# Curriculum Vitæ et Studiorum

Francesco Gullo



Date and Place of Birth:	March 17, 1982 — Cosenza, Italy
Citizenship:	Italian
Languages:	Italian (native), English (fluent), Spanish (conversant/proficient)
Address:	DISIM Department @UnivAQ, "Alan Turing" building, Room 203
	Via Vetoio, 41, 67100 Coppito (AQ) – Italy (office)
E-mail:	gullof@acm.org
ORCiD:	0000-0002-7052-1114
Web:	https://fgullo.github.io/
	LinkedIn ResearchGate Google Scholar
	Scopus GitHub Semantic Scholar

# Short Bio

Francesco Gullo is an associate professor of computer science at the University of L'Aquila (Italy), in the Department of Information Engineering, Computer Science, and Mathematics (DISIM). He received his PhD, in "Computer and Systems Engineering", from the University of Calabria (Italy), in 2010. During his PhD, he was an intern at the George Mason University (US), and a teaching/research assistant at the University of Catanzaro (Italy). After his graduation, he was a postdoc at the University of Calabria (Italy), a postdoc and a research scientist at the Yahoo Labs (Spain), a research scientist at the Fundació Barcelona Media (Spain), and a senior associate researcher at the UniCredit banking group (Italy).

His research falls into the broad areas of *artificial intelligence* and *data science*, with emphasis on *algorithmic* aspects. His recent interests include graph machine learning, graph data management, natural language processing, trustworthy AI. He is a co-author of 90+ publications, most of them in premier venues such as SIGMOD, VLDB, KDD, WWW, ICDM, CIKM, EDBT, WSDM, ECML-PKDD, SDM, TODS, TKDE, TKDD, MACH, DAMI, JCSS, TNSE, PR.

He has also been actively serving the scientific community: he was/is/will be Associate Editor of EPJ Data Science journal, Finance Chair of CIKM'24, Workshop Chair of KDD'24 and ICDM'16, Program co-Chair of MIDAS workshop @ECML-PKDD['16-'24], MultiClust symposium @SDM'14, MultiClust workshop @KDD'13, 3Clust workshop @PAKDD'12, as well as (senior) program-committee member of major conferences, including SIGMOD, KDD, WWW, IJCAI, AAAI, CIKM, SIGIR, ICDM, WSDM, SDM, ECML-PKDD, ECAI, ICWSM.

# CV IN A NUTSHELL

## ACADEMIA

- Italian National Scientific Habilitation (ASN), Associate Professor ("II fascia"), Sectors 01/B1 and 09/H1, since 2018 (valid until 2029).
- Indicators ("mediane") satisfied for ASN as a Full Professor ("I fascia"), for both Sectors 01/B1 and 09/H1.
- 18-year experience in *fundamental research* in *artificial intelligence* and *data science*, with focus on *algorithmic* aspects, and *graph machine learning/data management*, *natural language processing*, *trustworthy AI* as recent interests.
- Co-author of **90+ publications**, mostly in *premier* journals (TODS, PVLDB, TKDE, TKDD, MACH, DAMI, JCSS, TNSE, PR) and conferences (SIGMOD, VLDB, KDD, WWW, ICDM, CIKM, EDBT, WSDM, ECML-PKDD, SDM).
- H-index: 28; Total number of citations: 2401 (source: Google Scholar, Mar 25<sup>th</sup>, 2024).
- Top-2% Most Cited Scientist in 2022, according to Stanford University's standardized citation indicators.
- Recipient of the **best paper award** for the *BDTA'16 conference*.
- Wide network of (international) collaborations. Excluding editorship publications: 75 co-authors (source: dblp.org, Sep 4<sup>th</sup>, 2023), from Italian (Sapienza Univ., SNS, Unina, UniTrento, Uniba, Unical, Unicz, ISI Found, ENEA), US (Stanford Univ., CMU, NYU, Brown Univ., Boston Univ., USC, GMU, Boise State Univ.), and rest-of-the-world (ETH HKUST, HKU, NTU, Telecom Paristech, Aalto Univ., Aarhus Univ., AUEB) institutions; ~70% publications with at least one co-author from a different institution.
- Numerous teaching activities, carried out in both academia and industry.
- External Scientific Board Member of the ICT PhD Program of the University of Calabria, since 2017.
- UniCredit Scientific co-Referent for the Industrial Liason Program (ILP) of Sapienza University, since 2016.
- 30+ B.Sc./M.Sc./Ph.D. theses and internships (co-)supervised.
- Consolidated experience with externally-funded cooperative projects, both national and European.

#### SCIENTIFIC COMMUNITY SERVICE

- Finance Chair (role equivalent to General Chair) of the CIKM '24 conference.
- Workshop co-Chair of the ICDM '16 and KDD '24 conferences.
- **Program co-Chair** of several workshops/symposia/special sessions, including the well-established *MIDAS work-shop* @ECML-PKDD conference, now in its 8th edition.
- Stably part of the **program committee** of major conferences, including SIGMOD, KDD, WWW, IJCAI, AAAI, CIKM, SIGIR, ICDM, WSDM, SDM, ECML-PKDD, ECAI, ICWSM.
- Associate Editor of *EPJ DS* journal; Guest Editor of *ECML-PKDD Journal Track*; Distinguished Reviewer Board member of *TWEB* journal.
- Appointed Evaluator of several *cooperative project proposals*, funded by entities like the European Commission and the *Austrian Research Promotion Agency* (FFG).
- Best/outstanding reviewer award for KDD'23, WWW'23, ECML-PKDD'21, and WWW'17 conferences.

#### INDUSTRY

- 12-year experience in *industrial-research* R&D labs (Yahoo! Labs, Fundació Barcelona Media, UniCredit), closely interacting with industrial partners to make fundamental research impactful in an industry setting.
- Long-established portfolio of **industrial-research projects**, carried out until either a *proof-of-concept* or a *production-ready* stage, with focus on high *return on investment* (**ROI**) from different perspectives, including improved customer satisfaction/engagement, office automation, process optimization, decision support, cost saving.
- Co-inventor of 2 patents and co-author of 2 defensive publications.
- Recipient of the ABI IT Innovation Award 2018 ("Premio Innovazione IT Tecnologie Disruptive").
- Skilled in a high number of **technologies**, spanning domains such as general-purpose/domain-specific programming languages, versioning, build-automation, databases, virtualization, cloud computing, machine-learning frameworks, distributed computing, stream processing, search engines, web applications, etc.

# Work Experience (Appointments)

[Mar <sup>4</sup> 24 – present	] Associate Professor of Computer Sci University of L'Aquila DISIM Department L'Aquila, Italy	ence
[Sep'22 – Feb'24]	Researcher (senior associate) UniCredit banking group UniCredit (holding) company "Group Data & Intelligence" department Rome, Italy	<ul> <li>Design and implementation of cutting-edge, research-driven ICT solutions that are of interest to the banking industry. Main responsibilities:</li> <li>Business-driven, mid/short-term applied research</li> <li>Fundamental/applied research with longer-</li> </ul>
[Jul <sup>·</sup> 20 – Aug <sup>·</sup> 22]	Researcher (senior associate) UniCredit banking group UniCredit Services (controlled) company "AI, Data & Analytics ICT" department "Applied Research & Innovation" (ARI) unit Rome, Italy	<ul> <li>Fundamental/applied research with longer- term "transversal" applicability in multiple contexts</li> <li>"Quasi-production"-stage prototyping (i.e., development of SW prototypes that are ma- ture enough to be industrialized with mini- mal effort)</li> </ul>
[Jul <sup>•</sup> 15 – Jun <sup>•</sup> 20]	Researcher (associate) UniCredit banking group UniCredit (holding) company "Research & Development" (R&D) department Rome, Italy	<ul> <li>Supervising meeting</li> <li>Impact on the scientific community (papers, open-source SW libraries, collaboration with universities and research institutes, PC committees and editorial boards, organization of scientific events)</li> </ul>
[Mar'15 – Jun'15]	Research ScientistCarryi compaFundació Barcelona Media (now part of Eurecat)•Barcelona, Spain•	ng out fundamental and applied research that is impactful for ny's business needs. Main responsibilities: Fundamental/applied research research with longer-term "transversal" applicability in multiple contexts Business-driven, mid/short-term applied research, with the
[Sep'13 – Feb'15]	Research Scientist Yahoo Labs "Web Mining" group Barcelona, Spain	main goal of developing SW prototypes out of the conducted research, and interacting with production teams of the company to help industrialize them Supervising interns Getting patents out of the conducted research
[Sep'11 – Aug'13]	• Postdoctoral Researcher Yahoo Labs "Web Mining" group Barcelona, Spain	Participation to externally-funded research/innovation coopera- tive projects (both in the proposal-writing phase and in carrying out the projects' research activities) Impact on the scientific community (papers, collaboration with universities and research institutes, PC committees and edito- rial boards, organization of scientific events)
[Jan'10 – Aug'11]	<b>Postdoctoral Researcher</b> University of Calabria DEIS (now DIMES) department Cosenza, Italy	
[Apr'09 – Sep'09]	Short-term Visiting Scholar George Mason University "Data Mining & Machine Learning" group headed by Prof. Carlotta Domeniconi Fairfax, VA – USA	Carrying out fundamental research with the main goal of publishing papers; taking care of the experimental evalua- tion to support the conducted research; getting involved into teaching activities; helping in the supervision of students;
[Jan'06 – Aug'11]	Research & Teaching Assistant University of Calabria DEIS (now DIMES) department Cosenza, Italy	participation to externally-funded research/innovation coop- erative projects; participation to program committees of
[Jan'06 – Aug'11]	Research & Teaching Assistant "Magna Græcia" University of Catanzaro Deparment of Medicine and Surgery Catanzaro, Italy	3

## EDUCATION

[Jan'10]	<b>Ph.D.</b> in Computer and Systems Engineering [with highest honors] University of Calabria, Italy Thesis: Overcoming Uncertainty and the Curse of Dimensionality in Data Clustering Advisor: prof. Sergio Greco
[Dec'05]	M.Sc. in Computer Engineering [with highest honors] University of Calabria, Italy Thesis: Querying and Repairing Inconsistent XML Databases Advisor: prof. Sergio Greco
[Oct'03]	<b>B.Sc.</b> in Computer Engineering [with highest honors] University of Calabria, Italy Thesis: Semistructured Data and XML Advisor: prof. Sergio Greco

## RESEARCH ACTIVITY

His research falls into the broad areas of *artificial intelligence* and *data science*, with special emphasis on *algorithmic* aspects, i.e., on *formulating*, *theoretically characterizing*, and *designing effective yet efficient algorithms* for (novel) problems that are useful to gain insights/information/knowledge from data.

As far as data types, special emphasis has been given to graphs, text, and temporal data, but he has also dealt with tabular data, probabilistic data, and semistructured data. Large-scale data processing and combinatorial optimization are frequently-occurring keywords in his work.

More specifically, his recent research interests include:

- Graph machine learning [2, 3, 27]
  - graph representation learning; graph neural networks; machine/graph learning for combinatorial optimization; deep/reinforcement learning paradigms for graph clustering; knowledge graphs; graph classification;
- Graph data management [1, 4, 9, 10, 15, 17, 28, 29, 30, 33, 34, 37, 38, 39, 42, 43, 44, 45, 47, 48, 73, 83]
  - dense-subgraph discovery; graph summarization; reachability/distance queries on graphs; graph pattern mining; graph clustering; querying graph databases; link prediction;
- Natural Language Processing (NLP) [6, 35, 72, 75]
  - large language models; natural language generation; question answering; word embeddings; text classification; information extraction from text; sentiment analysis/transfer; text summarization; entity recognition and disambiguation;
- Ethics and trustworthiness in data management and AI (algorithmic fairness, explainable AI, interpretability, transparency, privacy, responsible AI, sustainable AI) [71]
- AI in finance [5, 32]

Other topics/problems he has focused on in the past include (*social*) Web mining (e.g., discovering polarized communities, social-network analysis, information propagation in online services, community search/detection, social-influence analysis, personalization of online services) [7, 11, 13, 16, 31, 36, 40, 46, 48, 77, 84, 86], mining high-dimensional/multifaceted data (i.e., projective/subspace clustering, clustering ensembles, projective clustering ensembles) [14, 19, 21, 49, 50, 52, 53, 54, 67, 76, 85], clustering uncertain data [12, 20, 22, 41, 51, 55, 74, 80, 89, 90], time-series data management [25, 78, 91], bioinformatics [23, 26, 79, 81, 82], XML data management [24, 70, 88].

### ORGANIZATION Finance Chair (role equivalent to General Chair)

32nd ACM International Conference on Information and Knowledge Management (CIKM '24)

#### Workshop co-Chair

30th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'24)

### Workshop co-Chair

16th IEEE International Conference on Data Mining (ICDM '16)

#### **Program co-Chair**

[1st, 2nd, 3rd, 4th, 5th, 6th, 7th, 8th] Workshop on MIning DAta for financial applicationS (MIDAS ['16, '17, '18, '19, '20, '21, '22, '23])

In conjunction with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECML-PKDD ['16, '17, '18, '19, '20, '21, '22, '23])

#### Industrial Session co-Chair

[2nd, 3rd] International Conference on Machine Learning, Optimization and Big Data (MOD ['16, '17])

#### Program co-Chair

Mini-Symposium on Multiple Clusterings, Multi-view Data, and Multi-source Knowledge-driven Clustering (MultiClust '14) In conjunction with the 201/ SIAM Int. Conf. on Data Mining (SDM (14))

In conjunction with the 2014 SIAM Int. Conf. on Data Mining  $({\bf SDM~'14})$ 

## Program co-Chair

4th MultiClust Workshop: Multiple Clusterings, Multi-view Data, and Multi-source Knowledge-driven Clustering In conjunction with the 19th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data

Mining (KDD '13)

## Program co-Chair

1st International Workshop on Multi-view data, High-dimensionality, External Knowledge: Striving for a Unified Approach to Clustering (3Clust '12)

In conjunction with the 16th Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD '12)

## EDITORIAL Associate Editor

BOARDS EPJ Data Science journal journal ranked Q1 according to https://www.scimagojr.com/ (Nov 23<sup>rd</sup>, 2023)

#### Associate Editor

Frontiers in Big Data journal Data mining and Management section

### **Guest Editor**

Journal Track of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD JT) [2020–2023]

## OTHER Distinguished Reviewer Board member

Boards

ACM Transactions on the Web (**TWEB**) Appointed for [2022–2024]

## Scientific Committee member

"Data Science for Economics and Finance: Methodologies and Applications" book Editors: S. Consoli, D. Reforgiato Recupero, M. Salsana. Publisher: Springer Nature. [2020] PROGRAMACM SIGMOD International Conference on Management of Data (SIGMOD)COMMITTEES(research track) [2022, 2024]

(MAJOR CONFERENCES) International World Wide Web Conference (**WWW**) (graph algorithms and learning for the Web track) [2014, 2016–2024]

ACM International Conference on Web Search and Data Mining (WSDM) [2013–2014, 2020–2024]

SIAM International Conference on Data Mining (SDM) [2012, 2014–2024]

ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD) (research track) [2014–2023]

IEEE International Conference on Data Mining (ICDM) [2014–2016, 2018–2023]

International Joint Conference on Artificial Intelligence (IJCAI) [2020–2023]

European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD) (research track) [2016–2019, 2021–2023]

European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD) (applied data science track) [2015–2018, 2021–2023]

European Conference on Artificial Intelligence (ECAI) (research track) [2023]

ACM International Conference on Information and Knowledge Management (CIKM) (research track, long papers) [2012–2022]

AAAI Conference on Artificial Intelligence (AAAI) [2021–2022]

International AAAI Conference on Web and Social Media (ICWSM) [2014, 2017–2022]

Int. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR) [2020–2021]

ACM International Conference on Information and Knowledge Management (CIKM) (research track, short papers) [2019–2020]

ACM International Conference on Information and Knowledge Management (CIKM) (industry/case studies track) [2017–2018]

 
 PROGRAM
 SEBD [2022–2023]; FRCCS [2023]; DSAA – special sessions track [2021–2022]; DSAA – appli-COMMITTEES

 COMMITTEES
 cations track [2019–2020]; DS [2022]; ASONAM [2015–2020]; NetSci [2020]; LOD [2017–2020]; CONFERENCES

 CONFERENCES
 EIDWT [2019]; SocInfo [2019]; KDIR [2016–2019]; ICoMS [2018]; ICBK [2017]; CIDM [2016]; EIDWT [2011]; Open-SOFOS Work.@ECML-PKDD [2022]; Soc2Net Work.@ASONAM [2019– 2020]; BSMDMA Work. @IEEE BigData and @IJCAI [2018–2019]; MATNet Work.@WWW SHOPS)

 SHOPS
 SoAPS Work.@ECIR [2018]; MultiClust Work.@SDM [2012]

INVITED	Network-based Receivable Financing
Talks	Scuola Normale Superiore
	Virtual Seminar, December 2021

	Dense Subgraph Discovery in Temporal Networks Workshop on Algorithmic Aspects of Temporal Graphs IV 48th Int. Colloq. on Automata, Languages, and Programming (ICALP '21) Virtual Event, July 2021
	Data Science on Information-rich Graphs: New Frontiers, Problems, and Methods University of Milan Milan, Italy, June 2019
	Mining Massive Complex Graphs: New Frontiers and Methods ISI Foundation Turin, Italy, September 2015
	From Patterns in Data to Knowledge Discovery: what Data Mining can do 3rd International Conference Frontiers in Diagnostic Technologies ( <b>ICFDT '13</b> ) Frascati, Italy, November 25-27, 2013
	Projective Clustering Ensembles Yahoo Labs Barcelona, Spain, September 2011
	Information-Theoretic Hierarchical Clustering of Uncertain Data George Mason University Fairfax VA, USA, April 2009
PhD Thesis Reviewing	<ul> <li>Stefano Piersanti</li> <li>Sapienza University – Dept. of Computer Engineering, Control and Management (DIAG)</li> <li>PhD Program in Engineering in Computer Science</li> <li>Thesis: Will it fail and why? A large case study of company default prediction with highly interpretable machine learning models</li> <li>Thesis advisor: Prof. Aris Anagnostopoulos</li> <li>XXXIII cycle</li> </ul>
External Reviewing	<b>Journals</b> : TKDE, TKDD, TPAMI, VLDBJ, MACH, DAMI, TIST, Information Systems, TWEB, TCSS, TBDATA, Pattern Recognition, Internet Mathematics, Nature Communications, AIRE, KAIS, Information Sciences, APNS, DKE, SAM, AI Communications, AI in Medicine, JOCS, Computational Intelligence, BDR, IJDSN, IJITDM, JINT, JIIS, EAAI, ADAC, Neural Networks, JDIQ, Information Fusion, IEEE Access, ISSI
	<b>Conferences</b> : SIGMOD, KDD, WWW, ICDE, ICDM, WSDM, CIKM, EDBT, SDM, ECML PKDD, PAKDD, DaWaK, CIDM, EDB, IDEAS, SEBD
Conference Participation	ECML-PKDD '22 [57, 58], CIKM '21, ECML-PKDD '21 [28, 59, 60], ECML-PKDD '20 [61], ECML-PKDD '19 [62], BigNOMICS '19, CIKM '18 [30, 31, 32], ECML-PKDD '18 [63], ECML-PKDD '17 [11, 64], WIMS '17 [72], IEEE BigData '16 [35], ECML-PKDD '16 [66], ECML-PKDD '14 [41], KDD '14 [42], EDBT '14 [44, 45], VLDB '12 [22], SEBD '11, SIGMOD '11 [50], ICDM '10 [52, 51], ICDM '09 [53], IDEAS '09 [78], SDM '09 [54], ICDM '08 [55], SUM '08 [80], CBMS '07 [82]

TEACHING COLLABORATIONS WITH UNIVERSITIES		<ul> <li>External Scientific Board Member</li> <li>Information and Communication Technologies PhD Program</li> <li>University of Calabria (Italy), Department of Computer Science, Modeling, Electronics and</li> <li>Systems Engineering (DIMES)</li> <li>Cycles: XXXVIII (years '22-'25), XXXVII (years '21-'24), XXXVI (years '20-'23), XXXV</li> <li>(years '19-'22), XXXIV (years '18-'21), XXXIII (years '17-'20)</li> <li>UniCredit Scientific co-Referent for the Industrial Liason Program (ILP)</li> <li>Sapienza University, M. Sc. in Data Science</li> <li>Years: 2016 – present</li> </ul>		
INDUSTRY COURSES Machine Learning I UniCredit – Advanced A Attendees: UniCredit em Description: theoretical f Syllabus: regression, clas learning Held on Jun-Jul 2021 Machine Learning II UniCredit – Advanced A Attendees: UniCredit em Description: theoretical f Syllabus: neural network networks, neural network Held on Oct-Nov 2021 Neural Natural Langu UniCredit – Group Data Attendees: UniCredit Em Description: theoretical f Syllabus: word embedding models, model pretraining Held on Jun-Jul 2023		hine Learning I (32 hours) redit – Advanced Analytics Lab, Data Skill Booster program idees: UniCredit employees (data scientists, machine-learning engineers) ription: theoretical foundations and hands-on sessions on machine-learning basics bus: regression, classification, decision trees, support vector machines, ensembles, unsupervised ing on Jun-Jul 2021		
		chine Learning II (32 hours) Credit – Advanced Analytics Lab, Data Skill Booster program ndees: UniCredit employees (data scientists, machine-learning engineers) ription: theoretical foundations and hands-on sessions on advanced machine-learning topics abus: neural networks, convolutional neural networks, recurrent neural networks, graph neural porks, neural networks for recommender systems on Oct-Nov 2021		
		ral Natural Language Processing (21 hours) Credit – Group Data & Intelligence, <i>Data Upskilling</i> program ndees: UniCredit Employees (data scientists, machine-learning engineers) cription: theoretical foundations and hands-on sessions on the latest advances in NLP abus: word embeddings, sequence models, large language models, attention models, autoregressive els, model pretraining 1 on Jun-Jul 2023		
University courses (Instructo	A Un R) De Ph Ac	dvanced Learning and Mining Problems in Big Graph Data (6 hours/year) niversity of Calabria (Italy) epartment of Computer Science, Modeling, Electronics and Systems Engineering (DIMES) hD Program in Information and Communication Technologies cademic year: '17-'18		
	C Uı Fa Ac	<ul> <li><i>computer Science</i> (55–84 hours/year)</li> <li>niversity of Calabria (Italy)</li> <li>aculties of Engineering, Pharmacy, and Political Sciences, B.Sc. 1st year (remedial class project)</li> <li>cademic years: '09-'10, '08-'09</li> </ul>		
	C "N Fa Ao	<i>Computer Science</i> (16–24 hours/year) Magna Græcia" University of Catanzaro, Crotone campus (Italy) aculty of Medicine and Surgery, B.Sc. in Nurse Sciences (3rd year) cademic years: '10-'11, '09-'10, '08-'09, '07-'08		
	La "N Fa Ao	aboratory Activity in Computer Science (16 hours/year) Magna Græcia" University of Catanzaro, Catanzaro campus (Italy) aculty of Medicine and Surgery, B.Sc. in Dental Health (3rd year) cademic years: '09-'10		

UNIVERSITY Foundations of Computer Science (18 hours/year) COURSES University of Calabria (Italy) (TEACHING Faculty of Engineering, B.Sc. in Computer Engineering (1st year) Assistant) Academic years: '10-'11, '09-'10, '08-'09 Computer Science (20 hours/year) University of Calabria (Italy) Faculty of Political Sciences, M.Sc. in Political Sciences (1st year) Academic years: '10-'11, '08-'09, '07-'08, '06-'07 Internet Algorithms and Cryptography (13 hours/year) University of Calabria (Italy) Faculty of Engineering, M.Sc. in Computer Engineering (2nd year) Academic years: '10-'11, '09-'10 Web-based Information Systems (13 hours/year) University of Calabria (Italy) Faculty of Engineering, M.Sc. in Computer Engineering (2nd year) Academic years: '09-'08, '08-'07, '06-'07, '05-'06 Data and Knowledge Bases (13 hours/year) University of Calabria (Italy) Faculty of Engineering, M.Sc. in Management Engineering (1st year) Academic years: '07-'08, '05-'06 Algorithms and Data Structures (13 hours/year) University of Calabria (Italy) Faculty of Engineering, B.Sc. in Computer Engineering (2nd year) Academic years: '06-'07 Foundations of Computer Science I (30 hours/year) "Magna Græcia" University of Catanzaro (Italy) B.Sc. interuniversity course in Computer and Biomedical Engineering (1st year) Academic years: '10-'11, '09-'10, '08-'09, '07-'08 Computer Science (8 hours/year) "Magna Græcia" University of Catanzaro, Catanzaro campus (Italy) Faculty of Medicine and Surgery, B.Sc. in Health Services (3rd year) Academic years: '09-'10 TRAINING Co-instructor of the "Knowledge Graph Completion" training camp (9 hours) CAMPS AND Sapienza University of Rome M.Sc. in Data Science HACKATONS Academic year: '20-'21 Liliana Martirano DOCTORAL ADVISORSHIPS University of Calabria Dept. of Computer Science, Modeling, Electronics and Systems Engineering (DIMES) PhD Program in Information and Communication Technologies Industrial Program ("Dottorati Innovativi con Caratterizzazione Industriale", PON R&I 2014-2020 Azione I.1) Thesis: Interpretable AI: Deep Learning for Knowledge Graphs XXXVI cycle (years 2020–2023)

B.Sc./M.Sc. Alessio Barboni

Thesis	Sapienza	a University of Rome, M.Sc. in Data Science (A.Y. '22-'23)
Advisorships	Thesis:	Temporal Graph Representation Learning meets Credit-risk Assessment

Giorgia Salvatori Sapienza University of Rome, M.Sc. in Data Science (A.Y. '21-'22) Thesis: Graph Summarization in the Financial Domain Francesco Ottaviani Sapienza University of Rome, M.Sc. in Computer Engineering (A.Y. '20-'21) Thesis: Characterizing Knowledge Graphs in the Embedding Space Dilara Isikli Sapienza University of Rome, M.Sc. in Data Science (A.Y. '20-'21) Thesis: Financial Sentiment Analysis Marco Moauro Roma Tre University, M.Sc. in Computer Engineering (A.Y. '20-'21) Thesis: Knowledge Graph Construction and Processing: a Case Study on the Italian Bourse Giacomo Lo Cascio Sapienza University of Rome, M.Sc. in Data Science (A.Y. '19-'20) Thesis: Representation Learning on Blockchain Networks, with an Application to Cryptocurrency Price Prediction Alessia Galli Sapienza University of Rome, M.Sc. in Data Science (A.Y. '19-'20) Thesis: Motif Extraction and Discord Discovery in Temporal Graphs Vlado Vukovic Sapienza University of Rome, M.Sc. in Data Science (A.Y. '18-'19) Thesis: Discord Discovery in Temporal Graphs Giulia Di Brango Sapienza University of Rome, M.Sc. in Data Science (A.Y. '18-'19) Thesis: Anomalous-Event Detection in Heterogeneous Graphs Lucia Gagliarducci Sapienza University of Rome, M.Sc. in Data Science (A.Y. '18-'19) Thesis: Relation-Extraction for Knowledge-Graph Completion Giacomo Legnaro Sapienza University of Rome, M.Sc. in Data Science (A.Y. '17-'18) Thesis: Efficient and Effective Streaming Algorithms for Network-based Invoice Factoring Giulia Gavazzi Sapienza University of Rome, M.Sc. in Data Science (A.Y. '16-'17) Thesis: Mining Dense and Non-redundant Subgraphs, with an Application to Event-graph Exploration Fabrizio Granieri University of Calabria, B.Sc. in Computer Engineering (A.Y. '09-'10) Thesis: Graph Partitioning for Clustering Ensembles Ronny Meringolo University of Calabria, B.Sc. in Computer Engineering (A.Y. '09-'10) Thesis: Graph Partitioning for Clustering Ensembles Antonio Senno University of Calabria, M.Sc. in Computer Engineering (A.Y. '08-'09) Thesis: Clustering Ensembles Methods Giuseppe Scrivano University of Calabria, B.Sc. in Computer Engineering (A.Y. '06-'07) Thesis: Multidimensional Time Series: Similarity Detection and Clustering

	Emanuele Forlano University of Calabria, B.Sc. + M.Sc. in Computer Engineering (A.Y. '05-'06) Thesis: Algorithms for Time Series Clustering
Research Intern Supervising	<ul> <li>Giulia Gavazzi, Giacomo Legnaro, Hugo Maldini</li> <li>M.Sc. students at Sapienza University of Rome</li> <li>Research Interns at UniCredit, Winter 2016 (Data Science Industrial Liaison Program)</li> <li>Project: News Collection for Emerging-Event Detection [72]</li> </ul>
	<ul> <li>Alessio Areni, Cristiano Di Crescenzo, Elena Troccoli</li> <li>M.Sc. students at Sapienza University of Rome</li> <li>Research Interns at UniCredit, Winter 2016 (Data Science Industrial Liaison Program)</li> <li>Project: Event Detection in Temporal Graphs [72]</li> </ul>
	Pranay Anchuri Ph.D. student at Rensselaer Polytechnic Institute Research Intern at Yahoo Labs Barcelona, Summer 2014 Project: <i>Mining Interesting Patterns from Uncertain Graphs</i>
	Daniele Ramazzotti Ph.D. student at University of Milano-Bicocca Research Intern at Yahoo Labs Barcelona, Summer 2014 Project: Social Influence Detection by Probabilistic Causation and Spatial Proximity [31]
	Edoardo Galimberti M.Sc. student at Politecnico of Milan Research Intern at Yahoo Labs Barcelona, Summer 2014 Project: <i>Efficient and Effective Community Search</i> [16]
	Natali Ruchansky Ph.D. student at Boston University Research Intern at Yahoo Labs Barcelona, Summer 2014 Project: The Minimum Wiener Connector Problem [39]
	Davide Mottin Ph.D. student at University of Trento Research Intern at Yahoo Labs Barcelona, Summer 2013 Project: <i>Query Reformulation in Graph Databases</i> [38]
	<ul> <li>Panos Parchas</li> <li>Ph.D. student at Hong Kong University of Science and Technology</li> <li>Research Intern at Yahoo Labs Barcelona, Summer 2013</li> <li>Project: Extracting Representative Instances of Uncertain Graphs [43, 15]</li> </ul>
	Lucrezia Macchia Ph.D. student at University of Bari Research Intern at Yahoo Labs Barcelona, Winter 2013 Project: <i>Mining Summaries of Propagations</i> [46]
	Arijit Khan Ph.D. student at University of California at Santa Barbara Research Intern at Yahoo Labs Barcelona, Summer 2012 Project: Fast Reliability Search on Uncertain Graphs [45]
	Charalampos Tsourakakis Ph.D. student at Carnegie Mellon University Research Intern at Yahoo Labs Barcelona, Summer 2012 Project: <i>Extracting optimal quasi-cliques with quality guarantees</i> [47]

# EXTERNALLY-FUNDED RESEARCH/INNOVATION COOPERATIVE PROJECTS

Participation as a Research- staff Member	<b>The Rome Technopole Ecosystem</b> Funding entity: European Union - NextGenerationEU, National Recovery and Resilience Plan (NRRP) (" <i>Piano Nazionale di Ripresa e Resilienza</i> " ( <i>PNRR</i> )) Innovation ecosystem number: ECS_00000024		
	<b>TYPES</b> (Towards transparencY and Privacy in the onlinE advertising business) Funding entity: EU-H2020 Project number: H2020-DS-2014-1 – Project ID 653449		
	<b>SUPER</b> (Social sensors for secUrity Assessments and Proactive EmeRgencies management) Funding entity: EU-FP7 Project number: FP7-Security-2013.6.1-1 (http://super-fp7.eu/)		
	Cenit Social Media Funding entity: Spanish Centre for the Development of Industrial Technology CENIT program, project CEN-20101037 (http://www.cenitsocialmedia.es/)		
	GeoPKDD (Geographic Privacy-aware Knowledge Discovery and Delivery) Funding entity: EU-FET-014915 Project number: IST-6FP-014915 (http://www.geopkdd.eu)		
	LogNET (An Innovative Network to Improve Logistics in Gioia Tauro) Subproject LOGICA (laboratory of LOGIstics in CAlabria) Funding entity: POR Calabria 2000-2006		
	<b>EUREKA!</b> An idea for energy Project: <i>Fraud Detection via Load Profiles</i> Funding entity: Enel University, area "Distribution Nets of Electricity"		
PARTICIPATION IN	<b>PDS-OBDA</b> : Protected Data Sharing with OBDA Target funding programme and entity: ICT-13-2018-2019, European Commission		
PROPOSAL DRAFTING (PROJECTS NOT ACCEPTED)	<b>ALGO-AUDIT</b> : Network for Training Algorithm Auditors for Privacy-Preserving, Fair and Transparent Data-Driven Decisions Target funding programme and entity: H2020-MSCA-ITN-2018, European Commission		
	<b>SeDaSOMA</b> : <u>Se</u> rendipitous, <u>Da</u> ta-as-a- <u>S</u> ervice-oriented, <u>O</u> pen big data <u>M</u> anagement and <u>A</u> nalytics framework Target funding programme and entity: H2020-ICT-16-2015, European Commission		
Appointed Evaluator of Project Proposals	<ul> <li>Funding entity: European Commission</li> <li>Funding programme: Horizon Europe</li> <li>Proposal category: Research &amp; Innovation Action (RIA)</li> <li>Call(s): 2022</li> </ul>		
	Funding entity: <b>European Commission</b> Funding programme: <b>Horizon 2020 ICT</b> Proposal category: Innovation Action (IA) Call(s): 2017		
	<ul> <li>Funding entity: Austrian Research Promotion Agency (FFG)</li> <li>Funding programme: ICT of the Future</li> <li>Proposal category: Industrial Research, Experimental Development</li> <li>Call(s): 4th (2015), 5th (2016), 6th (2017)</li> </ul>		

Funding entity: Austrian Research Promotion Agency (FFG)
Funding programme: Beyond Europe
Proposal category: Industrial Research, Experimental Development Call(s): 1st (2016), 2nd (2017)
Funding entity: Austrian Research Promotion Agency (FFG) & Chinese Academy of Sciences (CAS)

Funding programme: **ICT of the Future – Bilateral China Call** Proposal category: Industrial Research, Experimental Development Call(s): 1st (2017)

# HONORS AND AWARDS

## Italian National Scientific Habilitation (Abilitazione Scientifica Nazionale (ASN))

Granted by: Italian Ministry of Education, Universities and Research (*MIUR*) Role: Associate Professor (*II fascia*) Sector 01/B1 - Computer science (*Informatica*), valid from Aug 7<sup>th</sup>, 2018 to Aug 7<sup>th</sup>, 2029 Sector 09/H1 - Information processing systems (*Sistemi di elaborazione delle informazioni*), valid from Jul 26<sup>th</sup>, 2018 to Jul 26<sup>th</sup>, 2029

## KDD 2023 Certificate of Excellence in Reviewing

Granted to reviewers "in recognition of an outstanding contribution to the quality of KDD"

### WWW 2023 Best Reviewer Award

Granted to reviewers "selected by track chairs of the research track from the Program Committee and Senior Program Committee, based on the quality of their reviews, and their engaged participation during the discussion phase"

#### Top-2% Most Cited Scientist across all Scientific Disciplines

Updated science-wide author databases of standardized citation indicators Stanford University (https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/4)

Year: 2022

## ECML-PKDD 2021 Best Reviewer Award

Granted to selected reviewers for "their timely and insightful reviews and for actively participating in the discussions"

**ABI IT Innovation Award 2018** ("Premio Innovazione IT – Tecnologie Disruptive") Granted by "Associazione Bancaria Italiana" (ABI), the trade association of Italian banks Awarded system: UC Balance (an innovative system for network-based invoice factoring) Personal contribution: design, theoretical characterization, and implementation of the algorithmic core behind the system

## WWW 2017 Outstanding Reviewer Award

Granted to selected reviewers "who went well beyond their call of duty, provided exceptional reviews, and contributed a lot to discussions"

**BDTA 2016 Best Paper Award** Paper "Handling Uncertainty in Clustering Art-exhibition Visiting Styles" [74]

"BC" label ("paper where the results are backed by code") from the 2015 Arizona University Repeatability in Computer Science study (http://reproducibility.cs.arizona.edu/v2) Paper "Uncertain Centroid based Partitional Clustering of Uncertain Data" [22]

## SIGMOD 2011 Repeatability and Workability Evaluation (RWE) label

Paper "Advancing Data Clustering via Projective Clustering Ensembles" [50]

## SDM 2009 Student Travel Award

Paper "Diversity-based Weighting Schemes for Clustering Ensembles" [54]

Italian Innovation Award (Premio Nazionale per l'Innovazione) - I edition (2009) Fraud Detection via Load Profiles project within EUREKA! An idea for energy initiative

## (Italian) Habilitation to practice as an Engineer

Section A, Field of Information Technology Qualifying exam passed on October 2006 (1st 2006 session), no expiration date

## Tuition fee reimbursement (based *solely* on merit)

M.Sc. in Computer Engineering University of Calabria, 2003

## INDUSTRIAL ACTIVITY

Patents	H. Vahabi, F. Gullo Method and/or System for Recommender System Patent Granted [US10796239B2, 2020-10-06]
	X. Bai, B. B. Cambazoglu, F. Gullo, A. Mantrach, F. Silvestri Using Exogenous Sources for Personalization of Website Services Patent Granted [US9792372B2, 2017-10-10]
	F. Bonchi, A. Gionis, F. Gullo, A. Ukkonen Method and System for Computing Relationship-Constrained Shortest Paths in Large Networks Defensive Publication
	F. Bonchi, A. Gionis, F. Gullo, C. Tsourakakis Method and System for Extracting High-quality Subgraphs from Large Graphs Defensive Publication
Industrial Projects (selected)	Large-Language-Model-based Search on Internal Sharepoint Content Outcome: PoC Skills: large language models; semantic search; Technology: OpenAI; Azure Workbench
	Graph-based Credit-risk Assessment Outcome: PoC Skills: (temporal) graph representation learning; supervised learning Technology: Palantir Foundry, Python, Docker, PyTorch
	Prediction of Credit Card Installments Outcome: production Skills: supervised learning, sequential learning, reinforcement learning Technology: Palantir Foundry, Python, PySpark
	Building and Analyzing a Knowledge Graph of HW/SW Resources Outcome: PoC Skills: graph databases, (knowledge) graph analytics Technology: Neo4j, Python, Java

Building and Analyzing a Knowledge Graph of Corporates Product: PoC Skills: graph databases, (knowledge) graph analytics Technology: Neo4j, Python, Java

A Complete NLP Pipeline for Processing Garnishment Documents Outcome: production Skills: NLP, supervised learning, word embeddings, OCR Technology: Scala, Python, Weka, Stanford CoreNLP, SpaCy, OpenCV, Tesseract

Network-based Invoice Factoring Outcome: production Skills: combinatorial optimization, graph analytics, network flow, cycle enumeration Technology: Scala

Email Classification for Talent Acquisition Outcome: prototype Skills: NLP, supervised learning Technology: Java, Weka, JavaMail, EWS Java API

Identification of Crisis Events for Business Continuity Outcome: prototype Skills: NLP, entity recognition and disambiguation, sentiment analysis, text categorization Technology: Scala, Spark, HBase

Entity Recognition and Disambiguation for Financial News Outcome: PoC Skills: NLP, entity recognition and disambiguation, MinHash, locality-sensitive hashing (LSH) Technology: Scala, Spark, HBase

Detecting Corporate Group Structure from Unstructured Data Outcome: production Skills: web scraping, learning from positive and unlabeled examples, NLP, entity recognition and disambiguation, computer vision Technology: Scala, Spark, Weka, OpenCV

Exploiting Search History of Users for News Personalization Outcome: production Skills: personalization of online services, query-log analysis, distributed computing Technology: Java, Hadoop

Predicting the Next Application a User is Going to Use Outcome: PoC Skills: supervised learning, recommender systems, mobile data, distributed computing Technology: Java, Hadoop, Hive

Churn Analysis on Flickr Outcome: PoC Skills: churn analysis, supervised learning, distributed computing Technology: Java, Matlab, Hadoop

Optimal User Engagement through Social-Diffusion-aware Recommender Systems Outcome: production Skills: recommender systems, information diffusion in complex networks, distributed computing Technology: Java, Hadoop

Ticker Similarity Search Outcome: prototype Skills: online similarity search, time-series analysis, locality-sensitive hashing (LSH) Technology: Java Identification of Historical Trends in Organic Search via Query-log Analysis Outcome: prototype Skills: query-log analysis, distributed computing Technology: Java, Hadoop

OPEN SOURCE PyCCAlg - Correlation-Clustering Algorithms in Python https://github.com/fgullo/pyccalg Prototypes (CO-)DEVELOPED Skills: correlation clustering, linear programming (SELECTED) Technology: Python, SciPy, PuLP JPCE – Java-based Projective Clustering Ensembles tool http://people.dimes.unical.it/andreatagarelli/software/jpce/ https://github.com/fgullo/jpce Skills: clustering ensembles, subspace clustering, projective clustering ensembles Technology: Java JCLUdata – a Java software for clustering uncertain data https://github.com/fgullo/jcludata Skills: clustering, uncertain data, probability theory, information theory Technology: Java Hermes – A distributed-messaging tool for NLP http://hermes.rnd.unicredit.it:9603 Skills: NLP, graph analytics Technology: Scala, Spark, Kafka, Redis, ElasticSearch, Hbase MSPtool – Mass Spectra Preprocessing tool http://polifemo.deis.unical.it/~gtradigo/jnlp/msptool/ Skills: bioinformatics, mining of biological data, time-series analysis Technology: Java, JFreeChart, Java Web Start, J2EE TSCtool - Time Series Clustering tool Skills: time-series analysis, mining of spatio/temporal data Technology: Java, JFreeChart

## TECHNICAL SKILLS (NON-EXHAUSTIVE)

GENERAL-PURPOSE PROGRAMMING LANGUAGES: Java, Python, Scala, C/C++

Domain-specific languages:

MATLAB (matrix programming), SQL (relational DB querying), HTML (web-page markup), XML (machine-readable doc markup), Bash (shell scripting), LaTeX (typesetting), Markdown (text formatting), Prolog, Datalog (logic programming)

VERSIONING TECHNOLOGY: Git, SVN

BUILD-AUTOMATION TECHNOLOGY: SBT, Apache Maven

DATABASES:

 $MySQL,\ PostgreSQL\ ({\rm RDBMSs}),\ Neo4j\ ({\rm graph}\ {\rm DBs}),\ Redis,\ Couchbase\ Server\ ({\rm non-relational}\ {\rm DBs})$ 

VIRTUALIZATION TECHNOLOGY: Docker

CLOUD-COMPUTING SUITES: Google Cloud Platform, Microsoft Azure

MACHINE-LEARNING FRAMEWORKS: PyTorch, PyTorch Geometric

KNOWLEDGE-GRAPH-EMBEDDING FRAMEWORKS: PyKEEN, OpenKE

DISTRIBUTED-COMPUTING TECHNOLOGY: Apache Hadoop, Apache Spark, Apache HBase, Apache Hive

STREAM-PROCESSING TECHNOLOGY: Apache Kafka

DATA-MANAGEMENT PLATFORMS: Palantir Foundry

SEARCH ENGINES: Elasticsearch

INTERACTIVE-COMPUTING TECHNOLOGY: Jupyter Notebook

WEB-APPLICATION TECHNOLOGY: Java EE, Java servlet, JavaServer Pages (JSP), Jboss, Apache Tomcat

SOFTWARE LIBRARIES:

Stanford CoreNLP, Weka, JDBC, JavaMail, JFreeChart (Java); spaCy, NumPy, SciPy, pandas, NetworkX, OpenCV, scikit-learn, Hugging Face Transformers (Python)

OPERATING SYSTEMS: macOS, Windows, Unix/Linux

QUANTUM COMPUTING: IBM Qiskit SDK

# Selected Publications

#### Journals

- O. Balalau, F. Bonchi, T-H. Hubert Chan, F. Gullo, M. Sozio, H. Xie. Finding Subgraphs with Maximum Total Density and Limited Overlap inWeighted Hypergraphs. ACM Transactions on Knowledge Discovery from Data (TKDD), 18(4): 1-21, 2024 (DOI: https://doi.org/10.1145/3639410).
- [2] F. Gullo, D. Mandaglio, A. Tagarelli. A Combinatorial Multi-Armed Bandit Approach to Correlation Clustering. *Data Mining and Knowledge Discovery* (DAMI), 37(4): 1630-1691, 2023 (DOI: https://doi.org/10.1007/s10618-023-00937-5).
- [3] J. Layne, J. Carpenter, E. Serra, F. Gullo. Temporal SIR-GN: Efficient and Effective Structural Representation Learning for Temporal Graphs. *Proceedings of the VLDB Endowment* (**PVLDB**), 16(9):2075-2089, 2023 (DOI: https://dx.doi.org/10.14778/3598581. 3598583).
- [4] E. Galimberti, M. Ciaperoni, A. Barrat, F. Bonchi, C. Cattuto, F. Gullo. Span-core Decomposition for Temporal Networks: Algorithms and Applications. ACM Transactions on Knowledge Discovery from Data (TKDD), 15(1): 2:1-2:44, 2021 (DOI: https: //dx.doi.org/10.1145/3418226).
- [5] I. Bordino, F. Gullo, G. Legnaro. Advancing Receivable Financing via a Network-based Approach. *IEEE Transactions on Network Science and Engineering* (TNSE), 8(2):1328-1337, 2021 (DOI: http://dx.doi.org/10.1109/TNSE.2020.3005612).
- [6] I. Bordino, A. Ferretti, F. Gullo, S. Pascolutti. GarNLP: A Natural Language Processing Pipeline for Garnishment Documents. *Information Systems Frontiers* (ISFI), 23(1):101-114, 2021 (DOI: http://dx.doi.org/10.1007/s10796-020-09997-0).
- [7] M. Ciaperoni, E. Galimberti, F. Bonchi, C. Cattuto, F. Gullo, A. Barrat. Relevance of temporal cores for epidemic spread in temporal networks. *Scientific Reports* (SciRep), 10, Article number: 12529, 2020 (DOI: https://dx.doi.org/10.1038/s41598-020-69464-3).
- [8] E. Galimberti, F. Bonchi, F. Gullo, T. Lanciano. Core Decomposition in Multilayer Networks: Theory, Algorithms, Applications. ACM Transactions on Knowledge Discovery from Data (TKDD), 14(1):11:1-11:40, 2020 (DOI: http://dx.doi.org/10.1145/ 3369872).
- F. Bonchi, I. Bordino, F. Gullo, G. Stilo. The Importance of Unexpectedness: Discovering Buzzing Stories in Anomalous Temporal Graphs. Web Intelligence (WI), 17(3):177-198, 2019 (DOI: http://dx.doi.org/10.3233/WEB-190412).

- [10] A. Khan, F. Bonchi, F. Gullo, A. Nufer. Conditional Reliability in Uncertain Graphs. *IEEE Transactions on Knowledge and Data Engineering* (TKDE), 30(11):2078-2092, 2018 (DOI: http://dx.doi.org/10.1109/TKDE.2018.2816653).
- [11] A. Tagarelli, A. Amelio, F. Gullo. Ensemble-based Community Detection in Multilayer Networks. Data Mining and Knowledge Discovery (DAMI), 31(5):1506-1543, 2017 (DOI: http://dx.doi.org/10.1007/s10618-017-0528-8).
- [12] F. Gullo, G. Ponti, A. Tagarelli, S. Greco. An Information-Theoretic Approach to Hierarchical Clustering of Uncertain Data. *Information Sciences*, 402:199-215, 2017 (DOI: http://dx.doi.org/10.1016/j.ins.2017.03.030).
- [13] X. Bai, B. Barla Cambazoglu, F. Gullo, A. Mantrach, F. Silvestri. Exploiting Search History of Users for News Personalization. *Information Sciences*, 385-386:125-137, 2017 (DOI: http://dx.doi.org/10.1016/j.ins.2016.12.038).
- [14] I. Assent, C. Domeniconi, F. Gullo, A. Tagarelli, A. Zimek. MultiClust 2013: Multiple Clusterings, Multi-view Data, Multi-source Knowledge-driven Clustering. SIGKDD Explorations, 18(1):35-38, 2016 (DOI: http://dx.doi.org/10.1145/2980765.2980769).
- [15] P. Parchas, F. Gullo, D. Papadias, F. Bonchi. Uncertain Graph Processing through Representative Instances. ACM Transactions on Database Systems (TODS), 40(3):20:1-20:39, 2015 (DOI: http://dx.doi.org/10.1145/2818182).
- [16] N. Barbieri, F. Bonchi, E. Galimberti, F. Gullo. Efficient and Effective Community Search. Data Mining and Knowledge Discovery (DAMI), 29(5):1406-1433, 2015 (DOI: http://dx. doi.org/10.1007/s10618-015-0422-1).
- [17] F. Bonchi, A. Gionis, F. Gullo, C. Tsourakakis, A. Ukkonen. Chromatic Correlation Clustering. ACM Transactions on Knowledge Discovery from Data (TKDD), 9(4):34:1-34:24, 2015 (DOI: http://dx.doi.org/10.1145/2728170).
- F. Gullo. From Patterns in Data to Knowledge Discovery: What Data Mining Can Do. *Physics Procedia*, 62:18-22, 2015 (DOI: http://dx.doi.org/10.1016/j.phpro.2015. 02.005).
- [19] F. Gullo, C. Domeniconi, A. Tagarelli. Metacluster-based Projective Clustering Ensembles. *Machine Learning*, 98(1-2):181-216, 2015 (DOI: http://dx.doi.org/10.1007/s10994-013-5395-y).
- [20] F. Gullo, G. Ponti, A. Tagarelli. Minimizing the Variance of Cluster Mixture Models for Clustering Uncertain Objects. *Statistical Analysis and Data Mining* (SAM), 6(2):116-135, 2013 (DOI: http://dx.doi.org/10.1002/sam.11170).
- [21] F. Gullo, C. Domeniconi, A. Tagarelli. Projective Clustering Ensembles. Data Mining and Knowledge Discovery (DAMI), 26(3):452-511, 2013 (DOI: http://dx.doi.org/10.1007/ s10618-012-0266-x).
- [22] F. Gullo, A. Tagarelli. Uncertain Centroid based Partitional Clustering of Uncertain Data. Proceedings of the VLDB Endowment (PVLDB), 5(7):610-621, 2012.
- [23] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, P. Veltri. A Time Series Approach for Clustering Mass Spectrometry Data. *Journal of Computational Science* (JOCS), 3(5):344-355, 2012 (DOI: http://dx.doi.org/10.1016/j.jocs.2011.06.008).
- [24] S. Greco, F. Gullo, G. Ponti, A. Tagarelli. Collaborative Clustering of XML Documents. Journal of Computer and System Sciences (JCSS), 77(6):988-1008, 2011 (DOI: http: //dx.doi.org/10.1016/j.jcss.2011.02.005).
- [25] F. Gullo, G. Ponti, A. Tagarelli, S. Greco. A Time Series Representation Model for Accurate and Fast Similarity Detection. Pattern Recognition, 42(11):2998-3014, 2009 (DOI: http://dx.doi.org/10.1016/j.patcog.2009.03.030).

[26] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, P. Veltri. MaSDA: A System for Analyzing Mass Spectrometry Data. Computer Methods and Programs in Biomedicine (CMPB), 95(2 suppl.):S12-S21, 2009 (DOI: http://dx.doi.org/10.1016/j.cmpb.2009.02.011).

Conferences

- [27] M. K. Egger, W. Ma, D. Mottin, P. Karras, I. Bordino, F. Gullo, A. Anagnostopoulos. Relik: A Reliability Measure for Knowledge Graph Embeddings. In *Proceedings of the* ACM International Web Conference (WWW '24). Singapore, May 13-17, 2024.
- [28] D. Mandaglio, A. Tagarelli, F. Gullo. Correlation Clustering with Global Weight Bounds. In Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD '21). pp. 499-515. Bilbao, Spain, September 13-17, 2021.
- [29] D. Mandaglio, A. Tagarelli, F. Gullo. In and Out: Optimizing Overall Interaction in Probabilistic Graphs under Clustering Constraints. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '20), pp. 1371-1381. San Diego, CA – USA, August 23-27, 2020.
- [30] E. Galimberti, A. Barrat, F. Bonchi, C. Cattuto, F. Gullo. Mining (maximal) Spancores from Temporal Networks. In *Proceedings of the ACM International Conference on Knowledge and Information Management* (CIKM '18), pp. 107-116. Turin, Italy, October 22-26, 2018.
- [31] F. Bonchi, F. Gullo, B. Mishra, D. Ramazzotti. Probabilistic Causal Analysis of Social Influence. In Proceedings of the ACM International Conference on Knowledge and Information Management (CIKM '18), pp. 1003-1012. Turin, Italy, October 22-26, 2018.
- [32] I. Bordino, F. Gullo. Network-based Receivable Financing. In Proceedings of the ACM International Conference on Knowledge and Information Management (CIKM '18), pp. 2137-2145. Turin, Italy, October 22-26, 2018.
- [33] E. Galimberti, F. Bonchi, F. Gullo. Core Decomposition and Densest Subgraph in Multilayer Networks. In Proceedings of the ACM International Conference on Knowledge and Information Management (CIKM '17), pp. 1807-1816. Singapore, November 6-10, 2017.
- [34] N. Ruchansky, F. Bonchi, D. Garcia Soriano, F. Gullo, N. Kourtellis. To Be Connected, or Not to Be Connected: That is the Minimum Inefficiency Subgraph Problem. In Proceedings of the ACM International Conference on Knowledge and Information Management (CIKM '17), pp. 879-888. Singapore, November 6-10, 2017.
- [35] I. Bordino, A. Ferretti, M. Firrincieli, F. Gullo, M. Paris, S. Pascolutti, G. Sabena. Advancing NLP via a distributed-messaging approach. In *Proceedings of the 2016 IEEE International Conference on Big Data* (IEEE BigData '16), pp. 1561-1568. Washington D.C., USA, December 5-8, 2016.
- [36] H. Vahabi, I. Koutsopoulos, F. Gullo, M. Halkidi. DifRec: a Social-Diffusion-aware Recommender System. In Proceedings of the ACM International Conference on Knowledge and Information Management (CIKM '15), pp. 1481-1490. Melbourne, Australia, October 19-23, 2015.
- [37] A. Khan, F. Gullo, T. Wohler, F. Bonchi. Top-k Reliable Edge Colors in Uncertain Graphs. In Proceedings of the ACM International Conference on Knowledge and Information Management (CIKM '15), pp. 1851-1854. Melbourne, Australia, October 19-23, 2015.
- [38] D. Mottin, F. Bonchi, F. Gullo. Graph Query Reformulation with Diversity. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15), pp. 825-834. Sydney, Australia, August 10-13, 2015.
- [39] N. Ruchansky, F. Bonchi, D. Garcia Soriano, F. Gullo, N. Kourtellis. The Minimum Wiener Connector Problem. In Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '15), pp. 1587-1602. Melbourne, Victoria, Australia,

May 31-June 14, 2015.

- [40] O. Balalau, F. Bonchi, T-H. Hubert Chan, F. Gullo, M. Sozio. Finding Subgraphs with Maximum Total Density and Limited Overlap. In *Proceedings of the International Conference on Web Search and Data Mining* (WSDM '15), pp. 379-388. Shanghai, China, February 2-6, 2015.
- [41] F. Gullo, G. Ponti, A. Tagarelli. Be certain of how-to before mining uncertain data. In Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD '14) (Nectar Track), pp. 489-493. Nancy, France, September 15-19, 2014.
- [42] F. Bonchi, F. Gullo, A. Kaltenbrunner, Y. Volkovich. Core Decomposition of Uncertain Graphs. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '14), pp. 1316-1325. New York City, New York (USA), August 24-27, 2014.
- [43] P. Parchas, F. Gullo, D. Papadias, F. Bonchi. The Pursuit of a Good Possible World: Extracting Representative Instances of Uncertain Graphs. In *Proceedings of the ACM SIGMOD International Conference on Management of Data* (SIGMOD '14), pp. 967-978. Snowbird, Utah (USA), June 22-27, 2014.
- [44] F. Bonchi, A. Gionis, F. Gullo, A. Ukkonen. Distance oracles in edge-labeled graphs. In Proceedings of the International Conference on Extending Database Technology (EDBT '14), pp. 547-558. Athens, Greece, March 24-28, 2014.
- [45] A. Khan, F. Bonchi, A. Gionis, F. Gullo. Fast Reliability Search in Uncertain Graphs. In Proceedings of the International Conference on Extending Database Technology (EDBT '14), pp. 535-546. Athens, Greece, March 24-28, 2014.
- [46] L. Macchia, F. Bonchi, F. Gullo, L. Chiarandini. Mining Summaries of Propagations. In Proceedings of the IEEE International Conference on Data Mining (ICDM '13), pp. 498-507. Dallas, Texas (USA), December 7-10, 2013.
- [47] C. E. Tsourakakis, F. Bonchi, A. Gionis, F. Gullo, M. A. Tsiarli. Denser than the densest subgraph: extracting optimal quasi-cliques with quality guarantees. In *Proc. of the ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining* (KDD '13), pp. 104-112. Chicago, Illinois (USA), August 11-14, 2013.
- [48] F. Bonchi, A. Gionis, F. Gullo, A. Ukkonen. Chromatic Correlation Clustering. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '12), pp. 1321-1329. Beijing, China, August 12-16, 2012.
- [49] F. Gullo, A. K. A. Talukder, S. Luke, C. Domeniconi, A. Tagarelli. Multiobjective Optimization of Co-Clustering Ensembles. In *Proceedings of the Genetic and Evolutionary Computation Conference* (GECCO '12), pp. 1495-1496. Philadelphia, PA (USA), July 7-11, 2012.
- [50] F. Gullo, C. Domeniconi, and A. Tagarelli. Advancing Data Clustering via Projective Clustering Ensembles. In Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data (SIGMOD '11), pp. 733-744. Athens, Greece, June 12-16, 2011.
- [51] F. Gullo, G. Ponti, A. Tagarelli. Minimizing the Variance of Cluster Mixture Models for Clustering Uncertain Objects. In *Proceedings of the 10th IEEE International Conference* on Data Mining (ICDM '10), pp. 839-844. Sydney, Australia, December 14-17, 2010.
- [52] F. Gullo, C. Domeniconi, A. Tagarelli. Enhancing Single-Objective Projective Clustering Ensembles. In *Proceedings of the 10th IEEE International Conference on Data Mining* (ICDM '10), pp. 833-838. Sydney, Australia, December 14-17, 2010.
- [53] F. Gullo, C. Domeniconi, A. Tagarelli. Projective Clustering Ensembles. In Proceedings of the 9th IEEE International Conference on Data Mining (ICDM '09), pp. 794-799. Miami,

Florida (USA), December 6-9, 2009.

- [54] F. Gullo, A. Tagarelli, S. Greco. Diversity-based Weighting Schemes for Clustering Ensembles. In *Proceedings of the 9th SIAM International Conference on Data Mining* (SDM '09), pp. 437-448. Sparks, Nevada (USA), April 30-May 2, 2009.
- [55] F. Gullo, G. Ponti, A. Tagarelli, S. Greco. A Hierarchical Algorithm for Clustering Uncertain Data via an Information-Theoretic Approach. In *Proceedings of the 8th IEEE International Conference on Data Mining* (ICDM '08), pp. 821-826. Pisa, Italy, December 15-19, 2008.

BOOKS

[56] F. Bonchi, D. Garcia Soriano, F. Gullo. Correlation Clustering. Synthesis Lectures on Data Mining and Knowledge Discovery, Morgan & Claypool Publishers, 2022 (DOI: http: //dx.doi.org/10.2200/S01163ED1V01Y202201DMK019).

# OTHER PUBLICATIONS

EDITED [57] Koprinska et al. Machine Learning and Principles and Practice of Knowledge Discovery in VOLUMES Databases – International Workshops of ECML PKDD 2022, Proceedings, Part I. Communications in Computer and Information Science (CCIS) 1752, Springer, 2023, ISBN 978-3-031-23617-4. [58] Koprinska et al. Machine Learning and Principles and Practice of Knowledge Discovery in Databases – International Workshops of ECML PKDD 2022, Proceedings, Part II. Communications in Computer and Information Science (CCIS) 1753, Springer, 2023, ISBN 978-3-031-23632-7. [59] Kamp et al. Machine Learning and Principles and Practice of Knowledge Discovery in Databases – International Workshops of ECML PKDD 2021, Proceedings, Part I. Communications in Computer and Information Science (CCIS) 1524, Springer, 2021, ISBN 978-3-030-93735-5. [60] Kamp et al. Machine Learning and Principles and Practice of Knowledge Discovery in Databases - International Workshops of ECML PKDD 2021, Proceedings, Part II. Communications in Computer and Information Science (CCIS) 1525, Springer, 2021, ISBN 978-3-030-93732-4. [61] V. Bitetta, I. Bordino, A. Ferretti, F. Gullo, G. Ponti, L. Severini. Mining Data for Financial Applications - 5th ECML PKDD Workshop, MIDAS 2020, co-located with the 2020 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECML-PKDD 2020), Ghent, Belgium, September 18, 2020. Lecture Notes in Computer Science (LNCS) 12591, Springer, 2021, ISBN 978-3-030-66980-5. [62] V. Bitetta, I. Bordino, A. Ferretti, F. Gullo, S. Pascolutti, G. Ponti. Mining Data for Financial Applications – 4th ECML PKDD Workshop, MIDAS 2019, co-located with the 2019 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECML-PKDD 2019), Würzburg, Germany, September 16, 2019. Lecture Notes in Computer Science (LNCS) 11985, Springer, 2020, ISBN 978-3-030-37719-9. [63] C. Alzate, A. Monreale, L. Bioglio, V. Bitetta, I. Bordino, G. Caldarelli, A. Ferretti, R. Guidotti, F. Gullo, S. Pascolutti, R. G. Pensa, C. Robardet, T. Squartini. ECML PKDD 2018 Workshops – MIDAS 2018 and PAP 2018, Proceedings, co-located with the 2018 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECML-PKDD 2018), Dublin, Ireland, September 10-14, 2018. Lecture Notes in Computer Science (LNCS) 11054, Springer, 2019, ISBN 978-3-030-13462-4.

[64] I. Bordino, G. Caldarelli, F. Fumarola, F. Gullo, T. Squartini. Proceedings of the First

Workshop on MIning DAta for financial applicationS (MIDAS 2017), co-located with the 2017 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECML-PKDD 2017), Skopje, Macedonia, September 18, 2017. CEUR Workshop Proceedings Vol-1941, CEUR-WS.org, 2017.

- [65] C. Domeniconi, F. Gullo, F. Bonchi, J. Domingo-Ferrer, R. Baeza-Yates, Z.-H. Zhou, X. Wu 16th IEEE International Conference on Data Mining Workshops (ICDMW 2016), Barcelona, Spain, December 12, 2016. IEEE Computer Society, 2016, ISBN 978-1-5090-5472-5.
- [66] I. Bordino, G. Caldarelli, F. Fumarola, F. Gullo, T. Squartini. Proceedings of the First Workshop on MIning DAta for financial applicationS (MIDAS 2016), co-located with the 2016 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECML-PKDD 2016), Riva del Garda, Italy, September 19, 2016. CEUR Workshop Proceedings Vol-1774, CEUR-WS.org, 2016.
- [67] I. Assent, C. Domeniconi, F. Gullo, A. Tagarelli, A. Zimek. MultiClust '13: Proceedings of the 4th MultiClust Workshop on Multiple Clusterings, Multi-view Data, Multi-source Knowledge-driven Clustering, co-located with the KDD '13 conference, Chicago, Illinois (USA), August 11 - 14, 2013. ACM, 2013, ISBN 978-1-4503-2334-5.
- [68] T. Washio, J. Luo, P. Desikan, K.-W. Hsu, J. Srivastava, E.-P. Lim, M. Teisseire, M. Roche, C. Domeniconi, F. Gullo, A. Tagarelli, H. K. Tan, W. C. Onn. *Emerging Trends in Knowledge Discovery and Data Mining PAKDD 2012 International Workshops: DMHM, GeoDoc, 3Clust, DSDM, Kuala Lumpur, Malaysia, May 29 June 1, 2012, Revised Selected Papers.* Lecture Notes in Computer Science 7769, Springer, 2013, ISBN 978-3-642-36777-9.
- [69] F. Bonchi, F. Gullo, A. Kaltenbrunner Core Decomposition of Massive, Information-Rich Graphs. In Encyclopedia of Social Network Analysis and Mining - 2 edition, Springer, 2018 (DOI: https://doi.org/10.1007/978-1-4939-7131-2\_110176).
  - [70] F. Gullo, G. Ponti, S. Greco. Organizing XML Documents on a Peer-to-Peer Network by Collaborative Clustering. In XML Data Mining: Models, Methods, Applications, IGI Global, 2012, pp. 449-466 (DOI: http://dx.doi.org/10.4018/978-1-61350-356-0. ch018).
- Other International Conferences

Воок

CHAPTERS (REFEREED)

- [71] F. Gullo, L. La Cava, D. Mandaglio, A. Tagarelli. When Correlation Clustering Meets
   Fairness Constraints. In *Proceedings of the Int. Conf. on Discovery Science* (DS '22).
   Montpellier, France, October 10-12, 2022.
  - [72] C. Di Crescenzo, G. Gavazzi, G. Legnaro, E. Troccoli, I. Bordino, F. Gullo. HERMEVENT: A News Collection for Emerging-Event Detection. In *Proceedings of the Int. Conf. on Web Intelligence, Mining and Semantics* (WIMS '17). Amantea, Italy, June 19-22, 2017.
  - [73] F. Bonchi, I. Bordino, F. Gullo, G. Stilo. Identifying Buzzing Stories via Anomalous Temporal Subgraph Discovery. In *Proceedings of the IEEE/WIC/ACM International Conference* on Web Intelligence (WI '16), pp. 161-168. Omaha, Nebraska, USA, October 13-16, 2016.
  - [74] F. Gullo, G. Ponti, A. Tagarelli, S. Cuomo, P. De Michele, F. Piccialli, Handling Uncertainty in Clustering Art-exhibition Visiting Styles. In *Proceedings of the 7th EAI International Conference on Big Data Technologies and Applications* (BDTA '16). Seoul, South Korea, November 17-18, 2016.
  - [75] I. Bordino, A. Ferretti, M. Firrincieli, F. Gullo, M. Paris, G. Sabena. Hermes: A Distributed-Messaging Tool for NLP. In *Proceedings of the 2nd International Conference* on Machine Learning, Optimization, Big Data (MOD '16), pp. 402-406. Volterra, Italy, August 26-29, 2016.
  - [76] S. Romeo, A. Tagarelli, F. Gullo, S. Greco. A Tensor-based Clustering Approach for Multiple Document Classifications. In Proceedings of the International Conference on Pattern

Recognition Applications and Methods (ICPRAM '13), pp. 200-205. Barcelona, Spain, February 15-18, 2013.

- [77] A. Tagarelli, F. Gullo. Evaluating PageRank Methods for Structural Sense Ranking in Labeled Tree Data. In *Proceedings of the Int. Conf. on Web Intelligence, Mining and Semantics* (WIMS '12). Craiova, Romania, June 13-15, 2012.
- [78] F. Gullo, G. Ponti, A. Tagarelli, S. Iiritano, M. Ruffolo, D. Labate. Low-voltage Electricity Customer Profiling based on Load Data Clustering. In *Proceedings of the 13th International Database Engineering & Applications Symposium* (IDEAS '09), pp. 330-333. Cetraro, Italy, September 16-18, 2009.
- [79] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, P. Veltri. Hierarchical Clustering of Microarray Data with Probe-level Uncertainty. In *Proceedings of the 22th IEEE International Symposium on Computer-Based Medical Systems* (CBMS '09). Albuquerque, New Mexico (USA), August 3-4, 2009.
- [80] F. Gullo, G. Ponti, A. Tagarelli. Clustering Uncertain Data via K-medoids. In Proceedings of the 2nd International Conference on Scalable Uncertainty Management (SUM '08), pp. 229-242. Napoli, Italy, October 1-3, 2008.
- [81] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, P. Veltri. MSPtool: A Versatile Tool for Mass Spectrometry Data Preprocessing. In *Proceedings of the 21th IEEE International* Symposium on Computer-Based Medical Systems (CBMS '08), pp. 209-214. Jyväskylä, Finland, June 17-19, 2008.
- [82] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, P. Veltri. A Time Series Based Approach for Classifying Mass Spectrometry Data. In *Proceedings of the 20th IEEE International Sympo*sium on Computer-Based Medical Systems (CBMS '07), pp. 412-417. Maribor, Slovenia, June 20-23, 2007.
- [83] D. Mandaglio, A. Tagarelli, F. Gullo. Correlation Clustering: from Local to Global Constraints. In Proceedings of the 30th Italian Symposium on Advanced Database Systems (SEBD '22). Tirrenia (Pisa), Italy, June 19-22, 2022.
  - [84] X. Bai, B. B. Cambazoglu, F. Gullo, A. Mantrach, F. Silvestri. Improving News Personalization Through Search Logs. In *Proceedings of the 1st Workshop on Social Aspects in Personalization and Search* (SoAPS '18), in conjunction with the 40th European Conference on Information Retrieval (ECIR '18). Grenoble, France, March 26-29, 2018.
  - [85] C. Domeniconi, F. Gullo, A. Tagarelli. The Projective Clustering Ensemble Problem for Advanced Data Clustering. In *Proceedings of the 25th Italian Symposium on Advanced Database Systems* (SEBD '17). Squillace Lido (Catanzaro), Italy, June 25-29, 2017.
  - [86] A. Freire, M. Manca, D. Saez-Trumper, D. Laniado, I. Bordino, F. Gullo, A. Kaltenbrunner. Graph-Based Breaking News Detection on Wikipedia. In Proceedings of the 2016 International Wiki Workshop, in conjunction with the 10th International AAAI Conference on Web and Social Media (Wiki@ICWSM '16). Cologne, Germany, May 17, 2016.
  - [87] S. Greco, F. Gullo, G. Ponti, A. Tagarelli, G. Agapito. Clustering XML Documents: a Distributed Collaborative Approach. In *Proceedings of the 18th Italian Symposium on* Advanced Database Systems (SEBD '10), pp. 406-413. Rimini, Italy, June 20-23, 2010.
  - [88] S. Greco, F. Gullo, G. Ponti, A. Tagarelli. Collaborative Clustering of XML Documents. In Proceedings of the 1st International Workshop on Distributed XML Processing: Theory and Practice (DXP '09), in conjunction with the 38th International Conference on Parallel Processing (ICPP '09). Vienna, Austria, September 22-25, 2009.
  - [89] F. Gullo, G. Ponti, A. Tagarelli, S. Greco. Information-Theoretic Hierarchical Clustering of Uncertain Data. In Proc. of the 17th Italian Symposium on Advanced Database Systems (SEBD '09), pp. 273-280. Geneva, Italy, June 21-24, 2009.

Workshops & National Conferences

23

- [90] F. Gullo, G. Ponti. Hierarchical Clustering of Uncertain Data. In *Doctoral Symposium* in conjunction with the 14th GII Doctoral School on Advances in Databases, Cetraro, Italy, September 19, 2009.
- [91] F. Gullo, G. Ponti, A. Tagarelli, S. Greco. Accurate and Fast Similarity Detection in Time Series. In Proceedings of the 15th Italian Symposium on Advanced Database Systems (SEBD '07), pp. 172-183. Bari, Italy, June 17-20, 2007.