
XFML, Standard for Distributed Information Architecture

Aparajita Suman

Abstract

The unplanned dumping of information on WWW is the primary reason for the chaotic web of today. The way out of this WEB (in literal sense) seems to be only through planned organization of the burgeoning sources of information using faceted techniques. The role of XFML i.e. eXchangeable faceted metadata language gets defined in this context. The metadata language allows one to relate the topics of concern and to provide a meaningful way of accessing information. This paper tries to explore its nuances and suitability for defining distributed information architecture.

Keywords : Faceted Metadata Language, Metadata Information Architecture, XFML

0. Introduction

XFML, i.e. eXchangeable faceted metadata language is no doubt another addition to the series of XML based languages rather but has a different and distinct flavor. XFML is an XML format for exchanging metadata in the form of faceted hierarchies, which are also called 'taxonomies'. It helps in solving one aspect of metadata problem, i.e., interchange of faceted classification and indexing data.

Facets can be said to be clearly defined, mutually exclusive and collective exhaustive aspects or properties or characteristics of a specific subject or class [1]. It is quite similar to human way of thinking, when we use multiple facets to describe a subject rather than attempt to fit it into some type of taxonomy or hierarchy. That's why faceted classification (i.e., describing things by their characteristics rather than assigning them a universal category) always has advantage over standardized ontologies.

Whenever indexing and classification of facets is done for a particular subject area, then, it is always advantageous and easy for others to use the same data rather than going for duplication of effort. For example, let us imagine a situation where we have a digital library in a particular subject area. We post the entries in our tailor made categories, now, some other library may also have digital library service in the same subject area and it may post entries in its own categories. Here, possibility of categories being similar is greater and the difference might only be in the way they are addressed.

The role of faceted metadata language gets defined here very well, as both the libraries can publish their metadata in XFML format. It allows them to relate their categories, where they continue to have their respective names but the two library systems know that they are really the same and related. Actually, we import the XFML file of the other library and relate the categories, then, the DL software is configured to import the file daily to check for changes. So, the XFML compatible software can automatically generate links to new features about the same topic as the other digital library.

1. Need of XFML

Library science research has proved that the best way of organizing knowledge is to put topics in various facets because categorizing, browsing and searching becomes easy in this case. The content management systems serve the purpose to some extent but they are deficient in content mapping. Most CMs implement a certain level of metadata but the standard way of publishing this metadata remains missing.

XFML came up as a solution to the problems associated with imperfect and ever changing taxonomies. In content management systems, metadata is intertwined with content definition; so, adjusting the string of metadata involves a lot of work. But in XFML map, metadata and content is separate. Here, we create a map that exists on its own, regardless of whether there is any content that relates to the map or not. So, it provides a lot of power to the user by allowing her to work with the map as an entity of its own, import facets of other maps, merge topics etc. [2]

2. Definition

XFML is an XML format for publishing and sharing hierarchical faceted metadata and indexing efforts. It provides ways to build connections between topics, information that helps in designing tools to automate the sharing of indexing efforts [3]. Its primary goal is to allow us to publish our metadata categories on web in a standard format and connect our metadata with the metadata, other people might have published. So, finally it enables automatic link generation to related content on our own and other websites.

Here, faceted classification is helpful because instead of building one huge tree of topics, multiple smaller trees are used that can be combined by the users to find things more easily.

3. History and Evolution

On May 30, 2002, XFML 0.1 was published, and then came XFML 0.2 on July 6, 2002. It was introduced with a number of refinements compared with the version 0.1 i.e. elements which allow software to easily display singular or plural versions of topics, facets and occurrence types. The main outstanding issues were whether topics should be allowed to have multiple parents. The other additions till date are XFML core (Oct 08, 2002), revisions to XFML core specification (Dec 13, 2002) and classification to XFML core specification [1].

4. Features of XFML [2]

- ✍ XFML takes the distributed approach and at the same time, allows individual authors to connect their metadata schemes by merging topics. It is impractical even to think of creating a centralized metadata store for the web, so, we always need shared metadata to make the web easier to get around it. This distributed connected metadata network can be made using XFML and which will function very much like a centralized metadata store.
- ✍ Taxonomies can be easily shared and published using XFML. Publishing the taxonomy means that using it one can get relevant incoming links and there is no need to do the labor-intensive task of creating taxonomies.
- ✍ XFML is well equipped to deal with changes in taxonomies as metadata in XFML map can evolve more easily than in current content management systems.
- ✍ Faceted taxonomies are generally more powerful for websites than classic hierarchical taxonomies. So, use of XFML is more suitable as it is based on facets.
- ✍ XFML allows us to index anything on the web as no write-access is needed to a web page to index it in the topic map
- ✍ Finally, allowing meaningful metadata connections between separate systems is what makes XFML so powerful

5. Difference from other Standards (Metadata Related Technologies)

XFML is a very specific and focused format, as opposed to eXtensible Topic Map or Resource Description Format, which are generic metadata formats [4]. It is optimized for a specific goal i.e. to enable sharing and connecting faceted metadata between websites; moreover, it is easy to write code for XFML at the same time.

5.1 XFML and Dublin core

Dublin Core is a specification that is additional to XFML: these two specifications can work together. XFML indicates relationships between topics, but not what the topics mean. Dublin Core can be incorporated into XFML to do that.

5.2 XFML and RDF

RDF is all about adding meaning to the web documents while XFML explains the manner in which the facets are related on web. RDF also lets web pages publish their metadata but it is too complex to code while XFML has been designed for easier code implementation.

5.3 XFML and TopicMaps:

XFML is a subset of topicmaps. The topicmap specification was designed so that subsets of XTM could easily be created. It is very easy to turn an XFML document into an XTM document (a topicmap) but it doesn't work the other way round. So, one can work with XFML and later in the future when one needs to do things XFML doesn't provide, but topicmaps do, one can easily convert the XFML map to a topicmap.

6. XFML, Standard for Distributed Information Architecture

In due course of time, people felt the need of standards that will allow web sites to share data with respect to their categorization, organization, and labeling. Further, creating standards for distributed information architecture would have allowed for easier and more effective combination of content, resources, and metadata across sites [5].

To understand the problem better, let us imagine a scenario where we have five sets of photographs in five different places, and may be also labeled under five different headings, but actually they are not five different things rather five different ways of looking at the same thing. The photographs may be of the same person who is father at one place, friend at another, husband at the third place, brother at the fourth place and boss at the fifth place. So, the photograph of the same person will be addressed in five different ways at five different places but any change in the address or status of the person will affect all of them equally.

By utilizing a standard for distributed information architecture, one can store aggregated information about any topic irrespective of their different locations and descriptions. The basic philosophy of XFML is to make real a distributed, loosely connected web of metadata. It gives us the freedom to choose topics for our XFML map and how do we want to organize the information.

XFML can be considered as a standard because it has many features that facilitate the development of distributed information architectures. It creates a loosely coupled net of published taxonomies where authors themselves can create and share taxonomies, merge XFML documents, mutually define metadata and facets, and import XFML from other authors. It came up as a format specifically for publishing and connecting faceted metadata between websites.

7. How Does XFML Work

Before going to the detailed specification of XFML, let us understand its working in a broad manner [6]:

- ✍ First we find another site with similar topics which may be of our interest, this should also publish its data in XFML format to allow linking.
- ✍ Then we configure the XFML software to trace their map once, per specified-time-period.
- ✍ In our XFML map, we use the connect element to link topics on our map to identical / very similar topics on the map of the other website. For example, we may have a topic called "Classification" while they have one called "Knowledge Organization". The meaning is the same, so a link can be created between them.
- ✍ Now, whenever the other site publishes something on the topic on their site, then the software at our end automatically links to it as "related reading" on our site. This is how XFML works and facilitates metadata sharing across web sites.

8. XFML Specifications

It can be divided to three categories [7]:

- ✍ A set of concepts, i.e. a conceptual model.
- ✍ An XML format for expressing these concepts, and
- ✍ A set of processing instructions that explain how applications should work with XFML data.

The availability of one or more XFML documents on a website can be indicated by a particular logo indicating link to the XFML document. In case of multiple XFML documents, one can have multiple buttons, although a page explaining the differences is always a good idea.

An XFML document is a valid, well-formed XML document, and conforms to the XFML DTD and XFML Core specifications. Example of an XFML document:

```
<xfml version="1.0" url="http://xfml.org/spec/example.xml" language="en-us">
<!--This document conforms to XFML Core. See http://purl.oclc.org/NET/xfml/core/-->
<mapInfo>

<publisher>
  <name>DRTC</name>
  <url>http://drtc.isibang.ac.in</url>
</publisher>

<generator>
  <name>word processor document</name>
</generator>
</mapInfo>

  <facet id="Subject categories">subject categories</facet>
  <facet id="author">author</facet>

<topic id="Knowledge management" facetid="subject categories">
  <name>Knowledge management</name>
  <connect>http://othersite.com/xfml.xml#18753</connect>
```

```
<psi>http://www.cia.gov/cia/publications/KMresources.html</psi>
<description>Knowledge Management tools and techniques</description>
</topic>
```

```
<-<topic id="Knowledge management tools" facetid="subject categories" parentTopicid="Knowledge
management ">
<name>Knowledge management tools</name>
</topic>
```

```
<-<topic id="ADIS student" facetid="Authors">
<name>ADIS student</name>
</topic>
```

```
<-<page url="https://drtc.isibang.ac.in">
<title>Seminar volumes of DRTC</title>
<description>All the seminar volumes of DRTC can be accessed from here</description>
<occurrence topicid="authors" />
<occurrence topicid="Knowledge management" strength="1" />
</page>
</xfml>
```

9. Compatible Softwares and Formats

Cardinal XFML Parser: is an XFML Core compatible XFML processor implemented in Visual Basic 6.0 and built upon the MSXML 4 DOM implementation. Cardinal provides an XFML abstraction to simplify the development of tools to create and consume XFML documents.

FacetMap : Facetmap is a system for managing faceted hierarchies. It was the first fully XFML compatible application, and lets you import XFML documents.

Drupal : the popular Drupal CMs supports XFML export.

Compatible formats : XFML documents can be converted to other formats. With some formats, some of the information gets lost; with some formats, conversion can only go one way; with others, full, two-way conversion is possible.

XTM : It is the XML expression of topicmaps. XFML is a subset of XTM, any XFML document can be expressed in XTM, but not the other way around. Work is going on to develop a stylesheet that transforms XFML into XTM.

RDF : Research is going on for a RDF serialization of XFML. Presently any XFML document can be expressed as RDF, but not the other way round.

XFML libraries are existing libraries with facets and topics one can easily copy when creating a new XFML file. Example is IPTC library, which is categorized on the basis of subject, genre, media and news item types.

10. Limitations of XFML

- ✂ Only parent-child relationships between pieces of content are possible; it limits the types of associations that could be made.
- ✂ It is not possible to have multiple languages (real languages like English, French, Spanish—not programming or markup languages) within an XFML document as in the case of XHTML.
- ✂ It is still being developed, so, XFML has to be created manually and is not supported in any editing software, or on any existing Web sites, only demos or examples available.
- ✂ Moreover it is in very nascent stage, so, there is no guarantee that it will develop as full-fledged standard.

11. Conclusion

XFML is focused on sharing indexing efforts with faceted metadata, this is important because creating metadata vocabularies is really hard, and indexing lots of pages is even harder, but, sharing these efforts can be made possible by using <connect> and <psi> elements of XFML. As XFML Core is a frozen standard, so we can safely implement it. Work on XFML 2.0 is going on, but that is a long way off, and it will be a language with a different purpose than XFML Core. In addition to this, tool support is also taking off, so, it will be very easy to use it in near future and the websites will be able to talk to each other in real sense about their metadata.

12. References

1. Official publication of XFML core. <http://purl.oclc.org/NET/xfml/core/> (accessed on 20/9/04)
2. Exchangeable Faceted Metadata Language (XFML) as Fuzzy-Lightweight XTM and RDF. <http://xml.coverpages.org/ni2002-06-04-a.html> (accessed on 1/10/04)
3. eXchangeable faceted metadata language. Home Page. <http://xfml.org> (accessed on 2/10/04)
4. Software that supports XFML. <http://xfml.org/software.html#compatibleformats>. (accessed on 25/10/04)
5. Lash, Jeff. Standards for distributed information architecture. <http://www.digital-web.com/columns/ianythinggoes/ianythinggoes2002-09.shtml>. (accessed on 20/10/04)
6. Pilgrim, Mark. Mark goes XFML <http://simon.incutio.com/archive/2002/12/05/markGOesxfml> (accessed on 22/10/04)
7. Introduction to XFML. <http://xfml.org/spec/1.0html#introductiontoxfml>. (Accessed on 27/9/04)
8. Xfmlib, A python library for reading and writing XFML documents. <http://diveintomark.org/projects/xfmlib/>. (accessed on 20/11/04)
9. Dive into mark. <http://diveintoaccessibility.org/> (accessed on 12/11/03)

About Author



Aparajita Suman has done ADIS from DRTC, Indian Statistical Institute, Bangalore and is presently working as Junior Research Fellow at Defense Research and Development Laboratory, Hyderabad

E-mail : s_aparajitha@yahoo.co.in, aparajita@drtc.isibang.ac.in