



Kansalliskirjaston PAS-taival

4.5.2023

Juha Hakala

PAS-juhla, Kansalliskirjasto

Taustaa

- Kansalliskirjastossa pitkäaikaissäilyttäminen on ollut tärkeää vuodesta 1640, mutta e-aineistojen osalta tarve tuli akuutiksi vasta World Wide Webin myötä, 90-luvun puolivälissä
 - Ennen Webiäkin verkossa oli jo paljon aineistoa (<ftp.funet.fi>), mutta julkaisuja vain vähän, koska riittävän helppo jakelukanava puuttui
- Lähtötilanne oli kansalliskirjastoille haasteellinen: osaamista ja resursseja oli niukasti, eivätkä vapaakappalelait kattaneet e-aineistoja
- PAS-toiminta oli yleisestikin vielä lapsenkengissä: valmiita sovelluksia ei ollut, standardit puuttuivat ja perusratkaisuistakin oli kiistaa
 - Emulaatio vai migraatio?

Teoriasta käytäntöön

- Käytännön toimet e-aineistojen pitkäaikaissäilyttämiseksi käynnistyivät Kansalliskirjastossa kolmella tasolla:
 - Vapaakappalelain uudistaminen
 - Verkkoaineiston haravointi
 - PAS-toimintaan ja sen perusteisiin perehtyminen
- Nämä toimintalinjat tukivat ja edellyttivät toisiaan
 - Verkon haravointia ei voi sisällyttää lakiin, ellei tiedä ainakin perusasioita siitä, miten sitä tehdään
 - PAS-palveluja ei voi kehittää itse eikä niitä voi tilata muilta, jos ei tiedä mitä ne ovat tai mitä niiden pitäisi olla

Lainsäädännön kehittämistavoitteita

- Lain katteen muuttaminen julkaisutoiminnan muutoksen edellyttämällä tavalla
 - E-julkaisut niiden tiedostomuodosta riippumatta
 - Julkisesti tarjolla olevat verkkoaineistot sekä radio- ja tv-aineisto
- Lupa aineistojen kopioimiseen ja muokkaamiseen niiden säilymisen takaamiseksi (tuki migraatiolle)
- Mahdollisuus tallentaa aineistot kolmannen osapuolen ylläpitämään (keskitettyyn) PAS-järjestelmään
- Oikeus tarjota e-aineistot valvotusti käyttöön vapaakappalekirjastoista
- Oikeus saada metatiedot aineiston mukana

Lainsäädännön haasteita

- Haravakoneen ajolupa
 - Tekijänoikeusjärjestöjen edustaja edellytti aluksi luvan pyytämistä Web-julkaisijoilta, mutta pyörsi kantansa myöhemmin
- Kopiointi- ja muokkausluvan saanti
 - Todettiin, että tarvitaan tekijänoikeuslain ja –direktiivin muuttamista; jälkimmäinen hoitui niin, että ylikirjastonhoitaja Häkli kuvasi kirjastojen tarpeet Jukka Liedekselle, joka lisäsi ne direktiiviin
- Levytilan tarve
 - Kun päätettiin tallentaa kaikki radio- ja tv-aineisto eikä vain otoksia, tallennustilan tarve nousi niin suureksi, ettei levytilan hankkimiseen ollut varaa -> lain voimaan astuminen siirtyi muutamia vuosia eteenpäin

Kulttuuriaineistolaki

- Vuonna 2007 voimaan astunut laki toimii e-aineistojen osalta edelleen varsin hyvin
 - Monet muut maat eivät ole onnistuneet yhtä hyvin
- Verkon haravoitinta koskevat kohdat kirjoitettiin jo 90-luvulla, ja niistä otettiin mallia Uuden Seelannin vapaakappalelakiin, joka tuli voimaan 2003
- Valmisteluvaiheessa olisi voinut kiinnittää jo aiemmin huomiota tietosuojaan
 - Verkkoarkisto henkilörekisterinä?

Verkkoaineiston haravoinnista

- Kansalliskirjasto haravoi verkkoa jo ennen lain voimaantuloa
 - Pelättiin aineiston katoamista ennen kuin se saadaan talteen
 - Keräämämme aineisto on edelleen tallessa ns. pimeässä arkistossa
- Keruuta varten kehitettiin yhdessä CSC:n Mika Rissasen kanssa haravakone, joka oli kilpailevia ratkaisuja teknisesti parempi
 - Sama tiedosto tallennettiin vain kertaalleen, ja samalle palvelimelle lähetettävien pyyntöjen aikaväli muokkautui automaattisesti vasteaikojen perusteella
 - Web-palvelimien ja haravoitavien tiedostojen huono laatu iso haaste
- Omasta haravasta on luovuttu, ja nykyään käytämme samoja yhteisiä sovelluksia kuin muutkin kansalliskirjastot

PAS-toimintaan perehtyminen

- Kansalliskirjaston kannalta ratkaiseva vaikutus oli EU:n NEDLIB-hankkeella (Networked European Deposit Library, 1998-2000)
 - Pitkäaikaissäilytyksen viitemalli (ISO 14721, suomennos SFS 5972 2009) omaksuttiin PAS-toimintaa ohjaavaksi ajattelumalliksi; myös alan keskeiset käsitteet tulevat tästä standardista
 - Migraatio ja emulaatio menetelminä; kohdeyhteisön merkitys
 - Visio: PAS-järjestelmä on toteuttava Suomessa keskitetysti, koska OKM:n alaisilla muistiorganisaatioilla ei ole resursseja toteuttaa itse järjestelmiä, joka täyttäisivät PAS-vaatimukset
- Kansalliskirjasto ja Kansallisarkisto esittivät yhdessä OKM:lle keskitetyn PAS-palvelun toteuttamista -> KDK-hanke

Lopuksi

- Pitkäaikaissäilytys vaikuttaa kaikkeen toimintaan: ei riitä, että on olemassa erillinen PAS-palvelu, vaan prosessien on oltava PAS-yhteensopivia e-julkaisujen luonnista alkaen
 - Kun väitöskirja on valmis, pitää heti luoda tarkistussumma, jonka avulla voidaan jatkossa varmistaa tiedoston muuttumattomuus
- Tätä prosesseihin liittyvää kehittämistä tarvitaan lisää kaikkien pitkäaikaissäilytettävää aineistoa tuottavien tahojen kanssa
- Lisäksi tarvitsemme vielä sovelluskehitystä: merkityksen säilyttämiseen liittyvät PAS-toiminnot, ohjelmistorajapinnat, ...
- Mutta Suomi on paljon pidemmällä kuin esim. muut Pohjoismaat; niillä ei vielä ole mitään CSC:n PAS-palvelua vastaavaa ratkaisua



www.kansalliskirjasto.fi

Kysymyksiä?