



# Xiang Li

 [xiang-li.nl](https://xiang-li.nl) |  [LinkedIn](#)

---

## Education

### University of Amsterdam

Aug 2020 – 2023 (Expected)

MSc in Computational Science (joint-degree UvA/VU)

Amsterdam, the Netherlands

- GPA 7.76 (Up to Dec 20, 2022)

### Sun Yat-sen University

Sep 2014 – Dec 2019

BSc in Ecology / Statistics (Double Major)

Guangzhou, China

- GPA 3.7/5.0, Rank 12/40
- Honors: 3rd Class of SYSU Scholarships 2015&2016&2017 (Top 30% university-wide), Excellent League Member of SYSU 2015&2017, Exchange Study Scholarship from Wens Foodstuff Group Co. Ltd, 2017.

---

## Internship Experience

### Janssen Biologics B.V.

Jan 2022 – July 2022

Data Science Intern, MAM-ES Team

Leiden, the Netherlands

- Designed and implemented a process control pipeline for fault detection and online monitoring using manufacturing data from OSIPI system, covering the entire process from raw data collection to obtaining instructive reports.
- Built MPCA and MPLS model using historical batches, enabling real-time monitoring of in-progress batches. **Spark** code was written from the ground up to automate the computation of control parameters and their visualization.
- Contributed to the transition of the data analytics workflow to **Azure Databricks**. This project has been a groundbreaking exploration of using Databricks as a data analytics hub for Janssen, with more business cases to follow.

### KPMG (China) Advisory

Sep 2019 – Nov 2019

Risk Consulting Intern

Guangzhou, China

- Data collection and statistical analysis, reading corporate annual reports, conducting industry research and writing reports independently.
- Work with consultants to review the internal control risks of client companies to compile internal control manuals and audit reports.

---

## Projects

### Bitcoin Snapshots

Sep 2021 - Nov 2021

Large Scale Data Engineering course project

- Studied the Bitcoin blockchain using the *hadoopcryptoledger* library deployed on **Databricks** and the *Spark SQL* and *GraphX* libraries.
- Constructed temporal graphs of Bitcoin transactions and created their snapshots by time. Analyzed how transactions as well as wallets evolve over time by applying graph algorithms.
- Optimized the amount of computation required in building snapshots and queries by means of data engineering.
- Took on all technical work, responsible for all **Scala** code and report writing in a three-person team.

### Research on Sina Weibo hotspots based on webtext mining techniques

Jan 2019 - May 2019

Graduation thesis SYSU

- Designed and implemented a **web crawler** with **python** for Weibo (Chinese Twitter).
- **Natural language processing** techniques were applied to the text data. Finally the news hotspots over a period of time were obtained using cluster analysis.

---

## Skills

**Programming Languages:** Java, Python, R, SAS, Scala

**Tech Skills:** Data structure and algorithms, Azure Databricks, Spark, Bitcoin Blockchain, Web crawler

**Finance:** Chartered Financial Analyst (CFA) Level I Passed