3 July 2014

DOCUMENT
AC/324-D(2014)0008

ARCHIVES COMMITTEE

Directive on the Preservation of NATO Digital Information of Permanent Value

1. At annex please find the approved Directive on the Preservation of NATO Digital Information of Permanent Value. The Directive was developed by the Archives Committee and discussed during the December 2013 meeting, the May 2014 workshop and the meeting of experts in March 2014.

2. The Archives Committee also approved the public disclosure of the Directive to facilitate communication with industry during system procurement and development.

3. The Profile for the Long Term Preservation of NATO Digital Information of Permanent Value (Appendix 1 to Annex 1) will be updated regularly to reflect the development in preservation standards.

4. The Directive and the regular updates of Appendix 1 will be communicated to the NATO Command, Control and Consultation Board (C3B) for inclusion in Volume 3 of the NATO Interoperability Standards and Profiles (AC/322-N(2013)0197).

5. The NATO Archives will develop additional guidance for NATO Civil and Military Bodies related to the creation of submission and archival information packages.

(Signed)
Ineke Deserno

1 Annex

1 Appendix

Action Officer: Ineke Deserno
Original: English

ANNEX 1
AC/324- D(2014)0008

DIRECTIVE ON THE PRESERVATION OF NATO DIGITAL INFORMATION OF PERMANENT VALUE

References

a. AC/324-D(2012)0003, NATO Strategy for the Long Term Preservation of Digital Information
b. C-M(2007)0118, NATO Information Management Policy
c. C-M(2011)0043, NATO Records Policy
d. C-M(2009)0021, Policy on the Retention and Disposition of NATO Information

Introduction

1. The NATO Strategy for the Long Term Preservation of Digital Information (ref a) was developed to provide overall vision, goals, principles and initial steps for the long-term preservation and access of digital information within NATO. It builds upon and supplements the NATO Information Management Policy (ref b), NATO Records Policy (ref c) and the Policy on the Retention and Disposition of NATO Information (ref d)[1].

2. Under NATO policy, the Archives Committee and the NATO Archivist develop, approve, and implement policies and direction to support the retention and long-term preservation of NATO information in all formats.

3. This directive supports the long-term preservation[2] and access to digital information kept in systems supporting the NATO mission, and sets out related principles and requirements for NATO systems in accordance with ref (c).

4. The directive and the attached appendices provide an elaboration of ref (a) on how digital information should be submitted to the NATO Archives for long-term preservation. In addition it provides preservation format guidance for NATO Bodies that hold digital information of permanent value in their repositories.

5. This directive is published by the Archives Committee and is authorised for public disclosure.

Scope and Application

6. This directive applies to all NATO Civil and Military Bodies and NATO Operations who hold information of permanent value[3] and is recommended for NATO Nations and non-NATO Nations when participating in NATO-led Operations.

---

[1] The Strategy is based on the principles outlined in the international standard on long-term preservation and access to digital information, ISO 14721: 2003 Space Data and Information Transfer Systems - Open Archival Information System — Reference Model.

[2] The totality of processes and operations involved in the stabilization and protection of digital objects against damage or deterioration. The goal is to ensure continued access to digital objects or at least to the information contained in them, indefinitely. (definition taken from AC/324-D(2012)0003).

[3] Information of permanent value includes both structured and unstructured data. Structured data is characterized by a presence of a schema, whether explicitly defined, e.g. as an XML schema definition or a formal ADatP-3 message structure, or included implicitly e.g. through a database table structure. Unstructured data does not contain such a schema, e.g. free text descriptions.

7. Long-term preservation is part of the lifecycle for all information of permanent value. It has to be considered during the planning and incorporated during the design, procurement, and implementation of any NATO system.

8. Long-term preservation concerns the information held within NATO systems. The functionality of the systems does not have to be preserved. Contextual information should be included to document the environment from which the information originates.

Principles

9. The following principles apply to the development and implementation of systems:

a. Sustainability: Systems containing information of permanent value shall be developed and maintained taking into account the short, medium and long-term implications of managing such information as well as relevant preservation requirements.

b. Authenticity: Systems containing information of permanent value, and the processes developed to manage those systems, shall be developed and implemented so as to ensure that the authenticity and integrity of the records contained therein is maintained in accordance with ref (c).

c. Accessibility: Systems containing information of permanent value shall be developed and maintained to ensure the long-term accessibility of both the systems and the information which they contain. Any encryption or password protection shall be removed when the information is submitted for long-term preservation.

d. Lifecycle Management: Information of permanent value contained in NATO systems shall be kept in accordance with NATO's records retention and disposition principles and processes for records.

Roles and Responsibilities

10. It is the responsibility of all staff involved in system procurement, implementation and management of NATO systems, and long-term preservation of NATO's digital information to ensure that the principles and requirements established in this directive are adhered to.

11. More specifically, it is the responsibility of:

a. Archives Committee:

(1) to make recommendations and provide direction, on behalf of the North Atlantic Council, on the appropriate measures and mechanisms to preserve NATO digital information of permanent value, and

(2) to define and publish criteria for accepting digital information of permanent value from NATO Information Custodians for long-term preservation.

b. NATO Archivist:

(1) to provide further guidance on the implementation of this directive,

(2) to report to the Archives Committee on any issues which may arise regarding the long-term preservation of digital information,

(3) to ensure that information submitted by NATO Information Custodians is in compliance with the criteria set by the Archives Committee,

(4) to generate Archival Information Packages for digital information of permanent value submitted for long-term preservation (see Appendix 1),and

(5) to establish and manage standard submission and ingress processes for NATO digital information of permanent value.

c. NATO Information Custodians:

(1) to identify digital information of permanent value for long-term preservation, in accordance with ref (d),

(2) to generate Submission Information Packages for digital information of permanent value, following profile in Appendix 1 and any additional guidance provided by the Archives Committee and the NATO Archivist, and

(3) to ensure that any information under their custodianship submitted for long-term preservation is in accordance with the criteria set by the Archives Committee.

d. NATO Host Nations/Implementation Authorities:[4]

(1) to ensure that any systems they implement are in accordance with this directive, and

(2) to ensure that the requirements for long-term preservation of digital information of permanent value can be met by the procured and implemented system.

Requirements

12. The following shall be taken into account when developing and implementing systems containing digital information of permanent value:

a. Content: The system must be able to present digital information of permanent value in a sustainable format suitable for long-term preservation. Content for long-term preservation can be distinguished by seven major types (data sets, text, still images, moving images, sound, geospatial and web archive). The characteristics and the approved sustainable formats for these content types are discussed in Appendix 1.

b. Data Definition: For structured data, the semantics and structural relationships of the Content must be expressed in one of the sustainable formats described in Appendix 1.

[4] NISP projects are implemented by a Host Nation, which would normally be the country on whose territory the project is to be implemented, a NATO agency or a Strategic Command. A Host Nation is the entity which implements a project on behalf of NATO. (Source: NSIP Manual, 2011).

c. Metadata: All Content must be accompanied by descriptive and administrative metadata following a set of metadata elements in accordance with this directive and subsequent guidance issued by the NATO Archivist. Additional preservation metadata will be added by the NATO Archivist.

d. Visualization: A human readable representation or visualization of the Content in its original context can be included as an optional item. This can include screenshots or diagrams that illustrate the environment in which the Content was used.

e. Packaging: the Content, Data Definition and Visualization information shall be packaged into a single digital object, using an approved digital packaging format.

13. System documentation related to information of permanent value shall be maintained and shall be included as part of the package submitted for long-term preservation.

14. Information of permanent value shall be submitted by the NATO Information Managers in their role as Information Custodians to the NATO Archivist in one of the approved sustainable archival formats and packaged as described in Appendix 1.

15. The submission process for information of permanent value for long-term preservation is shown in figure 1.

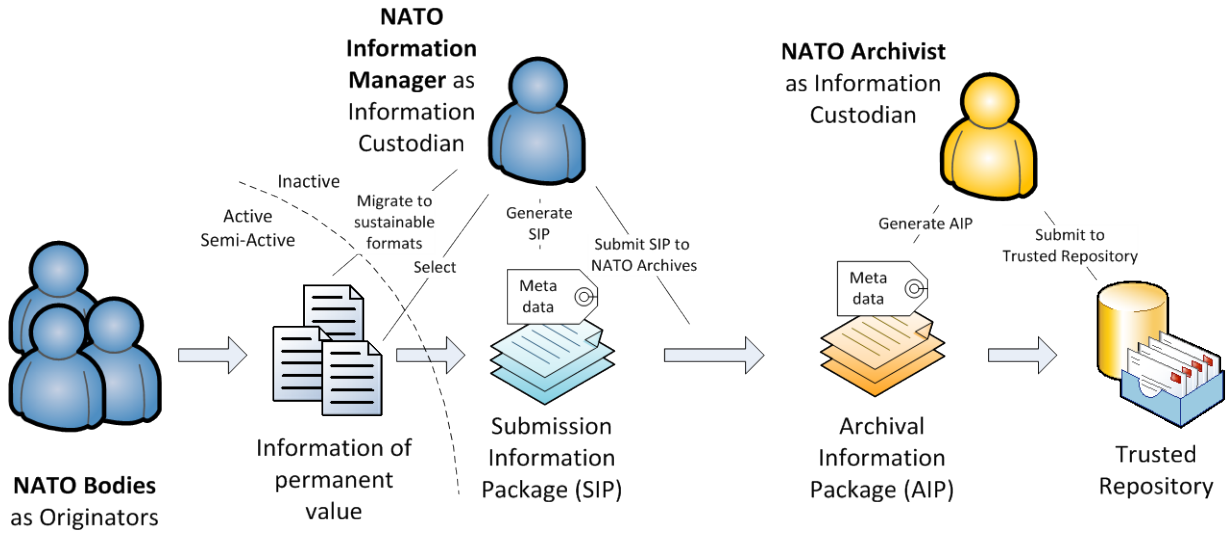

Figure 1: Long-term preservation

16. Appendix 1 contains the profile outlining the approved sustainable file formats for long-term preservation. This appendix will be updated and re-circulated as appropriate.

APPENDIX 1 to ANNEX 1
AC/324- D(2014)0008

PROFILE FOR THE LONG TERM PRESERVATION OF NATO DIGITAL INFORMATION OF PERMANENT VALUE

1. This profile outlines the file formats (Part I) and package structures (Part II) approved by the Archives Committee for the long-term preservation of NATO digital information of permanent value.

2. NATO information custodians shall provide information in these formats and structures to the NATO Archivist.

3. Further guidance on best practice will be issued in the near future. The contents of this profile shall become part of Volume 3 of the NATO Interoperability Standards and Profiles [AC/322-N(2013)0197, 2013].

Part I: File Formats for Long Term Preservation

4. The following sustainable file formats are approved by the Archives Committee for the long term preservation of NATO digital information of permanent value. The formats are ordered by content type. A brief characterization of the generic requirements for the preservation of content is included.

a. Data sets

Data sets are typically collections of individual values or larger coherent structures such as messages. The data set might be an export from a database or the results of an information exchange between systems.

There is typically a structure associated with the data set, either implicitly contained within the data set (e.g. a table structure of an Excel document or a database), or explicitly defined (e.g. as a schema definition)

Content	Requirements	Formats
Data sets (e.g. scientific data) and any structured information not fitting other content types	• Preserve structured and unstructured data for future analysis • Preserve logical structure of dataset as well as syntax and semantics of elements within the dataset • Preserve data types and data structures	• IETF RFC 4180:2005, Common Format and MIME Type for Comma-Separated Values (CSV) Files. • Extensible Markup Language (XML), v1.1 2nd Edition, W3C Recommendation, 29 September 2006. • XML Schema Definition Language (XSD) 1.1 Part 1: Structures and Part 2: Datatypes, W3C Recommendation, 5 April 2012.
Database content		• ISO/IEC 9075 (Parts 1 to-14):2011, Information technology -- Database languages -- SQL.

b. Text

Documents consisting primarily of textual descriptions are the most prevalent and important category of information of permanent value in the NATO context. Text documents might also include embedded diagrams, pictures, or other non-text material. These items shall not be separated from the text and kept as part of the document.

Content	Requirements	Formats
Text documents, including common MS Office document formats (docx, xlsx, pptx).	• Preserve integrity of text, diagram and figures, pagination and navigation (formatting) • Preserve document metadata • Inclusion of fonts, layout information, and indices	• ISO 32000-1 :2008, Document management – Portable document format – Part 1 : PDF 1.7, conformance level : PDF/A-2a
Email (e.g. MS Outlook PST files)	• Preserve email content, including attachments • Preserve complete mailboxes. Important messages might be exported and preserved as individual text documents.	• IETF RFC 4155 :2005, The application/mbox Media Type
Chat (e.g. JChat conversations)	• Preserve message content, including attachments • Preserve complete dialogs per user or multi-user chat room with time-stamps. • Preserve information about users and user groups	• ISO 32000-1 :2008, Document management – Portable document format – Part 1 : PDF 1.7, conformance level : PDF/A-2a • IETF RFC 4155:2005, The application/mbox Media Type

c. Still Images

Still images are visual representations, including photographs, graphs, and diagrams. Still images can be divided into two main types, bitmap (or raster) images and vector images. Bitmap images are typically photographs produced by scanners and cameras at a fixed resolution, while vector images consist of scalable objects. Both types can be combined, e.g. in course of action diagrams where a bitmap image of an area can have symbology vector overlays.

Content	Requirements	Formats
Bitmap/raster images	• Preserve resolution (clarity, colors), scalability, and ability of render the image • Preserve image metadata • Compressibility, preference for lossless compression • Preference for larger resolution	• ISO/IEC 15444-1 :2004, Information technology – JPEG 2000 image coding system, Part 1 (J2K_C_LL, Core Coding, Lessless Compression) • ISO/IEC 10918-1 :1994, Information Technology – Digital compression and coding

		of continuous-tone still images • Adobe TIFF UNC (uncompressed bitmap), part of TIFF 6.0 (1992) • Adobe TIFF G4 (compressed bitmap, part of TIFF 6.0 (1992)
Vector images		• W3C Scalable Vector Graphics (SVG) 1.1, 2011

d. Moving Images

Moving images are digital recordings of still images at a particular frame rate and resolution. A compression is often applied by only capturing the difference between adjacent frames. Moving images are typically combined with audio data and packaged into a common container.

Content	Requirements	Formats
Video files	• Preserve resolution (clarity, colors), scalability, and ability of video • Preserve video metadata, including timecodes and other tagging • Compressibility, preference for lossless compression • Preference for larger resolution and higher audio bitrates	• ISO/IEC 13818-2:2000, Information technology – Generic coding of moving pictures and associated audio information: video / ITU T H.262 (MPEG-2) • ISO/IEC 14496-2:2004, Information technology – Coding of audio-visual objects – Part 2: Visual / ITU-T H.263 (MPEG-4) • ISO/IEC 14496-10:2003, Information technology -- Coding of audio-visual objects -- Part 10: Advanced Video Coding / ITU-T H.264 (MPEG-4 AVC)

e. Sound

Sound files contain recordings of voice or other audio. This includes audio recordings from meetings if they contain information of permanent value.

Content	Requirements	Formats
Audio files	• Preserve resolution (sampling frequency) and depth • Preserve audio metadata	• European Broadcast Union Tech 3285 – Specification of the Broadcast Wave Format (BWF) – Version 2 (2011) (WAVE Audio with LPCM) • ISO/IEC 11172-3:1993. Information technology -- Coding of moving pictures and associated audio for digital

		storage media at up to about 1,5 Mbit/s -- Part 3: Audio. 256 KB/s or higher. - or – ISO/IEC 13818-3 (Second edition, 1998). Information technology -- Coding of moving pictures and associated audio information -- Part 3: Audio. 256 KB/s or higher. (MPEG Layer III Audio Encoding) •

f. Geospatial

Geospatial information is typically produced, used, and contained in geographic information systems (GIS). The information is related to the still image category, as geospatial information consists of bitmap or vector images plus additional attributes associated with particular locations depicted in the image data.

Content	Requirements	Formats
Geospatial information (e.g. GIS data)	• Preserve resolution and scalability • Preserve geospatial metadata	• OGC 07-147r2, Keyhole Markup Language (KML) 2.2.0, April 2008. • OGC 12-128r10, OGC GeoPackage Encoding Standard V1.0, 12 February 2014.

g. Web Archive

The web archive type concern the archival of entire web sites, portals, or parts of them. While some information might be contained in static web pages and is therefore easy to capture, other parts might be dynamically rendered.
Web archives typically contain structured textual descriptions as well as still and moving images.

Content	Requirements	Formats
Web sites and portals	• Preserve structure and content of web, including scripts • Inclusion of external content might be necessary • Preserve metadata associated with content • Dynamic/interactive or user-specific content is problematic	• ISO 28500:2009, Information and documentation -- WARC file format. • IETF RFC 2557, MIME Encapsulation of Aggregate Documents, such as HTML (MHTML)

Part II: Package Structures for Long Term Preservation

5. NATO digital information of permanent value shall be processed by their Information Custodians into single digital information items with associated metadata and packaged into submission and archival information package structures [ISO 14721, 2003].

Submission Information Package

6. NATO digital information of permanent value selected by Information Custodians for long term preservation should be delivered to the NATO Archivist as a Submission Information Package (SIP).

7. The SIP consists of two parts: the actual information packaged as a single digital information item and a set of metadata associated with this item (see Figure 1)Figure 1

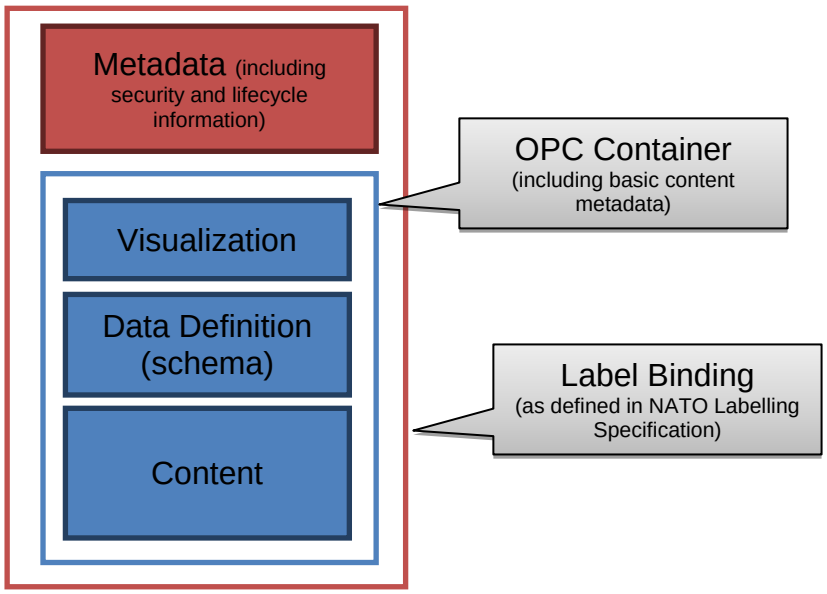

Figure 2: Submission Information Package structure

8. The single digital information item has the following structure:

- Content: Information of one of the seven types listed under 4a-g). For certain types of content, primarily data sets (4a), several pieces of information might be grouped. A schema provided as part of the Data Definition can be used to describe the structure of these groupings. For other types such as documents, images, or recordings, information items shall be included individually. Items might contain other objects that should also be preserved in a sustainable format. For example, an archived email message could have text documents as attachments that should be stored in the sustainable formats listed in 4b). Guidance on granularity and grouping will be provided by the Archives Committee.
- Data Definition: If the Content consists of structured data, a separate Data Definition shall be included that describes the logical structure of the Content. This is primarily

applicable to Content of the types Data Set (4a), Geospatial (4f), and Web Archive (4g). The format of the Data Definition shall be XML Schema 1.1.

- Visualization: A visualization and human readable representation of contextual information is optional. The format used for the context information shall be one of those listed under 4a–g).

9. The individual parts (Content, Data Definition and Visualization) shall be packaged as a single digital information item by using the Open Packaging Conventions [ISO 29500-2:2012, 2012] format.

10. The file name of the submission or archival information package (SIP or AIP) shall follow the NATO Guidance on File Naming [AC/322-N(2010)0025, 2010]. OPC does not define an extension; the .zip extension shall be used for packages for long term preservation.

11. The SIP or AIP shall contain a basic set of metadata for the container. OPC supports a subset of six Dublin Core metadata elements (creator, description, identifier, language, subject, and title) and two Dublin Core terms (created, modified). The elements shall be filled by the Information Custodian when the OPC container for the single digital information item is created. Note that this metadata refers to the container itself, not to its contents. For example, the creation date is the date the container was created, not the creation date of the content.

12. In addition to the OPC container metadata, the Information Custodian will generate a full metadata description for the content of the SIP, including the classification of the single digital information item.

13. The SIP metadata follows the NATO Core Metadata Specification (NCMS) [AC/322-D(2014)xxxx, 2014] and the NATO Labelling Specification [AC/322-D(2014)yyyy/zzz, 2014].. Values for all mandatory elements shall be assigned by the Information Custodian. The NATO Archivist shall reject all submissions with incomplete metadata.

14. No copies of information of permanent value packaged in a SIP and submitted by the Information Custodian shall be destroyed unless the SIP has been explicitly acknowledged and accepted by the NATO Archivist.

Archival Information Package

15. If the content of the SIP submitted by an Information Custodian for long-term preservation are accepted by the NATO Archivist, the SIP will be processed into an Archival Information Package (AIP).

16. The AIP consists of the same structure as the SIP, i.e. the single digital information item for long-term preservation packaged as an OPC container, and the NCMS-compliant metadata information bound to the container.

17. As part of the Ingest process, the metadata supplied with the SIP will be augmented by preservation metadata approved by the NATO Archivist. In addition, NATO Archivist shall become the custodian for the AIP.

18. The preservation metadata will be an extension to the NCMS metadata. The extension shall be based on the PREMIS metadata set [PREMIS Metadata, 2008].

References

[AC/322-D(2014)xxxx, 2014][5]:
Consultation, Command and Control Board (2014, in prep), NATO Core Metadata Specification, AC/322-D(2014)xxxx, NATO Unclassified.

[AC/322-D(2014)xxxx, 2014]:
Consultation, Command and Control Board (2014, in prep), Information Management Directive for Confidentiality Labelling of NATO Information, AC/322-D(2014)xxxx, NATO Unclassified.

[AC/322(2014)xxxx, 2014]:
Consultation, Command and Control Board (2014, in prep), Information Management Guidance for Confidentiality Labelling of NATO Information, AC/322(2014)xxxx, NATO Unclassified.

[AC/322-N(2013)0197, 2013]:
Consultation, Command and Control Board (2013), NATO Interoperability Standards and Profiles, Version 8 (NISP V8), AC/322-N(2013)0197, NATO Unclassified, Releasable to Australia/New Zealand/Singapore.

[AC/322-N(2010)0025, 2010]:
Consultation, Command and Control Board (2010), Guidance on File Naming, AC/3222-N(2010)0025, NATO Unclassified, Releasable to PfP.

[ISO 14721, 2003]:
ISO, Space data and information transfer systems – Open archival information system – Reference model, First Edition, ISO 14721:2003.

[ISO/IEC 29500-2:2012, 2012]:
ISO/IEC, Information technology -- Document description and processing languages -- Office Open XML File Formats -- Part 2: Open Packaging Conventions, ISO/IEC 29500-2:2012.

[PREMIS Metadata, 2008]:
PREMIS Editorial Committee, PREMIS Data Dictionary for Preservation Metadata, Version 2.0, March 2008.

[5] The 2014 AC/322 documents are in the process of being established. Once the documents are approved the applicable document reference number will be inserted under references.