

ISO 21812-1:2019(E)

INTERNATIONAL STANDARD

**ISO
21812-1**

First edition
2019-06

**Graphic technology — Print product metadata for PDF
files —**

Part 1: Architecture and core requirements for metadata

*Technologie graphique — Métadonnées des produits d'impression pour les fichiers PDF —
Partie 1: Architecture et exigences principales pour les métadonnées*

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 21812-1:2019

<https://standards.iteh.ai/catalog/standards/sist/5ef9d6e4-94cd-409f-afaf-eca2cafd0667/iso-21812-1-2019>

Deleted: /TC 130 N 3903

Deleted: TC130/WG2/TF5 N 070
ISO/FDIS

Deleted: :

Deleted: (E)

Deleted: ISO TC 130/WG 2/TF5
Secretariat: SAC

Date: 2019-01-14

Deleted: Digital data exchange —

Deleted: FDIS stage
Warning for WDs and CDs
This document is not an
ISO International Standard. It is
distributed for review and comment.
It is subject to change without
notice and may not be referred to as
an International Standard.
Recipients of this draft are invited to
submit, with their comments,
notification of any relevant patent
rights of which they are aware and
to provide supporting
documentation.

ISO 21812-1:2019(E)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 130, *Graphic technology*.

A list of all parts in the ISO 21812 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

PDF files represent content pages and do not normally contain information identifying the usage of these content pages in print production. Document part metadata is a simple mechanism that allows for the exchange of information regarding a set of content pages to aid the receiver of the PDF files in determining the intended use of those content pages in the final print product. By understanding the intended use of content pages, the receiver of the PDF file can make more informed decisions regarding the production process for the final print product.

Several Industry groups have initiated work in the area of workflow control and print product semantics for use with document exchange using PDF. These include CIP4, Ghent Workgroup, the PDF/VT Competence Center, and TC 130 WG 2.

A set of application notes for this document may be found at <http://www.printtechnologies.org/standards/tools--best-practices/>. In addition, pointers may be found on this site to development tools provided for the assistance of developers and users of applications prepared based on this document.

A standard set of such document part metadata is needed to allow composition system and pdf creation vendors to effectively allow their users to communicate with printing and finishing systems that will receive and act on the provided PDF content data. This document defines a standard for document part metadata keys for PDF and their meanings for the purposes of driving workflows or aiding the creation of print production job tickets such as JDF or XJDF.

The intent is to accomplish this through standardizing the document part metadata that can be provided by a document creator. This document builds on the initial CIP4 ICS-Common Metadata for Document Production Workflow published in 2010. This document focuses on defining standardized document part metadata for PDF files using the DPart syntax as defined in ISO 16612-1 (PDF/VT) and ISO 32000-2 (PDF 2.0).

This document is the first part of a series of international standards that define a set of metadata keys and their meanings for use in PDF files to identify printed products and their component pages, to describe their appearance and characteristics and to guide their production.

The structure of the metadata is intended to encapsulate sufficient information in a PDF file to guide the production of printed products without the creator needing to know the details of the production processes that will be used.

It is expected that additional parts of this document will be published that standardize additional print application specific metadata using the architecture defined in this document.

Graphic technology — Digital data exchange — Print product metadata for PDF files — Part 1: Architecture and core requirements for metadata

1 Scope

The document part metadata in a PDF file that conforms to this document can be used to communicate the intended appearance of print products and their components. Examples of intended use are: direct interpretation within a production process, creation of job tickets such as XJDF, or populating records in an MIS. This document builds on the DPart syntax as specified in ISO 16612-1 (PDF/VT) and ISO 32000-2 (PDF 2.0) which is designed for encoding metadata related to pages or groups of pages in PDF files.

NOTE The document part metadata provided in this document applies to individual document parts, whereas XMP metadata typically applies to the scope of the entire document. XMP can apply to the scope of an individual page or part of a page but this usage is very uncommon. Thus, XMP is not applicable for the case where metadata is required for sets of pages such as multiple recipients or binding information. For example, XMP is used within PDF/X for file conformance identification and is also used for additional file level information such as author.

This document defines standardized metadata to:

- provide product intent specifications such as paper media selection and binding information;
- identify the type of product that the content pages are intended to represent (e.g. a brochure, letter or postcard);
- identify the intended recipient of each of the content pages for variable document printing applications.

This document defines a base conformance level that includes the syntax of the metadata framework and the semantics of a core set of metadata.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 16612-2, *Graphic technology — Variable data exchange — Part 2: Using PDF/X-4 and PDF/X-5 (PDF/VT-1 and PDF/VT-2)*

ISO 32000-1:2008, *Document management — Portable document format — Part 1: PDF 1.7*

ISO 32000-2:2017, *Document management — Portable document format — Part 2: PDF 2.0*

ISO 12647-2:2013, *Graphic technology — Process control for the production of half-tone colour separations, proof and production prints — Part 2: Offset lithographic processes*

ISO 3166-1, *Codes for the representation of names of countries and their subdivisions — Part 1: Country codes*

LANGUAGE E.M. (XML) 1.0 (Second Edition), 6 October 2000, World Wide Web Consortium, Available from internet <<https://www.w3.org>>

ISO 21812-1:2019(E)

XJDF Specification, Release 2.0, 2018, CIP4 Organization, Available from internet <<https://www.CIP4.org>>

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <https://www.electropedia.org/>

3.1

JDF

job definition format

3.2

print product

outcome of the processing of a document through a print manufacturing process

Note 1 to entry: Examples include a perfect bound book or postcard.

3.3

product part

part of a print product

Note 1 to entry: Examples include the cover part of a saddle-stitched booklet.

3.4

recipient

person or institution that receives a print product

3.5

XJDF

simplified version of JDF as defined by XJDF Specification Release 2.0

4 Notation

4.1 Keywords

Glossary items are designated in **bold**.

EXAMPLE **recipient**.

Metadata keywords are designated in **bold font**.

EXAMPLE **CIP4_Root**.

Metadata values are designated in *italic font*.

EXAMPLE *true*.

4.2 Cardinality

Optional keys are labelled (Optional) in the description and required keys are labelled (Required).

4.3 Values of lists

This specification provides both open and closed value lists. Open value lists provide a list of suggested values that should be used. Open value lists are marked as (Extendable). Additional values may be

added in case no value in the list sufficiently matches the requirements of the conforming writer. Open lists are identified by specifying that one of the values should be used. Closed lists shall not be extended. Closed lists are identified by specifying that only values that are defined in the list shall be used. Closed value lists are marked as (Closed).

NOTE Some of the standardized metadata values have been defined as open lists of suggested values. The goal is to provide as much interoperability as possible without restricting the use of the standard to a limited set of use cases or print products. If extensions to these open lists are used, the correct interpretation of the extended values needs to be ensured.

4.4 XPath Notation

A notation that is based on XPath will be used to describe nested PDF dictionaries in the **DPart** hierarchy. Unless stated otherwise, no assumption is made whether the respective dictionaries are direct objects or indirect objects within the PDF structure. The root of any such XPath always specifies a child of a **DPM** dictionary. For instance, **CIP4_Root/CIP4_Metadata/CIP4_Conformance** specifies a key named **CIP4_Root** in a **DPM** dictionary that references a dictionary that contains a **CIP4_MetaData** key that references a dictionary that contains a key with the name **CIP4_Conformance**.

5 Conformance

This document specifies a base conformance level for the exchange of document part metadata in PDF files. The base conformance level defines the syntax and semantics of document part metadata properties.

Conforming document part metadata shall conform to all the technical requirements set out in Clauses 6 to 7 of this document. Conforming document part metadata shall include a conforming **CIP4_Root** dictionary at the root of the document part hierarchy of the document part metadata as defined in 7.2 of this document. A conforming writer is an application that shall write a conforming file according to the requirements specified in this document.

A conforming processor is an application that shall read and appropriately process the metadata encoded within a conforming file according to the requirements specified in this document.

A conforming file is a pdf file that contains document part metadata conforming to the requirements specified in this document and that also conforms to ISO 16612-2 (PDF/VT), ISO 32000-2 (PDF 2.0), or any file that is in accordance with ISO 32000-1, such as PDF/X-4 (ISO 15930-7) and that includes an extensions dictionary (ISO 32000-1:2008, 7.12) as follows. The prefix used for the name of the extension shall be **GTS_**, the value of the **BaseLevel** entry shall be 1.7 and the value of the **ExtensionLevel** entry shall be 1.

EXAMPLE In a PDF with only this extension, the extensions dictionary would look like:

```
<<
/GTS_ << /BaseLevel /1.7 /ExtensionLevel 1 >>
>>
```

6 Technical requirements

6.1 Encoding metadata keys

Each metadata key shall be encoded as a PDF name that consists of the second class name prefix of the metadata property followed by an underscore symbol and the name of the metadata property.

Elements and attributes that are defined in the XJDF namespace but not in this document may be used. They shall then be specified using the local name with a prefix of **CIP4**. A conforming writer wishing to add private metadata properties into the CIP4 hierarchy may do so but shall explicitly identify those

ISO 21812-1:2019(E)

private metadata properties and levels by specifying an alternate second class name prefix for that property.

NOTE ISO 32000-2:2017, Annex E contains the definition of second class prefixes.

EXAMPLE A vendor that is using the second class name prefix ACME that wishes to encode a value for a key named foobar in the CIP4_Root/CIP4_Recipient hierarchy will therefore use a metadata property called CIP4_Root/CIP4_Recipient/ACME_foobar.

6.2 Encoding metadata values

6.2.1 Mapping of the encoding of XJDF Intent

Explicit product definitions shall only be specified in the **CIP4_Root/CIP4_Intent** hierarchy. This hierarchy is based on the Intent resources that are defined in chapter 6, Product Intent Description of XJDF Specification, Release 2.0, 2018.

The key names in **CIP4_Intent** shall match the respective XJDF Intent element names. Any attributes on an XJDF Intent element shall be specified as keys in their respective parent level.

6.2.2 Encoding of XML

NOTE Most XJDF datatypes are specified in XML Schema Definition Language (XSD) 1.1 Part 2: Datatypes.

The data types of XML attributes shall be mapped according to Table 1 below.

Table 1 — XJDF datatypes

XJDF datatype	PDF datatype	Comments
integer	integer	
float	number	
List	array	Any list that is encoded in XJDF as a whitespace separated list of base type is encoded as an array of the respective base type, e.g. IntegerList will be encoded as an Array of integers.
Range	array	Any range is encoded as an array of 2 elements of the respective base type, e.g. IntegerRange will be encoded as an array of 2 integers.
Enumeration NMTOKEN ID	name	Computer readable values such as NMTOKEN, enumeration or ID are encoded as names.
Enumerations NMTOKENS	array	Lists of computer readable values such as NMTOKEN, enumeration or ID are encoded as array of names
boolean	boolean	
String	text string	NOTE 1 The encoding of text strings as UTF-8 is only valid in ISO 32000-2.
dateTime date	date string	NOTE 2 See 7.9.4 Dates in ISO 32000-2 for a definition of PDF date string.
Any other singular data type	string	This includes duration, etc.
XML elements with maximum cardinality of one	dictionary	XML elements that are specified in XJDF with a maximum cardinality of 1 shall be encoded as a metadata key whose value is a dictionary. The name of the metadata key shall be the local name of the element with a CIP4 second class name prefix. Any PDF dictionary that represents an XML element may contain an optional key with a name of Type and a value of the local name of the element with a

		CIP4 prefix. NOTE 3 This addition allows for identification of the dictionaries when they are encoded as indirect objects.
XML elements with a maximum cardinality of 2 or more	array	XML elements that are specified in XJDF with a maximum cardinality of 2 or more shall be encoded as a metadata key whose value is an array of dictionaries. The name of the metadata key shall be the local name of the element with a CIP4 second class name prefix. All other restrictions are identical to XML elements with a maximum cardinality of one.

6.3 Document part (DPart) hierarchy

Files conforming to this document shall contain a **DPartRoot** entry in the Catalog dictionary, the value of which shall be the root node of a hierarchy of **DPart** dictionaries (a document part hierarchy).

The hierarchy of **DPart** dictionaries and the **DPart** entries in page objects shall conform to 14.12 of ISO 32000-2:2017.

NOTE 1 The reference to 14.12 of ISO 32000-2 is included solely for the purpose of defining the document part hierarchy; there is no requirement that a file that complies with the ISO 21812 series need be a compliant 32000-2 file in other respects. See Clause 5 Conformance.

The root node of the **DPart** hierarchy shall contain a **DPM** key, and other **DPart** dictionaries may contain a **DPM** key.

NOTE 2 A **DPM** key in the root is necessary to carry the metadata required by Clause 5 Conformance.

If metadata in conformance with this document is to be associated with a node of the **DPart** hierarchy then the **DPart** shall reference a **DPM** dictionary that shall reference the **CIP4_Root** dictionary that contains the metadata.

6.4 Defining metadata within a DPart

Metadata properties defined for a given **DPart** shall be considered to apply to all **DParts** that are child nodes of that **DPart**. Metadata properties shall not be specified in **DParts** that are in the scope of parent **DParts** which already specify the same metadata properties. In accordance with ISO 16612-2 and ISO 32000-2, each **DPart** node may have at most one **DPM** containing a dictionary of one or more metadata properties from the common metadata hierarchy specified within it.

6.5 Registered second class name prefixes

Table 2 defines the list of registered second class name prefixes.

Table 2 — Registered Second Class Name Prefixes

Prefix	Namespace URI	Organization
GTS	https://www.npes.org/pdfx/ns/id/	NPES and ISO
CIP4	https://www.CIP4.org/PDFMetaData 2 0	CIP4

7 CIP4 Common metadata hierarchy

7.1 Background

The CIP4 Common metadata hierarchy is designed to associate metadata to individual pages or ranges of pages. Standard metadata definitions are provided by this document for use in describing:

- finished printed products or pages of printed products;
- summary information to aid in optimizing the production process;

ISO 21812-1:2019(E)

— recipient information for variable data jobs.

7.2 CIP4_Root hierarchy

At least one DPM dictionary of a conforming file shall have a CIP4_Root key whose value is a reference to a CIP4_Root dictionary.

The root dictionary of CIP4 metadata trees is CIP4_Root. Some types of metadata are restricted in scope to specific DParts. These restrictions are called out in the column labelled "Scope" in Table 3. Some metadata types can only occur at certain levels within the DPart hierarchy. These restrictions are called out in the column labelled scope. The following levels are defined.

- any: The metadata may occur at any level in the DPart hierarchy.
- root: The metadata shall occur only in the document root in the DPart hierarchy. The root DPart is defined as the DPart that is referenced from DPartRoot.
- record: The metadata shall occur only in the recipient level in the DPart hierarchy.

Table 3 — CIP4_Root

Name	Data type	Scope	Description
Type	name	any	(Required) The value of Type shall be CIP4_Root.
CIP4_DescriptiveName	string	any	(Optional) Human readable description of the DPart.
CIP4_ExternalID	name	any	(Optional) External identifier of the DPart.
CIP4_Intent	dictionary	any	(Optional) CIP4_Intent specifies the creator's view of a product or document.
CIP4_IntentSummary	dictionary	any	(Optional) CIP4_IntentSummary shall specify intent properties of a DPart that are in use within the scope of the DPart. If present, all references to specific intents from CIP4_Root/CIP4_Intent shall be indirect references to a specific intent that is referenced from CIP4_IntentSummary.
CIP4_Metadata	dictionary	root	(Required) The CIP4_Metadata dictionary contains metadata properties that provide information regarding the PDF document as a whole.
CIP4_Production	dictionary	any	(Optional) The CIP4_Production dictionary contains metadata properties that may be used to parameterize a job ticket or provide additional production information that is not available in CIP4_Root/CIP4_Intent.
CIP4_Recipient	dictionary	record	(Optional) The CIP4_Recipient dictionary contains metadata properties with information regarding the intended recipient of the pages. CIP4_Recipient shall not be specified in DPart levels other than those selected by the value of RecordLevel in DPartRoot.

7.3 CIP4_Metadata level

Table 4 defines the CIP4_Root/CIP4_Metadata level that shall contain metadata properties that provide information regarding the PDF document as a whole. The CIP4_Root/CIP4_Metadata shall not be defined in any DPart node other than the root DPart node.

Table 4 — CIP4_Metadata

Name	Data type	Description
Type	name	(Required) The value of Type shall be CIP4_Metadata.

CIP4_Accounting	dictionary	(Optional) CIP4_Accounting identifies the CIP4_Contact information of where to send the invoice for the production of the PDF data.
CIP4_Administrator	dictionary	(Optional) CIP4_Administrator identifies the CIP4_Contact information regarding the execution of the PDF data.
CIP4_Author	dictionary	(Optional) CIP4_Author identifies the CIP4_Contact information for the author of the PDF data.
CIP4_Conformance	array	(Required) CIP4_Conformance is an array of string that indicates the conformance to which the metadata in the PDF data adheres. A value of <i>CIP4_IntentBase_2.0</i> shall be used if no other more restrictive value applies. NOTE 1 The value of <i>CIP4_IntentBase_2.0</i> was chosen to indicate that current intent is a mapping from XJDF. Each ICS that restricts the use of metadata properties defined in this ICS should include a required value for this metadata property that uniquely identifies that ICS. That required value shall adhere to the requirements for XML name token.
CIP4_Creator	string	(Required) CIP4_Creator identifies the conforming writer of the metadata.
CIP4_JobID	name	(Optional) CIP4_JobID identifies the job or contract to which the PDF data as a whole belongs in the context of the originating system.
CIP4_ModificationDate	date string	(Optional) CIP4_ModificationDate identifies the date at which the PDF data was last modified or created. A conforming writer shall update CIP4_ModificationDate to the current date and time whenever the PDF file is modified. NOTE 2 CIP4_ModificationDate allows detection of the modifications to PDF data by a non-conforming writer. The PDF specification already encodes a last modification date but this modification date by itself is not necessarily sufficient to detect modifications relating to the metadata by a non-conforming writer.
CIP4_ProjectID	name	(Optional) CIP4_ProjectID identifies the project or group of jobs that the PDF data as a whole belongs to in the context of the originating system.
CIP4_Sender	dictionary	(Optional) CIP4_Sender identifies the CIP4_Contact information for the sender or originator of the PDF data.

7.4 Recipient level

Table 5 CIP4_Recipient contains metadata properties with information regarding the intended recipient of the pages.

Table 5 — CIP4_Recipient

Name	Data type	Description
CIP4_ExternalID	name	(Optional) The value of the CIP4_ExternalID property shall uniquely identify the recipient within this PDF document.
CIP4_Contact	dictionary	(Optional) The value of the CIP4_Contact property shall provide contact information about the recipient.

7.5 Intent level

7.5.1 Background

In XJDF from which this document has been derived, Product Intent specifies the creator's view of a product or document. Providing intent level information within the **CIP4_Root/CIP4_Intent** hierarchy of a PDF allows a PDF creator to specify additional properties that define how the respective pages that are referenced by the **DPart** shall be used in the context of a finished printed product.

7.5.2 Intent referencing

Each dictionary in Table 6 CIP4_Intent, should be an indirect reference to a dictionary that is referenced from CIP4_Root/CIP4_IntentSummary under the same name.

Table 6 — CIP4_Intent

Name	Data type	Description
Type	name	(Required) The value of Type shall be <i>CIP4_Intent</i> .
CIP4_ProductType	name	<p>(Optional, Extendable) The value of CIP4_ProductType property shall indicate the general product class that the DPart represents. The name should be one of the following:</p> <ul style="list-style-type: none"> — <i>BackCover</i>: The last page or sheet of a soft-cover book or magazine, commonly a heavier media. — <i>Body</i>: Generic content inside of a <i>Cover</i>. — <i>Book</i>: <i>Body</i> with a <i>Cover</i> and a <i>Spine</i>. — <i>BookBlock</i>: The assembled body of pages for a hard-cover book. — <i>BookCase</i>: The assembled covers and spine component of a hard-cover book, prior to "casing in" (attaching to the book block) — <i>Booklet</i>: <i>Body</i> with a <i>Cover</i> without a <i>Spine</i> (typically stapled). — <i>Box</i>: Convenience packaging that is not envisioned to be protection for shipping. — <i>Brochure</i>: A single folded sheet. — <i>BusinessCard</i>: A small card that displays contact information for an individual employed by a company. — <i>Cover</i>: A single sheet covering a side of a print product. — <i>CoverLetter</i>: A letter accompanying another print product. — <i>Envelope</i>: A folded paper container, with sealable flap, that encloses and protects a document or contents. — <i>FrontCover</i>: The first page or sheet of a soft-cover book or magazine, commonly a heavier media. — <i>Insert</i>: A product part intended to be inserted into a print product — <i>Jacket</i>: Hard cover case jacket — <i>Label</i>: A piece of paper or plastic that is attached to an object in