

ICS 01.020
CCS A 22



**NATIONAL STANDARD OF THE
PEOPLE'S REPUBLIC OF CHINA**

GB/T 44217.6-2024

**Language resource management - Semantic annotation
framework - Part 6: Principles of semantic annotation**

语言资源管理 语义标注框架 第6部分：语义标注原则

(English Translation)

Issued on: July 24, 2024

Implemented on: February 1, 2025

**Issued by: State Administration for Market Regulation, China National
Standardization Administration**

Table of Contents

Preface	
Introduction	
1 Scope	
2 Normative references	
3 Terms and Definitions	
4 Purpose and function	
4.1 Purpose	
4.2 Function	
5 Brief description	
6 Marking Principles	
6.1 Principles inherited from language annotation frameworks	
6.2 Other general marking principles	
6.3 Specific principles for semantic annotation	
7 Methodological foundations of SemAF	
7.1 Steps of designing a labeling scheme	
7.2 Metamodel	
7.3 Abstract Syntax, Concrete Syntax, and Semantics	
7.4 Steps and feedback in the design process	
7.5 Optional elements in annotation schemes	
8 Overlap between annotation schemes	
8.1 Semantic and terminological consistency	
8.2 Spatial and temporal relations as semantic roles	
8.3 Event	
8.4 Discourse relations in dialogue	
9 Semantic phenomena across multiple annotation frameworks	
9.1 Ubiquitous semantic phenomena	
9.2 Quantification	
9.3 Quantity and volume	
9.4 Negation, modality, facticity, and attributes	
9.5 Modification and Quantification	
Appendix A (Informative) Natural Language Quantitative Annotation Method	
Reference	

Foreword

This document is in accordance with the provisions of GB/T 1.1-2020 "Guidelines for standardization work Part 1: Structure and drafting rules for standardization documents" Drafting.

This document is part 6 of GB/T 44217 "Semantic Annotation Framework for Language Resource Management".

Lower part.

--- Part 6: Principles of semantic annotation;

--- Part 11: Measurable Quantity Information (MQI).

This document is modified to adopt ISO 24617-6:2016 "Semantic annotation framework for language resource management Part 6: Principles of semantic annotation" (SemAF Principles)".

This document adds a chapter on "Normative References".

The technical differences between this document and ISO 24617-6:2016 and the reasons are as follows.

--- Changed the explanation of why semantic annotations in digital corpora often support interpretation and inference to a declarative description (See 6.3), the standard does not need to explain the reasons;

--- Deleted the explanation of why feedback relations and dependency relations did not appear in the discourse relation annotation scheme, and deleted the discourse relation the ISO notation scheme should inherit these relationships from ISO 24617-2 (see 7.4 of ISO 24617-6:2016).

No need to explain the reasons;

--- Changed the marking method of values and quantities (see 9.3). Due to technological development, the marking method of values and quantities is adopted.

ISO 24617-11:2022 specifies the notation of values and quantities.

The following editorial changes were made to this document.

--- Changed the standard name and deleted the abbreviation of semantic annotation framework;

--- Changed the expression form of the three ways to achieve the purpose of this document (see 4.1) and changed it to a list according to the requirements of GB/T 1.1-2020 express;

--- Changed the expression form of the examples (see 4.2) to examples that conform to the Chinese expression form;

- Changed the contents of Example 1, Example 2 and Example 3 (see 4.2) to examples that conform to the Chinese expression;
- Changed the expression form of consistency between annotation schemes and completeness of annotation scheme sets (see Chapter 5). According to GB/T 1.1-2020 requirements, changed to itemized representation;
- Changed the expression form of the two functions of semantic annotation (see 6.3) and changed to item expression according to the requirements of GB/T 1.1-2020;
- Changed the expression form of the examples (see 6.3) to examples that conform to the Chinese expression form;
- Changed the interpretation of meaningful marking (see 6.3), the standard does not need to explain the reasons, instead of declarative description;
- Changed the content of Example 2 (see 6.3) to an example that conforms to the Chinese expression;
- Changed the contents of Example 4, Example 5 and Example 6 (see 7.2) to examples that conform to the Chinese expression;
- Changed the expression form of the examples (see 7.2) to examples that conform to the Chinese expression form;
- Change the expression of references in the text (see 7.2) to comply with the requirements of GB/T 1.1-2020;
- Changed the expression form of the steps of converting a given representation into another semantically equivalent representation (see 7.3) to a list of items;
- Change the expression of references in the text (see 7.4) to comply with the requirements of GB/T 1.1-2020;
- Changed the description of distance in Example 4 (see 8.2) to the unit of distance measurement that conforms to Chinese habits;
- Changed the two types of expressions of ISO dialogue behavior notation scheme (see 8.4) to list items;
- Changed Example 1 and related content (see 8.8);
- Changed the expression form of the examples (see 9.2) to examples that conform to the Chinese expression form;
- Changed the example content and expression form (see 9.4) to examples that conform to the Chinese expression method;
- Deleted the original example (24) and the corresponding description in the text (see 9.5.1, 8.5.1 of ISO 24617-6:2016).

Conform to Chinese customs;

--- Changed the example content and expression form (see 9.5.2) to conform to the Chinese expression example.

Introduction

Semantic annotation is one of the important technologies for deep processing of natural language by computers. It is to add semantic annotation to words or sentences in text.

The process of semantic labeling is to solve the problem. According to the general principles of annotation and the specific annotation objects, the standard is divided into different parts.

GB/T 44217 "Semantic Annotation Framework for Language Resource Management" is planned to consist of 12 parts.

--- Part 1: Time and events. The purpose is to provide a general method for describing time and events in text.

--- Part 2: Dialogue acts. The purpose is to provide a notation language for representing dialogue acts and a method for segmenting dialogues into the method of defining units.

--- Part 4: Semantic roles. The purpose is to provide a consensus annotation scheme for semantic roles.

--- Part 5: Discourse structure. The purpose is to provide a way to represent discourse realization and discourse content.

--- Part 6: Principles of semantic annotation. The purpose is to define semantic annotation methods characterized by a semantic annotation framework.

--- Part 7: Spatial information. The purpose is to provide a general method to describe the spatial information related to the expression of motion in natural language text.

Spatial information and spatiotemporal information.

--- Part 8: Semantic relations in discourse, core annotation framework. The purpose is to provide a framework for the representation and annotation of discourse relations.

plan.

--- Part 9: Reference annotation framework. The purpose is to provide a framework for the annotation and representation of references in natural language text and multimodal interactions.

Provide a comprehensive model.

--- Part 11: Measurable Quantity Information (MQI). The purpose is to provide a notation scheme for measurable quantity information.