NATIONAL GEOSPATIAL-INTELLIGENCE AGENCY

NGA.STND.0076_0.3_GIMI
2023-08-30

NGA STANDARDIZATION DOCUMENT

Geospatial-Intelligence (GEOINT) Imagery Media for Intelligence, Surveillance, and Reconnaissance (ISR) (GIMI)

Profile of International Standards Organization Base Media File Format (ISOBMFF)

(2023-08-30)

Version 0.3

(DRAFT)

NATIONAL CENTER FOR GEOSPATIAL INTELLIGENCE STANDARDS

CONTACTS

The following point of contact is provided for assistance in understanding the contents of this standard.

Geospatial-Intelligence Standards Working Group (GWG)
NGA CIOT/TAES
ATTN: NGIIS
7500 GEOINT Dr, Springfield, VA 22150
NCGIS-mail@nga.mil

Table of Contents

Table of Figures

Table of Tables

Foreword

This document is for use by all departments and agencies of the Department of Defense (DoD) and Intelligence Communities (IC). The Geospatial-Intelligence (GEOINT) Imagery Media for Intelligence, Surveillance, and Reconnaissance (ISR) Profile of ISOBMFF is a standard for formatting files of digital Still Imagery (including multispectral and hyperspectral multiple band image files), image sequences, Motion Imagery, associated metadata, and imagery related products and exchanging them among members of the DoD and IC as defined by the Executive Order 12333, and other United States Government departments and agencies.

The National Imagery Transmission Format Standard (NITFS) Technical Board (NTB) in collaboration with the Motion Imagery Standards Board (MISB) developed this document using modern technical industry standards and information. The DoD and other IC members are committed to the interoperability of systems using imagery, which include the formatting, transmitting, receiving, and processing of imagery and imagery related information.

This document describes a set of capabilities and requirements for a broad range of imagery forms and types. Resulting from a collaborative US Government and industry effort, it is a common standard for storing and exchanging media content comprising imagery, metadata, and audio information.

Numerous pictures in this document convey certain concepts, aspects, and features of the standard. Any pictures without an annotation for credit are from sources within the Government as part of the development of this document and credit goes to the MISB. Remaining pictures are annotated with their first occurrence within the document.

The NTB and MISB develop, coordinate, review, and plan for GEOINT still and motion imagery standards. They are consensus-based government/industry forums responding to the Geospatial Intelligence Standards Working Group (GWG). The GWG manages geospatial and imagery standards for the DoD and IC encompassed by the National System for Geospatial Intelligence (NSG). This document is under configuration control by the GWG, the NTB, the MISB, and the National Geospatial-Intelligence Agency (NGA). Errata, comments, suggestions, or questions should be addressed to:

National Geospatial-Intelligence Agency (NGA)
Architectural and Engineering Group
Attn: GEOINT & IT Standards Division, Mail Stop S83-T
7500 GEOINT Drive, Springfield, VA 22150

or emailed to ntbchair@nga.mil or misb@nga.mil.

Because contact information can change, use the ASSIST Online database (https://assist.dla.mil) to verify the currency of this address information.

Revision History

Table 1: TBD/TBR Log

Date	Description	Assignee	Date Addressed

Table 2: Change Log

Date	Version	Description	Developer
31 Aug 2023	0.3	*Draft standard with ISOBMFF tutorial content moved to the newly formed ISOBMFF Handbook.*	NGA/TAES
10 Jul 2023	0.2	*Draft standard with motion imagery content included, submitted for internal review and comment*	NGA/TAES
01 Mar 2023	0.1	*Initial draft standard submitted for review and comment to internal NTB & MISB SMEs.*	NGA/TAES

Table 3: Editors' Notes

Date	Change	Rationale

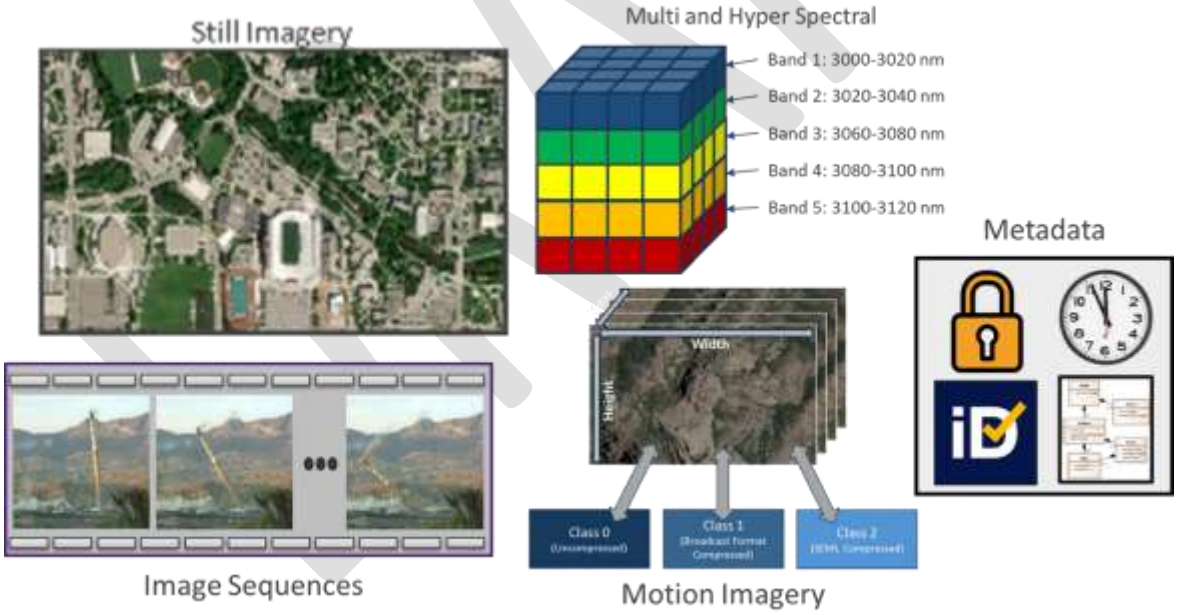
1 Introduction

This standard defines a profile of the International Standards Organization/ International Electrotechnical Commission (ISO/IEC) 14496-12 ISO Base Media File Format (ISOBMFF) video/audio standard [1] and the ISO/IEC 23008-12 High Efficiently Image File Format (HEIF) Still Imagery standard [2]. The Geospatial-Intelligence Standards Working Group (GWG) provides this standard to harmonize Still and Motion Imagery file formats for use within the GEOINT Enterprise Community (GEC).

1.1 Purpose

This standard provides a modular container format for Still Imagery and time-based multimedia data such as image sequences, Motion Imagery, metadata, and audio. Mandatory inclusion of International Atomic Time (TAI) nanosecond precision time stamps, information security marking, and content identification (Content ID) information supports enterprise-wide uniqueness, traceability, effective search/discovery, and releasability. The consistent labeling of time and identification with media items facilitate correlating individual pieces of content with other sources of information within the GEOINT ecosystem of ISR imagery and metadata.

This standard supports implementation of imagery solutions meeting stringent technical requirements in a common, interoperable industry format. At a high level, Figure 1 identifies the primary broad classes of media content covered by this standard.



'Still Imagery' credit: Google Maps, 'Image Sequences' and 'Motion Imagery' credit: White Sands Missile Range

Figure 1: Primary forms of media content covered in NGA.STND.0076.

This standard places constraints on ISOBMFF and HEIF with specific implementation details addressing GEOINT/ISR needs. Requirements for allocating and encoding of metadata, along with requirements restricting the types and usage of codecs per type of media content ensure files conformant to this standard maximize interoperability. NGA.STND.0076 uses the

ISOBMFF branding mechanism to specify file compatibility with this standard. To implement solutions based on this standard, understanding the ISOBMFF and HEIF standards is necessary. When developing software libraries and applications requiring full syntax of the format, the ISOBMFF and HEIF standards must be consulted.

ISOBMFF defines a "base" set of extensible file format capabilities addressing a wide variety of media use cases, but with a focus on video and other timed media. Other derivative documents extend the base to meet evolving needs of media workflows. The industry adoption, extensibility, and wide range of interoperable capabilities makes ISOBMFF the best common media standard for the NSG. Figure 2 lists the ISO/IEC industry standards addressing the primary functions of storage (both Motion Imagery and Still Imagery) and compression.

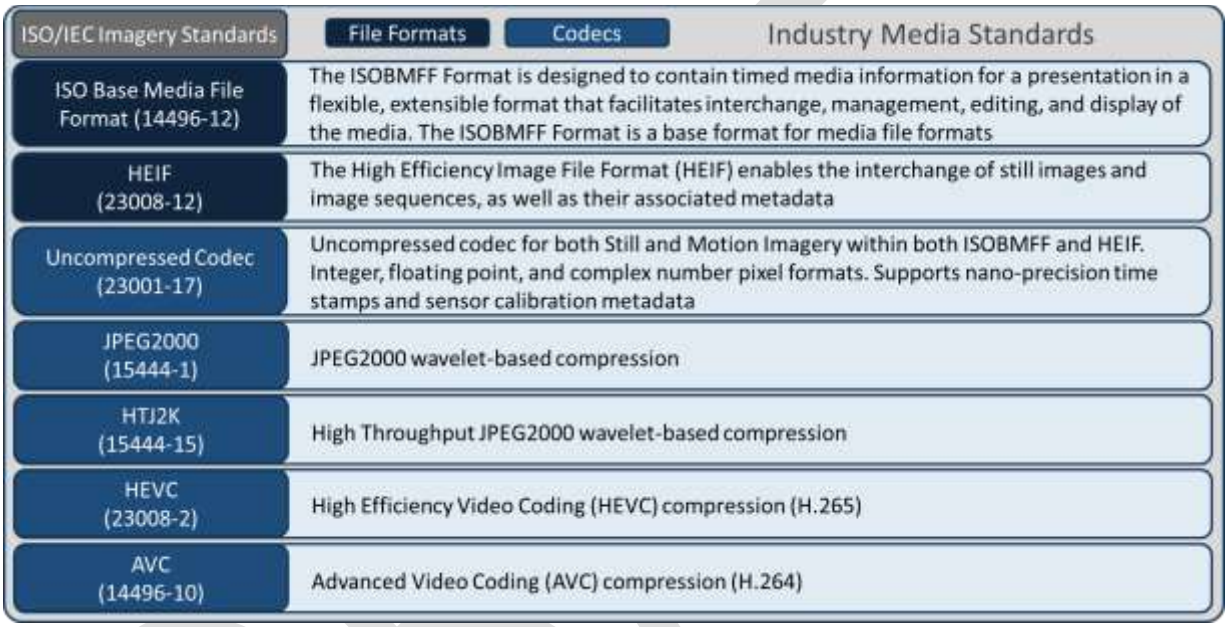


Figure 2: The broad extensibility of ISOBMFF makes it attractive for use in the NSG.

ISOBMFF meets a wide variety of GEOINT use cases, such as high-throughput recording, processing, exploitation, analysis, cloud-based integration, automation, dissemination, and efficient operation with common tools and applications.

ISOBMFF provides a foundational set of capabilities from which derivative profiles and extensions enable meeting application needs in the ever-expanding commercial media world. This ability to define derivative profiles enables the creation of a standard tailored to GEOINT applications. The ISOBMFF and its associated software libraries apply to a wide range of use cases, thereby simplifying the software infrastructure, and providing a framework for a high level of interoperability across many users and applications. Such a framework enables Still Imagery and Motion Imagery to merge into a unified container and workflow architecture.

This standard directs use of Key-Length-Value (KLV) encoding, based on SMPTE 336:2017 [3], for metadata to facilitate extensibility and bit efficient encoding. It encourages use of the Motion Imagery Metadata (MIMD) Model [4] where appropriate. In certain specific metadata implementations, the application may specify a different encoding scheme, such as the use of XML coding for information security marking metadata, which is derived from Office of the

Director of National Intelligence (ODNI) guidance related to the specification for ISM.XML [5].

1.2 Scope

The scope of this standard is to support the storage and containment of uncompressed and compressed forms of imagery, with supporting metadata, for all types of NSG ISR applications. The scope includes:

- Containment of imagery with a broad range of array sizes, frame rates, one-to-many components (bands), variability in color formats, and bit depths. NGA.SNTD.0076 addresses the following forms of "monolithic" two dimensional, rectangular forms of imagery:
  - Still Imagery (NTB defined)
    - Panchromatic
    - Color
    - Multispectral
    - Hyperspectral
  - Motion Imagery (MISP conformant)
    - Class 0 (uncompressed)
    - Class 1 (broadcast format compressed)
    - Class 2 (non-broadcast format compressed – machine vision, infrared, scientific)
  - Image Sequences
  - Image Collections
- Imagery codecs
  - NGA.SNTD.0076 constrains the use of imagery compression codecs for both Still Imagery and Motion Imagery
- Metadata
  - Methods for the encoding and carriage of metadata internal to a file
  - Methods for referencing metadata carried external to a file

The scope of this standard does not include:

- Introductory material for ISOBMFF and HEIF, which is addressed in the NGA Standard Information/Guidance document NGA.SIG.0045 ISOBMFF Handbook [6].
- Streaming
- Large Volume Motion Imagery (LVMI) - multiple "ganged" cameras. Currently addressed by NGA.STND.0044 MIE4NITF.

Finally, this standard supports Audio data but does not mandate its use.

1.3 Document Organization

NGA.SNTD.0076 supports the encapsulation and interchange of GEOINT Still Imagery, Motion Imagery, and associated metadata. It provides:

- Introductory information for context,
- A detailed description of the standard and associated GEOINT implementation details,
- Conformance requirements,
- Information necessary to understand conceptually how to implement this standard for GEOINT applications,
- Annexes providing additional information considered to be useful for developers and implementers.

2 Approach to Conformance

This standard utilizes the Easy Approach to Requirements Syntax (EARS) [7] to explicitly state the requirements for conformance to NGA.STND.0076. EARS addresses ambiguity in the creation of requirements and helps to facilitate conformance and testing in an efficient and effective manner.

A conformant implementation in accordance with this document is one including all mandatory provisions (requirement "shall" statements).

2.1 Order of Precedence

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows:

1. Normative prose in a requirement statement
2. Tables
3. Formal languages
4. Figures
5. Any other language forms.

In the event of a conflict between the text of this document (NGA.STND.0076) and text in either of the references for the ISOBMFF standard (ISO/IEC 14496-12) or the HEIF standard (ISO/IEC 23008-12), the ISO/IEC standards take precedence. Nothing in this document supersedes applicable laws and regulations unless a government program obtains a specific exemption.

2.2 Requirements Lists

The accumulation of all requirements defines conformance to this standard.

2.3 Conformance Testing & Certification

The collection of requirements listed in this document forms the scope of conformance testing and certification for this document. This document references multiple normative commercial media standards: therefore, to ensure maximum interoperability across systems and solutions, a complete conformance testing and certification process must also account for validating conformance to the relevant normative references for a specific implementation of this standard.

2.4 Conformance Principles

Developers of conformant tools and applications for the production, consumption, processing, editing, and display of imagery and related GEOINT metadata, test their capabilities for relevant functionality of conformant files within configurations defined by their program and acquisition requirements. NGA.SNTD.0076 contains a wide range of configurability, as well as dependency on other associated format, encoding, and metadata standards. Conformance testing must demonstrate conformance with requirements and successful interoperability of all relevant variations of content depending on intended implementation use.

Systems and applications generate, consume, and manipulate imagery and metadata to meet program specific requirements. Generated files must be conformant to all or a defined subset of

NGA.STND.0076 requirements and must not violate any requirements from this document. This standard does not require systems and applications to implement all capabilities delineated in this document.

For instance, a system producing only Motion Imagery does not need to claim conformance to NGA.STND.0076 Still Imagery requirements when a government program only needs Motion Imagery capabilities.

3 References

The following documents, in whole or in part, are normative references for this document as they are required for this application. Dated references indicate only that specific version of the document applies. Otherwise, if a reference is undated, the latest approved version, including amendments, of the reference applies.

[1] *ISO/IEC 14496-12, Information technology - Coding of audio-visual objects - Part 12: ISO base media file format.*

[2] *ISO/IEC 23008-12:2022, Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 12: Image File Format.*

[3] *SMPTE ST 336:2017 Data Encoding Protocol Using Key-Length-Value.*

[4] *MISB ST 1903.2 Motion Imagery Metadata (MIMD): Model, 2022.*

[5] *Office of the Director of National Intelligence. XML Data Encoding Specification for Information Security Markings (ISM.XML),* 2022.

[6] *NGA.SIG.0045 Standard Information/Guidance (SIG) ISO Base Media File Format (ISOBMFF) Overview for NGA Applications.*

[7] Mavin, *EARS (Easy Approach to Requirements Syntax),* Proceedings of the 2009 17th IEEE International Requirements Engineering Conference, 2009.

[8] *MISB ST 0603.5 MISP Time System and Timestamps (October, 2017).*

[9] *ISO/IEC 23001-17, Information technology — MPEG Systems technologies — Part17: Uncompressed video and images in ISO Base Media File Format.*

[10] *ISO/IEC 15444-16:2021, Information technology - JPEG 2000 image coding system — Part 16: Encapsulation of JPEG 2000 images into ISO/IEC 23008-12.*

[11] ITU-T H.265 (V8) (08/2021) High efficiency video coding.

[12] ISO/IEC 14496-10:2020 Information Technology - Coding of audio-visual objects - Part 10: Advanced Video Coding.

[13] ITU-T H.264 (V14) (08/2021) Advanced video coding for generic audiovisual services.

[14] ISO/IEC 14496-15:2019 Information technology - Coding of audio-visual objects - Part 15: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format.

[15] MISB ST 1902.2 Motion Imagery Metadata Model-to-KLV Instructions, Jun 2022.

[16] MISB ST 0107.5 KLV Metadata in Motion Imagery, Oct 2021.

[17] T. Edwards, "A Uniform Resource Name (URN) Namespace for the Society of Motion Picture and Television Engineers (SMPTE)," February 2008. [Online]. Available: https://www.rfc-editor.org/info/rfc5119.

[18] *NGA NGA.RP.0001_1.0.0 NSG Recommended Practice for Universally Unique Identifiers*, 2013.

[19] *RFC4122, A Universally Unique IDentifier (UUID) URN Namespace, July 2005, P. Leach, M. Mealling, R. Salz]*.

[20] *MISP-2023.1 Motion Imagery Standards Profile*, MISB, October 2022.

[21] *MISB ST 1603.3 Time Transfer Metadata, Oct 2022.*

[22] MISB, *MISP-2023.1 Motion Imagery Handbook, 2022*, 2023.

[23] *MISB ST 1001.1 Audio Encoding, Feb 2014.*

[24] ISO/IEC 14496-3:2019 Information technology - Coding of audio-visual objects - Part 3: Audio.

[25] *ISO/IEC 15444-1:2019, Information technology - JPEG 2000 image coding system: Part 1: Core coding system.*

[26] *ISO/IEC 15444-15, Information technology — JPEG 2000 image coding system — Part 15: High-Throughput JPEG 2000.*

[27] *ISO/IEC 23008-2, Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 2: High efficiency video coding.*

4 Terminology

The purpose of providing terminology is to clarify the meaning and intent of words used in the specification by giving the reader instruction and context on how they are to be interpreted, with an emphasis on terms that might otherwise be ambiguous.

4.1 Acronyms & Initialisms

The following lists the acronyms and initialisms used in this standard.

AVC	Advanced Video Coding
GEC	GEOINT Enterprise Community
GEOINT	Geospatial-Intelligence
HEIF	High Efficiency Image File Format
HEVC	High Efficiency Video Coding
HSI	Hyperspectral Imagery
HTJ2K	High Throughput JPEG 2000
ISO/IEC	International Standards Organization/ International Electrotechnical Commission
ISOBMFF	ISO Base Media File format
ITU	International Telecommunication Union
J2K	JPEG 2000
KLV	Key Length Value
LSB	Least Significant Byte
LVMI	Large Volume Motion Imagery
MIMD	Motion Imagery MetaData
MISB	Motion Imagery Standards Board
MISP	Motion Imagery Standards Profile
MPEG	Moving Picture Experts Group
MSB	Most Significant Byte
MSI	Multispectral Imagery
NAL	Network Abstraction Layer
NGIIS	Next Generation ISR Imagery Standards
NITF	National Imagery Transmission Format
NITFS	National Imagery Transmission Format Standard
NSG	National System for Geospatial-Intelligence
NTB	NITFS Technical Board

RDF	Resource Description Framework
SAI	Sample Auxiliary Information
SDO	Standards Development Organization
SEMI	Scientific/Engineering Motion Imagery
ST	Standard
TAI	International Atomic Time
URI	Uniform Resource Identifier
URN	Uniform Resource Name
W3C	World Wide Web Consortium

4.2 Terms and Definitions

Note: When multiple normative references define the same term but with different definitions, this document explicitly lists both definitions and intended usage of the term from each reference. This helps avoid confusion when definitions are not completely compatible. An example is the term 'sample'. MPEG uses sample to indicate the data associated with a single time, such as a collection of pixels in a frame of video. The MISB defines sample to refer to the measurement of a single pixel. The definitions for this situation qualify these terms as follow:

- Sample (MPEG): MPEG-sourced definition
- Sample (MISB): MISB-sourced definition

If there is a discrepancy between a definition listed here and the referenced source document, the definition of the source document takes precedence.

Terms relevant to NGA.STND.0076 (this document):

Term	Definition
4CC	An ASCII encoded "four-character code" indicating data formatting requirements. The code is a 32-bit big-endian unsigned integer with the first character in the MSB, last character in the LSB.
Advisory timing	In HEIF files containing sequences, timing may be associated with the timing at collection, or the timing intended for display/playout. HEIF viewers may ignore timing and display sequence content in a static gallery. MPEG refers to this nature of sequence timing as advisory.

Asynchronous	Samples not occurring at 'precisely' the same time. The time window defining "at precisely the same time" is problem dependent and determined by application requirements.
Asynchronous Tracks	Two tracks with samples that occur at different times on the movie timeline.
Band	Within an image, a two-dimensional (row/column) array whose elements represent one component of a tuple value for a pixel.
Block	A rectangular array of pixels comprising a sub-region of an image or a full image. (Synonymous with tile).
Box	(MPEG) An object-oriented container with a unique type identifier and length.
Class 0 Motion Imagery	(MISB) *Class 0 Motion Imagery* represents the collective requirements for non-compressed Motion Imagery, Metadata, Audio, and suitable Containers. The non-compressed Motion Imagery of Class 0 Motion Imagery allows for non-constrained Image Characteristics (see MISP Sec. 2.1.4.1), such as unrestricted Pixel Value Range, Number of Bands, Number of Pixels per Image, and Number of Images per second. Examples include 14-bit infrared, raw Bayer, 3-Band High Definition. Class 0 Motion Imagery converts to Class 1 Motion Imagery or Class 2 Motion Imagery by applying compression.
Class 1 Motion Imagery	(MISB) *Class 1 Motion Imagery* represents the collective requirements for compressed Motion Imagery, Metadata, Audio, and suitable Containers. Class 1 Motion Imagery is applicable when delivering monochrome and color Motion Imagery in cases where the transmission bandwidth prohibits the use of Class 0 Motion Imagery. MISP Section 2.1.4.1 constrains the Image Characteristics of Class 1 with limits in Pixel Value Range, Number of Bands, and Number of Images per second. Example: H.264/AVC compressed airborne Motion Imagery. Class 1 Motion Imagery, based on standards from commercial SDO's, uses three Bands of color.
Class 2 Motion Imagery	(MISB) *Class 2 Motion Imagery* represents the collective requirements for compressed Motion Imagery, Metadata, Audio, and suitable Containers. Unlike Class 1 Motion Imagery, Class 2 Motion Imagery allows for a non-constrained set of Image Characteristics. Examples include high frame rate scientific imaging, Large Volume Motion Imagery, high bit-depth compressed infrared. Class 2 Motion Imagery is based on standards from commercial SDO's and non-commercial standards unique to a governing organization.

Class 3 Motion Imagery	(MISB) *Class 3 Motion Imagery* represents sources external to the NSG, such as cell phones, mobile devices, and surveillance cameras, which may use formats, compression, containers, and other technologies that do not conform to the MISP. The requirements specified for Class 3 Motion Imagery address conversion of Motion Imagery, Metadata, Audio and Containers to meet Class 1 Motion Imagery requirements.
Collection	A set of items without any required relationships in time, space, and format.
Component	A homogenous element representing a specific aspect of measurement and comprising all or part of an image pixel. Frequently related to distinct wavelength regions and sometimes referred to as a band, layer, or channel of an image. A part of a sample (MISB) representing a specific color band (e.g., Red, Green, or Blue) or chroma channel.
Content ID	A universally unique identification parameter providing enterprise-wide search and discovery ability.
Entity	1) An item instantiated in a MetaBox 2) A track instantiated in a MovieBox.
Entity group	A grouping of items and/or tracks sharing a particular characteristic or having a particular relationship.
Frame	(MPEG) Two-dimensional rectangular array of pixels contained in the sample data. (MISB) A two-dimensional array of regularly spaced pixels in the shape of a rectangle indexed by rows and columns along with a start time and an end time of each pixel.
GIMI File	A file conformant to NGA.STND.0076.
Image collection	A set of images without any required relationships in time, space, and format.
Image Sequence	(MPEG) A sequence of coded images which may be associated with advisory timing and where images may use inter prediction.
International Atomic Time (TAI)	A high-precision time scale derived from hundreds of precise atomic clocks from around the world and maintained as closely as possible to the *Système International* (SI) second. Current practice achieves a maximum deviation of approximately one second every 100 million years.
Inter-prediction	A form of coding leveraging knowledge of form and location of content in one or more other coded images.
Intra-coded image	An image coded with dependencies only on the specific image.

Key-Length-Value (KLV)	A form of metadata encoding based on SMPTE 336:2017.
Large Volume Motion Imagery (LVMI)	A form of Class 2 Motion Imagery characterized by "very large" array sizes and frame rates that typically fall outside the norm of broadcast formats. Systems may generate the large arrays of image data either from single, large monolithic image sensors, or the compositing of multiple images from independent sensors.
Metadata	Data that describes and gives information about other data (including imagery), thereby providing context about the data.
Metadata Sets	Collections of individual metadata items addressing a specific topic of metadata coverage.
Motion Imagery	(MISB) A sequence of images, that when viewed (e.g., with a media player) must have the potential for providing informational or intelligence value. This implies the images composing the Motion Imagery are: (1) generated from sensed data, and (2) related to each other both in time and in space. The sequence of images enables detecting motion of the sensor and/or objects within.
Motion Imagery Metadata (MIMD)	Contextual information, including temporal, platform, payload, sensor, command, automated processes, exploitation, and security, about Motion Imagery.
Movie	(MPEG) The structural data describing a timed presentation.
Non-compressed	See uncompressed.
Panchromatic	A single band covering a broad range of EO wavelengths, usually used in context of collecting information from much of the visible spectrum and sometimes part of the near infrared spectrum.
Persistent Metadata	See Static Metadata.
Pixel	(MISB) A combination of one or more individual numerical values, where each value is a sample (MISB). (MPEG) Smallest element of an image, comprised of one or more components.
Raster	Images defined as a two-dimensional array of pixel values in row and column format.
Rate Partitioning	A method of grouping multiple sources of sample data into tracks with "like" sample rates. For instance, sampling and capturing independent data sources at 10 Hz, 30 Hz, and 50 Hz can be "rate partitioned" into three separate groups of tracks with update rates of 10 Hz, 30 Hz and 50 Hz.
Raw	Uncompressed imagery where no corrections or adjustments have been made to the pixel values.

Reader	An application that reads and processes or displays the media within a GIMI File
Sample	(MISB) Data representing a measured phenomenon such as light intensity. (MPEG) All the data associated with a single time.
SEMI	(MISB) Scientific/Engineering Motion Imagery (SEMI): A form of Motion Imagery that is typically Class 0, but can also be Class 2, where the Motion Imagery Sample Value Range may be more than 10 bits per Band, there may be more than three color/spectral Bands, or the Images per second may be in the hundreds or thousands. There may be a need for numerically lossless compression, as well as "perceptually lossless" compression, where the Motion Imagery compression does not inject artifacts detectable by human observation.
Sensor Data	Information captured and collected from a sensor representing the measurement of some physical phenomena, such as photons, time, gimbal angles, geodetic location, slant range, pressure, temperature, etc. In this document, the term "data" is synonymous with "sensor data." Sensor data can also be metadata if it helps to define additional meaning or context of other data.
Sequence	A set of related items in a specific order.
Sequence track	A sequence of items placed on a timeline, via a track.
Sidecar	A file containing only metadata and pertaining to imagery in a separate file.
Slide show	A series of potentially unrelated images played as a sequence.
Static Metadata	Metadata values that do not change over a time period deemed sufficient for a given use case. Also known as Persistent Metadata. Persistent metadata does not require a timestamp.
Still Image	(NTB) A two or more-dimensional rectangular array of pixels derived from sensed phenomena indexed by row and column.
Still Imagery	A single image lacking the repetitive temporal aspect of Motion Imagery. Without additional images, determining motion in a scene is much more difficult.
Synchronous	Samples occurring at 'precisely' the same time. The time window defining "at precisely the same time" is problem dependent, may depend on the bandwidth of the sensor sampling the phenomena, and is determined by application requirements.
Tile	(MPEG) Two-dimensional rectangular array of pixels within a frame. (Synonymous with block)

Timed Metadata	Dynamic metadata captured with a time stamp facilitating alignment with other media tracks.
Uncompressed Image	A form of imagery where the image content pixel values have a direct, one to one association with each sensor measurement location allowing direct query of each sample's (MISB) numerical value.
Unframed	Storing content in its raw form, without any control/structure bits or bytes encapsulating the content.
Video	An ordered series of image frames with each frame assigned an increasing presentation time, where the presentation time is a relative time.
Writer	An application that creates or edits a GIMI File

5 General Overview

This document defines a standard based on ISOBMFF and HEIF for use with Still Imagery (including multi and hyperspectral imagery), image sequences, Motion Imagery, and associated metadata within the NSG community. Together they provide a single foundation of general capabilities enabling this implementation of a common architecture for both Still Imagery and Motion Imagery.

5.1 Overview of the GEOINT Imagery Media for ISR Standard

NGA.SNTD.0076 provides an integral foundation for GEOINT ISR imaging workflows. It addresses file-based imagery activities and facilitates applications ranging from collection to dissemination and archive. Additional companion profiles, formats and containers address other aspects of media handling, such as streaming and metadata "sidecar" files, to create a comprehensive capability for the NSG.

NGA.STND.0076 supports containment of a wide range of imagery types, categorized by:

- Temporal makeup – still, motion, sequences, slide shows, flip books, etc.
- Spatial makeup – monolithic, tiled/gridded, overlays, etc.
- Spectral makeup – monochrome, panchromatic, color, patterned, MSI, HSI, ultraviolet, visible, near infrared, short wave infrared, mid-wave infrared, and long-wave infrared, etc.
- Numerical encoding – integer, floating point, complex number
- Compression – raw, uncompressed, lossless compression, lossy compression, intraframe compression, interframe compression

NGA.SNTD.0076 also supports the carriage of metadata to facilitate NSG needs. Conformant files must include:

1. Information security marking
2. Content Identification (Content ID)
3. High precision timing information

To meet security and access requirements a file includes information security marking. NGA.STND.0076 invokes the ODNI ISM.XML specification and directly encodes and carries security information defined by the specification. NGA.STND.0076 only specifies how to allocate and where to insert the security information into a file. The ODNI standard evolves independently, and existing tools supporting the ODNI standard can directly support NGA.STND.0076 applications.

Content IDs are Universal Resource Names (URNs), enabling enterprise-wide uniqueness for each piece of content within the GIMI file.

Media timestamps are Nano Precision Time Stamps with a TAI epoch and include clock pedigree and status metadata.

To associate metadata with imagery content, and to enable global search and discovery of content, the standard mandates attaching both Content ID and timing metadata to imagery and other metadata inside files. Additionally, these two pieces of information enable external

linking to GIMI file content. [8]External files and systems may link or reference (or "point") to data within the GIMI file.

References for metadata can be made to the entire contents of a file, Still Image items, image sequence and Motion Imagery tracks, Motion Imagery frames, groups of image items and/or tracks, image tiles, image pixels, and user defined regions of interest. Where appropriate, external referencing enables carriage of content outside the file.

NGA.SNTD.0076 optionally carries additional metadata in a default form as KLV sets and packs. The governing bodies within the GWG imagery domain define the KLV sets and packs. When this profile requires it, the profile allows additional encoding methodologies, such as XML coding of information security markings.

While the carriage of metadata in separate files is a primary design feature of the overall architecture, implementers can encode and carry metadata inside the image file where appropriate. This is desirable for certain use cases, such as product files, where imagery and metadata content support an intended playout of imagery with synchronized metadata and graphical overlays. In an archival case, a single file may include both raw sensor information and processed results to provide and document full provenance.

NGA.SNTD.0076 specifies how to implement and utilize ISOBMFF and HEIF capabilities for both producer and consumer applications. From a high-level perspective, this standard specifies the following primary forms of content storage:

Items:

- A static block of data, which may be an image or a set/pack of metadata.

Tracks:

- A timed sequence of related media samples.

Image Items:

- The storage of a single coded image or a single derived image.
- NGA.SNTD.0076 refers to and stores still images as items. It requires independently coded images with no decoding dependence on any other item.
- For tiling Still Imagery, the standard defines an "image grid" as a derived image item formed through the collection of independently coded images arranged in a 2D grid (or tiled) arrangement.
- A file can include any number of image items.

Metadata Items:

- The storage of static blocks of metadata.
- NGA.SNTD.0076 refers to and stores each separate block of static metadata as a metadata "item". These are declared and described inside a MetaBox and stored (typically) in the MediaDataBox ('mdat'). Metadata can also be stored in the ItemDataBox ('idat'), internal to the MetaBox or in a container external to a file.
- The file header (MetaBox) defines the encoding syntax and semantics for the metadata. This standard utilizes both Key-Length-Value and Extensible Markup Language encoding.
- The metadata describing an image item (e.g., camera ID, focal length, location information, etc.) links to a specific image item through a reference.

- The union of all metadata items referencing an image comprises the complete metadata for an image item. In addition, a single metadata item can be referenced to one or more image items, allowing common metadata to be stored once and applied to individual images and to groups of images.
- Metadata unique to a single image is its own separate, unique metadata item.

Image Collections:
- A set of image items stored within a single file.
- HEIF defined collection types include user-defined albums, a favorites list, an input list to a panorama image, and a slideshow group.

Metadata for Image Collections:
- Metadata items referenced to a collection of image items.
- Individual image items in a collection of images may have their own individual metadata items through unique references.

Tracks: A timed sequence of related media samples of the following types:
- **Motion Imagery:** traditional video type imagery, compressed or uncompressed, and captured at a constant frame rate.
- **Image Sequences:**
  - A collection of coded image items which may be associated with advisory timing to play in a desired timed playout manner.
  - Coded images may be dependent on other coded images when in a sequence, such as when using predictive coding in video compression.
  - Storing image sequences as tracks, which play as a timed sequence, like a video track, using the structures inherited from ISOBMFF.
  - Displaying images in a sequence track in a non-timed manner, such as in a tiled array or composited gallery.
- **Timed Metadata:**
  - Timed metadata describing images in a Motion Imagery or image sequence track.
  - The default method for encoding metadata as Key-Length-Value in a timed metadata track.
  - Using track referencing to match a timed metadata track to a specific motion imagery or image sequence track. Each timed metadata sample describing a timed image sequence sample is the time-aligned sample in the metadata track relative to the image sample.
- **Audio:** timed audio samples, frequently providing sound information for an accompanying video track.

Track Samples:
- Individual samples in a track, such as a single image frame, an audio sample, or a block of metadata at a specific time. In general, the data residing at a specific time instant on a track's timeline.

Track Sample Auxiliary Information:
- Metadata or other information associated with each sample in a track.

Track Metadata (static):

- Metadata that applies to an entire track of data.

Presentation Metadata (static):

- Metadata addressing a collection of tracks in a MovieBox.

When implementing NGA.STND.0076 conformant files in NSG applications, unless specifically restricted by a requirement in this document, features described in the ISOBMFF family of standards are available for implementation and use.

6 Detailed Description and Requirements

6.1 File Format Interoperability Indicators

6.1.1 FileTypeBox and Branding

The FileTypeBox carries branding information identifying the specifications to which the file complies. Brands dictate requirements for producers when generating files and for readers when decoding, interpreting, and presenting content. Brands define interoperability points through format version compatibility by indicating the use of specific image codecs and possibly metadata encoding inside a file.

The listing of brands indicates one or more specifications to which a file is conformant. To enable all features and content types within a file, a reader application must implement all features documented for the specific brands listed. Readers may ignore:

1. features not addressed by this specification and not relevant to NSG GEOINT applications, such as out-of-scope metadata encoding formats (MPEG-7, etc.).
2. features non-conformant to this specification, such as the use of codecs not approved for use with files conformant with this document.

The complete list of brands indicates to readers the full set of requirements and constraints when opening and accessing its content. This standard uses individual requirements to address branding requirements.

For acquisition purposes, file producers must generate files conformant to the requirements associated with features implemented within the file. This includes conforming to the requirements of the implemented features of each brand indicated within a specific file.

Readers claiming conformance to the GIMI standard must handle all requirements of this standard, as well as the brand's requirements. Reader may claim partial conformance to this standard by clearly stating conformance to a specified subset of the listed brands. As an example, an application utilizing JPEG 2000 encoded imagery only may claim conformance to the format when accessing and interacting with JPEG 2000 content.

Readers only need to be conformant with and recognize one of the brands to determine if it can make use of a file. Scanning the list of brands on file open enables reader applications to determine the level of interoperability expected with the content in the file. The MPEG Registration Authority (https://mp4ra.org/#/) maintains and manages code-points (i.e., brands as four-character codes – 4CC) on behalf of ISO in support of the MPEG and JPEG communities.

The MP4 Registration Authority website provides a list of all registered codes for each defined code-point within the file. Currently defined code points include brands, boxes, codecs, handlers, etc. As brands and other 4CC codes are generated specifically for this specification, they are submitted to the MP4RA for registration.

For this specification, the branding selected ensures the required capabilities within ISOBMFF will meet the needs of NSG GEOINT applications. As ISOBMFF and HEIF evolve, the brands required by this standard may be augmented to enable the exercising of new capabilities that emerge through newer versions of the ISOBMFF and HEIF standards.

ISOBMFF supports “dual-branding”, allowing for the carriage of both Still Imagery and Motion Imagery content within the same file. A dual-branded file contains both Still Imagery with its associated metadata and Motion Imagery with its associated metadata. Files containing multiple media types include the necessary brands for properly interpreting each type of media content.

To meet the requirements readers must support specific brands for Motion Imagery as listed in Annex E of ISO/IEC 14496-12, and specific brands for Still Imagery as listed in Section 10 of ISO/IEC 23008-12. These sections list the evolution of the brands and the features enabled with each new version. Requirements in this document specify current approved 4CC brands. Readers conformant with these NGA.STND.0076 requirements through the latest ISO versions are capable of reading and understanding files conformant with this standard.

Writers are free to generate files using capabilities defined by an appropriate set of brands based on the needs of the writer. Readers conformant to this specification may only be capable of deciphering content required by this standard.

Readers, as per the requirement given in ISOBMFF documentation, must ignore unrecognized boxes and media. Writers implementing features beyond the brands identified in this document, do so with the understanding that conformant readers may not be interoperable for those features implemented outside the range covered by the specific brand.

The GIMI document defines brands for the unique capabilities this document defines. When implementing a GIMI brand’s capabilities, the compatible brands list must include the GIMI brand, which then identifies a file as being conformant with NGA.STND.0076, or a specific subset of requirements within NGA.STND.0076.

Expect future versions of the GIMI standard to include additional brands to support specific implementation capabilities. For example, including high dynamic range compression via high profile codecs. This will define additional interoperable capabilities within the standard.

This standard defines all its brands beginning with the ‘geo’ prefix for files carrying still media and/or timed media. Following the ‘geo’ prefix is a version number. The first initial version is ‘geo1’ and following versions will increment by one until reaching 9, then the version digit will shift to a, b, c, etc. Applications include the ‘geo’ + version brand in the compatible brands list which signals the file is conformant to this specification which carries GEOINT media and metadata.

The scope of the ‘geo1’ brand includes monolithic, rectangular, uncompressed, and compressed forms of Still Imagery, including monochrome, color visible, panchromatic, MSI, HSI, rectangular uncompressed, compressed forms of Class 0 (raw/uncompressed), Class 1 (Broadcast), and Class 2 (SEMI) Motion Imagery. The presence of the ‘geo1’ brand ensures conformant software applications can operate with these delineated forms of Imagery. This branding approach improves the interoperability of applications across a wider number of use cases within the NSG.

From a baseline interoperability level, this version of the standard is based on the HEIF defined ‘mif2’ brand for Still Imagery, the HEIF defined ‘msf1’ brand for image sequences, and the ISOBMFF defined ‘isoa’ brand for Motion Imagery and ‘unif’ brand mandating unique IDs for items and tracks in a file. Table 4 summarizes these brands.

Table 4: Descriptions for Brands for NGA.STND.0076 Implementation

Brand	Source	Description
'mif2'	HEIF	The 'mif2' brand represents interoperability requirements for image and metadata items through the file level MetaBox within Edition 2 of the HEIF standard. It represents a baseline for Still Imagery support in this standard. The 'mif2' brand inherits the capabilities of the initial 'mif1' brand and adds additional conformance requirements related to referencing, alpha/depth maps, image scaling, etc. Specifics of the branding differences are documented in the HEIF standard.
'msf1'	HEIF	The 'msf1' brand indicates the presence of a HEIF defined image sequence.
'isoa'	ISOBMFF	The 'isoa' brand represents interoperability requirements for the base format as well as Motion Imagery requirements for this standard. The 'isoa' brand specifically adds item properties.
'unif'	ISOBMFF	The 'unif' brand indicates the unified implementation and handling of IDs across file-level MetaBox items, tracks, track groups, and entity groups.
'geo1'	NGA.STND.0076	The 'geo1' brand indicates a file is conformant to the version 1 of this standard (NGA.STND.0076_1.0). Future versions of the standard will increment the brand number accordingly.

In some cases, branding indicates codec specific brands. These brands then indicate to a reader the requirements for decoding and interpreting encoded content. Many codecs support branding for different features, such as higher-level profiles. Of note, the inclusion of uncompressed images and video, via ISO/IEC 23001-17 [9] does not require a special brand.

Table 5 shows an example of a FileTypeBox for a file containing a JPEG 2000 encoded image item and an HEVC Main Profile encoded image sequence.

Table 5: Example payload of a FileTypeBox

Data Type	Example Value	Element Name
unsigned int(32)	'mif2'	Major Brand
unsigned int(32)	0	Minor version
unsigned int(32) [ ]	'mif2', 'msf1', 'geo1', 'j2ki', 'hevc', 'unif'	Compatible Brands

The following are requirements related to the implementation of a FileTypeBox and branding.

Requirement(s)	
NGA.STND.0076_1.0-01	Applications shall generate files conformant with all applicable normative references.
NGA.STND.0076_1.0-02	Producer systems shall include the 'unif' brand, indicating the use of unique IDs as per ISO/IEC 14496-12.

NGA.STND.0076_1.0-03	Producer systems generating files conformant with this version of NGA.STND.0076 shall include the 'geo1' brand in the compatible brands list inside the FileTypeBox.
NGA.STND.0076_1.0-04	Systems consuming files with Still Imagery content shall be able to process files conformant to the 'mif2' brand.
NGA.STND.0076_1.0-05	Systems consuming files with image sequence content shall be able to process files conformant to the 'msf1' brand.
NGA.STND.0076_1.0-06	Systems consuming files with Motion Imagery content shall be able to process files conformant to the 'isoa' brand.

6.1.2 File Extensions

While file extensions frequently signal a specific format and interoperability requirements, an ISOBMFF file extension carries less definitive meaning. Brands in the FileTypeBox communicate interoperability requirements. When an ISOBMFF file includes multiple brands the ISOBMFF guidance is to select a file extension identifying the primary use of the file. For instance, files carrying both Still and Motion Imagery may use a file extension of HEIF or MP4. The writer application determines the best option for the extension. Reader applications utilize the brands to determine interoperability points with a specific file.

6.2 Imagery Content

The focus of this standard is the storage of Still Imagery, image sequences, Motion Imagery, and associated metadata, as well as providing mechanisms for labelling content and identifying defined relationships among the stored content. This section covers the nature of imagery in general and the forms for uncompressed and compressed imagery when stored in a file. The following sections cover the details of Still Imagery, image sequences, and Motion Imagery. Section 6.3 addresses metadata content and its association with imagery.

6.2.1 Still Imagery and Image Sequence Codecs

ISOBMFF and HEIF are agnostic to image encoding and supporting a wide variety of existing codecs. NGA.STND.0076 allows a select sub-set of codecs to balance the trade space of functional coverage, supporting legacy needs, and maximizing interoperability. Table 6 lists the approved codecs for Still Imagery and image sequence coding.

Table 6: List of codecs approved for use with Still Imagery and Image Sequences

Encoding Method	Codec Standard	Carriage in HEIF
Uncompressed Video and Images in ISOBMFF	ISO/IEC 23001-17	ISO/IEC 23001-17
JPEG 2000	ISO/IEC 15444-1	ISO/IEC 15444-16
High Throughput JPEG 2000	ISO/IEC 15444-15	ISO/IEC 15444-16
HEVC	ISO/IEC 23008-2	ISO/IEC 23008-12
AVC	ISO/IEC 14496-10	ISO/IEC 23008-12

The first column lists each encoding method, the second column identifies the associated codec standard, and the third column indicates the standard for carriage of the coded item within a HEIF file. NGA.STND.0076 addresses a broad range of imagery forms and types found in the

NSG community. This includes common one and three band imagery, multi and hyperspectral imagery, imagery with unusually high resolutions and/or high dynamic range, and imagery with extended definition storage formats, such as floating point and complex variable.

The ISO/IEC 23001-17 uncompressed codec, as well as numerically lossless implementations of other codecs, such as JPEG 2000, are for applications where preservation of content and quality of imagery is of primary concern. For situations where storage and transmission efficiency are of primary consideration, lossy compression options are available.

Different codecs, and their various profiles, place constraints on an input image. Additionally, codecs constrain the wide range of options for arranging component and pixel content, such as tiling, interleaving, padding, sub-sampling, and pixel alignment on defined byte boundaries. These options become limited based on increasing imagery bit depth and formatting constraints of an individual codec.

JPEG 2000 is a primary codec for GEOINT Still Imagery applications. The underlying wavelet technology supports high dynamic range, high density images, high numbers of components, and unique methods for accessing regions-of-interest without having to access the entire image. To support the need for greater processing efficiency, this standard supports the use of the J2K derivative codec High-Throughput JPEG 2000 (HTJ2K).

JPEG2000 and High Throughput JPEG2000 coding systems support Class 2 Motion Imagery applications with a primary benefit of high dynamic range and high-resolution support. Implementation as time sequences is via ISO/IEC 15444-16 [10]. ISOBMFF does not include sample entry support for the carriage of JPEG 2000 or HTJ2K in a video track with a ‘vide’ handler; therefore, writers implement image sequences using the tools of ISO/IEC 23008-12 and ISO/IEC 15444-16 to address JPEG2000 and HTJ2K Motion Imagery requirements.

While HEVC and AVC are principally video codecs, they do support still image compression. These codecs find use in applications requiring high interoperability, such as disaster relief, and dissemination to disadvantaged users, as well as applications for extracting and storing a single frame from a video file.

A visually-eccentric image is complex, large, or based on a still image encoded in an uncommon form (high bit depth, very high pixel count, number of bands greater than three, etc.). A visually-normal image is readily displayable with common image viewers, i.e., an RGB or monochrome image with a general device-friendly resolution and bit depth.

When a writer includes a visually-eccentric image item, it may be advantageous to create and include an additional alternate visually-normal image and provide a reference to the visually-eccentric image item to link them together, forming a group.

Grouping can combine alternate versions of the same content. In cases of an uncommon codec or profile, it may be advantageous to re-encode the image using a common codec, such as HEVC, and include this as an alternate version of an image item or track. This improves interoperability with commodity devices and software allowing a higher percentage of applications to view some version of imagery in the file. The utility of this approach is dependent on a variety of issues related to workflow, security, stakeholder resources, long term storage cost, etc.

Image sequences are supported with the following traits:

1. Sequences are made up of coded images.
2. Sequences ‘may’ be associated with advisory timing.
3. The coding of images may use inter-prediction.

Image sequences are composed of image items but are playable like a Motion Imagery track. They are identified by their handler type, which is set to ‘pict’. Sequences have more flexibility in their implementation than video tracks as their timing is less restrictive. The images in a sequence support playout as a movie, as a slideshow, with pauses and transitions, or displayable as a single composited frame. Image sequence allocation utilizes the MovieBox (‘moov’) and TrackBox (‘trak’) structures. This provides a configurable timing structure and options for attaching timestamps, metadata, and audio to a sequence. These tools provide an ability to playout a collection of image items in a timed presentation. The editing features available within a track facilitate presentation options, such as looping sequences.

Requirement(s)	
NGA.STND.0076_1.0-07	Files containing still image items shall contain a PrimaryItemBox with a primary item declared.
NGA.STND.0076_1.0-08	Files containing image sequences shall contain a PrimaryItemBox with a primary item declared.
NGA.STND.0076_1.0-09	Where a file contains only image sequences, the primary item shall be a single, renderable, intra-coded image item from a sequence.
NGA.STND.0076_1.0-10	Uncompressed encoded still image items shall conform to ISO/IEC 23001-17.
NGA.STND.0076_1.0-11	Uncompressed encoded image sequence content shall conform to ISO/IEC 23001-17.
NGA.STND.0076_1.0-12	Encoding of still images using JPEG2000 shall conform to ISO/IEC 15444-1.
NGA.STND.0076_1.0-13	Encoding of image sequences using JPEG2000 shall conform to ISO/IEC 15444-1.
NGA.STND.0076_1.0-14	Files containing JPEG2000 encoded image items shall conform to ISO/IEC 15444-16.
NGA.STND.0076_1.0-15	Files containing JPEG2000 encoded image sequences shall conform to ISO/IEC 15444-16.
NGA.STND.0076_1.0-16	Encoding of still images using HTJ2K shall conform to ISO/IEC 15444-15.
NGA.STND.0076_1.0-17	Encoding of image sequences using HTJ2K shall conform to ISO/IEC 15444-15.
NGA.STND.0076_1.0-18	Files containing HTJ2K encoded image items shall conform to ISO/IEC 15444-16.
NGA.STND.0076_1.0-19	Files containing HTJ2K encoded image sequences shall conform to ISO/IEC 15444-16.
NGA.STND.0076_1.0-20	Encoded forms of imagery shall conform to one of the approved codecs listed in Table 6.

NGA.STND.0076_1.0-21	Carriage of imagery content shall conform to one of the approved standards for carriage listed inTable 6.
NGA.STND.0076_1.0-22	Reader applications shall decode content encoded from one or more of the codecs listed in Table 6*

* *Note: programs may independently require supporting multiple or all codecs listed and may also constrain support to specific codec profiles, depending on program application needs.*

6.2.2 Motion Imagery Codecs

This standard allocates Motion Imagery in a file using the MovieBox and the TrackBox. To signal a track is carrying Motion Imagery, the handler type for the track is set to 'vide'. The SampleEntryBox signals the encoding method for the imagery. Approved codecs for Class 1 Motion Imagery applications include the HEVC (ISO/IEC 23008-2) and AVC (ISO/IEC 14496-10) codecs, with Levels and Profiles as defined by the MISP.

The approved codec for Class 0 Motion Imagery applications is the Uncompressed Video and Images in ISOBMFF (ISO/IEC 23001-17), and the approved codecs for Class 2 Motion Imagery are JPEG 2000 (ISO/IEC 15444-1) and HTJ2K (ISO/IEC 15444-15). These codecs support high image quality and generally operate in environments which support higher bandwidth. Implementation of Motion Imagery using JPEG 2000 and HTJ2K is via image sequences as described in Section 6.2.1.

Table 7, Table 8, and Table 9 lists codecs approved for use with Motion Imagery implementations.

Table 7: Codecs approved for use with Class 0 Motion Imagery

Encoding Method	Codec Standard
Uncompressed Video and Images in ISOBMFF	ISO/IEC 23001-17

Table 8: Codecs approved for use with Class 1 Motion Imagery

Encoding Method	Codec Standard
HEVC	ISO/IEC 23008-2
AVC	ISO/IEC 14496-10

Table 9: Codecs approved for use with Class 2 Motion Imagery

Encoding Method	Codec Standard
JPEG 2000*	ISO/IEC 15444-1
High Throughput JPEG 2000*	ISO/IEC 15444-15
HEVC	ISO/IEC 23008-2
AVC	ISO/IEC 14496-10

*Implemented as image sequences (see Section 6.2.1)

The following lists additional details for these codec options:

- ISO/IEC 23001-17 provides a flexible mechanism for the carriage of uncompressed Class 0 Motion Imagery within ISO Base Media File Format.
- High Efficiency Video Coding (HEVC/H.265): The successor to H.264 provides the benefits of enhanced capabilities and improved compression efficiency. HEVC profiles support bit depths of 8-bits and up to 16-bits at the highest profiles and levels. The MISP provides guidance for which profiles and levels of HEVC to use with a specific Motion Imagery Class and in specific applications. ISO/IEC 23008-2 and ITU-T H.265 [11] jointly specify the HEVC codec.
- Advanced Video Coding (AVC/H.264): a commercial media standard defined within the MPEG-4 suite of standards. AVC profiles support bit depths of 8-bits and up to 14-bits at the highest profiles and levels. The MISP provides guidance for which profiles and levels of AVC to use with a specific Motion Imagery Class and in specific applications. ISO/IEC 14496-10 (MPEG-4 Part 10) [12] and ITU-T H.264 [13] jointly specify the AVC codec.

An additional standard of interest when using AVC and HEVC is ISO/IEC 14496-15, which defines how to carry AVC and HEVC coded video in NAL unit form.

- ISO/IEC 14496-15 [14] specifies the formatting and carriage of HEVC and AVC content within ISOBMFF. It includes Sample Entry and box type definitions for handling both HEVC and AVC compressed video and describes how to encapsulate a video stream into Network Abstraction Layer (NAL) units within the ISOBMFF framework.

Requirement(s)	
NGA.STND.0076_1.0-23	Systems producing Class 0 Motion Imagery shall encode the imagery using the codec options listed in Table 7.
NGA.STND.0076_1.0-24	Systems producing Class 1 Motion Imagery shall encode the imagery using the codec options listed in Table 8.
NGA.STND.0076_1.0-25	Systems producing Class 2 Motion Imagery shall encode the imagery using the codec options listed in Table 9.
NGA.STND.0076_1.0-26	Where a file includes JPEG 2000 encoded Class 2 Motion Imagery, the imagery shall comply with ISO/IEC 15444-1.
NGA.STND.0076_1.0-27	Where a file contains JPEG 2000 Motion Imagery, it shall be implemented as an image sequence conformant with ISO/IEC 15444-16.
NGA.STND.0076_1.0-28	Where a file includes HTJ2K encoded Class 2 Motion Imagery, the imagery shall comply with ISO/IEC 15444-15.
NGA.STND.0076_1.0-29	Where a file contains HTJ2K Motion Imagery, it shall be implemented as an image sequence conformant with ISO/IEC 15444-16.

NGA.STND.0076_1.0-30	Where a file includes HEVC/H.265 imagery, the imagery shall comply with MISP-approved profiles and levels for Class 1 and Class 2 Motion Imagery.
NGA.STND.0076_1.0-31	Where a file includes HEVC/H.265 encoded imagery configured with Network Abstraction Layer (NAL) units, it shall be conformant with ISO/IEC 14496-15.
NGA.STND.0076_1.0-32	Where a file includes AVC/H.264 imagery, the imagery shall comply with MISP-approved profiles and levels for Class 1 and Class 2 Motion Imagery.
NGA.STND.0076_1.0-33	Where a file includes AVC/H.264 encoded imagery configured with Network Abstraction Layer (NAL) units, it shall be conformant with ISO/IEC 14496-15.
NGA.STND.0076_1.0-34	Where a system is using subsampled imagery, it shall be limited to 4:4:4, 4:2:2, and 4:2:0.

6.3 Metadata Content

In typical workflows, metadata critical to the imagery content resides within the same file as the content. This standard does, however, allow the referencing of additional associated metadata within one or more external "sidecar" files. This enables producing simplified imagery files with all the necessary and required metadata, while metadata sidecar files support optimized access, modification, and transmission of metadata content in certain applications and workflow infrastructures.

Regardless of the implementation (i.e., contained or sidecar metadata), NGA.STND.0076 requires three specific items of metadata to be present within a file containing the imagery content:

1. Information security marking (see Section 6.4.5 for details)
2. Content Identification (Content ID) labeling for all media items and tracks (see Section 6.4.3 for details)
3. Nano Precision Time Stamps (timestamp) for all media items and samples (see Sections 6.4.4 and 6.5.2 for details)

Content ID and timestamp items use different ISOBMFF methods of allocation for still items and timed media sample content.

For applications requiring the containment of metadata, in addition to the security marking, Content IDs, and timestamps inside a file, generalized methods permit the carriage of both static and temporally varying metadata. Metadata serves multiple purposes and there are multiple mechanisms for encoding and attaching metadata to imagery content.

The general ISOBMFF implementation form is metadata as defined by the format and carried within the box constructs of ISOBMFF. This standard defines a means to carry GEOINT metadata in a file using the generalized metadata carrying capabilities of ISOBMFF. The ISOBMFF format "box" data structure includes elements directly associated with the form, location, and intended display of imagery content. This design aligns with the needs of industry and consumer media applications, while also extensible to support GEOINT applications, if appropriate.

Box structured metadata addresses topics ranging from content storage, manipulation (such as image rotation), text labeling, and item properties. These boxes carry metadata with a specific intent and purpose, which in certain cases overlaps with GEOINT metadata requirements. Existing boxes contain metadata relating to pixel bit depths, pixel counts, and color information. These play a fundamental role in formatting, storing, accessing, and properly displaying media content identified within a file.

For GEOINT application metadata not consistent with the existing box hierarchy, built-in "plug-in" methods enable allocating application-specific metadata objects, and for locating, storing, indexing, grouping, and referencing (linking) the metadata to other media content. This metadata may be static or dynamic depending on its context and the media it describes. Although ISOBMFF provides the ability to signal a user-defined metadata encoding method, the encoding of information security markings is with XML. Content IDs are implemented as unframed metadata, and KLV is currently utilized for all other GEOINT metadata contained in a file.

The standard assumes KLV encoding for static metadata, synchronous and asynchronous timed metadata, and sample auxiliary information metadata. While it supports legacy KLV implementation, it encourages using the Motion Imagery Metadata (MIMD) Model [15] with KLV encoding for implementations, which do not require transcoding from instances of Motion Imagery (MPEG2-TS, etc.) that contain legacy KLV Local Sets and Packs.

Table 10 lists the major categories and scoping levels of ISOBMFF boxed and GEOINT metadata carried or referenced in a file.

Table 10: Categories and scoping levels of metadata in a file

Scoping of Metadata	Application of Metadata
Format Defined 'Boxed' Metadata [ISOBMFF file hierarchy boxed metadata, which carries structural metadata required to display, or otherwise carry and manipulate media content within the file]	
Format functionality	Supports file functionality, describes form and location of media content
Item Properties	Metadata about a declared item as defined in ISOBMFF and HEIF documentation. May include information affecting the displayed form of the source imagery.
File Level Metadata [Security Markings, Mission Information, etc., declared as metadata items in file level MetaBox]	
File	GEOINT metadata that addresses aspects of file media content in whole, such as security, summary information of content, etc.
Still Image Metadata [Declared as metadata items in file level MetaBox and referenced to still image content of interest]	
Image	Information about an image as a whole
Collection of images	Information about a declared set, or group, of images
Image tile	Information about a specific tile of an image

Image Region of Interest (ROI)	Information about a specific region in an image consisting of a point (pixel), polyline, rectangle, ellipse, or arbitrary pixel mask (list of pixels) in an image
Groups of defined ROIs	Information about a collection of defined regions
Full frame of pixels	Unique information about each pixel in a full image – cloud cover, geodetics, pixel quality, etc.
Presentation Metadata [Declared inside the MetaBox of the MovieBox]	
Collection of tracks	Information describing a collection of timed media tracks. Declared in the MetaBox inside the MovieBox.
Static Metadata for a Track [Declared in the MetaBox inside the TrackBox]	
Individual timed media tracks	Static information describing a single timed media track
Track Sample Auxiliary Information [Declaration and indexing via 'saiz' / 'saio' boxes in Sample Table]	
Individual samples in timed media tracks	Sample metadata, including timestamps and associated timing metadata for individual timed samples. Enables correlation of each sample with all other media in the NSG standards conformant ecosystem.
Timed Metadata Tracks [KLV encoded, declared as a timed metadata track]	
Timed metadata tracks	Timed metadata information stored in a track. May be synchronous or asynchronous to other media tracks, such as video, image sequences, and audio. Declared as a track and referenced to the track(s) it describes using the 'cdsc' reference type. Rate partitioning may be implemented by containing metadata at the same rates in separate tracks.

6.4 Static Metadata

Static metadata refers to metadata that either does not change or applies to a self-contained media item that exists at a single sample point in time. Central to the implementation of static metadata is the MetaBox, which enables the allocation of static GEOINT metadata items at the file level, and inside movie, movie fragment, track, and track fragment boxes. Unless specifically identified within this standard as using a different encoding, the encoding of metadata is in the Key-Length-Value format.

The high-level diagram in Figure 3 shows how to allocate static metadata for items at the file level and inside a track. In the figure, the ItemDataBox ('idat') stores the ISM.XML security marking metadata, with the MediaDataBox ('mdat') storing(idat) the remaining metadata content (mdat).

For all but the minimum metadata required in a file (security, timestamps, and identification), options exist for including additional GEOINT metadata either inside the file, outside the file (in a separate sidecar file), or a combination of both.

Metadata may be allocated and placed in a file container external to the immediate ISOBMFF file. In this case, a data reference provides the name and location of the file resource holding the metadata, which then maintains an ability to locate and utilize the metadata. In cases where the data referencing to the external file is not possible, not implemented, or not maintained, Content ID, Nano Precision Time Stamp, and other metadata information is used to determine the relationships between different media content.

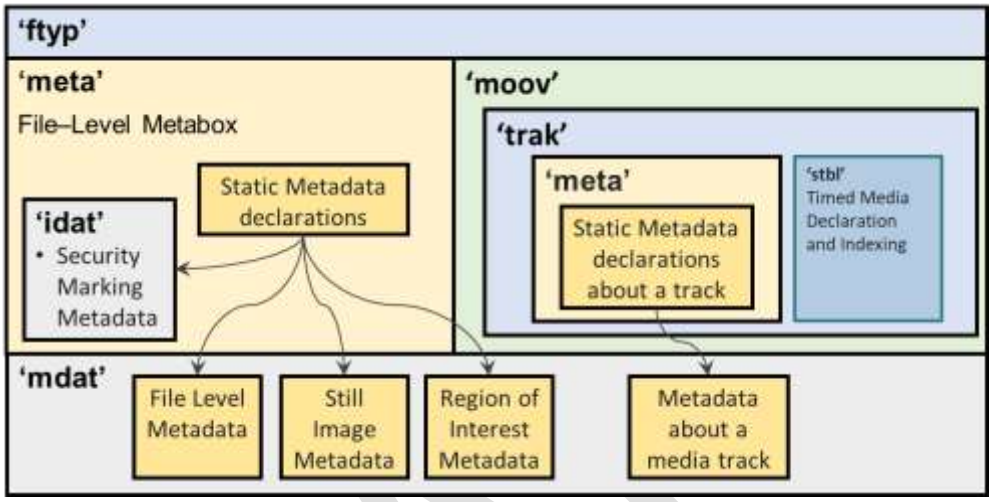


Figure 3: Declaration and storage of static metadata

Typical real-time collection applications produce and capture both static and dynamic metadata. In real-time situations, static metadata accompanies the Motion Imagery at a periodically rate defined by MISB KLV documentation, typically once every 10 or 30 seconds. When recording real-time data directly to a GIMI file, the recording system may not have knowledge on whether the metadata information is static or dynamic. As such, it is best to capture and record this data along with any dynamic timed metadata to a timed metadata track.

In situations where the recorder lies at the sensor the implementation should ideally capture and store static information separate from dynamic metadata. The ability to generate files in either situation dictates that downstream reader software check all possible locations for relevant metadata within the file. To facilitate easier and more consistent access to static metadata, processing systems downstream from initial capture may choose to extract static metadata from timed tracks and re-locate that data to static metadata locations. Where possible, it is recommended systems write static metadata to static metadata locations.

6.4.1 Static Metadata encoding

ISOBMFF and HEIF both support specific types of metadata encoding (EXIF, XMP, MPEG-7, etc.), as well as generalized methods for identifying other user-implemented encoding types (MIME and URI defined types, such as XML and KLV). While this provides great flexibility, it can hamper interoperability when allowing too many encoding methods. To maximize interoperability, the Key-Length-Value encoding method, adopted by the Motion Imagery Standards Board for Motion Imagery applications is the default encoding for metadata stored internal to NGA.STND.0076 conformant files. MISB ST 0107 [16] provides baseline requirements for implementing KLV Metadata in NSG applications. Additional MISB

standards define packs and sets of specific metadata elements addressing various application use cases. While many of these sets and packs support legacy implementation, moving forward, general metadata applications will utilize the MIMD suite of standards. Although developed to support Motion Imagery applications, MIMD is also applicable to Still Imagery applications and continues to evolve accordingly. When implementing metadata in external files, the Resource Description Framework (RDF) W3C standard is a preferred method of implementation. Carriage of metadata in RDF sidecar files is outside the scope of this document and addressed in separate documentation.

When implementing metadata in NGA.STND.0076, the encoding of static metadata is signaled in the allocation process for a metadata item. Individual blocks of metadata are allocated in the ItemInformationBox with the associated ItemInfoEntry configured with its item type set to either 'uri ' or 'mime' depending on the encoding type. For KLV encoding, the item type is set to 'uri ', which then requires an additional item URI type string (utf8string). Coded as a URN for a specific 16-byte MISB defined KLV key as 'urn:smpte:ul:[16-byte key]', declares the specific KLV Local Set or Pack. Note: the brackets are not included in the string but are shown for clarity of where the 16-byte value is placed. The item URI type string identifies the item_type as having a specific implementation of Key-Length-Value coding and utilizes the form of a SMPTE URN, as defined in RFC 5119 [17], where the 16-byte key is declared with each successive 4-bytes separated by a '.' (period). Figure 4 shows an example using the MIMD defined length pack key. "theURI" string in the figure shows a configuration as it is stored in a file.

ST1902 MIMD Defined Length Pack Key: 06.0E.2B.34.02.05.01.01.0E.01.05.04.00.00.00.00

theURI: 'urn:smpte:ul:060E2B34.02050101.0E010504.00000000'

Figure 4: Example of an 'item_uri_type' string generation using the MIMD key.

For a generated KLV metadata item, the 'key' of the KLV metadata is stored in the 'item_uri_type' string. The 'length' parameter of the KLV metadata pack (or set) is stored as the totality of 'extent_lengths' for the item in the ItemLocationBox. The 'value' portion of the KLV metadata is stored in the MediaDataBox, the ItemDataBox, or external to the file, depending on construction method chosen in the ItemLocationBox. The value portion contains the encoded metadata information from an approved KLV set or pack standard. This method for GEOINT metadata declaration applies to any MetaBox in a file.

Requirement(s)	
NGA.STND.0076_1.0-35	Files shall contain a file-level MetaBox ('meta').
NGA.STND.0076_1.0-36	Static metadata applicable to all media in a file shall be implemented using the MetaBox at the file level.
NGA.STND.0076_1.0-37	Static metadata applicable to single timed media tracks shall be implemented using the MetaBox inside the TrackBox ('trak').
NGA.STND.0076_1.0-38	No temporally varying metadata shall be allocated using static metadata boxes.

NGA.STND.0076_1.0-39	A MetaBox declaring MISB defined KLV metadata sets and packs shall set the ItemInfoEntry field item_type = 'uri '
NGA.STND.0076_1.0-40	A MetaBox declaring MISB defined KLV metadata sets and packs shall encode ItemInfoEntry field item_uri_type as a URN in the format of: 'urn:smpte:ul:16-byte key', with the 16-byte key following the QUADBYTE notation declared in RFC 5119.
NGA.STND.0076_1.0-41	Metadata utilizing Key-Length-Value encoding shall conform to MISB ST 0107.

6.4.2 File Level Metadata

Metadata applicable to all file content is allocated in the file level MetaBox and not referenced to any specific media content in the file. This standard mandates all files include security marking in the file level metadata, see Section 6.4.5.

The file level MetaBox may contain localized metadata which does not apply to the whole file. Applications allocate metadata in the file level MetaBox and include a reference to define the scope using association with specific entities of content. An example is associating metadata with one or more still image items; a reference between the metadata and each image item defines the scope of the metadata.

Metadata applicable to all file content is allocated in the file level MetaBox; such file-wide metadata does not reference specific media content in the file. NGA.STND.0076 requires including information security marking metadata, as described in Section 6.4.5, at the file level for all instances of this standard. Metadata, when not applicable across an entire file, may be allocated in the file level MetaBox with a reference used to define scope through association with specific entities of content. An example is metadata associated with one or more still image items, where a reference between the metadata and each image item defines the scope of the metadata.

6.4.3 Content Identification Labeling

To manage content strictly within a file, ISOBMFF incorporates an internal ID system using integer IDs to keep track of individual items (MetaBox items – at the file, MovieBox, and TrackBox levels), and tracks (image sequences and video) within a single file referred to as item_IDs and track_IDs, respectively. When combined in use, they are referred to as entity IDs. There are three scoping levels for IDs: the file level, the presentation level (MovieBox and MovieFragmentBox), and the track level (TrackBox and TrackFragementBox). These IDs provide the referencing mechanisms to associate metadata with image content.

ISOBMFF defines the generation, scoping, tracking, and storing of the entity IDs. Entity ID's are either 16-bit integers or 32-bit integers, depending on how many declared objects are in a file. The entity IDs facilitate relationships between separate pieces of content, such as metadata describing an image or a relationship between two instances of the same image coded with different codecs. The 'unif' brand, defined in ISOBMFF and mandated by this standard, simplifies the tracking of these IDs across the various types of content within a file.

During editing and other operations, applications may modify IDs making the entity IDs potentially transitory. This limits their use for maintaining long-term cataloguing of content for

enterprise level search and discover, etc. Because entity IDs may not persist over the life of a file, Content IDs, in the form of a uniform resource name (URN) and defined by NGA.STND.0076, provide persistent, universally unique IDs for content.

To facilitate the tracking and management of media content throughout a workflow process, over time, and across the GEOINT Enterprise, this profile mandates the use of Content IDs. Every allocated image item, metadata item, region item, and track in a file has an assigned Content ID. They provide a means to identify and locate content, and for linking to related ontology structured data contained in external/separate files (e.g., RDF structures). Traditional GEOINT metadata, such as sensor, platform, and mission data, contained in RDF or other sidecar files, link to image media using these Content IDs. To uniquely identify a single image frame in a Motion Imagery track, the combination of the track's Content ID and the Nano Precision Time Stamp for the frame provide global uniqueness.

Figure 5 shows an example for how content in an image file may be associated with content in a sidecar file using timestamps and Content IDs. In this case, the sidecar file is an RDF metadata file. RDF is well suited for implementation of GEOINT metadata "sidecar" files, enabling the ability to access, manipulate, and update small pieces of metadata content while minimizing interaction with bulky image files. Across the NSG, Content IDs uniquely identify a specific piece of content and timestamps allow for the alignment of different pieces of content temporally.

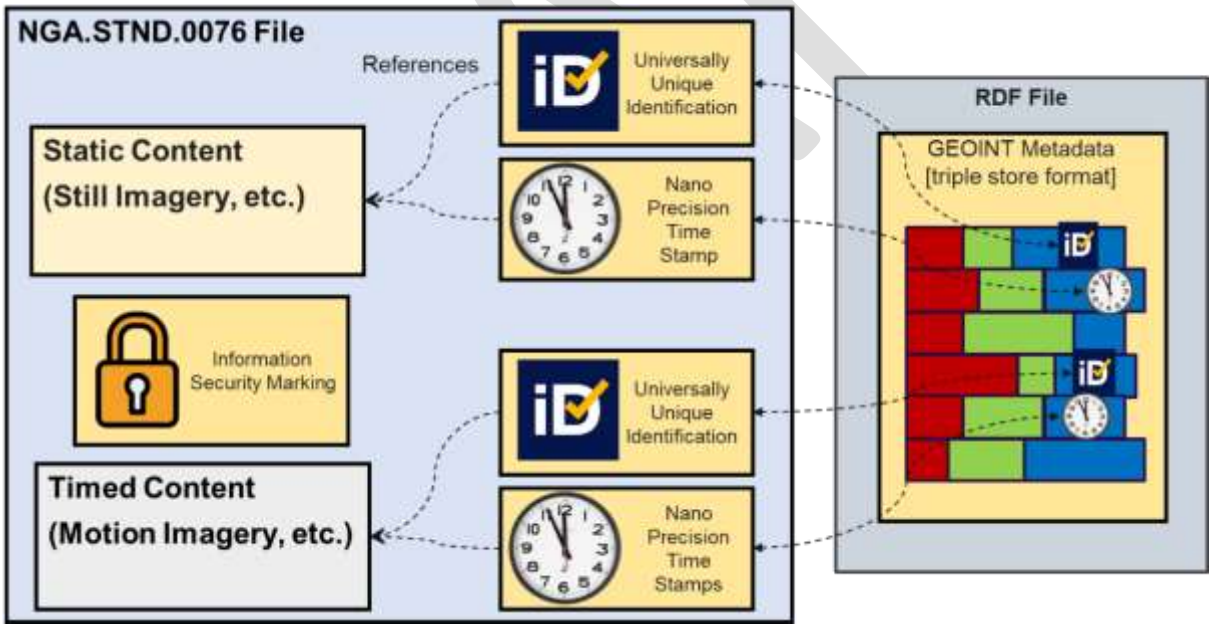


Figure 5: Content IDs and Nano Precision Time Stamps identify and associate content.

Content IDs for each unique piece of image, region, metadata, and track content in a file use existing ISOBMFF capabilities for creating user-defined metadata items in a MetaBox. Within an ItemInfoBox ('iinf'), an ItemInfoEntry ('infe') box instantiates a Content ID by setting the item_type to 'uri ' and the 'item_uri_type' to a defined value of 'urn:uuid:aac8ab7d-f519-5437-b7d3-c973d155e253', which then signals the allocation of a Content ID.

NGA.RP.0001 [18] defines the form and how to create UUIDs used as Content IDs for image and metadata items, and Motion Imagery, metadata, and audio tracks. Instances of Content IDs take the general form of 'urn:uuid:<uuid>' [19], where <uuid> is generated following the guidance provided in the NGA Recommended Practice. Figure 6 shows an example of the allocation of a Content ID.

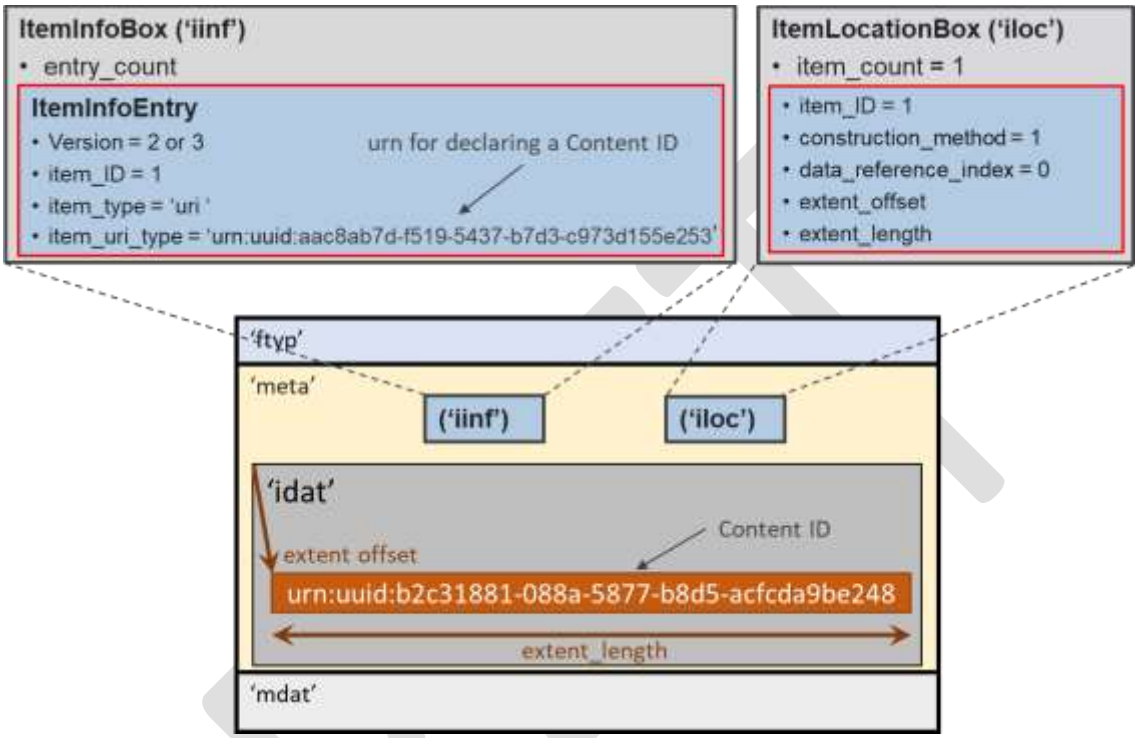


Figure 6: Allocation of a Content ID in a MetaBox.

For reference, the MISB maintains a Namespace UUID as the basis for creating UUIDs in MISB and other standards. Appendix C of the MISP [20] provides the Namespace UUID. Figure 7 shows the result of creating a version 5 UUID using the MISP Namespace UUID and the text "Content ID". When allocating Content IDs, this standard mandates the item_uri_type to be this value.

item_uri_type = 'urn:uuid:aac8ab7d-f519-5437-b7d3-c973d155e253'

Figure 7: item_uri_type for allocating a Content ID

A Content ID is unframed (i.e., without any wrapper content) when stored and placed in any location allowed by ISOBMFF, such as in a MediaDataBox or the ItemDataBox of the containing MetaBox. The exception to this is the Information Security Marking metadata, which must be located with the ISM.XML data in the ItemDataBox of the file level MetaBox.

When generating a Content ID for an item, a companion item reference of type 'cdsc' (content describes) indicates a reference from the Content ID to the specific item it is labeling. This provides the mechanism for attaching the Content ID to an item. Figure 8 shows an example of where a Content ID references an uncompressed image item.

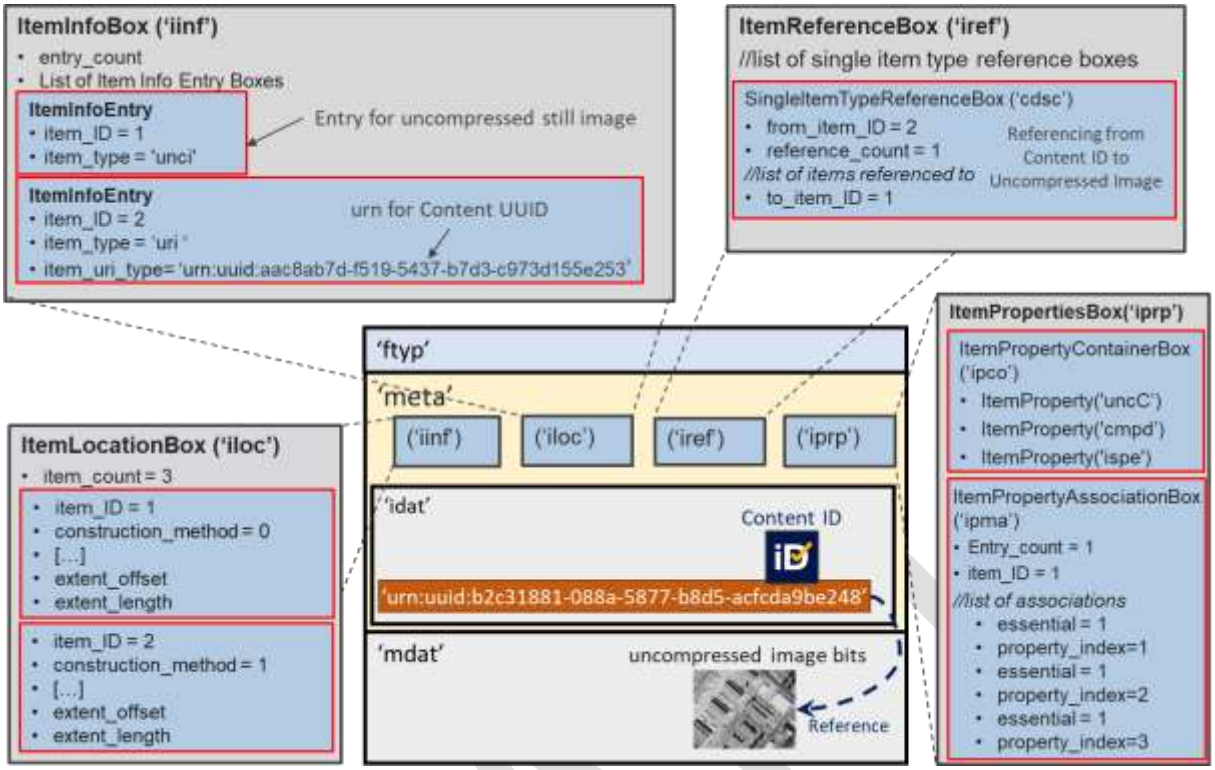


Figure 8: Referencing a Content ID to an image item.

When generating a Content ID for a Motion Imagery track (or other timed media track), the Content ID identifies the entire track and is allocated in the track's MetaBox. Since the Content ID is associated with the track through encapsulation, there is no need for a reference. For other items in the track, such as for static track metadata, writers generate Content IDs and they require references. As such, one, and only one, unreferenced Content ID is allocated within a MetaBox inside a TrackBox and it applies to the track.

Content IDs maintain a universally unique identifier for each piece of media content in a file and are immutable. The Content ID identifies a specific sequence of bits, which represent a piece of content, such as an image item, a metadata item, or a track. If the bits change in any way, including the addition of more bits, systems generate a new Content ID and reference the ID to the modified content. In this way, a provenance tracking system can maintain information showing the media with Content ID #2 derived from the media with Content ID #1 and how/why it differs. If an old form of content no longer exists, either through modification or deletion from the file, applications remove the associated Content ID from the file. Should both the old and new forms of content remain, both Content IDs remain and reference their respective content.

Requirement(s)	
NGA.STND.0076_1.0-42	Content IDs shall be allocated in an ItemInfoEntry box with an item type of 'uri '.
NGA.STND.0076_1.0-43	When allocating a Content ID, the 'item_uri_type' value in the ItemInfoEntry box shall be set to: 'urn:uuid:aac8ab7d-f519-5437-b7d3-c973d155e253'

NGA.STND.0076_1.0-44	Instances of Content IDs shall be generated in the form of 'urn:uuid:<uuid>', as per NGA.RP.0001_1.0.0.
NGA.STND.0076_1.0-45	When an image item is created, a Content ID shall be generated for the image.
NGA.STND.0076_1.0-46	When a Content ID is generated for an image item, an item reference of type 'cdsc' shall be implemented from the Content ID item to the image item.
NGA.STND.0076_1.0-47	When a generalized GEOINT metadata item is created, a Content ID shall be generated for the metadata item.
NGA.STND.0076_1.0-48	When a Content ID is generated for a metadata item, an item reference of type 'cdsc' shall be implemented from the Content ID item to the metadata item.
NGA.STND.0076_1.0-49	All tracks shall have a Content ID.
NGA.STND.0076_1.0-50	A Content ID for a track shall be allocated in the track's MetaBox.
NGA.STND.0076_1.0-51	A track's Content ID shall not have a reference.
NGA.STND.0076_1.0-52	There shall be one, and only one, Content ID without a reference in a track MetaBox.
NGA.STND.0076_1.0-53	When media content is deleted from a file, the media's referenced Content ID shall be removed.
NGA.STND.0076_1.0-54	When media content is modified, a new Content ID shall be generated and referenced to the modified content.
NGA.STND.0076_1.0-55	When media content is modified and the original content does not continue to exist, the original media's Content ID shall be removed.
NGA.STND.0076_1.0-56	When media content is copied to a new location, the same Content ID and reference associated with the original media shall be re-generated for the copy.

6.4.4 Time Annotation

Two separate forms of timing exist within a file.: ISOBMFF Time and Nano Precision Time Stamp ISOBMFF Time is an internal mechanism which ISOBMFF and HEIF use for labeling absolute time and for creating a timeline for timed media. The MISB defines the DoD/IC Nano Precision Time Stamp, which is traceable to TAI and GPS timing standards.

Absolute time labels in ISOBMFF enable storing the creation time and the latest modification time for a media item or group of entities using the CreationTimeProperty and ModificationTimeProperty, respectively. The MovieHeaderBox lists the creation time and the modification time, each time is a 64-bit absolute UTC time value with an epoch of midnight, Jan. 1, 1904. A shortcoming of these time values is UTC associated leap seconds, and a lack of metadata identifying source, status, and accuracy for the time values.

For media tracks, ISOBMFF defines user-configurable timescale. The 'timescale' parameter is an integer specifying the timescale for an entire presentation. ISOBMFF defines the timescale to be the number of time samples occurring over a period of one second. While this facilitates the playout of a movie or sequence, the standard timing mechanisms are not suitable for high-precision GEOINT applications.

To provide necessary time labeling for high performance, high accuracy applications, timing information conformant with DoD/IC standards is mandatory. Absolute timestamps for media, including still media, media sequences, timed metadata, and Motion Imagery, are formatted and implemented according to the MISB ST 1603 Time Transfer Pack [21], which is based on the MISP time system defined in MISB ST 0603 [8] and described in the MISP Handbook [22]. The MISP time system is synchronous with the TAI and GPS time systems and avoids problems associated with leap seconds, which occasionally occur in the UTC time system. Timestamps utilized for GEOINT measurements and calculations are not based on UTC time for this reason.

Software applications convert Nano Precision Time Stamps to UTC with the proper leap second corrections and adjustments for time zones, etc. for human readable display. Applications need to be mindful of properly conveying to users the type and source of time when displayed.

Sections 6.4.6.1 describes the formatting and carriage of Nano Precision Time Stamps for Still Imagery and Section 6.5.2 describes the Nano Precision Time Stamp formatting and carriage for Motion Imagery and other timed media tracks.

Requirement(s)	
NGA.STND.0076_1.0-57	For absolute time labeling, the MISP Time system, defined in MISB ST 0603, shall be the basis for time information.
NGA.STND.0076_1.0-58	MISB ST 0603 Nano Precision Time Stamps shall be used for absolute time labeling of media content.
NGA.STND.0076_1.0-59	Absolute time labeling shall be encoded using KLV as per MISB ST 1603.
NGA.STND.0076_1.0-60	Absolute time labeling shall include time quality metadata as specified in MISB ST 1603.
NGA.STND.0076_1.0-61	Labeling of timing variations related to rolling shutter, raster-scanning devices, and other sensor types shall conform to MISB ST 1507.
NGA.STND.0076_1.0-62	GEOINT calculations and functionality shall rely on timestamps based on the MISP Time System.

6.4.5 Security Metadata

To ensure writers properly and consistently mark files with security labeling, Information Security Marking metadata is allocated and carried in the file level MetaBox. Writers generate and encode the security metadata as per the ODNI XML Data Encoding Specification for Information Security Markings (ISM.XML). I security markings are encoded as an XML data object. ISOBMFF defines a mechanism for the allocation and carriage of general XML content via a MetaBox. In the ItemInfoBox a writer includes an ItemInfoEntry box to implement the allocation. The process involves assigning the following parameters in the ‘iinf’ box:

- an item ID
- setting the ‘item_type’ to ‘mime’
- setting the content type equal to ‘application/dni-arh+xml’

The writer assigns a storage location for the ISM.XML content using the ItemLocationBox. To ensure the ISM content is in a known and readily accessible location, this standard stores the

ISM content in the file level MetaBox's 'idat' box. Writers generate a Content ID (see in Section 6.4.3) and the writer generates an item reference to link the two together. This provides external metadata files with an ability to identify and locate the security information for the content inside the file.

Figure 9 shows the instantiation of information security marking metadata at the file level. The setting of the construction_method to '1' in the ItemLocationBox assigns the storage location to the ItemDataBox ('idat') of the file level MetaBox. Figure 10 shows the allocation of a Content ID for the security marking metadata and the generation of a reference from the Content ID to the ISM.XML information. The ISM.XML and associated Content ID are required for an implemented file to be conformant with this standard.

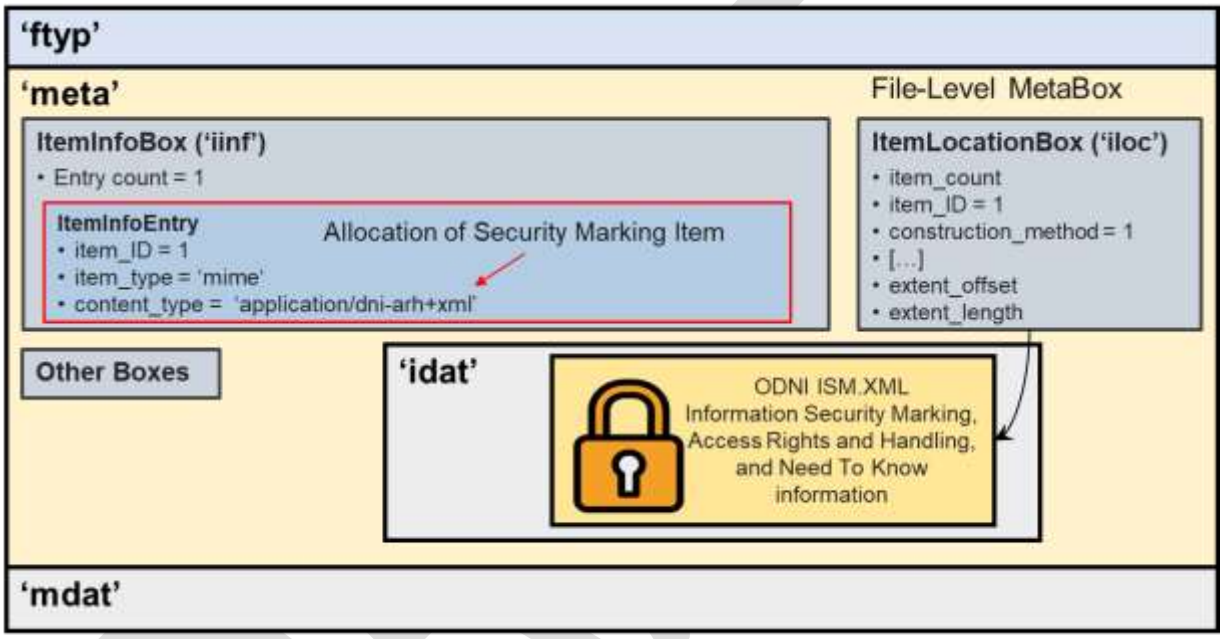


Figure 9: Declaration of Information Security Markings (ISM.XML).

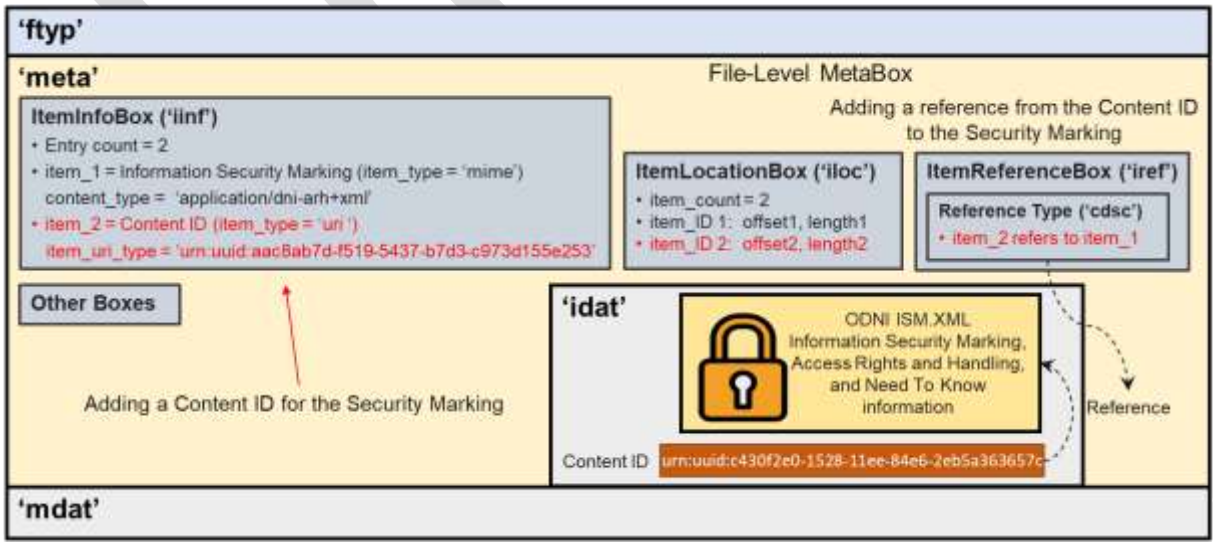


Figure 10: Adding a Content ID to the Information Security Marking metadata.

Requirement(s)	
NGA.STND.0076_1.0-63	Security markings shall conform with ODNI XML Data Encoding Specification for Information Security Markings (ISM.XML), Version 2021-NOVr2022-NOV.
NGA.STND.0076_1.0-64	Security markings shall be located in the file-level MetaBox.
NGA.STND.0076_1.0-65	A MetaBox declaring the ISM.XML metadata item shall set the ItemInfoEntry field item_type = ‘mime‘
NGA.STND.0076_1.0-66	A MetaBox declaring the ISM.XML metadata item shall encode the ItemInfoEntry field content_type as ‘application/dni-arh+xml’
NGA.STND.0076_1.0-67	The ISM.XML content shall be stored in the ItemDataBox in the file-level MetaBox.
NGA.STND.0076_1.0-68	When an Information Security Marking ISM.XML item is created, a Content ID shall be generated for the ISM.XML item.
NGA.STND.0076_1.0-69	The Content ID for the ISM.XML item shall be stored in the ItemDataBox of the file-level MetaBox.
NGA.STND.0076_1.0-70	An item reference from the Content ID to the ISM.XML item shall be implemented.
NGA.STND.0076_1.0-71	Where classification banners are included in an image overlay, they shall be at the highest level of the layering order.
NGA.STND.0076_1.0-72	Classification banners shall pertain only to the classification of the displayed overlay.

6.4.6 Still Imagery Metadata

This profile provides multiple techniques to implement metadata about a specific still image item. These fall into one of the following categories:

1. Item properties, defined in HEIF
2. Nano Precision Time Stamps
3. Content IDs
4. Information Security Marking
5. General application metadata about the image
6. General application metadata about defined regions of an image

The following sections describe details for including these types of metadata for an image item.

6.4.6.1 Attaching Timestamps to Still Images

To support GEOINT functionality, every still image is required to be associated with a high-resolution, absolute timestamp and associated metadata describing the source and quality of the timestamp. MISB ST 1603 Time Transfer Metadata defines the carriage of the timestamp and its metadata. The Nano Precision Time Stamp represents a sampled and quantized time value of the MISP Time System. The implementation allocates a KLV metadata item, based on the Nano Time Transfer Pack 16-byte key, and an item reference from the timestamp information to its associated still image item. Figure 11 identifies the item_type_uri entry for allocating a ST 1603 KLV Nano Time Transfer Pack, which can then reference a still image

item. When implementing the 16-byte key the ST 1603 document is the authoritative source, as versioning may cause the key value to change over time.

ST1603 Nano Time Transfer Pack Key: 06.0E.2B.34.02.05.01.01.0E.01.03.02.09.00.00.00
theURI: 'urn:smpte:ul:060E2B34.02050101.0E010302.09000000'

Figure 11: The item_uri_type for allocating a Nano Time Transfer Pack

Figure 12 shows the allocation of an uncompressed image ('unci') and timestamp metadata ('uri'), alongside the file level security metadata item of Figure 9, in the ItemInformationBox. The ItemLocationBox then defines the storage location of these items within the MediaDataBox. The ItemReferenceBox defines the linking of the timestamp information to the image using a content describes ('cdsc') reference type from the Nano Time Transfer Pack item to the image item. The items related to the allocation of the uncompressed image are colored green and the timestamp information are colored in red. Note the figure also shows the ISM.XML and its Content ID (as derived in Figure 10).)

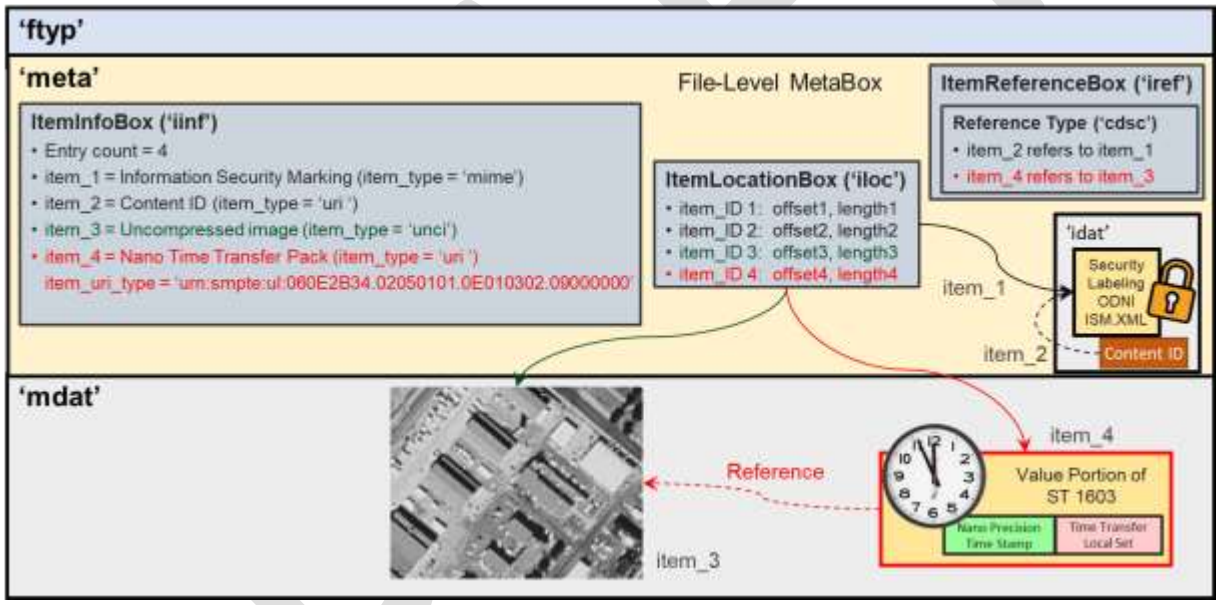


Figure 12: Allocation of a still image and its associated timestamp.

Requirement(s)	
NGA.STND.0076_1.0-73	MISB ST 1603 Nano Precision Time Stamps shall be used for absolute time labeling of media content.
NGA.STND.0076_1.0-74	Absolute time labeling shall be encoded using KLV as per MISB ST 1603.
NGA.STND.0076_1.0-75	Absolute time labeling shall include time quality metadata as specified in MISB ST 1603.
NGA.STND.0076_1.0-76	Each still image in a file shall have an item reference from a MISB ST 1603 Nano Time Transfer Pack metadata item.

NGA.STND.0076_1.0-77	Where MISB ST 1507 is used it shall reference an associated image item.
NGA.STND.0076_1.0-78	To avoid confusion and possible error, the creation time information property ('crtt') shall be omitted in a file.

6.4.6.2 Attaching Content IDs to Still Images

Assigning a ContentID as defined in 6.4.3 to individual image items uniquely identifies each image item, which is important in a distributed setting. Content IDs applied internally to a file are instantiated as a metadata item and referenced to an image item using a 'cdsc' reference type. Figure 13 shows the items related to the allocation of the identification information highlighted in red. In this example, adding Content ID (i.e., item_5) to the previously allocated content from the example of Figure 12, provides a reference to the uncompressed image item (i.e., item_3).

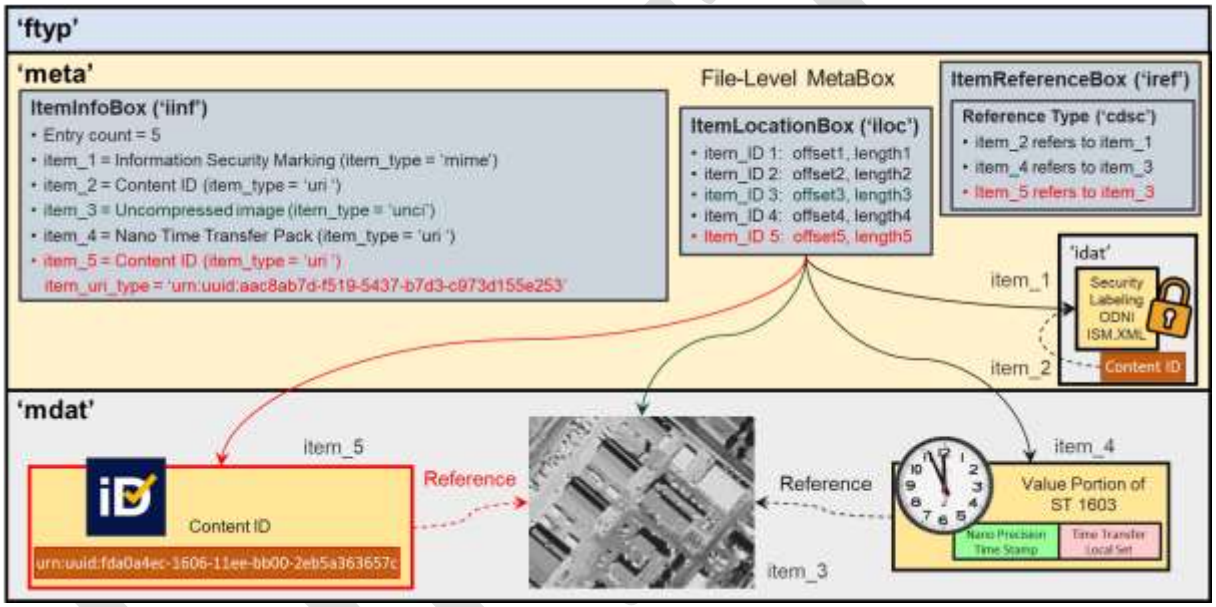


Figure 13: Applying a Content ID to an image item.

While HEIF provides an item identifier, defined as either a 16-bit or 32-bit integer, for each image in a file, the scope of the identifier is only within the file. In addition, there is no guarantee it remains constant over the life of a file or when exporting to another file. If two files contain one image, each with item_ID set to '1' and the images are both copied to a new file, the ID of at least one of the images needs to be changed. In such a case, the referencing for the item with the changed ID needs to be redirected to the new item_ID.

Requirement(s)	
NGA.STND.0076_1.0-79	Each still image in a file shall have a reference from an allocated Content ID, which provides a universally unique ID for the image.
NGA.STND.0076_1.0-80	A Content ID shall be referenced to a still image item using the content describes ('cdsc') reference type.

NGA.STND.0076_1.0-81	A Content ID attached to an image shall persist for that image until the image is deleted.
NGA.STND.0076_1.0-82	Where an image item is copied to another file, the same Content ID shall maintain the same relationship to the image.
NGA.STND.0076_1.0-83	Where an image item is modified, a new Content ID shall be generated for the new instance and form of the image item.

6.4.6.3 Attaching General Metadata to Still Images

For instances of encoding generalized GEOINT metadata and referencing to a specific still image item, the process previously described for allocating a KLV metadata item is the same. In the case shown in Figure 14, a MIMD defined length pack declared as item_6 references the image item declared as item_3. In similar fashion, any KLV set or pack included in a file can reference a declared image item. The MIMD model is the current standardized approach for including sensor and platform metadata in imagery files.

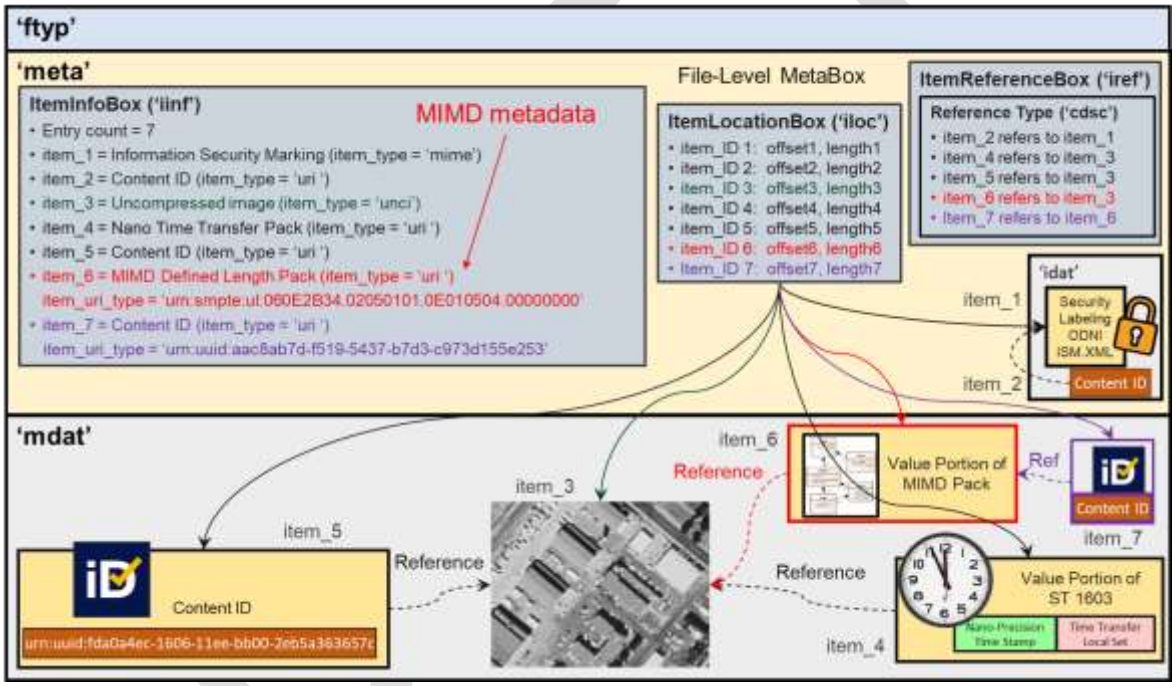


Figure 14: Attaching MIMD GEOINT metadata to an image item.

Requirement(s)	
NGA.STND.0076_1.0-84	Generalized metadata referenced to a still image shall use a MISB approved KLV set or pack.
NGA.STND.0076_1.0-85	Systems including sensor and platform metadata inside a file shall utilize the MIMD suite of standards.
NGA.STND.0076_1.0-86	Where MIMD metadata is included in a file, it shall be encoded in KLV as per MISB ST 1903.
NGA.STND.0076_1.0-87	All instances of generalized metadata shall be referenced (through a 'cdsc' reference) by a unique Content ID.

NGA.STND.0076_1.0-88	Generalized metadata shall be referred to one or more still image items using the content describes ('cdsc') reference type.
NGA.STND.0076_1.0-89	Region item metadata shall include the union of metadata referenced to the parent image item and metadata referenced to the region item.
NGA.STND.0076_1.0-90	Generalized metadata referred to a region item shall be encoded using a MISB-defined KLV set or pack.
NGA.STND.0076_1.0-91	Generalized metadata shall be referred to a region item using the content describes ('cdsc') reference type.
NGA.STND.0076_1.0-92	A Content ID shall be generated for each region item.
NGA.STND.0076_1.0-93	A 'cdsc' reference shall be generated from the Content ID to the region item.

6.5 Track Metadata

Tracks contain timed sequences of media, which may include imagery, audio, or GEOINT metadata. Metadata for tracks consist of both static metadata about a track as well as time varying metadata that may be synchronous or asynchronous with a track's dynamic sample media. In terms of metadata applied to tracks, the following levels of scoping apply:

1. File level static metadata, declared in the file level MetaBox – applies to all content in the file, including tracks. A specific example for this metadata is the ISM.XML security information.
2. Presentation level static metadata, declared in the MetaBox inside the MovieBox – applies to all tracks in the MovieBox. Currently, no specific presentation level metadata is defined in this standard.
3. Track specific static metadata, declared in the MetaBox inside the TrackBox – applies to the full track. Track specific metadata includes a Content ID for a track.
4. Sample auxiliary information (SAI) includes individual metadata elements mapped directly to individual media samples in a track using the sample auxiliary information boxes ('saiz' and 'saio') in the Sample Entry. SAI includes Nano Precision Time Stamps for track media samples.
5. Timed metadata, contained in a parallel track of time varying samples, applies to a track when it is referenced to the track using a track reference of type 'cdsc'.
6. Timed metadata tracks may be synchronous or asynchronous to another track it references.

6.5.1 Static Metadata for Tracks

Static GEOINT metadata describing a track is allocated as one or more metadata items in the MetaBox of the track. The union of these items defines the static metadata for the track. In addition to the Content ID for the track itself, individual Content IDs are generated for each allocated metadata item and referenced to those items. Content IDs are generated as described in Section 6.4.3.

KLV encoding is utilized for static GEOINT metadata when applied to tracks, unless otherwise specified, such as with Content IDs. This follows the same process for generating metadata items in the file level MetaBox. The allocation of a Content ID for a track is shown in Figure 15. Content IDs for each additional metadata item require a 'cdsc' reference from the Content ID to the metadata item. The track does not receive a time stamp as individual time stamps are

annotated with each sample in the track via Sample Auxiliary Information, as described in Section 6.5.2.

Additional static metadata addressing the track is allocated in the same manner as shown for the file level MetaBox (6.4.2). Referencing for metadata items describing the track is not necessary as the default scope for metadata allocated in this MetaBox is the track.

Tank Image credit: White Sands Missile Range

Figure 15: Attachment of a Content ID to a Motion Imagery track

Requirement	
NGA.STND.0076_1.0-94	Static metadata applying only to a track shall be instantiated as an item within the track's MetaBox.
NGA.STND.0076_1.0-95	Timed media tracks (motion imagery, image sequences, timed metadata, and audio) shall include a Content ID, instantiated in the track's MetaBox.
NGA.STND.0076_1.0-96	There shall be no reference for the Content ID of a track, as it is associated with the track through encapsulation.
NGA.STND.0076_1.0-97	The Content ID for a track shall be unframed.
NGA.STND.0076_1.0-98	Unless specified differently as a specific case, static GEOINT metadata about a track shall be encoded as KLV.

6.5.2 Sample Auxiliary Information

Sample auxiliary information (SAI) allows for carrying unique GEOINT metadata for each sample in a timed track. When software interrogates these boxes, the aux_info_type and aux_info_type_parameter variables signal how to interpret the information coded as sample auxiliary information. While the SAI boxes, in general, support user defined metadata encoding, this standard utilizes a default encoding method of Key-Length-Value. To implement KLV encoded SAI, the aux_info_type parameter is coded with the 4CC code 'klvm'. The aux_info_type_parameter variable is used to code an index to identify a pair of 'saiz'/'saio' boxes that go together and represent a specific "stream" of sample auxiliary information. Each unique stream of sample auxiliary information must have the same aux_info_type and aux_info_type_parameter values.

When implementing SAI as KLV encoded information about samples, configure the ‘saiz’ and ‘saio’ boxes with the settings shown in Table 11 for the aux_info_type and aux_info_type_parameter variables:

Table 11: Carriage of KLV encoded metadata as sample auxiliary information.

Parameter	Description
aux_info_type	A registered 4CC code set to ‘klvm’
aux_info_type_parameter	An unsigned 32-bit integer encoded with the following assignment of bits (with bit 31 as MSB and bit 0 as LSB): • bits 31 through 8: Reserved and set to 0. • bits 7 through 0: 8-bit integer identifying a specific stream of sample auxiliary information.

Setting the aux_info_type to ‘klvm‘ signals the sample auxiliary information is encoded using Key-Length-Value, with the information stored as a MISB documented KLV set or pack. Using KLV encoding for sample auxiliary information provides a general, extensible approach for carrying GEOINT metadata, which includes the required annotation of samples with Nano Precision Time Stamps. The default maximum size for sample auxiliary information is 255 bytes, dictated by the “unsigned int(8)” variable declarations for the default_sample_info_size and sample_info_size variables within the ‘saiz’ box, as shown in Figure 16.

Requirement(s)	
NGA.STND.0076_1.0-99	When annotating media samples with KLV encoded GEOINT metadata, the *aux_info_type* in ‘saiz’ and ‘saio’ boxes shall be set to ‘klvm‘.
NGA.STND.0076_1.0-100	The 8 least significant bits of the aux_info_type_parameter used to identify a stream of sample auxiliary information shall be set to the same value in both the ‘saiz’ and ‘saio’ box for a particular stream.
NGA.STND.0076_1.0-101	The 24 most significant bits of the aux_info_type_parameter are reserved for future use and shall be set to zero.
NGA.STND.0076_1.0-102	Motion Imagery samples in a track shall be annotated with ST 1603 Nano Time Transfer Packs encoded as sample auxiliary information
NGA.STND.0076_1.0-103	Image sequence samples in a track shall be annotated with ST 1603 Nano Time Transfer Packs encoded as sample auxiliary information
NGA.STND.0076_1.0-104	Timed metadata samples in a track shall be annotated with ST 1603 Nano Time Transfer Packs encoded as sample auxiliary information
NGA.STND.0076_1.0-105	KLV sets and packs for use as Sample Auxiliary Information shall be selected from the approved list in Table 12.

When encoding sample auxiliary information using KLV, the entire set triplet of KLV information is included in the media storage location. This is a departure from how KLV information implemented using the MetaBox holds the ‘Key’ and the ‘Length’, and the ‘Value’ is in the defined media storage location. Figure 16 shows a generalized example.

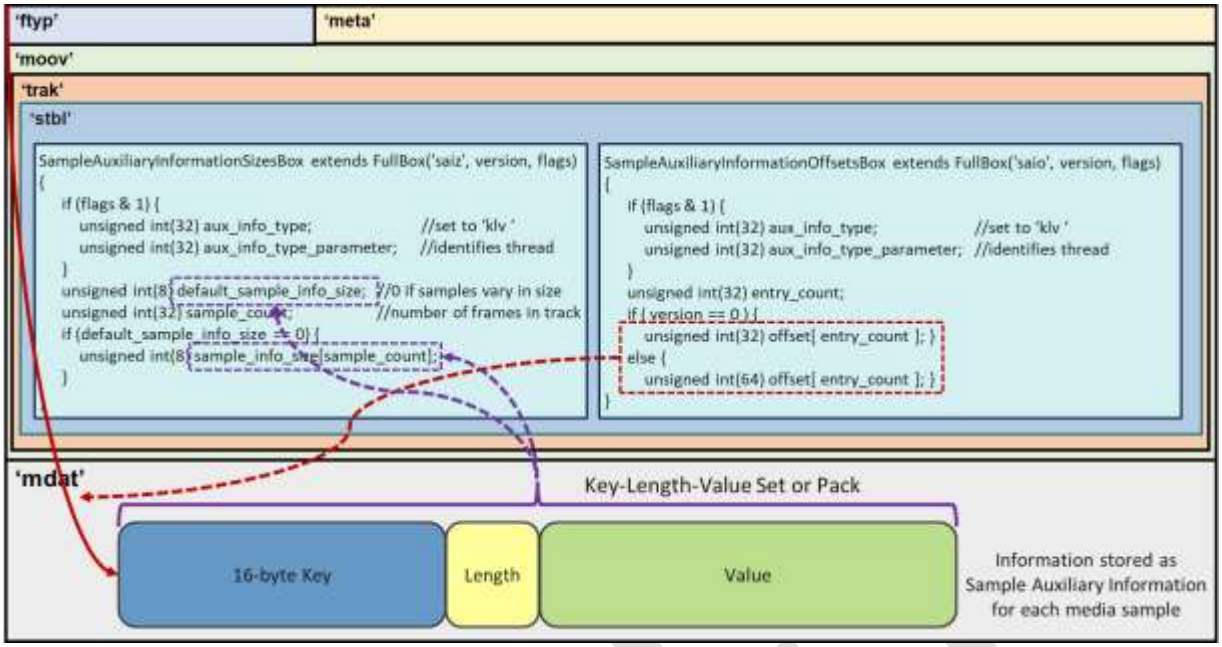


Figure 16: KLV encoding of sample auxiliary information.

Storage of the KLV encoded information representing the sample/frame metadata is typically in the MediaDataBox, although ISOBMFF allows other options, such as external to a file. Individual metadata items internal to each KLV metadata block are located and identified through the arrangement of their tags and lengths, as defined by the specific KLV set or pack standard.

The use of this generalized method for carrying MISB defined keys provides for extensibility. To manage interoperability, KLV sets and packs approved for use are limited by this profile. Table 12 lists those available for use. Carriage of ST 1603 timestamp information is mandatory; carriage of MIMD packs is allowed. The inclusion of MIMD provides the means to carry generalized GEOINT metadata synchronous with an image frame. The requirements for timestamps on all samples and the Content ID for the track facilitate traceability and interoperability on an Enterprise level. For GEOINT metadata not synchronous, or not directly applicable to a frame, carriage in a separate metadata track is the appropriate implementation.

Table 12: KLV sets and packs for use as sample auxiliary information

MISB Key	MISB Standard	KLV Local Set/Pack Name
06.0E.2B.34.02.05.01.01.0E.01.03.02.09.00.00.00 (CRC 58798)	ST 1603	Nano Time Transfer Pack
06.0E.2B.34.02.05.01.01.0E.01.05.03.00.00.00.00 (CRC 12649)	ST 1902	MIMD Defined Length Pack

Figure 17 shows an example of allocating Nano Time Transfer Packs using the ‘saiz’ and ‘saio’ boxes, which then associates them with the media samples in the containing track. In the example, a KLV block for a single sample is shown, with the entire KLV block stored in the

MediaDataBox. The block shown, when replicated for each sample in the track, provides each sample with both timestamp and time quality information. The value portion of the ST 1603 KLV block includes an 8-byte Nano Precision Time Stamp, and a Time Transfer Local Set with metadata describing the source, status, and quality of the timestamp.

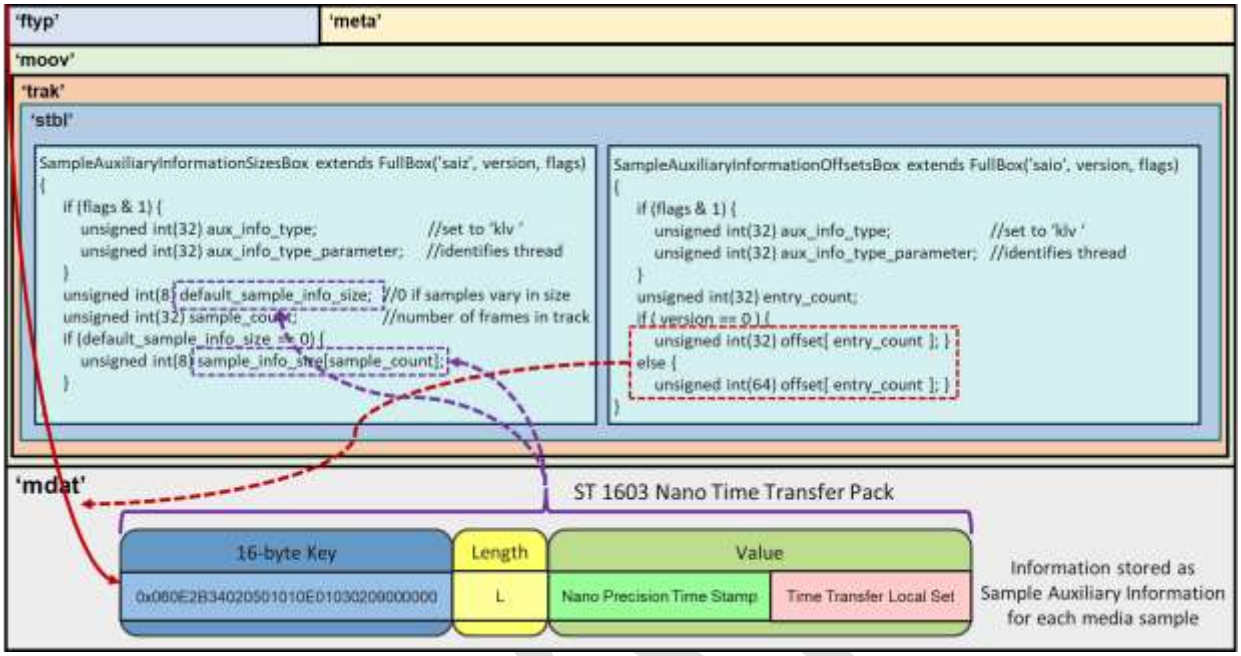


Figure 17: Carriage of ST 1603 timestamp information as SAI.

The carriage of GEOINT metadata as sample auxiliary information is identical for all forms of media tracks, including Motion Imagery, image sequences, and timed metadata. An image sequence and a timed metadata track have the same fundamental GEOINT metadata requirements as a video track, with a Content ID applied to the track and timestamps assigned to each sample. Figure 18 shows the implementation of an image sequence within a TrackBox. The Sample Auxiliary Information boxes ('saiz'/'saio') annotate each sample in the image sequence with Nano Time Transfer Pack information.

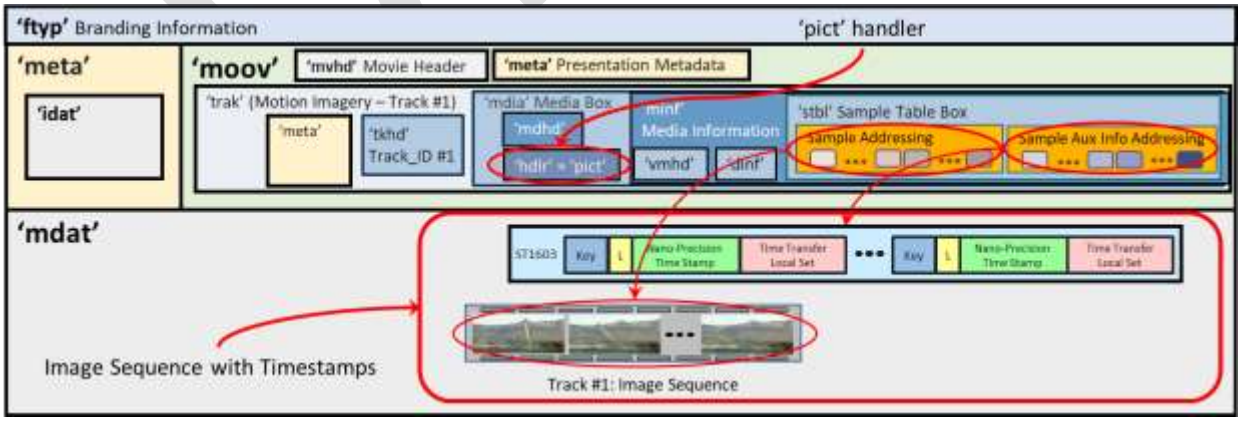


Figure 18: An image sequence track with Nano Precision Time Stamps.

Requirement(s)	
NGA.STND.0076_1.0-106	Timed imagery and metadata media in a track (Motion Imagery, Image Sequence, and Timed Metadata samples) shall include Nano Precision Time Stamp information via 'saiz'/'saio' boxes.
NGA.STND.0076_1.0-107	When written into a file, information in the Time Transfer Local Set shall be filled with known information when available and labeled with the options for 'unknown' when the required information cannot be determined.

6.5.3 Timed Metadata Tracks

Metadata collected or captured as a timed sequence of GEOINT information samples is stored as a media track either within a file or externally in a resource file. This information includes time varying sensor and platform metadata, such as gimbal angles, GPS location, zooming focal lengths, etc. and static metadata. Timed GEOINT metadata tracks in a 'geo1' branded file have the following features:

1. Timed metadata tracks stored in an ISOBMFF file are encoded using MISB-defined KLV encoding.
2. Timed metadata tracks stored outside an ISOBMFF file may be encoded in RDF and other user defined formats.
3. Timed metadata tracks include a Content ID.
4. Timed metadata tracks include Nano Precision Time Stamps stored as Sample Auxiliary Information
5. Zero or more timed metadata tracks can be stored in a file
6. Zero or more timed metadata tracks can be stored in external resource files and referenced from the motion imagery file.
7. Tracks can contain timed metadata that is synchronous or asynchronous with respect to referenced Motion Imagery tracks.
8. Timed metadata tracks may occur at rates independent of other tracks.

Implementing a timed metadata track is like a Motion Imagery track and shares the same timeline structure. When related to a separate media track, track referencing enables linking the separate tracks of dynamic information together to show the relationship. A content describes reference type ('cdsc'), implemented in the TrackReferenceBox of the timed metadata track signals the linkage of GEOINT metadata to Motion Imagery and image sequence tracks.

Timed metadata tracks may also be declared and defined in a Movie Fragment box for fragmented or segmented implementations. ISOBMFF rules defined for chunking and sample runs apply to timed metadata tracks.

A timed KLV metadata track is instantiated by including a TrackBox (inside the MovieBox) dedicated to the track. Implementation involves configuring and signaling a handler type of 'meta' and header information, in a NullMediaInfoHeaderBox ('nmhd'). Including this information indicates the track is not a Motion Imagery or audio track but instead a track for metadata.

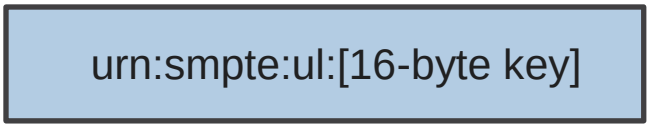


Figure 19: Format of URN sample entry to signal a timed MISB KLV metadata track

The URIMetaSampleEntryBox (‘urim’) inside the SampleDescriptionBox (‘stsd’) indicates the sample entry type through a URI string carried inside the URIBox (‘uri ‘). The string is formatted as a uniform resource name which includes the 16-byte MISB KLV key of interest as shown in Figure 19. The brackets are excluded from the actual URN and the 16-byte key is formatted following the QUADBYTE notation declared in RFC 5119. This is identical to the form of the URN described for implementation of a MISB KLV item type in a MetaBox for static metadata. In the context of a track, it identifies a KLV set or pack used to encode the timed metadata in the track. Current and future MISB-approved sets and packs encompass options available for the carriage of GEOINT metadata. MIMD is the prioritized method for encoding timed metadata tracks. Figure 20 shows the URI for a MIMD track. Of note, this URI may change as MIMD versions and associated keys evolve.

ST1902 MIMD Defined Length Pack Key: 06.0E.2B.34.02.05.01.01.0E.01.05.04.00.00.00.00

theURI: ‘urn:smpte:ul:060E2B34.02050101.0E010504.00000000’

Figure 20: A formatted URN string to signal a sample entry for timed KLV encoded MIMD metadata

In situations where the source of metadata is from more than one KLV set or pack, each distinct set or pack is carried within its own track. The declaration of the 16-byte set or pack outer key identifies the allowed metadata content within each track. For situations where a metadata track is independent and not referenced to another track, the implementation is the same, but without the ‘cdsc’ track reference to another track. In this way, a file with only timed metadata information is possible.

Within this standard the formatting of a MISB KLV track follows the implementation of the box structure shown in Figure 21 with items of note highlighted with a red box. In writing KLV metadata into a timed metadata track, both the 16-byte key and the length for each instance of a KLV packet are stored in the Sample Table box. The value portion of each KLV packet is stored in a chosen location, as defined by the Data Reference box. Choices include a MediaDataBox or an external file resource. This approach allows the Key to be stored efficiently, just once. Sample auxiliary information stores the Key with each piece of sample metadata, which is less efficient.

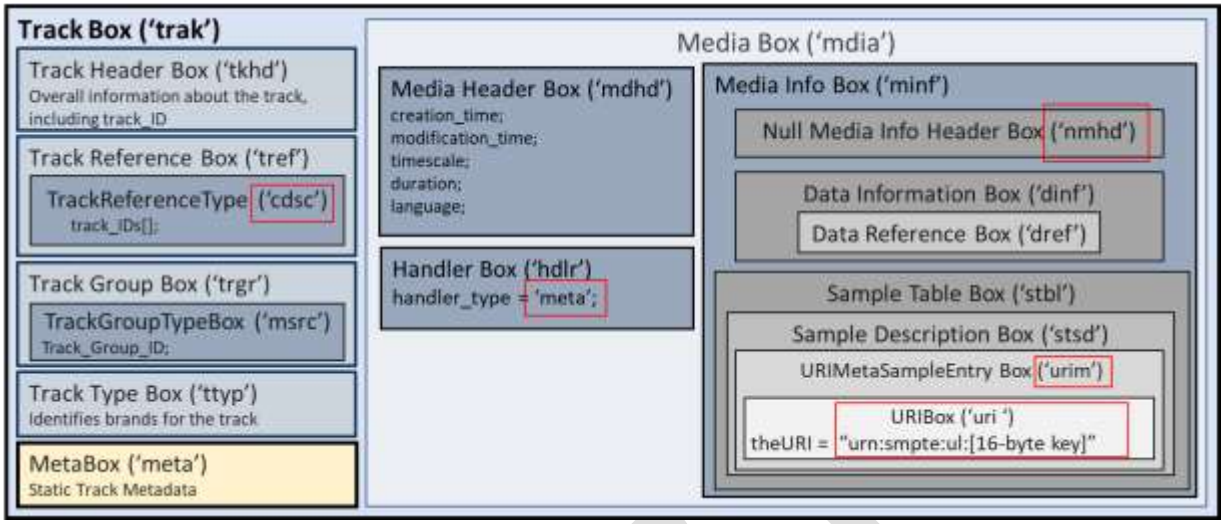


Figure 21: Configuration of a TrackBox for a KLV encoded timed metadata track

Figure 22 shows a Motion Imagery track with file level security metadata, a Content ID applied to the track, timestamps applied to each image in the track, and a separate, parallel track of timed MIMD metadata adding additional context for the Motion Imagery.

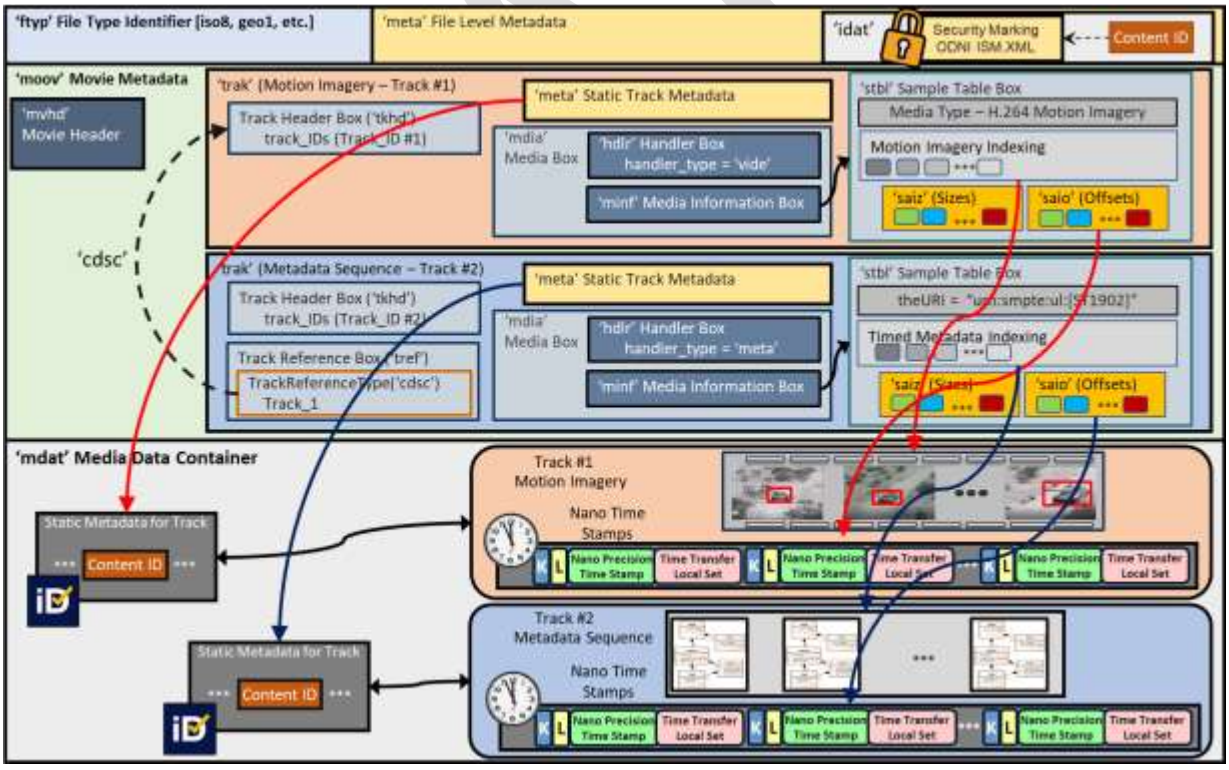


Figure 22: A populated file with Motion Imagery, timed MIMD, Content IDs, and Security marking.

Requirement(s)	
NGA.STND.0076_1.0-108	Metadata in a timed metadata track shall be sourced from a single MISB-approved KLV set or pack.
NGA.STND.0076_1.0-109	Timed metadata tracks shall utilize the URIMetaSampleEntryBox with a URI formatted as a URN in the format of: 'urn:smpte:ul:16-byte key', with the 16-byte key following the QUADBYTE notation declared in RFC 5119.
NGA.STND.0076_1.0-110	KLV metadata tracks shall contain data defined according to MISB specification.
NGA.STND.0076_1.0-111	Consistent with ISOBMFF track_ID labeling and track referencing, timed metadata tracks shall reference Motion Imagery track(s) to which they apply.
NGA.STND.0076_1.0-112	Consistent with ISOBMFF track_ID labeling and track referencing, timed metadata tracks shall reference image sequence track(s) to which they apply.
NGA.STND.0076_1.0-113	Consistent with ISOBMFF track_ID labeling and track referencing, timed metadata tracks shall reference groups of track(s) to which they apply.

6.5.3.1 Carriage of multiple timed KLV metadata tracks in a file

When a file contains timed metadata from multiple KLV sets or packs, the metadata from each unique pack or set is carried in its own timed metadata track. They are configured in the normal manner for a timed KLV metadata track. Each individual track carries a unique track ID, a unique Content ID, a URN defined by the 16-byte key of its set or pack, and sample auxiliary information with Nano Precision Time Stamps.

In situations where a complex set or pack with multiple data items fall on a single instance of time, the multiple data items require encapsulation into a single block of information. This meets the ISOBMFF requirement that no two samples within a track can share the same decoding time or the same composition time. If the data items cannot be merged into a single encapsulated block, multiple tracks may be required to carry the data.

6.5.3.2 Referencing timed KLV metadata to other media tracks

When a timed KLV metadata track is referenced to one or more other tracks (primarily Motion Imagery tracks), a track reference indicates the relationship. A track reference type of 'cdsc' (content describes) denotes the general relationship of how metadata provides additional information and context for another media track. In certain cases, one metadata track may apply to multiple Motion Imagery tracks. For example, a system with two cameras on a single gimbal or ball turret system generates two tracks of Motion Imagery, but the metadata generated from the gimbal and the platform is common to both. In this case, the single track of platform metadata references both Motion Imagery tracks.

Requirement(s)	
NGA.STND.0076_1.0-114	Timed metadata tracks shall be associated with the imagery tracks they apply using the TrackReferenceBox ('tref').

NGA.STND.0076_1.0-115	When referencing a timed metadata track to a Motion Imagery track, the Track Reference (‘tref’) type shall be ‘cdsc’ (content describes).
NGA.STND.0076_1.0-116	When referencing a timed metadata track to an image sequence track, the Track Reference (‘tref’) type shall be ‘cdsc’ (content describes).

6.5.3.3 Rate partitioning with KLV metadata tracks

Generating and storing separate tracks based on the sample rates of the underlying data can facilitate the organization of metadata. Proper sample rates dictated by the natural bandwidth of a particular sensor sub-system typically results in different sub-components sampled at different rates and offsets, resulting in asynchronous tracks. Using a rate partitioning strategy, sampling of metadata content produce tracks sampled at “like” rates. These can then be organized as separate groups of tracks using the ISOBMFF track grouping construct. Figure 23 shows a simple example of sensor system data sampled asynchronously and at different rates.

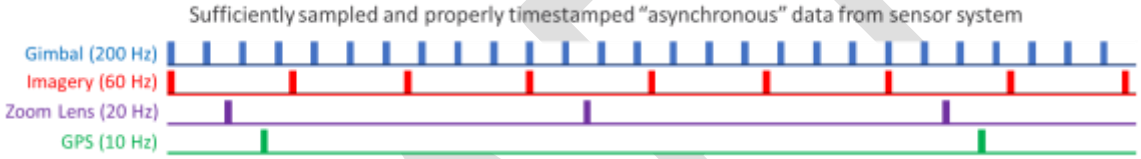


Figure 23: A rate partitioning example.

6.5.3.4 Generation of synchronized KLV metadata tracks

While many sensor systems collect sub-system data at multiple, unsynchronized rates, this creates a challenge for processing systems attempting to leverage this information when generating refined results through exploitation and analysis. A common strategy is to take natively sampled source information and process the data through filtering, interpolation, and re-sampling to generate a synchronous version of the data set, as shown in Figure 24. This processed metadata track added to a file can facilitate further exploitation and analysis of the overall dataset. This is especially useful with automated approaches, such as with AI/ML systems. When generated within the MovieBox of the same file as the original source track(s) they can reference the Motion Imagery track to which they apply. If included as a separate track, the content samples can be added to the same file or stored in an external file (such as RDF) and referenced using the Data Reference (‘dref’) box.

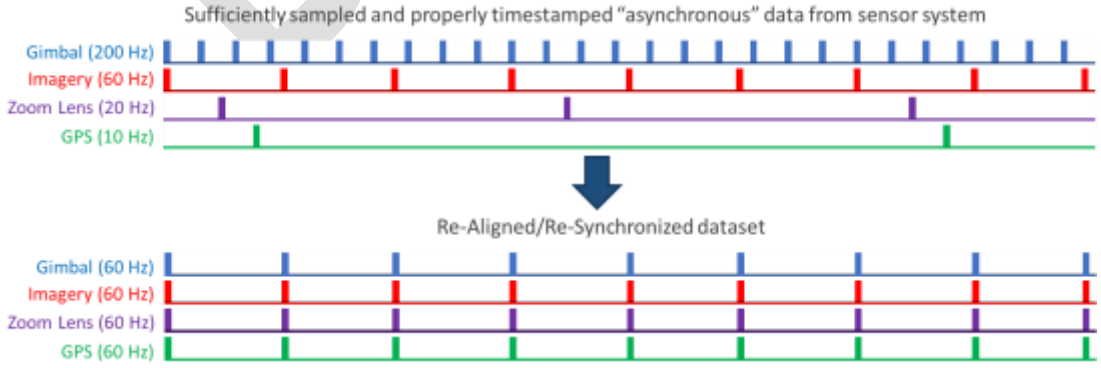


Figure 24: Generation of a synchronous metadata track via processing.

6.6 Audio tracks

Applications may require (or desire) the inclusion of audio content. This standard does not require but does allow for encoding and including audio content within ISOBMFF as per the codecs defined in MISB ST 1001 [23], and the associated formatting and carriage capabilities related to these codecs defined in ISO/IEC standards 14496-3 [24] and 14496-12.

Requirement(s)	
NGA.STND.0076_1.0-117	Where audio is included in a file, it shall utilize an approved audio codec per MISB ST 1001.
NGA.STND.0076_1.0-118	Where audio is included in a file, it shall be conformant with ISO/IEC 14496-3.
NGA.STND.0076_1.0-119	Where audio is included in a file, it shall be conformant with ISO/IEC 14496-12.

6.7 Imagery Access, Retrieval, Playout, Delivery and Display

File configuration and arrangement of content can have a significant impact on user experience when accessing imagery, especially when over network/cloud interfaces. Imagery products may be delivered and accessed through diverse use cases, such as:

- Local hard drive access
- Removable file media (Removable hard drive, CDROM, USB drive, etc.)
- Web services, Cloud, etc.
- File transfer and download services

A broad range of encoding options are available to facilitate performance when accessing image content across these delivery and access mechanisms, especially when files are large and/or complex. The following are primary techniques available for arranging content to improve access performance:

- Interleaving
- Tiling
- Overviews
- Encoding in one form for analysis and another for quick-look access and display
- Optimizing tiles and overviews based on cloud/network access capabilities

It is important to note that access performance and the arrangement of content may greatly depend on the portions of imagery needed for a certain type of analysis, display, etc. For instance, analyzing a sub-region of a large RGB image may dictate tiling and interleaving. Analyzing a hyperspectral image one band at a time, may instead dictate planar or component interleaving.

For maximum interoperability with commercial media applications and tools, coding an image item using the ‘heic’ brand is conformant with HEVC Main and HEVC Main Still Picture profiles. For image sequences, the ‘hevc’ brand provides a high level of interoperability using the same HEVC Main or HEVC Main Still Picture profiles. These offer interoperable options for image coding and may find use with disadvantaged users, disaster recovery situations, etc.

6.8 File Generation and Editing - Privacy and Security Considerations

When generating and editing files, files must address privacy and security considerations. As HEIF can store information not normally displayed by software, software applications require processes and tools to ensure and confirm only content intended and authorized to be in a file is contained within a file. The classification of files must address all content within a file, not just items which are viewable. Hidden images, for example, may have a raised classification level but may not be viewable using readily available software applications intended for commercial or personal use.

Removed images may have links to content which contains sensitive information. In addition, items, and tracks deleted from a MetaBox or the MovieBox require specific actions to delete the content from the MediaDataBox. When removing content from a file, care must be taken to ensure any references, groups, entities, etc. are updated accordingly. Deletion of content, in the context of this standard, involves either rewriting the file without the content to be deleted, or overwriting the content, conformant with required regulations, to ensure the deleted content is not retrievable.

Annex N of ISO/IEC 23008-12 provides an overview of file implementation and editing topics with suggestions for best practices related to privacy and security of image and metadata content.

Requirement(s)	
NGA.STND.0076_1.0-120	When deleting an item from a file, the media for the item shall also be removed (deleted, overwritten, etc.) from the file to ensure it is no longer present.
NGA.STND.0076_1.0-121	When deleting an item from a file, all associated references shall be updated/removed as appropriate.
NGA.STND.0076_1.0-122	When deleting an item from a file, all associated groups shall be updated to remove the deleted item listing from the group.
NGA.STND.0076_1.0-123	When deleting a track residing in a file, the media for the track shall also be removed (deleted, overwritten, etc.) from the file to ensure it is no longer present.
NGA.STND.0076_1.0-124	When deleting a track from a file, all associated references shall be updated/removed as appropriate.
NGA.STND.0076_1.0-125	When deleting a track from a file, all associated groups shall be updated to remove the deleted track listing from the group.
NGA.STND.0076_1.0-126	When deleting an image item from a file, all region items and associated references to the image shall be deleted from the file.

6.9 File Complexity

Rather than provide a list of file complexity categories with ranges of content parameters defined to estimate file complexity (large images, high dynamic ranges, many bands, etc.), this standard leverages branding to provide an upfront indicator of interoperability information to allow an application to easily probe a file for content type and makeup. Files may contain a single type and encoding of content, or a combination of independently usable content of varying types with the same or different encoding methods. Even though a file may be very

large and very complex, there may be individual pieces of content accessible by general software tools. Thus, compatible image and metadata items and tracks can be discovered, accessed, opened, and properly handled as capabilities allow. All other content the reader is not capable of reading and handling is simply avoided. While not a requirement, users are reminded, when appropriate, of the ability to include alternative versions of images encoded in non-commercial/non-consumer codecs. As an example, re-encoding an uncompressed or JPEG 2000 encoded image using HEVC or AVC, possibly in a lower resolution, to provide a visual of the image in a more highly interoperable form.

Annex A – Deprecated Requirements

N/A

Annex B – Codec Brands (Informative)

Branding is used in some cases to indicate codec specific implementation within a certain format standard, such as HEIF. For video tracks, codec signaling is through the Sample Entry code. The use of brands and sample entry codes indicates to a reader the requirements for decoding and interpreting encoded content. Table 13 lists a summary of brands for implementation of image items and image sequences. Additional derivative brands, outlined in referenced documentation for each codec, are available for addressing specific features such as higher-level profiles, etc.

Table 13: Common codec specific brands for Still Imagery and Image Sequences

Image Items		
Codec	Brand	Description
JPEG 2000, HTJ2K	‘j2ki’	The 'j2ki' brand identifies a file with a JPEG 2000 [25] or HTJ2K [26] coded image item. The item type for a JPEG 2000 image item is ‘j2k1’. The JP2 Header box is included as an essential item property ('j2kH') for the image item. J2K and HTJ2K codestreams are carried identically in HEIF. A codestream using HTJ2K is indicated in the CAP marker segment in the codestream header. HEIF metadata alone does not indicate a distinction.
HEVC	'heic'	The ‘heic’ brand identifies a file with an HEVC coded image item of type ‘hvc1’ and coded as per ISO/IEC 23008-2 [27]. It is associated with an essential item property ‘hvcC’.
HEVC	'heix'	The ‘heix’ brand identifies a file with an HEVC coded image item of type ‘hvc1’ and coded as per ISO/IEC 23008-2. The content of the item conforms to the Main 10 profile or any of the format range extensions profiles of HEVC. It is associated with an essential item property ‘hvcC’.
AVC	‘avci’	The ‘avci’ brand identifies a file with an AVC codec image item or type ‘avc1’ and coded as per ISO/IEC 14496-10. It is associated with an essential item property 'avcC'.
Image Sequences		
Codec	Brand	Description
JPEG 2000, HTJ2K	‘j2is’	The ‘j2is’ brand identifies a file with the encapsulation of a sequence of JPEG 2000 or HTJ2K codestreams, as an image sequence with a sample entry of type ‘j2ki’. The J2KSampleEntry is defined in ISO/IEC 15444-16:2021 [10].
HEVC	‘hevc’	The 'hevc' brand identifies a file with an HEVC image sequence track with ‘hvc1’ sample type with bitstream conforming to the Main profile of HEVC.
HEVC	‘hevx’	The ‘hevx’ brand identifies a file with an HEVC image sequence track with ‘hvc1’ sample type with bitstream conforming to the Main 10 profile or any of the format range extensions profiles of HEVC.
AVC	‘avcs’	The ‘avcs’ brand identifies a file with an AVC image sequence track with a sample entry of 'avc1'. The AVCSampleEntry is used as specified in ISO/IEC 14496-15.
Motion Imagery		
Codec	Brand	Description

HEVC	‘hvc1’	The ‘hvc1’ brand identifies a file with the encapsulation of a track of HEVC encoded video with a sample entry of type ‘hvc1’. The HEVCSampleEntry is defined in ISO/IEC 15444-15:2022.
AVC	‘avc1’	The ‘avc1’ brand identifies a file with the encapsulation of a track of AVC encoded video with a sample entry of type ‘avc1’. The AVCSampleEntry is defined in ISO/IEC 15444-15:2022.