



IPTC SPRING MEETING 2005

held at the
Holiday in on the Bay hotel
1355 Harbour Drive, San Diego, CA 92101, USA

3rd to 5th March 2005

Minutes of the News Content Working Party held on 4th March 2005

Document history				Document URN: urn:iptc:workdoc:nct:0502:1
Revision	Issue Date	Pages	Author (revised by)	Remark
unrevised	2005-03-19	6+3	Michael Steidl/	

Present:

Chairman: Henrik Stadler, Tidningarnas Telegrambyrå

Walter Baranger, New York Times Company

Scott Calder, Mainstream Data

Guthrie Collin, The Associated Press

Dave Compton, Reuters Limited

Jay Cousins, RivCom

Honor Craig-Bennett, PA NewsLtd

Arnaud Descamps, Relaxnews

Takahiro Fujiwara, EAST Co. Ltd

Stéphane Guérillot, Agence France Presse

Darko Gulija, HINA

Klaus Herwig, Deutsche Presse-Agentur GmbH

Niels Hojer-Pedersen, Ritzau Bureau I's

Rudi Horvath, Austria Presse Agentur

John Iobst, Newspaper Association of America

Hugh Johnstone, IPTC

Alan Karben, XML Team Solutions, Inc.

Dean Large, Business Wire

Michael Steidl, IPTC Managing Director as Secretary

Laurent Le Meur, Agence France Presse

Jaak Leenknecht, Agence de Presse Belga

Johan Lindgren, Tidningarnas Telegrambyrå

Harald Löffler, ifra

Jayson Lorenzen, Business Wire

Angelo Marrara, ANSA

John Minting, United Press International

Peter Müller, SDA/ATS

Stuart Myles, Dow Jones & Company

Karl Oanes, CCMMatthews

Jean-François Richard, Agence France Presse

Hitoshi Saito, Nihon Shinbun Kyokai

Hiroshi Shinotsuka, Kyodo News

Klaus Sprick, Deutsche Presse-Agentur GmbH

Charles Tichenor, The Associated Press

Misha Wolf, Reuters Limited

1 Chairman's introduction to the scope of the new Working Party

Henrik Stadler said as Geoffrey Haynes was not available to chair this meeting he stands in.

Then the Chairman emphasised this is the first meeting of the News Content Working Party after the Standards Committee reshaped the working structure of the IPTC.

The focus of this working party is now on any type of content and its markup. So this is a step beyond the former scope of "specialised content" to embrace content at a generic level and hence a "NewsML WG" was added to the WP.

The charter of this working party from the STA0409c paper on the new working structure is: "The News Content WP responsibilities are extended as it is required to provide full comprehensive requirements to the Common Components WG and take into account the components and related processing and holding structures as made available within the set of Common Components."



2 Minutes of the SPC Meeting held on 6th October 2004 (SPC0407)

It was proposed by John Iobst and seconded by Walter Baranger that the minutes be accepted as circulated. This was agreed unanimously.

3 Matters Arising

No matters were arising from the minutes not covered by this agenda.

4 EventsML Working Group (Johan Lindgren)

The WG chair said the EventsML developer groups is currently in some idle state: The group is waiting for the basic framework of the News Architecture and will then proceed with creating EventsML by using and implementing the building blocks.

The groups has set up a “iptc-eventsml-dev” Yahoo group for the internal communication, the existing “eventsml” group is maintained for public communication.

5 SportsML Working Group (Alan Karben)

5.1 Marketing update

* Yahoo!Group membership up to 340, with new inquiries such as:

- San Diego Padres

The local baseball team hereabouts

“We were just discussing an intermediate standard data format for all the information we'll be getting, and XML came up. I'll take a look at SportsML....”

- U.S. College Athletic Association

A “mini” NCAA

“Put together what might be an XML file for a single score. Is there any chance I can get feedback on it?”

* SportsML.org replaces “SportsML.com” as the primary domain for SportsML

5.2 XML Schema improvements

The goal is to improve the Controlled Vocabulary listings

- More sample values for “pick lists” should be provided.

- Better string-formatting specifications, for percentages, etc.

Schema Style Synchronization

Work with IPTC XML Schema Style Taskforce

5.3 Auto Racing Plugin

A new sport-specific plugin DTD/Schema (joining the current 7)

Steady demand reported by U.S sports data publishers; occasional murmurings on SportsML Yahoo!Group

To Be Determined: How similar / different are the structures for Auto Racing results & stats between the U.S. and European countries and other Auto Racing hotspots around the world?

DISCUSSION:

It was discussed what types of auto racing is covered by this plugin (e.g. Auto Rallye as well?) and possibilities to extend it to motorbike racing.



Licensing:

Alan Karben and Johan Lindgren reported an inquiry on the licensing terms for SportsML.

The Managing Director said this issue is on the agenda of the Management Committee meeting of the same day, the directors are aware of that.

Misha Wolf added a most interesting issue was raised: plugins can be added to SportsML – who owns these plugins as they are not part of the core specifications.

5.4 Upcoming issues

SportsML and the News Architecture

“What’s in it for SportsML?”

Quick SportsML overview:

```
<sports-content>
<sports-metadata>
...
</sports-metadata>
(
<statistics> OR <schedule> OR <standing>
OR <tournament> OR <article>
)*
</sports-content>
```

Sports metadata

```
<sports-metadata
  date-time="20050224T103900-0500"
  doc-id="xt.2634150-NBA-INJURIES"
  revision-id="xt.2634032-NBA-INJURIES"
  fixture-key="injuries"
  fixture-name="Injury Report"
  document-class="news"
  >
  <sports-title>NBA Injuries</sports-title>
  <sports-content-codes>
  ...
  </sports-content-codes>
</sports-metadata>
```

Sports Content Codes

```
<sports-content-codes>
  <sports-content-code code-name="XML Team"
    code-key="xmlteam.com" code-type="distributor" />
  <sports-content-code code-type="sport"
    code-key="15008000" code-name="Basketball" />
  <sports-content-code code-type="league"
    code-key="l.nba.com" code-name="National Basketball Association" />
  <sports-content-code code-type="priority" code-key="normal" />
</sports-content-codes>
```

Sports Metadata within NAR

Prime candidate for either:

- * Replication up into metadata common component (intermediate-term option)



* Replacement by
metadata common component
(long-term possibility)

SportsML <article> within NAR

Right now: <article> generally embeds an NITF article

In the future, one might envision an NAR wrapper containing an article just like NewsML documents would

Case Study: “Game Recap” document, which includes:

- * Article reporting on the game
- * <sports-event> containing the linescore

Case Study: Game Package:

- Case Study: Following many games, clients often wind up with:
 - ◆ Final Linescore
 - ◆ Final Box Score (player stats)
 - ◆ Game Writeup & Preview (editorially written articles)
 - ◆ Game Photos / Audio / Video
 - ◆ Etc.
- Clients often like getting these one at a time, as soon as each are available
 - ◆ All content pieces share a common “event-key”
- Some clients may want to just pick up a Game Package
 - ◆ Using NAR to relate all these items

DISCUSSION:

Laurent Le Meur said a first draft of the News Architecture Structure will be available by June 2005.

5.5 Next steps

- * Complete the Auto Racing plug-in
- * Revise and Re-issue Schema Draft
- * Ready SportsML 1.6 DTD and Schema for vote at next meeting

6 NewsML Working Group (Laurent Le Meur)

The WG chair made a presentation on the outline of the work for NewsML 2 - see slides 1 through 14 as Appendix 1 to these minutes.

DISCUSSION:

Re slide #8: (Content)

Arnaud Descamps: how to markup the Creator?

The WG chair said this will be part of a common component.

Misha Wolf asked how to include a non-IPTC content like XBRL into NewsML.



The WG chair said as this appears to be specific content a specific item type should be created for that purpose as currently no generic “anyXML”-wrapper is available.

Misha Wolf said he considers this to be an omission.

Darko Gulija proposed to think about adding an XBRL content component to the newsItem holding e.g. the article about a company.

The WG chair said he considers this to be a “back to NewsML1” approach – as NewsML2 will not allow packaging various content components inside a newsItem.

Stuart Myles said Dow Jones issues company reports having a text and a table in many cases and he considers this to be essentially the same information only in various formats.

Misha Wolf said XBRL is becoming more and more import for providing formalised information about a company.

Stuart Myles said he is concerned about treating XBRL as a completely different item – he emphasised to consider this formalised information only as a different rendition.

Dean Large said including XBRL as “alternate rendition” into an newsItem is a big deal as this has to be validated there and this could be tricky.

Dave Compton put emphasis on the packaging approach. It would be much easier to use the packaging mechanism to build complex content structure and the packaging item would allow telling explicitly if an XBRL item is an alternative to a newsItem.

Misha Wolf added also he is in favour of using packaging for this as e.g. a story would come from a reporter while the XBRL would come from an analyst and hence this is no alternative rendition anymore. Solutions could be using a generic wrapper and having a more generic item that allows to include any non-IPTC content.

Darko Gulija said all IPTC standards deal with news content what makes the item different is primarily the set of metadata assigned to the content to describe the content properly.

Misha Wolf asked which Working Group or Party will deal with “headlines” and “alert” items.

The WG chair said this must be discussed as he thinks these types are very specific and the issue is currently only raised by Reuters.

Darko Gulija said if this is a very specific set of content and metadata we should think about an “AlertML”.

Stuart Myles asked Misha Wolf if the alert item knows already about the type of the item which will follow and will contain content: if the identifier should be maintained for making a full featured newsItem out of the alert item this won’t work with the current news management outline.

Dave Compton said it must be possible to be reference these items as a chain.

Walter Baranger pointed at the short lifetime of an alert: it only indicates some news story will follow shortly – and has no real meaning after the story was sent.

Misha Wolf said the management issue is: how to indicate this news story replaces the alert.

Darko Gulija said the EventsML and SportsML groups should be asked what they think about alerts.

Alan Karben said in the scope of SportsML no special structure is required for an alert as only the required metadata are conveyed and assigning a higher priority to the item would indicate something urgent is coming up.



Johan Lindgren added we should evaluate if there is a difference between having a special alert item and the issue of how to alert people on the receiver side of a news feed.

The WG chair said he discussed the issue of the “alert chain” with Dave Compton more in detail (in a break) and he feels Reuters’ needs can be covered by IPTC news management facilities.

Dave Compton added: currently Reuters uses two IDs: a usage Id which is persistent from the stage of sending alert until the stage of sending full stories – while the management Ids for the alert and the full story are different.

The WG chair said we should look into having Ids which are not an identifier of the managed item but for e.g. a chain of items.

Stuart Myles agreed to consider this within the scope of the News Management Working Group.

The WG chair asked who else but Reuters has the requirement for sending a “Headline” structure to clients.

Stuart Myles responded saying Dow Jones sends gradual updates: a headline first, then a first paragraph, a second ... and so on.

Then Stuart Myles asked about the difference between an alert and a headline item.

Misha Wolf replied a headline is extended to a full story – while alerts are not – within the scope of the same identifier.

What they have in common: they have to be extremely compact.

7 Weather markup – update

John Iobst said not much work was done on that since the last meeting.

Alan Karben pointed at the availability of the domain “WeatherML.org”

8 Update on content markup standards

Misha Wolf said how to include several non-IPTC standards – including XBRL – into the News Architecture should be considered by the IPTC.


9 Any Other Business

There was no other business.

10 Date and Place of Next Meeting

6 – 9 June 2005, London, Bonnington Hotel in Bloomsbury


continued: Document Appendix 1 >>> (+ 3 pages)



NewsML WG

General News Content Introduction


Laurent Le Meur (AFP)
NewsML WG chairman, IPTC
San Diego / March 04, 2005



Charter & forum

- Charter: Developing a model for the representation of general news content.
- Dev forum at <http://groups.yahoo.com/group/iptc-newsml-dev>
- Keeps features from NewsML 1.x
- Timeline: will use News Architecture

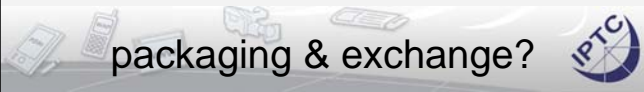
© IPTC – www.iptc.org 2



How to create NewsML 2

- The NewsML WG works on “general news content” (free format journalistic information)
- NewsML WG is part of the News Content WP
- NewsML2 will use the NAR
- NewsML WG members are members of the NAR development (->Laurent, Stuart ...)
- NewsML should define one item class :
 - NewsItem
- The NewsML WG
 - Chooses components from the common library
 - Creates specific components


© IPTC – www.iptc.org 3



packaging & exchange?

- The NewsML WG does not work on packaging & exchange anymore; packaging & exchanging news are ‘cross-content’ features
- It’s up to the IPTC to define in which standard Package & Message classes should be defined.


© IPTC – www.iptc.org 4



A NewsItem

Item	Management Component	Required, same for all IPTC standards
	Metadata Construct	Chosen in the common library
	Metadata Construct	description/labels/rights/pub
	News Metadata Construct	Specific to the domain of general news (?)
	News Content Component	Alternate content parts, specific to the domain of general news
	News Content Component	

© IPTC – www.iptc.org 5



Content

- Characteristics (physical/technical)
- XOR
 - Reference to an external resource
 - Plain text
 - XML (usually NITF or XHTML)
 - Binary
- *A personal opinion*: Content should be self-sufficient (should include title, byline, dateline) and not rely on the presence of external labels to be complete.

© IPTC – www.iptc.org 6

Simple NewsItem

- News metadata
 - Management properties (itemVariant = Text, Photo, Video, Audio)
 - Descriptive metadata
 - Rights metadata
 - Signature
- Set of alternative renditions of the “same” content
- Examples: news alert, news story, captioned picture, video clip

IPTC

© IPTC – www.iptc.org 7

Compound content

- Has content (any media type)
- Is not complete if some other items are not available at the same time
- Example : a multimedia story; text story, illustrated with pictures & graphics;
 - Illustrations have their own metadata, their own “life” = they are news items by themselves
 - They are “imported” in the compound news item


IPTC

© IPTC – www.iptc.org 8

Sample: compound doc

High Court Ends Death Penalty for Youths
By HOPE YEN, Associated Press Writer

WASHINGTON—A closely divided Supreme Court ruled Tuesday that it's unconstitutional to execute juvenile killers, ending a practice in 19 states that has been roundly condemned by many of America's closest allies.




AP Photo

The 5-4 decision throws out the death sentences of 72 murderers who were under 18 when they committed their crimes and bars states from seeking to execute minors for future crimes.

The executions, the court said, violate the Eighth Amendment ban on cruel and unusual punishment.

“The age of 18 is the point where society draws the line for many purposes between childhood and adulthood. It is, we conclude, the age at which the line for death eligibility ought to rest.” Justice Anthony Kennedy ([news](#) · [web site](#)) wrote.

The ruling continues the court's practice of narrowing the scope of the death penalty, which justices reinstated in 1976. Executions for those 15 and younger when they committed their crimes were outlawed in 1988. Three years ago justices banned death sentences for the mentally retarded.



AP Photo

High Court Ends Death Penalty for Minors

IPTC

© IPTC – www.iptc.org 9

Compound vs Composite

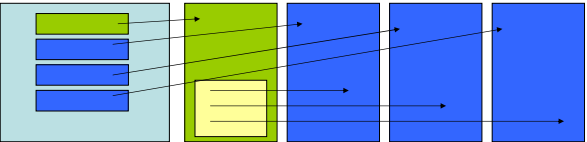
- Currently “composite” = a content component with several media types in it
- E.g. a pdf file

IPTC

© IPTC – www.iptc.org 10

Compound vs package

- A PackageItem may glue together independent items
- A compound news item includes dependencies = associations with other required items.



IPTC

© IPTC – www.iptc.org 11

Specific metadata values

- Management
 - itemClass = ‘News’
 - mediaType/itemVariant = ‘Text’ | ‘Photo’ | ‘Graphic’ | ‘Audio’ | ‘Video’ | ‘Composite’
 - structureType/representation = ‘Alert’ | ‘Headline’ | ‘Article’ ...
- Description
 - subject, audience, **language, genre, date of origin, location of origin, scene + named entity references** (person, location, event)
- Association with other items
 - For navigation purpose (news item, sports item, event item)
 - **Imported news items in compound news items (*)**

IPTC

© IPTC – www.iptc.org 12

Conformance levels



- Basic level
 - Follows general rules for management, structure and common components
 - maybe
 - Characteristics restrained to an IPTC defined set
 - XML content nif and xhtml only
 - Binary content base64 only
- Power level
 - No restriction on the characteristics (namespace)?
 - No restriction on the XML content (namespace)?
 - No restriction on the encoding used for binary content?

Thank you for your time



Check: <http://www.newsml.org>

Dev : [iptc-newsml-dev](#) on Yahoo!