

Kas ja kuidas säilib fail?

Lauri Leht

Rahvusarhiivi digitaalarhiiv

21.11.2014 Imaveres

Väide:



~



?

Lihtsustame:

fail => digitaalne dokument

digitaalne dokument õiguslikult = igasugune
dokument = paberdokument

Dokumendi olulised omadused:

- autentsus
- usaldusväärsus
- terviklus

autentsus = allkiri?

- allkiri = digitaalallkiri = DDOC (BDOC)



Digiallkirja säilitamine

Kui meil on:



Digiallkirja säilitamine

aga vaja on teha:



siis:

~~Digiallkiri~~

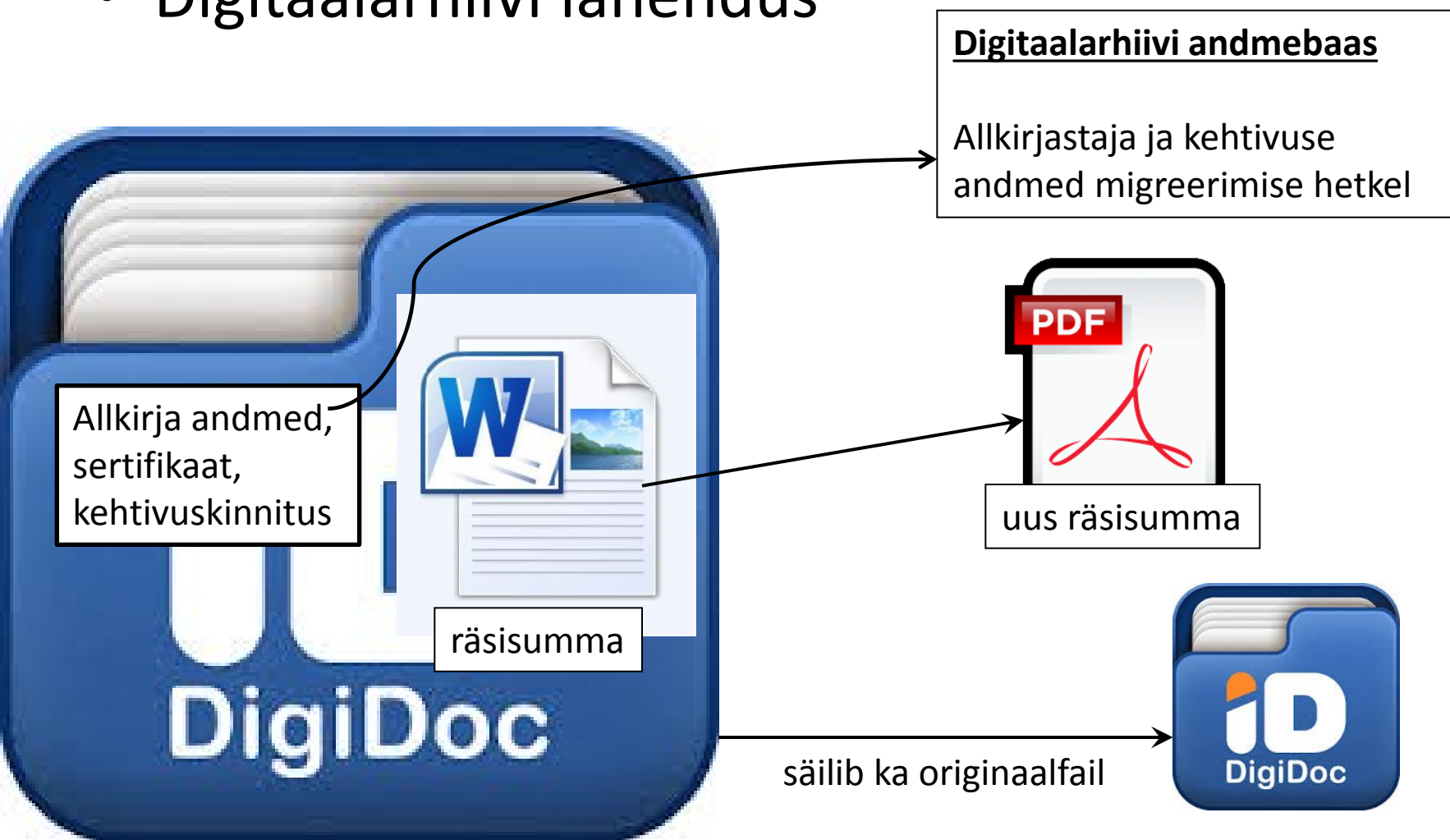
~~Räsisumma~~

~~Autentsus~~

Digiallkirja säilitamine



- Digitaalarhiivi lahendus



DDOCi probleem

- DDOC = aasta 2002 = vananenud tehnoloogia
- DDOC murtakse lahti 5-10 aasta jooksul
- BDOC = aasta 2013 tehnoloogia => valdav 2015 => kohustuslik 2016

- Mis saab vanadest DDOCidest???
 - Käsitsi üle allkirjastamine
 - Süsteemides automaatne ajatembeldus ja uus allkirjastamine
 - „nalja“ saab = kohtud hindavad tõestusväärtust

Terviklus

- Räsiumma (hash) = unikaalne kontrollsumma
- Erinevad meetodid: SHA-1, SHA-256, SHA-512, MD4, MD5
- MD5 (128 bit) näide:
0F05F18AAC48DFAA95AFBEBBC418F12AB
- Transpordil, säilitamisel

Vormingud

Tavalised vormingud:

DOC, XLS, PPT, RTF, JPG, DOCX jne

Arhiivipüsivad vormingud:

TXT, PDF, PNG, TIFF, XML

Mis vahe neil on?

Arhiivipüsiv vorming on ...

- avatud, ehk sisemise struktuuri kirjeldus peab olema avalik = *ei ole Microsofti oma*
- kadudeta ehk kompresserimata
- laialdaselt toetatud erinevate tarkvaratootjate poolt
- failide kasutamiseks (eelkõige sisu vaatamiseks) on olemas tasuta vahendid
- standardiseeritud (n ISO)

PDF vs PDF/A

- PDF/A → PDF-i kitsendus, mis jätab välja kõik pikaajaliseks säilituseks sobimatu (audio, video, käivitataavad failid, hüperlingid, krüpteerimine)
- Iga juba olemasolev PDF võib olla PDF/A ühilduv, kui loomisel on PDF/A nõudeid järgitud

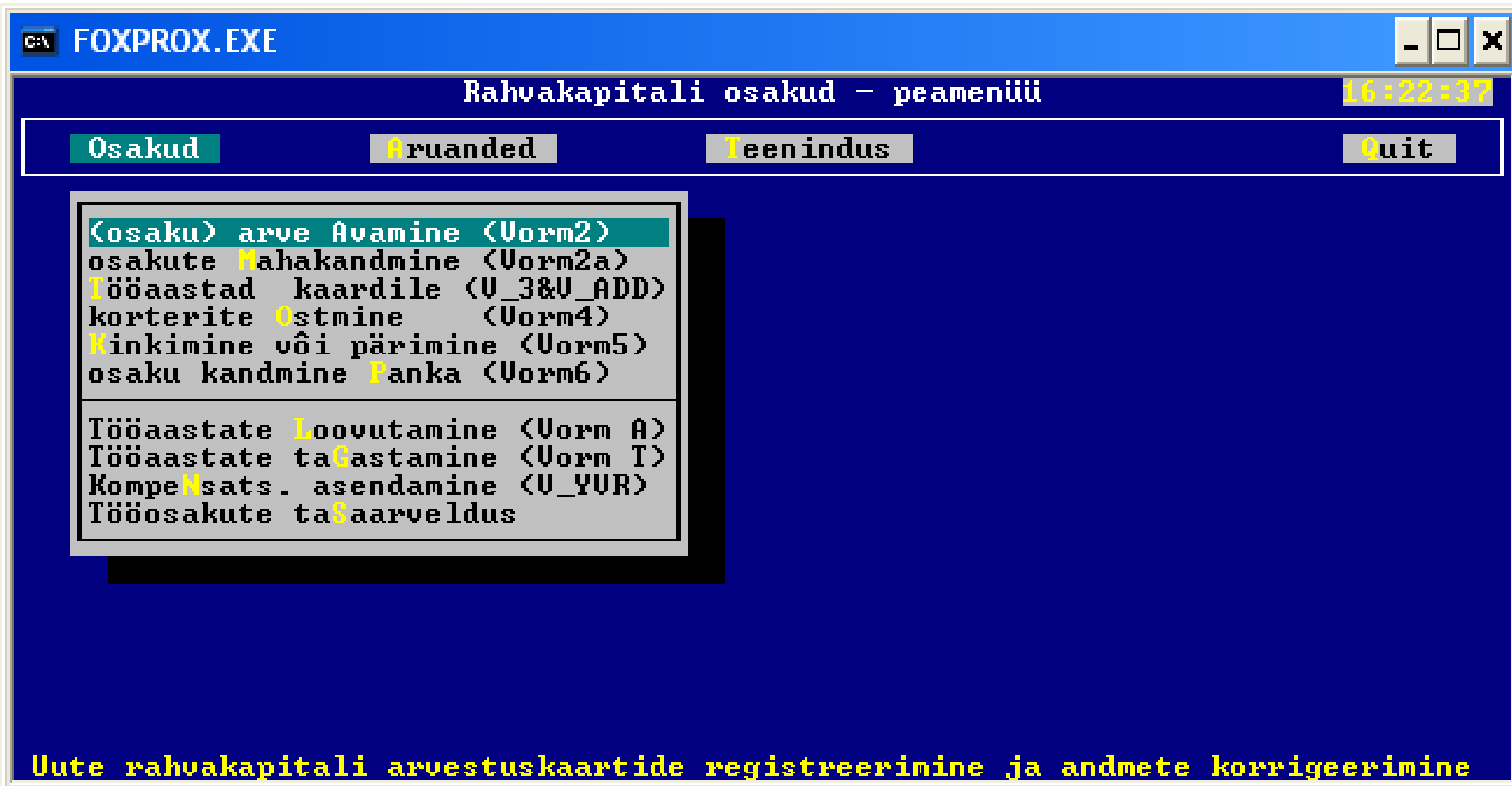
Lihtsalt teadmiseks:

- PDF/E = jooniste ja ruumiandmete jaoks
- PDF/X = graafika jaoks

Põhimõttelised strateegiad

- Migreerimine = fail ei säili, säilib info
- Emuleerimine = fail säilib, muutub kasutustarkvara
- Muuseum = säilib fail ja säilib tarkvara ja riistvara

Emuleerimine



Füüsilised andmekandjad

- Mitmel erineval andmekandjal, mitmes erinevas kohas
- LOCKSS = LOTS OF COPIES KEEPS STUFF SAFE
- Vaja migreerida uutele andmekandjatele
- Digitaalarhiivides: kettamassiiv + LTO-lindirobot

Kõvaketas vs lint



- 300 – 3000 GB
- 100 – 1000 €/tk
- 0,2-0,33 €/GB



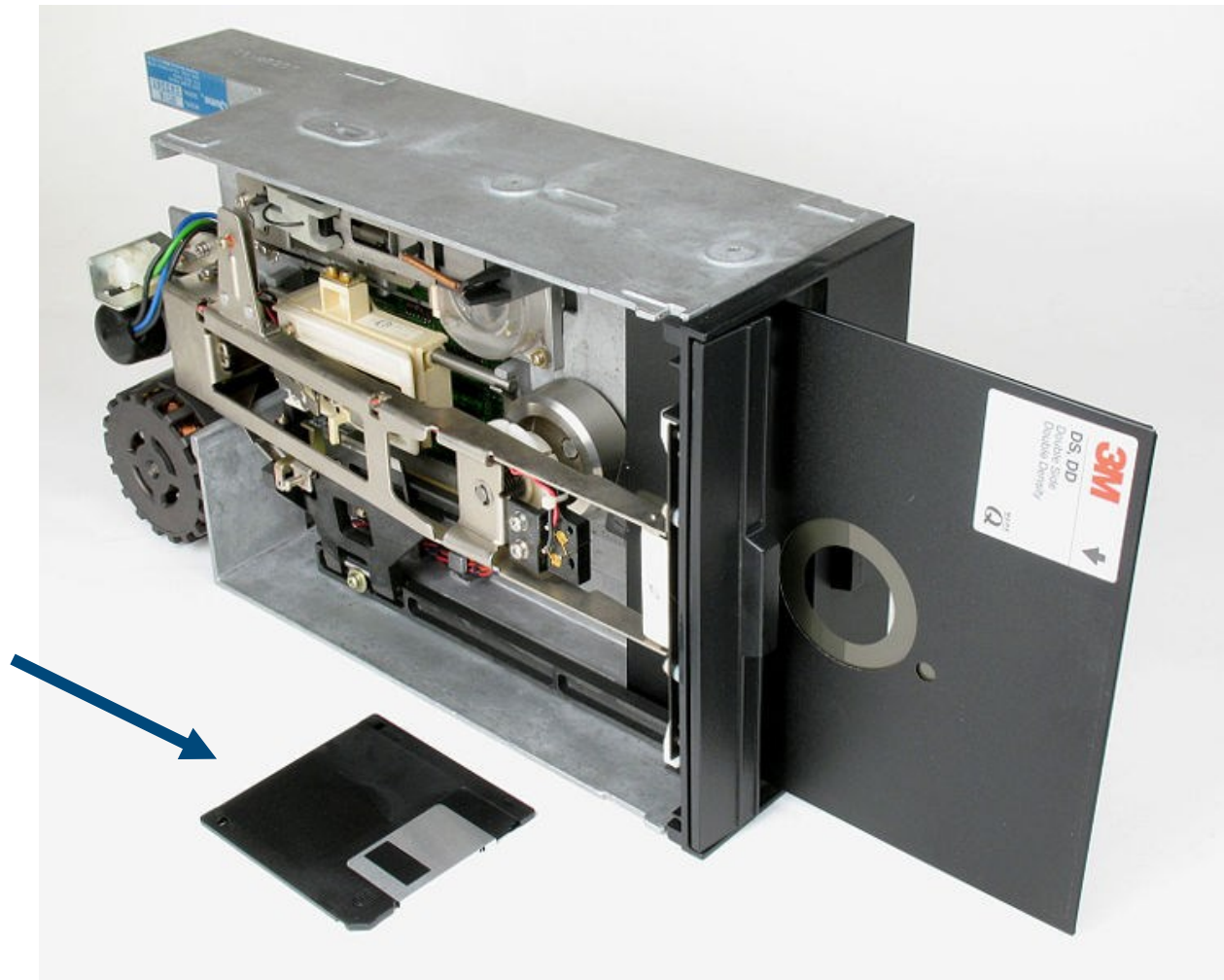
- 1500 GB
- 25 €/tk
- 0,016 €/GB

Andmekandjad aeguvad moraalselt

- Lugeva seadme puudumine
- Lugeva seadme külge ühenduva arvuti puudumine
- 8" diskett – ammu unustatud
- 5,25" diskett – u 15 a käibelt maas
- 3,5" diskett – u 5 a käibelt maas
- Lahtine magnetlint – ammu unustatud

8" diskett ja lugev seade

Mõõtude
võrdlemiseks
3,5" diskett

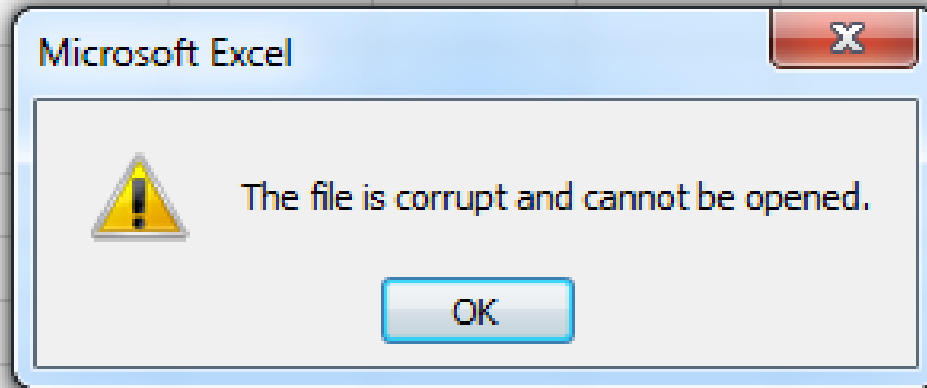


Viimase paari aasta arengud

- Windows XP lõpp = probleemid SCSI liidesega seadmetega
- Uutes arvutites reeglina pole optilist lugejat
- Blu-ray seadmed pole alati CD-le tagasiühilduvad



Failide konserveerimine?



Niisiis, kuidas?



~



Aitäh!