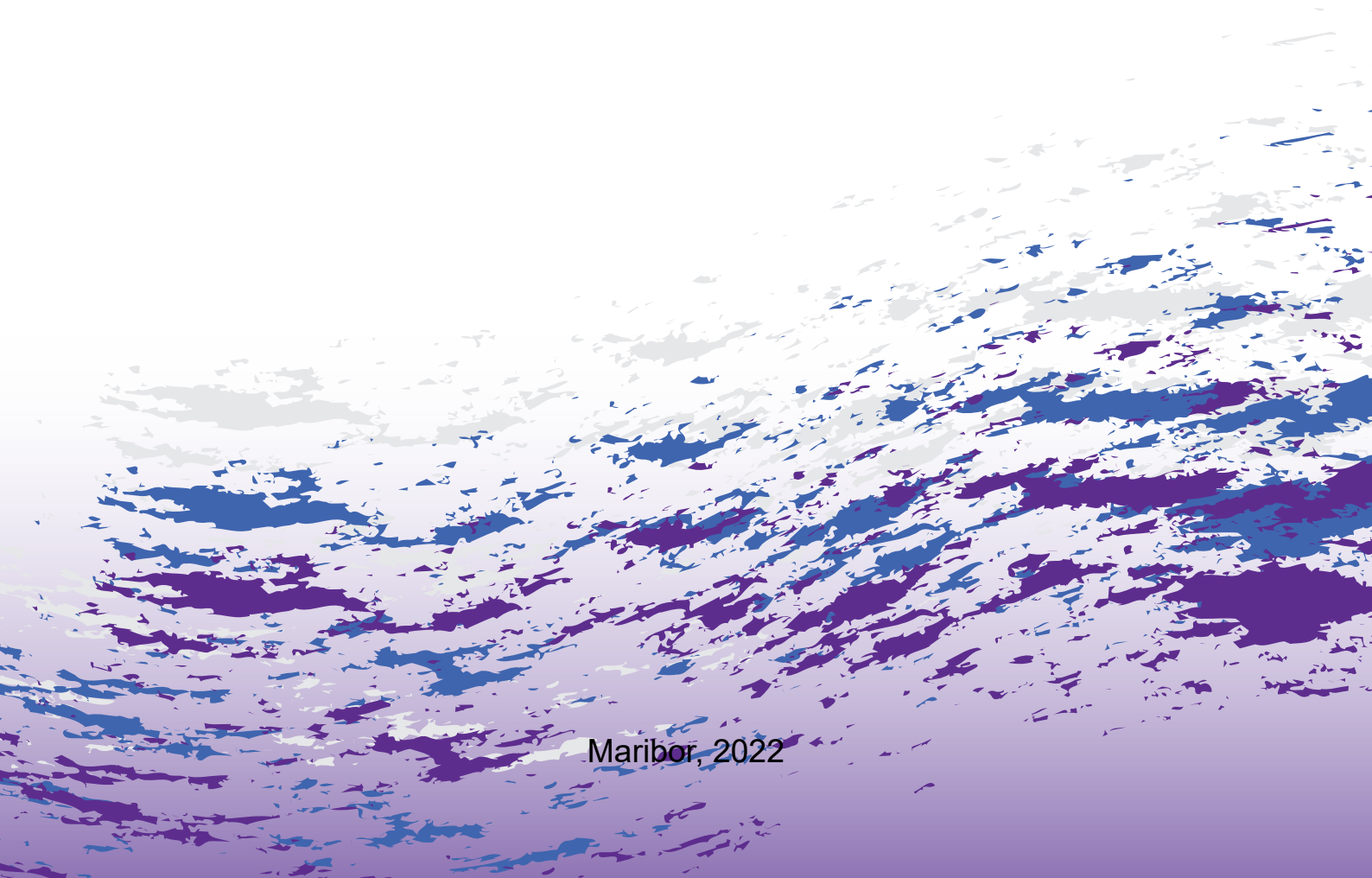
PAM Pokrajinski arhiv Maribor

Moderna arhivistika

Časopis arhivske teorije in prakse
Journal of Archival Theory and Practice

Letnik 5 (2022), št. 2 / Year 5 (2022), No. 2



Maribor, 2022

Prejeto / Received: 28. 06. 2022

1.03 Kratki znanstveni članek *1.03 Short scientific article*

https://doi.org/10.54356/MA/2022/IZAE9112

SOME CONSIDERATIONS ON THE ARCHIVAL STORAGE IN DIGITAL PRESERVATION OF RECORDS

Bogdan Florin POPOVICI, Ph. D.
National Archives of Romania, Romania
Archival science assistant professor, Alma Mater Europaea, ECM Maribor, Slovenia
bogdan.popovici@arhivelenationale.ro

Abstract:

Starting from the assumption that the carrier is no longer an intrinsic part of digital record (though still a necessary presence for the record to exist), the paper seeks to critically examine the role of storage in digital preservation of records. Is the carrier/storage area a topic of concern for archivists or it is rather a topic for IT professionals? Is longevity of carrier an important archival topic? Is the size of storage an important pre-requisite for a digital preservation system? The examination of these questions will consider both theoretical approaches and personal experiences of the author.

Key words:

digital preservation, storage system, digital medium

Izvleček:

Razmišljanja o arhivski hrambi pri digitalnem varovanju zapisov

Začenši s stališča, da medij ni več bistven del digitalnega zapisa (čeprav še vedno obvezen za obstoj zapisa), želi prispevek kritično raziskati vlogo hrambe pri digitalnem varovanju zapisov. Ali je medij/hramba še stvar obravnave za arhiviste ali spada k strokovnjakom informacijske tehnologije? Ali je trajnost medija pomembna arhivska tema? Ali je kapaciteta hrambe pomemben predpogoj za sistem digitalnega varovanja? Raziskovanje teh vprašanje bo zavzelo tako teoretične pristope kot tudi avtorjeve izkušnje.

Ključne besede:

Digitalna hramba, sistem hrambe, digitalni medij

1. The carrier – an important topic for archivists...

The carrier is one of the components for a document to exist. In a definition from 1988, a document was even defined as information on a medium (Peter 1988, 56)[1]. In classical Diplomatics, the carrier was part of the diplomatic analysis to determine the authenticity of records (Duranti 1998, 140); its characteristics served for understanding and authenticating the record. In archival science, preserving documents on paper equivalates, in many regards, the preservation of the carrier; the destruction of the carrier meant the destruction of the record itself. This made the interest for the carrier intrinsic to the profession of archivists and made understandable the interest of archivists to promote standards or pieces of legislation that impose certain parameters of quality for records, especially permanent records.

In the digital realm though, a shift from the traditional regard over records occurred. Luciana Duranti, in an influential article promoting the result of Interpares II project, admitted that the carrier is essential for the records to exist, but it is no longer a constituent part of it (Duranti and Thibodeau 2006, 18). That is, a content in a Word file cannot be considered as a document until it is saved (that is, inscribed on a carrier, the harddisk), but, at the same time, the change of carrier (moving that saved file on a USB key) does not alter the characteristics of the document.

Despite such change of perspective, the records carrier in digital world is still of great interest. I would like to start from reminding that storage is one of the functions of an archive in OAIS model. My course for digital archives includes a chapter about digital carriers. Some years ago, during IIAS Autumn school, Stefano Alegrezza gave a very interesting lecture about digital carriers (Alegrezza 2016). Commercially, a permanent challenge is to manufacture carriers with longer span dates. There are offers of optical discs with tested long-time durability (Millenial Disc (M-DISC n.d.) or Optical Disk Archive (Optical Disc Archive n.d.)); also, films are suggested as fitted for digital long term preservation (CPC London n.d.) (Piql n.d.), ceramics (What is MoM n.d.), DNA ((DNA-is-the-future-of-data-storage n.d.)), not to mention already classical magnetic tapes (Linear Tape Open n.d.).

2. …or not?

Despite all these facts, in the latest years I have concluded that, in many regards, storage for digital records should not be treated as a topic situated firmly within the realm of archival profession, except for the cases when the carrier has, in itself, a preservation value (Graves 2013). And while this may sound paradoxically, I shall try in the following to make a case supporting this idea.

Digital carriers serve several purposes. The main one, is, of course, to record digital data and provide their persistency. This is the mission of computers' HDD. A second purpose is to make information circulate. A USB key, a DVD serves this purpose. But it needs to make a difference between 1990s or 2000s, when the computer was at the core of digital activity and floppies and CDs were the rule, and present times, when technology emphasises systems, the network, the clouds. The slow removal of drives for floppies and, today, even for optical disks highlights this trend. How many users still transfer information via removable disks? Or more precisely, how many users rely on removable carriers to store the information in the long run? In my opinion, there are very

[1] *There is a common overlapping of meaning between medium and carrier, in the professional literature in English language. In this paper it the terms will be considered as equivalent, though there might be differences in other contexts.*

few. We can have the case of digital publications published on optical disks, or personal desire to have one more backup or convenient copies of working files, where the reliability/availability of the implemented systems may be a concern. But in big enterprises, in administrations—the network is a rule, and large enterprise level storage stacks are at the core of system design. And, anyway, as a best practice, except for optical disks, nobody would recommend a long term preservation on USB keys or even on external harddrives (Portable media n.d.).

While creators of records may not be very concerned about the durability (focusing on usability in the first place), organizations should obviously focus on having the preservation of digital records as a core duty. And while I have heard more than once a sole concern, the amount of space available ("How many pentabytes you host?"), I think the topic deserves a finer exploration.

As a matter of principle, digital records should not exist in a single copy. It is not only a security of information rule, but it is also designed in OAIS model, which makes a difference between the preservation copy (AIP) and research copy (DIP). Nothing new here, it was the same with microfilms. So, assuming someone will receive a removable disk in an accession (external HDDs, co mpact disks), it is very likely those disks will be at least duplicated, if not copied/migrated to different carrier types.

The use of portable media as archival carriers brings in several major disadvantages. Except for those devices connected through USB ports, all other disks are subject to severe risk, due to technological obsolescence. No matter how good the state of disk and information are on disk, if the operating system and the hardware will not support the drivers and the drive suited for that device, the information will be locked on disks, without the possibility of retrieval and read. But this is not the only technological intermediation, because handling large amounts of optical disks, for instance, may be cumbersome, so a jukebox may be employed. And here the issue of hardware and software obsolescence duplicates the initial problem, not to mention the supplementary issue of potential mechanical problems which, again, may hamper the access to the information stored. For these and other reasons (high dependency of hardware and software, small amount of data stored, rather high access time, fragility of the carrier), it would not be advisable to preserve electronic archives on portable media.

Alternative to such method of storing would be large disk storage systems. These can be easily scalable, by adding new drives when needed. Storage systems can be configured in various models of redundancy (RAID), ensuring bit level reliability. And, not least, such systems have a direct and higher speed access.

The disks-based storage systems have, however, their downsides. Direct access to digital records may expose them to various types of malwares, like ransomware, which may encrypt the whole storage and make the whole digital archive captive. Such a case is unlikely, if types of cold storage are used, that is, disconnected devices, like tapes or offline disks. Then, storage systems, at their turn, are subject to obsolescence. From the protocol of communication (like Samba) to the compatibility of various components (like controllers), storage systems need a way to be upgraded or replaced periodically. (In this regard, acquiring a large storage space at a given moment, without having a clear and realistic plan of using that capacity, means to make big investments in hardware that will be replaced in 5-10 years without having been used once.) Another downside of disk storage is that it is permanently plugged in, so the amount of electricity consumed and the need for cooling systems are high and these may raise ecological problems.

All these arguments above may seem confusing, since they do not lead to a clear conclusion about the best system. More, terminology and the argumentation is not necessary familiar to archivists, being technical terms specific to IT industry. But, in fact, this is precisely the conclusion of my arguments: due to its specialisation, storage of electronic records is no longer a field for archivists, but for the IT departments, which should offer the technical support for archivists in their activity. As in the case of microfilms, the archivists were not compelled to know how to develop a film and how to produce a positive from a negative – these were the tasks of photo laboratories.In a similar way, today, recordkeeping professionals should not be concerned with span dates of carriers or the solution for storage. This is, and should be, the task of technicians, of those specialists in hardware/software that are up to date with the latest developments in the industry and have the task of projecting, implementing and maintaining a safe storage architecture for digital objects. And they likely know better: for instance, while the archival texts about storage invoke frequently the famous Krogh's "3-2-1 rule", modern technologies are already overpassing those requirements by their performances and configurations (321-backup-rule n.d.).

3. Instead of a closing line…

I do believe that the archival profession has many challenges today with electronic records. From preserving the legal value of records until keeping them accessible in a fast-changing environment—these are large topics that need our attention. Our concern should be on logical facets of digital objects, not in designing storage systems and containers – read carriers—for them. I think it would be wise to allow IT staff to help us with their share of competences. Their input would be better than ours, probably.

4. References

"*321-backup-rule.* n.d. https://www.veeam.com/blog/321-backup-rule.html.

Alegrezza, Stefano. "Data security and reliability of digital storage." Trieste, 2016.

CPC London. n.d. https://www.cpclondon.com/.

"DNA-is-the-future-of-data-storage." *www.nordic-backup.com.* n.d. https://nordic-backup.com/blog/dna-is-the-future-of-data-storage/.

Duranti, Luciana. *Diplomatics. New Uses for an Old Science.* 1998.

Duranti, Luciana, and Kenneth Thibodeau. "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES." *Archival Science* 6, no. 1 (2006): 13–68.

Graves, Michael W. *Digital Archaeology: The Art and Science of Digital Forensics .* 2013.

Linear Tape Open. n.d. https://en.wikipedia.org/wiki/Linear_Tape-Open.

M-DISC. n.d. https://en.wikipedia.org/wiki/M-DISC.

Optical Disc Archive. n.d. https://en.wikipedia.org/wiki/Optical_Disc_Archive.

Peter, Walne. *Dictionary of Archival Terminology.* Paris, 1988.

Piql. n.d. https://www.piql.com/.

"Portable media." *www.dpconline.org.* n.d. https://www.dpconline.org/digipres/champion-digital-preservation/bit-list/portable-media.

"What is MoM." *Memory of the Mankind.* n.d. https://www.memory-of-mankind.com/what-is-mom/.

POVZETEK

RAZMIŠLJANJA O ARHIVSKI HRAMBI PRI DIGITALNEM OHRANJANJU ZAPISOV

Dr. Bogdan Florin POPOVICI
Romunski državni arhiv, Oddelek v Brasovu, Romunija
Docent programa Arhivska znanost, Alma Mater Europaea, ECM Maribor, Slovenija
bogdan.popovici@arhivelenationale.ro

Nosilec je eden ob nujnih pogojev, da document obstaja, tako iz stališča diplomatike in materialnega varovanja. V digitalnem svetu pa je nosilec še vedno potreben za obstoj dokumenta, ni pa več nujna komponenta. Standardizacija nosilcev je zmanjšala njihovo pomembnost pri ocenjevanju enkratnih lastnosti dokumenta. Te okoliščine poudarjajo, da je dokument enak, ne glede na nosilec.

Glede na očitno kontradiktornost tega argumenta, avtor preglejuje različne digitalne nosilce, of magnetnih in optičnih diskov, USB ključev do velikih profesionalnih sistemov za hrambo, ki so sestavljeni iz trdih diskov ali trakov. Vsak opisan sistem ima svoje slabosti, najbolj očitna je seveda nevarnost tehnološke zastarelosti. V takšnih primerih je Edina rešitev migracija, torej premik zapisov v nov sistem digitalne hrambe.

Iz avtorjevih zaključkov izhaja, da digitalni nosilec prazaprav ni nekaj kar zadeva arhiviste. V redkih primerih, kjer sam nosilec predstavlja neko vrednost, ga kaže ohraniti, vendar bi v tem primeru postal muzejski eksponat. Kot informacijski predmeti, se dokumenti zapisujejo na novejše nosilce, kar pa je naloga strokovnjakov s področja informacijske tehnologije. Trajnost nosilca nima več enake pomembnosti kot je to bilo v analognih časih, saj je migracija danes splošno pravilo. Hitrost dostopa, gostota hranjenih zapisov in varnostne kopije so delo informatikov, ne arhivistov.