UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Ontology enrichment from heterogeneous sources on the web

de Boer, V.

Publication date
2010

Link to publication

Citation for published version (APA):
de Boer, V. (2010). *Ontology enrichment from heterogeneous sources on the web*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:07 Oct 2026

CONTENTS

PREFACE vii

1 INTRODUCTION 1
- 1.1 Background . . . 1
- 1.2 Semi-automatic Ontology Enrichment . . . 2
  - 1.2.1 Ontology Learning, Population and Enrichment . . . 2
  - 1.2.2 The Information Extraction Task . . . 3
  - 1.2.3 Heterogeneity of the Web . . . 4
- 1.3 MultimediaN E-culture repositories and vocabularies . . . 6
- 1.4 Scalability issues . . . 6
- 1.5 Approach . . . 7
- 1.6 Research Questions . . . 8
- 1.7 Outline of the Thesis . . . 9

2 A REDUNDANCY-BASED METHOD FOR RELATION EXTRACTION 11
- 2.1 Introduction . . . 11
- 2.2 Relation Instantiation Task . . . 13
- 2.3 Redundancy-Based Relation Instantiation . . . 14
  - 2.3.1 Approach . . . 14
  - 2.3.2 Method Specification . . . 14
- 2.4 Extracting Artist-Art Style Relation . . . 17
  - 2.4.1 Cultural Heritage domain . . . 17
  - 2.4.2 Basic Method Experiments . . . 18
  - 2.4.3 Experiment 3: Iterative Experiment . . . 24
  - 2.4.4 Experiment 4: Ten Art Styles . . . 26
- 2.5 Experiments in a second domain: Football players . . . 30
  - 2.5.1 Task Description . . . 30
  - 2.5.2 Experiment Setup and Evaluation . . . 31
  - 2.5.3 Results . . . 31
- 2.6 Comparison to the Normalized Google Distance Method . . . 32
- 2.7 Related Work . . . 36
- 2.8 Conclusions and Further Research . . . 38

3 EXTRACTING HISTORICAL TIME PERIODS FROM THE WEB 41
- 3.1 Introduction . . . 41
- 3.2 Task and Method . . . 43
- 3.3 The Information Extraction Phase . . . 44
  - 3.3.1 Information Extraction method overview . . . 44
  - 3.3.2 Year extraction . . . 45
  - 3.3.3 Normalization . . . 45
  - 3.3.4 Model fitting . . . 46
  - 3.3.5 Experiments . . . 49
- 3.4 The Semantic Representation Phase . . . 57
  - 3.4.1 Discrete Start and End Dates . . . 58

3.4.2 Multiple discrete intervals . . . 58
3.4.3 Distribution . . . 60
3.4.4 Structured Vocabulary . . . 60
3.5 Related Work . . . 68
3.6 Conclusions . . . 69
4 ONTOLOGY ENRICHMENT USING A TEXT-ANALYSIS TOOL 75
4.1 Introduction . . . 75
4.2 Ontology Enrichment using patterns . . . 78
4.2.1 Task definition . . . 78
4.2.2 The tOKo tool and its pattern language . . . 78
4.2.3 Relation Instantiation using patterns . . . 80
4.3 Experiments . . . 85
4.3.1 Roman Gods . . . 86
4.3.2 Cultural Heritage Domain . . . 92
4.4 Conclusions and Future Work . . . 102
5 USING HETEROGENEOUS SOURCES TO IMPROVE ONTOLOGY ENRICHMENT 105
5.1 Introduction . . . 105
5.2 Problem definition . . . 106
5.3 Method description . . . 108
5.3.1 Using Background Knowledge . . . 109
5.3.2 The Information Integration step . . . 111
5.4 Experiments . . . 113
5.4.1 Roman Generals . . . 113
5.4.2 Art Style-Artist . . . 118
5.4.3 Regional Artists . . . 121
5.5 Related Work . . . 126
5.6 Conclusions . . . 127
6 CONCLUSIONS AND DISCUSSION 129
6.1 The Research Questions Revisited . . . 129
6.2 General Conclusions . . . 132
6.3 Discussion . . . 134

BIBLIOGRAPHY 137

SUMMARY 145

NEDERLANDSE SAMENVATTING 151

SIKS DISSERTATION SERIES 157