



# IPTC Mirror

## A Solid Architecture

**Providing the foundation for future generations of IPTC standards development of the new News Architecture is a major undertaking, but applied efforts ensured that the first stages were well underway by the Spring Meeting.**

Established to help achieve the aims behind the IPTC "Standards Tower" the **News Architecture Working Party - NAR** - (Chair Laurent Le Meur - AFP) is developing a generic architectural framework to manage and distribute all types of news related content. To do this the work has been split between four working groups (WG):

### Working Groups

**News Structure Working Group - NSTR** (Chair Laurent Le Meur - AFP): Developing a structural/conceptual model for the representation of all types of news-related content, to be adopted by all new major versions of IPTC standards.

**News Management Working Group - NMAN** (Chair Stuart Myles - Dow Jones): Developing a processing model for management of all types of news-related content, to be adopted by all new major versions of IPTC standards.

**Common Components Working Group - COCO** - (Chair Johan Lindgren - TT): Developing components common to all new major versions of IPTC standards.

**News Metadata Framework Working Group - NMDF** (Chair Mischa Wolf - Reuters): Specifying how metadata will be expressed, referenced and managed in all new major versions of IPTC standards.

### Goals

Opening the first session for the new Working Party, Chair Laurent Le Meur outlined the main goals for the NAR:

- make it simple - for basic needs;
- make it powerful - for high-level needs;
- make it easy to understand - with clear concepts and easy adoption;
- make it modular - so parts can be reused by other IPTC standards;
- make it interoperable - to allow standard processing;
- make it extensible - for providers;
- do not reinvent the wheel - use other standard where available;
- and think forward - to allow evolution.

Taken together it is hoped that

meeting these aims will help ensure that the standards produced using the new architecture will have world-wide applicability and acceptance.

### Development

Development of the NAR concept has drawn on a number of sources including the NewsML and EventsML requirements; the current generation of IPTC standards (NewsML 1, NITF and SportsML); and the work of other standards bodies (such as Atom, MPEG, PRISM, TopicMaps, and RDF/OWL).

This has resulted in substantial reshaping of the previous IPTC standards approach. It was emphasised that this meant that the NAR was not the same as NewsML2, as it has a much wider scope

Starting from well defined requirements, work will proceed with the development of conceptual and processing models. Use will be made of UML (Unified Modelling Language - <http://www.uml.org>) and an object-orientated approach.

Most of the work will be carried out within a set of development forums (for IPTC members only) with proposals and points raised by e-mail being discussed (and hopefully agreed) Continued on page 3

## Summary

**News Architecture Working Party** goals outlined - [Page 1](#). Working Groups established to deal with specific aspects of the new architecture. News Management - [Page 3](#). Common Components - [Page 3](#). News Structure - [Page 3](#). News Metadata Framework - [Page 3](#). Consultants commissioned to investigate the application of XML Schema technology to IPTC standards, and initial report presented - [Page 3](#).

**News Content Working Party** established - [Page 4](#). EventsML and WeatherML are waiting for the new architecture to become available - [Page 4](#). SportsML developing plug-in for Auto Racing - [Page 4](#) and investigating implications of new architecture - [Page 4](#). New Working Group to deal with general (unstructured) news content - [Page 4](#).

**News Metadata Working Party** sees further additions to

the Subject NewsCodes and progress in provision of translations - [Page 5](#). NewsCodes viewer - [Page 5](#).

**NewsML 1 Maintenance Working Party** - [Page 5](#). Adoption of NewsML as a JIS Standard - [Page 6](#).

News Industry Text Format Working Party working on V3.3 and draft Schema developed - [Page 6](#).

**IPTC4XMP** custom panels available for application of IPTCCore metadata in Adobe Photoshop - [Page 6](#).

**News Standard Summit 2005** - presentations on the major standards used by the news industry complemented by case studies from users - [Page 2](#).

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

# News Standards Summit 2005

24 May 2005 - Amsterdam RAI Centre, Netherlands

**Chair:** Misha Wolf, Standards Manager, Reuters; IPTC News Metadata Framework WG Chair

## Standards

**Keynote Presentation: Metainformation in XHTML 2.0**

Steven Pemberton, W3C HTML WG and Forms WG Chair

**NewsML and future News Architecture**

Laurent Le Meur, Technical Manager of Multimedia Development, AFP; IPTC News Architecture WP Chair

**NITF** - Hubertus Koehler, Chief Technical Officer, dpa-infocom GmbH

**RDF/OWL Semantic Technologies** - Ivan Herman, Head of Offices, W3C

**EventsML** - Arnaud Descamps, Chief Technical Officer, relaxnews

**PRISM** - Dianne Kennedy, IDEAlliance

**Photo Metadata** - XMP: Gunar Penikis, Adobe; IPTC Core: Michael Steidl, IPTC; DISC: Dianne Kennedy, IDEAlliance

**Radio and TV Metadata: Overview and TV-Anytime** - Jean-Pierre Evain, European Broadcasting Union

**Atom & RSS** - Bob Wyman, CTO, PubSub

**ICE 2.0 Web Services Syndication Protocol** - Dianne Kennedy, IDEAlliance

## Case Studies

**MINDS:** Hubertus Koehler, Chief Technical Officer, dpa-infocom GmbH

MINDS (Mobile Information and News Data Services for 3G) is a EU funded project that aims to identify, establish and maintain a common set of metadata, to enhance the quality of mobile services.

**Vodafone:** Kevin Smith, Architect Vodafone Global Technology

The Vodafone live! mobile portal brokers content from third party providers, then presents it in a format suitable for the requesting user and their handset. This use case discusses the differentiators between mobile and fixed-line web portals, how Vodafone have created plug-in extensions to the standard XML syndication formats, and how the syndicated data can then be presented to mobile users to generate content sales.

**Business Wire:** Jayson Lorenzen, Software Developer and Dean Large, Sr. Software Architect, Business Wire; IPTC NewsML 1 Maintenance WP Chair

Use of "Business Wire's Profile of NewsML" to integrate the flexibility of NewsML with structured delivery of news release documents. This helps consumers of Business Wire's NewsML documents to quickly and easily process the news feed.

**Austria Press Agency (APA):** Manfred Mitterholzer, Manager Research & Development and Rudolf Horvath, Technical Director APA, Managing Director, APA-IT Informations Technologie GmbH

The APA group operates the largest database host in Austria. With over 250 multimedia sources continuously updated in real time. These include the APA picture database and major German daily and weekly newspapers. For collecting this information standards have become indispensable. In addition there will be a short overview of the next generation fulltext-retrieval system APA-IT PowerSearch.

**PubSub:** Bob Wyman, CTO, PubSub

PubSub works with all the flavors of RSS and Atom as well as the NewsML and NITF formats that are used for press releases. Currently, they monitor over 8.5 million blogs, 50 newsgroups, press releases from the major US release distributors, all US SEC Edgar Financial filings, as well as earthquake notifications from around the world. They will show how they process these different formats and adapt them to users' needs.

**Reuters:** Dave Compton, Technical Specialist, News & Media, Reuters

Learn how Reuters uses NewsML for production, rendering and delivery of news.

**Agence France Presse (AFP):** Laurent Le Meur, Technical Manager of Multimedia Development, AFP

Overview of the different XML languages used by AFP, and a demonstration of the use of NewsML/NITF for the syndication of illustrated articles on the new AFP eCommerce platform - Magazine Forum.

## Open Industry Discussion

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

See <http://www.newssummit.org/2005>

during regular weekly conference calls.

### Working Groups

Following on from the Working Party overview, each of the Working Group Chairs outlined the scope and aims of the work being undertaken by their groups - applied effort prior to the Spring meeting had resulted in the production of initial discussion papers for each of the groups.

### News Management

News management is seen as an evolution of the NewsML 1 news management to cover all types of news related content, with all IPTC standards using the same approach. A proposed management component will govern the creation, evolution and destruction of items, and may be included in any item. Use of a unified management model will promote interoperability, and simplify implementation. Having a single management mechanism for all types of item should help encourage the adoption of a wider range of IPTC item types as it will make implementation easier. Any future changes to the management component and its model will apply to all IPTC item types



*Stuart Myles (Dow Jones) is responsible for the News Management Working Group.*

### News Structure

At the abstract level the "item" is the unit which will be dealt with by news management, and will contain components, which may be content or metadata. Classes will be extracted from the abstract level to the generic level (used by the standards) and these will include the Package Item and the News Message (for the exchange of items).



*As well as being the Chair for the News Architecture WP, Laurent Le Meur (AFP) heads the News Structure WG.*

Two conformance levels are envisaged, a core level giving simplicity and interoperability, and a power level offering the capability to meet the high level needs of the news industry.

### Common Components

Common Components can be considered as building blocks that will be used within the individual standards and integrated into them.



*Work of the Common Components Group is under the guidance of Johan Lindgren (TT)*

The 'anyComponent' will be the starting point for most of the Common Components, which will include Metadata components, (such as signature, rights, description, and label) and Topic related components (including person, location, contact, and resources for example)

### News Metadata Framework

Intended to provide a common way of processing metadata for all (new) IPTC standards the News Metadata Framework aims to: establish functional requirements for the ways in which metadata will be expressed, referenced and managed in all new major versions of IPTC standards; develop a model, satisfying these requirements; and establish the most appropriate approaches to expressing this model in XML.



*Mischa Wolf (Reuters) is taking care of the News Metadata Framework.*

### Consultants' report

In addition to the Working Groups the development process is making use of consultants to provide assistance and advice in the specific area of XML Schema implementation.

These consultants are Jay Cousins (RivCom) and Ulf Wingsted (Cnet), with their main tasks being to check the conceptual model against the capabilities of W3C

XML Schemas and to provide templates for XML schema adoption. Their initial conclusions were presented in a draft document 'Expressing the IPTC Standards Architecture using W3C XML Schema' and representing the consultants at the Spring Meeting Jay Cousins (RivCom) outlined the main features.

The approach taken had involved an in-depth study of the IPTC objectives and the conceptual model for the standards architecture, and investigations of alternative XML Schema representations of the architecture.

Key properties for a XML representation of the architecture were the support of object orientated inheritance and reuse; maintaining the distinction between the semantic layers of the news domain (these are exchange, management, content metadata and content); maintaining the distinction between structural layers of the news domain; and providing a granularity of common components that would simplify reuse.

### Component architecture

He went on to explain that the most thorough work on component architecture was ebXML (Electronic Business using eXtensible Markup Language - <http://www.ebxml.org>). Although this was complex it was likely that a simplified version would be applicable to the IPTC common component structure.

This provided a proposed structure, and XML Schema representation, for the common components with the lowest level being data types (such as string or date).

### Basic components

Next, basic components are a single property or piece of business data (name or age), and aggregate components consist of a specific set of properties composed from basic components and other aggregate components.

Namespaces provide a mechanism for identifying families of XML structures and their associated processing rules and could be used for individual standards and for common components.

Other factors considered included the requirements for extensibility of content and of the data model, and a look at the evolution of XML schema and an approach to versioning.

Proposals for introducing conformance levels included require-

ments that the structure should be easy to understand; include functionalities that logically belong together; and be layered so each level includes support for every lower level.

Finally Jay described a template for building common components with a consistent style using consistent process. The template requires a lot of data input, and provides a standard uniform format that makes it possible to specify specific characteristics for each component (such as identification, component type, naming, and constraints).



*An overview of the work being carried out by consultants on XML Schema implementation for IPTC standards was provided by Jay Cousins (RivCom).*

the consultants their work.

### Decision support document

During, and following, the Spring Meeting a concentrated effort was made to deal with these outstanding questions with the answers being supplied by mid March. This meant that the consultants were able to continue their work, and deliver a finalised decision support document with suggestions and samples; a draft template for description of common components; and an implementation support package reflecting the decisions and choices that have been made.

### Sustained efforts

A high level of activity has been maintained in News Architecture Wp since the Spring Meeting with a series of draft specifications haven been prepared by the Working Groups for consideration during the Annual General Meeting, along with the consultants' report.

## News Content Working Party

Taking over from the Specialised Content Working Party the News Content Working Party has the



*Vice-Chair of the News Content Working Party Henrik Stadler (TT).*

Group - the NewsML WG.

The Working Party also has a general responsibility to provide input to the News Architecture Working Party as appropriate and to make use of the output of the Common Components Working Group. Details of the activities of the Working Groups are as follows:

### EventsML

Activity on EventsML has generally been at a low level in recent months. However much of the work already carried out had been taken into the new structure with a series of requirements having been submitted to the Common Components, XML Schema Style and News Management WGs.

Work on EventsML itself will receive a fresh impetus when output is available from the News Architecture groups. The position was summarised as being that the group knew what they had to do, but needed to know how to do it (within the new IPTC standards architecture).

### WeatherML

Similarly the WeatherML project is on hold, waiting for the new architecture, though members were reminded that there is an active need for the standard.

### SportsML

Interest in SportsML continues to grow and Working Group Chair Alan Karben told delegates that the discussion group was up to 340 members. Recent enquiries had come from such parties as the San Diego Padres baseball team (appropriately the local team with the meeting being held in San Diego) and the US College Athletics Association.

Version 1.5 of the standard has been released with specifications and documentation, though examples were still to be produced.

same Chair - Geoffrey Haynes (PA) - and Vice-Chair Henrik Stadler (TT), who took the chair at the Spring Meeting session.

Scope of this WP has been extended to cover all types of news content and there is a new Working

A draft SportsML 1.5 XML Schema has been produced with improvements being considered including better formatting specifications for strings and improved controlled vocabulary listings. In addition the Schema style needs to be synchronised with the overall IPTC style.

### Auto Racing

A plug-in for Auto Racing is under development with the first step being the production of a sports-sheet to list the required properties. These will have to cover US, European and International structures for such aspects as results. Different types of motor sport would also have to be considered, including, for example, rallying, drag racing and various types of motorcycle sport. However, it was pointed out that since SportsML is used driven the initial coverage for Auto racing will depend on the user requirements.

### Architecture

The relationship between SportsML and the News Architecture work was also considered. It was explained that SportsML has a lot of statistics with a comparatively small metadata structure on top. In the short-term it might be possible to replicated the sports metadata in NAR common components. In the longer term it might be appropriate to replace the sports metadata by NAR common components.

Another possible use of the new architecture would be in the production of Game packages which could include a series of content items including a text write-up, scores and player performance statistics, photographs and audio and video.

### NewsML Working Group

This new group within the News Content Working Party is chaired by Laurent Le Meur, and is primarily concerned with the representation of general news content.

This is typically unstructured content - such as text and images, as opposed to, say, EventsML which will have structured content. It is intended to maintain features from NewsML1 as appropriate and NewsML 2 will make use of the NAR, using common components as appropriate and creating specific components as required.

It was pointed out that the NewsMLWG is not concerned with packaging and exchange (as was previously the case with NewsML

2) as these are cross-content features.

### NewsItem

Laurent went on to consider possible structure for a NewsItem, which would consist of:

- Management component - this will be required and the same for all IPTC standards.
- Metadata constructs - from the common components library with typical examples including description, labels and rights.
- News metadata construct - specific to the domain of general news.
- and News content components - which are specific to the general news domain.

In its simplest form a NewsItems would contain metadata and content. More complex versions might include several news items - for example a multimedia story with text illustrated by pictures and graphics - and would not be complete unless all the items were available. Another structure might be a package which could consist of a set of independent items, without any dependencies between the items.

During discussions the general case of including content that is structured by other Markup Languages - such as XBRL (eXtensible Business Markup Language), which is finding increasing use - was raised. Consideration was also given to the handling of alerts and headlines.

## News Metadata Working Party

Introduction of the new working arrangements has little effect on the activities of

the Metadata Working Party, under the Chair of John Minting (UPI) and Vice-Chair Honor Craig-Bennet (PA). However, it was pointed out during the session that it may be necessary to start considering new ways of handling NewsCodes to take advantage of the work under way to establish a Metadata Framework.



News Metadata Chair John Minting (UPI) - to left - and Vice-Chair Honor Craig-Bennet (PA).

A number of SubjectDetail entries had been approved by the fast track process - including new entries for international relations, peace negotiations and alliances under the SubjectMatter of diplomacy, with existing SubjectDetails under Politics being deprecated. Sports SubjectDetails entries for continuous and other billiards and for karate formal exercise were also deprecated.

### Translations

There has been further progress in the provision of translated versions of the NewsCodes with V15 of the SubjectCodes being available in Spanish (provided by UPI and AFP) while dpa have produced German versions of the Genre NewsCodes and Scene NewsCodes, in addition to the Subject NewsCodes.

### NewsCodes brand

As part of the process of introducing NewsCodes as a "brand" all of the controlled vocabularies have been given individual brand names such as 'Audiocoder NewsCodes' and 'Confidence NewsCodes'. Members were asked to use these names to help reinforce the identity of the IPTC metadata.

### NewsCodes Viewer

Development of the NewsCodes viewer (by Johan Lindgren) has been completed and members were given a demonstration of the main features, following which approval was given to its release for public use,

The viewer covers all of the IPTC NewsCodes which are handled as IPTC TopicSets. There are three main areas to the user interface: a list or tree view of the selected NewsCodes set; details of any selected NewsCode (formal name, name explanation and change information); and a display of any available translations of the selected NewsCode.

A download package including the viewerprogramme (Windows only) plus a user guide and the NewsCode files are available from <http://www.iptc.org/NewsCodes>.

### Non-sports qualifiers

John Minting informed delegates that the UPI editorial department had a requirement for a gender qualifier that could be applied to terms dealing with, for example, health, social issues and people - at present subject qualifiers are restricted to sports applications. He

said that a possible approach would be to use qualifiers from the generic "00" Subject for all codes other than sports, but was seeking to canvass opinion before making a formal proposal.

The Chair repeated his request for members to submit proposals for additions to the IPTC set of NewsCodes rather than developing their own systems. However it was pointed out that in some cases the metadata sets developed by members represented valuable intellectual property, and could form part of the contractual terms between an information provider and their customers. For example, this could be the case with financial news services.

## NewsML 1 Maintenance Working Party

Although development effort is

now concentrated on the new-generation standards architecture, NewsML 1 has a substantial, and growing, user base. To provide support for these users a new NewsML Working Parts has been established with Dean Large (Business Wire) as Chair and Takahiro Fujiwara (East Co) as Vice Chair.



Dean Large (Business Wire) has taken on the responsibility of Chair for the NewsML 1 Maintenance Working Party.

Main tasks of this Working Party were identified as being to: promulgate NewsML 1 while the new architecture is under development; implement any critical features while awaiting the new standard (though new developments should wait for NewsML2); and identify and promote a clear, painless, update path to NewsML2.

Public discussion of NewsML1 issues will be on <http://groups.yahoo.com/group/newsML>, while there is also an internal development group for IPTC members.

### RFC issues

Delegates were informed that some issues had been identified with the NewsML URN RFC, and it was pointed out that the RFC is now a

relatively old document that needed to be thoroughly reviewed, and a small working group is dealing with this. It was also agreed to submit a report to the RFC errata page and announce IPTC's intention of dealing with the problems.

### JIS Standard

Steps are under way to have NewsML adopted as a Japanese Industrial Standard (JIS), as it is



*Takahiro Fujiwara (East Co) outlined the progress being made in establishing NewsML as a JIS standard.*

now widely used in Japan. Takahiro Fujiwara briefed members on the background to this, and on the progress being made.

Initial work was carried out under the NSK with the NewsML and NewsML Support Teams, followed by a NewsML JIS

Working Group. Next stage was a NewsML JIS Working Party which went beyond the NSK activities and involved users (news agencies and newspapers), systems providers

and "neutral" parties including outside experts and academics.

Formal work on the JIS submission started in August 2004 with the production of a Japanese version of the NewsML specification. The proposal was accepted by the internal JIS process in early 2005, with formal approval by the Ministry of Economy Trade and Industry and release as a JIS Standard anticipated for later this year.

The process identified a number of minor errors in the NewsML 1 specification and these were reported to IPTC so they could be dealt with.

## News Industry Text Format Working Party

The NITF Working Party remains unchanged under the new working practices. A number of minor requests - all backwards compatible - have been made and a draft NITF 3.3 should be available for approval at the AGM.

### XML Schema

Working Party Chair Alan Karben (XML Team) told delegates that there was an increasing demand

from potential users for a NITF XML Schema. Accordingly, a draft has been developed and made available for comment (from [www.nitf.org/dev](http://www.nitf.org/dev)), with the intention of that it should be considered at the AGM. At this stage there are no functional differences from the DTD as no extra datatyping has been included.



*Alan Karben (XML Team) is Chair of the NITF Working Party and also heads the SportsML effort.*

This effort substantially preceded the main IPTC work on a common XML Schema style and it was stressed that any initial introduction of a

XML Schema for the NITF should be considered as experimental, does not affect the move to a common IPTC XML Schema style and will not necessarily represent future IPTC practice.

It is anticipated that the structure of the NITF XML Schema will change in line with IPTC recommendations and experience of their use.

## IPTC4XMP Custom Panels

**Custom panels to allow the use of "IPTC Core" metadata in Adobe Photoshop are now available for download from (<http://www.iptc.org/IPTC4XMP>).**

The complete package includes the IPTC Core 1.0 specification document; the 'custom panels' for Adobe Photoshop CS; a comprehensive user guide to these panel (with video tutorials on their use provided by David Riecks and the Stock Artists Alliance) example photos with IPTC Core metadata; and guideline for XMP software implementers

There are four custom panels which deal with: **Contact Information**; basic information about the **Content** of the photograph; more abstract **Image** information relating to the photograph; and **Status** information for image management, workflow and copyright. A section of the Image panel is given on the right, showing use of the IPTC Scene NewCodes.

Some problems have been reported with the use of these custom panels in Photoshop 7.1 and CS 1 with some data input not being saved properly. A "fix" to deal with this has been developed and is available from the IPTC4XMP pages on the IPTC web site.

There is also a discussion group and support forum at <http://groups.yahoo.com/group/iptc4xmp>

