

Lisänäkökohtia Diplomatarium Fennicum - tutkimustietokantahankkeen toteutukseen

Eljas Orrman tarttui *Historiallisen Aikakauskirjan* vuoden 2020 ensimmäisessä numerossa julkaistussa puheenvuorossa suomalaisen keskiajantutkimuksen kannalta erittäin tärkeään aiheeseen eli Diplomatarium Fennicum (DF) -tietokannan sisältöön ja sen kehittämiseen. Kommentoimme tässä vastauspuheenvuorossamme Orrmanin erinomaisia ehdotuksia ja avaamme DF-projektin nykytilaa ja tulevia mahdollisuuksia.

Vaikka Koneen Säätiön rahoitus Kansallisarkiston johtamalle DF-hankkeelle päättyikin jo muutama vuosi sitten, tietokannan kehittämisajatusta ei ole suinkaan haudattu, vaan Kansallisarkisto on nimittänyt hankkeelle toimituskunnan, jonka tarkoituksena on jatkaa ja kehittää tietokantaa jatkossakin.¹ Eljas Orrmanin puheenvuorossaan esiin tuomat seikat ovat tärkeää osviittaa tässä työssä, ja siksi me Diplomatarium Fennicumin toimituskunnan tämänhetkinä vetäjinä halusimme tarttua tilaisuuteen kommentoida Orrmanin sanomaa.

Orrman tarttui kirjoituksessaan useampaan tärkeään yksityiskohtaan ja DF-tietokannan kehittämismahdollisuuteen: hakumahdollisuuksien parantamiseen ja täydentämiseen, tietokannan kasvattamiseen uusia asiakirjoja lisäämällä, joukkoistamisen mahdollisuuteen tietokannan täydentämisessä ja tietojen korjaamisessa, tietokannan sisällön parempaa avaamista sen käyttäjille myös muilla kuin suomen kielellä sekä Kansallisarkiston rooliin keskiaikaisten lähteiden julkaisemisessa. Toimituskunnan näkökulmasta Orrmanin puheenvuoro on erityisen arvokas nostaessaan esiin juuri niitä kehityskohteita, jotka eniten parantaisivat tietokannan käyttömahdollisuuksia. Monia näistä kohteista harkittiin toteutettavaksi myös viimeisimmän DF-hankkeen (2015–2018) aikana. Hankkeessa jouduttiin kuitenkin resurssi- ja aikataulusyistä keskittymään tietokannan suurimpien kokonaisuuksien – kuten käyttöliittymän – toteutukseen, mikä lopulta pakotti jättämään monta tutkijoiden näkökulmasta yhtä tärkeää asiaa jatkokehitykseen. Tulevaisuudessa toimituskunnan keskeinen tehtävä onkin löytää kehityskohtiin toteutuskeinot. Jatkokehityksestä puhuttaessa tulee kuitenkin tiedostaa tämänhetkinen resurssitilanne: DF-tietokannan kehittämiseen ei ole pysyvää rahoitusta, eikä tällä hetkellä myöskään käytettävissä olevaa hankerahoitusta. Kehitystarpeet typistyvät siis helposti resurssikysymyksiksi. Toimituskunnan puolesta pyrimme kuitenkin vastaamaan esitettyihin toiveisiin muutoinkin kuin ”tekisimme toki, jos olisi rahaa”, ja avaamaan laajemmin tietokannan kehittämissuunnitelmia.

Puheenvuorossaan Eljas Orrman toteaa, että kirjamuotoisten julkaisujen, kuten *Finlands medeltidsurkunder* -julkaisun (FMU) tai *Turun tuomiokirkon mustan kirjan* painetun version (REA), lopussa olevat hakemistot ovat kätevä tapa etsiä tietoa teoksista. Internetissä olevassa tietokannassa tällaista mahdollisuutta ei ole, mutta kuten Orrman toteaa, DF:n hakutoiminnot korvaavat suurelta osin kirjamuotoisten julkaisujen hakemistojen tehtävän. DF-hankkeen alusta asti

¹ Tätä kirjoitettaessa toimituskuntaan kuuluvat dos. Tapio Salminen (Tampereen yliopisto), prof. Kirsi Salonen (Turun yliopisto) ja FM Lauri Leinonen (Helsingin yliopisto, sihteeri), jotka muodostavat toimikunnan työtä johtavan koordinaatiotyöryhmän. Muut jäsenet ovat FM Seppo Eskola (Helsingin yliopisto), prof. Tuomas Heikkilä (Helsingin yliopisto), FM Maria Kallio (Kansallisarkiston edustaja), prof. Anu Lahtinen (Helsingin yliopisto), prof. Harry Lönnroth (Jyväskylän yliopisto) ja FL Minna Sandelin (Turun yliopisto).

on ollut tarkoitus saada tietokantaan toimivat ja kattavat hakumahdollisuudet, joiden avulla olisi helppo löytää Suomen keskiaikaan liittyvät henkilöt, paikat ja asiat. DF-tietokannassa on tällä hetkellä hakutoiminto, jonka avulla on mahdollista löytää edellä todetut maininnat julkaistuista asiakirjoista, mutta olemme hyvin tietoisia siitä, että aivan kaikki henkilöt tai paikat eivät ole vielä hakutoiminnon piirissä, koska kaikkia lähes 7000 asiakirjaa ei ole indeksoitu hakua varten asianmukaisesti. Asiakirjojen antajatiedot kuitenkin löytyvät täydellisinä DF:n antajarekisteristä. DF-toimituskunta yrittää parhaansa mukaan edistää myös asiakirjojen indeksoimista, mutta tosiasiasa tämä työ vaatisi palkattua työvoimaa.

Sen sijaan Orrmanin konkreettinen ehdotus siitä, että FMU:n ja REA:n hakemistot voisi tallentaa tietokantaan pdf-muodossa helpottamaan osaltaan tutkijoiden hakumahdollisuuksia, on toteutettavissa, ja otamme sen huomioon tulevissa DF-tietokannan päivityksissä. Hakemistot ovat jo saatavilla pdf-muotoisina osana teosten kokonaisia digitoituja versioita, jotka ovat vapaasti saatavilla tietokannassa.² Hakemistojen ja kirjallisuusosion saatavuutta ja löydettävyyttä on siten suhteellisen helppo parantaa.

Orrman huomauttaa puheenvuorossaan tietokannan sisältöön liittyen, että tietokannasta puuttuu useita kymmeniä tutkijoiden käyttämiä ja muualla julkaisemia asiakirjoja, ja että olisi tärkeää saada myös nämä asiakirjat tietokantaan. Tässä Orrman on täysin oikeassa. DF-toimituskunnan seuraavana suurempana tavoitteena onkin tietokannan täydentäminen sieltä vielä puuttuvien keskiaikaisten Suomea tai suomalaisia koskevien asiakirjojen osalta. Tällä hetkellä tietokannan ytimen muodostavat Reinhold Hausenin FMU:ssa julkaisemat asiakirjat, joita on täydennetty REA:ssa olevilla editioilla. Lisäksi tietokantaan lisättiin päättyneen DF-hankkeen puitteissa Sara Risbergin ja Kirsi Salosen julkaisemat editiot Suomea koskevasta Vatikaanin arkiston penitentiariaattineiston merkinnöistä (DF 6727–6856) sekä Tapio Salmisen editiot yhdeksästä, pääosin aiemmin julkaisemattomasta Tallinnan Kaupunginarkiston Suomea koskevasta pergamenttikirjeestä (DF 6857–6865). Tietokantaa on myöhemmin päivitetty myös Nådendalsdiplomen-hankkeen puitteissa tehdyillä uusilla editioilla Naantalin luostaria koskevista keskiaikaisista asiakirjoista. Lisää uusia editioita julkaistaan syksyllä tehtävässä tietokannan päivityksessä.

Jatkossa olisi ensisijaisen tärkeää, että tietokantaan saadaan täydennettyä tällä hetkellä tunnetut, mutta sieltä vielä puuttuvat asiakirjat. Työtä tämän eteen on jo tehty. Viimeisimmän DF-hankkeen aikana kerättiin jatkokehitystä varten lista sellaisista julkaistuista keskiaikaisista asiakirjoista, jotka eivät sisälly FMU:hun tai REA:an. Toimituskunta pyrkii jatkossa saamaan nämä asiakirjat julkaistua myös DF:n kautta. Ilman tälle työlle kohdennettua henkilöresurssia asiakirjoja saadaan kuitenkin julkaistua vain tipoitain, sen verran suuri työ liittyy tekstien muuttamiseen digitaaliseen muotoon, tietokantamuutoksiin, sekä muun muassa julkaisulupien saamiseen. Samoin olisi tietenkin tärkeää, että tulevaisuudessa löytyvät uudet Suomea tai suomalaisia koskevat asiakirjat saataisiin viivytystä mukaan tietokantaan. Toimituskunta pitää aktiivisesti silmällä tutkimuskenttää, jotta uudet löydöt saataisiin tuoreeltaan DF:n piiriin. Myös tarkkaan kohdennettu aineiston kerryttäminen

² Projekti digitoi kokonaisuudessaan Reinhold Hausenin julkaisut *Finlands Medeltidsurkunder (1910–1935)*, *Registrum Ecclesiae Aboensis eller Åbo domkyrkas svartbok (1890)* ja *Finlands medeltidssigill (1900)*, sekä A. I. Arwidssonin teoksen *Handlingar till upplysning af Finlands häfder (1846–1857)* ja Edward Grönbladins *Nya källor till Finlands medeltidshistoria (1857)*. Nämä teokset ovat saatavilla DF-tietokannan Kirjallisuutta/Litteratur/Literature -osiossa (<http://df.narc.fi/info/literature>). Näiden teosten lisäksi kirjallisuuslistaan on koottu muualla digitoituja ja vapaasti saatavilla olevia keskeisiä teoksia.

on ollut esillä. DF-toimituskunnan aikeena onkin hakea rahoitusta pienemmille, selkeästi määritellyille alahakkeille, joiden puitteissa voitaisiin käydä yksi kerrallaan läpi pienempiä aineistoja ja lisätä ne yksi kerrallaan tietokantaan. Yksi tällainen hanke voisi olla Orrmanin mainitseman tuomiokirja-aineiston läpikäynti.

Toimituskunnan tavoitteena on myös hyödyntää DF:n täydentämisessä opiskelijoita sekä samalla perehdyttää heitä keskiaikaisen aineiston kanssa työskentelyyn. Tällä hetkellä DF-toimituskunnan puitteissa on menossa kokeilu, jossa Turun yliopiston opiskelijat käyvät läpi Tukholman kaupungin raadin oikeuspöytäkirjoja ja etsivät sieltä puuttuvia Suomea tai suomalaisia koskevia mainintoja. Löytyneiden asiakirjojen tuominen DF-tietokantaan vaatii kuitenkin muutakin kuin asiakirjojen paikallistamisen. Asiakirjat pitää editoida asianmukaisesti ja niiden sisällöstä pitää laatia yhteenveto eli regesta, mikä on paitsi aikaa vievää, myös vaatii pitkälle vietyä asiantuntemusta ja kielitaitoa. Kokeilun kautta saamme kuitenkin perehdytettyä innostuneita opiskelijoita lähteiden julkaisemisen saloihin ja lisättyä alan osaamista Suomessa.

Orrman ottaa julkaisussaan esille myös sen tärkeän seikan, että tutkijayhteisössä on olemassa paljon tietoa DF-tietokannassa olevista virheellisistä tai puuttuvista yksityiskohdista ja asiakirjoista, ja että olisi tärkeää saada tämä tieto DF:n toimituskunnan tietoon. Näin voisimme tietokannan päivitysten yhteydessä korjata löytyneitä virheitä säännöllisesti. DF-toimituskunta onkin tähän liittyen tehnyt päätöksen, että sivustolle tehdään uusi toimivampi palautekanava, jonka kautta tietokannan puutteita havainneet tutkijat voivat ilmoittaa havainnoistaan. Tehdyt havainnot kootaan yhteen, ja niitä käsitellään toimituskunnan kesken, jotta virheet ja puutteet saadaan korjattua säännöllisesti. Laajemman joukkoistamistyökalun toteuttaminen on ikävä kyllä toimituskunnan resurssien ulottumattomissa tällä hetkellä.

Eljas Orrman nosti puheenvuorossaan esille, että tällä hetkellä tietokannassa oleva ohjeistus on vain suomenkielistä ja että ulkomaisten tutkijoiden on siksi hyvin hankala käyttää tietokantaa. Tähän pitkään tiedossa olleeseen ongelmaan on tulossa parannus jo lähiaikoina, sillä tietokannan suomenkieliset tekstisisällöt on jo ehditty kääntää niin ruotsiksi kuin englanniksi, ja ne liitetään DF-sivustolle syksyn päivitysversiossa. Kieliongelma ei kuitenkaan ratkea täysin tällä päivityksellä, sillä tietokannassa tällä hetkellä olevissa FMU:sta ja REA:sta kopioituissa asiakirjoissa on pääsääntöisesti ainoastaan Reinhold Hausenin 1900-luvun alkupuolella ruotsiksi laatimat sisältötiivistelmät eli regestat. Jotta DF-tietokantaa saataisiin laajempaan käyttöön suomalaistenkin keskuudessa, nämä vanhentuneella ruotsilla laaditut tiivistelmät pitäisi sekä modernisoida nykyruotsiksi että kääntää suomeksi. Tämä olisi valtava työ, joka vaatisi ehdottomasti oman erillisrahoituksensa. Toimituskunnan keskuudessa on myös pohdittu sitä, että tiivistelmien kääntämisen yhteydessä niiden sisältöä olisi myös syytä päivittää, sillä Hausen ei ole aina tavoittanut tiivistelmissään kaikkien asiakirjojen sanomaa. Tiivistelmistä on myös monesti jätetty pois asioita tai henkilöitä, jotka nykyisten standardien mukaan pitäisi mainita tiivistelmissä. Tiivistelmien korjaaminen näiltä osin vaatisi siis myös alkuperäisten asiakirjojen lukutaitoa.

Orrman nostaa kirjoituksessaan esille myös, että tietokannan esittelysivulla ei avata riittävän hyvin sitä, mitä tietokannassa on, mitä sieltä puuttuu, mitä asiakirjoille on tehty ja miten siellä olevia asiakirjoja pitäisi tulkita ja käyttää. Tunnistamme hyvin tämän puutteen, ja asian korjaamiseen on ryhdytty. Tietokannan esittelyyn tullaan lisäämään juuri näitä tietoja, ja yritämme muutenkin ottaa paremmin huomioon tietokannan käyttäjien tarpeet sisällön ymmärtämiseksi.

Monet keskeiset kysymykset ovat siis ratkottavissa DF:n käyttäjien ja toimituskunnan paremmalla viestinnällä sekä tietokannan ajantasaisilla korjauksilla. Kun palautekanavat on saatu kuntoon, voi toimituskunta tehdä DF:ään pienimuotoisia korjauksia sitä mukaa, kun käyttäjät löytävät korjattavaa. Jatkossa myös tietokannan päivitys sujuu aiempaa jouhevammin. Kansallisarkisto on kehittämässä tietokantojensa hallintaa, ja osana tähän liittyvää keskustelua DF:lle on hahmoteltu ensi syksystä 2020 alkaen puolivuositaisia päivityksiä. Korjaukset sekä asiakirjojen lisäykset saadaan siis jatkossa suhteellisen nopeasti käyttäjien näkyville, ja toimituskunta pyrkii saamaan kuhunkin päivitykseen joitain korjauksia ja lisäyksiä tehtyä. Päivityksistä tullaan tiedottamaan DF-tietokannan etusivulla sekä Kansallisarkiston tiedotuskanavilla.

Kaikkein keskeisimpänä asiana on kuitenkin syytä vielä kerran nostaa esille DF:n kehittämisen taloudellinen puoli. Kuten jo monissa kohdin tätä vastinetta on todettu, tietokannan kehittäminen ja laajentaminen ei ole mahdollista ilman erillistä rahoitusta. Tarvittavan rahoituksen määrä riippuu luonnollisesti haluttujen muutosten laajuudesta. Tietokantaa on mahdollista täydentää esimerkiksi palkkaamalla tutkijoita käsittelemään rajattuja aineistokokonaisuuksia muutaman kuukauden työpanoksella. Suuremmat tekniset muutokset taas vaativat käytännössä suurempia monivuotisia hankkeita toteutukseen. Korjaukset sekä yksittäiset tietopäivitykset taas olisivat helpoimmin toteutettavissa pienimuotoisella jatkuvalla rahoituksella. DF-toimituskunta onkin ottanut tehtäväkseen ulkopuolisen rahoituksen etsimisen. Eteneminen pikkuhiljaa erilaisten alahankkeiden puitteissa on näkemysellemme mukaan tällä hetkellä hedelmällisin tapa toimia aineiston kerryttämiseksi, mutta haemme myös jatkuvaa rahoitusta toiminnan kehittämiseen. Eljas Orrman esittää kuitenkin puheenvuorossaan myös, että Kansallisarkiston palkkalistoilla pitäisi olla keskiajan lähteiden asiantuntija, jolla olisi riittävästi tietotaitoa lähteiden julkaisemiseksi. Tällainen mahdollisuus olisi DF-tietokannankin kannalta todella käyttökelpoinen, ja se toisi Suomen samalla tasolle Ruotsin, Norjan ja Tanskan kanssa, sillä niissä on kussakin omat, pysyvät instituutiot keskiajan lähteiden julkaisemiselle.

Lauri Leinonen, Tapio Salminen & Kirsi Salonen

DF-toimituskunnan koordinaatioryhmä