Aggregating media fragments into collaborative mashups: Standards and a prototype

Open Data on the Web - W3C Workshop
London, April 23-24, 2013

Philippe Duchesne, HighLatitudes
@pduchesne
Observation: data curators face issues such as

- **not owning** referred data sources
- need **only parts** of resources identified by URIs
- need to **organize** annotate store exchange the result of their curation as a **resource of its own**
CrossLinks project is grounded in several use cases:

• Work Context storage and sharing
  – Disaster management
• Story telling
  – Data journalism
• Metadata augmentation
  – Digital media archiving

It’s all about *stitching together* heterogeneous *data fragments* into meaningful and *contextualized* patchworks
1. Enable identification and linking of **fragments** of online media resources
   – Unstructured as video, audio, text, …
   – Structured as XML, tabular, …

2. Devise an abstract model and **representation** format(s) for data mosaics

3. Implement a **platform** for authoring, storage, sharing and visualization
• URI spec leaves the fragment part unspecified

• Several standardization efforts exist regarding specific mime types
  – MediaFragments (W3C Recommendation)
    • Focused on temporal (#t) and pixel space fragments (#xywh)
    • Dedicated to audio/video media
  – text/csv fragments (IETF Draft)
    • Defines #col, row, cell
  – text/plain fragments (IETF RFC 5147)
    • Defines #line
Other mainstream usages

- `#page` for PDF resources
- `#<id>` implicit id fragment used in HTML, RDF, …
• But no harmonized definition of fragments for all media types

• Also, need to decouple media types from fragment dimensions
  – XML, PDF, MPEG, CSV, … vs Text, Time, Spatial, …
Identified dimensions

• Textual
  – Line numbers and char offset

• Temporal

• Pixel space

• Tabular

• Geospatial space
  – Proposing \#bbox

• Tree path
  – Proposing \#xpath
<table>
<thead>
<tr>
<th>Id</th>
<th>Text</th>
<th>Time</th>
<th>Pixel</th>
<th>Spatial</th>
<th>Tabular</th>
<th>Tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>html</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>text/plain</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pdf</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>video/*</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>audio/*</td>
<td></td>
<td></td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>image/*</td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>geotiff</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>kml</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>csv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>json</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>
Abstract model basically a set of annotations organized in a folder-like structure

Annotations can link any resource fragment to another (or to inline content), with
- author
- semantic type
- free form tags
Data mosaics: representation

• JSON-based encoding
  – More convenient for browser-based apps
  – Easier to get acquainted with for target audience

• RDF-based representation also available
Data mosaics: representation
• Triple store
• REST-based interface  
  – JSON encoding
• SPARQL endpoint  
  – fine-grained queries of the data mosaics graphs
• web portal  
  – Authoring & Management
• Bookmarklets  
  – Overlays resources with annotations
- Seamless integration in regular browser flow
  - viewing: offers augmented view on annotated resources
  - authoring: helps users capture comments and train of thought while browsing
    → browse the annotation graph in a natural way
- Collaborative annotations
  - All entities (mosaics, resources, annotations) have authors and timestamps
## Supported media and fragments

<table>
<thead>
<tr>
<th></th>
<th>Id</th>
<th>Text</th>
<th>Time</th>
<th>Pixel</th>
<th>Spatial</th>
<th>Tabular</th>
<th>Xpath</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>html</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>text/plain</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pdf</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>video/*</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>audio/*</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>image/*</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>geotiff</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>kml</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(✓)</td>
</tr>
<tr>
<td></td>
<td>csv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>json</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Customization

• Platform meant to be extended
  – Custom viewers
  – Extra mime types
  E.g. YouTube, SlideShare

• Boils down to implementing
  – HTML5 portraying
  – Fragment resolution, display and capture
### Example - Authoring

#### CrossLinks

**Edit Link**

- **EventID**: pde2010112215310060_13
- **Version**: 1.3634E+12

**Edit Annotation**

- **Title**: Quake aftershocks
- **Source**: pde
- **EventID**: pde2010112215310060_13
- **Version**: 1.3634E+12

**Link it!**

- **EventID**: pde2010112215310060_13
- **Version**: 1.3634E+12

---

<table>
<thead>
<tr>
<th>DateTime</th>
<th>DateTime</th>
<th>DateTime</th>
<th>DateTime</th>
<th>DateTime</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-01-12T21:53:10.050+00:00</td>
<td>2010-01-12T22:00:41.490+00:00</td>
<td>2010-01-12T22:07:02.040+00:00</td>
<td>2010-01-12T22:09:11.610+00:00</td>
<td>2010-01-12T22:12:04.650+00:00</td>
</tr>
<tr>
<td>2010-01-12T22:52:51.260+00:00</td>
<td>2010-01-12T22:57:03.740+00:00</td>
<td>2010-01-12T23:12:01.990+00:00</td>
<td>2010-01-12T23:26:05.850+00:00</td>
<td>2010-01-12T23:27:34.860+00:00</td>
</tr>
<tr>
<td>2010-01-12T23:35:39.210+00:00</td>
<td>2010-01-12T23:47:36.100+00:00</td>
<td>2010-01-12T23:47:36.100+00:00</td>
<td>2010-01-13T00:10:19.150+00:00</td>
<td>2010-01-13T00:23:13.200+00:00</td>
</tr>
<tr>
<td>2010-01-13T00:23:55.580+00:00</td>
<td>2010-01-13T00:42:48.990+00:00</td>
<td>2010-01-13T00:43:27.280+00:00</td>
<td>2010-01-13T00:59:05.050+00:00</td>
<td>2010-01-13T01:05:46.610+00:00</td>
</tr>
<tr>
<td>2010-01-13T01:14:53.790+00:00</td>
<td>2010-01-13T01:16:56.980+00:00</td>
<td>2010-01-13T01:24:31.990+00:00</td>
<td>2010-01-13T01:32:44.210+00:00</td>
<td>2010-01-13T01:36:31.370+00:00</td>
</tr>
<tr>
<td>2010-01-13T01:50:00.040+00:00</td>
<td>2010-01-13T01:55:16.000+00:00</td>
<td>2010-01-13T01:57:34.550+00:00</td>
<td>2010-01-13T02:17:56.270+00:00</td>
<td>2010-01-13T02:17:56.270+00:00</td>
</tr>
<tr>
<td>2010-01-13T02:24:17.630+00:00</td>
<td>2010-01-13T02:26:33.670+00:00</td>
<td>2010-01-13T02:43:47.160+00:00</td>
<td>2010-01-13T02:54:15.140+00:00</td>
<td>2010-01-13T03:13:26.870+00:00</td>
</tr>
<tr>
<td>2010-01-13T03:17:11.500+00:00</td>
<td>2010-01-13T03:31:57.190+00:00</td>
<td>2010-01-13T03:53:45.370+00:00</td>
<td>2010-01-13T04:29:49.020+00:00</td>
<td>2010-01-13T04:37:50.890+00:00</td>
</tr>
</tbody>
</table>
Data mosaics: representation
Aftershock

By Bryan Walsh, Jay Newton-Small and Tim Padgett | Thursday, Jan. 21, 2010

Michaud Jonas returned to the ruins of the Palm Apparel factory to see if he could find his little sister’s body — and, possibly, a job. Hundreds of workers were buried under the rubble of this T-shirt-manufacturing plant in the Port-au-Prince suburb of Carrefour, and Jonas’ sister, 22, was one of them. The scent of decay around the neighborhood was overpowering. Yet though he mourned his loss — his brother and mother also died, when the family’s home collapsed — he looked ahead. “Here was the worst place hit, so maybe it’ll be the first place to recover,” he said. “I need to find a job so I can help what’s left of my family. They are depending on me.”

Haiti is a country in agony. More than a week after a 7.0-strength quake leveled Port-au-Prince and its surroundings, rescue teams kept combing the ruins for survivors — and some were miraculously found after days in the rubble — but hope was slowly dwindling. International help began to pour into Haiti, but with the capital and most political infrastructure destroyed, the aid wasn’t nearly enough. Emergency medical teams operating out of shipping containers worked to save limbs and lives, and desperate requests went out to U.S. medical schools for more volunteers. More than 1.5 million Haitians are homeless, and the full death toll — perhaps 100,000 or more — remains unknown and may be unknowable, as tens of thousands of corpses were delivered to mass graves without ceremony. As if that weren’t bad enough, a 6.1-magnitude aftershock rocked Haiti on Jan. 20, further rattling nerves. “This is really an unprecedented situation,” says U.N. Secretary-General Ban Ki-moon. He feels it personally. The U.N. lost more than 40 people in the quake, the worst disaster the global body has suffered in its 65-year history. “It is overwhelming — overwhelming.”

(See TIME’s exclusive photos of the earthquake’s destruction.)
MAGNITUDE 7.0

Aftershock
By Bryan Walsh, Jay Newton-Small and Tim Padgett | Thursday, Jan. 21, 2010

Michael Jonas returned to the ruins of the Palm Apparel factory to see if he could find his little sister’s body — and, possibly, a job. Hundreds of workers were buried under the rubble of this T-shirt-manufacturing plant in the Port-au-Prince suburb of Carrefour, and Jonas’ sister, 22, was one of them. The scent of decay around the neighborhood was overpowering. Yet though he mourned his loss — his brother and mother also died, when the family’s home collapsed — he looked ahead. “Here was the worst place hit, so maybe it’ll be the first place to recover,” he said. “I need to find a job so I can help what’s left of my family. They are depending on me.”

Haiti is a country in agony. More than a week after a 7.0-strength quake leveled Port-au-Prince and its surroundings, rescue teams kept combing the ruins for survivors — and some were miraculously found after days in the rubble — but hope was slowly dwindling. International help began to pour into Haiti, but with the capital and most political infrastructure destroyed, the aid wasn’t nearly enough. Emergency medical teams operating out of shipping containers worked to save limbs and lives, and desperate requests went out to U.S. medical schools for more volunteers. More than 1.5 million Haitians are homeless, and the full death toll — perhaps 100,000 or more — remains unknown and may be unknowable, as tens of thousands of corpses were delivered to mass graves without ceremony. As if that weren’t bad enough, a 6.1-magnitude aftershock rocked Haiti on Jan. 20, further rattling nerves. “This is really an unprecedented situation,” says U.N. Secretary-General Ban Ki-moon. He feels it personally. The U.N. lost more than 40 people in the quake, the worst disaster the global body has suffered in its 65-year history. “It is overwhelming — overwhelming.”

(See TIME’s exclusive photos of the earthquake’s destruction.)
MAGNITUDE 7.0
Aftershock
By Bryan Walsh, Jay Newton-Small and Tim Padgett  Thursday, Jan. 21, 2010

Michaud Jonas returned to the ruins of the Palm Apparel factory to see if he could find his little sister's body — and, possibly, a job. Hundreds of workers were buried under the rubble of this T-shirt-manufacturing plant in the Port-au-Prince suburb of Carrefour, and Jonas' sister, 22, was one of them. The scent of decay around the neighborhood was overpowering. Yet though he mourned his loss — his brother and mother also died, when the family's home collapsed — he looked ahead. "Here was the worst place hit, so maybe it'll be the first place to recover," he said. "I need to find a job so I can help what's left of my family. They are depending on me."

Haiti is a country in agony. More than a week after a 7.0-strength quake leveled Port-au-Prince and its surroundings, rescue teams kept combing the rubble, but some were miraculously found after days in the rubble — but hope was slowly diminishing. International help began to pour into Haiti, but with the capital and most political infrastructure destroyed, the aid wasn't nearly enough. Emergency medical teams operating out of shipping containers worked to save limbs and lives, and desperate requests went out to U.S. medical schools for more volunteers. More than 1.5 million Haitians are homeless, and the full death toll — perhaps 100,000 or more — remains unknown and may be uncountable, as tens of thousands of corpses were delivered to mass graves without ceremony. As if that weren't bad enough, a 6.1-magnitude aftershock rocked Haiti on Jan. 20, further rattling nerves. "This is really an unprecedented situation," says U.N. Secretary-General Ban Ki-moon. He feels it personally. The U.N. lost more than 40 people in the quake, the worst disaster the global body has suffered in its 65-year history. "It is overwhelming — overwhelming."

(See TIME's exclusive photos of the earthquake's destruction.)
MAGNITUDE 7.0

Aftershock

By Bryan Walsh, Jay Newton-Small and Tim Padgett | Thursday, Jan. 21, 2010

[Image of a map and a photo]

Haiti is a country in agony. More than a week after a 7.0-magnitude quake leveled Port-au-Prince and its surroundings, rescue teams kept combing the ruins for survivors — and some were miraculously found after days in the rubble — but hope was slowly dwindling. International help began to pour into Haiti, but with the capital and most political infrastructure destroyed, the aid wasn’t nearly enough. Emergency medical teams operating out of shipping containers worked to save limbs and lives, and desperate requests went out to U.S. medical schools for more volunteers. More than 1.5 million Haitians are homeless, and the full death toll — perhaps 100,000 or more — remains unknown and may be unknowable, as tens of thousands of corpses were delivered to mass graves without ceremony. As if that weren’t bad enough, a 6.1-magnitude aftershock rocked Haiti on Jan. 20, further rattling nerves. "This is really an unprecedented situation," says U.N. Secretary-General Ban Ki-moon. He feels it personally. The U.N. lost more than 40 people in the quake, the worst disaster the global body has suffered in its 65-year history. "It is overwhelming — overwhelming."

(See TIME’s exclusive photos of the earthquake’s destruction.)
Example
• RDF/SPARQL exposure using well-known vocabularies
  – Open Annotations
    • Proposes a detailed model for annotations, compatible with this project
• Include visualization-specific config in data mosaics (esp. for CSV)
• Allow annotations filtering based on author, dates, semantic tags
• Some features would benefit from becoming mainstream browser features
  – Fragment resolution for supported mimetypes
  – Fragment selection
• Seamless integration of authoring/viewing avoids breaking train of thought \(\Rightarrow\) favors engagement and better captures domain knowledge
• Data mosaics act as exchangeable “mind map” containers, capturing contextual metadata in a domain-specific perspective, for a given use case
Thank you

phd [at] highlatitud.es
@pduchesne