Interactive Authoring of Visual Timelines from Unstructured Text

TimeLineCurator

Johanna Fulda, Matthew Brehmer, Tamara Munzner,
University of British Columbia & University of Munich (LMU)
Setting the stage

- Experience creating infographics in newsroom
- **Timelines** are popular, but tedious to create
- How could we make that easier?
Introducing TimeLine Curator

Video:

vimeo.com/jofu/tlc
Creating Timelines

2 common approaches

• **Manual creation**
  Creating **static timelines**, e.g. for print

• **Structured creation**
  Creating **interactive timelines**, e.g. for online use

5 tasks to go through

- Browse
- Extract
- Format
- Show
- Update
Manual creation process

Browse
Extract
Format
Show
Update

**Creating Timelines**

**Manual Creation Process**

1. **Browse**
   - Explore available datasets and sources.

2. **Extract**
   - Extract data from the sources.

3. **Format**
   - Format the extracted data for visualization.

4. **Show**
   - Display the timeline visualizations.

5. **Update**
   - Update the timeline as new data becomes available.

---

**Historical Timeline**

- **1868 The Typewriter**
  - Invented by Christopher Sholes, typewriters quickly became indispensable tools for practically all writers, both their personal correspondence and for business correspondence in private homes.

- **1986 The Mouse**
  - Some additional information here.

- **2007 Multi Touch**
  - With the idea of using Multi-touch became a reality.

---

**Roles and Tools**

- **Speech Recognition**
  - Used to simplify and speed up data entry.

---

**Appendix**

**Manual /Bibliography**

- Various sources and references for creating timelines.
Structured creation process

Browse ➔ Extract ➔ Format ➔ Show ➔ Update

TimelineJS
timeline.knightlab.com/
Timeline Authoring Model

**Manual Creation**
- **Browse**: slow
- **Extract**: slow
- **Format**: slow
- **Show**: slow
- **Update**: slow

**Structured Creation**
- **Browse**: slow
- **Extract**: slow
- **Format**: slow
- **Show**: automated
- **Update**: fast

**TimeLine Curator**
- **Browse**: fast
- **Extract**: automated
- **Format**: automated
- **Show**: fast
- **Update**: fast
Previous Work

Unstructured Text

Structured Data

Vis Authoring Tools

**LeadLine, Jigsaw**

**LifeLines, Google News Timeline, TimelineJS**

**Tableau, Lyra**

**e.g. TERNIP**

**Structured Data**

<table>
<thead>
<tr>
<th>col1</th>
<th>col2</th>
<th>col3</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>birth</td>
<td>some more text</td>
</tr>
<tr>
<td>65</td>
<td>graduation</td>
<td>that belongs</td>
</tr>
<tr>
<td>74</td>
<td>first job</td>
<td>to the topic and</td>
</tr>
<tr>
<td>76</td>
<td>next job</td>
<td>was curated</td>
</tr>
<tr>
<td>80</td>
<td>discovery</td>
<td>very carefully to</td>
</tr>
<tr>
<td>90</td>
<td>nobel prize</td>
<td>give a lot of info</td>
</tr>
</tbody>
</table>

**TimeLineCurator**
Architecture

**Extract**

- Recognition
- Normalization

**Format**

- Expression
- Date
- Context

<table>
<thead>
<tr>
<th>expression</th>
<th>date</th>
<th>context</th>
</tr>
</thead>
<tbody>
<tr>
<td>yesterday</td>
<td>2014/10/30</td>
<td></td>
</tr>
<tr>
<td>last week</td>
<td>2014/10/25</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>annually</td>
<td>??</td>
<td></td>
</tr>
</tbody>
</table>

**Show**

- 2010
- 2011
- 2012
- 2013
- 2014
- 2015

**Curate/Update**

- 25.10.2014
- An interesting day
- And this is what happened that day...

- Present

- 2010
- 2011
- 2012
- 2013
- 2014
- 2015
Validation

• “Into the wild” feedback
  • Used for actual projects
    \textit{Deaf churches in England, company profile}
  • Promoted by Nick Diakopoulos at NICAR15
  • Mentioned in “Making Timelines“ from Lena Groeger (ProPublica)

• Ideas for further development
  • NLP-related: \textit{Including B.C. dates, Including other languages}
  • Integration with existing tools and workflows: e.g. \textit{Overview}
for analyzing large sets of documents: www.overviewdocs.com
Validation

• **User Experience Comparison**
  - Comparing TimelineJS approach with TLC
  - All participants strongly preferred visual authoring environment

• **Solicited potential users**
  - semi-structured interviews with 8 potential users
  - 7 journalists, 1 policy researcher

For the less geeky journalists who might be scared of timelines, this is a brilliant super-easy way to see what it might look like.

[A step to] break the barrier between the artiste writer and the data journalist.
Discussion

• Analysis use case turned out to be popular

• “Human in the loop” needed for curation and deciding what’s important and interesting
Interactive Authoring of Visual Timelines from Unstructured Text

Project page about.TimeLineCurator.org
Example gallery gallery.TimeLineCurator.org

Thanks to Francesco Escalona and Joshua Romero, Chad Skelton and Nick Diakopoulos, Michelle Borkin, Anamaria Crisan, Enamul Hoque, Sung-Hee Kim, and Narges Mahyar