

Meri Kytö

***Tiedekustantamisen tulevaisuus avoimen
julkaisemisen ja tekoälyn pinteessä***

FT, dos. Meri Kytö (meri.kyto@utu.fi) toimii musiikkitieteen yliopistonlehtorina Turun yliopistossa, Suomen musiikkitieteellisen seuran puheenjohtajana, ja on Musiikin päätoimittaja vielä tässä numerossa.

ORCID: <https://orcid.org/0000-0002-7877-5522>

DOI: [10.51816/musiikki.144765](https://doi.org/10.51816/musiikki.144765)

Tiedekustantamisen tulevaisuus avoimen julkaisemisen ja tekoälyn pinteessä

Meri Kytö

.....

Osallistuin Tiedekustantajien opintomatkalta London Book Fair -tapahtumaan maaliskuussa 2024. Nämä vuosittain järjestettävät messut eroavat kovasti esimerkiksi kotimaisista kirjamessuista siinä, että kohdeyleisönä ovat kustannusalan ammattilaiset. Näillä päivillä ei niinkään törmätä kirjoihin ja kirjailijoihin, mutta kustantajiin, kääntäjiin ja konsultteihin kyläkin. Paikalla oli kustantajatahojen edustajia paljon, ympäri maailmaa. Näytteilleasettajia oli yli tuhat.

Vaikka paikalla oli fiktion ja lastenkirjallisuuteenkin painottuvia kustantajia, oli omalle vierailulleni toki keskeisempänä kiinnostuksen kohteena tiedekustantajat. Läsnä olivat kaikki suuret kansainväliset toimijat Elsevieristä MIT:hin, Springer Naturesta eri yliopistopainoihin. Mukana oli myös paljon musiikintutkimusta julkaisevia kustantamoita, kuten Bloomsbury, Cambridge University Press ja jättimäinen Taylor & Francis. Ohjelmaksi oli laadittu useita pienempiä paneeleja ja keskusteluja sekä näiden lisäksi torstaille koko päivän tiedekustannusseminaari.

Tässä seminaarissa käsiteltiin kustannuskäytäntöjen muutoksia, tekoälyn vaikutuksia kentälle, avoimen julkaisemisen kansainvälisiä suuntaviivoja, tekijänoikeutta sekä kestäväää kehitystä. Oli hyvin kiinnostavaa päästä kuulemaan, miten suuret kustantajat puhuvat toisilleen, millainen sanasto heille on yhteistä ja mitä he havainnoivat tulevaisuuden kustannustyön muodoiksi. Taloudellisen kannattavuuden sanana käytettiin *sustainability* -sanaa, viitaten siis lähinnä vain taloudelliseen kestävyYTEEN. Tämä oli yksi toistuvimpia keskustelun teemoja, jota pyöriteltiin myös tiede-etiikan kehityksessä. "We don't exist for profit, but without profit we don't exist", kuten Nicola Ramsey Edinburgh University Pressistä muotoili tiedekustannusalan taloudelliseksi periaatteeksi. Tästä varmaankin olisi ollut myös poikkeavia näkemyksiä läsnä olevilla tahoilla, joille akateeminen kustantaminen on suuren suuri bisnes, mutta tämä oli keskustelun-

aihe, joka herätti huomiota lähinnä poissaolollaan ja *sustainability* -sanan toistelulla.

Monessa puheenvuorossa toistui toteamus siitä, että kustannusala elää jännittävää aikaa. Konsulttikielelle tuttu nopean reagoinnin ja koikeilun toimintatapa *failing fast, learning fast* tuntui olevan se keino, jolla alan muutokseen kannattaisi monen mielestä suhtautua. Jatkuvan epävarmuuden tuottaman epämiellyttävyyden kanssa toimiminen nivottiin myös osaksi kustantajan nykyistä työnkuvaa.

Tekoälyn teema kattoi sateenvarjon tavoin useamman paneelin spesifimmät keskustelunaiheet. Tämän vuoden tapahtumassa teemaa ruoti jopa noin 30 eri paneelia. Oppivien ohjelmistojen, tekstiä generoivien kielimallien ja erilaisten laadunvarmennustyökalujen kanssa on yhden panelistin mukaan ”mahdollisuus tehdä todella cooleja juttuja”. Tekoälyn kanssa painiminen (hyvässä, pahassa ja epäselvässä) on arkipäivää jo suomalaisissakin tiedelehdissä. Mainittakoon, että *Musiikkiinkin* on jo lähetetty käsikirjoituksia, joissa kerrotaan käytetyn luonnosteluapuna suurten kielimallien (LLM) generatiivista tekstintuottamista.

EU:n tulevat säännösesitykset aiheesta kiinnostivat monia kustantajia myös Euroopan ulkopuolella, koska ne tulevat ensimmäisinä määräämään sen rakenteen ja fokuksen, jolla keskustelu tekoälystä, kielimalleista, generatiivisuudesta, plagioinnista ja näiden kehittämisen ja käytön reguloinnista jatkuu. Mainittakoon tässä yhteydessä, että esimerkiksi *Musiikin* sisältöjä ei saa käyttää kielimallien opettamiseen kaupallisia tuotteita tehtäessä. *Musiikissa* kuten monessa muussakin kotimaisessa avoimen saatavuuden tiedelehdessä on käytössä CC BY-NC-ND -lisenssi, joka sen ainakin toistaiseksi kieltää. Tarkeimmin ei kuitenkaan ole tiedossa, onko niitä kuitenkin koneopettamiseen käytetty, kehittäjiltä kun ei ainakaan vielä vaadita listausta käytetyistä aineistoista. Isossa-Britanniassa tätä ilmoitusvelvollisuutta pyritään vielä tämän vuoden kuluessa säätämään, kertoi Richard Mollet konsulttiyhtiö RELX:istä. Oikeudenkäyntiprosesseja tekoälyn tiimoilta seurataan myös silmä tarkkana. Esimerkkinä mainittiin Nashville vs. Anthropic -lyriikkaoikeudenkäynti Yhdysvalloissa. Tästä tuli juuri viikko messujen jälkeen uutisointi *The Guardian* -lehdessä (22.3.2024). Päätös näyttää tuovan musiikintekijöille ja sanoittajille pienen hetken hengähtää ja turvaa suojella tekijänoikeuksiaan lauluääntä, musiikkia ja lyriikkaa generoivilta tekoälysovelluksilta.

Paneelissa ”Strategies and trends in academic publishing” keskusteltiin myös saavutettavuuden lisäämisen tärkeydestä. Monissa maissa on julkaistu tai kehitteillä avoimeen julkaisemiseen vaikuttavia linjavetoja. Japanilaista avoimeen julkaisemiseen kehottavaa mallia kutsuttiin ag-

gressiiviseksi, kiinalaista ”kiinnostavaksi”. Intiassa on suunnitelmia avata kaikki tiedejulkaiseminen avoimeksi kaikelle kansalle ilman embargoa. Euroopan komission ja Suomen Akatemiankin tukemaa Plan S:ää ei mainittu keskusteluissa, mikä kertoo ehkä siitä, että tällä hetkellä kustantajia kiinnostavin muutos on tapahtumassa edellä mainituissa maissa. Tämä on ymmärrettävää katsottaessa julkaisemisen volyymia: Kiinassa julkaistaan vuosittain yli puoli miljoonaa tiedeartikkelia, joista noin 30% on avoimia ja tästä artikkelin prosessointimaksun (APC) maksavia (Wang & Campbell 2023).

Tekoälystä toivotaan monella taholla apua itse toimitusprosesseissa. Springer Naturen tekoälyosaston johtaja Thomas Sütterlin huomautti, että kaikki käyttämämme tekoälysovellukset ja -työkalut ovat nyt huonoimpia mitä tulemme koskaan käyttämään, ja hän muistutti (ja toivoi) niiden paranevan ja monipuolistuvan nopeaakin vauhtia. Vertaisarvio on kustantajalle se eniten aikaa vievä osuus julkaisemisessa, johon toivottiin jouhevoittamista. Raikasta oli kuulla yhden panelistin kyseenalaistavan jouhevoittamisen päämäärän kiirehtimisenä, ja kysyvän josko loputon tehostus ja kiirehtiminen vaarantaa tiedekustantamisen tärkeimmäksi mainitun arvon eli luotettavuuden. Yksi monisäikeinen tekoälyn tuottama ongelma kustantajille ja toimittajille on misinformaation ja disinformaation erottelu: kuka kirjoittajista tietää ja tiedostaa miten käyttää kielimalleja ”oikein” ja milloin niiden käyttö on puolestaan ”väärin”, yhteisiä sääntöjä tähän kuin on vaikea pitää ajan tasalla kiivastahtisessa muutoksessa. Ajankohtaiseksi huoleksi julkilausuttiin myös tieteellisen tiedon väärinkäyttö ja manipulaatio poliittisiin tarkoituksiin, ja kuinka tärkeä rooli tiedekustantajilla tässä on. Tässä oli aistittavissa hieman sitäkin ajatusta, että tiedekustantajille on nyt siirtynyt sitä roolia minkä journalismi aiemmin tiedonjakamisella kantoi. Huolestuttavaksi kasvavaksi trendiksi nimettiin myös julkaisutehtailu (*paper mills*). Hyviäkin tekoälyn käyttömahdollisuuksia hahmoteltiin: heikommin englannin kieltä taitavat tutkijat voisivat keskittyä tutkimuksen sisältöön kieliopin ja muodon fiksaamisen sijaan. Käännösohjelma-avusteisesti on myös helpompi kommunikoida englanniksi toimittajan kanssa. Aikaa ei enää mene yksittäisten sanojen etsimiseen sanakirjoista, vaan toimittajan kanssa viestintään menee tutkijalta vähemmän aikaa.

Musiikki-lehden ja sen kustantajan eli Suomen musiikkitieteellisen seuran näkökulmasta keskustelu seminaareissa oli pääosin kiinnostavaa, mutta toimijoiden suuruusluokasta johtuen hieman kaukaista. Tuntui siltä, että joissain asioissa on etu olla kaltaisemme pieni ja nopeasti tarpeellisia muutoksia tekemään kykenevä toimija tiedekustannusalalla.

Tärkeimpänä asiana laatutyössä kaikki tiedekustantajat pitävät vertaisarviointiprosessia. Plagioinnintunnistustyökalut ja tästä edistyneemmät ”integrity checking” -työkalut (joiden automatisoidut toiminnot jäivät minulle vielä hieman epäselviksi) tuntuvat juoksevan generoivien työkalujen jäljessä. Näyttää siltä, että vaikka suuret kustantajat pystyisivätkin tuottamaan erilaisin oppivin ohjelmistosovelluksin julkaisutietokannoistaan monenlaista reaaliaikaista visualisointia, taulukkoa ja oppimateriaalia, haku- ja generointimahdollisuuksia ja tiivistelmien automatisoitua tekemistä, jää toiminnan ytimeen kuitenkin tieteellisen laadun vaaliminen vertaisarviointikäytännöin. On kuitenkin kiinnostavaa seurata, josko suurista kustantajista muotoutuu tulevaisuudessa enemmänkin alustatalouden mukaisia toimijoita, jotka tarjoavat maksua vastaan verkkojulkaisujen alustan ja markkinoinnin tiedeyhteisöille, jotka tuottavat sille sisällön.

Lähteet

The Guardian 22.3.2024. ”Tennessee becomes first US state to protect musicians from threat of AI”. Luettu 4.4.2024.

<https://www.theguardian.com/music/2024/mar/22/tennessee-bill-protect-musicians-ai>

Wang, Chongfang & Nick Campbell 2023. ”Why China is critical to the growth of open access”. Luettu 4.4.2024.

<https://www.springernature.com/gp/advancing-discovery/springboard/blog/blogposts-open-research/why-china-is-critical-to-the-growth-of-open-access/25293782>