

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Problem Statement . . . . .	2
1.2	Hypothesis . . . . .	5
1.3	Research Goals . . . . .	6
1.3.1	General Goals . . . . .	6
1.3.2	Specific Goals . . . . .	6
1.4	Contributions . . . . .	6
1.5	Publications . . . . .	8
1.6	Thesis outline . . . . .	9
<b>2</b>	<b>Preliminaries</b>	<b>11</b>
2.1	Definition of “entity” . . . . .	12
2.2	EL Formalisms . . . . .	13
2.3	Resource Description Framework . . . . .	15
2.4	RDF query languages . . . . .	18
<b>3</b>	<b>Related Work</b>	<b>20</b>
3.1	Entity Recognition . . . . .	20
3.1.1	NERL strategy . . . . .	21
3.1.2	End-to-End strategy . . . . .	22
3.2	Entity Disambiguation . . . . .	23
3.2.1	Candidate Entity Generation . . . . .	23
3.2.2	Candidate Entity Ranking . . . . .	25
3.2.3	Unlinkable Mention Prediction . . . . .	27
3.3	Entity Linking Evaluation Measures . . . . .	28
3.4	EL Benchmark Datasets . . . . .	30
3.5	EL Formats . . . . .	31
3.6	EL Design Issues . . . . .	34
<b>4</b>	<b>Consensus about Entity Linking</b>	<b>38</b>
4.1	Questionnaire on the Goals of EL . . . . .	39
4.2	Proposed Solution . . . . .	43

<b>5</b>	<b>Multilingualism</b>	<b>45</b>
5.1	Multilingual EL Systems . . . . .	46
5.1.1	Multilingual EL Datasets . . . . .	47
5.1.2	Non-English Entity Linking: Spanish use-case . . . . .	49
5.2	The VoxEL Dataset . . . . .	54
5.3	Multilingual EL performance . . . . .	56
5.3.1	Why not translate to English? . . . . .	58
5.4	Translating across languages . . . . .	61
<b>6</b>	<b>Fine-Grained Entity Linking</b>	<b>64</b>
6.1	Fine-Grained Categories . . . . .	64
6.1.1	Base Form . . . . .	65
6.1.2	Part of Speech . . . . .	66
6.1.3	Overlap . . . . .	67
6.1.4	Reference . . . . .	67
6.2	Fine-Grained EL Format . . . . .	68
6.2.1	Vocabulary . . . . .	68
6.2.2	Extending NIF . . . . .	69
6.3	Fine-Grained Entity Annotation . . . . .	73
6.4	NIFify . . . . .	75
6.4.1	NIF Construction . . . . .	76
6.4.2	Validation . . . . .	77
6.4.3	Result Visualization . . . . .	79
6.5	Relabeling KORE50, ACE2004 and VoxEL . . . . .	81
6.6	Fine-Grained Evaluation . . . . .	83
6.7	Fuzzy Recall and $F_1$ Measures . . . . .	86
6.7.1	Fuzzy Framework . . . . .	86
6.7.2	Fuzzy Evaluation . . . . .	88
<b>7</b>	<b>Fine-Grained Entity Linking Systems</b>	<b>91</b>
7.1	Adding Coreference Resolution . . . . .	91
7.2	Adding Word Sense Disambiguation . . . . .	92
7.3	Combined CR and WSD Results . . . . .	94
<b>8</b>	<b>Conclusions and Future Work</b>	<b>97</b>
8.1	Contributions and results . . . . .	98
8.2	Limitations and Future Work . . . . .	99
8.3	Outlook . . . . .	102
	<b>Bibliography</b>	<b>104</b>