Principles for long-term accessibility of information

eArchiving Reference Architecture

v1.0

CEF eArchiving Building Block, E-ARK3

CEF-TC-2019-3 eArchiving

# Cover Sheet

**Document Status:**

|  |
| --- |
| **Status** |
| Complete |

 **Document Approver(s)**

|  |  |
| --- | --- |
| Name | Role |
|  E-ARK team |   |

 **Document Reviewer(s)**

|  |  |
| --- | --- |
| Name | Role |
|  E-ARK team |   |

 **Summary of Changes**

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Date | Created by | Short Description of Changes |
| V1.0 | 17/01/2020 | István Alföldi | First full version |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Contributors**

|  |  |
| --- | --- |
| Name | Organisation |
| Aadi, Kaljuvee | National Archives of Estonia |
| Aas, Kuldar  | National Archives of Estonia |
| Alföldi, István | Poliphon |
| Holmelund, Frederik | The Danish National Archives |
| Jacobsen, Kristin  | The National Archives of Norway |
| Kruithof, Gijsbert | National Archives of the Netherlands |
| Langeland, Øivind | The National Archives of Norway |
| Lund, Audun | Swiss Federal Archives |
| Merenmies, Markus | National Archives of Finland |
| Nielsen, Anders Bo | The Danish National Archives |
| Proença, Diogo | University of Lisbon |
| Schlarb, Sven | Austrian Institute of Technology |
| Skofljanec, Jože | Archives of the Republic of Slovenia |
| Skog, Björn | ESS Solutions |
| Silva, Hélder | KEEP SOLUTIONS |
| Tømmerholt, Phillip | The Danish National Archives |
| Trani, Silvia | Central State Archives |

#

# Introduction

This document defines and describes the most important digital archiving principles that can serve as the basis of the reference architecture. The *Principles for long-term accessibility of information* document is part of the eArchiving Reference Architecture documentation.

**eArchiving Reference Architecture**

Implementing an effective digital archiving solution requires an understanding of an institution’s legal, organisational, and business context upon which technical solutions can be built. The Reference Architecture workgroup is creating an eArchiving Reference Architecture to simplify the eArchiving Building Block’s uptake. This is:

* aligned with the best practices established by the European Interoperability Reference Architecture (EIRA);
* inspired by Enterprise Architecture conceptual models in digital preservation and information governance (for example, the outputs of the EC-funded SHAMAN project).

The task will start by describing a neutral core eArchiving Reference Architecture. The architecture will map crucial strategy and business components of long-term accessibility and present a reference architecture connected to eArchiving specifications and Sample Software components.

Further, this task aims to contact the DG.DIGIT team responsible for EIRA and work together towards describing the CEF eArchiving Building Block as a component within EIRA. Similar actions are planned at the national level, where it is possible to work together with interested European states (e.g. Portugal, Denmark, Estonia, Norway, Finland) on establishing eArchiving as a component in national architecture.

**Background and purpose**

Principles are general rules that have a normative characteristic within a particular domain. We are aware that the principles do not necessarily reflect the current practice; they instead convey a conviction regarding how the desired situation can be reached. Having everyone following the same set of principles facilitates cooperation and sharing of ideas, practices and building blocks. The principles should be general and flexible enough to be stable over time and give room for variations, but at the same time specific enough to give guidance, have practical value and define some absolute boundaries. From a European perspective, it is vital to get this balance right to accommodate the differences between member states. The principles defined here apply to the domain of long-term accessibility of information. The principles in their current form may overlap and are not hierarchical.

**Our approach**

We are following the guidelines of the Open Group Architecture Framework (TOGAF®). The principles below are classified accordingly into business, information, and application categories. Although the principles are not hierarchical in general, there is an implicit hierarchy. Some of the information and application principles are consequences of the business principles and can be considered derived principles formulated from an information or application aspect.

We are also following the description structure recommended by TOGAF®. The TOGAF® recommendations related to principles can be found in chapter 20 of the [TOGAF® document](https://www.opengroup.org/togaf).

|  |  |
| --- | --- |
| **Name** | Should both represent the essence of the rule as well as be easy to remember. Specific technology platforms should not be mentioned in the name or statement of a principle. Avoid ambiguous words in the Name and in the Statement such as: "support", "open", "consider", and for lack of good measure the word "avoid", itself, be careful with "manage(ment)", and look for unnecessary adjectives and adverbs (fluff). |
| **Statement** | Should succinctly and unambiguously communicate the fundamental rule. For the most part, the principles statements for managing information are similar from one organisation to the next. |
| **Rational** | Should highlight the business benefits of adhering to the principle, using business terminology. Point to the similarity of information and technology principles to the principles governing business operations. Also describe the relationship to other principles, and the intentions regarding a balanced interpretation. Describe situations where one principle would be given precedence or carry more weight than another for making a decision. |
| **Implications** | Should highlight the requirements, both for the business and IT, for carrying out the principle — in terms of resources, costs, and activities/tasks. It will often be apparent that current systems, standards, or practices would be incongruent with the principle upon adoption. The impact to the business and consequences of adopting a principle should be clearly stated. The reader should readily discern the answer to: "How does this affect me?". It is important not to oversimplify, trivialize, or judge the merit of the impact. Some of the implications will be identified as potential impacts only and may be speculative rather than fully analysed. |

***Table 20-1****: Recommended Format for Defining Principles (Source TOGAF®)*

TOGAF® also defines an example set of architecture principles. We consider principles like *Principle 1: Primacy of Principles* or *Principle 7: Compliance with Law* as basic principles fundamental to any approach describing an information system, but we do not copy them into our set unless they are closely related to digital archiving. Even in this case, we do not use the original wording but reformulate it according to digital archiving concepts and language.

Our expected audience is not limited to traditional archives. One of the eArchiving Building Block’s declared goals is to broaden the scope and include a wider range of organisations dealing with mid- or long-term preservation issues such as commercial archives. Accordingly, we intend to use clear language that everybody can understand both within and outside the archival community.

Archival terms and concepts are defined in many related standards (e.g. ISO 16175, ISO 17068 ISO 15489, ISO 14721, ISO 16363), and (to say the least) they are not consistent with each other. We do not intend to create a new set of definitions; on the other hand, we would like to use the archival terms consistently throughout the reference architecture documents. So, in the Glossary, we explain the intended meaning of the most important or most ambiguous terms.

As this document’s title shows, we would like to move the focus from long-term preservation towards the long-term access and reuse aspects of digital archiving. However, that does not mean that the significance of the traditional archival approach or the importance of the long-term preservation aspect is decreasing. (And – to prevent misunderstanding – long-term access definitely would not mean the archive constantly or automatically provides access to the content). Instead, it is a change in the perspective and is in-line with our intention to cover a broader spectrum of institutions providing or using archival services.

**Status of the document**

Although the principles belong to the motivation aspect, we created a separate document with the title *Principles for long-term accessibility of information* because of the significant interest in the work within the community.

#

# Principles

Business Principles

1. Archiving is focussed on use
2. Archiving is preserving information over technologies
3. Archiving ensures the trustworthiness of information
4. Archiving is organising and designing information and processes
5. Archiving is pragmatic, efficient, flexible and sustainable
6. Archiving is an integral part of information management
7. Archiving takes into account rights, permissions and restrictions

Information Principles

1. Archival data are associated with sufficient metadata and documentation
2. Archived information is long-term retrievable
3. Archived information makes use of common vocabularies and definitions
4. Archival information is machine actionable

Application Principles

1. Archival systems should plan for and manage their own obsolescence
2. Archiving can be implemented within any information system
3. Archival systems and their components are interoperable

Business Principles

1. Archiving is focussed on use

|  |  |
| --- | --- |
| Statement | The choices made regarding collecting and preserving information are derived from the purpose for which that information might be used, now or in the future. |
| Rational | This is the most critical principle from which all other principles are derived. Archiving is essentially making choices. These choices determine which information is recorded by the data creators and must be kept accessible over time. The underlying question is what kind of information requires long-term accessibility. These choices are necessary to be able to target the finite resources needed to achieve sustainable accessibility. This explicitly includes answering the question of what time period the information must at least remain accessible. |
| Implications | This principle implies that the archival value of the information is appraised. The appraisal should be risk based, taking into account the needs of current, and if possible, future user groups (designated communities). Within the appraisal process, one should make explicit (as far as possible) what information is to be kept, the level of importance and for how long. When possible, the identified user groups should be involved in the appraisal process.  |
| Notes (if any) |  |

2. Archiving is preserving information over technologies

|  |  |
| --- | --- |
| Statement | The main goal of archiving is to ensure access to information, as long as necessary, without being reliant on originating systems, technologies and media. |
| Rational | Unlike paper-based records, digital records can hardly be kept accessible by default for a long time because systems are replaced, formats become obsolete and physical media deteriorate. To keep the digital records accessible, they must be migrated periodically to both new systems, formats, and media. Because the new system or format often provides different formatting, preservation, or presentation options than the original ones, each of those migration points represents a moment of risk. The goal of digital archiving is to tackle these moments of risk and to prevent losing context and information value over time. This principle should drive the archival decisions behind extracting and reformatting the information when migrating, so the context and information value is preserved in the next period. |
| Implications | This principle implies that when trying to achieve long-term accessibility, it must be clear how preservation actions will affect the context and information value of the preserved content and whether this poses an acceptable risk. More specifically, digital archivists must ensure that they document all aspects of the original technological environment, which are crucial for understanding and reusing the archived information. These are often lost in the process of archiving and long-term preservation (f.ex original database management platforms, GUIs, technical architectures, etc.). |
| Notes (if any) |  |

3. Archiving ensures the trustworthiness of information

|  |  |
| --- | --- |
| Statement | Users must be able to assess the trustworthiness of the preserved information. |
| Rational | Archiving should ensure that all preservation or management actions are transparent and sufficiently documented touphold the trustworthiness of the preserved information.These include actions carried out on both an archive as whole or individual pieces of data and metadata. |
| Implications | * Both preservation and data management activities should support the ability to trace changes in information packages and how those have changed the structure or content of information. Some examples of the documented changes are file format conversions or the splitting or merging of IPs.
* When designing information systems, steps must be taken to ensure that all relevant processes create metadata, ideally automatically
* Steps must be taken to ensure that the metadata and documentation created by people or systems remain available and usable as long as necessary
 |
| Notes (if any) | See information principle 8.) on how to underpin trustworthiness. |

4. Archiving is organising and designing information and processes

|  |  |
| --- | --- |
| Statement | During the design of work processes and information systems, choices about how information is archived are to be made. |
| Rational | One of the most crucial archiving tasks is appraisal - deciding how long information has to remain available. Ideally, these decisions are made during the design of business processes: the owner of the process analyses the pieces of data and metadata which are received or created during the process, organises the information into logical groups (records, entities, etc.) and determines the information value of these groups. Such analysis is the key to making appropriate decisions regarding information life cycle aspects in business process and information system design.  |
| Implications | * Records managers and archivists should be involved through all phases of designing work processes and information systems to ensure that the aspects of information value and lifecycle management are taken into account from the very beginning and considered in all relevant design decisions.
* The organisation of information according to its value allows process owners to implement information systems in a way that respects these decisions. For example, it is possible to introduce functionality to delete less relevant data as soon as possible or set up actions to export valuable long-term information into dedicated archival systems or introduce preservation actions inside the source system.
* The value of information is implementation-independent. Once the value of information has been decided, appropriate technical solutions need to be implemented in all systems where the information is held throughout its lifecycle.
* Not addressing archival considerations during the design of business processes and information systems can significantly increase the resources needed for the continuous maintenance of operations and legal compliance.
 |
| Notes (if any) | Archiving is primarily a designing activity focused on keeping information components of transactions over time considered as relevant. Archiving is focused on “replaying of sequences of actions in order to prove or disprove the state of something or the conclusions reached.” (Upward, Read, Oliver, Evans: Recordkeeping Informatics for a Networked Age, Monash University Publishing, 2018)Please also see application principles 12 and 13 for the implementation aspects.  |

5. Archiving is pragmatic, efficient, flexible and sustainable

|  |  |
| --- | --- |
| Statement | At any given point in time, archiving implements solutions that are fit for purpose, sufficient, adaptable, flexible and sustainable. |
| Rational | Practical archiving is a balance between providing the best possible access to information and, at the same time, is sufficient and efficient. Archival operations should be established in a way that takes into account available human, intellectual and material resources and that respects green sustainability considerations. This also means that an archive should not impose an absolute set of standards and guidelines but rather consider that the measures and processes implemented are pragmatic, sufficient, and politically, technically, organisationally and financially possible.One of the critical risks in digital archiving is the availability of staff skilled in the relevant aspects of archiving, digital preservation, and IT. Therefore, in most organisations, it is reasonable to automate archival processes to the largest possible extent.Another critical consideration regarding sustainability is that valuable long-term information needs to last longer than any given information system. As such, adaptable and flexible system architectures should be preferred, as these allow data owners to update and replace systems or individual components whenever needed efficiently.  |
| Implications | * Information should be categorised into risk and value categories. This allows for the definition of archiving efforts that are appropriate to the value of the information and the availability of resources (see also principle 4).
* Archiving can either be implemented as a component or set of functionalities within larger information systems or as a separate “digital archives”. Such architectural decisions must be made by weighing the cost and efficiency of either solution.
* Archiving should, as much as possible, make use of international standards, specifications, best practices, and tools to lower the total cost of ownership and enhance interoperability. Whenever available, tools which have been certified to support appropriate standards should be preferred.
* Archival decisions, including the ones about storage policies, continuous integrity checks and validation, should be done concerning green sustainability (i.e. the amount of electricity and other resources required to implement these decisions).
* Institutions implementing long-term availability measures must be aware of their internal skills and competencies. Pragmatic decisions need to be made about which archival tasks can be carried out internally versus which can and have to be outsourced from external archival service providers (for an overview of archival tasks, please see the business layer description of this Reference Architecture).
* Records management functionality and the transfer of content between different information systems (either different generations of the same system or between business information systems and archival components) can be a significant cost component. This relates to practically any data production and management systems with long-term preservation consequences at the data producer and at the archival institution sites. Institutions can save a lot when taking these needs into account while designing business processes and business information systems (see also Principle 6).
 |
| Notes (if any) |  |

6. Archiving is an integral part of information management

|  |  |
| --- | --- |
| Statement | Archiving is to be seen as a part of information (lifecycle) management. |
| Rational | Several tasks have to be carried out to ensure long-term accessibility. For example, the information has to be enriched with sufficient metadata, physical storage has to be reliable and accessible, and the file format risks and/or system dependencies have to be analysed and solved. Many of these tasks are not unique to archiving but also relevant to information management in general. As such, “good archiving” does not start once information is delivered to a separated archival repository. It begins before this, with the application of best practices during the creation and initial management of the information.  |
| Implications | * Archiving has to ensure that the standards and tasks it uses are in line with the standards and tasks used in information creation and early management.
* Archiving must collaborate with information management to discuss and propose best-practices that have to be implemented for the early stages of the lifecycle.
* There is no single “correct” way of integrating archiving and information management. Rather, it is about organisational and systems design choices if archiving is separated both functionally and architecturally, or indeed natively integrated into information management organisations, functions, and systems (either way, the two implications above remain valid).
 |
| Notes (if any) |  |

7. Archiving takes into account rights, permissions and restrictions

|  |  |
| --- | --- |
| Statement | Archiving respects legal restrictions to reuse information and ensures that information is accessible only when appropriate rights and permissions are in place. |
| Rational | There can be several different reasons why archived information not made publicly and freely accessible. In many cases, the restrictions fall into the categories of Intellectual Property Rights (IPR) and personal information (GDPR). Elsewhere, trade secrets and confidentiality can be relevant. Regardless of which restrictions apply, archiving has to ensure that the information containing restrictions is correctly identified and appropriate access mechanisms are in place.  |
| Implications | * Archiving has to ensure that its setup is sufficient and appropriate to the level of restrictions applicable to the information.
* In most cases, a security policy, governance processes and actors will be required; this must be capable of both avoiding compromises and reducing liabilities.
* Legislation that poses reasons for restricting information or prescribes methods for providing access can change over time. Archiving has to ensure that its processes regarding managing access rights and restrictions and appropriate metadata can be updated whenever necessary.
* In addition to the legally defined right, permissions, and restriction, an archive must also consider the ethical aspects, such as if sharing information might damage the human rights of those who have a stake in the information itself.
 |
| Notes (if any) |  |

Information Principles

8. Archival data are associated with sufficient metadata and documentation

|  |  |
| --- | --- |
| Statement | Archival data are associated with sufficient metadata and documentation about the data themselves, their organisational and business context, and activities impacting data throughout their lifecycle. |
| Rational | Documentation and metadata are relevant for archives for multiple purposes:* Ensure that archival data has value as information.
* Assess authenticity, integrity and reliability of information.
* Facilitate accessibility and usability of information.
* Enable sustainable preservation of information.
* Achieve transparency and trustworthiness of the archive.

This is achieved by gathering sufficient documentation and metadata along with the data themselves at the right time and maintaining it as long as necessary.  |
| Implications | * Take into account that documentation and metadata can be created throughout the whole lifecycle of information.
* A shared understanding is needed between data producers, archives, and solution providers about which standards, specifications, processes, and best practices should be used to have consistent documentation and metadata.
* Relevant metadata and documentation must be created at the right time, ideally automatically.
* Metadata elements must always be related to one or more purposes and must be captured, validated and stored in a way that makes it possible to support those purposes.
* Stakeholders within the lifecycle must align their processes and systems to the implications above to contribute to the creation and maintenance of documentation and metadata.
 |
| Notes | This principle underpins Principle 3 on the trustworthiness of information.The creation of relevant metadata is a socio-technical issue, including organisational measures such as training and awareness-raising, in addition to systems design and automation. |

9. Archived information is long term retrievable

|  |  |
| --- | --- |
| Statement | Archiving should ensure that information is identifiable, retrievable and renderable. |
| Rational | Digital information is vulnerable because of changes in information preserving technologies. Archiving should ensure that the information is obtainable despite all changes in technology or preservation processes (like migration). |
| Implications | The preservation processes, standards and tools are used to facilitate data interpretation as information for designated communities. Proactive preservation planning and sufficient technical metadata are needed to support migration activities. |
| Notes (if any) |  |

10. Archived information makes use of common vocabularies and definitions

|  |  |
| --- | --- |
| Statement | Archived information is defined consistently, and the definitions are understandable and available to all users. |
| Rational | Information should have an explicit common definition throughout the archive to enable information sharing. The use of common vocabularies will facilitate communication and allow dialogue to be effective.  |
| Implications | * Resources must be committed to this task. It is key to the success of efforts to improve the information environment.
* The archive administration must establish the initial common vocabulary, and these definitions will be used uniformly.
* Whenever a new data definition is required, the definition effort will be coordinated and reconciled with the archive “glossary” of data descriptions. The archive administrator can provide this coordination.
* Widely accepted definitions and vocabularies should be preferred over local and institutional ones.
* The parallel implementation of multiple vocabularies or definitions for the same type of entity or data (i.e. place names, keywords, periods) should be avoided.
* Responsibilities for managing vocabularies and definitions must be assigned.
* While the use of external sources (ontologies, vocabularies) is recommended, archives have to be aware that these external sources have to be managed as well, in the sense of long-term accessibility, versioning, etc. Should the external ontologies and vocabularies change or disappear, in the worst case, the archived information might become unusable or lose authenticity.
 |
| Notes (if any) |  |

11. Archival information is machine-actionable

|  |  |
| --- | --- |
| Statement | IT systems must be able to find, access, interoperate, and reuse information with none or minimal human intervention. |
| Rational | Humans rely on computational support to process information in a meaningful and scalable manner to cope with the increase in volume, complexity, and creation speed of information. This is especially relevant when defining and managing metadata.  |
| Implications | Information in archives should be Findable, Accessible, Interoperable and Reusable for both humans and machines. Specifically, the information in archives:* Is assigned a globally unique and persistent identifier.
* Is retrievable by their identifier using a standardised communications protocol.
* Use a formal, accessible, shared, and broadly applicable language for knowledge representation.
* Is released with a clear and accessible data usage license.

In many cases, archives also make an effort to ensure that appropriate viewers for all data and file formats are available either within their own systems or as “standard practice” within their users’ devices.  |
| Notes (if any) | Taken from FAIR <https://www.go-fair.org/fair-principles/> |

Application Principles

12. Archival systems should plan for and manage their own obsolescence

|  |  |
| --- | --- |
| Statement | Archiving must plan for the obsolescence of the software and hardware it uses and manage the replacement of technical components in a way that does not endanger the authenticity, availability and accessibility of the information it gathers, preserves and provides access to. |
| Rational | Digital archiving is technically implemented as a set of hard- and software. As such, a digital archive is similarly prone to IT obsolescence as any other information system. From the moment development begins, a digital archive must plan for such obsolescence and ensure that the system can be updated regularly in a sustainable and (resource-)effective way, all while ensuring the accessibility and authenticity of archival information are not endangered.Further, a digital archive must be aware that at some point, it might need to replace the whole archival system for a new one. For this case, a clear exit strategy is necessary, including details about how the bulk of data and metadata stored within the archive can be either exported to or accessed by external systems for system migration purposes.  |
| Implications | * An archive must ensure that its technical design and component architecture are well thought through and as future-proof as possible.
* A clear exit strategy must exist, describing how all the information stored in the archive can be migrated into other systems in a standardised and resource-efficient manner;
* Ideally, the exit strategy mandates the use of open standards (for example, the E-ARK specifications).
* An archival system’s infrastructure is reasonably modular and allows individual modules to be replaced or updated independently.
* Standardised communication / APIs are used to interact between different modules of an archive.
* Archived information is stored in a system-agnostic way.
* The technologies used within an archive are widely supported (to ensure that sufficient technical expertise exists).
 |
| Notes (if any) |  |

13. Archiving can be implemented within any information system

|  |  |
| --- | --- |
| Statement | Archival actions can take place within any information system, not only in dedicated digital archiving systems.  |
| Rational | Archiving is foremost about the needs of long-term accessibility, understanding the risks that endanger information accessibility and implementing appropriate measures and technical solutions to mitigate these risks. Traditionally “archiving” has been about transferring content physically to an archive which in turn is a wholly separate entity (i.e. “let’s bring this document to the archive in the cellar”). However, this paradigm is not always valid in the digital age. A growing number of institutions need to use content operationally for extended periods, which might not be possible if the content is transferred to and stored within a separate “digital archive. Institutions must therefore ensure that the risks of long-term accessibility are analysed across all their information systems, and that appropriate measures are implemented in all relevant systems (as opposed to only within a dedicated digital archive).  |
| Implications | * In extreme cases, digital archiving might be implemented entirely within producers’ live information systems.
* An information lifecycle-based risk analysis should be the standard practice in institutions while designing and developing new information systems.
* Appropriate decisions in regard to the technical architecture of digital archiving can only be made if the information has been well organised and its value has been assessed (see Principle 4).
* Systems must be documented in a way that makes it possible to verify that concerns and requirements regarding archiving are met.
 |
| Notes (if any) |  |

14. Archival systems and their components are interoperable

|  |  |
| --- | --- |
| Statement | Archival systems and their components must conform to standards and specifications to promote interoperability for data, metadata, applications, and technology. |
| Rational | Standards help to ensure consistency, simplify the management of information systems, improve user satisfaction, and protect existing investments. When following widely accepted standards, institutions also increase the availability of knowledge and support (i.e. it becomes easier to hire new staff or external consultancy with appropriate skills). Implementing technical standards helps institutions specifically replace systems or components and transfer information seamlessly between different systems (e.g. a live information system, a separate archival component and an external access portal). All of the above helps institutions to maximise return on investment and reduce costs. |
| Implications | * Interoperability standards and industry standards should be followed unless there is a compelling business reason to implement a nonstandard solution.
* Institutions should consider establishing a process for setting standards, reviewing and revising them periodically.
* The existing solutions must be identified and documented.
 |
| Notes (if any) |  |

**Glossary**

Archival terms and concepts are defined in many related standards (e.g. ISO 16175, ISO 17068 ISO 15489, ISO 14721, ISO 16363), many of which are not consistent with each other. We do not intend to create a new set of definitions; on the other hand, we would like to use the archival terms consistently throughout the reference architecture documents. So, in the Glossary, we explain the intended meaning of the most important or most ambiguous terms.

**Archive and Archiving**

In professional terminology, *Archive* can mean either a physical repository, a set of (archived) information or an (archival) organisation. The E-ARK Reference Architecture uses the term *Archive* to refer to any group of people and associated infrastructure which implements long-term digital accessibility measures. As such, an *Archive* can be anything from a single person to specific large archival organisations.

Accordingly, the term *Archiving* refers to the processes carried out by the *Archive,* which can be implemented either as small functional components or processes within larger information environments, as a specific large-scale digital preservation infrastructure or anything in between.

**Information versus Data**

*Information* is considered to be the main asset of archiving. The word *information* is used whenever the archived content and its technical and business context are referred to. *Data* is only used when we refer to the technical aspects and the content bitstreams and files managed by archiving.

**Archival information versus archived information**

The term *archival information* refers to any data element or piece of information anywhere throughout the entire archival process from pre-ingest to access. *On the other hand, archived information* only referred to data elements or pieces of information already stored in an archival repository (e.g. in an Archival Information Package).

**Available versus Retrievable versus Accessible versus Usable**

These terms are not easy to differentiate. They are synonyms in everyday language but used in very specific meanings in several standards. Unfortunately, these particular meanings are inconsistent in those documents. We have decided to use them in the following sense:

If *available*, then data in its purest sense are there. If you have methods and tools, you can *retrieve* them. If it is *available* and *retrievable*, then the information is *accessible* if you have the proper rights to access it (or, in general, if one could have the rights to access it). Finally, information is only *usable* if you can view and interpret it along with its context.

**Trustworthiness versus Reliability versus Integrity versus Authenticity**

One of the critical aspects of any archive is to ensure that the information it keeps can be trusted. More specifically, the E-ARK Reference Architecture understands *Trustworthiness* as a quality of both the Archive and the Information it holds.

For Archives, *Trustworthiness* is about carrying out their processes in a controlled and repeatable fashion.

For Information, *Trustworthiness* does have the sub-characteristics of *reliability*, *authenticity* and *accuracy (*[*https://interparestrust.org/terminology/term/trustworthiness*](https://interparestrust.org/terminology/term/trustworthiness)*)*:

* *Reliable* information is one whose content can be trusted as a full and accurate representation of the transactions, activities, or facts to which it attests and therefore can be depended upon in the course of subsequent transactions or activities.
* *Authentic* information is proven to be what it purports to be, created by the agency with which it is identified, and created when it claims to have been created.
* *Accurate* information is precise, correct, truthful, free of error or distortion, or pertinent to the matter.