Exploring Challenges in Embedding Metadata of Licence Information in Digital Work

Jing Liu - gusjingli@student.gu.se
Bruce Haoqing Yinhe - gusyinha@student.gu.se
1. Introduction
1.1 Research Questions
2. Theoretical Background
3. Methodology
4. Results
5. Discussion
6. Recommendations
Introduction

Key words:

Digital work,
Attribution of copyright,
Licence,
Metadata
1. Introduction
   1.1 Research Questions
2. Theoretical Background
3. Methodology
4. Results
5. Discussion
6. Recommendations
Research Questions:

● RQ1 - What are the technological and social challenges of building an end-user tool, which can embed metadata of licence information in digital work?

● RQ2 - What are the suggestions to address these challenges?
Introduction

Identifying and tracking the metadata:

1. “Which author is it?”
2. “How to attribute this author?”
3. “Under which licence was this work released?”
1. Introduction
2. Theoretical Background
3. Methodology
4. Results
5. Discussion
6. Recommendations
Theoretical Background

- **Licences**

- **Rights Expression Language (REL)**
  
  E.g.
  
  - Creative Commons Right Expression Language (ccREL)
  
  - Open Digital Rights Language (ODRL)

- **Digital Rights Management (DRM)**

  - Digital Rights Enforcement (DRE)
1. Introduction
2. Theoretical Background
3. Methodology
4. Results
5. Discussion
6. Recommendations
Methodology

● Qualitative case study

● Five semi-structured interviews
Methodology

- Sound recording
- Transcription
- Thematic analysis
- Prototype

Fig. 1 Modified version of data collection and analysis process (Runeson and Höst, 2009)
Methodology

- **Expert group (3 ppl.)**
  - ODRL Community Group
  - Creative Commons
  - Europeana

- **End user group (2 ppl.)**
  - Independent photographers
1. Introduction
2. Theoretical Background
3. Methodology
4. Results
5. Discussion
6. Recommendations
Results

● **7 Challenges**
  ○ The potential challenges in embedding metadata of licence information in digital work from the **experts’ perspectives**.

● **7 Problems**
  ○ The practical problems encountered in the current way in which attribution is handled from the **potential end users’ perspectives**.

● **11 Suggestions**
  ○ The suggestions to address such potential challenges and practical problems, given by **both experts and end users**.
### Fig. 2 Mapping of Suggestions to Challenges/Problems

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C1. Increasing end-users’ awareness</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2. Increasing the adoption of the end-user tool</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3. Integrating the end-user tool with existing</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>software</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4. Embedding licences with too many constraints</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5. Attributing multiple right holders in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>composite work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6. Rework effort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7. Verifying the authenticity of the source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>metadata</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1. Informal attribution standard</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2. Users not providing attribution</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3. Low awareness of licensing</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4. Lack of attribution support in CMS</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P5. Issues using a watermark</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P6. Issues keeping track the attribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P7. Verifying the authenticity of the source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>metadata</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1. Introduction
2. Theoretical Background
3. Methodology
4. Results
   4.1 Challenges
   4.2 Problems
   4.3 Suggestions
5. Discussion
6. Recommendations
Challenges

- C1. Increasing end-users' awareness
- C2. Increasing the adoption of the end-user tool
- C3. Integrating the end-user tool with existing software
Challenges

- C4. Embedding licences with too many constraints
- C5. Attributing multiple right holders in composite work
- C6. Rework effort
- C7. Verifying the authenticity of the source metadata
1. Introduction
2. Theoretical Background
3. Methodology
4. Results
   4.1 Challenges
   4.2 Problems
   4.3 Suggestions
5. Discussion
6. Recommendations
Problems

- P1. Informal attribution standard
- P2. Users not giving attribution
- P3. Low awareness of licensing
- P4. Lack of attribution support in CMS
Problems

- P5. Issues using a watermark
- P6. Issues keeping track the attribution
- P7. Verifying the authenticity of the source metadata
1. Introduction
2. Theoretical Background
3. Methodology
4. Results
   4.1 Challenges
   4.2 Problems
   4.3 Suggestions
5. Discussion
6. Recommendations
Suggestions

- S1. Embedding metadata
- S2. Reusing existing tools
- S3. Integrating with CMS
Suggestions

● S4. Improving the user experience using icons

● S5. Using GUI pop-ups

● S6. Embedding during content creation
Suggestions

- S7. Verifying the authenticity of the source metadata
- S8. Resolving conflicts in composite work
- S9. Storing metadata separately
Suggestions

● S10. DRM issues

● S11. Use simple constraints
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C1. Increasing end-users’ awareness</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2. Increasing the adoption of the end-user tool</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3. Integrating the end-user tool with existing software</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4. Embedding licences with too many constraints</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5. Attributing multiple right holders in composite work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6. Rework effort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7. Verifying the authenticity of the source metadata</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1. Informal attribution standard</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2. Users not providing attribution</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3. Low awareness of licensing</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4. Lack of attribution support in CMS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P5. Issues using a watermark</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P6. Issues keeping track the attribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P7. Verifying the authenticity of the source metadata</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2 Mapping of Suggestions to Challenges/Problems
Results

Software Requirements

RQ1. The application should be able to read *.xmp file into XMP Meta objects.
RQ2. The application should be able to embed XMP Meta objects into images (preliminarily *.jpg and *.png).
RQ3. The application should be able to read previously embedded XMP metadata in the image file.
RQ4. The application should support ccREL syntax.
RQ5. The application should be based on Adobe XMP Toolkit SDK.
RQ6. The application should be able to run locally (offline).
RQ7. The application should integrate with WordPress as a plug-in.
RQ8. The application should use icons.
RQ9. The application should use GUI pop-ups.
RQ10. The application should support embedding licence at content creation.
RQ11. The application should have a trust mechanism that verifies the authenticity of the source metadata.
RQ12. The application should resolve licence conflicts in composite work.
RQ13. The application should be implemented in a DRM-free environment.
RQ14. The application should avoid complex constraints on rights statements.
RQ1. The application should be able to read *.xmp file into XMP Meta objects.
RQ2. The application should be able to embed XMP Meta objects into images (preliminarily *.jpg and *.png).
RQ3. The application should be able to read previously embedded XMP metadata in the image file.
RQ4. The application should support ccREL syntax.
RQ5. The application should be based on Adobe XMP Toolkit SDK.
RQ6. The application should be able to run locally (offline).
RQ7. The application should integrate with WordPress as a plug-in.
RQ8. The application should use icons.
RQ9. The application should use GUI pop-ups.
RQ10. The application should support embedding licence at content creation.
RQ11. The application should have a trust mechanism that verifies the authenticity of the source metadata.
RQ12. The application should resolve licence conflicts in composite work.
RQ13. The application should be implemented in a DRM-free environment.
RQ14. The application should avoid complex constraints on rights statements.
Results Prototype

Licence: CC BY-SA 2.5 SE (http://creativecommons.org/licenses/by-sa/2.5/se/)
Title of work: TheTitle
Attribute work to name: TheAuthor
Attribute work to URL: http://theurl
More permissions URL: http://permissionurl
Results

Prototype

<?xpacket begin='' id=''?>
<x:xmpmeta xmlns:x='adobe:ns:meta/'>
  <rdf:RDF xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#'>
    <rdf:Description rdf:about=' ' xmlns:xmpRights='http://ns.adobe.com/xap/1.0/rights/'>
      ...  
      <xmpRights:UsageTerms>
        <rdf:Alt>
          <rdf:li xml:lang='x-default' >This work is licensed under a &lt;a rel=&#34;license&#34; href=&#34;http://creativecommons.org/licenses/by-sa/2.5/se/#34;&gt;Creative Commons Attribution-ShareAlike 2.5 Sweden License&lt;/a&gt;.&lt;/rdf:li>
        ...  
      </rdf:Alt>
      </xmpRights:UsageTerms>
    </rdf:Description>
    <rdf:Description rdf:about=' ' xmlns:dc='http://purl.org/dc/elements/1.1/'>
      <dc:title>
        <rdf:Alt>
          <rdf:li xml:lang='x-default' >TheTitle</rdf:li>
        </rdf:Alt>
      </dc:title>
    </rdf:Description>
    <rdf:Description rdf:about=' ' xmlns:cc='http://creativecommons.org/ns#'>
      <cc:license rdf:resource='http://creativecommons.org/licenses/by-sa/2.5/se/'/>
      <cc:attributionName>TheAuthor</cc:attributionName>
      <cc:morePermissions rdf:resource='http://permissionurl1'/>
    </rdf:Description>
  </rdf:RDF>
</x:xmpmeta>
<?xpacket end='r'?>
Results

Prototype

Fig. 3 Mapping of Suggestions to Challenges/Problems
1. Introduction
2. Theoretical Background
3. Methodology
4. Results
5. Discussion
6. Recommendations
Discussion

- Embedding metadata improves user's experience of identifying and attributing the original author.

- The lack of awareness of licence and attribution is a fundamental challenge. Increasing user's awareness of licence and attribution is a long-term process.
Recommendations

To consider in future research:

● Explore more challenges and investigate further from wider perspectives.

● Build and evaluate a complete end-user tool.
Recommendations

Future prototype development is suggested to consider:

● Support authentication
● Support multi-licences in composite work
● Compare interpreting licence information with different REL standards
● Develop strategies to increase awareness for attribution
● Transform image-watermarking solutions to embedding metadata
● Promote usage by integration with mainstream social network CMS.
We would like to dedicate this paper to the memory of our academic supervisor: Lars Pareto, who had been a constant source of support and inspiration.

We would also like to thank:

Jonas Öberg,
Commons Machinery AB,
and the anonymous participants who volunteered their time for the interviews.
References in this presentation:


- Creative Commons http://creativecommons.org

- CcREL http://wiki.creativecommons.org/CC_REL

Thank you!

Jing Liu - gusjingli@student.gu.se
Bruce Haoqing Yinhe - gusyinha@student.gu.se