

Greenstone

Presentation #1

27 Oct 2005

LIS 462

Overview

Overview

★ Ingest

Overview

☆ Ingest

★ Build

Overview

☆ Ingest

☆ Build

★ Render

Ingest

What are we ingesting, exactly?

Ingest

002.jpg



Content, of course...

Ingest

002.jpg



metadata.xml

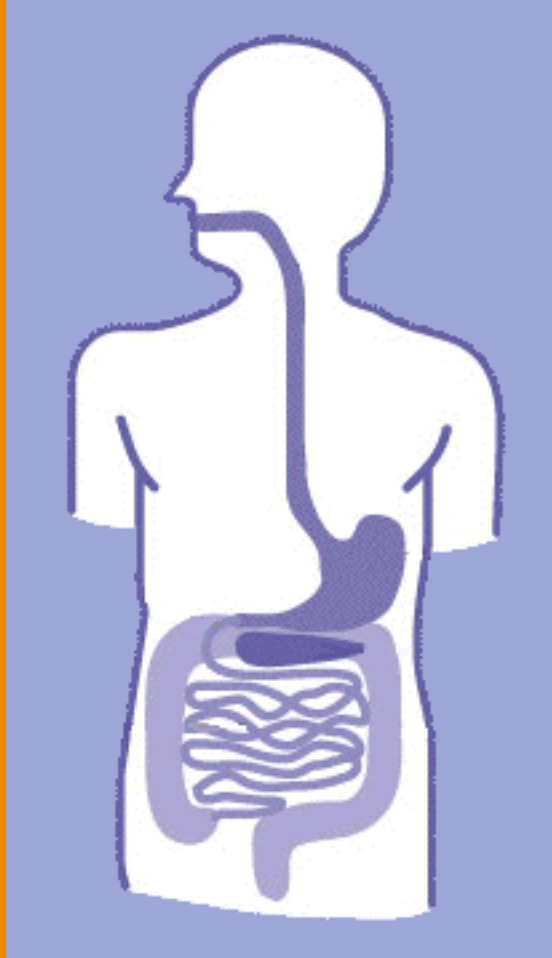
```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ...>
<DirectoryMetadata>
  <FileSet>
    <FileName>puppy.jpg</FileName>
    <Description>
      <Metadata name="dc.Title">Murphy at rest</Metadata>
      <Metadata name="dc.Creator">Lass E.</Metadata>
      <Metadata name="dc.Subject">Murphy</Metadata>
      <Metadata name="dc.Description">...</Metadata>
      <Metadata name="dc.Date">2004-07-22</Metadata>
      ...
    </Description>
  </FileSet>
</DirectoryMetadata>
```



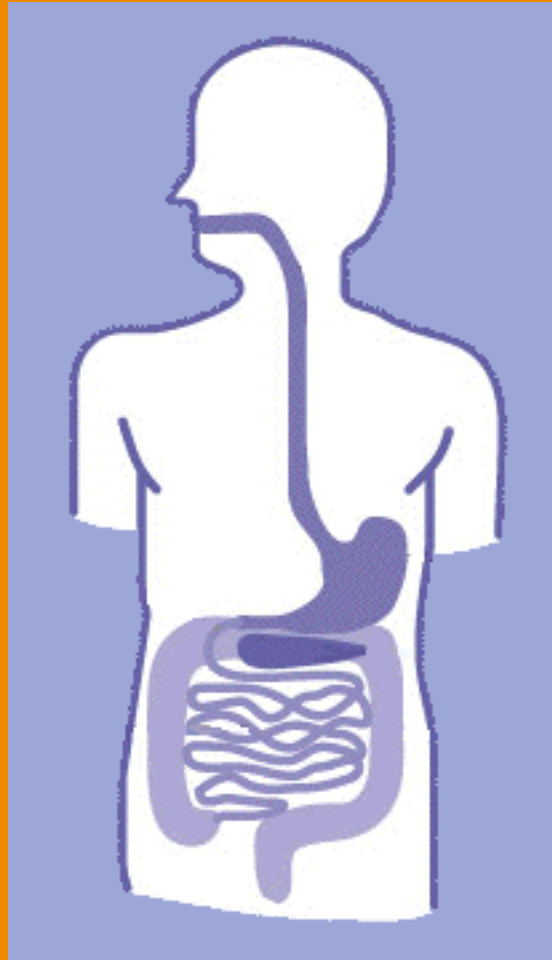
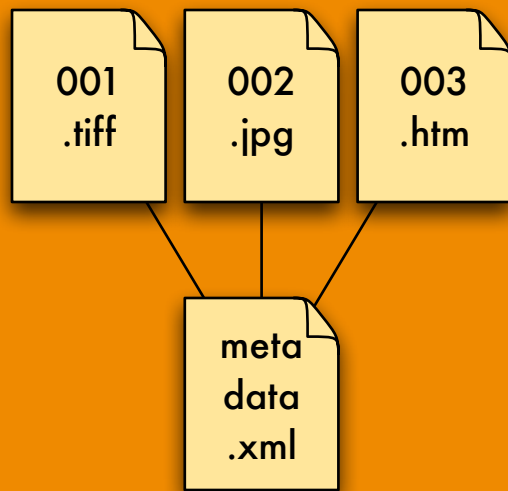
...and metadata.

Ingest

Ingest

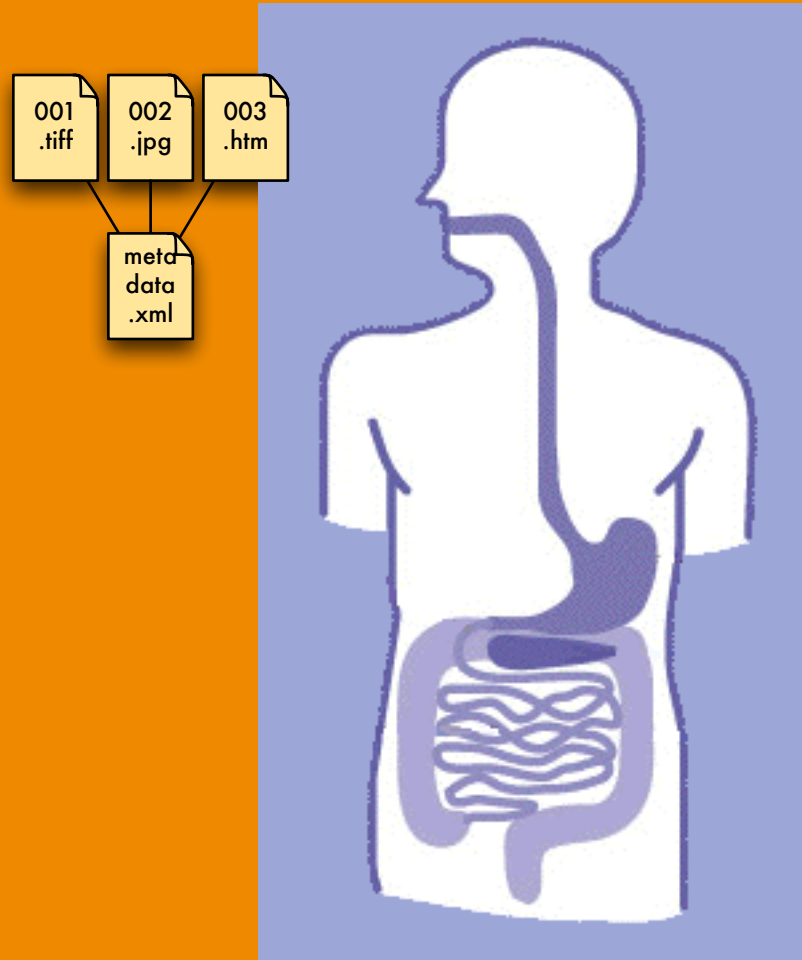


Ingest



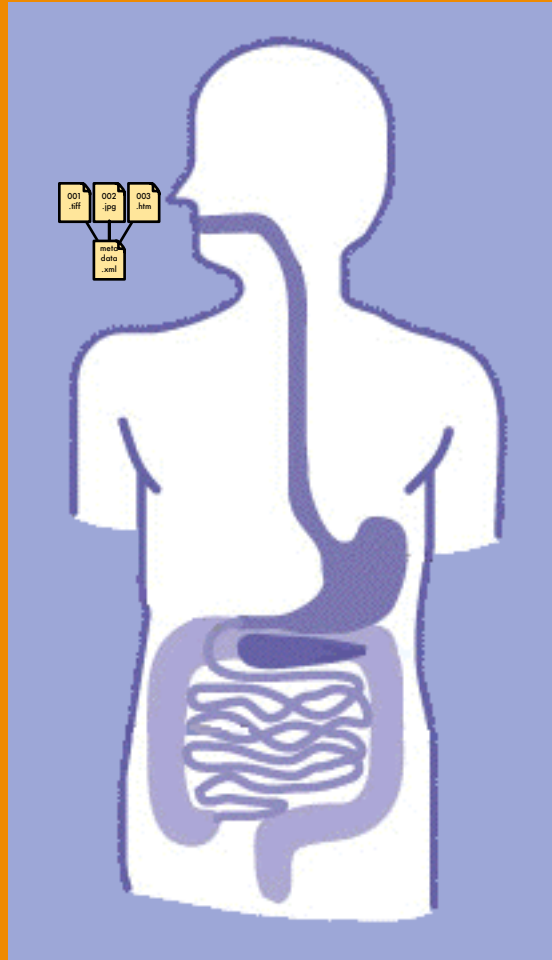
One metadata file, many content files

Ingest



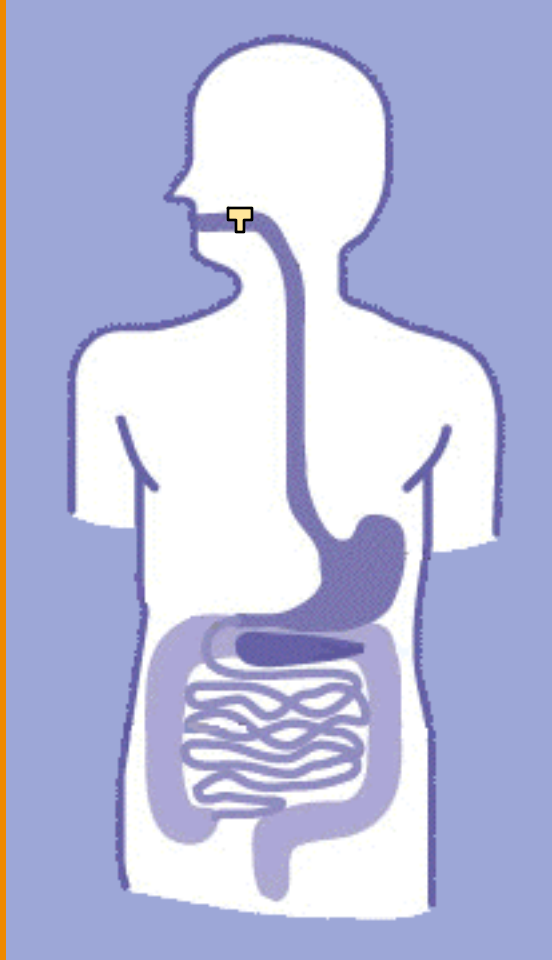
One metadata file, many content files

Ingest



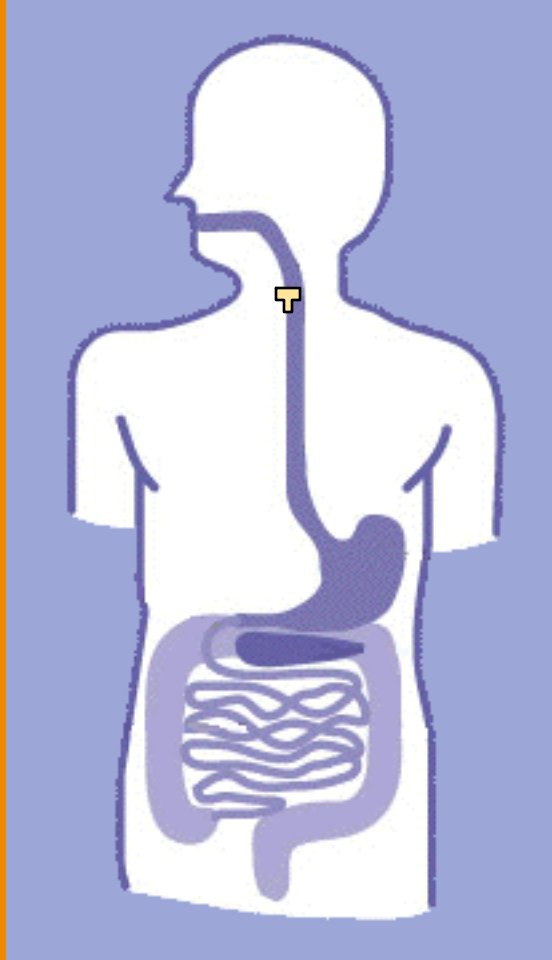
One metadata file, many content files

Ingest



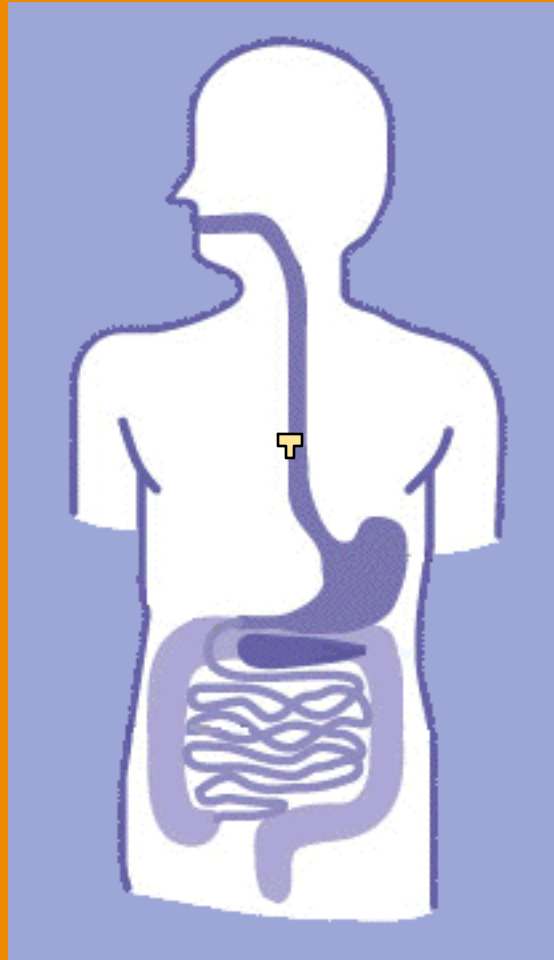
One metadata file, many content files

Ingest



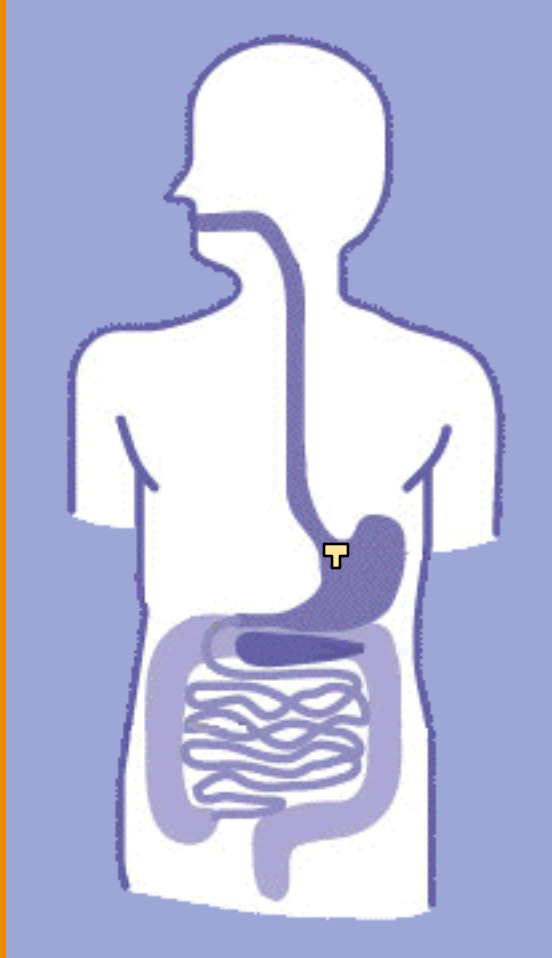
One metadata file, many content files

Ingest



One metadata file, many content files

Ingest



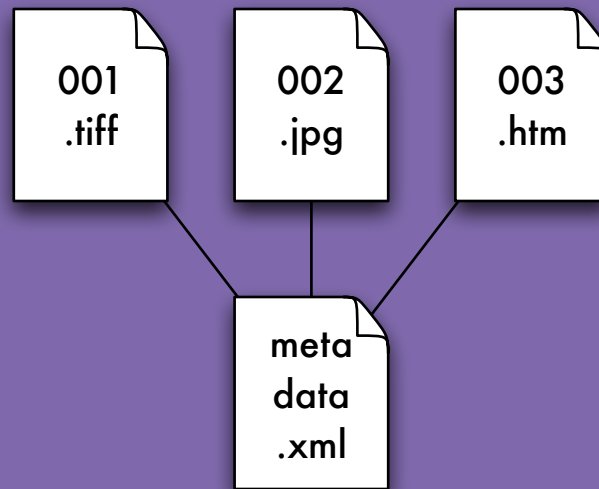
One metadata file, many content files

Ingest



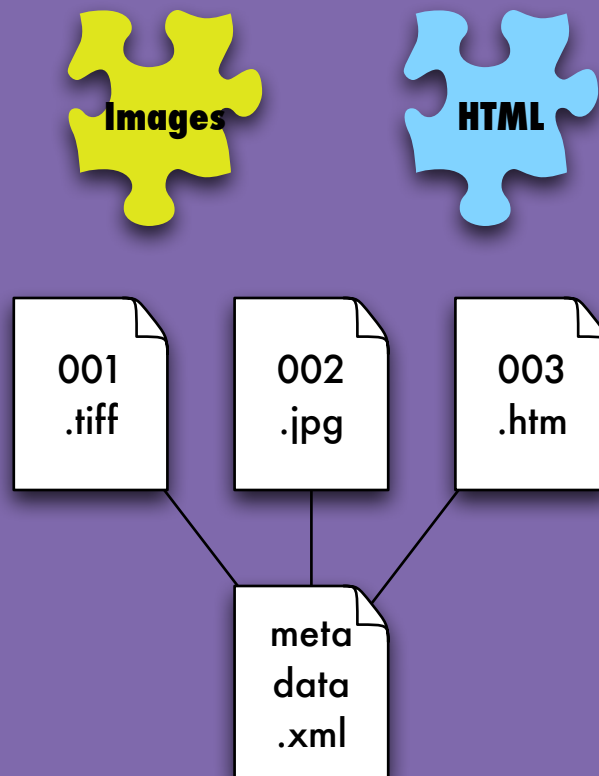
One metadata file, many content files

Ingest



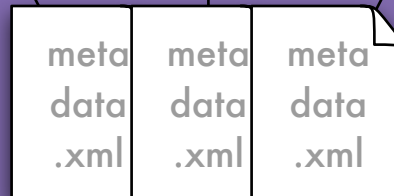
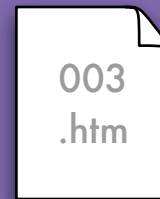
One metadata file, many content files

Ingest



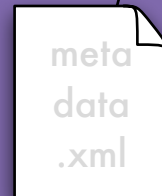
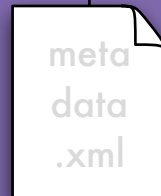
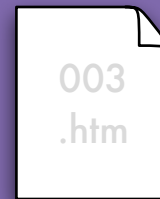
Plugins do the hard work

Ingest



Extract metadata

Ingest



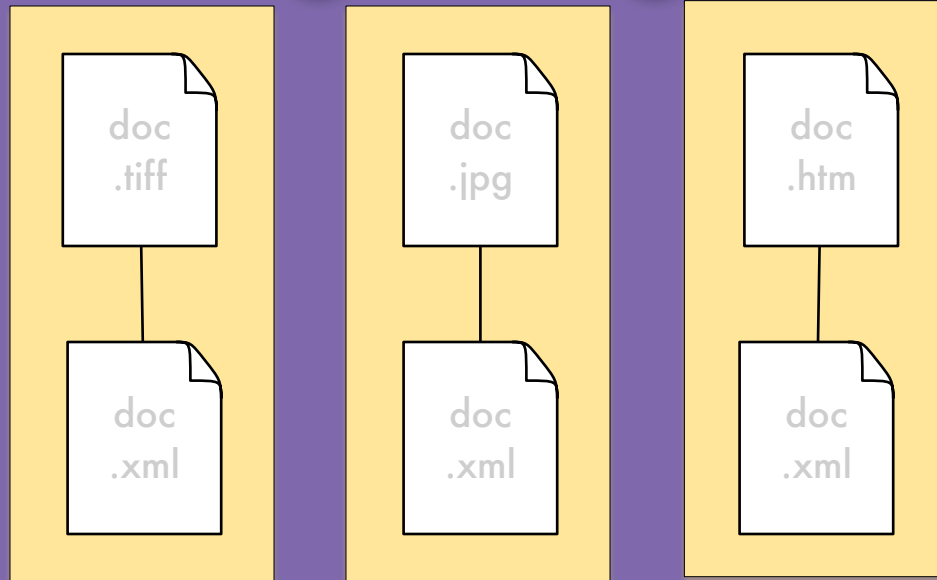
Extract metadata

Ingest



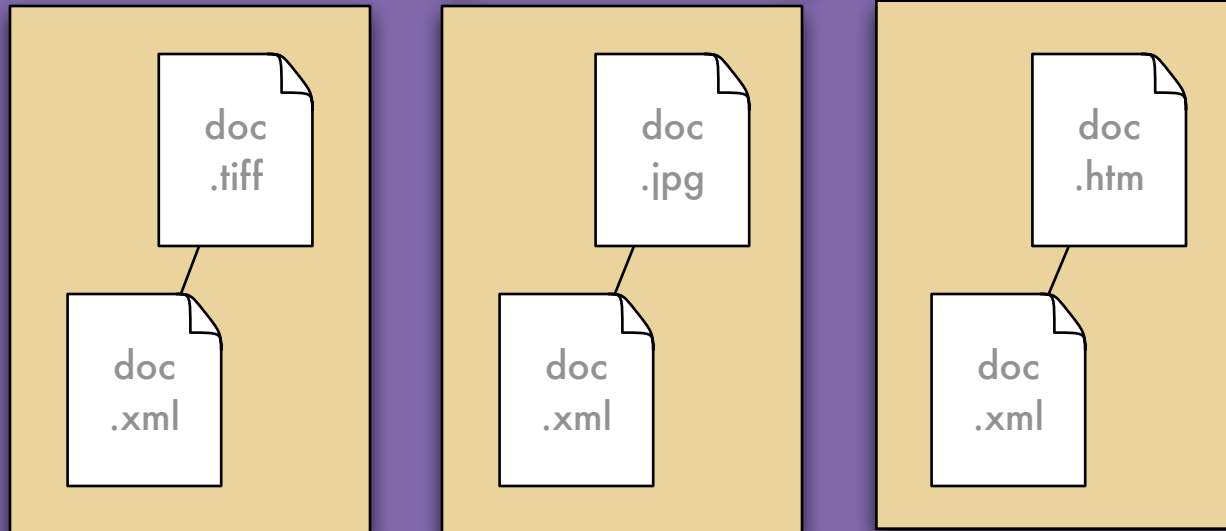
Extract metadata

Ingest



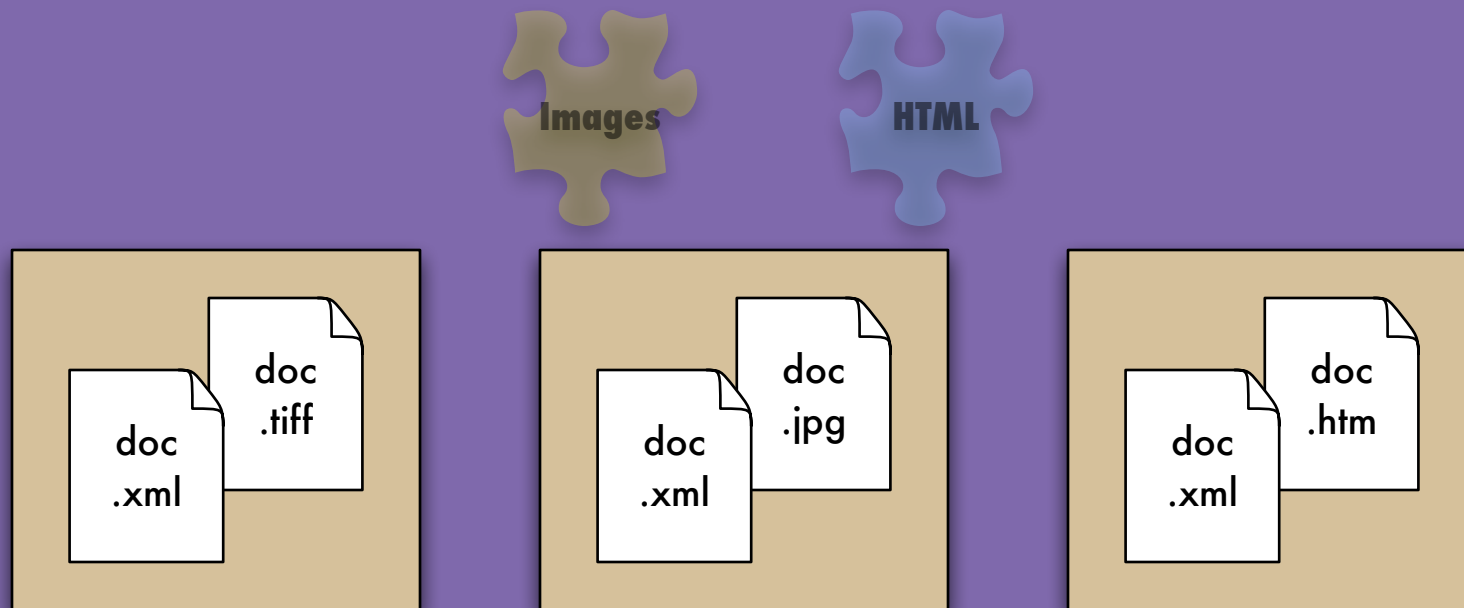
Rename files

Ingest



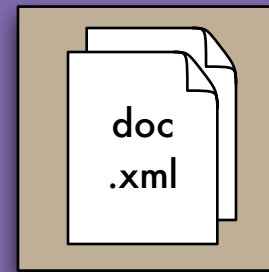
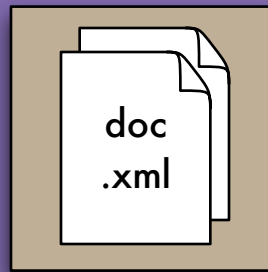
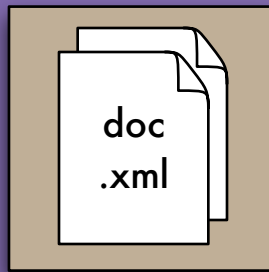
From here on, it's all about the XML

Ingest



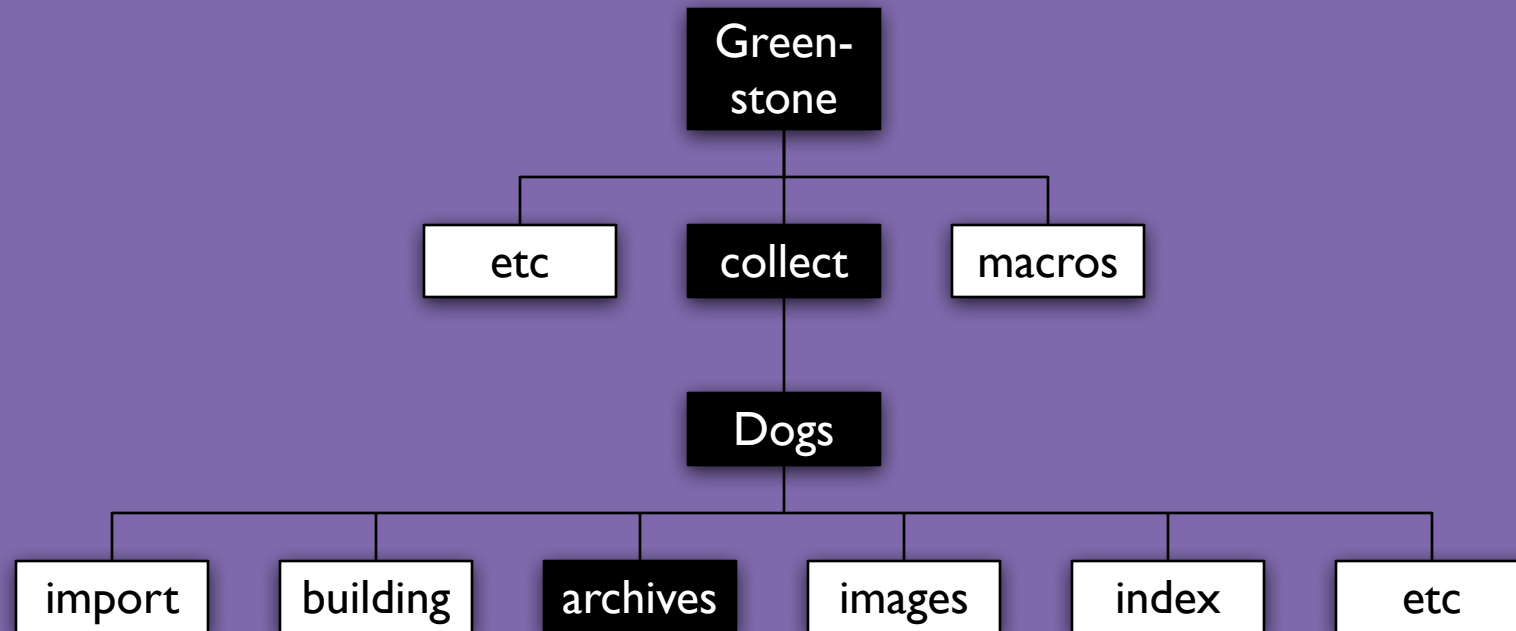
From here on, it's all about the XML

Ingest



From here on, it's all about the XML

Ingest

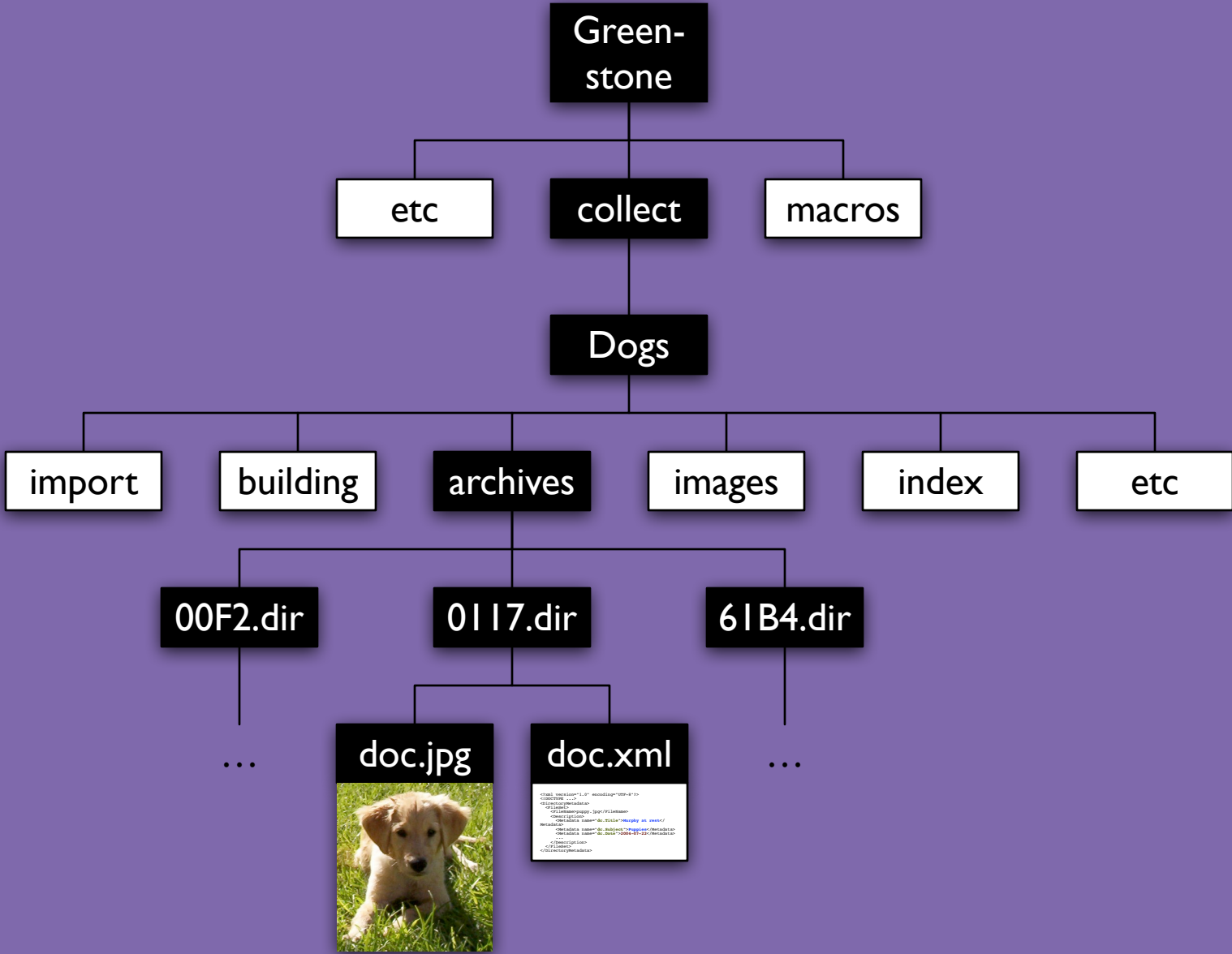


Where goes it go?

Ingest



Ingest



Build

Build



Build



Title: Murphy playing
Creator: Rin T. N. Tin
Subject: Murphy



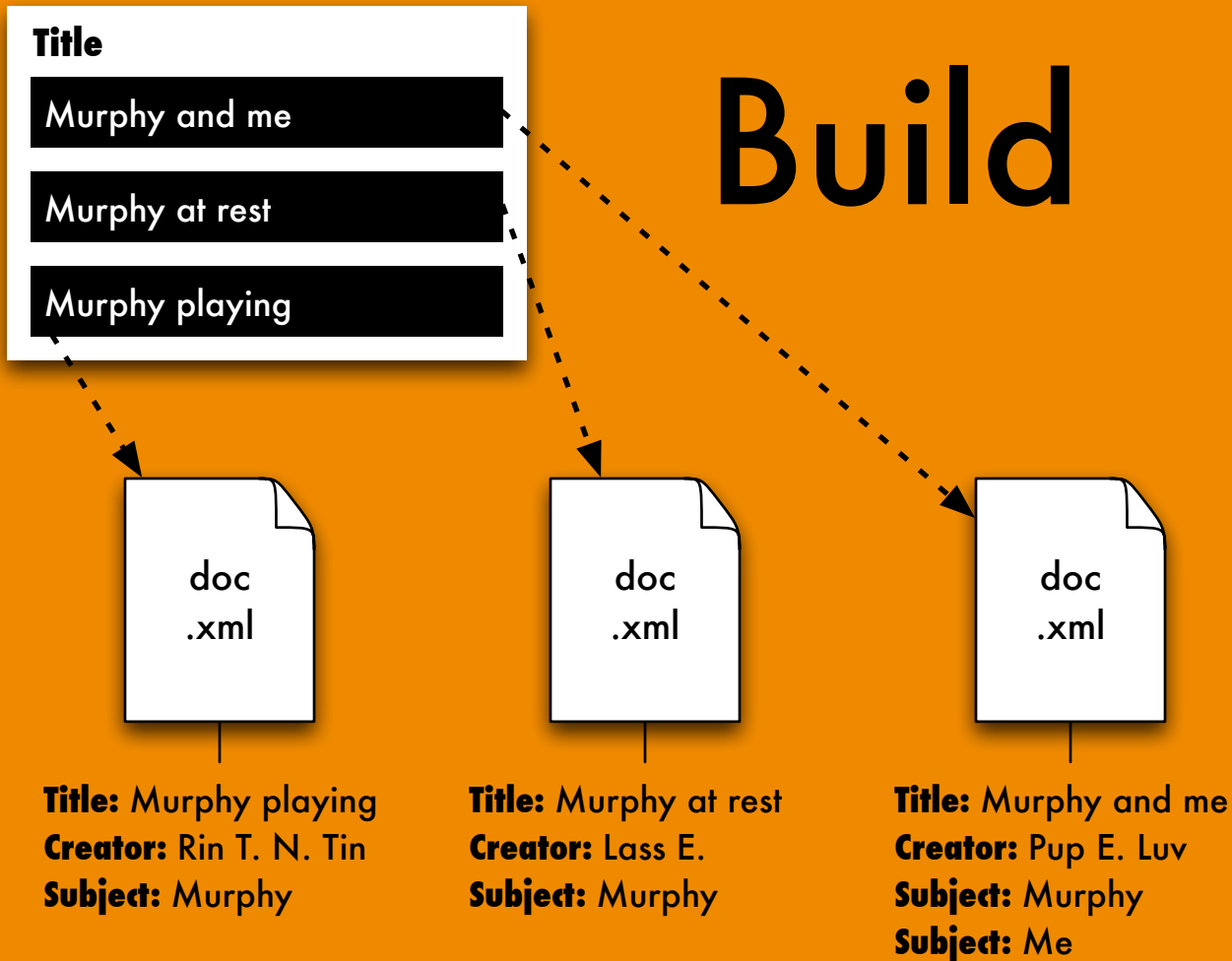
Title: Murphy at rest
Creator: Lass E.
Subject: Murphy



Title: Murphy and me
Creator: Pup E. Luv
Subject: Murphy
Subject: Me

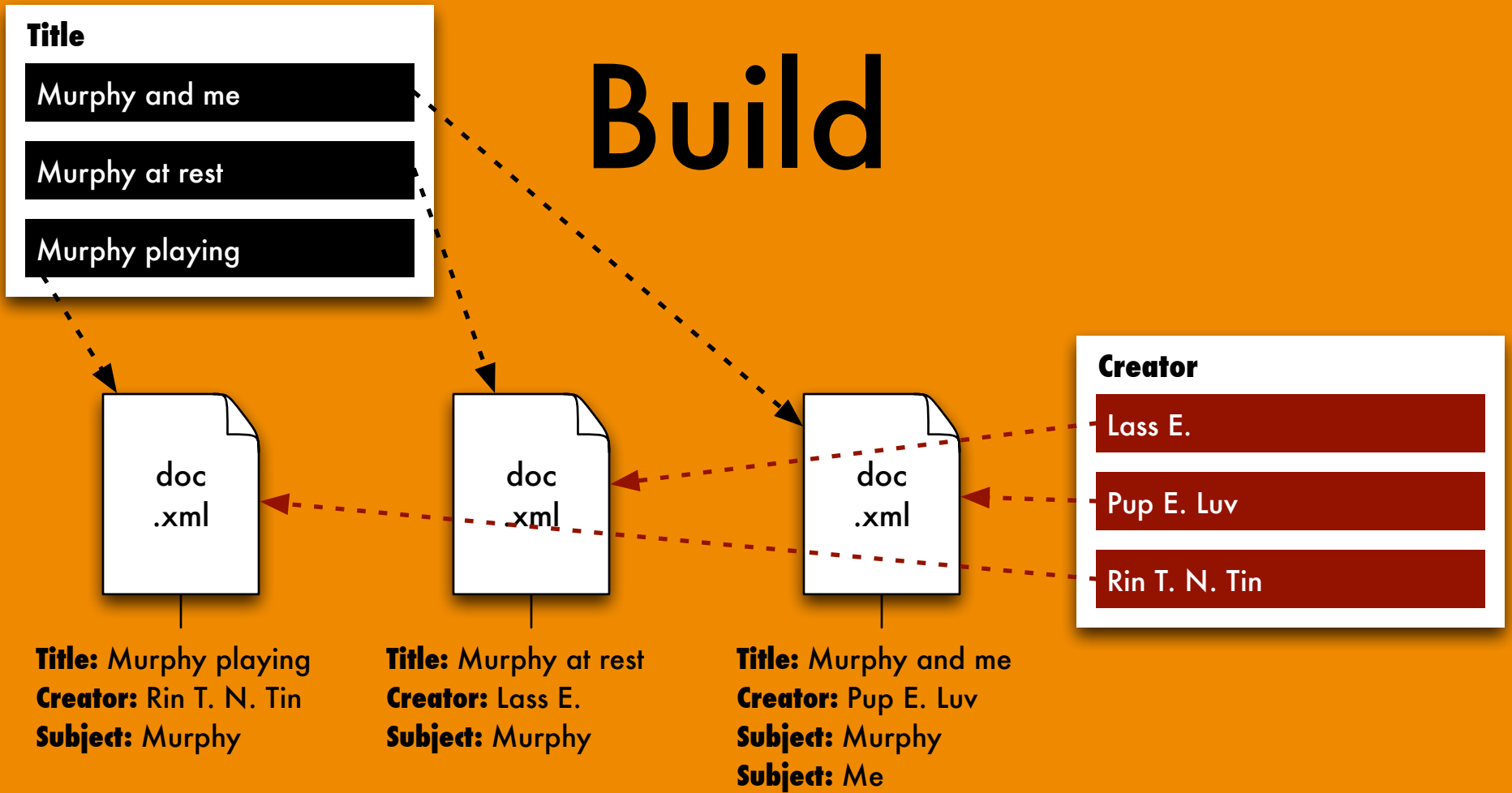
Build indexes for searching and browsing

Build



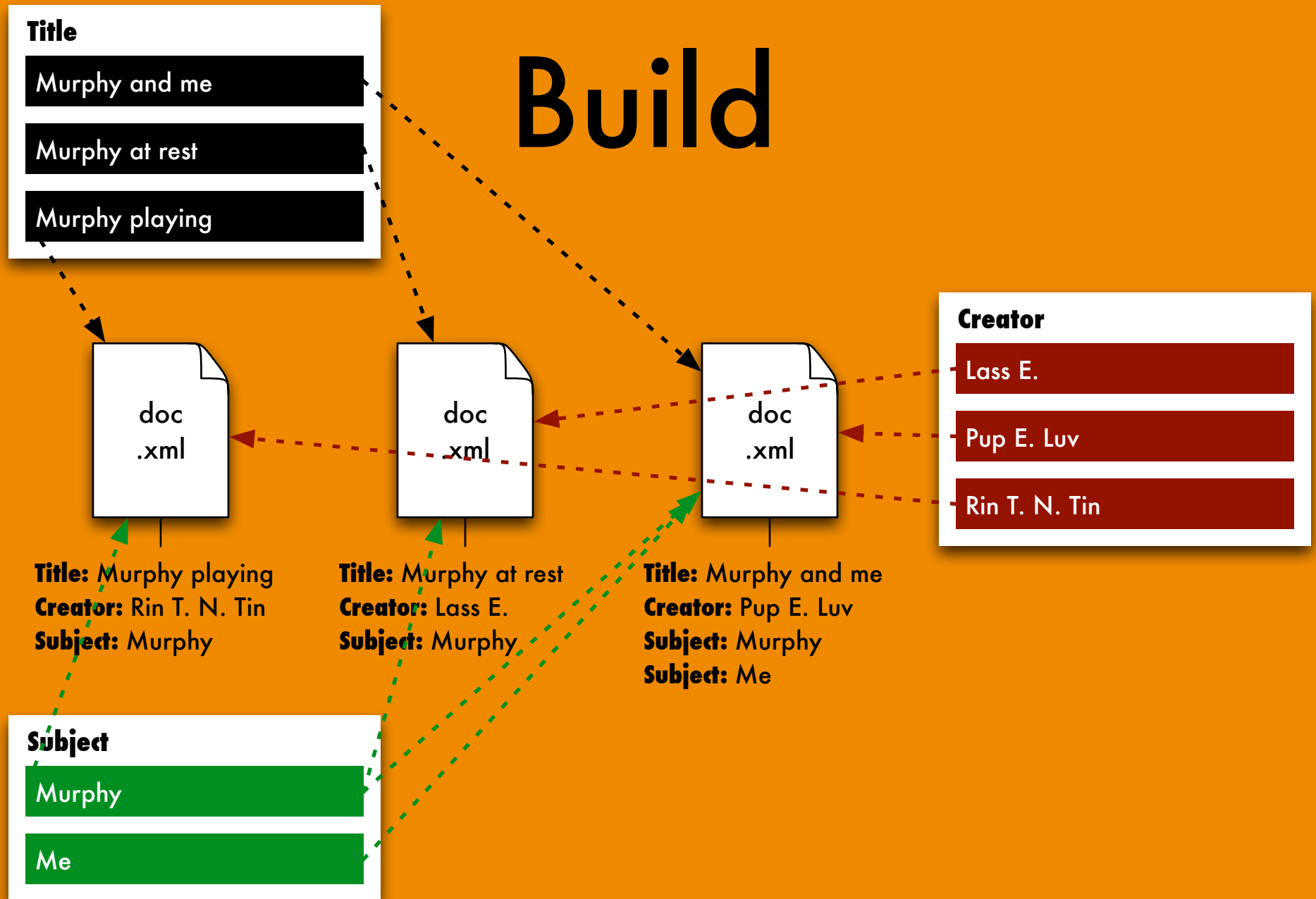
Build indexes for searching and browsing

Build



Build indexes for searching and browsing

Build



Build indexes for searching and browsing

Render

Render

Here's where things get **really** interesting...

Render

My dog pictures

Murphy at rest

by Lass E.

Here's a favorite of mine.
Murphy had been playing for a
long time and was dog tired.



Goal: Text and pictures for people to look at.

Render

```
<html>

<head>
<title>My dog pictures</title>
</head>

<body>

<h1>My dog pictures</h1>

<h2>Murphy at rest</h2>

<p class="byline">by Lass E.</p>

<p class="description">Here's a
favorite of mine. Murphy had been
playing for a long time and was dog
tired.</p>

<p></p>

</body>

</html>
```



My dog pictures

Murphy at rest
by Lass E.

Here's a favorite of mine.
Murphy had been playing for a
long time and was dog tired.



Means: Create HTML with references to images.

Render

```
<html>

<head>
<title>My dog pictures</title>
</head>

<body>

<h1>My dog pictures</h1>

<h2>Murphy at rest</h2>

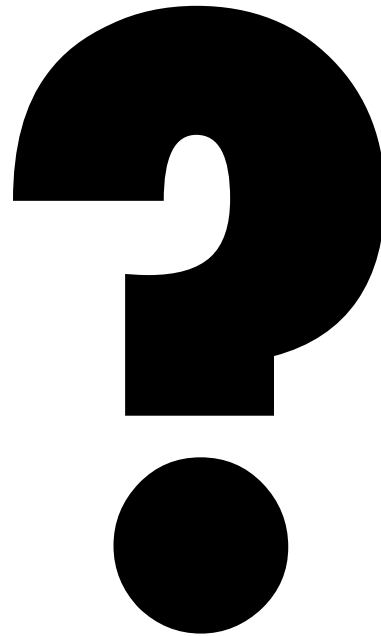
<p class="byline">by Lass E.</p>

<p class="description">Here's a
favorite of mine. Murphy had been
playing for a long time and was dog
tired.</p>

<p></p>

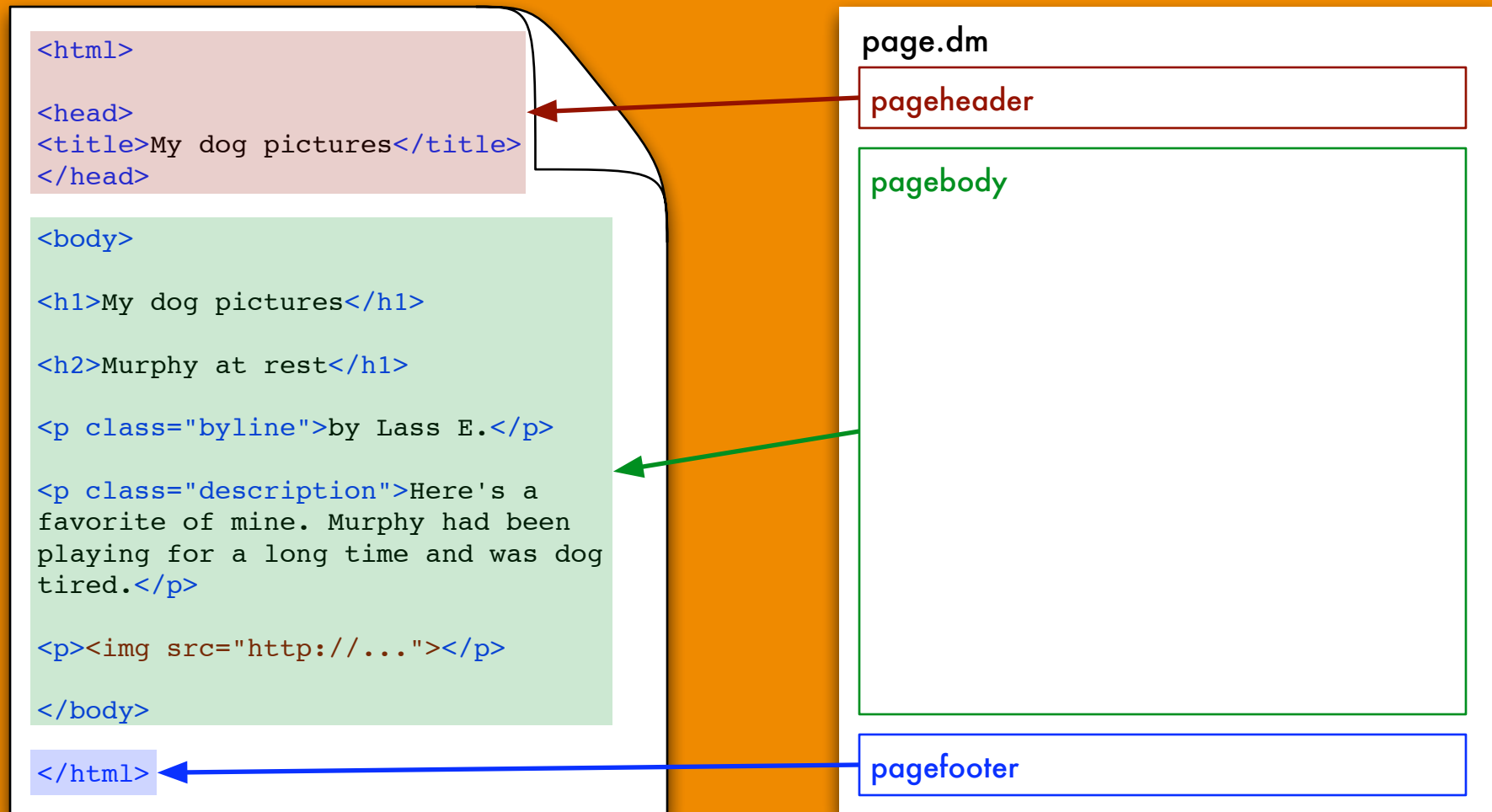
</body>

</html>
```



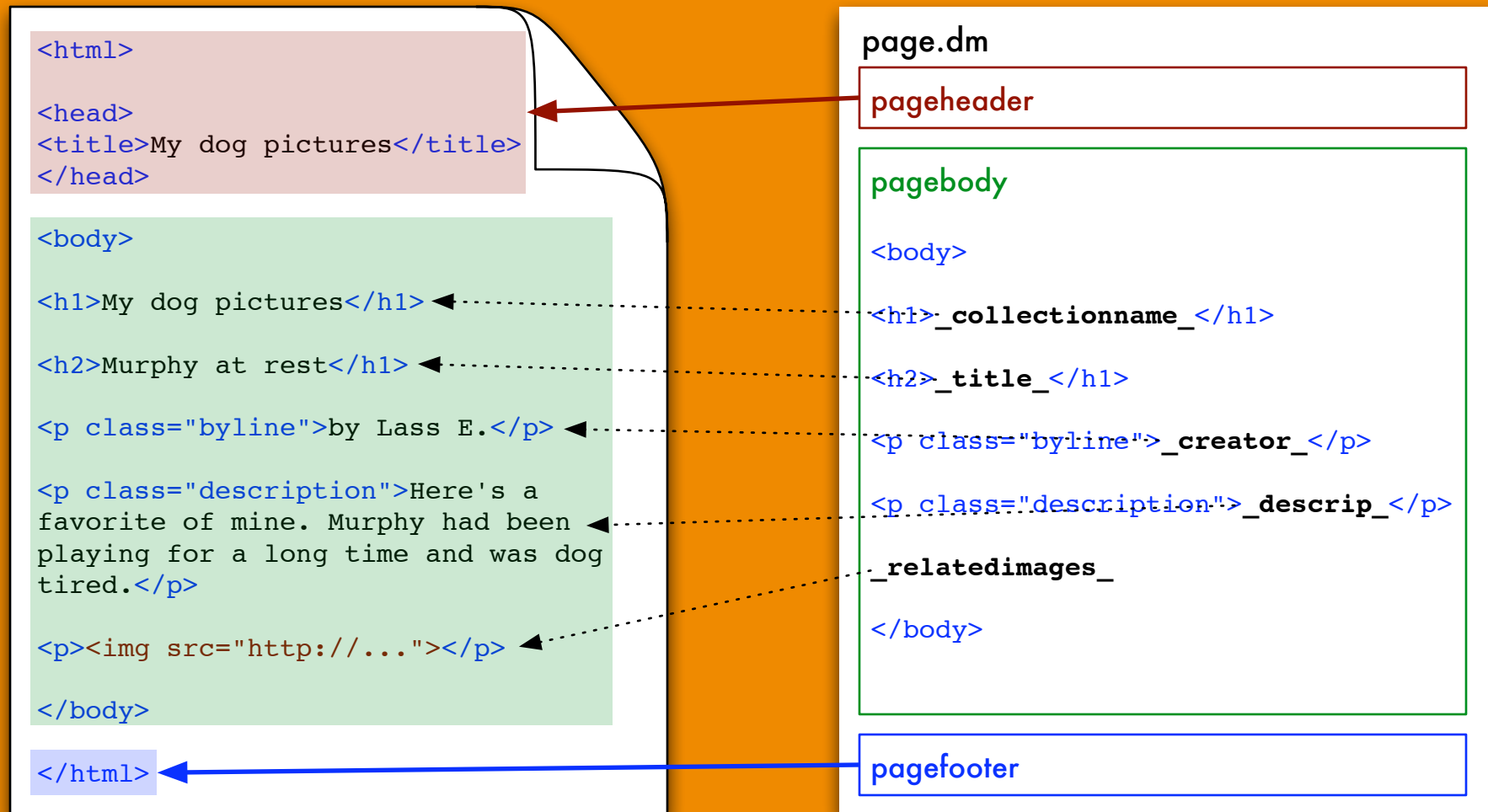
Q: How is the HTML created?

Render



A: Bit by bit using *macros*, which define page structure...

Render



A: ...and *formats*, which are like fill-in-the-blanks...

Render

collect.cfg

```
collectionname "My dog pictures"  
...
```

```
<html>  
  
<head>  
<title>My dog pictures</title>  
</head>  
  
<body>  
  
<h1>My dog pictures</h1>  
  
<h2>Murphy at rest</h2>  
  
<p class="byline">by Lass E.</p>  
  
<p class="description">Here's a  
favorite of mine. Murphy had been  
playing for a long time and was dog  
tired.</p>  
  
<p></p>  
  
</body>  
  
</html>
```

page.dm

pageheader

pagebody

```
<body>  
  
<h1>_collectionname_</h1>  
  
<h2>_title_</h2>  
  
<p class="byline">_creator_</p>  
  
<p class="description">_descrip_</p>  
</body>
```

metadata.xml

```
...  
<Metadata name="dc.Title">Murphy at rest</Metadata>  
<Metadata name="dc.Creator">Lass E.</Metadata>  
<Metadata name="dc.Subject">Murphy</Metadata>  
<Metadata name="dc.Description">...</Metadata>  
<Metadata name="dc.Date">2004-07-22</Metadata>  
...
```

A: ...and *metadata*, which fills in the blanks.

Recap



Ingest



Build



Render

Puppy

