

# Yandi Li (Boris)

☎ (+86) 18401596880 | ✉ liyandi007@163.com | 🏠 yandili | 📷 Yandi | 🌐 yandi.space  
📍 Beijing, China

## Work Experience

---

**Sina Weibo** 🌐 Beijing, China 11/2014 - 12/2025

◦ **Tech Lead / Senior Algorithm Engineer**

- Transitioned from Junior to Senior, serving as Tech Lead for core business lines since 2019 (later transitioned to a core Technical IC focusing on architecture evolution).
- Architected Weibo's foundational unified user profiling system, delivering a universal feature library powering platform-wide feeds, ads, and communities.
- Owned end-to-end engineering and business metrics for core recommendation scenarios (e.g., end-of-gallery image recommendation).

**Yahoo! Research** Barcelona, Spain 02/2014 - 07/2014

◦ **Research Intern**

- First author of a recommendation system paper published at WSDM 2017 (Top tier conference).

## Work Projects

---

**Unified User Profiling** Senior Algorithm Engineer 05/2022 - 12/2025

- **User Profiling:** Led the construction of Weibo's unified profiling system (demographics, interests, behavior) deployed to an online Feature Store. This foundational infrastructure serves multiple core businesses across the platform.
- **Deep Interest Modeling:** Built a multi-objective prediction system (100K+ vocabulary) mapping user behaviors into a dense vector space. Implemented SENet for dynamic feature weighting, Caser for cross-features, and RQ-VAE to improve generalization on long-tail data. Introduced "Serendipity" as a dynamic objective to penalize repetitive recommendations.
- **LLM-Based Automated Tagging:** Engineered a few-shot classification pipeline using LLMs for creator profiles. Abstracted Prompts into configurations stored in Hive for no-code iteration by operations teams. Implemented automated fallback, API retry tracking, and monitoring scripts to ensure pipeline stability.

**Interest Computation & Ad Targeting** Tech Lead 05/2019 - 05/2022

- **Real-Time Streaming:** Upgraded interest computation from offline Hive to Flink streaming, slashing latency to <10 mins and processing 3x more behavior logs. Optimized time decay strategies, boosting overall feed CTR by 5% and time spent by 3.4% (Granted [Patent CN115827966A](#)).
- **Long-Term Debiasing:** Designed an audience-normalization algorithm to mitigate the Matthew Effect. By calculating the deviation between click intensity and the platform average, reduced the Gini coefficient of head tags from 0.58 to 0.45, lifting single-channel CTR by +6%.
- **Ad Targeting Revamp:** Led the migration of the targeting system from legacy statistical models to LightGBM. Expanded core industry UV coverage by 6x and increased the revenue share of interest-based ads from 8% to 10%.

**End-of-Gallery Image Recommendation** Tech Lead 08/2017 - 04/2019

- **System Integration:** Led the post-view image recommendation system. Collaborated with client teams and guided the algorithm team to deploy an LR/FM ranking model and multi-channel Faiss recall (tags, semantics, popularity) natively within the existing C++ engine.
- **High-Concurrency Engineering:** Independently built a high-performance asynchronous gateway using Sanic and asyncio. Implemented multi-page Redis concurrent caching and failover mechanisms, supporting 5000+ QPS with a stable response time under 100ms.

## Content Understanding

Algorithm Engineer

03/2015 - 10/2017

- **Vision & Risk Control:** Developed CNN and MTCNN models for image aesthetic quality, smart cropping, and sensitivity recognition. Designed a robust **fake celebrity avatar** interception model to combat black-market activities.
- **NLP & Classification:** Built a TextCNN multi-model stacking framework for headline classification based on titles and summaries. Developed and deployed a Bi-LSTM **clickbait detection model**.

## Personal Projects

---

### Cyber Dingding (Multi-Agent Collaboration System)

12/2025 - 02/2026

- Engineered a WebSocket/SQLite-based multi-agent collaboration system supporting human-in-the-loop interventions.
- **Architecture:** Designed a multi-role chatroom where agents share global context via @-mentions.
- **Execution:** Integrated OpenInterpreter for computer use capabilities, enforcing private context isolation to hide execution logs from the planning agent, effectively preventing token pollution.

### KDD Cup (Authorship Disambiguation)

Lyon, France

03/2013 - 06/2013

- Disambiguated author names in a large academic database (19 million authors across 50 million Microsoft Academic publications), [\[Report\]](#).
- **Tech Stack:** R, PostgreSQL, Python, LaTeX, Multilingual text processing, LDA topic extraction.

## Education

---

### UPC & Université Lumière Lyon 2

Barcelona & Lyon

09/2012 - 09/2014

#### ◦ M.S. in Data Mining and Knowledge Management (DMKM)

- Erasmus Mundus Scholarship
- Key courses: Kernel-based Learning, Statistical NLP, Advanced Statistical Modeling

### Zhejiang University

Hangzhou, China

09/2008 - 06/2012

#### ◦ B.S. in Mathematics and Applied Mathematics

- GPA: 3.87/4.0 (Top 5%)
- Honors: Third Prize in ZJU ACM Programming Contest; First Prize in Provincial Higher Math Contest.

### UCLA

Los Angeles, USA

08/2009 - 09/2009

- **Summer Session** (Finance, English Writing)

## Publications

---

- **Apply Space Syntax to Online Mapping Tools**, *WSDM 2017* — [Yandi Li](#), Nicola Barbieri, Daniele Quercia.
  - Proposed a context-aware recommendation system for urban navigation. Pioneered the application of Factorization Machines (FM) and BPR to address high-dimensional sparse feature combinations of time, weather, and spatial syntax.
- **Different Implementations and Comparisons of the Chebyshev-Tao Method** — Undergraduate thesis, [\[Part of the Thesis\]](#), [\[Literature Review\]](#), [\[Defense\]](#).

## Skills

---

- **Languages:** Python, SQL
- **Frameworks:** PyTorch, LLM Deployment (Ollama, Huggingface), Flink, Spark, Hive, Docker
- **Hobbies:** Tennis🎾, Ping Pong🏓, Tai Chi🧘, Skiing🎿, DIY🔧