



Kaakkois-Suomen
ammattikorkeakoulu



South-Eastern Finland
University of Applied Sciences

**PLEASE NOTE! THIS IS A PARALLEL PUBLISHED VERSION /
SELF-ARCHIVED VERSION OF THE ORIGINAL ARTICLE**

This is an electronic reprint of the original article.

This version may differ from the original in pagination and typographic detail.

Author(s): Kosonen, Miia

Title: Mikkelissä luodattiin menneisyyden tulevaisuutta

Version: Publisher's PDF

Please cite the original version:

Kosonen, M. (2024). Mikkelissä luodattiin menneisyyden tulevaisuutta. Faili 4, 28 - 31.

HUOM! TÄMÄ ON RINNAKKAISTALLENNE

Rinnakkaistallennettu versio voi erota alkuperäisestä julkaistusta sivunumeroiltaan ja ilmeeltään.

Tekijä(t): Kosonen, Miia

Otsikko: Mikkelissä luodattiin menneisyyden tulevaisuutta

Versio: Publisher's PDF

Käytä viittauksessa alkuperäistä lähdettä:

Kosonen, M. (2024). Mikkelissä luodattiin menneisyyden tulevaisuutta. Faili 4, 28 - 31.

digitaalisten aineistojen määrää. Sitä on tärkeää kasvattaa myös alkujaan digitaalisen aineiston vastaanotolla.

Tekoälyä tulisi hyödyntää sisältöontunnistuksessa ja metatietojen tuottamisessa, mikä tehostaisi aineistojen hakua ja mahdollistaisi tutkijoille aineiston analysoinnin erilaisin keinoin. Helppokäyttöiset, tutkijoiden tarpeisiin mukautetut palvelut ovat avainasemassa. Käyttöliittymiä tulisi kehittää siten, että ne tukevat tehokasta tiedonhakua ja analyysiä moderneilla digitaalisilla työkaluilla. Lisäksi tarvitaan tiivistä yhteistyötä tutkimusyhteisön ja eri arkistotoimijoiden kanssa.

Yksityiset keskusarkistot 25 vuotta -paneelikeskustelun lopputulemana sanoisin, että digitalisaatio ja sen vaikutusten huomiointi ei ole vain mahdollisuus, se on välttämättömyys. Arkistojen tulee mukautua muuttuvaan tutkimukseen ja asiakkaiden odotuksiin. Riskit täytyy tunnistaa, mutta ne eivät saa olla muutoksen este. Digitaaliset palvelut, tekoälyn hyödyntäminen sekä monitieteisen digitaalisen tutkimuksen tukeminen ja mahdollistaminen tekevät arkistoista merkittävän voimavaran. Näin ne voivat vastata yhteiskunnan tiedonkäytön murrokseen.

TAPAHTUMAT

Mikkelissä luodattiin menneisyyden tulevaisuutta



Miia Kosonen
Tutkimuspäällikkö
Digitalia
Xamk

Tutkimuskeskus Digitalian perinteinen kesäkoulu kasvoi tänä vuonna uusiin mittoihin, jotka kaikki alkoivat nelosella. Tiedonhallinnan klusteri Memory Campuksen ja Digitalian yhdessä järjestämä Digiarkistoinnin tulevaisuuskuvat -tapahtuma kokosi syyskuun alussa Xamkin Mikkelin kampukselle ja verkkoyhteyksien päähän peräti 400 osallistujaa yli 40 eri maasta. Osallistujapalautteissa tapahtuma sai yleisarvosanaksi 4.4/5.

Digiarkistoinnin tulevaisuuskuvienväiteemoina olivat jaettu tieto, tekoäly ja klusteriyhteistyö. Kyseessä oli suurin Mikkelissä koskaan järjestetty alan tapahtuma.

"Viimeistään tämän jälkeen Mikkelin tunnetaan sekä Suomessa että maailmalla vahvana digitaalisen tiedonhallinnan ja arkistojen keskittymänä", tiivistä tapahtuman annin Suomen Elinkeinoelämän Keskusarkiston (Elka) arkistonjohtaja Olli Alm.

Tapahtuman pääjärjestäjä oli Memory Campus tiedonhallinnan

kansainväliseksi klusteriksi -hanke. Sitä rahoittaa Etelä-Savon maakuntaliitto Alueiden kestävä kasvun ja elinvoiman tukemisen määrärahasta (AKKE). Hanke on käynnissä vuoden 2024 ajan.

Kuten Xamkin tutkimusjohtaja Noora Talsi muistutti avauspuheen vuorossaan, projektin tärkein tavoite on saavutettu: Memory Campus on nyt rekisteröity eurooppalaiseksi klusteriksi. Avainsana on yhteisö ja sen toiminta. Ohjelmakoordinaattori Kati Saltiolla Xamkista kertoi tarkemmin Memory Campuksesta ja hankkeen tavoitteista. Klusteri vahvistaa tiedonhallinnan alan toimijoiden välistä yhteistyötä ja tukee uuden liiketoiminnan ja uusien innovaatioiden syntymistä mm. yhteisten TKI-hankkeiden kautta. Hankkeen aikana on myös saatu uusia yritysjäseniä mukaan klusteritoimintaan. Näistä kuulemme lisää Memory Campus -yhteisön seuraavassa tapaamisessa 17. joulukuuta.

Yhteistyöverkostoja, jaettava tietoa ja jatkuvaa oppimista

Johtaja Janet Anderson (Highbury Research & Development Ltd.) esitteli tapahtuman osallistujille Euroopan laajuista eArchiving-kokonaisuutta, joka on jo 10 vuoden ajan edistänyt arkistoprojektien harmonisointia Euroopan tasolla. Yksi konsortion jäsenistä on DLM Forum, jonka antimia *Failissakin* on ansiokkaasti tuotu esiin mm. Digitalia-kollega Anssi Jääskeläisen johdolla. Toteutukset liittyvät läheisesti



Kuva 1. Tapahtuman osallistujia Mikpolin edustalla. Kuva: Saara Vauhkonen.

eurooppalaisiin data-avaruuksiin, esimerkiksi kulttuuriperinnön ja median osalta.

E-ARK-projektiperheessä on tuotettu spesifikaatioita, yhteensopivia työkaluja, ohjeistuksia ja käytäntöjä, jotka edistävät eArchivingin missiota aineistojen hankkimisesta, säilyttämisestä ja uudelleenkäytön mahdollistamisesta. ”Interoperability is real, not just a buzzword”, kiteytti Janet tämän periaatteen. Hän kertoi olevansa ensimmäistä kertaa Mikkelissä ja arvosti erityisesti Xamkin hienoa kokoustilaa sekä minuutilleen ajoissa alkanutta tapahtumaa.

Musiikkiarkiston Juha Henriksson esitteli Europeanaa, Time Machinea ja eurooppalaisia data-avaruuksia. Europeanan ja Time Machinen tavoitteena on turvata eurooppalaisia kulttuuriperintöä ja parantaa digitaalisen kulttuuriperinnön saatavuutta. Kiitos Juhan aktiivisen osallistumisen näissä verkostoissa,

molempien vaiheet ovat tulleet *Faillinkin* lukijoille jo tutuiksi.

Allekirjoittaneelta erityiskiitokset Juhalle erinomaisesta syyskauteen sopineesta analogiasta: menneisyyden big data on kuin valtavan laaja sienirihmasto, josta dataa prosessoimalla ja jalostamalla kasvaa maan päälle eli nykyhetken informaation sieniä. Kenties vertausta voisi jatkaa toteamalla, että liian usein ohitamme niin kulttuuriperinnön kuin sienirihmastonkin itsestään selvyytensä, vaikka kyse on maailman suurista ihmeistä. Tiesitkö, että planeettamme maaperän ylimmän 10 cm:n kerroksen sienirihmaston kokonaispituus on yli 450 kvadriljoonaa kilometriä, eli noin puolet galaksimme leveydestä?

Elkan arkistonjohtaja Olli Alm peräänkuulutti puheenvuorossaan organisaatiokohtaisia digitointistrategioita. Mitä digitoinnilla haetaan ja miksi sitä tehdään? Mitkä ovat konkreettiset hyödyt? On demand

-tyylinen digitointi on monelle organisaatiolle järkevää. Elka digitoi av-aineistoja pelastamistarkoituksessa, ja asiakirjoja niiden käytettävyyden ja käyttöasteen nostamiseksi. Digitoinnin, optisen tekstintunnistuksen ja tekoälyn avustaman sisältöanalyysin yhdistelmät vievät löydettävyyttä ja käytettävyyttä eteenpäin, ja ovat osittain jo arkipäivää. Osaaminen olisi digitoinnissa helposti pullonkaula, mutta onneksi osajia koulutetaan paikan päällä Etelä-Savossa.

Osana Digiarkistoinnin tulevaisuus-*kuvat* -tapahtumaa lanseerattiin Joustavat koulutus- ja työelämäpolut tulevaisuuden digitointiosajille (JoDi) -hankkeessa kätilöity digitoinnin osaamisverkosto, josta kerroimmekin jo edellisessä *Failin* numerossa. Sitten viime lukeman verkostoon on liittynyt jo yli 150 jäsentä jakamaan tietoa ja kokemuksia mm. digitointiprojektien läpiviennistä. Tapahtumassa esiteltiin myös JoDi-hankkeessa rakennetut

digitointikoulutukset toteuttajille ja kehittäjille, yhteistyössä Samiedun ja Kansallisarkiston kanssa. Xamkin TKI-asiantuntija ja tiedonhallinnan suunnittelija Kirsi Hännisen alustuksesta osallistujat saivat kokonaiskuvan pitkästä ketjusta digitoinnin koulutushankkeita, joita Memory Campuksessa on viiden vuoden aikana toteutettu. Koulutukset on läpikäynyt kaikkiaan 130 henkilöä. JoDissa tuotettu Digitoinnin kehittäjä -opintojakso tulee Xamkin avoimen AMK:n tarjontaan syksyllä 2025.

Tekoäly KAM-sektorilla

Arkistoalan tekoälykonkretiaa toivat esiin Kansallisarkiston ylitarkastaja Majju Pohjola sekä Xamkin tutkimuspäällikkö Anssi Jääskeläinen omilla puheenvuoroissaan. Xamkin ja Kansallisarkiston yhteisissä TKI-hankkeissa on tuotettu ratkaisuja erityisesti massadigitoidun aineiston

laadunvalvontaan, automaattiseen kuvailuun, kuvien sisällön tunnistamiseen ja tiedon tiivistämiseen valtavista aineistomassoista, mutta kehittämisen tuloksia voidaan hyödyntää muuallakin kuin massadigitoinnissa. Ratkaisuja on koottu avoimesti jakoon Arkkiivi.fi-sivustolle.

Esimerkiksi dokumentin rakenteen tunnistamisessa ja automaattisen metadatan lisäämisessä opetusaineistona oli perintöveroon liittyviä asiakirjoja. Asiasanoittamisessa käytettiin Kansalliskirjaston kehittämää avoimen lähdekoodin Annify-työkalua. Pohjola listasi tulosten ohella tärkeimmät haasteet: komponentit eivät ole valmiita tuotteita, eri järjestelmiä on runsaasti, ja tästä aiheutuu paljon keskinäisriippuvuusia. Fokus tulee siirtää analytiikkainfraan, missä komponenttien käyttäminen on tarvittaessa mahdollista erillään, ja niitä voi jatkokehittää itsenäisesti.

Jääskeläinen nosti esiin Xamkissa rakennetun Memory Lab -infran ja erityisesti supertietokone Hipun. Sen kapasiteetti vastaa noin 20 000 tavallisen läppärin suorituskykyä ja soveltuu vaativampaankin suurteholaskentaan. Memory Lab digitaalisuuden kokeilu- ja innovaatiokulttuurin ajurina Etelä-Savossa (AIKO) -hankkeen yrityspilotit vietiin päätökseen viime keväänä ja tulokset on julkaistu Memory Labin sivustolla. Tekoälykehitystä jatkaa alkusyksystä käynnistynyt Etelä-Savon koulutettu tekoäly Memory Labissa eli ETKOT-hanke, jossa testataan mm. esiopettajien avointen kielimallien jatkokouluttamista suurilla arkistojen ja kirjastoaineistoilla. Työ kytkeytyy datatalouden toimintamalleja ja data-avaruuksia edistävään DAME-hankkeeseen, joka käynnistyi kesällä. Rahoittajalla on parhaillaan käsitellyssä myös laaja tekoälyyn liittyvä koulutushankkeemme, josta lisätietoja allekirjoittaneelta.



Kuva 2. Datan kulku tiedonlouhintaa käyttävässä tutkimuksessa. Kuva: Johanna Lilja.

Sodan ja rauhan keskus Muistin toimitusjohtaja Olli-Pekka Leskinen nosti museosektorilta tekoälyn mahdollisuuksina esiin tietokantojen tehokkaamman rakentamisen ja hyödyntämisen, uudet näyttelysisällöt ja sisällöntuotannon tuen, sekä vuorovaikutteisuuden lisäämisen näyttelyihin. Toisin sanoen museo-kävijä kysyy kysymyksiä ja näyttelyn oma chatbot vastaa. Aidot objektit – Muistin tapauksessa vaikkapa Molotov-Ribbentrop-sopimuksen kopio – ovat tietenkin tärkeitä, mutta niitä ei välttämättä tuoda esille näyttelyissä: tärkeintä on tuoda esiin se tieto, mitä kyseinen objekti edustaa.

Mindhive Oy:n toimitusjohtaja Ville Venäläisen alustuksessa näkökulma oli siinä, miten herättää arkistot eloon tekoälyä hyödyntäen. Villeltä onkin lähtöisin myös tämän artikkelin pääotsikko menneisyyden tulevaisuudesta. Generatiivinen tekoäly ei ole vielä edennyt taaperosta tohtoriksi, mutta se tulee tekemään sen nopeammin kuin allekirjoittanut kirjoittaa luomuälyllä 50 julkaisua ammattilehtiin, eli jo 1–5 vuoden sisällä. Puheenvuoron aikana nähtiin useampi demoesimerkki. Varsinkin tekoälystä puhuneen video-Mannerheimin piipahdus tämän alan tapahtumassa odotetusti jakoi osallistujien mielipiteitä: historiallinen henkilö oli laitettu sanomaan asioita, joita hän ei todellisuudessa ole koskaan voinut sanoa. Tuleeko tiedon alkuperän, autenttisuuden, eheyden ja luotettavuuden rajoja venyttää vain siksi, että se on mahdollista? Videopuheenvuoro oli joka tapauksessa erinomainen muistutus siitä, että tarvitsemme Digitalian kesäkouluihin entistä enemmän digitiikkaa käsitteleviä alustuksia.

Kulttuuriperintö tiedonlouhinnassa ja tekoälykehittämisessä

Tämä väliotsikko on tiivistetty Kansalliskirjaston palvelujohtaja Johanna Liljan esityksestä. Lilja tarjoili hyvän yhteenvedon datan kulun vaiheista tutkimuksen tiedonlouhinnassa (kuva 2) ja pureutui puheenvuorossaan myös edellä peräänkukulutettuun digiajan etiikkaan. Eettiselle keskustelulle on valtava tarve. Tekoälykehitys ja kielimallit ovat lisänneet huolta datan väärinkäytöstä mm. kustantajien keskuudessa. Generatiivista tekoälyä soveltavien globaalien yritysättien palveluissa haasteena on datan läpinäkyvyyden puute. Hannu Toivosen tuoreen kirjan (2023) mukaan tekoäly on lähinnä kielikuva, jolla ohjelmistojen toiminta rinnastetaan inhimilliseen älykkääseen toimintaan.

Palaute ja jatko

Palautteissa tapahtuma sai kiitosta monipuolisesta ja ajatuksia herättävästä sisällöstä, ”josta on hyvä ammentaa arkiston arkeen”. Useita mainintoja saivat myös aikataulujen pitävyys ja muutenkin erinomaiset järjestelyt. Kruununa ohjelmalle toimivat vierailut paikallisissa arkistoissa ja Muistissa, risteily Aino-laivalla Astuvansalmen kalliomaalausten äärelle, sekä loistava illallinen Tertin kartanossa.

Ison kansainvälisen tapahtuman järjestäminen oli myös Digitalialle tärkeä oppimiskokemus. Työmäärältään toteutus oli juuri sellainen kuin yliopistovuosien konferensseja järjestäneet kollegat ovat tienneet kertoa. Huomioitavia yksityiskoh-
tia on valtavasti ja yksinoikeutta

vain yhteen tehtävään keskittymiseen ei ole kenelläkään: ilmassa on jatkuvasti monta palloa yhtä aikaa. Onneksi digitalialaiset saivat erinomaista taustatukea ja työpanosta Juvenia-tutkimusyksikön Saara Vauhkoselta, joka on kokenut tapahtumajärjestäjä ja toimi myös tilaisuuden juontajana.

Lopuksi, Xamkin Digitalia-tiimi ja Memory Campus -klusterihanke haluavat kiittää lämpimästi kaikkia tapahtumaa toteuttamassa olleita yhteistyökumppaneitamme: Sodan ja rauhan keskus Muisti, Kansallisarkisto, Kansalliskirjasto, Elka ja Disec Oy. Kiitämme myös kaikkia puheenvuorojen ja alustusten pitäjiä ml. Mindhive Oy ja Musiikkiarkisto, sekä tietenkin tapahtuman osallistujia.

Digitalia täyttää pian pyöreitä vuosia, sillä 1.5.2025 tulee kuluneeksi 10 vuotta Digitalian perustamishankkeen käynnistymisestä. Mitä varmimmin luvassa on juhlaseminaari, tutulla digitaalisen tiedon kesäkouluun konseptilla täydennettynä. Nähdään siis tulevissa tapahtumissamme!

Tapahtuman tallenteet ja kuvat löydät Memory Campuksen sivuilta: <https://memorycampus.fi/digital-archiving-futures-2024-esitykset/>