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

**Part 9:
Reference annotation framework
(RAF)**

iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

[ISO 24617-9:2019](https://standards.iteh.ai/catalog/standards/iso/37ca9521-1d12-4bcd-810e-122f826b4a65/iso-24617-9-2019)

<https://standards.iteh.ai/catalog/standards/iso/37ca9521-1d12-4bcd-810e-122f826b4a65/iso-24617-9-2019>



iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

[ISO 24617-9:2019](https://standards.iteh.ai/catalog/standards/iso/37ca9521-1d12-4bcd-810e-122f826b4a65/iso-24617-9-2019)

<https://standards.iteh.ai/catalog/standards/iso/37ca9521-1d12-4bcd-810e-122f826b4a65/iso-24617-9-2019>



COPYRIGHT PROTECTED DOCUMENT

© ISO 2019

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
Fax: +41 22 749 09 47
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 Basic principles	2
5 Meta-model for reference annotation	3
5.1 Overview	3
5.2 Referring expressions	3
5.3 Data categories for referring expressions	4
5.4 Lexical relations	5
5.5 Discourse entities	5
5.6 Objectal relations	5
5.7 Metadata	5
6 Abstract syntax, concrete syntax, and semantics of annotations	6
6.1 Introduction	6
6.2 Abstract syntax	6
6.2.1 Conceptual inventory	6
6.2.2 Annotation structures: Entity structures and link structures	7
6.3 Semantics	8
6.3.1 Discourse entity structures and objectal relation links	8
6.3.2 Referential expression entity structures and lexical relation links	9
6.4 Implementing an XML serialisation compliant with the TEI P5 guidelines	10
6.4.1 Introduction	10
6.4.2 Namespace	10
6.4.3 Generic principles attached to a TEI compliant serialisation	10
6.4.4 Feature structures	11
6.4.5 General document architecture	12
6.5 Implementation of the <i>Referring expression</i> component	12
6.6 Implementation of the <i>Discourse entity</i> component	13
6.7 Implementation of referential relations	13
6.8 Objectal relations: grouping	14
6.9 Alternative linking: ambiguity	15
6.10 Multiple links	15
6.11 Representing referential chains	16
6.12 Bridging phenomena	16
Annex A (normative) Data categories for reference annotation	18
Annex B (informative) Complementary examples or partial examples referred to in the main text of the document	25
Bibliography	26

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. 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. 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, *Terminology and other language and content resources*, Subcommittee SC 4, *Language resource management*.

A list of all parts in the ISO 24617 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

This document is intended to complement the ISO 24617 series and to provide all the necessary conceptual and technical mechanisms for the annotation of referential phenomena in multimodal discourse. Reference phenomena are an essential component for the understanding and structuring of discursive mechanisms, ranging from very basic pronominal relation to complex bridging anaphora. Annotating such phenomena in an interoperable way improves the re-usability of language resources in such applications in language technology as named entity recognition, text understanding and synthesis, text summarization, information retrieval, automatic question-answering, man-machine dialogue, and machine translation.

The content of this document builds upon various projects and software platforms that have been dealing with reference annotation (RA), in particular the following References [9],[2],[16],[21],[26],[25],[22],[5],[15],[13] but also the TEI P5 guidelines. Based on these and other previous works, the Referential Annotation Framework (RAF) aims at providing a synthesized way of treating various reference phenomena in discourse. In continuity with most practices in the field, RAF focuses on marking up referring expressions in a discourse and the relations that hold between them and the corresponding entities, whether this is based upon employing crowd sourcing or machine learning strategies.

iTeh Standards (<https://standards.iteh.ai>) Document Preview

[ISO 24617-9:2019](https://standards.iteh.ai/catalog/standards/iso/37ca9521-1d12-4bcd-810e-122f826b4a65/iso-24617-9-2019)

<https://standards.iteh.ai/catalog/standards/iso/37ca9521-1d12-4bcd-810e-122f826b4a65/iso-24617-9-2019>