



IPTC Mirror

IPTC - INFORMATION TECHNOLOGY FOR NEWS

Autumn 2003 - Leipzig

While work on the new family of IPTC standards remains a priority, attention is also being paid to the existing standards. Efforts to promote IPTC, and IPTC standards continue, along with initiatives to promote co-operation with other standards organisations.

Standards

Opening the Meeting, Standards Committee Chairman Stéphane Guérillot (AFP) told members that it now appeared that the original Roadmap for new standards development might have been too challenging so that some slippage was likely to occur. At the moment it appeared that this would mainly involve NewsML V2.0 with the target date for a draft standard being moved back by one meeting to the 2004 AGM. A revised Roadmap would be produced to take account of developments at the Leipzig Meeting.

Survey

A number of factors had delayed production of the planned User Survey but it is now available on the IPTC web site www.iptc.org/survey. The survey is being undertaken with assistance from the NAA. Target respondents include the IPTC, IFRA and NAA membership, along with other contacts such as member's customers and system suppliers - any suggestions should be submitted to IPTC Managing Director Michael Steidl.

Respondents are asked to complete the survey before the 3rd December 2002. The importance of taking this survey seriously was stressed, as the results will have a major impact on future IPTC developments.

XML Schema Conversion

Henrik Stadler (TT) presented the results of work that had been carried out for TT and IPTC to look into the conversion of DTDs to XML Schemas, with particular reference to IPTC standards. The consultant - Ulf Wingstedt of CNet Svenska AB - pointed out that XML Schemas are now replacing DTDs as they have increased expressive power and offer stronger validation.

There are a number of commercial XML tools that can be used to give a starting point for conversion of a DTD into a Schema. However, manual editing is generally needed to correct errors and the full power of a XML Schema cannot be achieved with automatic conversion. Conversion of the existing IPTC DTDs can be undertaken successfully to give XML Schemas with equivalent content models, while stronger type checking can be added.

This presentation was followed by a wide-ranging discussion of the approach that should be taken to the use of XML Schemas for IPTC standards. Although there was general agreement that XML Schemas were the way ahead there were a series of other factors that had to be taken into account before they could be fully adopted.

Reference version

Although a formal decision has not yet been taken the probable approach will be to maintain DTDs as the reference for the current generation of standards releases and move to XML Schemas for the next generation (for example NewsML V2.0). The plan is for DTD versions of the new releases to be made available as well - conversion of an XML Schema to a DTD can probably be automated to a large extent, though there will be a loss of richness and validation.

Where possible current versions of the standards will also be offered in XML Schema form. It would be sensible to ensure that these have a common style - this would probably mean that any DTD conversions would have to be carried out using the same tool each time. Another factor was the amount of extra work that would be involved in generating and maintaining XML Schema versions of existing standards.

A NewsML namespace was registered in 2001 and it was later decided to extend it to handle other standards such as the NITF

and SportsML. The proposal was submitted to the IETF (Internet Engineering Task Force) last year but returned to IPTC with a request for further explanation of the structures being proposed - which were different according to the application.

Further investigation showed that while the original namespace was satisfactory for its original purpose of providing an identifier

Continued on page 3

Contents

Working Party Reports:

| | |
|----------------------------------|-----------|
| Standards Committee | P1 |
| Working Party Vice-Chairs | P3 |
| IPTC Namespaces | P3 |
| Standards Liaison | P4 |
| Project Management | P4 |

| | |
|---------------------------|-----------|
| NITF Working Party | P4 |
|---------------------------|-----------|

| | |
|--------------------------------------|-----------|
| News Management Working Party | P4 |
|--------------------------------------|-----------|

| | |
|--------------------------------|-----------|
| NewsML Support | P5 |
| Version 2 Brainstorming | P6 |

| | |
|---------------------|-----------|
| PR Committee | P5 |
|---------------------|-----------|

| | |
|-------------------------------|-----------|
| Metadata Working Party | P6 |
| Special Content | P6 |

| | |
|------------------------------|-----------|
| News Standards Summit | P2 |
|------------------------------|-----------|

| | |
|-------------------------------------|-----------|
| 2004 AGM | P6 |
| and Chinese NewsML Symposium | |

This PDF version of the IPTC Mirror has active links to both the content and to external web sites. Clicking on the red links - Standards - will take you to the appropriate page, while the blue links - www.iptc.org - will attempt to open your browser and connect to the appropriate site.

News Standards Summit

During the 2003 Annual General Meeting it was decided to try to get together with other interested parties and arrange a news summit to investigate the profiles of the various standards available for news exchange and look at how they interacted. This task was undertaken by Mischa Wolf (Reuters) and arrangements are now almost complete.

The News Standards Summit will be held on Monday, 8th December 2003, in conjunction with XML 2003 in Philadelphia, Pennsylvania, USA. The program features presentations from

the key news standards efforts balanced by a requirements overview from a panel of international users. It will also include an open session where productive ideas for moving forward will be proposed.

The Summit is a collaborative effort of organizations committed to open standards development and sponsors include:

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

IDEAlliance <http://www.idealliance.org>

IFRA <http://www.ifra.com>

NAA <http://www.naa.org>

OASIS <http://www.oasis-open.org>

Programme

Standards

Chaired by Stéphane Guérillot, Director for Technical International Affairs, AFP and IPTC Standards Committee Chair

Keynote: Chet Ensign, Director of Architecture & Development Services, LexisNexis

NewsML: Laurent Le Meur, Technical Manager of Multimedia Development, AFP; NewsML Working Party Chair

NITF: Alan Karben, President, XML Team Solutions; NITF Working Party Chair

SportsML and Other Payload MLs for News: Alan Karben, President, XML Team Solutions; SportsML Working Party Chair

Prism: Ron Daniel, Principal, Taxonomy Strategies; PRISM Specification Editor

RSS: Ben Hammersley, Author of "Content Syndication with RSS"

Atom: Sam Ruby, Senior Technical Staff Member, Emerging Technologies Group, IBM; VP, Apache Software Foundation

ICE: Dianne Kennedy, VP Publishing Technologies, IDEAlliance; Editor ICE 2.0 Specification

XMP: Gunar Penikis, Product Manager, Creative Professional Products, Adobe Systems Inc.

User Requirements

Chaired by Linda Burman, CEO, L. A. Burman Associates Inc.; PRISM Working Group Chair and Founder

A series of short presentations on requirements for News Exchange Standards by representatives of users and implementers, with particular emphasis on: News; Online content; Mobile communications; Newspapers and magazines; and Picture libraries.

Confirmed speakers include: **Reuters** - David Schlesinger, Global Managing Editor and Head of Editorial Operations

Dow Jones - Ken Ficara, Director, Content and Information Strategy; **Time Inc** - Peter Meirs, Director of Alternative

Media Technologies; **Associated Press** - Geoff Haynes, Manager of Product Development, AP Digital; IPTC

Specialized Content Working Party Chair; **Vodafone** - Dan Appelquist, Director of Content Operations; **NSK** - Japan

Newspaper Publishers & Editors Association; **National Information Standards Organization (NISO)** - Patricia R. Harris, Executive Director.

Open Discussion

Chaired by Misha Wolf, Standards Manager, Content Architecture Group, Reuters

Taking the theme "What steps have to be taken to get the best of all worlds for the sake of hassle free news exchange?"

Registration:

Pre-registration and a nominal meeting fee of 90 US\$ will be required to cover the cost of lunch, breaks, and room fees. Register online at www.idealliance.org/news-summit, where the complete programme is available. The event is open to all and it is hoped that a substantial number of IPTC members will take part. Members are also asked to forward details to other organisations they know are involved with news standards, such as media companies and publishers.

for a NewsItem (which contains NewsML) the proposed extensions would not be satisfactory as there was no clear branch point to allow identification of standards or other documents. In addition, it would have meant that the NITF, for example, would have appeared under the NewsML namespace even though it is a totally separate standard.

Namespace proposal

Since there is a need for a namespace that can be used to define all IPTC standards, IPTC Managing Director Michael Steidl has developed a general IPTC namespace that can be used to uniquely identify all IPTC created resources, including documents specifying and supporting standards and working documents - see below.

The proposal was supported by delegates and has been submitted to the IETF - see <http://www.ietf.org/internet-drafts/draft-steidl-iptc-urn-00.txt>. Assuming that the new URN is approved the first application will be to create new namespaces for the NITF, SportsML and ProgramguideML. For the time being the established URN will be used for NewsML, but a new "IPTC" namespace will be introduced for NewsML V2.0.

Two MIME types have already been registered (with IANA - the Internet Assigned Numbers Authority) by IPTC. These are *text/vnd.IPTC.NewsML* together with

Working Party Vice-Chairs

At the AGM it was proposed that Vice-Chairmen should be appointed for the various Working Parties to provide back-up for occasions when the Chairmen are unable to attend a meeting. Individual Working Party Chairmen were asked to make nominations for these positions and as first set of appointments were approved as follows: News Metadata - Honor Craig-Bennet (PA); Special Content - Henrik Stadler (TT) and NewsML Support - Stuart Myles (Dow Jones). Subsequent to the meeting Johan Lindgren (TT) agreed to take on the positions of Vice Chairman for the SportsML Sub-group.

From left to right below: Honor Craig Bennet , Henrik Stadler and Johan Lindgren



text/vnd.IPTC.NITF. However, these types do not make it clear that the standards are XML based and this omission may complicate processing. Replacement MIME Types of *text/vnd.IPTC.NewsML+xml* and *text/vnd.IPTC.NITF+xml* have been suggested and this proposal was agreed, with the use of the previous versions being deprecated.

Directory structure

Most of the established IPTC Standards were the result of individual development projects and although many of the documentation and reference files are specific to individual standards, there are a number of common files. Unfortunately the directory structures used are different and this results in unnecessary complications.

A proposed common directory structure was presented to members by Michael Steidl, resulting in extensive discussions. While the idea of a common structure was well received it also raised the more general question of the policy to be adopted for maintaining different versions of the standards and related documents.

The structure final developed is valid for all IPTC standards and is being used for all new IPTC file packages, while it is planned to update previously released files to the new structure.

In the new structure the root directory of "IPTC" contains:

- catalog** - mainly for NewsML applications;
- topicset** - for TopicSet files
- metadata** - such as the Subject Reference System;
- tools** - software tools provided by IPTC; and
- sub-directories for each standard**, each of which contains further sub-directories for each version of the standard. These sub-directories contain the relevant specifications (DTDs, XSDs), documentation and examples.

The root directory can be stored anywhere in the user's file system and is independent of the path syntax of the operating system being used. A naming convention for IPTC Standards documents has also been adopted, which conforms to the URN

IPTC URN Namespaces

The proposed IPTC URN namespace has three branches: **std**; **std-draft**; **workdoc**. For the std (standards) branch the structure is:

urn:iptc:std:{std-name}:{std-version}:{res-group}:{res-name}{:res-version}?

where:

"std-name" is a unique identifier for an IPTC standard..

"std-version" reflects the version of this standard.

"res-group": this field will take only two values:

"spec" for all resources used for specifying a standard,

"doc" for all resources used for additional documentation and to support the use of a standard..

"res-name" is an identifier for a resource.

"res-version" reflects the version of this resource - this value is optional.

Structure for the std-draft branch is essentially the same:

urn:iptc:std-draft:{std-name}:{std-version}:{res-group}:{res-name}{:res-version}?

The "workdoc" branch URNs are for IPTC resources not directly related to the standards but to the work of IPTC and will have the general structure:

urn:iptc:workdoc:{group-id}:{doc-id}:{doc-version}:{doc-descr}?

where

"group-id" is a unique identifier for working groups and constitutes a document group.

"doc-id" is a unique identifier for a document within a documentgroup.

"doc-version" reflects the version of this work document.

"doc-descr" is an optional concise description of the document content.

"doc-format-id" is an optional identifier for the internal file format of the document.

Namespace Specific String used in the proposed new IPTC namespace.

Approvals

A series of motions from the Working Parties resulted in approval for NewsML V1.2, and NITF V3.2 which have been formally released for use. An amended XML Schema for NewsML 1.2 will also be released but formal approval was not required as the XML Schema is considered as experimental and so can be modified as appropriate. However it was pointed out that the value validation that had been incorporated ought to be removed to ensure that documents accepted by the DTD - which is the standard for V1.2 - will not be rejected by the XML Schema version.

A series of additions to the Subject Reference System were also approved along with a new "Scene" TopicSet and the updated SRS Guidelines.

NITF

Proposals for a NITF V3.2 had been refined at the last meeting with details being posted on the NITF discussion group for comments. Working Party Chairman Alan Karben (XML Team Solutions) reminded delegates that the changes included a number of incremental adjustments designed to improve consistency for various properties, and the addition of an extension to handle "Ruby" (a text annotation used in some Asian languages). This extension is a standard module according to W3C guidelines. All of these V3.2 additions are fully backwards compatible. After a brief discussion Version 3.2 was approved by the Working Party and passed to the Standards Committee for approval and release.

Response to the request for more details on companies using the NITF has resulted in a first set of entries for the "Who's using NITF" section of the NITF web site (www.nitf.org). Each entry gives brief details of the organisation and how it is using the NITF, along with contact details (these are coded to minimise the possibility of spammers obtaining the addresses). Further entries would be welcome. There are now over 400 members of the NITF discussion group (<http://groups.yahoo.com/group/nitf/>) with 10 to 15 new applications a month.

News Management

A revised version of the News Management Guidelines had been circulated for comments in Spring 2003, but there had been no feedback. Introduction of NewsML V1.1 only involved minor changes to the handling of updates.

Delegates were reminded that the guide does not add anything to the standard but provided an integrated view of the features provided for news management implementation, and describes alternative implementation approaches. These are considered from both the news provider and the news consumer viewpoints. It was agreed that no further work was needed on the Guidelines and that they could now be incorporated in the new NewsML documentation package.

Processing Model

Working Party Chairman Stuart Myles (Dow Jones) went on to propose that the next task for the group should be the development of a Processing Model for NewsML V2. This would provide a guide to the ways in which applications should process NewsML documents, and the hope is that it will help promote adoption of the standard. In addition it should provide a mechanism for setting conformance levels for applications.

It was explained that NewsML was seen as being hard to implement. This is partly because the wide range of features means that there are often several ways of doing the same thing, so that in-depth knowledge may be needed to decide on the best approach for a given application.

Use of a Processing Model would help simplify the implementation process by using object models to interpret the data. For example in a simple case a NewsML document would be associated with one or more sets of semantics (typically a set of relationships and properties that describe how the document should be processed).

In a more sophisticated model the NewsML application might be split into two parts: "NewsML Process" would deal with operations on the document; and "NewsML Behaviour" would specify interactions with the rest of the software.

There is already an effort to develop a conceptual (object) model for NewsML and

Standards Liaison

Liaison with other standards bodies is seen as having increasing importance and outline functions for representatives undertaking this were agreed:

- 1) Convey details of IPTC decisions and actions.
- 2) Establish what other standards bodies are doing.
- 3) Carry out specific tasks as requested by the Standards Committee.

So far, Mischa Wolf (Reuters) has agreed to undertake technical liaison with OASIS and W3C, and Laurent Le Meur (AFP) with ATOM.

this would make it possible to associate processes with individual objects. Individual parts could be dealt with one at a time, and overall the aim would be to produce an abstract approach that would not force users to adopt any particular technology.

The suggested way of doing this is to define the requirements that the standard processing model would have to meet: create a NewsML data object model; define the processing and behaviour associated with each NewsML object; and then refine both the object model and the processing model until the requirements are met.

This would be a major project and as a starting point any member with an existing processing model was asked to consider making it available to the Working Party. Otherwise it was proposed to start work on a Processing Model for NewsML News Management and then extend it to cover other areas. There was general agreement that this producing a Processing Model would have significant benefits, and it was decided that a subgroup would start to establish an initial set of requirements.

Project Management

A proposal to adopt Project Management techniques for the development of new IPTC standards was put forward by Geoff Haynes (AP), who outlined the basic principles and benefits and described how the concepts could be applied to the IPTC Standards development process.

There was general approval that adopting Project Management would be beneficial, though it was also recognised that implementation could be a significant burden and it was important to establish a workable system. A more detailed implementation proposal - including template documents - will be developed and the process test run on a new project.

The Management Committee (meeting after the working sessions) agreed to establish a Project Review Committee to oversee the Project Management and the first project will be a system for handling Weather Data - see the Special Content Working Party report (to be published next month) for more information.

NewsML Support

Although some delays have been experienced it is still planned to produce NewsML V2 during 2004, though this will require substantial input from members.

Working Party Chairman Laurent Le Meur (AFP) said that he hoped that the challenging target would help avoid dilution of efforts. The new version was needed to help increase use of NewsML, and some delegates were already expressing an interest

in implementation of the new version.

The speed of adoption will depend on the wider perception of the standard - if users know about it and favour its use then providers and system vendors will implement it more rapidly. It was also pointed out that

Public Relations Committee

Concentrated efforts have been made to develop a new, integrated, style for all the IPTC web sites. Details of the results obtained, and some of the thinking behind them, were presented by PR Committee Chairman Walt Baranger (New York Times) who explained that much of the credit was due to Alan Karben and Michael Steidl.

The main IPTC site features all the standards with direct links to separate sites (to the same design) for each standard along with a link to IPTC metadata and the SRS.

A series of suggestions were made during the meeting and revised pages subsequently made available to members for comment, some fine-tuning was carried out, and the sites are now in operation - see www.iptc.org.

The new-look IPTC Web site includes a members only section which is accessed by password and provides access to Working Documents along with details of Meetings, including booking forms and draft schedules where appropriate (the schedule also includes information on the date by which papers for consideration have to be submitted). Members should contact the Managing Director for information on how to access this area.

Promotion of IPTC Standards is seen as major aim and the new web sites are one of the ways of doing this. The News Standards Summit - see page 2 - is another in what will be a continuing process, and it may be appropriate to follow up this summit with similar presentations at other events such as NEXPO and IFRA. It was pointed out that additional technical documentation may be needed to help with this.

During discussions the question of alternative web addresses was raised and it was agreed that consideration should be

given to registering at least some of the obvious alternatives. Subsequent to the meeting sportml.org/com have been registered and point to the established (but new look) www.sportsML.com pages. Domains for [newscode\(s\).org/com](http://newscode(s).org/com) have been registered and their use will be considered at the Spring 2004 Meeting (in Athens).

Outline contents for the press release were agreed - the release itself is available in the Press Release section at www.iptc.org. Members were also informed that members Business Wire had issued a release on the launch of News Express, an Internet platform that delivers multimedia content in NewsML format, while accommodating the ANPA requirements of legacy news systems.

All the web sites for IPTC standards have a similar structure. The main IPTC site is at the bottom and the IPTC Members Only site below. Clicking on either image should open your browser at the appropriate page

Address: <http://www.iptc.org/pages/index.php>



| UPDATE: | |
|-------------|---|
| 13 Nov 2003 | Metadata topic sets: one updated - a new one named "Scene" released |
| 12 Nov 2003 | IPTC News Standards Survey launched open until 3 December 2003 The world's news agencies, news publishers, and software developers are being asked to help shape the future of standards for sharing and publishing news. [Press Release] |
| 11 Nov 2003 | ProgramGuideML: 1.0 draft released |
| 28 Oct 2003 | IPTC invites to the "News Standard Summit" in Philadelphia, Pa (USA): 8 Dec 2003 |

The International Press Telecommunications Council (IPTC) was established to safeguard the telecommunications interests of the World's Press. Now its activities are primarily focussed on developing and publishing Industry Standards for the interchange of news data. At present the IPTC membership is drawn mainly from the major news agencies around the globe but it also has a strong representation from newspaper publishers, system

NewsML V2 is needed as the core element for implementation of the integrated family of IPTC Standards.

At the same time it is important to ensure that the current standard meets user's needs and it was decided to produce a maintenance (update) NewsML V1.2 that would be fully backwards compatible. It was explained that this would only incorporate small changes to the standard as with most efforts are directed to the next-generation NewsML V2.0.

A number of Change Requests had been received and, after due consideration, the following changes were agreed:

Multiple creators allowed in Administrative Metadata, and "Contribution" Element added to identify the nature of a contribution.

FormalName of a Party element made optional.

Values for the Language element to be controlled by RFC 3066 (rather than by a TopicSet). This means that evolving changes in language identifiers can be accommodated automatically, and makes the format the same as that used for xml:lang.

Another Change Request had been submitted by CCI to add a class and style attribute to NewsLines so that different types of headings could be identified. An alternative approach of including them in the content block had been identified during the NewsML V2 brainstorming sessions so the request was withdrawn.

A new "Scene" TopicSet has been developed by the Metadata Working Party and it was agreed that this could be attached using a Property element to the Description Metadata element.

The practice of having Topics without defining the relevant TopicSet was discussed in some detail and it became clear that this raised complex issues with regard to operation of the TopicSet mechanism. It was generally accepted that implementers may not always provide a TopicSet as a controlled vocabulary even though this is required by the specification. However, care is needed as it may result in the processing system showing an error - but it was also agreed that such error signals could be ignored. This area is one that will receive more detailed consideration for NewsML V.2.

Update of NewsML documentation is continuing and it is hoped to have a draft version, of some chapters at least, in the public domain by early December when the News Summit will take place. However, it is also seen as important to ensure that what is released is right and a controlled review will

be needed. Draft chapters are available on the IPTC members mailing list for review.

Work on the documentation has raised questions about implementation details. For example it is hard to be sure if any of the TopicSets are normative - the Newsagency Guidelines say that NewsItem Type, Status, Priority and Location are but it is not clear if this statement only applies to news agency applications or to general use of NewsML.

The Leipzig Meeting also included productive sessions for the Metadata and Special content Working Parties with the main activities including:

Metadata:

New "Scene" TopicSet for image descriptions

Modifications and additions to Subject Reference System (SRS)

SRS Guidelines approved.

Special Content:

EventsML Requirements reviewed; object model under production, possibility of working with OASIS to be investigated

SportsML - Update on applications

ProgramguideML - V1.0 draft released.

Meeting with European Broadcasting Union (EBU) on compatibility with TV-Anytime.

Weather Markup - New project started.

AGM 2004 and Chinese NewsML Symposium

A brief update on arrangements being made for the 2004 AGM - to be held in Hong Kong - was given by Luka Lo, Senior Project Manager for the Chinese NewsML Community (CINTEC) who are hosting the meeting. The meeting will be held from Tuesday 25 May to Friday 28 May and the venue is the Inter-Continental Grand Stanford Hong Kong.

It is planned to hold the 2nd Chinese NewsML Forum on the day before the AGM (Monday the 24 May) with the keynote speaker being an IPTC representative and the session including a Round Table Discussion. AGM delegates will be welcome at the Forum and it is hoped that as many as possible will be able to attend.

Full reports on these Working Party sessions will be included in the next issue of the IPTC Mirror.

NewsML Version 2 Brainstorming

A special brainstorming meeting was held in Leipzig to help refine the requirements for NewsML V2.0. The meeting was well attended with significant progress being made in identifying areas needing specific attention.

Originally a single session was scheduled for the Tuesday afternoon before the start of the main Autumn Meeting, but the high level of activity resulted in a further session being held on the Thursday evening (after the planned working sessions had been completed). The Tuesday session also benefited from the presence of Daniel Rivers-Moore from RivCom - the consultant responsible for the initial version of NewsML - who had been specially invited to attend (RivCom have since joined IPTC as associate members).

Results from the sessions are being collated and refined to provide a framework for further consideration - any additional thoughts and comments would be welcome. Preliminary indications are that although NewsML has been well received as a powerful and versatile standard it tends to be seen as being complex and potentially difficult to implement - partly due to limitations of the documentation.

It seems that some simplification would be welcome, making it easier for customers to understand and use. This might involve alternative extension mechanisms including XML Schemas and Topicmaps. Other features such as authentication and signature, accessibility, internationalisation, and rights handling would also have to be considered.

At this stage it is uncertain if V.2 needs to contain all the features of V.1 or if it would be possible to deprecate some of them. In any case it will be important to ensure that the transition from NewsML V1.x to NewsML V2 is as easy as possible.