



Content Management in an Unmanageable World

Bill Kasdorf

Vice President, Apex Content Solutions

General Editor, *The Columbia Guide to Digital Publishing*

Metadata Subgroup Lead, EPUB 3.0 & 3.0.1 WG

**“Start with the end in mind,”
aka “First, a word from our sponsor” . . .**

**EPUB 3 is useful for way more
than just putting books on e-readers.**

**This session will show you
how fundamentally EPUB 3 helps you . . .**



... get from this ...



... to this.



Thanks to Jake Zarnegar of Silverchair for the graphics.

It's all about content management.

**Which means managing
your **content**
not just
your **publications.****

Six Things “Content Management” Might Mean



Six Things “Content Management” Might Mean

1. **WEB CONTENT MANAGEMENT**

**It’s about
managing what’s on your website.**

Six Things “Content Management” Might Mean

2. **DIGITAL ASSET MANAGEMENT**

**It's about
managing all the stuff.**

Six Things “Content Management” Might Mean

3. **WORKFLOW MANAGEMENT**

**It's about
who does what, when.**

Six Things “Content Management” Might Mean

4. XML WORKFLOW MANAGEMENT

XML First?
XML Early?
XML Middle?

XML When it Works for You.

(The earlier the better.

Which is what we’re working for here.)

Six Things “Content Management” Might Mean

5. XML REPOSITORY MANAGEMENT

*Do you build the content of your repository
from your products/publications?*

*Or do you build the products/publications
from the content of your repository?*

You can do both.

Six Things “Content Management” Might Mean

- 1. WEB CONTENT MANAGEMENT**
- 2. DIGITAL ASSET MANAGEMENT**
- 3. WORKFLOW MANAGEMENT**
- 4. XML WORKFLOW MANAGEMENT**
- 5. XML REPOSITORY MANAGEMENT**
- 6. ALL OF THE ABOVE**

Three Things “Content Management” Might Mean but Doesn’t, Usually

Three Things “Content Management” Might Mean but Doesn’t, Usually

1. **TITLE MANAGEMENT**

It’s all about the metadata.

Acquisition and development

Scheduling and monitoring the workflow

Cost estimating, budgeting, invoice processing

Load balancing and vendor management

Marketing and sales

Three Things “Content Management” Might Mean but Doesn’t, Usually

1. TITLE MANAGEMENT

2. **DIGITAL ASSET DISTRIBUTION**

It’s about the products you send your customers.

EPUBs, apps, PDFs, POD

Variants for Kindle, Nook, Kobo, etc.

*Getting metadata to the retailers
& aggregators*

Three Things “Content Management” Might Mean but Doesn’t, Usually

1. TITLE MANAGEMENT

2. DIGITAL ASSET DISTRIBUTION

3. WEB HOSTING

*It’s about making your content
available online.*

You need to do all these things.



You need to do all these things.

**You're probably
already doing them.**



**But they're usually done in silos.
And the silos are getting a bit shaky.**

Three Fundamental Aspects of Content Management

1. METADATA

2. CONTENT MARKUP

3. COMPONENT MANAGEMENT

*Making these three aspects work in concert
is what results in effective
content management in today's ecosystem.*

COMPONENT MANAGEMENT

The more granular, the better?

A red tracked wood chipper is shown in a wooded area. The machine is positioned on a gravelly ground, and a large volume of wood chips is being ejected from its discharge chute, creating a wide, fan-shaped spray. The background consists of bare trees and a dense forest. The machine has a prominent red body, a black engine compartment, and a large black discharge chute. A black metal frame is visible behind the machine.

**Don't get carried away.
There's work required
to manage all the bits.**

THOUGHTFUL GRANULARITY

What are the pieces you need to:

SELL?

RE-ARRANGE?

REUSE?

RENDER?

FIND?

THOUGHTFUL GRANULARITY

What are the pieces you need to:

SELL?
RE-ARRANGE?
REUSE?

RENDER?

FIND?

**Typically
maintained
as separate
components.**

THOUGHTFUL GRANULARITY

What are the pieces you need to:

SELL?

RE-ARRANGE?

REUSE?

RENDER?

FIND?

**Best
managed via
metadata
and markup.**

Typically
maintained
as separate
components.

**Good metadata and markup
are key to content management.**

**Formal, standards-based schemas
and clear, thorough specifications
are invaluable.**

*[Hint: there's a really useful one.
It's called EPUB 3.]*

METADATA

IDENTIFIERS

*They have different purposes;
use them all, and use them properly.*

SUBJECT CODES

*PanThema: Rich international standard,
20+ national groups, 15+ languages;
beta draft expected in mid-April at LBF.*

ONIX 3.0

*DO IT! Big advance over ONIX 2.1;
2.1 will only be supported through 2014.*

CONTENT MARKUP

STRUCTURE

What are the pieces, and how do they relate?

SEMANTICS

What are the pieces for, what are they about?

RESOURCES

Images, multimedia, scripts, stylesheets, etc.

ASSOCIATIONS

Links, references, annotations, indexes, etc.

**The Lingua Franca of the Web
is HTML.**

But it's no longer just for the web.

**Consider UPFRONT XHTML
as the foundation for
content management.**

Why are people skeptical of Upfront XHTML?

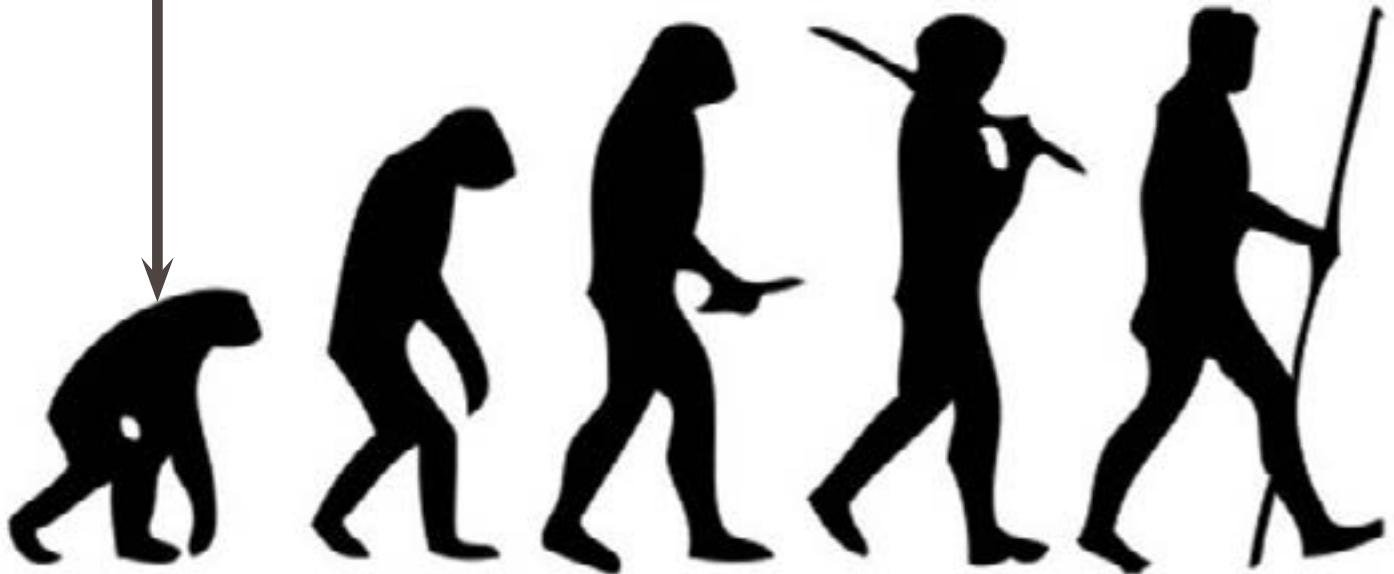
**The tangled, troubled
history of XHTML.**

**The big sea change
on presentation and semantics.**

THE EVOLUTION OF HTML

HTML1:

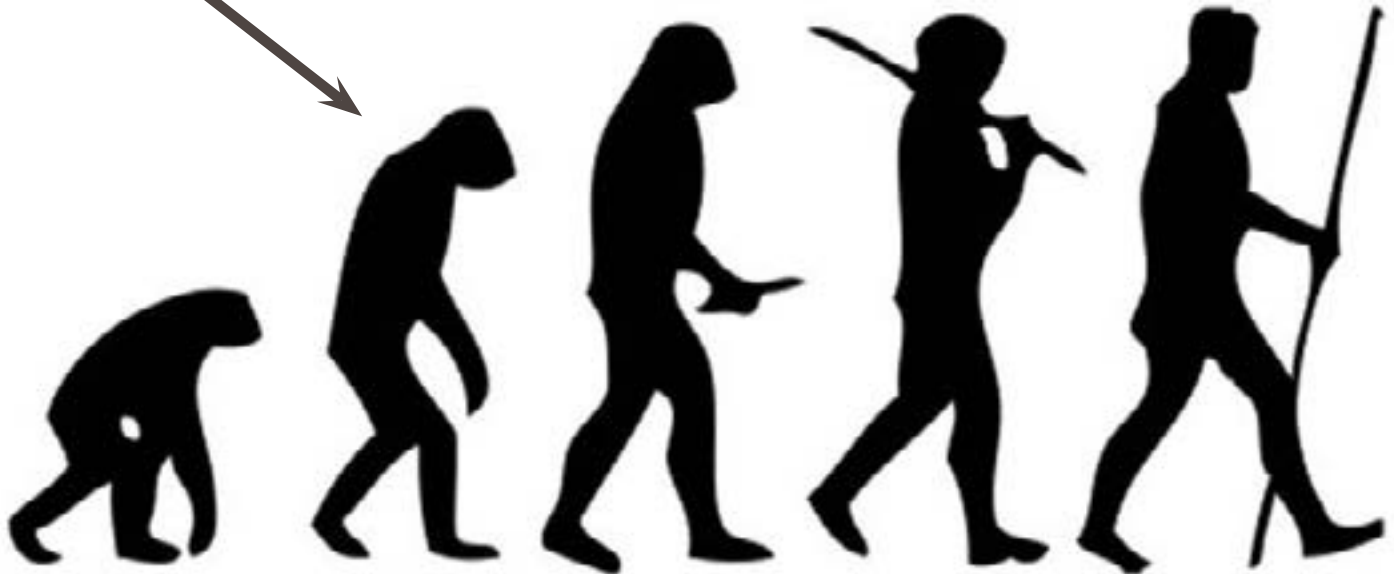
“Look, Ma,
I made a web page!”



THE EVOLUTION OF HTML

HTML2:

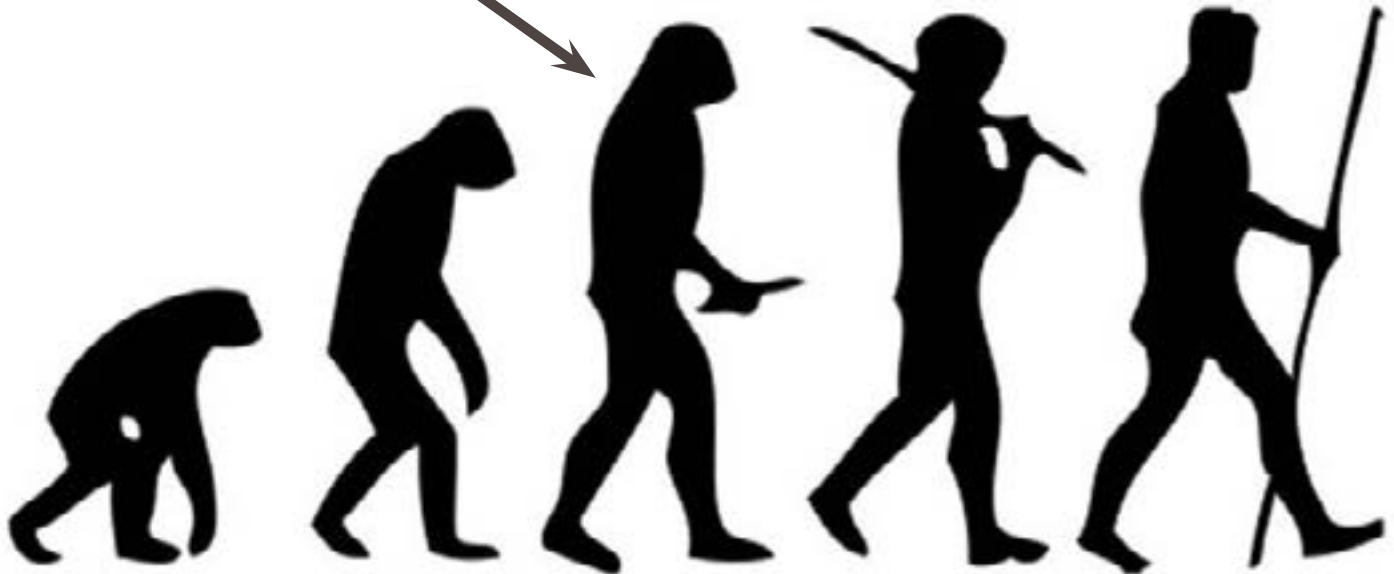
“Somebody actually PAID me to make a web page!”



THE EVOLUTION OF HTML

HTML3:

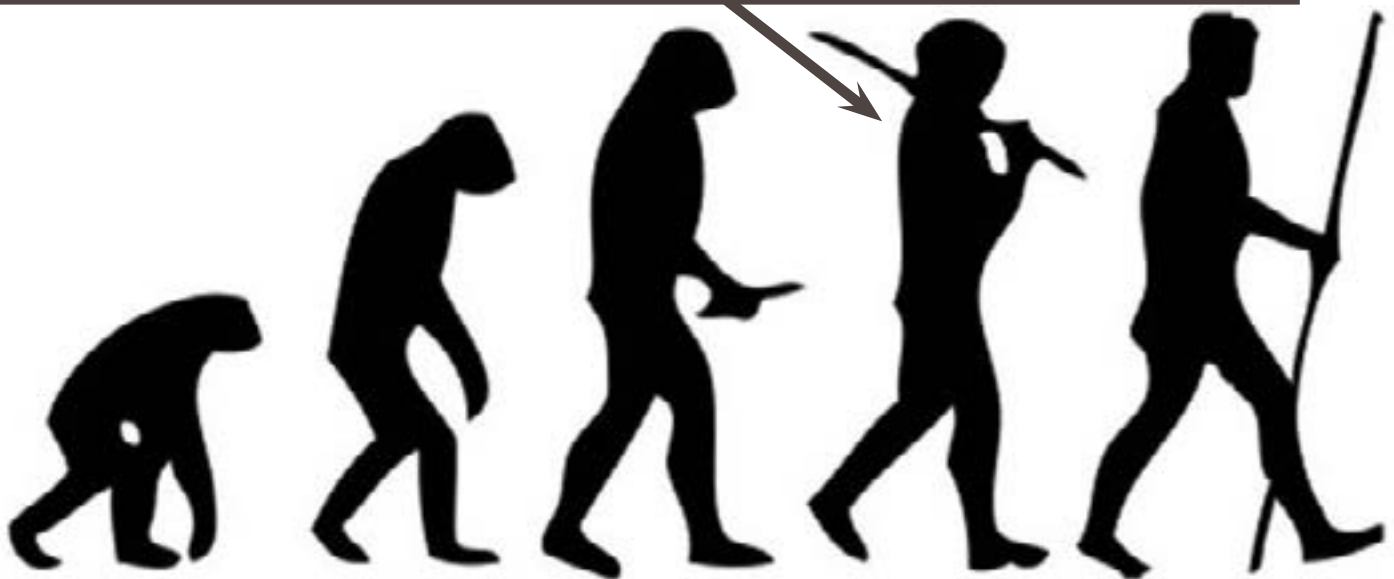
Early signs of
primitive tool use.



THE EVOLUTION OF HTML

HTML4:

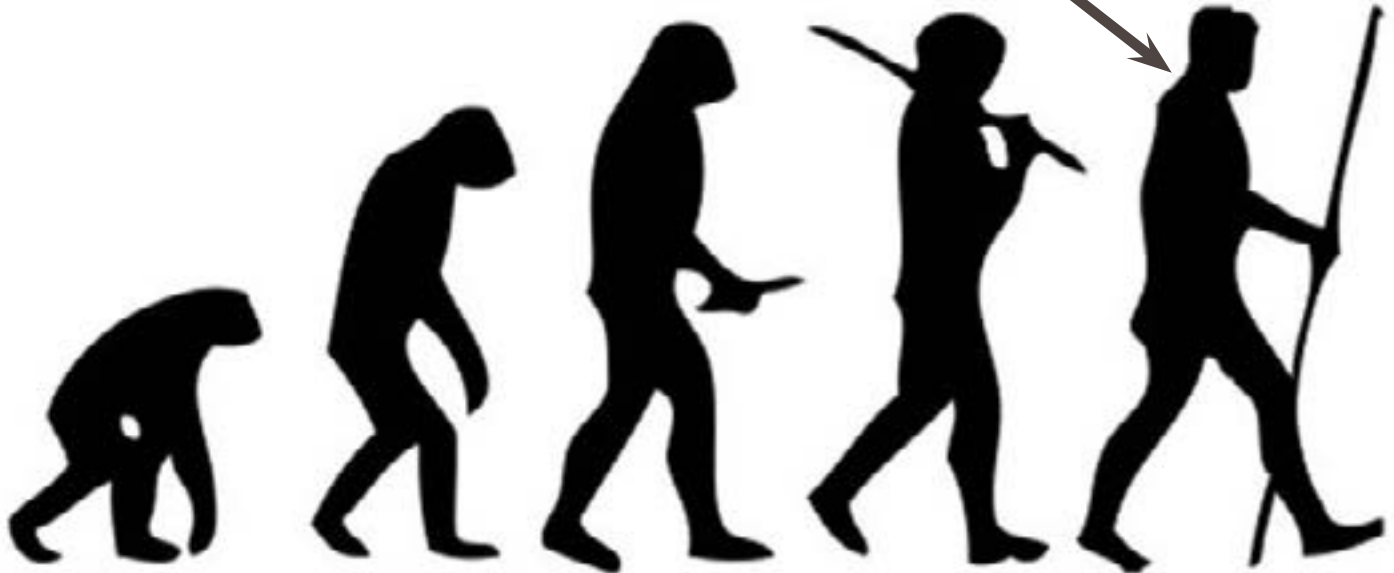
**Don't use presentational markup.
Unless you want to. [It's "deprecated."]**



THE EVOLUTION OF HTML

HTML5:

No presentational markup. Use CSS. Period.



**Okay, HTML isn't necessarily always XML.
(More on that in a moment.)**

**But no matter what XML model
you use for workflow,
repository, interchange**

**(e.g., DocBook, DITA, TEI,
NLM/JATS/BITS)**

**... you still always need
HTML.
Always.**

**But you still always need
HTML.**

Always.

Online? HTML.

**But you still always need
HTML.**

Always.

**Online? HTML.
EPUB? HTML.**

**But you still always need
HTML.**

Always.

Online? HTML.

EPUB? HTML.

Apps? HTML.

XHTML IS XML.



XML.





Models Made with XML.

**XHTML is a Model
Made with XML.**

XHTML IS XML.

**XHTML is HTML that
follows XML rules.**

XHTML IS XML.

**XHTML is HTML that
follows XML rules.**

XHTML IS XML.

It can be simple or complex.

It can be rigorously structured.

It can be richly semantic.

It can contain tons of metadata.

It can be HTML5.

EPUB 3 is XHTML5
with clear, specified semantics,
rich, flexible metadata,
and organized, documented components.

**XHTML5 is the framework
for the content documents.**

**EPUB 3 adds a framework for semantics,
metadata, and component documentation.**

**Within that framework,
granular markup for structure and
semantics enables you to
manage your content effectively
from creation to distribution.**

Basing your content management on EPUB 3 / XHTML5 enables:

Web-based authoring and editing.

Web-based proofing.

**Same basic XHTML for
print, online, EPUB, apps, interchange,
archiving, repurposing.**

EPUB 3 is useful for way more than just putting books on e-readers.

It's a well-thought-out, widely adopted specification based on Open Web standards that will make all aspects of managing your content easier and better.

Thanks!

Bill Kasdorf

bkasdorf@apexcovantage.com

+1 734 904 6252

@BillKasdorf

