INTERNATIONAL STANDARD

ISO 24627-3

First edition
2021-02

Language resource management — Comprehensive Annotation Framework (ComAF) —

Part 3: Diagrammatic semantic authoring (DSA)

Gestion des ressources linguistiques — Cadre global d'annotation (ComAF) —

Partie 3: Création sémantique diagrammatique (DSA)

This document is a preview generated by EVS

Reference number
ISO 24627-3:2021(E)

© ISO 2021

This document is a preview generated by EVS



COPYRIGHT PROTECTED DOCUMENT

© ISO 2021

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword ..... **iv**
Introduction ..... **v**
1 Scope **1**
2 Normative references **1**
3 Terms and definitions **1**
4 Specification **1**
Annex A (informative) **Examples** **3**
Bibliography **8**

This document is a preview generated by EVS

© ISO 2021 – All rights reserved

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 37, *Language and terminology*, Subcommittee SC 4, *Language resource management*.

A list of all parts in the ISO 24627 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.

This document is a preview generated by EVS

 © ISO 2021 – All rights reserved

Introduction

Graphs (diagrams consisting of nodes and links) have been used for decades to represent and visualize both documents (instance data) and data schemas. This document concerns graph-based representation (not visualization) of documents (not data schemas).

Graph-based representation and visualization of documents are addressed by concept maps,[15] mind maps, argument maps, and so on. Theoretical linguistics and artificial intelligence have also used graph-based content visualization associated with semantic network, mental space,[10] discourse representation structure,[13] and so forth.

Graph-based visualization of data schemas (or ontologies, terminologies, metamodels, etc.) is a more usual practice. Ontologies are often visualized as graphs in which nodes are classes (and datatypes) and links are properties (relations). ISO 24156-1 specifies a UML-based visualization of concept modelling. Other metamodels are usually represented as similar diagrams, too.

This document gives a data schema of graph documents to facilitate composition and comprehension by making logical document structure explicit. It neither covers visualizations or manipulations of graphs nor does it define annotations to existing documents, but rather it addresses graphical/diagrammatic representation of documents for the sake of semantic authoring: i.e., for people to directly view and manipulate syntactic/semantic structures on computer displays or their future alternatives. The linearity of traditional text documents is due to the linearity of speech languages, which constrains the interaction between people and documents, making it hard for people to read and write. DSA defines graphical/diagrammatic documents with more explicit structures than in text in order to make it easier for people to read and write. Documents based on DSA, together with some user interfaces involving appropriate visualizations and easy operations, can enhance collaborations among people and between people and machines.

DSA mainly deals with syntactic or document structures. It addresses some fragmentary semantic structures as well, but more systematic semantics (formal mapping between documents and their meanings or logical forms) can be provided by another specification so that machines better 'understand' DSA-based documents and thereby better assist information sharing and consensus building among people.

Figure 1 shows a workflow involving DSA and other types of documents. The DSA-based documents in the upper half can be automatically converted (while preserving propositional content) to and from machine-understandable documents based on appropriate standards on semantic representations and annotations. It is possible to automatically generate traditional text documents from these machine-understandable documents (while preserving the propositional content, too), though the inverse conversion cannot generally be automated. Since DSA-based documents (together with some appropriate user interfaces) are easier for people to compose and interpret than text documents, people can usually touch and see DSA-based documents whereas traditional documents could be used for legacy procedures (such as patent applications) and oral presentations.

This document is a preview downloaded by EVS

© ISO 2021 – All rights reserved

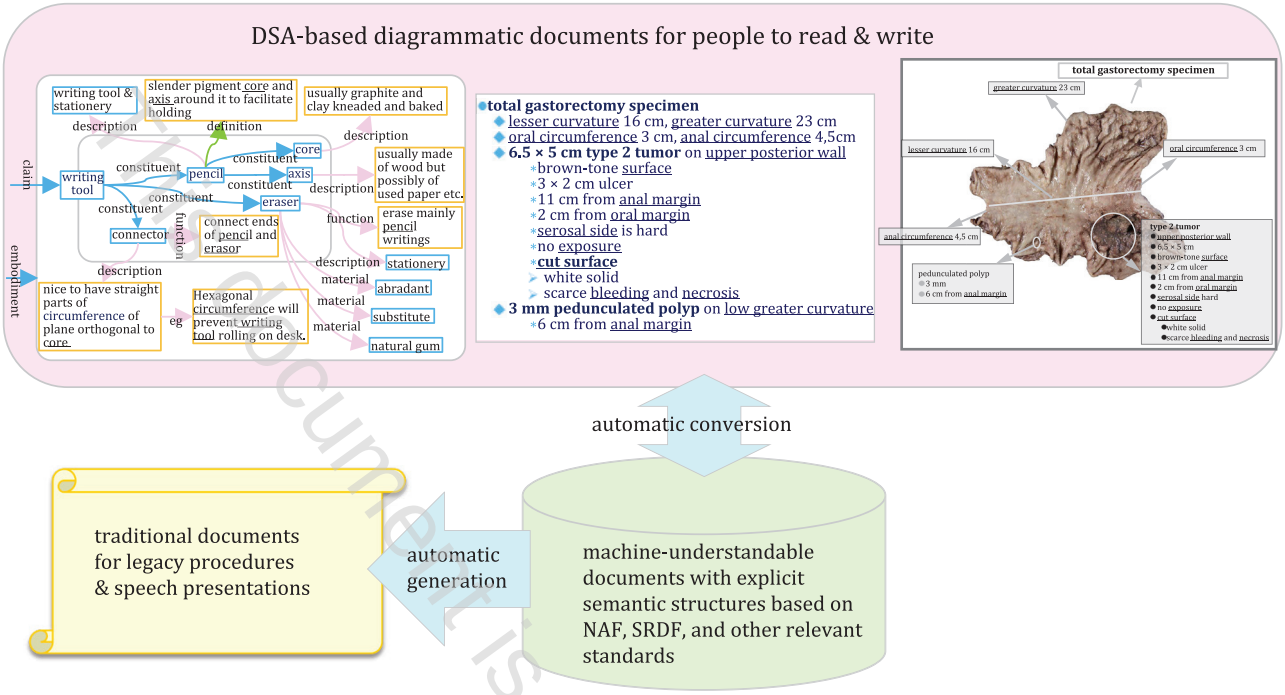


Figure 1 — Document workflow involving DSA

DSA is a minimal metamodel for ISO TS 24617-5 (SemAF-DS), which in turn is based on ISO/IEC 15938.5/Amd.1 (MPEG-7 MDS AMD1 — Linguistic description scheme). The machine-understandable documents in Figure 1 are assumed to use other standards including ISO 24615 (SynAF), ISO 24612 (LAF) and ISO 24617 (SemAF) while also incorporating insights from other relevant literature[1][8][9][10][11][12][13][14][15][16][17][18].

This document is a preview generated by EVS

© ISO 2021 – All rights reserved

Language resource management — Comprehensive Annotation Framework (ComAF) —

Part 3: Diagrammatic semantic authoring (DSA)

This document is a preview generated by EVS

1 Scope

This document specifies how to represent (not visualize) documents (instance data, not data schemas) as graphs. It does not specify how to visualize or operate on document data, but it aims at making documents easier for people to compose and comprehend by allowing for various graph-based flexible user interfaces, possibly incorporating document-visualization practices (see Introduction). In this connection, this document does not specify annotations to existing documents either, but rather it specifies a schema of documents with explicit logical structures.

2 Normative references

There are no normative references in this document.

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 http://www.electropedia.org/

3.1
hypernode
node which is a graph segment

3.2
segment
referenceable part of a DSA-based document, which is either a graph segment or a data segment (text, image, audio, video, etc.)

3.3
semantic authoring
composition of documents while making their logical structures explicit

4 Specification

Figure 2 illustrates DSA, which is a metamodel (ontology) of graph documents. Each markable (referenceable) part of a DSA-based document is called a segment. Each DSA-based document itself is a segment, too.

© ISO 2021 – All rights reserved