


INFORMAZIONI PERSONALI

Benotto Giulia

 Via Argine Vecchio 219, 56019 Vecchiano (Italia)

 3494295170

 g.benotto@me.com

Sesso Femminile | Data di nascita 22/07/1983 | Nazionalità Italiana

ESPERIENZA
PROFESSIONALE

10/2014–alla data attuale

Assegnista di Ricerca

ILC CNR, Pisa (Italia)

Attualmente collaboro allo sviluppo del Progetto TALMUD, un'applicazione web collaborativa per la traduzione del Talmud Babilonese in italiano. Sto svolgendo attività di ricerca e sviluppo nel settore della rappresentazione della conoscenza e dei sistemi di supporto alle decisioni per l'interpretazione di testi letterari.

07/2013–07/2014

Assegnista di Ricerca

IIT CNR, Pisa (Italia)

Nel corso di questa esperienza ho lavorato su un progetto riguardante l'estrazione di informazioni turistiche da testi, focalizzandomi in particolare sull'estrazione di relazioni semantiche e cercando di integrarvi il più possibile il mio lavoro di dottorato riguardante l'estrazione automatica e non supervisionata di relazioni semantiche paradigmatiche. Ho lavorato anche all'annotazione di dati da utilizzare per compiti di Named Entity Recognition.

11/2011–06/2013

Assegnista di Ricerca

Dipartimento di Informatica, Università di Pisa, Pisa (Italia)

Lo scopo principale del progetto riguardava la costruzione di una tassonomia di termini estratti da testi italiani che avevano per argomento l'arte, la linguistica e la letteratura. Il mio lavoro ha riguardato principalmente l'annotazione semantica di dati per l'implementazione di metodi di Named Entity Recognition supervisionati.

01/2010–11/2011

Analista nell'area Search and Advertising

Wind - Italia On Line, Pisa (Italia)

Nel corso di questa esperienza ho lavorato principalmente al progetto 24x7. 24x7 è un aggregatore di notizie che integra informazioni riguardanti le entità nominate presenti in ogni articolo, e si focalizza sul tema principale trattato in ciascun articolo. Mi sono occupata di annotare ed etichettare manualmente i dati estratti dagli articoli, utilizzando *features* sintattiche e semantiche, allo scopo di creare ed estendere la base di conoscenza contenente tutte le entità nominate.

08/2009–12/2009

Collaboratrice alla Ricerca

CELI Language and Information Technology, Torino (Italia)

Il mio lavoro ha riguardato principalmente l'estrazione di informazione. Mi sono occupata, in particolare, di studiare lo stato dell'arte riguardante paradigmi di estrazione di informazione e ricercarne di nuovi.

08/2007–12/2007

Stagista

N.U.R.C. Nato Undersea Research Center, La Spezia (Italia)

Nel corso della mia seconda esperienza presso il N.U.R.C., il mio scopo era quello di costruire una tassonomia di tutte le attività in corso al N.U.R.C. e di tutti gli strumenti utilizzati. A questo scopo, interagendo con gli esperti, ho dovuto sviluppare una profonda conoscenza del dominio. Inoltre, mi sono occupata dell'estrazione di informazione e dati da testi (principalmente pubblicazioni prodotte dal N.U.R.C.), della loro annotazione e revisione, in modo da rendere tali dati utilizzabili per la costruzione della tassonomia.

07/2006–09/2006 Webmaster

N.U.R.C. Nato Undersea Research Center, La Spezia (Italia)

Nel corso di questa esperienza lavorativa ho sviluppato il sito web interno (intranet) e quello esterno del N.U.R.C.

03/2005–08/2005 Stagista

ISTI CNR, Pisa (Italia)

Nel corso del mio stage ho lavorato allo sviluppo di un'ontologia nel dominio medico. Per fare questo, ho raccolto e annotato dati da corpora medici (principalmente cartelle cliniche dei reparti di cardiologia), e ho estratto le relazioni che intercorrevano tra tali dati per costruire manualmente l'ontologia.

ISTRUZIONE E FORMAZIONE

2012–alla data attuale

Dottorato di Ricerca in Discipline Umanistiche, Linguistica Generale, Storica, Applicata, Computazionale e delle Lingue Moderne

Università di Pisa, Pisa (Italia)

Il progetto di ricerca al quale sto lavorando riguarda l'esplorazione della capacità dei modelli distribuzionali riguardo all'estrazione e alla classificazione delle relazioni semantiche paradigmatiche da testi, con particolare attenzione alla relazione di iperonimia e alla relazione di antonimia. Il mio scopo è quello di sviluppare un metodo totalmente non supervisionato per l'estrazione e la classificazione di tali relazioni. Nel corso del mio dottorato ho già sviluppato e testato una misura per la classificazione della relazione di iperonimia che ha performance migliori rispetto alle misure allo stato dell'arte attuale. Al momento sto lavorando a un metodo non supervisionato per la classificazione della relazione di antonimia utilizzando approcci distribuzionali.

01/2006–06/2009 Laurea Specialistica in Tecnologie del Linguaggio, Informatica Umanistica

Università di Pisa, Pisa (Italia)

Durante il mio percorso di Specialistica, il mio interesse si è concentrato sulla Linguistica Computazionale, gestione della conoscenza, linguistica e text mining. Per la mia Tesi finale ho lavorato sotto la supervisione del professor Alessandro Lenci. Mi sono occupata di estrazione di relazioni semantiche, utilizzando un metodo semi supervisionato. Il titolo del mio lavoro è "*Semantic Relation Extraction and classification: Experiments on Wikipedia.it*". Il mio voto finale è stato 110/110 e Lode.

10/2002–12/2005 Laurea Triennale in Informatica Umanistica

Università di Pisa, Pisa (Italia)

Lo scopo principale del mio corso di Laurea triennale era imparare come trattare dati di tipo umanistico utilizzando tecnologie informatiche. A tale scopo, ho seguito corsi che spaziavano dalla linguistica generale, linguistica italiana e computazionale alla programmazione, web design e data mining. Ho da subito focalizzato il mio interesse nell'ambito della linguistica computazionale e nella rappresentazione della conoscenza. La mia Tesi triennale ha quindi riguardato la costruzione di un'ontologia medica. Il titolo era "*Ontologie Mediche*" ed è stata sviluppata sotto la supervisione del dottor Amedeo Cappelli. Il voto finale è stato di 110/110 e Lode.

09/1997–07/2002 Maturità Scientifica

Liceo Scientifico Statale Ulisse Dini, Pisa (Italia)

Il mio voto di maturità è stato di 90/100.

COMPETENZE PERSONALI

Lingua madre Italiano

Altre lingue	COMPRESIONE		PARLATO		PRODUZIONE SCRITTA
	Ascolto	Lettura	Interazione	Produzione orale	
inglese	C1	C1	C1	C1	C1
First Certificate of English, University of Cambridge ESOL Examinations C1					
francese	A2	B1	A1	A1	A1

Livelli: A1/A2: Livello base - B1/B2: Livello intermedio - C1/C2: Livello avanzato

[Quadro Comune Europeo di Riferimento delle Lingue](#)

Competenze informatiche

Capacità nella Programmazione:

- Conoscenza molto buona di **Perl** e **Espressioni Regolari**
- Conoscenza molto buona di **MySQL**
- Conoscenza buona di **OWL** e **RDF**
- Conoscenza molto buona di **UNIX shell scripting**
- Conoscenza di base di **Python** e **Java**

Programmazione Web:

- Conoscenza buona di **html**, **xhtml** e **css**

Applicativi:

- Conoscenza molto buona di **TeX (Latex, Bibtex, PSTricks)**
- Conoscenza molto buona di suite per l'ufficio (**Microsoft Office, OpenOffice, iWork**)
- Buona conoscenza di applicazioni per il data mining quali **Clementine** e **Weka** e di suite MySQL grafiche, quali **Navicat** e **Heidi SQL**. Ho anche una buona conoscenza del software per lo sviluppo di ontologie **Protégé**
- Ho una conoscenza molto buona dei seguenti sistemi operativi: **Windows 3.x, Windows 9x, Windows Xp, MacOS 10.x, Linux (Ubuntu - Debian)**

ULTERIORI INFORMAZIONI

Riconoscimenti e premi

Vincitrice del PREMIO ANTONIO ZAMPOLLI 2010 come miglior tesi di Laurea specialistica avente per argomento fenomeni di Elaborazione del Linguaggio Naturale (FORUM TAL 2010 - Roma, 29 Ottobre 2010) per la Tesi "Semantic Relation Extraction and classification: Experiments on wikipedia.it"

Corsi di Formazione

Dal 13 al 16 Luglio 2011 ho frequentato la UCREL Summer School in Corpus Linguistics, tenutasi presso l'Università di Lancaster, allo scopo di migliorare la mia conoscenza riguardo alla ricerca basata su grandi corpora testuali e alle tecniche da utilizzare.

Pubblicazioni

- Davide Albanesi, Andrea Bellandi, Giulia Benotto, Emiliano Giovannetti (2015). Translation, Annotation and Knowledge Modelling of the Babyolian Talmud: the Traduco System, To Appear in *Proceedings of the Digital Humanities Conference 2015 (DH 2015)*.
- Alessia Bellusci, Andrea Bellandi, Giulia Benotto, Amedeo Cappelli, Emiliano Giovannetti, Simone Marchi (2014), "Towards a Decision Support System for Text Interpretation", in *Proceedings of the First Italian Conference on Computational Linguistics CLIC-it 2014 & the Fourth International Workshop EVALITA 2014*, Pisa, 9th-10th December 2014 (ISBN:978-886741-472-7):58-62.

- Giulia Benotto (2013), "Modelli distribuzionali delle relazioni semantiche: il caso dell'iperonimia", in *Atti della conferenza CODISCO*, Roma, 24th-26th September 2012 (ISBN: 978-88-98138-06-7): 85-99.
- Alessandro Lenci, Giulia Benotto (2012), "Identifying hypemyms in distributional semantic spaces", in *Proceedings of *SEM 2012: The First Joint Conference on Lexical and Computational Semantics*, Volume 1, Montreal, 7th-8th June 2012 (ISBN: 978-1-937284-21-3): 75-79.
- Giulia Benotto (2010), "Semantic relation extraction and classification. Experiments on wikipedia.it", in *Atti della conferenza TAL 2010 - Il TAL come motore per l'innovazione*, Roma, 28th-29th October 2010.
- Giulia Benotto, Sara Bianchi, Sarah Corsini (2009), "Comunicazione Reale e Virtuale nel 'Villaggio Globale'", in *Parlare, Comprendersi, Interagire* a cura di Daria Carmina Coppola (ISBN 978-88-6019-237-0).