








John Bianchi

Professional CV

Last update
14/06/2024

 Male
 08/02/1995
 55100 Lucca (LU) - Italy

john.lenci.bianchi@gmail.com 
<https://bianchi-john.github.io/personalWebsite/> 
(+39) 3318732944 
ORCID: 0009-0006-2582-1480 

Education

- **National Ph.D. program in Cybersecurity (Ph.D.-CySec)** 01 Dec. 2023 – Ongoing
IMT School for Advanced Studies Lucca *Lucca, Italy*
 - Tools and Techniques for Scalable and Usable Cyber Ranges
 - **Contact person:** Gabriele Costa
 - **Curriculum:** Software, System and Infrastructure Security
- **MSc Digital humanities - *Language technology* — 101/110** 26 Sept. 2019 – 03 Feb. 2022
University of Pisa *Pisa, Italy*
 - LM-43 - 2nd level degree in Computer Science for humanities - Department of Philology, Literature and Linguistics. Training course in *Language technology*
 - **Dissertation/thesis title:** Automatic recognition of scientific names of species in texts for FAO document indexing. **Thesis supervisor:** Dr. Gianpaolo Coro
- **BSc Digital humanities — 98/110** 04 Sept. 2014 – 09 May 2019
University of Pisa *Pisa, Italy*
 - L-10 - 1st level degree in Humanities - Department of Philology, Literature and Linguistics
 - **Dissertation/thesis title:** Design and development of a low-cost platform for the acquisition and analysis of tourism related data. **Thesis supervisor:** Dr. Andrea Marchetti
- **Erasmus+ Programme** 24 Sept. 2020 – 11 Jan. 2021
King's College London *London, United Kingdom*
- **Italian High School Diploma — 67/100** Sep. 2009 – July 2014
Liceo Artistico "A. Passaglia" *Lucca, Italy*

Work Experience

- **Research Collaborator at** 01 Jan. 2023 - 01 Dec. 2023
IMT School for Advanced Studies Lucca *Lucca*
 - Member of the *TOFFEE* project (*TOols for Fighting FakeEs*): my research activities focus on detecting the trustworthy level of a news article, by exploiting deep learning techniques and explainable AI. In particular, I am adopting

- * data scraping software such as *Selenium* with *Chromedriver* to extract textual content from more than 40 news websites using a *pipenv*-managed development environment.
- * Google *Colab* and *Jupyter* Notebooks for data mining, using tools such as *Scikit-Learn*, *Pandas*, *Numpy*, *Matplotlib*, *Seaborn*, *PyTorch* and *TensorFlow*.
- * Bert-derived transformers for text classification and study of these techniques through explainability techniques (*LIME*, *EXBERT*, *SHAP*.)

- **Junior staff in consulting** 21 March. 2022 - 21 Dec 2022
Deloitte & Touche SpA *Milano, Italy*
 - Back-end e-commerce website development using *Salesforce commerce cloud*.
 - Development of an e-commerce platform for *Unicoop Firenze*'s online grocery shopping.
- **Trainee student** 01 Mar. 2021 - 01 Jun. 2021
Institute of Information Science and Technologies - ISTI CNR *Pisa, Italy*
 - NLP software development for the Named Entity Recognition task required by the *ASFA* division of *Food and Agriculture Organization of the United Nations* (FAO).
- **Web Developer** 01 Jun. 2020 - 01 Jan. 2021
Wally Communication *Lucca, Italy*
 - Website development using *Wordpress* with *Elementor* page builder.
- **ICT help desk agent** 01 Jun. 2020 - 01 Oct. 2020
Accademia di Belle Arti di Carrara *Massa-Carrara, Italy*
 - Technical support for examination and graduation session.
- **Digital audio transcriber** 15 Feb. 2019 - 15 Jul. 2019
Sigma Technologies S.L.U. *Madrid, Spain - Remote work*
 - Transcription of Google voice action digital audio.
- **Trainee student** 01 Feb. 2019 - 01 Mar. 2019
Institute of Informatics and Telematics - IIT CNR *Pisa, Italy*
 - Development of Web scraping software for the automated acquisition of data regarding tourism flow in four different areas of Tuscany.
- **Bartender** 15 Jun. 2016 - 30 Dec. 2019
Bar Ponte Del Giglio *Lucca, Italy*
 - - Dealing with customers, handling phone calls and payments..

Projects and publications

- **Evaluating Trustworthiness of Online News Publishers via Article Classification** April 8-12, 2024
39th ACM/SIGAPP Symposium on Applied Computing, Avila, Spain

- Authors: John Bianchi, Manuel Pratelli, Marinella Petrocchi, Fabio Pinelli
- Edited by: Jiman Hong, Juw Won Park
- Published by: ACM
- Pages: 671–678
- DOI: 10.1145/3605098.3636044
- BibTeX entry

• **Reviewer**

01 Jan. 2023 - Ongoing

IMT School for Advanced Studies Lucca

Lucca

- SIGIR'23 (Special Interest Group on Information Retrieval), 2023
- ROMCIR 2024 (Reducing online misinformation through credible information retrieval)
- ITASEC 2023 (Italian Conference on Cybersecurity)
- ESORICS 2023 (European Symposium on Research in Computer Security)
- EAI Intetain 2023 - 14th EAI International Conference on Intelligent Technologies for Interactive Entertainment
- ECIR 2024 (46th European Conference on Information Retrieval)
- The Web Conference 2024 - Web4Good

• **Website Development for *Lucca National Museums***

01 Jan. 2023 - Ongoing

IMT School for Advanced Studies Lucca

- Coordinators: Prof. Gabriele Costa Prof. Emanuele Pellegrini
- Adoption of web development technologies, including HTML, CSS, and JavaScript, to create an interactive and visually appealing website.
- Implementation of dedicated APIs using Node.js MySQL and Docker.

• **Automatic recognition of scientific names of species in texts for FAO document indexing**

Coordinator: Prof. Gianpaolo Coro

- Project carried out as dissertation for my master's degree.
- Timeline for software development: 01 Mar. 2021 - 01 Jun. 2021.
- Timeline for thesis text realization: 01 Aug. 2021 - 01 Dec. 2021
- Named entity recognition for recognizing scientific names of species in texts in a computationally efficient manner and integrated as a web service to ASFA. for automatic document indexing.
- Resources:
 - * Web service page - nlp.d4science.org/asfa/
 - * GitHub repository - github.com/cybprojects65/SpeciesNamesNER
 - * Documentation - johnbianchi.it/documents/Taxa_project_Documentation.pdf

• **NLP of Tweets about disasters**

Coordinator: Dr. Andrea Esuli

- This project has been conducted for the 'Text Analytics' course of my university studies.
- Development of two different models to classify the tweets: the first one, based on a traditional approach, manually implements feature engineering, the second one relies on the Bert language model, through Simple Transformers library.
- The project was developed for a competition on Kaggle.
- Timeline: Jun. 2021 - Jul. 2021

- Resources:
 - * GitHub repository - github.com/bianchi-john/Natural-Language-Processing-with-Disaster-Tweets/blob/main/Disaster_tweets.ipynb
 - * Kaggle competition - kaggle.com/c/nlp-getting-started

- **Data Mining Project, Data analysis on Carvana data**

Coordinator: Prof. Anna Monreale, Prof. Dino Pedreschi

- This project has been conducted for the ‘*Data mining: fundamentals*’ course of my university studies.
- Application of statistics data mining to analyze data concerning the transportation sector.
- Timeline: 15 Sept. 2019 - 01 Dec. 2019
- Resources:
 - * Project documentation - johnbianchi.it/documents/Data_Mining-I.pdf

- **Project for the course *Computational Linguistics II***

Coordinator: Prof. Simonetta Montemagni, Dr. Giulia Venturi

- Developer activity for the ‘*Computational Linguistics II*’ course of my university studies.
- Collection of a test corpus to evaluate the accuracy of a chain of automatic linguistic annotators.
- Timeline: 01 Feb. 2020 - 01 Jun. 2020
- Resources:
 - * Project documentation - johnbianchi.it/documents/Report_Linguistica_Computazionale-II.pdf

- **Project for the course ‘*Artificial Intelligence And Society*’**

Coordinator: Dr. Ashwin Mathew

- Final essay for the course “*Artificial Intelligence And Society*” at King’s College London.
- Timeline: 01 Sept. 2020 - 01 Jan. 2021
- Resources:
 - * Essay - johnbianchi.it/documents/The_price_of_automatic_communication.pdf

- **Design and development of a low-cost platform for the acquisition and analysis of tourism data**

Coordinator: Dr. Andrea Marchetti

- This project has been carried out as the dissertation for my bachelor’s degree.
- Timeline for software development: Sept. 2018 - Feb. 2019.
- Timeline for thesis text realization: 01 Feb. 2019 - 01 Apr. 2019
- Web scraping of TripAdvisor using a Raspberry Pi 3 Model B+ and development of a dedicated website for the visualization of the extracted data.
- Resources:
 - * Documentation - johnbianchi.it/documents/Documentation.pdf

”

- **Progetto Pro3. Borghi, paesi, aree interne: infrastrutture, sostenibilità e qualità della vita**

Coordinator: Prof. Fabio Pinelli

- Timeline for software development: Sept. 2023 - Dec. 2023.
- Project to evaluate the feasibility and accessibility of public services in the Middle Valley region
- Utilizing GTFS data, OpenTripPlanner, and Tableau software for data visualization

Foreign language skills

- **English IELTS Academic**
British Council.
 - Europass level B2.

01 Dec. 2018

I authorize the processing of personal data pursuant to Legislative Decree 196/2003.

John Bianchi