IPTC Standards "Framework"

A proposal how to build all IPTC standards in a common and unified way

Michael Steidl

31 October 2004

Preamble

The current discussion about the standardisation process at the IPTC focussed on terminology like "what exactly is a framework" – is NewsML a framework and is that what I proposed in previous documents DIR0407 and DIR0409 as a set of common elements to the IPTC standards a framework or a "toolbox".

I leave selecting the proper term for this "thing" to a future stage of our discussion and put the term "Framework" on the front page in quotation marks. I ask the reader to focus on the concepts behind the terms and to discuss these concepts first before starting to argue about terms.

Initial thoughts

- **All** IPTC standards are about **news exchange**

- = they deal with data about newsworthy information as content

- = they deal with distribution/exchange management issues

Initial thoughts

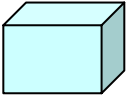- **NEWS CONTENT covered by IPTC standards:**
  - various types of content data (text, image, audio …)
  - focus on sub-regions of the broad term "news":
    - human expressions reporting about and commenting events
    - formal representation of (= metadata about):
      - events
      - sports
      - radio & tv programs
      - weather
      - elections
      - sub-items: persons, organisations, companies
      - sub-item: locations

Initial thoughts

- How to "box" news content:
  - Create a generic, basic and abstract "news box" that allows to be adapted to the needs of the "sub-regions"
    either by extension or by constraint
  - The basic "news box" is a container for any kind and type of content and facilitates a basic set of metadata describing the contained news related content
  - The basic "news box" can be expanded by optional elements to identify the content in a unique way, to associate the content with other content, and to facilitate content management.
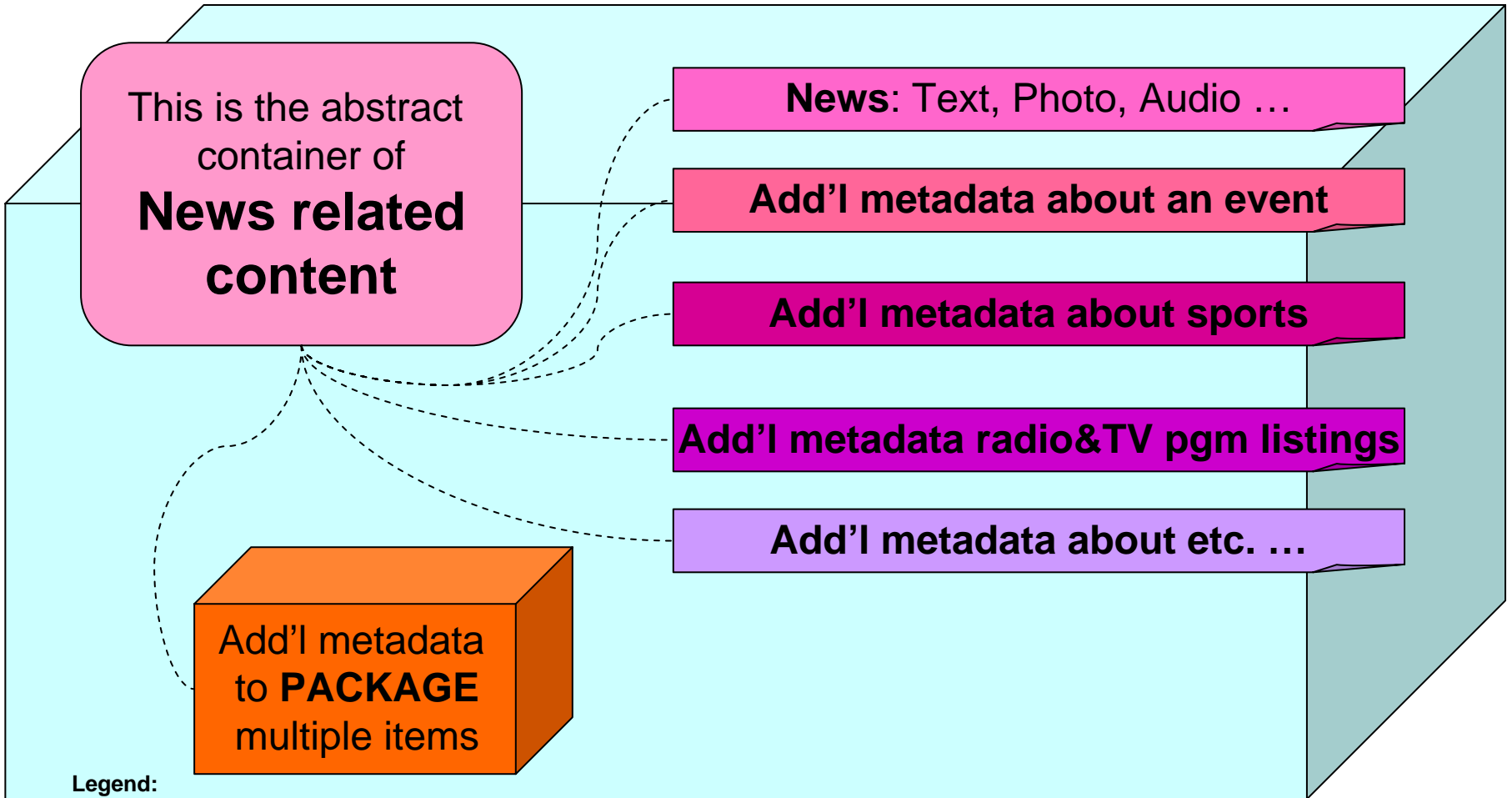
Initial thoughts

- What is in the content container?
  - For "human expressions about events" the data of this expression: text, image data, audio or TV data.
  - For "formal representations" the specialised set of metadata – optionally employing a controlled vocabulary for its values – required to express the representation in full extent.

Step by step

- The next slides show how to build an instance of any IPTC standard in a "standardised" way

- = Building an instance of an IPTC standard is to run through this process step by step. Each individual standard then specifies the options for each step:
  - is this step applicable to this standard
  - have metadata values to be preset for this standard (~ making them default or required values)

Step 1: Insert content

This is the abstract container of **News related content**

News: Text, Photo, Audio …

Add'l metadata about an event

Add'l metadata about sports

Add'l metadata radio&TV pgm listings

Add'l metadata about etc. …

Add'l metadata to **PACKAGE** multiple items

Legend:

- - - = alternative content

My comment on Step 1:

The three facets of news content:

1. **News content data** of various media types
2. Set of **metadata** to **describe** newsworthy information in an abstract and formal way
3. Set of **metadata** (= references to already existing news content of facet 1 or 2) to create a **package** of news content

Step 2: describe the content

News related content

DESCRIBE the content
-human readable: Labels
- -Headline
- -Keywords
- -Byline, Byline Title
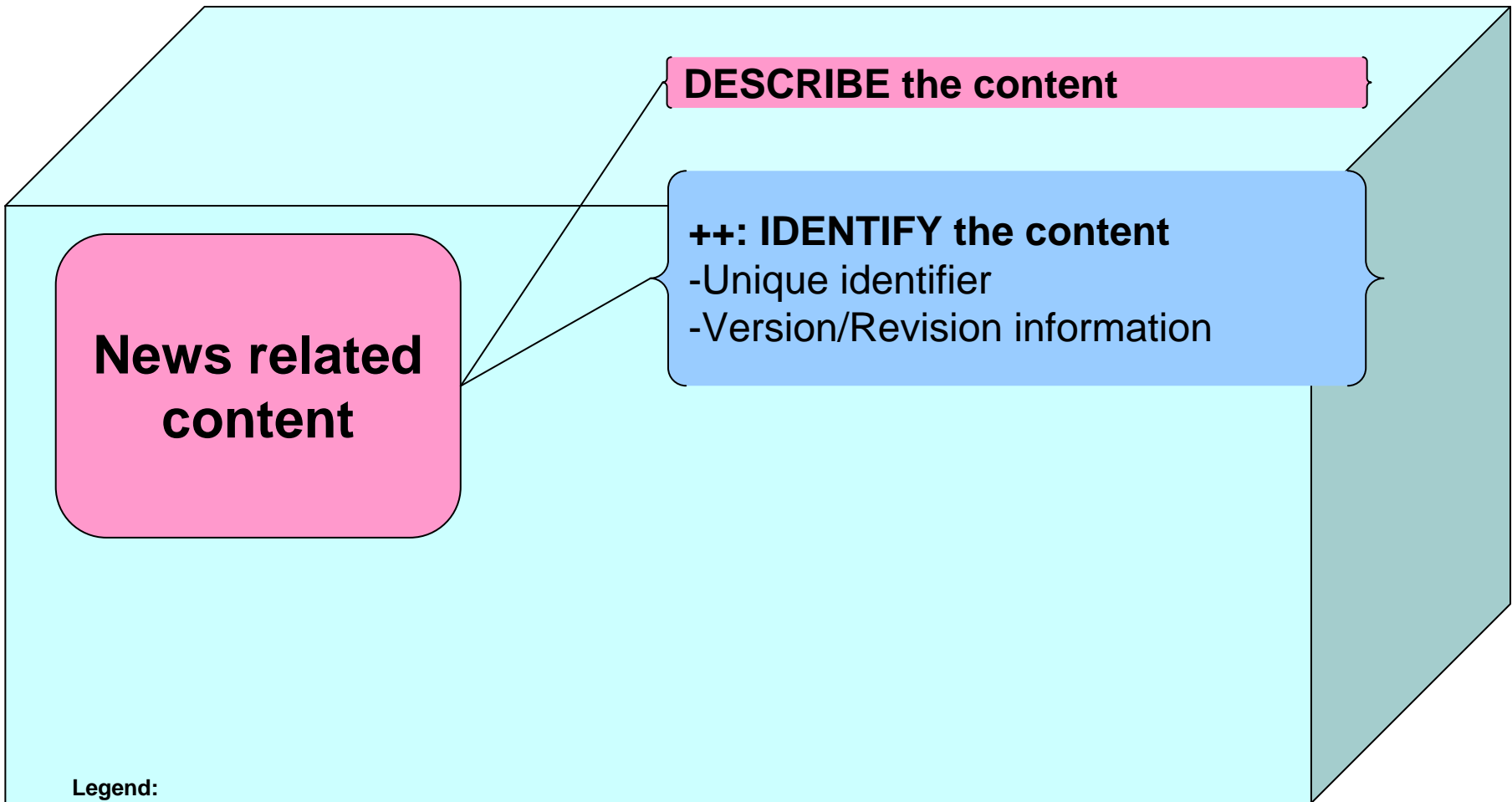- -Dateline
- - etc…
- formal:
- -Subject-NewsCodes
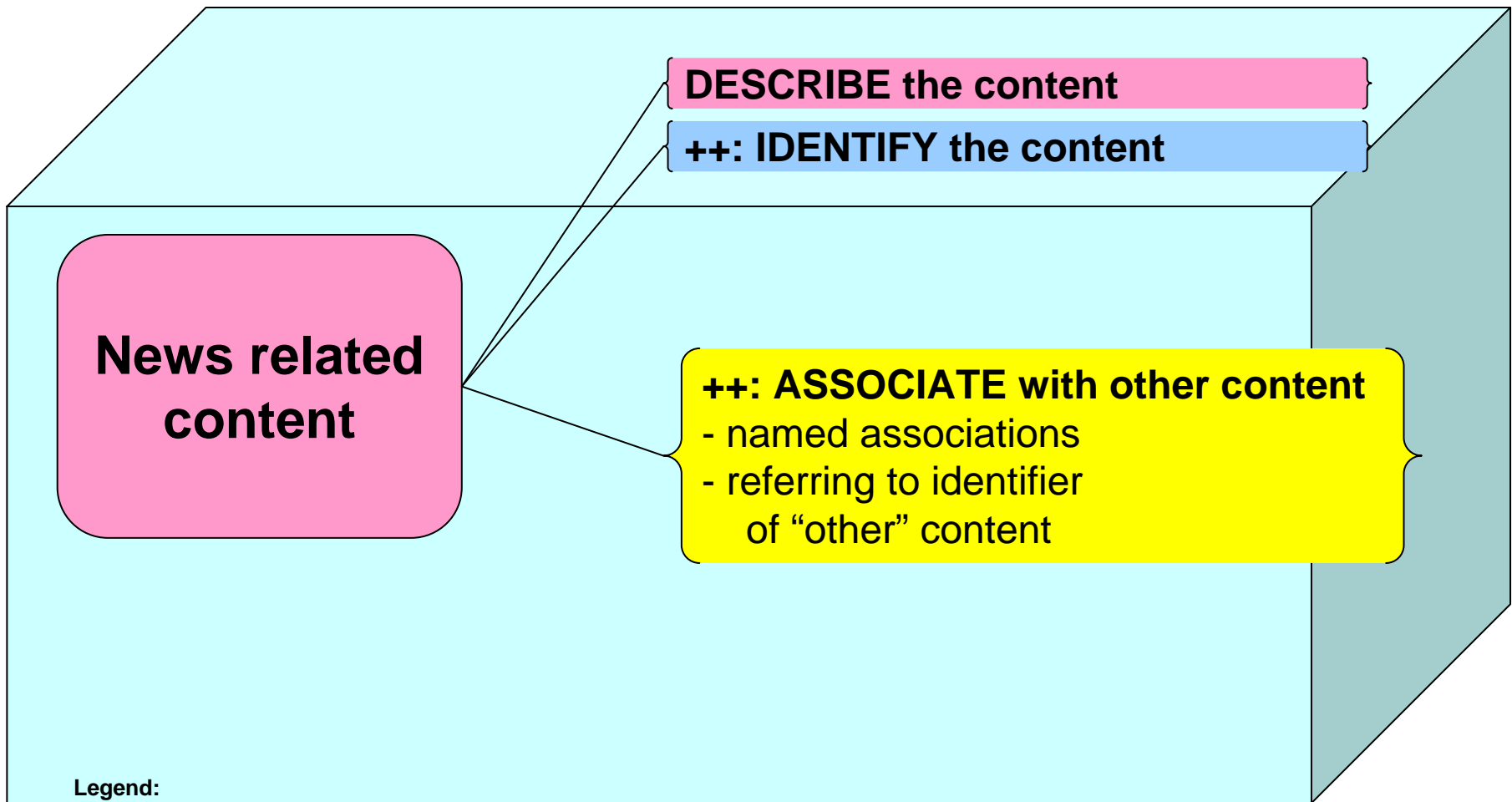- -Scene-NewsCodes
- -Creator
- -Copyrights
- -Usage Rights

DESCRIBE the content

++: IDENTIFY the content
-Unique identifier
-Version/Revision information

News related content

Legend:

++ = optional, additional metadata

DESCRIBE the content

++: IDENTIFY the content

News related content

++: ASSOCIATE with other content
- named associations
- referring to identifier
   of "other" content

Legend:

++ = optional, additional metadata

DESCRIBE the content

++: IDENTIFY the content

++: ASSOCIATE with other content

News related content

++: MANAGE the content
- Status & opt. Future Status
- Creation date & time
- Usage instructions
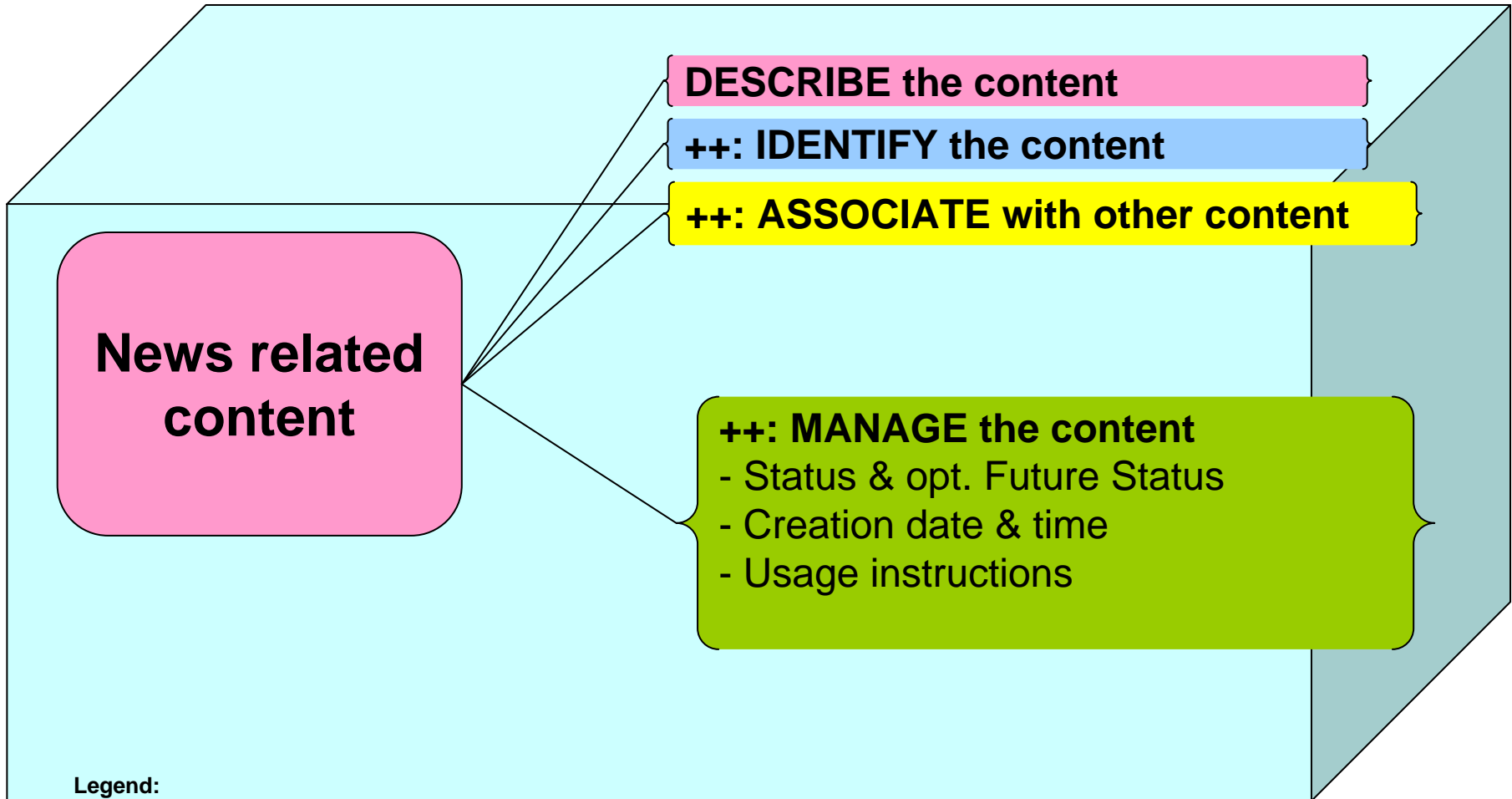
Legend:

++ = optional, additional metadata

Step 6: add distribution info

DESCRIBE the content

++: IDENTIFY the content

++: ASSOCIATE with other content

++: MANAGE the content

News related content

++: DISTRIBUTE the content
- Distribution channel information
- Distribution instructions
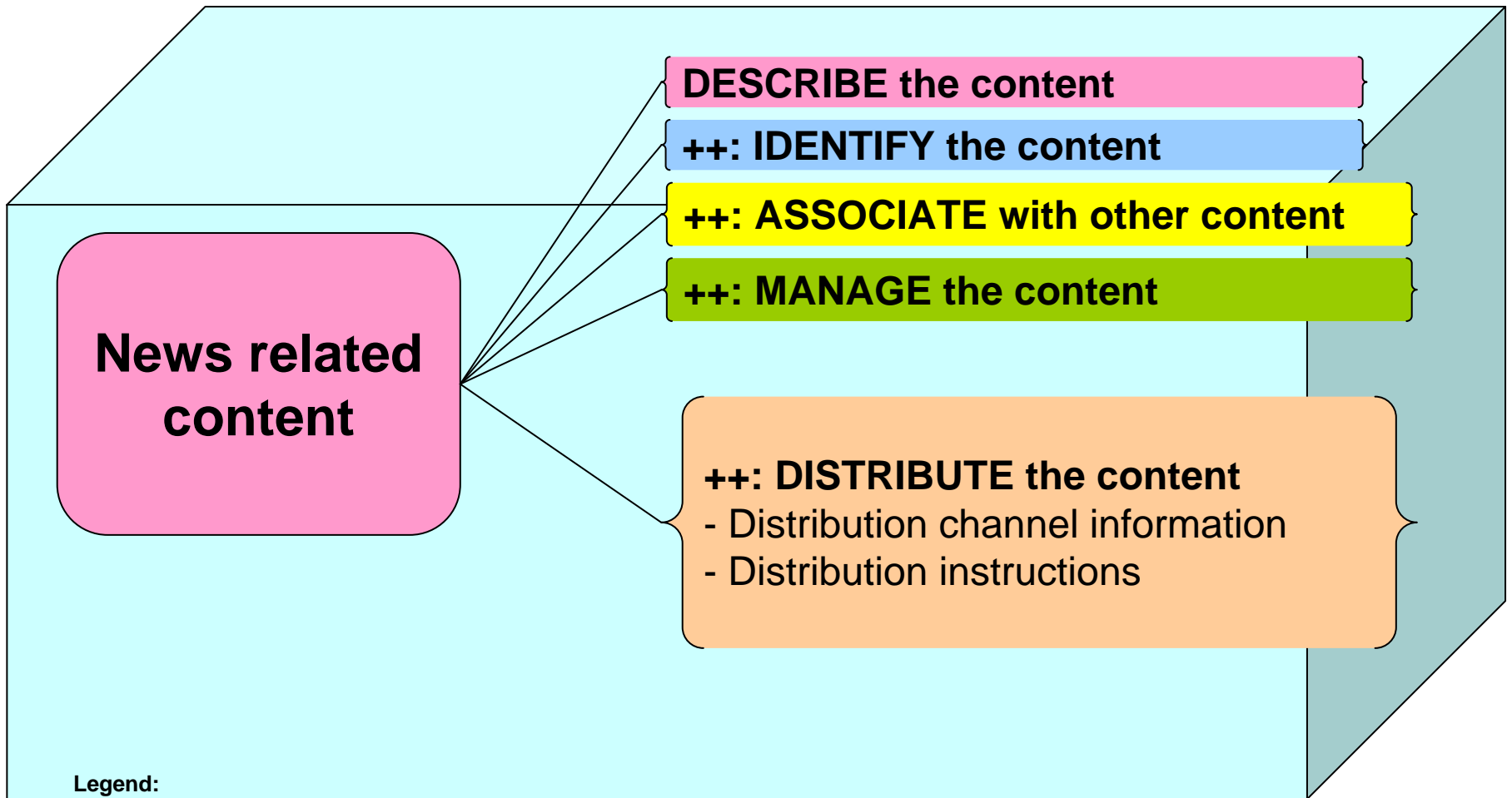
Legend:

++ = optional, additional metadata

My comments on all steps

Consider this process in a software development environment:

- Each coloured set of elements would form a class
- For each class attributes and operations are defined
- Relationship between classes are defined by inheritance
- Adapt a global class to specialised needs by creating a new class with inheritance from the global class.

Consider this process in the scope of XML specs by XML Schemas:

- Each coloured set of elements would be a module to adopt by "inclusion"
- Adapt a global type for specialised needs by employing the extension or constraint mechanism of XML Schemas.

How to develop standards

- A group works on the "news box" = work on
  - the structural model
  - the processing model

- A group works on a common model to
  - express metadata (implementation language)
  - manage topics
  - process metadata/topics

- Groups work on the individual standards:
  - additional specialised content metadata
  - what optional elements of the "news box" are adopted
  - additional processing model

Transition process

The current work setup:

- Working groups for many types of content
  - single item, text centric news: markup and metadata (NITF)
  - metadata about events and about the reporting about events
  - metadata for editorial assignments for news reporting
  - metadata about sport events
  - metadata about radio & TV broadcasting
  - metadata about weather – as reports and forecasts
- A working group about how to package and distribute content
- A working group about controlled metadata values

Transition process

A future work setup (1):

- "Global" working groups on:
  - Metadata expression and management (GMF)
  - Controlled metadata values (NewsCodes)
  - The "news box" = basic structure for news content

Charter:

To create a specialised standard/purpose agnostic specification that can be implemented/applied by all IPTC standards – as they are **all** dealing with news exchange.

Transition process

A future work setup (2):

- "Specialised" working groups on:
  - packaging content of various media types and content types
  - Content type "Events"
  - Content type "Sports"
  - Content type "radio & TV program listings"
  - Content type "Weather"

Charter:

To create a specifications focussed on the specialised purpose by reusing/implementing/applying the "global" framework and adding facilities required for the specific purpose.

Transition process

Transition, step by step:

1. The "specialised" groups submit a set of requirements to the "global" groups they consider to be beyond the specific focus of their standard.

2. The "global" groups assess the submitted sets and build specifications based on the intersection of the sets.

3. The "specialised" groups adapt their specifications to implement the global elements and to extend them as required to their specific needs.

Transition process

A few obvious pitfalls – and challenges:

– What drives what – what has to align to what:
It has to be discussed
- if there is a hierarchical structure having only the abstract "news box" as parent to all IPTC standards
- or if the "news box" specs are identical with NewsML and hence NewsML is the "King of all IPTC standards".
This was insinuated by Laurent and others in recent postings but as I know not all member delegates share this view this has to be discussed.
Or in other words - who governs the "news box": is it owned by the "NewsML WP" or is it a result of collaborative work of several working groups. Despite being a new approach to the standardisation work at the IPTC I ask to consider the second option too.

Transition process

A few obvious pitfalls – and challenges:

- – More group-interaction required:
  Currently a group does its work without much consideration of the "outside"
  → Implement another communication structure
  → Implement straightforward communication rules
- – Work synchronisation:
  If a specialised group wants to extend the global specification they have to wait until this is approved and implemented by the global group
  → Implement a decision making structure to enable approving specs not only at IPTC meetings to speed up "spec making".

happy

The End