

COMPARING HOW TO TAKE CARE OF HUMANS' AND BIT-STREAMS' LIVES

Eld Zierau

Royal Danish Library

Denmark

elzi@kb.dk

ORCID: 0000-0003-3406-3555

Abstract – This paper investigates a number of practices to ensure safety for human lives and compare them with practices to ensure the 'lives' of bit-streams. The selected practices for human lives are: A. common emergency preparedness practices for securement of places like shopping centers, B. safety critical systems commonly used for airplanes, space crafts and nuclear power plants, and C. pandemic preparation planning.

The results of the comparison are used to illustrate how human security precautions can be used in similar ways in a bit preservation case involving methods and systems on both the technical and the organizational level.

Keywords – Bit preservation, Safety critical systems, Safety Procedures, Pandemics, Risk management

Conference Topics – Exploring the New Horizons; Covid-19 and Digital Preservation

I. INTRODUCTION

The aim of this paper is to find out what bit preservation practices can learn from the comparison with human life safety. Bit preservation is here defined as the required activities to ensure that the bit-streams remain intact and readable [1]. The human life safety practices included in the comparison are:

- Emergency practices, which cover how companies organize emergency preparedness and response processes to “minimise adverse effect on the health and safety of [people]” [2].

17th International Conference on Digital Preservation

iPRES 2021, Beijing, China.

Copyright held by the author(s). The text of this paper is published under a CC BY-SA license (<https://creativecommons.org/licenses/by/4.0/>).
DOI: 10.1145/nnnnnnnn.nnnnnnnn

- Safety critical systems practices, in which a safety critical system is defined as “A system in which any failure or design error has the potential to lead to loss of life” [3]).
- Pandemic preparedness planning practices, which cover “Advance planning and preparedness to ensure the capacities for pandemic response are critical for countries to mitigate the risk and impact of a pandemic” [4].

There are many similarities between keeping humans alive and keeping bit-streams 'alive'. An obvious and important common element is risk management focused on 'risks on loss of lives'.

There are of course also differences, where the biggest differences are the ability to replicate and the life expectancy.

Concerning the ability to replicate, a human being may be cloned, but the result will not be the same individual, whereas bit-streams are easily replicated one-to-one. For example, a person who dies in a plane crash cannot be replaced by a healthy clone made before the crash. However, in the case of a bit-stream, you can always replace it, as long as there are healthy copies left.

Concerning life expectancy, a human life is relatively short compared to the life expectancy for bit-streams that represent cultural heritage, which must survive for many generations. Therefore,

there will be fewer disaster events happening during a person's life than events happening in the expected lifetime of a preserved bit-stream. Examples of such events are natural disasters, solar storms and pandemics. These events can be destructive, but the derived consequences of the events (e.g. broken supply chains) can be destructive as well.

This paper investigates the similarities and differences between life precaution practices in more detail. To illustrate the outcome of the analysis, there will be a description of the Danish bit preservation solution case, from which the findings are discussed.

II. BASIC PRINCIPLES FOR BIT PRESERVATION

To understand the basis for comparison, this section provides a short summary of bit preservation practices.

Bit preservation principles are addressed in different ways in the different models used in digital preservation. The Open Archival Information System (OAIS) Reference Model also addresses such principles implicitly [5]. OAIS could be misinterpreted to say that bit preservation is identical to the *Archival Storage* Functional Entity (OAIS Functional Entities are illustrated in Figure 1).



Figure 1 OAIS Functional Entities.

However, a closer look reveals that all Functional Entities including *Administration*, *Data Management* and *Preservation Planning* need to be in play for bit preservation. The coming version of OAIS will make this more clear [6]. Another important point in OAIS is that it is not only a question of technology, but just as much a question of the surrounding organization.

During the last decade, the challenges of bit preservation have been more widely accepted and various community initiatives have been launched, e.g. the Preservation Storage Criteria group [7].

One of the starting points for this increasing acceptance of bit preservation challenges was David Rosenthal paper from the iPRES 2008 conference: "Bit Preservation: A Solved Problem?" [8], which indirectly identifies three basic elements to take into account when doing bit preservation: Number of copies, Independence between copies and Frequent integrity checking.

A **number of copies** (greater than one) is needed to be able to replace faulty or missing copies. It is obvious that one copy is not enough, since errors or events destroying this copy will mean loss of the only copy. There have been arguments that the more copies you have, the bigger the chance of survival. This is the original thought behind LOCKSS (Lots of Copies Keep Stuff Safe) [9]. This cannot however stand alone, but needs to be seen along with the other principles as described below.

Independence between copies is needed in order to ensure that the same incident cannot harm multiple copies in such a way that data is lost. Figure 2 illustrates such a scenario in which all copies are placed under the same active volcano and thereby will be harmed in an eruption.

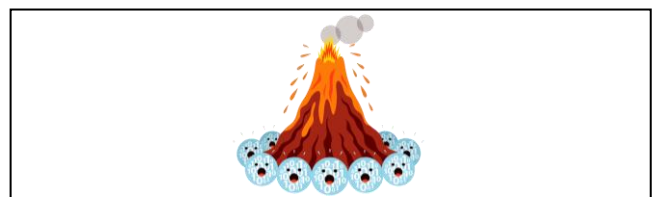


Figure 2 All copies will be destroyed by an eruptive volcano, Illustration from digitalbevaring.dk.

A real life example in which inadequate independence almost lead to loss of data was found in the Distributed Digital Preservation project. Here, an organization lost 7 out of 10 copies of data due to the same hardware error [10].

Frequent integrity checking both locally for one copy and globally comparing all copies is needed in order to detect faulty or missing copies. If copies with errors are not found in due time, there is a risk that different events will destroy each and every copy before the errors are discovered, and consequently the data will be lost. Even for backup systems, there are usually local checks of checksums for files in order to see whether the file has changed for some reason. Such checks can uncover lots of errors, but in Denmark, we have experienced software upgrades causing errors,

which were not discovered until we made a cross check with other independently placed copies. Non-detectable local errors can be caused by e.g. software errors or by malware that changes both files and checksum.

A simple general view can be made for solutions that serve the three basic elements. This is illustrated in Figure 3.

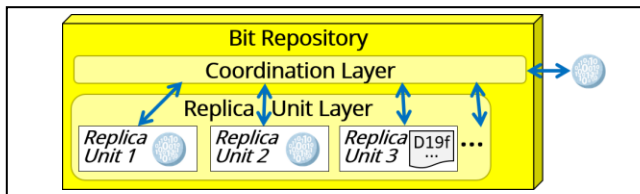


Figure 3 General structure of a bit repository.

All such solutions include replication of the data and some sort of coordination of communication via the Coordination Layer [1]. The replica units are the units where data is placed. They consist of a specific media in a technical and organizational environment. In Figure 3, Replica Units 1 and 2 contain full copies, while Replica Unit 3 only contains a checksum copy, which serves as an extra vote in case an error occurs in one or more of the full copies. The whole bit repository consists of both technical and organizational elements as well.

To which extent the three basic elements are used will depend on a **risk analysis** on how well a combination will mitigate different risks leading to loss of bit-streams. **Risk management** is in general relevant for all parts of bit preservation, including the technical, the organizational and the geographical level.

III. CASES AND COMPARISONS TO SAFETY OF HUMAN LIVES

This section describes each of the three human safety practices followed by a comparison to bit preservation, highlighting findings in bold/italic.

A. *Emergency Preparedness Practices*

There are many standards and local legislation for emergency preparedness and response in buildings with humans like a shopping center (e.g. ISO 45001: Clause 8.2 [2]). A common element is that there must be implemented emergency procedures with frequent drills to test them. The

following real life story provides an example where precautions were not in place with consequences that could have been fatal.

1. *The Shopping Center story*

I was in a shopping center buying new shoes for my children. While the children tried different shoes, it was announced on the speakers, that all people should go to the nearest exit. This happened in 2001 shortly after the 9/11 incident, so my mind was not set on fire alarm, but considered whether this was a bomb threat. The children were quickly helped to get into their own shoes, so they could run more quickly. When they were ready, all shop personnel had disappeared. We rushed out to find an emergency exit, but it was locked. People started to panic around us. We finally found our way to the escalators. All people were rushing to the escalator going to the roof. I had a strong urge to follow them, but stopped for a second thinking that the roof would be the worst place to be, no matter what caused the alarms. At the same time, I saw personnel rushing towards another emergency exit that I had overlooked. My logic now won over my urge to follow the crowd, - no doubt, the personnel were more knowledgeable about which exits to use. Soon after, we found out why people had rushed to the roof; they were getting their cars. Since the emergency exit path crossed the runway from the car parking, we had difficulties in passing cars with panicking drivers, who couldn't get anywhere, since the road infrastructure was not dimensioned to empty a roof full of cars in a short time.

I complained about the incident to the shopping center. The answer was that all personnel had guided people to the emergency exits, and that there had been guards to prevent people from going to the roof. Their answer showed that they did actually have procedures in place to limit casualties at real emergencies, but these procedures were obviously never implemented, so it was lucky that this was just a small fire and nobody was harmed.

2. *Learnings from Emergencies Preparedness*

The learnings from this story are many, and scarily with many similarities to bit preservation. One learning is to **have safety procedures**, which the shopping center might have had, but they had

certainly not been implemented. Another learning is consequently that *safety procedures must be implemented*. Procedures are no good if they are not carried out, e.g. because of lack of knowledge or training. This is very true for bit preservation as well. *Audits* of existence and implementation of safety procedures are essential for security in a shopping center, as well as of security for bit-streams.

A third learning is that *reactions based on following the crowd can be fatal*. The fact that there was no loss of lives in the shopping center case is obviously not an argument for abolition of safety procedures. In bit preservation, I have often heard comments to safety precautions like "Is that really necessary? We haven't seen any loss yet using less security" or "Why should we worry about that, when comparable archives do not seem to bother?". Especially, considering that our cultural heritage has a life expectancy of more than 100 years, we need to take into account what can happen regardless of whether this has happened in the past 20 years or not. Reactions should be based on scientific evidence, if possible, - otherwise use evaluated experience, and always supplied with reasoning.

B. Safety Critical Systems

Safety Critical Systems concerns systems where failures have high consequences, typically for human lives. It is an area that became important at an early stage when IT technology evolved and became part of systems where human lives could be endangered. Typical areas are space industry, systems supporting air traffic, driverless trains and nuclear power plants, but also more common areas like automatic parts of cars.

1. Safety Critical Systems Cases

An important strategy for safety critical systems is to have duplication and independence between critical 'components'. Examples are emergency power in power plants, duplication of flight altitude sensors or having co-pilots who can take over if the captain 'fails' to work. A similar strategy is to let different programs calculating the basis for decisions individually, which enables final decision of action based on voting between the different results. The point is to make sufficient risk analysis for duplication, meaning that you should not place e.g. emergency power supply in a basement that

can be flooded by a tsunami as was seen during the Fukushima case.

An important technique for increasing the reliability in safety critical system is to use formal development methods. Formal development methods provide mathematically based languages for specifying software systems and proof systems enabling verification that the program acts as specified [11]. The use of these methods are motivated by the fact that normal testing strategies can only cover a set of individual input, while proofs cover whole outcome spaces. Furthermore, formal development methods discover errors early in the development process, which influence quality and costs to the better [12]. Examples of formal languages are Z [13] and RAISE Specification Language (RSL) [11].

2. Learnings from Safety Critical Systems

The *duplication and independence strategy* for safety critical systems can be compared to the principle of having independence between replica units with different data copies. On the software level it corresponds to having *different software* serving different replica units. The use of co-pilots is similar to having *different staff* operating the different replica units. Likewise, duplication of independent components can be compared to having *different hardware*. It can even be compared to the *bit-stream level*, since one of the key differences between bit-streams and humans is that bit-streams can be made into equally worthy copies.

Comparing results from different components/software/operations corresponds to *Integrity checks across data copies*, since the checksum for a damaged copy will differ from checksums for the other copies. In this way, the integrity check supports error finding and repairing of faulty copies with non-faulty 'cloned' copies.

Using formal development methods has not yet been seen as a method applied to digital preservation. The challenge is that such methods require understanding of mathematics, at least logic and set theory, and the methods themselves introduce a more time-consuming development process. In the example of RSL in the 90's, it took 3-12 months to master RSL [14], but it is probably

close to 3 months for less complex languages or subsets of RSL¹. Development of 'bit safety critical systems' for our cultural heritage is unfortunately not as well funded as 'human safety critical systems'. Therefore, I doubt that we will ever get a budget that could support such formal development methods. Instead, we need to be extra cautious about the use of other techniques to minimize the error rate, e.g. *development methods using code reviews*, pair programming, automated testing etc. The good side of the story is that bits can be 'cloned', so if a clone 'dies', it is possible to replace it with a clone that existed independently from the events that destroyed the 'deceased' clone.

C. Pandemic Preparedness Planning Practices

For years and years, epidemiologists have pointed to the danger of a global pandemic. In 2005 WHO had a publication "WHO checklist for influenza pandemic preparedness planning" [15], which has a preface story "Some time in the future" with an example of an outbreak of a respiratory pandemic. The checklist is focused on being prepared to take the needed actions in case of an outbreak. Several science fiction movies have also illustrate the danger (e.g. The Outbreak), and there are numerous examples of appearing potential pandemics like Ebola and SARS.

1. The Covid-19 Case

In Asia, many countries were prepared for the Coronavirus disease (Covid-19), because these countries have handled other epidemics within the past 20 years. In the start of 2020, it seemed like Europe did not believe that there was a need to be concerned about the virus, since it was happening far away and no deadly virus had hit Europe for the past 100 years. Consequently, many countries were too late in their reactions to keep the virus under control.

It quickly became obvious that European countries were not sufficiently prepared for the pandemic. Most countries hardly had any protective equipment for the health personnel, and respirators for treatment of the patients with Covid-

19 were in shortage too. The plans that did exist in some European countries were far too insufficient to be of much help.

Especially during the first virus wave in spring 2020, there were many different approaches. Some of them took their own approach, like the Swedish approach. Another tendency was that the countries 'followed the crowd' in regulations, as for instance regarding the closure of borders. Some initiatives were based on experience and practices from Asia, like lockdown. Other experience, like use of surgical masks by the public, was only adopted at a much later stage. The reason was missing scientific evidence for masks to work, and such evidence had to be in place before the recommendations were given.

The pandemic waves also affected daily operations in society in many different ways, because the consequences of a lengthy lockdown, resulted in lack of staff and delivery shortage. At the Royal Danish Library, one of the effects was that we were not up to speed with migration of tapes, since this operation would require personnel to go to the premises. Furthermore, we are right now changing a replica unit for our web collections, since the old replica unit has hardware which will reach end-of-life in August 2021. At the time of writing, we struggle to meet the deadline, since many terabytes of data must be transferred to a new platform, and both delivery of hardware and assistance in setup are delayed due to other consequences of the pandemic.

2. Learnings from Pandemics

One learning, as for any loss giving events, is to be prepared, and for bits this need to include the perspective of what can happen over a very long period of time. In other words, we need to be ***prepared for the worst possible scenario***, regardless of whether we have experienced this within the last century or not. One example is hardware errors resulting in loss in all instances of the same hardware. Another example is a big batch of manufactured magnetic tapes has an error resulting in a shorter tape lifetime, which can lead to loss of all data on such tapes. There are also examples like natural disasters or solar storms harming all data (or all data on magnetic media) in a large area.

¹ Informal experienced based estimate by associate professor Anne Haxthausen who teach in this area.

Another learning is to *pay attention to experience and expert knowledge*. If the warnings from epidemiologists and WHO had been heard, *creation of plans* would have been in place in due time. This would enable *quick reaction to plans*. Such plans would involve a supply of needed equipment stored for such events. Plans for implementing restrictions to keep the virus under control could have been effectuated, already when the first warnings of Covid-19 appeared. Many lives would have been saved and restrictions would not have needed to be as strict, since avoiding a larger outbreak from getting out of control requires less restrictions than stopping an outbreak which is already out of control.

A third learning is to *listen to experience* from someone who has actually been in the situation before, for instance regarding lockdown and wearing surgical masks. There is a challenging balance here, since there is a grey zone between following experience and following crowd behavior. Science is obviously the best basis, but if there is no scientific basis, there is a good chance that long experience (re-evaluated by logical sense) is worth considering, e.g. for the surgical masks.

A fourth important learning is that we need to be *prepared for periods with broken supply chains*. This is a general learning, not just from the Covid-19 pandemic, but also from the tsunami in 2004 and many other events [16]. A small scale example leading to staff shortage is the lockout, as we experienced it in Denmark in 2018 [17].

IV. SYSTEMS SUPPORTING BIT PRESERVATION

This section will present a short overview of systems for bit preservation, and then dive into a single case of Danish bit preservation solution in order to illustrate and discuss how the findings fit with bit preservation solutions.

A. Various solutions

Different institutions have various solutions for bit preservation. The reason for the variations can be anything from different requirements to acceptance of different risks of loss of data.

There are various technical software solutions to support bit preservation, where each of them can

be setup to meet different requirements. One example is LOCKSS, which is based on on-line caches. LOCKSS is good for publicly accessible data, but can be a problem for confidential material where off-line media is required. Another example is DuraCloud², which is mainly developed for Cloud solutions, and has a relatively weak authorization and authentication implementation for confidential materials. Other solutions use built-in systems supporting bit preservation like Archivematica³, Preservica⁴, Arkivum⁵ or Libnova⁶. Each of these solutions can fulfill different but not all types of requirements. Several organizations have built their own solutions to fit their specific purposes. In Denmark, we use the BitRepository.org⁷ framework. Common for all these systems is that the systems depends on their physical instantiation and the surrounding organization, which both are just as important as the system itself.

B. The Danish Bit Preservation Case

This section describes the Danish bit preservation case to illustrate how the learnings can be interpreted for a bit preservation solution in practice. Other bit preservation solutions may be evaluated in the same way.

The solution in Denmark is based on a technological framework developed in the early 2010s and has been in use since 2012. However, the framework is just a piece in the puzzle to meet requirements for bit preservation, since the organization and choices for implementation must be included to illustrate all the learnings. Therefore, the description also includes these elements.

1. The BitRepository.org framework

The aim of the framework is to provide a basis for secure large-scale ingest, storage, access, audit and advanced integrity check of bit-streams. The motivation for building the framework was that none of the known existing systems could meet the

² <http://www.duracloud.org/>

³ <https://www.archivematica.org/>

⁴ <https://preservica.com/>

⁵ <https://arkivum.com/etmf-archiving-preservation/>

⁶ <https://www.libnova.com/>

⁷ <http://BitRepository.org/>

requirements from both the libraries and archives with differing requirements regarding bit safety, confidentiality, access, and possibility for cheap storage solutions. For sustainability purposes, the framework was required to be independent of technology stacks and storage platforms.

In order to meet these requirements, the architecture of the bit repository takes into account; A: Components of the system must have *no* direct knowledge of each other's implementation, B: All communication is based on a common message protocol, and C: All communication is asynchronous. The design also views the bit repository as an OAIS on its own⁸, including elements of all OAIS Functional Entities, not just *Archival Storage*.

Finally, the framework supports execution and monitoring of bit preservation actions like checks of missing files, consistency checks of checksums across all involved copies, and recalculation time for checksums, as well as the possibility of replacing faulty copies, supplied by various monitoring operations. The architecture of the framework is as shown in Figure 4.

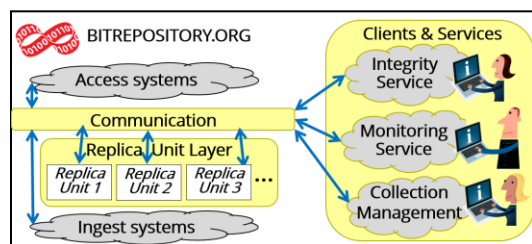


Figure 4 Architecture of BitRepository.org.

In this architecture, the Coordination Layer (from Figure 3) is split into a Communication Layer and a Clients & Services Layer. The backbone of the design is a message protocol, which is known and respected by each component in the system. The messages are specified in XML following an XSD scheme⁹. In this way each component only needs to know which collections it serves and which communication lines it must use. The protocol also defines adequate behavior to missing response or wrong responses, to avoid that such cases can result in delays or errors of the components.

The Communication Layer has only one function: to coordinate exchange of information between Clients & Services and Replica Units. Such information will be either data (exchanged via dedicated data transmission areas) or messages respecting the message protocol (via a message broker). In this way, the layer is made independent of any persistent information, or any special features of the components. In order to facilitate the asynchronous communication operations, the protocol has a specific pattern for a set of messages used in each operation. This is illustrated for the 'Get File' operation in Figure 5.

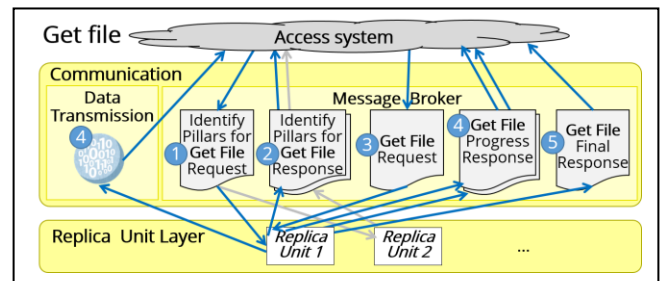


Figure 5 Messages for the Get File operation.

This set of messages consists of two parts; the first two messages will identify a replica unit that can deliver the requested data, by asking all relevant replica units whether they can deliver. This identification part ensures that the services do not need any information about the replica units, and that unavailable replica units will not necessarily result in an inability to get data. The remaining messages perform the actual 'get operation'. There can be several progress responses before completion, since replica units with off-line media like tapes need to carry out several steps to provide the data.

The Replica Unit Layer consists of independent replica units, where each replica unit stores authoritative and complete knowledge of the data placed inside it. All communication with the other bit repository components are solely based on the communication protocol. This architecture enables an arbitrary number of replica units to be implemented with separate types of platforms. This enables shifting and/or adding replica units as new storage technologies appear.

The Clients & Services Layer contains components to facilitate integrity services and the

⁸ According to the OO-IO model [17].

⁹ XSD and examples can be found at BitRepository.org

bit preservation surveillance as well as preservation actions like replacement of copies with errors. Furthermore, it facilitates monitoring of all the components in the system, and also collection management, like ingest, access and deletion. The services uses different operations defined by the protocol, which are *PutFile* and *GetFile* for ingest and access of objects, *GetChecksum* and *GetFileID* to get basis information for cross integrity check, *ReplaceFile* and *DeleteFile* to assist in repair of faulty objects, *GetStatus* and *AlarmMessage* to assist in daily operation, and finally *GetAuditTrails* enabling extraction of provenance of objects

The framework is scalable by allowing any component to be parallelized to enable scaling and avoid a single-point-of-failure. Security can be set at several places to different degrees by using certificates for messages and data transmission.

2. The Danish Bit Preservation Solution

The use of the BitRepository.org framework includes different technology stacks for the replica units. Each replica unit has special characteristics with respect to e.g. geographical location, operating systems, software, media and organization in charge of operation. This enables collections data copies to be placed in a replica unit combination, which fulfill its requirements best. To illustrate this, Table II lists three of the Royal Danish Library's collections with different requirements to bit safety, confidentiality and storage pricing.

TABLE II
Requirements for the three selected collection








	Material	Bit safety	Confidentiality	Storage
	Web archive	High	High	Normal
	Radio & TV	High	Low	Cheap
	Digitized material	Medium	Low	Normal

Table III lists a subset of our replica units and some of the independence characteristics, as they will appear in 2021, after improving our bit preservation (excluding checksums for simplicity).

TABLE III
Characteristics for Replica Units

	Location	Operation	Software	OS	Media
	Abroad	3 rd party	.NET/Acronis	Windows	Tape LTO9
	Aarhus	IT Aarhus	Java/Netbackup	Linux	Tape LTO6
	Skejby	IT Aarhus	Java/Netbackup	Linux	Tape LTO 8
	Cph	IT Cph	Java	Linux	DiscIsilon

The final setup for the three collections is illustrated in Figure 6. The replica units are displayed as houses in the colors pictured in Table III, while the collections are displayed as colored bit-heads as in Table II.

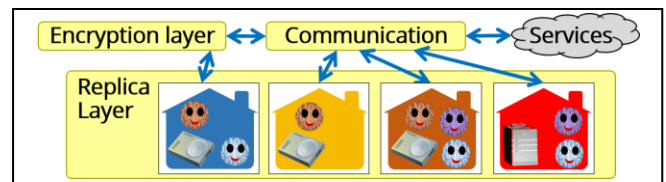


Figure 6 Future setup for the three selected collections.

Compared to today's setup, the framework supports a number of improvements, which are not directly a part of the framework: A: allowing encryption of one or more copies, B: new .NET software for off-line media replica unit, and C: new 3rd party replica unit.

The new encryption layer is needed to allow replica units to be placed at 3rd parties without violation of jurisdiction, if the replica unit is placed abroad. This layer is not directly a part of the framework, since neither the Communication Layer nor the Replica Unit Layer will act differently if there had been no encryption.

The new .NET software is needed, since there are dependencies on operating system, software and backup software for all tape replica units. This could lead to errors resulting in losses for collections only having tape copies, e.g. the Radio & TV collection.

The move to a 3rd party is needed to introduce an independent operator, better geographical diversity between the copies and better software independence.

At the Royal Danish Library we have a group of specialists responsible for daily operation and future evolvement of digital preservation. This also includes formalization and implementation of needed procedures for bit preservation, as well as self-auditing. Examples of daily operation are; integrity checking, checking age of calculation of checksums (3 months for disks, 7 years for tapes), maintenance and execution of procedures for operation and technology watch, and reacting quickly to incidents putting the safety of the bits at

risk (e.g. endangered copy of data due to planned blasting in the surroundings of a copy [17]). Examples of future evolvement is maintenance of policies and strategies, followed up by involvement in planning of future actions to fulfill the policies and strategies.

The status for the bit preservation work is that we are still in the process of establishing common procedures after the merge of two libraries into one, and regrettably, we have not got to the point of carrying out self-audit on the aligned setup.

V. BIT PRESERVATION IMPLEMENTATION VS. FINDINGS

At the Royal Danish Library, we are very much aware of the importance of having *implemented, maintained and regularly audited* procedures. We are far in establishment of procedures, and plan for an audit, but obviously we are behind schedule.

We have to a great extent succeeded in implementing bit preservation *based on science and evaluated experience*, and we have so far *avoided decisions based on crowd reactions* and politics, e.g. argued against less independence between our replica units, and convinced the government not to outsource bit preservation [17].

Use of *duplication and independence strategy* and *comparing results from different components/software/operation* are areas that are very much present in our bit preservation policies and strategies. Therefore, these aspects are also implemented to a high degree. The bitrepository.org has proved to be a good framework supporting sustainability by letting us exchange replica units with new ones allowing improvement of a range of independence criteria concerning, organization, location and technology in a way that is independent from the framework itself. We do, however, face a challenge regarding large collections with copies only on tape. The challenge is that the number of tape providers on the world market is narrowing, and soon there will only be LTO7 to LTO9 tapes on the market. The LTO tapes are produced by only two companies using the same formula, i.e. faulty batches of tapes can lead to losses after a few years, if all copies were stored on tapes from this batch.

With the described enhanced bit preservation solution, we will be *prepared for the worst possible scenarios* in many situations, including considering changing technology over time. However, there are still issues like the tape technology, and depending on the final location of the new 3rd party replica unit, there may be lack of independence regarding threats like war and natural disasters. Concerning *quick reaction to plans*, there is still a need for improvement. This is especially true for large amounts of data that are hard to move or migrate within a short period of time, which could be needed if the company chosen for the new 3rd party replica unit goes bankrupt. So far, we have not been well *prepared for periods of time with broken supply chains*, which have been seen in the case of the lockout described in [17] and in the current case with changes of the web archive platform.

Concerning risk related to components in the BitRepository.org framework, these are admittedly not developed using *formal development methods*. There can be no doubt that robust code is needed, although the risks for bits can be lowered compared to risks for humans, since bits can be 'cloned' to identical 'clones' and used for replacement. The BitRepository.org framework has proved to serve its purpose by being flexible, although we are aware that there is a risk in being the only ones using it. BitRepository.org is open source, but until now, we haven't gotten around to making real effort to include other partners.

VI. CONCLUSION

This paper has shown that there are many similarities between safety of human lives and bits' 'lives', in spite of the differences that cultural heritage bits have a much longer life expectancy and that bit streams can be 'cloned' on-to-one.

From the emergency practices, we have been reminded of the importance to have implemented safety procedures, that are regularly audited on both the technical and the organizational level. Combined with pandemic preparedness planning practices, we saw that decisions should be based on science if available, - otherwise use experience with care, but do not rely on crowd reactions. The pandemic preparedness planning practices

combined with the life expectancy of preserved bit-streams made it clear that we need plans for the worst possible scenario (100 years' events and broken supply chains). We also saw from the web archive platform case, that quick reaction to such plans can be challenging for large amounts of data. Furthermore, from the safety critical systems practices, we were reminded of the importance to replicate 'components' behavior (implemented in different ways), in order to avoid losses caused by wrong decisions or actions. Finally, we have seen that software code should be as error proof as possible.

The bit preservation example showed a case where many of the finding can be supported. The most important learning from this example, is that bit preservation is far from implemented by the framework/system itself, since the organization, placement, physical environment and local software stacks for each copy are elements with just as big importance as the system itself. Furthermore, the coming challenge regarding tape technology being on few hands, illustrates that bit preservation solutions will continuously be challenged and in need of change in an ever changing world.

ACKNOWLEDGMENT

A huge thank you to associate professor Anne Haxthausen, who I worked with in research projects about RAISE back in the 1990's, and who has given valuable updates on Safety Critical System. Also, a big thank you to my colleagues who have helped me make this paper possible.

REFERENCES

- [1] E. M. O. Zierau, A Holistic Approach to Bit Preservation, Doctoral Dissertation, Copenhagen University, 2011.
- [2] ISO 45001:2018, "Occupational health and safety management systems - Requirements with guidance for use".
- [3] *Oxford Dictionary of Computing (7 ed.)*, A. Butterfield, G. E. Ngondi, A. Kerr, Eds., Oxford University Press, 2016, DOI: 10.1093/acref/9780199688975.001.0001.
- [4] Essential steps for developing or updating a national pandemic influenza preparedness plan, Web archive: archive.org, archival date: 2021-02-01T15:45:30Z, archived URI: <https://apps.who.int/iris/bitstream/handle/10665/272253/WHO-WHE-IHM-GIP-2018.1-eng.pdf?ua=1>

- [5] ISO 14721:2012, "Space data and information transfer systems - Open archival information system (OAIS) - Reference model". 2012.
- [6] Giaretta, D., Garrett, J., Conrad, M., Zierau, E., Longstreth, T., Hughes, J.S., Hemmje, M., Engel, F. OAIS version 3 Draft Updates. Proceedings of the 16th International Conference on Preservation of Digital Objects, 2019.
- [7] E. Zierau, S. Schaefer, N.Y. McGovern, A. Goethals, "An Overview of the Digital Preservation Storage Criteria and Usage Guide", Proceedings of the 16th International Conference on Preservation of Digital Objects, 2019, pp. 276-281.
- [8] D. S. H. Rosenthal, "Bit Preservation: A Solved Problem?", Proceedings of the 5th International Conference on Preservation of Digital Objects, London, Great Britain, 2008, pp. 274-280.
- [9] V. Reich, D. Rosenthal, "Distributed Digital Preservation: Lots of Copies Keep Stuff Safe". 2009. Web archive: archive.org, archival date: 2021-04-25T14:10:51Z, archived URI: https://web.stanford.edu/group/lockss/resources/2009-03_Distributed_Digital_Preservation.pdf.
- [10] E. Zierau, M. Schultz, "Creating a Framework for Applying OAIS to Distributed Digital Preservation", Proceedings of the 10th International Conference on Preservation of Digital Objects, 2013, 78-83.
- [11] C. George, A. E. Haxthausen, "The Logic of the RAISE Specification Language", Computing and Informatics Vol. 22, 2003, pp. 323-350.
- [12] P. Liggesmeyer, M. Rothfelder, M. Rettelbach, T. Ackermann. "Qualitätssicherung software-basierter technischer systeme - problembereiche und Lösungsansätze". Informatik Spektrum, Vol. 21 no. 5., 1998, pp. 249-258.
- [13] V. Ruhela, "Z Formal Spedfication Language - An Overview", International Journal of Engineering Research & Technology, Vol. 01, No. 6, August 2012.
- [14] Zierau, E. "Use of the Formal Method RAISE in practice", Proceedings of the 13th International Conference on Computer Safety, Reliability and Security, Anaheim, California, USA, 1994, pp. 31-40.
- [15] WHO checklist for influenza pandemic preparedness planning, Web archive: archive.org, archival date: 2021-01-25T01:28:43Z, archived URI: <https://www.who.int/csr/resources/publications/influenza/FluCheck6web.pdf>.
- [16] B. Anstey, C. Bayazit, Y. Malik, A. Padhi, N. Santhanam, S. Tollens, "Why now is the time to stress test your industrial supply chain", 2020, Web archive: archive.org, archival date: 2020-11-27T01:49:24Z, archived URI: <https://www.mckinsey.com/business-functions/operations/our-insights/why-now-is-the-time-to-stress-test-your-industrial-supply-chain>.
- [17] E. Zierau, "The Rescue of the Danish Bits". Proceedings of the 15th International Conference on Preservation of Digital Objects, 2018, DOI: 10.17605/OSF.IO/U5W3Q.