IPTC Standards

NewsML™ G2

(version 2.21)

Quick Start Guide
to Conveying Text

Quick Start: Text

Introduction

One of the most fundamental needs of a news organisation is to handle text. This Quick Start Guide covers the basics of a simple NewsML-G2 News Item containing text content.

We recommend reading the Quick Start Guide to G2 Basics
before this Quick Start Guide to Text.

Example

Below is an example story and supporting information as might be displayed on the journalist's editing screen at a fictional news provider, Acme News and Media (ANM):

Acme News and Media – Content Editing System		
Slugline	US-Finance-Fed	
Created on	2015-11-21 15:21:06	
Source	ANM	
Author	mjameson	
Latest edit	2015-11-21 16:22:45	
Latest editor	moiras	
Categories	economy, finance, business, central bank, monetary policy	
Headline	Fed to halt QE to avert "bubble"	
Byline	By Meredith Jameson	
Location / Date	Washington	21/11/2015
Body Text	Et, sent luptat luptat, commy nim zzriureet vendreetue modo dolenis ex euisis nosto et lan ullandit lum doloreet vulla feugiam coreet, cons eleniam il ute facin veril et aliquis ad minis et lor sum del iriure dit la feugiamcommy nostrud min ullapat velisl duisismodip ero dipit nit utpatum sandrer cipisim nit lortis augiat nulla faccum at am, quam velenis nulput la auguerostrud magna commolore eliquatie exerate facilis modiamconsed dion henisse quipit at. Ut la feu facilla feu faccumsan ecte modoloreet ad ex el utat.
.	

This screen contains nearly all of the information needed to create a valid NewsML-G2 document.

Document structure

The building blocks of the text document are the <newsItem> root element, with additional wrapping elements for metadata about the News Item (itemMeta), metadata about the content (contentMeta) and the content itself (contentSet). The top level (root) element <newsItem> attributes are:

```
<newsItem xmlns="http://iptc.org/std/nar/2006-10-01/"
   guid="urn:newsml:acmenews.com:20151121:US-FINANCE-FED"
   version="8"
   standard="NewsML-G2"
   standardversion="2.21"
   conformance="power"
   xml:lang="en-US">
```

This is followed by references to the Catalogs used to resolve QCodes in the Item, and Rights information:

```
<catalogRef
   href="http://www.iptc.org/std/catalog/catalog.IPTC-G2-Standards_27.xml" />
<catalogRef
   href="http://catalog.acmenews.com/news/ANM_G2_CODES_2.xml" />
<rightsInfo>
   <copyrightHolder uri="http://www.acmenews.com/about.html#copyright" >
      <name>Acme News and Media LLC</name>
   </copyrightHolder>
   <copyrightNotice>Copyright 2015 Acme News and Media LLC</copyrightNotice>
</rightsInfo>
```

Item Metadata <itemMeta>

Note the three mandatory child elements of the mandatory<itemMeta>:

- ❖ Item Class
- ❖ Provider
- ❖ Version Created

In addition, a publication status is also mandatory, but the <pubStatus> element that expresses it may be omitted if the publication status is "usable", the default. It is recommended that the publication status is explicitly carried as in this example. As Acme News & Media is fictional, the Provider property does not use one of the IPTV Provider NewsCodes, and so it expressed using a URI:

```
<itemMeta>
    <itemClass qcode="ninat:text" />
    <provider uri="http://www.acmenews.com/about.html" />
    <versionCreated>2015-11-21T16:25:32-05:00</versionCreated>
    <pubStatus qcode="stat:usable" />
</itemMeta>
```

Content Metadata <contentMeta>

Administrative Metadata

The administrative properties of the example text story are:

```
<contentCreated>2015-11-21T15:21:06-05:00</contentCreated>
<contentModified>2015-11-21T16:22:45-05:00</contentModified>
```

The place that the content was created uses the <located> element:

```
<located qcode="geoloc:NYC">
    <name>New York, NY</name>
</located>
```

(Note that this is where the story was written, not the place where the subject of the story took place. That would be expressed using <subject>, part of Descriptive Metadata.)

The author of the article is expressed using the <creator> element:

```
<creator uri="http://www.acmenews.com/staff/mjameson">
    <name>Meredith Jameson</name>
</creator>
```

The Information Source for the article is also given. When used without a @role, <infoSource> is used to denote the person or party that provided the original information on which the content is based. This is the relationship to be expressed here:

```
<infoSource qcode="is:AP">
    <name>Associated Press</name>
</infoSource>
```

The default language for the content is given as U.S. English:

```
<language tag="en-US" />
```

Descriptive Metadata

In the example, the Subject properties use QCodes from the Controlled Vocabulary of Media Topics NewsCodes that are owned and maintained by the IPTC and expressed as QCodes. Thus:

```
<subject qcode="medtop:04000000">
    <name>economy, business and finance</name>
</subject>
<subject qcode="medtop:20000350">
    <name>central bank</name>
</subject>
<subject qcode="medtop:20000379">
    <name>money and monetary policy</name>
</subject>
```

The <slugline> property contains the value of the "Slugline" field of the story:

```
<slugline>US-Finance-Fed</slugline>
```

In a similar fashion, the <headline> property will contain the value of the "Headline" field:

```
<headline>Fed to halt QE to avert "bubble"</headline>
```

Complete Content Metadata

```
<contentMeta>
    <contentCreated>2015-11-21T15:21:06-05:00</contentCreated>
    <contentModified>2015-11-21T16:22:45-05:00</contentModified>
    <located qcode="geoloc:NYC">
        <name>New York, NY</name>
    </located>
    <creator uri="http://www.acmenews.com/staff/mjameson">
        <name>Meredith Jameson</name>
    </creator>
    <infoSource qcode="is:AP">
        <name>Associated Press</name>
    </infoSource>
    <language tag="en-US" />
    <subject qcode="medtop:04000000">
        <name>economy, business and finance</name>
    </subject>
    <subject qcode="medtop:20000350">
        <name>central bank</name>
    </subject>
    <subject qcode="medtop:20000379">
        <name>money and monetary policy</name>
    </subject>
    <slugline>US-Finance-Fed</slugline>
    <headline>Fed to halt QE to avert "bubble"</headline>
</contentMeta>
```

Text content choices

Inline XML

The content of the NewsML-G2 document is enclosed by the <contentSet> wrapper. In the example, the IPTC news mark-up language NITF (News Industry Text Format) is used to format the text content. As an XML standard, it is contained in an <inlineXML> child element of <contentSet>, and uses @contenttype to denote the XML-based standard, using the IANA MIME type.

XHTML is also a popular text mark-up choice among G2 providers. As alternatives, the contents of <inlineXML> may be any XML language that can express generic or specialised news information, including SportsML-G2 and EventsML-G2. Other languages such as XBRL (Extended Business Reporting Language) may also be used. The content inside <inlineXML> must be valid XML, in other words, it could stand alone as a valid XML document in its own namespace.

```
<contentSet>
    <inlineXML contenttype="application/nitf+xml">
        <nitf xmlns="http://iptc.org/std/NITF/2006-10-18/">
            <!--STORY CONTENT HERE -->
        </nitf>
    </inlineXML>
</contentSet>
```

Inline data

The <inlineData> element can contain plain text, and in this case MUST be identified by the IANA MIME type of "text/plain" thus:

```
<contentSet>
    <inlineData contenttype="text/plain">
        Et, sent luptat luptat ...
    </inlineData>
</contentSet>
```

Putting it together

Code Listing: NewsML-G2 Text Document

All QCodes used in listing below are from IPTC NewsCodes vocabularies, except for the following values: *geoloc, is,*

```xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<newsItem xmlns="http://iptc.org/std/nar/2006-10-01/"
  guid="urn:newsml:acmenews.com:20151121:US-FINANCE-FED"
  version="8"
  standard="NewsML-G2"
  standardversion="2.21"
  xml:lang="en-US">
  <catalogRef
    href="http://www.iptc.org/std/catalog/catalog.IPTC-G2-Standards_27.xml" />
  <catalogRef
    href="http://catalog.acmenews.com/news/ANM_G2_CODES_2.xml" />
  <rightsInfo>
    <copyrightHolder uri="http://www.acmenews.com/about.html#copyright">
      <name>Acme News and Media LLC</name>
    </copyrightHolder>
    <copyrightNotice>Copyright 2015 Acme News and Media LLC</copyrightNotice>
  </rightsInfo>
  <itemMeta>
    <itemClass qcode="ninat:text" />
    <provider uri="http://www.acmenews.com/about/" />
    <versionCreated>2015-11-21T16:25:32-05:00</versionCreated>
    <pubStatus qcode="stat:usable" />
  </itemMeta>
  <contentMeta>
    <contentCreated>2015-11-21T15:21:06-05:00</contentCreated>
    <contentModified>2015-11-21T16:22:45-05:00</contentModified>
    <located qcode="geoloc:NYC">
      <name>New York, NY</name>
    </located>
    <creator uri="http://www.acmenews.com/staff/mjameson">
      <name>Meredith Jameson</name>
    </creator>
    <infoSource qcode="is:AP">
      <name>Associated Press</name>
    </infoSource>
    <language tag="en-US" />
    <subject qcode="medtop:04000000">
      <name>economy, business and finance</name>
    </subject>
    <subject qcode="medtop:20000350">
      <name>central bank</name>
    </subject>
    <subject qcode="medtop:20000379">
      <name>money and monetary policy</name>
    </subject>
    <slugline>US-Finance-Fed</slugline>
    <headline>Fed to halt QE to avert "bubble"</headline>
  </contentMeta>
  <contentSet>
    <inlineXML contenttype="application/nitf+xml">
      <nitf xmlns="http://iptc.org/std/NITF/2006-10-18/">
        <body>
          <body.head>
            <hedline>
              <hl1>Fed to halt QE to avert "bubble"</hl1>
            </hedline>
            <byline>By Meredith Jameson, <byttl>Staff Reporter</byttl></byline>
          </body.head>
          <body.content>
            <p>(New York, NY - November 21) Et, sent luptat luptat, commy Nim
              zzriureet vendreetue modo dolenis ex euisis nosto et lan ullandit
              lum doloreet vulla. </p>
            <p>Ugiating ea feugait utat, venim velent nim quis nulluptat num
              Volorem inci enim dolobor eetuer ercin utpatio dolorpercing.</p>
          </body.content>
        </body>
      </nitf>
    </inlineXML>
  </contentSet>
</newsItem>
```

NewsML-G2 Implementation Guidelines and Specification

For more comprehensive information about G2 implementation than is covered by these Quick Guides, the full Guidelines for NewsML-G2 Implementers is available for download by visiting the IPTC website https://iptc.org/standards/ and following the link to NewsML-G2.

This includes more detailed "How To" topics to help implementers with more complex needs, and also covers subjects such as creating and managing Catalogs and Controlled Vocabularies, conveying multiple G2 Items in News Messages, and using Planning and EventsML-G2 for news management and fulfilment.

The IPTC website also contains a link to the complete technical description of all G2 properties and help with processing models documented in the NewsML-G2 Specification.

Terms of Use

If you have any questions about the terms, please contact the managing director of the International Press Telecommunication Council. Contact details of the IPTC are listed below.

While every care has been taken in creating this document, it is not warranted to be error-free, and is subject to change without notice. Check for the latest version of this Document and applicable G2 Standards and Documentation by visiting https://iptc.org/standards/ and following the link to NewsML-G2. The version of NewsML-G2 covered by this document is 2.21.

Contacting the IPTC

IPTC, International Press Telecommunications Council
Web address: www.iptc.org
Follow us on Twitter: @IPTC
Email: office@iptc.org
Business address
25 Southampton Buildings
London WC2A 1AL
United Kingdom

The company is registered in England at 10 Portland Business Centre, Datchet, Slough, Berks, SL3 9EG
as Comité International des Télécommunications de Presse
Registration No. 1010968, Limited by Guarantee, Not Registered for VAT