



# Discrete Entities

*How should they be defined, named,  
and described in educational content?*

---

**Bill Kasdorf**

VP and Principal Consultant, Apex Content Solutions



**AKA chunks.**

# Proposed Goals of This Initiative\*

---

**Define or adopt a vocabulary  
(terms and clear definitions)  
for all appropriate sizes  
of discrete entities.**

*\*Thanks to Nick Brown of Ingram*

# Proposed Goals of This Initiative\*

---

**Define or adopt a vocabulary  
(terms and clear definitions)  
for all appropriate sizes  
of discrete entities.**

**Define or adopt any metadata required  
to manage those discrete entities.**

*\*Thanks to Nick Brown of Ingram*

# Proposed Goals of This Initiative\*

---

**Define or adopt a vocabulary  
(terms and clear definitions)  
for all appropriate sizes  
of discrete entities.**

**Define or adopt any metadata required  
to manage those discrete entities.**

**Define any needs beyond  
the EPUB 3.0.1 spec to properly package,  
manage, and utilize discrete entities.**

*\*Thanks to Nick Brown of Ingram*

# Proposed Goals of This Initiative\*

*Are there any intended benefits beyond the general utility of a **shared vocabulary** across the industry?*

*E.g., “facilitate redistribution of constituent discrete entities contained within an **EPUB**?” or “enable reading systems to **directly address content chunks** at a more granular level than each content document”?*

*\*Thanks to Nick Brown of Ingram*

# Discrete Entity Issues Discussed So Far

---

**How big (or small) can they be?**

# Discrete Entity Issues Discussed So Far

---

How big (or small) can they be?

Can they nest? Overlap? One or many files?

# Discrete Entity Issues Discussed So Far

---

**How big (or small) can they be?**

**Can they nest? Overlap? One or many files?**

**Do they have to be able to exist  
outside their original context?**

# Discrete Entity Issues Discussed So Far

---

**How big (or small) can they be?**

**Can they nest? Overlap? One or many files?**

**Do they have to be able to exist  
outside their original context?**

**What are they called?**

# Discrete Entity Issues Discussed So Far

---

**How big (or small) can they be?**

**Can they nest? Overlap? One or many files?**

**Do they have to be able to exist  
outside their original context?**

**What are they called?**

**Do they need IDs? If so, what kind?**

# Discrete Entity Issues Discussed So Far

---

How big (or small) can they be?

Can they nest? Overlap? One or many files?

Do they have to be able to exist outside their original context?

What are they called?

Do they need IDs? If so, what kind?

**Should they be required to be accessible?**

# Discrete Entity Issues Discussed So Far

---

**How big (or small) can they be?**

**Can they nest? Overlap? One or many files?**

**Do they have to be able to exist outside their original context?**

**What are they called?**

**Do they need IDs? If so, what kind?**

**Should they be required to be accessible?**

**What metadata should be associated with them?**

**How big  
(or how small)  
can they  
be?**



*[ Arnold  
Stang, the  
“Whatta  
Chunka  
Chocolate!”  
guy ]*

**How big  
(or how small)  
can they  
be?**

**How small is  
too small?  
Paragraph?  
Phrase? Word?**



*[ Arnold  
Stang, the  
“Whatta  
Chunka  
Chocolate!”  
guy ]*

**How big  
(or how small)  
can they  
be?**

**How small is  
too small?  
Paragraph?  
Phrase? Word?**

**How big is  
too big?  
Chapter?  
Book?  
Course?**



*[ Arnold  
Stang, the  
“Whatta  
Chunka  
Chocolate!”  
guy ]*

# Elsevier's Cisco ILSG Learning Hierarchy

---

**Curriculum:** Collection of learning events traversed by *tracks*

**Course:** Collection of *modules* or *lessons*

**Module:** Group of *lessons* or Reusable Learning Objects (*RLOs*) that support a single module objective

**Section:** A grouping of *lessons*

**Lesson:** Coll. of *topics* or *RIOs* (Reusable Information Objects)

**Topic:** Collection of *blocks* and *assets* (*RIO*)

**Block:** Collection of *assets* that present an idea (*subtopic*)

**Asset:** Smallest standalone *object*

# Observation from Yonah Hirschman:

---

**From a content perspective,  
publishers will focus on:**

**Learning Object**

**Module**

**Section**

**Lesson**

**Topic**

**Block**

**Asset**

# Pearson's Proposed EDUPUB Hierarchy:

---

**Product:** Highest level, required in product.xhtml;  
uses @class="product\_book" vs. @class="product\_course"

**Volume:** Optional value for multi-volume books;  
may be sold separately or as a group

**Part:** Optional level for grouping Chapters

**Chapter:** A collection of modules

**Module:** Collection of topics, assignments, and  
assessment items

**Card:** The content visible to a student on a screen/page  
at a given time during a lesson

**Can they nest?  
Overlap?  
One or many  
files?**

**Is this one  
Discrete  
Entity?  
Or four?**



**Can they nest?  
Overlap?  
One or many  
files?**

**Is this one  
Discrete  
Entity?  
Or four?  
Or two?  
Or all of  
the above?**



**Can they nest?  
Overlap?  
One or many  
files?**

**Is this one  
Discrete  
Entity?  
Or four?  
Or two?  
Or all of  
the above?**

**Text**

**Image**

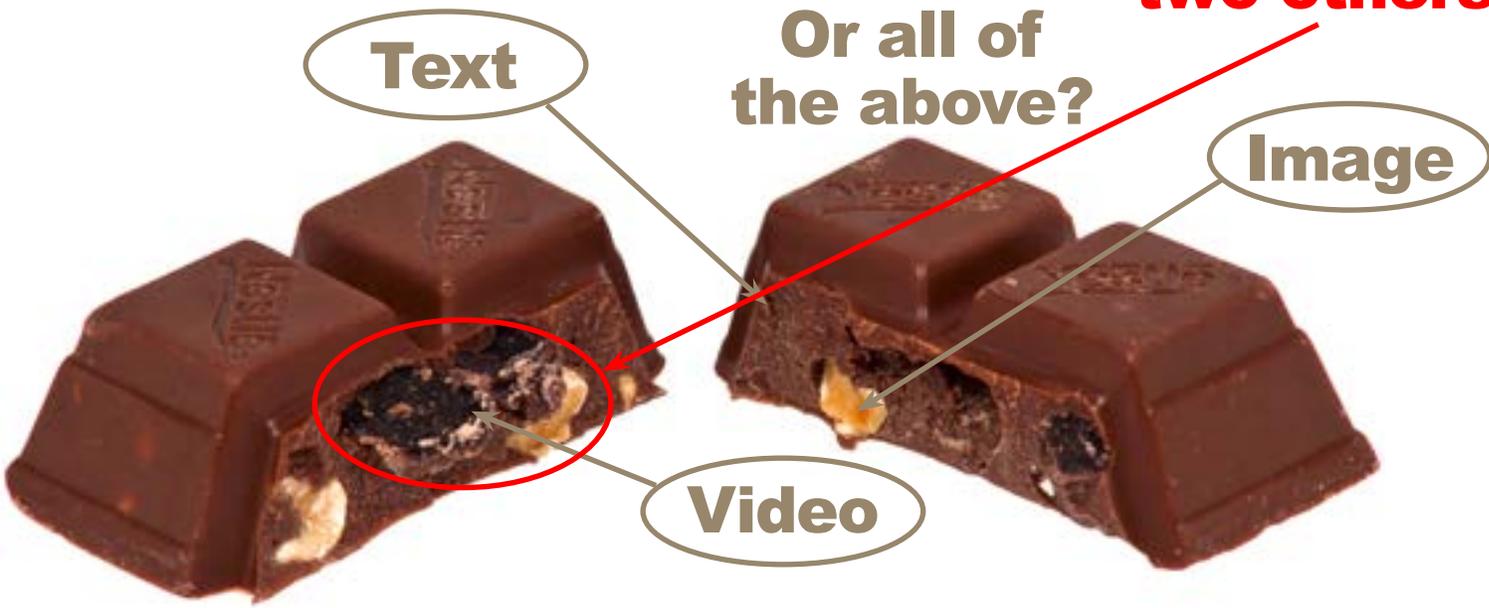
**Video**



**Can they nest?  
Overlap?  
One or many  
files?**

**Is this one  
Discrete  
Entity?  
Or four?  
Or two?  
Or all of  
the above?**

**Is the video  
a Discrete  
Entity?  
(Note that  
it overlaps  
two others.)**



**Do they have to  
be able to exist  
outside their  
original  
context?**

**General  
consensus:**

**Yes.**

**(Note  
implications  
for IDs and  
metadata.)**



**What are they called?**

*Semantically Generic*

*Semantically Specific*



**Cards**  
**Manageable**  
**Objects**  
**Chunks**

**Learning**  
**Objects**  
**Assessments**  
**Tips**

*... and lots more ...*

# Discrete Entity Issues Discussed So Far

---

How big (or small) can they be?

Can they nest? Overlap? One or many files?

Do they have to be able to exist outside their original context?

What are they called?

**Do they need IDs?** If so, what kind?

Should they be required to be accessible?

*Of course!*

What metadata should be associated with them?

# Discrete Entity Issues Discussed So Far

---

How big (or small) can they be?

Can they nest? Overlap? One or many files?

Do they have to be able to exist outside their original context?

What are they called?

Do they need IDs? **If so, what kind?**

Should they be required? Possible?

*Of course!*

*Ummm . . .*

What metadata should be associated with them?

# Discrete Entity Issues Discussed So Far

---

How big (or small) can they be?

Can they nest? Overlap? One or many files?

Do they have to be able to exist outside their original context?

*Unanimous consensus:*  
**YES!**

Do they have to be a certain kind?

**Should they be required to be accessible?**

What metadata should be associated with them?



# Discrete Entity Issues Discussed So Far

---

How big (or small) can they be?

Can they nest? Overlap? One or many files?

Do they have to be able to exist outside their original context?

*Unanimous consensus:*  
**YES!**

Do they have to be of a certain kind?

**Should they be required to be accessible?**

*But what does that mean, exactly?*



# Discrete Entity Issues Discussed So Far

---

How big (or small) can they be?

Can they nest? Overlap? One or many files?

Do they have to be able to exist outside their original context?

What are they called?

Do they need IDs? If so, what kind?

Should they be discoverable?

*Haven't done much on this yet either.*

**What metadata should be associated with them?**

# Where Are We at This Point?

---

**Clear consensus:**

**One or many files**

**Not just HTML docs\***

**Able to stand alone**

**Pedagogical purpose**

**Size is a function of purpose\***

**Must have IDs and metadata\***

**Must be accessible\***

*\*But wait, there's more . . .*

# Where Are We at This Point?

---

**Clear consensus:  
One or many files**

**Not just HMTL docs\***

*Must there be at least one HTML doc?*

*Must it be deliverable as an EPUB?*

*Must it conform to EDUPUB spec?*

**Must have IDs and metadata\***

**Must be accessible\***

*\*But wait, there's more . . .*

# Where Are We at This Point?

---

**Clear consensus:**

*How small is too small?  
How about a single Learning Objective?  
How big is too big?  
Whole chapter? Book? Course?*

**Size is a function of purpose\***

**Must have IDs and metadata\***

**Must be accessible\***

*\*But wait, there's more . . .*

## Where Are We at This Point?

---

*Require standard ID syntax? URI?*

*Looking at LRMI/schema.org,  
IEEE LOM (Learning Object Metadata),  
IMS Content Packaging,  
accessibility metadata needs.*

*(Align with LRMI + Accessibility Metadata  
Integration Guide?)*

**Must have IDs and metadata\***

**Must be accessible\***

*\*But wait, there's more . . .*

# Where Are We at This Point?

---

**Clear consensus:  
One or many files**

*What are minimum  
accessibility requirements?*

*Accompanied by what metadata?  
(Align with LRMI + Accessibility Metadata  
Integration Guide?)*

**Must be accessible\***

*\*But wait, there's more . . .*

# Where Are We at This Point?

---

**Active discussion:**

**Terminology**

**Structure**

**Modeling**

*(Can't really separate these . . .)*

**Yet to address in depth:**

**Metadata specifics**

**Alignment with EDUPUB model**

# Where Are We at This Point?

---

## Active discussion: Terminology

*One collective name for  
EDUPUB chunks in general,  
different names for different kinds,  
or both? Controlled vocab or not?*

*Must use existing HTML5 elements and  
attributes? If so, how?  
(E.g., <article> discussion)*

# Where Are We at This Point?

---

**Active discussion:**

**Terminology**

**Structure**

*As part of source vs.  
independent deliverable object?*

*Must be well formed? Overlap okay?  
Okay to be non-well-formed in source  
(cf. <milestone/>s in JATS/BITS)  
but well-formed as deliverable?*

# Where Are We at This Point?

---

**Active discussion:**

Terminology

Structure

**Modeling**

*Just leave it to EDUPUB?*

*More permissive? If so, how?*

**Alignment with EDUPUB model**

# Thanks!

---

**Bill Kasdorf**

[bkasdorf@apexcovantage.com](mailto:bkasdorf@apexcovantage.com)

+1 734 904 6252

@BillKasdorf

