

# Andrea Giorgio Paglia

## Data Scientist

Milan, Italy • +39-339-5434969 • andrea6692@gmail.com • <https://www.linkedin.com/in/andrea-paglia>

### WORK EXPERIENCE

---

#### Data Scientist

March 2024 – Present

Whatwapp | Industry: Mobile Gaming | Milan, Italy

- Managed the entire data stack (ETL, warehousing, and reporting), standardizing and maintaining all principal business monitoring dashboards. Reduced dashboard loading time by up to 80% and lowered query engine costs by 30% through the implementation of a new caching layer.
- Partnered with Product Managers to synthesize deep user behavior data, guiding key development decisions and feature prioritization.
- Developed and executed A/B testing strategy in partnership with Product, while optimizing the testing framework through automation for increased throughput (dozens of tests monthly).
- Developed and deployed a scalable anomaly detection solution that automatically tracks over 50 KPIs across multiple user dimensions, enabling proactive monitoring and substantially reducing problem resolution latency.
- Deployed an autonomous AI agent workflow for investigating anomalies and conducting root cause analysis, significantly reducing problem resolution time via automated reports to Slack and Notion.
- Co-developed a specialized data analyst agent to automate internal data querying and initial analysis, reducing the manual effort and time spent on data exploration for new projects by over 50%.

#### Data Scientist

July 2020 – February 2024

Everli | Industry: E-commerce | Milan, Italy

- Defined and analyzed dozens of A/B tests to optimize user experience and drive conversion, within the customer-facing website and mobile apps.
- Created an attribution model for the marketing team, used to assess the effectiveness of the marketing channels.
- Led the implementation of an analytics tracking system employing GTM and Mixpanel, used to gather customer usage data; leveraged data-driven insights to evaluate the performance of new product features.
- After one year, moved to the Supply product team, focusing on logistics and courier services:
  - Built many ETL processes to convert structured and unstructured data into AWS Athena tables, enabling large-scale analysis.
  - Developed an ML model to predict when a courier is likely to revoke an availability, used to reduce by 40% problematic orders.
  - Implemented an A/B testing framework for switchback experiments, which was employed to evaluate dozens of new iterations of predictive models.
  - Developed a model to predict probability of couriers' assignments acceptance. The model was used to improve the assignment of orders and resulted in an 8% improvement of acceptance rate.
  - Created a CLI to easily deploy new models on AWS sagemaker, as endpoints or batch inference jobs. Currently ~20 models have been migrated to AWS using this CLI, with more in progress.
  - Developed a churn prediction model for couriers, improving the f1-score from the previous model by 16%.

#### Business Analyst

May 2017 – June 2020

Everli | Milan, Italy

- Leveraged MySQL database and other data sources to produce in-depth analyses and recurring reports, enabling the business to make informed decisions.
- Built and maintained the main reports for the Marketing, Business Development and Catalog departments. Streamlined reporting processes through automation, reducing report generation time by 50%.

- Produced the main recurrent and ad-hoc reports for brands advertising on the platform.

### **SAP BI Consultant intern**

**December 2016 – April 2017**

IBM | Milan, Italy

- Collaborated with clients to produce key reports, using SAP Business Data Warehouse and BI solutions.

### **Business Analyst intern**

**January 2016 – June 2016**

Zalora (Rocket Internet Group) | Kuala Lumpur, Malaysia

- Built reports and analyses for different departments, extracting and cleaning data from databases.

## **PERSONAL PROJECTS**

---

### **Sense Feed**

<https://www.sensefeed.io/>

- An AI-powered platform that provides subscribers with summaries of content from YouTube, RSS feeds, and Twitter/X to manage information overload. The service uses multimodal AI to transcribe and extract key insights, which are delivered as a formatted digest. Features include content organization into categories, control over notification frequency, a distraction-free reader with highlights, and an interactive chat to query the source material.

## **CERTIFICATIONS**

---

- [Deep Learning Specialization](#) 5 courses, Issued by Coursera
- [Machine learning in Python with scikit-learn](#) Issued by Inria
- [Machine Learning Specialization](#) 4 courses, Issued by Coursera
- [Statistical Inference](#) Issued by Coursera
- [Python Programming](#) Issued by Udemy

## **EDUCATION**

---

### **Copenhagen Business School, Copenhagen, Denmark**

**September 2014 – September 2016**

MSc in Accounting, Strategy and Control

### **Università Commerciale L. Bocconi, Milan, Italy**

**September 2011 – September 2014**

BSc in Economics

## **SKILLS**

---

**Tools:** SQL, Python, AWS, Web Analytics tools (Google Analytics, Google Tag Manager, Mixpanel), Segment Customer Data Platform, Business Intelligence tools (Domo, Quicksight, Google Data Studio, Hex Notebooks), Excel  
**Languages:** Italian (Native), English (Fluent)