

Tiedonhausta informaatiolukutaitoon - henkilökohtaisia tutkimuspolkuja

Eero Sormunen

Tampereen yliopisto

eero.sormunen@staff.uta.fi

<https://orcid.org/0000-0002-1885-5045>

Professori Eero Sormusen jäähyväisluento Tampereen yliopistossa 22.11.2017.

Asiasanat: tutkimustyö; tiedonhaku; tiedonhankinta; informaatiolukutaito; esitelmät

Arvoisa rehtori, Arvoisa dekaani, Hyvät kollegat ja ystävät!

Kun dekaani esitti ajatuksen jäähyväisluennon pitämisestä, en ollut kovin innostunut asiasta. Jäin kyllä 1.4.2017 eläkkeelle professorin tehtävästä mutta jatkan osa-aikaisena tutkimusjohtajana kahdessa Suomen akatemian rahoittamassa hankkeessa vuoteen 2019 saakka. Kun olen vahvasti sitoutunut hankkeissa tehtävään tutkimukseen, jäähyväisluento ei vielä tuntunut ajankohtaiselta. Toisaalta pyyntö herätti kiinnostuksen koota ja reflektoida erityisesti tutkimukseen liittyviä kokemuksia akateemisen uran varrelta.

Esittelen luennossani henkilökohtaisia, enemmän tai vähemmän tietoisesti valittuja tutkimuspolkuja tiedonhaun ja tiedonhankinnan erityisalueilla sekä näihin valintoihin vaikuttaneita tekijöitä. Yritän myös taustoittaa valintojani informaatio- ja mediaympäristössä tapahtuneilla muutoksilla, jotka toisaalta vaikeuttivat työtä aiemmin valitulla tutkimuslinjalla mutta toisaalta loivat mahdollisuuksia aloittaa

jotain uutta. Sivuan esityksessä myös professionaalisen ja akateemisen uran vuorovaikutusta, joka on ollut itselleni tärkeä kysymys.

Tarkastelen tutkimuspolkuihini liittyviä kysymyksiä viidellä vuosikymmenellä. 1970-lukua voi pitää tässä esihistoriallisena kautena jolloin opiskelin ja valmistuin diplomi-insinööriksi. Sain määräaikaisen työn hervantalaisesta T3-kirjastosta, mikä vaikutti olennaisesti urani jatkokehitykseen. 1980-luvun työskentelin informaattikkona Valtion teknillisen tutkimuskeskuksen (VTT) teknillisessä informaatiopalvelulaitoksessa. Tehtävänimike oli tutkija ja työ sisälsi jossain määrin myös tutkimustehtäviä. Pääosan 1990-lukua hoidin apulaisprofessorin sijaisuuksia Tampereen yliopiston kirjastotieteen ja informatiikan laitoksella ja suoritin työn ohessa lisensiaatti- ja tohtoriopintoja. Heti 2000-luvun alussa väittelin, sain pian nimityksen tiedonhaun professuuriin ja vuosikymmenen loppuvuodet kuuluivatkin laitosjohtajana. 2010-luku on painottunut tutkijan tehtäviin kahden sapattivuoden ansiosta ja keväästä alkaen Kevan maksaman eläkkeen ja akatemian hankerahojuksen turvin. Voi sanoa että neljän vuosikymmen valmistautumisen jälkeen olen päässyt keskittymään tutkimukseen.

1970-luku: esihistoria

1970-luvun lopulla edistyneimmissä tieteellisissä kirjastoissa ja tietopalveluissa oli otettu käyttöön online-tiedonhakujärjestelmät. Hervannan T3-kirjastossa sain katsoa vierestä, kun informaattikko teki hakuja tekniikan alojen kirjallisuuden viitetietokannoista. Tietokannat sijaitsivat pääasiassa USA:n länsirannikon datakeskuksissa ja niitä käytettiin kohtuullisen alkeellisten tietoliikenneverkkojen välityksellä. Maksimissaan ns. paperipäätteen yhteysnopeus nousi jopa (!) 30 merkkiin sekunnissa. Ehkä kuitenkin olennaisempaa oli että kirjastonjohtaja Vesa Kautto antoi minulle tehtävän F.W. Lancasterin 'Paperless information systems' -kirjan innoittamana. Laskin arvion siitä, kuinka paljon kirjaston aikakauslehtikokoelman digitointi vaatisi tietokoneelta tallennuskapasiteettia ja minkälaisia ongelmia toteutukseen mahdollisesti liittyisi. Kirjoitin tuloksista tieteellisen kirjastoseuran Signum-lehden 7/1979 numeroon artikkelin, josta tuli ensimmäinen tutkimusartikkelin kaltainen tekstini (Sormunen, 1979). Nyt digitaalisten lehtikokoelmien käyttö on arkipäivän rutiinia, mutta toteutusmuoto on hiukan erilainen kuin 1970-luvulla ajattelimme.

1980-luku: asiantuntijapolulla

Aloitin 1980 VTT:n tietopalvelussa Espoossa, jatkoin seuraavana vuonna Ouluun perustamaan palveluyksikköä ja edelleen Tampereelle vastaavaan tehtävään vuonna 1984. Tietopalvelun kentällä pääsin hyvin perehtymään tutkijoiden tiedontar-

peisiin ja hankintaan eri laboratorioissa, sekä tekemään monipuolisesti kirjallisuuden ja patenteihin liittyviä tiedonhakuja. Tämä antoi hyvän pohjan tiedonhaun tutkimukselle myöhemmin. Informaatiopalvelulaitoksen johtaja Sauli Laitinen antoi myös vapaat kädet tehdä tutkimusta, kunhan hankin siihen rahoituksen. Yleensä tutkimustehtävät olivat uuteen tietopalvelutekniikkaan liittyviä selvityksiä. Kahdella vahvemmin tutkimuksellisella hankkeella oli erityistä merkitystä myöhemmin tutkijauralla. Yhdessä tietojenkäsittelytekniikan laboratorion kanssa kehittelimme Tekes-rahoitteisessa projektissa tiedonhaun älykästä käyttöliittymää. Hanke liittyi 80-luvun tekoäly- ja tietämysjärjestelmien tutkimuksen boomiin. Tutkimus auttoi saamaan otteen tiedonhaun tietämyksen mallintamisesta ja siitä tehty sivugradu oli lähtölaukaus jatko-opinnoille (Sormunen, 1989). Toinen tärkeä aihe oli suomenkielisten tekstitietokantojen hakutekniikat (Sormunen & Alkula, 1990). Aihepiiri oli hyvin keskeisessä roolissa Tampereen yliopiston tiedonhakututkimuksessa seuraavalla vuosikymmenellä.

Pidän asiantuntijatehtävissä hankittua kokemusta arvokkaan tutkijan uralle suuntautuessa sillä se antaa hyvän kokonaiskuvan ilmiöiden luonteesta käytännön kannalta. Toisaalta selvitysluonteiset tai puolivillaisiin menetelmin tehdyt empiiriset tutkimukset eivät anna vielä osaamispohjaa tehdä tutkimusta akateemisen yhteisön edellyttämällä järjestelmällisyydellä. Tähän tarvitaan tutkijakoulutusta.

1990-luku: tutkijakoulutuksessa

Kun tulin hoitamaan ensimmäistä määräaikaista tiedonhaun apulaisprofessorin sijaisuutta Tampereen yliopiston kirjastotieteen ja informatiikan laitokselle, oppiaineen akateeminen profiili oli jo määritetty vanhempien kollegojen Vakkarin ja Järvelinin linjausten mukaan pitkin 1980-lukua käydyssä debatissa. Tutkimukselle oli määritelty painopistealueet, joista tiedonhakua ja tiedonhankinta osoittautuivat vahvimiksi. Itse en ammatilliselta kentältä tullessa kokenut oppialan syvimmän teoreettisen olemuksen pohtimista kovin tärkeäksi, vaan huomioni kiintyi ensisijaisesti kiinnostavien ongelmien tunnistamiseen ja millä lähestymistavoilla niitä pystyttäisiin ratkomaan. Tässä mielessä olen ollut jatkuvasti oppiaineen suhteen opportunisti: tärkein tutkimusta motivoiva tekijänä oli ja on edelleen merkittävältä tuntuva yhteiskunnallinen tilaus, ei niinkään oppiaineen tutkimustradition vaaliminen.

Laitoksella tehtiin merkittävä investointipäätös lukuvuonna 1990–91 perustamalla tiedonhaun tutkimuslaboratorio, jossa luotiin palvelintietokoneelle oma tutkimusympäristö. Itse olin keskeisesti mukana 55000 suomenkielistä sanomalehtiartikkelia sisältävän TUTK-testitietokannan kokoamisessa, johon suunnittelimme 35 testihakuaihetta ja jäljitimme kuhunkin aiheeseen liittyvät relevantit artikkelit. Tiedonhakulaboratorion perustaminen oli linjassa tiedonhaun kansainvälisen tutki-

muksen päätrendin kanssa. Samaan aikaan rakennettiin USA:ssa kansainvälisenä yhteistyönä TREC-testikokoelmaa, joka toimi yhteisenä foorumina vuosittaisissa hakualgoritmien 'evaluointikilpailuissa'.

Keskeinen ero TUTK- ja TREC-testikokoelmien välillä oli se, että meidän koelmassamme dokumenttien relevanssi oli arvioitu neliportaisella asteikolla, kun kansainvälisessä kokoelmassa käytettiin binääristä relevantti-/epärelevantti-asteikkoa. Sovelsimme moniportaista relevanssiarviointia myöhemmin myös TREC-aineistoon (Sormunen, 2002), minkä raportointi sai merkittävää huomiota kansainvälisessä tutkijayhteisössä. Moniportaiset relevanssiarviot mahdollistivat monipuolisemmat koeasetelmat ja tuki myös uusien tiedonhaun tehokkuusmittarien kehittämisen, jotka levisivät laajaan kansainväliseen käyttöön (ks. Järvelin & Kekäläinen, 2002, s. 2875 viittausta 20.11.2017 Google Scholarin mukaan). 1990-luvun lopulla testikokoelmien käyttöliittymäksi kehitettiin tiedonhakupeli, joka oli varsin mielenkiintoinen oppimis-, demonstrointi- ja tutkimusympäristö (Halttunen & Sormunen, 2000).

Olin vuosikymmenen lopulla myös rakentamassa kuvatiedonhaun testiympäristöjä, jolla alueella meillä oli yhteinen 'Digitaalinen media' akatemiahanke Tampereen teknillisen korkeakoulun kanssa. Samaan teemaan liittyen olimme jäsenenä eurooppalaisessa multimediatiedonhaun MIRA-verkostossa (Sormunen, Markkula, & Järvelin, 1999).

Tiedonhaun tutkimuslaboratorio on hyvä esimerkki tutkimukseen liittyvistä investoinneista: rahaa sijoitetaan tekniikan lisäksi osaamisen kehittämiseen, jonka varassa tutkimusta pystytään viemään eteenpäin tehokkaasti ja pitkäjänteisesti. Kaikkiin investointeihin, niin myös tutkimusinvestointeihin, liittyy riskejä. Kaikki investoinnit sitovat resursseja valittuun tutkimuslinjaan pitkähköksi aikaa, mutta vain onnistuneet investoinnit tuottavat tuloksia. Tiedonhaun tutkimuslaboratorio testikokoelmineen mahdollisti laitoksen tiedonhakututkimuksen tason noston kansainväliselle tasolle. Valitettavasti jokaisella investoinnilla on oma rajallinen elinkaarensa. Yleiskäyttöisiin testikokoelmiin perustuvan tutkimuksen paras puhti alkoi omalta kannaltani hiipua jo 2000-luvun alkupuolella. Tärkein syy testikokoelmien alamäkeen oli tiedonhakuympäristön radikaali muutos internetin ja sen hakukoneiden kehittyessä. Tiedonhaun tutkimus pirstoutui kapeisiin erikoisaloihin ja tämä olisi vaatinut merkittäviä investointeja uudentyyppeisiin tutkimusympäristöihin. Testikokoelmien ja tiedonhakupelin ideat palasivat kuitenkin omaan tutkimukseeni, kun aloimme tutkia tiedonhaun opetusinterventioiden vaikuttavuutta ennen-jälkeen testeillä (Sormunen ym., 2017).

2000-luku: laitosjohtajuus

Väittelin vuonna 2000 ja sain nimityksen professoriksi vuoden 2003 alussa. Saman tien tulin valituksi informaatiotutkimuksen laitoksen johtajaksi, missä roolissa olin yli kuusi vuotta. Yliopiston hallinnon toimintamallien tiivistyminen ja myöhemässä vaiheessa tapahtunut laitosfuusio hypermedialaboratorion kanssa sitoi melkoisesti henkilökohtaisia voimavaroja. Tutkimustyö painottui pääsääntöisesti projektien ohjaamiseen ja toisen tai kolmannen kirjoittajan rooliin. Tämä tuottaa tietysti tutkimuksesta kiinnostuneelle turhauttavia kokemuksia.

Tiedonhaun oppimiseen ja tiedonhakupeliin kohdistunut mielenkiintoni kohdentui 2000-luvun alkuvuosina internetin tuomiin haasteisiin kehittää yleissivistävän koulutuksen kykyä ohjata oppilaista itsenäisiä ja kriittisiä tiedonhakijoita ja -käyttäjiä. Linjasin tutkimusintressini professorin virkaanastujaisesityksessäni vuonna 2003. Samalla totesin tutkimusaiheen vaativan hakeutumista yhteistyöhön kasvatustieteilijöiden kanssa. Aihe kiinnosti mitä ilmeisimmin rahoittajakin, sillä sain ensi yrittämällä yhdessä Esa Poikelan kanssa Suomen akatemian nelivuotisen rahoituksen WebSeal-hankkeeseen, jonka teemoina olivat nettitiedonhaku, informaatiolukutaito ja oppiminen. Pian ilmeni että kasvatustieteiden puolella ei ollut juurikaan aiempaa, tiedonhakemiseen liittyvää tutkimusta. Pääsimme kuitenkin jyvälle tiedonhaun ja oppimisen välisestä problematiikasta mutta tutkimussuunnitelmamme oli liian hajanainen jotta toimivaa poikkitieteellistä tutkimusta olisi syntynyt (Sormunen & Poikela, 2008).

2010-luku: varttuneena tutkijana

Vuonna 2010 loppui ns. oksanreikävuoroni laitoksen johtajana ja viime vaiheessa tutkinto-ohjelman vetäjänä. Olin vuoden 2011 Suomen akatemian rahoittamalla tutkimusvapaalla ja vapauduttuani viiden vuoden hakusulusta sain vuodeksi 2016 säätiöiden professoripoolin tukeman tutkimusvapaan. Akatemiassa meni myös läpi Web-Seal -hankkeen jatkoksi Eero Ropon kanssa tehty Know-Id-konsortiohakemus vuosille 2010–2013 (Ropo, Sormunen, & Heinström, 2015). Know-Id -hanke sai jatkoa kahdesta akatemiahankkeesta: ARONI (2015–19), jonka konsortiossa ovat mukana Miika Marttusen ryhmä Jyväskylän yliopistosta ja Auli Toomin ryhmä Helsingin yliopistosta, sekä iFuCo (2016–18), jossa ovat mukana Jyväskylän lisäksi, Marja Vauraksen ryhmä Turun yliopistosta sekä tutkimusryhmät kahdesta yliopistosta Chilessä. Voin todeta että nyt on tilaisuus tehdä mielenkiintoista tutkimusta ja onneksi siihen on lopulta aikaakin.

Useat konsortiohankkeet antavat mahdollisuuden reflektoida tieteidenvälisen

yhteistyön kokemuksia. Kun kaksi tutkimusalaa ja paradigmaa kohtaavat, on ensimmäinen aika tuskainen vaihe yrittää ymmärtää, mitkä ovat naapurin intressit, tavoitteet, teoreettiset lähtökohdat, ydinkäsitteiden tulkinat, menetelmälliset käytännöt, julkaisemiseen liittyvät konventiot, jne. Puhutaan usein myös yhteisestä kielestä. Vastaavasti on huolena ymmärtääkö naapuri oman tutkimusalamme vastaavat piirteet. Kokemukseni on että tuottavaan tieteidenväliseen tutkimukseen pääseminen edellyttää että osapuolilla on lähtökohtaisesti yhteinen käsitys tutkimuskohteesta. Toisin sanoen on tunnistettu yhteiskunnallisesti merkittävä ongelma, johon tutkimuksella haetaan yhteistä ratkaisua. Jos tämä ehto täyttyy, vaaditaan vielä paljon keskusteluja ja yhdessä tekemistä ennen kuin löytyy riittävän integroitu tutkimussuunnitelma, jonka toteutuksessa yhteistyö ja toisiaan täydentävä osaaminen alkaa tuottaa hedelmää.

Käynnissä olevissa ARONI- ja iFuCo-hankkeissa informaatiotutkimuksen ja kasvatustieteiden tutkijoiden roolit ja vuorovaikutus onnistuttiin jo hakemuksen tutkimussuunnitelmassa määrittelemään hyvin. Konsortion jäsenillä on omia osahankkeita, jotka kehittävät aineistonkeruun instrumentteja yhteiseen päätutkimukseen tai palvelevat sen pilotteina. Lopulta integraatio muodostuu yhteisten aineistojen pohjalta työskentelystä, jossa osapuolet voivat soveltaa vahvinta osaamistaan. Odotan suurella mielenkiinnolla, käynnissä olevien hankkeiden tuloksia. Alku on hyvin lupaava (Sormunen ym., 2017).

Muutama sana oppiaineen tulevaisuudesta Tampereen yliopistossa

Kansainvälisissä konferensseissa tai vierailukohteissa on helppo huomata että Tampereen yliopiston tiedonhaun ja tiedonhankinnan tutkimus on laajalti tunnettua ja arvostettua. Tutkimus on lähellä kansainvälistä kärkeä myös yliopiston oman ulkoisen arvioinnin, Suomen akatemian selvityksen ja useiden kansainvälisissä joulleissa julkaistujen arviointien perusteella. Viimeksi tänä vuonna julkaistiin laaja kansainvälinen analyysi informaatiokäyttäjien tutkimusimpaktista alan tutkimuksessa (Soheili, Khasseh, & Mousavi-Chelak, 2017). Kuudella vaikuttavuuden eri puolia edustavilla indikaattoreilla lasketun kokonaisarvion perusteella Savolainen, Järvelin ja Vakkari sijoittuivat 15 kärkitutkijan joukkoon eli 20 prosenttia informaatiokäyttäjien huippututkijoista on tästä yliopistosta. Koulutuksen osalta ei ole saatavilla vastaavia riippumattomia arvioita, mutta se voidaan todeta varmuudella että Tampereen yliopisto on volyymiltään kirjasto- ja informaatioalan suurin akateeminen kouluttaja Suomessa.

Tilanne ei näytä tulevaisuuden osalta yhtä valoisalta. Vuoden 2018 alkupuolella kaikki oppiaineen neljä professoria ovat siirtyneet eläkkeelle. Koko ydinjou-

kon eläköitymisen nopea tahti aiheuttaa ongelmia, jotka ovat olleet ennakoitavissa jo vuosien ajan. Olemme useaan otteeseen yrittäneet saada huomiota asialle sekä edellisessä tieteenalayksikössä että nykyisessä tiedekunnassa mutta laihoihin tuloksin. Yhdestäkään vapautuvasta tehtävästä ei ole käynnistetty tehtävän määrittely- ja rekrytointiprosessia. Meille on esitetty perusteluna säästötarve ja budjetin niukuus. Budjetin tasapainotuksen lähtökohtana on karsia kustannuksia missä se on helpointa eli vapautuvista tehtävistä. Tällöin oppiaine, joka on sattumalta meidän kaltaisessamme erityisessä tilanteessa, joutuu rankimman päälle maksajaksi. Vuosibudjetointi ohjaa nyt strategisia päätöksiä vaikka suunnan pitäisi olla juuri päinvastainen.

Arvoisa rehtori ja arvoisa dekaani!

Ehdotan että nimeätte pikavauhtia kriisiryhmän etsimään ratkaisua informaatiotutkimuksen rekrytointiongelmaan. Ratkaisun viivyttäminen vuoden 2018 yli johtaa mielestäni tilanteeseen, jossa voimme heittää hyvästit informaatiotutkimukselle vakavasti otettavana toimijana Tampereen yliopistossa.

Vaikka en voikaan olla tyytyväinen oppiaineen saamaan kohteluun, haluan omien työsuhteeni osalta kiittää yliopistoa menneiden vuosien tarjoamista haasteista ja mielenkiintoisista tehtävistä sekä mahdollisuudesta jatkaa tutkimuksen parissa. Kiitän myös lähintä työyhteisöäni, tutkijoita, opettajia ja hallinnon väkeä näistä vuosista. Tässä porukassa on ollut mukava työskennellä.

Lähteet

- Halttunen, K., & Sormunen, E. (2000). Learning Information Retrieval through an Educational Game. Is Gaming sufficient for learning? *Education for Information*, 18(4), 289–311. <https://oadoi.org/10.3233/EFI-2000-18403>
- Järvelin, K., & Kekäläinen, J. (2002). Cumulated Gain-based Evaluation of IR Techniques. *ACM Trans. Inf. Syst.*, 20(4), 422–446. <https://oadoi.org/10.1145/582415.582418>
- Ropo, E., Sormunen, E., & Heinström, J. (2015). *Identiteetistä informaatiolukutaitoon: tavoitteena itseenäinen ja yhteisöllinen oppija*. Tampere: Tampere University Press. <http://urn.fi/URN:ISBN:978-951-44-9920-3> (luettu 16.12.2017)
- Soheili, F., Khasseh, A. A., & Mousavi-Chelak, A. (2017). The most influential researchers in information behaviour: An integrative view on influence indicators. *Aslib Journal of Information Management*, 69(2), 215–229. <https://oadoi.org/10.1108/AJIM-01-2017-0027>
- Sormunen, E. (1979). Tieteellisen informaation sähköinen tallennus-, haku- ja siirtojärjestelmä. *Signum*, 12(7), 127–132.
- Sormunen, E. (1989). *An analysis of online searching knowledge for intermediary systems* (s. 81). Espoo: Technical Research Centre of Finland.
- Sormunen, E. (2002). Liberal Relevance Criteria of TREC - Counting on Negligible Documents? Teoksessa *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Deve-*

- lopment in Information Retrieval* (ss. 324–330). New York, NY, USA: ACM. <https://oadoi.org/10.1145/564376.564433>
- Sormunen, E., & Alkula, R. (1990). *Suomenkielisten tekstitietokantojen tallennus- ja hakutekniikkojen kehittäminen : esitutkimusraportti*. Espoo: Technical Research Centre of Finland.
- Sormunen, E., & Poikela, E. (Toim.). (2008). *Informaatio, informaatiolukutaito ja oppiminen* (s. 266). Tampere: Tampere University Press.
- Sormunen, E., González-Ibáñez, R., Kiili, C., H.T. Leppänen, P., Mikkilä-Erdmann, M., Erdmann, N., & Escobar, M. (2017). A Performance-based Test for Assessing Students' Online Inquiry Competences in Schools. Teoksessa S. Špiranec, S. Kurbanoglu, J. Boustany, E. Grassian, D. Mizrachi, L. Roy, & D. Ko (toim.), *Fifth European Conference on Information Literacy (ECIL) - Abstracts*. Saint-Malo, France. <http://bit.ly/ecil-2017-abstracts> (luettu 30.12.2017)
- Sormunen, E., Markkula, M., & Järvelin, K. (1999). The Perceived Similarity of Photos - Seeking a Solid Basis for the Evaluation of Content-based Retrieval Algorithms. Teoksessa *Proceedings of the MIRA Final Conference*. Glasgow, UK. <https://www.researchgate.net/publication/2541411> (luettu 17.12.2017)