



# IPTC AUTUMN MEETING 2005

held at the  
Capitol Hotel Milan  
Via Cimarosa 6, 20144 Milan (Italy)

**24 to 26 October 2005**

## Minutes of the NewsCodes Working Party held on 25 October 2005

Document history				[Document URN: urn:iptc:workdoc:ncd:0515:1]
Version	Issue Date	Pages	Author (revised by)	Remark
1	2005-11-11	5	Michael Steidl/John Minting	

### Present:

Chairman: John Minting, United Press International

Walter Baranger, New York Times Company

Dave Compton, Reuters Limited

Honor Craig-Bennett, PA NewsLtd

Takahiro Fujiwara, EAST Co. Ltd

Walter Grolimund, Keystone

Stéphane Guérillot, Agence France Presse

Darko Gulija, HINA

Geoffrey Haynes, The Associated Press

Klaus Herwig, Deutsche Presse-Agentur GmbH

Niels Hojer-Pedersen, Ritzau Bureau I's

Rudi Horvath, Austria Presse Agentur

John Iobst, Newspaper Association of America

Hugh Johnstone, IPTC Editor

Olli Kemppainen, Oy Suomen Tietotoimisto

George Landau, NewsEngin, Inc.

Dean Large, Business Wire

Laurent Le Meur, Agence France Presse

Johan Lindgren, Tidningarnas Telegrambyrå

Michael Steidl, IPTC Managing Director as Secretary

Harald Löffler, ifra

Jayson Lorenzen, Business Wire

Takis Mantis, ANA, Athens News Agency

Angelo Marrara, ANSA

Hironori Mizoguchi, Kyodo News

Peter Müller, SDA/ATS

Stuart Myles, Dow Jones & Company

Hitoshi Saito, Nihon Shinbun Kyokai

Hiroshi Shinotsuka, Kyodo News

Henrik Stadler, Tidningarnas Telegrambyrå

Charles Tichenor, The Associated Press

Tadashi Tsuji, Nihon Shinbun Kyokai

Evi Varsou, Athens Technology Center S.A.

Misha Wolf, Reuters Limited

As Guest:

Paul Kelly, XML Team Solutions

Pauli Tölli, Oy Suomen Tietotoimisto

### 1 Minutes of the Meeting held on 7 June 2005 (NCD0511.1)

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 there had been an increasing amount of activity regarding NewsCodes in recent months.

While there is a lot of interest both within and outside the information provision area it is felt that the IPTC vocabularies are considered as being inflexible and outdated by parties outside the IPTC, hence the NewsCodes need an urgent review. The NCD WP has also to consider the issues of integrating NewsCodes into the upcoming News Architecture.



John Iobst said he fully supports any efforts to investigate new technologies and adopting them for the IPTC.

Walter Baranger encouraged the WP to study the current situation – the New York Times decided not to use the Subject NewsCodes because of their limitations. He pointed at the architectural constraints as limiting the expression of cross relationships and synonyms.

Misha Wolf said the architectural considerations pertain to the work of the News Metadata Framework WG – so please send any requirements to design relationships between topics properly to the NMDF Yahoo list (<http://groups.yahoo.com/group/iptc-metadata-dev> )

Geoffrey Haynes said the AP has hired Sarah Goldman, a taxonomist, and she is willing to contribute to the work of the IPTC.

The chairman said he has spoken to LexisNexus representatives about specific vocabulary issues – and he has asked them to join the IPTC again to move things forward.

George Landau reminded the WP of his categorisation presentation at the Spring Meeting 2005 and proposed that IPTC think about a public server providing a service to categorise news content. From a web interface users could input their text obtaining a set of IPTC category codes relative to that text .

The Managing Director said the IPTC policy is to support any kind of implementation of IPTC standards by any means, but there is no budget for such a project.

George Landau responded that he didn't expect financial support from IPTC, but asked that IPTC should think about how a project such as this could work.

#### **4 Approval of submissions / Report about submissions approved by the Fast Track Jury**

The chairman said UPI has submitted 20 international organisations to be added to the Subject NewsCodes – but this had been rejected by some members because these are entities and not categories.

At a conference call of the Fast Track Jury it was proposed to create a new specific TopicSet for organisations. The chairman said he thinks this should have a similar treelike hierarchy as the Subject NewsCodes.

Walter Baranger said dealing with entities opens up a very complex work.

The chairman said entities like these are required for the work of news agencies like UPI.

Laurent Le Meur said we have to be aware of the principles of the News Architecture, – not to add entities if this is something completely different from the abstract concepts of the NAR. He added Organisation NewsCodes is something very interesting.

The chairman said AFP has suggested links between the “International Organisation”-Subject NewsCode and this Organisation NewsCodes – how could this be achieved within XML

Misha Wolf pointed out Reuters uses only the Property structure of NewsML 1 and this allows any kind of relationship between topics – he showed some examples of this.

John Iobst questioned a new Organisation vocabulary with the same outdated hierarchical structure as the Subject NewsCodes.



## 5 NewsCodes translations – update

Currently the main international languages – French, Spanish, German – are up to date, at least up to the version before the current reference version.

Darko Gulija asked if HINA has to make a full translation to be accepted by the IPTC. The chairman said this is not required for a translation for local use only.

## 6 Document specifying “deprecating topics”

The chairman said he made a first draft, which was reviewed by the UPI editorial and then sent to the Managing Director who sent comments on this draft:

1<sup>st</sup> Version

Deprecation:

In a world that continues to change, there will be a continuing need to remove, relocate or replace existing terms in the IPTC Topic Sets.

In order to support legacy content items that may contain a topic set term that needs to be changed, the original term, its numeric code and explanation must remain as part of the topic set, but must be labelled as “Deprecated – not to be used after dd/mm/yyyy”

UPI Editorial’s version

Deprecation:

In a world that continues to change, there will be a continuing need to remove, relocate or replace existing terms in the IPTC Topic Sets.

In such cases, the original term, its numeric code and explanation must remain as part of the topic set, but must be labelled: “Deprecated – not to be used after dd/mm/yyyy”

This version with comments from the Managing Director was reviewed in depth by the WP and changes were proposed. Finally the motion below was agreed:

MOTION to the Standards Committee:

Deprecation:

In a world that continues to change, there will be a continuing need to remove, relocate or replace existing terms in the IPTC Topic Sets.

In order to support legacy content items that may contain a topic set term/code that needs to be changed, the original term, its numeric code and explanation must remain as part of the topic set, but must be labelled in the “Change Comment” field as “Deprecated – shall not be assigned after yyyy-mm-dd”.

## 7 NewsCodes maintenance: reviewing terms and adding explanations (see NCD0514.1)

The Managing Director explained document NCD0514.1.

He said many of the inputs come from NSK which has translated all TopicSets into Japanese and found a number of minor and major issues with the original TopicSet in English.

**Action** ● It was decided to go over the list and to assign members to review specific NewsCode sets:

- \* xyzCoders/Codecs: the Managing Director should ask the EBU
- \* Provider: Walter Baranger will send a request to register to all IPTC members
- \* NewsItemType: Dean Large and Laurent Le Meur
- \* TopicType: to be reviewed by the new WG (see item 8 of these minutes)
- \* MimeType: to be replaced by a rule based vocabulary controlled by IANA



- \* Notation: ask members if anyone is using this
- \* Location: Honor Craig-Bennett
- \* Labeltype: ask members if anyone is using this
- \* HowPresent: ask members if anyone is using this and how. .
- \* Format: Chuck Tichenor
- \* Colourspace: ask members if anyone is using this

## 8 Discussion on building controlled vocabularies for the News Architecture

Misha Wolf, News Metadata Framework WG lead, made a presentation of DOs and DON'Ts for adopting the NDMF for NewsCodes:

For Thesauri:

- DO use schemes
- DO continue to use the existing Topic codes
- DON'T continue to treat the codes as representing a hardwired hierarchy
- DO develop thesauri of uneven depth, to suit the tasks in hand
- DO feel free to relocate nodes in the thesaurus
- DO consider creating sub-thesauri for distinct groups of concepts (eg Sports or Finance)
- DO allow nodes to have multiple parents where appropriate (eg Mexico, Nicosia, Russia, Turkey)
- DO make use of typed relationships for linking Topics

For NewsItems:

- DO include ancestor Topics in the News Item (replacing the hardwired hierarchy)
- DO make use of the <bag> element for grouping concepts

Two code examples:

How to express a Subject NewsCode in a NewsItem.

```
<subject code="nc:15062000"
  creator="..." created="..." type="..."
  confidence="..." relevance="..." why="...">
  <title>...</title>
  <sameAs code="...">
  <childOf code="...">
</subject>
```

How to group Subject NewsCodes in a NewsItem to replace "Qualifiers".

```
<subject>
  <bag>
    <bit code="nc:15000002"/>
    <bit code="nc:15062000"/>
  </bag>
  <title>Women's Swimming</title>
</subject>
```

Discussion:

Laurent Le Meur asked how we can obtain input from professional taxonomists to the News Architecture for designing the correct structure to express values from thesauri.

The chairman said we should try to get these professionals to one of our meetings.

Misha Wolf added we should also try to get them on our development mailing lists.



Walter Baranger proposed to create a Working Group to study the nature and structure of metadata values in relation to the News Architecture work.

The goal of this group should be figuring out what is ideal and to make recommendation how to achieve this.

**MOTION:** To establish a Working Group to study possible changes to the structure of IPTC NewsCodes, including the architecture and relationship of the Subject Codes.

The Working Group will make its first report at the IPTC Spring Meeting 2006.

The motion was seconded by Geoffrey Haynes.

**DISCUSSION:**

Walter Baranger proposed to hold a few face to face meetings in the US for a start.

The chairman added that initial conference calls would allow for international participation and pique the interest of non-members of IPTC .

**VOTE:** On vote by voice the motion was agreed unanimously.

The chairman asked who of the member representatives present wanted to participate:

Laurent Le Meur, Misha Wolf, John Iobst and George Landau showed interest.

## **9 Any Other Business**

There was no other business.

## **10 Date and Place of Next Meeting**

27-29 March 2006, Vancouver (CAN), Marriott Pinnacle Downtown Hotel

=== END of document ===