



IPTC Standards



Specification Version 1.0
Core Conformance Level

Document Revision 2

International Press Telecommunications Council
Copyright © 2008. All Rights Reserved.
www.iptc.org



Copyrights

Copyright © 2008 by IPTC, the International Press Telecommunications Council. All Rights Reserved.

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.

Acknowledgements

This documentation is the result of a team effort by members of the International Press Telecommunications Council, with input and assistance from other contributors.

The document reflects and includes contributions from work by the following persons (ordered by family name):

The effort to develop EventsML-G2 was lead by Johan Lindgren (Tidningarnas Telegrambyrå) and these persons contributed: Dominic Chan (CNW Group), Arnaud Descamps (Relax News), Darko Gulija (HINA), Geoffrey Haynes (AP), George Landau (NewsEngin), Laurent Le Meur (AFP), Michele Mottini (Tera DP), Michael Steidl (IPTC Managing Director), Charles Tichenor (AP).

The Technical Writer of this document was Scott Meltzer.

About the Standard

Specification Versioning History

Version	Issue Date	Approved by	Remarks
1.0	2008-01-31	IPTC Standards Committee	

Document Revision History

Revision	Issue Date	Author (revised by)	Remarks
1	2008-01-31	Michael Steidl	
2	2008-02-25	Michael Steidl	Errata fixed.

About this Document

This document specifies the IPTC news exchange standard EventsML-G2 which is a conceptual and processing model making freely available the IPTC knowledge of the most effective ways to structure, describe, manage and exchange newsworthy information about events.

Status of this Document

This document is under the governance of the IPTC EventsML-G2 Working Group (EventsML-G2 WG) of the IPTC News Content Working Party (NCT WP).

This is a specification document which was endorsed by the IPTC members and may be updated, replaced or obsoleted by other documents at any time.

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

<http://www.iptc.org/std/EventsML-G2/1.0/>

Public comments should be sent to the forum and mailing list at:

<http://tech.groups.yahoo.com/group/eventsml-g2>

A page with all errata not covered by the latest version of the EventsML-G2 specification is available at:

<http://www.iptc.org/goto?EventsML-G2-2.0Errata>



The Full Set of Specification Documents

The full set of specification documents for EventsML-G2 1.0 consists of (file names are added, # is to be substituted by the most current document revision number):

- ◆ This Specification document - EventsML-G2_1.0_Specification_#.pdf
- ◆ These XML Schema files:

For the Core Conformance Level (see [Conformance Levels](#) on page 12):

- EventsML-G2_1.0-spec-Framework-Core.xsd
- EventsML-G2_1.0-spec-Framework-Core_#.xsd
- EventsML-G2_1.0-spec-ConceptItem-Core.xsd
- EventsML-G2_1.0-spec-ConceptItem-Core_#.xsd
- EventsML-G2_1.0-spec-KnowledgeItem-Core.xsd
- EventsML-G2_1.0-spec-KnowledgeItem-Core_#.xsd
- NAR_1.1-spec-Framework-Core.xsd
- NAR_1.1-spec-Framework-Core_1.xsd

For the Power Conformance Level (see [Conformance Levels](#) on page 12):

- EventsML-G2_1.0-spec-Framework-Power.xsd
- EventsML-G2_1.0-spec-Framework-Power_#.xsd
- EventsML-G2_1.0-spec-ConceptItem-Power.xsd
- EventsML-G2_1.0-spec-ConceptItem-Power_#.xsd
- EventsML-G2_1.0-spec-KnowledgeItem-Power.xsd
- EventsML-G2_1.0-spec-KnowledgeItem-Power_#.xsd
- NAR_1.1-spec-Framework-Power.xsd
- NAR_1.1-spec-Framework-Power_1.xsd

All files above can be obtained from:

<http://www.iptc.org/std/EventsML-G2/1.0/specification/>

XML Schemas for the generic G2 Package Item and the News Message can be obtained from:

<http://www.iptc.org/std/NAR/1.1/specification/>

Note on the XML Schema File Names

XML Schemas are revised for two reasons:

- ◆ The EventsML-G2 specifications have been changed: this results in a new version of the standards, this will be reflected by a new path to files and a new standard version number like EventsML-G2_1.9
- ◆ The XML Schema has been edited a) to fix errors and b) to change non-normative parts, like the wording of an element's annotation: this is reflected by a new revision number at the end of the file name like the "8" in EventsML-G2_1.0-spec-Framework-Core_8.xsd.

The XML Schema files without the document revision number (e.g. "_8") at the end of the file name are true copies of the latest document revision. This allows to apply a persistent reference to the latest XML Schema file version regardless of any edits of the document.



License Agreement

This document is issued under the *Non-Exclusive License Agreement for International Press Telecommunications Council Specifications and Related Documentation*

Important: International Press Telecommunications Council (IPTC) standard specifications for news (the Specifications) and supporting software, documentation, technical reports, web sites and other material related to the Specifications (the Materials) including the document accompanying this license (the Document), whether in a paper or electronic format, are made available to you subject to the terms stated below. By obtaining, using and/or copying the Specifications or Materials, you (the licensee) agree that you have read, understood, and will comply with the following terms and conditions.

1. The Specifications and Materials are licensed for use only on the condition that you agree to be bound by the terms of this license. Subject to this and other licensing requirements contained herein, you may, on a non-exclusive basis, use the Specifications and Materials.

2. The IPTC openly provides the Specifications and Materials for voluntary use by individuals, partnerships, companies, corporations, organizations and any other entity for use at the entity's own risk. This disclaimer, license and release is intended to apply to the IPTC, its officers, directors, agents, representatives, members, contributors, affiliates, contractors, or co-venturers acting jointly or severally.

3. The Document and translations thereof may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the copyright and license notices and references to the IPTC appearing in the Document and the terms of this Specifications License Agreement are included on all such copies and derivative works. Further, upon the receipt of written permission from the IPTC, the Document may be modified for the purpose of developing applications that use IPTC Specifications or as required to translate the Document into languages other than English.

4. Any use, duplication, distribution, or exploitation of the Document and Specifications and Materials in any manner is at your own risk.

5. NO WARRANTY, EXPRESSED OR IMPLIED, IS MADE REGARDING THE ACCURACY, ADEQUACY, COMPLETENESS, LEGALITY, RELIABILITY OR USEFULNESS OF ANY INFORMATION CONTAINED IN THE DOCUMENT OR IN ANY SPECIFICATION OR OTHER PRODUCT OR SERVICE PRODUCED OR SPONSORED BY THE IPTC. THE DOCUMENT AND THE INFORMATION CONTAINED HEREIN AND INCLUDED IN ANY SPECIFICATION OR OTHER PRODUCT OR SERVICE OF THE IPTC IS PROVIDED ON AN "AS IS" BASIS. THE IPTC DISCLAIMS ALL WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, ANY ACTUAL OR ASSERTED WARRANTY OF NON-INFRINGEMENT OF PROPRIETARY RIGHTS, MERCHANTABILITY, OR FITNESS FOR A PARTICULAR PURPOSE. NEITHER THE IPTC NOR ITS CONTRIBUTORS SHALL BE HELD LIABLE FOR ANY IMPROPER OR INCORRECT USE OF INFORMATION. NEITHER THE IPTC NOR ITS CONTRIBUTORS ASSUME ANY RESPONSIBILITY FOR ANYONE'S USE OF INFORMATION PROVIDED BY THE IPTC. IN NO EVENT SHALL THE IPTC OR ITS CONTRIBUTORS BE LIABLE TO ANYONE FOR DAMAGES OF ANY KIND, INCLUDING BUT NOT LIMITED TO, COMPENSATORY DAMAGES, LOST PROFITS, LOST DATA OR ANY FORM OF SPECIAL, INCIDENTAL, INDIRECT, CONSEQUENTIAL OR PUNITIVE DAMAGES OF ANY KIND WHETHER BASED ON BREACH OF CONTRACT OR WARRANTY, TORT, PRODUCT LIABILITY OR OTHERWISE.

6. The IPTC takes no position regarding the validity or scope of any Intellectual Property or other rights that might be claimed to pertain to the implementation or use of the technology described in the Document or the extent to which any license under such rights might or might not be available. The IPTC does not represent that it has made any effort to identify any such rights. Copies of claims of rights made available for publication, assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of the Specifications and Materials, can be obtained from the Managing Director of the IPTC.

7. By using the Specifications and Materials including the Document in any manner or for any purpose, you release the IPTC from all liabilities, claims, causes of action, allegations, losses, injuries, damages, or detriments of any nature arising from or relating to the use of the Specifications, Materials or any portion thereof. You further agree not to file a lawsuit, make a claim, or take any other formal or informal legal action against the IPTC, resulting from your acquisition, use, duplication, distribution, or exploitation of the Specifications, Materials or any portion thereof. Finally, you hereby agree that the IPTC is not liable for



any direct, indirect, special or consequential damages arising from or relating to your acquisition, use, duplication, distribution, or exploitation of the Specifications, Materials or any portion thereof.

8. Specifications and Materials may be downloaded or copied provided that ALL copies retain the ownership, copyright and license notices.

9. Materials may not be edited, modified, or presented in a context that creates a misleading or false impression or statement as to the positions, actions, or statements of the IPTC.

10. The name and trademarks of the IPTC may not be used in advertising, publicity, or in relation to products or services and their names without the specific, written prior permission of the IPTC. Any permitted use of the trademarks of the IPTC, whether registered or not, shall be accompanied by an appropriate mark and attribution, as agreed with the IPTC.

11. Specifications may be extended by both members and non-members to provide additional functionality (Extended Specifications) provided that there is a clear recognition of the IPTC IP and its ownership in the Extended Specifications and the related documentation and provided that the extensions are clearly identified and provided that a perpetual license is granted by the creator of the Extended Specifications for other members and non-members to use the Extended Specifications and to continue extensions of the Extended Specifications. The IPTC does not waive any of its rights in the Specifications and Materials in this context. The Extended Specifications may be considered the intellectual property of their creator. The IPTC expressly disclaims any responsibility for damage caused by an extension to the Specifications.

12. Specifications and Materials may be included in derivative work of both members and non-members provided that there is a clear recognition of the IPTC IP and its ownership in the derivative work and its related documentation. The IPTC does not waive any of its rights in the Specifications and Materials in this context. Derivative work in its entirety may be considered the intellectual property of the creator of the work. The IPTC expressly disclaims any responsibility for damage caused when its IP is used in a derivative context.

13. This Specifications License Agreement is perpetual subject to your conformance to the terms of this Agreement. The IPTC may terminate this Specifications License Agreement immediately upon your breach of this Agreement and, upon such termination you will cease all use, duplication, distribution, and/or exploitation in any manner of the Specifications and Materials.

14. This Specifications License Agreement reflects the entire agreement of the parties regarding the subject matter hereof and supersedes all prior agreements or representations regarding such matters, whether written or oral. To the extent any portion or provision of this Specifications License Agreement is found to be illegal or unenforceable, then the remaining provisions of this Specifications License Agreement will remain in full force and effect and the illegal or unenforceable provision will be construed to give it such effect as it may properly have that is consistent with the intentions of the parties.

15. This Specifications License Agreement may only be modified in writing signed by an authorized representative of the IPTC.

16. This Specifications License Agreement is governed by the law of United Kingdom, as such law is applied to contracts made and fully performed in the United Kingdom. Any disputes arising from or relating to this Specifications License Agreement will be resolved in the courts of the United Kingdom. You consent to the jurisdiction of such courts over you and covenant not to assert before such courts any objection to proceeding in such forums.

IF YOU DO NOT AGREE TO THESE TERMS YOU MUST CEASE ALL USE OF THE SPECIFICATIONS AND MATERIALS NOW. IF YOU HAVE ANY QUESTIONS ABOUT THESE TERMS, PLEASE CONTACT THE MANAGING DIRECTOR OF THE INTERNATIONAL PRESS TELECOMMUNICATION COUNCIL. AS OF THE DATE OF THIS REVISION OF THIS SPECIFICATIONS LICENSE AGREEMENT YOU MAY CONTACT THE IPTC at www.iptc.org.

License agreement version of: 30 January 2006



Contents

1 Introduction to EventsML-G2	11
1.1 Overview	11
1.1.1 What is EventsML-G2?	11
1.1.2 Business Advantages of Using EventsML-G2	11
1.1.3 What is an Event – to be represented by EventsML-G2	11
1.2 Definitions	12
1.2.1 Event Information	12
1.2.2 Coverage Information	12
1.2.3 The Data Model	12
1.3 Conformance Levels	12
1.4 EventsML-G2 and iCalendar	13
2 Events	15
2.1 The Core Information about Events	15
2.2 Event Information in Items	16
2.2.1 Identification and Versioning of Items	16
2.2.2 Events in a NewsItem	16
2.2.2.1 News Metadata	17
2.2.3 An Event in a Concept Item	17
3 Representing News - newsItem	19
3.1 Description	19
3.2 Indication of Compliance with a Standard and Conformance Level	19
3.3 Identification and Versioning	19
3.4 Catalog of Controlled Vocabularies	20
3.5 Rights Information	20
3.6 Item Metadata	21
3.6.1 Management Metadata	21
3.6.2 Processing the Publish Status of an Item	21
3.6.2.1 State Transition Diagram	22
3.6.2.2 Use Cases Associated with a Status of Withheld	22
3.6.2.3 Processing Model on the Recipient Side	22
3.6.3 Processing of versionCreated	22
3.7 Item Links	23
3.7.1 Processing Links	23
3.8 News Content Metadata	23
3.8.1 Administrative Metadata	23
3.8.1.1 Dates Processing Model	24
3.8.1.2 Audience Processing Model	24
3.8.2 Descriptive Metadata	25
3.8.3 Other Content Metadata	25
3.9 newsItem Content	25
3.10 News Content Characteristics	26
4 Representing Concept Information - concept Component	27
4.1 Concept Component	27
4.2 Relationships Between Concepts	27
4.3 Details on Specific Entities	27
4.3.1 Contact Information	28
4.3.2 Postal Address	28
5 Managing Individual Concepts - conceptItem	29



- 5.1 Description 29
- 5.2 Structure of a conceptItem 29
- 5.3 Item Metadata 29
- 5.4 Concept related Metadata 29
- 5.5 conceptItem Content 29
- 6 Managing Sets of Concepts - knowledgeItem 30
 - 6.1 Description 30
 - 6.2 Structure of a knowledgeItem 30
 - 6.3 Item Metadata 30
 - 6.4 Knowledge Related Metadata 30
 - 6.5 knowledgeItem Content 30
- 7 Packaging Items - packageItem 31
 - 7.1 Structure of a packageItem 31
 - 7.2 Item Metadata 31
 - 7.3 Package Related Metadata 31
 - 7.4 packageItem Content 31
- 8 Dealing with Controlled Values 33
 - 8.1 {scheme, code} Pair, Scheme URI and Concept URI 33
 - 8.2 Qualified Code (QCode) 34
 - 8.2.1 Processing QCodes 34
 - 8.3 Processing Catalogs 34
 - 8.3.1 Structure of a Catalog 34
 - 8.3.2 Processing Remote Catalogs 35
 - 8.3.3 Caching a Catalog 35
 - 8.3.4 Checking a Catalog 35
 - 8.4 Processing Schemes 35
 - 8.4.1 Evolution of Scheme URIs 35
 - 8.4.2 Retrieving All Terms of a Scheme 35
 - 8.5 Qualified and Typed Properties 36
 - 8.6 Flexible Properties 36
- 9 Dealing with Labels and Blocks 37
 - 9.1 Introduction 37
 - 9.2 Internationalization Attributes 37
- 10 Exchanging Items - newsMessage 38
 - 10.1 Message Information 38
 - 10.2 About Using Schemes in a newsMessage 38
- 11 Specification Reference 39
 - 11.1 Introduction to the Common Components 39
 - 11.2 General Specifications 39
 - 11.2.1 XML Namespaces 39
 - 11.2.2 MIME Types 40
 - 11.2.3 Extension Points in XML 40
 - 11.3 Implementation Design Rules 40
 - 11.4 Processing Model Terminology 40
 - 11.5 Component Structure Format 41
 - 11.6 Element Definitions 42
 - 11.6.1 Access 42
 - 11.6.2 Access Status 43



11.6.3 Accountable Person 44

11.6.4 Address Line 45

11.6.5 Affiliation 46

11.6.6 Audience 47

11.6.7 Broader 48

11.6.8 By 49

11.6.9 Capacity 50

11.6.10 Catalog 51

11.6.11 Channel for News Message 52

11.6.12 Concept 53

11.6.13 Concept Definition 54

11.6.14 Concept Identifier 55

11.6.15 Concept Item 56

11.6.16 Concept Name 57

11.6.17 Concept Set 58

11.6.18 Confirmation 59

11.6.19 Contact Information 60

11.6.20 Content Metadata for Concept Items 61

11.6.21 Content Metadata for Knowledge Items 62

11.6.22 Content Metadata for News Items 63

11.6.23 Content Metadata for Package Items 64

11.6.24 Content Provider 65

11.6.25 Content Set 66

11.6.26 Contributor 67

11.6.27 Copyright Holder 68

11.6.28 Copyright Notice 69

11.6.29 Country 70

11.6.30 Country Area 71

11.6.31 Creator 72

11.6.32 Date Content Created 73

11.6.33 Date Content Modified 74

11.6.34 Date Item Embargo Ends 75

11.6.35 Date Item First Created 76

11.6.36 Date Item Version Created 77

11.6.37 Date of Birth 78

11.6.38 Date of Death 79

11.6.39 Date of Dissolution 80

11.6.40 Date of Foundation 81

11.6.41 Date of Transmission 82

11.6.42 Dateline 83

11.6.43 Dates 84

11.6.44 Description 85

11.6.45 Destination 86

11.6.46 Duration 87

11.6.47 Editorial Note 88

11.6.48 Editorial Service 89

11.6.49 Email Address 90

11.6.50 End Date/Time 91

11.6.51 Event 92

11.6.52 Event Details 93

11.6.53 Event Location 94

11.6.54 Events Wrapper 95

11.6.55 Exclusion Date 96

11.6.56 Exclusion Rule 97

11.6.57 Facet 98

11.6.58 Fax Number 99

11.6.59 File Name 100

11.6.60 Genre 101



11.6.61 Geographic Position	102
11.6.62 Geopolitical Area Details	103
11.6.63 Group	104
11.6.64 Group Reference	105
11.6.65 Group Set	106
11.6.66 Headline	107
11.6.67 Inline Data (NewsML-G2 Specific)	108
11.6.68 Inline XML (NewsML-G2 Specific)	110
11.6.69 Instant Messaging Address	112
11.6.70 Information Source	113
11.6.71 Item Class	114
11.6.72 Item Metadata	115
11.6.73 Item Reference	116
11.6.74 Item Set	117
11.6.75 Item Title	118
11.6.76 Knowledge Item	119
11.6.77 Language	120
11.6.78 Line Break	121
11.6.79 Link	122
11.6.80 Link Title	123
11.6.81 Locality	124
11.6.82 Located	125
11.6.83 Location	126
11.6.84 Location Details	127
11.6.85 Message Header	128
11.6.86 Narrower	129
11.6.87 News Coverage	130
11.6.88 News Item (NewsML-G2 Specific)	131
11.6.89 News Message	132
11.6.90 Note	133
11.6.91 Occurrence Status	134
11.6.92 Opening Hours	135
11.6.93 Origin	136
11.6.94 Organisation Details	137
11.6.95 Organiser	138
11.6.96 Package Item	139
11.6.97 Participant	140
11.6.98 Participation Requirement	141
11.6.99 Person Details	142
11.6.100 Phone Number	143
11.6.101 POI Details	144
11.6.102 Postal Address	145
11.6.103 Postal Code	146
11.6.104 Priority	147
11.6.105 Property Value Name	148
11.6.106 Publish Status	149
11.6.107 Recurrence Date	150
11.6.108 Recurrence Rule	151
11.6.109 Registration	152
11.6.110 Related Concept	153
11.6.111 Remote Catalog Reference	154
11.6.112 Remote Content (NewsML-G2 Specific)	155
11.6.113 Rights Information	157
11.6.114 Role in the Workflow	158
11.6.115 Same As	159
11.6.116 Scheme Declaration	160
11.6.117 Sender	161
11.6.118 Slugline	162



11.6.119 Start Date/Time	163
11.6.120 Subject	164
11.6.121 Timestamp	165
11.6.122 Transmission Identifier	166
11.6.123 Type of a Concept	167
11.6.124 Urgency	168
11.6.125 Usage Terms	169
11.6.126 Web Address	170
11.7 Element Group Definitions	171
11.7.1 Recurrence Group	171
11.7.2 Concept Relationships Group	171
11.7.3 Entity Details Group	171
11.7.4 Administrative Metadata Group	172
11.7.5 Knowledge Descriptive Metadata Group	172
11.7.6 Descriptive Metadata Group	172
11.7.7 Item Management Group	173
11.8 Datatype Definitions	174
11.8.1 Any Item Type	174
11.8.2 Approximate Date and Time Property Type	175
11.8.3 Block Type	176
11.8.4 Concept Identifier Type	177
11.8.5 Content Metadata Type	178
11.8.6 Date and Optional Time Property Type	179
11.8.7 Date and Optional Time Type	180
11.8.8 Date and Time Property Type	181
11.8.9 Electronic Address Type	182
11.8.10 Flexible Property Type	183
11.8.11 Integer 0 to 100 Type	184
11.8.12 Integer 1 to 9 Type	185
11.8.13 International String Type	186
11.8.14 IRI Type	187
11.8.15 Label Type	188
11.8.16 Link Type	189
11.8.17 QCode Property Type	190
11.8.18 QCode Type	191
11.8.19 Qualified Property Type	192
11.8.20 Recurrence Rule Type	193
11.8.21 Truncated Date and Time Property Type	194
11.8.22 Truncated Date and Time Type	195
11.8.23 Typed Qualified Property Type	196
11.9 Attribute Group Definitions	197
11.9.1 Internationalization Attributes Group	197
11.9.2 News Content Attributes	197
11.9.3 News Content Characteristics	197
12 Glossary	201
13 References	204
14 Index	205
14.1 Elements	205
14.2 Datatypes	207



1 Introduction to EventsML-G2

EventsML-G2 specifications are a member of the Family of IPTC G2-Standards which is built on a common structural and function framework called the IPTC News Architecture (NAR). For this reason many components of EventsML-G2 are common with other members of the G2-Standards, e.g. NewsML-G2.

To better understand the terminology the IPTC used for the G2-Standards specification we recommend the **Glossary** (page 201) as reference, as it provides an extensive set of terms and their definitions.

1.1 Overview

1.1.1 What is EventsML-G2?

- ◆ EventsML-G2 is a standard for conveying event information in a news industry environment.
- ◆ EventsML-G2 is a member of the Family of IPTC G2-Standards, this family shares a lot of common specifications (the IPTC “News Architecture - NAR”) for the exchange of news items and knowledge about topics and concepts.
- ◆ EventsML-G2 may be used for:
 - Receiving all facts about an event from the event organiser,
 - Publishing all facts about a specific event by a news provider,
 - Publishing all or only a subset of the facts of one to many events by event listings,
 - Adding information regarding the coverage of an event by a news provider to the distributed event facts, e.g. for daybooks,
 - Storing facts about knowledgeable events in archives.

1.1.2 Business Advantages of Using EventsML-G2

EventsML-G2 are:

- ◆ Comprehensive (= many types of events may be covered).
- ◆ Flexible (= copies of substructures may be used many times, e.g. all the people appearing at an event).
- ◆ Extensible (= news provider specific data may be added) data structure to capture facts about events.

EventsML-G2 may express facts and information about events either by literal text (free text) or by codes from controlled vocabularies.

EventsML-G2 provides very flexible date types:

- ◆ year, month, day, optionally plus time
- ◆ year and month only or even year only
- ◆ approximative dates = a date range

EventsML-G2 reuses building blocks from the common News Architecture allowing to reuse software components, making their implementation cheaper.

EventsML-G2 makes use of industry standards: allows processing with standard tools. The EventsML-G2 syntax is built on XML, the Extensible Markup Language of the W3C, furthermore, EventsML-G2 makes use of W3C XML Schema and complies with the basic notion of the Semantic Web, the Resource Description Framework (RDF). This allows an easy transfer of EventsML-G2 structures to other XML-based standards and the integration of information about an event into the Semantic Web.

1.1.3 What is an Event – to be represented by EventsML-G2

An event is “something that happens” by definition. For the news industry, it is “something that happens and is subject to news coverage.” All the events in a day make up an “agenda,” which can be a marketable product sold to clients or simply an internal daybook used by editors to organise their work.



An event is planned or unplanned, with breaking news capable of overshadowing everything on the schedule.

Breaking news can generate a series of planned events; it becomes part of the daily news agenda the moment a decision is made to cover it. For agencies, this occurs with the first advisory announcing plans to provide coverage. For broadcasters it comes with the dispatch of a news team to the site; for newspapers it is when space is reserved for the story or page makeups rearranged.

Automated systems need to store and exchange information about news events. This is currently done in an ad-hoc manner, leading to overly-specialized formats and incompatible exchange by models. From that the IPTC learned that the industry would benefit from an event information interchange standard. Such a standard would facilitate the efficient exchange of event information, and the creation of better tools to support event management.

1.2 Definitions

1.2.1 Event Information

The event information describes a particular event in detail. This includes the “who”, “what”, “when”, and “where” information for the event along with identification and publication (news management) information. The event information also includes facilities for relating events to each other and relating news items (both complete and incomplete) to the event information.

1.2.2 Coverage Information

The coverage information describes the plan of news coverage for this event. Each event may have zero or more assignments containing this coverage information. This information can be used internally within a news organisation for assignment of resources, planning of coverage, etc. It can also be used to publish information about expected coverage, so that consumers of the news coverage can plan their own news coverage accordingly.

1.2.3 The Data Model

The data model for EventsML-G2 has to cover two different facets of event information which relate to a basic distinction made for all G2 standards:

- ◆ Topical News: is typically volatile information in the sense of “nothing is older than yesterday's news”.
- ◆ Persisting Knowledge: is information which is remembered and referenced to for a long time.

For EventsML-G2 this is reflected by two different data models:

- ◆ Volatile information about an event is represented by an “event” structure which is plugged into a NewsML-G2 news item as its content. A single news item may include one to many event structures. This kind of event information cannot be referenced as persisting information from any other item. Find details about this data model in section [Events in a NewsItem](#) (page 16).
- ◆ Persistent information about an event is represented by an EventsML-G2-Concept Item which is a generic NAR structure for concepts extended by a set of detailed information specific to an event. As any other kind of a Concept Item also this specific one for events can be referenced by its Concept Identifier.

The same applies to KnowledgeItems: a variant with event specific extensions is available, in particular event details are added to the concept structure inside the KnowledgeItem.

Find details about this data model in section [An Event in a Concept Item](#) (page 17).

The most important thing to note about the EventsML-G2 data model is that the core structures holding information about an event are identical for both the content plugged into a News Item and the extension of a Concept Item. Hence it is very easy to build a single EventsML-G2 processor for topical and persisting information about an event.

1.3 Conformance Levels

Another feature EventsML-G2 inherits from the NAR are the two conformance levels “Core” and “Power”.

Different conformance levels are defined in the model, each of them related to a level of complexity (at the conceptual and processing level) of the related Items. This feature adds modularity to the model.

The current model defines two conformance levels named “core conformance level” (CCL) and “power conformance level” (PCL). The core conformance level is focused on simplicity and interoperability. The power conformance level is a superset of the core conformance level which gives more flexibility to providers who choose it, but the recipient processors are more complex to program to comply with PCL and interoperability is lower than for CCL as not all recipients will implemented the power level.

A compliant processor must therefore assert supporting either “core” or “power” functionality.

As the “power” features are only an extension of the “core” features, a “core” compliant processor SHOULD process “power” Items by simply ignoring the information pertaining to the “power” level.

EventsML-G2 specifies does not specify in its own scope data structures which are different for the two conformance levels, but it inherits specifications of datatypes of properties and attributes from the NAR which are different at the two conformance levels. The data types corresponding to the conformance level are defined in the specification tables in the [Specification Reference](#) (page 39).

1.4 EventsML-G2 and iCalendar

A well known and widely used standard for events data is “iCalendar” which is specified by RFC 2445.

EventsML-G2 compares very well to it as it covers virtually all features of an iCalendar Event Component:

Table 1. iCalendar-to-EventsML-G2 Component Mapping

iCalendar Event Component (Alphabetically)	Corresponding EventsML-G2 Component
attach	“link” property of a G2-item
attendee	“participant” property
categories	“subject” property
class	Access management functionality, no direct equivalence in EventsML-G2
comment	“note” property (under “event” for news and “concept” for a concept)
contact	“contactInfo” property (under eventDetails)
created	“contentCreated” property (in contentMeta) of a G2-item for news or a concept.
description	“definition” property (under “event” for news and “concept” for a concept)
dtend	“end” property (under eventDetails/dates)
dtstamp	“contentCreated” property (in contentMeta) of a G2-item for news or a concept.
dtstart	“start” property (under eventDetails/dates)
duration	“duration” property (under eventDetails/dates)
exdate	“exDate” property (under eventDetails/dates)
exrule	“exRule” property (under eventDetails/dates)
geo	“position” property (under eventDetails/location/geoAreaDetails)
last-mod	“contentModified” property (in contentMeta) of a G2-item for news or a concept.
location	“location” property (under eventDetails)



Table 1. iCalendar-to-EventsML-G2 Component Mapping (Continued)

iCalendar Event Component (Alphabetically)	Corresponding EventsML-G2 Component
organizer	“organiser” property (under eventDetails)
priority	As this iCalendar property reflects the priority for a calendar of an individual no equivalent exists in EventsML-G2.
rdate	“rDate” property (under eventDetails/dates)
recurid	No direct equivalence in EventsML-G2, assigned functionality may be replicated by G2-item means.
related	No direct equivalence, but relationships can be expressed by other G2-item means
resources	Not covered by EventsML-G2 1.0, planned for future versions.
rrule	“rRule” property (under eventDetails/dates)
rstatus	Scheduling protocol functionality is not covered by EventsML-G2
seq	“version” attribute of the G2-item's root element
status	“confirmation” (under eventDetails/dates) reflects the status of confirmation of the dates of the event, while “occurStatus” (under eventDetails) reflects the overall status of the event.
summary	“name” property (under “event” for news and “concept” for a concept)
transp	Not covered by EventsML-G2
uid	“guid” attribute of the G2-item's root element
url	No direct equivalence. For G2-items it may be defined individually by each news provider how to resolve the identifier of an G2-item to an accessible location.
x-prop	EventsML-G2 provides “Extension points” for this purpose.



2 Events

2.1 The Core Information about Events

Regardless whether the information about an event is topical or persistent (see [The Data Model](#) on page 12) the same structure is used to mark it up.

The information about an event includes, first a set of more generic properties:

- ◆ A natural language **name** (page 57) for the event. This name should be rather concise and could be expressed in different languages.
- ◆ A natural language **definition** (page 54) for the event and it could be more extensive than the name, it could explain facets in detail. It can also be expressed in different languages.

The *role* attribute of a definition could be used to provide variants of an explanation, e.g. a short one for overviews and a rather extensive one for a detailed presentation.

- ◆ A natural language **note** (page 133) about the event. This could be an explanation of details or background information regarding the definition. Again this note can be expressed in different languages and can be qualified by a *role* attribute.
- ◆ A **facet** (page 98) property to further qualify the nature of the event. The facet can take either literal values or values from a controlled vocabulary and could be used to express e.g. that this event is a concert, a hockey game or a press conference.
- ◆ The properties **sameAs** (page 159), **broader** (page 48), **narrower** (page 129) and **related** (page 153) can be used to define a relationship with this event to another event or concept.

In particular broader may be used to express that this event is a sub-event to another event, e.g. a break-out session of a big conference, one competition of the Olympic Games or one of the concerts of a festival.

Then a set of rather event-specific properties - wrapped by the **eventDetails** (page 93) property:

- ◆ A **dates** (page 84) sub-structure to express the start date and the end date or duration of the event. This includes using the “approximative dates”, i.e. a range of dates, and on date in this range as a kind of best guess or most likely date.

If this event is recurring this can be expressed by means of recurrence properties which align to equivalent properties of the iCalendar standard RFV 2445, see more below.

- ◆ An **occurStatus** (page 134) to indicate the status of the occurrence - if this is a unplanned or planned event, and if it is planned how likely it is to occur.
- ◆ A set of **registration** (page 152) information which may be used to define how persons have register for the event, this may include the accreditation of journalists.
- ◆ A set of **accessStatus** (page 43) information.
- ◆ A set of **participationRequirement** (page 141) properties. This could be used e.g. for expressing age limits - think of required parental guidance for movies - or for formal requirements for training course events.
- ◆ A set of **subject** (page 164) properties to express what the event is about. Be aware of the difference between a facet and a subject: a facet should indicate the nature of the event, what the event is, while a subject indicates applicable categories for what the event is about. For example, “concert” is a facet, while “music” or “Wolfgang Amadeus Mozart” is a matching subject.
- ◆ A set of **location** (page 126) properties. In most cases it will be the only location of where the event will take place - but e.g. festivals could have more than one location.
- ◆ A set of **participant** (page 140) properties to list all kinds of parties appearing in different roles at the event - the particular role can be expressed by the *role* attribute.



- ◆ A set of **organiser** (page 138) properties to list all parties involved in organising the event - the particular role can be expressed again by the *role* attribute
- ◆ A set of **contactInfo** (page 60) properties for the event. Be aware that the location, the participant and the organiser properties may contain contactInfo structures, but they pertain only to this particular property while this contactInfo is to be used for the event as a whole.
- ◆ A set of **language** (page 120) properties to reflect all languages which will be spoken at the event.
- ◆ A set of **newsCoverage** (page 130) properties. They can provide all necessary information about the planned coverage for this event by a news provider. To express this planned coverage either a free-text property can be used or provider specific additional properties.
- ◆ As for many wrapping elements in G2-Standards, the information about an event can also be extended by provider-specific properties.

2.2 Event Information in Items

2.2.1 Identification and Versioning of Items

It is possible to positively identify any kind of an item specified by the G2-Standards as it moves through the news workflow, and is transferred from place to place and from system to system.

Each and every item – including News Items, Concept Items, Knowledge Items and Package Items – **MUST** have a *guid* attribute, that is a persistent and globally unique identifier. The guid is required to be in the form of an IRI. Any IRI capable of acting as a globally unique identifier is accepted.

Note: the IPTC will provide the newsml-URN for this purpose, specified by a successor of RFC-3085.

A newsItem **MAY** have a *version* attribute, and this version **MUST** be incremented when the content of the Item is updated. The first version **MUST** be numbered 1: if the version is not explicitly set, the value "1" must be assumed as default by the recipient of the Item.

The *standard* attribute must reflect the G2-Standard which governs the structure:

- ◆ For topical events in a News Item this is NewsML-G2.
- ◆ For persisting information about an event in a Concept Item this is EventsML-G2.

The *standardversion* attribute must reflect the version of the standard as it is implemented by the corresponding XML Schema.

Samples:

```
<newsItem standard="NewsML-G2" standardversion="2.0"
  guid="urn:newsml:iptc.org:20071231:newssample" version="2"
  xmlns="http://iptc.org/std/nar/2006-10-01/" >
</newsItem>
```

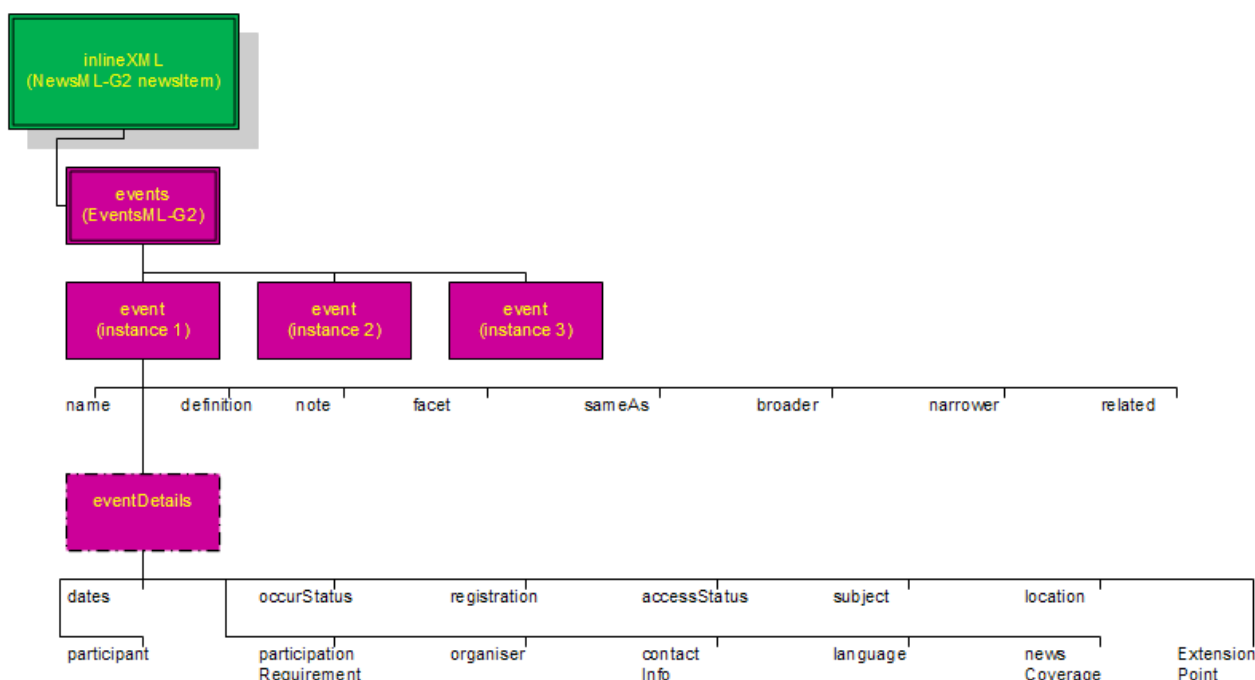
```
<conceptItem standard="EventsML-G2" standardversion="1.0"
  guid="urn:newsml:iptc.org:20071231:eventsample" version="4"
  xmlns="http://iptc.org/std/nar/2006-10-01/" >
</conceptItem>
```

2.2.2 Events in a NewsItem

Topical event information must be conveyed by using the NewsML-G2 NewsItem (see **Representing News - newsItem** on page 19) as a wrapping item instance. The structure of a NewsItem defines a special node where to attach content plug-ins, the inlineXML element.

For EventsML-G2 an **Events Wrapper** (page 95) element will be added as child to **Inline XML (NewsML-G2 Specific)** (page 110) and it acts as a wrapper of one-to-many **Event** (page 92) elements, each representing the topical information of a single event.

Figure 1. Event Information in a News Item



The event element wraps a group of more generic descriptions and a couple of details about an event. The first group is made of a short name which can be displayed as a one-liner, a more comprehensive definition of the event and a note with supplemental information.

A sibling to this group is eventsDetails, it wraps all the details of the event, when and where it happens, who is involved and how to get there.

Finally optional information about the planned news coverage of this item may be added.

2.2.2.1 News Metadata

In general the News Metadata section of a NewsItem - wrapped by the **Content Metadata for News Items** (page 63) element - should be populated and used as specified for NewsML-G2.

Further to this general recommendation these event specific considerations apply:

- ◆ If more than a single event is conveyed by a NewsItem the content metadata apply to the set of events as a whole. In most cases this set will be selected from a larger repository by some rules, like “events of next week”, or “music events”. This could be reflected by e.g. the headline, the description or even the subject property.
- ◆ Genre property: an appropriate value should be applied, like “almanac” or “daybook” from the IPTC Genre NewsCodes
- ◆ Language property: be aware of the difference between the language property of the content metadata - it reflects the languages used in the content, in this case in the description of the events - and the language property of the event structure - it reflects a language which is used at an event.

2.2.3 An Event in a Concept Item

The persisting knowledge facet of event information is represented by the EventsML-G2 Concept Item, an extended variant of the generic G2 Concept Item (see **Managing Individual Concepts - conceptItem** on page 29), and subsequently the EventsML-G2 Knowledge Item, an extended variant of the generic G2 Knowledge Item (see **Managing Sets of Concepts - knowledgeItem** on page 30).

Any Concept Item or Knowledge Item provides a group of generic definitions and a set of details specific to a kind of concept, in this case specific to an event.

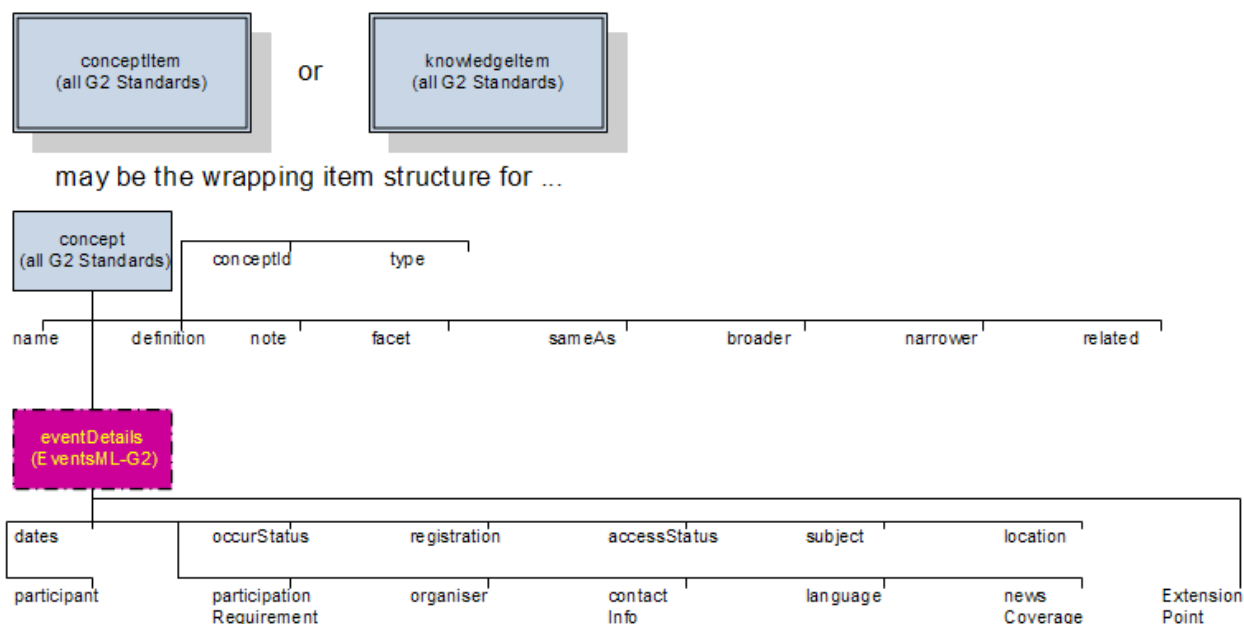
Generic definitions are part of the basic Concept Item or Knowledge Item and include:

- ◆ The Concept Identifier for this event.
- ◆ Relationships to other events.
- ◆ A name, a definition, explanatory notes and refining facets.

Find more about the generic part in section **Representing Concept Information - concept Component** (page 27).

In EventsML-G2 Concept Items the value of the type of a concept (conceptItem/concept/type) must be set to concept URI of <http://cv.iptc.org/newscodes/cpnature/event> which may translate to a QCode of cpnat:event.

Figure 2. Event Information in a concept element



The event specific details are expressed by an eventDetails structure plugged into the “concept” of a Concept Item or a Knowledge Item. The eventDetails used there are completely identical to the structure with the same name used for the “event” element in the content set of a News Item.

The Concept Identifier of an event may be used by other items (either News Items or Concept Items) to reference this event. On a purely technical level this Concept Identifier can be used for any “qcode” qualifier of a property. On a semantic level the only prerequisite are reasonable semantics of the property to reference an event – e.g. a property not limited to persons or locations by its semantics.

Examples are:

- ◆ Using an event's Concept Identifier as QCode for the “subject” property of a News Item. This indicates that the content of the News Item is about this event, the News Item's content may be a text, a photo, audio or video covering the event.
- ◆ Using an event's Concept Identifier with the “sameAs”, “broader”, “narrower” and “related” properties of another Concept Item. By these means a structure or network of events can be created, e.g. to link individual performances with a cultural festival or different talks to a conference.



3 Representing News - newsItem

An XML Schema file corresponding to the specifications for this item is available (see [The Full Set of Specification Documents](#) on page 3).

3.1 Description

A newsItem aims to convey news with the sense of the reporting of a newsworthy event or fact. Its content is gathered by journalists, presented with a journalistic style, and updated according to the progression of the story.

Examples of newsItems are a news report, a picture, a graphical illustration of some event, a video clip or an illustrated biography.

Typical characteristics of a newsItem are:

- ◆ Its content may be of any media type or format, e.g., the thumbnail, preview and high definition renditions of a picture.
- ◆ It can also convey more structured news information, e.g., information about companies, sports events and general events, in instances when this information is related to an event or fact.
- ◆ Its content is of short term interest: newsItems are volatile, and interest in them fades as time passes (“nothing is older than yesterday’s news”).
- ◆ It is expressed via a set of alternative renditions of some media content.
- ◆ It will usually be updated only for a short period of time, as long as the covered event evolves, and then may be archived.
- ◆ It refers to an arbitrary set of concepts and entities.
- ◆ It may be associated with other newsItems or Web resources via typed links.

3.2 Indication of Compliance with a Standard and Conformance Level

The IPTC newsItem *standard* attribute MUST be set to “NewsML-G2”.

The schema version to which the newsItem conforms MUST be indicated as an attribute. The current version is identified by the string “2.0”.

The IPTC conformance level to which the newsItem conforms in this specification MAY be indicated by the value “core”, which is the default value for this attribute.

Sample:

```
<newsItem standard="NewsML-G2" standardversion="2.0"
  xmlns="http://iptc.org/std/nar/2006-10-01/" >
</newsItem>
```

3.3 Identification and Versioning

It is possible to positively identify a newsItem as it moves through the news workflow and is transferred from place to place and from system to system.

A newsItem MUST have a *guid* attribute, which is a persistent and globally unique identifier. The guid is required to be in the form of an IRI. Any IRI capable of acting as a globally unique identifier is accepted.

Note: The IPTC will provide the newsml-URN for this purpose, specified by a successor of RFC-3085.

A newsItem MAY have a *version* attribute, and this version MUST be incremented when the content of the Item is updated. The first version MUST be numbered 1: if the version is not explicitly set, this value must be assumed by the recipient of the Item.

The *standardversion* attribute must reflect the version of the standard as it is implemented by the corresponding XML Schema.

Sample:



```
<newsItem standard="NewsML-G2" standardversion="2.0"
  guid="urn:newsml:iptc.org:20071231:sample" version="2"
  xmlns="http://iptc.org/std/nar/2006-10-01/" >
</newsItem>
```

3.4 Catalog of Controlled Vocabularies

NewsML-G2 recommends the use of controlled values for most properties. Each news provider is free to use their own taxonomies of subjects, genres, geopolitical areas, organisations etc., and to use any value scheme it decides in the Items it provides. A provider must therefore indicate the list of the schemes he is using.

Cataloguing information **MUST** be included at the top of each Item.

A catalog is defined as a set of scheme declarations in use by a news provider for a given Item.

Due to the large number of schemes potentially used in a single Item, and knowing that bandwidth is very important to the News industry, the catalog may be stored remotely.

A remoteCatalog **MUST** have a *href* attribute which contains the URI of a remote catalog. A remote catalog takes the form of an XML file with a catalog element as root.

The URI of a remote catalog acts both as a locator and a global identifier, therefore:

- ◆ The URI of a remote catalog **MUST NOT** be relative.
- ◆ If a remote catalog is functionally changed, the IRI used to access it **MUST** be changed. Functional changes are:
 - the addition or removal of a scheme declaration,
 - a change to a scheme alias,
 - a change to a scheme URI.

One or more additional title for a catalog or catalogRef **MAY** be provided in different languages and variants.

As some required properties take a QCode as a value, at least one catalog or remoteCatalog **MUST** be present.

In general, a given provider will define a unique catalog of all used schemes, store it in a central repository and reference it from all Items it provides. A provider **MAY** declare several catalogs in the same Item. This may be especially useful for an aggregator which uses property values from different sources, but requires a way to avoid scheme alias clashes. In this case, catalog and catalogRef elements **MAY** appear in any order, and their order is not relevant.

Detailed information on the structure of catalogs and their processing is given in [Dealing with Controlled Values](#) (page 33).

Sample:

```
<newsItem standard="NewsML-G2" standardversion="2.0"
  guid="urn:newsml:iptc.org:20071231:sample" version="2" xmlns="http://
  iptc.org/std/nar/2006-10-01/" >
  <catalogRef href="http://aprovider.com/cv/newsml-g2-catalog-4.xml"/>
  ....
</newsItem>
```

3.5 Rights Information

The content of a newsItem is bound to a set of copyrights and licensing information.

A rightsInfo wrapper element acts as a container for a set of properties related to rights, which offer a basic expression of the copyright and usage conditions associated with an Item.

This set is limited to an accountable person, a copyrightHolder and a set of copyrightNotices and usageTerms.

The order of the properties is flexible: The non-repeatable properties **MUST** come first, then the repeatable properties **MAY** be inserted in any order.

Each provider may add a set of metadata properties which have to be defined in a non-IPTC namespace. See also **XML Namespaces** (page 39) and **Extension Points in XML** (page 40).

3.6 Item Metadata

Such information is wrapped in the itemMeta wrapper element and split between news management metadata and Item links.

3.6.1 Management Metadata

Management metadata is bound to the Item as a whole and reflects its processing in a professional workflow.

The order of the properties in this set is imposed by the W3C XML schema.

Table 2. Item Management Group Elements

Element Title	Element Name	Card	Described on Page
Item Class	itemClass	(1)	114
Content Provider	provider	(1)	65
Date Item Version Created	versionCreated	(1)	77
Date Item First Created	firstCreated	(0..1)	76
Date Item Embargo Ends	embargoed	(0..1)	75
Publish Status	pubStatus	(0..1)	149
Role in the Workflow	role	(0..1)	158
File Name	filename	(0..1)	100
Editorial Service	service	(0..unbounded)	89
Item Title	title	(0..unbounded)	118
Editorial Note	edNote	(0..unbounded)	88

The IPTC provides a standardised scheme applicable to the itemClass property of a newsItem, identified by the URI <http://cv.iptc.org/newscodes/ninature/>.

Each provider may add a set of metadata properties which have to be defined in a non-IPTC namespace. See also **XML Namespaces** (page 39) and **Extension Points in XML** (page 40).

3.6.2 Processing the Publish Status of an Item

The IPTC makes these values normative for the exchange of Items between a provider and its customers:

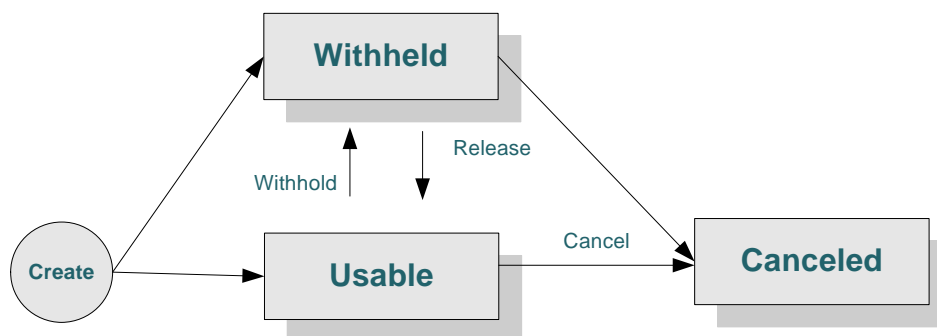
- ◆ Usable: The Item **MAY** be published without restriction.
- ◆ Withheld: Until further notice, the Item **MUST NOT** be published or used under any circumstances. If the Item has been published the publisher **MUST** take immediate action to withdraw or retract it.
- ◆ Canceled: The Item **MUST NOT** be published or used under any circumstances. If the Item has been published the publisher **MUST** take immediate action to withdraw or retract it.

Embargoes are managed by the embargoed property, which indicates the date of end of an embargo. This property is optional; if present, the date and time it contains must be compared with the current date and time before the Item is used. The Item must not be published before this time. The embargo overrides the usability of the Item, inferred from its status. If embargoed does not exist the item is not embargoed.

3.6.2.1 State Transition Diagram

This depicts the state transition diagram reflecting the ways in which the pubStatus values are intended to be used. Thus, upon creation of an Item, allowed statuses are usable and withheld. It is possible to withhold a usable document; it is possible to release a withheld document; it is possible to cancel a usable or withheld document. Once an Item has had its status set to canceled, it has reached a final state.

Figure 3. State Transition Diagram



3.6.2.2 Use Cases Associated with a Status of Withheld

Use Case 1: A provider distributes a story as a newsletter (version 1) with the status usable. At a later stage he learns that there may be a problem with the information included in the Item. He sends a new version of the newsletter (version 2) with a status set to withheld. All recipient systems must display a warning on this newsletter, and recipient publishers must postpone the publication of the information contained in the newsletter until further notice. The news provider has confirmation that the information is false and decides to set the status to canceled (version 3).

Use Case 2: An eCommerce system proposes a large collection of illustrated articles managed as news items. The publisher managing the system sees that the information included in a newsletter (version 1) is not up to date anymore, and decides to hide this Item from its customers until it is properly revised. He set then its status to withheld (version 2), edits the newsletter and set its status back to usable (version 3).

3.6.2.3 Processing Model on the Recipient Side

Here is the processing model on the recipient side and relies on the pubStatus and embargoed properties:

```

Test pubStatus = canceled:
    The Item must not be used, ever. Any usage of the Item must be prohibited, if needed by the way of alerts.
    Else: next
Test pubStatus = withheld:
    The Item must not be used until further notice. Any usage of the Item must be prohibited, if needed by the way of alerts.
    Else: next
Test pubStatus = usable:
    Test embargoed is not past: if yes, the live status is embargoed
    Else: The Item is usable and may be safely published
  
```

3.6.3 Processing of versionCreated

If the value provided by any date/time field does not conform to the appropriate syntax (e.g. format "YYYY-MM-DDTHH:MM:SS[+-]MM:SS") it MUST be considered as being not existent.

In the case of the mandatory versionCreated property the full Item MUST be considered as being void.



3.7 Item Links

A powerful feature of NewsML-G2 is the capability to associate Items via links. It is therefore possible to create a network of news resources, for management and navigation purposes.

The link element offers a generic mechanism for linking Items within the NAR framework as well as creating links from Items to other Web resources.

The semantic of the link *MAY* be refined via a relationship attribute (*rel*). In the absence of such indicator, the implied meaning of the link is “see also” (i.e. a navigation link).

The IPTC provides a recommended scheme of link relationships identified by the URI <http://cv.iptc.org/newscodes/relation/>.

If the target resource is an Item, the *guidref* attribute *SHOULD* identify the target Item by its globally unique identifier and a *version* attribute *MAY* indicate the target Item version; in the absence of version information, the target resource is the latest version available. The *href* attribute *MAY* additionally indicate the location of the target resource.

If the target resource is a Web resource, the *href* attribute *MUST* indicate the locator of the target resource.

The content type, a.k.a. IANA MIME type of the target resource *MAY* also be indicated by a *hreftype* attribute.

In order to ease the processing of a link, the size of the target resource *MAY* be added as an attribute. Such feature is especially useful if the target on the link is a Web resource. If the target resource is an Item, the size which is given here *MUST* be the size of the XML representation of the Item.

A title child element describing the link *MAY* be added for display to the users.

3.7.1 Processing Links

Link processing rules:

Link.1: Processor on the consumer side: If a guid and a version are provided, check whether the specific version of the Item is accessible using this information.

Processor on the provider side: If a guid and a version are provided deliver only the item version with the requested version number.

Link.2: Processor on the consumer side: If only a guid is available and no version, check whether an item is delivered by the provider. Consider a delivered version of the item as being the latest one.

Processor on the provider side: if only a guid is requested and not version, check if any version of the item exists, and if yes provide the one with the highest version number.

Link.3: Check whether the value of the *href* attribute allows some direct retrieval of the target resource via the Web (e.g. if the scheme is http: or ftp:), or an implicit resolution mechanism (e.g. DOI).

Link.4: Signal an error or ignore the link.

3.8 News Content Metadata

News Content Metadata is directly associated with the news information conveyed by the Item, independently of the processing of the Item in a professional workflow. Such information is wrapped in the *contentMeta* wrapper element and split between administrative and descriptive metadata.

3.8.1 Administrative Metadata

This is a set of properties associated with the administrative facet of content, i.e. data that cannot be inferred from “consuming” (reading, listening to, watching) the content.

All properties are optional. The order of the properties in this set is flexible: the non-repeatable properties MUST come first and then the repeatable properties may be inserted in any order.

Table 3. Administrative Metadata Group Elements

Element Title	Element Name	Card	Described on Page
Urgency	urgency	(0..1)	168
Date Content Created	contentCreated	(0..1)	73
Date Content Modified	contentModified	(0..1)	74
Located	located	(0..unbounded)	125
Information Source	infoSource	(0..unbounded)	113
Creator	creator	(0..unbounded)	72
Contributor	contributor	(0..unbounded)	67
Audience	audience	(0..unbounded)	47

3.8.1.1 Dates Processing Model

Two optional dates are associated with the content of an Item.

contentCreated and contentModified processing rules:

DatesGeneral.1: If the value provided by any date/time field does not conform to the appropriate syntax (e.g. format "YYYY-MM-DDTHH:MM:SS[+-]MM:SS") it MUST be considered as being not existent.

DateValues.1: If contentCreated is present it MUST NOT be later than versionCreated.

Error handling if it is later: at the creator's site an error alert should be issued, on the receiver's site it should be set to versionCreated.

DateValues.2: If contentModified is present contentCreated SHOULD be present as well.

In this case contentModified MUST NOT be earlier than contentCreated.

Error handling if it is earlier: at the creator's site an error alert should be issued, on the receiver's site it should be set to contentCreated

DateValues.3: If contentModified is present it MUST NOT be later than versionCreated.

Error handling if it is later: at the creator's site an error alert should be issued, on the receiver's site it should be set to versionCreated.

DateProcessing.1: The recipient processor MUST first check if a contentModified element is present.

DateProcessing.2: If not it MUST check if a contentCreated element is present.

DateProcessing.3: If not it SHOULD assume that the content was created at the time indicated by versionCreated element in itemMeta.

3.8.1.2 Audience Processing Model

Audience processing may be used to form ad hoc groups of recipients for which the Item is particularly significant or to filter out some users from the list of intended recipients of an Item.

The audience is expressed as a set of "positive" values (audience) and a set of "negative" values (exclAudience). The logic is to make the content easy to find to the audience identified by the positive values, but keep this content away from the audience identified by the negative values. An attribute of each property may indicate the expected significance of the content for this specific audience, and acts as a threshold for recipient filters.

The model for the audience processing is only a part of the overall filter that is used to determine whether a particular recipient is entitled to have access to the Item. It could be combined with the processing of other properties to further narrow the number of Items that match the recipient profile.

The processing rule has to be considered as a function which returns TRUE to indicate the recipient is entitled to receive the content, FALSE in case he is not entitled and NULL if the item does not contain any audience statements that apply to the Recipient.

Audience processing rules:

Audience.1: If any of the exclAudience properties applies to the recipient: return FALSE

Audience.2: If any of the audience properties applies to the recipient: return TRUE.

Audience.3: Return NULL.

3.8.2 Descriptive Metadata

This is a set of properties associated with the descriptive facet of news content, i.e. data that can be inferred from “consuming” (reading, listening to, watching) the news.

All properties are optional, repeatable and may be inserted in any order.

Table 4. Descriptive Metadata Group Elements

Element Title	Element Name	Card	Described on Page
Language	language	(0..unbounded)	120
Genre	genre	(0..unbounded)	101
Subject	subject	(0..unbounded)	164
Slugline	slugline	(0..unbounded)	162
Headline	headline	(0..unbounded)	107
Dateline	dateline	(0..unbounded)	83
By	by	(0..unbounded)	49
Description	description	(0..unbounded)	85

3.8.3 Other Content Metadata

Each provider may add a set of metadata properties which have to be defined in a non-IPTC namespace. See also [XML Namespaces](#) (page 39) and [Extension Points in XML](#) (page 40).

3.9 newsItem Content

Content may be included by value or by reference, and useful characteristics are represented along with such content, in order to facilitate its processing.

Alternative renditions of the news content, i.e. different technical representation of the same logical content, are wrapped by a contentSet wrapper element. Their order of appearance in contentSet is of no relevance. Their presence is optional: this allows for a lightweight and progressive representation of information.

Each rendition SHOULD be defined by a *rendition* attribute.

All alternative renditions SHOULD be derived from an original rendition by a software processor. For example: images in different resolutions, vector graphics and alternative bitmap images, text in different formats (e.g. NITF and PDF). The rendition from which all other renditions originate is indicated by the *original* attribute of contentSet; this attribute takes as a value the local identifier (id) of the original content component included in the contentSet.

They are three kinds of content components, inline XML, inline data and remote content:

- ◆ The inlineXML wrapper element holds XML content which is directly embedded in the element. The root element of this structure must be the root element of the language. Content may belong to any XML language capable of expressing generic or specialized news information, e.g. NITF, XHTML, SportsML or XBRL. The XML vocabulary is identified by a content type attribute (*contenttype*).



- ◆ The `inlineData` wrapper element holds plain-text or base64 encoded content. Plain text or CDATA content **MUST** be identified by the “text/plain” content type. Binary content, like images, audio clips or even PDF or Word documents may be exchanged after proper encoding, but it is strongly recommended to use this structure for small assets only. The encoding algorithm **MAY** be indicated using the *encoding* attribute. In the absence of this attribute, the content must be plain text, and the content type must be set accordingly. Encoding is constrained to base64 at this level of conformance.
- ◆ The `remoteContent` wrapper element may be used for representing any kind of media and data format. The data is stored independently of the `newsItem` and is referenced via a hyperlink (`href`). The size in bytes of the remote content **MAY** be indicated. The same rendition of content **MAY** be present at different remote locations. In such a case the same value of the *rendition* attribute **MAY** be given to several `remoteContent` elements.

The description of the content in each content component **MAY** be complemented by a content type (`contenttype`), a format acting as an optional refinement of the content type plus additional news content characteristics.

3.10 News Content Characteristics

News Content Characteristics are these physical properties of media content like the height and width of a picture, the word count of a news story or the duration of an audio clip, that help making selections among alternate renditions of news content.

The characteristics defined by the IPTC are a small and typical set of properties. Individual providers may add more characteristics they consider reasonable, i.e. audio data for professional broadcasting may require a different set from audio content for a podcast.

4 Representing Concept Information - concept Component

4.1 Concept Component

Concepts fall in two broad categories: named entities and generic (or abstract) concepts. Generic concepts range from themes (e.g. soccer) to emotions (e.g. smiling, love); they have no specific property defined, beyond generic properties. Named entities are people, organisations, geographical areas and points of interest for which a specific set of properties is defined for the purpose of a refined definition and improved search and processing capabilities.

The concept element provides a set of properties shared by all concepts.

A concept can be identified in different schemes by different controlled values, this is why a concept identifier must be unambiguous, but cannot be unique: for example, a company is commonly identified by different ticker symbols. In the case of abstract topics, the strict sameness of two concepts may be subject to discussion, and therefore a notion of equivalence of concepts is preferred.

A concept **MUST** have a concept identifier, expressed as `conceptId` child elements.

The `conceptId` element **MUST** have a `qcode` attribute. It **MAY** have a `created` attribute and a `retired` attribute which limit the usage of the concept identifier in time.

A concept **MAY** have a `type` child element. The type of a concept reflects its nature, e.g. generic, person, organisation, geopolitical area, point of interest etc...

A concept **MAY** have a name and **MAY** be further defined in natural-language by a definition and note. These labels are repeatable and **MAY** be specified in multiple languages.

More detailed facets of a concept (e.g. that the concept "is" an artist, listed company, city, restaurant) **MAY** be expressed by a specific facet property. The facet property **MUST** have a `rel` attribute which specifies the exact facet which is described.

A concept **MAY** have a set of alternative identifiers, expressed as `sameAs` child elements.

The `sameAs` element **MUST** have a `qcode` attribute.

4.2 Relationships Between Concepts

A concept **MAY** be associated with other concepts in the form of a taxonomy (i.e. a hierarchy of concepts) or thesaurus (i.e. a set of concepts associated via standard relationships). A concept **MAY** have an indication of most standard relationships as `broader`, `narrower` and `related` child elements.

The `broader`, `narrower` and `related` properties **MUST** have a `qcode` attribute which identifies a concept. They **MAY** have a `type` attribute which reflects the nature of the associated concept, and **MAY** have one or more names.

4.3 Details on Specific Entities

Details associated with specific entities **MAY** additionally be defined. All have been chosen for their potential usefulness in the news industry.

`personDetails` include a date of birth (`born`) and date of death (`died`) a repeatable indication of affiliation with an organisation and contact information (`contactInfo`).

`organisationDetails` include a date of foundation (`founded`) and date of dissolution (`dissolved`), a repeatable location and contact information (`contactInfo`).

The registered address of an organisation is indicated as part of its contact information; in such a case this address may not be used for making direct contact with this company.

`geoAreaDetails` include the geographic coordinates (`position`) of the place.

The `position` **MUST** have `latitude` and `longitude` attributes. It **MAY** have an indication of the altitude above the zero elevation reference level.



It MAY have an indication of coordinate reference system (*gpsdatum* attribute) expressed as a string. In the absence of this attribute, the WGS84 system is assumed.

POIDetails include the geographic coordinates (position) of the place, plus practical information like opening hours (*openHours*), capacity, access information, details on the location (e.g. room number, stair number etc.), and contact information (*contactInfo*).

4.3.1 Contact Information

Contact information (*contactInfo*) is repeatable in the definition of a person, an organisation and a Point of Interest, and each set of properties supports a *role* attribute which makes possible to group together all information belonging of the same nature.

Contact information include email addresses, instant messaging addresses (*im*), international phone numbers, international fax numbers, web addresses and postal addresses.

E-mail and instant messaging addresses, phone and fax number are all electronic addresses. These are qualified by a *role* attribute which specifies the nature of the address, e.g. home or work.

4.3.2 Postal Address

The definition of a Postal Address includes free-text lines (in the format expected by a recipient postal service), the indication of a locality (i.e. city, town, village etc...), a subdivision of a country (area), a country and a postal code (*postalCode*).

A postal address is structured as a set of properties likely edited and displayed as a form. The relative order of its properties is not universal, and if used for traditional postal mail, presentation algorithms are to be developed which depend on the source and recipient countries.

The city, country area and country may be indicated as a name or as a controlled value. The use of an ISO compliant country code is particularly recommended.

5 Managing Individual Concepts - conceptItem

An XML Schema file corresponding to the specifications for this item is available (see [The Full Set of Specification Documents](#) on page 3).

5.1 Description

A conceptItem aims to convey knowledge about a unique concept (a named entity such as an organisation or an abstract notion such as a news subject). Typically a conceptItem itself holds only rather short and structured information about the concept and about its relationships with other concepts.

Typical characteristics of a conceptItem are:

- ◆ It focuses on a single concept or entity.
- ◆ It will usually be updated infrequently but over a long period of time, when the information about the concept evolves.
- ◆ Its content is of long term interest.
- ◆ It may be referenced by other items.

Different conceptItems, managed by different providers, may contain structured information about the same concept.

5.2 Structure of a conceptItem

The model of a conceptItem is very similar to the model of a newsItem. Both share the same Indication of compliance with a standard and conformance level, Identification and versioning, Signature, Rights Information, Item Metadata, Item links. Please review the corresponding specification of a newsItem for more information.

Note that the globally unique guid of a conceptItem, which is used for management purposes, must not be confused with the unambiguous concept identifier (conceptId) defined by an authority and conveyed as part of the content of the Item.

5.3 Item Metadata

The IPTC provides a standardised scheme applicable to the itemClass property, identified by the URI: <http://cv.iptc.org/newscodes/cinature/>.

5.4 Concept related Metadata

The set of administrative metadata is common to all classes of Items. Please review [Representing News - newsItem](#) (page 19) for more information.

A conceptItem does not support descriptive metadata.

5.5 conceptItem Content

The content of a conceptItem is a concept component.



6 Managing Sets of Concepts - knowledgeltem

An XML Schema file corresponding to the specifications for this item is available (see [The Full Set of Specification Documents](#) on page 3).

6.1 Description

A knowledgeltem bundles a set of concept components which are managed and exchanged as a whole. A knowledgeltem is used best where a provider wants to circulate a snapshot of a set of entries from one or more controlled vocabularies.

The concepts represented in a knowledgeltem can be of different types, and their identifiers may come from different schemes. A “scheme definition” is therefore a particular case of structure, where all concepts support a concept identifier from a same specific scheme.

Examples of knowledgeltems are the taxonomy of IPTC Subject NewsCodes or an authority list of people’s descriptions maintained by a given provider.

Typical characteristics of a knowledgeltem are:

- ◆ It contains a meaningful set of concepts components.
- ◆ It will usually be updated infrequently but over a long period of time, for example when a controlled vocabulary evolves.
- ◆ Its content is of long term interest.

6.2 Structure of a knowledgeltem

The model of a knowledgeltem is very similar to the model of a newsItem. Both share the same Indication of compliance with a standard and conformance level, Identification and Versioning, Signature, Rights Information, Item Metadata, Item links. Please review [Representing News - newsItem](#) (page 19) for more information.

6.3 Item Metadata

The IPTC provides a standardised scheme applicable to the itemClass property, identified by the URI <http://cv.iptc.org/newscodes/cinature/>.

6.4 Knowledge Related Metadata

The set of administrative metadata is common to all classes of Items. Please review the corresponding specification of a newsItem for more information.

The set of descriptive metadata is listed below. All properties are optional, repeatable and may be inserted in any order.

Table 5. Knowledge Descriptive Metadata Group Elements

Element Title	Element Name	Card	Described on Page
Subject	subject	(0..unbounded)	164
Description	description	(0..unbounded)	85

Each provider may add a set of metadata properties which have to be defined in a non-IPTC namespace. See also [XML Namespaces](#) (page 39) and [Extension Points in XML](#) (page 40).

6.5 knowledgeltem Content

A conceptSet wrapper element contains a set of concept components. Their order of appearance in conceptSet is not relevant.

Note: All concept definitions share the same catalog of schemes, declared at the top of the knowledgeltem.

