

# KANSALLISKIRJASTON TUTKIJAPALVELUITA KEHITETÄÄN VERTAISOPPIMALLA

Kansalliskirjasto rakentaa digitaalisten aineistojen käyttöä edistäviä tutkijapalveluita Digitaalinen avoin muisti (DAM) -hankkeessa. Hanke parantaa digitaalisten aineistojen saatavuutta, saavutettavuutta ja käytettävyyttä. Samalla kehitetään tutkimusyhteistyötä ja oman henkilöstön osaamista teknisissä ja tutkimuksellisissa kysymyksissä. Artikkelit kertoo hankkeen kiinnostavasta eurooppalaisten yliopistojen ja library labien benchmarkkaustyöstä.

**D**igitaalisuus on haastanut kirjastoja eri tavoin tähän mennessä jo vuosikymmeniä. Kirjastojen on mukautettava digitaalisuuden aiheuttamiin muuttuneisiin käyttäjien tarpeisiin. Näin on myös Kansalliskirjastossa, jossa paine aineistojen digitointiin ja kaikenlaisten digitaalisten aineistojen avaamiseen käyttäjille on kova.

Digitaalisissa muodoissa tarjottavaa aineistoa vaaditaan käytettäväksi yhä enemmän ja vapaammin, mutta aineistojen vapaaseen käyttöön liittyy runsaasti lainsäädännöstä riippuvaisia näyttö- ja käyttörajoitteita. Näitä on Kansalliskirjastossa helpotettu neuvottelemalla tekijänoikeusjärjestöjen kanssa ja avaamalla joitakin aineistoja rajoitetusti tutkimuskäyttöön (Tutkain-hanke).

Kansalliskirjaston digitaalisia aineistoja tutkitaan kevään 2020 tutkijoille suunnatun kyselyn mukaan enimmäkseen perinteisin menetelmin, mutta kasvussa olevat tiedonlouhin-

nan menetelmät edellyttävät digitaalisten aineistojen muuttamista dataksi. Kokoelmat dataksi eli *Collections as data* -ajattelulla viitataan aineistojen analysoitavuuteen tietokoneavusteisesti. Kaikki digitaalinen aineisto ei ole aina sellaisenaan sopivaa tietokoneellisesti luettavaksi, vaan tarvitsevat omat formaattinsa. Näitä formaatteja voivat olla esimerkiksi MARC-muotoiset metatiedot tai datapakettien METS/ALTO. Formaattit rakenteellistetaan osana digitointiprosessia, jotta raakadatan muuttaminen tietokoneluettavaksi ja edelleen käytettäväksi digitaalisilla menetelmillä onnistuu.

## Kehittämishankkeessa haetaan mallia eurooppalaisista verrokeista

Haastattelimme ja havainnoimme DAM-hankkeessa kevään ja kesän 2020 aikana, millaisia toimintoja muualla Euroopassa on kehitetty isojen digitaalisten aineistojen tutkimuskäytön ym-



pärille ja millaisia datapalveluja kunkin maan kansalliskirjastoihin verrattavat organisaatiot tarjoavat tutkijoilleen. Vertaisoppimiseen perustuvaa tiedonhankintaa kutsutaan benchmarkkaukseksi.

Benchmarkkaushaastattelumme koostuivat seitsemän eri organisaation jäsenistä ja kaikkiaan haastateltavia oli 12. Haastateltavamme olivat Saksasta, Ranskasta, Luxemburgista, Belgiasta, Espanjasta, Tanskasta ja Ruotsista. Kirjoitimme raportin haastatteluista, jonka lähetimme myös kommentoitavaksi haastattelemllemme henkilöille (Näpärä & Liukkonen 2020). Lisäksi hyödynsimme runsaasti esimerkiksi Alankomaista ja Britanniaasta saatavaa tietoa digitaalisten aineistojen yhteyteen luoduista tutkijapalveluista.

Havainnoimme myös laajemmin digitaalisten ihmistieteiden tutkimuksen kenttää, tutkijapalveluiden internetsivuja sekä osallistuimme erilaisiin virtuaalitapahtumiin ja -keskusteluihin. Internetissä tapahtuva havainnointi korostui erityisesti maaliskuusta 2020 alkaen, koska benchmarkkauksen aloitustapahtumaksi suunniteltu Riian digitaalisten ihmistieteiden DHN2020-konferenssi siirrettiin. Myös moni muu tapahtuma siirtyi verkkoon, mikä osittain mahdollisti osallistumisen myös sellaisiin tapahtumiin, joihin ei olisi muutoin ollut mahdollista osallistua.

### Labit tutkimuksen kumppanina

Benchmarkkauksemme kohteena olevilla *library labeilla* ei ole vain yhtä

määritelmää ja toimintamallit vaihtelevat hieman, mutta niistä on löydettävissä runsaasti myös yhteisiä piirteitä. Haastatteluissamme labeilla saatettiin viitata konkreettiseen fyysiseen tilaan, jolloin lab tarkoitti lähinnä digitaalisia tutkimusmenetelmiä testattavaa kokeellista laboratoriota. Joillekin haastateltaville lab oli enemmän vain www-sivu, jonne koottiin yhteen digitaaliset aineistot sekä niiden analyysiin tai visualisoitiin tarvittavat työkalut. Kuitenkaan pelkkää internetsivua ja siellä olevaa yhteydenottomahdollisuutta ei koettu riittävänä sille, että palvelu olisi määritelty haastateltavien puheessa labiksi, vaan ehtona pidettiin aitoa vuorovaikutusta tutkijoiden ja aineistojen välillä.

Benchmarkkauksemme perusteella labit nähdään kohtaamispaikkana kirjastojen digitaalisille aineistoille ja tutkimukselle. Tärkeänä pidettiin tutkijoiden tuottamaa hyötyä digitaalisille aineistoille ja tutkijan mahdollisuuksia käyttää aineistoja omassa tutkimuksessaan. Ihanteellisessa tapauksessa molemmat hyötyvät yhteistyöstä, mitä voi oikeastaan pitää kaikenlaisen yhteistyön peruseriaatteena. Monilla labeilla on käytössä erilaisia kriteereitä, joiden perusteella tutkimuksen yhteistyökumppanit valitaan erillisissä arviointiryhmissä, joissa arvioidaan hankkeen hyötyjä suhteessa kirjaston omiin aineistoihin ja strategiaan tavoitteisiin.

Tutkimusyhteistyökumppanuuden muotoja on kuitenkin erilaisia. Niistä intensiivisin yhteistyön muoto on tutkijaresidenssi. Tässä kontekstissa merkittävimmät tutkijaresidenssit si-

jaitsevat Lontoon British Libraryssa ja Alankomaiden Koninklijke Bibliotheekissa. Niihin haetaan erityisesti uransa alkuvaiheessa olevia väitöskirjatutkijoita. Tilaa tarjotaan Alankomaissa puoleksi vuodeksi ja Briteissä jopa useiksi vuosiksi, jolloin heistä tulee oikeastaan osa kirjaston tutkimusyhteisöä.

Residenssissä on mahdollista hyödyntää sellaisia aineistoja, joita ei saa viedä esimerkiksi tekijänoikeuksien ja aineistojen harvinaisuuden vuoksi kirjaston ulkopuolelle. Residenssityöskentelyssä olennaista on päästä työskentelemään aineistojen parissa pitkään ja keskeytyksettä, jolloin tutkijat voivat kehittää digitaalisten aineistojen kontekstissa niiden käsittelyyn ja analyysiin uusia menetelmiä. Tavoitteena on tarjota digitaalisten aineistojen käyttöön kehitetyt työkalut myös seuraaville residenssitutkijoille.

Tutkijoille tarjotaan labeissa runsaasti aineistoja myös internetin kautta. Tärkeimpiä aineistoja näyttävät olevan kunkin maan digitoituidet sanomalehdet. Suurin osa tutkijoiden käyttöön tarjottavasta digitaalisesta aineistosta ja datasta sisältää vain tekijänoikeudesta vapaita aineistoja, mutta erityisehdoin labeissa on saatavissa tutkimuskäyttöön myös tekijänoikeuksia sisältävää dataa. Erilaiset tulkinnot tekijänoikeuksien alaisesta datasta vaikuttavat kansainvälisiin yhteistyömahdollisuuksiin, mutta eivät tee sitä mahdolltomaksi. Tästä ovat esimerkkinä sanomalehtiaineistojen kehittämiseen keskittyvät NewsEye- ja Impresso-hankkeet sekä muut Euroopassa maantieteelliset rajat ylittävät yhteiskuviot.

## Henkilöstön osaaminen

Tarkasteltaessa labeihin haettavan henkilöstön osaamis pohjaa tarvittavan osaamisen voi jakaa viiteen kategoriaan 1) tutkimuksen tuntemus, 2) aineiston tuntemus, 3) IT-osaaminen, 4) juridinen osaaminen ja 5) projektinhallinta- ja johtamistaidot. Osaamiskategoriat ovat limittäisiä ja yhdellä henkilöllä käytännössä pitää olla osaamista eri kategorioissa, sillä labeissa työskenteli täysipäiväisesti usein vain 1–3 henkilöä.

Suurimmalla osalla labeissa työskentelevillä henkilöillä oli itsellään tutkijatausta ja monella oli tohtorin tutkinto. Lisäksi osalle tutkimuksen tekeminen kuului osa-aikaisesti työkuvaan. Henkilöiden tieteenalaedustus oli varsin laaja. He olivat esimerkiksi kirjasto- ja informaatioalan ammattilaisia, tietojenkäsittelijöitä (esim. data scientist, IT specialist) ja digitaalisten ihmistieteiden edustajia. Joidenkin tausta painottui humanismiin ja joillakin tekniseen osaamiseen, mutta he muistivat korostaa kirjastoaineistojen keskeytyttä molemmista tulokulmista.

Kokonaisuudessaan havainnoisamme korostui labien monitieteisyys ja moniammatillinen osaaminen. Se nähtiin etuna, koska digitaalisten ihmistieteiden ja muiden tietokoneisiin tukeutuvien tutkimussuuntausten kehittämällä uusilla teknisillä ratkaisuilla on mahdollista nähdä kauemmas myös perinteisissä ihmistieteellisissä kysymyksissä. Menetelmällisesti puhutaan usein kaukoluvusta (distant reading), joka vertautuu ihmistieteilijöille tuttuun lähilukuun. Tutkimuskysymyk-

siin vastaamiselle pyritään molemmissa selittämään ihmisen toimintaa eri tilanteissa ja aikakausina, jolloin ymmärrys eri asioista kasvaa.

Tutkimuskysymyksiin vastasivat kuitenkin pääasiassa labin aineistoja käyttävät tutkijat – ei välttämättä labin koko henkilöstö. Suurin osa aineistojen parissa toimivista henkilöistä tarjosi kuitenkin osaamistaan tutkimushankkeisiin ja labin toimintaan liittyi runsaasti henkilöitä, jotka eivät virallisesti kuuluneet henkilöstöön. Heillä oli esimerkiksi aineistoihin ja tekniseen puoleen tarvittavaa asiantuntemusta, joka henkilöstöllä itsellään ei ollut. Näin osaaminen on verkostomaista.

## Uudenlaiset tutkijapalvelut kehitteillä

DAM-hankkeessa kehitetään Kansalliskirjaston tutkijapalveluita käyttäjälähtöisesti, mutta samalla myös kohti library lab-konseptin mukaisia datalähtöisiä esikuviaan. Hankkeen aikana ja lähitulevaisuudessa sen jälkeen on odotettavissa, että digitaalisten aineistojen käyttöön tulee uusia ominaisuuksia, jotka vastaavat käyttäjiensä tarpeeseen. Hankkeessa on esimerkiksi pohdittu, miten koodaustaidottomille tutkijoille voisi tarjota isojen aineistomassojen lataamista. Lisäksi pohditaan aineistojen parempaa löydettävyyttä, haettavautta

ja tulosten esittämisen tapoja. Hankkeen tutkijakärjestä huolimatta myös muut käyttäjät hyötyvät nyt tehtävästä kehittämistyöstä.

Jo ennen DAM-hanketta Kansalliskirjasto on tarjonnut tutkijoille digitaalisia aineistoja tutkimuskäyttöön, myös datan muodossa. Käyttäjälähtöisten palvelujen luominen vaatii työtä ja myös resursseja. Tämä tarkoittaa huolellista suunnittelua siitä, miten nykyinen Kansalliskirjaston osaaminen jaetaan uusien tutkijapalveluiden luomiseen. Lab-malliset tutkijapalvelut haastavat olemassa olevaa osaamista, mutta tyhjästä ei tarvitse lähteä kehitystyössä liikkeelle. Kehitys tuo kuitenkin mukanaan myös uuden osaamisen hallinnan tarvetta. Se edellyttää uuden oppimista ja tiedonvaihtoa sekä kollegoiden kesken että laajemmissa verkostoissa niin kotimaassa kuin ulkomailla. Kansalliskirjaston tarjoamien palveluiden on pysyttävä ajanmukaisina ja vastattava tutkijoiden tarpeeseen tänään ja vielä huomennakin mutta uudesta näkökulmasta.

DAM-hanketta rahoittaa Euroopan aluekehitysrahasto Vipuvoimaa EU:lta 2014–2020. Hanke toteutetaan yhteistyössä Kaakkois-Suomen ammattikorkeakoulun (Xamk) kanssa. Hanke käynnistyi syksyllä 2019 ja jatkuu syyskuun 2021 loppuun asti.

## Lähteet

NÄPÄRÄ, LIISA & LIUKKONEN, ERNO, 2020. Report on the benchmarking interviews in the Digital Open Memory project. Zenodo. <http://doi.org/10.5281/zenodo.4285836>.

**Kirjoittaja:**  
LIISA NÄPÄRÄ  
Kansalliskirjasto  
[liisa.napara@helsinki.fi](mailto:liisa.napara@helsinki.fi)