

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: UniCredit, Via Molfetta, 101, 00171 Rome - Italy (*office*)
E-mail: gullof@acm.org
ORCID: 0000-0002-7052-1114
Web: <https://fgullo.github.io/>
<https://www.linkedin.com/in/francesco-gullo-2717a17>
<https://www.researchgate.net/profile/Francesco-Gullo3>
<http://scholar.google.it/citations?user=ueZsy3wAAAAJ&hl=en>
<https://www.scopus.com/authid/detail.uri?authorId=22034355400>
<https://github.com/fgullo>
<https://bit.ly/2UgZtgu>

SHORT BIO

Francesco Gullo is a researcher at UniCredit, R&D department. He received his PhD from the University of Calabria, Italy, in 2010. During his PhD, he was an intern at the George Mason University, Fairfax VA, USA. Before joining UniCredit, he spent 1.5 years in the University of Calabria, Italy (as a postdoc), and 4 years in the Yahoo Labs, Spain (as a postdoc first, and as a research scientist then).

His research is focused on *algorithmic data science*, i.e., on *formulating novel problems* to gain insights/information/knowledge from data, *theoretically characterizing* them, and *designing/analyzing effective yet efficient algorithms* for their solution. As far as data types, special emphasis has been given to *graphs*, *text*, and *time series*, but he has also dealt with Euclidean data, probabilistic data, and semistructured data. *Large-scale data processing* and *combinatorial optimization* are frequently-occurring keywords in his work. His research has been published in premier venues in the areas of *data mining*, *machine learning*, *databases*, *Web data management*, and *NLP*, including SIGMOD, VLDB, KDD, ICDM, CIKM, EDBT, WSDM, ECML-PKDD, SDM, TODS, TKDE, TKDD, Machine Learning, DAMI, JCSS, Pattern Recognition.

He has also been active in serving the data-science scientific community, by, e.g., being Workshop Chair of ICDM'16, organizing workshops/symposia (MIDAS workshop @ECML-PKDD[16-19], MultiClust mini-symposium @SDM'14, MultiClust workshop @KDD'13, 3Clust workshop @PAKDD'12), or being part of the program committee of major conferences (KDD, WWW, IJCAI-PRICAI, ICDM, WSDM, ECML-PKDD, CIKM, SIGIR, ICWSM, SDM).

WORK EXPERIENCE

[Jul'15—Present]	Researcher UniCredit, Rome, Italy R&D department	<p>Design and implementation of novel, research-driven technological solutions that are of interest to the banking industry:</p> <ul style="list-style-type: none">• Business-driven, mid/short-term applied research• Fundamental research/applied research with longer-term “horizontal” applicability (i.e., potentially applicable in multiple contexts)• “Quasi-production”-stage prototyping (i.e., development of SW prototypes that are mature enough to be industrialized with minimal effort)• Supervising interns• 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 (conferences, workshops, sessions))
[Sep'13 – Jun'15]	Research Scientist Yahoo Labs, Barcelona, Spain Web Mining group	<p>Carrying out fundamental research and applied research that can be impactful for the business needs of the company:</p> <ul style="list-style-type: none">• Fundamental research/applied research with longer-term “horizontal” applicability (i.e., potentially applicable in multiple contexts)• Business-driven, mid/short-term applied research, with the 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• Participation to externally-funded research/innovation cooperative 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 editorial boards, organization of scientific events (conferences, workshops, sessions))
[Sep'11 – Aug'13]	Postdoctoral Researcher Yahoo Labs, Barcelona, Spain Web Mining group	
[Jan'10 – Aug'11]	Postdoctoral Researcher University of Calabria, Italy Dept. of Electronics, Computer, and Systems Science (DEIS, currently DIMES)	
[Apr'09 – Sep'09]	Short-term Visiting Scholar George Mason University, Fairfax, US Data Mining & Machine Learning group headed by Prof. Carlotta Domeniconi	<p>Carrying out fundamental research with the main goal of publishing papers; taking care of the experimental evaluation to support the conducted research; getting involved into teaching activities; helping in the supervision of students; participation to externally-funded research/innovation cooperative projects; participation to PC committees; helping in the organization of scientific events.</p>
[Jan'06 – Dec'09]	Research & Teaching Assistant University of Calabria, Italy Dept. of Electronics, Computer, and Systems Science (DEIS, currently DIMES)	

EDUCATION

- [Jan'10] **Ph.D.** 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.Sc.** in Computer Engineering [*with highest honors*]
University of Calabria, Italy
Thesis: *Semistructured Data and XML*
Advisor: prof. Sergio Greco

RESEARCH ACTIVITY

His research is focused on *algorithmic data science*, 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 time series*, but he has also dealt with relational data, probabilistic data, and semistructured data. *Large-scale data processing and combinatorial optimization* are frequently-occurring keywords in his work.

More specifically, he has primarily contributed to the areas of *data mining, machine learning, databases, Web data management, and NLP*, on the following main topics/problems:

- Graph analytics, graph querying, graph mining [3, 4, 9, 11, 21, 23, 24, 25, 28, 29, 30, 33, 34, 35, 36, 38, 39, 57]
 - graph clustering; dense-subgraph discovery; mining/querying information-rich graphs (e.g., uncertain/multi-layer/temporal/heterogeneous graphs); reachability/distance queries on graphs; querying graph databases; graph pattern mining;
- Natural Language Processing (NLP) [1, 26, 56, 59]
 - entity recognition and disambiguation; distributional semantic embeddings
- (Social) Web mining [5, 7, 10, 22, 27, 31, 37, 39, 61, 68]
 - information propagation in online services; community search/detection; personalization of online services; social-influence analysis; ranking and centrality
- Mining high-dimensional and multifaceted data [8, 13, 15, 40, 41, 43, 44, 45, 52, 60, 67]
 - projective/subspace clustering; clustering ensembles; projective clustering ensembles; document clustering leveraging multiple views/classifications
- Uncertainty in data mining & machine learning [6, 14, 16, 32, 42, 46, 58, 64, 71, 72]
- Time-series data management [19, 62, 73]
- Bioinformatics [17, 20, 63, 65, 66]
 - querying/mining biological networks; modeling/mining proteomic data; modeling/mining gene expression data
- XML data management [18, 55, 69, 70]
 - querying and repairing inconsistent XML databases; XML document clustering

INDUSTRIAL ACTIVITY

PATENTS

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]

H. Vahabi, F. Gullo
Method and/or System for Recommender System
Published Patent Application [US20170061310A1, 2017-03-02]

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)

A Complete NLP Pipeline for Processing Garnishment Documents

Product: UniCredit PIG/C2E

Skills: NLP, supervised learning, word embeddings, OCR

Technology: Scala, Python, Weka, Stanford CoreNLP, SpaCy, OpenCV, Tesseract

Network-based Invoice Factoring

Product: UniCredit Balance

Skills: combinatorial optimization, graph analytics, network flow, cycle enumeration

Technology: Scala

Email Classification for Talent Acquisition

Product: UniCredit Aladin

Skills: NLP, supervised learning

Technology: Java, Weka, JavaMail, EWS Java API

Identification of Crisis Events for Business Continuity

Product: UniCredit Hermes Crisis

Skills: NLP, entity recognition and disambiguation, sentiment analysis, text categorization

Technology: Scala, Spark, HBase

Entity Recognition and Disambiguation for Financial News

Product: UniCredit Hermes News

Skills: NLP, entity recognition, entity disambiguation, MinHash, locality-sensitive hashing (LSH)

Technology: Scala, Spark, HBase

Detecting Corporate Group Structure from Unstructured Data

Product: UniCredit EGM (Enhanced Group Mapping)

Skills: web scraping, learning from positive and unlabeled examples, NLP, entity recognition, entity disambiguation, computer vision

Technology: Scala, Spark, Weka, OpenCV

News Categorization for Crisis Event Detection

Product: UniCredit Hermes Crisis

Skills: NLP, entity recognition, entity disambiguation, text mining

Technology: Scala, Spark

Exploiting Search History of Users for News Personalization

Product: Yahoo News

Skills: personalization of online services, query-log analysis, big data processing

Technology: Java, Hadoop

Predicting the Next Application a User is Going to Use

Product: Yahoo Aviate

Skills: machine-learning predictive analytics, recommender systems, mobile data management, feature engineering, big data processing

Technology: Java, Hadoop, Hive

Churn Analysis on Flickr

Product: Flickr

Skills: churn analysis, predictive analytics, big data processing

Technology: Java, Matlab, Hadoop

Optimal User Engagement through Social-Diffusion-aware Recommender Systems

Product: Tumblr

Skills: recommender systems, information diffusion in complex networks, big data processing

Technology: Java, Hadoop

Ticker Similarity Search

Product: Yahoo Finance

Skills: online similarity search, time series analysis, time series motif discovery, locally-sensitive hashing (LSH)

Technology: Java

Recommendation of Items and Streams in Heterogeneous Information Networks

Product: Flickr

Skills: recommender systems, heterogeneous information networks

Technology: Java, Hadoop

Identification of Historical Trends in Organic Search via Query-log Analysis

Product: Yahoo Sales

Skills: query-log analysis, big data processing

Technology: Java, Hadoop

OPEN SOURCE/PUBLIC
PROTOTYPES
(CO-)DEVELOPED
(SELECTED)

Hermes – A distributed-messaging tool for NLP

<http://hermes.rnd.unicredit.it:9603>

Skills: NLP, graph analytics

Technology: Scala, Spark, Kafka, Redis, ElasticSearch, Hbase

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

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

SCIENTIFIC COMMUNITY SERVICE

ORGANIZATION	<p>Workshop co-Chair <i>16th IEEE International Conference on Data Mining (ICDM ‘16)</i></p> <p>Program co-Chair <i>[1st, 2nd, 3rd, 4th] Workshop on Mining Data for financial applications (MIDAS [‘16, ‘17, ‘18, ‘19])</i> <i>In conjunction with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECML-PKDD [‘16, ‘17, ‘18, ‘19])</i></p> <p>Industrial Session co-Chair <i>[2nd, 3rd] International Conference on Machine Learning, Optimization and Big Data (MOD [‘16, ‘17])</i></p> <p>Program co-Chair <i>Mini-Symposium on Multiple Clusterings, Multi-view Data, and Multi-source Knowledge-driven Clustering (MultiClust ‘14)</i> <i>In conjunction with the 2014 SIAM Int. Conf. on Data Mining (SDM ‘14)</i></p> <p>Program co-Chair <i>4th MultiClust Workshop: Multiple Clusterings, Multi-view Data, and Multi-source Knowledge-driven Clustering</i> <i>In conjunction with the 19th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD ‘13)</i></p> <p>Program co-Chair <i>1st International Workshop on Multi-view data, High-dimensionality, External Knowledge: Striving for a Unified Approach to Clustering (3Clust ‘12)</i> <i>In conjunction with the 16th Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD ‘12)</i></p>
EDITORIAL BOARDS	<p><i>Journal Track of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD JT)</i> [2020]</p>
SCIENTIFIC COMMITTEES	<p><i>“Data Science for Economics and Finance: Methodologies and Applications” book</i> Editors: S. Consoli, D. Reforgiato Recupero, M. Salsana. Publisher: Springer Nature. [2020]</p>
PROGRAM COMMITTEES	<p><i>ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD)</i> [2014–2020]</p> <p><i>International World Wide Web Conference (WWW)</i> [2014, 2016–2020]</p> <p><i>IEEE International Conference on Data Mining (ICDM)</i> [2014–2016, 2018–2020]</p> <p><i>ACM International Conference on Information and Knowledge Management (CIKM)</i> (Research Track, long papers) [2012–2020]</p> <p><i>International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)</i> [2020]</p> <p><i>SIAM International Conference on Data Mining (SDM)</i> [2012, 2014–2020]</p>

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

International AAAI Conference on Web and Social Media (ICWSM)
[2014, 2017–2020]

International Joint Conference on Artificial Intelligence – Pacific Rim International Conference on Artificial Intelligence (IJCAI-PRICAI) (Special Track on AI in FinTech)
[2020]

IEEE/ACM Int. Conf. on Advances in Social Networks Analysis and Mining (ASONAM)
[2015–2020]

IEEE International Conference on Data Science and Advanced Analytics (DSAA)
(Application Track) [2019–2020]

International Conference on Network Science (NetSci)
[2020]

International Conference on machine Learning, Optimization and big Data (LOD)
[2017–2020]

International Workshop on Modeling and mining Social-Media-driven Complex Networks (Soc2Net), in conjunction with the *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*
[2019–2020]

ACM International Conference on Information and Knowledge Management (CIKM)
(Research Track, short papers)
[2019]

European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD) (Research Track)
[2016–2019]

International Conference on Computational Data and Social Networks (CSoNet)
[2019]

International Conference on Social Informatics (SocInfo)
[2019]

International Conference on Knowledge Discovery and Information Retrieval (KDIR)
[2016–2019]

International Workshop on Big Social Media Data Management and Analysis (BSMDMA), in conjunction with the *International IEEE Conference on Big Data (IEEE BigData)* and *International Joint Conference on Artificial Intelligence (IJCAI)*
[2018–2019]

ACM International Conference on Information and Knowledge Management (CIKM)
(Industry/Case Studies Track)
[2017–2018]

European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD) (Applied Data Science Track)
[2015–2018]

International Conference on Mathematics and Statistics (ICoMS)
[2018]

International Workshop on Mining Attributed Networks (MATNet), in conjunction with the *International World Wide Web Conference (WWW)*
[2018]

Workshop on Social Aspects in Personalization and Search (SoAPS), in conjunction with the *European Conference on Information Retrieval (ECIR)*
[2018]

IEEE International Conference on Big Knowledge (ICBK)
[2017]

IEEE Symposium on Computational Intelligence and Data Mining (CIDM)
[2016]

3rd MultiClust Workshop: Discovering, Summarizing and Using Multiple Clusterings, in conjunction with the *SIAM International Conf. on Data Mining (SDM '12)*

Int. Conf. on Emerging Intelligent Data and Web Technologies (EIDWT '11)

INVITED TALKS

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 (**ICFDT '13**)
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

EXTERNAL REFEREEING

Journals: TKDE, TKDD, TPAMI, VLDBJ, Machine Learning, DAMI, TIST, Information Systems, TWEB, TCSS, Pattern Recognition, Internet Mathematics, 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

Conferences: SIGMOD, KDD, WWW, ICDE, ICDM, WSDM, CIKM, EDBT, SDM, ECML PKDD, PAKDD, DaWaK, CIDM, EDB, IDEAS, SEBD

CONFERENCE PARTICIPATION

ECML-PKDD '19 [47], BigNOMICS '19, CIKM '18 [21, 22, 23], ECML-PKDD '18 [48], ECML-PKDD '17 [5, 49], WIMS '17 [56], IEEE BigData '16 [26], ECML-PKDD '16 [51], ECML-PKDD '14 [32], KDD '14 [33], EDBT '14 [35, 36], VLDB '12 [16], SEBD '11, SIGMOD '11 [41], ICDM '10 [43, 42], ICDM '09 [44], IDEAS '09 [62], SDM '09 [45], ICDM '08 [46], SUM '08 [64], CBMS '07 [66]

EXTERNALLY-FUNDED RESEARCH/INNOVATION COOPERATIVE PROJECTS

PARTICIPATION AS A RESEARCH-STAFF MEMBER

TYPES (*Towards transparencY and Privacy in the onlinE advertising business*)
Funding entity: EU-H2020
Project number: H2020-DS-2014-1 – Project ID 653449

SUPER (*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

EUREKA! An idea for energy

Project: *Fraud Detection via Load Profiles*

Funding entity: Enel University, area “Distribution Nets of Electricity”

EVALUATOR OF
PROJECT PROPOSALS

Funding entity: **European Commission**

Funding programme: **Horizon 2020 ICT**

Proposal category: Innovation Action (IA)

Call(s): 2017

Funding entity: Austrian Research Promotion Agency (**FFG**)

Funding programme: **ICT of the Future**

Proposal category: Industrial Research, Experimental Development

Call(s): 4th (2015), 5th (2016), 6th (2017)

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)

TEACHING ACTIVITY

BOARDS

External Scientific Board Member

Information and Communication Technologies PhD Program

University of Calabria (Italy), Department of Computer Science, Modeling, Electronics and Systems Engineering (DIMES)

Academic years: '19-'20 (XXXV cycle), '18-'19 (XXXIV cycle), '17-'18 (XXXIII cycle)

COURSES TAUGHT
AS AN INSTRUCTOR

Advanced Learning and Mining Problems in Big Graph Data (6 hours/year)

University of Calabria (Italy)

Department of Computer Science, Modeling, Electronics and Systems Engineering (DIMES)

PhD Program in Information and Communication Technologies

Academic year: '17-'18

Computer Science (55–84 hours/year)
 University of Calabria (Italy)
 Faculties of Engineering, Pharmacy, and Political Sciences, B.Sc. 1st year (remedial class project)
 Academic years: '09-'10, '08-'09

Computer Science (16–24 hours/year)
 “Magna Græcia” University of Catanzaro, Crotona campus (Italy)
 Faculty of Medicine and Surgery, B.Sc. in Nurse Sciences (3rd year)
 Academic years: '10-'11, '09-'10, '08-'09, '07-'08

Laboratory Activity in Computer Science (16 hours/year)
 “Magna Græcia” University of Catanzaro, Catanzaro campus (Italy)
 Faculty of Medicine and Surgery, B.Sc. in Dental Health (3rd year)
 Academic years: '09-'10

COURSES TAUGHT
 AS A TEACHING
 ASSISTANT

Foundations of Computer Science (18 hours/year)
 University of Calabria (Italy)
 Faculty of Engineering, B.Sc. in Computer Engineering (1st year)
 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

THESIS
 (CO-)ADVISING

Giacomo Lo Cascio
 Sapienza University of Rome, M.Sc. in Data Science (A.Y. '19-'20)
 Thesis: *In-depth Analysis and Representation Learning of Financial Networks*

Alessia Galli
 Sapienza University of Rome, M.Sc. in Data Science (A.Y. '19-'20)
 Thesis: *A Unifying View of Motifs and Discords 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
 (CO-)SUPERVISING

Giulia Gavazzi, Giacomo Legnaro, Hugo Maldini
 M.Sc. students at Sapienza University of Rome
 Research Interns at UniCredit, Winter 2016 (Data Science Industrial Liaison Program)
 Project: *News Collection for Emerging-Event Detection* [56]

Alessio Areni, Cristiano Di Crescenzo, Elena Troccoli
 M.Sc. students at Sapienza University of Rome
 Research Interns at UniCredit, Winter 2016 (Data Science Industrial Liaison Program)
 Project: *Event Detection in Temporal Graphs* [56]

Pranay Anchuri
 Ph.D. student at Rensselaer Polytechnic Institute
 Research Intern at Yahoo Labs Barcelona, Summer 2014
 Project: *Mining Interesting Patterns from Uncertain Graphs*

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* [22]

Edoardo Galimberti
 M.Sc. student at Politecnico of Milan
 Research Intern at Yahoo Labs Barcelona, Summer 2014
 Project: *Efficient and Effective Community Search* [10]

Natali Ruchansky
 Ph.D. student at Boston University
 Research Intern at Yahoo Labs Barcelona, Summer 2014
 Project: *The Minimum Wiener Connector Problem* [30]

Davide Mottin
 Ph.D. student at University of Trento
 Research Intern at Yahoo Labs Barcelona, Summer 2013
 Project: *Query Reformulation in Graph Databases* [29]

Panos Parchas
 Ph.D. student at Hong Kong University of Science and Technology
 Research Intern at Yahoo Labs Barcelona, Summer 2013
 Project: *Extracting Representative Instances of Uncertain Graphs* [34, 9]

Lucrezia Macchia
 Ph.D. student at University of Bari
 Research Intern at Yahoo Labs Barcelona, Winter 2013
 Project: *Mining Summaries of Propagations* [37]

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* [36]

Charalampos Tsourakakis
 Ph.D. student at Carnegie Mellon University
 Research Intern at Yahoo Labs Barcelona, Summer 2012
 Project: *Extracting optimal quasi-cliques with quality guarantees* [38]

HONORS AND AWARDS

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 7th, 2018 to Aug 7th, 2024

Sector 09/H1 - Information processing systems (*Sistemi di elaborazione delle informazioni*), valid from Jul 27th, 2018 to Jul 27th, 2024

ABI IT Innovation Award (“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 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*” [58]

“**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*” [16]

SIGMOD 2011 Repeatability and Workability Evaluation (RWE) label

Paper “*Advancing Data Clustering via Projective Clustering Ensembles*” [41]

SDM 2009 Student Travel Award

Paper “*Diversity-based Weighting Schemes for Clustering Ensembles*” [45]

Italian Innovation Award (*Premio Nazionale per l’Innovazione*) - I edition (2009)

Fraud Detection via Load Profiles project within *EUREKA! An idea for energy* initiative

Habilitation to practice as an Engineer

Section A, Field of Information Technology

Qualifying exam passed on October 2006 (1st 2006 session), no expiration date

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

RDBMSs: MySQL, PostgreSQL

NON-RELATIONAL DBs: Redis, Couchbase Server

VIRTUALIZATION TECHNOLOGY: Docker

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

STREAM-PROCESSING TECHNOLOGY: Apache Kafka

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, pandas, NetworkX, OpenCV (Python)

OPERATING SYSTEMS: macOS, Windows, Unix/Linux

JOURNALS

- [1] I. Bordino, A. Ferretti, F. Gullo, and S. Pascolutti. GarNLP: A Natural Language Processing Pipeline for Garnishment Documents. *Information Systems Frontiers (ISFI)*, 2020, ONLINE (DOI: <http://dx.doi.org/10.1007/s10796-020-09997-0>).
- [2] E. Galimberti, F. Bonchi, F. Gullo, and T. Lanciano. Core Decomposition in Multilayer Networks: Theory, Algorithms, and Applications. *Transactions on Knowledge Discovery from Data (TKDD)*, 14(1):11:1-11:40, 2020 (DOI: <http://dx.doi.org/10.1145/3369872>).
- [3] F. Bonchi, I. Bordino, F. Gullo, and 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>).
- [4] A. Khan, F. Bonchi, F. Gullo, and 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>).
- [5] A. Tagarelli, A. Amelio, and 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>).
- [6] 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>).
- [7] 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>).
- [8] I. Assent, C. Domeniconi, F. Gullo, A. Tagarelli, A. Zimek. MultiClust 2013: Multiple Clusterings, Multi-view Data, and Multi-source Knowledge-driven Clustering. *SIGKDD Explorations*, 18(1):35-38, 2016 (DOI: <http://dx.doi.org/10.1145/2980765.2980769>).
- [9] 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>).
- [10] N. Barbieri, F. Bonchi, E. Galimberti, and 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>).
- [11] F. Bonchi, A. Gionis, F. Gullo, C. Tsourakakis, and A. Ukkonen. Chromatic Correlation Clustering. *Transactions on Knowledge Discovery from Data (TKDD)*, 9(4):34:1-34:24, 2015 (DOI: <http://dx.doi.org/10.1145/2728170>).
- [12] 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>).
- [13] F. Gullo, C. Domeniconi, and 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>).
- [14] F. Gullo, G. Ponti, and 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>).

- [15] F. Gullo, C. Domeniconi, and 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>).
- [16] F. Gullo, and A. Tagarelli. Uncertain Centroid based Partitional Clustering of Uncertain Data. *Proceedings of the VLDB Endowment (PVLDB)*, 5(7):610-621, 2012.
- [17] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, and 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>).
- [18] S. Greco, F. Gullo, G. Ponti, and 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>).
- [19] F. Gullo, G. Ponti, A. Tagarelli, and 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>).
- [20] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, and 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

- [21] E. Galimberti, A. Barrat, F. Bonchi, C. Cattuto, F. Gullo. Mining (maximal) Span-cores 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.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] 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.
- [26] 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.
- [27] 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.

- [28] 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.
- [29] D. Mottin, F. Bonchi, and 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.
- [30] N. Ruchansky, F. Bonchi, D. Garcia Soriano, F. Gullo, and 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.
- [31] O. Balalau, F. Bonchi, T-H. Hubert Chan, F. Gullo, and 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.
- [32] F. Gullo, G. Ponti, and 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.
- [33] F. Bonchi, F. Gullo, A. Kaltenbrunner, and 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.
- [34] P. Parchas, F. Gullo, D. Papadias, and 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.
- [35] F. Bonchi, A. Gionis, F. Gullo, and 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.
- [36] A. Khan, F. Bonchi, A. Gionis, and 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.
- [37] L. Macchia, F. Bonchi, F. Gullo, and 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.
- [38] C. E. Tsourakakis, F. Bonchi, A. Gionis, F. Gullo, and M. A. Tsiarli. Denser than the densest subgraph: extracting optimal quasi-cliques with quality guarantees. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '13)*, pp. 104-112. Chicago, Illinois (USA), August 11-14, 2013.
- [39] F. Bonchi, A. Gionis, F. Gullo, and 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.

- [40] F. Gullo, A. K. A. Talukder, S. Luke, C. Domeniconi, and 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.
- [41] 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.
- [42] F. Gullo, G. Ponti, and 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.
- [43] F. Gullo, C. Domeniconi, and 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.
- [44] F. Gullo, C. Domeniconi, and 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.
- [45] F. Gullo, A. Tagarelli, and 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.
- [46] 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.

OTHER PUBLICATIONS

EDITED VOLUMES

- [47] 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.
- [48] 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.
- [49] 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.

- [50] 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.
 - [51] 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.
 - [52] I. Assent, C. Domeniconi, F. Gullo, A. Tagarelli, and A. Zimek. *MultiClust '13: Proceedings of the 4th MultiClust Workshop on Multiple Clusterings, Multi-view Data, and 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.
 - [53] 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, and W. C. Onn. *Emerging Trends in Knowledge Discovery and Data Mining - PAKDD 2012 International Workshops: DMHM, GeoDoc, 3Clust, and 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.
- BOOK CHAPTERS
(REFEREED)
- [54] 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).
 - [55] F. Gullo, G. Ponti, and S. Greco. Organizing XML Documents on a Peer-to-Peer Network by Collaborative Clustering. In *XML Data Mining: Models, Methods, and Applications*, IGI Global, 2012, pp. 449-466 (DOI: <http://dx.doi.org/10.4018/978-1-61350-356-0.ch018>).
- OTHER
INTERNATIONAL
CONFERENCES
- [56] C. Di Crescenzo, G. Gavazzi, G. Legnaro, E. Troccoli, I. Bordino, and 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.
 - [57] F. Bonchi, I. Bordino, F. Gullo, and 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.
 - [58] 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.
 - [59] 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, and Big Data (MOD '16)*, pp. 402-406. Volterra, Italy, August 26-29, 2016.
 - [60] S. Romeo, A. Tagarelli, F. Gullo, and 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.

- [61] A. Tagarelli, and 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.
- [62] F. Gullo, G. Ponti, A. Tagarelli, S. Iiritano, M. Ruffolo, and 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.
- [63] 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.
- [64] F. Gullo, G. Ponti, and 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.
- [65] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, and 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.
- [66] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, and P. Veltri. A Time Series Based Approach for Classifying Mass Spectrometry Data. In *Proceedings of the 20th IEEE International Symposium on Computer-Based Medical Systems (CBMS '07)*, pp. 412-417. Maribor, Slovenia, June 20-23, 2007.
- [67] C. Domeniconi, F. Gullo, and 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.
- [68] A. Freire, M. Manca, D. Saez-Trumper, D. Laniado, I. Bordino, F. Gullo, and 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.
- [69] S. Greco, F. Gullo, G. Ponti, A. Tagarelli, and 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.
- [70] S. Greco, F. Gullo, G. Ponti, and 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.
- [71] F. Gullo, G. Ponti, A. Tagarelli, and 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.
- [72] F. Gullo, and 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.

- [73] F. Gullo, G. Ponti, A. Tagarelli, and 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.