

Toimittajan palsta

Ajatuksia konferenssiartikkeleista

Tietojenkäsittelytieteessä konferenssiartikkeleita arvostetaan julkaisuina toisin kuin joillakin muilla tieteenaloilla. Tästä seuraa joskus ongelmia sellaisissa arvioinneissa, joita on suorittamassa monien tieteenalojen edustajia.

Alallamme on yleisesti hyväksytty tapaa julkaista tutkimus tavallaan kahteen kertaan ensin kokousartikkelina ja sen jälkeen lehtiartikkelina. Tyypillisesti lehtiartikkeli on laajempi ja perusteellisempi kuin alkuperäinen kokousartikkeli, mutta joskus on samasta kokousartikkelista kirjoitettu useitakin lehtiartikkeleita. Toisinaan lehtiartikkelissa viitataan alkuperäiseen kokousartikkeliin, minkä käytännön toivoisin vahvistuvan, jotta tutkimustulosten historian selvittäminen olisi jälkikäteen helpompaa.

Vuosikymmeniä sitten tutkija saattoi julkaista monen vuoden työnsä laajana monografiana, mikä ei ole enää muodissa nykyisenä tulosvastuun aikakautena. Nopea julkaiseminen on toki tulosten hyödyntäjien kannalta hyvä asia, ja rahoittajat tuntuvat arvostavan enemmän määrää kuin laatua. Taitava julkaisija pilkkookin työnsä pienimmiksi julkaistavissa oleviksi yksiköiksi (engl. LPU = least publishable unit). Jos tutkimus on monitieteellinen tai liittyy useaan tietojenkäsittelytieteen haaraan, voidaan tulokset julkaista useamman alueen julkaisukanavia käyttäen. Tällöin kuitenkin liikutaan tulosten alkuperäisyyden suhteen harmaalla alueella. Myöhemmissä julkaisuissa on tärkeää viitata ensimmäiseen julkaisuun, jottei lukijalle jää asiasta väärää käsitystä.

Vielä nykyisenä Internetin aikana tulee vastaan tapauksia, joissa yritetään tar-

jota jo julkaistua artikkelia toiseen kokoukseen. Joskus tämä voi johtua tietämättömyydestä. Referoijan tähtihetkiä on bongata verkosta lähes identtinen aikaisempi julkaisu, johon ei viitata. Kerran ollessani ohjelmatoimikunnassa saimme hyväksymisviestiin vastauksen: ”Artikkelimme on hyväksytty kahteen kokoukseen, emme ole vielä päättäneet kumpaan matkustaisimme.” Kirjoittajat eivät onneksi tavoitelleet tässä tapauksessa kaksoisjulkaisua mutta rikkoivat alkuperäisyyden vaatimusta vastaan. Ohjelmatoimikunnan kannalta mahdolliset poisjäännit olisivat kiusallisia.

Uusia tieteellisiä kokouksia syntyy tiuhaan tahtiin. Miten sitten arvioida kokousten tasoa? Yleisenä perusteena on se, että kaikki tarjotut artikkelit referoidaan. Tällöin olisi toivottavaa, ettei kaikkia lähetettyjä artikkeleita hyväksytä. Vaikka kaikki artikkelit hyväksyttäisiinkin, voi kokous silti olla hyvätasoinen, mutta laajaa arvostusta sellainen kokous tuskin saa. Hyväksymisprosenttikan ei aina kuvaa kokouksen tasoa, sillä parhaimpiin kokouksiin saatetaan tarjota itsesensuurin vuoksi vähemmän artikkeleja. Keskimääräinen impaktiluku on hyvä mitta, mutta toisaalta viittauskäytännöt eivät ole samanlaisia tietojenkäsittelytieteenkään eri osa-alueilla. Uuden kokoussarjan ensimmäisen kokouksen tasoa on usein erityisen hankala arvioida.

Kokousten tieteellisyyteen kuuluu myös se, että artikkeleiden hyväksymisprosentti julkaistaan. En voi ymmärtää joidenkin ohjelmatoimikuntien politiikkaa pitää hyväksymisprosentti salaisena. Kokousjulkaisussa julkaistujen posterien hyväksymispolitiikan tulisi myös olla julkinen.

On tiedostettava, että kokouksissa on hyväksymisrajan läheisyydessä harmaa alue, jossa artikkelin hyväksyminen tai hylkääminen riippuu siitä, keille referoijille artikkeli on sattunut joutumaan. Sama pätee tietysti muunkinlaiseen hajautettuun arviointiin. Eräässä tutkimuksessa ilmiötä testattiin lähettämällä joukko artikkeleita kymmenille referoijille (H. Walker et al.: Variability of referees' ratings of conference papers. Proc. ITiCSE '02, ACM, 2002). Tästä syystä valveutunut ohjelματοimikunnan puheenjohtaja uhraa eniten aikaa hyväksymisrajan läheisyydessä oleviin artikkeleihin, kun taas epäpätevä kollega saattaa käyttää kylmästi keskiarvoa artikkelin saamista arvosanoista, jolloin yksikin arvioija saattaa pudottaa artikkelin rajan alapuolelle. Usein osa arvosoista voi olla vähemmän tieteellistä sisältöä mittaavia, kuten arvio englannin kielen hallinnasta.

Koska periaatteessa kuka tahansa voi perustaa uuden kokoussarjan, emme voi välttyä ikäviltä lieveilmiöiltä. On perustettu kokouksia, joiden päätarkoituksena on kerätä rahaa tai toimia jonkin kansallisuuden kokoontumisen kulissina tai joissa referointia ei ole asianmukaisesti järjestetty. Viimeksi mainittu ongelma on

tieteellisyyden kannalta hankala. Jotkut tutkijat ovatkin julkaisseet verkossa mustia listoja tällaisista bogus-kokouksista, mutta näitä listoja on jouduttu sulkemaan suoranaisten uhkausten takia. Kopioita joistakin listoista löytyy vielä, esim. www.inesc-id.pt/~aml/trash.html arkistossa web.archive.org.

Ryhmä MIT:n jatko-opiskelijoita päätti paljastaa yhden tällaisen bogus-kokouksen. He tekivät ohjelman, joka generoi ulkonaisesti oikealta näyttäviä artikkeleita. Yksi tarjotuista bogus-artikkeleista hyväksyttiin, mikä osoittaa, ettei kukaan ollut sitä edes silmäilyt, vaikka ero on tietysti havaittavissa. Generointiohjelmaa, joka toimii www.pdos.csail.mit.edu/scigen/, kannattaa kokeilla. Samassa osoitteessa on blogi, jossa kerrotaan generointiohjelman "saavutuksista".

Miten sitten voitaisiin rajoittaa bogus-kokousten toimintaa? Eräs tapa voisi olla antaa kansainvälisille järjestöille, kuten ACM, IEEE ja IFIP, valvontatehtävä ja sopia yhteisesti, että vain hyväksytyt kokoukset katsotaan tieteellisiksi.