

# Valerio Basile

aggiornato il 16/4/2018

Sito Internet: <http://valeribasile.github.io/>

Linkedin: <https://www.linkedin.com/in/valeribasile>

Github: <https://github.com/valeribasile/>

## Istruzione

### Istruzione principale

- PhD presso Università di Groningen (Dicembre 2015)
- Laurea magistrale in Informatica presso Università di Bologna (Marzo 2010)
- Laurea triennale in Informatica presso Università di Bologna (Ottobre 2008)

### Istruzione ulteriore

- Corso online Coursera *Computing for Data Analysis* (2012)
- Corso online Coursera *Machine Learning* (2012)
- European Summer School in Logic, Language and Information (ESSLI 2011)
- Corso per PhD su Computational Ontologies presso l'Università di Bologna (Febbraio 2010)

## Esperienza lavorativa

Al momento occupo una posizione di ricerca postdoc presso il dipartimento di Informatica dell'Università di Torino.

Precedentemente, ho svolto attività di ricerca come postdoc presso Sapienza Università di Roma, finanziato dal progetto ERC MOUSSE, occupandomi di parsing semantico del linguaggio naturale in molteplici lingue, e Inria Sophia Antipolis Méditerranée, finanziato dal progetto europeo CHIST-ERA **ALOOF (Autonomous Learning of the Meaning of Object<sup>1</sup>)**. In ALOOF, il mio lavoro ha riguardato l'estrazione di conoscenza di senso comune da testo in linguaggio naturale, come ad esempio proprietà e relazioni tra oggetti domestici. La procedura di estrazione della conoscenza che ho sviluppato coinvolge **parsing semantico del linguaggio naturale, disambiguazione lessicale, entity linking, modello vettoriali di parole e concetti, rappresentazione della conoscenza in linked open data, e ragionamento automatico basato su RDF**. ALOOF è una collaborazione internazionale con gruppi di ricerca attivi in aree quali robotica e computer vision.

Ho ottenuto il dottorato di ricerca presso l'Università di Groningen, Paesi Bassi, sotto la supervisione del Prof. Johan Bos (**Semantica Computazionale**), all'interno del progetto *D-MAP: the Deep Meaning Annotation Project*, dal 1/1/2011 al 30/6/2015. Durante il dottorato ho lavorato principalmente nel campo della Semantica Computazionale. Ho partecipato alla creazione del **Groningen Meaning Bank<sup>2</sup>**, una raccolta estesa di testi annotati con informazione semantica, e di **Wordrobe<sup>3</sup>**, una collezione di giochi di linguistica. Per la mia tesi di

---

1 <http://project.inria.fr/aloof>

2 <http://gmb.let.rug.nl/>

3 <http://www.wordrobe.org/>

dottorato, ho sviluppato un sistema in grado di generare espressioni in linguaggio naturale a partire da formule logiche (**Generazione del Linguaggio Naturale**), supportato da un algoritmo originale per la scelta degli elementi lessicali.

Precedentemente (1/4/2010 – 31/12/2010) ho lavorato con un assegno di ricerca presso il CIRSFD, Università di Bologna, nel campo della Informatica Legale.

Ho svolto ricerca nell'ambito dell'**analisi del sentimento**, in collaborazione con le Università di Bologna, Torino e Valencia. Ho creato *TWITA*<sup>4</sup>, la raccolta su grande scala di messaggi di Twitter in lingua italiana, il primo approccio all'analisi del sentimento sull'Italiano, ed all'organizzazione della prima campagna di valutazione di Sentiment Polarity Classification sulla lingua italiana.

## Pubblicazioni

### Estrazione della conoscenza

- S. Jebbara, **V. Basile**, E. Cabrio , P. Cimiano: *Distributional and Neural Models for Extracting Manipulation-Relevant Relations from Text Corpora*. Semantic Web Journal (2017)
- **V. Basile**, S. Jebbara, E. Cabrio , P. Cimiano: *Populating a Knowledge Base with Object-Location Relations Using Distributional Semantics*. Proceedings of the 20th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2016)
- **V. Basile**, E. Cabrio , F. Gandon: *Building a General Knowledge Base of Physical Objects for Robots*. Proceedings of the 13th European Semantic Web Conference, poster (ESWC 2016)

### Robotica e percezione

- J. Young, **V. Basile**, M. Suchi, L. Kunze, N. Hawes, M. Vincze, B. Caputo: *Making sense of indoor spaces using semantic web mining and situated robot perception*. Proceedings of the Extended Semantic Web Conference (ESWC 2017)
- J. Young, L. Kunze, **V. Basile**, E. Cabrio, N. Hawes, B. Caputo: *Semantic Web-Mining and Deep Vision for Lifelong Object Discovery*. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2017)
- J. Young, **V. Basile**, L. Kunze, E. Cabrio, N. Hawes. *Towards Lifelong Object Learning by Integrating Situated Robot Perception and Semantic Web Mining*. Proceedings of the European Conference on Artificial Intelligence (ECAI 2016)

### Semantica Computazionale

- J. Bos, **V. Basile**, K. Evang, N. J Venhuizen, J. Bjerva: *The Groningen MeaningBank*. Book chapter in "Handbook of Linguistic Annotation" (2017)
- **V. Basile**, E. Cabrio , C. Schon: *KNEWS: Using Logical and Lexical Semantics to Extract Knowledge from Natural Language*. Proceedings of the European Conference on Artificial Intelligence (ECAI demo 2016)
- **V. Basile**, J. Bos, K. Evang, N. Venhuizen: *Ugroningen: Negation detection with discourse representation structures*. Proceedings of the First Joint Conference on Lexical and Computational Semantics (\*SEM 2013)
- N. Venhuizen, **V. Basile**, K. Evang, J. Bos: *Gamification for word sense labeling*. Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)
- **V. Basile**, J. Bos, K. Evang, N. Venhuizen: *Developing a large semantically annotated Corpus*. Proceedings of the eighth international conference on Language Resources and Evaluation (LREC 2012)
- **V. Basile**, J. Bos, K. Evang, N. Venhuizen: *A platform for collaborative semantic annotation*. Proceedings

---

4 <http://valeribasile.github.io/twita>

of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2012)

### Generazione del Linguaggio Naturale

- **V. Basile**: *A Repository of Frame Instance Lexicalizations for Generation*. Proceedings of the 2nd International Workshop on Natural Language Generation and the Semantic Web (WebNLG 2016)
- **V. Basile**: *From Logic to Language: Natural Language Generation from Logical Forms*. PhD Thesis (2015)
- **V. Basile**: *A Lesk-inspired Unsupervised Algorithm for Lexical Choice from WordNet Synsets*. First Italian Conference on Computational Linguistics (CLiC-it 2014), 48
- **V. Basile**, J. Bos: *Aligning Formal Meaning Representations with Surface Strings for Wide-coverage Text Generation*. Proceedings of the 14th European Workshop on Natural Language Generation (ENLG 2013), 1
- **V. Basile**, J. Bos: *Towards generating text from discourse representation structures*. Proceedings of the 13th European Workshop on Natural Language Generation (ENLG 2011), 145-150

### Analisi del Sentimento

- **V. Basile**, N. Novielli, D. Croce, F. Barbieri, M. Nissim, V. Patti: *Sentiment Polarity Classification at EVALITA: Lessons Learned and Open Challenges*. Transactions on Affective Computing (submitted, 2018)
- P. Basile, **V. Basile**, M. Nissim, N. Novielli, V. Patti: *Sentiment Analysis of Microblogging Data*. Chapter in the Encyclopedia of Social Network Analysis and Mining (2017)
- F. Barbieri, **V. Basile**, D. Croce, M. Nissim, N. Novielli, V. Patti: *Overview of the EVALITA 2016 SENTiment POLarity Classification Task*. Proceedings of Third Italian Conference on Computational Linguistics (CLiC-it 2016)
- P. Basile, **V. Basile**, M. Nissim, N. Novielli: *Deep tweets: from entity linking to sentiment analysis*. Proceedings of the Italian Computational Linguistics Conference (CLiC-it 2015)
- G. Attardi, **V. Basile**, C. Bosco, T. Caselli, F. Dell'Orletta, S. Montemagni, V. Patti, M. Simi, R. Sprugnoli: *State of the Art Language Technology for Italian Text: Problems, Solutions and Applicative Impact*. Special issue of Intelligenza Artificiale on "AI for Society and Economy" (2015)
- **V. Basile**, A. Bolioli, M. Nissim, V. Patti, P. Rosso: *Overview of the Evalita 2014 SENTiment POLarity Classification Task*. Proceedings of the 4th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA 2014)
- **V. Basile**, M. Nissim: *Sentiment analysis on Italian tweets*. Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2013)

### Informatica Legale e Argumentation Mining

- P. Basile, **V. Basile**, E. Cabrio, S. Villata: *Argument Mining on Italian News Blogs*. Proceedings of the Third Italian Conference on Computational Linguistics (CLiC-it 2016)
- **V. Basile**, E. Cabrio, S. Villata, C. Frasson, F. Gandon: *A Pragma-Semantic Analysis of the Emotion/Sentiment Relation in Debates*. Proceedings of the 4th International Workshop on Artificial Intelligence and Cognition (AIC 2016)
- A. Wyner, J. Bos, **V. Basile**, P. Quaresma: *An Empirical Approach to the Semantic Representation of Laws*. Proceedings of the The 25th International Conference on Legal Knowledge and Information Systems (JURIX 2012), 177-180

## Varie

- **V. Basile**, S. Peroni, F. Tamburini, F. Vitali: *Topical Tags vs. Non-topical tags: Towards a Bipartite Classification?* Journal of Information Science (2015)
- K. Evang, **V. Basile**, G. Chrupala, J. Bos *Elephant: Sequence Labeling for Word and Sentence Segmentation*. Conference on Empirical Methods in Natural Language Processing (EMNLP 2013), 1422-1426

## Presentazioni e Seminari

- *Semantic web-mining and deep vision for lifelong object discovery*, Data Science Meetup, Learning Center of Sophia Polytech, 20/6/2017
- *Building a default knowledge base of objects (and other stories of robots)*, Università di Torino, 16/11/2016
- *Meaning Banking and Beyond*, University of Montreal, 18/11/2015
- *Sentiment Analysis on Twitter: The Case of Italian Language*, EURECOM, Agosto 2015
- Seminari su Wordrobe a Bologna e Pisa ad Aprile 2014 e Trento a Dicembre 2014
- *Bad habits in academic code*, PyGrunn, Groningen, May 2014
- *Giochi linguistici per annotazione semantica* Digital Culture seminar, Università di Pisa, 28/4/2014
- *You tweet, I vote. A computer-assisted analysis of the political communication of 2012 Dutch political election* Digital Humanities pitches, Università di Groningen, 17/4/2014
- *The Groningen Meaning Bank*, Seminario presso TALN, Universitat Pompeu Fabra, 27/11/2013 durante un breve periodo come ospite presso il gruppo di ricerca TALN nel Novembre 2013.

## Attività Accademica e Organizzativa

### Comitati di programma

- Language Resources and Evaluation Conference (LREC 2017, 2018)
- International Joint Conference on Artificial Intelligence (IJCAI-ECAI 2018)
- ACM Symposium On Applied Computing (SAC 2018)
- AAIL Conference on Artificial Intelligence (AAIL-18)
- Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT 2018)
- Annual Meeting of the Association for Computational Linguistics (ACL 2018)
- Symposium on Emotion Modeling and Detection in Social Media and Online Interaction (Emotions@AISB 2018)
- European Semantic Web Conference (ESWC 2017, 2018)
- International Semantic Web Conference (ISWC 2017)
- International Workshop on Application of Semantic Web technologies in Robotics (AnSWeR 2017)
- International Joint Conference on Artificial Intelligence (IJCAI 2017)
- Joint Conference on Lexical and Computational Semantics (\*SEM 2016 and 2017)
- European Chapter of the Association for Computational Linguistics (EACL 2017)
- IEEE International Conference on Robotics and Automation (ICRA 2017)
- International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017)
- Italian Conference on Computational Linguistics (CLiC-it 2016 and 2017)
- International Workshop on Mining and Reasoning with Legal Text (MIREL) at JURIX 2016
- International Workshop on Natural Language Generation and the Semantic Web (WebNLG 2015 and 2016) workshop at INLG

- PEOPLES workshop, COLING 2016
- International Conference on Knowledge Engineering and Knowledge Management (EKAW 2016)
- AIC workshop, BICA 2016.
- International Semantic Web Conference, Applications Track (ISWC 2016)
- Special Session on Emotion and Sentiment in Intelligent Systems and Big Social Data Analysis (SentiData) of the IEEE International Conference on Data Science and Advanced Analytics (IEEE DSAA 2015 and 2016)
- Workshop on Semantic Sentiment Analysis (SSA) at ESWC 2014.

### **Organizzazione scientifica**

- Organizzatore del Workshop on Contextual Representation of Objects and Events in Language (CREOL 2017) presso Joint Ontology Workshops (JOWO 2017)
- Responsabile di area per Conferenza Italiana di Linguistica Computazionale (CLiC-it 2016 e 2017)
- Responsabile delle pubblicazioni per Joint Conference on Lexical and Computational Semantics conference (\*SEM 2014)
- Organizzatore di Sentiment Polarity Classification shared task (Evalita 2014 e 2016).

### **Revisione per riviste scientifiche**

- Information & Processing Management
- Computers in Industry
- Computer Speech & Language
- Language and Cognition
- Interacting with Computers

### **Varie**

- Guest editor per la rivista Applied Ontology, numero speciale su Meaning in Context: ontologically and linguistically motivated representations of objects and events., 2018
- Organizzatore locale di International Conference on Legal Knowledge and Information Systems (JURIX 2016)
- Amministratore della mailing list SIGSEM (2013-2015).
- Membro del comitato organizzativo degli eventi per il Center for Language, Logic and Cognition Groningen, University of Groningen, dal 2011 al 2014.
- Guest editor per la rivista Data & Knowledge Engineering, numero speciale su Natural Language Processing and Information Systems (NLDB 2012), Marzo 2014
- Membro del comitato organizzatore della conferenza internazionale su Applications of Natural Language Processing to Information Systems (NLDB, Groningen, Giugno 2012).

### **Premi**

- Prix d'Excellence d'Université Côte d'Azur 2017 (Prix doctorants et post-doctorants)
- Best Paper Award, International Workshop on Application of Semantic Web technologies in Robotics (AnSWeR 2017), Portoroz, Slovénie
- Primo premio alla IBM Watson Services Challenge (Evalita 2016) con la mia app *Stockle*  
<http://www.evalita.it/2016/tasks/ibm-challenge>

### **Insegnamento**

- Modulo di *Information Retrieval* nel corso di *Knowledge Engineering*, Università della Costa Azzurra

(2016)

- Assistente per il corso *Advanced Programming II*, Università di Groningen (2012)

## Competenze Informatiche

Ho esperienza come programmatore e sviluppatore software, avendo lavorato con diversi linguaggi di programmazione e tecnologie tra cui C, Python, Java, PHP, Javascript, Octave, R, shell script, Lisp e Prolog. Inoltre ho esperienza con servizi Web REST, database relazionali, amministrazione di sistema, crowdsourcing e gamification.

### Pacchetti Software

- **KNEWS** (Knowledge Extraction With Semantics)  
Pipeline NLP per estrarre istanze di frame da testo in linguaggio naturale  
<https://github.com/valeribasile/learningbyreading>
- **CoSM** (Content Analysis of Social Media)  
Piattaforma di raccolta dati e annotazione di social media  
<https://github.com/valeribasile/cosm>
- **Stockle**  
App e API che analizza il sentimento dei mercati finanziari  
<http://gingerbeard.alwaysdata.net/stockle/>
- **C&C/Boxer HTTP API**  
API per facilitare l'accesso a [C&C tools](#)  
<http://valeribasile.github.io/candcapi/>
- **ListNet**  
Implementazione di un algoritmo listwise per Learning to Rank.  
<https://github.com/valeribasile/listnet>

### Data sets

- **DeKO** (Default Knowledge about Objects)  
Serie di dataset RDF di conoscenza di senso comune su oggetti e luoghi  
<https://project.inria.fr/aloof/data/>
- **Sentix**, lessico affettivo di Italiano  
<http://valeribasile.github.io/twita/sentix.html>
- Raccolta di tag di Del.icio.us  
<http://valeribasile.github.io/delicious/>

## Competenze Varie

Ho una passione per la matematica, in particolare teoria dei numeri, statistica, automi cellulari, topologia e geometria frattale. Sono inoltre un grafico competente (senza educazione formale), ed ho creato loghi e siti Web per conferenze e progetti software (<http://valeribasile.github.io/gfx/>). Ho competenze basilari di editing video (<https://www.youtube.com/watch?v=PrNRZUSAB8c>, <https://www.youtube.com/watch?v=pLCHmzvEPug>). Supporto il movimento per il Software Libero.

## Lingue

**Italiano:** Madrelingua

**Inglese:** Conoscenza professionale completa

**Spagnolo:** Conoscenza professionale

**Francese:** Conoscenza professionale

**Olandese:** Conoscenza lavorativa limitata

## Collegamenti

La lista completa delle mie pubblicazioni è disponibile all'indirizzo: <http://valeribasile.github.io/> e sul mio profilo Google Scholar <https://scholar.google.nl/citations?user=5VCe4aAAAAAJ&hl=en>