

Digitointi ja sähköisen aineiston vastaanotto yksityisarkistosektorilla

6.9.2024

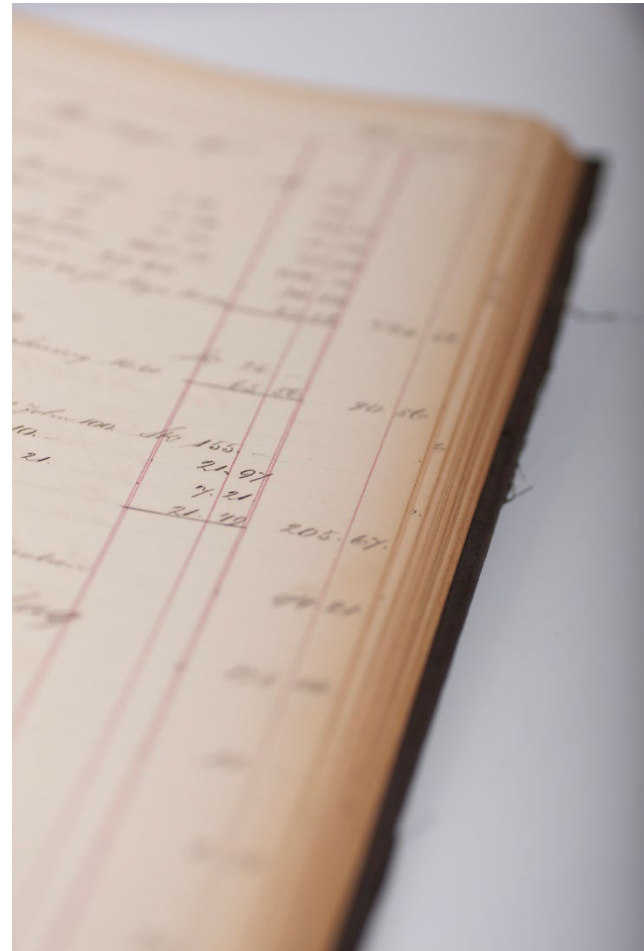
Olli Alm



SUOMEN
ELINKEINOELÄMÄN
KESKUSARKISTO

Sisältö

- Lyhyesti Elkasta
- Missä mennään, mitä on tehty
- Digitointi
- Syntysähköisten arkistointi: miksi tämä ei etene?
- Järjestämisen ja luetteloinnin tulevaisuus



Suomen Elinkeinoelämän Keskusarkisto

- Valtakunnallinen yritysten ja elinkeinoelämän järjestöjen asiakirjojen säilytys- ja tutkimuspaikka
- Aloitti vuonna 1981
- Henkilökuntaa tällä hetkellä 13
- Aineistoa noin 35 hkm
- Elkan yhteydessä toimii Designarkisto, joka on suomalaisen muotoilun ja käyttäjälähtöisen suunnittelun erikoisarkisto
- Elkan palveluita yrityksille mm. analogiset ja digitaaliset säilytyspalvelut, niihin liittyvä tietopalvelu sekä digitointi
- Tutkijoille tutkijasali- ja kaukolaina-, selvitys-, digitointi- ja verkkopalvelut.



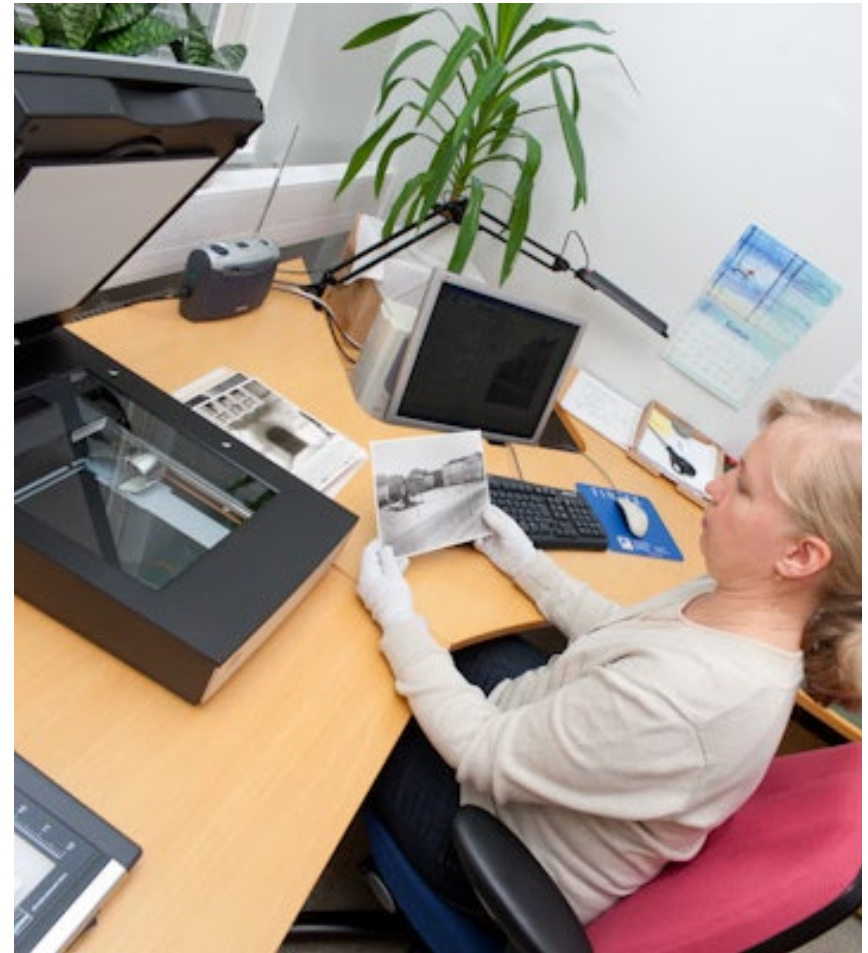
Mistä kaikki lähti liikkeelle

- 1984 ensimmäinen PC vuonna
- 1986 ensimmäinen arkistotietokanta
- 1988 Valtakunnallinen yritysarkistorekisteri (painetut kirjat Yritysarkistot I ja II)
- 1989 erikoismateriaalirekisterit
- 1990-luvun alussa arkistojen luettelointi Paradox pohjaiseksi
- 1993 ensimmäinen kirjastotietokanta
- 1996 sisäverkko ja ensimmäinen palvelin
- 1998 *www.elka.fi* eli Internet-sivut, jossa jo arkistotietokanta mukana
- 2002 *Elma*-arkistotietokanta, johon yhdistettiin aikaisempi tietokanta + erillisrekisterit
- 2006 *ElkaD*-digiarkisto av-aineistoille (sis. digitaalinen tutkijasali)



Digitointia jo neljännesvuosisata

- Valokuvien digitointi aloitettiin 1999 ja ne tulivat esille Internet-sivuillakin olleeseen valokuvarekisteriin
- Aluksi pieniä esikatselukuvia. Asiakkaille tehtiin tilauksesta korkearesoluutioisia
- Liikkuvan kuvan ja äänitteiden osalta iso urakka tehtiin 2004 – 2006 ElkaD-projektissa
- Nykyisin äänitteistä noin 80% ja Liikkuvasta kuvasta noin 40% on digitoitu
- Asiakirjallisen aineiston digitointi on käynnistynyt todenteolla vasta 2010-luvulla
- Nykyisin tehdään lähes kaikkea negatiiveista A0 –piirustuksiin saakka
- AV-aineistot ja mikrofilmit ulkoistettuna



Työvälineitä



Digitoinnin syvin olemus

- ”Aineen olomuodon muuttaminen kuluttaa energiaa”
- Tiedon säilytysmuodon muuttamiseen pitää olla hyvät perusteet
 - Säilyvyys, käytettävyys, käyttöaste, jakelu (edellä mainituista johdettu rahansäästö)
- Kannattaa tehdä organisaatiokohtainen digitointistrategia.
 - Mitä haemme digitoinnilla?
 - Mitä konkreettisia hyötyjä saamme?
- Kirkastaa itselle ns. iso kuva. Miksi teemme digitointia ja mitä siitä hyödyimme. Elkan tapauksessa:
 - AV-aineistoja digitoidaan pelastamistarkoituksessa
 - Asiakirja-aineistoja digitoidaan käytettävyyden ja käyttöasteen nostamiseksi.
- On demand –tyylinen digitointi lienee monelle organisaatiolle järkevää



Mistä resurssit digitointiin?

- Digitointia tehdään ns. virkatyönä, mutta tällä vauhdilla nykyisen aineiston digitointi vie vielä 875 vuotta
- 2008 – 2010 paikkeilla oli ”tekemisen meininkiä” ja jaettiin digitointimiljoonia myös arkistosektorille
- Yksityisillä keskusarkistoilla oli kaksi ääniprojektia ja Pikadigi-niminen hanke, jossa saatiin lähes miljoona euroa aineistojen digitointiin
- Samoihin aikoihin käynnistyi Kansallinen Digitaalinen Kirjasto – hanke
- ...jonka muuten piti olla valmista vuonna 2011, mutta esim. Elka sai aineistonsa Finnaan vasta 2023.



Muita haasteita ja näkökulmia

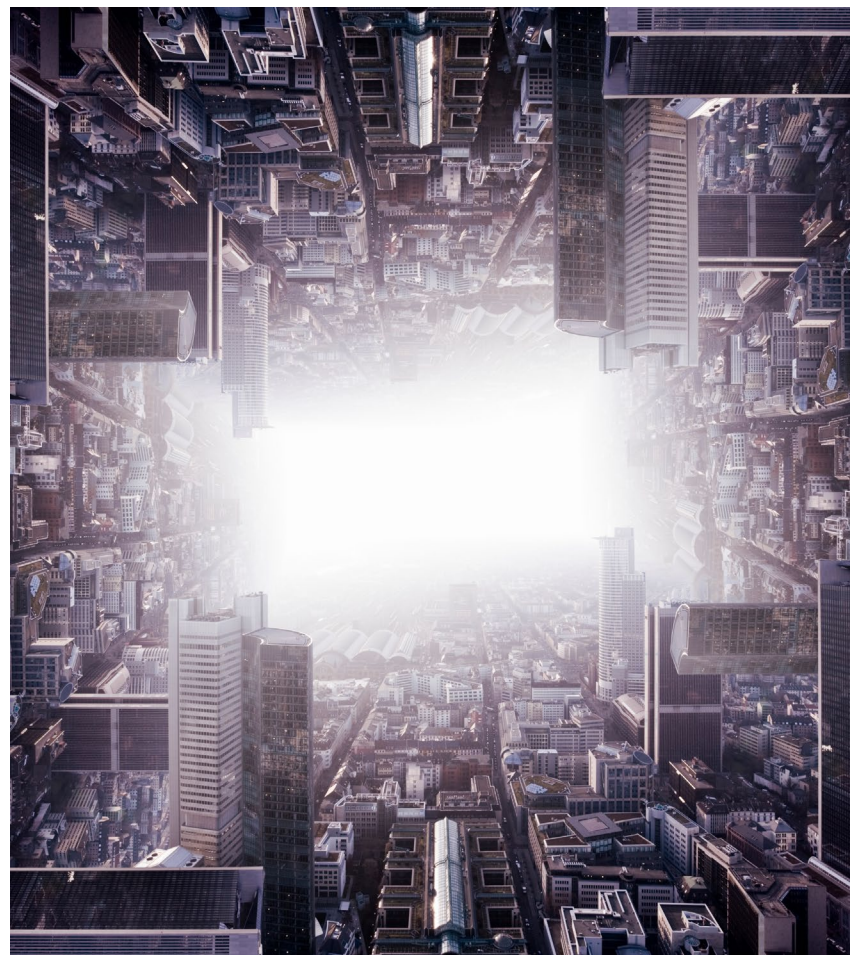
- Pienillä toimijoilla haasteita kalliiden ammattilaistason skannereiden hankinnassa
- Moni organisaatio käyttää järjestelmäkameroita ja jopa monitoimikoneita digitointiin
- Myös osaaminen voi olla pullonkaula. Mikkelissä sinänsä hyvä tilanne, koska täällä alan koulutusta
- Suuret digitointikeskukset olisivat varmasti kustannustehokkaampia
- ...toisaalta skannaus on korkeintaan noin 1/10-osa koko digitointiprosessista
- Eli osataanko näitä kustannuslaskelmia tehdä kaikki kustannukset huomioiden.



Syntysähköiset. Mikä siinä kestää?

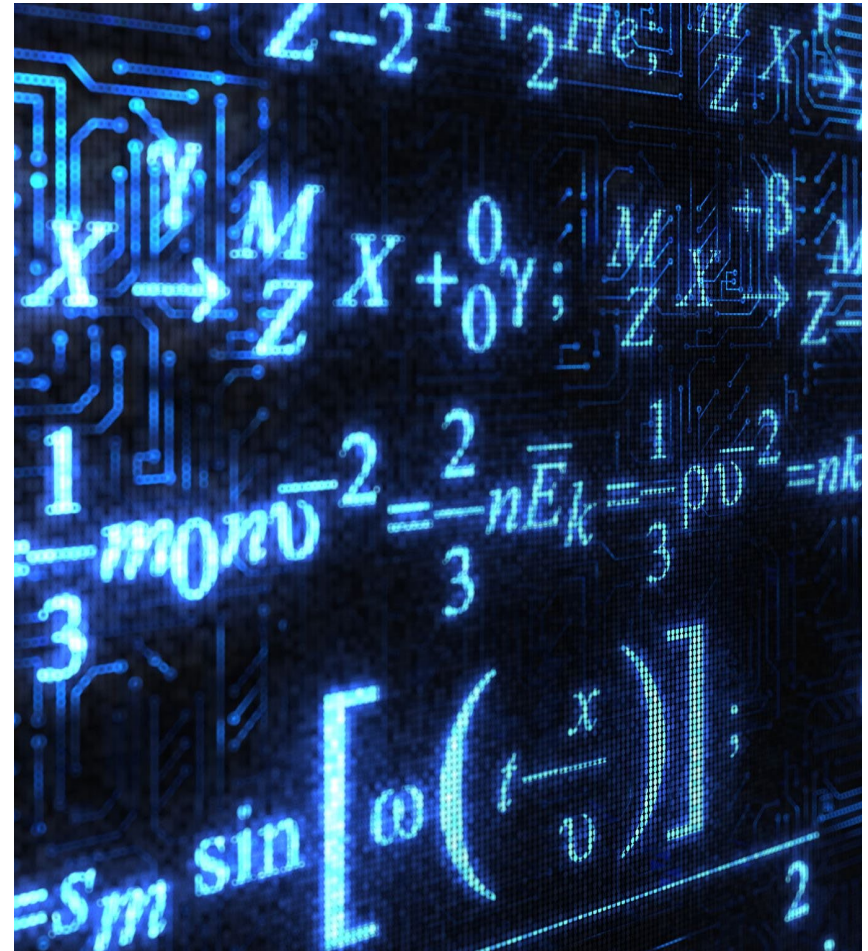
- ”Eräät yksityiset keskusarkistot – etunenässä Suomen Elinkeinoelämän keskusarkisto – varautuvat jo sähköisen aineiston vastaanottoon ja sähköiseen arkistointiin.” – Pääjohtaja Kari Tarkiainen Oulun arkistopäivillä 1998
- ”No niinpä” – Olli Alm Mikkelin DAF-seminaarissa 6.9.2024
- Olin kuvitellut, että ensin ratkaistaisiin syntysähköisen aineiston ongelma, ja siten katkaistaisiin digitaalisen ajan mustan aukon laajeneminen, ja vasta sen jälkeen keskityttäisiin digitointiin...mutta asiat näyttääkin menevän päinvastoin*
- Ei kuitenkaan pelkoa etteikö meistä jäisi merkkejä tulevien sukupolvien historiantutkijoille ja arkeologeille...

*Tarkoitan päätearkistoja, en eri organisaatioiden operatiivisia järjestelmiä



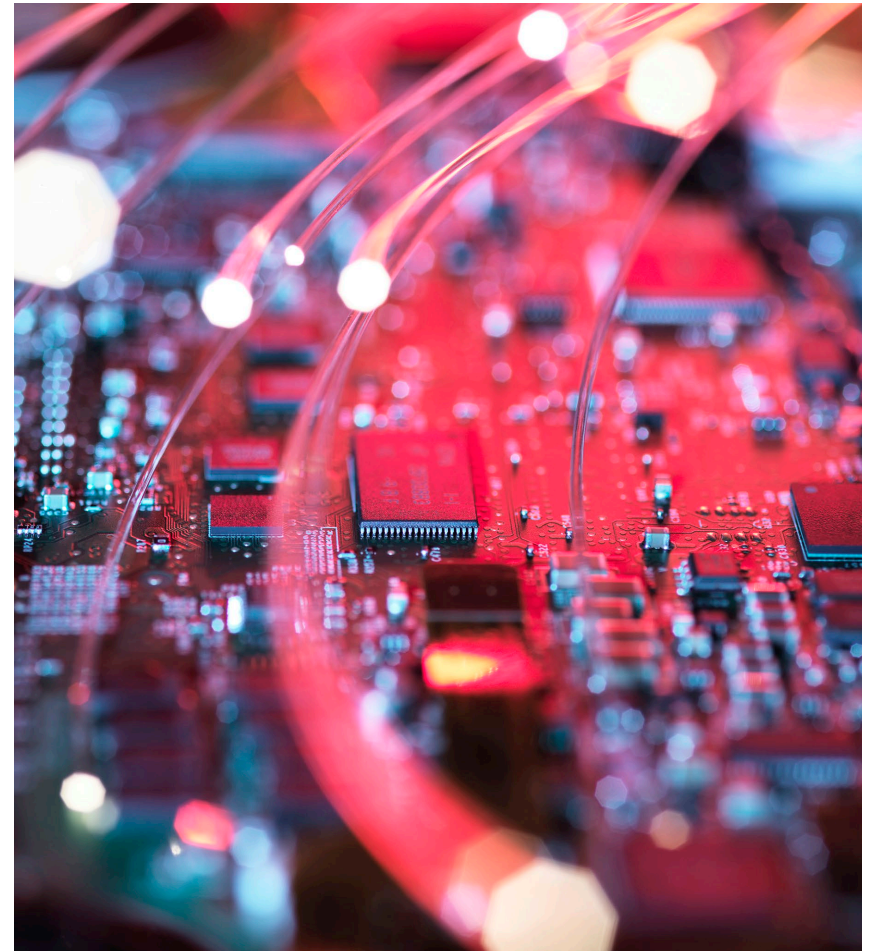
Haasteita

- Digitoimalla tuotetut tiffit ja PDF:t on helppoa ja homogeenista aineistoa
 - Ja varsinkin jos ne on arkistojen itsensä digitoimia
 - Tilanne kentällä on toinen. Suomessa on n. 560 000 yritystä, n. 140 000 yhdistystä, jne.
 - Tiedostoformaatteja tuhansittain, erilaisia tietojärjestelmiä tuhansittain, metatiedot hyvin vaihtelevia
- ➔ Määrämuotoisten vastaanottopalveluiden rakentaminen hankalaa
- ➔ Vaatii lähes aina kustomointia, määppämistä ja muuta käsityötä



Mitä on saatu aikaan

- Syntysähköisiä aineistoja on kuitenkin vastaanotettu jo 20+ vuotta, jos huomioidaan arkistojen mukana tulleet CD- ja DVD-levyt, muistitikut, yms.
- Myös hallitumpia FTP-siirtoja on tehty viimeisten 10+ vuoden aikana useampia
- YKSA –arkistotietokanta soveltuu mainiosti syntysähköisten aineistojen säilytykseen
- Siirrot ovat aina kuitenkin ”tapauskohtaisia” ja ne vaativat paljon käsityötä
- Tavoitetilä on pysynyt muuttumattomana, eli rakentaa joustava vastaanottolaituri, jonne asiakas voi helposti tallentaa itse aineistojaan
- Toivottavasti Memoriaali-järjestelmä lopulta toteuttaisi tämän tavoitteen



Arkistojen järjestämisen tulevaisuuskuva

- Klassisen luetteloinnin ja kuvailun kulta-aika lienee jo takana päin
- Henkilöresurssit eivät yksinkertaisesti riitä niin tarkkoihin arkistokuvailuihin ja –historioihin kuin 10 tai 20v sitten.
- Pääpaino on selkeämmin aineiston luetteloinnissa arkistoyksikkötasolla ja siitä ylöspäin
- Järjestäminen nähdään entistä enemmän digitoinnin esivaiheena
- Digitointi + optinen tekstintunnistus + tekoälyn avustama sisältöanalyysi vievät aineiston löydettävyyttä ja käytettävyyttä eteenpäin (osittain arkipäivää jo 2024)
- ChatGPT –tyylinen keskusteleva botti auttane loppuasiakasta jo muutaman vuoden päästä



Tulevaisuuden varjokuvat

- Ruusuinen tulevaisuuskuva koskee tietysti vain digitoituja ja syntysähköisiä aineistoja, jotka ovat murto-osa kokonaisuudesta
- Todennäköisesti valtaosa tutkimuksesta kohdistuu tuohon marginaaliseen, ja ehkäpä satunnaisesti valikoituneeseen, aineistojoukkoon
- Pahimmillaan johtaa merkittäviin vääristymiin ja laadun heikentymiseen esim. historiantutkimuksessa
- Myös päätöksenteko voi heikentyä, jos hukataan ”tietoisuus” jostakin tietoaaineistosta
- Tiedon eheys ja luotettavuus. Jo nyt netti pullollaan tekoälyn tuottamaa valetietoa. Kuinka estetään pääsy arkistoon
- Tietoturva, myös pahantahtoisia valtiollisia toimijoita vastaan



Kiitoksia!

Kysymyksiä,
kommentteja?

Olli Alm

044 321 3422

olli.alm@elka.fi

