

VAATENURK

KUIDAS PABERI KÕRVAL HAKATI BITTE JA BAITE SÄILITAMA? DIGITAALSE INFO SÄILITAMISE ARENGUJOONI

Kurmo Konsa

Inimühiskondade kujunemine ja areng on lahutamatu seotud informatsiooni ja kommunikatsiooniga. Info loomise, levitamise, kasutatavaks tegemise ja säilitamisega tegelevad kõikides teadaolevates ühiskondades spetsiaalsed inimesed, organisatsioonid ja institutsioonid. Suulisel infol baseeruvates ühiskondades olid nendeks inimesed ja inimeste grupid, kes tegelesid narratiivide meelepidamise, esitamise ja edasiandmisega. Suulisele tekstile lisandusid visuaalsed kujundid, maastikumärgid jms, mis kõik moodustas ühtse kommunikeeritava teadmise.

Läbi ajaloo võime me jälgida ühelt poolt järjest spetsiifilisemate ülesannetega organisatsioonide ja institutsioonide teket, teiselt poolt aga ka erinevate funktsioonide ühtesulamist. Pärast trükikunsti leiutamist oli sageli üks inimene või ettevõtte nii kirjastaja, trükkali kui ka raamatulevitaja ülesannetes. Veelgi enam, tihti oli ta ka ise autoriks. Hiljem toimus funktsioonide eraldumine ning kaasajal on need jagunenud mitte ainult ülesannete, vaid ka trükiste tüüpide alusel (ajalehekirjastus, teadusraamatute kirjastus, lasteraamatute kirjastus jne). Alates digitaalse kirjastamise, trükkimise ja levitamise kasutuselevõtust võib aga jällegi üks inimene olla nii info looja, levitaja, kasutatavaks tegija kui ka säilitaja.

Digitaalne kommunikatsioon on kaasaegse ühiskondliku infosüsteemi aluseks. E-kirjad ja muidu internetiteenused, mobiilitehnoloogiad, raadio ja televisioon, filmid, kujutised – kõik see baseerub digitaalsel tehnoloogial, millest suur osa jääb tavakasutajatele tabamatuks, taandudes tehnilistesse

süsteemidesse. Enamik tehnilistest seadmetest on tänapäeval varustatud digitaalsel põhimõttel töötavate mikroprotsessoritega. Lisaks inimestevahelisele kommunikatsioonile on ilmunud inimeste ja masinate vaheline kommunikatsioon ja masinate endi vaheline kommunikatsioon.

Ühiskondlike infosüsteemide üheks funktsiooniks on teabe säilitamine, selle olemasolu ja kasutatavuse tagamine. Seejuures saab info säilitamine olla vaid osa laiemast kommunikatsiooniprotsessist, mis määrab säilitamise funktsiooni. Seega on teabe säilitamine alati aktiivne tegevus. Säilitamist võiks siinkohal määratleda kui funktsiooni, mis annab infosüsteemile pidevuse. Enamikku ühiskonnas kasutatavast teabest loomulikult ei säilitata, samas aga pööratakse osa informatsiooni säilitamisele spetsiaalset tähelepanu. Kuid tänapäevases infoühiskonnas on süvenemas soov võimalikult enama info säilitamise järele. Osaliselt on selle põhjuseks ilmselt ka digitaalne teave ise – seda on võrratult lihtsam salvestada. Digitaalsete infoobjektide korral on seos info kandja (meediumi) ja info enda vahel märksa efemeersem. Digitaalsel kujul esitatud info sõltub tervest ahelast kodeerimis- ja dekodeerimisetappidest ning kasutajale esitatav info luuakse selle taasesitamise hetkel.

Digiinfo korral saab eristada kolme klassi objekte: füüsilised, loogilised ja kontseptuaalsed objektid.¹ Füüsilised objektid on meediumile (magnetlint või -ketas, optiline kompaktplaat vms) kantud märgid. Kuidas need märgid sinna kantakse ja mil viisil nende lugemine toimub, sõltub meediumist. Objektile kantud füüsilised märgid ei ole veel bitid! Tegemist on lihtsalt mingite märkidega mingil füüsilisel objektil või füüsilises keskkonnas. Füüsilise objekti tasand on täiesti sõltumatu bittide tähendusest ning füüsiliste märkide bittidena käsitlemiseks on vajalik nende töötlemine arvutisüsteemi poolt. Pärast seda, kui vastav riist- ja tarkvara on need märgid välja lugenud ja töödeldud, muutuvad füüsilised objektid loogilisteks objektideks. Inimesele esitatakse digitaalne infoobjekt kontseptuaalse objekti kujul. Kontseptuaalseks infoobjektiks võib olla näiteks dokument, foto, raamat, intervjuu või film. Kontseptuaalse objekti sisu ja struktuur sisalduvad loogilises objektis või objektides, aga ei ole sellele üks-üheselt taandatav. Nii näiteks võib mingi digitaalne dokument esineda nii tekstitöötlusdokumendina kui ka pildifailina. Seega, ühesugune kontseptuaalne objekt võidakse saada täiesti erinevatest loogilistest ja

¹ Kenneth Thioboedeau, "Overview of technological approaches to digital preservation and challenges in coming years", *The state of digital preservation: an international perspective*, Conference proceedings documentation abstracts, inc. Institutes for Information Science Washington, D.C. April 24–25, 2002 (2002).

füüsilistest objektidest. See näide illustreerib kahte digiobjektide eripära, mis on olulised ka säilitamise seisukohalt. Ühel kontseptuaalsel objektil võib olla mitu digitaalset kodeeringut (loogilist objekti) ja need säilitavad kõik kontseptuaalse objekti olulised tunnused.

Digitaalse objekti säilitamiseks peame esiteks teadma seoseid füüsiliste, loogiliste ja kontseptuaalsete objektide vahel. Teiseks ei seisne digitaalne säilitamine ainult füüsiliste objektide säilitamises, vaid tuleb säilitada võimet neid objekte esitada. Digitaalsel kujul esineva teabe säilitamine on ühelt poolt seotud tehnoloogiate arenguga, kuid teisalt ka sotsiaalse ja kultuurilise kontekstiga. Kui säilitamise tehnilise poolega on laialdaselt tegeledud, siis selle seostamine laiemate ühiskondlike protsessidega on jäänud varju. Käesoleva artikli eesmärgiks on käsitleda digitaalse info säilitamise kujunemist ning tuua välja peamised kontseptuaalsed ideed ja põhimõtted, millest on selle käigus lähtutud. Selline ülevaade on ühelt poolt vajalik teabe kasutajatele, kelle ülesanne on hinnata, millist teavet säilitatakse ja kuidas seda tehakse, teisalt aitab see kaasa säilitussüsteemide edasisel kujundamisel.

Kui arvutite ja digitaalsete kommunikatsioonivahendite arengust on kirjutatud põhjalikke ülevaateid,² siis digiinfo säilitamise vallas ei ole midagi sellist seni ilmunud. Põhjuseks on ilmselt asjaolu, et säilitusspetsialistid tegelevad digisäilitamise enda, mitte selle ajaloo uurimisega ning ajaloolaste silmis ei ole tegemist piisavalt huvitava teemaga.

² Vt nt Gerard O'Regan, *Brief history of computing* (London: Springer, 2012); Michael R. Williams, *A history of computing technology*, second edition (Los Alamitos: IEEE Computer Society Press, 1997); Georges Ifrah, *The universal history of computing: from the abacus to the quantum computer* (New York, Chichester et al.: 2001); Alfred D Chandler, Takashi Hikino, Andrew Von Nordenflycht, *Inventing the electronic century: the epic story of the consumer electronics and computer industries* (Cambridge, London: Harvard University Press, 2005); *The first computers – history and architectures*, ed. by Raúl Rojas and Ulf Hashagen (Cambridge, London: MIT University Press, 2000); Paul E. Ceruzzi, *A history of modern computing* (Cambridge MA: MIT Press, 1998); René Moreau, *The computer comes of age: the people, the hardware, and the software* (Cambridge MA: MIT Press, 1984); Jon Agar, *The government machine: a revolutionary history of the computer* (Cambridge, London: MIT University Press, 2003); Martin Campbell-Kelly, William Aspray, Nathan Ensmenger, Jeffrey R. Yost, *Computer: a history of the information machine*, third edition (Boulder: Westview Press, 2014); Charles J. Bashe, Lyle R. Johnson, John H. Palmer, Emerson W. Plugh, *IBM's early computers* (Cambridge, London: MIT Press, 1986).

Masinloetavate infokandjate kasutuselevõtt

Alates 19. sajandi keskpaigast võeti kasutusele mitmed uut tüüpi tehnoloogiad, mis muutsid oluliselt kommunikatsioonisüsteeme. Töökindel elektriline telegraaf leiutati 1840. aastatel, 1876. aastal patenteeriti telefon ning 1877. aastal fonograaf, 1895. aastal võeti kasutusele kinofilm. Tänapäevase automaatse digitaalse andmetöötluse eelkäijaks võib aga pidada perfokaartidel põhinevat mehaanilist andmetöötlussüsteemi, mille lõi ameeriklane Hermann Hollerith 1880. aastatel. Andmete säilitamine ja töötlemine toimus perfokaartide abil, mis kujutasid endast standardiseeritud kujuga kartongist kaarte. Teave kanti perfokaartidele kindlate positsioonide mulgustamise teel.

Kõigile eelpool loetletud kommunikatsioonivahenditele on iseloomulik see, et info edasiandmiseks, salvestamiseks ja kasutamiseks on vajalikud tehnilised süsteemid, varasemast ajast pärinevate infosüsteemide puhul, nagu näiteks suuline kõne, numbrisüsteemid, kiri, pildid jms, ei ole selliseid tehnilist vahendussüsteemi vaja. Raamatu lugemiseks tuleb küll osata keelt ja tunda kirja, kuid seda saab teha ilma tehniliste seadmete abita. Fonograafirulli, grammofoniplaadi või perfokaardi korral on selle kasutamine ilma vastava tehnilise seadmestikuta aga võimatu. Sellistes süsteemides kasutatavaid infokandjaid nimetati masinloetavateks (*machine readable records, machine dependent records*), viidates nende kasutamiseks vajaminevatele tehnilistele seadmetele.³ Kuidas suhtuda aga nende säilitamisse – põhinesid ju kõik senised info säilitamiseks kujunenud süsteemid (raamatukogud ja arhiivid) n-ö klassikalistel infokandjatel ning kogemusi uute infokandjatega oli veel vähevõitu?

Arhiivinduslikust seisukohast loeti perfokaardid koos kinofilmide, helisalvestiste, kaartide, plaanide jms dokumentideks juba 1939. aastal Ameerika Ühendriikides vastu võetud *Federal Records Act*'iga (ka *Records Disposition Act*).⁴ Nende arhivaalideks (st dokumentideks, millel on arhiiviväärtus ning mida tuleb säilitada) tunnistamine oli märksa vastuolulisem. Ameerika Ühendriikide rahvusarhiivi nõuandva komisjoni otsusega samast aastast jäi otsustusõigus selle üle, kas perfokaartidel on ajalooline väärtus, mille tõttu tuleks neid säilitada, riigiasutustele endile. Üldiselt otsustasid asutused perfokaardid pärast andmete töötlemist hävitada.⁵

³ Mõiste ise võeti kasutusele 1950. aastatel, vt Margaret O. Adams, "Punch card records: precursors of electronic records", *The American Archivist*, 58 (Spring 1995), 182–201 (187).

⁴ *An act for the disposition of certain records of the United States government*, 5 August 1939 (53 Stat. 1219–1221).

⁵ Adams, "Punch card records", 182–201.

Seejuures tugineti argumendile, et perfokaarte kasutati peamiselt andmetöötluks, säilitamisele kuuluvad aga algandmed ja nende töötlemise tulemused aruannete jms näol. Arhivaaride seas oli levinud arvamus, et perfokaardid on vaid vahendid andmete töötlemiseks ja iseseisvalt puudub neil arhiiviväärtus, seega pole ka põhjust neid säilitada.⁶⁷

Elektronarvutite leiutamine 1930. aastate lõpul ja laiem kasutuselevõtt riigiasutustes, äriettevõtetes ning teadusasutustes 1950.–60. aastatel infoandjate säilitamise seisukohast muutusi kaasa ei toonud. Elektronarvutite juures kasutati info sisestamiseks ja programmide salvestamiseks perfokaarte ja perfolinte, viimased võisid olla nii paberist kui ka metallist. Arvutuste tulemused trükiti välja paberile või väljastati perfolintidel. Seega käsitleti perfolinte ja perfokaarte vahepealsete andmekandjatena ning töödeldud informatsiooni säilitati enamasti kirjalikult erinevate dokumentide kujul.

Digitaalseid andmeid koguti erinevates tööstusvaldkondades (ravimistööstus, ehitus) ning lisaks sotsiaalteadustele ka teistes teadusvaldkondades (meteoroloogia, geofüüsika, tuumafüüsika). Selliseid andmeid säilitati reeglina samades asutustes, kus neid ka koguti ja töödeldi. Mõned üksikud institutsioonid kogusid andmeid erinevatelt loojatelt, et luua arhiiv, mida võisid kasutada erinevad uurijad.⁸ Üldjuhul kogutud andmete pikaajalist säilitamist ei kavandatud, neid hoiti ainult jooksvaks kasutamiseks.

Tõsiseks probleemiks kujunes kosmoseuuringutega seotud info säilitamine. 1952. aastal esitas International Council of Scientific Unions idee rahvusvahelise geofüüsika-aasta korraldamiseks 1957. aasta juulist 1958. aasta detsembrini. Selle käigus kogutavate andmete säilitamiseks ja jagamiseks loodi rahvusvaheline World Data Center (WDC).⁹ Teiseks väga oluliseks kosmose ja maa uurimisega seotud andmete kogujaks ja säilitajaks kujunes Ühendriikide rahvuslik aeronautika- ja kosmosevalitsus NASA. 1964. aastal alustas Goddard Space Flight Center (GSFC), mis vastutas satelliitidelt kogutava teabe säilitamise eest, vastava arhiivisüsteemi väljatöötamist.¹⁰ Töötati välja telemeetriaandmete arhiivi funktsionaalne skeem,

⁶ Adams, "Punch card records", 182–201.

⁷ Vt Meyer H. Fishbein, "Appraising information in machine language form", *The American Archivist*, 35:1 (January 1972), 35–43 (36).

⁸ Sellistest andmeid koondavatest asutustest võiks nimetada näiteks Social Science Research Council (SSRC) Data Bank'i Suurbritannias Essexi ülikooli juures, vt Patricia Sleemann, "It's public knowledge: The National Digital Archive of Datasets", *Archivaria. The Journal of the Society of Canadian Archivists*, 58 (2004), 173–200 (178).

⁹ WDC tegutseb tänapäevani, vt <<http://www.icsu-wds.org/>> (19.11.2014).

¹⁰ A. M. Demmerle, R. G. Holmes, W. B. Poland, "The GSFC scientific data storage problem", *NASA Technical Reports Server (NTRS)* (January 1967)

kuid arhiivi tegelik rakendumine ei toimunud. NASA arendas välja terve infrastruktuuri erinevate kosmoseandmete kogumiseks ja säilitamiseks.

Digitaalsete andmete salvestamiseks kasutati esmakordselt magnetlinti 1951. aastal Eckert-Mauchly poolt loodud üldotstarbelise elektronarvuti UNIVAC 1 (lühend nimetusest Universal Automatic Computer I) juures. Magnetlindina kasutati 12,7 mm laiust niklikihiga kaetud metall-linti. IBM arvutitel võeti samuti 1950. aastael kasutusele magnetlindid, aga need olid juba raudoksiidiga kaetud atsetaatselluloosalusel. Sellest ajast alates hakkasid magnetlindid üha enam asendama perfokaarte ja perfolinte masinloetava info säilitamisel. Magnetlintide sobivus andmete pikaajaliseks säilitamiseks ei olnud aga selge. Magnetlintide endi säilitamist arhiivis ei peetud otstarbekaks. Magnetlinte soovitati lugeda “vahepealseks meediaks” (*interim media*) ja infot säilitada endiselt paberdokumentide kujul.¹¹ See pidi tagama info säilimise juhul, kui lindiga midagi juhtub.¹²

Võrreldes perfokaartide ja perfolintidega võimaldasid magnetlindid märksa suuremat andmetihedust ja kiiremat andmevahetust, nii asendaski 1960. aastatel magnetlint perfokaardid ja perfolindid elektronarvutite sisend- ja väljundseadmetes ning andmete välissalvestistes. Perfokaartidel olevatest andmetest kanti osa üle magnetlintidele. Seda tehti nii andme-arhiivides, organisatsioonides, kus neid andmeid koguti ja kasutati, ning ka arhiivides. Magnetlintidele ülekantud andmete hulk oli üldiselt siiski suhteliselt väike.¹³

Ameerika Ühendriikide *Federal Records Act* laiendas dokumendi mõistet masinloetavatele materjalidele (*machine-readable materials*) küll juba 1950. aastal,¹⁴ kuid endiselt jätkusid arutelud selle üle, kas arvutiinfo on dokument ja kui vastus on jaatav, siis kas sellel võib olla arhiiviväärtust. Suures osas oli paljude arhivaaride kahtlev seisukoht seotud arvutiinfo enda olemusega. Tegemist oli ju andmetega (*data*), statistiliste ja muude uuringute tulemuseks olid täidetud küsitlusvormid, millelt kanti andmed masinloetavasse vormi nende hõlpsama töötlemise eesmärgil. Samuti kujutas raamatupidamises ja muudes ärivaldkondades kasutatav elektrooniline info endast peamiselt andmefailde äriprotsesside kohta. Andmete pidamine dokumentideks ja veegi enam – arhivaalideks – oli paljude jaoks küsitav.

¹¹ Fishbein, “Appraising information in machine language form”, 37.

¹² *Ibid.*, 35–36.

¹³ Margaret O’Neill Adams, “Analyzing archives and finding facts: use and users of digital data records”, *Archival Science*, 7 (2007), 21–36 (25).

¹⁴ Linda J. Henry, “Appraisal of electronic records: traditional principles endure”, *Thirty years of electronic records*, ed. by Bruce I. Ambacher (Lanham: Scarecrow Press, Inc., 2003), 25–41 (28).

Ühendriikide arhiiviteoreetik Theodore R. Schellenberg kirjutas oma raamatus "Appraisal of modern public records" (1956) perfokaartide kohta, et need ei oma arhiiviväärtust, kuna nende kasutamine sõltub mehaanilisest või elektroonilisest seadmestikust ning neid loonud asutused on kogu kogutud andmestiku juba "ära kasutanud" ja sellel puudub teisene kasutusväärtus.¹⁵ Andmed on seega ajutise väärtusega ning nende põhjal loodud kokkuvõtavad dokumendid püsiva väärtusega. Andmete säilitamine suurendaks ainult andmemahutu ja raskendaks olulise info eristamist ebaolulisest.¹⁶

Hoopis teistsugune suhtumine masinloetavatesse teabesse valitses nendes asutustes, kus kogu töödeldav ja säilitatav teave oligi masinloetaval kujul. Sellist teavet hakkasid alates 1930.–40. aastatest koguma erinevad uurimisasutused ja projektid. Algselt olid need seotud valdavalt sotsiaalteadustega.¹⁷ Kogutud andmeid oli vajalik edasisteks uurimisteks säilitada ning selleks loodi vastavad andmekogud, mida üldnimetusena kutsutakse sotsiaalteaduste andmearhiivideks (*social science data archives*). Esimesed sellised andmearhiivid loodi sõltumatult traditsioonilistest arhiiviasutatsioonidest. Nii näiteks asutati 1946. aastal eraalgatuse korras Roperi keskus (The Roper Public Opinion Research Center, Williams College). Tegemist oli andmearhiiviga, kus säilitati kommertslike avalike küsitluste tulemusi alates 1936. aastast.¹⁸ 1950.–60. aastatel tekkinud andmearhiivid olid seotud akadeemiliste uurimisasutustega.¹⁹

Esimene sotsiaalteaduste andmearhiiv Ida-Euroopas loodi 1985. aastal Ungaris, samal aastal alustati andmearhiivi (Sotsiaalteaduste andmepank) loomisega ka Nõukogude Liidu Teaduste Akadeemia Sotsioloogia

¹⁵ T. R. Schellenberg, "The appraisal of modern records", *Bulletins of the National Archives*, 8 (October 1956), <<http://www.archives.gov/research/alic/reference/archives-resources/appraisal-of-records.html>> (19.11.2014).

¹⁶ Sellise seisukoha esitas nt J. E. Thexton, "Archival potential of machine-readable records in business", *The American Archivist*, 37:1 (1974), 37–42 (38, 41).

¹⁷ Vt Carolyn L. Geda, "Social science data archives", *The American Archivist*, 42:2 (1979), 158–166 (158–159).

¹⁸ Adams, "Punch card records", 196.

¹⁹ Esimeste andmearhiividena võib mainida: Social Systems Research Institute Wisconsin ülikooli juures (1950); Zentralarchiv für Empirische Sozialforschung Kölnis (1960); Inter-University Consortium for Political Research (ICPSR) Ann Arboris Michiganis (1962); International Data Library and Reference Service California ülikoolis Berkeleys (1963); DATUM Bad Godesbergis (1964); Steinmetzarchief Amsterdami ülikoolis (1964); Louis Harris Political Data Center Põhja-Carolina ülikoolis Chapel Hillis (1965) ning Social Science Research Council (SSRC) Data Bank Essexi ülikoolis Suurbritannias (1967).

Instituudi juurde.²⁰ Eesti Sotsiaalteaduslikku Andmearhiivi asuti looma 1993. aastal Tartu ülikooli juurde. Ülikooli finantsilisel toetusel alustati Eesti Raadio Arvutuskeskuses säilinud sotsiaalteaduslike uuringuandmestike konverteerimisega magnetlintidelt diskettidele. 1994. aasta suvel pöördui andmepanga loomise projektiga Avatud Eesti Fondi poole, kust vastati taotlusele positiivselt ning aastateks 1994–96 eraldati tegevustoe-tuse grant. Toetusgrandi abil õnnestus kaitsta hävimise eest üle 200 aastatel 1975–94 läbiviidud uuringu andmestiku, mis olid säilinud Eesti Raadio Arvutuskeskuses ja Tartu ülikooli arvutuskeskuses, ning viia need üle PC-formaati ja korrastada.²¹

Sotsiaalteaduste andmearhiivid olid niisiis esimesed, mis hakkasid koguma ja säilitama digitaalset teavet. Andmearhiivides säilitati andmeid esialgu perfokaartide kujul. Enne 1950. aastate lõppu ei tegeletud tõsiselt andmete säilitamise probleemidega. 1960. aastatel asendus perfokaart salvestusmediumina järk-järgult magnetlindiga ning ka olemasolevaid andmeid kanti nende säilivuse tagamiseks üle magnetlintidele. Selline info ülekandmise vajadus sõltuvalt infokandjate ja nende kasutamiseks vajaliku riist- ning tarkvara arengust määras ära andmearhiivide säilitusstrateegia. Nimelt leiti, et andmete säilitamise ja kasutamise tagab kõige paremini nende hoidmine riist- ja tarkvarast võimalikult sõltumatul kujul.²² Sellist lähenemist kergendas oluliselt asjaolu, et uuringute andmefailid olid reeglina lihtsa struktuuriga ning neid oli võimalik kergesti muuta nn lameandmebaasideks (*flat file databases*)²³, mis on küllaltki tarkvarasõltumatud. Andmed pandi kirja ASCII märgistikus ning arhiivifailivorminguna kasutati paljudes andmearhiivides statistikapaketi OSIRIS²⁴ failitüüpe. Samuti kasutati teiste levinumate statistikapakettide nagu SAS, SPSS, NSD-Stat failivorminguid.²⁵ Andmete füüsilise kandja säilitamine oli seega andmearhiivide vaatenurgast ebaoluline probleem.

²⁰ Brigitte Hausstein, Evelyn Brislinger, "Data and information transfer between Eastern and Western Europe", *Information dissemination and access in Russia and Eastern Europe: problems and solutions in East and West*, ed. by Rachel Walker, Marcia Freed Taylor (Amsterdam: IOS Press, 1998), 77–87 (80).

²¹ Vt Eesti Sotsiaalteaduslik Andmearhiiv ESTA, <<http://www.psych.ut.ee/esta/>> (19.11.2014).

²² Denise Lievesley, "Increasing the value of data", *History and electronic artefacts*, ed. by Edward Higgs. Oxford: Clarendon Press, 1998), 253–264 (256–257).

²³ Lameandmebaasiks nimetatakse ühetabelilise struktuuriga andmebaasi, mille struktuur ei toeta hierarhilisi suhteid kirjete vahel.

²⁴ IBM suurarvutitel kasutatud statistikatarkvara.

²⁵ Peter Doorn, "Research data archives and public electronic record-offices: what can we learn from each other?", *Archives in cyberspace: electronic records in East and West*, ed. by Peter Doorn, Irina Garskova and Heiko Tjalsma (Moscow: Moscow University

Kuna andmearhiivid tegelesid just masinloetava teabe kogumise, töötlemise ja säilitamisega, siis kujunesidki nendes välja vastavad infosüsteemid ja nende kasutamise praktikad. Tegemist on hea näitega sellest, kuidas uue tehnoloogia kasutuselevõtt toimub kiiremini ja ilma põhimõtteliste vastuoludeta süsteemides, kus need luuakse n-ö tühjalt kohalt. Omandatud teadmised ja kogemused ei levinud laiemalt enne 1990. aastaid.

Digiinfo säilitamise paradigma muutus – põhjused ja tagajärjed

Arvutite laiem kasutuselevõtt tööstuses, panganduses ja riigiasutustes algas 1960. aastatel. Neid kasutati palgamaksmisel, inventarinimestike, kaupade transpordi ning varade vastuvõtmise nimekirjade, arvete jms koostamisel. Ilmselt arvutite järjest ulatuslikuma kasutamise tõttu oli 1960. aastate keskpaigaks suhtumine masinloetavatesse dokumentidesse juba selgelt muutunud. Samas tuleb tõdeda, et info säilitamisega tegelevatel organisatsioonidel ja spetsialistidel ei olnud infotehnoloogiatega kujundamisele mingit reaalselt mõju. Nii tehnoloogilised lahendused kui ka info kasutamine reaalses organisatsioonides kujunesid nendest sõltumatult.²⁶

1965. aastal töötati välja juhised, millele tuginedes valida välja püsiva (alalise) väärtusega masinloetavaid dokumente Ameerika Ühendriikide Rahvaloendusbüroo dokumentide hulgast. Tegemist on esimese kavaga masinloetavate dokumentide väärtuse määratlemiseks ja säilitamiseks.²⁷

1960.–70. aastatel oli rõhuasetus endiselt masinloetavatel andmekandjatel kui füüsilistel objektidel²⁸ ning digiinfosse suhtuti ikkagi lähtudes paberdokumentide säilitamisest. Digitaalsete andmete säilitusprobleemi lahendust nähti ennekõike võimalikult pikaajalistel andmekandjate kasutuselevõtmises. Peamiseks probleemiks oli andmekandjate, peamiselt magnetlintide eluiga ja seda mõjutavaid tegureid. Palju tähelepanu pöörati magnetlintide seisundile, säilivusele ja hoitingimustele.²⁹

Press, 2004), 98; Hans-Jørgen Marker, "Data conservation at a traditional data archive", *History and electronic artefacts*, ed. by Edward Higgs (Oxford: Clarendon Press, 1998), 294–303 (296–298).

²⁶ Vt nt Margaret Hedstrom, "Understanding electronic incunabula: a framework for research on electronic records", *The American Archivist*, 54:3 (1991), 334–354 (336).

²⁷ Vt Fishbein, "Appraising information in machine language form", 41.

²⁸ Kasutatakse näiteks terminit "arvuti lindifail" (*computer tape files*), vt nt Charles M. Dollar, "Appraising machine-readable records", *The American Archivist*, 41:4 (1978), 423–430.

²⁹ Vt nt Gerald J. Rosenkrantz, "National archives mass storage requirements – 1975 to 1980", *IEEE Transactions on Magnetics*, 7:4 (1971), 843–847; J. E. Thexton, "Archival

Magnetlintidele oli võimalik kirjutada andmeid korduvalt. See oli ühelt-poolt olulist kokkuhoidu võimaldavaks meetmeks, samas läksid aga ülekirjutamise käigus varsemad andmed kaduma. 1969. aastal anti Ühendriikides välja esimene juhise magnetlintide kogumiseks ja säilitamiseks arhiivis.³⁰ Juhendis olid määratletud peamised tehnilised nõuded magnetlintide vastuvõtmiseks arhiivi (lindi vorming, andmete loetavus), nõuded säilitustingimustele (temperatuur ja õhuniiskus) ning hooldusnõuded. Vastuvõetavatest magnetlintidest valmistati säilitus- ja tagavarakoopiaid. Kooptiaid hoiti turvalisuse kaalutlustel erinevates hoonetes. Tegemist on selliste põhimõtetega, mis on masinloetavate infokandjate säilitamise aluseks tänase päevani. 1973. aastal andis Ameerika Ühendriikide rahvusarhiiv välja juhise magnetlintide säilitamiseks.³¹

1970. aastatel kasutati teabeasutustes arvuteid juba üsna laialdaselt, kuid digitaalse teabe säilitamise probleemidega tegelesid siiski vaid üksikud asutused. 1968. aastal loodi Ameerika Ühendriikide rahvusarhiivis esimene kava arvutidokumentidega (*computerized records*) tegelemiseks (Data Archives Staff). 1970. aastate alguses koostasid mitmed rahvusarhiivid (Ameerika, Suurbritannia, Rootsi, Kanada) juhiseid masinloetavate infokandjate väärtuse määramiseks ja vastuvõtmiseks arhiividesse.³² Esimesed digitaalsed arhivaalid võeti Ameerika Ühendriikide rahvusarhiivi vastu 16. aprillil 1970. Tegemist oli NASA poolt edastatud andmetega, mis salvestati veealuse kosmosesimulatsiooni "Tektite I" käigus.³³

Digitaalsete dokumentide korraldamise meetodid võeti üle pabermaterjalidelt ning kohandati vastavalt uutele tehnilistele nõuetele. Olulisele kohale tõusis dokumentide kasutatavuse nõue. Teiseks küsimuseks oli piisava dokumentatsiooni olemasolu – kas on olemas piisavalt teavet dokumentide töötlemiseks ja kasutamiseks? Kui dokumendid sisaldavad küll olulist teavet, aga neid ei ole võimalik kasutada, siis muutub ka nende säilitamine mõttetuks.

1970. aastatel hakati pöörama tähelepanu sellele, et infokandjate vanemise kõrval on ka seadmete ja programmide pidev muutumine oluliseks

potential of machine-readable records in business", *The American Archivist*, 37:1 (1974), 37–42 (38).

³⁰ *A procedure for accepting digital and analog magnetic tape for archival storage* (Washington DC: NARS, Data Archives Staff, 1969).

³¹ *Recommended environmental conditions and handling procedures for magnetic tape* (Washington DC: NARS, 1973).

³² Harald Naugler, *The archival appraisal of machine-readable records: a RAMP study with guidelines* (Paris: UNESCO, 1984), 3.

³³ Th. E. Brown, "History of NARA's...", *Thirty years of electronic records*, ed. by Bruce I. Ambacher (Lanham, Md.: Scarecrow Press, 2003), 1–23 (1).

ohuks digiinfo säilitamisele.³⁴ Säilitatavate objektide füüsiliste omaduste kõrval muutus oluliseks nende loomise ja kasutamise kontekst. Erinevate riist- ja tarkvarasüsteemide probleemi lahendamiseks muutis Ameerika Ühendriikide rahvusarhiiv alguses arhiivi vastuvõetava digitaalse dokumendi riist- ja tarkvarast sõltumatusse vormingusse. Alates 1976. aastast aga nõuti, et arhiivimoodustajad esitaksid failid juba ise vastavas vormingus.³⁵

Lisaks erinevatele magnetkandjatele, mis olid peamisteks infosalvestusvahenditeks, ilmusid 1970. aastate keskel kasutusse ka esimesed optilised andmekandjad. Hakati arutama nende kasutusvõimalusi info säilitamisel.³⁶ Laiemalt tulid optilised andmekandjad kasutusele siiski märksa hiljem – 1980. aastatel.

1980. aastatel olid arvutid erinevates teabeasutustes juba laialt levinud. Algselt kasutusel olnud suurarvutid (*mainframe*) asendusid küllaltki kiiresti personaalarvutitega.³⁷ See tõi kaasa digiinfo hulga kiire suurenemise, veelgi tõsisemaks probleemiks kujunes selle info mitmekesisus. Andmete kõrval hakati tähelepanu pöörama ka tekstilistele digitaalsetele dokumentidele ning geograafilistele infosüsteemidele (GIS). Tekstidokumentide hulga suurenemine ja selle tõttu ka nendele tähelepanu pööramine on seotud personaalarvutite laialdase kasutuselevõtuga töökohtades. Probleemide hulka kerkis relatsiooniliste andmebaaside säilitamine. 1980. aastatel, kui digitaalselt hakati looma lisaks varasematele andmetele ka igasugust muud infot (tekstid, pildid, heli, filmid jms), muutus eristus digitaalse ja muu info vahel ebaoluliseks. Digitaalne tehnoloogia võimaldas töödelda ja säilitada igasugust infot, sõltumata selle liigist.

Personaalarvutite levik tähendas ka seda, et erinevad asutused võtsid kasutusele erinevaid andmete haldamise süsteeme. 1980. aastate lõpus oli suur osa digiteabest personaliseeritud, mittestandardsetes andmesüsteemides.³⁸

³⁴ Thexton, "Archival potential of machine-readable records in business", 38.

³⁵ L. J. Henry, "Appraisal of electronic records", *Thirty years of electronic records*, ed. by Bruce I. Ambacher (Lanham, Md.: Scarecrow Press, 2003), 25–41 (35).

³⁶ Vt nt Sam Kula, "Optical memories: archival storage system of the future, or More pie in the sky", *Archivaria*, 4 (1977), 43–48.

³⁷ Vt nt Richard M. Kesner, Don Hurst, "Microcomputer applications in archives: a study in progress", *Archivaria*, 12 (1981), 3–19; Michael Cook, *Archives and the computer* (London, Boston: Butterworths, 1980).

³⁸ Grant Blank, Karsten Boye Rasmussen, "The data documentation initiative: the value and significance of a Worldwide Standard", *Social Science Computer Review*, 22:3 (Fall 2004), 307–318 (309).

Mõiste “masinloetavad dokumendid” hakkas asenduma elektrooniliste või digitaalsete dokumentide mõistega.³⁹ Termin “elektroonilised dokumendid” võeti kasutusele 1990. aastate alguses ning see asendas senini kasutusel olnud “masinloetavaid dokumente”, tähistades elektronarvutite poolt loodud ja nende abil kasutatavaid dokumente. Muutus näitas seda, et senise ikkagi vaid üksikute masinloetavate dokumentide käsitlemise asemel hakati rõhutama kogu traditsioonilise meedia üha suuremat muutumist elektrooniliseks.

1980. aastate esimesel poolel levisid ideed, et kuna digitaalne info on hävimisohus füüsilise kandjate lagunemise, ebapiisava metaandmestiku ning vananeva riist- ja tarkvara tõttu, tuleb säilitamiseks kanda teave mikrofilmile (*computer output microfilm* ehk COM).⁴⁰ Selline lähenemine masinloetavate infokandjate säilitamisele lükati küll kohe tagasi, kuna oli selge, et masinloetavate infokandjate olulisus ja väärtus seisnebki selles, et need on masinloetavad. Kui neil säilitatav info ei ole enam masinkasutatav, siis kaotab see teave olulise osa oma väärtusest. Selline säilitamise idee tugines masinloetava andmekandja kui materiaalse objekti käsitlusele. Säilitamiseks valiti mikrofilm, kuna tollel ajal loeti pikaajaliseks säilitamiseks sobivateks materjalideks ainult paberit ja mikrofilmi.⁴¹

1990. aastate alguseks olid ainult suuremates teabeasutustes olema vastavad programmid elektrooniliste dokumentide vastuvõtmiseks ja säilitamiseks. Näiteks võib siin tuua Ameerika Ühendriikide ja Kanada rahvusarhiivid. Teistes asutustes aga alles tegeleti selliste kavade väljatöötamisega.⁴² Suurbritannia rahvusarhiivis alustati elektrooniliste dokumentide säilitamisega seotud programmide väljatöötamist 1990. aastate keskel.⁴³ Põhja-maadest oli digitaalse info säilitamine seadusandlikult reguleeritud kõige varem Rootsis ja seda juba 1970. aastatel. Soomes ja Norras kehtisid vastavad regulatsioonid 1980. aastatest. Teistes Euroopa riikides toimus vastavate regulatsioonide väljatöötamine valdavalt 1990. aastatel. Prantsusmaal oli vastav seadusandlus olemas juba 1978. aastast, kuid tegelikkuses

³⁹ Meyer H. Fishbein, “Recollections...”, *Thirty years of electronic records*, ed. by Bruce I. Ambacher (Lanham, Md.: Scarecrow Press, 2003), xiii–xix (xviii).

⁴⁰ Kõige mõjukamaks selleteemaliseks arvamuseavaldajaks oli Ameerika Ühendriikide Ameerika Ühendriikide rahvusarhiivi säilituskomitee, mis esitas 1984. aastal vastavateemalise aruande “Strategic technology consideration relative to the preservation and storage of human and machine readable records”, vt J. C. Mallinson, “Preserving machine-readable archival records for the millennia”, *Archivaria*, 22 (Summer 1986), 147–152.

⁴¹ Mallinson, “Preserving machine-readable archival records for the millennia”, 147.

⁴² Terry Cook, “Easy to byte, harder to chew: the second generation of electronic records archives”, *Archivaria*, 33 (1986), 202–216 (202).

⁴³ Sleemann, “It’s public knowledge”, 174.

ei rakendunud see enne 1982. aastat, mil hakati Prantsuse rahvusarhiivi kaasaegsete arhiivide keskusesse võtma esimesi magnetlinte digitaalsete andmetega.

Samas hakkas 1980. aastate algupoolel selguma tõsiasi, et andmekandjate füüsiline eluiga on digitaalse informatsiooni säilitamise seisukohalt küllaltki ebaoluline tegur. Peamiseks probleemiks hakati pidama hoopiski andmekandjate kasutamiseks ettenähtud seadmete kättesaadavuse tagamist. Lisaks tehniliste seadmete vananemisele rõhutati ka erinevate andmevormingute olemasolu ja küllaltki kiiret vaheldumist.⁴⁴ Tõdeti, et oluline on käsitleda infosüsteeme terviklikena ja teha kindlaks, millistes süsteemide osadest tuleks digitaalset teavet koguda ja säilitada. Sellega seoses räägitakse isegi “meediumi türanniast vabanemisest”, st säilitamisel on keskseks informatsioon, mitte selle erinevad füüsilised kandjad.⁴⁵ Elektrooniliste dokumentide pikaajalise säilimine tagatakse nende pideva kopeerimisega uutele andmekandjatele.⁴⁶ Tegemist on vägagi olulise kontseptuaalse nihkega kogu säilitamisvaldkonnas. Tunnistatakse asjaolu, et informatsioon on seotud tervete infosüsteemide tööga, leidub tervetes süsteemides ja seda, et digitaalne info on oma konkreetse füüsilise kandjaga seotud märksa vähem võrreldes n-ö klassikaliste infokandjatega, nagu raamatud, dokumendid, fotod jms. Digitaalse teabe säilitamine seda vastavalt andmekandjate või kasutussüsteemide vananemisele ümber kopeerides leidis esmakordselt fikseerimist Briti standardis 1988. aastal.⁴⁷

1990. aastate alguseks oli tehnoloogiate vananemine muutunud digisäilitamise keskseks teemaks.⁴⁸ Säilitamise paradigma nihkus selgelt infokandjate säilitamiselt (mis iseloomustab n-ö klassikaliste materjalide säilitamist) informatsiooni pikaajalise kättesaadavuse tagamisele (*maintaining the accessibility of digital records over time*). Digiinfo pikaajalise kättesaadavuse tagamine on seotud info loetavusega (*readability*), kasutatavusega (*retrievability*) ja arusaadavusega (*intelligibility*). Loetavus tähendab seda, et info on kasutatav ka teiste kui selle loomiseks ja hetkel säilitamiseks kasutatavate arvutisüsteemide poolt. Kasutatavus tähendab seda, et vastavad

⁴⁴ Mallinson, “Preserving machine-readable archival records for the millennia”, 148–149, 151.

⁴⁵ Dorothy Ahlgren, John McDonald, “The archival management of a Geographic Information System”, *Archivaria*, 13 (1981/82), 59–65 (63–64).

⁴⁶ Sue Gavrel, “Preserving machine-readable archival records: a reply to John Mallinson”, *Archivaria*, 22 (1986), 153–155 (154).

⁴⁷ British Standard BS4783, Part 2 (1988).

⁴⁸ Vt nt Charles M. Dollar, “Archivists and records managers in the information age”, *Archivaria*, 36 (1993), 37–52 (45); *Structured glossary of technical terms. The impact of digital technologies*, <<http://www.clir.org/pubs/reports/lynn/impact.html>> (19.11.2014).

failid on võimalik identifitseerida ja olemasoleva tarkvara abil töödelda. Arusaadavus viitab sellele, et teave on kasutajatele arusaadav.⁴⁹ Alles mõne aja möödudes saadi aru, et digitaalne säilitamine ei seisne mitte niivõrd andmekandjate füüsilises säilitamises, vaid digitaalse info vastavas loomises ja säilitamises. Tegemist on ühelt poolt küll tehnoloogilise, aga teisalt ka sotsiaalse fenomeniga, seega peituvad ka lahendused nii tehnoloogias kui ka infokorralduses laiemalt.⁵⁰

Jõuti arusaamisele, et oluline on kogu säilitamise süsteemi loomine. Tuleb kindlaks määrata, millised institutsioonid hakkaksid vastutama digitaalse info säilitamise eest, luua vastav seadusandlik keskkond. Tehnilised probleemid on tegelikult teisejärgulised. Informatsiooni terviklikkuse ja kasutatavuse tagamine on märksa olulisemad küsimused võrreldes konkreetsete andmekandjate vastupidavuse ja elueaga. Juba 1980. aastate teisest poolest hakati looma infotehnoloogilisi standardeid, mille laiem kasutuselevõtt oli aluseks digitaalse teabe kogumiseks, säilitamiseks ja kasutamiseks.⁵¹ Oluliseks hakati pidama digitaalsete objektide loomise ajal võtta arvesse nende säilitamise vajadust. 1996. aastal ilmus raport "Preserving digital information", kus rõhutati asjaolu, et digitaalse säilitamise eesmärgiks on infoobjektide terviklikkuse (*integrity*) säilitamine. Selleks on vajalik määratleda nende sisu (*content*), püsivus (*fixity*), seosed (*reference*), päritolu (*provenance*) ja kontekst (*context*).⁵²

1990. aastate alguses aduti tõsiasi, et digitaalse info edukaks säilitamiseks peavad säilitamisega tegelejad mõjutama infotehnoloogia valmistasid ja kasutajaid, et need arvestaksid tehnoloogiate loomisel ja nende

⁴⁹ Dollar, "Archivists and records managers in the information age", 45–46.

⁵⁰ Hedstrom, "Understanding electronic incunabula", 338.

⁵¹ Nt U.S. National Archives and Records Administration, Archival Research and Evaluation Staff, *A National Archives strategy for the development of standards for the creation, transfer, access, and long-term storage of electronic records of the Federal Government*, National Archives Technical Information Paper, 8 (June 1990); Margaret H. Law, Bruce K. Rosen, *Framework and policy recommendations for the exchange and preservation of electronic records* [report prepared by the National Computer Science Laboratory, National Institute of Standards and Technology for the National Archives and Records Administration] (March 1989); Canadian Bureau of Management Consulting, *Data and document interchange standards and the National Archive* [Project No. 1-6465] (Ottawa, June 1987); Protocols Standards and Communications, Inc., *The application of ODA/ODIF standards* [Prepared for the National Archives of Canada] (Ottawa, 1989); and Protocols Standards and Communications, Inc., *Application portability* [prepared for the National Archives of Canada] (Ottawa, 28 December 1989).

⁵² Donald Waters, John Garrett, *Preserving digital information: final report of the task force on archiving of digital information* (The Commission on Preservation and Access and The Research Libraries Group, 1996).

kasutamisel digiinfo säilitamise vajadusi.⁵³ Sellisel juhul saaks mõjutada säilitamist juba digiobjektide loomise ajal, mitte alles tagantjärele.

1990. aastate lõpus tõusis digisäilitamise keskpunkti ka metaandmete küsimus.⁵⁴ 1990. aastate keskpaigast muutusid oluliseks ka digitaalse info metaandmetega seotud teemad.⁵⁵ Esimene säilitusmetaandmete süsteem loodi Austraalia rahvusraamatukogus 1990. aastate lõpus.⁵⁶ Sellele järgnesid mitmed teised metaandmete süsteemid.

Digitaalsete dokumentide säilitamine jäi tavaliselt arhiivide ülesandeks. Mõnes riigis aga otsustati anda digitaalsed dokumendid lepingu alusel üle arvutuskeskustele. Nii näiteks sõlmis Soome rahvusarhiiv 1987. aastal lepingu Soome riikliku arvutuskeskusega, kus digitaalsed dokumendid võeti lühiajalisele säilitamisele. 1996. aastal võttis rahvusarhiiv õigused, mis on seotud avaliku sektori digitaalsete dokumentide säilitamisega, taas endale. Soome ja Islandi rahvusarhiivid sõlmisid 1995. aastal lepingud vastavate arvutuskeskustega digitaalsete dokumentide säilitamiseks, aga selline lahendus ei osutunud jätkusuutlikuks, kuna arvutuskeskused vastasid küll tehnilistele, aga mitte arhiivinduslikele kriteeriumidele digitaalsete dokumentide säilitamisel.⁵⁷ Suurbritannias tegeleb valitsusasutuste andmekogude säilitamisega alates 1997. aastast vastavalt rahvusarhiiviga sõlmitud lepingule Londoni ülikooli arvutuskeskus, kuhu loodi 1998. aastal The National Digital Archive of Datasets (NDAD). Austraalia rahvusarhiiv otsustas 1995. aastal digitaalseid dokumente mitte arhiivi üle võtta ja jätta need asutustesse, kus nad olid loodud. Selline “mitte-üleandmise strateegia” (*distributed custody, non-custody practice, post-custodial*) tähendas seda, et arhiiv säilitas vaid intellektuaalse kontrolli digitaalsete dokumentide üle. Pärast dokumentide hindamist ja neile säilitustähtaja määramist jäid nad edasi neid loonud asutusse ja arhiiv pakkus nende säilitamiseks ja kasutamiseks vaid nõustamisabi. Selline praktika lõpetati

⁵³ Vt nt Hedstrom, “Understanding electronic incunabula”, 337.

⁵⁴ Vt nt Susan S. Lazinger, *Digital preservation and metadata: history, theory, practice* (Libraries Unlimited, 2001).

⁵⁵ Michael Day, “Preservation metadata”, *Metadata applications and management*, ed. by G. E. Gorman, Daniel G. Dorner, International Yearbook of Library and Information Management 2003–2004 (London: Facet Publishing, 2004), 253–273.

⁵⁶ M. Phillips, D. Woodyard, K. Bradley, C. Webb, *Preservation metadata for digital collections: exposure draft* (National Library of Australia, 2001).

⁵⁷ Matti Pulkinen, Tom Quinlan, “Nordic archives and electronic records: preservation of electronic records in Nordic Countries”, *For the record: data archives, electronic records, access to information and the needs of the research community*, ed. by Rena Lohan et al. (Dublin: Institute of Public Administration, 1996), 49.

2000. aasta kevadel ja sellest ajast alates võeti digitaalsed arhivaalid säilitamiseks Austraalia rahvusarhiivi.

1990. aastatel teatasid paljud arhiivid ja muud asutused, et nad suudavad säilitada digitaalset teavet. See kutsus esile ka skeptitsismi ja nõude kehtestada standardid, mis võimaldaksid hinnata asutuste võimet digitaalse info pikaajaliseks säilitamiseks.⁵⁸

Raamatukogudes seostus digitaalse teabe säilitamine esmalt elektrooniliste ajakirjadega. Paljud raamatukogud, eriti need, mis tegutsesid ülikoolide ja uurimisasutuste juures, hakkasid üha enam paberkanajatel ajakirjade asemel tellima nende elektroonilisi versioone. Nende materjalidele pikaajalise ligipääsu tagamine sõltub aga muutuvast tehnoloogiast ja välistest organisatsioonidest. Need probleemid olid aga hoopiski erinevad traditsioonilistest raamatukogu puudutavatest probleemidest. Peagi lisanud ka muude digitaalsete ressursside säilitamise probleemid. 1980 aastatel hakati digitaliseerima ja internetis kättesaadavaks tegema üksikuid kollektsioone. 1990. aastatel algatati paljudes teabeasutustes suuremahulisi digiteerimisprojekte, mille eesmärgiks oli võimalikult suure hulga teabe kättesaadavaks tegemine. Euroopas seostus see näiteks e-Euroopa agenda.⁵⁹ Projektid olid enamuses omavahel koordineerimata ning nende läbiviimisel kasutati erinevaid tehnilisi ja organisatsioonilisi lähenemisi. Lähenemiste ühtlustamiseks alustati rahvusvahelisi projekte. 4. aprillil 2001. aastal kohtusid Rootsis Lundis (Rootsi eesistumisajal) Euroopa Nõukogu ja Euroopa Liidu liikmesriikide esindajad ja eksperdid, et arutada, kuidas koordineerida ja väärtustada rahvuslikke digiteerimisprogramme Euroopa tasandil. Kohtumise tulemusena avaldati üldpõhimõtete kogum avalike digiteerimisalaste ettevõtmiste juhtimise ja koordineerimise kohta, mida tuntakse Lundi põhimõtetena.⁶⁰

Kokkuvõte

Digisäilitamise arengus võib täheldada paralleele muude objektide säilitamisega. Esialgu keskenduti digiinfot kandvatele objektidele ning nende säilivusele. Samuti tekitas uuel kujul esitatud informatsiooni säilitamine esialgu kahtlusi, kaheldi selle võimalikkuses. Uute meediate suhtes ollakse

⁵⁸ Vt Bruce Ambacher, *The development of a standard for digital repository certification*, <<http://www.ukoln.ac.uk/events/pv-2005/pv-2005-final-papers/016.pdf>> (19.11.2014).

⁵⁹ e-Europe Action Plan, <http://ec.europa.eu/information_society/eeurope/i2010/archive/eeurope/index_en.htm> (19.11.2014).

⁶⁰ Vt <ftp://ftp.cordis.europa.eu/pub/ist/docs/digicult/lund_principles-en.pdf> (19.11.2014).

teabeasutustes alati üsna konservatiivselt meelesstatud ning nende säilitamisel lähtutakse kord juba väljakujunenud praktikatest. Situatsioon on võrreldav näiteks paberi kasutuselevõtuga. Pika aja kestel peeti seda märksa halvemaks materjaliks võrreldes pärgamendiga. Tegemist oli materjaliga, mille säilivuse kohta puudus teave ja seega suhtuti sellesse õigustatud ettevaatusega. Sarnaselt olid paljud säilitamisega tegevad spetsialistid kuni 1970. aastate keskpaigani veendunud, et elektrooniline teave sobib säilitamiseks ainult siis, kui selle sisu on kantud paberile või mikrofilmile. Olu-kord oli erinev nendes asutustes – andmearhiivides –, kus tegeleti algusest peale vaid masinloetavate infokandjatega. Kuna andmearhiivid tegelesidki ainult masinloetava teabe kogumise, töötlemise ja säilitamisega, siis kujunesid vastavad infosüsteemid ja nende kasutamise praktikad nendes välja kiiremini ja ilma põhimõtetlike vastuoludeta, kuna tekkisid n-ö tühjale kohale. Hiljem saadi aru, et on vajalik kogu süsteemi ümberkorraldamine. Nii näiteks tuleb teha otsus digiinfo säilitamise kohta selle elutsükli märksa varasemas etapis võrreldes paberkandjal infoga.

Digitaalse info loomise, haldamise ja kasutamisega on ühiskonnas seotud terve rida institutsioone – raamatukogud, arhiivid, muuseumid, riis-tavara ja tarkvaratootjad, teadusasutused, valitsusasutused jne. Kõikide nende institutsioonide konkreetsed huvid ja tegevusvaldkonnad on vägagi erinevad ning nende ühitamine ja koordineerimine pole lihtne. Alguses proovisid erinevad institutsioonid reguleerida digisäilitamist oma konkreetsemas valdkonnas. Hiljem hakati välja töötama rahvuslikke ja rahvusvahelisi kavasisid. Kõik need institutsioonid tegelevad väljakujunenud õiguslikus, kultuurilises ja sotsiaalses keskkonnas, mille kiire muutmine ei ole võimalik. Samas sõltub ühiskonnale olulise digiteabe säilitamine erinevate institutsioonide ja üksikisikute koostöös kujunevast infosüsteemist.

ABSTRACT: How to start preserving bits and bytes in addition to paper? Development of preserving digital information

Digital communication is the foundation of the contemporary social information system. One of the functions of social information systems is the

preservation of information and guaranteeing its existence and usability. Preservation could hereby be defined as a function that provides continuance to the information system. On the one hand, preservation of digitally presented information is related to technological developments, however, on the other hand, it also relates to social and cultural processes. Whereas the technical aspect of preservation has been widely dealt with, then its relation to wider social processes has been neglected. Also, parallels can be drawn between digital preservations and preservation of other objects. First, the focus was on objects carrying digital information and their preservability. Initially there were also doubts regarding the preservation of information presented in a new way. Many preservation specialists were convinced up to the mid-1970s that electronic information was suitable for preservation only if its content had been saved in paper or microfilm. Information establishments are always quite conservatively-minded about new mediums and proceed from previously established practices when preserving them. The situation was different in the establishments which dealt with computer-readable information carriers, i.e. in data archives. As data archives only focused on collecting, processing and preserving computer-readable information, then it was there where corresponding information systems and the practices for using them developed. It is a good example of how a new technology is implemented faster and without any principal contradictions in the systems where it is founded from the scratch. The knowledge and experience gathered in data archives did not start to spread until the 1990s. At the beginning of the 1980s, it became clear that the physical life span of data carriers is a rather insignificant factor from the point of view of preserving digital information. What was perceived as the main problem, however, was ensuring the availability of devices needed to use the data carriers. In addition to the aging of technical devices, the existence of different data search formats and their relatively fast alteration were emphasised. It was understood that information system should be treated as a whole and it needed to be made sure from which parts of the system digital data should be collected from and preserved. In this context, it was even discussed as “liberation from the tyranny of the medium”; i.e. in preservation, what becomes the focus is the information itself, not its different physical carriers. It is a very significant conceptual shift in the entire field of preservation. It is recognised that information is related to the functioning of all information systems, it can be found in all systems, and digital information is considerably less connected to its specific physical carrier, in comparison with the so-called classical information carriers, such as

books, documents, photos, etc. In case of digital information objects, the relation between the carrier (medium) and information itself is notably more ephemeral. Information presented in digital format is dependent on an entire chain of encoding and decoding stages and the information presented to a person is recreated at the moment of its presentation. Digital computing and communication technologies have developed rapidly and for the time being, problems related to preserving information have not played an important role in this development. There are several institutions in the society engaged in creating, managing and using digital information: libraries, archives, museums, hardware and software manufacturers, research institutions, government institutions etc. All these establishments have very different specific interests and fields of specialty, which makes it very difficult to match and coordinate them. At the same time, the preservation of digital information relevant for the society is dependent on the information system which is formed in the course of cooperation between different institutions and individuals.

KEYWORDS: information systems, digital information, digital preservation, history of digital preservation.

KURMO KONSA (b. 1965) is Associate Professor at the Department of Archival Studies at the University of Tartu, and Professor of Conservation at Tartu Art College.*

* Correspondence: Institute of History and Archaeology, University of Tartu, Ülikooli 18, 50090 Tartu, Estonia. E-mail: Kurmo.Konsa@ut.ee