

# **E-thesis – Helsingin yliopiston verkkojulkaisut**

## **Raportti toiminnasta 1999-2003**

Elokuu 2003  
Laatinut: Jyrki Ilva,  
suunnittelija  
Helsingin yliopiston kirjasto,  
tietokantapalvelut

## SISÄLLYSLUETTELO

<b>1. Helsingin yliopiston verkkojulkaisut.....</b>	<b>3</b>
<b>2. Verkkojulkaisuja jo vuodesta 1999 – E-thesisin lyhyt historiikki.....</b>	<b>3</b>
2.1 Julkaisupankki-hankkeesta E-thesis-palveluun.....	4
2.2. Muutos työnjaossa.....	5
<b>3. Julkaisuprosessi.....</b>	<b>6</b>
3.1. Julkaisua varten tarvittava aineisto .....	7
3.2. Teoksen julkistaminen .....	9
3.3. Julkistamisen jälkeen .....	10
<b>4. Tekniset ratkaisut .....</b>	<b>11</b>
<b>5. Aineisto.....</b>	<b>12</b>
5.1. Väitöskirjat.....	13
5.2. Lisensiaatintyöt.....	14
5.3. Pro gradut.....	15
5.4. Sarjajulkaisut.....	16
<b>6. E-thesis ja käyttäjät .....</b>	<b>17</b>
6.1. E-thesis ja julkisuus .....	18
6.2. E-thesis ja tiedonhaku .....	20
<b>7. Johtopäätöksiä.....</b>	<b>22</b>

## 1. Helsingin yliopiston verkkojulkaisut

E-thesis-palvelun avajaisista tulee syyskuun lopulla 2003 kuluneeksi neljä vuotta. Näiden neljän vuoden aikana palvelun sivuilla on ilmestynyt jo toista tuhatta Helsingin yliopistossa hyväksyttyä opinnäytettä tai erilaisten laitos- tai tiedekuntasarjojen julkaisua. Vaikka E-thesis on jo pitkään toiminut melko matalalla profiililla, ja sen näkyvyydessä yliopiston omillakin sivuilla on toivomisen varaa,<sup>1</sup> Yliopistossa tehtävään tutkimustyöhön käy kuitenkin sen kautta päivittäin tutustumassa satoja sekä yliopiston omasta piiristä että eri puolilta maailmaa tulevia tiedonhakijoita.

E-thesis-palvelun toiminnasta ei ole sen alkuvaiheiden jälkeen kirjoitettu yhtenäistä kuvausta sen omista lähtökohdista käsin,<sup>2</sup> ja tämä raportti yrittää omalta osaltaan paikata tätä aukkoa. Tarkoitus on antaa mahdollisimman yksityiskohtainen ja monipuolinen kuva sekä palvelun tähänastisesta toiminnasta että neljän vuoden kuluessa vakiintuneista prosesseista, jotta yliopiston julkaisutoiminnan tulevaisuutta suunniteltaessa voidaan hyödyntää myös sen puitteissa kertyneitä kokemuksia.

## 2. Verkkojulkaisuja jo vuodesta 1999 – E-thesiksen lyhyt historiikki

Ajatus tieteellisten julkaisujen levittämisestä elektronisessa muodossa verkon kautta alkoi saada tuulta siipiensä 1990-luvun puolivälin tienoilla. Osittain tämä kehitys oli sidoksissa maailmanlaajuisen Internet-verkon kehitykseen ja sen mahdollistamiin utopistisiin visioihin, osittain kyse oli aivan konkreettisista perinteisiin painetun julkaisutoiminnan väyliin kohdistuneista paineista, jotka pakottivat miettimään uudenlaisia ratkaisuja myös akateemisten opinnäytteiden julkaisukäytännöissä.

Suomessa väitöskirjojen määrän nopea lisääntyminen oli johtanut 1990-luvun kuluessa siihen, että monet perinteiset kotimaiset tieteelliset julkaisukanavat olivat tukkeutumassa yhä pahemmin, kun väitöskirjojen ”pakkojulkaiseminen” söi aina vain suuremman osan resursseista. Lisäksi perinteinen väitöskirjavaihto koettiin monilta osin epätydyttäväksi – huolimatta yliopiston kirjastolle luovutettavien ja maailmalle lähetettävien väitöskirjojen määrästä<sup>3</sup> se ei taannut niille merkittävää julkaisuutta. Päinvastoin useimpien väitöskirjojen kohtalona oli päätyä kunkin vastaanottajakirjaston varaston täytteenä muun erittäin vähän kysytyn aineiston joukkoon. Niinpä väitöskirjojen levittäminen elektronisessa muodossa verkon kautta alkoi näyttäytyä yhä luontevampana väitöskirjojen julkaisuprosessia keventävänä ratkaisuna.

---

<sup>1</sup> Helsingin yliopiston kotisivuilla ei ole tällä hetkellä (elokuu 2003) E-thesis-palveluun johtavia linkkejä sen paremmin opiskelijoiden tai henkilökunnan ”etusivuilla” kuin Yliopiston kirjastojen yhteiseltä etusivullakaan, vaan verkkojulkaisuja etsivän täytyy joko seurata huomattavasti mutkikkaampia reittejä (esim. Kirjastot → Verkkoaineistot → E-opinnäytteet → E-thesis) tai sitten oikaista käyttämällä hakukonetta.

<sup>2</sup> Edellinen laajempi kuvaus sisältyy FinELibin rahoituksella keväällä 2000 laadittuihin *Elektronisen julkaisemisen teknisiin ohjeisiin* (<http://www.lib.helsinki.fi/finelib/eleinfo.html>), jossa esitellään Helsingin lisäksi myös Jyväskylän, Oulun ja Tampereen yliopistojen silloisia toimintamalleja. Yliopistojen elektronisen julkaisutoiminnan tilanteesta on vuosien kuluessa tehty useita muitakin yleisemmän tason selvityksiä, mm. Tuija Sonkkilan kaksi kartoitusta (19.11.1999 ja 20.12.2000) ja Désirée Räsänen *Selvitys elektronisesta julkaisemisesta ja sisällöntarjonnan nykytilasta Suomen yliopistoissa ja korkeakouluissa* (syksy 2002).

<sup>3</sup> Painettujen väitöskirjojen vähimmäispainos on 160 kappaletta (+ 6 lakisäateistä vapaakappaletta), joista sata päätyy nimenomaan väitöskirjavaihtoon.

## 2.1 Julkaisupankki-hankkeesta E-thesis-palveluun

Helsingin yliopiston kirjastotoimikunta asetti keväällä 1997 työryhmän valmistelemaan suunnitelmaa yliopiston elektronisen julkaisutoiminnan käynnistämiseksi. Työryhmä sai seuraavan vuoden kuluessa aikaan muistion ja hanke-esityksen yliopiston kirjaston koordinoimaa elektronisen julkaisemisen kokeilua varten. Helsingin yliopiston elektroninen julkaisupankki -hanke sai rahoitusta vuosiksi 1998 ja 1999 ja se käynnistettiin 24.3.1998.

Yliopiston kirjaston yhteistyökumppaneina Julkaisupankki-projektissa olivat mukana tiedekunnat, Yliopistopaino, yliopiston atk-osasto, yliopiston tiedotus ja yliopiston hallinto. Lisäksi yliopiston kirjaston koordinoimana tuolloin jo parin vuoden ajan toiminut kansallinen Elektra-julkaisuprojekti tarjosi runsaasti synergiaetuja ja toimintamalleja, sillä sen puitteissa oli jo aiemmin selvitetty useita tieteellisen aineiston elektroniseen julkaisemiseen liittyviä kysymyksiä, ja myös väitöskirjojen verkkojulkaisemisen kehittäminen oli alkuvaiheessa kuulunut projektin tavoitteisiin.<sup>4</sup>

Helsingin yliopiston elektroninen julkaisupankki -projektin tärkeimpänä tavoitteena oli sellaisen palvelun luominen, jonka avulla Helsingin yliopiston opinnäytteitä ja julkaisusarjoja julkaistaan elektronisessa muodossa ja jaellaan verkon välityksellä. Projektin puitteissa luotiin myös sopimusmekanismeja opinnäytteiden tekijänoikeuksien hallinnoimiseksi, selvitettiin elektronisen julkaisemisen aiheuttamia muutoksia opinnäytteiden aiempiin julkaisu-, levitys-, luettelointi- ja arkistointikäytäntöihin sekä kokeiltiin erilaisia julkaisuteknisiä ratkaisuja.

Toista vuotta kestäneen selvitys- ja suunnittelutyön jälkeen hankkeen tuloksena syntynyt E-thesis-palvelu aloitti virallisesti toimintansa syyskuun lopulla 1999.<sup>5</sup> Toisin kuin monissa muissa yliopistoissa, Helsingissä yliopiston kirjastolla ei ollut aiemmin ollut aktiivista roolia väitöskirjojen julkaisemisessa, vaan eri tiedekunnissa ja eri laitoksilla oli hyvin erilaisia julkaisutraditioita. Tämä mutkisti omalta osaltaan E-thesiksen kaltaisen keskitetyn, koko yliopiston kattavan julkaisupalvelun suunnittelua. Projektin alkuvaiheessa jouduttiinkin tekemään myös aivan perustason kartoitusta yliopiston siihenastisesta julkaisutoiminnasta ja solmimaan yhteyksiä tiedekuntiin ja yliopiston hallintoon.

Julkaisupankki-projektin käytännön tavoitteena oli luoda nopeasti sellainen kevyt organisaatio, joka kykenisi hallitsemaan opinnäytteiden julkaisu- ja jakeluprosessin. Elektroninen julkaiseminen jätettiin ainakin toistaiseksi vapaaehtoiseksi, eikä siitä - osittain jo opinnäytteiden suuren määränkin vuoksi - pyritty kehittämään mitään pakollista tai kaikenkattavaa järjestelmää. Optimistisena oletuksena oli, että elektronisen julkaisemisen suosio lisääntyisi vähitellen itsestään, kun palvelu saavuttaisi riittävän kriittisen massan. Tätä kehitystä pyrittiin tukemaan sillä, että julkaiseminen pidettiin tekijöiden kannalta ilmaisena ja mahdollisimman vaivattomana.

Aluksi E-thesiksen julkaisutoiminta oli opinnäytteiden tekijöiden kannalta täysin ilmaista, sillä yliopiston kirjasto huolehti alkuperäistiedostojen muuntamisesta julkaistavaan muotoon. Yliopiston konsistorin 5.5.1999 tekemän päätöksen mukaisesti väitöskirjan julkistaminen yliopiston palvelimella oikeutti yliopistolle luovutettavien kappaleiden määrän merkittävään alentamiseen: aiemman 160 kirjan sijasta elektronisessa muodossa julkaistusta väitöskirjasta täytyi luovuttaa

---

<sup>4</sup> Tavallaan Elektra-projekti toimi lähtökohtana väitöskirjojen verkkojulkaisemisen aloittamiselle sekä Helsingin että Oulun yliopistoissa – myös Oulun yliopiston kirjasto kehitti omaa SGML-pohjaista julkaisumalliaan aluksi nimenomaan Elektra-projektin yhteydessä. Elektrasta ks. <http://www.lib.helsinki.fi/elektra/> ja Oulun yliopiston elektronisista julkaisuista ks. <http://www.kirjasto oulu.fi/julkaisutoiminta/elektroniset/>

<sup>5</sup> ”E-thesis”-nimi otettiin käyttöön vasta hieman ennen palvelun avajaisia – sitä ennen työnimenä oli ollut mm. ”Elbanco”.

(tiedekunnasta riippuen) enää 30-50 painettua kappaletta. Väitöskirjojen julkaiseminen verkossa yleistyikin hyvin nopeasti – jo pian E-thesiksen virallisten avajaisten jälkeen syksyllä 1999 lähes puolet uusista väitöskirjoista ilmestyi myös elektronisessa muodossa.

Väitöskirjojen lisäksi E-thesis-palvelu pyrki luomaan toimintamallia myös pro gradujen ja lisensiaatintöiden julkaisemista varten. Julkaistavia töitä kerättiin tuoreille, hyvän arvosanan saaneille graduntekijöille lähetetyllä kirjeellä. Lisäksi kukin tiedekunta määritteli julkaistaville pro gradu –töille arvosanarajan - osa tiedekunnista antoi luvan kaikkien opinnäytteiden julkaisemiseen, osa päätyi asettamaan alarajaksi cum lauden tai magnan. Maaliskuun loppuun 2000 mennessä palvelussa julkaistiinkin 139 väitöskirjan lisäksi myös 77 pro gradua ja neljä lisensiaatintyötä. Tiedekunta- ja laitossarjojen saattaminen elektronisen julkaisutoiminnan piiriin ei sen sijaan vielä tässä vaiheessa ennättänyt päästä liikkeelle, ja koko E-thesis-palvelun tilanne muuttui ennen kuin asiaan ennätettiin paneutua.

## 2.2. Muutos työnjaossa

E-thesiksen pitkäjänteisempää kehittämistä vaikeutti se, että yliopiston kirjaston Julkaisupankki-hanketta varten yliopistolta saama rahoitus päättyi vuoden 2000 alussa, ja koko palvelun tulevaisuus oli tämän vuoksi pitkän aikaa hämärän peitossa. Myös Julkaisupankki-hankkeen johtoryhmä lopetti toimintansa tässä epäselvässä tilanteessa. Lopulta E-thesiksen toimintaa muutettiin yliopiston johdon päätöksestä siten, että verkkojulkaisujen tuottaminen siirrettiin huhtikuun 2000 alusta lähtien Yliopistopainolle, jolloin siitä samalla tuli maksullista toimintaa.

Uuden työnjaon mukaan Yliopistopaino huolehti alkuperäisten taitto-, tekstinkäsittely- tai tulostustiedostojen muuntamisesta asiaankuuluviin elektronisiin muotoihin ja se toimitti valmiit julkaisut yliopiston kirjastoon. Yliopiston kirjasto puolestaan ylläpiti E-thesis-palvelinta ja palvelun verkkosivuja sekä huolehti töiden julkaisemista. Tämän maksullisen väylän rinnalle opinnäytteiden tekijöille jätettiin mahdollisuus toimittaa teoksensa myös suoraan yliopiston kirjastoon, edellyttäen että se oli muokattu valmiiksi verkkojulkaisuksi. Yliopiston kirjastolla ei kuitenkaan enää ollut oikeutta muuntaa julkaistavia tiedostoja PDF- tai HTML-muotoon, eli väittelijän oli joko tehtävä muunnos itse tai käytettävä Yliopistopainon palveluita. Lisäksi julkaisun HTML-muotoisen nimiösivun teettäminen muuttui maksulliseksi.<sup>6</sup>

Samalla myös pro gradut ja lisensiaatintyöt jäivät aiemmasta poiketen lähes kokonaan E-thesiksen julkaisutoiminnan ulkopuolelle. Kun opinnäytteiden muuntaminen PDF:ksi yliopiston kirjastossa ei enää ollut mahdollista eikä graduja pyritty enää muutenkaan aktiivisesti hankkimaan julkaistavaksi, koko prosessi jäi pelkästään tekijöiden oman aktiivisuuden varaan. Tilannetta vaikeutti se, ettei läheskään kaikilla graduntekijöillä ollut käytettävissään verkkojulkaisun laatimiseen tarvittavia ohjelmistoja tai edes riittävää teknistä osaamista.

Projektirahoituksen loputtua E-thesiksen käytössä olevat resurssit yliopiston kirjastossa vähenivät dramaattisesti. Palvelua pyörittämässä oli syksyllä 1999 ollut parhaimmillaan peräti viisi koko- tai osa-aikaista työntekijää, joista kaksi työskenteli kokopäiväisesti nimenomaan E-thesiksen parissa. Kevään 2000 jälkeen palvelulla ei enää ollut lainkaan kokopäivätoimista henkilökuntaa, vaan sitä pyritettiin vastuuhenkilöiden muiden töiden ohessa. Myöhemmätkään yritykset rahoituksen

---

<sup>6</sup> Kesällä 2000 käyttöön otettu nimiösivujen tekemiseen tarkoitettu automaattinen ohjelma (ks. <http://ethesis.helsinki.fi/ohjeet/nimiosivunteko.html>) mahdollisti kuitenkin sen, että väittelijät saattoivat tehdä nimiösivun helposti myös omin voimin.

saamiseksi eivät ole johtaneet tulokseen, ja niin koko palvelun toiminta on syksystä 2000 lähtien ollut yhden henkilön oman varsinaisen toimensa ohessa tekemän työn varassa.<sup>7</sup>

Näin jälkikäteen on selvää, että palvelun pyörittämisen yliopiston kirjastossa vaatimaa työmäärää ja siihen liittyviä tehtäviä aliarvioitiin tässä vaiheessa merkittävästi. Vaikka julkaisujen muuntaminen PDF- ja HTML-muotoon siirtyi Yliopistopainon maksulliseksi toiminnaksi, suurin osa palveluun liittyvästä työstä jäi edelleen yliopiston kirjastoon. Resurssipula ja koko palvelun epäselvä tulevaisuus ovat tämän jälkeen johtaneet käytännössä siihen, ettei E-thesiksen toimintaa ole juuri ollut mahdollista kehittää saati laajentaa, ja lähes kaikki voimavarat ovat hupenneet juokseviin asioihin ja päivittäisen toiminnan pyörittämiseen. Muutamista E-thesikseen aiemmin liittyneistä tehtävistä (mm. julkaisujen Dublin Core –muotoiset metatiedot ja gradu-julkaisujen luettelointi tietokantoihin) on myös jouduttu luopumaan, ja yhteydenpito yliopiston sisäisiin sidosryhmiin on jäänyt käytännössä minimiin.

Joitakin pieniä parannuksia E-thesikseen on resurssipulasta huolimatta tehty. Merkittävin kesän 2000 jälkeen tehty uudistushanke on E-thesiksen www-sivujen englanninkielisen navigoinnin käyttöönotto tammikuussa 2001. Lisäksi sivujen navigointia ja ulkoasua on muutenkin kehitetty vähitellen, vaikka suurempiin sivu-uudistuksiin ei ole ollut mahdollisuutta. Samoin E-thesiksen ohjeistusta on päivitetty hiljalleen tarpeen mukaan, ja mukaan on lisätty käyttäjäpalautteissa kaivattuja ohjeita ja toimintoja. Kevästä 2002 lähtien palvelun sivut ovat sisältäneet myös kuukausittaisia tilastotietoja suosituimmista julkaisuista.<sup>8</sup>

### 3. Julkaisuprosessi

E-thesiksen julkaisut sijaitsevat Helsingin yliopiston kirjaston ylläpitämällä palvelimella, jonka verkko-osoite on <http://ethesis.helsinki.fi>. Julkaisut ovat vapaasti kaikkien kiinnostuneiden luettavissa, ja ne voi myös tulostaa omaan käyttöönsä. Julkaisujen lisäksi palvelun verkkosivuilla on myös runsaasti opinnäytteiden tekijöille tarkoitettuja ohjeita, ja sivut ovatkin monille väittelijöille ja graduntekijöille tärkein tietolähde E-thesiksen toiminnasta ja julkaisemisen pelisäännöistä.<sup>9</sup>

Käytännössä merkittävä osa E-thesis-palvelun puitteissa tehtävästä työstä liittyy nimenomaan julkaisijoiden neuvomiseen ja opastamiseen sekä toisaalta erilaisiin aineiston käyttäjiltä tuleviin kysymyksiin vastaamiseen. Huolimatta palvelun sivuilla olevista yksityiskohtaisista ohjeista suurella osalla väittelijöistä ja graduntekijöistä on jotain sellaista kysyttävää, johon he kaipaavat vastausta tai varmennusta. Enimmäkseen näihin kysymyksiin vastaaminen on hyvin rutiininomaista, sillä niissä toistuvat saman asiat, ja useimpien vastausten pääasiallinen tarkoitus on vakuuttaa, että kaikki on todellakin kunnossa ja väitöskirja ilmestyy ajallaan. Välillä vastaan tulee kuitenkin myös huomattavasti hankalampia periaatteellisia kysymyksiä, jotka koettelevat säännösten asettamia rajoja ja vaativat enemmän pohtimista.

---

<sup>7</sup> Vertailun vuoksi voi mainita, että esim. Uppsalan yliopiston verkkojulkaisuista vastaavalla, vuonna 2000 perustetulla *Enheten för digital publicering* –yksiköllä on tällä hetkellä yhteensä kahdeksan työntekijää. Tämän Uppsalan yliopiston kirjastossa toimivan yksikön tehtävät ovat tosin jonkin verran E-thesistä laajemmat. Yksikön esittely löytyy osoitteesta <http://publications.uu.se/epcentre/>

<sup>8</sup> Tilastot löytyvät osoitteesta <http://ethesis.helsinki.fi/arkisto/tilasto/>.

<sup>9</sup> E-thesiksen sivujen lisäksi Yliopistopaino tarjoaa tietoa väitöskirjojen elektronisesta julkaisemisesta sekä omilla www-sivuillaan että painettuna kirjasena julkaistavassa *Väittelijän oppaassa*. Yliopiston viestintäosastolla on oma *Tie tohtoriksi* –sivustonsa, ja tiedekunnilla on myös omia väittelijöille tarkoitettuja ohjemonisteitaan.

### 3.1. Julkaisua varten tarvittava aineisto

E-thesiksen toiminnan peruseriaatteena on se, että opinnäytteiden julkistaminen E-thesis-palvelimella on tekijälle ilmaista. E-thesis-palvelu ei kuitenkaan muunna tiedostoja julkaistavaan muotoon (PDF-tiedostoksi), sillä tämä työ on annettu Yliopistopainon tehtäväksi. Tekijä voi siis valita käyttääkö hän PDF-tiedoston ja julkaisun nimiösivun tekemiseen Yliopistopainon tarjoamia maksullisia palveluita vai tekeekö hän työn itse.<sup>10</sup> Yliopistopainon tarjoama vaihtoehto on tekijän kannalta vaivattomin, mutta käytännössä monilla väittelijöillä on itselläänkin riittävästi verkkojulkaisemisen tekemiseen tarvittavaa osaamista.

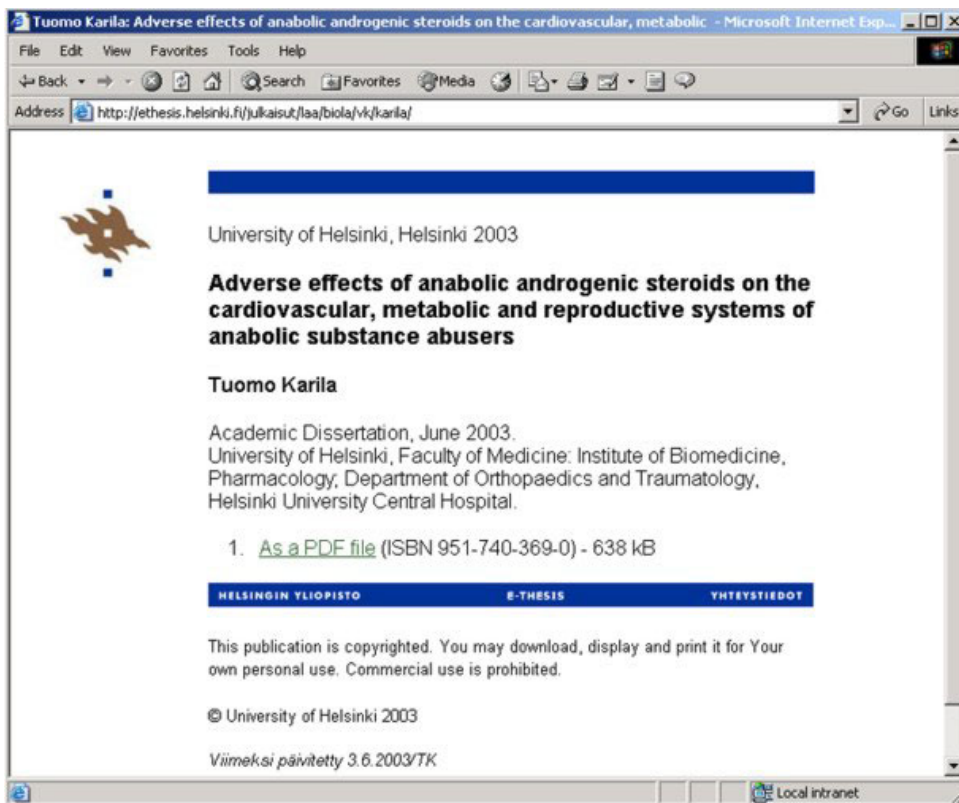
Viime aikoina Yliopistopainon osuus kaikista elektronisista väitöskirjoista on ollut hieman yli puolet. Vuoden 2003 ensimmäisen puoliskon 123 väitöskirjasta 66 tuli julkaistavaksi Yliopistopainon kautta ja 57 suoraan tekijöiltä. Matemaattis-luonnontieteellisessä tiedekunnassa selkeä enemmistö väittelijöistä (25-12) teki teoksensa verkkoversion itse, kun taas esim. lääketieteellisessä (14-26) tai valtiotieteellisessä tiedekunnassa (3-7) Yliopistopainon osuus oli selkeästi suurempi. Huomattava osa myös suoraan tekijöiltä tulleista julkaisuista on kuitenkin painettu Yliopistopainossa, eli tällaisissa tapauksissa kyse on yleensä siitä, että tekijä on jo toimittanut teoksen PDF-muodossa Yliopistopainoon, eikä ole enää halunnut maksaa erikseen verkkoversion tekemisestä. Käytännössä näistä suoraan tekijöiltä tulleista tiedostoista puuttuvat usein kirjanmerkit (*bookmarks*) eikä niitä välttämättä ole optimoitu yhtä hyvin verkkokäyttöön soveltuviksi kuin Yliopistopainon toimittamia tiedostoja, mutta nämä sinänsä harmilliset puutteet eivät kuitenkaan heikennä ratkaisevasti julkaisujen käytettävyyttä.

Silloin kun väitöskirjan elektroninen versio tehdään Yliopistopainossa, Yliopistopaino huolehtii valmiiden tiedostojen toimittamisesta yliopiston kirjastoon. Jos taas väittelijä tekee tiedostot omin voimin (tai teettää ne muualla), hänen täytyy toimittaa tiedostot itse, joko sähköpostitse tai jollain tallennusvälineellä (CD-R, levyke, zip-levyke). Pro gradujen, lisenssiaatintöiden ja sarjajulkaisujen tiedostot on yleensä toimitettu suoraan yliopiston kirjastoon, eli niiden tuottamisessa Yliopistopainolla ei ole ollut roolia.

Varsinaisen julkaistavan PDF-tiedoston lisäksi julkaistavaa opinnäytettä varten tarvitaan erityinen HTML-muotoinen nimiösivu, joka sisältää julkaisun keskeiset tiedot ja linkin itse julkaisuun:

---

<sup>10</sup> Väittelijöillä on mahdollisuus hakea yliopistolta teoksensa julkaisukuluja varten enintään 840 euron suuruista painatustukea, joka voi korvata myös elektronisesta versiosta aiheutuneita kustannuksia.



Yliopistopaino toimittaa elektronisten väitöskirjojen mukana myös valmiin nimiösivun. Jos tekijä tekee elektronisen julkaisun itse, hän voi käyttää nimiösivun laatimiseen E-thesiksen sivuille sijoitettua ohjelmaa, joka luo nimiösivun automaattisesti käyttäjän syöttämien tietojen pohjalta.<sup>11</sup>

Väitöskirjojen nimiösivuilla mainitaan myös PDF-version ISBN-numero, joka tekijän täytyy hankkia Suomen ISBN-keskuksesta. Painettujen, kansallisbibliografiaan luetteloitavien julkaisujen elektronisilla versioilla tulee kullakin olla oma ISBN-numeronsa – näin esim. saman julkaisun PDF- ja HTML-versiolla on eri numerot ja lisäksi vielä painetulla omansa. Painamattomille pro graduille ja liseniaatintöille ei sen sijaan anneta ISBN-numeroa, vaan ne saavat verkkojulkaisuille tarkoitetun URN-tunnuksen. URN:in (Universal Resource Name) tekijä saa luotua automaattisesti yliopiston kirjaston palvelimella sijaitsevalla ohjelmalla.<sup>12</sup>

Tiedostojen lisäksi opinnäytteen tekijän täytyy toimittaa yliopiston kirjastoon myös allekirjoitettu verkkojulkaisusopimus – sopimus solmitaan kirjaston kanssa silloinkin kun väitöskirja toimitetaan Yliopistopainon kautta. Sopimuksessa tekijä antaa kirjastolle luvan teoksensa julkistamiseen verkossa. Tekijänoikeus säilyy tekijällä, eikä sopimus rajoita teoksen mahdollista myöhempää julkaisemista muussa muodossa. Tekijä on toisaalta vastuussa siitä, että hänellä on tarvittavat julkaisuoikeudet kaikkeen materiaaliin (ml. kuvat ja kartat). Sopimuslomakkeen saa tulostettua E-thesiksen verkkosivuilta ja sen voi toimittaa allekirjoitettuna postitse kirjastoon, joko yhtenä tai kahtena kappaleena riippuen siitä haluaako tekijä paluupostissa itselleen oman kappaleen sopimuksesta. Pro graduja varten tarvitaan myös erillinen saatelomake, josta käy ilmi mm. opinnäytteen hyväksymispäivä ja arvosana.

<sup>11</sup> Petri Heliniemen kesällä 2000 laatima ohjelma kaipaisi kipeästi päivittämistä, sillä se ei enää toimi joillakin uusimmilla selainversioilla.

<sup>12</sup> URN-tunnusten luontiohjelma ks. <http://www.lib.helsinki.fi/cgi-bin/urn.pl>



### 3.2. Teoksen julkistaminen

Väitöskirjat julkaistaan E-thesiksen sivuilla normaalisti riiputusajan alkaessa kymmenen päivää ennen väitöstilaisuutta. Osa väittelijöistä julkistaisi teoksensa mieluusti jo aiemmin, mutta tämä selkeä aikataulu on kuitenkin osoittautunut palvelun toiminnan kannalta käytännölliseksi. Tässä aikataulussa on myös se hyvä puoli, että riiputusajan alkamisen jälkeen voidaan ryhtyä kyselemään mahdollisten puuttuvien väitöskirjojen perään, jos niiden tulo vain on ollut etukäteen tiedossa. Ennakkotiedot tulossa olevista elektronisista väitöskirjoista kerätään E-thesiksen www-sivujen uutuuslistaan, sitä mukaa kun väitöskirjan tiedot ilmestyvät yliopiston viestintäosaston ylläpitämään väitöskalenteriin tai kun julkaisusopimus tai julkaistava aineisto saapuvat yliopiston kirjastoon.<sup>13</sup>

Yliopistopainon kautta tulevien töiden kohdalla on toisinaan ollut ongelmia sen vuoksi, ettei yliopiston kirjastolla ole monissa tapauksissa lainkaan suoraa kontaktia väittelijään itseensä. Tämän vuoksi osa julkaisusopimuksista on jäänyt saamatta, ja joissakin tapauksissa myös tiedostot ovat myöhästyneet sen vuoksi, ettei niiden perään ole osattu kysellä.<sup>14</sup> Väitöskirjan virallinen riiputuskappale on edelleen painettu versio, eli aikataulusta lipsuminen ei sinällään ole mikään katastrofi. Gradujen, lisensiaatintöiden ja sarjajulkaisujen kohdalla tällaista tiukkaa aikataulua ei normaalisti ole, ja ne on muiden kiireiden salliessa yleensä julkaistu välittömästi sen jälkeen kun kaikki tarvittava materiaali on tullut perille ja tietojen paikkansapitävyys on saatu tarkistettua.

Jos kaikki on muuten kunnossa, tiedostojen siirtäminen verkkolevyiltä E-thesis-palvelimelle on nopea operaatio – normaalisti yhden väitöskirjan tai gradun julkaiseminen ja linkittäminen E-thesiksen verkkosivuille vie noin 10-15 minuuttia. Tässä vaiheessa tarkistetaan, että nimiösivulla olevat tiedot ovat oikein ja linkit toimivat sekä korjataan mahdolliset virheet. Julkaisemisen jälkeen kaikki julkaisuun liittyvät tiedostot kopioidaan yliopiston kirjaston Piika-verkkopalvelimella sijaitsevaan erilliseen arkistohakemistoon, josta ne todennäköisesti aikanaan siirretään pitkäaikaissäilytystä varten erilliselle arkistopalvelimelle tai digitaalisen aineiston hallintaa varten tarkoitettuun järjestelmään. Arkistoinnin yhteydessä julkaisun tiedot kirjataan myös E-thesiksen omaa sisäistä kirjjanpitoa varten tarkoitettuun Word-tiedostoon.

Palvelimelle siirretyt julkaisut päivitetään välittömästi E-thesiksen www-sivujen julkaisulistoihin, joiden kautta ne ovat lukijoiden löydettävissä. Julkaisut on sivuilla listattu sekä tekijän nimen että tiedekunnan/laitoksen mukaan. Väitöskirjoilla ja graduilla on myös omat uutuuslistansa, minkä lisäksi väitöskirjat ovat riiputusajan esillä palvelun etusivulla. Julkaisujen rekisteröityminen hakukoneisiin vie jonkin verran pidempään – esim. yliopiston oma hakukone indeksoi sivut yleensä viikossa tai parissa, ja Google päivittää indeksiansä noin puolentoista kuukauden välein. Myös väitöskirjojen luettelointi kirjastotietokantoihin tapahtuu vasta myöhemmin. Yleensä tämä viive vaihtelee elektronisen version kohdalla muutamasta päivästä pariin kuukauteen.<sup>15</sup> Kunkin väitöskirjan tietueeseen sisältyy suora hyperlinkki E-thesis-palvelimella sijaitsevaan julkaisuun.

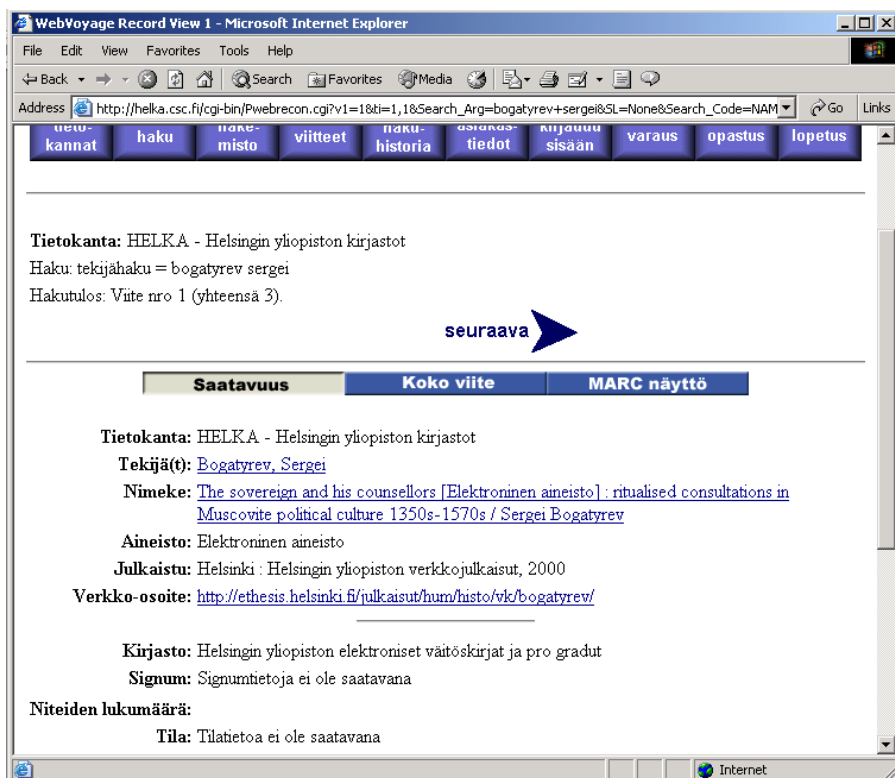
---

<sup>13</sup> Yliopiston väitöskalenterissa mainitaan yleensä jos kyseessä on elektroninen väitöskirja, mutta kaikkien verkkojulkaisujen kohdalla tätä tietoa ei kuitenkaan ole. Lisäksi aivan kaikki väitökset eivät koskaan päädy väitöskalenteriin.

<sup>14</sup> Yliopistopaino on normaalisti toimittanut tiedostot sähköpostitse muutamaa päivää ennen julkaisupäivää, mutta ajoittain osa tiedostoista ei ole tullut ajoissa, ilmeisesti Yliopistopainon sisäisessä tiedonkulussa tapahtuneista katkoksista johtuen. Parissa tapauksessa elektroninen väitöskirja on julkaistu vasta useita kuukausia väitöstilaisuuden jälkeen.

<sup>15</sup> Ajoittaiset viiveet ovat aiheuttaneet myös jonkin verran sekaannuksia, kun tiedekuntakirjastot ovat kiirehtineet luetteloimaan myös väitöskirjojen elektronisia versioita, minkä vuoksi kaikkien julkaisujen luettelointitiedot eivät aina ole keskenään yhteismitallisia.

Vuodesta 2001 lähtien E-thesiksen väitöskirjojen luettelointi on tehty Helsingin yliopiston kirjaston sisältöpalveluissa kansallisbibliografian (Fennican) luetteloinnin yhteydessä. Fennicasta väitöskirjan tietue kopioidaan myös Helka- ja Linda-tietokantoihin. Sisältöpalveluiden kanssa tehty sopimus ei ole kuitenkaan kattanut pro graduja ja lisensointitöitä, sillä niitä ei luetteloida kansallisbibliografiaan. Elektronisten väitöskirjojen tietueet on luotu painettujen kirjojen tietueiden pohjalta, eli kuvailutiedot on yleensä voitu kopioida valmiina. Käytännössä tämä tarkoittaa tosin myös sitä, että elektronisen version luettelointi voidaan yleensä tehdä vasta painetun version luetteloinnin jälkeen.



### 3.3. Julkistamisen jälkeen

Julkaisusopimuksen mukaisesti väitöskirjan tulee olla sisällöltään identtinen tiedekunnalle jätetyn version kanssa, eikä sisältöä voida enää jälkikäteen muuttaa, vaikka se sisältäisi virheellistä tietoa. Väittelijä on kuitenkin halutessaan voinut liittää väitöskirjansa yhteyteen erillisen errata-sivun, jolla virheet on eritelty ja korjattu. Sama sääntö on koskenut myös E-thesiksessä julkaistuja graduja ja lisensointitöitä – myös niiden tulee olla sisällöltään identtisiä tiedekunnassa hyväksytyn version kanssa.<sup>16</sup> Toisaalta julkaisujen ulkoasu saattaa poiketa huomattavasti painetusta versiosta, ja täydellisen samankaltaisuuden saavuttaminen on usein jo teknisistä syistä käytännössä mahdotonta.

E-thesiksen julkaisujen pysyvyys ei ole periaatteista huolimatta muutenkaan aina ollut aivan rikkumatonta: monissa tapauksissa PDF-tiedostoissa on ilmennyt myöhemmin erilaisia

<sup>16</sup> Käytännössä tätä identtisyyttä ei ole mahdollista kontrolloida ainakaan kattavasti, eli koko systeemi perustuu pohjimmiltaan luottamukseen. Tavallaan tässä tekstien identtisyyttä ja muuttamattomuutta korostavassa politiikassa on omat ongelmansa, sillä jos mahdollisia teksteihin sisältyviä virheitä ei jälkeenkäin saa korjata, ne saattavat jatkaa elämäänsä verkossa ja olla satunnaisten, tekstien kontekstia tuntemattomien lukijoiden kannalta hämääviä. Lisäksi on mahdollista argumentoida sen puolesta, että opinnäytteen muokkaaminen vielä hyväksymisen jälkeen paremmin verkkojulkaisuksi soveltuvaan muotoon voisi olla hyödyllinen osa oppimisprosessia. Tämä edellyttäisi kuitenkin laajempia muutoksia nykyisiin opinnäytteen ohjaus- ja hyväksymisprosesseihin.

latautumishäiriöitä tai muita vain tietyissä tilanteissa ilmeneviä teknisiä ongelmia, jotka ovat vaatineet korjaamista. Tällaisissa tapauksissa esim. tiedoston koko ja myös ulkonäkö saattaa muuttua merkittävästi, vaikka sen intellektuaalinen sisältö pysyy ennallaan. Lisäksi muutamassa tapauksessa on jälkeinpäin havaittu, että julkaistu tiedosto on ollut jonkin erehdyksen vuoksi sisällöltään paperiversiosta poikkeava, ja tällöin se on saatettu korjata jälkikäteen alkuperäisversion kaltaiseksi.

E-thesiksen julkaisusopimus ei rajoita opinnäytteiden myöhempää julkaisemista muualla, ja osa väitöskirjoista ilmestyykin vasta myöhemmin joko sinällään tai muokattuna jonkin kaupallisen kustantajan kautta. Muutamassa tapauksessa tämä on johtanut verkkoversion poistamiseen jälkikäteen E-thesiksen palvelimelta. Jos julkaisusopimus irtisanotaan alle vuoden kuluessa väitöstilaisuudesta, väittelijä joutuu luovuttamaan yliopiston kirjastolle väitöskirjavaihtoa varten sata kappaletta väitöskirjastaan. Tällöinkin yliopiston kirjastolla on oikeus säilyttää tiedostot osana vapaakappalekokoelmaansa.

#### 4. Tekniset ratkaisut

Helsingin yliopiston elektronisen julkaisupankin kaikkien julkaisujen oli alunperin tarkoitus olla saatavilla sekä PDF- että HTML-muodossa. Jo projektin varhaisvaiheessa vuonna 1999 alkoi kuitenkin käydä selväksi, että monien opinnäytteiden muuntaminen riittävän laadukkaaksi HTML:ksi oli hyvin työläs prosessi, jopa niin ettei se useissa tapauksissa maksanut vaivaa.<sup>17</sup> Lopullisesti HTML syrjäytyi sen jälkeen kun julkaisujen tuottaminen siirtyi yliopiston kirjastolta Yliopistopainon maksulliseksi toiminnaksi. HTML-version tuottaminen oli niin kallista, ettei väittelijöillä ollut halukkuutta kustannusten maksamiseen. Lisäksi laatuksittavat HTML-version tuottaminen oli sen verran vaativa operaatio, ettei kovin monella väittelijällä tai graduntekijällä ollut itsellään riittävää motivaatiota tai osaamista muunnoksen tekemiseen.<sup>18</sup>

E-thesiksen pääasiallinen julkaisuformaatti on ollut siis PDF – muutamaa yksittäistä poikkeusta lukuunottamatta kaikki julkaisut ovat saatavilla tässä muodossa. Amerikkalaisen Adobe-yhtiön kehittämä PDF onkin tällä hetkellä elektronisen julkaisemisen ja julkaisujen verkkolevityksen de facto -standardi. PDF sopii erinomaisesti paperilla julkaistavaksi tarkoitettujen tekstien sähköiseen levitykseen, sillä PDF-tiedosto säilyttää julkaisun ulkoasun ja mahdollistaa sen tulostamisen paperille alkuperäisen kaltaisena. Lisäksi PDF-tiedoston tuottaminen on automaattista ja useimmissa tapauksissa ongelmatonta (olettaen että tarvittavat ohjelmat ovat käytettävissä), eli se ei vaadi sen paremmin kohtuutonta työ määrää kuin kovin merkittävää teknistä erityisosaamista.

Käytännössä PDF:n tuottamiseen tarvittavien ohjelmistojen saatavuus on kuitenkin ollut jonkinasteinen ongelma toiminnan kannalta. Adoben oma PDF-tiedostojen tuottamiseen tarkoitettu Acrobat-ohjelma on maksullinen, ja edullisista yliopistolisensseistä huolimatta se ei ole kaikissa tiedekunnissa ollut kovin kattavasti sen paremmin väittelijöiden kuin opiskelijoidenkaan saatavilla. Lisäksi esim. yliopiston atk-asemilla Adoben maksullisesta Acrobat Distiller –ohjelmasta on ollut

---

<sup>17</sup> E-thesis-palvelun HTML-julkaisuissa opinnäytteiden ulkoasua pyrittiin muokkaamaan nimenomaan verkkojulkaisemisen ehdoilla toimivaksi. Koska tekstin alkuperäisen ulkoasun säilyttäminen HTML-muodossa ei ollut mahdollista, HTML-versioissa keskityttiin ulkoasun sijasta rakenteen ja sisällön säilyttämiseen. Tekstit jaettiin alalukujen mukaan erillisiin tiedostoihin, jotka sitten sijoitettiin vakioimuotoiseen kehysrakenteeseen ja yhdistettiin toisiinsa hyperlinkeillä. Automaattisten käännösohjelmien tekemä koodi ei ollut sinällään julkaisukelpoista, ja etenkin kuvien, kaavojen, taulukoiden ja erikoismerkkien esittämiseen liittyvien ongelmien vuoksi julkaisukelpoisen HTML-version teko saattoi joissakin tapauksissa viedä useita työpäiviä. Kaikkiaan HTML-versioita ennätettiin tehdä vuosina 1999-2000 hieman yli sata; tämän jälkeen HTML-tiedostoja on julkaistu vain satunnaisesti.

<sup>18</sup> Lisäksi hakukone Google alkoi talvella 2001 indeksoida myös PDF-tiedostoja, jolloin yksi keskeisimmistä HTML-versioiden tarpeellisuuden perusteista poistui päiväjärjestyksestä.

tarjolla ainoastaan vanhoja versioita, joiden käyttöä varten ei ole ollut olemassa kunnollista ohjeistusta.<sup>19</sup> Tässä suhteessa tilanne lienee kuitenkin paranemassa sitä mukaa kun asemien konekantaa ja ohjelmistovalikoimaa uudistetaan. E-thesiksen omilla sivuilla tarjotun ohjeistuksen kehittämistä on kuitenkin haitannut tarjolla olevien ohjelmaversioiden ja teknisten ympäristöjen kirjavuus.

Julkaisujen pitkäaikaissäilytyksen kannalta PDF ei ole optimaalinen ratkaisu, mutta ainakin E-thesiksen nykyresursseilla se on käytännössä ainoa realistinen mahdollisuus. PDF säilyttää sekä tekstin sisällön että ulkoasun, mutta toisaalta sitä on hankala enää kääntää johonkin muuhun, pitkäaikaissäilytyksen kannalta mahdollisesti soveltuvampaan muotoon.<sup>20</sup> E-thesiksen julkaisuista on pyritty keräämään pitkäaikaissäilytystä varten myös alkuperäisiä tekstinkäsittely- tai taittotiedostoja, mutta näitä ei kuitenkaan ole projektin alkuvaiheiden jälkeen saatu talteen kuin vain pienestä osasta julkaisuja. Esim. Yliopistopainon tekemistä julkaisuista ei ole toimitettu muuta kuin PDF-versio – tosin oletettavasti Yliopistopainolla on itsellään alkuperäistiedostot tallennettuna. Alkuperäistiedostojen kohdalla ongelmana on myös se, ettei niiden automaattiseen muuntamiseen johonkin muuhun muotoon ole välttämättä tulevaisuudessakaan tarjolla yhtään sen parempia välineitä, ja erilaisten käyttöjärjestelmien, ohjelmistojen ja tallennusformaattien lukumäärän vuoksi kyseessä olisi väistämättä hyvin raskas ja työläs operaatio. Tieteellisten tekstien kaltaiset kompleksiset dokumentit soveltuvat lisäksi usein huonosti automaattiseen käännökseen, tai ainakaan tällä hetkellä ei ole saatavilla sellaisia ohjelmistoja, jotka pystyisivät tekemään riittävän luotettavia ja laadukkaita käännöksiä erilaisista tiedostoformaateista. E-thesiksen aineiston ”lopulliseksi” tallennusformaatiksi on tarjottu esim. XML:ää, mutta sitäkin varten ei ole olemassa riittävän hyviä muunnostyökaluja.<sup>21</sup>

Monista muista vastaavista palveluista poiketen E-thesiksellä ei ole käytössään mitään tietokantaratkaisua, vaan julkaisut sijaitsevat omissa hakemistoissaan tavallisella linux-palvelimella, ja palvelun www-sivut ovat aivan tavallisia käsin koodattuja HTML-sivuja. Alkuperäisten Julkaisupankki-projektin suunnitelmien mukaan palvelulle oli tarkoitus kehittää myös oma tietokanta, mutta palvelun jäätyä vuonna 2000 ilman rahoitusta tällaiseen kehittämistyöhön ei enää ollut resursseja. HTML-pohjainen sivusto on sinällään toiminut hyvin eikä sen päivittämisessä ole ollut suurempia ongelmia, mutta tämä tekninen ratkaisu rajoittaa kuitenkin vakavasti sivuille koottujen metatietojen käyttömahdollisuuksia. Sivuilla olevia julkaisulistoja ei näet voi järjestellä uudestaan muuten kuin käsityönä, ja hakumahdollisuudet rajoittuvat pelkkään sanahakuun. Julkaisulistojen muuntaminen johonkin tietokantaformaattiin ei sekään todennäköisesti onnistu muuten kuin kirjoittamalla ne alusta lähtien uudestaan, eli tästäkin prosessista saattaa tulla tarpeettoman työläs. Jos tilanteesta haluaa etsiä hyviä puolia, kiinteät HTML-sivut ovat soveltuneet hyvin hakukoneiden robottien haravoitavaksi, eli julkaisujen löydettävyyttä tätä kautta on kuitenkin ollut hyvä.

## 5. Aineisto

E-thesis-palvelussa on elokuuhun 2003 mennessä julkaistu hieman yli tuhat nimekettä. Suurin osa julkaisuista (noin 850) on väitöskirjoja, mutta mukaan mahtuu myös noin 150 gradua ja jonkin

---

<sup>19</sup> Esim. syksyllä 2002 Fabianinkadun atk-aseman Windows NT/Zen-mikroilla tehdyn kokeilun perusteella Acrobat Distiller 4.0 –ohjelman käyttöön saamiseen ei ollut tarjolla riittävää ohjeistusta, vaan se onnistui osaavaltakin käyttäjältä ainoastaan pitkällisen yrityksen ja erehdyksen kautta edenneen prosessin jälkeen.

<sup>20</sup> Myös PDF-formaattia on tosin kehitetty vähitellen, ja se mahdollistaa jo esim. julkaisun metadatan sisällyttämisen osaksi dokumenttia

<sup>21</sup> XML:stä tallennusmuotona ks. Jyrki Ilva: *XML:n mahdollisuuksista Helsingin yliopiston julkaisu-toiminnassa (selvitys, 15.9.2000)*. Teksti on saatavilla PDF-muodossa osoitteessa [http://www.lib.helsinki.fi/eva/xml\\_selvitys.pdf](http://www.lib.helsinki.fi/eva/xml_selvitys.pdf)

verran myös lisensiaatintöitä ja erilaisia sarjajulkaisuja. Seuraavassa taulukossa on eritelty vuosittain ja julkaisun tyyppin mukaan palvelussa ilmestyneet nimekkeet:

Vuosi	Väitöskirjat	Lisensiaatintyöt	Pro gradut	Sarjajulkaisut
1999	78	3	56	-
2000	197	2	28	2
2001	215	3	18	4
2002	229	2	31	4
2003*	123	1	19	17

\* vuoden 2003 lukuihin sisältyvät vain kesäkuun loppuun mennessä ilmestyneet julkaisut.

Julkaisujen vuosittainen kokonaismäärä on siis pysynyt alusta lähtien melko vakaana. Verkkojulkaisuna ilmestyvien väitöskirjojen osuus kaikista väitöskirjoista on tosin hiljalleen kasvanut, mutta muutos ei ole ollut millään lailla dramaattinen. Gradujen julkaiseminen on jäänyt palvelun alkuvaiheiden jälkeen taka-alalle ja niiden määrä on pysynyt potentiaaliseen volyyymiin verrattuna vähäisenä, noin yhdessä prosentissa kaikista vuosittain hyväksyttävistä graduista. Gradujulkaisujen määrässä on tosin tapahtunut hienoista kasvua vuosien 2000-2001 aallonpohjan jälkeen, ja on ilmeistä, että määrä voisi kasvaa hyvinkin nopeasti, jos palvelua markkinoitaisiin graduntekijöille. Sarjajulkaisujen osalta alkuvuonna 2003 tapahtunut määrän lisääntyminen johtuu käytännössä kokonaan yhden uuden, aiempia aktiivisemmän sarjan siirtymisestä verkkojulkaisuksi.

Seuraavassa tarkastellaan yksityiskohtaisemmin eri julkaisutyyppisiä ja niiden erityispiirteitä.

## 5.1. Väitöskirjat

Alkuvuoden 2003 aikana E-thesis-palvelussa julkaistut 123 väitöskirjaa jakautuvat tiedekunnittain seuraavasti:

Tiedekunta	lukumäärä
Lääketieteellinen	40
Matemaattis-luonnontieteellinen	37
Maatalous-metsätieteellinen	12
Valtiotieteellinen	10
Kasvatustieteellinen	9
Humanistinen	8
Eläinlääketieteellinen	4
Teologinen	2
Erillislaitokset	1

Suurin osa elektronisessa muodossa ilmestyneistä väitöskirjoista on siis lääketieteellisestä ja matemaattis-luonnontieteellisestä tiedekunnasta, joissa muutenkin julkaistaan vuosittain eniten väitöskirjoja. Elektronisten julkaisujen jakauma ei silti vastaa suoraan kaikkien julkaistujen väitöskirjojen tiedekuntakohtaista jakaumaa, vaan eri tiedekuntien välillä on huomattavia eroja siinä miten suuri osa väitöksistä päätyy elektroniseen muotoon.

Enemmistö E-thesiksessä julkaistuista väitöskirjoista on tieteellisissä lehdissä aiemmin ilmestyneistä artikkeleista koostuvia ns. artikkeli- tai nippuväitöskirjoja. Artikkeliväitöskirjoista palvelussa julkaistaan tekijänoikeussyistä ainoastaan yhteenveto-osa – vaikka julkaisuluvan saaminen kustantajalta olisi sinällään useimmissa tapauksissa mahdollista, tekijänoikeuksien ja

mahdollisten kustantajien erityisvaatimusten hallinnointi olisi hyvin työlästä.<sup>22</sup> Tästä säännöstä on poikettu vain muutamassa tapauksessa erityisen perustelluista syistä, ja tällöinkin tekijän on täytynyt toimittaa E-thesis-palvelulle kirjallisena kustantajien antamat julkaisuluvat. Käytännössä tämä on tarkoittanut sitä, että monet E-thesiksessä ilman artikkeleita julkaistuista väitöskirjoista ovat hyvinkin suppeita, lyhimmillään jopa vain kymmenkunta sivua. Artikkeliväitöskirjaan tulee kuitenkin sisältyä alkuperäisten artikkelien julkaisutiedot, joiden perusteella lukija voi saada ne käsiinsä.<sup>23</sup>

Artikkeliväitöskirjat ovat erityisen tavallisia etenkin lääketieteessä ja luonnontieteissä, kun taas monografiat ovat edelleen tyypillisiä mm. humanistisissa ja yhteiskuntatieteissä. Tämä jako on monilta osin yhteneväinen myös E-thesiksen julkaisujen tiedekuntajakauman kanssa, sillä lähes kaikki lääketieteen ja luonnontieteiden väitöskirjat ilmestyvät verkossa, kun taas esim. humanistisessa tai teologisessa tiedekunnassa verkkojulkaiseminen on yhä selkeästi harvinaisempaa. Osaltaan tämä liittyy siihen, että suuri osa humanististen ja yhteiskuntatieteiden väitöksistä ilmestyy erilaisten kaupallisten kustantajien kautta, ja tällöin niiden saaminen vapaaseen verkkolevitykseen on vähemmän todennäköistä (joskaan ei täysin mahdotonta). Kovin suurista kaupallisista intresseistä ei kaupallisten kustantajienkaan julkaisemien väitöskirjojen kohdalla yleensä ole kyse, sillä myyntiluvut ovat pääsyyntä tai tenttikirjoja kenties lukuunottamatta harvoin muutamaa sataa kappaletta suurempia. Tähänastisten kokemusten perusteella näyttää siltä että verkossa saatavilla oleva PDF-versio usein pikemminkin lisää väitöskirjaan kohdistuvaa kiinnostusta kuin kilpailee painetun version kanssa. Tästä kertoo sekin, että suuri osa E-thesis-palveluun tulevasta palautteesta on nimenomaan kyselyitä siitä, mistä kirjaa mahdollisesti saa ostaa.<sup>24</sup>

Verkkojulkaisemiseen on monessa yhteydessä liitetty visioita hypertekstimuotoisista väitöskirjoista, jotka hyödyntäisivät audiovisuaalista aineistoa ja kokeilisivat muutenkin uuden tekniikan ja ilmaisukeinojen rajoja. Tähänastisen kokemuksen perusteella näyttää kuitenkin siltä, että sähköisen julkaisemisen avantgarde on jossain aivan muualla kuin akateemisissa väitöskirjoissa. Väitöskirjojen konsepti on jo niiden akateemisen funktion vuoksi korostetun konservatiivinen – kun tärkeintä on toisaalta intellektuaalinen sisältö ja toisaalta perinteisten muodollisten kriteerien noudattaminen, muutokokeiluihin tai riskien ottamiseen ei yleensä riitä motivaatiota. Niinpä verkossa julkaistut teokset ovat pysyneet uskollisina perinteiselle painetun julkaisun formaatille, mitä myös tekniseksi julkaisuformaatiksi vakiintunut PDF-muoto palvelee hyvin.

## 5.2. Lisensiaatintyöt

Lisensiaatintöiden julkaiseminen E-thesiksen kautta on toistaiseksi ollut melko harvinaista, eikä niille vähäisen määrän vuoksi ole laadittu sen paremmin omaa ohjeistusta kuin omaa uutuuksilistaakaan.

---

<sup>22</sup> Suomalaista yliopistoista toistaiseksi ainoastaan Teknillinen korkeakoulu on julkaissut verkkopalvelussaan myös väitöskirjoihin sisältyvät artikkelit.

<sup>23</sup> Tulevaisuudessa julkaisulistasta voi kenties olla OpenURL-tekniikan mahdollistama linkki suoraan artikkeliin niissä tapauksissa joissa lukijalla on käyttöoikeus johonkin artikkelin sisältävään verkkoaineistoon.

<sup>24</sup> Kysymys onkin monissa tapauksissa aiheellinen, sillä Helsingin yliopistossa tilanne on tältä osin hyvin sekava. Osa väitöskirjoista ilmestyy laitossarjoissa joiden yhteystietoja voi joutua metsästäämään kissojen ja koirien kanssa, osa taas on tekijöiden omakustanteita, joita ei välttämättä myydä missään. Aikanaan Julkaisupankki-hankkeen projektisuunnitelmassa kaavailtiin myös tarvepainatusjärjestelmää yliopiston julkaisuja varten, mutta julkaisutoiminnan hajanaisuudesta johtuen tällaisen järjestelyn luominen olisi käytännössä hankalaa.

Lisensiaatintöitä ei ole aiemmin yleensä julkaistu painettuna. Niinpä E-thesiksessä on sovellettu niihin (arvosanarajoja lukuunottamatta) samoja sääntöjä kuin graduihin. Lisensiaatintyö voidaan julkaista verkossa sen jälkeen kun sen hyväksymisestä on päätetty tiedekuntaneuvostossa.

### 5.3. Pro gradut

E-thesiksen julkaisemien gradujen määrä on alkanut hiljalleen lisääntyä vuoden 2001 aallonpohjan jälkeen, vaikka palveluun päätyy edelleen vain häviävän pieni otos kaikista vuosittain hyväksyttävistä graduista. Kovin paljon suurempia määriä ei E-thesiksen nykyresursseilla olisi edes mahdollista käsitellä, ja tämän vuoksi palvelua ei ole sen ilmeisestä potentiaalista huolimatta markkinoitu opiskelijoille muuten kuin E-thesiksen omilla kotisivuilla olevan ohjeistuksen avulla. Kaikki julkaistavaksi tarjotut, teknisesti riittävän hyvätasoiset opinnäytteet on kuitenkin haluttu ottaa mukaan. Julkaiseminen onkin perustunut tyystin opiskelijoiden - tai joissakin tapauksissa heidän opettajiensa - omaan aktiivisuuteen.

Käytännössä suuri osa elektronisesta muodossa julkaistuista graduista on ollut huomattavan tasokkaita – joukossa on runsaasti laudaturin ja eximian graduja. Näyttää siltä, että hyvän arvosanan saaneilla opiskelijoilla on usein muutenkin erityisen suuri motivaatio työnsä julkistamiseen. Vaikuttaa myös siltä, että tieto julkaisumahdollisuudesta leviää usein tuttavapiirin kautta – tietyiltä laitoksilta gradujulkaisuja on huomattavasti enemmän kuin toisilta, ja usein julkaistavat opinnäytteet tulevat pikemminkin parin-kolmen työn ryppäissä kuin yksitellen. Erityisen runsaasti gradujulkaisuja on mm. kasvatustieteellisen tiedekunnan eri laitoksilta sekä humanistisesta tiedekunnasta historian laitokselta. Gradujulkaisujen tieteenalakohtainen profiili poikkeaaakin merkittävästi väitöskirjoista: siinä missä lääketiede ja luonnontieteet dominoivat väitöskirjoissa, suurin osa verkossa julkaistuista graduista on ns. pehmeiltä tieteenaloilta. Väitöskirjoista poiketen enemmistö graduista on myös suomenkielisiä.

Tähänastisten kokemusten perusteella gradujen julkaiseminen verkossa vaikuttaa sinällään mielekkäältä toiminnalta. Jos palvelun markkinointiin, graduntekijöiden opastamiseen ja julkaistavan materiaalin käsittelyyn vain olisi resursseja, julkaistavien töiden määrää olisi todennäköisesti mahdollista kasvattaa hyvinkin nopeasti. Sekä käyttötilastoista että käyttäjiltä tulevista kyselyistä päätellen gradut kiinnostavat lukijoita, ja niitä toivotaan saataville lisää. Samoin tiedekuntakirjastojen päivittäisen kokemuksen mukaan monien oppiaineiden gradut ovat olleet hyvinkin kysytyä aineistoa. Graduihin kohdistuvasta mielenkiinnosta kertoo sekin, että jo nelisenkymmentä Helsingin yliopiston oikeustieteellisen tiedekunnan gradua on ilmestynyt Edita Publishing Oy:n ylläpitämässä Edilex-palvelussa.<sup>25</sup>

Toisaalta on myös ilmeistä, että gradujen asema ja merkitys vaihtelevat huomattavasti eri tiedekunnissa ja eri oppialoilla. Lääketieteellisessä tiedekunnassa varsinaisia pro gradu -töitä ei edes tehdä. Niinpä graduja koskevaa yhtenäistä, koko Helsingin yliopiston kattavaa poliittista tahtoa voi olla huomattavasti vaikeampaa löytää kuin väitöskirjojen kohdalla. Tästä ehkä kertoo sekin, ettei yliopiston keskushallinnon suunnalta ole ainakaan toistaiseksi ilmennyt kovin suurta mielenkiintoa gradujen julkaisemista kohtaan.

Tiedekuntatasolla kiinnostusta tuntuu sen sijaan olevan huomattavasti enemmän. Mm. kasvatustieteellisestä tiedekunnasta on tullut kyselyitä gradujen julkaisemisesta, ja valtiotieteellisellä tiedekunnalla on jo muutaman vuoden ajan ollut oma

---

<sup>25</sup> Edita on tehnyt gradujen julkaisemisesta sopimuksen oikeustieteen opiskelijoiden tiedekuntajärjestön Pykälä ry:n omistaman Paragraaffi Oy:n kanssa.

opinnäytetiivistelmätietokantansa.<sup>26</sup> Vastaavia esimerkkejä löytyy myös muualta Suomesta. Jyväskylän yliopistossa paikallinen yliopiston kirjasto on skannannut lähes kaikki gradut jo vuodesta 1997 lähtien, ja ne ovat nykyään PDF-muotoisina vapaasti luettavissa ja tulostettavissa JYKDOK-tietokannan kautta.<sup>27</sup> Samoin Tampereen yliopiston kirjastolla on oma gradutietokantansa – Tampereella julkaiseminen perustuu tosin Helsingin tavoin vapaaehtoisuuteen.<sup>28</sup>

Gradujen julkaisemisessa on siis runsaasti potentiaalia, mutta poliittista tahtoa on toistaiseksi puuttunut. Tällä hetkellä elektronisessa muodossa julkaistuja graduja ei edes luetteloida kirjastotietokantoihin. Gradujen suuren määrän ja tieteenalakohtaisten erojen vuoksi tuntuisi luontevalta hajauttaa niitä koskevaa toimintaa osittain tiedekuntiin ja kampuksille, vaikka itse aineisto kannattaa toki kerätä yhteen koko yliopiston kattavaan järjestelmään. Nykyinen vapaaehtoisuuteen ja graduntekijöiden omaan aktiivisuuteen perustuva järjestelmä voisi jatkaa entisellään, mutta tiedekunnat voisivat sen lisäksi halutessaan päättää myös omien opinnäytteidensä laajemmasta julkaisemisesta.

Vaikka gradujen nykyistä laajempi julkaiseminen kokotekstimuodossa on tavoiteltavaa, ensisijaisena lyhyemmän tähtäimen tavoitteena voisi ehkä pitää sitä, että edes opinnäytteiden tiivistelmät olisivat saatavilla elektronisessa muodossa. Helsingin yliopiston gradujen ja lisensointitöiden tiivistelmiä on aiemmin julkaistu myös tiedekunnittain koottuna monistesarjana, mutta tämän sarjan ilmestyminen näyttää katkenneen lähes tyystin 90-luvun kuluessa. Tämän sarjan toiminnan jatkaminen tavalla tai toisella elektroniseen muodossa laajentaisi merkittävästi sekä tiivistelmien että itse gradujen käyttömahdollisuuksia, etenkin jos tiivistelmien julkaiseminen voidaan integroida osaksi muita tiedonhakuympäristöjä.

## 5.4. Sarjajulkaisut

E-thesis on hyväksynyt julkaistavaksi kaikki sille tarjotut laitos- ja tiedekuntasarjojen julkaisut. Koska Helsingin yliopisto ei ole ainakaan toistaiseksi määritellyt omaa politiikkaansa verkossa ilmestyvien sarjajulkaisujen suhteen eikä sarjoilla yleensä ole erikseen nimettyjä toimituskuntia, on katsottu kunkin laitoksen/tiedekunnan olevan omalla alallaan riittävä auktoriteetti arvioimaan julkaisujen tasoa. Lisäksi on syytä huomata, että osa julkaisuista on erilaisia oppaita ja raportteja, joiden tasoa ei välttämättä muutenkaan tarvitse arvioida tieteellisin kriteerein.

Tähän mennessä on käynyt ilmi ainakin se, että nykymuotoinen E-thesis soveltuu ensisijaiseksi pysyväksi tarkoitettujen teosten julkaisemiseen. Sen sijaan opinto-oppaat ja esim. jatkuvasti päivitettävät opetusmonisteet sopivat paremmin laitosten tai tiedekuntien omille kotisivuille. Vaikka sama teos voi toki ilmestyä uudestaan uutena painoksena, tällainen olisi usein päivitettävän julkaisun kohdalla kohtuuttoman hankalaa ja aiheuttaisi helposti turhia sekaannuksia, jos lukijat voivat edelleen päätyä teoksen vanhan version pariin. Lisäksi tekstien päivittäminen on oleellisesti helpompaa, jos ne ovat suoraan tekijöiden ulottuvilla – E-thesis-palvelimella sijaitsevia tiedostoja ei ole tällä hetkellä mahdollista päivittää muualta kuin yliopiston kirjastosta käsin.

Sarjajulkaisuista ei ole tehty erillisiä teoskohtaisia julkaisusopimuksia kirjoittajien kanssa, vaan elektronisesta julkaisusta sopiminen on jätetty tekijän ja laitoksen väliseksi asiaksi. Oletettavasti

---

<sup>26</sup> Ks. <http://www.valt.helsinki.fi/kirjasto/tiivisttk.htm>

<sup>27</sup> Jyväskylän opinnäytteet olivat aiemmin luettavissa verkon kautta ainoastaan heikkotasoisina jpg-kuvina, ja tulosteet täytyi tilata Jyväskylän yliopiston kirjastosta. Järjestely oli kuitenkin sekä hankala että taloudellisesti kannattamaton. Julkaisujen nykyinen etusivu löytyy osoitteesta <http://kirjasto.jyu.fi/showtext.php?lang=fin&keyword=kokoelmat-opinnaytteet>

<sup>28</sup> Ks. <http://tutkielmat.uta.fi/>



käytännöt vaihtelevat tässä suhteessa eri laitoksilla ja eri tiedekunnissa. Huomattava osa E-thesiksessä julkaistuista väitöskirjoista on käytännössä ilmestynyt myös sarjajulkaisuna, mutta niiden osalta toimitaan joka tapauksessa väitöskirjoja koskevien sääntöjen mukaisesti, eli verkkojulkaisemisen kannalta sarjalla ei ole tällöin merkitystä.

Sarjajulkaisut ovat tulleet pääasiassa valmiiksi PDF-muotoisina suoraan laitoksilta. Julkaisujen vaihtelevan formaatin vuoksi niiden nimiöisivut ja muu esillepano ovat vaatineet joissakin tapauksissa huomattavaa soveltamista verrattuna väitöskirjoihin. Esim. helmikuussa 2003 Metsävarojen käytön laitoksen julkaisusarjassa ilmestynyt *ECOWOOD studies made at the University of Helsinki* sisältää koosteen kaikista yhden tietyn projektin yhteydessä laitoksella tehdyistä tutkimuksista. Lukuisten artikkelien ja raporttien lisäksi julkaisuun sisältyy myös kolme tietokoneohjelmaa.<sup>29</sup>

Julkaisujen määrän puolesta merkittävin E-thesiksen julkaisusarjoista on Kansantaloustieteen laitoksen *Discussion papers*, joka siirtyi helmikuussa 2003 painetusta muodosta verkkojulkaisuksi. Sarjassa ilmestyy vuosittain arviolta 40-50 nimekettä. Suurin osa muista sarjajulkaisuista on ilmestynyt sekä painettuna että verkkoversiona, eli niiden osalta kyse on rinnakkaisjulkaisemisesta. Yleensä elektroniset julkaisut ovat ilmestyneet vanhan painetun sarjan osana tai jatkona (samalla ISSN-numerolla), joissakin tapauksissa niitä varten on perustettu kokonaan uusi, pelkästään elektronisille julkaisuille varattu sarja. Suomen ISSN-keskuksen kanta on, että molemmat käytännöt ovat hyväksyttäviä.

## 6. E-thesis ja käyttäjät

E-thesis-palvelun käytöstä on kerätty kuukausittaisia palvelimen lokitietoihin perustuvia tilastoja syksystä 2000 lähtien. Yleisenä trendinä on ollut se, että sekä kävijämäärät että julkaisujen käyttö ovat lisääntyneet tasaisesti, ainakin jos vuodenajoista johtuvat vaihtelut suodatetaan tilastoista pois. Yksittäisten julkaisujen keskimääräinen lukijamäärä on pysynyt lähes vakiona (tai korkeintaan lisääntynyt hieman), mutta julkaisujen määrän jatkuva lisääntyminen on kuitenkin kasvattanut kokonaisvolyymia huomattavasti. Siinä missä palvelimelta lokakuussa 2000 ladattiin tiedostoja 4,5 GB:n edestä, määrä oli kaksi vuotta myöhemmin kasvanut jo 29,7 GB:iin. Samana aikana palvelun suomenkielisen pääsivun kuukausittaisten latauskertojen määrä kasvoi noin neljästä tuhannesta yli yhdeksään tuhanteen.

Keväästä 2002 alkaen E-thesiksen www-sivuille on koottu tilastoja myös suosituimpien yksittäisten julkaisujen kävijämääristä.<sup>30</sup> Suurimmat kuukausittaiset kävijämäärät ovat ylittäneet muutamaan sataan, ja keskiarvo lienee tällä hetkellä noin 40-50 kävijän tienoilla.<sup>31</sup> Vaikka graduja on vain noin 15% kaikista E-thesiksen julkaisuista, niiden osuus 50 suosituimmasta julkaisuista on kuitenkin huomattavasti suurempi, lukukausien aikana ajoittain jopa lähes puolet. Osittain tämä epäsuhta johtuu varmaankin siitä, että gradut painottuvat erilaisille tieteenaloille kuin väitöskirjat, ja monet niistä käsittelevät keskivertoväitöskirjaa yleistajuisempia aiheita, joilla on mielenkiintoa myös ns. suuren yleisön keskuudessa. Vastaava ilmiö on nähtävissä myös muiden suomenkielisten

<sup>29</sup> Julkaisu löytyy osoitteesta <http://ethesis.helsinki.fi/julkaisut/maa/mvaro/publications/31/>

<sup>30</sup> Ks. <http://ethesis.helsinki.fi/arkisto/tilasto/>. Tilastot perustuvat palvelimen lokitiedoista kerättyihin julkaisujen nimiöisivujen hakukertoihin, ei siis esim. PDF-tiedostojen lataamiskertoihin. Vaikka nimiöisivulla käyminen ei kerrokaan sitä moniko kävijöistä lataa koneelleen itse julkaisun saati sitten lukee sen, nämä luvut ovat kuitenkin paremmin vertailukelpoisia keskenään kuin PDF-tiedostojen latauskerrat, joiden tilastointiin liittyy teknisiä ongelmia

<sup>31</sup> On tosin ilmeistä, että sivustoa säännöllisesti haravoivat hakurobotit aiheuttavat jo sinällään sen verran liikennettä, että niiden aiheuttaman ”taustakohinan” vuoksi luvuista voi leikata pois arviolta noin parikymmentä kuukausittaista kävijää. Hakurobotit aiheuttavat nykyään noin 20-30% E-thesis-palvelimen kaikesta liikenteestä.

julkaisujen kohdalla: vaikka englanninkielisiä julkaisuja on monin verroin enemmän, suomenkieliset tekstit keräävät keskimäärin selvästi enemmän lukijoita. Se, onko tämä positiivinen vai negatiivinen asia, riippuu suurelta osin näkökulmasta. Toisaalta suomenkielisten julkaisujen suhteettoman suuri suosio kertoo siitä, ettei palvelu ole välttämättä tavoittanut kovin suurta kansainvälistä yleisöä, toisaalta kotimaassa kiinnostusta tuntuu löytyvän myös yliopisto- ja tiedemaailman ulkopuolelta.<sup>32</sup>

## 6.1. E-thesis ja julkisuus

E-thesiksen kaltaisella *Open Access* -palvelulla on väistämättä useita erilaisia yleisöjä. Sen julkaisut ovat yhtä lailla yliopiston oman väen, kansainvälisen tiedeyhteisön kuin aivan satunnaisten verkkokäyttäjien löydettävissä ja luettavissa. Kenelle palvelu oikeastaan on suunnattu? Ovatko kaikki sivuille päätyneet tiedonhakijat yhtä merkittäviä? Tällaisille kysymyksille ei välttämättä ole helppoa löytää aivan yksiselitteisiä vastauksia. Yhtä lailla myös yliopistolla itsellään on useita erilaisia rooleja sekä yhteiskunnallisena vaikuttajana että kansainvälisen tiedeyhteisön osana.

Tilastojen mukaan enemmistö E-thesiksen julkaisujen käyttäjistä tulee edelleen Suomesta, vaikka myös ulkomaisten käyttäjien osuus on vähitellen lisääntynyt. Useat seikat viittaavat siihen, että merkittävä osa kotimaisista kävijöistä tulee katsomaan E-thesis-palvelua kokonaisuutena ja etsimään sieltä kiinnostavia julkaisuja, kun taas ulkomaiset käyttäjät tupsahtavat E-thesis-palvelimella sijaitseviin tiedostoihin usein enemmänkin sattumanvaraisesti, jonkin tekemänsä nettihaun kautta. Tämä lienee jokseenkin väistämätöntä sen vuoksi, että kansallisella tasolla E-thesis on jo sinällään merkittävä palvelu, kun taas harvoja ulkomailta tulevia tiedonhakijoita kiinnostaa, mitä jostakin asiasta on kirjoitettu juuri Helsingin yliopistossa. Toisaalta useimpien tiedonhakijoiden kannalta lienee yhdentekevää minkä nimisellä tai missä maassa sijaitsevalla palvelimella heidän löytämänsä dokumentti sijaitsee, jos se sisältää heidän etsimänsä informaation tai on muuten mielenkiintoinen.

Toukokuussa 2003 E-thesis-palvelimen käyttö jakaantui Internetin maakohtaisten domain-päätteiden mukaan jaoteltuna seuraavasti (järjestysnumero, pyyntöjen määrä, päätte):

```
1: 236563: .fi (Finland)
2: 132075: [unresolved numerical addresses]
3: 71303: .com (Commercial)
4: 58277: .net (Networks)
5: 18259: .edu (USA Higher Education)
6: 13519: .uk (United Kingdom)
7: 7780: .de (Germany)
8: 6966: .it (Italy)
9: 6482: .au (Australia)
10: 6471: .fr (France)
11: 5642: .ca (Canada)
12: 5507: .jp (Japan)
13: 5483: .nl (Netherlands)
14: 4277: .se (Sweden)
15: 3643: .es (Spain)
16: 3541: .be (Belgium)
17: 3497: .br (Brazil)
18: 3122: .pl (Poland)
19: 2229: .tr (Turkey)
20: 2109: .ch (Switzerland)
21: 1953: .tw (Taiwan)
22: 1888: .pt (Portugal)
23: 1799: .org (Non Profit Making Organisations)
24: 1789: .ee (Estonia)
25: 1684: .dk (Denmark)
```

---

<sup>32</sup> Mielenkiintoisena yksityiskohtana voi mainita, että E-thesis-palveluun on tullut useita kyselyitä olisiko jokin yksittäinen englanninkielinen julkaisu saatavilla suomenkielisenä versiona.

Maakohtaiset domain-päätteet eivät aina kerro aivan luotettavasti mistä kukin käyttäjä on todella tullut, mutta ne antavat kuitenkin viitteitä siitä, että E-thesiksen sivuille riittää kävijöitä lähes kaikista maailman maista. Täydellisessä listassa olisi mukana kaikkiaan 128 verkkotunnusta, ja listan loppupäästä löytyy useita sellaisia kehitysmaita joissa Internetin käyttö ei todennäköisesti ole kovin arkipäiväistä huvia. Suurin osa kävijöistä tulee tästä huolimatta toki Suomesta, USA:sta ja muista läntisistä teollisuusmaista. Listassa mainitut pyyntöjen (*request*) absoluuttiset määrät eivät sinällään ole tärkeitä, mutta niitä voi kuitenkin käyttää eri maista tulleen liikenteen suhteellisen osuuden vertailemiseen.

Käyttötilastojen lisäksi myös erilaiset sähköpostitse ja puhelimitse tulevat kyselyt antavat viitteitä aineiston käytöstä ja käyttäjien toiveista. Tavallisimmista palveluun tulevista kysymyksistä saa jonkinlaisen kuvan esim. E-thesiksen *Usein kysytyt kysymykset* –sivulta.<sup>33</sup> Erilaisia teknisiä toimintahäiriöitä koskevien valitusten jälkeen yleisin kyselyjen aihe on se, mistä kirjojen painettuja versioita voi ostaa. Ulkomailta tulee ajoittain kyselyitä jonkin E-thesiksessä julkaisseen tutkijan yhteystiedoista tai jonkin yksittäisen julkaisussa mainitun artikkelin saatavuudesta. Toisinaan kysellään myös olisiko vanhempia vuosikymmenten takaisia julkaisuja saatavilla sähköisessä muodossa.

E-thesiksen julkaisujen markkinointi on toistaiseksi ollut hyvin vähäistä. Suurelta osin se on jäänyt julkaisujen tekijöiden omalle vastuulle – osa tekijöistä on mainostanut julkaisuaan aktiivisesti, osalle taas riittää, että se on ylipäänsä saatavilla. On todennäköistä, että verkkojulkaisemisen arkipäiväistyessä mielenkiinto yksittäisiä verkkojulkaisuja kohtaan on vähentynyt. Siinä missä uuden julkaisun ilmestyminen saattoi vielä viisi vuotta sitten olla uutinen, nyt uutuusjulkaisuja koskevien viestien lähettäminen vaikkapa tieteenalakohtaisille sähköpostilistoille saatettaisiin jossain vaiheessa jo tulkita *spammaamiseksi*. Tiede- ja kulttuuriuutisia kirjoittavilta toimittajilta on tullut jonkin verran kyselyitä ajankohtaisista väitöskirjoista, mutta tällaisissa sinänsä laajalle leviävissä suppeissa jutuissa ei välttämättä mainita lainkaan sitä, että teos on saatavilla myös E-thesiksen kautta. Sinällään E-thesikseen viittaava linkki sopisi ainakin verkkomediassa julkaistuihin juttuihin hyvin, sillä se antaisi uutisen lukijalle mahdollisuuden päästä suoraan alkuperäisen tekstin ääreen.

Väitöstilaisuuksien osalta tiedotuksesta huolehtii yliopiston viestintäosasto, joka ylläpitää yliopiston tapahtumakalenterin osana myös väitöskalenteria. Lisäksi viestintäosasto julkaisee väitöskirjojen tiivistelmiä erillisellä sivulla.<sup>34</sup> Aiemmin tiivistelmät julkaistiin myös Yliopisto-lehdessä, mutta tilanpuutteen vuoksi tästä luovuttiin pari vuotta sitten. Tiivistelmien näkyvyys yliopiston verkkosivuilla ei ole tällä hetkellä kovin hyvä, ja talvella 2002 toiminut yliopiston väitöstietojen keräämistä pohtinut työryhmä esittikin tiivistelmien julkaisemisen siirtämistä E-thesiksen yhteyteen. Työryhmän ehdotus jäi rahoituksen puuttuessa toteutumatta, mutta tiivistelmien julkaisemisen siirtäminen E-thesikselle saattaa olla pitkällä tähtäimelläärkevin ratkaisu, etenkin jos ne saadaan samalla integroitua osaksi suunnitteilla olevia laajempia tiedonhakupäijärjestelmiä. Tiivistelmien julkaiseminen muuttaisi tosin jonkin verran koko E-thesiksen konseptia, sillä tiivistelmät pitäisi saada myös sellaisista väitöskirjoista jotka eivät ilmesty verkkoversiona. Lisäksi tässä yhteydessä pitäisi määritellä se, miten suhtaudutaan erilaisiin tiivistelmiin – julkaistaisiinko verkossa väitöskirjan suomen- tai ruotsinkielinen populaari tiivistelmä, laajempi tieteellinen tiivistelmä vai kenties molemmat?

Aiemmin pelkästään tiedekunta- tai kampuskirjastoon päätyneiden pro gradujen ja liseniaatintöiden osalta verkkojulkaiseminen avaa runsaasti uusia mahdollisuuksia.

<sup>33</sup> Ks. <http://ethesis.helsinki.fi/kysymykset.html>

<sup>34</sup> Kuluvan vuoden väitöstiiivistelmät löytyvät osoitteesta <http://savotta.helsinki.fi/halvi/tiedotus/vaitokset.nsf>

Todennäköisesti mahdollisuus tekstin julkistamiseen voisi joissakin tapauksissa lisätä jopa opiskelijoiden kirjoitusmotivaatiota. Vaikka monien tieteenalojen gradut ovat kysytyä aineistoa kirjastoissa, niiden tavoittama yleisö on useimmiten ollut hyvin rajallinen, ja monet potentiaalisesti hyvinkin kiinnostavat tekstit ovat hautautuneet varaston uumeniin. Graduihin ja lisensiaatintöihin sisältyvän tiedon leviämistä auttaisi jo sekin, jos niiden tiivistelmät olisivat väitöskirjojen tavoin saatavilla verkon kautta.

## 6.2. E-thesis ja tiedonhaku

E-thesiksen nykyiset verkkosivut tarjoavat käyttäjilleen vain melko rajallisia selailumahdollisuuksia. Palvelun sivuilla on listattu julkaisut toisaalta tiedekunnan ja laitoksen mukaan ja toisaalta tekijän nimen mukaan. Lisäksi väitöskirjat ja gradut on listattu myös ilmestymisjärjestyksessä. On kuitenkin todennäköistä, ettei mikään näistä jaoista vastaa kovin hyvin E-thesiksen sivuja käyttävien tiedonhakijoiden tarpeita. Laitosten mukaan järjestetty lista tarjoaa toki yliopiston organisaatiota tuntevalle tiedonhakijalle mahdollisuuden päästä suurin piirtein oikean tieteenalan julkaisujen äärelle, mutta tekijän nimen mukaan aakkostetusta listasta lienee yleisluontoisen selailun lisäksi hyötyä vain silloin kun halutun julkaisun tekijä on jo valmiiksi tiedossa. Ilman dynaamisten sivujen tuottamiseen soveltuvaa tietokantaratkaisua monipuolisempien julkaisulistausten tekeminen olisi kuitenkin hyvin työlästä.<sup>35</sup>

Tilannetta kohentaa oleellisesti se, että julkaisuihin on mahdollista päästä käsiksi useita eri reittejä. Palvelun omien kotisivujen julkaisulistojen lisäksi E-thesiksen julkaisemat väitöskirjat on näet luetteloitu myös Helka-, Fennica- ja Linda-tietokantoihin, joissa kunkin väitöskirjan tietueesta on hyperlinkki suoraan itse julkaisun nimiösivulle. Näiden kotimaisten kirjastotietokantojen lisäksi julkaisut rekisteröityvät moniin verkkoa haravoivien hakurobottien keräämiin tietokantoihin, ja ne ovat löydettävissä myös näiden kautta. Kansainvälisistä hakukoneista ainakin Google ja AlltheWeb indeksoivat myös PDF-tiedostojen sisällön, eli tiedonhakija voi löytää PDF-muotoiset julkaisut myös silloin kuin hakutermi sijaitsee ainoastaan tiedoston sisällä. Kotimaisista hakukoneista on syytä mainita Helsingin yliopiston oma, yliopiston sivuihin kohdistuva hakupalvelu (<http://www-index.helsinki.fi/>), joka on myös käytettävissä E-thesiksen omilla sivuilla palvelun julkaisuihin rajattuna versiona.<sup>36</sup> Lisäksi E-thesiksen julkaisuihin ja E-thesis-palveluun itseensä on epälukuinen määrä linkkejä myös erilaisilla muilla www-sivuilla, eli osa tiedonhakijoista päätyy E-thesiksen sivuille myös niiden kautta.

E-thesis-palvelimen keräämään *referrer*-informaation pohjalta voi tehdä päätelmiä palvelimella olevaan aineistoon päätyneiden hakujen luonteesta. Käytännössä lähes kaikki suosituimmat E-thesikseen päätyneet hakutermit ovat hyvin asiallisia ja spesifejä, usein jonkin tieteenalan erikoissanastoa. Tämä johtuu siitä, että E-thesiksen sisältämät dokumentit ovat suurimmaksi osaksi erikoistuneita ja kapea-alaisia aiheita käsitteleviä tieteellisiä tekstejä, eikä sen sisältämästä aineistosta todennäköisesti löydy sen paremmin Britney Spearsin nakukuvia kuin vastauksia sellaisiin yleisiin kysymyksiin kuin "mikä on elämän tarkoitus". Niinpä voi olettaa, että suurin osa haun tuloksena löytyneistä dokumenteista on ainakin jossain määrin relevantteja. Tällaisten spesifien hakujen osalta E-thesis sijoittuu usein aivan kärkipäähän esim. Googlen hakutuloksissa, kun taas kaikkein yleisimmillä (ja kilpailluimmilla) hakusanoilla sitä ei aina löydy edes kymmeneltä ensimmäiseltä tulossivulta, vaikka palveluun sisältyisi ko. hakusanat sisältävää materiaalia.

<sup>35</sup> Dynaamisten sivujen toteuttaminen ei sinällään välttämättä vaatisi kovin järeää tietokantaratkaisua, sillä se olisi mahdollista esim. XML:n ja XSL:n avulla tai sitten jollain PHP:n kaltaisella tekniikalla. Vanhojen julkaisujen tietojen takautuva syöttäminen järjestelmän vaatimaan muotoon olisi kuitenkin todennäköisesti tehtävä käsityönä.

<sup>36</sup> Ks. <http://ethesis.helsinki.fi/search.html>. Googlasta poiketen tämä hakukone ei kuitenkaan kykene PDF-tiedostojen indeksointiin.

Hakutermin sijainnilla on tosin merkitystä - jos termi on esim. sivun otsikkona (*title*), se saattaa kaikesta huolimatta sijoittua hakutuloksissa yllättävän hyvin.

Seuraavassa listassa on lueteltuna 20 hakupalvelua joiden kautta tuli eniten E-thesiksen aineistoihin kohdistuvia pyyntöjä (*requests*) toukokuun 2003 aikana (tiedot on koottu Webalizer-tilasto-ohjelmalla palvelimen keräämän *referrer*-informaation pohjalta):

17778	<a href="http://www.google.com/">http://www.google.com/</a>
10702	<a href="http://www.google.fi/">http://www.google.fi/</a>
6410	<a href="http://search.yahoo.com/">http://search.yahoo.com/</a>
1455	<a href="http://www.google.de/">http://www.google.de/</a>
1358	Helsingin yliopiston hakukone
1258	<a href="http://www.altavista.com/">http://www.altavista.com/</a>
1057	<a href="http://www.google.ca/">http://www.google.ca/</a>
989	<a href="http://www.scirus.com/">http://www.scirus.com/</a>
977	<a href="http://search.msn.com/">http://search.msn.com/</a>
942	<a href="http://www.google.fr/">http://www.google.fr/</a>
868	<a href="http://www.google.com.au/">http://www.google.com.au/</a>
819	Helka-tietokanta
724	<a href="http://haku.suomi24.fi/">http://haku.suomi24.fi/</a>
721	<a href="http://web.ask.com/">http://web.ask.com/</a>
678	<a href="http://www.google.co.uk/">http://www.google.co.uk/</a>
601	<a href="http://www.google.it/">http://www.google.it/</a>
513	<a href="http://aolsearch.aol.com/">http://aolsearch.aol.com/</a>
409	<a href="http://www.google.com.br/">http://www.google.com.br/</a>
408	<a href="http://www.google.nl/">http://www.google.nl/</a>
363	<a href="http://www.google.co.jp/">http://www.google.co.jp/</a>

Kuten tilasto tuo hyvin esiin, valtaosa kävijöistä on päätenyt E-thesiksen sivuille jonkin verkon yleisen hakukoneen kautta. Googlen ja sen kansallisten versioiden ylivoimainen suosio on erityisen silmiinpistävä.<sup>37</sup> Toisaalta kirjastotietokantojen osuus näyttää suhteessa lähes masentavan pieneltä. Jopa Helsingin yliopiston kirjastojen yhteisen Helka-tietokannan kautta tulleiden hakujen määrä on vain pieni murto-osa Googlen luvuista.<sup>38</sup> Muiden kirjastotietokantojen osalta luvut ovat lähinnä mitättömiä:

19	Linda-tietokanta
11	Fennica-tietokanta

Osittain tämä epäsuhta johtuu tietysti E-thesis-palvelun erityisluonteesta, sillä toisin kuin jossain tietyssä kirjastossa sijaitsevilla painetuilla julkaisuilla, verkon välityksellä vapaasti saatavilla olevilla kokotekstijulkaisuilla on aivan toisella tavalla potentiaalista kansainvälistä mielenkiintoa. Google ja muut laajassa käytössä olevat hakukoneet tavoittavat aivan toisenlaisen yleisön kuin Helkan kaltaiset paikalliset tai kansalliset tietokannat, eikä niitä tämän vuoksi kannata ajatella toistensa kilpailijoina. Lisäksi on mahdollista (tai ainakin toivottavaa), että esim. Helkan kautta E-thesiksen sivuille päätyneet tiedonhakijat ovat ehkä todennäköisemmin onnistuneet tekemässään tiedonhaussa ja löytäneet tarpeidensa kannalta hyödyllisen dokumentin.

Googlen ja muiden Internetin hakukoneiden kautta tehtyjen hakujen ongelma on se, että ne ovat pohjimmiltaan enemmän tai vähemmän selkeästi sanahakuja. Google ei hyödynnä dokumentteihin liitettyä metadataa tai asiasanoja, vaan se indeksoi suoraan niiden sisällön ja arvottaa niitä sen mukaan, miten paljon ja millaisia linkkejä sivuihin on muilta www-sivuilta. Tämän vuoksi tulokset ovat toisinaan melko sattumanvaraisia, ja jokin relevantti dokumentti saattaa jäädä puuttumaan

---

<sup>37</sup> Tämä saattaa tosin johtua epäsuorasti siitä, että Google on indeksoinut myös E-thesiksen sisältämät PDF-tiedostot: tietyistä teknisistä syistä johtuen PDF-tiedostot tuottavat usein yhdellä latauskerralla useampia palvelimen lokitiedostossa näkyviä pyyntöjä, eli latauskertoja ei ole todellisuudessa ole läheskään niin paljon kuin miltä *referrer*-tietojen pohjalta näyttää.

<sup>38</sup> Kannattaa myös huomata, että Helkan etusivulta (<http://helka.linneanet.fi>) on suora linkki E-thesiksen pääsivulle, eli osa Helkan kautta tulleista kävijöistä on todennäköisesti tullut tätä kautta, tekemättä hakua itse tietokannassa.

hakutuloksesta, jos siinä ei mainita oikeaa hakusanaa. Googlesta poiketen kirjastotietokannoissa hyödynnetään myös metadataa, ja niissä on mahdollista tehdä hakuja asiasanojen avulla. Tästä on etua etenkin yleisempien hakutermien kohdalla, olettaen tietysti, että dokumenttien luettelointi ja asiasanoitus on tehty kunnolla, ja tiedonhakija osaa hyödyntää niiden tarjoamia mahdollisuuksia.

E-thesiksen varhaisimpien opinnäytteiden nimiösivuille on liitetty Dublin Core -muotoinen metadata,<sup>39</sup> mutta käytännössä tällä ei ole ollut toistaiseksi suurempaa merkitystä, sillä mikään tietokanta tai hakukone ei ole osannut tai halunnut hyödyntää tätä informaatiota.<sup>40</sup> Tulevissa kirjastojen ylläpitämässä julkaisujärjestelmissä, tietokannoissa ja hakuportaaleissa Dublin Core tulee kuitenkin saamaan keskeisen roolin. Toinen E-thesiksenkin kannalta lupaava ilmiö on Open Access Initiative (OAI), jonka päämääränä on mahdollistaa eri palvelujen sisältämien dokumenttien metadatan kokoaminen tietyn standardoidun protokollan avulla yhteisiin tietokantoihin. Tämä tekee toivon mukaan ennen pitkää mahdolliseksi sellaisten kansallisten ja kansainvälisten yhteistietokantojen kokoamisen, joissa voisi tehdä hakuja esim. kaikista jollain alalla, jossain maassa tai jopa koko maailmassa tehdyistä väitöskirjoista. Ainakin paperilla tällainen järjestelmä kuulostaa akateemisten tiedonhakijoiden kannalta lähes ihanteelliselta, mutta sen maailmanlaajuiseen toteuttamiseen tarvittavan poliittisen tahdon löytymisestä ei valitettavasti välttämättä ole varmoja takeita, etenkin kun asiaan liittyy monenlaisia osin ristiriitaisia kaupallisia ja kansallisia intressejä.

## 7. Johtopäätöksiä

Alkuperäisen Helsingin yliopiston elektroninen julkaisupankki –hankkeen tavoitteiden mukaisesti E-thesis-palvelu on pystynyt vakiinnuttamaan toimintansa Helsingin yliopiston väitöskirjojen sähköisenä julkaisukanavana. Yliopiston kirjaston ja Yliopistopainon välille luotu työnjako on myös omien rajojensa sisällä toiminut pääosin tyydyttävästi. Rahoituksen ja riittävien henkilöresurssien puuttuessa E-thesiksen toiminta on kuitenkin jähmettynyt keväällä 2000 saavutetulle tasolle, eikä sen kehittämiseen tai laajentamiseen ole ollut mahdollisuuksia. Vaikka asenteet elektronista julkaisemista kohtaan ovat todennäköisesti kehittyneet positiiviseen suuntaan (sekä E-thesiksen saama palaute että erilaiset juhlapuheet ja strategiat viittaavat tähän), palvelun tueksi ei kuitenkaan ole toistaiseksi löytynyt konkreettista poliittista tahtoa. Käytännön tasolla tämä on johtanut siihen, että elektronisen julkaisemisen eteneminen Helsingin yliopistossa on ollut monilta osin paljon hitaampaa kuin olisi toivottavaa.

Vaikka myös E-thesiksen tekninen infrastruktuuri kaipaasi monilta osin uudistamista ja ajanmukaistamista, suurempi ongelma on kuitenkin se, ettei palvelulle ole määritelty selkeää mandaattia toimintansa organisoimiseen ja kehittämiseen. Elektroninen julkaiseminen on jäänyt irralleen muusta yliopiston toiminnasta omaksi erilliseksi saarekkeekseen, ja monien siihen liittyvien varsin ilmeistenkin potentiaalisten kehittämisvaihtoehtojen tarjoamat mahdollisuudet ovat jääneet hyödyntämättä. Resurssien puutteen ja koko palvelun epäselvän hallinnollisen aseman vuoksi yhteydenpito yliopiston sisäisten sidosryhmien kesken on jäänyt vähäiseksi, eikä E-thesiksen ympärille ole toistaiseksi voinut syntyä sellaista yliopiston eri instansseja yhdistävää jatkuvaa yhteistyötä tai keskustelua, joka vielä projektia luotaessa näytti mahdolliselta.

---

<sup>39</sup> Dublin Coresta ks. [http://www.lib.helsinki.fi/dublin\\_core/](http://www.lib.helsinki.fi/dublin_core/)

<sup>40</sup> Hakukoneiden näkökulmasta tavallisten verkkosivujen sisältämät metadatakentät ovat käytännössä menettäneet merkityksensä, sillä sivutekijöiden itsensä tuottama metadata on usein ollut harhaanjohtavaa tai suorastaan valheellista - monissa tapauksissa ainoana päämääränä on ollut päästä mahdollisimman korkealle hakutuloksissa riippumatta siitä ovatko sivut ollenkaan relevantteja tiedonhakijan tekemän haun kannalta.

Tällä hetkellä yliopistokirjastojen yhteistyönä on tekeillä useita hyvin kunnianhimoisia kehittämishankkeita, jotka tulevat mullistamaan yliopistolaisten käyttämän tiedonhakuympäristön ja tarjoavat samalla myös uusia mahdollisuuksia elektronisten aineistojen käsittelyyn ja julkaisemiseen. Vuosien 2003 ja 2004 aikana tulevat käyttöön mm. uusi kansallinen (Metalib-ohjelmistoon perustuva) Nelli-tiedonhakuportaali ja etenkin elektronisten vapaakappaleiden käsittelyä varten tarkoitettu digitaalisten objektien hallintajärjestelmä DOMS eli käytännössä todennäköisesti Endeavor Inc:in tuottama ENCompass-ohjelmisto.<sup>41</sup> Olisi toivottavaa, että E-thesis voisi hypätä mahdollisimman varhaisessa vaiheessa tämän aallon mukaan, sillä etenkin ENCompass tarjoaa mielenkiintoisia mahdollisuuksia elektronisen julkaisemisen teknisen infrastruktuurin luomiseen. Samalla palvelulle olisi syytä antaa myös riittävät resurssit toimintansa organisoimiseen. Näitä mahdollisuuksia hyödyntämällä E-thesis voisi integroitua luontevaksi osaksi muita kansallisia – ja ennen pitkää mahdollisesti myös kansalliset rajat ylittäviä – informaatiojärjestelmiä.

---

<sup>41</sup> Sopimusneuvottelut Endeavorin kanssa ovat tämän raportin laatimisaikana vielä kesken.