
ELEONORA CICERI

CURRICULUM VITAE

PERSONAL INFORMATION

Name Eleonora Ciceri
Date of birth October 3, 1987
Place of birth Erba (Italy)
E-mail Italian
Work address eleonora.ciceri@polimi.it
Web page <http://www.eleonoraciceri.it/>
Languages Italian (mother tongue), English (fluent), French (basic)

EMPLOYMENT

ACADEMIA

2/2016 - present Fondazione Centro San Raffaele (Italy). Research collaborator
1/2015 - 1/2016 Politecnico di Milano (Italy). Research collaborator
11/2011 - 12/2011 Politecnico di Milano (Italy). Research collaborator

EDUCATION

2012 - 2015 Politecnico di Milano (Italy), Dept. of Electronic Engineering, Information Sciences and Bioengineering.
Ph.D. in Computer science and Engineering
Thesis title: Humans in the loop: Optimization of active and passive crowdsourcing
Advisor: Prof. Piero Fraternali
Co-Advisors: Prof. Davide Martinenghi, Prof. Marco Tagliasacchi

2009 - 2011 Politecnico di Milano (Italy), Dept. of Electronic Engineering, Information Sciences and Bioengineering.
M. Sc. *cum Laude* as Computer Engineer (“voto 110/110 e lode”)
Thesis title: Diversificazione dei risultati di ricerca in presenza di limiti di accesso
Advisor: Prof. Piero Fraternali
Co-Advisor: Prof. Marco Tagliasacchi

2006 - 2009 Politecnico di Milano (Italy), Dept. of Electronic Engineering, Information Sciences and Bioengineering.
B. Sc. *cum Laude* as Computer Engineer (“voto 110/110 e lode”)
Title: Una nuova interfaccia grafica per la carrozzina autonoma LURCH
Advisor: Matteo Matteucci

2001 - 2006 Istituto Tecnico Industriale Statale “Magistri Cumacini”, Como, Italy
Perito Industriale Capotecnico (specializzazione “Informatica”)

PARTICIPATION TO LOCAL AND EUROPEAN RESEARCH PROJECTS

<i>Acronym</i>	<i>Call</i>	<i>Title</i>	<i>Period</i>	<i>Covered topics</i>	<i>Role</i>
CUBRIK	FP7-ICT-2011-7	Human-enhanced time-aware multimedia search	2011-2014	Crowdsourcing, multimedia	Scientific investigator
SmartH2O	FP7-ICT-2013-11	An ICT platform to leverage on Social Computing for the efficient management of Water Consumption	2015	Social media analysis, influence, trust, multimedia, crowdsourcing	Research investigator
Proactive	Regional project	Protezione del territorio con infrastrutture ICT avanzate, cittadinanza attiva e reti sociali	2015	Social media analysis, influence, trust, multimedia, crowdsourcing	Research investigator
WITDOM	H2020-ICT-2014-1	Empowering privacy and security in non-trusted environments	2016-2017	Privacy and security of sensitive data, genomic analysis and bioinformatics	Research investigator, software engineer, team management
PRISMACLOUD	H2020-ICT-2014-1	Privacy and security maintaining services in the cloud	2016-2018	Privacy and security of health data, big data anonymization	Software engineer
OPERANDO	H2020-DS-2014-1	Online privacy enforcement, rights assurance and optimization	2017	Privacy enforcement in the health field, nutrition	Software engineer

RESEARCH INTERESTS

Top-K queries, incomplete data, crowdsourcing, multimedia, machine learning, big data

CROWDSOURCING

Our aim is to improve the performance of state-of-the-art (multimedia and textual) content analysis algorithms by merging automatic computation and human computation. On the one hand, by applying active crowdsourcing techniques (i.e., outsourcing of tasks to a large group of people called workers), we aim at reducing the uncertainty in ranked result sets. On the other hand, by applying passive crowdsourcing techniques (i.e., leveraging of online activity of users for task resolution), we aim at identifying the most relevant content and influencers for a specific topic of interest, discarding the huge amount of non-relevant, offensive and non-authoritative content generated by users.

Related publications: (Bozzon, Catallo, Ciceri, Fraternali, Martinenghi, & Tagliasacchi, 2012), (Ciceri, Fraternali, & al, The CUBRIK project: Human-enhanced time-aware multimedia search, 2012)

TOP-K QUERIES

When submitting a query, a user is often interested in identifying the most relevant results matching her user needs. Usually, relevance is expressed as a scoring function combining attribute values and attribute relevance to the query. In our research work, we studied: i) how to retrieve relevant and diversified results; ii) how to deal with uncertainty in the scoring function.

Related publications: (Catallo, Ciceri, Fraternali, Martinenghi, & Tagliasacchi, 2013) (Ciceri, Fraternali, Martinenghi, & Tagliasacchi, 2015)

MULTIMEDIA PHYLOGENY

On a daily basis, new image and video content is published on the web, even though the diffusion of this information is not always authorized. Multimedia files are indeed copied, resized, re-encoded and finally reposted. We studied techniques for the identification of original content, which reconstruct the process that brought to the creation of such copies, and identify users that maliciously replicated content.

BIOINFORMATICS

Gene annotation databanks are used to store biological knowledge about the functionalities of each gene. A gene annotation associates each gene with feature terms denoting its characteristics.

In a first research work (in cooperation with Politecnico di Milano), we created algorithms for the automatic integration of available data sources providing such annotations, so as to identify potential annotation errors and infer new annotations based on the similarity between feature terms.

In a second research work (in cooperation with San Raffaele Scientific Institute), we developed a software that helps operators to deal with time-evolving variant annotations. The software answers to the need of genomic laboratories to keep up with their constantly changing knowledge about variants and annotations, so as to reclassify periodically patients' variants and provide them with up-to-date reports.

Related publications: (Chicco, Ciceri, & Masseroli, 2014) (Catallo, et al., 2017)

SOCIAL MEDIA

In the last years, social media have attracted millions of users and have been integrated in people's daily practices. They enable users to create and share content or to participate in social networking. User-generated content, i.e., the various forms of media assets publicly available and created by end-users, is published every day on the Web and mostly in social media at a massive scale, either in the form of textual documents or in the form of multimedia items. The analysis of such content allows us to discover interesting insights, e.g., user interests, trends and behavior, ongoing events.

Related publications: (Ciceri, Identifying topical content and experts in Twitter using text and visual content, 2015) (Ciceri, Catallo, Martinenghi, & Fraternali, 2015) (Ciceri, Fedorov, Umuhoza, Brambilla, & Fraternali, 2015)

COOPERATION WITH OTHER RESEARCH GROUPS

May 2015 Guest at TU Wien, Cartography group
Preparation of "Emotional Mapping Party" in Como during FOSS4G Conference

ACTIVE PARTICIPATIONS TO CONFERENCES

2015 FOSS4G Conference in Como
"Emotional Mapping Party" Committee

TEACHING ACTIVITY

AS TEACHING ASSISTANT

2011-2014	Digital image processing (prof. Marco Tagliasacchi), Politecnico di Milano, Polo di Como.
2012-2015	Web Technologies (prof. Piero Fraternali), Politecnico di Milano, Polo di Como. In Italian.
2013-2015	Fondamenti di Informatica ("Computer science principles", prof. Piero Fraternali), Politecnico di Milano, Polo di Como. In Italian.
2013-2015	Mobile Applications (prof. Davide Martinenghi), Politecnico di Milano, Polo di Como.

AS TUTOR

2012-2013	Cross Boundary Processes (prof. Michela Arnaboldi and prof. Fabio Salice), Politecnico di Milano, Polo di Como.
2013-2014	"CorDonare" project, Politecnico di Milano, Polo di Como.

TECHNICAL SKILLS

TOPICS

- **Machine learning:** SVM classification, both on textual data and multimedia data
- **Social media analysis:** user-generated content classification, user influence analysis, scalability on real-time data feed analysis
- **Crowdsourcing:** optimization of active crowdsourcing processes to maximize information retrieval under budget constraints
- **Web development**
- **Information retrieval**
- **Mobile application development:** Android (proficient) and iOS (basic)
- **Bioinformatics:** basic knowledge of NGS technology and related genomic analysis pipeline (sequence alignment, variant calling, variant annotation & reclassification)

PROGRAMMING LANGUAGES

<i>Java</i>	Fluent
<i>Node.js</i>	Fluent
<i>HTML</i>	Fluent
<i>MATLAB</i>	Fluent (used for algorithm prototyping)
<i>JavaServlet & JSP</i>	Fluent
<i>C++</i>	Working knowledge
<i>JavaScript</i>	Working knowledge
<i>CSS</i>	Working knowledge
<i>C#</i>	Basic knowledge
<i>PHP</i>	Basic knowledge
<i>Spring</i>	Prior experience

DATABASE EXPERIENCE

<i>SQL</i>	Fluent
<i>No-SQL</i>	MongoDB: fluent

ARTICLES FROM JOURNALS AND CONFERENCES

Bozzon, A., Catallo, I., Ciceri, E., Fraternali, P., Martinenghi, D., & Tagliasacchi, M. (2012). A framework for crowdsourced multimedia processing and querying. Lyon, France: CrowdSearch workshop, WWW.

Catallo, I., Ciceri, E., Fraternali, P., Martinenghi, D., & Tagliasacchi, M. (2013). Top-K diversity queries over bounded regions. *Transactions On Database Systems (TODS)* .

Catallo, I., Ciceri, E., Stenirri, S., Merella, S., Sanna, A., Ferrari, M., et al. (2017). An open-source tool for managing time-evolving variant annotation. *Proceedings of Computational Intelligence methods for Bioinformatics and Biostatistics (CIBB)*. Cagliari.

Chicco, D., Ciceri, E., & Masseroli, M. (2014). Correlation of gene function annotation lists through enhanced Spearman and Kendall measures. *Proceedings of Computational Intelligence methods for Bioinformatics and Biostatistics (CIBB)*. Cambridge.

Ciceri, E. (2015). Identifying topical content and experts in Twitter using text and visual content. *HCOMP*.

Ciceri, E., Catallo, I., Martinenghi, D., & Fraternali, P. (2015). When food matters: identifying food-related events on Twitter. *KDWEB workshop*.

Ciceri, E., Fedorov, R., Umuhoza, E., Brambilla, M., & Fraternali, P. (2015). Assessing online media content trustworthiness, relevance and influence: an introductory survey. *KDWEB workshop*.

Ciceri, E., Fraternali, P., & al. (2012). The CUbRIK project: Human-enhanced time-aware multimedia search. Lyon, France: World Wide Web conference (WWW).

Ciceri, E., Fraternali, P., Martinenghi, D., & Tagliasacchi, M. (2015). Crowdsourcing for top-K query processing over uncertain data. *TKDE*.