

Foundations for Extension

NewsML Support

Although no major changes to NewsML have been needed since its introduction in October 2000, there has been a continuing process of refinement and amplification, which was continued during the Paris session. This process was given added impetus by release of the News Agency Guidelines, which provide a "good practice" guide for the application of NewsML.

Included amongst the changes were some minor modifications to the Functional Specification which should help to minimise application problems and improve the compatibility of NewsML with other standards.

For example, when setting the language the xml:lang attribute used is based on RFC3066. This allows for both two letter (eg US) and three letter (eg USA) country codes, but the language TopicSet for NewsML is derived from ISO 639-1 which

M. Bertrand Eveno, the chairman of Agence France Presse (AFP), took advantage of the Spring Meeting being held in Paris to make a visit and discuss some of the current developments. Shown here are Management Committee members Klaus Sprick (dpa) and Stéphane Guérillot (AFP) to left and right respectively, with M. Eveno, and IPTC Managing Director David Allen (centre right).



Solid efforts to improve the usability of the main standards and lay the foundations for new releases were the major feature of the 2002 Spring Meeting, which was held in Paris at the end of February. As normal the working sessions were complemented by some informative presentations.

has two letter codes (the three letter codes are in ISO 639-2). Accordingly it was agreed to modify the functional specification to make it clear that the alternative approach of using ISO3166 country codes - in either two or three letter forms - is also permissible, depending on the user's preference. The TopicSet will continue to use the two character ISO639-1 codes, with the possibility of future extension to the three later codes if this becomes compatible with W3C practice.

Similarly NewsML is designed to be neutral to the delivery method, and this means that any references to external content should also be neutral. The *Href* attribute is used for such references and the wording in the Functional Specification has been changed so the reference is to the general URI (Universal Resource Indicator) as opposed to the more specific URL (Universal Resource Indicator).

Date and Time information is handled according to ISO 8601 and this allows

Summary

Spring Meeting Report:

Committee of the Whole updates members on new MD recruitment, new Chairman, and publication changes - **P3**

NewsML Support work continues on improving the Functional Specification and preparing additions for a new NewsML version - **P1**

News Management Guide to News Management under development, procedure for handling deletions - **P3**

RadioTV-NewsML Japanese-led initiative to develop a new standard for television and radio listings - **P4**

Specialised Content Updates on Election reporting, Financial Data, Government Data, Public Relations and SportsML - **P5**

NITF Maintenance detailed changes to maintain compatibility with NewsML, discussion of newsgroup issues, handling of copyright - **P7**

Standards Committee procedure for handling NewsML updates, processing news flow via e-mail - **P8**

PR Committee need for NewsML examples and training package, overview of NewsML activities in Japan - **P7**

Presentations:

Xyleme News and Publishing aims to help users exploit XML for news applications - **P2**

NAMIC - automated multi-lingual categorisation and browsing - **P6**

Information and content Exchange (ICE) update - **P8**

NewsML editor - **P8**

Rising to the challenge

Xyleme (www.xyleme.com) is a spin-off from the INRIA IT Research Laboratory and was established in 2000 following market-research, with the Xyleme News and Publishing system being released in 2001.

Introducing the organisation, Patric Ferran explained that xylem is the tissue that carries sap up in plants and by analogy Xyleme helps raise information to where it can best be used. He went on to outline a number of challenges faced by users wanting to exploit XML for news applications: How to store large volumes of XML and support large numbers of users; How to obtain relevant and precise information and improve navigation; How to control information heterogeneity; How to be aware of changes in a large corpus of information; how to improve content distribution; and how to migrate legacy data.

A description of the way Xyleme addresses these challenges was then provided by Serge Abiteboul. He explained that traditional databases are not adapted to storing trees, so a native XML repository has been developed with a cluster (PC) architecture. This database has been

extensively tested with NewsML data and can be extended to terabytes of data.

Both plain text searching and database queries are combined in an integrated system. There is an arbitrary level of granularity so searching is possible, for example at both XML element and attribute level, with the arrangement being said to offer both efficiency and scalability.

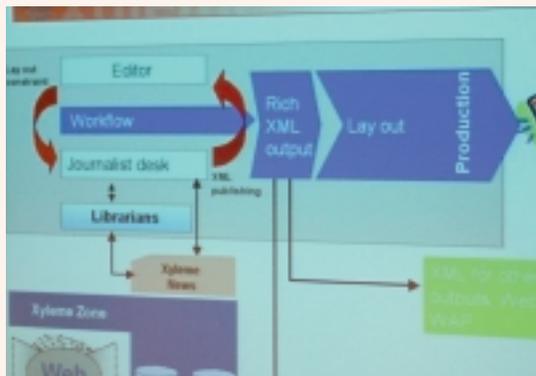
Navigation through the data is achieved by a system of customised screens backed by a XMLQuery language, while the system is designed to be compliant with the planned Xquery standard. A semantic mapper is used to provide uniform views - using an abstract DTD - of information produced using different DTDs. The system also includes features for monitoring internal data, data from licensed sources and the Web, so users can be notified when relevant changes occur.

Migration of legacy data is a tough challenge and there is no general-purpose approach, but tools can be produced to suit specific applications - for example by analysing layout patterns in the documents.

This outline was followed by a brief demonstration of the system in operation, based on sports feeds. Starting point was an interview, displayed on the Xyleme XML editor, which included a number of incomplete references. Opening a sub-screen made it possible to search for the missing information in a wire report

database and bring the results up for inclusion in the article. In a second search, looking for "red card" incidents during the (football) World Cup found the relevant information from

one source, and complemented it with biographical details for the named player from a second source.



Copy flow in the Xyleme system



Patric Ferran (above) explained the background to Xyleme News and Publishing, while Serge Abiteboul showed how the challenges were met.



for a number of different formats. During discussions it became clear that several members were using the "Z" suffix to show that the time corresponds to UTC (Universal Coordinated Time), so it was agreed to amend the specification to allow this.

Such changes can affect the tools available for Conformance Checking and there was a brief review of these tools. Some upgrading work is in progress, and more detailed coverage will be given in a future issue of the IPTC Mirror.

Versioning

The approach taken to establish a NewsML namespace and identify different versions was reviewed in some detail, resulting in a proposal that the namespace will be fixed, with a version attribute being added to the DTD. Every revision to the DTD would become a new NewsItem with the reversion level being kept at 1, while the NewsItemID version would be changed as necessary - such as V1.01 to V1.02.

Including an attribute to identify the version would make it easier for subscribers to use a locally stored version for validation - if required. To allow this publishers should keep the relevant version of the DTD available on their systems. However it was important to note that incorrect use of the version attribute could lead to problems, and it was also important to indicate which version of the DTD is being used.

There is a need to identify the date and time a ContentItem (or DataContent) was created, when this is not the same as that of the NewsItem. Details were given of two approaches, one having the information in the ContentItem and the other in the NewsComponent. Both of these use the Property mechanism and it was agreed to establish the FormalName of the property as OriginalContentCreationDate, which can be used according to the individual publishers practices.

Byline structure

A number of suggestions have been made for extending the Byline Structure (and the Dateline) including: adding property elements to the AdministrativeMetadata; extension of the NewsLine structure or the use of nested origin elements; or by modification of the DTD. It was felt that this would need careful consideration with both Technical and Journalistic inputs. Accordingly, it was decided to establish a small ad-hoc group to look into the options for modification of the DTD, with particular reference to ByLine and DateLine extensions, the possibility of including a ByLineTitle and the potential introduction of an EventLocation element.

Most of the changes agreed during the session have now been taken into the lat-

News Management

Chairman Stuart Myles (WSJ) has been preparing a "Guide to News Management with NewsML" and he gave a brief progress report to members. The Guide is described as dealing with both NewsItem replacement and NewsItem manipulation and makes use of the Duid (document-unique identifier) and Euid (element-unique identifier). When the Guideline is complete copies will be distributed to members for comments before it is formally submitted for discussion and approval.

Some questions have been raised about the way deletions should be handled - specifically where it is intended to totally retract a previously transmitted story. Use of the "Cancelled" status involves retransmission of the entire NewsManagement element (and any children) and there was a suggestion that this might involve too much information. There might also be a need to clarify exactly what a status of "Cancelled" means - ie what action should be taken by the subscriber receiving the information.

Status

Similar considerations apply to the other status values and to help clarify things it was agreed that explicit explanations would be added to the News Agency Guidelines to cover the status values of Usable, Embargoed, Withheld and Cancelled - at the moment there is just a brief explanation in the TopicSet. However, it was recognised that the usage could be different for non News Agency applications so it was important to make it clear that the Guidelines were specifically intended for News Agency applications.

There was also some discussion of the need to include News Management features in the other IPTC standards. So far as SportsML was concerned, Alan Karben said that he thought it was likely that SportsML content would be transmitted within NewsML so the existing mechanisms could be used. Essentially this means that it is possible to update anything that is identified by a Duid.

In addition there was a feeling that it might be useful to have status elements in the NITF (mainly for when the NITF is being used on its own) and a suggestion to look at this was passed on to the NITF Maintenance Working Party.

Feedback was requested from any users of the NewsML news management features - specifically are the mechanisms adequate and efficient? and are there any other scenarios for news management?

Committee of the Whole

A special session of the Committee of the Whole was held, under the chairmanship of Peter Müller (SDA/ATS) to update delegates on a number of general issues concerning the running of IPTC itself.

Foremost among these was the progress being made in finding a replacement for IPTC Managing Director David Allen, who will be leaving early in 2003. The Management Committee have produced a document giving a profile of the position and outlining the required capabilities. An advertisement has been produced (see below) for use in appropriate publications and other areas - such as the NewsML mailing list.

The intention is to try to identify suitable candidates by the end of June so that members can be informed during the AGM. Interviews will then be held in late summer and the hope is that it will be possible to make the appointment in sufficient time to allow some overlap before David Allen leaves. Initial approaches should be made via David Allen, who will forward details to Peter Müller for consideration by the Management Committee. Members were informed that appropriate provision has been made to cover the costs of recruiting a new Managing Director and the transition. No increase in membership costs were seen for 2003, but in the worst case a small increase in subscriptions might be needed to cover additional administrative costs - at the moment these are limited by the Managing Director working from home, but this might not be the case in the future.

IPTC Chairmen generally serve for a three-year term, and Peter Müller told members that he would be standing down at this year's AGM. The outgoing Chairman normally has the privilege of proposing a successor and Mr Müller felt that he had identified an excellent candidate in the person of John Iobst (NAA), who has agreed to take on the responsibility (subject to agreement of the members at the Annual General Meeting). In addition all members of the Management Committee have expressed their willingness to stand for re-election - though additional nominations can also be made.

Much of the business of IPTC is now handled electronically and it was agreed that this should be extended to the IPTC Mirror Newsletter, with the distribution (to members) being changed to e-mail in the near future. However, the Annual Spectrum meets a different need and will continue to be distributed in printed form to both members and non-members.

Managing Director

The IPTC, the International Press Telecommunications Council, is seeking a new Managing Director for January 2003 to replace the retiring Managing Director. IPTC is the major international body setting information exchange standards for the News Business.

The Managing Director (MD) is the only IPTC employee and responsible for administering all of the Council's work and organising its meetings. The working language of Council is English. The MD is also responsible for liaison with external organisations and for representing the Council through public speaking and other activities. The MD is also expected to contribute to the technical work of the IPTC by drawing upon appropriate professional knowledge.

The ideal candidate is not expecting a career progression but wishing to consolidate and broaden skills and grow in the job. Applicants must be willing to travel, preferably from a European base and, ideally, operate out of a home office.

An attractive remuneration package is on offer subject to agreement over full-time or flexible-part-time working. Applicants should in the first instance obtain more information by contacting the current MD at m_director_iptc@iptc.org

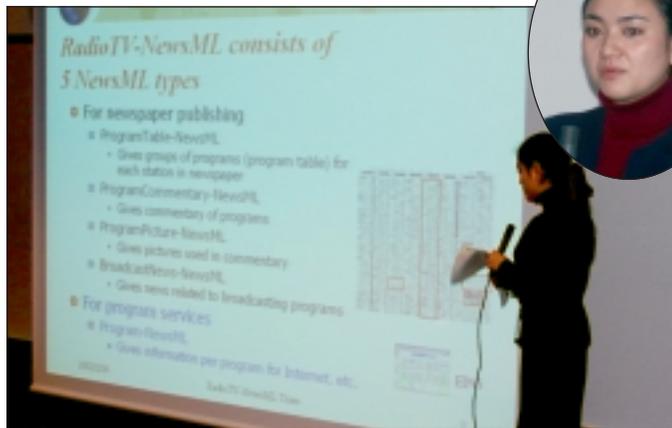
RadioTV-NewsML

A major initiative is under way in Japan to develop a system for handling television and radio programme information within the NewsML framework. Presenting the project at the Spring Meeting, Ryugo Onishi (Tokyo News) said that the team is developing RadioTV-NewsML for use in Japan, but would also like to propose RadioTV-NewsML as a uniform, world-wide approach that IPTC can promote as a global standard.

He went on to outline the background to TV program listings in Japan. Many newspapers and magazines use the services of specialised delivery-service companies which take information from the programme tables, edit it into programme tables, add supplementary information, and then deliver to the publications. Currently the information is handled in a fixed text format which has a number of problems; it is not unified, is slow and inflexible. This means that it would be very difficult to expand the service.

The aim is to use RadioTV-NewsML for delivery, so the publishers receive the content in a structured form. The new system can also be used for other applications such as Electronic Program Guides (EPG) which presents programme data in a digital form that can be interfaced with information terminals and allow easy control of recordings. Main customers for the delivery services are newspaper and magazine companies and electronic media, while information is delivered in three main styles - daily, weekly and by program category.

Mr Onishi then introduced Shoko Okudaira (NEC) who started by

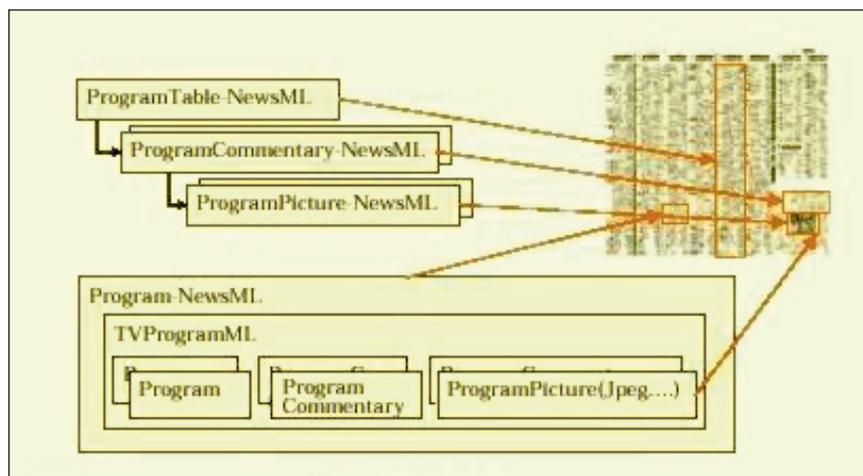


Shoko Okudaira describes the structure of RadioTV-NewsML.

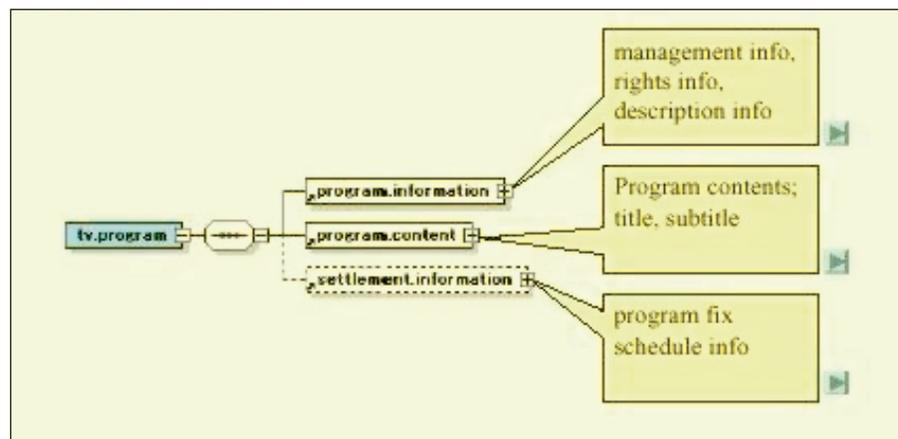
describing a typical (Japanese) programme listing. This has a series of components; the Programme Table lists the programmes by station and by time; specific Programme information, such as time, title, subtitle and credits; Programme Pictures; and Programme Commentary.

To handle these components RadioTV-NewsML has five parts: four for newspaper publishing (ProgramTable-NewsML, ProgramCommentary-NewsML, ProgramPicture-NewsML and BroadcastNews-NewsML) and one (ProgramNews-NewsML) for program services. All of these comply with the basic NewsML specification with metadata (to be defined) and DataContent.

Probable metadata - handled in TopicSets - includes the broadcasting service name and the planned broadcast dates, for example. The intention is to use the IPTC Subject



Top left. Relationship between the different parts of RadioTV-NewsML, showing how they are used to produce a page of TV listings. Individual items on the page, such as the table column containing the entries for a particular station are handled by the appropriate sub-programs with NewsML. AssociatedWith elements being used to assemble the page.



Bottom left. TVProgramML carries the TVProgram DTD which is used to provide complete descriptions for individual programs and has provision for a comprehensive set of metadata.



Ryugo Onishi

Reference System for categorisation even though there is already an ARIB GenreCode for the Japanese broadcast industry - but the ARIB codes will be included in the DataContent.

For newspaper applications, managing the components individually gives flexibility - for example programme pictures and commentaries can be sent separately from the schedule information, while the layouts may need to vary by region of television station.

Program-NewsML also conforms to the NewsML specification but the DataContent consists of a separate DTD - TVProgramML - which handles the information for an individual programme. Typical content might include the title, start time, the name of the director, a cast list and so on, along with pictures and reviews. It can also include control information such as the charging structure (Pay per View) and Parental Guidance information. Coding all this information in XML gives a 'one source, multi use' package, allowing easy selection by content (sports only, or movie only).

Miss Okudaira pointed out that there was already an XMLTV programme set, which might appear to meet the same needs as TVProgramML. However comparisons showed that TVXML was essentially intended for the user, while TVProgramML was aimed at the information delivery companies. The possibility of achieving compatibility between the standards had been investigated, but this would be difficult due to the different design concepts.

A RadioTV-NewsML mailing list has been established with open membership to interested parties - see <http://groups.yahoo.com/group/radiotv-newsml/join>. There is also an experimental RadioTV-NewsML web site at <http://www.newsml.jp/radiotv/>.

Initial response from delegates was limited but seemed to indicate that no serious shortcomings had been encountered so far.

Specialised Content

Emphasis in this group is on handling structured content (other than plain text) which is often in tabular form, and deals with specific areas. The interest is primarily for use in connection with NewsML, although the applications may also be appropriate for stand-alone use. Work in hand includes some specific developments by IPTC members, along with monitoring of other relevant initiatives.

Election reporting: Geoff Haynes (AP) the Working Party Chairman, gave a brief overview of the two main initiatives in this area - OASIS and VNS. OASIS (www.oasis-open.org) acts as an XML clearing house and has formed an Election and Voter Services Technical Committee. This has a broad scope with the main aim appearing to be the development and sale of voting systems to governments. The first meeting was held in May 2001 and a draft V0.1b Schema was released in February 2002.

The Voter News Service (VNS) (www.vnusa.org) is a consortium founded by the US Television Networks and AP and is a major source of election night results reporting and exit polls for US elections. At the moment the output is not in a form suitable for immediate publication, but they are now moving towards an XML-based system. Progress so far includes identification of the vocabulary for a Schema with the Schema planned for release during the summer, so it will be available for use for the delivery of results in the November 2002 US elections.

IPTC are monitoring both these developments, with particular attention on the outputs available for reporting election reports. By the end of the year it should be possible to see what has been achieved and decide what approach to take - in addition to the above initiatives both AP and Canadian Press have already produced their own DTDs for handling elections.

Weather markup: Little activity has been reported in this area.

Financial data: There are a number of initiatives in this area that appear relevant to IPTC members and Irving Levine (Reuters) provided a brief summary. A first formal draft Schema for MDDL (Market Data Definition Language - www.mddl.org) was released at the end of 2001 and deals with equities and mutual funds, providing end-of-day and snapshot reports. Since the standard is new it has not yet found significant use.

There is a new release of XBRL (Extensible Business Research Language -

www.xbrl.org) with some major revisions. It is based on the use of X-link to describe the relationships between terms and efforts are under way to develop the local taxonomies that are needed to meet, for example, local accounting standards.

Some contacts have been made with the RIXML (Research Industry Extensible Markup Language - www.rixml.org) Group, which seems to be mainly concerned with reporting research results from investment organisations. Detailed metadata has been developed but, so far, there does not seem to be much structure for the content.

Government data: Efforts are being made to encourage Government and other official organisations to present information in easy-to process forms - a paper dealing with this and produced by Walt Barranger (New York Times) is being reviewed and it is hoped that it will be ready for presentation (and approval) at the San Francisco AGM.

Members were also informed that the UK Government has announced plans to release statistical data in XML format though timing and technical details are not yet available.

Until recently regulatory announcements in the UK had to pass through the Stock Exchange, but following recent changes a group of companies are offering services in this area. These include a number of IPTC members, with the information being provided in NewsML format.

Public relations: Contact has been made with the UK-based XPRL (extensible Public Relations Markup) group - see <http://www.lmu.ac.uk/lbs/pr/xprl.htm> - who have produced a draft business outline and are starting work on a Schema.

SportsML: Activity in the SportsML group has been relatively low key since release of Version 0.5 at the Autumn Meeting, but Group Chairman Alan Karben (Screaming Media) said that the next stage of the development process should soon get under way - the SportsML mailing list (sportsml.yahoo-groups.com) now has some 45 members. SportsML has a core DTD containing basic common information, which can be complemented by plug-in DTDs to provide additional information for individual sports (see www.sportsml.com).

Some trial implementations have been undertaken over the past months - for example Screaming Media (who are information aggregators) have experience in converting AP sports feeds to the SportsML format. Similarly the Swedish News Agency TT have been using a previously developed XML markup to meet their customer's needs and output from the Winter Olympics was converted into SportsML for evaluation purposes. The conversion proved fairly straightforward, and it was possible to

Categorising the News

NAMIC (News Agency Multilingual Information Categorisation) is a EU funded project which is now reaching its conclusion. A number of IPTC members - and IPTC itself - are participants and a briefing on the project and its achievements was provided at the Spring Meeting by Dr Roberto Basili from the University of Rome, and Maria Vittoria Marabello of Knowledge Stones.

Aim of the project was to apply Natural Language Processing (NLP) technologies to carry out content-driven classification of news items, and to produce multilingual hypertext browsing systems. Work has been carried out using three main languages

-English, Italian and Spanish - with a common processing architecture.

News Agencies are seen as the core users, at least for the first stage of development, with the following specific business needs: receiving news from geographically distributed sources, possibly in several languages; and categorising news items using content-dependent processes. The news input can be enriched by deriving information that can be used for later retrieval (such as keywords, names, and industrial sectors), and by deriving relationships among different documents. This could identify

documents describing the same information

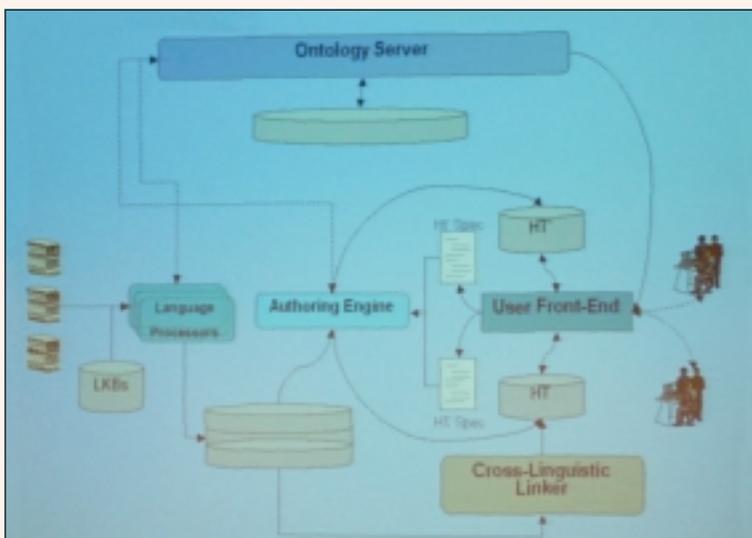
or the same event, for example.

Specific services that could be offered include: the retrieval of news in a specified language, usually according to a subject category (this represents some of the input from IPTC); the browsing of news databases following specific search parameters; and accessing items written in other languages. Agencies would also be able to integrate information from outside sources into their own repositories even though the items were originally produced in different languages and classified using different standards.

A prototype system has been produced and tests carried out on sample news items gave a typical classification accuracy of better than seventy per cent. However much higher accuracies were achieved in some areas such as Sports. The categorisation system seems to be the component that has the most immediate application, and Knowledge Stones are currently working on a implementation package.

Further information on the Namic project is available on <http://namic.itaca.it/>.

Schematic outline of the NAMIC system



Dr Roberto Basili



Maria Vittoria Marabello

cover all events using the core DTD. However, some problems were encountered in handling the Medal Tables and this needs to be addressed.

Input from, and co-operation with, the major sporting bodies is seen as valuable for the development of SportsML and a number of efforts have been made in this direction. Geoffrey Haynes and Alan Karben held a meeting with the Major League Baseball authorities, who are looking to move to XML based systems.

Although there appears to be some common ground there are also potential problems in that the professional leagues like to own the data (score and other facts) on their games. At the moment AP, and other agencies, obtain their reports from stringers who provide real-time feeds but this is currently subject to legal disputes.

In other areas, attempts by David Allen to contact the PGA and FIFA have not yet had any success.

Looking ahead, the aim is to move towards a V1.0 release of SportsML in the Autumn; to look at covering other sports - individual members are encouraged to take on the production of specific vocabularies for individual sports not already dealt with; and to continue with efforts to involve national and international sports organisations.

TV listings: A major effort is under way in Japan to develop a structured format for Radio and Television listings and there was a detailed presentation of the progress that has been made - see the panel on page 4.

News Metadata

A number of requests have been made for additions to the NewsML vocabularies, with the aim of increasing the scope and usability of this 'standard' metadata. Members were reminded that these vocabularies are intended to cover areas that are of general interest so the entries will find broad application. However, there is a danger that continued extension could result in the TopicSets becoming large and unwieldy, while requiring a substantial continuing maintenance effort.

To help avoid this it is generally assumed that individual users will develop vocabularies to meet the specific requirements of their services. There is also a need to agree on which TopicSets will be regularly maintained by IPTC, and which have been provided (in the initial NewsML release) for illustrative purposes. Taking these factors into account agreement was reached on the following additions:

MediaType - ComplexData added to deal with file formats other than the standard media types, for example Adobe PDF and Microsoft Word. It was also pointed out that at

present there is no entry to cover multimedia, but again this can be handled by ComplexData - which has now been defined as "Software application or other composite data".

OfInterestTo - new Topic Set to be developed based on world regions, generally corresponding to regional division used by the United Nations. Where more detailed information is called for the OfInterestTo TopicSet can be repeated, and so used to reference provider-specific sets in addition to the IPTC ones.

ColourSpace - additional entry 'Bi-Level' to cover the monochrome space typically used for facsimile transmissions.

Geopolitical entities - following a request from the NewsML Support group it was agreed to investigate the possibility of producing a new TopicSet to cover political/geographic groupings that are typically larger than city but smaller than a country. A paper outlining a possible approach has since been produced by David Allen and circulated as NMA02/08.

Subject Reference System

Turning to the Subject Reference System (SRS), the updated and reworked Guidelines were approved for release, and are now available on the IPTC Web site.

Members were reminded that proposals had been made for a general tidy-up, with particular emphasis on rationalising the use of capital letters, singular and plural entries, and the use of English and/or US-English.

Working Party Chairman Stéphane Guérillot pointed out that the list had seen a significant growth with time so it might be appropriate to look at the possibility of carrying out some standardisation. An offer (by Walt Barranger, of the NYT) to produce a revised list was accepted, and since the list is maintained as a TopicSet using XML, it should be possible to deal with language variations using xml.lang attributes.

A number of additional entries for the SRS had been submitted for consideration. Many of these had come from UPI and John Minting explained that his organisation had been making significant use of the Subject Reference System over the past six months, and their requests were the result of practical experience from the editorial departments in applying the codes to news items. There was general agreement that most of the proposed new entries were appropriate, but that care was needed to ensure that they were added at the right level and in the right areas.

After some discussion it was agreed to include *Medical Specialisations* as a Subject Matter heading, with Subject Details to include *Geriatrics, Pediatrics, Reproduction and Genetics*. *Healthcare Policy* was also

included as a Subject Matter. Although *Ethics* had also been proposed for inclusion in the medical area it was decided that this had a much broader application and will need to be reconsidered. Under *Science and Technology* additional entries were made for *Electronics* and *Information Technology*, while *Global Warming* was included under *Environment*.

An ad-hoc meeting has been arranged for May (in Washington) to review the revised list, and consider additional submissions with the aim of presenting an updated version of the SRS for consideration at the AGM.

NITF Maintenance

Following the approach adopted by the NewsML Support Group (and to maintain a constant approach in the standards) it was agreed to modify the date time format to allow use of the 'Z' suffix to indicate that is no time offset in the UTC format - a similar approach will be adopted for SportsML. In addition where no time information is available - for example when formatting old data - the entry 000000 (*hmmss*) can be used.

Another issue related to NewsML was the inclusion of news management features in the NITF, mainly for when the NITF is being used on its own. This will be achieved by adding a 'management status' element to docdata, having allowed values of Embargo, Cancelled, Withdrawn and Usable, with the default value being Usable.

However, it was recognized that this approach also raises general points about the relationship between the two standards and so will be looked at in more detail. One possibility is to include a guide to make it clear that when the NITF is used with NewsML, priority will be given to NewsML operations. Formal approval of these changes is planned for the Annual General Meeting in July.

Although the development of a NITF Schema has been proposed it was decided that this would not be a particularly high priority, but will remain under consideration.

A number of points have been raised by members of the NITF newsgroup, including a request for extensions to improve the handling of obituary information. This was generally seen as an appropriate addition and will be dealt with by the addition of a 'gender' attribute to 'person'. Other points

PR Committee

Promotion of NewsML is a continuing activity and Johan Lindgren (TT) help to highlight the need. He explained that there is a lot of interest in NewsML from Swedish news organisations and educational organisations - but most of them still think it is a complete news solution that includes (for example) text handling features. To help overcome such misconceptions it is important for IPTC (and individual members) to produce application examples showing how the content is structured separately - say by using the NITF.

Reuters have had a NewsML training package produced for their editorial and technical staff and intend to make copies of the package available to IPTC members. The modular learning package may also be available commercially from Wide Learning, who produced the package for Reuters.

Some reliability problems had been encountered with the NewsML site, but it is hoped that these have been overcome by moving to a fully managed Web service using improved hardware.

In a separate presentation Naoki Onodera (NSK) described some of the efforts being made to promote NewsML in Japan. Activities by NSK started in 2000 and there are now more than 760 participants directly involved, with a further 50 on the mailing list. These include some 25 newspapers and news agencies as well as equipment manufacturers, vendors and application developers, together with some involvement of broadcasters.

As well as direct application development, it has been found that the meetings are a valuable way of helping vendors understand the specific needs of the news production process.

The first step was production of a set of Guidelines for Japanese NewsML users - at the time the only specific news format in use in Japan was the NSK Tiff image format. The requirement was emphasised by the fact that Kyodo Press were planning a new XML based system and the Guidelines were seen as being important for users planning to comply with the Kyodo developments.

Current developments include: Mainichi Press delivery of NewsML to Internet users; the Yomiuri newspaper has been using NewsML since 2001; and Kyodo will be moving to full NewsML delivery in 2003, with internet delivery this year.

looked at included: mapping IPTC Subject Codes to the Dublin Core; the handling of section titles within documents (as for magazine applications); use of the 'regsrc' attribute (for locating the source of data); and inclusion of straps, which appear before a headline.

The handling of copyright was examined in some detail, starting from the question 'What is the meaning of copyright in the NITF context?'. Rights information can appear in a number of places in the standard and there may be a number of different rights (creators, distributors and so on) in a single item. It is seen as particularly important with regard to re-processing and archiving and within documents that have been put together from a number of sources.

Discussion resulted in agreement to start looking at a rights model that will take account of current practice. This would also have to take allow for regional differences. It was pointed out that similar considerations will apply to NewsML.

A spreadsheet to allow mapping of components from NITF to NewsML is still under development, as is the coding of an XSLT stylesheet for NITF conversion to NewsML. These will be circulated for consideration, with the aim of reaching agreement on the approach which can then be refined ready for publication as the approved conversion method..

Members were informed that an object-orientated PERL module designed to let users manage NITF 3.0 documents without any specialised XML or NITF knowledge has been released - see <http://cpan.perl.org/authors/d/B/BQ/BQUINNS>. Where possible links to such developments and NITF publishers will be included on the NITF Web site (www.nitf.org), while there are also plans to add more sample documents.

Standards Committee

Summing up progress at the start of the Standards Committee session - as normal, the final one of the Meeting - Chairman

John lobst (NAA) said that many of the other sessions had been concerned with groundwork so there were no substantial developments for review or approval on this occasion.

It was reported that contact had been made by a non-member who was having some problems with news material that was being provided from a number of sources in the form of mails. These had very little structure so processing was difficult and the enquirer wanted to know if IPTC could provide some help or guidance to improve the news flow.

Although the request was from a non-member, John lobst said that he thought that IPTC's aim of providing technical standards for the news industry means that there is some responsibility in this area. Accordingly he suggested that, if possible, it might be appropriate for IPTC to offer some

assistance. This would take the form of a simple tool (or template) which could be run on a PC and convert text into NITF ready for transmission by e-mail. It was suggested that members should be approached to see if any of them may have suitable software that they could make available.

With NewsML well established, with systems in use and under development by both IPTC members and outside organisations there is a need to establish a policy for handling updates.

As part of this a notice was posted on the NewsML mailing list to solicit proposals, with the aim of moving towards a new release (possibly at the Autumn Meeting). However, for proposals to be considered they would have to have the support of at least one IPTC member - it was also pointed out that interested parties can become more involved by joining IPTC.

Other Presentations

A brief update on developments in the ICE (Information and Content Exchange) Protocol was provided by Alan Karben, following a meeting of the ICE 2.0 Authoring Group. He explained that ICE is an XML based standard that can provide an automated syndication handling for on-line businesses.

However, it appears that the standard is facing strong competition from other web service standards (such as UDDI and SOAP), with syndication becoming a big business. The recent ICE Authoring Group session concentrated on the relationship between ICE and Web services, with work having started on an ICE 2.0 standard that will make use of Schemas and Namespaces. See www.icestandard.org for more information on the standard.

Revenue Generation and Rights Management was the theme of an interesting presentation by John Moroney from RightsCom. This subject is gaining increasing importance - for both NewsML and the NITF - and more detailed coverage will be included in a future issue of the IPTC Mirror.

Dr Angelo Marrara (ANSA) gave a brief demonstration of work in progress on a NewsML editor incorporating an update feature. News items are converted from the original format into NITF and then put into NewsML as the delivery container. With photographs the IIM record is also converted for use. Once the original item has been delivered the system makes it possible to gradually build the story, adding analysis, images and related material as well as updates to the main news item as the coverage develops. Coverage can be built up over different timescales to suit the story, such as a day or a week.

Update

Latest papers issued

CAT 02/03	SRS update
CWC 02/01	Paris Minutes
DIR 02/01	Paris Schedule
MAN 02/06	Paris Minutes
NM 02/06	Conversion Ssheet
NMLS 02/10	UPI Dateline
NMA 02/08	Gewopolitical TSet
NMG 02/02	Paris Minutes
PRC 02/02	Paris Minutes
SPC 02/06	SportsML TSet
STA 02/04	RightsCom

Published by the International Press

Telecommunications Council
Royal Albert House, Sheet Street,
Windsor, Berkshire SL4 1BE
England

Managing Director: David Allen

Editor: Hugh Johnstone

TEL: +44(0)1753 705051

Fax: +44(0)1753 831541

Email: m_director_iptc@iptc.org

editor_iptc@iptc.org

Web Pages: www.iptc.org; www.newsml.org;
www.nitf.org

IPTC Documents

IPTC-NAA Information Interchange Model
and
IPTC-NAA Digital Newsphoto
Parameter Record

IPTC Recommended Message Format 7901

News Industry Text Format

IPTC Satellite Systems for Press Use