

ISO/IEC JTC 1/SC34 N0406



ISO/IEC JTC 1/SC34

Information Technology --

Document Description and Processing Languages

TITLE: Draft Alignment of N0396 with N0393
SOURCE: [Jan Algermissen](#)
PROJECT: Topic Maps
PROJECT EDITORS: [Michel Biezunski](#), [Martin Bryan](#), [Steven R. Newcomb](#)
STATUS: Editor's Draft, Revision 1.3
ACTION: For review and comment
DATE: 27 April 2003
SUMMARY:
DISTRIBUTION: SC34 and Liaisons
REFER TO:
SUPERCEDES:
REPLY TO:

Dr. James David Mason
(ISO/IEC JTC1/SC34 Chairman)
Y-12 National Security Complex
Information Technology Services
Bldg. 9113 M.S. 8208
Oak Ridge, TN 37831-8208 U.S.A.
Telephone: +1 865 574-6973
Facsimile: +1 865 574-1896
E-mail: <mailto:mxm@y12.doe.gov>
<http://www.y12.doe.gov/sgml/sc34/sc34oldhome.htm>

Ms. Sara Hafele Desautels, ISO/IEC JTC 1/SC 34 Secretariat
American National Standards Institute
25 West 43rd Street
New York, NY 10036
Tel: +1 212 642-4937
Fax: +1 212 840-2298
E-mail: sdesaute@ansi.org

27 April 2003

JTC 1/SC 34 N 0406

Draft Alignment of N0396 with N0393

Version 1.3, 2003/04/27

Go to <http://www.isotopicmaps.org/TMMM/TMSM-latest.html> to view or contribute to the current editors' working revision.

Table of Contents

[0 Introduction](#)

[1 Comparison of N0396 and N0393](#)

[1.1 Similarities](#)

[1.2 Differences](#)

[2 Overview of the alignment](#)

[3 What properties in N0396 are intended to define the subject identity of information items?](#)

[4 How can the different kinds of relationships between subjects \(names, occurrences, and so on\) be expressed using the assertion structure of N0393?](#)

[4.1 Proposed change to N0396:](#)

[4.2 What assertion types and role types need to be defined in the TMA definition?](#)

[4.3 How can the properties in N0396 that are not used for subject indication be defined as OPs in the TMA definition?](#)

[4.4 Remaining issues](#)

[Annex A Annex A](#)

0 Introduction

This document is an effort to align N0396 (the current draft of the "Standard Application Model") with N0393 (the current draft of the "Reference Model"). For such alignment it is necessary to separate the structural material of N0396 from the material that defines semantics (e.g., 'Occurrence-ness') over and above those defined in N0393. Once this semantic material is identified, it can be expressed as a Topic Map Application (TMA) definition as described in section 6 of N0393. Annex A provides a draft of such a TMA definition using an illustrative XML syntax.

The alignment of the variant name structure is omitted, because it follows logically from the alignment of the other structures.

I hope that this document will contribute to a better understanding of N0396 and N0393 and that it will help to resolve open issues in both documents.

1 Comparison of N0396 and N0393

1.1 Similarities

N0396 is similar to N0393 in that it uses surrogates for subjects and typed connections between these surrogates to provide a representation of information.

It is similar in the way it uses properties of the items to identify what the subject is that an item surrogates.

1.2 Differences

The data model of N0396 uses typed surrogates for subjects, which prohibits the merging of surrogates (of different types) even if they are known to surrogate the same subject. N0396 introduces the [reifier] and [reified] properties to solve this situation.

Typed surrogates add unnecessary complexity to the underlying data model, although the concept of names, occurrences and associations as types might be useful when designing an API on top of N0393.

The model defined in N0393 does not use typed surrogates, thus enabling subjects to be always represented by a single surrogate. This approach is less complex, while it does not constrain the design of APIs for interacting with a topic map.

2 Overview of the alignment

In order to align N0396 with N0393 the following issues need to be resolved:

1. What properties in N0396 are intended to define the subject identity of information items? These properties can be directly expressed as SIDPs in the TMA definition.
2. How can the different kinds of relationships between subjects (names, occurrences, and so on) be expressed using the assertion structure of N0393? What assertion types and role types need to be defined in the TMA definition?
3. How can the properties in N0396 that are not used for subject indication be defined as OPs in the TMA definition?

3 What properties in N0396 are intended to define the subject identity of information items?

The properties [subject addresses], [subject identifiers] and [source locators] are properties that define the subject that a given surrogate represents. They map directly to SIDP definitions in the TMA.

4 How can the different kinds of relationships between subjects (names, occurrences, and so on) be expressed using the assertion structure of N0393?

The data model defined in N0396 does not enable the representation of naming and occurrence characteristics as assertions, since the subjects that play the name role or the occurrence role are not recognized as such and are therefore not surrogated. In the case of names, the name is just represented as a string valued property; in the case of occurrences, the information resource is just represented as a locator valued property.

However, the prose of N0396 says that both, name and occurrence characteristics are indeed associations:

"Occurrences are essentially a specialized kind of association, where one participant in

the association must be an information resource." (Section 3.7)

"Essentially, a base name is a specialized kind of occurrence." (Section 3.5)

Below I propose a small change to the model defined by N0396 that enables the integration of two item types 'occurrence item' and 'name item' with the item type 'association item'. This change does not only makes the alignment of N0396 with N0393 more efficient, it also reduces the complexity of the model of N0396 itself.

4.1 Proposed change to N0396:

Change the value of the [name] property on the name item to be a topic (a topic that surrogates the information resource that is the name string) and change the value of the [resource] property on the occurrence item to hold the topic that surrogates the information resource that plays the role of occurrence. Introduce a new property on the topic item, called [subject data] that can hold the data content of an information resource. Finally, drop the [value] property from the occurrence item since that value can now be 'stored' in the [subject data] property of the topic that represents the information resource.

This change provides the additional benefit that each information resource addressed from within the topic map will be represented as a topic and thus all information about that resource will be readily accessible from this single location, again in aid of the Subject Location Uniqueness Objective.

With this change it is possible to integrate the different types of topic characteristic with the concept of association, thus significantly reducing the complexity of the model.

By simply providing the appropriate assertion types and role types, all relationships between subjects defined in N0396 can be expressed with the assertion structure defined by N0393.

4.2 What assertion types and role types need to be defined in the TMA definition?

For the topic-name, topic-occurrence, and class-instance relationships one assertion type and two role types each have to be defined in the TMA definition. The semantics of these relationships include that there is only one player playing each role.

Assertion types and role types for the supertype-subtype relationship must also be included.

4.3 How can the properties in N0396 that are not used for subject indication be defined as OPs in the TMA definition?

N0396 defines one property ([scope]) that needs to be expressed as an OP in the TMA definition. As in N0396, the TMA-defined property has as its value the set of topics that are the scope.

4.4 Remaining issues

The [type] property could also be expressed as an OP (having the typing topic as its value) but since the notion of class-instance-ness will be eventually part of the SAM semantics anyway, I chose to map the [type] property to an assertion of type class-instance.

The remaining properties of N0396 are entirely structural properties, and hence redundant given N0393. Some of them map directly to the 'TMM defined properties' of N0393 (e.g. [roles played]) while the others become obsolete (e.g. [reifier] and [reified]).

Annex A Annex A

This Annex provides a TMA definition for the semantic material of N0396 using an XML syntax devised to illustrate such a definition.

```
<?xml version="1.0"?>
<!DOCTYPE tma
[
<!ENTITY base "http://www.isotopicmaps.org/psi/">
]>

<tma>
<name>IS13250-N0396-1.0</name>
<description>
  A TMA definition of N0396.
</description>
<requiredModels />

<valueTypes>
  <valueType>
    <name>topic</name>
    <description>
      A value of the type topic is a surrogate for a subject. The exact
      nature of a topic is entirely in the realm of the application. It
      could be an integer value providing a unique ID, the address of a
      data structure, or a data structure, etc.
    </description>
    <equality>
      Two values are equal if the application considers them to be equal.
    </equality>
    <merging>
      Values may not be merged.
    </merging>
  </valueType>
  <valueType>
    <name>locator</name>
    <description>
      A value of the type locator consists of a pair of non-empty strings, one
      containing the name of the notation, the other containing the
      reference.
    </description>
    <equality>
      Values are equal if the first strings of the pairs are byte identical and
      the second strings of the pairs are byte identical.
    </equality>
    <merging>
      Vales may not be merged.
    </merging>
  </valueType>
  <valueType>
    <name>text</name>
    <description>
      A value of type text is an arbitrary sequence of bytes.
    </description>
    <equality>
```

```

    Two texts are equal if they are represented by the same sequence of
    bytes.
  </equality>
  <merging>
    Vales may not be merged.
  </merging>
</valueType>
<valueType>
  <name>set</name>
  <description>
    A set is a unordered collection of values. All values must have the same
    value type and no two values may be equal.
  </description>
  <equality>
    Two sets are equal if the contain exactly the same elements.
  </equality>
  <merging>
    The result of merging two set values is the union of the two sets.
  </merging>
</valueType>
</valueTypes>

<properties>
  <property>
    <name>PR_SubjectAddresses</name>
    <valueType>set(locator)</name>
    <type>SIDP</type>
    <description>
      A set of locator items. The locators, if present, refer to the
      information resource that is the subject of this topic. If the
      set contains more than one locator this implies that the locators
      all address the same information resource.
      The fact that in some information management systems (e.g., the Web)
      an information resource is not directly addressable is ignored.
    </description>
    <mergingCondition>
      Two topics exhibit a value for this property and the
      intersection of the two sets is not empty.
    </mergingCondition>
  </property>

  <property>
    <name>PR_SubjectIdentifiers</name>
    <valueType>set(locator)</name>
    <type>SIDP</type>
    <description>
      A set of locator items. The locator items refer to the subject indicators
      of this topic.
    </description>
    <mergingCondition>
      Two topics exhibit a value for this property and the
      intersection of the two sets is not empty.
    </mergingCondition>
  </property>

  <property>
    <name>PR_SourceLocators</name>
    <valueType>set(locator)</name>
    <type>SIDP</type>
    <description>
      If the existence of a topic has been demanded by the presence of
      some syntactical construct (a node demander), the address of that
      syntactical construct is said to be a source locator of that topic.
      Merging can cause the situation that a single topic has multiple
      source locators, thus this property is a set.
    </description>
    <mergingCondition>
      Two topics exhibit a value for this property and the
      intersection of the two sets is not empty.
    </mergingCondition>
  </property>

  <property>

```

```

<name>PR_SubjectData</name>
<valueType>text</name>
<type>OP</type>
<description>
  A topic that exhibits a value for this property is the surrogate
  for an information resource. The data content of the surrogated
  information resource is the value of this property.
  This is an OP, because the data content of an information resource
  is not used for subject discrimination. Two information resources
  can have the same data content and still be different information
  resources.
</description>
</property>

<property>
  <name>PR_Scope</name>
  <valueType>set(topic)</name>
  <type>OP</type>
  <description>
    Topics that exhibit a value for this property must be surrogates for
    relationships (that is, they must exhibit a value for the property
    IS13250::a-sidp). The set of topics that is the value is defined
    to represent the context in which the represented relationship is
    considered to be valid.
  </description>
</property>
</properties>

<assertionTypes>
  <assertionType>
    <name>AT_NamedName</name>
    <subjectIdentity builtInTopic="#t1" />
    <roleType>
      <name>RO_Named</name>
      <subjectIdentity builtInTopic="#t2" />
      <rolePlayerConstraint>
        There may only be one player of this role in a single assertion.
      </rolePlayerConstraint>
    </roleType>
    <roleType>
      <name>RO_Name</name>
      <subjectIdentity builtInTopic="#t3" />
      <rolePlayerConstraint>
        There may only be one player of this role in a single assertion.
        Players of this role must exhibit a value for at least one of the
        two properties PR_SubjectAddress and PR_SubjectData, that means,
        they must be surrogates for information resources.
      </rolePlayerConstraint>
    </roleType>
  </assertionType>

  <assertionType>
    <name>AT_OccurringOccurrence</name>
    <subjectIdentity builtInTopic="#t4" />
    <roleType>
      <name>RO_Occurring</name>
      <subjectIdentity builtInTopic="#t5" />
      <rolePlayerConstraint>
        There may only be one player of this role in a single assertion.
      </rolePlayerConstraint>
    </roleType>
    <roleType>
      <name>RO_Occurrence</name>
      <subjectIdentity builtInTopic="#t6" />
      <rolePlayerConstraint>
        There may only be one player of this role in a single assertion.
        Players of this role must exhibit a value for at least one of the
        two properties PR_SubjectAddress and PR_SubjectData, that means,
        they must be surrogates for information resources.
      </rolePlayerConstraint>
    </roleType>
  </assertionType>

```

```

<assertionType>
  <name>AT_ClassInstance</name>
  <subjectIdentity builtInTopic="#t7" />
  <roleType>
    <name>RO_Class</name>
    <subjectIdentity builtInTopic="#t8" />
    <rolePlayerConstraint>
      There may only be one player of this role in a single assertion.
    </rolePlayerConstraint>
  </roleType>
  <roleType>
    <name>RO_Instance</name>
    <subjectIdentity builtInTopic="#t9" />
    <rolePlayerConstraint>
      There may only be one player of this role in a single assertion.
    </rolePlayerConstraint>
  </roleType>
</assertionType>

<assertionType>
  <name>AT_SuperclassSubclass</name>
  <subjectIdentity builtInTopic="#t10" />
  <roleType>
    <name>RO_Superclass</name>
    <subjectIdentity builtInTopic="#t11" />
    <rolePlayerConstraint>
      There may only be one player of this role in a single assertion.
    </rolePlayerConstraint>
  </roleType>
  <roleType>
    <name>RO_Subclass</name>
    <subjectIdentity builtInTopic="#t12" />
    <rolePlayerConstraint>
      There may only be one player of this role in a single assertion.
    </rolePlayerConstraint>
  </roleType>
</assertionType>
</assertionTypes>

<builtInTopics>
  <topic id="t1">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;at-named-name</locator>
      </set>
    </property>
  </topic>
  <topic id="t2">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;role-named</locator>
      </set>
    </property>
  </topic>
  <topic id="t3">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;role-name</locator>
      </set>
    </property>
  </topic>

  <topic id="t4">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;at-occurring-occurrence</locator>
      </set>
    </property>
  </topic>
  <topic id="t5">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">

```



```

        <set>
          <locator notation="URI">&base;role-occurring</locator>
        </set>
      </property>
    </topic>
  <topic id=t6">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;role-occurrence</locator>
      </set>
    </property>
  </topic>

  <topic id=t7">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;at-class-instance</locator>
      </set>
    </property>
  </topic>
  <topic id=t8">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;role-class</locator>
      </set>
    </property>
  </topic>
  <topic id=t9">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;role-instance</locator>
      </set>
    </property>
  </topic>

  <topic id=t10">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;at-superclass-subclass</locator>
      </set>
    </property>
  </topic>
  <topic id=t11">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;role-superclass</locator>
      </set>
    </property>
  </topic>
  <topic id=t12">
    <property model="IS13250-N0396-1.0" name="PR_SubjectIdentifiers">
      <set>
        <locator notation="URI">&base;role-subclass</locator>
      </set>
    </property>
  </topic>

</builtInTopics>

<propertyValueComputationRules>
  <!--
    This TMA does not define any rules for conferring values on
    properties on the basis of the roles that a topic plays.
  -->
</propertyValueComputationRules>

</tma>

```