-alaista muutosta pohjaksi.

TR: No meidänhän on tavallaan turha ennustaa jotain ja et on nimenomaan tarkoitettu, että tätä kuunnellaan viidenkymmenen vuoden päästä, niin nythän on huono strategia oikeastaan lähtee mitään ennustamaan. Ehkä minäkin tyydyn olemaan aika varovainen ja lähetän vaan sellaisia terveisiä, että

ottakaa nyt te edes opiksi historiasta. Teillä on nyt 50 vuotta enemmän perspektiiviä siellä vuodessa 2071, että varmasti moni asia on mennyt pieleen, siihenkin mennessä on tullut monenlaisia ikäviäkin yllätyksiä matkalla, että kannustan heitäkin sitten vaan viisastumaan niistä virheistä, joita me tulemme seuraavan viidenkymmenen vuoden aikana tekemään. Ettei meidän virheet ja vastoinkäymiset menisi hukkaan. Eli olkaa hyvä. Me tulemme tekemään paljon virheitä ja mokia. Se on vaan teidän tulevaisuuden ihmisten etu. Oppikaa niistä.

AK: Kyllä, just se et kun mietitään, että mitä tekoäly tulee mullistamaan, niin sehän tulee mullistaa sitä mitä ihmiset haluaa sillä tehdä. Ja tietysti mä toivon, että sillä halutaan mieluummin vaikka ratkaista ilmastonmuutosta eikä voittaa sotia, mutta sen nyt sitten aika näyttää, että mikä se painopiste tulee olemaan.

TR: Toivotaan, että me osataan käyttää sitä fiksusti ja jos ei osatakaan, niin tulevaisuuden ihmiset voi sitten oppia näistä meidän huonoista päätöksistä.

AK: Toivotaan, että se rauhankone tehdään tässä vielä.

TR: Toivotaan, toivotaan näin. Ehkä näihin sanoihin, näihin tunnelmiin on hyvä päättää meidän turinatuokio. Todennäköisesti jatketaan sitten seuraavan viidenkymmenen vuoden aikana vielä muidenkin keskusteluiden yhteydessä. Toivottavasti päästään seuraamaan sitä, että miten se LUMI käynnistyy. Mutta siis lämmin kiitos Aleksii, että sain jutella sun kanssa.

AK: Kiitos paljon. Oli erittäin mukava istunto kyllä. Tämä podcastien tekeminenhän on aika hauskaa puuhaa, että ehkä jatkossakin voisın ruveta. Kiitoksia myös kuulijoille.

TR: Todellakin... kiitoksia kaikille kuulijoille ja mukavaa päivänjatkoa.

Outro:

Tämä podcast tallennetaan CSC:n 50-vuotisjuhlan kunniaksi aikakapseliin, joka avataan seuraavan kerran 100-vuotisjuhlan yhteydessä vuonna 2071.

Kiitos, kun kuuntelit For The Better Future -podcastia!