



Personal information

Surname(s) / First name(s)	Di Caro, Luigi
Address(es)	Turin, Italy
Telephone(s)	+39 0116706708
Email(s)	luigi.dicaro@unito.it
Nationality(-ies)	Italian
Date of birth	01/10/1982
Gender	Male

Work experience

Current position	Associate Professor University of Turin
Past positions	Assistant Professor University of Turin, October 2014 – September 2022 Post-doc University of Turin, January 2011 – September 2014

Education and training

Turin, 2011 PhD	University of Turin, 2007 – 2011 in Computer Science
Turin, 2007 Master	University of Turin, 2003 – 2007 in Computer Science (110 cum laude + recommendation for publication)

National recognitions

Date	05/12/2017 – 05/12/2023
Qualification	National Scientific Qualification (Abilitazione Scientifica Nazionale) as Associate Professor in Computer Science

Institutional Roles

Date	2017 – now
------	------------

Doctoral Board

Member of the academic board of the Joint International Doctoral (Ph.D.) Degree in Law, Science and Technology (<https://www.unibo.it/en/teaching/phd/2018-2019/law-science-and-technology-1>)

Researcher Public IDs

ORCID <http://orcid.org/0000-0002-7570-637X>
Scopus <https://www.scopus.com/authid/detail.uri?authorId=57207954874>
WoS <https://www.webofscience.com/wos/author/record/29514>
DBLP https://dblp.org/pers/hd/c/Caro:Luigi_Di
Google Scholar <https://scholar.google.com/citations?user=XQAvqi4AAAAJ>

Personal skills and competences

Brief Track Record

Luigi Di Caro is Associate Professor at the Department of Computer Science of the University of Turin. He has a Master degree and a **Ph.D. in Computer Science**, and his main interests include Artificial Intelligence (AI), Natural Language Processing (NLP), Data Mining (DM), Machine Learning (ML), Legal Informatics (LI) and related interdisciplinary interactions with Cognitive Sciences (CS) and social-impact applications.

Luigi Di Caro started working on Data Mining techniques applied on text sources since his master thesis in 2007, then continued during his Ph.D, defending in 2011 the **internationally-supervised PhD thesis** supervised by prof. Maria Luisa Sapino (University of Turin) and prof. Kasim Selcuk Candan (Arizona State University).

He published more than 100 papers in several research areas and he has an **H-Index of 22** on Google Scholar (17 on Scopus).

Luigi Di Caro has active international **collaborations with more than 50 people** in different countries, and leads the NLP activities within the “*Social Computing*” research group of the Department of Computer Science, where he is / has been **supervisor of 13 PhD students**.

Luigi Di Caro is also in the **academic board of the Joint International Doctoral Degree** in Law, Science and Technology.

About the technological transfer, he **co-produced a patent** with Telecom Italia Lab and entered as **partner in a spinoff** of the University of Turin (Nomotika s.r.l.) on NLP-based supporting tools for the work of legal practitioners.

He is **senior member of prestigious programme committees** such as IJCAI (CORE=A++), AAAI (CORE=A++) and ACL (CORE=A++), EMNLP, ECAI (CORE=A) and **reviews papers for leading journals** such as the Journal of Artificial Intelligence (IF: 3.034), the Journal of Information Sciences (IF: 4.305), and the Journal of Knowledge and Information Systems (IF: 2.247).

He currently coordinates/coordinated **8 European research projects** at different degrees: 1 at project-level (total budget: 868,719 euros), 6 at partner-level (total budget: 358,000 euros), and 1 at task-level. He also coordinates/coordinated **3 national projects** with a total budget of 324,000 euros.

Scientific Production

Luigi Di Caro, in his 12 years of research after the PhD program, has been author of a large number of papers in more than ten different research areas, published in major international journals such as

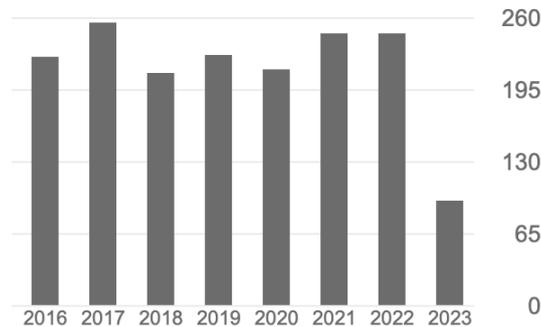
- TIST (**IF=3.19**), Scientometrics (**IF=2.173**), Computer Standards and Interfaces (**IF=1.465**), JIIS (**IF=1.107**), JLVIC (**IF=0.971**), AI&LAW, JODS;

and in peer-reviewed conferences such as

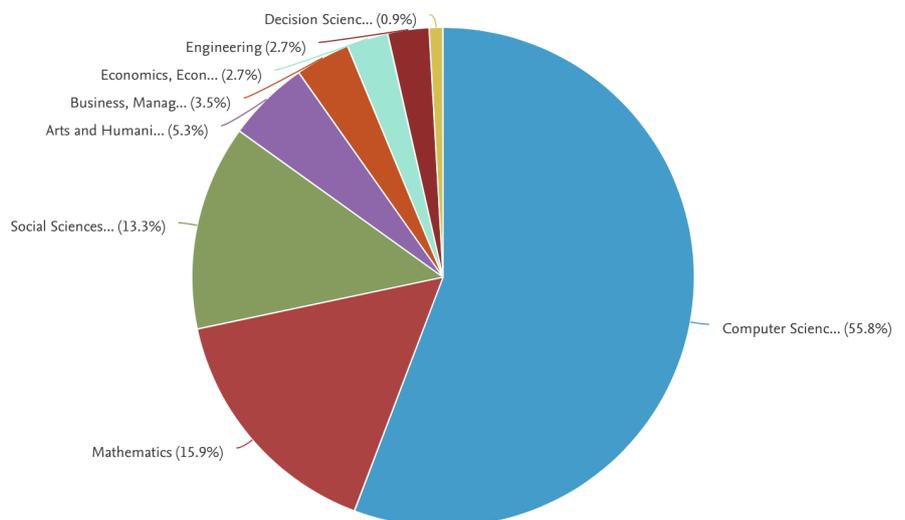
- ACL (CORE:**A++**, SHINE:**A++**), KDD (CORE:**A++**, SHINE:**A++**), EDBT (CORE:**A**, SHINE:**A**), ECML-PKDD (CORE:**A**, SHINE:**A**), PAKDD (CORE:**A**), SAC (CORE:**B**, SHINE:**B**), EKAW (CORE:**B**, SHINE:**B**), ICAIL (CORE:**C**, SHINE:**B**), LREC (CORE:**C**, SHINE:**A**) and others.

In particular,

- he has around **100 papers listed on DBLP and on Scopus**), which were mostly published after the PhD title without his supervisors;
- he has a **H-Index of 22 on Google Scholar and of 17 on Scopus**;
- he has a **i10-Index of 45** on Google Scholar;
- his papers have more than **2.2k citations** on Google Scholar and more than **1k on Scopus**;
- in the ACM Digital Library, he is author of the **3rd most downloaded research paper (and 5th most cited)** belonging to the University of Turin out of more than 2k total papers, with a total of around 5k downloads (<https://dl.acm.org/institution/60012259> - update: July 2023).



Citation trend on Google Scholar, update: July 2023



Interdisciplinarity

Papers by subject area on Scopus

Research Collaborations

Luigi Di Caro made several research experiences in both academic and business-oriented institutions abroad and in Italy. In 2007, he was **visiting student at the Arizona State University**, where he carried out a research work on Text Mining for the Master thesis, which he then published in 2008 at the International Conference on Knowledge Discovery and Data Mining (CORE:A++, MAS:A++, SHINE:A++).

In 2008, Luigi Di Caro spent 6 months at TiLab (a **research centre of Telecom Italia**), working on TV contents classification and recommendation. In this experience, Luigi Di Caro has been able to produce a patent on a system for content classification (patent US 20110264699 A1).

In 2009, Luigi Di Caro made **internship at Telefonica Research Madrid** where he combined Data Mining approaches to User Modeling applications, under the supervision of Alejandro Jaimes and Nuria Oliver.

Luigi Di Caro worked with **more than 50 people** both in the academic and in the business area, and from different countries:

Italy (out the University of Turin)

- Monica Palmirani: work on European projects BO-ECLI (JUST-JACC-2014) and MIREL (H2020 MSCA-RISE-2015), and research works on Legal Informatics (main reference: **[19]**).
- Roberto Navigli: research works on Lexical Semantics (main reference: **[73]**).

France

- Serena Villata: collaboration in the European project MIREL (H2020 MSCA-RISE-2015) and research works on legal informatics (main reference: **[3]**).
- Mario Cataldi: co-organization of the *Sideways* workshops, and research works on Bibliometric Analysis and Social Media Mining (main references: **[43][55]**)
- Marie-Aude Aufaure: research works on Bibliometric Analysis (main reference: **[53]**).

Spain

- Alejandro Jaimes, collaboration on research works on data visualization and user modeling, within the internship in Telefonica Research Madrid (main reference: **[64]**).
- Enrique Frias-Martinez and Vanessa Frias-Martinez: collaboration on research works on data visualization and user modeling (main reference: **[61]**).

Luxembourg

- Leon van der Torre: collaboration in the European project MIREL (H2020 MSCA-RISE-2015) and research works on legal informatics (main reference: **[21]**).
- Martin Theobald: research on big data management and semantic similarity (main reference: **[18]**).

Wales - Livio Robaldo: collaboration in the European project MIREL (H2020 MSCA-RISE-2015) and research works on Natural Language Processing and legal informatics (main references: **[28][46]**).

Netherlands

- Marc van Opijnen: collaboration in the European project BO-ECLI (JUST-JACC-2014). (main reference: **[19]**).
- Rohan Nanda: research on semantic similarity techniques applied on legal texts. (main references: **[49][58][65][69]**)

Bulgaria

- Hristo Konstantinov: collaboration in the European projects EUCases (10-SME-DCA 2013), MIREL (H2020 MSCA-RISE-2015), and InterLex (JUST-JCOO-AG-2017). (main reference: **[10]**).

U.S.A

- K. Selçuk Candan: research works on Data Mining, during the PhD program (main references: **[54][60]**).
- Valeria de Paiva: collaboration in the European project MIREL (H2020 MSCA-RISE-2015), and research works on legal informatics (main reference: **[2]**).

European Projects
Coordination at
project-level, unit-level
and **task-level**
(chronological order)

POLINE, Luigi Di Caro is **unit coordinator** of the European research project *POLINE: Principles Of Law In National and European VAT* of the call EU JUST-AG. The project will start in December 2023. **ROLE**: Responsible and local coordinator of the project, with a **personal budget of 75,000 euros**, leading the Natural Language Processing activities.

FACILEX, Luigi Di Caro is **unit coordinator** of the European research project *FACILEX: Analytics and Capacity building Information LEgal eXplainable tool to strengthen cooperation in the criminal matter* of the call EU JUST-2022-JCOO. The project start on December 2022. **ROLE**: Responsible and local coordinator of the project, with a **personal budget of 45,000 euros**, leading the Natural Language Processing activities (<https://site.unibo.it/facilex/en>).

ADELE, Luigi Di Caro is **unit coordinator** of the European research project *ADELE: Analytics for DEcision of LEgal Cases* of the call EU JUST-JCOO-AG-2019. The project started in February 2021, and aims at applying legal analytics to judicial decisions. **ROLE**: Responsible and local coordinator of the project, with a **personal budget of 70,000 euros**, leading the Natural Language Processing and Machine Learning activities (<https://site.unibo.it/adele/en>).

InterLex, Luigi Di Caro was co-author and **project coordinator** of the European research project *InterLex* of the call EU JUST-JCOO-AG-2017. The project started in September 2018, and is about an Advisory and Training system for Internet-related Private International Law. The total **project budget is 868,719 euros**, while the owned **part as coordinator is 190,866 euros**. A part from the role of coordinator and the management of the project, Luigi Di Caro is directly involved in all technical work packages (<http://www.interlexproject.eu>).

Cross-Justice, Luigi Di Caro was the **unit coordinator** (partner: University of Turin) of the European research project *Cross-Justice* of the call EU JUST-AG-2018. The Cross-Justice project aims at devising a free online platform to offer knowledge and legal advice on criminal procedural rights, directed to legal professionals and accessible to EU citizens. **ROLE**: Responsible and local coordinator of the project, with a **personal budget of 41,000 euros**, working on the automatic annotation and enrichment of national legislative instruments and court decisions for advanced semantic access. (<https://crossjustice.eu/>).

MIREL, Luigi Di Caro was the **unit coordinator** (partner: University of Turin) of the *MIREL* project of the call H2020 MSCA-RISE-2015. This project has the goal of creating an international and inter-sectorial network to define a formal framework and to develop tools for MIning and REasoning with Legal texts, with the aim of translating them into formal representations that can be used for querying norms, compliance checking, and decision support. **ROLE**: Responsible and local coordinator of the project, with a **personal budget of 81,000 euros**, working on text mining and ontology learning approaches in the legal domain (<http://www.mirelproject.eu>).

BO-ECLI, Luigi Di Caro was the **unit coordinator** (partner: University of Turin) of the BO-ECLI project of the call EU JUST-JACC-2014. This project aimed at improve accessibility of case law by automatically extracting legal references expressed by the judge by way of textual citations using Information Extraction and Natural Language Processing technologies. **ROLE**: Responsible and local coordinator of the project, with a **personal budget of 46,000 euros**, working on the automatic identification and completion of legal identifiers by means of rules and statistical Natural Language Processing techniques (<http://bo-ecli.eu/conference>).

EuCases, Luigi Di Caro was the **responsible of a task** in the research project supported by the Seventh Framework Programme (FP7) funding (call 10-SME-DCA 2013). The project developed a unique pan-European law and case law Linking Platform transforming multilingual legal open data into linked open data after semantic and structural analysis. **ROLE**: Responsible of the automatic classification task in the EuCases platform (<http://eucases.eu/start.html>).

National Projects
Coordination
(chronological order)

KURAMI, Luigi Di Caro is the **Principal Investigator** of a PRIN-2022 research project on Natural Language Processing, transparency and explainability. This project has a **budget of 250,000 euros**.

SemBurst, Luigi Di Caro was the **Principal Investigator** of a research project of a Compagnia San Paolo young researchers' Call, in 2014. This project had a **budget of 49,000 euros**, and ended in March 2017 with successful results in terms of scientific production and reached objectives.

KnowYouAll, Luigi Di Caro was the **Principal Investigator** of the *KnowYouAll* research project on Computational Linguistics, that has been awarded with a **25,000 euros grant** by Telecom Italia in a national competition. This project produced techniques and software that have been integrated into the Telecom Italia Lab research centre (for educational purposes). See the paper published at the International Conference ECML-PKDD 2014 for the details https://link.springer.com/chapter/10.1007/978-3-662-44845-8_30.

Projects Participation
(chronological order)

LAST-JD-RIoE (Law, Science and Technology, Rights of Internet of Everything) is a recently-funded Horizon-2020 research project of the call MSCA ITN (Marie Skłodowska-Curie actions European Joint Doctorates). Its main aim is to investigate the *Internet of Everything* topic under different perspectives: technical, legal, economical, ethical, and philosophical.

ITxLaw proposed a new methodology based on ontologies for connecting legal texts with normative meaning, taking into account the interpretation of the norms and their evolution (project financed by Compagnia di San Paolo call2013, 2013 - 2015). **ROLE:** Development of computational solutions for automatic constructions of ontologies and relative legal text classification.

RVILP. Ordinary experience suggests that lexical competence, i.e. the ability to use words, includes both the ability to relate words to the external world as accessed through perception and the ability to relate words to other words in inferential tasks. This project was about a functional neuroimaging study of 'The Role Of Visual Imagery In Lexical Processing' (call 2012, funded by Compagnia di San Paolo, 2012 - 2014). **ROLE:** Development of computational models for correlation analysis of words in vector-based conceptual spaces.

ICT4LAW was a large interdisciplinary research project involving several university departments and industry partners (call 'Bando regionale' on converging technologies). The goal was to create novel services for citizens, enterprises, public administration and policy makers (2010 - 2013). **ROLE:** Legal data analysis for classification of laws and automatic ontologies learning.

ATLAS was a project on the automatic translation from Italian to Italian Sign Language (LIS) of deaf people (call 'Bando regionale' on converging technologies). The aim of the project is to create applications to improve inclusion of deaf people by providing contents in their language using virtual characters (2010 - 2013). **ROLE:** Application of Data Mining standard techniques to support syntactic parsing in the translation process.

DynamicTV. The project consisted of a system to classify and recommend TV contents based domain ontologies and texts. In fact, large text corpora, like newspapers archives, used to contain dynamic data from a cultural and a linguistics point of view. Domain ontologies, on the other hand, represent a static and domain-expert knowledge. The project aimed at bridging the gap between these different sources of information for improving content classification and recommendation. **ROLE:** Application of Data Mining techniques for automatic content clustering, classification, and recommendation based on user profiles.

Member of Doctoral
(Ph.D.) Degree

Luigi Di Caro is an academic **member of the Joint International Doctoral (Ph.D.) Degree in Law, Science and Technology**, initially funded by the EU Erasmus Mundus EACEA initiative (2012-2016), and now supported by the recently-funded *LAST-JD-RIoE* project (Marie Skłodowska-Curie actions European Joint Doctorates).

Research group and supervision of PhD Students

Luigi Di Caro works within the **Social Computing** research group of the Department of Computer Science (<http://beta.di.unito.it/index.php/english/research/groups/social-computing/about>), where **he leads all projects and activities related to Natural Language Processing and Legal Informatics**.

In this context, he has been supervising the work of the following **13 PhD students**:

[S1] Eniafe Festus Ayetiran, PhD student of Erasmus Mundus International Joint Doctorate in Law, Science and Technology, working on Word Sense Disambiguation (PhD obtained in 2017);

[S2] Adebayo John Kolawole, PhD student of Erasmus Mundus International Joint Doctorate in Law, Science and Technology, working on deep learning and legal informatics (PhD obtained in 2018);

[S3] Alice Ruggeri, PhD student in Cognitive Science, working on cognitive modeling of linguistic phenomena and common-sense knowledge (PhD obtained in 2018);

[S4] Rohan Nanda, PhD student of Erasmus Mundus International Joint Doctorate in Law, Science and Technology, working on semantic similarity of legal documents (PhD obtained in 2019);

[S5] Michele Fioravera, PhD student in Mathematics, working on Natural Language Processing and ontology learning in educational virtual community environments (PhD obtained in 2019);

[S6] Giovanni Siragusa, PhD student in Computer Science working on Open Information Extraction, Topic Modeling, Summarization and Neural Networks (PhD obtained in 2020);

[S7] Valentina Leone, PhD student of Erasmus Mundus International Joint Doctorate in Law, Science and Technology, working on semantic representations, legal ontologies, and design patterns (PhD obtained in 2021);

[S8] Francesca Alloatti, industrial PhD student in Computer Science (PhD programme in apprenticeship), working on Computational Linguistics (PhD obtained in 2022);

[S9] Roger Ferrod, PhD student in Computer Science, working on Deep Learning, Computational Linguistics and biomedical applications (PhD program is **running**, started in **2021**);

[S10] Davide Audrito, PhD student of Erasmus Mundus International Joint Doctorate in Law, Science and Technology, working on legal ontologies (PhD program is **running**, started in **2021**);

[S11] Francesca Grasso, PhD student in Computer Science, working on Applied Linguistics and Ecolinguistics (PhD program is **running**, started in **2022**).

[S12] Giuseppe Ruggero, industrial PhD student in Computer Science (PhD programme in apprenticeship), working on NLP, Speech and Deep Learning technologies (PhD program is **running**, started in **2022**).

[S13] Giulia Ruffini, PhD student in Computer Science (PhD programme in apprenticeship), working on Process Mining (PhD program is **running**, started in **2023**).

Research Periods Abroad

[2016], one month - Visiting researcher at **Stanford University**. Topics: Computational Linguistics, Legal-informatics. Invited by Prof. Cleo Condoravdi, Center for the Study of Language and Information, Stanford University.

[2009], six months - Internship at **Telefonica Research in Madrid**, Spain. Topics: User Modeling, Interactive Query Systems, Data Visualization, under the supervision of Alejandro Jaimes and Nuria Oliver (now Director of Research in Data Science at Vodafone).

[2008], six months - Collaboration with **TiLab (Telecom Italia Lab)** on TV content classification and recommendation, automatic enrichment of domain ontologies and analysis of contexts of use, invited by Fabrizio Antonelli and Dario Mana.

[2007], four months - Visiting student at the **Arizona State University**, under the supervision of prof. K. Selcuk Candan (Arizona State University) and prof. Maria Luisa Sapino (University of Torino), working on Data Mining, Latent Semantic Analysis, and Ontology Learning.

Technological Transfer

Luigi Di Caro **co-produced a patent** with Telecom Italia Lab (TiLab), concerning the automatic classification of text content (Patent ID: United States - US20110264699A1).

Luigi Di Caro has been invited to enter as **partner in a spinoff of the University of Torino** (Nomotika s.r.l.) specialized in research and development of information technology solutions for the daily work of legal practitioners, working on Natural Language Processing pipelines.

Luigi Di Caro **developed software and tools** for scientific research such as TMine, CoSeNa, ImmEx, D-Index, PhC, PdView (all published in conference proceedings and journals); some of the software has been revised and **integrated in business institutions** (TiLab, Telefonica Research, Nomotika s.r.l., Augeos s.p.a).

Luigi Di Caro **is supervisor** of two industrial PhD programs, with Celi (<https://www.celi.it/en/>) s.r.l. and Cerence (<https://www.cerence.com>).

Luigi Di Caro has been **scientific tutor** of two **industrial research contracts** with Celi s.r.l. (<https://www.celi.it/en/>):

- one PhD grant in apprenticeship (see PhD student **[S8]** Francesca Alloatti); - one research contract in apprenticeship (held by the Data Scientist Maria Luisa Gabrielli).

Luigi Di Caro **is co-holder** of an industrial agreement between the Department of Computer Science of the University of Turin and Schneider Electric (with Prof. Claudio Schifanella).

Awards

[A1] Luigi Di Caro won the best paper **award** at the 23rd International Conference on Knowledge Engineering and Knowledge Management (EKAW 2022), with the paper titled "MultiAlignNet: Cross-Lingual Knowledge Bridges between Words and Senses". **[86]**

[A2] Luigi Di Caro won the best paper **award** at the International Conference on Hybrid Intelligent System (HIS) 2016, with the paper titled "Text Segmentation With Topic Modeling And Entity Coherence". **[14]**

[A3] Luigi Di Caro collaborated in the ICT project 'First Life', **awarded** by MIUR in a national competition about innovative ideas called "'Smart Cities and Social Innovation and Community'".

[A4] Luigi Di Caro has been **short-listed** with a project on Question Answering in the national competition "'Working Capital 2011'" for innovative ideas. Only 12 out of 846 ICT projects have been selected up to this phase.

[A5] Luigi Di Caro has been **awarded** with one research grant in the national competition "'Working Capital 2012'" for innovative ideas (only 20 out of around 1000 presented projects have been awarded). The title of the project is "'KnowYouAll'", a cloud-based semantic search engine which has been then used by Telecom Italia Lab in an Education-oriented scenario. A published paper (**[34]**) at the International Conference ECML-PKDD (2014) describes the system https://link.springer.com/chapter/10.1007/978-3-662-44845-8_30

[A6] Luigi Di Caro won the **prize** 'Best Master Thesis in Information and Communication Technology' (involved regions : Piemonte, Valle d'Aosta, Liguria). Premio Zucca 2008, Italy.

Organized Workshops

[SIDEWAYS 2015, 2016, 2017, 2018, 2019, 2020, 2022, 2023] Co-founder and organizer of eight editions of the International Workshop on Social Media World Sensors <http://linc.iut.univ-paris8.fr/sideways/index.html>.

[KM4Law 2022, 2023] Co-founder and member of the organization of the Workshop on Knowledge Management for Law, held in conjunction with EKAW 2022 <https://km4law.di.unito.it>.

[GENERAL 2023] Co-founder and member of the organization of the Workshop on GENERative, Explainable and Reasonable Artificial Learning, held in conjunction with CHIItaly 2023 <https://www.evilscrip.t.eu/general/>.

[MIREL 2017] Member of the organization of the Workshop on Mining and Reasoning with Legal Text, held in conjunction with ICAIL 2017 <http://www.iaail.org/?q=article/mirel-2017-workshop-mining-and-reasoning-legal-texts>.

[DC 2014] Member of the organization of the AI*IA 2014 Doctoral Consortium in Artificial Intelligence, coordinated by Prof. Roberto Navigli (University of Sapienza, Rome).

[DWA I 2013] Co-chair of the First Doctoral Workshop in Artificial Intelligence, held in conjunction with the 13th International Conference of the Italian Association for Artificial Intelligence (AI*IA 2013).

Programme Committees

IJCAI (2017, 2018, 2019, 2020, **2021***, 2022, 2023) - International Joint Conference on Artificial Intelligence. ***Senior PC member**.

CORE:A++, MAS:A++, SHINE:A++

AAAI (2020, **2021***, 2023) - International Conference on Artificial Intelligence. ***Senior PC member**

CORE:A++, MAS:A++, SHINE:A++

ACL (2018, 2019, **2021***, 2023*) and **ACL Rolling Review** - Annual Meeting of the Association for Computational Linguistics (*three tracks*: World level semantics; Resources and evaluation; Information extraction e text mining). ***Area Chair** for the NLP Applications track.

CORE:A++, MAS:A++, SHINE:A++

EMNLP (2018, 2019, 2020, 2021, **2022***) - International Conference on Empirical Methods in Natural Language Processing. ***Main and Demo tracks**

CORE:A, MAS:B, SHINE:A+

ECAI (2023*) - European Conference on Artificial Intelligence ***Senior member / meta-reviewer**

CORE:A, MAS:A-, SHINE:A-

EACL (2021, *2023) - Conference of the European Chapter of the Association for Computational Linguistics. *** 4 tracks: i) Lexical Semantics, ii) Information Extraction, iii) Language resources and Evaluation and iv) Document Analysis.**

CORE:A, MAS:A-, SHINE:A-

ECML-PKDD (2014, 2015, 2016) - European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases

CORE:A, MAS:A, SHINE:A

HyperText (2022) - ACM International Conference on Hypertext and Social Media

CORE:A, MAS:A, SHINE:A

SAC (2018, 2019, 2020, 2021, 2022, 2023, 2024) - ACM/SIGAPP Symposium On Applied Computing (**three tracks: Data Mining; Knowledge and Language Processing; Machine Learning Applications**)

CORE:B, MAS:B, SHINE:B

AAACL-IJCNLP (2020, 2022) - The 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 9th International Joint Conference on Natural Language Processing

NAACL-HLT (2021) - Annual Conference of the North American Chapter of the Association for Computational Linguistics

COGSCI (2016, 2017, 2018, 2019) - Annual Meeting of the Cognitive Science Society
CORE:A, SHINE:C

LREC (2016, 2018, 2020, 2022) - Biennial Language Resources and Evaluation Conference

CORE:C, MAS:B, SHINE:A

ISMIS (2017, 2018, 2020) - International Symposium on Methodologies for Intelligent Systems

CORE:C, MAS:B, SHINE:B-

ICAIL (2021, 2023*) - International Conference on Artificial Intelligence and Law (***main track and doctoral consortium**)

CORE:C, MAS:B, SHINE:B

WEBIST (2016, 2017, 2018, 2019, 2020, 2021, 2022) - International Conference on Web Information Systems and Technologies

CORE:C, MAS:C, SHINE:C

ASONAM (2017) - International Conference series on Advances in Social Network Analysis and Mining

SHINE:C

Unclassified conferences and Workshops:

SEM (2019) **NLLP** (2019/20/22) **MOD/LOD** (2015/16/17/18) **JURIX** (2018/19/20)
COINS (2019/21) **BIDMA** (2016/19) **DMNLP** (2015/16/17)

Track chair

HyperText (2019) - ACM International Conference on Hypertext and Social Media
CORE:A, MAS:A, SHINE:A

Reviewer for International
Journals

An updated profile with the certified reviews is available at the *Publons* public profile:
<https://publons.com/a/3393932>

Journal of Knowledge-based Systems

IF: 5.101, SJR-2018: 1.460

Journal of Information Sciences

IF: 4.305, H-index: 142, SJR-2017: 1.635

Information Processing and Management

IF: 3.892, H-index: 88, SJR-2017: 1.043

Journal of Artificial Intelligence

IF: 3.034, H-index: 129, SJR-2017: 0.88

Transactions on Knowledge Discovery from Data

IF: 2.538, H-Index: 44, SJR-2018: 0.629

Journal of Knowledge And Information Systems

IF: 2.247, H-index: 52, SJR-2017: 0.672

Journal of Computer Standards and Interfaces

IF: 1.465, H-index: 55, SJR-2017: 0.378

Journal of Behaviour and Information Technology

IF: 1.380, H-index: 60, SJR-2017: 0.676

Journal of Intelligent Information Systems

IF: 1.107, H-index: 47, SJR-2017: 0.481

Journal of Natural Language Engineering

IF: 0.8, H-index: 48, SJR-2017: 0.264

Methods of Information in Medicine

IF: 1.024

The International Journal of Urban Policy and Planning

IF: 3.853, SJR-2017: 1.44

Journal of Artificial Intelligence and Law

H-index: 44, SJR-2017: 0.492

Journal Plos One

IF: 2.740

Journal of SN Applied Sciences

(new journal)

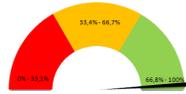
Trattamento del Linguaggio Naturale (Natural Language Processing)

Hours of frontal lessons: 24 (per year)

Editions: 2017-18, 2018-19, 2019-20, 2020-21, 2021-22, 2022-23

Avg number of examined students: 45 (per year)

Max student satisfaction Index: 98.88 (EduMeter system, Univ. of Turin)

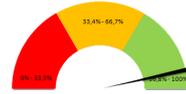
**Data Mining e Analisi Multivariata (Data Mining and Multivariate Analysis)**

Hours of frontal lessons: 36 (per year)

Editions: 2018-19, 2019-20

Avg number of examined students: 80 (per year)

Max student satisfaction Index: 93.48 (EduMeter system, Univ. of Turin)

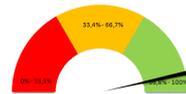
**Linguaggi Formali e Traduttori (Formal Languages and Compilers)**

Hours of frontal lessons: 42 (per year)

Editions: 2019-20, 2020-21, 2021-22, 2022-23

Avg number of examined students: 90 (per year)

Max student satisfaction Index: 93.72 (EduMeter system, Univ. of Turin)

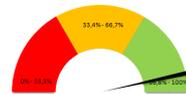
**Informatica per l'Economia e la Statistica (Computer Science for Economics and Statistics)**

Hours of frontal lessons: 18 (per year)

Editions: 2018-19, 2019-20, 2020-21, 2021-22, 2022-23

Avg number of examined students: 300 (per year)

Max student satisfaction Index: 93.44 (EduMeter system, Univ. of Turin)

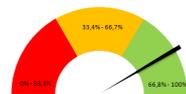
**Laboratorio di Programmazione I (Programming Lab)**

Hours of frontal lessons: 30 (per year)

Editions: 2022-23

Avg number of examined students: 100 (per year)

Max student satisfaction Index: 83.86 (EduMeter system, Univ. of Turin)

**Informatica e Sistemi Informativi (Informative Systems)**

Hours of frontal lessons: 36 (per year)

Editions: 2014-15, 2015-16, 2016-17, 2017-18

Avg number of examined students: 300 (per year)

Max student satisfaction Index: 91.92 (EduMeter system, Univ. of Turin)

**Sistemi Operativi (Operative Systems)**

Editions: 2014-15, 2015-16

Avg number of examined students: 100 (per year)

Max student satisfaction Index: 90.91 (EduMeter system, Univ. of Turin)



Research Interests

Computational Linguistics. Luigi Di Caro is working on the integration of different Lexical Semantics resources (dictionaries, multilingual repositories, common sense knowledge and association norms) towards novel (explainable and interpretable) encoding mechanisms. However, he approaches NLP by means of a wide range of perspectives, ranging from linguistic annotation to Neural Networks.

Ontology Learning. The automatic construction of ontologies has been the first research task faced by Luigi Di Caro in his Master Thesis. Since then, he started applying Natural Language Processing techniques and Pattern Mining approaches to improve accuracy and fit domain-specific scenarios and real applications.

Topic Extraction and Topic Trend Analysis. In the first years of his career, Luigi Di Caro developed novel algorithms based on Vector Space Models and Latent Semantic Analysis to automatically extract structured knowledge from plain text. Moreover, he proposed several visualization techniques to explore topics evolution over time.

Text Classification. Luigi Di Caro built a sound background on Machine Learning techniques applied on textual data for automatic classification. He both created new algorithms and applied them on real cases (i.e., for Legal text).

Information Retrieval and Navigation. Within this field, Luigi Di Caro published several works touching different aspects. In particular, he put particular attention to alternative exploration strategies where users become aware of the content by directly browsing within it (orienteering).

Social Network Analysis and Social Media. Luigi Di Caro worked on textual data extracted from Twitter developing novel algorithms to extract and compose emerging topics, proposing novel approach to this research field based on the Aging Theory.

Content Recommendation. In the business area, Luigi Di Caro worked on the project Dynamic TV (TiLab) on recommendation of TV contents. From a technical point of view, I collaborated in the development of the techniques for automatic classification of contents into ontology concepts. This work produced a Patent on the field.

Bibliometric Analysis. Luigi Di Caro's interests touched other Data Mining and Social Network-related research areas like the analysis of research products and the researchers' level of independence. As first contributor, he developed a new algorithm to find latent dependencies of authors within their research communities, as a tool for supporting standard quantity and quality factors like H-index.

Named Entity Recognition. Luigi Di Caro won a national competition with the project KnowYouAll, i.e., a system which included an identification and linking process of named entities (people, organizations, locations, and topics) into temporal frames to support advanced search in text collections.

Syntactic Parsing One of the research challenge faced by Luigi Di Caro in his recent works is the integration of syntactic information into standard word-based approaches for Data Mining on text. In particular, he recently developed a novel technique to feed Support Vector Machines with syntactic knowledge for automatic ontology learning.

Word Sense Disambiguation Luigi Di Caro is currently interested in new challenges related to the Word Sense Disambiguation task. In detail, he is studying techniques to automatically extract word senses based on the actual use of natural language.

Sentiment Analysis Luigi Di Caro deeply studied this novel research area proposing new algorithms and new formalization schemes.

Legal Text Enrichment Luigi Di Caro is also working on the task of text navigation and enhancement, in the legal domain. This problem has been both faced in research scenarios as well as for business solutions.

Interdisciplinary Integration of Data Mining, Data Visualization, User Modeling, Legal approaches Informatics, Human-Computer Interaction and Cognitive Science Luigi Di Caro has published papers on more than 10 research areas, demonstrating his particular attention to interdisciplinary techniques and methodologies.

12 Selected Publications
(chronological order)

[SEL-1] Luigi Di Caro, K. Selçuk Candan, and Maria Luisa Sapino. "Navigating within news collections using tag-flakes." *Journal of Visual Languages and Computing* 22.2 (2011): 120-139.

Impact Factor = 0.971, H-Index = 11, SJR-2011 = 0.543, citations = 19

[SEL-2] Luigi Di Caro, Mario Cataldi, and Claudio Schifanella. "The d-index: Discovering dependences among scientific collaborators from their bibliographic data records." *Scientometrics* 93.3 (2012): 583-607.

Impact Factor = 2.173, H-Index = 90, SJR-2012 = 1.28, citations = 16

[SEL-3] Livio Robaldo and Luigi Di Caro. "OpinionMining-ML." *Computer Standards and Interfaces* 35.5 (2013): 454-469.

Impact Factor = 1.465, H-Index = 55, SJR-2012 = 0.647, citations = 40

[SEL-4] Luigi Di Caro and Matteo Grella. "Sentiment analysis via dependency parsing." *Computer Standards and Interfaces* 35.5 (2013): 442-453.

Impact Factor = 1.465, H-Index = 55, SJR-2012 = 0.647, citations = 95

[SEL-5] Candan, K. Selçuk, Luigi Di Caro, and Maria Luisa Sapino. "PhC: Multiresolution visualization and exploration of text corpora with parallel hierarchical coordinates." *ACM Transactions on Intelligent Systems and Technology (TIST)* 3.2 (2012): 22:1-36.

Impact Factor = 3.19, H-Index = 40, SJR-2013 = 3.927, citations = 13

[SEL-6] Cataldi, Mario, Luigi Di Caro, and Claudio Schifanella. "Personalized emerging topic detection based on a term aging model." *ACM Transactions on Intelligent Systems and Technology (TIST)* 5.1 (2013): 7.

Impact Factor = 3.19, H-Index = 40, SJR-2013 = 3.927, citations = 48

[SEL-7] Guido Boella, Luigi Di Caro, Livio Robaldo. "Learning from syntax generalizations for automatic semantic annotation." *Journal of Intelligent Information Systems* 43.2 (2014): 231-246.

Impact Factor = 1.107, H-Index = 47, SJR-2014 = 0.386, citations = 41

[SEL-8] Boella, G., Di Caro, L., Humphreys, L., Robaldo, L., Rossi, P., and van der Torre, L. Eunomos, a legal document and knowledge management system for the web to provide relevant, reliable and up-to-date information on the law. *Artificial Intelligence and Law*, 24(3) (2016): 245-283

Cite-score-2017 = 2.07, H-Index = 30, SJR-2016 = 0.364, citations = 91

[SEL-9] Gianmaria Ajani, Guido Boella, Luigi Di Caro, Livio Robaldo, Llio Humphreys, Sabrina Praduroux, Piercarlo Rossi, Andrea Violato. "The European Taxonomy Syllabus: A multi-lingual, multi-level ontology framework to untangle the web of European legal terminology." *Applied Ontology* 11.4 (2016): 325-375.

Cite-score-2017 = 1.80, H-Index = 24, SJR-2016 = 0.447, citations = 36

[SEL-10] Alice Ruggeri, Luigi Di Caro, and Guido Boella. "The Role of Common-Sense Knowledge in Assessing Semantic Association." *Journal on Data Semantics* (2019): Volume 8, Issue 1, pp 39–56.

Cite-score-2017 = 2.16, H-Index = 11, SJR-2017 (last) = 0.602, citations = 3

[SEL-11] Guido Boella, Luigi Di Caro. Supervised learning of syntactic contexts for uncovering definitions and extracting hypernym relations in text databases. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML-PKDD)*. Springer, Berlin, Heidelberg, 2013:64-79 (full paper).

CORE:A, SHINE:A, citations = 9

[SEL-12] Luigi Di Caro, Vanessa Frías-Martínez, Enrique Frías-Martínez. Analyzing the Role of Dimension Arrangement for Data Visualization in Radviz. *Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)* 2010: 125-132 (full paper).

CORE:A, MAS:B, citations = 63

All Publications

- journal
- proceedings

- **[88]** Federico Torrielli, Amon Rapp, Luigi Di Caro. How Shall a Machine Call a Thing? International Conference on Natural Language Processing and Information Systems (NLDB): pp 546–557 (2023).
DOI: https://doi.org/10.1007/978-3-031-35320-8_41
- **[87]** Davide Audrito, Emilio Sulis, Llio Humphreys, Luigi Di Caro. Analogical lightweight ontology of EU criminal procedural rights in judicial cooperation. Artificial Intelligence and Law: 31, pages 629–652 (2023).
DOI: <https://doi.org/10.1007/s10506-022-09332-9>
- **[86]** Grasso, F. Lovera Rulfi, V. Di Caro, L. MultiAligNet: Cross-Lingual Knowledge Bridges between Words and Senses. 23rd International Conference on Knowledge Engineering and Knowledge Management (EKAW), 2022. **[Best paper award]**
DOI: https://doi.org/10.1007/978-3-031-17105-5_3
- **[85]** Sulis, E. Humphreys, L. Vernerio, F. Amantea, I. Audrito, D. Di Caro, L. Exploiting co-occurrence networks for classification of implicit inter-relationships in legal texts. Information Systems, Volume 106.
DOI: <https://doi.org/10.1016/j.is.2021.101821>
- **[84]** Llio Humphreys, Guido Boella, Leendert van der Torre, Livio Robaldo, Luigi Di Caro, Sepideh Ghanavati, Robert Muthuri: Populating legal ontologies using semantic role labeling. Artificial Intelligence and Law 29(2): 171-211 (2021).
DOI: <https://doi.org/10.1007/s10506-020-09271-3>
- **[83]** Ferrod, R. Brunetti, E. Di Caro, L. Di Francescomarino, C. Dragoni, M. Ferrod, R. Ghidini, C. Marinello, R. Ronzani, M. Sulis, E. A Support for Understanding Medical Notes: Correcting Spelling Errors in Italian Clinical Records. SMARTERCARE@AI*IA 2021: 19-28.
REF: <http://ceur-ws.org/Vol-3060/paper-3.pdf>
- **[82]** Aringhieri, R. Boella, G. Brunetti, E. Di Caro, L. Di Francescomarino, C. Dragoni, M. Ferrod, R. Ghidini, C. Marinello, R. Ronzani, M. Sulis, E. Leveraging Structured Data in Predictive Process Monitoring: the Case of the ICD-9-CM in the Scenario of the Home Hospitalization Service. SMARTERCARE@AI*IA 2021: 48-60
REF: <http://ceur-ws.org/Vol-3060/paper-6.pdf>
- **[81]** Roberto Aringhieri, Boella, G. Brunetti, E. Di Caro, L. Di Francescomarino, C. Dragoni, M. Ferrod, R. Ghidini, C. Marinello, R. Ronzani, M. Sulis, E. Towards the Application of Process Mining for Supporting the Home Hospitalization Service. ITBPM@BPM 2021: 33-38
REF: http://ceur-ws.org/Vol-2952/paper_295a.pdf
- **[80]** Grasso, F., Di Caro, L.: A Methodology for Large-Scale, Disambiguated and Unbiased Lexical Knowledge Acquisition Based on Multilingual Word Alignment. CLiC-it 2021, 2021.
REF: <http://ceur-ws.org/Vol-3033/paper18.pdf>
- **[79]** Ferrod, R., Cena, F., Di Caro, L., Mana, D., Simeoni, R.: Identifying Users Domain Expertise from Dialogues. 29th Conference on User Modeling, Adaptation and Personalization (UMAP) 2021 (LBR paper), 2021.
DOI: <https://doi.org/10.1145/3450614.3461683>
- **[78]** Ferrod, R., Di Caro, L., Schifanella, C.: Structured Semantic Modeling of Scientific Citation Intents. Extended Semantic Web Conference (ESWC) 2021, to appear (2021).
DOI: https://doi.org/10.1007/978-3-030-77385-4_27
- **[77]** Alloatti F., Cena, F., Di Caro, L., Ferrod, R., Siragusa, G.: Towards Mental Model-driven Conversations. Theoretical and Methodological Challenges in Intelligent Conversational User Interface Interactions (Workshop at the 2021 ACM Intelligent User Interface (IUI) Conference), 2021.
REF: <http://speechinteraction.org/IUI2021/docs/programme.pdf>

All Publications

- journal
- proceedings

- **[76]** Siragusa, G., Robaldo L., Di Caro, L., Violato, A.: Textual Entailment for Cybersecurity: An Applicative Case. *Journal of Logics and their Applications (FLAP)*, 975-992, (2021).
Open Access: <https://collegepublications.co.uk/ifcolog/?00046>
- **[75]** Humphreys L., Boella G., Di Caro L., Robaldo L., van der Torre L., Ghanavati S., Muthuri R.: Populating Legal Ontologies using Semantic Role Labeling. *LREC 2020*: 2157-2166
REF: <https://www.aclweb.org/anthology/2020.lrec-1.366.pdf>
- **[74]** Alloatti F., Di Caro L., Sportelli G.: Real Life Application of a Question Answering System Using BERT Language Model. *SIGdial 2019*: 250-253
DOI: <https://doi.org/10.18653/v1/W19-5930>
- **[73]** Leone V., Siragusa G., Di Caro L., Navigli R.: Building Semantic Grams of Human Knowledge. *LREC 2020*: 2991-3000
REF: <https://www.aclweb.org/anthology/2020.lrec-1.366.pdf>
- **[72]** Siragusa G., Di Caro L., Tosalli M.: Mining annotators' common knowledge for automatic text revision. *International Journal of Metadata, Semantics and Ontologies (IJMSO)* 13(3): 254-263 (2019).
DOI: <https://doi.org/10.1504/ijmsso.2019.099837>
- **[71]** Leone V., Di Caro L., Villata S. Taking stock of legal ontologies: a feature-based comparative analysis. *Artificial Intelligence and Law* 28.2 (2020): 207-235.
DOI: <https://doi.org/10.1007/s10506-019-09252-1>
- **[70]** Di Caro L., Ruggeri A. Unveiling middle-level concepts through frequency trajectories and peaks analysis. *34th ACM/SIGAPP Symposium On Applied Computing (SAC 2019)*: 1035-1042.
DOI: <https://doi.org/10.1145/3297280.3297383>
- **[69]** Nanda R., Siragusa G., Di Caro L., Boella G., Grossio L., Gerbaudo M., Costamagna F. Unsupervised and supervised text similarity systems for automated identification of national implementing measures of European directives. *Artificial Intelligence and Law* 27(2) (2019): 199-225.
DOI: <https://doi.org/10.1007/s10506-018-9236-y>
- **[68]** Boella G., Di Caro L., Leone V. Semi-automatic knowledge population in a legal document management system. *Artificial Intelligence and Law* 27(2) (2019): 199-225.
DOI: <https://doi.org/10.1007/s10506-018-9239-8>
- **[67]** Ruggeri, A., Di Caro, L., Boella, G. The Role of Common-Sense Knowledge in Assessing Semantic Association (2019) *Journal on Data Semantics*, 8 (1), pp. 39-56.
DOI: <https://doi.org/10.1007/s13740-018-0094-2>
- **[66]** Ferrod R., Schifanella C., Di Caro L., Cataldi M. Disclosing Citation Meanings for Augmented Research Retrieval and Exploration. *Extended Semantic Web Conference (ESWC) 2019*: 101-115. DOI: https://doi.org/10.1007/978-3-030-21348-0_7
- **[65]** Siragusa, G., Nanda, R., De Paiva, V., Di Caro, L. Relating Legal Entities via Open Information Extraction (2019) *Communications in Computer and Information Science*, 846, pp. 181-187.
DOI: https://doi.org/10.1007/978-3-030-14401-2_17

- journal
- proceedings

- [64] Leone V., Di Caro L., Villata S. Legal Ontologies and How to Choose Them: the InvestigatiOnt Tool. International Semantic Web Conference (P&D/Industry/BlueSky) 2018.
- [63] Marchisio, M., Di Caro, L., Fioravera, M., Rabellino, S. Towards Adaptive Systems for Automatic Formative Assessment in Virtual Learning Communities (2018) Proceedings - International Computer Software and Applications Conference, 1, art. no. 8377801, pp. 1000-1005.
DOI: <https://doi.org/10.1109/COMPSAC.2018.00176>
- [62] Cataldi, M., Di Caro, L., Schifanella, C. Improving search and navigation user experience by making use of social data (2018) Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications, 4, pp. 2132-2156.
DOI: <https://doi.org/10.4018/978-1-5225-5191-1.ch095>
- [61] Boella, G., Di Caro, L., Leone, V. Semi-automatic knowledge population in a legal document management system. Artificial Intelligence and Law 27.2 (2019): 227-251.
DOI: <https://doi.org/10.1007/s10506-018-9239-8>
- [60] Cataldi, M., Di Caro, L., Schifanella, C. All for one or one for all? Analyzing collaboration patterns in research environments (2018) CEUR Workshop Proceedings, 2080, pp. 76-79.
- [59] Barana, A., Di Caro, L., Fioravera, M., Marchisio, M., Rabellino, S. Ontology development for competence assessment in virtual communities of practice (2018) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10948 LNAI, pp. 94-98. DOI: https://doi.org/10.1007/978-3-319-93846-2_18
- [58] Nanda, R., Siragusa, G., Di Caro, L., Boella, G., Grossio, L., Gerbaudo, M., Costamagna, F. Unsupervised and supervised text similarity systems for automated identification of national implementing measures of European directives. Artificial Intelligence and Law 27.2 (2019): 199-225
DOI: <https://doi.org/10.1007/s10506-018-9236-y>
- [57] John, A.K., Di Caro, L., Boella, G. Siamese network with soft attention for semantic text understanding (2017) ACM International Conference Proceeding Series, 2017-September, pp. 160-167.
DOI: <https://doi.org/10.1145/3132218.3132236>
- [56] Nanda, R., Konstantinov, H., Hristov, H., Di Caro, L., Tyankov, T., Costamagna, F., Boella, G., Traykov, D., Humphreys, L., Robaldo, L., Romano, M. A unifying similarity measure for automated identification of national implementations of european union directives (2017) Proceedings of the International Conference on Artificial Intelligence and Law, pp. 149-158.
DOI: <https://doi.org/10.1145/3086512.3086527>
- [55] John, A.K., Caro, L.D., Boella, G. Solving bar exam questions with deep neural networks (2017) CEUR Workshop Proceedings, 2143.
- [54] Siragusa, G., Di Caro, L., Tosalli, M. Automatic extraction of correction patterns from expert-revised corpora (2017) Communications in Computer and Information Science, 755, pp. 134-146.
DOI: https://doi.org/10.1007/978-3-319-70863-8_13
- [53] John, A.K., Di Caro, L., Boella, G. Text segmentation with topic modeling and entity coherence (2017) Advances in Intelligent Systems and Computing, 552, pp. 175-185. **[Best Paper award]**
DOI: https://doi.org/10.1007/978-3-319-52941-7_18
- [52] Siragusa, G., Leone, V., Di Caro, L., Schifanella, C. Discovering relational phrases for qualia roles through open information extraction (2017) Communications in Computer and Information Science, 786, pp. 69-84.
DOI: https://doi.org/10.1007/978-3-319-69548-8_6

All Publications

- journal
- proceedings

- **[51]** Kolawole John, A., Di Caro, L., Robaldo, L., Boella, G. Textual inference with tree-structured LSTM (2017) *Communications in Computer and Information Science*, 765, pp. 17-31.
DOI: https://doi.org/10.1007/978-3-319-67468-1_2
- **[50]** John, A.K., Di Caro, L., Robaldo, L., Boella, G. Legalbot: A deep learning-based conversational agent in the legal domain (2017) *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10260 LNCS, pp. 267-273.
DOI: https://doi.org/10.1007/978-3-319-59569-6_32
- **[49]** Nanda, R., Siragusa, G., Di Caro, L., Theobald, M., Boella, G., Robaldo, L., Costamagna, F. Concept recognition in European and national law (2017) *Frontiers in Artificial Intelligence and Applications*, 302, pp. 193-198.
DOI: <https://doi.org/10.3233/978-1-61499-838-9-193>
- **[48]** Agnoloni, T., Bacci, L., Peruginelli, G., Van Opijnen, M., Van Den Oever, J., Palmirani, M., Cervone, L., Bujor, O., Arsuaga Lecuona, A., Boada García, A., Di Caro, L., Siragusa, G. Linking European case law: BO-ECLI parser, an open framework for the automatic extraction of legal links (2017) *Frontiers in Artificial Intelligence and Applications*, 302, pp. 113-118.
DOI: <https://doi.org/10.3233/978-1-61499-838-9-113>
- **[47]** John, A.K., Di Caro, L., Boella, G. A supervised keyphrase extraction system (2016) *ACM International Conference Proceeding Series*, 13-14-September-2016, art. no. 2993323, pp. 57-62.
DOI: <https://doi.org/10.1145/2993318.2993323>
- **[46]** Boella, G., Caro, L.D., Humphreys, L., Robaldo, L., Rossi, P., van der Torre, L. Eunomos, a legal document and knowledge management system for the Web to provide relevant, reliable and up-to-date information on the law (2016) *Artificial Intelligence and Law*, 24 (3), pp. 245-283.
DOI: <https://doi.org/10.1007/s10506-016-9184-3>
- **[45]** Di Caro, L., Boella, G. Automatic enrichment of WordNet with common-sense knowledge (2016) *Proceedings of the 10th International Conference on Language Resources and Evaluation, LREC 2016*, pp. 819-822.
- **[44]** Di Caro, L., Boella, G. Semantic similarity reasoning (2016) *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9577, pp. 127-138.
DOI: https://doi.org/10.1007/978-3-319-33500-1_11
- **[43]** John, A.K., Di Caro, L., Boella, G. NORMAS at SemEval-2016 task 1: Sem-sim: A multi-feature approach to semantic text similarity (2016) *SemEval 2016 - 10th International Workshop on Semantic Evaluation, Proceedings*, pp. 718-725.
- **[42]** Nanda, R., Di Caro, L., Boella, G. A text similarity approach for automated transposition detection of European Union directives (2016) *Frontiers in Artificial Intelligence and Applications*, 294, pp. 143-148.
DOI: <https://doi.org/10.3233/978-1-61499-726-9-143>
- **[41]** Adebayo, K.J., Boella, G., Di Caro, L. Neural reasoning for legal text understanding (2016) *Frontiers in Artificial Intelligence and Applications*, 294, pp. 175-178.
DOI: <https://doi.org/10.3233/978-1-61499-726-9-175>
- **[40]** Cataldi, M., Di Caro, L., Schifanella, C. Ranking researchers through collaboration pattern analysis (2016) *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9853 LNCS, pp. 50-54.
DOI: https://doi.org/10.1007/978-3-319-46131-1_11
- **[39]** Ajani, G., Boella, G., Di Caro, L., Robaldo, L., Humphreys, L., Praduroux, S., Rossi, P., Violato, A. The European Taxonomy Syllabus: A multi-lingual, multi-level ontology framework to untangle the web of European legal terminology (2016) *Applied Ontology*, 11 (4), pp. 325-375.
DOI: <https://doi.org/10.3233/A0-170174>

- journal
- proceedings

- **[38]** Boella, G., Di Caro, L., Graziadei, M., Cupi, L., Salaroglio, C.E., Humphreys, L., Konstantinov, H., Marko, K., Robaldo, L., Ruffini, C., Simov, K., Violato, A., Stroetmann, V. Linking legal open data: Breaking the accessibility and language barrier in European legislation and case law (2015) Proceedings of the International Conference on Artificial Intelligence and Law, 08-12-June-2015, pp. 171-175.
DOI: <https://doi.org/10.1145/2746090.2746106>
- **[37]** Di Caro, L., Cataldi, M., Lamolle, M., Schifanella, C. It is not what but who you know: A time-sensitive collaboration impact measure of researchers in surrounding communities (2015) WWW 2015 Companion - Proceedings of the 24th International Conference on World Wide Web, pp. 995-1000.
DOI: <https://doi.org/10.1145/2740908.2742023>
- **[36]** Cataldi, M., Di Caro, L., Schifanella, C. Twitter as a personalizable information service (2015) Multimedia Data Mining and Analytics: Disruptive Innovation, pp. 61-92.
DOI: https://doi.org/10.1007/978-3-319-14998-1_3
- **[35]** Di Caro, L., Ruggeri, A., Cupi, L., Boella, G. Common-sense knowledge for natural language understanding: Experiments in unsupervised and supervised settings (2015) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9336 LNCS, pp. 233-245.
DOI: https://link.springer.com/chapter/10.1007/978-3-319-24309-2_18
- **[34]** Humphreys, L., Santos, C., Di Caro, L., Boella, G., Van Der Torre, L., Robaldo, L. Mapping recitals to normative provisions in EU legislation to assist legal interpretation (2015) Frontiers in Artificial Intelligence and Applications, 279, pp. 41-49.
DOI: <https://ebooks.iospress.nl/publication/419751>
- **[33]** Cataldi, M., Di Caro, L., Schifanella, C. Improving search and navigation user experience by making use of social data (2014) Harnessing the Power of Social Media and Web Analytics, pp. 231-253.
DOI: <https://doi.org/10.4018/978-1-4666-5194-4.ch010>
- **[32]** Di Caro, L., Boella, G. Mining meaning from text by harvesting frequent and diverse semantic itemsets (2014) CEUR Workshop Proceedings, 1202, pp. 17-32.
- **[31]** Di Caro, L., Robaldo, L., Bersia, N. KnowNow: A serendipity-based educational tool for learning time-linked knowledge (2014) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8726 LNAI (PART 3), pp. 440-443.
DOI: https://doi.org/10.1007/978-3-662-44845-8_30
- **[30]** Boella, G., Di Caro, L., Violato, A., Robaldo, L. Exploiting networks in Law (2014) Proceedings of the 9th International Conference on Language Resources and Evaluation, LREC 2014, pp. 1654-1658.
- **[29]** Boella, G., Caro, L.D., Ruggeri, A., Robaldo, L. Learning from syntax generalizations for automatic semantic annotation (2014) Journal of Intelligent Information Systems, 43 (2), pp. 231-246.
DOI: <https://doi.org/10.1007/s10844-014-0320-9>
- **[28]** Boella, G., Di Caro, L., Robaldo, L., Violato, A. Network analysis to support navigation and use of ever-changing legislation (2014) CEUR Workshop Proceedings, 1105.
- **[27]** Cataldi, M., Lamolle, M., Caro, L.D., Schifanella, C. Tell me who your friends are and I'll tell you who you are: Studying the evolution of collaborations in research environments (2014) ASONAM 2014 - Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, art. no. 6921703, pp. 964-967.
DOI: <https://doi.org/10.1109/ASONAM.2014.6921703>

All Publications

- journal
- proceedings

- [25] Ghanavati, S., Humphreys, L., Boella, G., Di Caro, L., Robaldo, L., Van Der Torre, L. Compliance with multiple regulations (2014) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8824, pp. 415-422.
- [24] Cataldi, M., Caro, L.D., Schifanella, C. Personalized emerging topic detection based on a term aging model (2013) ACM Transactions on Intelligent Systems and Technology, 5 (1), art. no. 2542189
DOI: <https://doi.org/10.1145/2542182.2542189>
- [23] Boella, G., Di Caro, L. Supervised learning of syntactic contexts for uncovering definitions and extracting hypernym relations in text databases (2013) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8189 LNAI (PART 2), pp. 64-79.
DOI: https://doi.org/10.1007/978-3-642-40991-2_5
- [22] Boella, G., Di Caro, L., Rispoli, D., Robaldo, L. A system for classifying multi-label text into EuroVoc (2013) Proceedings of the International Conference on Artificial Intelligence and Law, pp. 239-240.
DOI: <https://doi.org/10.1145/2514601.2514635>
- [21] Robaldo, L., Di Caro, L. OpinionMining-ML (2013) Computer Standards and Interfaces, 35 (5), pp. 454-469.
DOI: <https://doi.org/10.1016/j.csi.2012.10.004>
- [20] Di Caro, L., Grella, M. Sentiment analysis via dependency parsing (2013) Computer Standards and Interfaces, 35 (5), pp. 442-453.
DOI: <https://doi.org/10.1016/j.csi.2012.10.005>
- [19] Boella, G., Di Caro, L., Robaldo, L. Semantic relation extraction from legislative text using generalized syntactic dependencies and support vector machines (2013) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8035, pp. 218-225.
DOI: https://doi.org/10.1007/978-3-642-39617-5_20
- [18] Ruggeri, A., Di Caro, L. How affordances can rule the (computational) world (2013) CEUR Workshop Proceedings, 1100, pp. 88-99.
- [17] Ruggeri, A., Di Caro, L. Linguistic affordances: Making sense of word senses (2013) CEUR Workshop Proceedings, 1100, pp. 136-139.
- [16] Boella, G., Di Caro, L. Extracting definitions and hypernym relations relying on syntactic dependencies and support vector machines (2013) ACL 2013 - 51st Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference, 2, pp. 532-537.
- [15] Robaldo, L., Di Caro, L., Antonini, A. SentiTagger - Automatically tagging text in OpinionMining-ML (2013) CEUR Workshop Proceedings, 1096, pp. 177-180.
- [14] Schifanella, C., Di Caro, L., Cataldi, M., Aufaure, M.-A. D-INDEX: A web environment for analyzing dependences among scientific collaborators (2012) Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1520-1523.
DOI: <https://doi.org/10.1145/2339530.2339770>
- [13] Candan, K.S., Di Caro, L., Sapino, M.L. PhC: Multiresolution visualization and exploration of text corpora with parallel hierarchical coordinates (2012) ACM Transactions on Intelligent Systems and Technology, 3 (2), art. no. 22, .
DOI: <https://doi.org/10.1145/2089094.2089098>
- [12] Di Caro, L., Cataldi, M., Schifanella, C. The d-index: Discovering dependences among scientific collaborators from their bibliographic data records (2012) Scientometrics, 93 (3), pp. 583-607.
DOI: <https://doi.org/10.1007/s11192-012-0762-1>

All Publications

- journal
- proceedings

- **[11]** Boella, G., Di Caro, L., Humphreys, L., Robaldo, L., Van Der Torre, L. NLP challenges for eunomos, a tool to build and manage legal knowledge (2012) Proceedings of the 8th International Conference on Language Resources and Evaluation, LREC 2012, pp. 3672-3678.
- **[10]** Boella, G., Di Caro, L., Lesmo, L., Rispoli, D., Robaldo, L. Multi-label classification of legislative text into eurovoc (2012) Frontiers in Artificial Intelligence and Applications, 250, pp. 21-30.
DOI: <https://doi.org/10.3233/978-1-61499-167-0-21>
- **[9]** Cataldi, M., Di Caro, L., Schifanella, C. ImmEx: IMMersive text documents EXploration system (2011) Proceedings - International Workshop on Content-Based Multimedia Indexing, art. no. 5972511, pp. 1-6.
DOI: <https://doi.org/10.1109/CBMI.2011.5972511>
- **[8]** Di Caro, L., Sapino, M.L., Selçuk Candan, K. Unraveling multi-dimensional data using pDView (2011) ACM International Conference Proceeding Series, pp. 570-573.
DOI: <https://doi.org/10.1145/1951365.1951443>
- **[7]** Di Caro, L., Selçuk Candan, K., Luisa Sapino, M. Navigating within news collections using tag-flakes (2011) Journal of Visual Languages and Computing, 22 (2), pp. 120-139.
DOI: <https://doi.org/10.1016/j.jv1c.2010.11.001>
- **[6]** Di Caro, L., Frias-Martinez, V., Frias-Martinez, E. Analyzing the role of dimension arrangement for data visualization in Radviz (2010) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6119 LNAI (PART 2), pp. 125-132.
DOI: https://doi.org/10.1007/978-3-642-13672-6_13
- **[5]** Cataldi, M., Di Caro, L., Schifanella, C. Emerging topic detection on Twitter based on temporal and social terms evaluation (2010) Proceedings of the 10th International Workshop on Multimedia Data Mining, MDMKDD '10, art. no. 4, .
DOI: <https://doi.org/10.1145/1814245.1814249>
- **[4]** Cataldi, M., Schifanella, C., Candan, K.S., Sapino, M.L., Di Caro, L. CoSeNa: A context-based search and navigation system (2009) Proceedings of the International Conference on Management of Emergent Digital EcoSystems, MEDES '09, pp. 218-225.
DOI: <https://doi.org/10.1145/1643823.1643864>
- **[3]** Di Caro, L., Jaimes, A. ClusTR: Exploring multivariate cluster correlations and topic trends (2009) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5782 LNAI (PART 2), pp. 722-725.
DOI: https://doi.org/10.1007/978-3-642-04174-7_49
- **[2]** Candan, K.S., Di Caro, L., Sapino, M.L. Creating tag hierarchies for effective navigation in social media (2008) International Conference on Information and Knowledge Management, Proceedings, pp. 75-82.
DOI: <https://doi.org/10.1145/1458583.1458597>
- **[1]** Di Caro, L., Candan, K.S., Sapino, M.L. Using tagFlake for condensing navigable tag hierarchies from tag clouds (2008) Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1069-1072.
DOI: <https://doi.org/10.1145/1401890.1402021>

Update

July 2023