Sudarshan Murthy, David Maier, Lois Delcambre
Department of Computer Science, Portland State University

http://sparce.cs.pdx.edu/mash-o-matic/
mailto:smurthy@cs.pdx.edu

Supported in part by US National Science Foundation grants 0511050 and IIS-0534762
Mash-ups

• Web Applications that combine information from multiple sources [Wikipedia]

• Several map-based mash-ups exist
• Some non-map mash-ups also exist
  - Google News, Yahoo! News
• Portland State University (PSU) Campus Map demo
Motivation

• Data for a mash-up may come from several sources
• The sources may vary in terms of format, structure, and location
• Data may need to be elaborated and transformed before use
• Likely use only fragments of a source’s contents
• May require considerable effort to select, extract, elaborate, and transform data
Portland State University Campus Map

- 45 markers, 53 landmarks
  - Marker: Balloon on map
  - Landmark: Building, department, ...

- Information comes from 188 fragments in 58 web pages

- Fragments selected manually

http://sparce.cs.pdx.edu/cmap/
Portland Metro Food Markets

- 154 markers, 154 landmarks
- 154 fragments harvested programmatically from 4 MS Word documents
- Developed for the Oregon Department of Agriculture (ODA)

http://sparce.cs.pdx.edu/mash-o-matic/oda-1.1/
Outline

• Introduction
• The mash-up production process
• Mash-o-matic components
  - Superimposed information and applications
  - Transforming bi-level information
  - Non-map mash-ups
• Producing map-based mash-ups
  - Model and benefits
• Conclusions
The Mash-up Production Process

Collect fragments from different sources

Structure, restructure, elaborate fragments

Refine and re-purpose information

Display re-purposed information
E.g., Producing a Map-based Mash-up

Obtain landmark names and addresses

Classify landmarks: Housing, dining, …

Geo-code, generate XML/KML data

Display markers using a map API

Select and Aggregate

Organize and Classify

Clean and Transform

Display
How Mash-o-matic Helps

Select and Aggregate → Organize and Classify → Clean and Transform → Display

Services

DBMS → Docs

Framework to reference and retrieve fragments; apps to organize and elaborate fragments; Transform retrieved fragments, elaborations

External function

12-Oct-06

Mash-o-matic
Outline

• Introduction
• The mash-up production process
• Mash-o-matic components
  - Superimposed information and applications
  - Transforming bi-level information
  - Non-map mash-ups
• Producing map-based mash-ups
  - Model and benefits
• Conclusions
Superimposed Information (SI)

- SI is new information overlaid on existing *base information*
  - Add new data
  - Impose new schema, model
- *Mark*: A reference to a fragment
- **Benefits**
  - Multiple, simultaneous organizations
  - Make new connections among base fragments
  - Preserve context
Select and Aggregate → Organize and Classify → Clean and Transform → Display

Select Information

- Use plug-ins/add-ins developed for base applications to create marks
  - MS Office, Acrobat, IE, Windows Media Player
  - New base types may be supported easily
- Marks may also be created programmatically

Using an IE plug-in to create a mark into an HTML fragment
Aggregate Information

- **Sidepad**: a superimposed application (SA); a scratchpad tool
- **Item**: An SI structure with a name, comment text, and an optional mark
- **Group**: A collection of items and other groups
- **Document**: A collection of items and groups

Referencing a document fragment from a Sidepad (superimposed) document
Organize and Classify Information

- Schema by convention
  - A landmark’s information is placed inside a group
  - Item with name ‘Type’ indicates landmark type
  - A marker’s information is placed inside a group, including sub-groups for landmarks

- Structure and restructure information
  - Easily move information between groups
Clean and Transform Information

- **Bi-level information**: Referenced fragments and the superimposed elaborations, together
- Transform a Sidepad document to an HTML outline document
  - Excerpts of referenced fragments shown in italics
  - Excerpts shown for items that have no comments

*Display tool*: HTML browser
Sidepad ⇒ Map*

* Generated using Mash-o-matic

Display tool: Google Maps
Timeline drawn using an XML transformation based on the work of Nicolas Kruchten

**Sidepad ⇒ Timeline**

Course information transformed to a timeline (includes references to HTML, Word, PDF, and PowerPoint selections)

*Display tool: HTML browser*
A strand map about “Relational Database Normalization” transformed to Scalable Vector Graphics format (includes references to PowerPoint selections)

Display tool: Acrobat SVG Viewer
A concept map with references to the IR text “Modern Information Retrieval” transformed to the GetSmart format.

Display tool: GetSmart, a concept mapping tool.
Transformation Mechanics

<!--Mark/Context/Content/Text
XPath expression to retrieve excerpts of referenced fragments:

//Mark/Context/Content/Text-->

<Document name="PSU Campus Map">
  <Group name="Simon Benson">
    <Item name="Type">
      ...
    </Item>
    <Item name="Name">
      <Mark ID="HTMLMark...">
        ...
      </Mark>
    </Item>
    <Item name="Info">
      ...
    </Item>
  </Group>
</Document>

<!--Mark/Context/Content/Text
XPath expression to retrieve excerpts of referenced fragments:

//Mark/Context/Content/Text-->

<Document name="PSU Campus Map">
  <Group name="Simon Benson">
    <Item name="Type">
      ...
    </Item>
    <Item name="Name">
      <Mark ID="HTMLMark...">
        ...
      </Mark>
    </Item>
    <Item name="Info">
      ...
    </Item>
  </Group>
</Document>
Outline

• Introduction
• The mash-up production process
• Mash-o-matic components
  - Superimposed information and applications
  - Transforming bi-level information
  - Non-map mash-ups
• Producing map-based mash-ups
  - Model and benefits
• Conclusions
Refresh: The Mash-up Production Process

Services

Select and Aggregate

Organize and Classify

Clean and Transform

Display

DBMS

Docs
Producing Map-based Mash-ups

Geo-coder

Markers

Google Maps API

JavaScript (Ajax)

Landmarks

Bi-level Transformation

HTML (UI)

Geo-coder

JavaScript

Markers

Landmarks

HTML (UI)
Using Yahoo! Maps Simple API

Markers → Transformation

12-Oct-06

Mash-o-matic 24
Using Google Earth KML

Markers → Transformation
Benefits of Using Mash-o-matic

- Time to develop Mash-o-matic: 35.5 hours
  - Already had: Sidepad, infrastructure to create marks and retrieve context, and bi-level query processor
  - Developed: Sidepad document convention, XML schema, transformations; HTML, JavaScript for Google Maps API
- Time to create PSU campus map: 5.5 hours
- Time to create ODA mash-up: 16.5 hours
  - Several customizations after interactions with customer; automatic geo-coding
- Transform to use Yahoo! Maps Simple API: 1 hour
- Transform to Google Earth KML: 2 hours
Outline

• Introduction
• The mash-up production process
• Mash-o-matic components
  - Superimposed information and applications
  - Transforming bi-level information
  - Non-map mash-ups
• Producing map-based mash-ups
  - Model and benefits
• Conclusions
Conclusions

• Mash-o-matic demonstrates integration of heterogeneous technologies, components, and information to produce web applications
• Shows ability to *declaratively* express integration and transformation of document fragments and their elaborations
• Sidepad not the best SA to organize large amounts of data
• SA not needed, but using one can help
Sudarshan Murthy, David Maier, Lois Delcambre
Department of Computer Science, Portland State University

http://sparce.cs.pdx.edu/mash-o-matic/
mailto:smurthy@cs.pdx.edu

Supported in part by US National Science Foundation grants 0511050 and IIS-0534762
Albert Cuypstraat en Van Woustraat
De Pijp, Amsterdam

Adres: Van Woustraat 7 NL-1074 AA Amsterdam
Questions?

Download Mash-o-matic: http://sparce.cs.pdx.edu/mash-o-matic/

Web search: “Mash-o-matic”