

**D**igital information and digital work processes are increasingly used in order to cut costs and gain efficiency. It is, however, far from clear which processes should be put into place to keep digital records alive, reliable and trustworthy. So far, focus on long-term information retention has been determined by two main issues:

- Legal compliance – demanding approximately 10 years retention;
- Digital preservation – related to libraries and archives, where information records are in principle ‘frozen’ for eternity, i.e. only format transitions are required.

However, for a work process to exist for decades, supporting information must not only be preserved but also maintained and kept alive. Long-term work processes are particularly relevant to physical objects that have a lifetime spanning decades, for example:

- A ship needs certificates from a class society which are first issued when newly built and then renewed regularly in connection with inspections or modifications;
- Documentation to support a patient in order to receive proper treatment for his/her lifetime.

In the event of an incident onboard the vessel a class society must document that all work processes and inspection routines were followed correctly. The investigation of an incident will rely on many, often old, documents. An investigation team will typically wish to examine documentation such as:

- Regulations and requirements for the building of the object (those in place at the time of construction);
- Original construction drawings including revisions;
- Modification drawings;



### Records management

# Reading, understanding and trusting 30 year old information records – achievable or utopia?

These days, digital information and digital work processes are used more often to cut costs and gain efficiency. But which processes should be adopted in order to keep digital records alive and reliable?

By **Olga Cerrato, Inger-Mette Gustavsen, Thomas Mestl, Mick Potterton and Jon Øines**

- Inspection reports made by a class society that prove compliance with the class rules.

Furthermore, a class society will need to produce evidence of the proper authorisations at the time of

an incident. Also, the information must be complete, reliable, authentic and sufficiently protected against unauthorised modification.

For normal daily routines, such as a ship inspection, the ship surveyor should have all necessary information

readily available. Updating should be easy and reliable. This can include material such as text, photographs, video and results from various measurement equipment, which may all be in digital form.

practitioner; specialists; possibly several hospitals and their departments. In addition to the availability requirements, this raises other concerns such as authenticity, accuracy and confidentiality

20 years time. Today's main direction is to ensure that information is stored in well-documented formats that allow conversion to whatever the future format may be. Of course, these conversion processes will carry a risk of failure that can lead to loss of information. Furthermore, the conversion process might affect the evidential value (authenticity) of the information. However, keeping old technology alive or emulating old technology on modern equipment is not seen as viable by most experts. Note that not only the format of the information object itself (the 'document') is subject to conversion but also the formats of metadata, semantic information, and presentation information. In many cases maintaining a consistent presentation of the information may not be important since only the content matters whereas in other cases the presentation could be the essential part.

---

**The challenges in information management are often related to the fact that information is the element that has the longest lifespan (apart from possibly a physical object). IT systems may have an expected lifespan of up to 15 years.**

---

The challenges in information management are often related to the fact that information is the element that has the longest lifespan (apart from possibly a physical object). IT-systems may have an expected lifespan of up to 15 years. Exceptions can be found where systems live for more than 20 years but experience shows that the cost of maintaining older systems increases dramatically. Over time, system changes are inevitable, as are changes to organizations, procedures, personnel, roles, ownership, regulatory environments and essentially all parameters related to technology.

In the example of a patient; the journal system, the applied medical systems, formats and other information (for example, images) will be updated over time. Medical staff will have been replaced, hospital departments and even entire hospitals may have been reorganized. Progress in medicine will lead to changes in treatments and the natural developments of society will lead to revisions of patient and treatment laws including information handling (for example, privacy). Throughout the changes, the entire medical history of the patient must be available.

To complicate issues, information is expected to be shared between various actors, such as the patient's general

where potentially several copies of the information may be created over time, with the need for a common understanding of the information.

#### **The main challenges associated with long-term information management**

If we assume that information is stored as records (i.e. the information itself and metadata about the information) and make no further assumptions on technology, organisational structure or work processes, then the following challenges can be pointed out:

##### *Media lifetime:*

Digital preservation (the need to ensure that an information record maintains its availability) is a challenge, yet this can be solved. Storage media and equipment become obsolete and media has a limited lifetime, but moving of records from old to new technologies is mainly a logistics problem. It is important that effective back-up systems are in place.

##### *Format lifetime:*

Another challenge is how to ensure information records can be read over time. The problem of obsolete formats has already been experienced by many who have attempted to access old word processing files. Word processors might not provide retrospective access or compatibility in

##### *Process and organisational lifetime:*

If the objective is to achieve independence from technology, organization and work processes, then a modular structure of the IT-systems might be preferred. Information records are stored in repositories, which offer well-defined interfaces to the systems that support the work processes. Thus, the repository technology, including formats of records, may change as long as the interfaces are kept. Work processes and supporting technology can change, provided they relate to the interfaces. Additionally, repositories should be allowed to exchange information such as when a patient is moved from one hospital to another. Another example may be to extract defined subsets of the information records for export, for example, on sell-out of a part of a company. The Open Archive Initiative (OAI, <http://www.openarchives.org>) works on specifications on these as well as other issues related to repositories, notably also metadata harvesting.

*Preservation of evidential value:*

The evidential value of an information record is dependent upon whether or not its authenticity and correctness can be trusted. These properties are threatened by errors, mistakes and failures, and potentially by intentional

implies a risk of loss of information. Safeguarding information may be the most important aspect but securing this information against unauthorised actions may turn out to be the most challenging. An external attack may not be the real threat, but rather legitimate users

that constituted the evidential value in its original context. This should be migrated to the new format and/or system. The previous evidential value should be stored as part of the metadata of the new information record. The use of a neutral, trusted third party may be advantageous here, leading to an outsourcing and maintenance of repository functionality in order to preserve the evidential value.

---

**As the amount of digitally stored information and the pace of technology development increase, the urgency to address the problems surrounding long-term records management must evolve accordingly.**

---

*Preservation of semantic value:*

Not only the meaning of the information but also the context, in which it is supposed to be interpreted, can change such as rules, laws and regulations. Recently, a project in the Norwegian health care sector aimed at standardising medical terms encountered eight definitions for the seemingly simple term 'hospitalization'. This led to a risk of misinterpretation. All but one of those definitions were altered, illustrating the fact that semantic definitions evolve over time. A patient journal entry that is twenty years old must be interpreted relatively to the terms and definitions in use at that time. Similarly, it should be possible to translate to the corresponding terms and definitions in use today. Structuring information, using ontologies (in some contexts termed as taxonomies or classification schemes) or topic maps could be applied to minimize the deterioration of the semantic value. Ontologies may be defined by an organization itself, but the use of established definitions from standardization bodies, such as the International Organisation for Standardisation (ISO), or common business sector classification schemes are preferred. It is necessary that definitions be shared between involved parties for a successful information exchange. In the long-term perspective an organisation could either rely on external actors to maintain and keep available evolving ontology versions, or do it itself or, alternatively, choose a combination of these. Dependent on the



attacks. Furthermore, information must be collected (and protected) to ensure accountability of events related to the record: either in the information itself, in the metadata, or separately in logs. Other aspects may also have to be considered, such as confidentiality, intellectual property rights, privacy requirements and access. Evidential value will decrease over time. Any activity taken on an information record carries risks because it involves elements such as software, hardware and people: even storage over time alone

attempting to manipulate information for personal gain.

Continuous preservation of evidential value as technology changes is difficult. For example, logs are rarely migrated as old systems are replaced. Format conversion needs to be trustworthy and secure as certain elements may not survive, such as digital signatures. A digital signature is linked to the format of the information signed and this link can be invalidated by a format conversion. In this case, one has to record the information

solution chosen, risk analysis should be carried out.

*Search, retrieval and verification:*

Efficient search functionality is essential in order to retrieve information from a repository. A search may be done based on content, metadata and semantic information, where availability of the latter two can greatly improve the search results. The main challenge in searching for information is identifying the best keywords. New terms are constantly being created whereas older ones may have changed or even disappeared.

**What the future should be**

Reliable records management for long-term work processes is complicated and a proper design of a repository together with specification of interfaces and file formats may constitute a first step. Work processes and their supporting IT-systems must be changed to adhere to the requirements as and when information is stored or updated, notably with respect to collection of metadata, reference to semantic information, and delivery of information using formats that can be directly used or reliably converted into. Similarly, search, retrieval

new tools and methodologies will be developed that might make it easier to manage the records, but unfortunately, new technology will also have new, and currently unknown, challenges.

It becomes increasingly clear that long-term records management will require high focus on developing essential support processes within organisations and that not all organisations (or private persons) will have the resources nor the desire to establish them. Results from the InterPARES 2 project show that in order to fulfil such requirements in daily work, organizations must already be at a mature level with respect to information management. Making the leap from unorganised information handling to a proper long-term records management in one step is simply not possible.

As the amount of digitally stored information and the pace of technology development increase, the urgency to address the problems surrounding long-term records management must evolve accordingly. ■

---

**Reliable records management for long-term work processes is complicated and a proper design of a repository together with specification of interfaces and file formats may constitute a first step.**

---

Thus, preservation of semantic value is essential in order to enhance search efficiency. Some argue that good search functionality is actually all you need. However, many work processes rely on structure of the relationships between information records in order to obtain a complete picture.

and use of the information must be appropriately implemented in the work processes. Legal compliance requirements must be met where required.

It is obvious that there can be no quick fixes to solve the problems associated with long-term records management. As technology evolves,

---

*This paper is written by the DNV members of LongRec project team (<http://research.dnv.com/longrec>) partly funded by the Norwegian Research Council. The LongRec project also constitutes the Norwegian Team of the InterPARES 3 project (<http://www.interpares.org>).*



## need reprints of IK articles?

Please e-mail Graeme Burton with your requests on [gburton@ark-group.com](mailto:gburton@ark-group.com) or telephone +44 (0)20 8785 5957

Published by:

**ark**  
G R O U P