



**IPTC Standards DRAFT**

---

# **NewsML 2 Architecture**

**Version 1.0**

**Documentation**

**Glossary**

**Document Revision 1**



Copyright © 2005 by IPTC, the International Press Telecommunications Council. All rights reserved.

IPTC address:

Postal mail: Royal Albert House, Sheet Street, Windsor, Berkshire, SL4 1BE, United Kingdom

Web: [www.iptc.org](http://www.iptc.org)

email: [office@iptc.org](mailto:office@iptc.org)

This project intends to use materials that are either in the public domain or are available by the permission for their respective copyright holders.

Permissions of copyright holder will be obtained prior to use of protected material.

All materials of this IPTC standard covered by copyright shall be licensable at no charge.

---

**Document file name** (+ Word file extension “.doc”): **DRAFT-NAR\_1.0-doc-Glossary\_1.doc**

**Document URN: urn:iptc:std-draft:NAR:1.0:doc:Glossary:1**

(More information on IPTC URNs in RFC 3937)

### Specification Versioning History

Version	Issue Date	Approved by	Remark
1	TBD	IPTC Standards Committee	Currently under review

### Document Revision History

Revision	Issue Date	Author (revised by)	Remark
1	2005-12-16	Michael Steidl	moved from NAR Tech Spec, document version 18



## Table of contents

<b>1</b>	<b>About this document</b> .....	<b>3</b>
1.1	Abstract .....	3
1.2	Status of this document .....	3
<b>2</b>	<b>The Glossary</b> .....	<b>4</b>

## 1 About this document

### 1.1 Abstract

This document provides the generic glossary of terms used by the IPTC in the scope of its “NewsML 2” family of news exchange standards which is based on the “NewsML 2 Architecture”.

The IPTC NewsML 2 Architecture is primarily intended for, but not restricted to, use in the news industry.

### 1.2 Status of this document

This document is development by the IPTC NewsML 2 Architecture Working Party (NAR WP).

Comments from IPTC members which are intended to be visible to IPTC members only should be sent to the mailing list described at:

<http://groups.yahoo.com/group/iptc-news-architecture-dev/>

Public comments should be sent to the mailing list described at:

<http://groups.yahoo.com/group/newsml-2/>

Public versions of this document and of related IPTC documents are available at:

<http://www.iptc.org/dev/>



## 2 The Glossary

Term	Definition
alias	See <a href="#">scheme alias</a> .
anonymous controlled vocabulary	A <a href="#">controlled vocabulary</a> that is not a <a href="#">scheme</a> .
catalog	A file containing information about <a href="#">schemes</a> .
code	A character sequence which forms a member of a <a href="#">controlled vocabulary</a> .
compact URI	A URI <i>represented</i> by a string of the form sss:ccc, where sss is a <a href="#">scheme alias</a> and ccc is a <a href="#">code</a> . Examples are iso4217:USD, rfc3066:zh-Hant, nc:15062000, nasdaq:msft and cusip:594918104. A compact URI is not the same as a QName (qualified name) [W3C: Namespaces in XML ( <a href="http://www.w3.org/TR/REC-xml-names/">http://www.w3.org/TR/REC-xml-names/</a> )], though there are substantial similarities. The two main differences are: (i) the code does not have to be a valid XML name (eg, can start with a digit), and (ii) the scheme alias does not have to be declared using a namespace declaration.
concept	Anything that one may wish to refer to, eg Diplomacy, Paris, the Euro, OECD, the Japanese language, the IMF, Oil, Madonna, Olympic Games. Thus concept here has a broader meaning than is usual. This is because we are dealing with the idea of Paris, rather than with Paris itself, the idea of Oil, rather than Oil itself, and so on.
concept type	A concept type allows the logical grouping of all similar <a href="#">concepts</a> , regardless of the <a href="#">schemes</a> the concepts belong to. Examples of concept type might be: Person, Organisation, Language, Business Sector, News Subject or Geography. A concept type is itself a concept and, as such, is represented by a code in a scheme.
concept URI	A URI which identifies a <a href="#">concept</a> . A concept URI is obtained by appending the <a href="#">code</a> representing this concept to the <a href="#">scheme URI</a> corresponding to the <a href="#">scheme</a> to which the code belongs.
conformance level	A layer of functionality defined by a standard. The NewsML 2 power conformance level is a superset of the NewsML 2 core conformance level, both in terms of structure and processing.
controlled vocabulary	A set of <a href="#">codes</a> , managed by some authority (eg a person or an organisation), employing some mechanism (eg an XML Schema, a Web page, an RFC, or a collection of <a href="#">TopicItems</a> ). A controlled vocabulary is either a <a href="#">scheme</a> or is anonymous (ie an <a href="#">anonymous controlled vocabulary</a> ). Each code in a controlled vocabulary represents a <a href="#">concept</a> . NOTE: There are some cases where codes within controlled vocabularies do not represent concepts, but these are unlikely to be used by IPTC standards. An example is a controlled vocabulary containing permitted names for babies.
constrained metadata	A <a href="#">metadata container</a> which <i>either</i> accepts only <a href="#">codes</a> of a specified <a href="#">concept type</a> <i>or</i> accepts only codes from a specified <a href="#">controlled vocabulary</a> (which may be an



<b>container</b>	<a href="#">anonymous controlled vocabulary</a> or a <a href="#">scheme</a> ).
<b>description</b>	A human-readable string, held within a <a href="#">TopicItem</a> , which describes the <a href="#">concept</a> which the TopicItem represents. Descriptions will be implemented using <a href="#">labels</a> .
<b>directed graph</b>	A <a href="#">graph</a> connected using <a href="#">links</a> .
<b>edge</b>	A line connecting two <a href="#">nodes</a> of a <a href="#">graph</a> . An edge may be directed or undirected.
<b>formal metadata element</b>	A metadata element designed to hold data that is not <a href="#">free-form text</a> , eg <a href="#">codes</a> . Such data is usually consumed by software. An example of such an element is <i>subject</i> . An example value of <i>subject</i> is “nc:15062000”.
<b>free-form metadata element</b>	A metadata element designed to hold <a href="#">free-form text</a> . Such data is usually consumed by humans. An example of a free-form metadata element is <i>title</i> . An example value of <i>title</i> is “Ian Thorpe makes a splash”. Free-form metadata elements are based on the <a href="#">label</a> datatype.
<b>free-form text</b>	Arbitrary text, ie text which does not consist of <a href="#">codes</a> drawn from a <a href="#">controlled vocabulary</a> . A headline is an example of free-form text.
<b>globally unique identifier</b>	An identifier that is unique, unambiguous, and persistent. Being unique and unambiguous means that there is a 1:1 relationship between the identifier and the identified object. Being persistent means that the identifier never changes as time passes, and that it is never reused as an identifier for another object even if the original object disappears. See also <a href="#">persistent identifier</a> , <a href="#">unambiguous identifier</a> , <a href="#">unique identifier</a> .
<b>graph</b>	A set of <a href="#">nodes</a> , optionally connected by lines, called <a href="#">edges</a> . The connecting lines can be directed or undirected.
<b>identifier</b>	A string used to identify a specific <a href="#">resource</a> . See <a href="#">persistent identifier</a> , <a href="#">unambiguous identifier</a> , <a href="#">unique identifier</a> , and <a href="#">globally unique identifier (GUID)</a> .
<b>label</b>	A datatype designed to hold <a href="#">free-form text</a> .
<b>link</b>	A directed <a href="#">edge</a> (also called arc or arrow).
<b>metadata</b>	Data which asserts something about some other data.
<b>metadata container</b>	A location (eg an element or an attribute) in a data structure, designed to hold <a href="#">metadata</a> .
<b>metadata element</b>	An XML element, which is either a <a href="#">formal metadata element</a> or a <a href="#">free-form metadata element</a> .
<b>named entity</b>	A named entity may be a person, place, event, organization, product name, object name or any other news-related real life entity.
<b>news provider</b>	A provider of news content. May be a news agency, a syndication company, a newspaper, a magazine ... or a blogger.
<b>node</b>	An object (ie point) in a <a href="#">graph</a> .
<b>ontology</b>	See <a href="#">taxonomy</a> .



<b>persistent identifier</b>	An identifier which is associated with the same resource for all time. See also <a href="#">unambiguous identifier</a> , <a href="#">unique identifier</a> , and <a href="#">globally unique identifier (GUID)</a> .
<b>processor</b>	An application that supports the handling and processing of Items. Also known as a user agent.
<b>property</b>	An (XML) element or attribute.
<b>provider</b>	See <a href="#">news provider</a>
<b>publish</b>	Make available to other parties involved in the news exchange process, according to the business practices of the provider.
<b>representation</b>	The physical form of something
<b>representation of a TopicItem</b>	A manifestation of a given <a href="#">TopicItem</a> that is suited for some particular purpose. The various representations of a given TopicItem may differ, for example, in whether they are verbose or concise, or in which language(s) they use for descriptions.
<b>resource</b>	A resource is anything that has identity.
<b>scheme</b>	A <a href="#">controlled vocabulary</a> which is not an <a href="#">anonymous controlled vocabulary</a> . Each scheme is identified by a <a href="#">scheme URI</a> .
<b>scheme alias</b>	A character sequence which is used as an abbreviation for a <a href="#">scheme URI</a> . A scheme alias is similar to an XML Namespace prefix.
<b>scheme URI</b>	The URI which identifies the <a href="#">scheme</a> .
<b>synonym</b>	Synonyms are { <a href="#">scheme alias</a> , <a href="#">code</a> } pairs that refer to the same <a href="#">concept</a> . Synonymy is a symmetric relationship, which means that if A is synonymous with B, then B is also synonymous with A. An example of synonyms is "cemetery" and "graveyard".
<b>target</b>	The data being described by the <a href="#">metadata</a> . We have chosen to use the term <a href="#">target</a> rather than subject (the term used by RDF [ <a href="http://www.w3.org/RDF/">http://www.w3.org/RDF/</a> ]), as subject has a special meaning in the context of News.
<b>taxonomy</b>	In a broad sense, taxonomy is the science of classification, but is often taken to mean a particular classification. In the context of the NewsML 2, a taxonomy is a collection of <a href="#">concepts</a> , with associated <a href="#">codes</a> . A taxonomy may support typed relationships between concepts. Such a taxonomy is sometimes known as an <a href="#">ontology</a> or <a href="#">thesaurus</a> .
<b>thesaurus</b>	See <a href="#">taxonomy</a> .
<b>TopicItem</b>	A specialised data structure containing data representing a <a href="#">concept</a> and, optionally, providing information such as status, descriptions, relationships, etc. A TopicItem is identified by a { <a href="#">scheme alias</a> , <a href="#">code</a> } pair. The reverse relationship does not necessarily hold. In other words, there is no requirement that each { <a href="#">scheme alias</a> , <a href="#">code</a> } pair has a corresponding TopicItem. See also <a href="#">representation of a TopicItem</a> .
<b>tuple</b>	A set of values. The word tuple is a generalisation of the sequence: couple, triple, quadruple, quintuple, sextuple, etc. Tuples are conventionally written as a comma-separated list of items, enclosed within braces, eg { <a href="#">scheme alias</a> , <a href="#">code</a> }.



- type** See [concept type](#).
- unambiguous identifier** An identifier is unambiguous if it identifies one and only one object (but an object may have several different unambiguous identifiers). See also [globally unique identifier](#).
- unconstrained metadata container** A [metadata container](#) that accepts [codes](#) from any [controlled vocabulary](#) and of any [concept type](#).
- unique identifier** The only identifier of a resource. See also [persistent identifier](#), [unambiguous identifier](#), and [globally unique identifier \(GUID\)](#).
- Web resource** The data content that can be retrieved from a Web server using a Web-compliant transport protocol.  
See also [resource](#).

=== END of document ===