Trusting semi-structured web data
Ceolin, D.

2014

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 07. Feb. 2021
Contents

Preface vii

Introduction 1

Context 1
State of the Art 3
Research Questions and Thesis Statement 8
Contributions 12
Publications 13
Thesis Structure 15

Preliminaries 17

Introduction 17
Subjective Logic 18
Subjective Logic Operators for Combining Opinions 27
Conjugate Priors 32
Dirichlet Process 33
Semantic Similarity Measures 35

I Using Web Data to Make Trust Assessments 39

1 Estimating Trust in Annotations using Web Data 40
1.1 Introduction 40
1.2 Approach and Related Work 41
1.3 Trust Model 41
1.4 Case Study: Naturalis Data 46
1.5 Discussion 53
1.6 Conclusion 54

2 Estimating Trust in Confidence Heuristics 55
2.1 Introduction 55
2.2 Georeferencing Challenges 56
2.3 Related Work 57
II Uncertainty Reasoning for Assessing Trust

3 Uncertainty Reasoning for Handling Web Data
3.1 Introduction .......................... 77
3.2 Scope of this Chapter .................. 78
3.3 Modeling Categorical Web Data .......... 79
3.4 Subjective Logic Extensions for the Semantic Web . . . 80
3.5 Discussion ................................ 86
3.6 Conclusion .............................. 96

4 Reliability Analyses of Web Data
4.1 Introduction .......................... 100
4.2 Related Work .......................... 101
4.3 Comparing Closed and Open Data .......... 102
4.4 Analyzing Open Data .................... 104
4.5 Case Study - Police Open Data Analyses . . . . 107
4.6 Conclusions ............................. 119

III Provenance Analyses for Assessing Trust

5 Provenance Analyses for Trust Assessment .......................... 125
5.1 Introduction ................................ 126
5.2 Related Work .......................... 127
5.3 Mapping SEM and OPM .................... 127
5.4 Subjective Logic for Trusting AIS Messages .......... 130
5.5 Estimating the Trust in AIS Messages Using Provenance . . . 131
5.6 Algorithm Application .......................... 135
5.7 Results .............................. 137
5.8 Discussion ............................. 139
5.9 Conclusion .............................. 140

6 Combining Provenance with User Reputation for Trust Assessment 141
6.1 Introduction .......................... 142
6.2 The Waisda? Dataset .................... 143
### IV Semantic Similarity to Improve Trust Estimation

#### 7 Assessing Annotation Trustworthiness Using Semantic Similarity

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1 Introduction</td>
<td>158</td>
</tr>
<tr>
<td>7.2 Related Work</td>
<td>159</td>
</tr>
<tr>
<td>7.3 System Description</td>
<td>159</td>
</tr>
<tr>
<td>7.4 Evaluation</td>
<td>166</td>
</tr>
<tr>
<td>7.5 Discussion</td>
<td>171</td>
</tr>
<tr>
<td>7.6 Algorithm - Interactive Version</td>
<td>173</td>
</tr>
<tr>
<td>7.7 Conclusion</td>
<td>176</td>
</tr>
</tbody>
</table>

#### 8 Provenance-based Assessment of Annotations Trustworthiness

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Introduction</td>
<td>177</td>
</tr>
<tr>
<td>8.2 System Description</td>
<td>178</td>
</tr>
<tr>
<td>8.3 Evaluation</td>
<td>185</td>
</tr>
<tr>
<td>8.4 Conclusion</td>
<td>189</td>
</tr>
</tbody>
</table>

**Conclusion and Discussion**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Research Questions Revisited</td>
<td>192</td>
</tr>
<tr>
<td>Discussion and Future Work</td>
<td>198</td>
</tr>
</tbody>
</table>

**Bibliography**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>205</td>
</tr>
</tbody>
</table>

**Summary**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>221</td>
</tr>
</tbody>
</table>

**Samenvatting**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>225</td>
</tr>
</tbody>
</table>

**SIKS Dissertatiereks**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>229</td>
</tr>
</tbody>
</table>