

Sara Geremia

Curriculum Vitae

[Applied Data Science & Artificial Intelligence](#)

Department of Mathematics and Geosciences

University of Trieste

Via Alfonso Valerio 13, Trieste - Italy



Education

2022/12/31 - current: **Ph.D.** in Applied Data Science & Artificial Intelligence, University of Trieste - IT; Curriculum in Economy and Society; Topic: Network Analysis Methods to evaluate industrial performance, supervisor: Prof. D. De Stefano.

2021/10/29: **M.Sc.** (*Magna cum Laude*) in Statistical Sciences - Curriculum Data Science (LM-82), Alma Mater Studiorum University of Bologna; title of the thesis: Automatic assessment of English proficiency levels via language complexity metrics, advisor: Prof. L. Anderlucci.

2021/03/01 - 2021/09/10: **Dissertation Intership Abroad**, National University of Ireland, Galway - School of Mathematics, Statistics and Applied Mathematics; topic: A Supervised Learning Model for the Automatic Assessment of Language Levels in learner English, supervisor: Dr. A. Simpkin;

2019/07/11: **B.Sc.** (*Magna cum Laude*) in Statistical Sciences - Curriculum Biodemography (L-41), Alma Mater Studiorum University of Bologna; title of the thesis: Uno studio case-crossover sull'associazione tra inquinamento atmosferico da particolato e visite al pronto soccorso pediatrico di Padova per bronchiolite (*A case-crossover study on the association between particulate air pollution and visits to the Pediatric Emergency Department of Padua for bronchiolitis*), advisor: Prof. R. Miglio.

Working Activity

2023/10/15 - 2023/12/31: **Tutoring university students**, Department of Political and Social Sciences - University of Trieste;

2022/02/23 - 2022/12/31: **Data cleaning, statistical analyses and reports production**, Centro collaboratore dell'OMS dell'IRCCS materno infantile "Burlo Garofolo": *Carrying out the project: 5 per mille, anno 2014 "Lo stato di salute delle donne e dei bambini in Friuli Venezia Giulia e la valutazione dei suoi determinanti"*;

2021/03/01 - 2021/08/31: **Statistical analyses and static and dynamic reports production**, National University of Ireland, Galway - School of Mathematics, Statistics and Applied Mathematics: *Taking part as a trainee in a research project aimed at the automatic assessment of language levels in learner English*;

2019/01/14 - 2019/02/15: **Data cleaning and statistical analyses**, University of Padova - Unit of Biostatistic, Epidemiology and Public Health: *Taking part as a trainee in a research project aimed at assessing the effect of particulate pollution on the occurrence of respiratory diseases in children*.

Conferences

2023/12/16: **CFE 2023**, Collaborator to the abstract: *Investigating the impact of content similarity on density-based clustering of social networks*;

2023/09/12: **CLADAG 2023 Scientific Meeting**, Contributed Talk: *Visualization of proximity and role-based embeddings in a regional labour flow network*;

2019/08/21: **ISES ISIAQ 2019 Joint Annual Meeting**, Collaborator to the abstract: *Air pollution particulate matter is associated with Pediatric Emergency Room admissions for bronchiolitis*;

2019/10/19: **YMS 2019 Young Statisticians' Meeting**, Contributed Talk: *Increasing Risk Of Emergency Room Admission For Bronchiolitis In Infants Exposed To Air Pollution*.

Publications

S. Geremia, T. Gaillat, N. Ballier, A. Simpkin. **Automatic assessment of English proficiency levels via language complexity metrics**. Submitted to *Journal of Quantitative Linguistics* (September 2023);

S. Geremia, E.P. Valente, I. Mariani, P. Dalena, M. Lazzerini. **Women's suggestions on how to improve the quality of maternal and newborn care during the COVID-19 pandemic in Italy: A co-occurrence network analysis**. *J Glob Health* (May 2023);

R. Costa et al. **Quality of maternal and newborn care around the time of childbirth for migrant versus nonmigrant women during the COVID-19 pandemic: Results of the IMaGiNE EURO study in 11 countries of the WHO European region**. *International Journal of Gynecology & Obstetrics* (December 2022);

E. Gallo, S. Bressan, S. Baraldo, D. Bottigliengo, S. Geremia, A.S. Acar, L. Zagolin, G. Marson, L. Da Dalt, D. Gregori. **Increased risk of Emergency Department presentations for bronchiolitis in infants exposed to air pollution.** *Risk Analysis* (August 2022).

Editorials and Reviews

S. Geremia, F. Morea, D. De Stefano. **Visualization of proximity and role-based embeddings in a regional labour flow network.** *CLADAG 2023 Book of Short Papers* (September 2023);

S. Geremia, E.P. Valente, I. Mariani, M. Lazzerini. **Prevailing emotions in women's suggestions on how to improve the quality of care around the time of childbirth during the COVID-19 pandemic.** Accepted by *Nursing Open* (August 2023).

Language skills

Italian (*Mother tongue*); **English** (*Listening: C1, Reading: C1, Spoken interaction: C1, Spoken production: C1, Writing levels: C1*); **Portuguese** (*Beginner*).

Digital skills

General skills

Information processing (*Proficient: import, tidying, analyses of health data*);
Communication (*Independent: production of static, dynamical and interactive R Markdown documents, Shiny App development*);

Program languages

R (*Proficient: Statistical and Machine Learning, Data mining*);
Python (*Independent*);
Stata (*Independent*).

Operating System and software

Tidyverse R packages suite (*Independent*); **Shiny** (*Competent*); **RStudio IDE** (*open Proficient*); **PyCharm IDE** (*open Independent*); **GitLab/GitHub** (*Independent*); **MS Windows XP** (*Independent*); **Suite MS Office** (*Proficient*); **Linux** (*Beginner*)

Certificated courses

2021/03: K. Medeiros (Medidata Solutions): Building Web Applications with Shiny in R - DataCamp, 100%.

2019/02: R. Scavetta (Science Craft): Introduction to Data Visualization with ggplot2 - DataCamp, 100%.