Preserving an endangered society: the case of Maldives

David Giaretta
Giaretta Associates
Dorchester, United Kingdom
david@giaretta.org

Ahmed Asmin
National Archives of
Maldives
ahmed.asim@archives.gov.mv

***Abstract*— A society is a complex entity. The threats of Climate Change mean that societies on low lying island and coastal nations are likely to be submerged within the century. Some of these nations have been offered a land of their own elsewhere in the world in such a catastrophic eventuality. The Maldives, an island country, is building a futuristic floating city to combat climate change and to keep up with increasing sea levels. Such ground-breaking innovations may be crucial in assisting atoll countries—like the Maldives—as combating the effects of climate change may not bring any sustainable difference in how such countries will be affected.**

World Meteorological Organization (WMO) estimates that we have seen 15 centimeters of sea-level rise during the 20th century and projecting 30-60 centimeters by 2100 even if greenhouse gas emissions are sharply reduced and global warming is controlled at or below 2 degrees Celsius. Such a calamitous scenario will be particularly concerning for the Maldives, the lowest-lying country in the world, with 80% of the island nation lying less than one metre above sea level.

Even if Maldives choses to be a floating island nation or in the inevitable climatic circumstances, environmental refugees, what must be transported to such a new home if the society is to survive? Clearly people and whatever physical possessions that can be transported must be transferred. In addition, there are many items of digitally encoded information, including, for example scans of physical buildings, mosques, historical monuments, artefacts and landscapes just to mention a few.

This paper addresses challenges which some of the governmental, legal, societal, commercial and personal information present, for preservation and transfer. A new issue which will be investigated is that of where usability and Representation Information overlap strongly with that of relationships between multiple objects. This may be looked at from the OAIS point of view as a single Information Object or, from a Records Management point of view as a "fond". From both these points of view one needs to consider, capture, preservation and transfer the relationships between objects. One complication is that the information objects are contained in multiple organizations, each with its own classification system, which needs to be interconnected with other classification systems.

Although much of the information is open, there will be confidential data, including governmental, commercial, health and personal, and these will be governed by strict security policies. The archive for the information must be OAIS conformant and should be ISO 16363 certifiable and so besides demanding a technological solution which fully supports the OAIS Information Model and Mandatory responsibilities. This paper will also describe the work which has been done to capture and preserve the society's intellectual lifeblood, so that it is ready for transfer to new home to re-create that society.

Keywords—information, interoperability, data, reuse, digital, trust, repository

I. INTRODUCTION

Maldives has the lowest terrain of any country in the world making it very vulnerable to sea-level rise. Much of it is likely be uninhabitable by 2050 [1]. The population would have to move, for example, to artificial floating platforms or elevated land purchased in other countries. Governmental, legal, societal, commercial and personal information, currently in many organisations, which is the lifeblood of Maldives society, must be preserved for use in the new location. This paper summarises the challenges and the way in which they are being addressed.

The National Archives of Maldives (NAM) has a statutory responsibility to preserve government and other information, just as happens in many other countries. Just as in those other countries there can be a gap of 25 years before the preservation responsibility for the information is transferred to the National Archives. The Director General of NAM has a good deal of discretion as to what should be preserved by NAM, and when the responsibility should be taken.

This transfer of responsibility would be necessitated when, for example, the originating organisation is simply not able to preserve the digitally encoded information for 26 years.

The threat of climate change, to the nation and in particular its digital life blood, in far less than 25 years, makes the issue even more pressing.

NAM does not, at the moment, have a digital repository, and so must create one. There are a number of challenges which must be addressed. This paper describes these challenges and the plans to address them.

II. Description of the Challenges

The first challenge is to understand the types of, and the volume of, information to be preserved. To address this challenge, a very broad survey has been conducted to find the volume and types of information, and much else besides. This has allowed us to scope the archive system as described in detail in section IV.

The second challenge is to specify, procure and configure the archive system. The information to be preserved is not limited to simple documents. In addition there is a great variety of other types of information ranging from proprietary medical records to simple spreadsheets. An archive system which conforms to OAIS [3] and is likely to be ISO 16363 certifiable [4] will be able to preserve such a variety of information. Very few commercial systems can support these requirements. Details are provided in section V.

Thirdly the information must be collected and transferred to NAM. This requires more simply transferring or copying the bits. It requires a high level of co-operation of the staff at each organisation with NAM, because those staff know about the information, including the semantics and software which is currently needed to extract and use information from the bits. In other words they must help supply the Representation Information associated with the digital objects i.e. the bits. They also are likely to know details of the Provenance, Access Rights and Context, to complement the information which can be extracted automatically. Fixity Information, as defined in OAIS, is not simply a hash code, it includes a description of the processes and techniques used to ensure that the bits have not been altered in an undocumented manner.

A fourth challenge is to create a Records Management system in order to support the relationship between pieces of information. Such a Records Management system is legally required and is likely to be key to developing value added services which will help to justify the resources required for the preservation system. Some ideas for this system are described in section VI.

III. Summary of the Survey Results

In order to understand the types and volumes of digital information which will need to be preserved an online survey was created. It asked about digital as well as physical holdings. The survey was sent to several hundred organisations in Maldives, including government ministries, statutory bodies, commercial companies, hospitals and health centres, courts, parliament and atoll councils.

Although details as specified in the OAIS Reference Model were wanted, the questions were phrased in more easily understood terminology.

Each organisation provided a single response but within each response up to 8 information collections could be reported, half physical and half digital collections. By the end of the survey the results were:

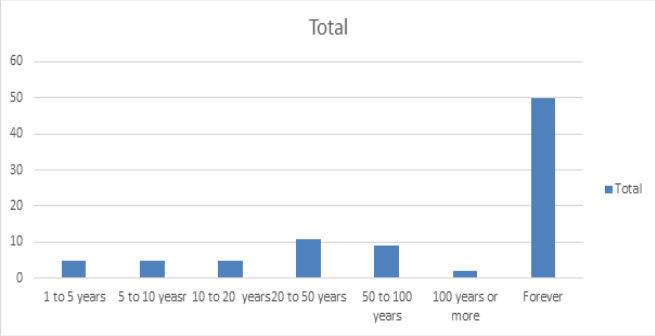
- 65 organisations responded;
- 96 digital information sets were reported on;
- 148 physical information sets were reported on.

The response rate was about 10%, except in the case of statutory bodies of which 70% responded.

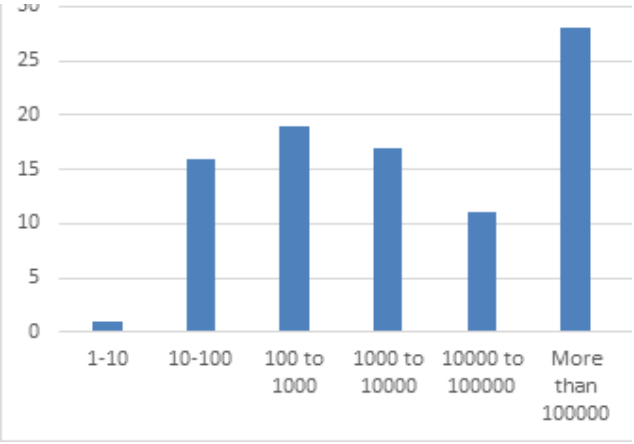
IV. Requirements for Digital Information Preservation

A number of important questions were clearly answered by the survey.

The length of time the information should be preserved shows a range from a few year to “forever”, with the latter being the clear favourite, as shown in the following figure.



The size of the datasets in MB were distributed as follows:



As will be discussed later, all repository systems can deal with a variety of formats, and particularly objects that are simply rendered i.e. displayed, printed or played. The responses about which formats the information were held in clearly shows there are many MS Office documents, videos, images, audio and emails. The majority of the text found in documents is in Dhivehi, the main language used in the Maldives Islands, and English but Arabic and other languages are also used.

But there were large numbers of special scientific formats and proprietary formats, with which many repository systems do not deal well.

From the survey responses one can estimate that at least 40 TB of digitally encoded information will have to be preserved. In addition, the survey showed that over 100 million pages may need to be scanned, creating 100 TB of images.

Following the survey a number of interviews were held with selected organisations. Each of the interviews revealed details of the holdings in much greater detail than did the survey. Many revealed many 10s of TBs which had not been reported in the survey because they were too complex, or which were about to be collected. These facts suggest that the repository system should be able to cope with at least 1 PB.

V. Overview of Potential Preservation Software Systems

Rather shockingly most commercial systems cannot preserve even simple spreadsheets, because one does not need simply to render the table but one must also know what the columns mean and what the units are, otherwise the spreadsheets cannot be used in any sensible way. In a similar way the survey shows that the great majority of datasets use special terminology (semantics).

Knowing that the information to be preserved includes objects that are not simply rendered, and has semantics other than a natural language, means that the software to be used by the archive must be carefully chosen, to ensure that it can support these requirements.

There are two broad categories of archive software, commercial and non-commercial. There are costs associated with each, for the hardware, software and software support. Commercial software is supported by the vendor, who can also arrange for hardware support, often using commercial cloud compute and storage. Preservica (https://preservica.com/), Libnova (https://www.libnova.com/), Arkivum (https://arkivum.com/), PIQL (https://www.piql.com/) and RODA (https://www.keep.pt/en) were examined. All provide Cloud based solutions, and some provide on-premises solutions also. Most Cloud based solutions use commercial cloud systems such as Amazon, Google, or Microsoft, which allows one to select the region, or regions in which the information is located.

Open-source software is freely available but then the cost arises from the need for hardware, which the archive may own, or may rent, for example through commercial cloud systems. Although the software is open source and therefore freely obtainable, there are costs associated with installing and configuring the system, and in some cases may require the executables to be built from the source code. These require expertise which is likely to be quite costly.

VI. Costs

Examination of commercial/proprietary software suggests that the cost for 300TB holdings would be in the region of $100K/year, with one “hot” copy and two “cold” i.e. slower access copies. In many cases the majority of the annual fee is the license for the software.

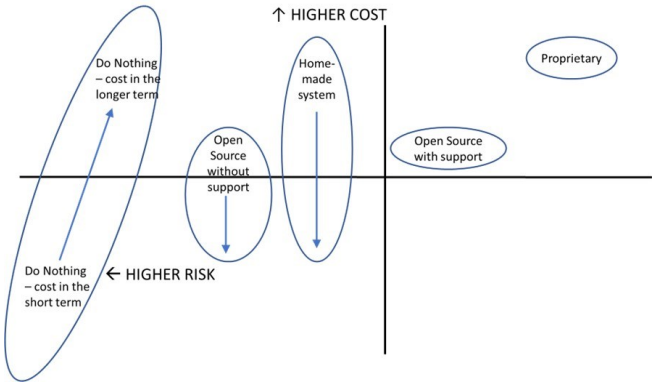
The costs associated with the use of open-source software is more difficult to estimate. There will be no license fee, but software experts will need to be employed.

The various costs have associated risks, some of which evolve over time. These are discussed next. For the proprietary solutions, the vendor support reduces the risk, but the license and storage costs mean these are the costliest, but least risky choice. It can be put in place rapidly.

At the other extreme, doing nothing is, initially, the least costly but most risky initially, but eventually will become the costliest in terms of the impact of loss of information and desperate, costly, attempts to recover.

The use of open-source software may be done in two ways because although the software will be free, using and keeping it running requires effort. It may be possible to purchase this effort from the creators of the software, which adds a certain cost, but reduces risk. On the other hand, the effort may be developed locally, which may reduce the cost, although this is not certain, but large initial risk, which decreases as experience is developed.

The final option is to develop the software in-house, however this is likely to be very costly initially, and take considerable time, to develop the software, but in the longer term the cost will decrease because there will not be a license fee.



VII. Support for the OAIS Information Model

An important point which must be borne in mind is that the software mentioned above are, with one exception, incapable of supporting OAIS conformance except for very simple digital documents in a language known to the Designated Community. In these cases, additional software will have to be created to enable full OAIS conformance, which will add to the nominal costs.

VIII. Records Management

ISO 15489-1:2016 [5] defines a "record" as information created, received and maintained as evidence and as an asset by an organization or person, in pursuit of legal obligations or in the transaction of business. Records management is defined as the field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use and disposition of records, including processes for capturing and maintaining evidence of information about business activities and transactions in the form of records.

These definitions are important for two reasons. Firstly, much of the information identified in the survey concerns legal obligations or transactions, for example the operation of government or commercial enterprises and so are records, which must be efficiently and systematically controlled, not least to gain value from that information. Secondly NAM has a legal obligation to ensure that records are managed properly.

The archive system which the survey has helped to scope must be able to support the records management system, as well as preserve the records and many other types of information. Of course, there are some existing records management and, unless all that information is retrofitted to any new national records management organisation.

IX. Plans for the Next Steps

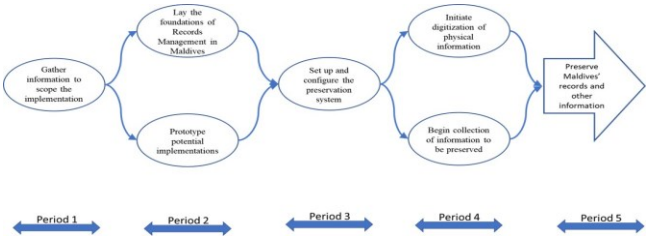
Having developed a good understanding of the information ecosystem in Maldives, it is now possible to propose a practical strategy to enable NAM to fulfil its legal mandate, preserve information in Maldives and secure the information on which Maldives society depends.

The strategy is described in terms of a number of Strategic Objectives, which in turn have associated activities 5 periods which should be approximately 6 months each; the first period covers the work done in the current contract and is included for completeness.

The Strategic Objectives are:

1. Gather information to scope the implementation (the current work reported on here)
2. Lay the foundations of Records Management in Maldives
3. Prototype potential implementations.
4. Set up and configure the preservation system.
5. Begin collection of information to be preserved.
6. Initiate digitization of physical information
7. Preserve Maldives' records and other valuable information.

The relationship between these objectives is illustrated in the following diagram.



X. Conclusions

This paper has described the work which has been done to scope the requirements for a digital archive for NAM, followed by a brief description of the plans for the next few years.

References

[1] USGS, Many Low-Lying Atoll Islands Will Be Uninhabitable by Mid-21st Century, 2018, https://www.usgs.gov/news/many-low-lying-atoll-islands-will-be-uninhabitable-mid-21st-century

[2] Finney, A., Maldives reveals "world's first true floating island city" to cope with rising sea levels, 2022, De Zeen, https://www.dezeen.com/2022/06/27/maldives-floating-city-climate-change-sea-levels/

[3] Reference Model for an Open Archival Information System (OAIS), also known as ISO 14721:2012, is available from ISO (www.iso.ch) and also from the CCSDS site (http://public.ccsds.org/publications/archive/650x0m2.pdf) – updated

[4] Audit and Certification of Trustworthy Digital Repositories, also known as ISO 16363:2014, is available from ISO (www.iso.ch) and also from the CCSDS site (https://public.ccsds.org/Pubs/652x0m1.pdf) or later version

[5] ISO 15489:2016 (Information and documentation —Records management —Part 1: Concepts and principles, reaffirmed in 2021)