

ICS 35.240.30

CCS L 60



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

GB/T 45949-2025

**Specification for semantic structuring annotation of Chinese
news**

中文新闻语义结构化标注

Issued on: June 30, 2025

Implemented on: June 30, 2025

**Issued by: State Administration for Market Regulation;
Standardization Administration of the PRC**

Table of Contents

Preface	III
Introduction	IV
1 Scope	1
2 Normative references	1
3 Terms and Definitions	1
4 Overview of News Semantic Structured Annotation	2
4.1 Labeling Process	2
4.2 Annotation Paradigm	3
5 Preprocessing	4
5.1 Setting the annotation object identifier	4
5.2 News Release Metadata Collection	4
6 Keyword annotation	5
6.1 Marking content	5
6.2 Marking Template	5
6.3 Marking method	5
7 Semantic Template Annotation	6
7.1 Marking content	6
7.2 Marking Templates	14
7.3 Marking method	15
8 Semantic Unit Annotation	16
8.1 Marking content	16
8.2 Marking Templates	17
8.3 Marking method	18
Appendix A (Informative) Application Scenario Examples	23
A.1 Structural management of news content	23
A.2 Generating a Knowledge Base of Manuscript Content	23
A.3 Building a high-quality knowledge graph	23
A.4 Low-cost and high-efficiency training of artificial neural network models	23
Appendix B (Informative) Complete Example of News Semantic Structured Annotation ...	24
B.1 Original	24
B.2 Preprocessing	24
B.3 Keyword Annotation	26

B.4 Semantic Template Annotation	27
B.5 Semantic Unit Annotation	29
Reference	32

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.

Please note that some of the contents of this document may involve patents. The issuing organization of this document does not assume the responsibility for identifying patents.

This document was proposed and coordinated by the National Technical Committee for Standardization of Chinese News Information (SAC/TC352).

This document was drafted by: Wenling Technology (Beijing) Co., Ltd., Beijing Language and Culture University, Southeast University, Xinhua News Agency, China Internet News News Center, China Academy of Information and Communications Technology, Zhejiang Daily Newspaper Group, and Baoding Data Bureau (National Data Labeling Base).

The main drafters of this document are: Wang Nan, Rao Gaoqi, Yang Peng, Ding Desheng, Wang Yi, Fu Rong, Wang Guirong, Wang Jianping, Huang Jing, Li Sun, Jia Chengxi, Zhang Chuang, Li Bing, Yang Dongmei, Zhou Tianwai, and Song Yongsheng.

Introduction

With the widespread application of artificial intelligence technology in the news field, establishing standardized Chinese news data processing standards has become an important part of industry development.

This document is formulated to build a unified Chinese news semantic annotation system, through a standardized data processing process to achieve The following two major goals. one is to provide support for the systematic construction of news knowledge base and knowledge graph; the other is to improve the training of artificial neural network models Improve data quality and efficiency, and promote the intelligent transformation of the news industry.

Semantic structured annotation of Chinese news

1 Scope

This document specifies the preprocessing, keyword annotation, semantic template annotation and semantic unit annotation of Chinese news semantic structured annotation.

method.

This document is suitable for news content providers, media applications and research institutions to manually annotate news articles, and can also be used for machine automation. Annotation.

This document is not intended for semantically structured annotation across manuscripts.

2 Normative references

GB/T 20092.2-2022

GB/T 20093-2022

3 Terms and Definitions

The following terms and definitions apply to this document.

3.1 keyword

Words that represent specific semantics in news articles.

Note. "Specific semantics" refers to the content of the news article that the annotator considers important, including the article's theme, important people, important events, and important common sense.

3.2 semantic unit

The basic unit that expresses complete semantics in news releases.

Note. Semantic units are mainly divided into entities, things, events and other types, and are the basis for complex semantic structures including knowledge graphs.

3.3

Describes or defines the characteristic quantity of a semantic unit.

NOTE. Semantic unit parameters include attributes, properties, aspects, capabilities (purposes), categories, limitations and states.

3.4 stateofsemanticunit

The set of values of semantic unit parameters in a specific time and space.

Note. The state of a semantic unit is used to describe the dynamic nature of the development and change of things. When marking, it is marked as a parameter of the semantic unit.

3.5 semantictemplate

A fixed structure that expresses semantics.

Note. It consists of core elements, limiting elements and their relationships.

chinastandards.org · PREVIEW