Sticky Stuff
An Introduction to the Burr Metadata Framework

Brad Collins
Chenla Laboratories

Abstract

This paper presents BMF (The Burr Metadata Framework), an XML based framework for creating integrated libraries of metadata, and encoded documents. In large part BMF is an extension and expansion of the FRBR (Functional Requirements for Bibliographic Data) model proposed by the IFLA, which uses standard thesaurus relationships to create complex, scalable, hierarchical structures.

Sticky Stuff

An Introduction to the Burr Metadata Framework

Table of Contents

A Conference of IDEAlliance

Extreme Markup
2006 Languages®

Montréal, August 7–11, 2006

Sticky Stuff
An Introduction to the Burr Metadata Framework
Brad Collins

§ 1 Conventions Used In This Paper

BMF [Burr Metadata Framework] uses a shorthand notation to map out the hierarchical relationships between Burrs (the basic record level building block in BMF).

Each line represents a single Burr, concept or term in a hierarchy. Boldface text indicates the logical, locus focus of the map with lines above it being broader terms, lines at the same level of indentation below it being equivalent. Lines below it with a greater indentation (depth is indicated by period for each level of depth in the hierarchy) are narrower or related terms.

Each line is made up of four fields separated by whiteface.

```
        PT  per  .. Dick, Philip Kindred.
```

Where :

- The first field, *PT* is a relationship code (see Relationship Codes, below)
- The second field, *per* is an entity-type code (see Entity-type Codes, below). The third field uses zero or more periods indicating the level of the hierarchy with one "." period used for each additional level.
- The fourth field, *Dick, Philip Kindred*, is a descriptor which provides a label for the item.

For example.

```
BTI   work  Ann Charter's Intro to "The special view of history".
PT    expr  . original text
NTP   div   .. body of text.
```

Note: The entity code field may be omitted in examples discussing relationships between terms but are required when discussing relationships between Burrs.

§ 2 Relationship Codes

The following relationship types are defined in ANSI Z39.19: Guidelines for the Construction, Format and Management of Monolingual Thesauri. [Z39.19]

All codes should use uppercase characters

BT	broader term.
BTG	broader term (generic).
BTI	broader term (instance).
BTP	broader term (partitive).
GS	generic structure.
NL	node label.
NT	narrower term.
NTG	narrower term (generic).
NTI	narrower term (instance).
NTP	narrower term (partitive).

PT	primary term.
RT	related term.
TT	top term.
U	use.
UF	used for.
UF+	used for ... and ...

In addition to the above relationships defined in Z39.19, BMF also uses the following codes :

BTR	broader term (responsibility).
NTR	narrower term (responsibility).

which are used in Burrs like a chapter or a story which are collected into a compound document.

PRE	previous node.
NEX	next node.

§ 3 Entity-type Codes

All codes should use lower case characters and be three to four characters in length.

3.1 *Please Note:*

* equivalent to same FRBR Group 1 Entity.

** equivalent to same FRBR Group 2 Entity.

*** equivalent to same FRBR Group 3 Entity.

3.2 *Bibliographic Group*

work	work*
expr	expression*
man	manifestation*
item	item*

3.3 *Agents Group*

per	person**
cor	corporate body**
fam	family**

3.4 *Semantic & Lexical Group*

con	concept***
sym	symbol
sys	system

3.5 *Temporal Group*

date	date
per	period
thrd	thread
evn	event***

3.6 *Physical Group*

obj	object***
des	design
mod	model
item	item
mat	material

3.7 *Locus Group*

wrld	world
loc	locus
plc	place*
feat	feature

3.8 *Content Group*

div	division
sch	scull
mess	message
day	day

3.9 *Documentation Group*

doc	documentation
ref	reference

3.10 *BMF Group*

gram	bramble
top	topicspace
stub	stub

§ 4 Preface

> If you do not work on an important problem, it's unlikely you'll do important work. It's perfectly obvious.
>
> —Richard Hamming, *You and Your Research* [HAMMING]

In late 1997 I was sitting in Osaka, at a cramped desk on the top floor of a musty, cold, cluttered office, stinking of stale cigarettes, when I read the following:

> There is no useful distinction between the representational needs of data and metadata. The kinds of information that need to be represented in metadata and data are very similar. Furthermore, every item of information, without exception, is likely to be regarded by some applications as ancillary and never to be displayed, and by others as core content that needs to be formatted, printed, or searched.
>
> —*Meta Content Framework Using XML* [GUHA]

I knew at that moment, that this was the insight which would be at the core of, maybe not the next generation of the Web, but perhaps the one after.

At that time the Dot-Com bubble was ready to pop, and nowhere was this more keenly felt than in Japan which was still stinging from the collapse of a monster bubble economy that nearly wrecked the country some years before. The Web had no business or revenue models at that time. It was all just smoke and mirrors.

So I packed it in, tried to prepare for the crash, moved to the backwaters of Thailand and turned my efforts to the next generation of the Internet, an Internet which would have oodles of bandwidth into every home and office, an Internet with a browser which could support powerful and mature applications that could get real work done, an Internet with a business and revenue model that you could make real money from.

It was in this context that I latched onto the idea that information has a dual nature, like the particle-wave nature of light.

The lack of any real metadata and cataloging of Web resources was such an obvious problem, that at the time it seemed that if you could crack the problem of providing a universal metadata system, you'd have everything.

I wasn't alone. It looked like Tim Berners Lee over at the W3C was thinking along the same lines. But his approach with the Semantic Web, although brilliant, didn't feel right. It felt like a cop out.

Just because the problem of adding metadata was difficult, people gave up working on it. Everyone threw up their collective hands and said "We'll never get people to do metadata so let's try to find a way of automating the process and let the machines distill meaning from chaos". The short comings of metadata systems were brilliantly summed up by Cory Doctorow in his essay "Meta Crap".[DOCTOROW]

But that was six years ago, and as they say in New England, if you don't like the weather, wait ten minutes. This is especially true on the Internet.

We now have Wikipedia[1], Distributed Proofreaders[2], Del.icio.us[3], Flickr[4], and Technorati[5].

Metadata is a matter of priorities, not how much work it takes. If you can get tens of thousands of people to volunteer everyday to proofread mind numbingly dull texts like lists of copyright renewals, nothing is impossible.

What the Semantic Web crowd was really missing, was that automated organization, sorting and uncovering patterns in collections of data is not an end in itself. Search is not everything. It's the process of organizing, sorting, abstracting and cataloging that leads to meaning and ultimately to understanding. In other words, it's the process that results in knowledge which we use to make decisions.

BMF is designed not only to be a content or a metadata framework, but a infrastructure for the process of learning, creating, sharing, collaborating and remembering.

That's about as important a problem and design goal as you can hope for.

§ 5 Introduction

> The tree is already the image of the world, or the root the image of the world-tree. This is the classical book, as noble, signifying, and subjective organic interiority (the strata of the book). The book imitates the world, as art imitates nature: by procedures specific to it that accomplish what nature cannot or can no longer do.
>
> —Gilles Deleuze, *Rhizome Versus Tree* [DELEUZE]

The small compass in which the eye can see clearly is little more than a knothole through which we are continuously taking a series of snapshots the brain uses to form a composite image, tricking us into thinking that we live in a panorama of clarity.

Memory is the mid-day light cast through the canopy of a grove of birch on a clear August day, coloring and mellowing the carpet of yellow leaves rustling and crunching beneath our passing feet. It is not the world, it is just what our feeble senses can take in, and even that is more than the brain can process and store.

So if the book, as Deleuze said, is an imitation of the world, it is an imitation twice removed from the world it seeks to ape. And if art imitates nature, it also captures our perception of nature so that others, twice more removed can see with another's eyes what has been mulched by another's mind.

Much of our lives are spent, sorting, organizing and picking out patterns in this cacophony of distorted information which is interleaved with the clear, the fuzzy, and a whole lot of line noise in between.

Every and all is schlepped onto the scales and weighed. All so that we can decide what to do.

> I would be an historian as Herodotus was, looking for oneself for the evidence of what is said.
>
> —Charles Olson, *Maximus Poems*, Letter 23

Isn't this exactly what we do with information? The evidence, comes to us directly through observation, but also again removed, as hearsay, tales told in bars through the amber lens of a pint glass. It's the oral tradition, the most immediate form of human intercourse.

But noise is added with each remove. And it's that noise that man has worked so hard to minimize. So we actualize human language through writing systems, but the duplication and distribution of that writing introduced a different kind of noise which Caxton's ink stained fingers finally remedied with blocks of lead viced together into a mirror image of what was wrote.

We have learned to capture light and fix it on paper. We have given it the illusion of motion by exceeding the brain's ability to discern change. We have fixed sound by reducing it to a groove etched in wax, which can be reanimated on a whim by pushing a paper cone fourteen thousand times a second.

So armed we can now transpose the garbage our senses take in, and the garbage our brains pass out and fix it all into something that is nothing short of magic! A click of a shutter, the ball on the point of a pen applying a smooth ellipse of ink on a piece of paper and we can teleport our memory and experience through any measure of space or passage of time.

Think of it. The words that Homer fixed in *his* present can become *anyone's* present so long as his words are not lost. Homer is our contemporary, as is anyone who has fixed some fragment of mind, no matter how trivial the catharsis, and passed it into the physical world.

We are what we pass on, in body, memory, experience and mind. But we are also the by-product of what others have passed on to us.

But the noise still bugs us. And with each reduction in noise, the bombardment of information is stepped up a notch.

Cut the noise and you are punished for your innovation with not just an equal but exponential bombardment of new information.

"On the Internet no one knows you're a dog," or at least so the tagline went in the early 90's. Cyberspace was thought of as being disconnected from the physical world. The interfaces were so abstract, and the few people on the Net were so geographically spread out across the planet, it certainly did feel that way. But what we were forgetting was that Cyberspace only existed because it's entire population were pounding away at keyboards in darkened rooms which were unquestionably still in meatspace.

Ideas are not physical in any sense. The products of intellectual and creative work are not property, but shadows cast by the mind as part of a process of taking in the world through the senses and then trying to make sense, identify, label, define and eventually understand in order to take action.

So cyberspace is a collective tapestry of our mind's interpretation of what our senses have gathered, overlayed and interwoven through the world.

The process of digitalization can be thought of as a technological consolidation of all our different technologies for fixing what we experience and our interpretation of them into a single system where all forms of writing, and recording of images and sound are interoperable. The network revolution is a complementary consolidation of communications, broadcasting and publishing.

But what has not yet happened is the corresponding shift in how we use this new medium.

We live in turbulent times, much like the end of the 19th century as the horse was rudely goosed to the side of the road by a puff of steam. But it wasn't steam that displaced the horse, it was the internal combustion engine which finally did that.

Before the telegraph, communication was a form of time travel where most information described events after the fact. An earthquake in Tokyo was something that happened in the past to someone living in London and vice versa.

The telegraph transformed communication so that everything that happened, happened everywhere simultaneously. Think of that. All those dots and dashes, tapped out across the wires, a beat you could almost tap your foot to and about as abstract in the moment as Bancusi's "Symbol for Joyce" (and if you got that one, I am truly sorry for you).

Like steam, the telegraph changed communications, but did not transform them for the average man. This feat was accomplished by the telephone, radio and television which brought us together into the same room.

Collapsing time changes our perception of space. In the global village, everyone is your neighbor. And every day, people all over the world turn on their televisions and see their new neighbors and mumble under their breath, "there goes the neighborhood..."

The PC revolution, the Windows GUI and the Office Suite are a lot like steam. They are clunky, transitional technologies which got people to adopt them, but aren't as revolutionary as they like to think of themselves.

Information in 19th century was based on paper. Communications, entertainment, business, government and even organized religion all used paper as the means of creating, organizing and controlling information. And, as has been said, the way an organization organizes information is the way it organizes power in that organization. Since all information was on paper, paper became synonymous with information.

The computer came along with a new way of creating and organizing information, but most people couldn't imagine information without paper, so there was little interest or adoption of computers as personal tools until the Desktop GUI, Word Processor, Spreadsheet, and presentation software gave us paper metaphors for using computers to work with information.

That paper crutch is beginning to show it's age and it's time for us to begin moving to a new conceptual framework for finding, creating, organizing and sharing information to replace it, just as the internal combustion engine replaced the steam engine.

Only when this happens will we truly have begun to live in the networked computer age.

This is where we begin.

§ 6 Required Reading

This paper assumes the reader has a working knowledge of XML and the basic concepts in the IFLA's FRBR [Functional Requirements for Bibliographic Recrods] [6] and ANSI Z39.19 [Guidelines for the Construction, Format, and Management of Monolingual Thesauri] [7].

It's strongly suggested that the reader keep copies of these papers as companions to this paper.

§ 7 Status of BMF

At the time of writing (April 2006) BMF has stable core feature set. A working schema is in place as well as a usable alpha version of a BMF browser and development environment.

In August, 2006 the BMF Guidelines (which will be a greatly expanded version of this paper), will be released for public comment together with the BMF schema, a comprehensive set of BMF encoded content for testing applications, and a content browser and development environment running in the Emacs text editor.

BMF will be released as an open specification under a free license.

§ 8 Concepts

A tragic sigh. "Information. What's wrong with dope and women? Is it any wonder the world's gone insane, with information come to be the only real medium of exchange?"

"I thought it was cigarettes."

"You dream." ....

—*Gravity's Rainbow*, pg. 258.

BMF [Burr Metadata Framework] is built on a number of core concepts which, taken together, form a vision for the next generation of the Internet, digital content and communications.

These concepts are the consequence of the two central trends which have sparked the twin Digital and Network revolutions.

Just to get these out of the way, these are:

• The process of converting all text, sound, images and video into native digital formats.

- The use of TCP/IP to network all services, forms of digital content and communications.

These two trends have a number of consequences, many of which we are already aware and others we are just beginning to recognize.

Much of this can be described in terms of information having a dual nature which is discussed in the next section but can summed up with the following five assumptions which BMF is built on:

1. There is no useful difference between data and metadata which describes it.
2. There is no useful difference between a single book and the library it is collected in.
3. There is no useful difference between the parts of a work (which can stand on their own) and the work itself.
4. There is no useful difference between a document and commentary made about that document.
5. There is no useful difference between text and code (computer software code).

BMF also draws on a number of other key concepts which include:

- LOD (Level of Detail) borrowed from the the world of 3D modeling.
- REPL [Read Evaluate Print Loop] from Lisp, which embodies idea of creation and organization as a process rather than an end in itself.
- Multi-pass markup at the center of the creation and editing process which makes it easy to move from the general to the specific.
- The idea of catalogers using folksonomies (tags) in the same way that lexicographers treat new words which have entered a language.
- And finally, that BMF data structures can be described as an ideal embodiment of Delueze and Guarrti's Rhizome metaphor.

§ 9 The Dual Nature of Information

Physical media comes with a lot of baggage. In many respects, since Caxton, mankind has increasingly based whole civilizations on this baggage.

Digitization and networking have all but removed the limitations that physical media impose though most people haven't realized this yet. Centuries of living within those confines have led us to believe that they are universal laws which can't be challenged.

The limits of physical media are physical — you can only fit so many words on a page, only bind so many pages into a book before it gets too big to handle.

Once you have divided words into volumes you need a means of organizing the information in each volume. It's practically impossible for a library to create a single index for every keyword in every book in the collection, or to create a single table of contents, so these navigational devices were created only at the level of single volumes. Library catalogs could practically only seek to treat each volume as an item, so the catalogs stopped at the covers of the books.

Significant physical resources are required to duplicate and distribute physical media and economics favors larger volumes which contained a lot of information rather than smaller publications. So smaller texts were collected into larger volumes, individual songs were collected into LP's (long playing record albums) etc.

After you strip away the paper from a text, the vinyl from a record album, or the film from an image, one of the first things that starts to become apparent is that those divisions are indeed artificial and that when they are removed information begins to behave as if it has a dual nature like the dual particle-wave nature of light.

BMF is based on five general principles for how this dual nature applies to information.

1. There is no useful difference between data and metadata which describes it.
2. There is no useful difference between a single book and the library it is collected in.
3. There is no useful difference between the parts of a work (which can stand on their own) and the work itself.

4. There is no useful difference between a document and commentary made about that document.

5. There is no useful difference between text and code (computer software code).

9.1 *Data and Metadata*

The idea that data and metadata are interchangeable is both natural and astonishing at the same time.

We think of metadata as a description of something else, in the way that a card in a library catalog is an external description of a resource in a library.

But a collection of bibliographic data on a particular subject becomes a bibliography which is a work in it's own right. The title page in a book, the liner notes in an album or a telephone directory all can be thought of as data in one context or metadata in another.

If metadata and data are indeed interchangeable, then metadata is not inherently external. This leads us to a very different concept of metadata.

Metadata is not simply a description of data, but a less detailed view of that data. Metadata is data seen at a distance.

9.2 *Library and Book*

> For our purposes, the document and the library are essentially the same. In other words, the traditional library-document dichotomy can be viewed as a smooth spectrum, which we consider as a whole.
>
> Towards one end of the spectrum, the number of authors decreases and the topics under discussion become more integrated, and the information artifacts look more document-like. Towards the other end, the number of authors grows and the semantic gaps between topics increase, and the information artifacts become more library-like.
>
> —*A Scholia-based Document Model for Commons-based Peer Production*, Joseph Corneli and Aaron Krowne [CORNELI]

The illusion of the distinction between document and library is in large part a by-product of the limits of physical media and Caxton's printing press.

Before Caxton, the distinction between a work and library was far less clear as was authorial ownership of documents and all sorts of other assumptions that we take for granted today. We'll come back to this point again later.

Once you have digitized all the works in a library and placed them within a single framework, the distinction is far less clear.

For example, in a digital library you can have one index rather than a different index at the end of every document. The table of contents, which is a tree, can be merged together with all of the other table of contents of all works in the library into a single tree.

The library catalog can be merged with all of the works they describe so that a bibliographic record is a description of a work at a distance.

Links between documents can lead directly to any part of any other document without the reader having to open the document like the cover of a book, work out the organization of the work and only then find the passage that was being referenced.

9.3 *Work and Part*

Many books and sound recordings are not mutually exclusive, but are collections of a number of smaller documents or songs which could easily stand on their own.

In some cases, the collection itself has value as a work in it's own right, but this does not take away from the fact that the parts could stand on their own.

Encyclopedia articles, main entries in dictionaries, newspaper stories and even chapters in many books could stand on their own without the reader needing to see any other part of the collection.

Many collections are for the sole purpose of making the amount of content that is sold on physical media viable as a commercial product. Sound recordings are well known for including songs of dubious quality

to make a album with a few popular singles long enough to sell as an album and justify a themed concert tour.

But the MP3 revolution and more recently iTunes and the iPod have brought back a new age of singles. iTunes are the digital equivalent of old 45rpm records which were the backbone of the recording industry during the 50's and 60's when radio was the chief marketing vehicle for music.

The first decade of the World Wide Web was based in large part on the idea of a Web Site being a mutually exclusive collection of information. In effect, Web Sites were treated as self-contained works like a physical book. Imposing the limits of physical media on electronic media is a theme which has been repeated over and over.

For the Web, RSS [Rich Site Syndication Format] blew this idea out of the water by breaking up content so that individual articles on the Web could stand on their own, irrespective of the Web Site which published it.

9.4 *Text and Commentary*

The relationship between text and commentary is probably as old as texts themselves.

Commentary can take all sorts of forms, such as foot-notes, glosses scribbled in the margins of a book, or notes made while reading a book for a class. Commentary can be as small as a single word or a multi-volume work composed by an army of scholars.

The commentary made by an authoritative person with lots of letters tagged on the end of their name and published along with a document, are not functionally or practically any different than notes scribbled by a high school student doing their homework on the kitchen table.

Such commentary is often a marketing function for a publisher, who is trying to add value to a work (which might be in the public domain) to try to coax readers to purchase their edition over another.

This is not to say that such commentary is not useful or important. It is enormously important to provide context and insight into texts which were based on common knowledge used within a narrow discipline or general knowledge from a past age.

Once commentary is understood to be simply a text, which has as a subject another text, irrespective of who wrote it or how it is published, then all commentary becomes an extension of and part of a work and by extension, the collective content of a library.

It could be said that the Internet itself is all commentary. Email between friends, or in a discussion group on Usenet or on a list-server, threaded comments on Slashdot[8], tags and comments about images on Flickr, bookmarks on del.icio.us, reviews on Amazon Books, and of course the entire blogsphere is all a relentless tidal current of commentary that ebbs and flows across the planet as each timezone passes from day into night.

9.5 *Text and Code*

Everything in Lisp is a list. There is no useful distinction in Lisp between the code and the data it is processing.[9]

The expression `(+ 2 2)` which is the way you write "2 + 2" in Lisp is a list with three elements where the first item is a symbol which represents a function ("+" is the name of a function which adds numbers together) and the second and third items are the numbers "2" and "2".

Documents which are marked up as Lisp data structures can be thought of in one context as a document, and in another as a program which can be evaluated (or invoked) to get a result.

To understand this, think of Harry Potter who lives in a world where magic is real. In Harry Potter's world, a device like a wand, is used to invoke spells which are spoken. This results in some kind of action which can be anything from levitating a chair, to erasing someone's memories.

Among other things, magic is based on the premise that human language, when used by someone with the appropriate skill and innate ability, has the power to effect the physical world around us. Speaking, or incanting a spell invokes unseen powers which can move and manipulate physical objects.

This belief is as old as humanity. Written texts in some contexts are believed to have magical powers in their own right. Sacred texts like the Bible are thought by believers to have the power to protect them from evil, and invoke supernatural powers.

I am writing this paper using Emacs, a text editor written in Lisp. I can move my cursor next to the expression `(+ 2 2)` on the screen and invoke the expression with a tap of my wand (by holding down the Control key and typing "x e"). The number "4" is returned in a window at the bottom of the frame.

A hypertext link on a Web page behaves in a similar way. When you click on a link and the browser opens up another page, you are invoking the link made between two documents.

The distinction between text and code will gradually fade. Twenty years from now, we could well have a generation of children who will have a difficult time thinking of a text as being an inert chunk of information permanently stamped on physical media.[10]

§ 10 The Legacy of Paper

When content recorded on physical media has been digitized and placed in a larger framework, you have in fact ripped the covers off of all books and tossed all of the jewel cases and album sleeves (if you are old enough to remember those) into the bin.

The PC revolution was based on convincing people that computers were just electronic versions of what they already knew. And what people knew was paper.

The desktop metaphor at the heart of the graphical user interface is based on manipulating and managing pieces of paper.

The now ubiquitous "Office Suite" is little more than a metaphor for it's paper counterparts. Word processors are typewriters, spreadsheets are ledgers, and presentation software like Powerpoint is foam core on an easel.

The Web too is built on paper metaphors. The Browser Wars were driven, at least in part, by the addition of proprietary features by Netscape and Microsoft that people were demanding to make Web pages look and feel more like paper based documents, magazines and catalogs.

Many traditional publishers who established Web sites brought with them the same territorial attitude that they had about physical media. They wanted people to first visit their home page before seeing any other content on the site in the same way that you have to see dust jacket of a book before seeing what's inside.

§ 11 Ripping off the covers

The consequence of the digitalization and networking of all content and communications is to erase the illusion of each work being a self-contained universe which is created by the limits of physical media.

The first major crack in the paper legacy was with the widespread adoption of P2P. Napster so completely destroyed the music record album as a mutually exclusive unit of content that the recording industry was left dumbstruck and it was left to companies like Apple with iTunes and Musicmatch to cash in on the new era of music singles.

The second great fissure was RSS which pulled content from millions of blogs into a breathtaking interconnected Web of content, rather than just a network of Web Sites.

Much of the anguish and beating of breasts by publishers and authors when Google Print was launched have nothing to do with copyright violations. What really scared them, though they probably didn't know it, was that Google had violated the sacred covers of the book and replaced the index at the back of the book with an index which could be used for all books ever written. Google had ripped off the covers and shattered the illusion that a book was a self-contained universe which can't be messed with.

This was as rude a shock to the publishing world as P2P was to the film and music industry. It never occurred to anyone to think that something as sacred as the sanctity of the covers of a book could be violated. The novelist John Updike recently summarized these sentiments in an anti-ebook rant in the New York Times, heavily laden with nostalgic memories of bookshops. [UPDIKE]

This same process will be repeated again and again at all levels of the information hierarchy until everything has been digitized and assimilated into a single global fabric of information containing all of mankind's experience and memory.

§ 12 Creation and Organization as Process: The Read Evaluate Print Loop

The Lisp concept of the REPL [Read Evaluate Print Loop] is all around us. Any process that collects information, requires you to do something with it and then take some kind of action with it, is an instance of the REPL.

The term *REPL* comes from the process used to write Lisp programs. But it is also a good way of thinking about more general and practical issues of how humans work and process information.

12.1 *The Lisp REPL*

Lisp is a programing language which has been around since 1958. In fact the only programming language older than Lisp which still in active use is Fortran. Lisp was far ahead of it's time. Many of it's most powerful features have only been introduced into more popular languages like Perl and Python in the last few years. Many people still consider Lisp to be more powerful than any other programming language. The Read, Evaluate, Print Loop (REPL) is a part of the Lisp development environment for writing Lisp programs.

> Lisp languages are frequently used with an interactive command line, which may be combined with an integrated development environment. The user types in expressions at the command line, or directs the IDE to transmit them to the Lisp system. Lisp *reads* the entered expressions, *evaluates* them, and *prints* the result. For this reason, the Lisp command line is called a "read-eval-print-loop", or *REPL*.
>
> —*Wikipedia: Lisp programing language* [WIKIPEDIA-LISP]

So why are we using the term REPL? After all, we could just as easily call it the "Search, Process, Publish Cycle" or SPPC. Is there a reason for using such obscure hardcore geek terminology? Well, yes.

The REPL embodies both the human process, as well as the machine process and keeps in mind our fifth principle that there is no useful distinction between text and code.

12.2 *A Simple Example of the REPL*

One of the most simple and elegant examples of the REPL is found in practically every office on every desk in the form of the ubiquitous, *in-tray*, *pending-tray* and *out-tray*.

Information is dropped into your *in-tray*. In many offices there is a cover note which indicates where the information came from, who sent it, what action you are required to make and then a list of other people who are expected to receive the information.

You take a look at it, evaluate it. Then you either deal with it right away, perhaps by just reading it, and marking on the note that you've seen it. You then drop it into your *out-tray* and it is picked up and filed or passed on to the next person in the chain.

If you can't evaluate something right away it is then put in a *pending-tray* to be evaluated at a later time.

Many people also keep in and out-trays on their desk at home, but in many cases (including myself) they tend to fill up without things ever moving out the in-try. Over time, pending and out-trays eventually just becoming holders for the overflow when the in-tray has reached capacity.

The reason for this is that there mechanism like the cover or action-note attached to items to keep information flowing and no-one to pickup things from the out-tray and pass them on to others.

We will come back to this point later.

12.3 *Developing Ideas*

An idea often proceeds and triggers the REPL which can be anything from something funny in an email which you want to remember, or a news story about a new product which you think you might be interested in. Any information you find of interest which you want to remember, or know more about, or might be of interest to someone you know is all fodder for the REPL.

Sometimes this will lead to an action, or something that lead to writing a report, or proposal, making a purchase or changing jobs.

The REPL represents a process which employs any number of techniques and approaches. It's worth looking at each step in the loop.

12.4 *The Read Process (search and collect)*

The *read process* of the loop includes searching, collecting and remembering information that we are looking for or that we come across.

Searching and collecting information is a continuous, ongoing process. Sometimes this is done deliberately done, and other times information may be sent in an email or in dropped in the inbox and kept until it can be evaluated at a later date. It's common to collect information on specific topics over days, weeks, months and even years before there is enough information to be acted on.

12.5 *Evaluate (edit, sort, organize)*

In many respects, the *evaluation process* is the most important part of the REPL and oddly enough, it's the part that has received the least attention from software developers.

The *evaluation process* uses what we find and collect to make sense of it and decide on actions to take based on what we come up with, the *evaluation process* includes a wide variety of techniques which are used in any number of ways by each person depending on their preferences and the job at hand.

This requires a set of tools which can be as simple and general or as complex and fine-grained as is needed. The evaluation process is as much a creative process as much as a formal process and tools should be flexible enough to work with whatever information you are working with, rather than imposing limits on how you can display, edit, sort or publish that information.

12.6 *Print (share, exchange, publish)*

The print process involves editing the results of the evaluation process into a format that can be understood by others and distributing it.

The most informal way to do this is through email. Email allows us to easily exchange information with other people or groups (mailing lists).

In the past few years, blogs have emerged to fill a need for publication which is more formal than an email, but it is far less formal than something that has been published. A Blog is what we come up with after going through the initial evaluation process. It's a means of getting feedback on ideas in progress and whatever other half-baked stuff that is going on in our brains.

For formal publication there are Journals, Newspapers and Books which have traditionally been paper based but are increasingly being replaced with Web-based services.

But formal publication is not just a matter of a single person making something public. Publication is a collaboration which requires intermediate steps including peer-review (or review by some authority), copy editing, and formating which follow conventions that makes it easier for people to understand what is being published.

The print process is a means of travel, both through space and time. The added steps taken for formal publication are important in order for works to become part of mankind's collective knowledge.

12.7 *Loop*

The feedback we get from the print process is then fed back into the beginning of a new loop to be read again. This will then spark new ideas which prompt us to search, collect and remember new things which are passed on to be evaluated again.

This process is used to understand change, to make decisions, and to contribute new information for publication and addition to mankind's collective knowledge and memory.

12.8 *Integrating Spheres of Information*

There is still the problem of exchanging information between people and people, groups and groups and between people and groups.

Each person or group is surrounded by a sphere of information which is processed using the REPL. This information sphere is made up of all of the email, notes, addresses, receipts, images, media, publications and other information which has been collected by the REPL process and makes up the base of information which is used to understand the world around us and to make decisions on how to deal with the world as things change.[ENGELBART]

There is no single way of accomplishing this. The way we organize information and put it in context determines the value and meaning of that information. Everyone uses a different process to collect information for different purposes, and evaluate and organize that information in different ways.

If you send information to another person or group without the context and structure that gives it meaning, the person or group receiving it will spend a large amount of time pulling that information apart and organizing it and putting it into context that they can understand and be used by their own REPL process.

What is missing is a means of making it easier for each person or group to easily integrate information sent to them into their own REPL process without having to strip everything down to bare wood. This has the potential of saving an enormous amount of time and resources in the exchange of information.

§ 13 Level of Detail

We have already touched briefly on the idea of metadata as data at a distance. In 3D modeling and animation, there is a similar concept called LOD [Level of Detail].

A 3D model is made up of polygons. The more polygons you have, the more detailed the model. And the more detailed the model, the more clock cycles your computer will have to burn to render them on screen.

The model for King Kong in the recent remake of the movie is likely composed of millions of polygons. And in complex scenes Kong will have to share the stage with any number of other high polygon models including dinosaurs, giant cockroaches, buildings etc.

For close ups you need all of those polygons to create a realistic image, but if the shot is from a distance, most of the detail is wasted. Your computer is computing polygons which will never be seen.

LOD is used to reduce the number of polygons in a model the farther away it is viewed. This saves an enormous amount of computational power that can be put to better use in models which are close up.

The same principles can be applied to a book or even a library.

If you are standing across the room from a book on a shelf you are looking at the book from a distance. All you might be able to read is the title, author and publisher on the spine. Walk up to the book, take it off the shelf and open to the title page which shows metadata describing the book in more detail. Go to the table of contents and you are closer still.

```
- list display -- title, author
 - scope note -- one or two lines describing the item
  - detailed metadata and scope note.
   - introductory note or synopsis
    - detailed introduction or analysis
     - table of contents
      - chapter synopsis
       - text of chapter
```

This hierarchy of detail is not simply a convenient means of organizing and finding information, it is an important part of the creation process.

Creating information within a framework which incorporates LOD is far more flexible in how and what you can create. This is covered in more detail in the next section.

§ 14 Extend, Expand & Refine

If the REPL represents the larger repeated process of acquisition, evaluation and publication, then what is happening in each iteration of the the loop?

Knowledge advances through the advancement of increasingly more complex and accurate systems which build one atop the next without canceling out what came before. Quantum mechanics was built on Einsteinian Relativity which was built on Newtonian Physics which had been built on Copernicus' model of planetary motion.

This principle goes to the heart of the process of creation.

14.1 *Top-down, Bottom-up.*

In programing there are two great design methodologies, top-down and bottom-up. Top-down favors the prepared, while bottom up favors the prepared mind.

A top-down approach to writing a novel might be to define the setting for the novel, outline the characters, and then writing an outline for each chapter. When the outline is complete you simply write every chapter according to your outline.

Top-down is favored by large organized projects and is perfect for projects like bridges, rockets and dams which need to have all the kinks worked out beforehand or there could be some nasty consequences.

In terms of what we've been talking about, a top-down approach starts by describing something from a distance and then approaching what you are creating, by creating increasingly more detailed descriptions until you are finished.

A bottom-up approach might start with writing a simple sentence like "Marley was dead: to begin with" without a clue as to who Marley was or how, when or why he was dead. From there you just continue writing and let, to paraphrase Tolkien, "the tale grow in the telling."

Bottom-up is an organic meandering experimental learning, process, full of blind allies, wild epiphanies and a lot of mistakes along the way allowing you to create things that you hadn't intended when you set out.

Bottom-up can start anywhere, from a distance or right smack in the middle. From there you can work your way closer by adding detail, adding new threads, lengthening existing ones and unraveling bits that you don't like as you go along.

In practice, people tend to use a mix of both top-down and bottom-up, a combination of planning peppered with taking advantage of the unexpected encountered along the way.

Collections of information, no matter how large or small must reflect both top-down and bottom-up methodologies. An electronic library should be able to represent works in progress, aborted drafts and anonymous fragments as transparently as it can handle polished published masterpieces.

14.2 *Multi-pass Markup*

> The <hi> element is used to mark words or phrases which are highlighted in some way, but for which identification of the intended distinction is difficult, controversial or impossible. It enables an encoder simply to record the fact of highlighting, possibly describing it by the use of a rend attribute, as discussed above, without however taking a position as to the function of the highlighting. This may also be useful if the text is to be processed in two stages: representing simply typographic distinctions during a first pass, and then replacing the <hi> tags with more specific tags in a second pass.
>
> —*TEI Guidelines*, 6.3.2.2 Emphatic Words and Phrases [TEI5]

The process of creating complex, semantic markup and metadata is hard work which takes time, and a lot of thought. In a world of exponential change, all of these things seem in short supply.

Depending on the task at hand, people won't adopt a system which is too difficult to do simply things or is too simple to do complex things with.

Even if your ultimate goal is to create something rich and complex, if it takes too much effort to start the process, not many people will get very far.

So an important design goal for BMF is to make it as simple as possible to jot down a note which enters the system with little or no thought and then at a later time that note can be added to, linked to other related terms, and eventually develop into as complex and dense a structure as is needed.

This can be accomplished by doing composition and markup in multiple passes without there being any requirement for anything to be more complex than it is in order to become part of the larger collection.

The idea of marking up texts in multiple passes is certainly nothing new, but it hasn't had a lot of attention lavished on it either.

To be clear, we are talking about markup here, not application user interfaces which hides the markup and presents a relatively simple interface to the user. Everything we will discuss in this section should be possible using a good text editor with basic syntax hi-lighting.

No one syntax will accomplish this, so instead we will use three different syntaxes which build one on top of each other. And just as importantly can gracefully degrade as well.

Let's now use an example to start with the most simple encoding to the most complex semantic markup possible.

Our example is a simple entry from the Dictionary of Angels[11].

At the bottom of the ladder is structured plain text. We prefer to use UTF-8 for all text, but for this example let's use basic ASCII.

```
Omael -- an angel who multiplies species, perpetuates
races, influences chemists etc.  Omael is (or was) of the
order of dominations and is among the 72 angels bearing the
mystical name of God Shemhamphorae.  Whether Omael is fallen
or still upright is difficult to determine from the data
available.  He seems to operate in both domains (Heaven and
Hell. [Rf. Amberlain, La Kabbale Pratique.]
```

Plain text has a lot going for it. Basic structural formating like paragraphs and sentences, lists etc can be easily indicated and there are a wide variety of tools for processing plain text.

But it is difficult to unequivocally indicate sections, headers, bold or italic text. To do this we can use a Wiki Markup language[12].

```
**Omael** -- an angel who multiplies species, perpetuates races,
influences chemists etc.  *Omael* is (or was) of the order of
dominations and is among the 72 angels bearing the mystical name
of God Shemhamphorae.  Whether *Omeal* is fallen or still upright is
difficult to determine from the data available.  He seems to
operate in both domains (Heaven and Hell). [Rf. Amberlain, La
Kabbale Pratique.]
```

The wiki markup is simple and easily converted into HTML or in our case, simple BMF. Block level and inline markup in BMF is based on TEI, so the following markup may look familiar.

```
<p><hi>Omael</hi> -- an angel who multiplies species,
perpetuates races, influences chemists etc. <hi>Omael</hi> is (or
was) of the order of dominations and is among the 72 angels
bearing the mystical name of God Shemhamphorae.  Whether
<hi>Omeal</hi> is fallen or still upright is difficult to
determine from the data available.  He seems to operate in both
domains (Heaven and Hell). [Rf. <hi>Amberlain, La Kabbale
Pratique</hi>.]</p>
```

But now we might want to mark this up more carefully, identifying each name and title and treating this as a formally marked up text in a division entity within an expression entity representing the book.

```
<p><pn>Omael</pn> -- an <top>angel</top> who multiplies species,
perpetuates races, influences chemists etc. <pn>Omael</pn> is (or
was) of the order of <top>dominations</top> and is among the 72
angels bearing the mystical name of God <pn>Shemhamphorae</pn>.
Whether <pn>Omeal</pn> is fallen or still upright is difficult to
determine from the data available.  He seems to operate in both
domains (<pl>Heaven</pl> and</pl>Hell</pl>). <ref>[Rf.
<tit>Amberlain, La Kabbale Pratique</tit>.]</ref></p>
```

Proper names have been marked up with the <pn> element, concepts with the topic <top> element and titles of works with the title <tit> element.

This is as far as most document-based markup languages will go. But BMF can then even go a step further by turning this entry into a proper record for the Angel named Omael.

First we'll use BMF Wiki shorthand to outline the Burr. The following markup is used in Emacs Burs, a BMF browsing and development environment. The Wiki syntax is based on Emacs Muse-Mode wiki syntax and is still in development.

```
              * hierarchy
$TT  top  Dictionary of Angels (topicspace)
$BTI top  beings (mythical & legendary)
$BTI top  dominions (angelic order)
$PT  per  Omael (angel; fallen or upright)
* terms
$PT  Omael   (angel; fallen or upright)
$UF  Shemhamphorae (used for the angel, Omael)
```

```
meta:
 ## entityType  : person
 ## PersonalName : Omael
 ## Affiliation : Heaven; Hell.
 ## Roles       : Angel.
     * scope
       An angel who multiplies species, perpetuates races,
       influences chemists etc.  Omael is (or was) of the order of
       dominations and is among the 72 angels bearing the mystical
       name of God Shemhamphorae.  Whether *Omeal* is fallen or
       still upright is difficult to determine from the data
       available.  He seems to operate in both domains (Heaven and
       Hell).
   * references
       - Dictionary of Angels. pg 212.
       - Amberlain, La Kabbale Pratique
```

We could can then mark this up it using BMF XML syntax. This example is simplified and shortened to make it more readable.

```
<BURR typ="per">
 <sec typ="hierarchy">
  <i r="TT"  e="top"  l="Dictionary of Angels" q="topicspace" />
  <i r="BTI" e="top"  l="beings"  q="mythical & legandary"
  <i r="BTI" e="top"  l="dominions" q="angelic order" />
  <i r="PT"  e="per"  l="Omael" q="angel; fallen or upright" />
 </sec>
 <sec typ="terms">
  <i r="PT"  l="Omael" q="angel; fallen or upright" />
  <i r="UF"  l="Shemhamphorae"  q="used for the angel, Omael" />
 </sec>
 <sec typ="meta">
  <entityType   l="person" />
  <personalName l=Omael" />
  <affiliation>
    <i l="Heaven;" />
    <i l="Hell." />
  </affiliation>
  <roles>
    <i typ="preferred" l="Angel." q="preferred"/>
  </roles>
 </sec>
 <sec typ="scope">
 <p><pn r="PT">Omael</pn> -- an <top r="BTG">angel</top> who
 multiplies species,
 perpetuates races, influences chemists etc. <pn>Omael</pn> is (or
 was) of the order of <top>dominations</top> and is among the 72
 angels bearing the mystical name of God <pn>Shemhamphorae</pn>.
 Whether <pn>Omeal</pn> is fallen or still upright is difficult to
 determine from the data available.  He seems to operate in both
 domains (<pl r="RT">Heaven</pl> and<pl r="RT">Hell</pl>).</p>
 </sec>
 <sec typ="reference">
 <i id="DOA" r="BTP" l="Dictionary of Angels"
   <a>Dictionary of Angels</i><b>/ Gustav Davison.
   - Toronto, Collier-Macmillan, 1967. - pg. 212.</b>
 </i>
 <i id="AMBERLAIN"  r="BT" l="La Kabbale Pratique">
   <a>La Kabbale Pratique</a><b>/ Robert Amberlain.
   - Paris: Editions Niclaus, 1951.</b>
 </i>
 </sec>
```

Multiple pass markup may not be painless, but it should at least ease the pain as much as possible.

14.3 *Folksonomies*

> info-civilians are remarkably cavalier about their information. Your clueless aunt sends you email with no subject line, half the pages on Geocities are called "Please title this page" and your boss stores all of his files on his desktop with helpful titles like "UNTITLED.DOC."
>
> This laziness is bottomless. No amount of ease-of-use will end it. To understand the true depths of meta-laziness, download ten random MP3 files from Napster. Chances are, at least one will have no title, artist or track information — this despite the fact that adding in this info merely requires clicking the "Fetch Track Info from CDDB" button on every MP3-ripping application.
>
> —Cory Doctorow, *Meta Crap* [DOCTOROW]

The problem with getting people to create metadata is that it's usually presented to users as something to do after the fact in the same way that librarians catalog material after it has been published.

If metadata is done in this way, few people will bother. They would rather save what they are working on and and kill time on #emacs or go for a beer. This is because metadata after the fact is something extra which can be put off to a later time. In most cases that time will never come.

This attitude towards metadata as something after the fact has changed dramatically with the introduction of folksonomies by Web services like Flickr, Technorati and Del.icio.us[WIKIPEDIA-FOLK].

A folksonomy is made up of tags. A tag is a keyword which is associated with an image, blog or web page. There is no rhyme or reason to creating tags, and more often than not they are created using nothing more sophisticated than free association.

The problem with tags is that they are flat. All tags are created equal, they are not hierarchical, or grouped or even spelled consistently. This limits what can be done with them.

The "wisdom of crowds" camp claim that collectively, people will create tag sets which are equal to or even superior to formal taxonomies. Others dismiss tags out of hand as being fuzzy and of limited use and will never be a replacement for formal taxonomies.

Both sides are missing the fact that folksonomies and formal taxonomies actually compliment each other and that the two approaches could be combined.

As a general rule, tagging is used for new content. Cataloging using formal taxonomies is done a bit later down the road when the dust has settled and the longer term value of that content has been deemed worth preserving.

Tags can also be useful of as a quick and dirty mnemonic to place something in context, so that we can remember the who what when where or why that information is worth remembering which a formal taxonomy normally wouldn't be able to provide.

Tags can be thought of as rough, first pass cataloging which can later be refined, defined and organized into a catalog record which is organized using a formal taxonomy.

Catalogers should treat tags as a lexicographer treats new words which they are considering for inclusion in a dictionary.

In this view tags aren't discarded in the formal taxonomy, but are incorporated into the process of cataloging and developing taxonomies.

For example, if you look up the tags used for Apple Computer's Web site you might see something like the following:

> computers, macs, apple applecomputer, osx, macintosh, itunes, powerbook, macos, lisa, appleii, stevejobs, ipod

These tags could later be used to create a record for Apple Computer.

```
BT   computers
PT   . Apple Computers (Computer manufacturer)
UF   . apple (tag)
UF   . applecomputers (tag)
NT   .. Macintosh (computer brand)
NT   .. OSX (computer operating system).
NT   .. Liza (computer model)
NT   .. Powerbook (computer laptop product series)
NT   .. iPod (electronic music player)
NT   .. iTunes (electronic music service)
NT   .. Macintosh OS (computer operating system)
RT   .. Steve Jobs (Am. co-founder of Apple Computer)
```

The tags which are directly used for Apple computer are included as *Used For* terms, and the other tags are replaced by preferred terms in the taxonomy, and included as *Used For* terms in the record for each of these terms.

```
        PT
        Macintosh (computer brand)
UF  mac (tag)
UF  macintosh (tag)
```

So tagging becomes the wiki markup of taxonomies, making it easier for people to include metadata as part of the process of creating something, not something done after the fact.

For those creating formal taxonomies and catalogs, tags provide source material for creating records based on consensus.

In addition, even when an item as been formally cataloged, tags may be used a mnemonic to help place content in a personal context.

§ 15 Rhizome Versus Trees

> The French Arthurian prose cycle with is various ramifications was not an 'assemblage of stories', but a singularly perfect example of thirteenth-century narrative art, subordinate to a well-defined principle of composition and maintaining in all of its branches a remarkable sense of cohesion. It was an elaborate fabric woven out of a number of themes which alternated with one another like the threads of a tapestry; a fabric whose growth and development had been achieved not by a process of indiscriminate expansion, but by means of a consistent lengthening of each thread.
>
> —Eugene Vinaver, Introduction *Works of Thomas Malory* [VINAVER]

In Vinaver's longer introduction to the three volume edition of the same work (which I don't have access to as I write this) he went on to make the argument that the Arthurian Prose Cycle was not, as critics like Sir Walter Scott would say, a badly written novel. It was not a novel at all, it was an entirely different form of narrative all together, which embodied a vegetable metaphor coined by Deleuze and Guattari (1976) called a rhizome. Umberto Eco would later summarize a rhizome as having the following properties.

> A rhizome is a tangle of bulbs and tubers appearing like "rats squirming on top of the other." The characteristics of a rhizomatic structure are the following: (a) Every point in the rhizome can and must be connected with every other point. (b) There are no points or positions in a rhizome; there are only lines (this feature is doubtful: intersecting lines make points). (c) A rhizome can be broken off at any point and reconnected following one of its lines. (d) The rhizome is antigenealogical. (e) The rhizome has it's own outside with which it makes another rhizome; therefore is not a calque but an open chart which can be connected with something else in all of it's dimensions.; it is dismountable, reversible, and susceptible to continual modifications. (g) A network of trees which open in every direction can create a rhizome (which seems to us equivalent to saying that a network of partial trees can be cut out artificially in every rhizome). (h) No one can provide a global description of the whole rhizome; not only because the rhizome is multidimensionally complicated, but also because its structure changes through the time; moreover, in a structure in which every node can be connected with every other node, there is also the possibility of contradictory inferences: if p, then any possible consequence of p is possible, including the one that, instead of leading to new consequences, leads again to p, so that it is true at the same time both that if $p, then q$ and that *if p, then non-q*. (i) A structure that cannot be described *globally* can only be described as a potential sum of *local* descriptions. (j) In a structure without outside, describers can look at it only by the inside\; as Rosenstiehl (1971, 1980) suggests, a labyrinth of this kind is a *myopic algorithm*\; at every node of if no one can have the global vision of all its possibilities but only the local vision of the closest ones: every local description of the net is a *hypothesis*, subject to falsification, about is further course\; in a rhizome blindness is the only way of seeing (locally), and thinking means to *grope one's way*. This is the type of labyrinth we are interested in. This represents a model (a Model Q) for an encyclopedia as a regulative semiotic hypothesis.[ECO]

Caxton's printing press effectively put a stop to this older form of narrative by permanently fixing a text in the form it was printed in. Making it increasingly difficult to add to the various threads that made up a prose cycle.

As we begin to see the end of the limits that physical media impose on creative and intellectual works, we will see new narrative models emerge which will increasingly look more like the thirteenth century Arthurian prose cycles.

I bring this up because on first reading about the rhizome metaphor in the 1970's I became obsessed with the idea of writing an electronic rhizomatic text. This was long before the Web or Ted Nelson's Xanadu. But the system that I envisioned at the time laid the groundwork for what would eventually become BMF.

The problem I faced at the time, was that at first glance a rhizomatic structure appears to be chaotic, without center and of limited value for structuring or modeling collections of information.

But this is not the case.

BMF is based on the idea of hierarchical relationships which are also links, like in a thesaurus.

Every Burr (the atomic unit in BMF) describes a mutually exclusive concept which is related to other concepts through broader, narrower, equivalent or related relationships with other Burrs.

Everything in BMF is treated as such a relationship, and every relationship can potentially point to another Burr which defines the concept.

This means that each Burr in BMF includes a partial tree that anchors it in relation to other Burrs in the collection.

Collections of Burrs, called Topicspaces, in turn are collected together to form Brambles, which also contain partial trees. But the partial trees in each Burr are not excerpts of the larger trees used by a Topicspace or Bramble. They are independent of the larger structure and may overlap with other structures.

In *Semiotics and the Philosophy of Language* [ECO], Umberto Eco quotes D'Alembert at length about his criteria for the *Encyclopedie*. The entire quote sheds light on BMF's rhizomatic structure.

> The general system of the sciences and the arts is a kind of labyrinth, a torturous road which the spirit faces without knowing too much about the path to be followed.

> But the disorder (however philosophical it be for the mind) would disfigure, or at least would entirely degrade an encyclopedic tree in which it would be represented. Our system of knowledge is ultimately made up of different branches, many of which have a simple meeting place and since in departing from this point it is not possible to simultaneously embark on all the roads, the determination of the choice is up to the nature of the individual spirit... However, the same thing does not occur in the encyclopedic order of our knowledge which consists in reuniting this knowledge in the smallest possible space and in placing the philosopher above this vast labyrinth in a very elevated point of perspective which would enable him to view with a single glance his object of speculation and those operations which he can perform on those objects to distinguish the general branches of human knowledge and the points dividing it and uniting it and even to detect at times the secret paths which unite it. It is a kind of world map which must show the principle countries, their positions and their reciprocal dependencies. It must show the road in a straight line which goes from one point to another; a road often interrupted by a thousand obstacles which might only be noticed in each country by travelers and its inhabitants and which could only be shown in very detailed maps. These partial maps will be the different articles in the encyclopedia and the tree or the figurative system will be its world map. Yet like overall maps of the world on which we live, the objects are more or less adjacent to one another and they present different perspectives according to the point of view of the geographer composing the map. In a similar way, the form of the encyclopedic tree will depend on the perspective we impose on it to examine the cultural universe. One can therefore imagine as many different systems of human knowledge as there are cartographical projections.

When you look at a Topicspace (collection of Burrs sharing the same id-space) from an "elevated point of perspective" you can see the Topicspace which holds all of the Burrs which are "more or less adjacent to each other" in a single tree. This is much like D'Alembert's world map.

Also, Like D'Alembert's articles in his encyclopedia, Burrs are partial maps which can only be seen close up revealing detail which can not be seen from a distance.

So let's go through Eco's definition of a rhizomatic structure point by point and see how it compares with BMF.

> *Every point in the rhizome can and must be connected with every other point.*

At the heart of BMF are two XML attributes which can be applied to almost every element\; the defined-by `d` and relationship `r` attributes.

```
<p><qt><pn r="RT" d="aut:TCS8-5152" l="King Kong">He</pn>
really did love <pn r="RT" d="aut:MAD1-6875" l="Fay Wrey">
her</pn> folks</qt></p>
```

This quote from *Gravity's Rainbow* which is marked as a paragraph shows that He (King Kong) is a related term which is defined by the Burr with the id aut:TCS8-5152 and that She (Fay Wrey) is also a related term which is defined by the Burr with the id aut:MAD1-6875.

It's important to remember that the relationships are between the concepts described by each Burr. So if the text of *Gravity's Rainbow* is encoded as an expression Burr, then King Kong and Fay Wrey are both related terms to the expression called *Gravity's Rainbow*.

```
             PT   expr   Gravity's Rainbow
RT   per     . King Kong
RT   per     . Fay Wrey
```

So any concept anywhere in any Burr can be mapped to any other Burr along with the relationship between the two Burrs.

> *There are no points or positions in a rhizome; there are only lines.*

In BMF links are the lines connecting every Burr at every point to any other Burr. And despite Eco's misgivings that lines intersect creating points, BMF's links can be thought of as molecular bonds between Burrs. So molecular structures which are directly linked to each other will not intersect, there are only lines bonding them together via relationships to each other.

> *A rhizome can be broken off at any point and reconnected following one of its lines.*

In BMF any Burr contains it's own detailed, partial tree which is independent of the larger macroscopic structure or structures which claim it as it's own (topicspaces and brambles).

For example a topicspace created for a library might categorize *Gravity's Rainbow* as a narrower term of a term called Post *Modern Novels*:

```
             PT   con   Post Modern Novels
NTI  work  . Gravity's Rainbow
NTI  work  . The Recognitions
```

Another topicspace used for bibliographies might categorize *Gravity's Rainbow* as a work by Thomas Pynchon.

```
             PT   per   Thomas Pynchon (Am. Novelist)
NT   work  . Gravity's Rainbow.
NT   work  . The Crying of Lot 49
NT   work  . V.
NT   work  . Vineland.
NT   work  . Mason & Dixson.
```

But if you open the Burr for Gravity's Rainbow you might see:

```
BT   work  Gravity's Rainbow.
PT   expr   . original 1973 text.
NTI  man    .. New York, Viking Press, 1973. - 1st ed.
                  - issued in hardcover and trade pbk.
RT   per    .. King Kong.
RT   per    .. Fay Wrey.
```

So any line can be broken off, and reconnected in context with the perspective of the local map.

> *The rhizome is antigenealogical.*

Deleuze wrote:

> There is always something genealogical about a tree. It is not a method for the people. A method of the rhizome type, on the contrary, can analyze language only by decentering it onto other dimensions and other registers. A language is never closed upon itself except as a function of impotence. [DELEUZE]

While a Burr might be tree-like (genealogical) seen up close, it cannot be mapped globally as a tree because any local tree can and must map to any other part of any other local map as well as any other world-map (another Burr or any topicspace that links to it).

> *The rhizome has it's own outside with which it makes another rhizome; therefore is not a calque but an open chart which can be connected with something else in all of it's dimensions; it is dismountable, reversible, and susceptible to continual modifications.*

Each local-map which defines a Burr, as well as any larger maps which define a topicspace, or a map of maps which defines a Bramble is open. But any tree which is defined, can be dismounted, reversed and changed anywhere and at any time. BMF is not a static framework, but designed to be in constant flux. Each local collection of Burrs will be different from every other collection of Burrs, each forming a rhizomatic structure, which can be turned inside out and form another rhizomatic structure.

> *A network of trees which open in every direction can create a rhizome (which seems to us equivalent to saying that a network of partial trees can be cut out artificially in every rhizome).*

BMF is a collection of macro and micro level trees, but they are completely open and have no center. Each tree starts in the middle and leaves off in the middle. You can only know the top term locally in the Bramble which you are using.

Since any topicspace in any other Bramble can be added to your local Bramble, the top term in a Bramble is just a placeholder for a top term — a horizon which recedes infinitely as you approach it.

> *No one can provide a global description of the whole rhizome; not only because the rhizome is multidimensionally complicated, but also because its structure changes through the time.*

This could be interpreted in several ways...

Monolithic structures controlled by a central authority can be frozen in time, and no matter how complex the structure you could theoretically create a global description of it. But decentralized structures can and are changed by anyone without the permission or even knowledge of anyone else.

You could also argue that a Bramble is a bit like Schroedinger's Cat which has both survived and died at the same. Only the act of opening the box to see what has happened to the cat will fix that outcome one way or another.

Every time you open a Burr, you establish relationships with all the other Burrs it is mapped to. Your Burr browser should then automatically download those Burrs and added them to your Bramble.

The act of opening a Burr determines what the Burr is, but it also changes your Bramble by updating Burrs and downloading new ones.

So no global description is possible because the structure changes as part of the process of observing it.

This leads us to the last two requirements which are more observations about rhizomatic structures than requirements.

> *A structure that cannot be described globally can only be described as a potential sum of local descriptions.*

> *In a structure without outside, describers can look at it only by the inside; ... at every node of if no one can have the global vision of all its possibilities but only the local vision of the closet ones: every local description of the net is a hypothesis, subject to falsification, about is further course; in a rhizome blindness is the only way of seeing (locally), and thinking means to grope one's way.*

In BMF only local descriptions are possible, but even they are not static, add a Burr to a Bramble and the entire structure changes. Add a note, an image, a book and it effects the global, as well as local structure of the collection and indirectly any other collection it is linked to.

§ 16 The Burr Metadata Framework

> The major difficulty, and it can be discouraging, is the large amount of reference needed to populate a poem that seeks to occupy and extend a world.
>
> —George F. Butterick, *A Guide to The Maximus Poems of Charles Olson*

Before going into detailed descriptions of all of the elements that make up BMF, it's helpful to provide a short overview of the framework and how the parts fit together.

16.1 *Burrs*

A Burr is the atomic unit in BMF. Each Burr describes a single mutually exclusive concept. A concept can be practically anything from a record describing a person, place, event, book, object or topic.

Burrs can be stuck together into molecular-like structures to create compound records or documents. This sticky nature of Burrs gave them their name.

Burrs come in a number of flavors called *Entities*. BMF entities are modeled on entities in the FRBR. Entities represent defined classes of types of metadata used to describe different classes of records and have nothing to do with SGML/XML entities.

16.2 *Topicspaces & BIXDs*

Burrs are collected together into collections which share a single id-space which are called topicspaces.

Topicspaces work on the same principle as namespaces in XML by using a unique URL (yes, that's a URL, not a URI) which uniquely and globally identifies a collection of ids which are unique within the topicspace and the location of the collection that it has come from.

All topicspaces used in a Burr are declared in the identity section of the Burr in a similar way that XML declares namespaces.

```
<topicspace>
<i pfx="aut" typ="http://chenla.org/bram/aut" />
<i pfx="evn" typ="http://localhost/bram/evn" />
<i pfx="evn" typ="file:///~/bram/evn" />
</topicspace>
```

In this example, the topicspace called "authority" has the URL **http://chenla.org/bram/aut**, which is the location where the topicspace can be found.

Topicspace prefixes are combined with an id system called the Burr Exchange ID or BXID. A BXID must begin with three characters from the Roman Alphabet, followed by a single number (0-9), a dash and then a four digit number (0000-9999).

BXIDs are mostly commonly used in the defined-by (d) attribute which is used throughout BMF to indicate Burrs which describe the concept that the element is describing.

```
<qt spk="Bill">Babes who've done time are so hot.</qt>
<qt spk="Ted">Yeah, like <pn d="aut:CIM8-4872">Martha Stewart</pn></qt>
```

Ids used to identify Burrs are made up of a topicspace prefix followed by the BXID:

```
aut:EBX5-1244
dov:JGL0-4847
```

The prefix is expanded to provide the URL, and the BXID itself is expanded to provide the directory structure for the Bramble. So aut:EBX5-1244 would be expanded into:

```
http://chenla.org/bram/aut/E/B/X/5/1/2/4/4/
```

The file name of the Burr is the BXID with a file extension of .wik or .xml. Processing applications should check first for an XML file, and if it doesn't exist look for a Wiki version before signaling an error or prompting the user if they want to create a new Burr with this id.

A topicspace can be used to collect any numbers of Burrs for any purpose. They can be used as an archive of email, a collection of day pages holding task lists, schedules and notes. They can be used for holding records of persons, an encyclopedia, a book, a dictionary, an inventory of medical images.... etc.

16.3 *Brambles*

Local copies of complete or incomplete topicspaces are kept in a Bramble.

BMF uses a distributed content model which is modeled on version control systems like CVS [Concurrent Version Control System] and Subversion, but in particular Arch[13].

While CVS and Subversion is based on the existence of a single repository. Users check-out a copy of the repository to work with locally and then check-in changes when they are finished if they have write access to the repository.

A copy can then become a branch of the main repository, and continue on as a separate project.

Arch has no centralized repository so everything is a branch which you have local write access to. If you want to submit a change to the maintainer you send it to her and she merges that change into her copy which acts as the central repository. Functionally, every repository is a central repository in Arch.

This is a good way of understanding Brambles.

When you browse a Bramble on your own local machine, you are looking at a file which is local. If you follow a link which points to a Burr you don't have a local copy of, the application you are using should grabs grab a copy of the Burr you want to see, as well as all immediate Burrs that the Burr links to and store local copies in your Bramble.

This approach might seem strange from a World Wide Web perspective, but BMF is a read-write framework, allowing anyone to annotate, edit and add content to their own personal collection of data.

Access restrictions can be placed on individual Burrs and even whole Topicspaces and Brambles so that publicly available material is not mixed up with private material and ensure that only the owner of a Burr is able to change or overwrite their own content.

16.4 *Links in BMF*

Links in BMF come in three flavors,

external	External links point to a resource outside of BMF. External links are the same as you would find in html, and most XML languages. External links use the `src` (source) attribute in many different elements. The value may be any valid URL.
pointers	Links within a Burr use the `ptr` (pointer) attribute which should point to a unique value in an id.
defined-by	The most common and important link in BMF is the `d` (defined-by) attribute which links to another Burr which defines the concept that the element is marking up.

16.5 *Relationships*

Most links in BMF specify the type of relationship that the link represents. These include equivalence, hierarchical, associative, responsibility, and sequential types of relationships between the concepts represented by the Burrs at each end of the link.

A large part of BMF is concerned with establishing relationship-links between the concept described by a Burr with related concepts described by other Burrs.

16.6 *Sections*

Burrs are divided into sections using the `<sec>` element. Sections come in a number of different types. A single section type can only be used once in a Burr.

For example, a simple Burr might include the following sections:

```
Burr
. Hierarchy Section
. Terms Section
. Meta Section
. Scope Section
. Reference Section
. Identity Section
. History Section
```

Sections can be roughly grouped into a couple of categories. These include the hierarchy sections and sections which extend it, the metadata sections which provide named fields for metadata, notes sections

for different kind of descriptive material, and finally sections for describing the Burr itself (version numbers, creation dates etc).

16.7 *Entities*

Burrs are divided into some forty different flavors called Entities. BMF entities (which should not be confused with XML and SGML entities) are based on the FRBR entity-relationship model.

BMF entities are clustered into entity groups which define relationships between in each group.

There are entities for persons, temporal concepts, physical objects, places, intellectual and creative works, symbols and words, concepts and technical documentation.

16.8 *XML serialization*

Burrs are serialized using XML syntax which can easily be parsed into SXML for use in Lisp data structures.

§ 17 Burr Structure

17.1 *XML Declaration*

Like all XML documents, a Burr begins with a XML declaration which must be the first thing in a file. No whitespace is allowed before the declaration.

```
<?xml version="1.0"  encoding="utf-8"?>
```

17.2 *The Document Element for locating the schema*

Optionally, Burrs may contain a documentElement containing the a uri attribute to locate the schema.

```
<documentElement prefix="xml" uri="bmf-1.rnc"/>
```

17.3 *The Structure of a Typical Burr*

The Root element for all Burrs is the <BURR> element. The <BURR> element is required to contain a typ (type) attribute which identifies the Burr's entity type.

A Burr is made up of <sec> (section) elements. There is no limit to the number of sections which can be included but sections can not be nested and only a sigle instance of each type of section is allowed.

The <sec> element requires a typ (type) attribute which identifies the type of section.

The overall structure of a typical Burr might look like this:

```
<BURR typ="entity-type">
<sec typ="hierarchy"> [ hierarchy list ...    ] </sec>
<sec typ="terms">     [ terms list ...        ] </sec>
<sec typ="meta">      [ meta fields ...       ] </sec>
<sec typ="scope">     [ scope note ...        ] </sec>

    [ other sections ... ]

<sec typ="reference"> [ cited sources list... ] </sec>
<sec typ="identity">  [ burr metadata info... ] </sec>
<sec typ="history">   [ change log entries ...] </sec>
</BURR>
```

This structure helps ensure that each Burr describes a single mutually exclusive concept. Keeping the structure relatively flat and simple makes it easier to keep each Burr focused on a single concept.

Compound Burrs can be created through the use of sections like the TOC (table of contents) section type which uses the source src attribute, to pull together any number of Burrs into complex molecular-like structures.

17.4 *BMF Schema*

The BMF schema uses the compact form of RELAX NG. The recommended tool for editing XML serialized Burrs is nxml-mode in Emacs which validates documents using a Compact RELAX NG.

The schema is heavily commented and designed to provide interim documentation until the BMF reference manual is ready for initial release.

§ 18 Relationships

The `d` (defined-by) attribute link is usually used in combination with the `r` (relationship) attribute. These two attributes together provide the glue which binds Burrs and BMF together.

```
<i r="BT" e="t" d="top:HGW6-7648" l="person"
   q="human being; living or dead" />
```

The relationship attribute uses a standard set of relationship codes which are found in all standard thesauri. A list of these codes can be found at the beginning of this paper.

18.1 *Kinds of Relationships*

BMF uses five kinds of relationships. The first three are drawn from ANSI Z39.19. [Z39.19]

- the equivalence relationship
- the hierarchical relationship
- the associative relationship
- the responsibility relationship
- the sequential relationship

Every relationship has the property of reciprocity, i.e. every relationship between term A and term B has a corresponding relationship from term B to term A.

18.2 *the equivalence relationship*

When two or more terms are used to describe the same concept, one term is selected as the preferred term, i.e., the descriptor. The equivalence relationship describes the relationship between preferred and non-preferred terms which describe the same concept.

The Preferred term should be indicated using the code `PT` (Preferred Term). Non-preferred terms should be indicated using the code `UF` (Used For). In an index, non-preferred terms point to the preferred term using the `USE` code.

```
              PT
              Charles Dickens
UF Boz (pseudonym for Charles Dickens).
```

```
Boz USE Charles Dickens
```

```
              PT
              Squid
UF Calamari
```

```
Calamari USE Squid
```

`UF+` (USED FOR . . . AND . . .) is used for non-preferred compound terms The + (plus) sign indicated that this is a compound term which is accompanied by an `AND`.

```
              PT
              coal
UF+ coal
AND mining
```

For compound terms which are made from two mutually exclusive terms to form a single concept, `USE +....` `AND` is used to indicate that both terms must be used together in an index.

```
coal mining USE+ coal AND mining
```

```
ferromagnetic films USE+ ferromagnetic materials AND films
```

18.3 *the hierarchical relationship*

The hierarchical relationship is what separates the men from the boys in BMF. It is the hierarchical relationship which turns the relationships and links in a Burr into a local tree.

The relationship is very basic and indicates if a term is broader (superordinate) or narrower (subordinate). The codes used to indicate this are BT (Broader Term) and NT (Narrower Term).

```
BT  Beings (real or imaginary creature)
PT  . Fairies (magical creatures)
NT  .. Brownies (magical creatures)
NT  .. Leprechauns (magical creatures)
NT  .. Goblins (magical creatures)
NT  .. Gnomes (magical creatures)
NT  .. Elves (magical creatures)
NT  .. Kobolds (magical creatures)
NT  .. Pixies (magical creatures)
```

18.4 *instantive hierarchical relationships*

This last example is okay, but Brownies, Goblins and Pixies are not simply narrower terms, they are instances of different types of Fairies.

The instantive relationship uses the codes BTI (Broader Term Instance) and NTI (Narrower Term Instance).

```
BT   Beings (real or imaginary creature)
PT   . Fairies (magical creatures)
NTI  .. Brownies (magical creatures)
NTI  .. Leprechauns (magical creatures)
NTI  .. Goblins (magical creatures)
NTI  .. Gnomes (magical creatures)
NTI  .. Elves (magical creatures)
NTI  .. Kobolds (magical creatures)
NTI  .. Pixies (magical creatures)
```

18.5 *The generic relationship*

The generic relationship indicates the relationship between a class of concepts and its members or species. The codes used are NTG (Narrower Term Generic) and BTG (Broader Term Generic).

```
BTG  rodents
PT   . mice
```

```
          PT
          rodents
NTG   . rats
NTG   . mice
NTG   . squirrels
NTG   . porcupines
NTG   . skunks
```

18.6 *Whole part hierarchical relationships*

The whole-part relationship is used to indicate that a term is part of a larger whole. The codes used are NTP (Narrower Term Partitive) and BTP (Broader Term Partitive).

```
          PT
          Crazy Horse (Musical Group)
NTP  . Niel Young
NTP  . Frank "Poncho" Sampedro
NTP  . Billy Talbot
NTP  . Ralph Molina
```

18.7 *Polyhierarchical relationships*

In some instances, a concept may belong to more than one category. When this cannot be avoided, the relationship can be indicated using multiple Broader terms.

```
BT  Peanut Butter (candies & sweets)
BT  Chocolate (candies & sweets)
PT  . Reces's Peanut Butter Cups
```

```
BTG  bones
BTP  head
PT   . skull
```

18.8 *Associative relationships*

When a relationship is neither equivalent or hierarchical but is associated with the concept the Burr is describing, the link should be made using the associative relationship which uses the code RT (Related Term).

```
BT Roman Catholic Church
PT  . Popes (officially recognized)
RT  .. Anti Popes (persons who claimed the title)
```

```
BT  pinup models
PT  . Pamela Anderson
RT  .. Bay Watch (Am. television series)
```

18.9 *Responsibility relationships*

The responsibility relationship is used to indicate when the concept represented by a term is responsible in whole or part for it's existence (such as a creative work) or taking place (such as an event). The codes used are BTR (Broader Term Responibility) and NTR (Narrower Term Responsibility).

```
BTR  per  Charles Dickens (person)
PT   wrk   . A Christmas Carol (work)
```

```
BTR  cor  3M Corporation
PT   evn   . Bhopal Chemical Spill
```

It should be noted that responsibility relationships are not a standard thesauri relationship type.

18.10 *Sequential relationships*

Another departure that BMF makes with standard thesaurus relationships is the addition of sequential relationships.

These are used in div (Division Entities) to indicate nodes which come before or after. In a full table of contents for a document all sections of the document are NTP (Narrower Term Partitive) terms of the document. But when you are looking at a single node, it is helpful to indicate which nodes precede the present node and which ones follow.

The codes used are PRE (Previous Node) and NEX (Next Node).

```
BT  A Christmas Carol (expression)
PRE Stave One (chapter)
PT   . Stave Two (chapter)
NEX  .. Stave Three (chapter)
```

18.11 *Node Labels*

Node labels are used in hierarchical displays to help show principle divisions which are helpful for the display but not intended to be used as index terms. The code used is NL (Node Label).

```
                          PT
                          Charles Dickens (Eng. novelist, 1812-1870)
        NL
                    major works
        NT    . The Pickwick Papers (1836)
        NT    . Oliver Twist (1837-1839)
        NT    . Nicholas Nickleby (1838-1839)
        NT    . The Old Curiosity Shop (1840-1841)
        NT    . Barnaby Rudge (1841)
        NT    . Martin Chuzzlewit (1843NT .1844)
        NT    . Dombey and Son (1846-1848
        NT    . David Copperfield (1849-1850)
        NT    . Bleak House (1852-1853)
        NT    . Hard Times (1854)
        NT    . Little Dorrit (1855-1857)
        NT    . A Tale of Two Cities (July 11, 1859)
        NT    . Great Expectations (1860-1861)
        NT    . Our Mutual Friend (1864-1865)
        NT    . The Mystery of Edwin Drood (unfinished) (1870)
        NL
                    Christmas books
        NT    . A Christmas Carol
        NT    . The Chimes
        NT    . The Cricket on the Hearth
        NT    . The Battle of Life
        NT    . The Haunted Man
```

18.12 *Top terms*

The top term in BMF uses the relationship code `TT` (Top Term) are treated somewhat differently from standard thesauri.

Top Terms in BMF are reserved for Topicspaces and Brambles. A Bramble is always a local Bramble, made up of what is on your computer, or the computer you are connecting to on the Internet.

So a Bramble is it's own top or root term or descriptor.

```
        TT    Bramble
        NTP   . topicspace 1
        NTP   . topicspace 2
        NTP   . topicspace 3
```

Topicspaces are the top term in their own hierarchy, so any Burr in a topicspace will point to the topicspace it belongs to as its top term.

```
        TT    Topicspace 1
        PT    . Burr 2
```

If you then go to the Burr which defines the topicspace you will see the local Bramble it belongs to.

```
        TT    Bramble
        PT    . Topicspace 1
        NTP   .. Burr 1
        NTP   .. Burr 2
```

This is very practical, because it allows you to download and keep a local copy of a Burr without having to remap the top term of your copy of the Burr.

§ 19 Entities

If the basic building blocks of BMF are Burrs, Entities are the paint making them different colors. They are still the same size and shape, and can be stacked together in any combination you choose, but you can tell at a glance what they are and sort them in useful ways.

BMF entities can be thought of in three ways:

- As types of entity-relationships as defined in the FRBR.
- As root facets or top terms in a thesaurus.
- As defining which metadata fields should be used to describe a common class or category of information.

Entities are distinct classes of collections of Burrs. Each type of entity has a different set of recommended fields and subfields associated with it and are considered to be a general enough class of information to require a distinct data structure.

Another useful way of thinking of BMF Entities is as a structural device of grouping records along the lines of who, what, when, where and which. Practically anything can be organized and placed in context by knowing the who, what, when, where, which and perhaps why of something. BMF asks this question at many different overlapping levels and scales.

19.1 *Entity Groups*

BMF attempts, whenever possible to expand on the FRBR model rather than change it. By expanding on FRBR instead of altering the model, it is relatively easy for BMF Burrs to map to any catalog system which adopts the FRBR model. [FRBR]

This should not be difficult since the places where BMF expand on FRBR are confined to parts of the model which are beyond the scope of traditional cataloging systems (ie. FRBR Entity Groups 2 and 3). Another advantage of this approach is that BMF can use the wealth of tools and experience that has been invested in National Bibliographic Records from around the world.

BMF has adopted the FRBR Group 1 almost verbatim, but expands FRBR Groups 2 and 3 into eight groups, for a total of nine entity groups and some forty entity-types (some of which are provisional).

Bibliographic Group	For describing creative and intellectual works
Agents Group	For describing Named Individuals, including persons, families and corporate bodies.
Semantic & Lexical Group	For concepts and lexical units including characters, words and phrases.
Temporal Group	For events and periods of time.
Physical Group	For Objects and Physical attributes.
Locus Group	For places.
Content Group	For encoding content.
Documentation Group	For self documentation of BMF.
BMF Group	For structural entities which make up BMF, namely Topicspaces and Brambles.

§ 20 Bibliographic Entity Group

> And to hook on here is a lifetime of assiduity. Best thing to do is to *dig one thing or place or man* until you yourself know more abt that than is possible to any other man. It doesn't matter whether it's Barbed Wire or Pemmican or Paterson or Iowa. But *exhaust* it. Saturate it. Beat it. And then U KNOW everthing else very fast: one saturation job (it might take 14 years). And you're in forever.
>
> —Charles Olson, Bibliography on America for Ed Dorn

20.1 *Introduction*

The entities in the first group can be thought of as a hierarchical representation of the different aspects of intellectual or artistic creations. This group is modeled on FRBR Group 1 Entities.

BMF uses FRBR's four different types of Group 1 entities to describe the different aspects of creative works.

20.2 *Entities*

work	A concept representing a distinct intellectual or artistic creation. Example: Miles Davis' Album. *Kind of Blue*
expression	The intellectual or artistic realization of a work, which is reflected in intellectual or artistic content. Example: The original Columbia Recording of *Kind of Blue* in New York, 1958 would be an expression. A live recording

of the same songs from *Kind of Blue* in Stockholm a few years later would be a second expression.

manifestation   The physical embodiment of an expression of a work. Example: the first LP Release By Columbia Records in 1958 would be one expression, and digitally remastered release on CD in 2000 would be a second manifestation of the same expression.

item   A single exemplar (physical copy or instance) of a manifestation. Example: the copy in the Library of Congress Reading Room would be an item or instance. The copy playing on your stereo next to your computer would be another instance. A copy of the CD for sale on EBay would be another instance.

20.3 *Relationships*

A *work* may share the same parent-work or story with one or more than one *work*, and may be *realized* through one or more than one *expression*.

```
                PT
                work
NTI    . expression
```

An *expression*, is the *realization* of one and only one *work*, but may be *embodied in* one or more than one *manifestation*.

```
BTI   work
PT    . expression
NTI   . manifestation
```

A *manifestation* may *embody* one or more than one expression, but may be *exemplified* by one or more than one item.

```
BTI expression
PT   . manifestation
NTI  .. item
```

An item may *exemplify* one and only one manifestation.

```
BTI manifestation
PT  . item
```

20.4 *BMF Bibliographic and FRBR Group 1*

Now that we have introduced BMF bibliographic entities we will compare the FRBR Group 1 with BMF Entity Groups in some detail as this concerns a number of important problems that were faced in developing BMF.

20.4.1 My God, It's Full Of Entities!

BMF has not been content with what many people in the cataloging community believe is an already complex FRBR/FRAR model. Nope, BMF broadens the scope of the FRBR far beyond the world of bibliographic data to create an infrastructure which is designed to describe, populate and extend whole worlds of both the conceptual and the physical.

The prospect of converting a hundred million or more WORLDCAT records in thousands of libraries into FRBR entities seems to have made more than a few folks wince, throw up their hands at the magnitude of the task and say that it's only possible to do it automatically.

The compromisers are already stepping in. Rather than expanding on the FRBR to truly make it powerful and flexible enough to take us to the end the century, they will try to water it down.

Converting all of those records, like the digitalization of works in print will measured in decades, or even generations, not years. So it's more important to do this right than to do it fast.

There are a number of proposals floating around which would do away with the expression altogether. This is wrong.

The manifestation is the working end of the FRBR for describing classes of physical objects, but it is the expression which will form the core of electronic libraries.

20.4.2 Unresolved issues with Group 1 Entities in FRBR

Should electronic documents be treated as expressions or manifestations? Trivial changes trigger new expressions (aka more work).. To quote the FRBR:

> Strictly speaking, any change in intellectual or artistic content constitutes a change in expression. Thus, if a text is revised or modified, the resulting expression is considered to be a new expression, no matter how minor the modification may be.

So even if the main text is the same in two documents but but they have different introductions it would trigger a new expression. In one sense this defeats the ability of expressions to group like texts together.

The format for titles for expressions have not been defined.

There are work titles which the spec says is the uniform title for all expressions beneath it. But there is no definition of what the expression title should be. From examples given in the spec it would appear that an expression title is very similar to "other title information".

FRBR treats records only at the item level.

Many of the troubles with applying the FRBR seem to be with the expression, but this is not strictly true. The bigger problem is that FRBR work-level entities are treated the same as item-level entities. This is mostly a bias based on the practicalities of managing records describing physical media.

But once you've introduced the idea of a work as a concept, you have tossed out those physical objects at least at the work and expression levels.

Manifestations describe classes of objects, and items describe physical objects. But work and expression entities won't work as long as they are treated as item level records.

A work is not a book. If a book contains an introduction by one author and the body of text from another author, the introduction and text should be treated as two distinct works. Each item must be examined to see if it contains a single work or is a composite of different works.

20.4.3 Summary of tweaks and clarifications to the FRBR Group 1 entities in BMF.

Works are concepts which are not restricted to the complete contents of published items. A published document may be made up of any number of distinct works which may be combined to form new works and expressions.

A master electronic markup of a work should be included as part of expression-level records. A master markup is meant for structural and semantic description of a document (BMF, TEI Docbook) and is not suitable for display.

An electronic markup of a work for display which is duplicated and distributed as a distinct file should be described at the manifestation level. Display markup languages include HTML, PDF Wiki and Plain Text. This includes links to any external Web resource.

Dynamically created temporary markup generated from master markup files in expressions used in displays should be treated as virtual manifestations which do not have records associated with them.

The titles for works, use a Uniform Title which is used for all expressions, manifestations and items under them. A responsibility statement may be used to distinguish the work from other works using the same title.

```
A Christmas Carol / Charles Dickens.
```

Titles for expressions are descriptive qualifiers which are concatenated with the (uniform) title of the work. When the expression is also used for the master markup of the text of a document, a responsibility statement should be included.

```
A Christmas Carol : original text and illustrations.
```

```
              / Charles Dickens; with illustrations by John Leech.
```

20.4.4 An example of FRBR Group 1 Entities in BMF

So for Charles Olson's *Special View of History* which was first published in 1970 with an original introduction by Ann Charters, we define Olson's work and Charter's Introduction as two separate works which are then combined into an expression combining the two.

The master markup is encoded in div (division) entities which are narrower partitive terms of the expression. This allows us to reuse the master markup in any other expressions which is based on the same original texts.

For the expression describing the original text of Olson's work, the work is a broader instance of the expression and the markup is broken into multiple division entities as narrower parts of the expression.

```
BTI   work  Olson's The special view of history
PT    expr   . original text
NTP  div    .. First page
NTP  div    .. Quotations
NTP  div    .. History, a definition.
[ ... rest of the chapters]
```

For the expression representing Charter's text of her original introduction. The work is a broader instance, and a single division is used to encode the text.

```
BTI   work  Ann Charter's Introduction to The special view of
              history.
PT    expr   . original text
NTP  div    .. body of text.
```

For the expression representing the book which combined Charter's introduction with Olson's text, the two works which make up the book are represented as broader instances of the expression and then the expressions of those works become narrower parts of the expression. By extension, the division entities become part of the combined expression as well.

```
BTI  work  Olson's The special view of history
BTI  work  Ann Charter's Introduction to The special view of
                history.
     PT    expr   . The special view of history: original text
                    with an introduction by Ann Charter.
     NTP  expr   .. original text of the introduction by Ann Charters
     NTP  expr   .. original text of of The special view of history.
```

This approach allows anyone in the future to create a new edition of "The Special View of History" with a new introduction which might also include notes and commentary. The notes and commentary would be encoded as Scholia entities (which are described elsewhere).

Finally, a manifestation for the 1970 Oyez edition in which the combined expression is a broader instance, and item entities are used to describe an individual copy of the book sitting on my desk as well as another copy in a library.

```
BTI   expr   The special view of history: original text
               with an introduction by Ann Charter.
     PT    man    . The special view of history / Charles Olson
                     with an introduction by Ann Charters.
                     - Berkeley, Oyez, 1970.
                     ISBN 0-520-04015-5.
     NTI   item   .. copy on my desk.
     NTI   item   .. Black Mountain College Library.
                      catalog number : PS3555.L850
```

This approach is a departure from the FRBR, but it resolves many of the problems which have been reported in trying to implement the model.

There is no question that our solution represents a significant number of additional records, but the flexibility of the approach justifies it.

§ 21 Agents Entity Groups

21.1 *Introduction*

Entities in the Agent group fall into two broad categories, people and corporate bodies.

21.2 *Entities*

person Any individual person or creature which is living, dead, mythical, legendary or fictional. Example: James Murray, chief editor of the Oxford English Dictionary; the fictional character James T. Kirk, Captain of the Star Ship Enterprise; the mythical giant called Maushop who figures in many legends in native tribes in New England; or even Hanno, Pope Leo X's White elephant who died in Vatican City in 1516.

corporate body Includes any specific, named, formal, informal, perceived or fictional group, company or organization. Example: The Free Software Foundation, the Irish Republican Army, The Walt Disney Corporation, The Kingdom of Thailand, The Republican Party, Hogworts School of Magic, Witchcraft and Wizardry (from the Harry Potter series of stories).

family Any group of people related to each other, usually by blood.

21.3 *Relationships*

A *person* may be a *member* of one or more *corporate body* or family.

```
BTP  corporate body
BTP  family
PT    . person
```

A family has two or more persons as narrower parts and if a family owns a business, a corporate body may be a narrower term.

```
              PT
              family
NTP  . person
NTP  . person
NTP  . corporate body
```

A *corporate body* may be a *member* of one or more other *corporate bodies* and one or more *persons* may be *members* of that *corporate body*.

```
BT corporate body
PT  . corporate body
NT .. corporate body
NT .. person
```

§ 22 Semantic & Lexical Entity Group

22.1 *Introduction*

Entities in the Semantic & Lexical Group are concerned with concepts (topics, subjects and ideas) and how they are labeled, identified and used. The entities in this group form the foundation for both category or subject trees but also for dictionaries and glosses.

22.2 *Entities*

concept Used for topics and subjects is the most abstract entity in BMF which describes a mutually exclusive idea.

symbol Used for specific words, phrases, characters, taxons and symbols (icons etc) which are used to represent concepts.

system Used for human languages, writing systems, encoding systems (xml), computer languages (C, Perl, Lisp etc). System entities may include any

number of concepts, symbols and forms. But concepts, symbols, forms and faces are not required to belong to any system.

22.3 *Relationships*

A symbol is always a narrower instance of a concept.

```
                PT
                concept
NTI   .  symbol
```

A system is made up of symbols which are narrower parts of a system, and based on a concept which is a broader instance of that system.

```
BTI concept
PT    . system
NTP  .. symbol
```

A symbol does not have to be a narrower term of a system, but usually is so the both of the following are correct:

```
BTI concept
PT   . symbol
```

```
BTI concept
BTP system
PT    . symbol
```

§ 23 The Temporal Entity Group

23.1 *Introduction*

Each of us lives in the present which is made up of the instance in which change occurs and the scope of that change in physical space. Modern man conceives of time as being a thread made up of events which is left behind as the present relentlessly extends it.

We think of the past as a snapshot of the world in a previous present but in fact the past is more of a portrait made up of artifacts and recollection from previous presents. The past does not exist, except as a mental image we have in the present which is changing from moment to moment as change occurs in the present.

The model used by the Temporal Entity Group takes this into account by establishing a single division of time which is used as a baseline overlayed by other calendar systems.

Specific instances in time are defined as relative to this baseline using the date entity which represents a specific duration of time.

The event entity describes an instance of a date in which something changes over the duration of that instance within specific spacial boundaries relative to an observer.

The other temporal entities in the group, periods and threads are different ways of describing specific changes and causative chains.

This allows any number of events to be attached to the same date. These different events may be concurrent local events which are separated geographically, but they also may be different descriptions of the same event from different perspectives within different contexts.

So the Crash of the Hindenburg, might have many different event entities which are narrower parts or instances of the crash. The news reel footage of the crash together with the reporters famous commentary is the event described from the perspective of an outside observer who was present during the event. Another event could be used to describe what happened from the perspective of a passenger who survived the crash. Still another event could be used to hypothesize the perspective of other passengers who did not survive, or an artificial perspective which pulls many different perspectives together.

A more mundane example would be a wedding, where different observers such as the father of the bride, the bride, a guest at the wedding and the caterer would all provide different perspectives of the same wedding within overlapping but not identical spacial boundaries.

In this way, there is no such thing as a single objective global description of an event, period or thread. Everything is the subjective description from the perspective of an observer anchored to a single shared baseline division of time.

23.2 *Entities*

date Just as a locus is a place without reference to time, A date is a time without reference to place, it has duration but no locus defining the scope of events being described.

Dates may be cyclical, like a holiday (every December 25th). And they may be open ended, but they are not infinite. Cyclical dates have a beginning and are assumed to have an end even if the end will take place at some unspecified time in the future.

period Age, era or named time.

A period is defined by both the duration and scope (locus) of what it is defining relative to an observer. A period is a pattern, which is observed in a series of threads and events which have something in common. The bronze Age is a period of time in a specific part of the world which was characterized by the use of bronze tools and weapons and the technological knowledge of working with this type of metal.

In many instances, events and threads can only be understood in context with the era or period of time that they take place in. This is true of geological periods or Eras, but also true for historical periods like *The Victorian Period*. Threads and events may be, but are not required to be linked to periods.

thread Grouped events which form a single complex event.

Threads are a collection of events which are grouped by theme or some other criteria. Threads can be nested into a hierarchy, but they can also be rhizomatic and cross other threads, branch off and rejoin at different times. A Thread can be as small as a group of messages in an Usenet News group or as large as World War II.

Threads, are defined by both the duration and scope (locus) of the collective events which are part of it, relative to an observer.

event Events are mutually exclusive actions or occurrences.

An event is defined by it's date, (including its duration), it's scope (location), a description of the changes which took place in the event and a description of the observer which the event was relative to.

Events can be as small as listing the date, location and agent who unlocked a door in a security log, or as big as a calamity like an earthquake. However, if an event, like a hurricane has detailed information associated with it — it should be classified as a thread which is broken down into discrete events. In many cases, little or nothing will be known about large events in the past which would be classified as a thread if more was known about it, so there is no hard and fast rule for when a event is bumped up to a thread level entity. Threads may be part of a period or other threads but this is not required..

23.3 *Relationships*

Periods, threads and events will always be narrower instances of a date.

```
                PT
                Date
NTI    . Period
NTI    . Thread
```

```
NTI   . Event
```

Threads and events will always be narrower parts of periods and instances of broader dates.

```
BTI   Date
PT    . Period
NTP   .. Thread
NTP   .. Event
```

Events will always be an instance of a broader date and may be broader parts of Periods or Threads.

```
BTI   Date
BTP   Period
BTP   Thread
PT    . Event
```

§ 24 Physical Entity Group

24.1 *Introduction*

BMF approaches man made objects in a similar way that bibliographic entities are described. In the object model, object entities are roughly equivalent to object entities, design entities are equivalent to expressions, models to manifestations and items to items.

In you understand the bibliographic group model then the physical entity group shouldn't be difficult.

```
          object
          electric guitars
design    . Gibson Stratocaster
model     .. Hendrix Stratocaster
item      .. Frank Zappa's Hendrix Strat.
```

The material entity has been added in order to describe materials which can be natural or man made.

```
material   air
material   water
material   snow
```

Materials may be agricultural products like rice, or soybeans.

```
          object
          rice (grain)
material   . Jasmine rice
```

The entity can be used alone, as a concept representing a natural material

```
          object
          ink
material   . Indian ink
model      .. Acme India Ink no.433
```

24.2 *Entities*

object A concept representing a class of material things created or fashioned or collected intentionally intervention, past or present, imagined, mythical or fictional. An object is a concept in the same way that a work is a concept.

Objects can be anything from a vehicles, buildings, devices, furniture. Nearly all objects are man-made, but not all. An ant-hill or bee-hive or even the pride and joy created by a dung beetle could all be classified as objects.

design A design is a specific expression of an object, including it's shape, parts, function etc. But it does not include color, options, period of manufacture etc.

> **model** A model is a manifestation of a design, including a defined range of colors it was made in and options which were offered with it etc.

> **material** A substance in which every part is the same as every other part. Materials can be natural (air or oak), refined or purified by man (aluminum, gasoline). Agricultural products (rice, wheat, barley) or a type of manufactured object like India ink.

> **item** This is the same entity used in the bibliographic group, and represents a specific physical instance of an object with specific options and colors.

24.3 *Relationships*

An object will always have designs as a narrower instance.

```
              PT
              object
NTI   . design
```

A design must have an object as a broader instance and a model as a narrower instance.

```
BTI object
PT    . design
NTI  .. model
```

A model must have a design as a broader instance and items as narrower instance.

```
BTI   design
PT    . model
NTI  .. item
```

Materials do not have objects as broader terms, but may have items, which include quantities as narrower terms.

```
BTG   object
PT    . material
NT   .. item
```

§ 25 Location Entity Groups

25.1 *Introduction*

> ...I will proceed with my history, telling the story as I go along of small cities no less than of great. For most of those which were great once are small to-day; and those which used to be small were great in my own time. Knowing therefore, that human prosperity never abides long in the same place, I shall pay attention to both alike.
>
> —Herodotus, *The Histories*

There are several excellent thesauri for place names including the Getty Thesaurus of Geographic Names and Project Alexandria which can be used for free online.

The Getty Thesaurus of Geographic Names used the following hierarchy for the city of Troy.

```
World (top term)
. Asia (continent)
.. Turkey (nation)
... Marmara (region)
.... Çanakkale (province)
..... Troy (deserted settlement)
```

This is a proper hierarchy for Troy (Hisarlik) today, as a deserted settlement in Turkey. But if you are looking up Troy in context of the Iliad, the hierarchy might be as follows:

```
World (top term)
. Asia (continent)
.. Asia Minor (region)
```

```
        ... Troad (legendary province)
        .... Troy (legendary inhabited settlement)
```

It makes sense to break places down into a concept representing the place and then define specific manifestations of that place during different periods in history.

So both the legendary Troy as well as the ruins that are thought to be Troy today might be represented as follows:

```
BTP  wld   World
BTP  loc   . Asia
PT   loc   .. Troy (legendary, historical and deserted settlement)
NTI  plc   ... Troy (Hellenic Troy; legendary settlement)
NTI  plc   ... Troy (Turkey; deserted settlement)
```

This is overly simplified — there are over 9 Troys starting in the Bronze Age and lasting till the end of the Roman Empire.

There are some very real advantages to this approach. References made to places in a work like William Bradford's "Of Plymouth Plantation" would point to locations which today are part of the United States of America.

By defining Plymouth as a locus with places including Plymouth (Pilgrim settlement), Plymouth (colony town) and then Plymouth (settlement in Massachusetts) references made by different people during different periods, could be understood in context with places, events people and content contemporary to that period.

At first glance the `world` entity might appear to be an odd choice. But it has proved to be invaluable for differentiating fictional worlds from the real world allowing us to describe whole fictional universes as systems.

```
BTP   wrld  middle earth (fictional past earth)
PT    loc   . Shire (fictional region)
```

Any stellar body can be described with the world entity.

```
BTP  wrld  milky way (galaxy)
PT   wrld  . sol (star)
NTP  wrld  .. mercury (planet)
NTP  wrld  .. venus (planet)
NTP  wrld  .. earth (planet)
NTP  wrld  .. mars (planet)
```

```
          PT   wrld  earth (planet)
NTP  wrld  . luna (moon)
```

```
          PT   cor   Federation of Planets (fictional government)
NTP  wrld  . Earth  (fictional future earth)
NTP  wrld  . Vulcan (fictional planet)
```

25.2 *Entities*

world	Either real worlds (eg. Earth, Mars etc), or fictional (eg. Middle Earth, Known Space etc) with a defined coordinate system, or at least potentially could have a coordinate system.
locus	A concept representing an inhabited place, building or physical feature. A place has physical coordinates.
place	A particular embodiment of a place during a specific period of time.
feature	Physical feature - river, lake, mountain.

25.3 *Relationships*
A locus or feature will always be narrower parts of a world.

```
                PT
                world
NTP   . locus
```

A locus will always be a broader part of a world, and a place and or feature as a narrower instance.

```
NTP   world
PT    . locus
NTI   .. place
NTI   .. feature
```

§ 26 Content Entity Group

26.1 *Introduction*

The content group is different from other entity groups we have discussed so far in that there is no hierarchical relationships between the different kinds of entities within the group.

The `div` (division) entity is very general, and could have been used for other kinds of content entities, but `scholia`, `messages` and `day` (day pages) will be used in such large numbers and be treated differently from other types of division entities that that have been given their own types.

The `scholia` and `glossa` entities are somewhat unique in BMF because they are the only entity types to be linked directly to inline elements within other Burrs rather than using the relationship and defined-by attributes used universally throughout BMF.

26.2 *Entities*

division	Used as a generic content container for marking up mutually exclusive divisions of content.
scholia	Used for commentary about another concept or block level text.
glossa	Used for glosses, notes and semantic markup of inline level text.
message	Typically a text message such as an email.
day	Day page — tasks & notes and lists.

26.3 *Relationships*

The division entity (div) is a container for holding master markup of texts and collections of media. At the moment the only place that it is expected to be used is as a narrower part of an expression entity.

```
                PT
                expression
NTP   . div
NTP   . div
NTP   . div
```

A scholia is an independent entity which may be part of an expression as a div is used, but it may also be used for commentary on other commentary.

```
                PT
                expression
NTP   . scholia
NTP   . scholia
```

```
BTP   expression
PT    . scholia
NTP   .. scholia
```

```
BTP   expression
PT    . glossa
```

A message or day page is typically associated with a concept which is used to define a group of emails, and or with a date. Scholia may be attached to messages as well.

```
BTP   concept (group name)
BT    date
PT    . message
NTP   .. scholia
```

```
BTP   concept (group name)
BT    date
PT    . day
NTP   . scholia
```

§ 27 Documentation Entity Group

27.1 *Introduction*

The entities in the Documentation group may be used to create documentation for any markup language, programing language, technical specification, reference manual or tutorial.

The `doc` (documentation) entity is similar to an expression entity. The body of text for each section or chapter uses division entities, and then reference entities are used to define elements, attributes, functions, enumerated values etc.

Documentation can then be converted into different formats for distribution (info, man, html, TeX, postscript etc) and be described using manifestation entities.

27.2 *Entities*

documentation   Root expression of a documentation document.

reference   Used for providing reference documentation for parts of a markup or programing language.

27.3 *Relationships*

A `doc` entity is a broader instance of a work, and then may have any number of narrower parts made up of div, element, attribute, enumerated, entity, section, function, class, variable and module entities.

```
BTI   work
PT    . doc
NTP   .. div
NTP   .. reference
```

§ 28 BMF Entity Group

28.1 *Introduction*

The Burr not only is the atomic unit in BMF, it is also used to describe larger structures.

In the early days of BMF development, it was assumed that special structures would be developed for pulling Burrs into larger structures, but when Emacs Burs was first written, Burrs were used to represent topicspaces and brambles.

It became obvious that the Burr structure was perfect for defining Brambles which are little more than lists of topicspaces and for topicspaces which are little more than lists of Burrs.

28.2 *Entities*

topicspace   Used to describe a BMF Topicspace

bramble   Used to describe a local BMF Bramble

stub   Used as a placeholder for a Burr which has not been written or is incomplete.

28.3 *Relationships*

A bramble is always the top term. Topicspaces are narrower parts of a bramble.

```
            TT
            bramble
NTP  . topicspace
NTP  . topicspace
```

```
            TT
            bramble
PT   . topicspace
```

§ 29 Sections

Burrs are divided into sections using the `<sec>` tag. Sections are structured as lists or notes.

Each type of entity uses a different combination of sections depending on it's purpose.

29.1 *Hierarchy section group*

At base, the hierarchy is a simple list of terms which are related to the concept the Burr is describing.

Some Burrs might have literally hundreds of terms, so in order to keep things manageable, the hierarchy section is broken into a number of different sections, but conceptually, they are treated as extensions of the hierarchy section.

The only element allowed in the any sections in the hierarchy section group is the item `<i>` element.

Items come in two different flavors, simple and bibliographic. This simple format looks like this:

```
<i r="BT" e="t" d="top:HGW6-7648" l="person"
   q="human being; living or dead" />
```

All items should include the following attributes:

r (relationship) indicating the relationship between the term and the concept represented by the Burr. Values must be from the Relationship Code (Thesaurus Code) list.

e (entity type) the type of entity used to encode the Burr described by the item. Values must be from the Entity-type Code list.

d (defined-by) points to a Burr which defines the term represented by the item. Values must be a resolvable BXID.

l (label) a string providing a label for the term described by the item. Typically this should use the preferred term defined by the Burr. The value provided by the label attribute

q (qualifier) a string providing a parenthetical qualifier which differentiates the term described by the item from other items.

er (extended relationship) provides a complex relationship type from an list of enumerated values. Different section types use different lists of enumerated values.

The bibliography and reference sections use a more complex approach which provides enough information to identify and understand the reference and then point to a Burr which defines it.

A bibliographic reference must serve two purposes, a) information which places the reference within the larger hierarchy of the Burr, and optionally, a more detailed bibliographic entry which:

- allow the title to be separate from the rest of the entry so that processing applications can easily turn the title into a hyperlink.
- include a separate URL link for online resources like Web pages.
- include a separate notes section for additional information and comments related to the information.

So an item will have the following structure:

```
<i r="relationship" e="entity-type" er="ext relationship"
   d="defined-by"  l="label" q="qualifier">
 <a>title proper</a> <b> [General material Designation] / responsibility.
  - city : publisher, date. </lb>
  URL: </b> <ref src="http://www.example.org">www.example.org</ref>
 <m>descriptive note</m>
</i>
```

The tags, <a> title proper, <b> rest of title and other information, and <m> descriptive note, are based on MARC tags.

Formating and punctuation should follow something along the lines of ISBD(G) rules. There are a number of subtle differences between ISBD and BMF as well as all sorts of record types which aren't covered by isbd at all, so BMF will have it's own full blown format guidelines which are documented as fully as ISBD.

For example, the GMD (General Material Designation) which is used to describe media types in ISBD is used to describe the form of the work or expression described.

29.2 *Sections*

29.2.1 Hierarchy

The hierarchy section is typically the first section a Burr and is required in every entity type. The hierarchy provides a hierarchical list of broader, narrower, and related terms.

```
<sec typ="hierarchy">
  <i r="TT"  e="t" d="aut:AAA0-0000" l="Authority"
     q="Librarium topicspace" />
  <i r="BT" e="t" d="top:HGW6-7648" l="person"
     q="human being; living or dead" />
  <i r="PT" e="p" d="aut:UJA7-6676" l="Dickens, Charles John Huffman"
     q="Eng. novelist, 1812-1870" />
</sec>
```

29.2.2 Terms

The terms section is used for equivalent terms which are used for the concept the Burr is describing. This includes alternate spellings, nicknames, translations,

Provides a list of terms used for or are equivalent to the preferred term for the concept the Burr is describing.

This information could easily be included in the hierarchy section, but a list of often obscure and seldom used equivalent terms was of more important for indexing engines and was more often than not a distraction in displays.

Because every term in the terms section, by definition is an equivalent term, entity-type and defined-by attributes are not required.

```
 <sec typ="terms">
  <i r="PT" l="Charles Dickens" q="Eng. novelist, 1812-1870" />
  <i r="UF" l="Charles John Huffam Dickens"
     q="full name for Charles Dickens"/>
  <i r="UF" l="Boz" q="psued. of Charles Dickens" />
  <i r="UF" l="Karol Dickens" q="used for Charles Dickens" />
  <i r="UF" l="C`arlz Dikensi" q="used for Charles Dickens" />
  <i r="UF" l="Charl'z Dikkens" q="used for Charles Dickens" />
  <i r="UF" l="Charlz Dikens" q="used for Charles Dickens" />
  <i r="UF" l="Charlz Dikkens" q="used for Charles Dickens" />
 </sec>
```

The terms section is optional and is not required if there are no equivalent terms. However, it is suggested that the section is used in every Burr so that it is clear there are no alternate terms that have been identified.

29.2.3 Related

The related section is only allowed in person and corporate body entities and is used to provide a list of human relationships including parents, children, friends, lovers etc.

The values for the extended-relationship attribute are taken from an enumerated list of values used for human and corporate relationships.

Defined-by attributes are not required in the related section, as it is impractical to require that records be created for every person, group or creature related to the person described by the Burr.

```
         <sec typ="related">
           <i r="NT" er="parent" l="John Dickens"
              q="Eng. 1785-1851" />
           <i r="NT" er="parent" l="Elizabeth Barrow"
              q="Scot. 1789-1863" />
           <i r="NT" er="sibling" l="Frances Elizabeth Fanny Dickens"
              q="Eng. 1810-48" />
           <i r="NT" er="sibling" l="Alfred Dickens"
              q="Eng. b. &amp; d. 1814" />
           <i r="NT" er="sibling" l="Letitia Mary Dickens"
              q="Eng. 1816-74" />
           <i r="NT" er="sibling" l="Harriet Ellen Dickens"
              q="Eng. b. &amp; d. 1819" />
           <i r="NT" er="sibling" l="Frederick William Dickens"
              q="Eng. 1820-68" />
           <i r="NT" er="sibling" l="Alfred Lamert Dickens"
              q="Eng. 1822-60" />
           <i r="NT" er="sibling" l="Augustus Dickens"
              q="Eng. 1827-68" />
           <i r="NT" er="spouse" l="Catherine Thompson Hogarth"
              q="Scot. 1815-79 married 1836" />
           <i r="NT" er="child" l="Charles Culliford Boz Dickens"
              q="Eng. 1837-96" />
           <i r="NT" er="child" l="Mary Angela Dickens"
              q="Eng. 1838-1896" />
           <i r="NT" er="child" l="Kate Macready Dickens"
              q="Eng. 1839-1929" />
           <i r="NT" er="child" l="Walter Landor Dickens"
              q="Eng. 1841-63" />
           <i r="NT" er="child" l="Francis Jeffrey Dickens"
              q="Eng. 1844-86" />
           <i r="NT" er="child" l="Alfred Tennyson Dickens"
              q="Eng. 1845-1912" />
           <i r="NT" er="child" l="Sydney Smith Haldimand Dickens"
              q="Eng. 1847-72" />
           <i r="NT" er="child" l="Henry Fielding Dickens"
              q="Eng. 1849-1933" />
           <i r="NT" er="child" l="Dora Annie Dickens"
              q="Eng. 1850-51" />
           <i r="NT" er="child" l="Edward Bulwer Lytoon Dickens"
              q="Eng. 1852-1902" />
           <i r="NT" er="lover" l="Maria Beadnell (1810-86" />
           <i r="NT" er="lover" l="Ellen Lawless Ternan"
              q="Eng. actress; 1839-1914" />
           <i r="NT" er="grandchild" d="aut:INX7-1885" l="Monica Dickens"
              q="Eng. novelist 1915-1992" />
           <i r="NT" er="pet" l="Grip"
              q="first pet raven" />
           <i r="NT" er="pet" l="Grip"
              q="second pet raven" />
           <i r="NT" er="pet" l="Sultan"
              q="pet dog; St.Bernard-bloodhound" />
           <i r="NT" er="pet" l="Timber"
              q="pet dog; white spaniel" />
           <i r="NT" er="pet" l="Turk"
              q="pet dog; mastiff" />
           <i r="NT" er="pet" l="Linda"
              q="pet dog; St.Bernard" />
           <i r="NT" er="pet" l="Don"
              q="pet dog; Newfoundland" />
           <i r="NT" er="pet" l="Bumble"
              q="pet dog; Newfoundland" />
           <i r="NT" er="pet" l="Mrs.Bouncer"
              q="pet dog; white Pomeranian 1859-74" />
           <i r="NT" er="pet" l="Williamina"
              q="pet cat" />
           <i r="NT" er="pet" l="Dick"
              q="pet canary" />
           <i r="NT" er="pet" l="Newman Noggs"
              q="pet pony" />
         </sec>
```

29.2.4 Bibliography

The bibliography section is used for a bibliographic list of works by the subject of the Burr, or related works about the concept of the Burr itself.

Unlike other sections in the hierarchy section group, a bibliography is more than simply a list of related terms. Bibliographies are more useful if they are organized by category, date etc. and should allow the inclusion of introductions.

At the same time, if there is no detailed bibliographic information available or the author only wants to include a placeholder for the item until a later time, a simpler form may be used.

```
<sec typ="bibliography">
 <div>
  <hd>Christmas books</hd>

  <p>In the brief preface to the collected <w>Christmas Books</w>,
  describes them as <qt>a whimsical kind of masque intended to awaken
  loving and forbearing thoughts.</qt></p>

  <p>The first of the series, <w>A Christmas Carol</w> was quickly
  written late in 1843 as a means of raising some quick money.  This
  is not to say that this was the only motivation.  <cit><qt>[T]he
  idea of Christmas as a season of good feeding and good feeling was
  congenial to all Dickens's best characteristics, though it may have
  slightly encouraged some of his weaknesses.</qt> <ref
  ptr="CHEAL">CHEAL</ref></cit>.</p>

  <p>The enormous popularity of <sc>The Carol</sc>, as it became known
  fueled calls for a sequel the following Christmas.  Dicken's obliged
  with <w>The Chimes</w> (1843) and the series continued with <w>The
  Cricket on the Hearth</w> (1845) and finally <w>The Haunted Man</w>
  (1848). Taken together these have become known as <sc>The Christmas
  Books</sc>.</p>

  <li>
   <i r="NT" e="e" er="author" d="bib:IUT4-2844" l="A Christmas Carol" q="1843">
    <a>A Christmas Carol</a> <b>[novella] with illustrations by John
    Leech. - London : Chapman & Hall, 1843.</b>
   </i>
   <i r="NT" e="e" er="author" d="bib:QGY1-3372" l="The Chimes" q="1844">
    <a>The Chimes</a> <b>; A Goblin Story, [novella] with
    illustrations by John Leech; Daniel Maclise; Richard Doyle;
    Clarkson Stanfield. - London : Chapman & Hall, 1844.</b>
   </i>
   <i r="NT" e="e" er="author" d="bib:COD4-0230" l="The Cricket on the Hearth
      q="novella, 1845" />
   <i r="NT" e="e" er="author" d="bib:NGC1-4171" l="The Battle of Life"
      q="novella, 1846" />
   <i r="NT" e="e" er="author" d="bib:MBP0-5042" l="The Haunted Man"
      q="novella, 1848" />
  </li>
 <div>
</sec>
```

29.2.5 TOC (Table of Contents)

The toc (table of contents) section is used to include links to the master markup of the parts of a document contained in division entities and to provide an electronic table of contents for an expression.

This should not be confused with the markup of a table of contents from an item which has already been published. These should be marked up and included in a division entity with the rest of the text.

Like the bibliography section, a toc section may be broken into sections with named headers and introductory material, or it may be a simple list of all of the parts that make up the composite expression.

```
<sec typ="toc">
  <i r="NTP" e="div" d="bib:XPX7-8381" l="Title Page" />
  <i r="NTP" e="div" d="bib:KLU8-0011" l="Incipit" />
  <i r="NTP" e="div" d="bib:KXC5-8788" l="Prayer" />
  <i r="NTP" e="div" d="bib:BGD4-0725" l="Preface" />
  <i r="NTP" e="div" d="bib:PLX1-3532" l="Contents" />
  <i r="NTP" e="div" d="bib:GGS1-0271" l="Chapter One" />
  <i r="NTP" e="div" d="bib:IWI0-1454" l="Chapter Two" />
</sec>
```

29.3 *Metadata Section Group*

At present, there is only a single section (meta) in this group. But it is expected that, like the hierarchy group, the meta section could be extended in a similar manner.

For example, the fields in a person Burr are presently designed to document public figures (both historical and living). But a person Burr for a business contact would include contact information, as well as purchase and billing information which might be more useful in a separate section.

A person Burr for yourself might include everything from medical information, personal information, educational, financial and family information. Some of this information you might want to selectively share with others and some would be strictly private. Breaking important metadata into different sections would make this easier.

29.3.1 Meta

The metadata section is an all purpose container for structured, named metadata elements. Each entity type defines which fields are required and allowed.

Only a single instance of an element is allowed in a section. Multiple values are included through the item <i> element.

Elements in the meta section are different from the rest of BMF in that long descriptive names are allowed.

As a general rule, all element and attribute names in BMF are 1-3 characters in length. It was felt that because only single instances of each element is allowed, that more descriptive names would make the section more readable.

```
<sec typ="meta">
  <entityType l="person" />
  <personalName sur="Dickens" giv="Charles" add="John Huffman"
            l="Dickens, Charles John Huffman"/>
  <dates>
   <i typ="birth" dt="1812-02-07"
      l="Born at Landport, near Portsmouth, England, 7 Feb., 1812;" />
   <i typ="death" dt="1870-06-09"
      l="died at Gadshill, near Rochester, England, 9 June, 1870;" />
   <i typ="burial" dt="1870-06-14"
      l="buried in Poet's Corner, Westminster Abbey, 14 June, 1870." />
  </dates>
  <roles>
   <i typ="preferred" l="novelist;" q="preferred" />
   <i l="journalist;" />
   <i l="editor." />
  </roles>
  <affiliation>
   <i typ="preferred" l="England" q="national, preferred" />
  </affiliation>
  <locus>
   <i l="England;" />
   <i l="Chatham;"/>
   <i d="geo:PDF8-7270" l="Portsmouth." />
  </locus>
  <gender typ="m" l="male" />
</sec>
```

29.4 *Notes Section Group*

Notes sections revolve around the idea of breaking records down into parts which allow you to start with the general and move in to progressively more detailed information. This is the embodiment of LOD (Level of Detail) which was discussed earlier.

29.4.1 Scope

The scope note is required in all types of entities, and is used to describe or even define the concept which a Burr is representing.

Scope notes are a short one to two paragraph prose description of the concept the Burr is describing. The only block level element allowed is the paragraph element. Divs, headers, lists, and lists are not allowed.

In most displays, the scope note is included with the data from the meta section, but there are any number of other purposes the scope note could be used for, including a descriptive passage in search results.

```
<sec typ="scope">
 <p>In full <pn>Charles John Huffam Dickens</pn> English
 novelist, generally considered the greatest of the Victorian
 era.  His many volumes include such works as <w>A Christmas
 Carol</w>, <w>David Copperfield</w>, <w>Bleak House</w>, <w>A
 Tale of Two Cities</w>, <w>Great Expectations</w>, and <w>Our
 Mutual Friend</w>.</p>
</sec>
```

29.4.2 Intro

The introduction section is longer and more detailed than the scope note and is used to provide a short encyclopedia style article describing the concept described by the Burr.

Divs, Headers, lists, tables are allowed in introduction sections, but if the text is long enough to be broken into more than one div it is probably too long to be in the intro section. The full article should be used in a macro section and then a shortened version of the article should be included in the intro section.

```
    <sec typ="intro">
      <hd>Charles Dickens Eng. novelist, 1812-1870</hd>

      <p>In full <pn>Charles John Huffam Dickens</pn> English
      <rl>novelist</rl>, generally considered the greatest of the
      Victorian era.</p>

      <p>He was the son of <pn>John Dickens</pn>, who served as a
      clerk in the navy pay-office and afterward became a newspaper
      reporter. Dickens' received an elementary education in private
      schools, served for a time as an attorney's clerk , and in
      1835 became reporter for the <ser>London Morning
      Chronicle</ser>. In 1833 he published in the <ser>Monthly
      Magazine</ser> his first story, entitled <w>A Dinner at Poplar
      Walk</w>, which proved to be the beginning of a series of
      papers printed collectively as <w>Sketches by Boz</w> in
      1836. He married <pn>Catherine</pn>, daughter of <pn>George
      Hogarth</pn>, in 1836. In 1836-37 he published the <w>Pickwick
      Papers</w>, by which his literary reputation was
      established. He became editor of <ser>Household Words</ser> in
      1849, and of <ser>All the Year Round</ser> in 1859, and
      visited America in 1842 and 1867-68.</p>

      <p>His chief works include <w>Pickwick Papers</w> (1837),
      <w>Oliver Twist</w> (1838), <w>Nicholas Nickleby</w>
      (1838-39), <w>Master Humphrey's Clock</w>, including <w>Old
      Curiosity Shop</w> and <w>Barnaby Rudge</w> (1840-41),
      <w>American Notes</w> (1842), <w>Christmas Carol</w> (1843),
      <w>Martin Chuzzlewit</w> (1843-44), <w>Chimes</w> (1844),
      <w>Cricket on the Hearth</w> (1845), <w>Dombey and Son</w>
      (1846-48), <w>David Copperfield</w> (1849-50), <w>Bleak
      House</w> (1852-53), <w>Hard Times</w> (1854), <w>Little
      Dorrit</w> (1855-57), <w>Tale of Two Cities</w> (1859),
      <w>Uncommercial Traveler</w> (1860), <w>Great Expectations</w>
      (1860-61), <w>Our Mutual Friend</w> (1864-65), <w>Mystery of
      Edwin Drood</w> (unfinished).</p>
    </sec>
```

29.4.3 macro

The macro note is based on Encyclopedia Britannica's concept of a micropedia and macropedia. Micropedia entries are equivalent to BMF scope notes. Intro section articles are more like the articles found in a single volume encyclopedia like the Columbia Encyclopedia. Macro notes are equivalent to macropedia articles which provide long, detailed referenced articles about the concept described by the Burr.

Macro articles are optional and not allowed in all types of entities. However, in order to include a macro article, you must also include an intro article. This is another instance of how LOD is used in BMF.

```
    <sec typ="macro">
      <div id="1">
        <hd>Childhood</hd>

        <p>Dickens was born in Portsmouth, England, to <pn>John
        Dickens</pn>, a naval pay clerk, and his wife <pn>Elizabeth
        Barrow</pn>. When he was five, the family moved to Chatham,
        Kent. When he was ten, the family relocated to Camden Town in
        London.</p>

        <p>His early years were an idyllic time for him. He described
        himself then as a <qt>very small and
        not-over-particularly-taken-care-of boy</qt>. He spent his
        time in the out-doors, reading voraciously with a particular
        fondness for the picaresque novels of Tobias Smollett and
        Henry Fielding. He talked later in life of his extremely
        strong memories of childhood and his continuing photographic
        memory of people and events that helped bring his fiction to
        life.</p>

        <p>His family was moderately well off and he received some
        education at a private school but all that changed when his
        father, after spending too much money entertaining and
        retaining his social position, was imprisoned for debt. At the
        age of twelve Dickens was deemed old enough to work and began
        working for 10 hours a day in <cor>Warren's boot-blacking
        factory</cor> located near the present Charing Cross railway
        station. He spent his time pasting labels on the jars of thick
        polish and earned six shillings a week. With this money he had
        to pay for his lodging and help support his family who were
        incarcerated in the nearby Marshalsea debtors' prison.</p>

        <p>After a few years his family's financial situation
        improved, partly due to money inherited from his father's
        family. His family was able to leave the Marshalsea but his
```

mother did not immediately remove him from the boot-blacking
factory which was owned by a relation of hers. Dickens never
forgave his mother for this and resentment of his situation
and the conditions working-class people lived under became
major themes of his works. Dickens wrote, <qt>No advice, no
counsel, no encouragement, no consolation, no support from
anyone that I can call to mind, so help me God!</qt></p>

<p>In May 1827 Dickens, began work as a law clerk, a junior
office position with potential to become a lawyer. He did not
like the law as a profession and after a short time as a court
stenographer he became a journalist, reporting parliamentary
debate and traveling Britain by stagecoach to cover election
campaigns. His journalism informed his first collection of
pieces <w>Sketches by Boz</w> and he continued to contribute
to and edit journals for much of his life. In his early
twenties he made a name for himself with his first novel,
<w>The Pickwick Papers</w>.</p>

<p>On April 2, 1836 he married <pn>Catherine Hogarth</pn>,
with whom he had ten children. In 1842 they traveled together
to the United States; the trip is described in the short
travelogue <w>American Notes</w> and is also the basis of some
of the episodes in <w>Martin Chuzzlewit</w>.</p>

<p>Dickens' writings were extremely popular in their day and
were read extensively. His popularity allowed him to buy
<pl>Gad's Hill Place</pl>, in 1856. This large house in
Rochester, Kent was very special to Dickens as he had walked
past it as a child and had dreamed of living in it. The area
was also the scene of some of the events of Shakespeare's
<w>Henry IV, part 1</w> and this literary connection pleased
Dickens.</p>
</div>
<div id="2">
<hd>Later life</hd>

<p>Dickens was a prolific writer who was almost always working
on a new installment for a story and rarely missed a
deadline.</p>

<p>Dickens separated from his wife in 1858. In Victorian times
divorce was almost unthinkable, particularly for someone as
famous as he was. He continued to maintain her in a house for
the next twenty years until she died. Although they were
initially happy together, Catherine did not seem to share quite
the same boundless energy for life which Dickens had. Her job
of looking after their ten children and the pressure of living
with and keeping house for a world famous novelist certainly
did not help. Catherine's sister Georgina moved in to help her
but there were rumors that Charles was romantically linked to
his sister-in-law. An indication of his marital
dissatisfaction was when in 1855 he went to meet his first love
<pn>Maria Beadnell</pn>. Maria was by this time married as well
but she seems to have fallen short of Dickens' romantic memory
of her.</p>

<p>On the 9th June, 1865 while returning from France to see
<pn>Ellen Ternan</pn>, Dickens was involved in the
<ev>Staplehurst train crash</ev> in which the first six
carriages of the train plunged off of a bridge that was being
repaired. The only first-class carriage to remain on the track
was the one Dickens was in. Dickens spent some time tending the
wounded and dying before rescuers arrived; before finally
leaving he remembered the unfinished manuscript for <w>Our
Mutual Friend</w> and he returned to his carriage to retrieve
it.</p>

<p>Dickens managed to avoid an appearance at the inquiry into
the crash, as it would have become known that he was traveling
that day with Ellen Ternan and her mother, which could have
caused a scandal. Ellen, an actress, had been Dickens'
companion since the break-up of his marriage and as he had met
her in 1857 she was most likely the ultimate reason for that
break-up. She continued to be his companion, and probably
mistress, until his death.</p>

<p>Although unharmed he never really recovered from the crash,
which is most evident in the fact that his normally prolific
writing shrank to completing <w>Our Mutual Friend</w> and starting
the unfinished <w>The Mystery of Edwin Drood</w>. Much of his time
was taken up with public readings from his best-loved
novels. The shows were incredibly popular and on December 2,
1867 Dickens gave his first public reading in the United States
at a New York City theatre. The effort and passion he put into
these readings with individual character voices is also thought
to have contributed to his death.</p>

<p>Exactly five years to the day after the Staplehurst crash,
on June 9, 1870, he died. Contrary to his wish to be buried in

Rochester Cathedral, he was buried in the <pl>Poets'
Corner</pl> of Westminster Abbey. The inscription on his tomb
reads: <qt>He was a sympathizer to the poor, the suffering, and
the oppressed; and by his death, one of England's greatest
writers is lost to the world.</qt></p>

<p>In the 1980s the historic Eastgate House in Rochester, Kent
was converted into a Charles Dickens museum, and an annual
Dickens Festival is held in the city. The Eastgate House was
closed in 2005 by Medway Council as an economy measure, but a
<qt>Dickens World</qt> theme park is scheduled to open in nearby
Chatham in 2007. The house in Portsmouth in which Dickens was
born has also been made into a museum.</p>
</div>
<div id="3">
<hd>Novels</hd>

<p>Dickens' writing style is florid and poetic, with a strong
comic touch. His satires of British aristocratic snobbery — he
calls one character the <sc>Noble Refrigerator</sc> — are wickedly
funny. Comparing orphans to stocks and shares, people to tug
boats, or dinner party guests to furniture are just some of
Dickens' flights of fancy which sum up situations better than
any simple description could.</p>

<p>The characters themselves are among some of the most
memorable in English literature. Certainly their names are. The
likes of Ebenezer Scrooge, Fagin, Mrs. Gamp, Micawber,
Pecksniff, Miss Havisham, Wackford Squeers and many others are
so well known they can easily be believed to be living a life
outside the novels, but their eccentricities do not overshadow
the stories. Some of these characters are grotesques; he loved
the style of 18th century gothic romance, though it had already
become a bit of a joke (see <pn>Jane Austen's</pn> <w>Northanger
Abbey</w> for a parodic example). One character most vividly
drawn throughout his novels is London itself. From the coaching
inns on the out-skirts of the city to the lower reaches of the
Thames, all aspects of the capital are described by someone who
truly loved London and spent many hours walking its streets.</p>

<p>Most of Dickens' major novels were first written in monthly
or weekly installments in journals such as <w>Master Humphrey's
Clock</w> and <w>Household Words</w>, later reprinted in book
form. These installments made the stories cheap and more
accessible and the series of cliff-hangers every month made each
new episode more widely anticipated. Part of Dickens? great
talent was to incorporate this episodic writing style but still
end up with a coherent novel at the end. The monthly numbers
were illustrated by, amongst others, <sc>Phiz</sc> (a pseudonym
for Hablot Browne).</p>

<p>Among his best-known works are <w>Great Expectations</w>,
<w>David Copperfield</w>, <w>The Pickwick Papers</w>, <w>Oliver
Twist</w>, <w>Nicholas Nickleby</w>, <w>A Tale of Two
Cities</w>, and <w>A Christmas Carol</w>. <w>David
Copperfield</w> is argued by some to be his best novel — it is
certainly his most autobiographical. Lesser known, <w>Little
Dorrit</w> is a masterpiece of acerbic satire masquerading as a
rags-to-riches story.</p>

<p>Dickens' novels were, among other things, works of social
commentary. He was a fierce critic of the poverty and social
stratification of Victorian society. Throughout his works,
Dickens retained an empathy for the common man and a skepticism
for the fine folk.</p>

<p>Dickens was fascinated by the theatre as an escape from the
world, and theatres and theatrical people appear in <w>Nicholas
Nickleby</w>. Dickens himself had a flourishing career as a
performer, reading scenes from his works. He traveled widely in
Britain and America on stage tours.</p>

<p>Much of Dickens' writing seems sentimental today, like the
death of Little Nell in <w>The Old Curiosity Shop</w>. Even
where the leading characters are sentimental, as in <w>Bleak
House</w>, the many other colorful characters and events, the
satire and subplots, reward the reader. Another criticism of his
writing is the unrealistic and unlikeliness of his plots. This
is true but much of the time he was not aiming for realism but
for entertainment and to recapture the picaresque and gothic
novels of his youth. When he did attempt realism his novels
were often unsuccessful and unpopular. The fact that his own
life story of happiness, then poverty, then an unexpected
inheritance, and finally international fame was unlikely shows
that unlikely stories are not necessarily unrealistic.</p>

<p>All authors incorporate autobiographical elements in their
fiction, but with Dickens this is very noticeable, particularly
as he took pains to cover up what he considered his shameful,
lowly past. The scenes from <w>Bleak House</w> of interminable

```
       court cases and legal arguments could only come from a
       journalist who has had to report them. Dickens' own family was
       sent to prison for poverty, a common theme in many of his books,
       in particular the Marshalsea in <w>Little Dorrit</w>. Little
       Nell in <w>The Old Curiosity Shop</w> is thought to represent
       Dickens' sister-in-law, Nicholas Nickleby's father and Wilkins
       Micawber are certainly Dickens' own father and the snobbish
       nature of Pip from <w>Great Expectations</w> is similar to the
       author himself.</p>
   </div>
   <div id="4">
     <hd>Legacy</hd>

     <p>Charles Dickens was a well known personality and his novels
     were immensely popular during his lifetime. His first full novel
     <w>The Pickwick Papers</w> brought him immediate fame and this fame
     continued right through his career. He maintained a high quality
     in all his writings and although never departing greatly from
     his typical <sc>Dickensian</sc> style he did experiment with different
     themes, moods and genres. Some of these experiments were more
     successful than others and the public's taste and appreciation
     of his various works have varied over time. He was usually keen
     to give his readers what they wanted and the monthly or weekly
     publication of his works in episodes meant that the books could
     change as the story proceeded at the whim of the public. A good
     example of this are the American episodes in *Martin Chuzzlewit*
     which were put in by Dickens in response to lower than normal
     sales of the earlier chapters. In <w>Our Mutual Friend</w> the
     inclusion of the character of Riah was a positive portrayal of a
     Jewish character after he was criticised for the depiction of
     Fagin in <w>Oliver Twist</w>.</p>

     <p>His popularity has waned little since his death and he is
     still one of the best known and most read of English authors. At
     least 180 movies and TV adaptations based on Dickens? works help
     confirm his success. Many of his works were adapted for the
     stage during his own lifetime and as early as 1913 a silent film
     of <w>The Pickwick Papers</w> was made. His characters were
     often so memorable that they took on a life of their own outside
     his books. Gamp became a slang expression for an umbrella from
     the character Mrs Gamp and Pickwickian, Pecksniffian and
     Gradgrind all entered the dictionary owing to Dickens' perfect
     portrayal of these kind of people. Sam Weller was an early
     superstar perhaps better known than his author at first and
     other characters have had their lives expanded upon by
     subsequent authors. It is likely that <w>A Christmas Carol</w>
     is his best known story with new adaptations almost every
     year. This simple morality tale with humour and pathos, for
     many, sums up the true meaning of Christmas and eclipses all his
     other Christmas stories.</p>

     <p>At a time when Britain was the major economic and political
     power of the world Dickens highlighted the life of the forgotten
     poor and disadvantaged at the heart of empire. Through his
     journalism he campaigned of specific issues such as sanitation
     and the workhouse but his fiction was probably all the more
     powerful in changing opinion. He revealed the harsh lives of the
     poor and satirized the people who allowed abuses to continue,
     all in the context of a good-humoured, entertaining story which
     sold widely.  His works seem to have inspired many more people
     to address problems and inequalities, even though he poked fun
     at these well meaning philanthropists, and his influence is
     often credited with having the Marshalsea and Fleet Prisons shut
     down.</p>

     <p>Dickens may have hoped for the foundation of a literary
     dynasty through his ten children and he named some of them after
     past writers but it would have been difficult for them to be
     anywhere near as successful as their father and some of them
     seem to have inherited their grandfather?s lack of financial
     acumen. Several of his children wrote of their memories of their
     father or prepared his surviving correspondence for publication
     but his great-granddaughter, <pn>Monica Dickens</pn>, would
     follow in his footsteps as a writer of novels.</p>

     <p>His works, with their vivid descriptions of life at the time,
     mean that the whole of Victorian society is often simply
     described as Dickensian.  Following his death in 1870 a greater
     degree of realism entered literature probably in reaction to
     Dickens' own tendency towards the picaresque and
     ridiculous. Late Victorian novelists such as <pn>Samuel
     Butler</pn>, <pn>Thomas Hardy</pn> and <pn>George Gissing</pn>
     all clearly owe much to Dickens but their works are usually much
     grittier and less sentimental.  Writers continue to be
     influenced by his books and although his many faults are
     criticized few other writers can match his blend of
     characterization, gripping plots, social commentary, popular,
     critical and financial success, and his sense of humour.</p>

     <p>Dickens enjoyed unparalleled world-wide popularity in his
```

```
        lifetime. This was as much due to his skills as a storyteller as
        the introduction of cheap mass printing and distribution to a
        literate audience which now included, for the first time in
        history, a large number of the working class; the first mass
        media market. Dickens dominated <t>Victorian fiction</t> in a
        similar way that <pn>Charlie Chaplain</pn> dominated silent
        pictures, becoming not so much a man as an institution and
        mythological figure surpassing anything seen since.</p>

        <p>Dickens' most popular works were his early novels <w>The
        Pickwick Papers</w>, <w>Oliver Twist</w>, <w>Martin
        Chuzzlewit</w>, <w>A Christmas Carol</w>, and <w>David
        Copperfield</w>. But contemporary critics did not approve of his
        later, darker and more symbolic works, and the loss of the freer
        comic spirit of his early work. <pn>F.R. Leavis</pn>, in 1948,
        summed up the general consensus, asserting that Dickens'
        <qt>genius was that of a great entertainer</qt>. Typecasting
        Dickens as merely being a popular author, <cit><qt>writing in
        the least disciplined of all literary genres in the most lawless
        literary milieu of the modern world</qt><ref>eb1911</ref></cit>
        explains, why little serious attention was given to his work for
        70 years after his death. The literary genre which came to be
        known as the <t>Victorian Novel</t> was summarily dismissed as
        popular entertainment, similar to the way <t>Science Fiction</t>
        was ignored in the 1950's and 60's.</p><p>But Dickens'
        significance had never been in question to the average reader to
        whom his work was <cit><qt>instinctively felt to be true,
        original and ennobling.</qt> The oft quoted exclamation of a
        costermonger's girl in 1870 says it all, <qt>Dickens dead?  Then
        will Father Christmas die too?</qt> <ref>eb1911</ref></cit>.</p>

        <p>The importance of Dickens' work continued to grow during the
        twentieth century through innumerable stage, television and film
        adaptations, transcending the world of Victorian England in
        which they were set and establishing many of his stories and
        characters as part of Western consciousness.</p>
      </div>
    </sec>
```

29.5 *Content Section Group*

29.5.1 Body

The body section includes the master markup of an encoded text of expressions in division Burrs. It is also used for notes in day pages, and the body of a message in message entities. The body section is not allowed in most other entity types.

```
        <sec "body">
          <div typ="verse">
           <ver ren="center">
            <ln>John Watts took</ln>
            <ln>salt - and shal-</ln>
            <ln>lops, from </ln>
            <ln>the <em ren="underline">Zouche Phoenix</em></ln>
            <ln>London's supplies</ln>
            <ln>10 Lb Island</ln>
           </ver>
          </div>
        </sec>
```

29.5.2 Gallery

The gallery section is used for the creating a collection of images, music or video.

Images and other binary media might be included with a Burr as a descriptive element, like the inclusion of a portrait of a person in a person Burr. Such images are not treated as distinct works and have no associated metadata except for a label and qualifier.

When images are treated as works, they will use bibliographic entities which will include detailed metadata. When an image like this is included in a gallery, a defined-by attribute is used instead of the src attribute. Processing applications are then required to download the Burr describing the image at the same time as the Burr containing the gallery.

The format for the gallery section has not been finalized and may well change, but it will likely look something like the following:

```
        <sec typ="gallery">
          <i r="NTP" src="ned-01.png" l="Uncle Ned""
             q="in front of his house" />
          <i r="NTP" src="ned-02.png" l="Uncle Ned"
```

```
        q="in the side of his house" />
    <i r="NTP" src="ned-03.png" l="Uncle Ned"
        q="in the back of his house" />
    <i r="NTP" d="mon:BET4-5250" l="The Spanish Inquisition" />
</sec>
```

29.5.3 Tasks

Task lists may be treated as extensions of the hierarchy, so rather than including them in a body section, they have been given their own type of section.

Task lists are similar to hierarchy lists but include two new attributes:

pri (priority) used to indicate the priority of an item.

sta (status) used to indicate the status of an item.

The format for the tasks section has not been finalized and may well change, but it will likely look something like the following:

```
<sec typ="tasks">
    <i r="NTP" pri="A" sta="p" l="call bank about car payment" />
    <i r="NTP" d="dpg:PPH4-8437"
        pri="A" sta="x" l="finish 2nd draft of paper" />
    <i r="NTP" d="bib:FPX2-2026"
        pri="B" sta="o" l="read Catch-22" />
</sec>
```

29.6 *Documentation Section Group*

29.6.1 Usage

The usage section is used for a proscriptive passage describing how the concept the Burr is describing should be used, including examples.

Allowed markup is the same as in the intro section.

29.6.2 Context

Used for additional cultural, linguistic or historical information which is assumed to be understood in the *scope*, *intro* or *macro* sections.

Allowed markup is the same as in the intro section.

29.6.3 Schema

The schema section is used to include schema definitions (DTD, XML Schema, RELAX-NG etc. The section is used in reference entities providing reference documentation for markup languages.

The schema section treats the contents literally, preserving whitespace and formating.

The idea is to incorporate concepts from literate programming into BMF, so that. Placing schema definitions in their own section makes it easier to generate a complete schema for a markup language by pulling out all of the schema sections in a topicspace and concatenating them together.

The format for the schema section has not been finalized, so it may change in the future.

```
<sec typ="schema">
# =========================================
# Person Entity
# -----------------------------------------

    ## Used for all named individuals; living, dead, fictional,
    ## legendary, mythical.  An individual can be a human, animal,
    ## alien, robot, god, ghost or spirit.
BMF.entity.person =
attribute typ { "person" }
& BMF.element.person.sec*

    ## section element for person entities.
BMF.element.person.sec = element sec {
(BMF.section.hierarchy?
| BMF.section.terms?
| BMF.section.related?
| BMF.section.person.meta?
| BMF.section.scope?
| BMF.section.intro?
| BMF.section.macro?
```

```
          | BMF.section.context?
          | BMF.section.usage?
          | BMF.section.bibliography?
          | BMF.section.reference?
          | BMF.section.identity?
          | BMF.section.history?)* }?
        </sec>
```

29.7 *Burr Metadata Section Group*

29.7.1 References

Used for a list of works and sources consulted in researching and creating the Burr. This section also serves the purpose of establishing a Burr as an authority record by citing where terms, dates and other information in the Burr had been originally found.

```
<sec typ="references">
  <i r="author-of" e="m" d="bib:GOW5-8744">
    <a>The Diamond Age</a> <b>[novel] / Neal Stephenson.
    New York : Bantam Books, 1995.</b></i>

  <i r="author-of" e="m" d="bib:UAA5-1783">
    <a>People of the talisman</a>
    <b>; The secret of Sinharat [novel] / Leigh Brackett.
    - New York : Ace Books, cop. 1964.</b></i>

  <i r="consulted" e="c" d="aut:GVE3-6267">
    <a>Slashdot.org</a>
    <b>; news for nerds, stuff that matters [web site] <lb />
    URL:<url src="http://slashdot.org">http://slashdot.org</url>.
    <lb />
    <m>An important news source for hi-tech news and discussion,
    for the technically literate, geek sub-culture.</m></i>
</sec>
```

29.7.2 Identity

The identity section is used for metadata describing the Burr, including the BXID, topicspaces ownership, permissions, version information and refresh periods.

```
<sec typ="identity">
  <descriptor>Charles Dickens</descriptor>
  <qualifier>Eng. Novelist, 1812-1870</qualifier>
  <exchangeID>aut:UJA7-6676</exchangeID>
  <topicspace>
    <i pfx="aut" url="http://chenla.org/bram/aut" />
    <i pfx="evn" url="http://chenla.org/bram/evn" />
  </topicspace>
  <replacedBy></replacedBy>
  <version>aut:UJA7-6676-4.xml</version>
  <previous>aut:UJA7-6676-3.xml</previous>
  <refresh>2005-09-16T12:37</refresh>
  <owner>brad@chenla.org&lt;Brad Collins&gt;</owner>
  <copyright>Public Domain</copyright>
  <created>2004-07-08T1732, brad@chenla.org</created>
</sec>
```

29.7.3 History

Used to provide a basic change-log of all changes made to the Burr.

```
<sec typ="history">
  <log>
    <stamp>PR3, 2005-09-16T10:33, Brad Collinsbrad@chenla.org</stamp>
    <comment>added type attributes to topicspace and changed
    refresh from period to date.</comment>
  </log>
  <log>
    <stamp>PR.2, 2005-07-19T15:02, brad@chenla.org</stamp>
    <comment>Massive overhaul of everything.  Starting proper
    version control with this version.</comment>
  </log>
  <log>
    <stamp>E1, 2004-07-08T1732, brad@chenla.org</stamp>
    <comment>Created Burr</comment>
  </log>
</sec>
```

§ 30 Larger Structures : Sticking Burrs Together

So far we have talked about the nuts and bolts of individual Burrs which was needed to understand how BMF pulls together Burrs into larger structures.

One of the most important design goals for BMF is for it to be a read-write system which allows people to keep what they create private, share it with select groups of people or make it available publicly.

The only practical way of doing this is if Burrs are downloaded and local copies are used. This is different from the model used by the World Wide Web which is based on centralized web servers, and browsers which only keep files that are viewed as long as is needed to view them.

Keeping local copies also makes it easier to create local organizational structures which pull together what you have found on online and what you have personally created.

Since BMF is distributed, we have borrowed many concepts from Version Control systems, especially ARCH.

ARCH is different from other version control systems like CVS and Subversion in that it doesn't have a central repository, everything is a branch.

It's not enough to just download copies of Burrs, in order to make this work we need unique identifiers and versioning of Burrs, so that it is easy for applications to check if there is a newer version of a file than the copy you have locally, and also to ensure that any changes you make to the file aren't confused with the original.

BMF achieves this through the use of unique identifiers called Burr Exchange Ids, topicspaces for keeping content from a single source unique and brambles which are local repositories for all Burrs you have created, downloaded or edited.

§ 31 Topicspaces

Each BXID is required to belong to a BMF topicspace. Topicspaces are unique URIs for a collection of BXIDs.

This is not much different from the World Wide Web. A Web Site is a collection of Web pages. The URL provides a unique id/address for all the Web pages on a Web site.

BMF Brambles are often accessed via Web Servers, but unlike the Web, BMF makes the assumption that you are making a local copy of a file which will remain on your computer after you have read it. Unless you explicitly make a local copy of a file, a browser only keeps the page in it's cache, which will automatically be deleted from your computer some time in the future.

This is an important distinction between BMF and the Web which makes it easy to create a broad range of features and services which are difficult or downright impossible to provide on the Web. These include making local notes and annotations and recombining content found on the Internet into new structures.

Topicspaces are similar to namespaces in XML which were introduced to solve a similar problem. How can you ensure that names of elements in XML languages are unique? If two languages that uses the element "date" you need some means of understanding which language is being used when you mix data together. This is done by declaring a namespace for each XML language. For example, let's say there is a language which was developed for Joe's Pizza Cafe. He could then use a prefix "pizza" to indicate all elements defined in his language. so that date, becomes pizza:date. Then when he sends his accounting data to his accountant who uses a different language with the namespace prefix "wong" (for Wong's Accounting Service) his date tag, which would be prefixed with his own namespace so that "wong:date" wouldn't get confused with "pizza:date".

A topicspace works the same way, except that where a namespace uniquely describes a collection of XML elements, a topicspace uniquely describes a collection of files (or more specifically, Burrs).

topicspace	topicspace element
i	item element
pfx	prefix attribute
src	url attribute

Topicspaces are declared in the |identity| section of a Burr. A typical declaration might look like this:

```
<topicspace>
  <i pfx="aut" url="http://chenla.org/bram/aut" />
  <i pfx="evn" url="file:///d/deerpig/work/evn" />
  <i pfx="bib" url="http://chenla.org/bram/bib" />
  <i pfx="geo" url="http://chenla.org/bram/geo" />
</topicspace>
```

Then, when you use a BXID anywhere in the Burr, you only need to use the topicspace's prefix as we saw the in the examples in the previous section:

```
aut:SBY7-8413
geo:CED2-8235
aut:CED2-8235
evn:CED2-8235
evn:GUL6-0282
```

The other difference between Topicspace and namespace declarations is that namespace declarations are not required to point to anything. In fact the standard only recommends that a namespace declaration point to some sort of definition or documentation.

BMF uses topicspace declarations to point to the master copy or mirror of the Bramble. This makes it possible for processing applications to check if a local copy of a Burr is the most recent, and that the copy has not been corrupted in any way.

31.1 *Reserved BXIDs*

We had originally used a default file-name for defining Brambles and topicspaces in the root directory for the bramble and in the directory for each topicspace.

But when developing applications we found it was far easier to treat the definitions for brambles and topicspaces just like any other Burr.

So we hit on the idea of reserving one BXID in each bramble and topicspace for use in defining the contents of brambles and topicspaces. The reserved id is:

```
XXX0-0000.
```

Each bramble has a directory called `root' which is reserved for holding the Root Burr for the Bramble.

So the path for the root Burr for a bramble will always be:

```
[BRAMBLE-ROOT]/bram/root/X/X/X/0/0/0/0/0/XXX0-0000.xml
```

And the path for the root Burr in the authority topicspace which uses the prefix `aut' would be:

```
[BRAMBLE-ROOT]/bram/aut/X/X/X/0/0/0/0/0/XXX0-0000.xml
```

This approach might seem strange at first, but this makes referencing the root Burr for a Bramble or Topicspace no different from any other Burr.

31.2 *Defining Topicspaces*

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet href="./xml.css" type="text/css"?>
<BURR typ="topicspace">
  <sec typ="hierarchy">
    <i r="TT" e="t" d="aaa:AAA0-0000" l="Bramble"
       q="Local Bramble" />
    <i r="PT"  e="t" d="aut:AAA0-0000" l="Authority"
       q="Librarium authority topicspace" />
    <i r="NT"  e="t" d="top:HGW6-7648" l="person"
       q="human being, living or dead" />
    <i r="NT"  e="t" d="top:BVI6-1681" l="person"
       q="fictional character" />
    <i r="NT"  e="t" d="top:RYT0-1638" l="person"
       q="mythical or legendary character" />
  </sec>
  <sec typ="terms">
```

```
      <i r="PT" typ="index" l="Authority" q="Librarium topicspace" />
      <i r="UF" typ="abbrev" l="aut" q="Authority; topicspace prefix" />
    </sec>
    <sec typ="meta">
      <entityType l="topicspace" />
      <topicName a="authority topicspace"
                 l="authority topicspace"/>
      <dates>
        <i typ= "created" dt="2005.04.15" pl=""
           l="created 15 April.2005 in Bangkok" />
      </dates>
      <responsibility l="Brad Collins" />
    </sec>
    <sec typ="scope">
     <p>Used for authority records for all specific persons and
     creatures (alive, dead, fictional or mythical).</p>
    </sec>
    <sec typ="usage">
     <p>This topicspace should not be used for records describing a
     species, only only specific instances of individuals.</p>

     <p>So the term <sc>Skunk</sc> should not be included in the
     authority topicspace, but <sc>Pepe Le Pui</sc> which is the name of
     a fictional, cartoon skunk character Looney Toons (a series of
     animations from Warner Brothers) should be included.</p>

     <p>The <sc>President of the United States</sc> is an office and
     title so shouldn't be included, but <sc>Richard Nixon</sc> who was
     President of the United States in the early 1970's should be
     included.</p>
    </sec>
    <sec typ="identity">
      <descriptor>Authority</descriptor>
      <qualifier>Librarium topicspace</qualifier>
      <exchangeID>aut:AAA0-0000</exchangeID>
      <topicspace>
        <i pfx="aut" url="http://chenla.org/bram/aut" />
      </topicspace>
      <replacedBy></replacedBy>
      <version>aut:AAA0-0000--D1.xml</version>
      <previous></previous>
      <refresh>7500000</refresh>
      <owner>brad@chenla.org&lt;Brad Collins&gt;</owner>
      <copyright>Public Domain</copyright>
      <dates>
        <i v="15.04.2005 13:11" rsp="brad@chenla.org" />
      </dates>
    </sec>
    <sec typ="history">
     <log>
       <stamp>D2, 2005-04-15T12:11, Brad Collins brad@chenla.org</stamp>
       <comment>created XML Burr from plan page.</comment>
     </log>
    </sec>
</BURR>
```

§ 32 Brambles

Burrs are collected into Brambles, so a Bramble is nothing more or less than a collection of Burrs which has been given a unique id called a topicspace.

A Bramble can be public, private or shared. Brambles hold copies of all Burrs that you read from other Brambles as well as your own personal collection of Burrs that you have created.

A Bramble can also be made publicly available like a Web Site so anyone can browse and download content from it.

32.1 *Bramble Directory Structure*

Brambles are typically kept in a directory called *bram*, called *bramble root* and are located either in a users home directory, *~/bram/* or in document root in a Web server or FTP site.

Inside bramble root, burrs are sorted into folders for each topicspace they are from. One folder for each topicspace. In this way it is easy to keep Burrs from different topicspaces separated from each other.

In each topicspace folder, Burrs are organized into directories based on their BXID. So, for example the BXID *aut:BKV4-6537* would be found in the the following directories:

By using a common directory structure for all Brambles, we refer to a Burr using little more than its BXID. So the BXID: aut:BKV4-6537 with a topicspace declaration: file:///home/user/bram/aut">file:///home/

user/bram/aut, for a private Bramble, or **http://example.com/bram/aut** for a public Bramble would respectively expand to:

```
~/bram/aut/B/K/V/4/6/5/3/7/BKV4-6537.xml
```

or

```
http://example.com/bram/aut/B/K/V/4/6/5/3/7/BKV4-6537.xml
```

32.2 *Defining Brambles*

Brambles are defined through a special Burr entity-type (more on entity-types below) called, as you might have guess, a "bramble".

A single Bramble Burr is used to include all the topicspaces in the local bramble.

```xml
<?xml version="1.0" encoding="utf-8"?>
<BURR typ="bramble">
  <sec typ="hierarchy">
   <i r="TT" e="t" d="root:XXX0-0000" l="Bramble"
      q="Local Bramble" />
   <i r="NT" e="t" d="dpg:XXX0-0000" l="Deerpig"
      q="Personal topicspace" />
   <i r="NT" e="t" d="aut:XXX0-0000" l="Authority"
      q="Librarium authorities topicspace" />
   <i r="NT" e="t" d="bib:XXX0-0000" l="Bibliographic"
      q="Librarium bibliographic topicspace" />
   <i r="NT" e="t" d="dic:XXX0-0000" l="Dictionary"
      q="Librarium dictionary topicspace" />
   <i r="NT" e="t" d="evn:XXX0-0000" l="Events"
      q="Librarium events topicspace" />
   <i r="NT" e="t" d="geo:XXX0-0000" l="Places"
      q="Librarium Geographic topicspace" />
   <i r="NT" e="t" d="top:XXX0-0000" l="Topics"
      q="Librarium Topic topicspace" />
   <i r="NT" e="t" d="bmf:XXX0-0000" l="Burr Metadata Framework"
      q="BMF documentation topicspace" />
  </sec>
  <sec typ="terms">
   <i r="PT" typ="index"  l="local bramble"
      q="local bramble on host bulma" />
  </sec>
  <sec typ="meta">
   <entityType l="topicspace" />
   <topicName a="Local Bramble on bulma"
              l="Local Bramble on bulma"/>
   <dates>
     <i typ="created" dt="2005.04.13" pl=""
        l="created on 13 April 2005 in Bangkok" />
   </dates>
   <responsibility d="" l="Brad Collins" />
  </sec>
  <sec typ="scope">
   <p>This is the local bramble for the user <sc>deerpig</sc> on the
   host <sc>bulma</sc>.</p>
  </sec>
  <sec typ="usage">
   <p>The local bramble collects together all Burrs kept by a single
   user/account or on a single machine or site. Burrs are sorted into
   directory trees with a separate directory tree for each
   topicspace.</p>

   <p>Topicspaces may be local and private, or local copies of public
   or shared topicspaces other places.</p>

   <p>The topicspace <sc>aaa</sc> is reserved as the default
   topicspace for defining local Brambles in the same way that
   <sc>localhost</sc> is the default hostname on Unix-like
   systems.</p>
  </sec>
  <sec typ="identity">
   <descriptor>Bramble</descriptor>
   <qualifier>local bramble</qualifier>
   <exchangeID>aaa:AAA0-0000</exchangeID>
   <topicspace>
     <i pfx="aut" url="http://chenla.org/bram/aut" />
     <i pfx="evn" url="http://chenla.org/bram/evn" />
     <i pfx="bib" url="http://chenla.org/bram/bib" />
     <i pfx="geo" url="http://chenla.org/bram/geo" />
   </topicspace>
   <replacedBy></replacedBy>
   <version>aaa:AAA0-0000--D1.xml</version>
   <previous></previous>
```

```
    <refresh>7500000</refresh>
    <owner>brad@chenla.org &lt;Brad Collins&gt;</owner>
    <copyright>Public Domain</copyright>
    <dates>
      <i r="created" v="2005.04.13"  rsp="brad@chenla.org" />
    </dates>
  </sec>
  <sec typ="history">
      <log>
        <stamp>D1, 2005-04-13, Brad Collins brad@chenla.org</stamp>
        <comment>created XML Burr.</comment>
      </log>
  </sec>
</BURR>
```

§ 33 BXID (Burr Exchange IDs) and Topicspaces

BMF is a distributed content system which encourages everyone to keep local copies of everything they read. This is the opposite to the World Wide Web which is designed so that only one copy of a file is kept on a Web Server which anyone is accessible from anywhere on the Internet.

Because the original file could be changed by the owner at any time, it is difficult to know if the copy of a file sent to you by a friend is the most recent version or where to get the new version.

To meet these needs, BMF has adopted a unique ID system which is called a Burr Exchange ID or BXID (pronounced Bix Eye Dee) which is tied to a Topicspaces.

Topicspaces provide a unique id for a collection of BXID's in a similar way that the URL for a Web Site provides a unique id for all of the pages on a web site.

The following sections give an overview of how BXIDs and Topicspaces work.

33.1 *The BXID: Easy on the Eye, Ear and Memory*

For BMF to work we need a system which is globally unique, but also is easy to remember and quickly and accurately written down in awkward situations, like on the street or in a crowded bar (ie. the five pint rule: if you write it down in a bar after drinking five pints and can still read it in the morning it's a good system).

In a recent paper on the same topic, Sheldon Brahms wrote,

> Most phone systems decades ago used a combination of letters and numbers in subscriber identification (phone numbers). It was common to see phone "exchanges" such as "Liberty 2" (LI 2) or Walnut 5 (WA 5). Instead of all numbers, a phone number would be expressed as WA 5-3491.

> This gave a sing-song, almost rhythm or lilt to the way a phone number was said. Often, advertisers used this phenomenon in jingles, which were also very popular in years past. Phone numbers were sung along with the rest of the lyric for client identification, and it lent itself very well to memory. Businesses would pay extra for phone numbers that rhymed or otherwise went well with this effect. [BRAHMS]

He recommended development of a similar system. Since we needed a practical solution for us to continue development of BMF we sat down and began trying out a number of ideas, most of which quickly became un-usable, in order to create enough possible IDs to build a global URI system.

Each BXID which is made up of three letters from the roman alphabet followed by a number then separated by a dash and then a four digit integer. GOL2-1023 could be remembered as Golum Ten Twenty-three. This produces over 175 million possible ids.

Here are a few examples of BXIDs:

```
    SBY7-8413
    CED2-8235
    GUL6-0282
    SKD3-0655
    QLV8-2404
```

In order for a BXID to be unique, every BXID must belong to a topicspace. A three letter prefix for a topicspace is prefixed to a BXID in the same way that namespace prefixes are used in XML namespaces.

Here are a few examples of BXIDs with topicspaces:

```
aut:SBY7-8413
geo:CED2-8235
aut:CED2-8235
evn:CED2-8235
evn:GUL6-0282
```

In the example above, note that the BXID CED2-8235 is used by the geo, aut and evn topicspaces. These are all valid universally unique ids because they belong to different topicspaces.

§ 34 Commentary

The separation of texts from the commentaries about them is a core concept in BMF. This was inevitable as soon as you make the decision to break up item level books and other media into multiple works where they occur.

So if you have a book with an intro by Donna, a story by Jill and textual notes by Jane, you can break the intro and story into two different works. But what about Jane's notes? It's difficult to treat them as a work as is defined by the bibliographic entity group because they can not stand on their own, they are woven together and through another expression.

If you treat this as a new work which is a combination of Jill's story and Jane's notes, then if another book uses Jill's story with notes by Bob you would have to create another work, even though the story is exactly the same.

BMF resolves this through the use of the Scholia entity. A Scholia is any type of external commentary, note, reference, analysis or gloss. Scholias can be thought of as a transparency overlaying the pages of a text which contain notes and commentary.

```
NTI work  Jill's story
PT   expr  . original text of Jill's story.
NTP sch   .. Jane's notes
NTP sch   .. Bob's notes
NTP sch   .. Nadia's class notes
```

This approach makes it possible for everyone to create commentary, Nadia who is a high school student who took notes on Jill's story for a class placing it on the same footing as the authoritative notes by Jane and Bob which had previously been published with Jill's story.

34.1 *Scholia*

Scholia can be external commentary, but marginal as well (literally in the margins). Anything in BMF can have Scholia attached to them, including other scholia — you can make comments about comments.

A Scholia must have a target (tar) attribute which points to a Burr and optionally the section, div and paragraph or line in the Burr using xpath notation.

A BMF scholia may only use the following sections — note, reference, identity, and history. The identity section may be automatically generated by an application.

```
<?xml version="1.0" encoding="utf-8"?>
<BURR typ="scholia" tar="aut:VTW0-0877:intro/">
  <sec typ="note">
    <hd>mega crap</hd>
    <p>I think this whole section is rubbish.</p>
  </sec>
  <sec typ="reference">
    <i ref="eb2004" d="bib:WOL0-3010" l="Enclopædia Britannica 2004." />
  </sec>
  <sec typ="identity">
    <descriptor>mega crap</descriptor>
    <exchangeID>myn:KLL2-5605</exchangeID>
    <topicspace>
      <i pfx="aut" typ="http://chenla.org/bram/aut" />
      <i pfx="myn" typ="http://example.org/bram/myn" />
    </topicspace>
    <created>2005-12-20T13:48, brad@chenla.org</created>
  </sec>
</BURR>
```

Notice that there are two topicspace definitions, one for the Scholia which has the prefix "myn" and the other for the Burr that the scholia is commenting on with the prefix "aut".

This approach allows comments from multiple authors, private, shared or public.

In this example there is no owner attached to the scholia so ownership is inherited from the topicspace it belongs to.

34.2 *Glossa*

Not all comments and annotations are block level elements. An annotation can be a translation, comment or explanatory text about a word, phrase, paragraph in a text, in other words a gloss (or perhaps we should use the Med. Latin *glossa*).

If this is all there was to a Glossa we could just use a Scholia, (which can serve as footnotes). But what if you want to include additional semantic markup to multiple parts of a passage? Or create running commentary along side a text.

Glossa are still very much in early development, but the problem raises a number of issues which go to the heart of a distributed shared read/write REPL.

The simplest approach would be to just create a local copy of a Burr and then mark it up however you wanted to. But this is a branching (as used in version control system) of the Burr rather than making commentary or glossing an existing text.

A Glossa can be thought of as an external layer of markup on top of another text.

I can only see two ways of accomplishing this. You can use Xpath to indicate all the text that you want to markup, and have a distinct path for each item linked to the changed text.

This is the approach we are using for Scholia, but if you are adding many small tags to a text this quickly becomes a nightmare both from a authoring and processing point of view.

The second approach is to duplicate the body of the text and then adding or changing inline markup.

So if you have the following marked up text:

```
<sec typ="scope">
  <p>In full <hi ren="italics">Charles John Huffam Dickens</hi>,
  English novelist, generally considered the greatest of the Victorian
  era.  His many volumes include such works as <hi ren="underline">A
  Christmas Carol</hi>, <hi ren="italics">David Copperfield</hi>, <hi
  ren="italics">Bleak House</hi>, <hi ren="italics">A Tale of Two
  Cities</hi>, <hi ren="italics">Great Expectations</hi>, and <hi
  ren="italics">Our Mutual Friend</hi>.</p>
  </sec>
[... the rest of the Burr ...]
</BURR>
```

First strip out the inline markup, leaving only block level markup, so you have the following:

```
<p>In full Charles John Huffam Dickens English novelist, generally
considered the greatest of the Victorian era.  His many volumes
include such works as A Christmas Carol, David Copperfield, Bleak
House, A Tale of Two Cities, Great Expectations, and Our Mutual
Friend.</p>
```

Then import the passage to a Glossa and add new markup:

```
<?xml version="1.0" encoding="utf-8"?>
<BURR typ="glossa" tar="aut:VTW0-0877:scope">
  <sec typ="scope">
    <p>In full <pn>Charles John Huffam Dickens</pn> English
    novelist, generally considered the greatest of the Victorian
    era.  His many volumes include such works as <w>A Christmas
    Carol</w>, <w>David Copperfield</w>, <w>Bleak House</w>, <w>A
    Tale of Two Cities</w>, <w>Great Expectations</w>, and <w>Our
    Mutual Friend</w>.</p>
  </sec>
[... the rest of the Burr ...]
</BURR>
```

The processing could be as simple as toggling between the original and the glossa, seeing both texts side by side, or could be as fancy as merging the two versions together into a single view (using an xml diff and merge utility).

For this to work, it will have to operate on whole sections, not just a single paragraph in a longer passage. It is also important that block level markup is retained, unless you are proposing complete, alternate versions of a section (say an earlier draft of a poem).

I tend to think this should be possible but not the best way of offering alternate passages of a text which could be better done by creating a different div Burr for each text.

Processing applications can include different options for creating Glossa. So you could strip all inline tags, or keep some or all.

34.3 *Glossa and textual corruption.*

A Glossa is not allowed to alter the original text. After a glossa has been created, a validator should be used which strips out the markup and runs a diff against the original text to insure that the text has not been altered. If it has, the text must be changed back to match the original before the glossa is validated.

34.4 *Scholia, Glossa and versions of Burrs.*

Arch treats everything as a branch; there is no difference between a local copy and the copy that everyone commits changes to. BMF is designed the same way.

So it's important that both Scholia and Glossa include version numbers of Burrs which they are commenting on. In this way, if the text in the Burr is changed, the commentary and markup can still be used, using a earlier version of the Burr

34.5 *Advantages of Scholia and Glossa*

One of the most powerful things about this approach is it's flexibility. Scholia could be added to the bottom of a section as threaded comments from multiple people. They could be added as private footnotes to a text you are studying for a paper, or they could be sent to someone else as comments as part of an editing or review process.

Multiple sets of glossa could be used to publish alternate views of a text, so that you could publish an edition of Darwin which had commentary and glossa from both a scientific point of view as well as a religious, creationist rant against it.

The reader could toggle between the different commentaries as well as see them side by side or merge them into a combined chaotic flamewar.

This approach would make a project like Wikipedia far easier, in that each article could use a base text which is used by multiple parties to add their own comments and commentary which reflect their point of view and provide an alternate to directly editing the text.

34.6 *X-Path for Scholia*

BMF uses X-Path to indicate block level items which Scholia are attached to [XPATH].

For example, to indicate the third paragraph (numbering starts with 0) in the intro section of a Burr we could use an X-Path like the following:

```
//sec[@typ='intro']/p[2]
```

Since the structure of Burrs are very regular, it might be better to provide simplified notation:

```
//intro/p[2]
```

But what about items like a item in a related section?

```
//related/i[3]
```

or more specifically

```
//related/i[@l='Alfred Dickens']
```

The second is better in that it would work even if the list were reordered, but it would be difficult to automate because we would have to establish rules for each type of section.

So here is a complete example of what a scholia Burr type about part of the macro note in the Charles Dickens Burr:

```
<BURR typ="scholia" ptr="aut:UJA7-6676//macro/div[3]/p[1]">
  <sec typ="hierarchy">
    <i r="TT" e="t" d="dpg:AAA0-0000"
       l="Deerpig" q="topicspace" />
    <i r="BT" e="d" d="dpg:GSO5-2743"
       l="2005-01-31" q="day page" />
    <i r="BT" e="p" d="aut:UJA7-6676"
       l="Dickens, Charles John Huffman" q="Eng. novelist, 1812-1870" />
    <i r="PT" e="s" d="dpg:BGL5-3748"
       l="Robert Louis Stevenson on Dickens's Christmas Books" q="scholia" />
  </sec>
  <sec typ="body">
    <cit>
     <quo>
      <p>I wonder if you have ever read Dicken's <sc>Christmas
      Book's</sc>... they are too much perhaps.  I have only read two
      yet, but I have cried my eyes out, and had a terrible fight not to
      sob.  But, oh, dear God, they are <em>good</em> -- and I feel so
      good after them -- I shall do good and lose no time -- I want to
      go out and comfort someone -- I <em>shall</em> give money.  Oh,
      what a jolly thing it is for man to have written books like these
      and just filled people's hearts with pity.</p>
     </quo>
     <bib>
      <pn>Robert Louis Stevenson</pn> to an unidentified
      correspondent.<ref>1</ref>
     </bib>
    </cit>
  </sec>
  <sec typ="reference">
    <i ref="1">Quoted in the <ser>Dickensian</ser>, vol. 16, 1920, p.200.</i>
  </sec>
  <sec typ="identity">
    <exchangeID>aut:UJA7-6676</exchangeID>
    <topicspace>
      <i pfx="aut" typ="local" />
    </topicspace>
    <owner>brad@chenla.org&lt;Brad Collins&gt;</owner>
    <created>2006-01-31T17:03, brad@chenla.org</created>
  </sec>
</BURR>
```

A potential problem with this is that the `ptr` address points to a Burr but not a specific version number of a Burr. If the Burr changes, the Scholia could become a dead link.

Comments may be lost and not carried forward if the text is changed.

So the spec will make it mandatory that the Burr won't be broken if a scholia's pointer doesn't match with say, the paragraph, it should first check for previous versions of the Burr for a match. If previous versions are not available, then it should match with the div, if not the div, it should match with the section, and if not with the section than with the Burr. Only then would the scholia be considered a dead link.

Scholia should always point to the most general block level element possible. It's the Benjamin Franklin approach to linking. Don't link to a line or a paragraph if you can point to a div, don't point to a div if you can point to a section, and don't point to a section if you can point to a Burr.

§ 35 Documentation features in BMF

35.1 *Documenting BMF*

The concept of documentation in BMF is somewhat different from other documentation systems which people are used to.

First of all, documentation on the BMF framework and markup language is integrated with BMF, it is interwoven within the framework itself.

In most languages, documentation for the language is provided in a separate file format like Texinfo, HTML or Docbook. Help files are kept in yet another file format and accessed through separate help applications or browsers.

Emacs is the exception. Emacs lisp (Elisp) encourages the inclusion of detailed documentation as part of the source code itself. These document strings in Lisp are not simply comments which are ignored, they provide invaluable context based documentation.

This is similar to the idea of literate programming. Literate programs combine source and documentation in a single file. Literate programming tools then parse the file to produce either readable documentation or compilable source. For compiled languages like C, this approach is about as close as you'll be able to get to what is available in Lisp.

BMF documentation is treated like any other content, so accessing the documentation for the language is just part of whatever other content that is available locally or over a network.

This also means that enumerated values used in BMF are defined in BMF in relation to Burrs defining concepts behind each value.

```
BTG con   parent (father or mother)
BT  sym   . parent (Eng. father or mother)
PT  enum  .. parent
UF+ con   .. father
OR  con   .. mother
```

So when you look up the extended relationship value for parent it leads you to definitions of the broader generic term for the concept for parent, the broader term which is the English word for parent, and indicates that father AND mother can be used for the term parent.

In this way, BMF documentation is interwoven with all content which is encoded in BMF.

35.2 *Self documentation for everything else*

BMF extends the concept of self documentation to any type of content encoded in a Burr. There is no reason why there shouldn't be documentation for all types of content. At first this might sound a bit strange, but any type of information could benefit from documentation.

A dictionary is documentation for human languages providing definitions, pronunciation and usage. Historical dictionaries also provide some context. Encyclopedias can also be seen in some contexts as documentation for persons, places, events and concepts.

The *context* and *usage* sections are provided exactly for this purpose.

One of the design goals for BMF was for it to be suitable for encoding libraries designed to last hundreds or even thousands of years.

But one of the biggest problems with information is that once you loose the context that it was written in, along with unspoken assumptions of facts and usage shared by everyone at the time of writing you have lost the ability to understand what an author was trying to communicate.

Adding documentation about etiquette, popular culture, slang, superstitions, urban legends related to a concept or topic would be very helpful to people from other cultures and countries, as well as for future generations.

§ 36 Conclusion

> More generally, programs that mediate between the user and the rest of the universe notoriously attract features. This includes not just editors but Web browsers, mail and newsgroup readers, and other communications programs. All tend to evolve in accordance with the Law of Software Envelopment, aka Zawinski's Law: "Every program attempts to expand until it can read mail. Those programs which cannot so expand are replaced by ones which can".
>
> Jamie Zawinski, inventor of the Law (and one of the principal authors of the Netscape and Mozilla Web browsers), maintains more generally that all really useful programs tend to turn into Swiss Army knives.
>
> —Eric Raymond, *The Art of Unix Programming* [RAYMOND]

Because of time and space constraints this paper was not able to cover a number of important features in BMF including:

- The SBFD (Standard Burr Format Definition) which is based on the ISBD group of bibliographic formating standards, which defines eye-readable formating of Burr entity records which do not depend on any markup or encoding.

  The SBFD definitions for each entity will be completely integrated with the formal definitions for each Entity type.

- The networking, exchange, version control and access control models for BMF which are still in the process of being formally defined.

- The extensibility of BMF, allowing anyone to extend the language locally while still preserving overall compatibility with any other implementations.

- An open API for BMF allowing anyone to integrate BMF encoded content into their applications and for BMF content to be interoperable with other encoding systems including Dublin Core Metadata, TEI encoded documents, Docbook, HTML, and MARC catalog records.

- Integration of messaging into BMF data structures and an generalized messaging model which pulls together all forms of text based messages and correspondence into a single model.

- An integrated Lisp syntax for encoding Burrs which is integrated with processing code, to make Burrs self indexing, updating, clustering and searchable in the same way that Skribe documents are able to convert themselves into HTML.

- The design features pertaining to BMF being used as a first generation markup language for the Long Now Foundation's proposed 10,000 Library both as a long term electronic format, and a format for preservation on physical archival media like Norsam's High Density Rosetta which provides microscopic analog storage of information and images on nickel plates that last for thousands of years. Disks.

So what is our last word on BMF? BMF is a large complex system which is still in it's very early days. There will be many people who will dismiss BMF as being too big and too complex to be widely adopted. But, like TEI (another very large markup language) there is no reason why smaller and more simple subsets of BMF can be created for more general purposes (BMF-Lite anybody). So the complexity and features will be there if you need them.

Providing a framework for electronic libraries is not a trivial problem. Mankind's collective memory and experience, which used to be locked in countless millions of paper tomes in brick and mortar libraries will gradually be digitized and placed on the Internet where they will be woven into the daily fabric of our lives. But the need to collect, preserve and organize that information will not be replaced by search services, no matter how good they become. BMF is an attempt to meet those needs well into our new century.

Notes

1. Wikipedia (**http://en.wikipedia.org**).

2. As their name implies Distributed Proofreaders (**http://pgdp.net**), is a group of thousands of volunteers who proof read books online for Project Gutenberg (**http://gutenberg.org**) a project which has been publishing free electronic editions of books which are in the public domain.

3. Del.icio.us (**http://del.icio.us**) who was recently bought by Yahoo provides a service for people to share their bookmarks with other people on the Internet.

4. Flickr (**http://flickr.com**) is another Yahoo aquisition which allows people to upload images and share them as galleries with people online.

5. Technorati (**http://technorati.com**) is an index and search engine for Blogs.

6. *Functional Requirements for Bibliographic records*: Final Report / IFLA Study Group on the Functional Requirements for Bibliographic Records, International Federation of Library Associations and Institutions. München:K.G.Saur, 1998. **http://www.ifla.org/VII/s13/frbr/frbr.htm** and **http://www.ifla.org/VII/s13/frbr/frbr.pdf**

7. *Guidelines for the construction, format, and management of monolingual thesauri* / developed by the National Information Standards Organization — (National information standards series, ISSN 1041-5653; ANSI/NISO Z39.19-1993). **http://www.niso.org/standards/resources/Z39-19.pdf**

8. Slashdot (**http://slashdot.org**) a tech news and information site for Geeks.

9. It's worth mentioning Skribe, which is written in Scheme (a Lisp langugage) which integrates data and code structures into a single language.[SERRANO]

10. It's expected that the second release of BMF will include a broad integration between text and code. Unfortunately, this paper is already far too long and this discussion will have to wait for another time.

11. Gustav Davidson, A Dictionary of Angels, The Free Press, 1967. page 212.

12. This example used Wiki Markup used by emacs muse-mode. Other Wikis like Wikipedia use different a different syntax.

13. Arch seems to attract a lot of Emacs and Lisp projects, but then Lispers tend to walk a slightly different path. A good intro to Arch can be found at **http://regexps.srparish.net/tutorial-tla/arch.html**

Acknowledgements

I would like to thank Ruben Seja who was the first to truly believe in the potential of BMF and has done so much to help keep me alive on the opposite side of the planet while I've been developing BMF.

Bibliography

[BERNERS-LEE] T. Berners-Lee, R. Fielding, L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax and Semantics", 05 Nov 1997. **http://www.ics.uci.edu/pub/ietf/uri/draft-fielding-uri-syntax-00.txt**

[BRAHMS] Sheldon Brahms, "The U.S. Phone Nomenclature System as Applied to Dynamic Knowledge Repositories", July 9, 2001. **http://www.hastingsreaseach.com/net/05-nomenclature.shtml**

[CORNELI] Joseph Corneli, Aaron Krowne, "A Scholia-based Document Model for Commons-based Peer Production" August 4, 2005 (draft)

[DELEUZE] Gilles Deleuze, ed. and intro by Constantin Boundas, The Deleuze Reader, "Rhizome Versus Tree", Columbia University Press, 1993.

[DOCTOROW] Doctorow, Cory, "Metacrap: Putting the torch to seven straw-men of the meta-utopia", The Well, 2001. **http://www.well.com/~doctorow/metacrap.htm**

[ECO] Umberto Eco, Semiotics and the philosophy of language Bloomington, Indiana University Press, 1984.

[ENGELBART] Douglas C. Engelbart, "Knowledge-Domain Interoperability and an Open Hyperdocument System." Proceedings of the Conference on Computer-Supported Cooperative Work, Los Angeles, CA, October 7-10, 1990, pp. 143-156 (AUGMENT,132082,).

[FRAR] IFLA UBCIM Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), "Functional Requirements for Authority Records Draft", International Federation of Library Associations and Institutions, 2005-06-15.

[FRBR] IFLA Study Group on the Functional Requirements for Bibliographic Records, "Functional Requirements for Bibliographic records: Final Report", International Federation of Library Associations and Institutions, 1998. **http://www.ifla.org/VII/s13/frbr/frbr.htm** and **http://www.ifla.org/VII/s13/frbr/frbr.pdf**

[GUHA] R.V. Guha; Tim Bray, Meta Content Framework Using XML - World Wide Web Consortium, 1997. **http://www.w3.org/TR/NOTE-MCF-XML-970624** NOTE: Submitted to W3C 6 June 97.

[HAMMING] Richard Hamming, "You and Your Research", Talk at Bellcore, 7 March 1986. **http://www.paulgraham.com/hamming.html**

[KINDEL] Charlie Kindel, "The uuid: URI scheme", Nov, 24 1997. **http://www.ics.uci.edu/pub/ietf/uri/draft-kindel-uuid-uri-00.txt**

[RAYMOND] Eric Steven Raymond, The Art of Unix Programming, 2003.

[SERRANO] Manuel Serrano, Erick Gallesio, "This is Scribe!" http://www-sop.inria.fr/mimosa/fp/Scribe/doc/scribe.html

[TEI5] C.M. Sperberg-McQueen, Lou Burnard, TEI P5 Guidelines for Electronic Text Encoding and Interchange (revised). The Association for Computers and the Humanities, 2005.

[UPDIKE] Updike, John, "The End of Authorship" The New York Times, June 25, 2006. http://www.nytimes.com/2006/06/25/books/review/25updike.html

[VINAVER] Eugene Vinaver ed., The Works of Thomas Malory. Second edition. Oxford, Oxford University Press, 1971.

[WIKIPEDIA-FOLK] Wikipedia, "Folksonomy". Accessed 2005. http://en.wikipedia.org/wiki/Folksonomy

[WIKIPEDIA-LISP] Wikipedia, "Lisp Programing Language". Accessed 2005. http://en.wikipedia.org/wiki/Lisp_programing_Language

[XPATH] Anders Berglund, Scott Boag, Don Chamberlain, etal, "XML Path Language (XPath) 2.0", World Wide Web Consortium, 2003. http://www.w3.org/TR/xpath20/

[Z39.19] National Information Standards Organization,"Guidelines for the construction, format, and management of monolingual thesauri" (National information standards series, ISSN 1041-5653; ANSI/NISO Z39.19-1993). http://www.niso.org/standards/resources/Z39-19.pdf

The Author

Brad Collins
Founder
Chenla Laboratories
Ban Kwao
Chaiyapoon
Thailand
brad@chenla.org

Brad Collins, a native New Englander, has lived and worked in the Far East since 1989. In 1992 he co-founded one of the first ISPs in Hong Kong and was a pioneer in early Web development. He moved to Japan in 1997, and then to Thailand a year later where he established a networked multi-media consultancy. He spent a year in Beijing developing network and Web based content for a major video on demand network project. Other work includes video and computer animation projects including a gig as the animation director for an MTV Canto-pop music video. He has spent most of the last four years working on BMF development and a startup which is deep in stealth mode.