
**Language resource management —
Semantic annotation framework —**

**Part 6:
Principles of semantic annotation
(SemAF Principles)**

*Gestion des ressources linguistiques — Cadre d'annotation
sémantique —*

Partie 6: Principes d'annotation sémantique (SemAF Principles)



This document is a preview generated by EBS



COPYRIGHT PROTECTED DOCUMENT

© ISO 2016, Published in Switzerland

All rights reserved. Unless otherwise specified, 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
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword	iv
1 Scope	1
2 Terms and definitions	1
3 Purpose and motivation	2
3.1 Purpose.....	2
3.2 Motivation.....	2
4 Overview	3
5 Annotation principles and requirements	4
5.1 Principles inherited from the Linguistic Annotation Framework.....	4
5.2 Other general annotation principles.....	5
5.3 Principles specific to semantic annotation.....	5
6 The methodological basis of SemAF	7
6.1 Steps in the design of an annotation scheme.....	7
6.2 Metamodels.....	8
6.3 Abstract syntax, concrete syntax and semantics.....	10
6.4 Steps forward and feedback in the design process.....	12
6.5 Optional elements in an annotation scheme.....	14
7 Overlaps between annotation schemes	15
7.1 Semantic and terminological consistency.....	15
7.2 Spatial and temporal relations as semantic roles.....	15
7.3 Events.....	17
7.4 Discourse relations in dialogue.....	18
8 Semantic phenomena that cut across annotation schemes	18
8.1 Ubiquitous semantic phenomena.....	18
8.2 Quantification.....	18
8.3 Quantities and measures.....	19
8.4 Negation, modality, factuality, and attribution.....	20
8.5 Modification and qualification.....	21
8.5.1 Modification and quantification.....	21
8.5.2 Qualification.....	22
8.5.3 Other issues.....	23
Annex A (informative) An approach to the annotation of quantification in natural language	24
Bibliography	28

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 on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: [Foreword - Supplementary information](#).

The committee responsible for this document is ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 4, *Language resource management*.

ISO 24617 consists of the following parts, under the general title *Language resource management — Semantic annotation framework (SemAF)*:

- *Part 1: Time and events (SemAF-Time, ISOTimeML)*
- *Part 2: Dialogue acts (SemAF-Dacts)*
- *Part 4: Semantic roles (SemAF-SR)*
- *Part 5: Discourse structures (SemAF-DS)* [Technical Specification]
- *Part 6: Principles of semantic annotation (SemAF Principles)*
- *Part 7: Spatial information (ISOspace)*

The following parts are in preparation:

- *Part 8: Semantic relations in discourse (SemAF DR-core)*
- *Part 9: Reference (ISOref)*

Language resource management — Semantic annotation framework —

Part 6: Principles of semantic annotation (SemAF Principles)

1 Scope

This part of ISO 24617 specifies the approach to semantic annotation characterizing the ISO Semantic annotation framework (SemAF). It outlines the SemAF strategy for developing separate annotation schemes for certain classes of semantic phenomena, aiming in the long term to combine these into a single, coherent scheme for semantic annotation with wide coverage. In particular, it sets out the notions of both an abstract and a concrete syntax for semantic annotations, mirroring the distinction between annotations and representations that is made in the ISO Linguistic Annotation Framework. It describes the role of these notions in relation to the specification of a metamodel and a semantic interpretation of annotations, with a view to defining a well-founded annotation scheme.

This part of ISO 24617 also provides guidelines for dealing with two issues regarding the annotation schemes defined in SemAF-parts: a) conceptual and terminological inconsistencies that may arise due to overlaps between annotation schemes and b) the treatment of semantic phenomena that cut across SemAF-parts, such as negation, modality and quantification. Instances of both issues are identified, and in some cases, direction is given as to how they may be tackled.

2 Terms and definitions

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

NOTE In addition, the terms ‘event’ and ‘eventuality’ are used (as synonyms) as defined in ISO 24617-1 as *something that can be said to obtain or hold true, to happen or to occur*.

2.1

primary data

electronic representation of text or communicative behaviour

EXAMPLE Digital representations of text, transcriptions of speech, gestures or multimodal dialogue.

Note 1 to entry: ISO 24612 defines primary data as the ‘electronic representation of language data’. This definition is unsatisfactory for this part of ISO 24617 as semantic annotation may relate to non-verbal or multimodal data, such as stretches of spoken dialogue with accompanying gestures and facial expressions, and even gestures and/or facial expressions without any accompanying speech.

2.2

annotation

linguistic information added to *primary data* (2.1), independent of its representation

[SOURCE: ISO 24612:2012, 2.3]

2.3

semantic annotation

annotation (2.2) which contains information about the meaning of a segment or region of *primary data* (2.1)