

Avoimesti tieteestä! - Tarmo Toikkanen: Tätä saa muutkin käyttää

Äänitteen kesto: 20 min

Litterointimerkinnät

Haastattelija: Henriikka Mustajoki
Vastaaja: Tarmo Toikkanen

sa- (sana) sana jää kesken
epävarmasti kuultu jakso puheessa tai epävarmasti tunnistettu puhuja

(-) sana, josta ei ole saatu selvää
(--) useampia sanoja, joista ei ole saatu selvää

, . ? : kieliopin mukainen välimerkki tai alle 10 sekunnin tauko puheessa

Intro: Avoimesti tieteestä. Keskustelua tieteen, tutkimuksen ja oppimisen avoimuudesta ja vastuullisuudesta.

Henriikka Mustajoki: Tervetuloa Avoimesti tieteestä-podcastiin. Minä olen Henriikka Mustajoki ja tänään seurani on Tarmo Toikkanen, jolla on vahva tausta tutkijana, joka tällä hetkellä toimii Open Knowledge Finlandin toiminnanjohtajana ja vastaa creative commons-lisenssin Suomen toiminnoista. Tervetuloa, Tarmo.

Tarmo Toikkanen: Kiitos, Henriikka.

Henriikka Mustajoki: Koska sun työ on tällä hetkellä liittyen vahvasti avoimiin lisensseihin, niin mikä on saanu sut vakuuttuneeksi, että avoimet lisenssit on meille, tiedeyhteisönä ja tutkimisyhteisönä, tärkeitä?

Tarmo Toikkanen: Mä muistan sillon 90-luvulla minä ja kaverit ja yhtiökumppanit oltiin vähän allergisia tämmöselle suljetun lähdekoodin ohjelmistolle niinku Windows ja muut, ja sieltä se harrastuksen kautta lipui työelämään oikeestaan 2001 joka on se sama vuosi jolloin ensimmäiset creative commons-lisenssit julkaistiin. Silloin aloitin työt Taideteollisessa korkeakoulussa, joka on nykyään osa Aalto-yliopistoa ja siellä (-) [0:01:11] tutkimusryhmässä mitä tehtiin, kaikki tutkimukset pyrittiin julkaisemaan mahdollisemmin avoimesti, kaikki prototyypit mitä tehtiin julkaistiin avoimella lähdekoodilla ja tehtiin mahdollisimman avoimesti kaikkea. Siitä on tullu vähän semmonen asennevamma. Oikeestaan kaikki mitä tässä ammatillisesti on elämän aikana tehnyt, niin kaiken on pyrkinyt järkevästi avoimesti lisensoimaan.

Henriikka Mustajoki: Sulla on siis pitkä historia lisenssien käyttäjänä. Jos joku ei oo koskaan kuullut tutkimuskoodien tai tuotosten tai muiden lisensoinnista, niin miksi

sitä kannattaisi sun mielestä harkita? Oikeestaan sen takia, että ne tekijänoikeudet tulee meille väkisin. Eli tämänkin podcastin tekijöinä minulla ja sinulla tulee tekijänoikeudet meidän omiin puheisiimme, koska ne ovat teoksia ja tälle tuottavalle organisaatiolle tulee tekijänoikeudet, yksinoikeudet tähän äänitallenteeseen. Nää tulee väkisin, ei ole mahdollisuutta sanoa, etten halua näitä. Eli kun tutkija julkaisee artikkelin tai tekee jonkun algoritmin tai koostaa jonkun tietoaaineiston, nää kaikki on automaattisesti hänelle yksinoikeuksin suojattuja. Ja ei voi sanoa ei, joten se toinen mahdollisuus on sanoa, että hei tätä saa muutkin jollain tavalla käyttää. Ja se tapa jolla se sanotaan on käyttöluva eli lisenssi. Nyt 20 vuoden ajan se lisenssi, jolla tämmösiä asioita kannattaa sanoa on creative commons. Se on se mitä maailmalla eniten käytetään, myös Suomessa. Silloin ku tekee jotain josta haluaa että muu voi sitä ikään kuin suhteellisen vapaasti hyödyntää, jatkojalostaa, niin silloin se kannattaa avoimella lisenssillä merkitä. Koska muuten lähtökohta, ihan sama oletko julkaissut sen netissä, kirjassa tai missä vaan, lähtökohta on se et yksinoikeudet on sinulla ja kukaan muu ei niihin koske muuta kuin ehkä yksityishenkilönä.

Henriikka Mustajoki: Miks just Creative Commons?

Tarmo Toikkanen: Siinä on tavallaan tää brändiasiakkin, että käytetään nimiä ja lisenassejä jotka on tuttuja. Jos nyt kattoo esimerkiks Suomessa miten avointa dataa ja tietoa on julkaistu, siellä alkuvuosina oli kaikennäköistä. Jokaisella oli oma lisenssi. Helsingillä oli oma lisenssi. Jos aattelee, että Helsingistä julkaistaan vaikka jotain tietoaaineistoja, tilastoja ja sitten jos ne kiinnostaa ulkomailla, niin ihanko oikeesti ne lukee jotain suomen kielellä ja ruotsin kielellä julkaistu Helsinki open licensea tai Helsingin avointa lisenssiä ja ymmärtää mitä se tarkoittaa ja luottaa, et se on jotain järkevää? Eiks siin oo pieni kynnyks lähteä, että hetkinen, voiko tähän nyt oikeesti luottaa. Toinen on ihan tämmönen tekninen yhteensopivuusasia eli se mitä etenkin tietoaaineistolle tehdään, on se että niitä jalostetaan ja yhdistetään. Otetaan eri kaupunkien väestödataa ja yhdistetään siitä jotain uutta, saadaan uutta ymmärrystä näkyviin.

Eli se on se yks syy miksi on 20 vuoden aikana ajauduttu siihen, että on todettu että okei, kannattaa olla ikään kuin siinä samassa isossa meressä polskimassa, eikä että jokaisella on oma pieni lätäkkö missä sitä julkaistaan. Näitten lisenssien taustalla on siis usalainen non-profit organisaatio, lakialan professori Lawrence Lessig oli sitä perustamassa. Se syntyhistoria, jos aattelee vuosituhannen vaihdetta niin siinä tietotekniikka alkoi hiljaksiin yleistä ja syntyi tää uus termi remix-kulttuuri, eli ikään kuin se et digitaalista dataa, aineistoa tai mediaa ja uudelleenyhdistellään siitä uutta. Se nähtiin vuosituhannen vaihteessa, et se on yleistymässä, se on kasvamassa ja tavallaan tää kun bitit ei enää maksa mitään, että niitten säilöminen ja siirtäminen on aina vaan halvempaa. Se tekee kopioimisen helpoksi. Ja sit nähtiin, et sil on myös kulttuurista arvoa, uutta luovaa arvoa, sillä että remixausta voidaan tehdä.

Ajatus oli, että Creative Commons olis tällöinen väliaikainen laastari tähän rikkinäiseen tekijänoikeuslakiin. Jossain vaiheessa tekijänoikeuslait korjaantuu ja remixaus muuttuu mahdolliseksi ja sitten näitä lisenssejä ei enää tarvittaisi, se oli alkuperäinen tavoite. Mutta nyt 20 vuotta myöhemmin voidaan todeta, että eihän näin oo missään nimessä käymässä. Tekijänoikeuslakien kehitystä kontrolloivat isot mediayhtiöt. Tavallaan se yhteiskunnan vapaan keskustelun ja uuden luomisen oikeudet on jäänyt vähän jalkoihin. Kyl se nyt näyttää, et Creative Commons on aika pitkään meidän kanssa. Ja se on ikään kuin tässä 20 vuodessa muuttunu, se ei oo enää vaan tällöinen käytännön standardi, että yhä useampi sitä käyttää, vaan se on ihan virallinen standardi.

Se on EU:ssa virallinen standardi, se on Suomessa virallinen standardi. Eli kun Suomessa julkishallinto julkaisee tietoa, niin ne pitäis julkaista Creative Commons-lisenssillä. Eli sillä tavalla tämä kehityskulku ikään kuin on tässä nyt 15-16 vuodessa käyty lävitse ja nyt ollaan siinä tilanteessa, että meillä on tällainen avoimen tiedon valtameri, joka määrittyy sillä et siihen kuuluva tieto on Creative Commons-lisensseillä merkattu avoimeksi. Ja sitä sitten saa peuhata ja uudelleen käyttää. On se käyttö sit tutkimusta, journalismia, taidetta tai mitä se sitten onkaan.

Henriikka Mustajoki: Miten sä näät et Creative Commons-lisenssi muuttaa sitä miten tutkimusyhteisö tavallaan omalla tavallaan tuottaa uudenlaista remix-kulttuuria? Se tuntuu kauheen luontevalta osalta tutkimustyötä.

Tarmo Toikkanen: Kyllä, kyllä. Joku minua viisaampi on sanonut, että standing on the shoulders of giants. Sitähän se tutkimus on, että aiemmin tehdyn tutkimuksen päälle rakennetaan uutta. Tavallaanhan tämä tekijänoikeus ei onneksi suoja tietoa, eli tieto on vapaata. Vaikka kirjalla on tekijänoikeudet, niin on sallittua lukea se kirja, oppia siitä ja hyödyntää sitä tietoa tai samalla tavalla lukea vaikka tieteellinen artikkeli ja ymmärtää että mitä siinä sanottiin ja voi sen päälle rakentaa uutta. Tähän tieteen perinteeseen, siihen tää ei sinäänsä vaikuta. Koko ajan rakennetaan sen jo julkaistun tiedon päälle. Täs on kaks juttuu.

Toinen on tavallaan tää tieteellisten artikkelien julkaisukulttuuri ja sen digitalisoituminen ja toinen on sitten taas tietoa ja niiden jakaminen, joka on ihan uus juttu. Tieteelliset journalit tavallaan perustellusti aikoinaan, kun ainoa tapa julkaista tietoa tai levittää tietoa oli se, että niitä printattiin kirjasiksi ja jaeltiin sitten ympäri maailmaa yliopistojen kirjastoihin, niin siinä on tiettyjä kuluja. Ja jollain tavalla ne pitää kattaa. Eli se on kaupallista toimintaa hyvin pitkälti. Ja sit täs on ikään kuin tän vuosituhatvuotisen aikana yhä noussut kritiikki tälle, että hetkinen, nää isot journalit itse asiassa tekee tosi hyvää tulosta, yliopistokirjastot maksaa valtavia summia, siis Suomessakin oli laskettu, Open Access Finlandissa me jossain vaiheessa tehtiin tietopyyntö ja selvitettiin, että kuinka monta miljoonaa suomalaiset yliopistot maksavat vuosittain käytännössä amerikkalaisille ja englantilaisille journali-yhtiöille siitä, että saadaan tiedettä katsella, tieteellisiä

tuloksia. Mutta kun ne journalit on nykyään digitaalista pdf:ää, niin siinä on julkaisu, painokulut on hävinny.

Se ainoa iso työ on oikeestaan se vertaisarviointi, mikä tietysti on iso työ, joka vaatii kovia asiantuntijoita, mut se tehdään vapaaehtoistyönä. Monet tutkijat on näissä rooleissa, eli se tehdään talkootöinä ja journalit sitten vaan kerää rahat ja ikään kuin jakelee pdf:iä. Tää on tavallaan tää kritiikki, ja siitä on ikään kuin nousut tää Open Access-liike eli OA. Ajatus siitä, et hei miksi tiede tehdään kenties julkisin varoin tai yliopiston varoin ja sitten se vertaisarvioidaan kollegoiden vapaaehtoistalkootöillä, niin miksi se sitten laitetaan maksumuurin taakse, jonkun kaupallisen firman arkistoon? Eiks se vois olla avoimesti saatavilla?

Open Access on koko ajan lisääntymässä ja se mitä on tutkittu, mikä on tavallaan nyt kiinnostavaa avoimen lisenssin ja avoimen remix-kulttuurin kannalta, on se että mitä vähemmän on maksumuureja, niin sitä enemmän sitä siteerataan ja se on tietenkin se peli mitä akatemiassa pelataan, että paljonko tulee siteerauksia, eli se että sun artikkeli on avoimesti saatavilla mahdollistaa esimerkiksi sen et tutkijat kehittyvissä maissa joilla ei välttämättä ole tilauksia näihin kalliisiin journaleihin, pääsee tähän sun tieteeseen käsiksi ja pystyy sen päälle rakentamaan uutta ja jotenkin aattelis et se ois ikään kuin tutkijankin etu, että hänen tuloksiaan mahdollisimman paljon, mahdollisimman nopeasti voidaan hyödyntää. Eli tää oli se Open Access-puoli. Sinäänsä Open Access ei nyt välttämättä edellytä et siel on joku Creative Commons-lisenssi.

Periaattees et ne artikkelit on jossain verkkosivulla julkaistuna riittää siihen, että siihen tietoon päästään käsiksi. Mutta tietenkin se, että jos ja kun siellä on avoin lisenssi mahdollistaa esimerkiksi sen että niitä artikkeleita ihan laillisestikin voit ikään kuin tutkijana tai esimerkiksi kirjaston henkilökuntana, tai miten voitte ikään kuin koostaa kokoelmia näistä artikkeleista, eikä vaan että ne on linkkiloja, vaan et ihan oikeesti ne artikkelit voi olla jossain yliopiston omassa varannossa tai voit jakaa kollegoille pdf:iä näistä artikkeleista kun niis on avoin lisenssi. Silloin tää jakaminen on mahdollista. Sillä tavalla sujuvoittaa tiedon leviämistä. Ja sitten se kokonaan puoli on nää tietoaineistot, eli se data, tietokannat, aineistot mitä monessa tutkimuksessa on siellä taustalla, joissain tietysti se voi olla hyvin tämmöstä haastatteludataa, joissain taas numerista dataa.

Tavallaan se liittyy taas tieteen integriteettiin, siis siihen että se on nähty tässä, että tiedettä välillä tehdään aika kevyillä eettisillä perusteilla eli siis roskaa on aika paljon. Hiljaksiin aletaan nähdä, että tutkija tai yliopisto ikään kuin julkaisee tieteellisen artikkelin, ni se ei ole riittävä tae siitä et se olisi kunnan tiedettä, eikä ne vertaisarvioijatkaan ku ei heilläkään oo perinteisesti pääsyä siihen lähdeaineistoon, he vaan lukee sen artikkelin ja kattoo et näyttääks se että se on kuvattu oikein ja siinä ei oo mitään puutteita. Se on tavallaan tää seuraava askel, että hei, mitäs jos edellytettäisiin ja esimerkiksi EU:n komissio edellyttää, että heidän rahoittamassaan tutkimuksessa tietoaineistot pitää julkaista. Eli jos sä saat

EU:n rahaa, niin ei riitä, että sä julkaiset open accessillä sen tutkimustuloksissa, sun pitää myös julkaista se tietoaaineisto ja avoimella lisenssillä.

Eli tää tietoaaineistojen lisensointi, tää on tämmönen eurooppalainen juttu, koska siis muualla maailmassahan tietokannat eivät ole suojattu. Kun sulla on tietokanta tai tietoaaineisto, niin jotta joku muu voisi esimerkiksi kopioida sen joku muu tutkija ja vaikka analysoida ja julkaista siitä jotain jalostettua versioita tai hyödyntää tietoaaineistoa omassa tutkimuksessaan. Ja siihen tarvitaan taas lupa, eli lisenssi eli Creative Commons-lisenssi. Mut siinä tavoitteena on ikään kuin se tutkimuksen laadun parantaminen. Ikään kuin että muut voi itsenäisesti verifioida, jos tää on sun data ja tässä sun artikkelissa on kuvattu tää käsittely, jos mä ajan sen saman jutun ni saanko mä samat numerot. Liian paljon on esimerkkejä siitä, että kun se data on myöhemmin saatu käsiin ja ajettuna todettu, että jaa, no eihän tästä tullut näitä numeroita mitä sinä sanoit, et mistäs tässä on kyse.

Henriikka Mustajoki: Samaan aikaan kun moni tunnistaa, et näil on keskeinen laadun tae, epäilemättä säkin oot kuullu paljon erilaisia huolia siitä, et jos minä, erityisesti nyt aineistoon liittyen, laitan nämä avoimella lisenssillä, niin miten käy minun etuoikeuteni ja mahdollisuuteni käyttää sitä aineistoa? Mitä jos joku väärinkäyttää sitä aineistoa, niin mikä on minun vastuu? Miten sä aattelet näistä huolista ja niiden suhteesta tähän lisenssiin?

Tarmo Toikkanen: Tavallaan täs on se toimintakulttuuri, kun se muuttuu ni se aina vähän huolettaa ja nää on ihan ymmärrettäviä. Mut nää aika hyvin pystyy taklaamaan. Tietysti tämä etuoikeus esimerkiks johonkin tutkijan dataan, jos se julkaistaan avoimel lisenssillä, niin sihän se ei enää ole sinun siipesi suojassa, vaan kuka tahansa voi sitä hyödyntää. Eli tavallaan se ainoa etu mikä sinulla on, on se että sä oot sen datan jo kerran möyhentänyt lävitse. Tää on oikeestaan näkökulma, ei pelkästään tieteessä vaan missä tahansa toiminnassa, että missä tilanteessa kannattaa lisensoida avoimesti? Oikeestaan se tiivistyy siihen, et jos on aikeita myydä rahaa vastaan sitä dataa, on se sitten valokuvia, videota, elokuva, tietokonepeli tai tietoaaineisto, jos haluaa myydä sitä rahaa vastaan, niin silloin ei kannata laittaa avointa lisenssiä, koska se tekee rahan ansaitsemisen kovin vaikeaksi.

Mutta jos se liiketoiminta-arvo, tai mistä saat palkkasi, on jokin sen datan ympärillä oleva lisäarvopalvelu, esimerkiks opettajalla voisi ajatella että se oppimateriaali ei ehkä ole se tärkein anti minkä hän tuottaa, vaan se opetus ja asiantuntemus ja ohjaaminen jonka hän siihen tuo. Vois ajatella, että hän ehkä voisi itse tekemäänsä oppimateriaalia julkaista avoimesti ilman että se haittaa hänen omaa tulovirtaansa. Tai ehkä se tutkijallakin lopulta on se vuosien, vuosikymmenien aikana kertyvä asiantuntemusosaaminen, ymmärrys siitä alasta, joka on ikään kuin se myyntivaltti jolla hän saa työtä toisesta yliopistosta tai mikä se sitten onkaan tai pystyy hakemaan rahoitusta tutkimusryhmälleen, eikä niinkään se että hän on torilla kaupittelemassa et täs ois vitosella tutkimus, se ei

oo se liiketoimintamalli. Se on se yks ajatus, sen terävöittäminen itselle, et mikä on se mitä minä ikään kuin myyn.

Onko se se data tai se joku artikkeliteksti vai onko sen ympärillä oleva oma osaaminen, tutkimusryhmän kehittyvä, kasvava osaaminen, mikä on ikään kuin se jolla on arvoa yhteiskunnalle, josta joku ulkoinen rahoittaja on valmis maksamaan? Tosiaan, semmonen data joka ei ole ikään kuin rahan arvoista, niin se on tavallaan alikäytöllä, jos se on sinun omassa (arkisto-) [0:14:59], eikä kukaan pääse siihen käsiksi. Sil on mahdollisuus tuottaa, parantaa, koko ihmiskunnan tieteen tasoa, jos se on avoimesti saatavilla. Väärinkäyttö on toinen.

Eli se on juuri tämä et jos minulla on tietoaaineisto josta saa jotain johtopäätöksiä, mitä jos joku ottaa siitä kopion ja kun kerran siinä on avoin lisenssi, niin hänellä on sitten oikeus ikään kuin muuttaa sieltä vaikka numeroita ja julkaista tämmönen muutettu tietoaaineisto, josta kun ajaa analyysit, niin saa vastakkaisen johtopäätöksen. Ensinnäkin, todetaan nyt se että nää CC-lisenssit edellyttää, että lähde mainitaan ja kun muunnellaan on pakko mainita mikä se alkuperäinen on. Eli jos nyt minä menen ottamaan väyläviraston merikartat ja teen oman version siitä, josta puuttuu pari saarta Helsingin edustalta ja julkaisen sen, niin ensinnäkin mun pitää sanoa että minä sen tein ja että se perustuu tuohon väyläviraston aineistoon, joka on osoitteessa x.

Jos mä jätän tän tekemättä, niin siinä vaiheessa mä rikoin sitä lisenssiä, jolloin siinä vaiheessa mä oon automaattisesti rikkonu tekijänoikeuksia, puhumattakaan kaikista muista mitä siinä on, petosta ja liikenteen vaarantamista. Eli tavallaan hyvä ymmärtää, että tekijänoikeus ei oo ainoa laki mikä meillä on, meillä on kaikki muutkin lait voimassa, et kaikkien lakien mukaan pitäis toimia. Mut lopultahan se vastuu on tietysti myös sillä käyttäjällä. Jos nyt sitten Silja Line lataa Tarmon tekemän väyläkartan kun se on uudempi ja se on varmaan tosi hyvä, niin kyllä se vastuu on ikään kuin katsojalla tai käyttäjällä [naurahtaa], et ehkä niitten kannattaa ladata se virallinen kartta, eikä jonkun huithapelin tekemää, jotain mistä ei oo mitään takeita et mitä se pitää sisällään.

Täs on taustalla se, että kun tekijänoikeus on vähän ollu semmonen yleispätevä, semmonen kielto mekanismi, nyt avoimet lisenssit poistaa sen. Sit ihmiset on vähän silleen hädissään, et hetkinen, että miten mä voin estää että joku väärinkäyttää mun aineiston. Muitakin lakeja, et jos joku vaarantaa liikennettä tai vaarantaa kansanterveyttä julkaisemalla jotain huuhaa-tutkimuksia, vetoaa johonkin itsetekemäänsä väärennettyyn data-aineistoon, joka perustuu sun aineistoon, niin siihen voi puuttuu muidenkin lakien avulla. Plus että siinä data-aineistossa lukee, että sen on tehnyt huithapeli x ja se perustuu siihen oikean tutkijan tekemään aineistoon. Kuka tahansa voi sitten kattoa, että hetkinen, miten nää kaks tietoaaineistoa eroaa toisistaan.

Totean vielä sen, että Creative Commonsissa, siellä on varauduttu siihen, että voi olla sellaisia käyttötarkoituksia jotka on ikään kuin tekijälle jotenkin moraalisesti hyväksymättömiä tai sillä tavalla, että ne ei oo lainvastaisia, mut ikään kuin se tekijä ei halua olla assosioituna tähän käyttötapaan. Tässä Creative Commons-lisenssissä on itse asiassa semmonen pykälä, et jos tulee tällöinen tilanne, että sinun aineistoasi käytetään jossain asiayhteydessä missä et halua tulla mainituksi, niin sulla on oikeus vaatia, että sun nimi poistetaan näkyvistä.

Henriikka Mustajoki: Nyt jos joku ois kovin kiinnostunut oppimaan lisää Creative Commons-lisensseistä ja aloittamaan niiden käytön, niin mistä lähteä liikkeelle?

Tarmo Toikkanen: Helpointa varmaan on creativecommons.fi-verkkosivusto, et sieltä löytyy lähtöohjeita. Mut sit myöskin täällä avoimen tieteen puolella, etenkin tutkijoille avointiede.fi on hyvä lähtökohta. Sieltä löytyy erillisiä lyhyitä, 10-20 sivun oppaita tieteellisille kustantajille, tutkijoille, miten tähän voi suhtautua, missä tilanteessa voi lisensoida, miten se lisensointi käytännössä tehdään. Tietysti jos englannin kieli taipuu, niin creativecommons.org on kansainvälisen järjestön sivusto mistä löytyy hyvinkin paljon ohjeita, tietoa ja hyviä käytäntöjä siitä miten tää homma toimii.

Henriikka Mustajoki: Mulle jäi tästä monta asiaa ja jotenkin se ajatus siitä, että jos me emme lisensoi, niin silloin me ollaan semmosessa oletustilassa, jolloin muu ei pysty hyötymään mistään ja sen lisensoinnin oleellinen rooli siinä, että jos meillä on halua jakaa, avata ja luoda sitä tiedon merta, niinku sä kuvasit, niin lisensointi on siihen aivan oleellinen askel. Nyt kiitän, Tarmo, kovasti ajastasi ja ajatuksistasi tämän podcastin yhteydessä ja toivotan hyvää talven jatkoa.

Tarmo Toikkanen: Hyvä. Kiitos, Henriikka, tää oli oikein hauskaa.

[Lopputunnusmusiikki.]