

40th IPTC ANNUAL GENERAL MEETING 2005

held at the

The London Bonnington Hotel in Bloomsbury 92 Southampton Row, London WC1B 4BH, United Kingdom

6rd to 9th June 2005

Minutes of the News Architecture Working Party held on 6th June 2005

Document history				[Document URN: urn:iptc:workdoc:nar:0508:1]
Version	Issue Date	Pages	Author (revised by)	Remark
1	2005-07-14	6+18	Michael Steidl/	

Present:

Chairman: Laurent Le Meur, Agence France Presse

Walter Baranger, New York Times Company

Scott Calder, Mainstream Data Inc. Dominic Chan, CNW Group Ltd.

Shun-Ping Chen, Deutsche Presse-Agentur GmbH

Dave Compton, Reuters Limited Honor Craig-Bennett, PA NewsLtd Takahiro Fujiwara, EAST Co. Ltd Andy Gleeson, BBC Monitoring

Stéphane Guérillot, Agence France Presse

Darko Gulija, HINA

Paul Ian Harman, PA NewsLtd

Geoffrey Haynes, The Associated Press Kelvin Holland, Harris and Baseview Rudi Horvath, Austria Presse Agentur Chris Hugh-Jones, Fingerpost Ltd

John Iobst, Newspaper Association of America

Hugh Johnstone, IPTC

Olli Kemppainen, Oy Suomen Tietotoimisto Zoltan Kirchknopf, Magyar Távirati Iroda Rt

Michael Steidl, IPTC Managing Director as Secretary

Dean Large, Business Wire

Harald Löffler, ifra

Jayson Lorenzen, Business Wire

Angelo Marrara, ANSA

John Minting, United Press International

Peter Müller, SDA/ATS

Stuart Myles, Dow Jones & Company

Jean-François Richard, Agence France Presse

Hitoshi Saito, Nihon Shinbun Kyokai Tetsuya Sato, Nihon Shinbun Kyokai Masaki Satsuka, Nihon Shinbun Kyokai Wade Sendall, New York Times Company

Hiroshi Shinotsuka, Kyodo News

Henrik Stadler, Tidningarnas Telegrambyrå

Tom Taylor, PR Newswire

Evi Varsou, Athens Technology Center S.A.

Misha Wolf, Reuters Limited

1 Minutes of the Meeting held on 3rd March 2005 (NAR0502.1)

It was proposed by Honor Craig-Bennett and seconded by Walter Baranger that the minutes be accepted as circulated. This was agreed unanimously.

2 Matters Arising

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

3 Chairman's report

The Chairman said he gave a presentation about the News Architecture at the News Standards Summit in Amsterdam in May.

NAR0508.1-AGMminutes.doc Page 1 of 6



4 Overview of the current state of development work

The Chairman gave a presentation on the current work – see "Appendix 1"

Discussion:

Walter Baranger said for the testing stage an intermediate cut-off date should be set to get first results in time to adapt the specifications and the documentation for final approval. The Chairman agreed.

The Managing Director made clear after the work on NAR implementation issues over the summer there will be deliveries to be approved at the Autumn meeting in October. Therefore the work of the NAR WP can be considered as being completed by this Autumn meeting. Based on that approved specifications NewsML 2 and EventsML can be built further down the timeline.

The Chairman agreed.

Walter Baranger asked whether people from outside of the IPTC should be invited to test the NewsML 2 and EventsML implementations – but this has to be decided at the Autumn meeting.

The Chairman said he intents to release drafts of the NAR model documents on the public Yahoo-Groups by the end of June to solicit for comments. But he expects not much of them as implementers are primarily interested in specifications like XML Schemas.

Geoffrey Haynes asked whether the standard will be approved at the Spring meeting without a documentation?

The Chairman said the time of the approval and the public release have to be disambiguated: at the Spring meeting the technical specifications should be available for approval, but for the public release some supplemental documentation is required.

And we have to disambiguate what documentation will come from the NAR – it will be adopted by standards inheriting from its framework, like NewsML 2 and EventsML – and which part of the documentation will be specific to a standard.

But the only documentation to the outside will be the one delivered with a standard.

Misha Wolf said as NMDF chair he thinks about releasing NMDF as a technology because we want people to consider building NMDF into their standards.

The Chairman made clear a technology building block is different from a standard – the IPTC will only approve a standard as a whole, not its building blocks.

Misha Wolf disagreed saying anything approved by the IPTC can be considered as a standard as the IPTC is a standardisation body.

5 News Management Working Group

The lead of the News Management WG, Stuart Myles, gave a presentation on the current work – see "Appendix 2".

Regarding slide 7 he asked the members present about that exchange models they use:

- Push/Broadcast: "pretty much"
- Pull/Syndication/Conversation: "less"
- Both: "a lot"

(No counting of votes was done)

Finally the lead asked members to get involved, to join the News Management Yahoo group and – not to forget – read the existing postings there – and to return comments.

And he pointed at the conference calls as a forum of ongoing NAR discussion.

NAR0508.1-AGMminutes.doc Page 2 of 6



The Managing Director emphasised the work on News Management is crucial now: defining structure is only half of the job that has to be done, the News Management group has to define a processing model and the NAR specification can't be implemented without having this defined thoroughly.

Misha Wolf showed the "Item matrix diagram" (see "Appendix 7") which is an approach to disambiguate the different use scenarios for items and hence different processing model. It was agreed this model requires further discussion but would be of help to specify the item specific processing models.

6 News Structure Working Group (Laurent Le Meur)

[see: DRAFT-NAR_1.0-spec-NewsStructure-Model_8.pdf]

The lead of the News Structure WG, Laurent Le Meur, gave a presentation on the current work – see "Appendix 3".

Discussion:

Stuart Myles asked why Laurent Le Meur considers other news exchange format models as more complicated.

Laurent Le Meur said he had NewsML 1 and MPEG 21 in mind and they are definitely not that refined as the current NAR model.

On the other hand Atom is a much simpler model of a "feed" but this is a NAR PackageItem without eg. any ordering feature and the Atom entries are less powerful compared to the NAR items, for instance they only can reference one piece of content, and they are not able to manage any revisions.

Dave Compton recommended creating a comparison chart of features for standards like RSS, Atom and NewsML 2 to show to the outside where the IPTC provides more features than others. And further a document should deal with interoperability of RSS/Atom and the new IPTC standards to explain how to integrate them into one system.

Walter Baranger said regarding RSS/Atom feeds vs. IPTC standards professional news companies like the New York Times have software engineers able to write software for all kinds of incoming feed but small companies can't do this.

7 News Metadata Framework

[see: DRAFT-NAR_1.0-spec-NMDF-BusReq_30.pdf]

The lead of the News Metadata Framework, Misha Wolf, gave a presentation on the current work – see "Appendix 4".

Misha Wolf said his intention is to integrate NMDF into the world wide semantic web. Then he continued with giving an overview of the current version of the NMDF requirements document, version 32.

Laurent Le Meur commented on how the group created these requirements. It started with basic requirements from Alan Karben, Misha Wolf and Arnaud Descamps. Then it became clear the NDMF has to cover two areas: how to contain metadata values but also how to maintain taxonomies and thesauri by adding information about a concept, that's the reason for having TopicItems. Tying these TopicItems together can form a semantic web like it is intended by Misha Wolf.

NAR0508.1-AGMminutes.doc Page 3 of 6



Then technical issues regarding a decent syntax for the NMDF model were discussed and Mark Birbeck, present as guest, proposed to consider the draft RDF/A notation.

Walter Baranger asked how the IPTC can make sure the industry will follow our decision on the NMDF implementation.

Laurent Le Meur said if the IPTC follows W3C there will be tools available to process the syntax.

Darko Gulija pointed at the controversial requirements for "a minimum in size" and being "well structured". The fancy XML Misha Wolf provided in draft version 32 merges several RDF statements into a single structure.

8 Common Components Working Group (TBA)

[see: DRAFT-NAR_1.0-spec-CommonComponents-Model_10.pdf & DRAFT-NAR_1.0-spec-CommonComponents-CoColist1.pdf & DRAFT-NAR_1.0-spec-CoCo-ManageGL_1.pdf]

As stand-in for the Common Components (CoCo) WG lead Johan Lindgren, Jayson Lorenzen gave a presentation on the current work on CoCos – see "Appendix 5".

Discussion:

The Managing Director explained the two stages of discussion and approval implemented for the CoCo discussion:

- firstly for a metadata container its semantics and cardinality are defined and a name is proposed.
- secondly a datatype is assigned which can either be a NMDF class or another CoCo, like for persons or locations.

This was done to sort out the discussions on "what this metadata container should cover" and on the container's technical implementation.

Regarding Rights Component:

Walter Baranger proposed to seek advise from an attorney to avoid any pitfalls while implementing the rights component.

The NAR Chair said currently there is no intention to create an elaborate rights management component, we rather think about adopting something from the outside.

Walter Baranger responded saying then the four most prominent rights management schemas should be assessed for adoption.

Dean Large asked how the group came to the current components: did it survey all the existing components from previous standards or did it start from scratch?

Jayson Lorenzen replied he thinks the group covered all structures from previous standards.

The NAR Chair said the group is open to any suggestions.

The Managing Director reminded of the working process of the CoCo group:

The group only acts on request from either other NAR WGs or from content standard WGs. The CoCo group already received a couple of requests for CoCos from the EventsML group like

for person and location.

The NAR Chair added the group also considers adopting standards from the outside, like it was said for a Rights Component but also e.g. for a Signature Component.

NAR0508.1-AGMminutes.doc Page 4 of 6



Stuart Myles said CoCo work is exposed to being more and more extended by proposed components and he asked whether the large area of work on CoCo needs prioritisation.

The NAR Chair said he expects the components currently under discussion will be the ones proposed for approval at the Autumn meeting.

Geoffrey Haynes asked if the CoCo group will have a full set of recommendations for key CoCos like rights, persons, locations?

The NAR Chair said he is pretty sure a draft for all components will be available by October – except for rights as this is a very tricky issue and maybe there will only be placeholder like a string to put there a copyright notice.

Geoffrey Haynes made clear for developing EventsML having approved CoCos available is a prerequisite.

Darko Gulija added the same applies for NewsML 2, it can't be built without having CoCos like Description or Labels.

Geoffrey Haynes said a group had a discussion at lunch this day and he proposed their issue for further consideration:

- to include a "secure identifier" into the content of an item
- this would make it possible to identify the content when it "floats around" outside the expected distribution paths
- and this would enable an aggregator system to cancel such a story on request.

The NAR Chair said this is something which is under consideration regarding the Signature component. As a signature would guarantee that the content was not changed since the originator signed it and if a syndicator changes the content to include this content into its feed the content has to be signed again by the syndicator. So it would be perfectly clear where the content comes from.

Then the source of the problem was discussed: is the problem "stealing" content (like scrapping content from a website), or is it the "leaking" of distribution channels when providers send content to receivers it should not go to.

The Managing Director said any current security system is based on a combination of static information – like an id or key – and processing modules for this information. This applies e.g. to all digital rights management systems. Hence we have to think of not only specifying a static id but also a processing model for implementing this level of security.

Walter Baranger added the IPTC has to think about how to enforce such a system and has to be prepared to back it up on a legal level.

The NAR Chair said the IPTC should not forget about the scope of this problem: a news agency could circulate an item to thousands of receivers, where should they get hold of the "leakage" of items to the outside of the business relations.

Paul Harman pointed at the nature of usage rights: they are defined by individual contracts between the provider and the consumer and hence they can't be defined in an item as "absolute" rights.

9 News Architecture Implementation:

[see: NAR_1.0-doc-ArchitectureImplementationGuidelines_2.zip]

NAR0508.1-AGMminutes.doc Page 5 of 6

RIC)

IPTC Document: NAR 0508.1

The Chairman showed a proposed statement of the NAR WP regarding the delivered Architecture Implementation Guidelines – see "Appendix 6".

He went over the statement.

As nobody raised objections this statement was considered by the Chairman as being agreed unanimously.

10 Any Other Business

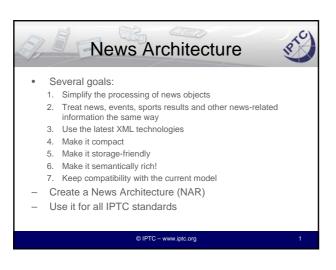
There was no other business.

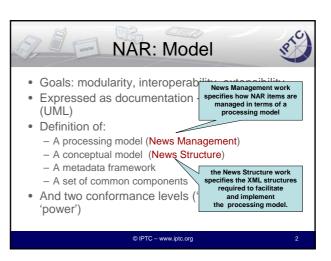
11 Date and Place of Next Meeting

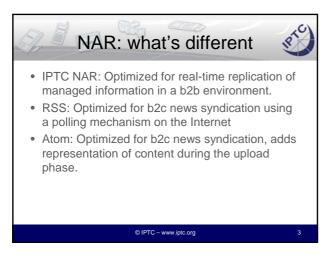
24 – 26 October 2005, Milan, Capitol Millenium Hotel

=== END of document ===

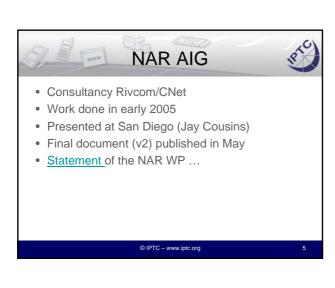
NAR0508.1-AGMminutes.doc Page 6 of 6

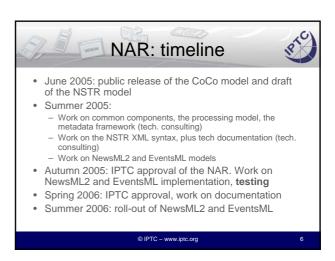












NAR News Management

IPTC ANNUAL GENERAL MEETING Monday, June 6, 2005

Stuart Myles
Dow Jones & Company Inc
smyles @WSJ.com

NAR: News Management

- Agenda
 - News Management in the NAR
 - Different Conformance Levels
 - Lots of Unresolved Topics
 - Next steps and how you can help

2

News Management

- Makes use of the Management Component
 - Identifiers
 - Type and class
 - Status and state
- Relations between items
- News envelope

3

Different Conformance Levels

- News processing is where the differences come in to play for
 - Core profile
 - Power profile
 - Extensions

,

Unresolved Topics

- We have established lots of areas within processing and management that need work
- Feedback on importance and relevance?
- We need your help in resolving them!

5

Versioning

- What are the boundaries of versioning?
- How are identifiers used in versioning?
- When does something change so much that it becomes something different?

6

Exchange Models

- What are the "exchange models" that we should support?
- Push / Broadcast
- Pull / Syndication / Conversation, e.g. ATOM, ICF
- Different flavours of each?
- How does aggregation come into play? E.g. can an aggregator version a news item? Can they add their name?
- Survey says?

Rights and Signature

- How are digital rights and digital signatures to be handled in the processing model?
- What meaning should be ascribed to the rights and signatures?
- To what extent must the processing model be specified in these areas?

8

Error processing

- Invalid structure
- Invalid metadata
 - Metadata that is unknown to a recipient
 - Metadata that is known to be "bad"
- Different error handling in different exchange models?

9

Item type specific processing

- News
 - How to choose between alternative versions of news items
- Metadata
 - Resolution of references
 - Validation of metadata against a controlled vocabulary

10

Item type specific processing

- Events
- Persons
- Alerts
 - Are Events, Persons and Alerts special types of News Items?
 - Are they processed differently than other kinds of News Items?
- Package Item processing

11

We need your help

- The processing model is important
- By definition, it can't be the work of just one person or organization
- Let's work through each of the unresolved areas together
- Pick an area that you would like to help with

12

What we will end up with

- A unified management model for all Item types and exchange models
- A description of how to process news
- Simpler implementation for both producers and consumers
- Wider adoption of IPTC standards
- More customers...

13

Ways to Get Involved

- Join the email lists
- Send email to the lists
- Attend the working calls
- Volunteer to become the "vice chair" of the News Management group

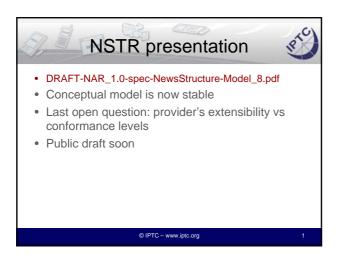
14

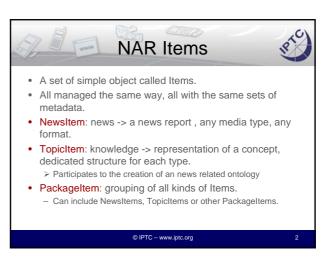
NAR News Management

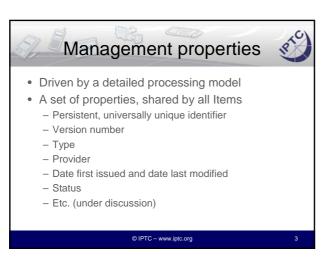
■ Discussion?

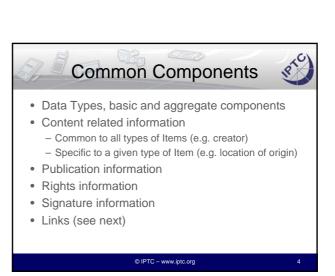
■ Volunteers?

15

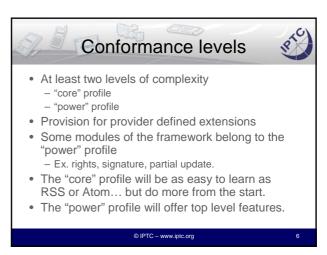


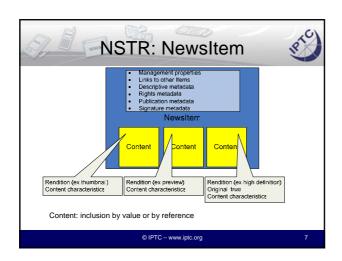


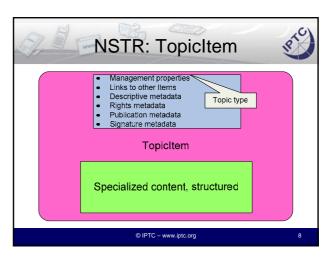


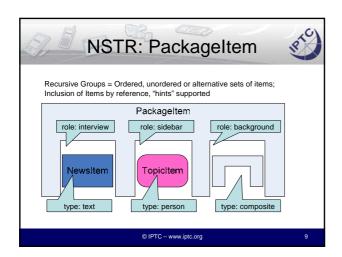




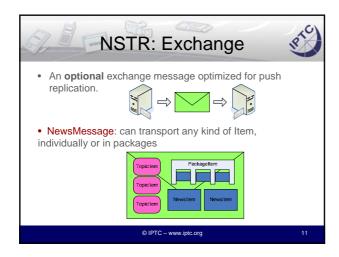


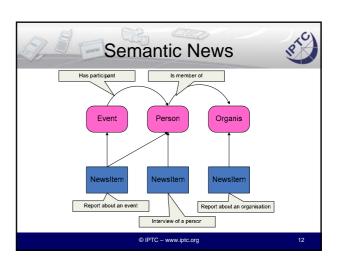












NMDF Goals

- · Be easy to understand and use
- Support a spectrum of needs, from lightweight to heavyweight
- Form a part of the World Wide Semantic Web



Metadata Classes

- · A plain string value (eg "Camille Claudel")
- A {scheme, code} pair representing a concept in a taxonomy (eg "iso4217" and "USD")
- A structured record or a pointer to such a record (eg info structured using the NewsML 1 Property element)



Metadata Classes

- Strings representing a date, time, duration, size, etc (eg "1864-12-08")
- Codes representing concepts in a taxonomy and controlled through an XML schema language (eg "F" or "M")



Metadata Containers

- Containers explicitly defined by various IPTC standards, eg a Creator element
- Containers created through the markup of text within a label, eg <foo bar="...">Bill Clinton</foo> has been appointed ...



Metadata Targets

- Targets explicitly defined by various IPTC standards, eg the content component of a NewsItem
- Targets created through the marking-up of text within a label, eg <foo bar="...">Bill Clinton</foo> has been appointed ...



{Scheme, Code} Pairs

- · A scheme is a taxonomy
- It may be very simple (a flat list) or very sophisticated (a thesaurus with typed relationships between nodes)
- Each scheme is identified by a URI



{Scheme, Code} Pairs

- As URIs are long and quite difficult to use in XML, for each scheme URI we define a scheme alias
- For example:
 - http://www.iptc.org/schemes/iso4217#
 - iso4217



{Scheme, Code} Pairs

- Each node in a taxonomy is represented by a code (eg "3", "USD", "XYZ92", "Interview")
- Each node is represented by a URI, obtained by concatenating the code to the scheme URI, eq:
 - http://www.iptc.org/schemes/iso4217#USD



QNames

- · QNames are Qualified Names
- They are introduced by Namespaces in XML and are included in the XML Schema specification
- An example of the QName syntax is "iso4217:USD"



Fantasy Examples

- <dc:subject val="SRS:15062000"/>
- <dc:subject val="SRS:15062000" xml:lang="en">Swimming</dc:subject>
- <dc:subject val="SRS:15062000" parent="SRS:15000000" xml:lang="en">Swimming</dc:subject>



Fantasy Examples

<dc:subject val="SRS:15062000"
 parent="SRS:15000000"
 xml:lang="en">Swimming</dc:subject>
 <dc:subject val="SRS:15000001"
 qualifies="SRS:15062000"
 xml:lang="en">Men's</dc:subject>



Other Properties

 <dc:subject val="SRS:15062000" assignee="afp:llm" confidence="50" relevance="50" when="2005-06-05T12:34:56" >Swimming</dc:subject>







Common Components

(DRAFT-NAR_1.0-spec-CommonComponents-Model_10.pdf)





WWW





State of the work

- DRAFT-NAR_1.0-spec-CommonComponents-CoColist1.pdf
 - Data Types for use by all Common Components, recommended by the Consultants.
 - Component Template Proposed by the Consultants
 - List of Components identified at this time
- Component Discussion Documents

Data-types

3.1 Description

This is a list of datatypes that we see a need for and a description of what they need to cover. With further work it might be that some are covered by the same technical solution.

3.2 List of suggested data types

Date Only date information.

DateRange A range with from and to dates

DateTime A complete date and time description.

DateTimeRange A range with a from and to both date and time

LanguageString A string with information of the language used.

Multiple Choice A choice from some valuelist that is not changing

and not IPTC-specific. Like yes/no.

Named String A string with an attribute stating what the string value represent

NonEmptyString A string value but not empty.

Number Various types of numerical values, like integer, decimal,

non-negative integer etc.

PartialDateTime A date time description where maybe only part of the

information is known.

String A string of any length String xx String of a maximum

(and minimum) length. Not known at the moment exactly

which combinations are needed?

Time Only time information.

TimeRange A range with a from and to time

Unit Value A numerical value with information about the measurement unit.

Value list A list of values probably maintained by IPTC and seldom updated.



Component Discussions

- Identified Components assigned to a volunteer
- Discussion Documents produced

http://groups.yahoo.com/group/iptc-common-components-dev/files/CoCo4Discussion/

- Docs Discussed during weekly Conf. Calls
- Docs Discussed via COCO yahoo group
- Documents evolve into Component's Specification.



Discussion Documents produced or with an owner assigned

- SignatureComponent (Jayson)
 CoCo-Discuss-Signature_1-2.doc
- RelationshipComponent / AssociationComponent (Darko)
- LabelComponent (Jayson, Takahiro)
 CoCo-Discuss-Label 7.doc
- DescriptionComponent (Michael) CoCo-Discuss-Descr 4.doc
- MetadataAssignmentComponent (Misha) (now called CommonAttributes)



Discussion Documents produced or with an owner assigned (Continued...)

- Topic related structures
 - Event (Johan)
 - Person (Johan)
 - Organisation (Johan)
 - Contact information (Johan)
 CoCo-Discuss-Cont 1.doc
 - Location (Johan)CoCo-Discuss-Loc_1.doc
 - Address (Johan)CoCo-Discuss-ADDR_1.doc



What will happen next?

You Join the Common Components Mail List and help create Common Components



http://groups.yahoo.com/group/iptc-common-components-dev/

Email: johan.lindgren@tt.se

Statement of the NAR WP about the NAR Implementation Guidelines (AIG)

Overview

The guidelines, written by Jay Cousins (RivCom) and Ulf Wingstedt (CNet) are found in the document: NAR_1.0-doc-ArchitectureImplementationGuidelines_2.doc referenced as: urn:iptc:std:NAR:1.0:doc:ArchitectureImplementationGuidelines:2 This document (second version) has been released on the 2005-05-03.

This document introduces the NSTR model, and studies the implementation of the NSTR model in XML.

It details a proposed model for common components, and design rules for implementing common components in XML.

It also details proposals in the area of validation (validation of news messages), extensibility, schema evolution and versioning, conformance levels.

It offers templates of XML schemas for the core classes of the model.

It describes a spreadsheet template for the definition of common components.

Statement

The NAR WP acknowledges that the AIG meets IPTC expectations.

The NAR WP has decided to adopt the model of common components proposed in this document. Therefore the part of the AIG corresponding to this model is directly reflected in the CoCo model: DRAFT-NAR_1.0-spec-CommonComponents-Model_10.doc.

The NAR WP has decided to adopt several recommendation of the AIG:

- The use of a specific namespace for each specific item class.
- The use of a unique namespace for common components.
- The schema version management

The NAR WP has decided not to use, at least for the short term, the CoCo spreadsheet template, which seems requesting too many details.

The NAR WP will walk through the AIG in the course of its work, and decide which recommendations to adopt, and which need some adaptation.