Compilation Methodology and Postprint Implementation of the Xinhua News Agency News Information Resources Catalog

Authors: Chai Yanyan

Date: 2023-10-08T00:00:00+00:00

Abstract

To effectively enhance the level of integration and sharing of national government information resources, there is an urgent need to systematically organize government information resources—including news information resources—and compile information resource catalogs. This paper, combining the actual situation of Xinhua News Agency and starting from its news information system, proposes a survey methodology for news information resources, delineates six major categories of Xinhua News Agency' s news information resources, establishes encoding rules for these resources, completes the first edition of the news information resource catalog, forms a "master ledger" of Xinhua News Agency' s news information resources, and establishes a foundation for information resource sharing and data openness.

Full Text

Preamble

Title: Method and Implementation of News Information Resource Cataloging at Xinhua News Agency
Author: Xinhua News Agency, Beijing 100803

Abstract: To effectively improve the integration and sharing of national government information resources, there is an urgent need to systematically organize government information resources, including news information resources, and compile information resource catalogs. This paper examines Xinhua News Agency' s practical experience, beginning with its news information systems. It proposes an investigation methodology for news information resources, establishes a six-category classification system for Xinhua' s news information resources, defines encoding rules, and completes the first edition of the news information resource catalog. This work creates a comprehensive "master ledger"

of Xinhua' s news information resources, laying the foundation for information resource sharing and data openness.

Keywords: news information; government affairs; information resource encoding; catalog compilation
Classification Code: G252.7
Document Code: A
Article ID: 1671-0134(2021)02-064-03
DOI: 10.19483/j.cnki.11-4653/n.2021.02.017
Citation Format: Chai Yanyan. Method and Implementation of News Information Resource Cataloging at Xinhua News Agency [J]. China Media Technology, 2021(02):64-65+99.

Government information resources refer to various information resources, including documents, materials, charts, and data, that are produced or obtained by government departments in the course of performing their duties and recorded and preserved in certain forms [1]. As the national news agency, Xinhua News Agency produces and distributes massive amounts of text, images, charts, audio, video, and multimedia content, representing a vital component of national government information resources. Government information resource catalogs are collections that describe the characteristics of these resources, organized according to specific classification and encoding methods. Catalog compilation constitutes a critical link in government information resource sharing and openness. As the national news agency, Xinhua actively participates in national government information resource sharing and openness, making the compilation of news information resource catalogs an urgent priority.

1. Current State of Government Information Resource Catalog Development

The standardized development, utilization, and sharing of government information resources represent key objectives of China' s government informatization construction. In September 2016, the State Council promulgated the "Interim Measures for the Management of Government Information Resource Sharing," the first guiding document on government information resource sharing since the founding of the People' s Republic of China. In June 2017, the National Development and Reform Commission and the Central Cyberspace Affairs Commission formulated the "Guidelines for Government Information Resource Catalog Compilation (Trial)" (hereinafter referred to as the "Guidelines") [2], which clarified catalog compilation standards and methods to guide government departments in compiling information resource catalogs and improve the integration and sharing of government information resources nationwide. Various national ministries and local government departments have conducted effective explorations into the theoretical and practical aspects of building government information resource catalog systems.

Xinhua News Agency' s news business systems currently classify and standardize

news information according to national standards "Classification and Codes for Chinese News Information" and "Chinese News Markup Language," providing a foundation for catalog compilation. However, these systems have not been standardized according to the "Guidelines," making it difficult for other government departments' information systems to recognize Xinhua' s news information resources and resulting in challenges such as sharing difficulties and insufficient development and utilization.

2.1 Information Resource Investigation and Organization

There are two primary methods for investigating and organizing government information resource catalogs. The first involves understanding an organization' s business operations and aggregating business-related information. The second involves investigating and organizing data resources from existing government information system function menus and databases to form information resource inventories [3]. Since information systems do not correspond one-to-one with business processes, the first method yields less practical inventories that present greater difficulties for subsequent data sharing and exchange. Given that Xinhua' s news information systems are relatively mature, this paper adopts the second method, which facilitates easier data resource organization, sharing, and exchange.

This study conducted detailed investigations and systematic organization of ten application systems supporting Xinhua' s news operations and daily office functions, including the data service platform, all-media collection-editing-distribution system, reporting command system, news hotspot and influence analysis system, news information statistical monitoring system, and daily office systems. The investigation extracted attribute information such as system names, basic data resource conditions (including database names, table counts, total storage volume, and update frequency), and database table details (field names, descriptions, data types, lengths, etc.), resulting in a comprehensive information resource inventory.

2.2 Information Resource Classification

To enhance the sharing and utilization of government information resources, Xinhua has developed a scientific classification method following the "Guidelines" to enable classified management, rapid retrieval, and precise location of information resources. The classification is performed hierarchically according to different information systems and system user interface menu names, adopting a "what you see is what you get" approach that effectively shields users from the complex underlying database structures while remaining simple and practical.

In accordance with relevant national standards and Xinhua' s actual circumstances, Xinhua' s news information resources are divided into six categories: class, item, sub-item, detailed sub-item, fifth-level detailed sub-item, and sixth-

level detailed sub-item. The first three categories are mandatory: department-level information resource catalog, State Council, and Xinhua News Agency. The latter three are optional. The detailed sub-item represents information system names, with ten systems identified, such as the data service platform and all-media collection-editing-distribution system. The fifth-level detailed sub-item corresponds to primary menu items in the system interface, with 50 categories such as Xinhua' s finished product database and newspaper database. The sixth-level detailed sub-item corresponds to secondary menu items, with 170 categories such as Xinhua News and Xinhua Economic Information. The information resource name represents the final-level menu item in the system, comprising 702 entries such as the Wire News Line and Xinhua Viewpoint.

2.3 Information Resource Encoding

Based on the news information classification described in Section 2.2 and incorporating identifiers such as information resource provider codes, Xinhua has established an information resource encoding rule that assigns a unique identifier —an information resource code—to each resource. The information resource code consists of three components: a prefix code, a separator "/" , and a suffix code, forming a 20-character string. The prefix code is Xinhua' s institutional code, a six-character string: 303083. The suffix code is a 13-character string comprising a 2-character detailed sub-item classification code, a 2-character fifth-level detailed sub-item classification code, a 3-character sixth-level detailed sub-item classification code, and a 6-character information resource name code. An example of the Wire News Line code is shown in [Figure 1: see original paper].

2.4 Information Resource Metadata

Metadata is a set of data that describes the characteristics of information resources. Using metadata to describe information resources enables rapid discovery and confirmation of resources, provides retrieval points, and facilitates data exchange between different systems [4]. Xinhua' s news information resource directory metadata expands upon and improves the requirements specified in the "Guidelines" issued by the National Development and Reform Commission, establishing 32 core metadata elements (partial metadata shown in [Figure 2: see original paper]).

3. Xinhua' s Catalog Compilation Practice

To standardize the management, sharing, and release of government information resources and facilitate the submission of government information resource catalogs, Xinhua employs a Government Information Resource Catalog Management System (hereinafter referred to as the Catalog Management System) developed in compliance with the "Guidelines" to assist in catalog compilation and sharing. This system features functions such as catalog visualization, automatic catalog encoding, validation, information resource organization, and catalog file export,

enabling internal catalog registration, catalog review, and catalog release, as well as exporting catalog files in national standard format for publication on the unified national catalog platform.

Xinhua' s catalog compilation process primarily comprises four steps: catalog compilation, catalog verification, catalog registration, and catalog release. Catalog compilation itself includes two stages: information resource inventory aggregation and catalog creation.

3.1 Information Resource Inventory Aggregation

Based on the information resource investigation, the assignment of values for the 32 core metadata elements described above is completed to form the information resource inventory. Using the data service platform as an example, the information resource inventory is shown in [Figure 3: see original paper].

3.2 Catalog Compilation

The automatic encoding and visualization functions of the Catalog Management System can effectively improve the efficiency and accuracy of catalog compilation. First, the information resource inventory is imported into the Catalog Management System for initialization. Next, automatic information resource encoding is performed, completing validation for data legality, consistency, completeness, and duplication.

Following catalog registration, review, and release, online management and retrieval of information resource catalogs are achieved, completing the catalog compilation work. This edition of the news information resource catalog completed the compilation of ten information system catalogs, totaling 702 information resource entries with 6,571 information items. The catalog retrieval interface is shown in [Figure 4: see original paper].

The compilation of Xinhua' s news information resource catalog represents pioneering work, marking the first time that the agency's internal news information resource catalog has been compiled according to national standards. This effort has clarified the full scope of the agency' s news information resources, with some catalogs already published on the national catalog platform. Based on the unified national e-government network, real-time sharing of government information resources has been initially realized with promising results. Future catalog compilation work should strengthen the application of new technologies. For instance, blockchain technology could enable real-time data collection and analysis, ensuring the timeliness of catalog services, while improving guidance and traceability systems to guarantee that data resources are single-sourced and accurate.

Note: Figure translations are in progress. See original paper for figures.