Best Archiving Practice Guidance

This document has been published under the auspices of the EU Telematics Implementation Group - electronic submissions (TIGes)

Please note that this document has been published with the aim that agencies and applicants can gain practical experience of archiving eCTD and NeeS electronic Submissions. This document provides agencies with recommendations on archiving submissions, and introduces applicants to preparing quality archivable eSubmissions.

The TIGes considers that through this process authorities and applicants will gain valuable experience of what is required to ensure sustainability of electronic Submissions in the future.

Version 1.0

June 2013

Table of Contents

1. Executive summary

Survey on archives of European medicinal products agencies conducted in 2010 showed that majority of EU National Competent Authorities have not established policies, processes and systems for long-term preservation of electronic submissions. That is why Telematic Implementation Group on electronic Submission (TIGes) in 2011 gathered a specialised subgroup with the mandate of producing Best Archiving Practice guidance.

Best Archiving Practice provides fundamental requirements for long-term preservation of electronic submissions in the National Competent Authorities and European Medicines Agency, and other recommendations that complete archival framework for long-term preservation. Fulfilment of basic requirements shall ensure long life of agencies' archival holdings in the context of rapidly changing technological environment. It is scalable enough to serve as reference framework that meets various national demands, and it serves to harmonise various organisational practices within stated archival framework. In addition, it should provide direction for applicants who are preparing and submitting electronic submissions to the medicinal products agencies.

When applying this guidance, National Competent Authorities and European Medicines Agency can store e-submission in a way which enables long-term preservation in digital archival information systems. These systems can be their own or those provided by third parties, e.g. national archiving authorities.

Digital archival information systems are systems capable of preserving digital objects for long and pre-defined term. Digital archival information systems are not just repositories for storing digital documents, but complete solutions for proactive archival management of (complex) digital objects, their organisational context and provenance, with a goal of continued usage of these objects as authentic and trusted source of information and evidence.

This document contains an overview of archiving guidelines for eSubmissions. Its Appendix 1 describes detailed specifications and recommendations regarding format of documents to be archived, archiving packages and archival information system. Document and archival package guidelines are applicable as such in all agencies whereas solutions to implement an archival information system vary.

2. Purpose of this document

The purpose of this document is to provide agencies and applicants with practical experience of archiving eCTD and NeeS electronic Submissions. The document contains recommendations on archiving submissions, and introduces applicants to preparing quality archivable eSubmissions.

Ad-hoc Drafting Guidance Group Best Archiving Practice considers that through this process agencies and applicants will gain valuable experience of what is required to ensure sustainability of electronic Submissions in the future.

Best Archiving Practice can be adapted according to national legislation and norms. Minimum requirements can be applied in all National Competent Authorities. Full guidance covers also final archiving in digital archival information systems which may not be applicable in all agencies.

Appendix 1 contains detailed description of recommended archival information system, archival packages and the format of documents to be archived.

2. Guidance objectives

An archival framework for long-term preservation is needed in order to create and store e-submissions and other electronic documentation in accordance with international standards. The framework should also establish archiving policies, processes and systems which assure basic compliance with National Competent Authorities' and European Medicines Agency's requirements. The common framework, when applied throughout the submission life cycle, reduces conversion and reformatting efforts and thus work and time needed for archiving.

Expected benefits from Best Archiving Practice are the following:

- Provision of fundamental and common requirements for long-term preservation
- Provision of other recommendations that meet organisational and national differences and harmonise archival practices
- Provision of the framework for long-term preservation archival holdings related to registration of medicinal products
- Model of archival package with mandatory and optional metadata for implementation.

3. Approach

3.1. General

Proactive archival management has become a necessity owning to obsoleting of hardware, media carriers and storage systems as well as operating systems and applications for creating, processing and viewing, review systems, file types and their versions. Entities of origin of archival holdings cannot rely on hardware, software and media stability due to the information and communication technology progress and competition on ICT markets, as well as on perennial support of software vendors' previous and obsolete versions of file types. Furthermore, content prepared for long-term preservation (LTP) should be accessible for digital migration or typed object model conversion, performed by the system.

In order to provide quality archival framework for various national and organisational practices, Best Archival Practice leans on international standards such as

- global and European eCTD specifications
- TIGes' eCTD and NeeS guidance documents
- international standard for describing produced by International Council of Archives
- ISO 14721:2003, ISO 14721:2012 Space data and information transfer systems - Open archival information system - Reference model

- ISO standards for records management and related standards (record management, metadata for records, work process analysis for records)
- Model Requirements for Electronic Records Management.

The basic standard in medicinal products agencies' environment that Basic archival practice is referring to, as well as elaborating it, is ISO standard and reference model for Open Archival Information Systems. In order to have one common system for electronic exchange of regulatory information the authorities need to harmonize the process (including legislation) of e-submission. Preferably open standards should be implemented in order to have the information systems working well for both authorities and for industry.

During the active part of their lifecycle submissions are stored as submission information packages, e-submissions. This is also the format for short-term archiving in the agencies' record and case management systems. For long-terms archiving they are packaged into archival packages. The packages are transferred to an archival information system. It can be the agencies' document management system or some other solution, e.g. a national long-terms archiving service.

Digital archival information systems should create archival information package and its metadata, manage it, enable its long-term preservation, analyze its content for preservation purposes, and preform selected long-term preservation method. All preservation activities should be carefully designed, authorized, documented and tracked.

Preservation activities can affect the integrity of archival packages and in that case system should re-establish it. With technological changes it is expected to migrate digital objects (or to use some other LTP method) and with usage of stated methodologies digital object cannot stay intact. For digital objects the inevitable changes are occurring at the physical or logical level, depending on the selected LTP method, etc. That is why integrity of digital records differs from integrity of paper records, and authenticity should be assured by the system in advance.

Preservation of digital object is a pro-active activity. Integrity of archival information packages means that entity of origin can prove that archival information package was not modified without authorization or accidentally. Documented modifications for preservation purposes are allowed. Authenticity should be assured by the entire system as first and immediate context of the content that we are preserving, as well as by documented activities of transfer to any other adequate and well-established context (e.g. external archive repository or service).

Reference Model for Open Archival Information Systems is a model for information systems capable of long-term preservation of archival holdings. OAIS model is directed towards the development of specific archival systems. OAIS model does not provide a technological solution to long-term preservation issues, but it directs entity of origin to a proactive approach in developing trustworthy digital archival environment.

Open Archival Information System (Open Archival Information System, OAIS) is a standardized (digital) archival information system, or "an archive, consisting of an organization of people and systems, which has accepted the responsibility to preserve information and make it available for a Designated Community. It meets a set of responsibilities [...] that allows an OAIS archive to be distinguished from other uses of the term 'archive'.

The term 'Open' in OAIS is used to imply that this Recommendation and future related Recommendations and standards are developed in open forums, and it does not imply that access to the archive is unrestricted (ISO 14721:2003, 1-11). This Best archiving practice document was created as a concretization of the ISO OAIS model in the domain of regulation of medicinal products. Digital archival information system, with its characteristics of stability in today's dynamic environment, will help current and prospective users obtain confidence in entities of origin of archival holdings. It creates a climate of trust and confidence.

Repackaging of digital object which was submitted into archivable information package that can be preserved over long period of time is the basic aspiration of OAIS-based archive, as well as protection of information authenticity. Agencies as entities of origin of medicinal product authorisation processes' dossiers combine information packages submitted by the marketing applicants and documents that registration processes create and accumulate.

3.2. Current Situation

Archiving practices among EU National Competent Authorities differ a lot. Most have no established policies, processes and systems for long-term preservation of electronic submissions. National archiving legislation and thus practices vary from quite liberal to very detailed.

Survey of conversion of PDF files from e-submissions to PDF/A files, conducted by Ad-hoc Guidance drafting group in 2012 showed various problems in keeping e-submissions. The major issues were presence of variety of copyrighted fonts and impracticality of creating valid PDF/A files. In some cases it was impossible to create such files. From the archival point of view today's e-submissions are not sustainable over long-term. In other words, it cannot be guaranteed that records in e-submissions will be usable in the future and most agencies have legal obligations to keep the records for longer period of time.

3.3. Vision

After the implementation of this Best Archiving Practice guidance each National Competent Authority and the EMA have a system which fulfills the requirements of applied archiving standards. e-submissions are transferred to the archiving system in a such a format which is compliant with or enables conversion to a long-term preservation format. Marketing Authorization Holders can, on the basis of this guidance, help in achieving this target by means of applying the guidance when creating e-submissions.

Some National Competent Authorities may themselves take care of long-term archiving and have a full scale archiving system. Their systems are compatible with Open Archival Information System definitions. Some authorities may only take care of short-term archiving and transfer e-submissions to another authority, e.g. national archiving authorities. They produce and transfer archivable information packages that can be preserved over long period. The guidance can be adapted to local legislations and practices.

Best Archiving Practice guidance acts as benchmark when auditing and evaluating National Competent Authorities in the BEMA process.

Both the European Medicines Agency after Centralised Procedure and Referent Member State in Mutual Recognition Procedure and Decentralised Procedure procedures should keep submissions as archival information packages according to the life cycle of the medicinal product. It should notify involved agencies and Concerned Member States about the final destination of packages (destruction/other), and, finally, it should offer possibility of transfer of packages to their digital archival information systems or other systems. Interoperability should be ensured by using common minimal set of metadata and standardised PDF files.

3.4. Critical Success Factors

- Comprehensive endorsement among pharmaceutical industry and National Competent Authorities
- Applied throughout the e-submission life cycle

- Compliance with international archiving standards and basic national requirements in EU countries
- Adaption of relevant standards and guidelines
- Comprehensive support for proposed file formats
- Compliance with other European s-submission guidelines, validation criteria and practices
- Covers both human and veterinary e-submissions

3.5. Summary of Best Practice Recommendations

Topic	Recommendation
Reference metadata	• Identification related metadata (11 pcs) • Context related metadata (10 pcs) • Content and structure related metadata (2 pcs) • Conditions and access of use (3 pcs) • Every organization may have more elements of metadata that must be used
Content of e-submissions	• Use PDF 1.4 from the new ICH M2 recommendations (April 2011) and the ICH eCTD 3.2.2 specification • or have archival PDF/A (ISO 19005-1, ISO 19005-2, ISO 19005-3) contained in submission information package • no Javascript, 3D or dynamic content (audio, video, other)
PDF files in e-submissions	• Preferably generated from electronic source • PDF version must be at least 1.4 to at most 1.7 • PDF version can be archival PDF/A derived from PDF 1.4 to 1.7 • The maximum size of a file is 100 MB
Fonts used in e-submissions	• Use 14 standard fonts addressed in ISO 32000-1 in section 9.6.2.2 and in the ICH eCTD v3.2.2 specifications: Times-Roman, Helvetica, Courier, Symbol, Times-Bold, Helvetica-Bold, Courier-Bold, ZapfDingbats, Times-Italic, Helvetica-Oblique, Courier-Oblique, Times-BoldItalic, Helvetica-BoldOblique, Courier-BoldOblique. • Use non-copyrighted fonts metrically identical to standard fonts • Base fonts are to be embedded

4. About this document

4.1. Document location

EMA eSubmission website

4.2. Definitions, acronyms, and abbreviations

Term	Explanation
AIP	Archival Information Package is the object of long-term preservation in the digital archival information system
Authenticity, authentic	Authentic means worthy of acceptance as conforming to source claimed. In this sense, authenticity is an attribute of well-preserved archival information packages
Conversion	File format conversion, transfer of content from one file format to another format or format version for preservation and other purposes
Digital archive/digital archival information system	Organisational unit or group of trained employees set of documented and executed policies and procedures, document and records management software and archival hardware, in the service of providing sustainable digital archival holdings of medicinal products agency
DIP	Dissemination information package is a form in which user is provided with information (out of scope of this guidance)
(electronic) Document management system, DMS	System, consisting of software and hardware, capable of managing electronic document. In narrower term: software for electronic document management. It includes functions such as check-in/out, versioning, workflow and other functions required for manipulation of active documents
(electronic) Records management system, RMS	System that upgrades DMS with archival functions. It includes functions such as recordkeeping, adding archival metadata, managing retention periods and automation of retention policy
electronic Submission (eSubmission)	Electronic medicinal products dossier containing sequences and files structured according to ICH and TIGes eCTD and NeeS specifications and guidance documents. In this context eSubmission is a submission information package
Enterprise content management system, ECMS	System that combines functions of document management, records management, collaboration, capture and managing of web content, business process

	management
Font	Complete character set of a single style. Fonts can be non-copyrighted or protected with copyright
Ingest	Function of accepting valid submission information package into digital archival information system of the agency
Integrity	Applied mechanism that ensures that the content of the package was not altered in an unauthorized manner
ISAD (G) General International Standard Archival Description	Standard for archival description issued by International Council on Archives
Long-term preservation, LTP	Long-term preservation - the act of maintaining information, in a correct and independently understandable form, over the long-term (ISO 14721:2012 definition); archiving of (digital) objects in the manner that objects remain sustainable over long-term
Migration	Transfer of information in its entirety for preservation purposes. Types of migration according to ISO standard 14721:2003 and 14721:2012 are • Refreshment; replacement of media (copying) with another media of the same generation • Replication; replacement of media (copying) with another media of different generation • Repackaging with some changes on physical level • Transformation with changes on physical and logical level
Open archival information system, OAIS	“An OAIS is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community.”, ISO standard 14721:2003 and 14721:2012 definition
PDF and PDF/A	Portable Document Format is the file format for representation of two dimensional documents invented by Adobe Inc. in 1993 and released as ISO standard in 2008. PDF/A is ISO standard of PDF file format for long-term preservation of electronic documents
SIP	Submission Information Package, sequence, sequences or dossiers submitted by applicants to agency

4.3. Referenced documents

- OAIS ISO Standard 14721:2003 and 14721:2012
- General International Standard Archival Description – ISAD(G)
- ICH M2 recommendations (April 2011) and the ICH eCTD 3.2.2 specification
- ISO 19005-1:2005, ISO 19005-2:2011

4.4. Document history

Version	Date	Authors
0.3		Arian Rajh Pieter Vankeerberghen Jaana Pohjonen
0.9		Arian Rajh Pieter Vankeerberghen Jaana Pohjonen
0.9F	23.1.2013	Juha-Pekka Nenonen Jaana Pohjonen
1.0	30 May 2013	TIGes

4.5. Appendices

Appendix 1 Best Archiving Practice Specifications