

# Emulation as an Alternative Preservation Strategy – Use-Cases, Tools and Lessons Learned

Dirk von Suchodoletz, Klaus Rechert, Isgandar Valizada and Annette Strauch

Albert-Ludwigs University Freiburg  
Professorship in Communication Systems

Ulm University  
Communication- and Information Center

**Abstract:** Emulation has evolved into a mature digital preservation strategy providing, inter alia, functional access to a wide range of digital objects using their original creation environments. In contrast to format migration strategies a functional, emulation-based approach requires a number of additional components. These can be provided by Emulation-as-a-Service, implemented and developed as a distributed framework for various emulation-based services and technologies for long-term preservation and access. This paper presents three distinct applications of the emulation-strategy to preserve complex scientific processing systems, to render complex interactive and dynamic digital objects, and to implement a universal migration-workflows utilizing the original environments in which objects were created.

## 1 Introduction

It can be stated that today's memory institutions are faced with an extensive range of increasingly complex digital objects (DOs) from a wide range of domains. Business processes but also (commercial) research and development are being carried out by using computer systems producing solely digital results. The material to be made accessible over longer periods includes not only the traditional digital objects such as PDFs or images of digitized items. These objects are complemented by primary research data, all types of software artifacts, ranging from educational material to computer games, digital art as well as machines of famous people. In terms of DOs, even 10 years are a huge timespan when trying to keep those objects accessible and usable. Access to individual artifacts or complete software environments as well as reconstructing and re-enacting related processes may be required at some point.

At the moment, migration is the method most often deployed and trusted by memory institutions for long-term archiving and access of digital objects. This strategy takes digital objects through the constantly changing digital environment, made up of changing hardware and software configurations. It usually requires translation of the objects inner structures to an up-to-date schema. Although these translations make it possible to use and render digital objects in actual computer environments, such an approach unnecessarily limits

the number of object types that can be archived. Moreover, suitable migration tools, are usually not available for dynamic and interactive digital objects. A further problem that mostly concern businesses and research is the data-centric view of a migration strategy. Modern research involves not only data but also heavily relies on dedicated workflows and tool-chains. Thus, a pure data-centric strategy misses important pieces that are necessary for an authentic re-enactment, for instance to replicate research results or to reproduce a complex digital environment with its complete internal context to be preserved.

Emulation as a long-term digital preservation strategy shifts migration to the hardware level. It complements migration, as it is able to provide a valuable service: by using emulators, a more accurate reproduction of current digital objects and their processing environment can be verified and, to a certain extent, long-term access is guaranteed. While emulation or format-migration strategies may also be applied in the future (i.e. to save costs, react only if required) authenticity may be difficult to verify.

This paper deals with three applications for emulation strategies. Each application is underpinned by a practical use-case that has been examined during the bwFLA project.<sup>1</sup> One of the project aims is, to provide emulation services in an easily understandable and easily usable fashion. The preservation of a complex scientific processing system is discussed as a first concrete use-case. The workflow and supporting software framework is explained in order to to safeguard ingest and access to these environments. A second use-case shows tools and workflows for rendering complex digital objects, presenting an example for digital art. Finally, a format migration-workflow is demonstrated, utilizing emulation as a scalable and universal migration tool. The respective past desktop environment is wrapped in such a way, that every human interaction is abstracted by an automated procedure. Each achieved result and the various problems encountered during the involved process are discussed and evaluated.

## 2 Emulation-based Preservation Strategies

For more than 15 years, there has been a vital debate on using emulation as a strategy to ensure long-term access to digital records. Although emulation has always been an essential complement of migration, especially for dynamic and interactive artifacts, it is often considered to be too expensive and to require a good deal of technical skills to be a viable solution. As emulation does not require changes to the object or its structure, but migrates the underlying hardware, the original state of the DO and its authenticity is preserved. While technical issues have been mostly solved at present, tackling the costs and scalability is now a major challenge. [BT13]

Emulation does not operate on the object directly but rather addresses the environment in which it was used to create the object. This means, for example, the replication of software and/or hardware through other software. In the best case, it will not make any difference whether the object is handled through an emulated or original environment. Emulators, i.e. specialized software applications running in digital environments, preserve or alter-

---

<sup>1</sup>bwFLA – Functional Long-Term Access, <http://bw-fla.uni-freiburg.de>, (20. June 2013).

natively replicate original environments. Research on emulation as a long-term archiving strategy has matured since the first reports on archiving of digital information in 1996 [oPAG96], fundamental experiments with emulation executed by Rothenberg [Rot95] and the theoretical and practical work within the long-term preservation studies of IBM and the Netherlands National Library [vDS02]. The next phase was reached by the EU projects PLANETS and KEEP. The former looked into the inclusion of emulation into preservation workflows [vdHvS09, BKK<sup>+</sup>09], while the latter was focused primarily on media transfer, emulators and emulation frameworks [PAD<sup>+</sup>09, LKMvdH11].

Digital objects cannot be used on their own. They require a suitable digital working environment in order to be accessed. This context must combine suitable hardware and software components so that its creation environment or a suitable equivalent is generated, depending on the type of the DO. No matter which emulator is being chosen, the contextual information of the original environment of the digital artifact in which it has been created is always required. Questions such as "for which operating systems is WordPerfect 6.0 compatible with?" or "which tool generated this specific statistical analysis presented in that paper" are less obvious today than years ago. To overcome this knowledge gap, a formalization process is needed to compute the actual needs for an authentic rendering environment. Several concepts like view paths [vDS02, vdHvS09] or specific technical metadata schemata [DA12, DPDC12] were proposed.

While the technical challenges developing emulators are not considered in this paper, usability and accessibility of emulators for non-technical users are crucial. Emulation technology usually resembles a specific computer system. Since the number of different ancient and current computer systems is limited, the number of required emulator-setups is also limited. In order to allow non-technical individuals to access deprecated original environments, the tasks of setting up and configuring an emulator, injecting and retrieving digital objects in and from the emulated environment have to be provided as easy-to-use services. Making these services web-based allows for a large and virtually global user base to access and work with emulated systems. As a result of the bwFLA project, "Emulation-as-a-Service"[vSRV13] was developed which includes a framework and workflows to build novel cost-effective services for digital objects and associated processes so that long-term access is guaranteed with a predictable quality [RVvL12].

### 3 Preservation of Complex Machines

In certain situations, preserving the full original machine is inevitable. Also, there are several good reasons for making images of entire computer environments and for maintaining the ability to render them over a period of time. For instance, researchers can be provided with the ability to experience individual users' or representative users' old information environments such as politicians', artists' and other famous persons information environments. Another example is to get an average/representative user's desktop from a particular time period for accessing old web pages or certain documents and spreadsheets. Scientific research environments produce additional examples as computer workstations installations have grown and have been modified over time. During peren-

nial research projects, fluctuation of staff is common and consequently system knowledge is volatile. Complex software workflows have been created, usually involving tailored software components and highly specialized tool-chains paired with non-standard system tweaks. Rebuilding the system from scratch is an often complicated and time-consuming task involving manual labor and requiring technically-skilled personnel. [MRN<sup>+</sup>12]

As another example, information stored in Electronic Document and Records Management Systems (EDRMS) can be ruined from an archival perspective if they are taken out of their EDRMS as their context can be lost. Here, it would theoretically be possible to define a metadata standard and preserve sufficient metadata to capture the context that the file came from; however in practice this is extremely difficult. Thus, preserving the complete workstation seems to be more economical and, full functionality of the system can be retained with minimal effort if carried out properly. [CvSC13]

In a very direct application of emulation strategy physical machines are migrated into virtual or emulated ones by making a direct image of a computer's hard disk. This image file, representing a virtual disk, can then be run again attached to emulated hardware. Full system preservation can be described as a migration on the hardware layer. This transformation optimally preserves the original system and contained objects with all context in such a way that everything "behaves" in the emulated environment as it did on the original system [Lof10, vSC11]. Limitations in emulators may affect the result, but in many preservation scenarios few technical limitations like a limited screen resolution or color depth do not necessarily decrease the quality and usability of the emulated environment.

Practical aspects of a system imaging workflow are being described in [WRCv12]. During the bwFLA project a first prototype of a semi-automated workflow got implemented which tries to simplify the several involved preservation steps [RVvL12]. The full system-preservation ingest workflow is split into three stages:

**Preparation & Characterization** In a first step the user characterizes the computer system to be preserved (in the following denoted as "preservation target") by determining the original operating system and computer architecture to select a purpose-made imaging mini environment to boot on the original hardware. Using the original hardware ensures disk adaptor compatibility and allows to gather additional technical metadata such as information on the CPU, amount of memory and original peripherals deployed. To actually perform the imaging process on the original hardware, the preservation target requires certain technical capabilities, e.g. USB-port or optical disk drive (CD-/DVD) and the ability to boot removable media. Furthermore, a (standard) network adapter is required to transfer the image data into the image archive. To ensure the necessary conditions, the user is interactively questioned if the preservation target meets these requirements. Depending on the choices made, the imaging process is prepared either to be carried out on the preservation target, or on a different (modern) computer system. The latter option requires a dismantling of the preservation target and the removal of the hard disk.

A knowledge base on different operating systems regarding their compatibility with emulators and hardware dependencies is part of the framework. The user is presented with known issues based on his previous selections and step-by-step guides describing user-actions to be carried out on the preservation target. Such tasks may include reconfiguration

of the system to default graphic drivers or disabling of external dependencies during the boot-process (e.g. mounting network shares, connections to license servers, etc.). External dependencies may be restored in the emulated environment in a post-processing step. Finally, a specially tailored bootable image is generated and made available for download. The bootable image is either to be written on a USB pen-drive or CD/DVD. The medium is able to boot the preservation target using a preconfigured system that contains necessary configuration and credentials to connect to the repository backend to upload at a later point the generated image.

**System Imaging** By using the software provided, the target machine is activated and the preservation workflow is executed. An automated process launches the imaging process, the gathering of relevant hardware information about the target machine and the uploading of this data to the frameworks data backend. At the moment, the only interactive choice allows the user to select the drive to be preserved, if multiple options are available. By default, the standard boot-device is chosen. Currently, only complete block device preservation is supported. Single partitions and multiple disk configurations including special setups like RAID are planned for inclusion in future work.

**Verification & Submission** In a last step, the generated disk image is post-processed to be used in an emulator. This may include steps to pre-load required peripheral drivers for the new emulated hardware or the disabling of the original driver configuration. Finally, an emulation component, part of the bwFLA framework, is invoked with the preserved system image. The result is presented to the user for approval. If approval is granted, the image is submitted together with generated technical metadata to a dedicated repository for further processing. The workflow ends with the opportunity for users to update or amend the frameworks software knowledge base.

With integration of full system-preservation workflows into a distributed digital preservation framework, a common knowledge base on preparing, imaging and re-enacting ancient computer systems can be built, providing step-by-step instruction even to non-technical users. Due to the integrated feedback loop, the owner of a machine, subject to preservation is able to describe and test desired properties of the system for later access. Furthermore, external dependencies, either functional or data services are identified. Both, interfaces to the services as well as their availability should be documented and monitored as part of preservation planning.

### 3.1 Use-Case OS/2-DB2-based Scientific Environment

At the Linguistics Department at the University of Freiburg, a long-running research project was finally shut down. It had started in the 1970ies with documentations into local dialects of the south-west region of Germany.<sup>2</sup> Later on, it became one of the first projects in this particular field of the humanities to move from laborious paper based evaluation to a computerized environment. The collected data, entered into a relational database, was

---

<sup>2</sup>The project was relevant enough to get added to the permanent exhibition of the Uniseum, the Freiburg University Museum, <http://www.uniseum.de>.

then used to create customized dialect maps depending on various input parameters which were published in numerous publications and theses. Several specific workflows and even custom font and symbol shapes were created to produce PostScript output from a TeX file source. Many researchers and PhD students contributed to the project and refined the workflows over time. The system can still be run and even now produces up-to-date language maps from the data source. Unfortunately, no actual user of the system has a full understanding of how the system was setup and configured.

The system was put together in 1993 and consists of one x86 server machine running OS/2 version 2.1, running a IBM DB2 database and six x86 clients, offering access to the database over a now deprecated LAN infrastructure. Various workflows for map generation could be executed on the clients from network shares. By the end of the project, the server was still fully functional and at least three clients were working completely, two of them more partially.

It was not possible to boot the imaging mini-environment on the original hardware for several reasons such as missing proper boot devices or incompatible network adaptors. Thus, the hard disks had to be removed and connected to a suitable imaging machine. Any pre-processing like resetting the disk driver to general IDE or the hardware specific SVGA driver to compatible VGA was not considered for the lack of deeper OS/2 configuration knowledge.

The SCSI disk got easily imaged to a container file. Additionally, this file back was written back onto a newer SCSI disk as a backup for the (by then) still in use database machine. This test served at the same time as proof of an identical copy of the original. The replacement worked as expected in the original server. The installed system started exactly as it was previously shut down. Different to expectations as most people would assume, IDE to be more widespread and compatible, the dumping of parallel port IDE disks was more challenging compared to the SCSI counterparts. A working IDE adaptor was required which not only correctly recognizes the disk but also produces reliable disk images.

## **3.2 Discussion**

Both the server and the client environments were finally re-run successfully after the relevant hardware drivers for disk, network and VGA had been modified for the new environment. Each machine setup regarding the relevant components like the DB2 database were exactly in the same state in which the original machine was shut down before imaging. Nevertheless, the involved tasks to revive OS/2 in its new environment were rather complex. Not only that the disk imaging process ran into unexpected problems, additionally a couple of post-processing steps were required to boot the original operating system completely. There are two virtual machines and emulators available (QEMU and VirtualBox) supporting OS/2. A couple of changes to the original configuration were to be made to successfully complete the system migration and start the environment. The original desktop screen resolution was demoted to standard VGA as no compatible driver is available to support the original screen resolution. Another issue was the network connection, which

is required to access the database and shared folders from the clients. The original Token Ring network was migrated to Ethernet as no equivalent was available in any emulator.

Since the object considered for system imaging is much more complex, the measurement of success and completeness of the workflows is not completely clear yet and it remains a topic of ongoing research. The process can be seen as a kind of migration affecting certain parts and aspects of the object. Usually, lower layer components of the hardware-software-stack like the CPU or drivers are affected. These should not influence the object of interest but can definitely do so. The presented method is a suitable and efficient preservation strategy for highly complex and deprecated systems where detailed knowledge of the system is not available anymore. With a full system-preservation, a device is preserved as a "black box" with somewhat limited utility for future use since details on inner mechanics and constructions are not covered by this process and hence potentially lost. If the system setup is completely known and comparably easy to reproduce, a slightly different method can achieve better results with a smaller footprint regarding the object size.

## 4 Preserving Environments and Processes

The electronic collections of libraries, museums and archives are growing and have an increasingly relevant role in their holdings. These objects are increasingly complex and may require certain software environments to run or render properly. Standard digital preservation methods can lose important parts of DOs and can not address properties like non-linearity or interaction as required, e.g. for electronic teaching material, encyclopedias, multimedia objects, computer games or digital art.

Often, the formats of those DOs are outdated and can no longer be run or rendered on today's systems. Emulation can provide the required digital environments suitable for a given object type. In order to deal with the different classes of objects, and also to cope with their special requirements, emulation can be applied on original system-environments to arbitrarily render DOs. To re-enact a digital object in its original system-environment, a number of additional components and configurations are required.

A traditional method to discover a digital object's runtime dependencies is querying a file-type database like PRONOM [BCHB07]. Several tools have been proposed to resolve software dependencies such as DROID<sup>3</sup> makes use of "file-magic" fingerprints, while other tools utilize system library resolving mechanisms [Jac11]. While these tools and techniques provide valuable information to users, they do not guarantee generation of a suitable rendering environment regarding, for instance, completeness, quality and conflicting dependencies. Identification of file-type and linking applications is only the first step. This information needs to be extended to a viewpath description [vdHvS09] with required additional software, a suitable operating system and hardware emulator. In order to preserve a re-enactable rendering environment, any dependencies from interactive applications to operating system and hardware components need to be identified.

---

<sup>3</sup>The DROID Project, <http://droid.sourceforge.net/>, (20. June 2013).





object's rendering environment and is aware of the objects significant properties and expected behavior. Finally, a dedicated software archive infrastructure is necessary.

In a first step, the digital object is imported from an archive, metadata and object manifestation is normalized so that it is useable within the bwFLA framework (WF-I.0 in Fig. 1). In the next step, a query for a suitable rendering environment is executed. The software archive suggests the known rendering environments, which the contributor is able to choose from (WF-I.1). If no suitable rendering environment is available, the contributor is redirected to the software archive ingest procedure (WF-I-SW). If a suitable rendering environment has been identified, it has to be instantiated, i.e. all software components defined by the descriptive viewpath have to be installed and configured. This task has to be carried out manually if no automated installation routine or pre-configured (i.e. cached) image is available (I-WF.2). The final steps of the workflow allow the user to fine-tune and configure the object's environment (I-WF.4) and to assess the rendering quality (WF-I.3) [GR12]. If the rendering quality is approved, metadata is generated and made available for further processing or storing.

The software archive ingest workflow starts with importing a single software component, e.g. a rendering application (WF-I-SW.0). During this process the user is able to provide detailed descriptive information about the object. This description is used as archival metadata for indexing and search purposes. Furthermore, this description contains information on the license model and intellectual property rights, as well as pointers to additional information like documentation and manuals (WF-I-SW.1). In a second stage of the workflow, the software component's dependencies are determined. If a required dependency is not known or not available in the software archive, it must first be ingested into the software archive by using a recursive invocation of the ingest workflow for this missing dependency. Through an installation and test procedure (WF-I-SW.3), the software component's functionality and completeness of the identified viewpath is verified. For each successfully ingested dependency object the viewpath is extended accordingly. The resulting viewpath then represents a suitable, manually tested and confirmed software environment. The generated metadata information might also include user feedback about the quality and/or costs of the produced technical metadata.

## 4.1 Use-Case Dynamic and Interactive Objects

Memory institutions like the "transmediale"<sup>4</sup> archive require versatile strategies to preserve, curate and display complex digital objects like digital art or similar types of objects which cannot be directly migrated. For this task, additional software components need to be preserved and enriched with additional information (metadata) like operation manuals, license keys, setup How-Tos and usage knowledge. Furthermore, each software component defines its own soft- and hardware dependencies. To ensure long-term access to digital objects through emulation, not only availability of technical metadata (e.g. TOTEM schema viewpath descriptions [ADP10]) is required, but these viewpaths also need to be

---

<sup>4</sup>It is an art festival based in Berlin running for over 25 years which forms new connections between art, culture and technology.

tested and evaluated by users who are aware of the digital object’s environment properties and performance. This can be achieved by the bwFLA framework with which one can perform a structured installation process of a reference workstation.

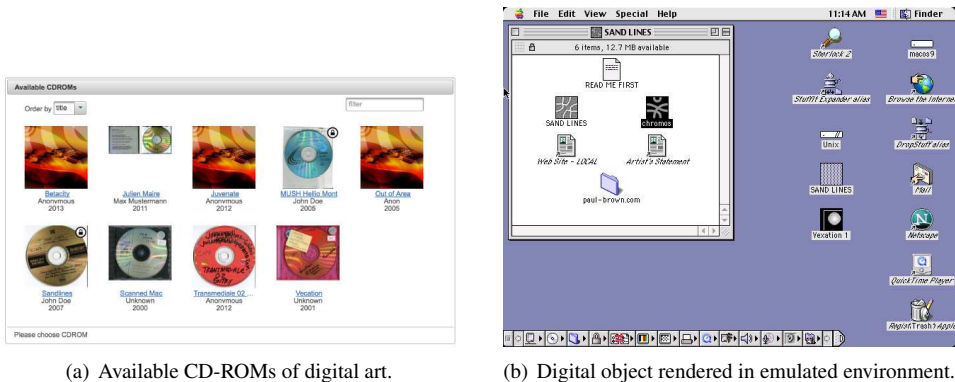


Figure 2: bwFLA: Functional access to digital art collection.

The bwFLA access workflow is able to provide a base technology for this task in order to keep digital artifacts alive. In bwFLA, different basic system-environments for manifold purposes were created. The base images include several Apple Macintosh systems offering Mac OS 7.5, 8.5 and 9.0 and various DOS/Windows environments, including the full range of the relevant early Windows versions from 3.11 till 98. On top of the base images, several access scenarios can be run. At the moment, the most advanced is a curation and access workflow for digital art. The use-case offers access to CD-ROM art which was produced, presented and distributed to end-users at the "transmediale" art fair (Fig. 2).

## 4.2 Discussion

The direct access to many DOs and interaction still offers the most complete and authentic experience. Especially dynamic and/or interactive objects like digital art and many other multimedia material cannot be (easily) preserved using format migration. The proposed workflow requires significant manual user interaction and seems costly and time-consuming at first sight. However, regarding preservation of current digital objects, the basic rendering environment is quite stable, concerning software and hardware dependencies. Usually, the main differences can be found on the top layer of the viewpath description, i.e. only a few additional steps are required if the software archive already contains suitable viewpath descriptions of today’s common digital objects. The ingest workflow could be further accelerated by employing caching strategies on created software images and by automation of installation tasks. However, the extra costs in terms of manual interaction during object ingest may reduce the long-term preservation planning costs since only the bottom layer (i.e. emulator) of the viewpath needs to be taken into account.

## 5 Migration-through-Emulation

Accessing objects directly in their original environment is desirable for authentic reproduction, but can be too laborious and costly if used for manual format migration like required for object normalization in archive ingest. If no suitable migration tools are available for a certain object type, the original application developed by the software producer is the best candidate for handling a specific artifact. Migration-through-Emulation (MtE) describes the concept of using the original or a compatible environment of a designated DO running in a virtual machine and thus replacing the original hardware and/or software stack. This approach avoids the often impossible alteration and adaptation of outdated software to present-day environments. A virtual machine runs within the host environment which contains the selected original system-environment suitable for handling a certain type of digital objects. The original system-environment is either reproduced from original software stored in the software archive or cloned from a prototypical original system.

To make MtE deployable in large-scale preservation scenarios without relying on user interaction, the user's function is replaced by a workflow execution engine [RvSW<sup>+</sup>09]. This requires appropriate interfaces in order to use emulators [vS10]. In contrast to simple command-line input-output migration tools, a MtE service builds on aforementioned technologies, e.g. using system emulation, a controlled rendering environment for certain types of digital objects but also an abstract description of all interactive commands to be carried out in order to perform a certain migration. Such a description consists of an ordered list of interactive input actions (e.g. key strokes, mouse movements) and expected observable output from the environment (e.g. screen- or system-state) for synchronization purposes (Fig. 3).

The *migration component* (MC) is the main module visible to the end-user, by exposing a simple *migrate* interface for (possibly) complex DO migration from format  $fmt_A$  to format  $fmt_B$ . The user requests a migration by providing a (set of) digital object(s) to be migrated, the requested final format, and a set of parameters. These may restrict the migration path length set quality or cost criteria for the migration process. The individual migration steps are identified. Figure 3 illustrates the general mode of operation of a MC. Based on the resulting identified migration path, the MC instantiates each node as a single migration unit. Beside this, the MC takes care of intermediate results and, if necessary, error reporting and recovery.

### 5.1 Use-Case Migration of PPT 4.0 to PDF

Migration tools working in today's environments are not available for all file formats. A good example is the Power Point 4.0 format which was used by the end of the 1990ies. Trying to open it in recent office suites produces a "format not understood"-error. For the evaluation a emulation based migration bwFLA workflow was produced, which transforms a PPT 4.0 input into a PDF file output. The workflow is implemented as a service accessible through the web. Another use case could be (chained) migrations to move the

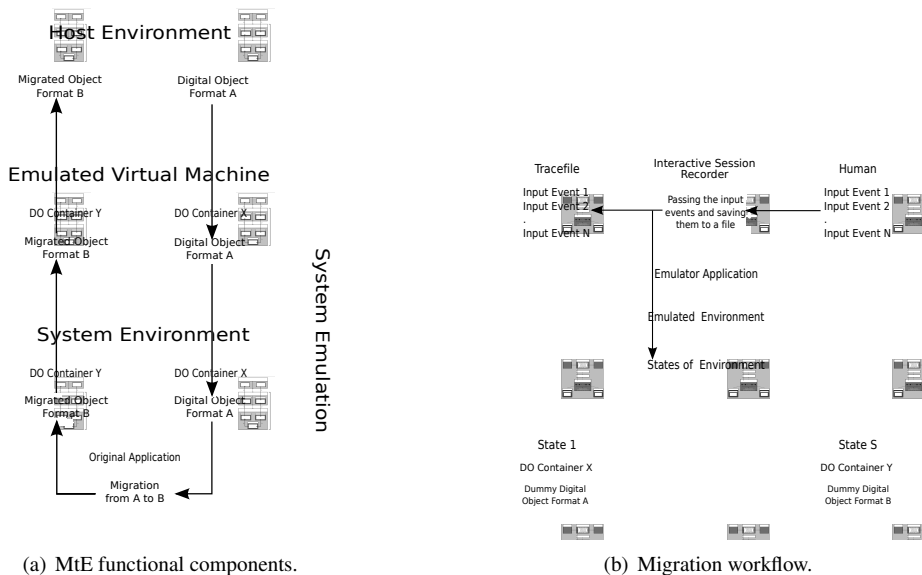


Figure 3: Abstract and generalized MtE workflow in the bwFLA framework.

outdated format up to an actual version. Several MtEs could be coupled to read the object in a newer version of the producing application and saving it in that one. Beside keeping the format, a wide range of different output formats supported by the producing applications could be generated. Thus a Microsoft Word 97 file, for example could be migrated to RTF, ASCII text or PDF at the same time for different purposes, such as viewing, indexing or further processing in actual software. Further workflows taking different inputs or producing different outputs can be created. Such procedures can then be used to e.g. normalize objects during archival ingest routines.

## 5.2 Discussion

To verify reliability of the functional encapsulated workflows, especially interactive workflow recordings a simple large-scale format migration process has been created as a test-case. Input format is a Microsoft DOC file version 8.0. Output format is RTF. The application used to migrate input files was MS Word 97 running on Windows 98. After recreation of the software environment, a single DOC file was injected into the emulation component with the help of a virtual floppy image and the system was started. In the next step, automated user interface interaction was carried out. In this case, the injected file was opened with the MS Word application and the file was exported as in RTF file format. After shutting down the emulated system, the migrated result was then available to the user.

The workflow has been carried out on 997 different objects of the same file type. In the evaluation 892 (89,47%) completed migrations and 105 failures were observed. Execution of a single instance took about 4 minutes. To investigate the reasons of failed migration workflows further, the experiment got re-run on the 105 objects and analyzed screenshots of 30 randomly chosen failure states. In these failure cases, three types of modal windows appeared on the screen. Note, the interactive workflow recording system is an independent platform and thus relies on graphical screen output and emulator stat. The first error category was due to a notification that there is no free space on the hard drive. Originally, the size was set to 30 MByte but due to the migration of embedded images into specific format MS Word 97 (BMP) more disk-space is required. This type of errors happened only when objects contained graphical content. The second and third error category was due to a MS Word specific warning (file was marked as write-protected and the file contained macros).

## 6 Conclusion and Outlook

Future users of digital assets significantly benefit from accessible data and user-friendly functional toolsets both in the scientific and cultural heritage as well as in the commercial domain. Emulation services for digital preservation can help to bridge outdated working environments for a wide range of objects and original environments onto today's devices.

After a number of successful national and international initiatives and projects on digital preservation and access it is time to take these results to memory institutions. The bwFLA project on functional long-term archiving, has started implementing and integrating different workflows in "Emulation-as-a-Service". These provide a range of different environments to deal with various curatorial and archival tasks using emulation. For original environments, the bwFLA EaaS supports at the moment 8 different emulators being able to run 15 distinct legacy computer platforms. The platforms range from MacOS 7 running on a Motorola 68K system emulator, PPC based platforms to various x86-based platforms. In a distributed EaaS model the costs of archiving secondary digital objects like operating systems and popular applications can be shared. With mutual specialization niches and specific areas can be covered without losing generality. The shift of the usually non-trivial task of the emulation of obsolete software environments from the end user to specialized providers can help to simplify digital preservation and access strategies. EaaS makes emulation and emulators more easy to handle. In combination with distributed access infrastructure structure, EaaS, preservation planning and preservation costs are fixed, determined only by the number of emulators and emulated systems.

While an emulation approach has technical limitations (e.g. due to external (network) dependencies, digital rights management, license dongles, etc.), the proposed workflows allow to uncover such issues and indicate risks w.r.t. to long-term preservation. By integrating emulation-based services, memory institutions are able to acquire new users and to provide new types of services. In cooperation with today's object creators, memory institutions can not only gain knowledge on future preservation challenges, but are also able to build a solid knowledge base on current formats.

## Acknowledgments

The work presented in this publication is part of the *bwFLA – Functional Long-Term Access* project funded by the federal state of Baden-Württemberg, Germany.

## References

- [ADP10] David Anderson, Janet Delve, and Dan Pinchbeck. Towards a workable, emulation-based preservation strategy: rationale and technical metadata. *New review of information networking*, (15):110–131, 2010.
- [BCHB07] Tim Brody, Leslie Carr, Jessie M.N. Hey, and Adrian Brown. PRONOM-ROAR: Adding Format Profiles to a Repository Registry to Inform Preservation Services. *International Journal of Digital Curation*, 2(2), 2007.
- [BKK<sup>+</sup>09] Christoph Becker, Hannes Kulovits, Michael Kraxner, Riccardo Gottardi, Andreas Rauber, and Randolph Welte. Adding quality-awareness to evaluate migration web-services and remote emulation for digital preservation. In *Proceedings of the 13th European Conference on Digital Libraries (ECDL09)*, 2009.
- [BT13] Daniel Burda and Frank Teuteberg. Sustaining accessibility of information through digital preservation: A literature review. *Journal of Information Science*, 2013. broad literature overview on actual research in the field of DP in business sphere, identification of research questions.
- [CvSC13] Euan Cochrane, Dirk von Suchodoletz, and Mick Crouch. Database Preservation Using Emulation – A Case Study. *Archifacts*, (April):80–95, 2013.
- [DA12] Janet Delve and David Anderson. *The Trustworthy Online Technical Environment Metadata Database – TOTEM*. Number 4 in Kölner Beiträge zu einer geisteswissenschaftlichen Fachinformatik. Verlag Dr. Kovač, Hamburg, 2012.
- [DPDC12] Angela Dappert, Sbastien Peyrard, Janet Delve, and Carol C.H Chou. Describing Digital Object Environments in PREMIS. In *9th International Conference on Preservation of Digital Objects (iPRES2012)*, pages 69–76. University of Toronto, 2012.
- [GR12] Mark Guttenbrunner and Andreas Rauber. A Measurement Framework for Evaluating Emulators for Digital Preservation. *ACM Trans. Inf. Syst.*, 30(2):14:1–14:28, May 2012.
- [Jac11] Andrew N. Jackson. Using Automated Dependency Analysis To Generate Representation Information. In *Proceedings of the 8th International Conference on Preservation of Digital Objects (iPRES2011)*, pages 89–92, 2011.
- [LKMvdH11] Bram Lohman, Bart Kiers, David Michel, and Jeffrey van der Hoeven. Emulation as a Business Solution: The Emulation Framework. In *8th International Conference on Preservation of Digital Objects (iPRES2011)*, pages 425–428. National Library Board Singapore and Nanyang Technology University, 2011.
- [Lof10] Mary J. Loftus. The Author’s Desktop. *Emory Magazine*, 85(4):22–27, 2010.

- [MRN<sup>+</sup>12] Rudolf Mayer, Andreas Rauber, Martin Alexander Neumann, John Thomson, and Gonalo Antunes. Preserving Scientific Processes from Design to Publication. In George Buchanan, Edie Rasmussen, and Fernando Loizides, editors, *Proceedings of the 15th International Conference on Theory and Practice of Digital Libraries (TPDL 2012)*, Cyprus, September 23–29 2012. Springer.
- [oPAG96] Commission on Preservation, Access, and The Research Libraries Group. Report of the Taskforce on Archiving of Digital Information. <http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>, 1996. [online; last accessed 10.06.2013].
- [PAD<sup>+</sup>09] Dan Pinchbeck, David Anderson, Janet Delve, Getaneh Alemu, Antonio Ciuffreda, and Andreas Lange. Emulation as a strategy for the preservation of games: the KEEP project. In *DiGRA 2009 – Breaking New Ground: Innovation in Games, Play, Practice and Theory*, 2009.
- [Rot95] Jeff Rothenberg. Ensuring the Longevity of Digital Information. *Scientific American*, 272(1):42–47, 1995.
- [RvSW<sup>+</sup>09] Klaus Rechart, Dirk von Suchodoletz, Randolph Welte, Maurice van den Dobbelen, Bill Roberts, Jeffrey van der Hoeven, and Jasper Schroder. Novel Workflows for Abstract Handling of Complex Interaction Processes in Digital Preservation. In *Proceedings of the 6th International Conference on Preservation of Digital Objects (iPRES2009)*, pages 155–161, 2009.
- [RVvL12] Klaus Rechart, Isgandar Valizada, Dirk von Suchodoletz, and Johann Latocha. bwFLA – A Functional Approach to Digital Preservation. *PIK – Praxis der Informationsverarbeitung und Kommunikation*, 35(4):259–267, 2012.
- [vdHvS09] Jeffrey van der Hoeven and Dirk von Suchodoletz. Emulation: From Digital Artefact to Remotely Rendered Environments. *International Journal of Digital Curation*, 4(3), 2009.
- [vDS02] Raymond van Diessen and Johan F. Steenbakkens. *The Long-Term Preservation Study of the DNEP project - an overview of the results*. IBM Netherlands, Amsterdam, PO Box 90407, 2509 LK The Hague, The Netherlands, 2002.
- [vS10] Dirk von Suchodoletz. A Future Emulation and Automation Research Agenda. In Jean-Pierre Chanod, Milena Dobрева, Andreas Rauber, and Seamus Ross, editors, *Automation in Digital Preservation*, number 10291 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2010. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Germany.
- [vSC11] Dirk von Suchodoletz and Euan Cochrane. Replicating Installed Application and Information Environments onto Emulated or Virtualized Hardware. In *Proceedings of the 8th International Conference on Preservation of Digital Objects (iPRES2011)*, pages 148–157, 2011.
- [vSRV13] Dirk von Suchodoletz, Klaus Rechart, and Isgandar Valizada. Towards Emulation-as-a-Service – Cloud Services for Versatile Digital Object Access. *International Journal of Digital Curation*, 8:131–142, 2013.
- [WB10] Kam Woods and Geoffrey Brown. Assisted Emulation for Legacy Executables. *International Journal of Digital Curation*, 5(1), 2010.
- [WRCv12] Ian Welch, Niklas Rehfeld, Euan Cochrane, and Dirk von Suchodoletz. A Practical Approach to System Preservation Workflows. *PIK – Praxis der Informationsverarbeitung und Kommunikation*, 35(4):269–280, 2012.