

# Theseus-käsittelyn erityiskysymyksiä ja vastuunjakoa

Tiina Tolonen, Antti Nyqvist & Minna Marjamaa

# Vastuukysymykset

Vastuu Theseuksessa jakautuu seuraavasti:

- ▶ Kansalliskirjasto vastaa käyttöliittymästä.
- ▶ AMKIT-konsortio ja Theseus-toimisto vastaa ja koordinoi sisältöjä.
- ▶ Käytännössä kukin ammattikorkeakoulu vastaa itse omista sisällöistään omana oikeustoimihenkilönään.
- ▶ Opiskelija vastaa viime kädessä omasta opinnäytetyöstään.

Tämä koskee mm. opinnäytetöiden sisältöjä ja esimerkiksi tutkimusetiikkaa ja henkilötietojen käsittelyä, opinnäytetyöpohjia, saavutettavuutta ja rinnakkaistallennusta.



# Käyttörajattu kokoelma

- ▶ Kahdeksan ammattikorkeakoulua lähtenyt mukaan pilottiin: LAB, Laurea, Oamk, Samk, SeAMK, Savonia, Tamk ja KAMK.
- ▶ Käyttörajattua kokoelmaa koskeva sopimus hyväksytty Arenessa
- ▶ Lisenssityöryhmä on kokoontunut kahdesti ja työstänyt lisenssitekstit valmiiksi. Työryhmä kartoitti muiden julkaisuarkistojen käyttörajattujen kokoelmien lisenssimalleja taustaksi. Muutokset olivat pieniä. Uudet lisenssitekstit liitteenä.
- ▶ Pilottiammattikorkeakouluissa on tarkoitus saada palvelu käyttöön syksyllä
- ▶ Ei vaikuta niihin amkeihin, jotka eivät lähde mukaan pilottiin.
- ▶ Käyttörajatun kokoelman voi ottaa myöhemmin käyttöön.

# Puuttuva haka-attribuutti

- ▶ Käyttörajatun Theseus-kokoelman yhteydessä pyrimme saamaan myös Haka-tunnistuksen sellaiseksi että kirjautumisen jälkeen listauksessa olisivat vain oman amkin kokoelmat. Tästä hyötyvät myös muut kuin käyttörajattun kokoelman pilottiin lähtevät amkit.
- ▶ Jotta Theseus-kirjautuminen voisi johtaa suoraan oman amkin kokoelmaan Theseus-tallennuksessa, täytyy kaikkien amkien ottaa käyttöön Haka-tunnistautumisessa käyttöön attribuutti: [eduPersonScopedAffiliation](#)
- ▶ Tämä vaatii kuitenkin, että kaikki amkit ottavat tämän attribuutin käyttöön. Toimintoa ei voida muuten toteuttaa.
  - ▶ Seuraavista se puuttuu:  
Centria, DIAK, Humak, Karelia, LapinAMK, Laurea, Metropolia, Novia, Polamk, VAMK
- ▶ Vaikuttakaa omien amkienne tietohallintoon tässä asiassa!
- ▶ Attribuutin voi testata myös itse kirjautumalla Haka:n testisivulle <https://rr.funet.fi/attribute-test/> Haka-tunnuksin.

# Huoltokatkolle parempi ajankohta?

- ▶ Theseuksen huoltokatko on maanantaisin klo 13–14 välisenä aikana.
- ▶ Amkeista on tullut palautetta, että se on huono ajankohta, koska se on parhaana töiden käsittelyaikana.
  - ▶ ”Ruuhka-ajan ulkopuolella sillä ei ole niin väliä, mutta kyllä se vähän pipoa kiristää, kun järjestelmä ilmoittaa katkosta kesken käsittelyn ja poolissa odottaa kymmenittäin opinnäytteitä. Joten jos vaan mitenkään mahdollista, toivoisin käyttökatkon ajankohtaan muutosta. Kiitos!”
- ▶ Kansalliskirjastolle ei sovi perjantai, mutta muuten he ovat avoimia ehdotuksille.
- ▶ Olisiko esim. ke 8–9 sopiva aika?



# Wihi-Theseus -putki

- ▶ Yhteinen Wihi-kokous 7.5.2021 Kansalliskirjaston, AMKIT-konsortion ja Wihi-kehittäjien välillä
- ▶ Wihistä ollaan valmistelemassa putki Theseus-tallennukseen
- ▶ Theseuksen metadata on tarkoitus saada Wihin käyttöön opinnäytetyötä tallennettaessa
- ▶ Asiasanoitus yson pohjalta, myös Annifista keskusteltiin
- ▶ Keskeinen kysymys on, kuka käsittelee työt, jos työt vapautetaan Wihistä suoraan Theseukseen
- ▶ Wihin käyttöönotto etenemässä seuraavissa amkeissa: Haaga-Helia, SAMK, TAMK, Xamk, Humak, Karelia, LapinAMK, JAMK, HAMK
- ▶ Tiedotamme lisää, kun työ etenee

# Some-nappien poistot



- ▶ GDPR-vaatimusten kiristyminen:  
Hallinto-oikeus kumosi huhtikuussa kaksi Liikenne- ja viestintävirasto Traficomien tekemää päätöstä, jotka koskivat verkkosivustojen tapoja pyytää käyttäjän suostumusta evästeiden käyttöön. Ratkaisuisaan hallinto-oikeus toteaa, että evästeiden käyttöön annettavan suostumuksen tulee olla yleisen tietosuoja-asetuksen suostumusta koskevien säännösten mukainen. Hallinto-oikeus katsoo, ettei suostumusta voida antaa pätevästi verkkoselainten asetuksilla.
- ▶ Katso: [Hallinto-oikeus: internetselaimen asetusten kautta annettava suostumus evästeiden käytölle ei täytä yleisessä tietosuoja-asetuksessa säädettyjä suostumuksen vaatimuksia - Tietosuojavaltuutetun toimisto](#)
- ▶ Theseukseen ei ole tulossa cookieiden hyväksymiseen liittyvää pop upia
- ▶ Theseuksesta tietoa siirtäviä nappeja: LinkedIn, Twitter, Facebook, RefWorks, Altmetrics
- ▶ Selvitämme, onko mahdollista poistaa varsinaiset some-napit ja jättää RefWorks ja Altmetrics: Altmetrics ei ainakaan nopean tarkistuksen perusteella näytä käyttävän evästeitä, eikä käytä Refworksikaan.
- ▶ Google Analyticsiä on ajateltu jatkettavan käytössä.

# Tiedonlouhinta Theseuksesta

- ▶ Theseuksen tiedonlouhintaan liittyy kaksi periaatetta:
  - ▶ Theseuksen metadata on avointa ja siihen voi kohdistaa tiedonlouhinta.
  - ▶ Sen sijaan Theseuksen lisensseissä ei ole kysytty lupaa tallentajilta kokotekstin käyttöön tiedonlouhinnassa ja lisäksi suurin osa opinnäytetöistä on lisensoitu All rights reserved -periaatteella. Nykyisen tekijänoikeuslain puitteissa kokotekstistä tehtävään tiedonlouhintaan ei näin ollen ole lupaa.
- ▶ Prosessiin liittyy monia tekijänoikeuslain kannalta ongelmallisia vaiheita, kuten louhittavan aineiston kopiointi, muuntelu ja tallentaminen yleisön saataville saatettavaksi tietokannaksi. Tekijänoikeuslain mukaan kaikkiin näihin toimenpiteisiin on saatava tekijänoikeuden haltijan lupa.
- ▶ Katso tarkemmin: [Kohti avointa tiedettä ja tutkimusta - IPRinfo](#)



# Theseuksen tiedonlouhinta jatkuu...

- ▶ Oleellista on se, mihin louhimalla saatuja tietoja käytetään ja miten asia on informoitu rekisteröidyille. Theseuksen tietosuojailmoitus ei sisällä muuta käsittelyä kuin julkaisemisen ja käyttäjähallinnan. Jos kyse on alkuperäisen käyttötarkoituksen kanssa yhteensopimattomasta jatkokäytöstä, siihen pitäisi löytyä tietosuoja-asetuksessa mainittu oikeutus eli asetus rajoittaa jatkokäyttöä.
- ▶ DSM-direktiiviin perustuvassa kansallisessa tekijänoikeuslain uudistuksessa tänä kesänä yksi osa-alue on tiedonlouhinnan mahdollistaminen tutkimusorganisaatioissa ja kulttuuriperintölaitoksissa (Tiedon louhinta artikla 3). Odotamme tekijänoikeuslain muutosta.
- ▶ Theseus-toimisto tai AMKIT-konsortion kannanotto: Emme tällä voi antaa lupaa tiedonlouhintaan.

# PDF-upotukset

- ▶ Videota voi upottaa Theseukseen tallennettaviin PDF-tiedostoihin, mutta siinä ongelmaksi tulee tiedoston koko, joka aiheuttaa ongelmia työn lataamisessa ja muussa käytössä.
- ▶ Maksimikoko 512 MB
- ▶ Theseukseen voi tallentaa myös HTML-tiedostoja, niitä ei ole estetty.
- ▶ HTML-tiedostojen osalta täytyy huomioida myös saavutettavuus eli HTML:n on täytettävä WCAG 2.1-kriteeristön A- ja AA-taso
- ▶ Kriteeristö nähtävillä täällä: <https://papunet.et/saavutettavuus/wcag-21-ohjeet>
- ▶ Suositellaan HTML-tiedostoa PDF-tiedoston sijaan.

# Viiteboksi

- ▶ Kansalliskirjastolla ryhmä, joka selvitti viittausboxin käyttöönottoa vuonna 2019. Luotiin suositus, mutta se ei toiminut kaikissa tapauksissa.
- ▶ Määritykset eri julkaisutyyppien viiteboksin muodostumisesta löytyy täältä <https://www.kiwi.fi/pages/viewpage.action?pagelId=110465015>
- ▶ Viiteboksin korjaamisesta tullut toiveita, keskusteltiin viimeksi KK:n kanssa 12.5.2021.
  - ▶ Tällä hetkellä ei ole aktiivista kehitystä meneillään tämän osalta, mutta nämä tarpeet ovat tiedossa ja pidemmällä aikavälillä tavoite on, että näitä toimintoja saadaan toteutettua. Mutta tällä hetkellä siis viittauksien kanssa on ongelmaa ja sen selvittäminen vaatisi vielä tarkempaa metadataa, jotta metadatan perusteella saatavista viittauksista saataisiin salonkikelpoisia. Toinen, ehkä realistisempi vaihtoehto olisi laskea rimaa ja määritellä yksinkertaisempi ja yleisluontoisempi viittausformaatti. Tämä tie on kuitenkin hidas, koska noita määrityksiä on tehty yhteistyössä usean organisaation kesken ja tällä hetkellä työ ei ole käynnissä sen osalta.
  - ▶ Tavoite siis olisi, että jos viittausten generointitoiminto otetaan käyttöön, se kävisi kaikille arkistoille, jotka noudattavat kansallista metadataformaattia. Siinä sitten vain lisänä pitää määritellä myös tiettyjä pakollisuuksia vielä tiukemmin syöttöpuolella, jotta pystytään muodostamaan oikeaoppinen viite.
- ▶ Eli valitettavasti nopeaa ratkaisua ei ole näköpiirissä, mutta tällaisen toiminnon saaminen on kehitettävien asioiden listalla.
- ▶ Parempi vaihtoehto tällä hetkellä poistaa viiteboksi käytöstä?
- ▶ Metadatasuositus <https://www.kiwi.fi/display/Julkaistuarkistopalvelut/Metadatasuositus+julkaisuarkistojen+tekstiaineistolle>

# Kumpaa suositellaan opiskelijalle: PDF vai PDF/A?

- ▶ Toistaiseksi Theseukseen voi tallentaa työn niin PDF- kuin PDF/A-muodossa.
- ▶ Theseuksessa päätettiin kesällä, että toistaiseksi ei oteta PDF/A-validaattoria Theseukseen, joka estäisi työn tallennuksen, jos tiedosto on väärässä formaatissa.
- ▶ Tämä seuraavista syistä:
  - ▶ PAS-siirto on koko ajan siirtynyt
  - ▶ sitten kun PAS-siirrot alkavat, sillä hetkellä Theseuksessa olevat työt massakonvertoidaan PDF/A-versioon
  - ▶ Theseukseen ei toistaiseksi haluttu tehdä pullonkaulaa tallennukseen eli suurta härdelliä siihen, miksi työtä ei saa tallennettua ollenkaan. Sama koskee myös saavutettavuustseikkausta.
- ▶ PDF/A-tallennukseen on kuitenkin hyvä opetella ohjeistamaan jo nyt, ennen kuin se on aivan pakollista.
- ▶ Tiedossamme ei toistaiseksi ole olemassa muuta tapaa tarkistaa työn formaattia kuin ladata se ulos Theseuksesta ja tarkistamalla työn properties-kohdasta.

# Kuka vastaa työn saavutettavuudesta?

- ▶ Theseuksen saavutettavuuden vastuu jakautuu seuraavasti: Kansalliskirjasto vastaa käyttöliittymän saavutettavuudesta; AMKIT-konsortio ja Theseus-toimisto vastaa ja koordinoi sisältöjä. Käytännössä kukin ammattikorkeakoulu vastaa itse omien sisältöjensä saavutettavuudesta.
- ▶ Opinnäytetöiden osalta vastuu jakautuu seuraavasti:
  - ▶ AMKIT-konsortio vastaa Theseuksen yleisistä ohjeista ja sivustolle tuotettavasta saavutettavuuden check lististä
  - ▶ Ammattikorkeakoulu vastaa opinnäytetyön mallipohjan saavutettavuudesta ja opiskelijan ohjauksesta saavutettavan PDF/A-tiedoston tekemiseksi.
  - ▶ Opiskelija toteuttaa ohjeiden mukaisesti saavutettavan PDF-dokumentin ja tallentaa sen PDF/A-muotoon esim. MS Wordissä ja testaa, onko opinnäytetyö saavutettava.
- ▶ Theseus-töiden tarkistaja ei ole vastuussa opinnäytetyön saavutettavuudesta vaan opiskelija itse sekä opinnäytetyön ohjaaja. Saavutettavuus on pitkä ohjausprosessi; saavutettavuutta ei voida rakentaa enää tarkistusvaiheessa.  
→ Sama koskee myös esim. opinnäytetyön tutkimusetiikkaa.



# Mitä työstä pitää tarkistaa?

- ▶ Koska jokainen amk on itsenäinen oikeustoimihenkilö, kukin taho on itse määritellyt mitä julkaisuvaiheessa tarkistetaan.
- ▶ Tässä kohtaa on hyvä huomioida se, mitkä asiat ovat sellaisia, jotka liittyvät suoraan opinnäytetyön ohjaukseen ja myös se, kuka vastaa esim. saavutettavuuden, tietosuojan tai tutkimusetiikan ohjeistamisesta – työn tallennusvaihe ei voi olla se vaihe, jossa aletaan opetella saavutettavuutta tai tieteellistä kirjoittamista.
- ▶ Ammattikorkeakoulujen rehtorit allekirjoittivat **kesäkuussa 2008 Theseus-verkkokirjastoon liittyvän sopimuksen**. Sen mukaan opiskelijan tehtävä oli päättää opinnäytetyönsä julkistamisesta. Opinnäytetyön ohjaajan tuli perehdyttää opiskelija verkkojulkaisemiseen liittyviin asioihin, kuten julkinen ja salassa pidettävä tieto, tekijänoikeudet, asiasanat, julkistamiseen liittyvä suostumus sekä Creative Commons -lisensointi. Kirjaston tehtäväksi jäi **tarkastaa opinnäytetyön bibliografiset tiedot** ennen julkistamista.
- ▶ Wihi–Theseus -yhteys tuo mahdollisuuden saada metadatta suoraan Pepistä, mm. koulutusala  
→ vähemmän tarkistettavaa Theseus-tarkistajille?

# Töiden poistot

- ▶ Theseuksen toimintamalliin kuuluu, että opiskelija voi pyytää työnsä poistoa ilman perusteluja. Poisto on tehty ammattikorkeakouluissa eri prosessein.
- ▶ Viimeisen vuoden sisällä on tullut enenevässä määrin poistopyyntöjä. Osa pyynnöistä on liittynyt julkisuudessa esillä olleisiin väärinkäyttöksiin: mm. opiskelija on halunnut poistaa kaikki tietonsa internetistä jouduttuaan Vastaamo-tietomurron kohteeksi tai opinnäytetyötä on käytetty huijausviestien tieteellisenä perusteluna.
- ▶ Kerran verkkoon ladattua pdf-tiedostoa voi olla mahdotonta varmuudella poistaa kaikkialta: Theseusta on haravoitu laillisin periaattein Coreen ja OpenAIReen ja toisaalta Googlen indekseissä tiedosto voi säilyä pitkään. Lisäksi opinnäytetöitä löytyy usein Doc Playeristä, jonka palvelimet sijaitsevat Venäjällä sekä Internet Archivesta.
- ▶ Coresta, OpenAIReesta, Googlen indekseistä ja Internet Archivesta voi pyytää poistoa jälkikäteen.



# Theseuksen replikoituminen Finnoihin

- ▶ Haaga-Heliassa havaittu, että Finna näyttää Theseukseen rinnakkaistallennetut julkaisut kirjoina, riippumatta julkaisutyypistä. Tämä vaikuttaa Finnan hakutuloksiin.
- ▶ Asia on käsitelty Finna-webinaarissa toukokuun alussa, jätetty kommentti Finna-foorumille. Finnan käytettävyyssuunnittelijat tutkii asiaa mutta vaatii varmaan seuranta.
- ▶ Kannattaa tarkistaa asiaa omassa Finnassa.



# Theseus-käsittelijöille lisää yhteisöllisyyttä?

- ▶ Postituslista Theseus-opinnäytetyökäsittelijöille:  
[theseus.yhteydenpito@lists.metropolia.fi](mailto:theseus.yhteydenpito@lists.metropolia.fi)
- ▶ Postituslista amk-rinnakkaistallentajille:  
[amk-rinnakkaistallentajat@lists.metropolia.fi](mailto:amk-rinnakkaistallentajat@lists.metropolia.fi)
- ▶ Jos haluat listalle, lähetä postia osoitteeseen:  
[theseus.rukkaset@lists.metropolia.fi](mailto:theseus.rukkaset@lists.metropolia.fi)

säännöllistä  
yhteydenpitoa

Postituslista  
yhdyshenkilöille

**Hyvää  
meininkiä!**

# Kiitos!

YHDESSÄ  
TEHDÄÄN



THESEUS