SHORTREF.ORG
Making URLs Easy-to-Cite

Jozef Mišutka
Background

- purpose and advantages of using PIDs
- in our context, PIDs are used for citing and referencing of articles and data
Motivation

- writing a research paper and want to reference/cite
  - a service used to query data
  - a service used to visualise data
  - long and unreadable URL
• is converting visualisation to picture always reliable and sufficient?

https://lindat.mff.cuni.cz/services/pmltq/#!/treebank/hamledt_cs/query/IYWgdg9gJgpIBAbTgXSA/result/svg
Issues with URLs

- decay
- not enough information
  - version of software, data
- return different content in time
- too long (not only formatting issue)
- proper citing/referencing
shortref.org solution

• instead of the URL use a **persistent identifier - handle** from shortref.org
  
  • can be updated to new target location
  
  • contains additional metadata
  
  • is monitored
Solution for users

- go to http://shortref.org
- fill out target url and metadata
- use the PID and save the token
Solution for services

- use REST API provided by shortref.org
- fill out target url and metadata
- return the PID to user and save the token
Summary

1. Creation
   - User
     - creates short URLs
   - shortref.org
     - Stores metadata
   - Handle Server

2. Monitoring
   - Icinga Monitoring Server
     - check short handles
   - Handle Server
     - Mark faulty
     - Error 4xx/5xx
     - Success
     - Mark OK

3. Resolution
   - User
     - redirects to resource
   - Handle Server
     - access working handle
       - returns Error page with metadata
     - access non-working handle
   - User
Persistency

- secret token can be used to update target URL
- keeping history of changes
  - internally
Monitoring

- all handles monitored from icinga
- not reachable or 4xx/5xx (except 401) error code returned for N consecutive times
- changed target URL to http://hdl.handle.net/11346/SHORTREF-PR6O#hdl =XXX
- http://hdl.handle.net/11346/fff-TN7H
The URL **http://ffff** your handle resolves to is **down since 2016-08-24T07:53:19+02:00**. Until it comes up again, use the information provided below to gather details about what this URL was.

<table>
<thead>
<tr>
<th>TITLE</th>
<th>fff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repository</td>
<td>LINDAT/CLARIN digital library at Institute of Formal and Applied Linguistics, Charles University in Prague</td>
</tr>
<tr>
<td>Submit date</td>
<td>2016-08-24T13:16:09Z</td>
</tr>
<tr>
<td>ORTEMAIL</td>
<td>fff</td>
</tr>
</tbody>
</table>

In case you have any question about why you are seeing this page please contact our support.

Powered by LINDAT/CLARIN

Copyright (c) 2015 UFAL MFF UK. All rights reserved.
Sustainability

- dedicated prefix - **11346**
- can be completely migrated from our infrastructure with zero downtime
• shortref.org
  • [https://github.com/ufal/lindat-shortener](https://github.com/ufal/lindat-shortener)
• handle-server
  • bundled into lindat-dspace as
    • "storage_type" = "CUSTOM"
    • "storage_class" = "org.dspace.handle.HandlePlugin"
  • [http://hdl.handle.net/11346/IRYS](http://hdl.handle.net/11346/IRYS)
• monitoring
  • [https://github.com/ufal/lindat-monitoring](https://github.com/ufal/lindat-monitoring)
  • [http://hdl.handle.net/11346/CCKR](http://hdl.handle.net/11346/CCKR)
Future work

- define the minimal metadata schema precisely
- citing format exports for handles based on their metadata
  - crossref for handles
- standalone library or handle server REST API
Thank you

Q&A