



IPTC Mirror

Planned Releases

It had been hoped to give formal approval to the first of the G2-Standards - NewsML-G2 and EventsML-G2 - during the Autumn Meeting (held in Prague during October). As IPTC Chairman Stéphane Guérillot pointed out, it is considered important to ensure that the new standards are released in a fully professional way.

In particular this means that full documentation has to be available, and producing this is proving time consuming so the release has been delayed. However, technical development of the standards is substantially complete, with the Model specification and associated XML Schema files for each of the standards having been adopted.

Work is continuing on the documentation and it is hoped that complete packages for both standards will be available for consideration during January. It is then intended to hold a special teleconference meeting of the Standards Committee to approve the standards for release.

At the same time work is under way on new versions of both the NITF and SportsML, which will offer G2-Standards compatibility.

The News Architecture (NAR) is the foundation for the G2-Standards family with all the standards sharing common NAR features, while the information payload is specific to each of the individual standards.

NAR 1.0 was formally approved at the 2007 Annual General Meeting (in Tokyo last May), and subsequent work - particularly on NewsML-G2 and EventsML-G2 - resulted in a number of refinements. These include a change to the rights implementation and the introduction of two new elements

for in-line markup. In addition a number of minor errata have been identified and corrected.

A new version NAR 1.1 taking in these changes was formally adopted at the Autumn 2007 meeting.

Further work is in hand to finish the NAR specification documentation, refine the processing model and complete the controlled vocabularies, while the original NAR design document will be updated. Creation of an OWL (Web Ontology Language) representation of

The business office of IPTC is moving from Windsor to central London in December. The new address is:

International Press
Telecommunications
Council
20 Garrick Street,
London WC2E 9BT
United Kingdom

Telephone and Fax details to follow when available.

the NAR has also been proposed.

An important feature of the NAR is the use of QCodes, each of which represents a URI identifying a concept that can be applied as the value of a property. Each QCode has to belong to a controlled vocabulary - a scheme - and each scheme also has to be identified by its' own, valid, URI.

URI resolution

In use it should be possible to resolve both of these types of URI (to

Summary

News Architecture

Minor refinements give NAR 1.1 - **Page 1**. IPTC URI resolution policy - **Page 1**. URI termination - **Page 2**.

News Content

Standards release held back for final documentation - **Page 2**. EventsML-G2 structure fixed - **Page 2**. NewsML-G2 structure fixed - **Page 3**. Open-source NewsML transform - **Page 3**. SportsML 2.0 planned, with G2-Standards compatibility - **Page 3**. Timescale for standards adoption - **Page 2**. Recommended set for standards documentation - **Page 5**.

NewsCodes

Taxonomy management system updated - **Page 2**. Core Subject NewsCode set proposed - **Page 4**. Emotions NewsCode set - **Page 4**.

NITF

Proposals for NITF 4.0 with G2-standards compatible "core" profile and full "power" profile - **Page 4**.

NewsML 1

Functional specification revisions - **Page 5**. Namespace problems - **Page 5**.

IPTC business office move - **Page 1**.

2008 Spring Meeting arrangements - **Page 2**.

MOS Protocol - **Page 3**.

Rudi Horvarth - an appreciation - **Page 4**.

Czech News Agency - **Page 6**.

The IPTC Mirror has hyperlinks for [web addresses](#) and for [page references](#).

give human, or machine, readable information about the concept) and during the Autumn Meeting consideration was given to how this would be done - initially by the NAR Working Party, and then by the PR Committee and finally the Management Committee.

This resulted in a formal policy *"the IPTC cannot maintain a web server for the purpose of resolving IPTC QCode URIs for validation by the production system of any G2-Standards implementer. It is recommended to download the scheme and validate the QCodes locally"*.

This policy will be published on IPTC web sites and included in all human and machine readable information about a concept delivered by the resolving mechanism.

URI termination

Another issue with scheme URIs is the choice of terminating character, with the NAR specification allowing "!", "?" and "#". The merits of each of these approaches are the subject of considerable discussion within the W3C community, while the choice also depends on the technical characteristics of the web server being

used. A "Scheme URI Terminator Working Group" has been established by the Standards Committee to investigate this issue and make recommendations.

It is intended to change the NAR namespace URI to the style now adopted for IPTC standards (<http://iptc.org/std/NAR/xxx-xx-xx> - in the current version the "NAR" is lower case) for the next major revision of the NAR. This change is not being made immediately to ensure backwards compatibility.

News Content

The first members of the G2-Standards family are gradually nearing completion. Experimental phases for both NewsML-G2 and EventsML-G2 were completed during the summer, leading to the production of final specifications.

In both cases the Model specification and associated XML Schema files were available for consideration at the IPTC Autumn Meeting in mid-October, and were formally adopted.

However, final release of the standards has been held back for the production of the detailed documentation - with Core Conformance Level (CCL) and Power Conformance Level (PCL) Specifications for both NewsML-G2 and EventsML-G2 all being derived from a master document, which will also be used as the basis for the documentation of future G2-standards.

Making this documentation as precise as possible is seen as important to ensure consistent application of the standards - and so the maximum compatibility between systems. This is a time-consuming operation and it is anticipated that the process will not be completed until early January 2008. The planned timescale is given in the accompanying panel.

EventsML-G2

The EventsML-G2 experimental package was made publicly available in mid June, with a related teleconference presentation attracting a good level of participation.

Although there was little direct feedback from the Experimental Phase, discussions on the EventsML Yahoo group along with other comments was taken into account to give a version 4 of the Model, along with a set of XML Schema files. These were the documents adopted, with the de-

2008 Spring Meeting

Beijing, China

31 March - 3 April

The IPTC Spring 2008 Meeting will take place over four-days with a full day being devoted to a meeting with the Media Operations department of the China Olympic Committee (for the 2008 Olympic Games), including a venue visit. In addition to the normal working sessions there will be a half-day meeting and discussion session with the Xinhua News Agency.

Booking will open on the 10th December 2007 and has to be made by the 12th

January 2008. This is to allow time for the Xinhua News Agency to send delegates formal invitation letters, which are needed to obtain the Business Visas required for entry to China, and for delegates to apply for their visas.

sign and structure now being fixed for EventsML-G2 1.0.

News coverage

The standard is intended to allow structured information about events to be included in news item as content, and in concept items as knowledge. At the moment there is only limited provision for dealing with news coverage (which was one of the initial requirements). This consists of a "free-text" note and an extension point which can accommodate user information (such as assignment details) in a user namespace. It is anticipated that coverage information may be extended using G2-structures for a future release.

Application

An example of how EventsML will be used was provided by Johan Lindgren (TT) during the Prague Meeting. At present TT have a general calendar database that available to their customers, while an internal system uses the same da-

Timeline for G2-Standards

October 2007

NewsML-G2 and EventsML-G2 Model specifications and XML Schema files adopted at the Prague Autumn Meeting.

November 2007.

Final discussion on the first two standards, followed by work on the final specification documents in association with the technical author.

2 to 12 January 2008.

Specification documents finalised by technical author.

15 January 2008.

Final NewsML-G2 and EventsML-G2 specification documents circulated to IPTC members.

29 January 2008

Standards Committee meeting (by teleconference) to vote an adoption and release of NewsML-G2 and EventsML-G2.



John Evans, Managing Director of systems suppliers Transtel Communications (www.Transtel.com) gave a brief outline of the MOS Protocol and its use in news broadcasting at the Autumn Meeting.

MOS Protocol

The MOS (Media Object Server) Protocol (www.mosProtocol.com) is used in the broadcast industry as a standard way of exchanging information between Newsroom Computer Systems (NCS) and Media Object Servers (MOS).

From the IPTC viewpoint it is important because agency news feeds are often combined with other material for broadcast, and the standard controls what happens to the associated metadata.

For example, starting with a (text) news item on some sporting success, the broadcast editor might commission street interviews or the public reaction, and then put together a radio package consisting of the news item (read by an announcer)

and the audio interview files.

This process can be achieved using MOS, which is XML based and uses SOAP call and response commands. A set of MOS Profiles provide increasing functionality with Profile 0 dealing with basic communication and Profile 1 the Basic Object Workflow.

Media objects are described in Profile 1 using a series of messages, with a number of descriptive tags, including the <MOSExternal Metadata>. This is a mechanism to support other metadata standards - and may be used, for example, to contain metadata associated with a NewsML-G2 news item. Profile 2 allows the real-time creation and alteration of a Playlist, from the media objects.

Other standards with similar application are:

MXF (Media eXchange Format) - an open file format for the interchange of AV material along with associated data and metadata - www.ebu.ch/trev_291-devlin.pdf.

and JuMMp (Joined Up Metadata for Media Playback) - www.juMMp.net - a recently formed industry group working to join up the metadata chain.

As in other areas there is a conflict between the need to improve the immediate production process by simplifying the amount of data, and increasing the long term value of news items by keeping (and so having to move) the metadata.

database for such applications as coverage planning. Consideration is now being given to converting this database into EventsML-G2, with the first phase being a test projection in association with a customer who has already expressed an interest in the use of EventsML-G2.

NewsML-G2

Feedback from the Experimental phase resulted in a number of modifications to NewsML-G2, while various items that had been previously identified were also dealt with, resulting in a new Model specification (and associated XML Schema files) that were adopted at the Autumn Meeting. As with EventsML-G2 this adoption means that the design and structure for the

launch version are now fixed.

Examples

As reported in issue 141 of the IPTC Mirror several of the organisations directly involved the development of NewsML-G2 have already started work on implementations and a number of examples resulting from this work were provided during the Autumn Meeting. These included:

- An EBU sample dealing with video stream information, including a significant amount of rights information.
- An illustrated article from AFP with the text content in-line and details of a remote image in three alternative renditions.
- The way Reuters are creating packages of text news items, with

the inclusion of photographs.

- Use of SportsML data as news item content.

An important feature of these examples is that all of the providers concerned have a common approach - effectively the examples have the same structure - so giving a high level of interoperability.

NewsML transform

A transform from NewsML 1 to NewsML-G2 has been created by AFP and made available. The transform handles the AFP version of NewsML 1, but since it is open source other users can adapt it to deal with their NewsML 1 versions. This is necessary because there can be considerable differences between NewsML 1 applications (something that NewsML-G2 has been specifically designed to avoid).

SportsML

It is intended that the next release of SportsML - v2.0 - will be suitable for use as both a standalone standard maintaining full backwards capability, and for use with G2-Standards.

There will be a series of additions to the core to give more detailed information on players, sites, and events, along with a series of additions to the plug-ins covering (mainly) specific play actions for basketball, baseball, ice-hockey, and American football. There is also a new plug-in for rugby which has been designed to cover both rugby union and rugby league.

G2 compatibility

Relatively few changes are needed for G2 compatibility, which will offer SportsML users more flexible and extensive metadata and packaging capabilities. Changes include:

- Making the existing <sports-metadata> element - which holds metadata about the item as a whole - optional so the information is carried by the NAR structures instead.
 - Handling text content outside of the SportsML <article> element.
 - Redefining attributes that can be expressed as QCodes as controlled strings - and the structure of SportsML means that there are a lot of potential QCodes.
- Draft XML Schema, DTD and samples for SportsML 2.0 are at www.sportsml.org/dev/2.0, and includes a draft guide for the use of SportsML use with G2-Standards. It is hoped that a finished package

will be available for consideration - and approval - at the Spring 2008 Meeting (to be held in Beijing).

NewsCodes

Initial reports from the core NewsCodes development group indicate that the SchemaLogic taxonomy management system is now working reasonably well (following a number of initial problems).

All of the existing NewsCodes have been loaded into the system and a first set of revisions considered by the group. Further software updates for the system were provided during October and it is considered prudent to check the usability of the latest version with commonly used web browsers.

Revised governance settings are being developed and proposals will be circulated to delegates. Again, checks are being made to ensure that the governance module meets the specific IPTC requirements.

As soon as it has been confirmed that the system is fully operational (with agreement on the governance arrangements) arrangements will be made for members to have access to the system so they can take part in the development process - interested delegates should contact IPTC Managing Director Michael Steidl.

Core NewsCodes

A proposal has been made (by AFP) to produce a core set of NewsCodes that would be easy to apply for generic news applications in a busy editorial environment. This should consist of 200 to 300 terms, using the existing 17 main Subject headings, each with around 10 to 15 subheadings. These would be complemented by more extensive classifications, for applications requiring greater detail. This approach was seen as offering a number of benefits and will be taken into account in development of guidelines for the new set of Subject NewsCodes.

The importance of maintaining the established NewsCode sets (in particular the Subject NewsCodes) was stressed at the Autumn Meeting and it was confirmed that there should be no problems in maintaining the existing sets alongside the new versions.

Emotions

In response to a request from the Photo Metadata Group, a draft set of "emotion" NewsCodes was considered at the Autumn Meeting.

After considerable discussion it was decided that the proper application of such terms - for example "anger and "happiness" would be difficult and highly subjective, leading to inconsistency in their application. It was further suggested that the underlying requirement might be better met by a controlled vocabulary for expressions of emo-

tions - such as smiling and crying. This suggestion was referred back to the Photo Metadata Group.

NITF

Proposals for NITF 4.0 describe two profiles: a "core" which includes inline and structural markup, but does not have any metadata that

Rudi Horvath - an appreciation

Rudi Horvath (APA Technical Director and Managing Director APA-IT) stood down from the IPTC Management Committee at the 2007 AGM in Tokyo.

Klaus Sprick – formerly dpa Technical Director and a long-time member of the IPTC Management Committee (and a past IPTC Chairman) reflects on Rudi's contribution to the news industry as a whole, and to IPTC in particular.

I have known Rudi for a long time, namely 25 years.

Frankly speaking I do not really remember the day or year when we met first, because there are so many options. We had certainly met around the time he joined APA in 1983 in our respective capacities as Technical Directors of APA and dpa, in Vienna, Hamburg or elsewhere. One of the last major projects I did at dpa was to work out a service contract between APA and dpa. Rudi was the APA counterpart in the negotiations and I remember with pleasure his fighting hard for the issues but always in a friendly and solution oriented manner, respecting the position of the counterpart.

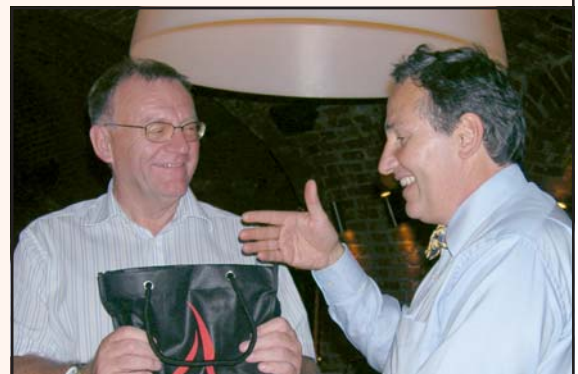
Because Rudi and APA became more and more active on an international level, there were at least three more organisations where we met and worked together.

- *The IPTC working parties, committees and later on in the Management Committee*
- *The EANA, the Alliance of European News Agencies, in the Technical Committee*
- *The epa, the European Pressphoto Agency, in the Technical Committee.*

As a Technical Director and Managing Director one has enough work on the plate and does not need to look out for something more to do. So what makes a man like Rudi to engage in voluntary work in the IPTC, EANA and epa? Of course, that is part of the business role in the agency, but that is only part of the answer. I am sure motivation comes from the conviction, that the entire industry we serve benefits from co-operation amongst us and by such work as for instance on information standards.

You know, that in the short term, to serve in such bodies, is work on top of the burden you already shoulder. In the long term, it helps the industry to flourish, and you and your agency amongst them. And one does not get more dumb doing it. There is a strong community of people in the IPTC with that very vision and Rudi was one of the outstanding individuals in there.

The Prague meeting was Rudi Horvath's last as a delegate and during a farewell dinner IPTC Chairman Stéphane Guérillot (right) made a presentation in recognition of his long service to IPTC.



would conflict with the G2-Standards metadata structure; and a "power" which would offer all existing NITF functionality. See <http://tinyurl.com/ywzawr> for the proposed split.

A mapping has been made between the proposed NITF power profile and G2-Standards metadata (see <http://tinyurl.com/2rgfx6>). This shows that most of the NITF elements and attributes map in a satisfactory manner, with the main exception being the NITF and publication data and revision history. This exception is not surprising since the NewsML-G2 is primarily intended for distribution applications. In addition the IIM (Information Interchange Model) is included in the NITF, but does not map to the G2-Standards.

Naming

In current discussions the different versions of the NITF are being referred to as "core" - for the subset and "power" - for the full version. However it is recognised that this usage might not be fully consistent with use of the same terms for the G2-Standards, so alternative names such as "standalone" and "embedded" are being considered.

G2 applications

Overall it does not appear that the NITF will not need any extensive changes for use with the G2-Standards, though these are some issues that have to be dealt with.

One of these is dealing with embedded metadata - NITF articles may include media items such as images or audio clips which has its own metadata. Three ways approaches have been identified:

- Create a separate news item for the media object and include it with the NITF article on a package item.
- Use links in the NewsItem header to reference the media objects.
- Leave the media items embedded in the NITF.

This problem does not only apply to the NITF, the same considerations will apply to other text formats, and the Reuters approach is to use a package with item references.

QCode use

Another consideration is the possible use of QCodes with the NITF. This could be achieved by the use of QCode-like values for existing NITF attributes. However, this would require detailed explanations (to customers) and might lead to limited compatibility between appli-

cations.

An alternative approach is to use the NAR inlineRef mechanism which makes it possible to apply QCodes to any element that has a XML id attribute - and this is the case with all NITF elements.

Future development

During the Prague meeting Working Party Chairman Stuart Myles outlined his views for the future development of the NITF, pointing out that at the moment it is a strong, stable and widely adopted standard, but that it faces emerging challenges.

Availability of different models for text articles raises interoperability problems (and increases costs) so there are significant advantages in encouraging wider NITF adoption.

One way of doing this would be to create an explicit article object model and generating mappings between this model and various

available text formats, while promoting the NITF as the preferred format for news applications.

NewsML 1

A number of revisions have been made to the Functional Specification Document for NewsML 1.2, including:

- Removing references to IETF RFC 3066 (Tags for the Identification of Languages) which is obsolete, and replacing them with references to the IETF BCP 47.
- Updating the date identification from "CCYY" to "YYYY".
- Modifying the ProviderID/domain name to allow the inclusion of a sub-domain.

Namespace problems

Efforts to have continued to try and resolve problems caused when documents without a namespace (typically created using a DTD for

Standard Standards Documentation

Introduction of the G2-Standards will be accompanied by a set of specification documents and supplementary material. There is already a substantial set of documentation for existing standards including the NITF and SportsML, both of which are moving towards the introduction of new major versions.

To help ensure consistency across the growing family of IPTC standards the Chairs of Working Parties and Working Groups dealing with the exchange standards recommend that a consistent set of documentation should be made available for each standard, as follows:

Specification documents

- Data model
- Processing model
- Technical specification
- An implementation of the standard in a widely used formal language such as XML Schema, DTD or RDF Schema

Examples and tutorials

- Full example XML documents (it is anticipated that there will also be fragment examples in the specification documentation).
- Tutorials on practical application
- Best Practices
- Frequently Asked Questions

Supplementary material

- XSLT for transforming XML documents between news exchange formats
- Software libraries (for Java, C# and other languages) to access XML documents by an API
- XPaths for all/most of the elements and attributes of the standard

Recommended formats for the documentation are Word/PDF and/or structured HTML documents with navigation features. DTDs and Schemas should include annotations to specify the semantics of properties and elements. The set of documents for each standard should be available as a compressed package with a single reference URL for easy access.

validation) are imported into a database using the NewsML 1.2 XML Schema.

The approach adopted was to use a style sheet to add a namespace where necessary, and consultants CNet were contracted

to do this. Initial tests show that this approach is satisfactory, and further tests by a number of system vendors are under way in Japan (it was these vendors that identified the original problem).

It is hoped that a complete set of

documentation, including the NewsML 1.2 XML Schema and associated XSLT style sheet, and the revised NewsML 1.2 function specification (along with a revised DTD) will be available for approval at the IPTC Spring 2008 Meeting.



Above: CTK Strategy Manager Jan Stejskal.

Below: Jan Kodera, CTK Technical Director.



become an important business area and includes 3D interactive infographics which are marketed internationally.

Jan Stejskal explained that CTK is now evolving a new strategy with the creation of innovative new solutions for their clients, while the independence of CTK gives them a lot of freedom. In this area Neris is a stand-alone subsidiary founded in 1999 and concentrating on new media. It provides a range of news services for the Internet along with SMS, MMS and WAP services, and operates its own news servers for general, financial and sports news.

The CTK IT department has a staff of 30 and includes sections to deal with development, operations, satellite networks and customer support and network and systems management and security. Internal software development has been carried out since the 1970s with recent IT milestones including:

- Infobank, a multimedia database system deployed in 1997 with CTK full text archives from 1988, an events database and factual databases.
- DVB-S satellite distribution in 2000.
- The new integrated editorial system was developed in house with the main design goal being to achieve a fully integrated workflow. It is based on XML and provides true multimedia production, including rich and flexible metadata.

Česká tisková kancelář (CTK)

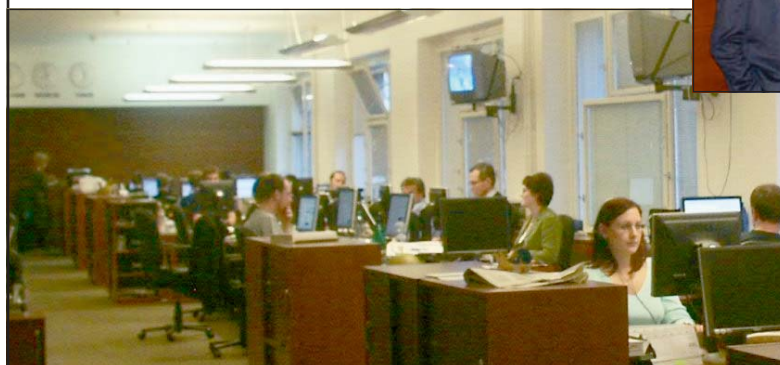
During the Autumn Meeting delegates had the opportunity to visit the offices of the Czech News Agency CTK. Jan Stejskal, CTK Strategy Manager, outlined the development and aims of the agency, while Jan Kodera, CTK Technical Director, gave an introduction to the organisation's approach to IT technology.

CTK is a public organisation, which has no owner, and is politically and economically independent (it has been in profit since 1997). Main clients are the media, business and financial institutions and state and government agencies. The principal news sources are a network of correspondents, international news wires and government sources.

Output includes text news in Czech, Slovak and English along with picture, audio and video news and infographics. Infographics has



Above: IPTC delegates being shown round the CTK offices by Jan Stejskal.



Left: The new newsroom was created in 2005 - this was a difficult task as it had to be fitted into the two existing buildings (one of which is a protected structure).

Published by the **International Press Telecommunications Council**
Royal Albert House, Sheet Street, Windsor, Berkshire SL4 1BE, England.
Managing Director: Michael W Steidl (mdirector@iptc.org) Editor: Hugh Johnstone (editor@iptc.org)
Tel: +44(0)1753 705051 Fax: +44(0)1753 831541 Web Site: www.iptc.org