# Arjang Talattof, Ph.D.

Building robust & scalable data systems. US Citizen and UK ILR.

## TECHNICAL SKILLS

- Python, Rust, C++, SQL, Cloud infras (AWS, GCP), Distributed systems, Graph Databases

## EXPERIENCE

**META** **Bay Area, CA**

*Senior Staff (IC7) and Data Lead, AI Research*
*FAIR*

- Advancing the state-of-the-art in artificial intelligence through open research.
- Specializing in scaling data pipelines and ensuring data safety for LLMs, VLMs, and next-generation reasoning models
- https://ai.meta.com/research/

*RealityLabs Research*

- Developing the next generation of computation in Reality Labs Research, building the digital interface for the physical world via Project Aria (see https://about.meta.com/uk/realitylabs/projectaria/)
- **Semantic Search** – *Python*
  - Enable discoverabilty and reusability of multi-modal (text, audio, video, images) datasets by developing a unified search API and powering semantic queries through the use of large transformer (e.g. CLIP, ImageBind, Whisper) models to extract embeddigns
- **Geospatial Ingestion and Indexing** – *Python, C++*
  - Creator and maintainer of pipeline, enabling geographic search for raw data
- **Distributed Computing** – *Python, C++*
  - Enabled distributed computing to run large-scale optimization libraries on Meta cloud infrastructure

**SCAPE TECHNOLOGIES** **London, UK**

*Machine Learning Engineer* *Acquired by Facebook January 2020*

- Building core data flow and analysis pipeline for location-based recognition, allowing devices to see and remember their surroundings and augment the world around them. Cloud infrastructure allows ordinary mobile devices to enhance the world around them by overlaying digital items onto the physical world, both indoors and outdoors, using machine vision and artificial intelligence.
- **Large-Scale Structure-from-Motion Pipeline** – *Python, C++, PyTorch, AWS, PostgreSQL, Redis*
  - Deploying and optimizing large computer vision end-to-end processing pipeline.
  - Developing and optimizing code base to improve runtime and significantly reduce costs.
  - Deep learning-based global image feature extraction and large-scale image retrieval.
  - Custom cloud-based graph database deployment for geospatial image data used in pipeline to build 3D geometric models.

**DELOITTE CONSULTING LLP** **Roslyn, VA, USA**

*Senior Consultant*

- Mission Analytics in Business Model and Transformation/Strategy and Operations.
- Supporting senior government executives in the development of the organization's strategy and business process
- **Distributed Graph Database Analytics** – *Scala, Java, Kafka, Cassandra/NoSQL*
  - Leveraging Cassandra and Spark for large-scale graph networks and analyses including:
    * Migration of client data from on-prem to cloud (AWS)
    * Building tools to explore and analyze graph data in a distributed system
    * Developing machine learning algorithms and automation of real-time entity resolution (data disambiguation) at scale.
    * Revenue increase from $1.8M to $6.0M; internal investment by firm ($0.5M) to generalize new capability based on client deliverable (see https://www2.deloitte.com/content/dam/Deloitte/de/Documents/operations/knowledge-graphs-pov.pdf)

**US FOOD AND DRUG ADMINISTRATION** **Silver Spring, MD, USA**
*Quantitative Research Fellow*

- Division of Quantitative Methods and Modeling in the Office of Research and Standards within the Office of Generic Drugs.
- Applying mathematical analysis to physiological/molecular based models for drug absorption, bioavailability, distribution and effectiveness. Using large data sets to improve the prediction and regulatory decision making for generic drugs.

## EDUCATION

**University of Michigan** **Ann Arbor, MI, USA**
*PhD Pharmaceutical Sciences* *Sep 2009 – May 2015*

- *Mechanistic Analysis and Quantification of Gastrointestinal Motility: Physiological Variability and Plasma Level Implications*

**New York University** **New York, NY, USA**
*MSc Computational Biology* *Sep 2007 – May 2009*

**University of Arizona** **Tucson, AZ, USA**
*BSc Mathematics* *Aug 2002 – Jun 2006*

## SELECTED ARTICLES

F. Bordes et al. An introduction to vision-language modeling. *arXiv*, page 2405.17247, may 2024. doi: 10.48550/arXiv.2405.17247. URL https://arxiv.org/abs/2405.17247.

J. Engel et al. Project aria: A new tool for egocentric multi-modal ai research. *arXiv*, page 2308.13561, aug 2023. doi: 10.48550/arXiv.2308.13561. URL https://ui.adsabs.harvard.edu/abs/2023arXiv230813561E/.

A. Talattof and G. L. Amidon. Pulse Packet Stochastic Model for Gastric Emptying in the Fasted State: A Physiological Approach. *Molecular Pharmaceutics*, 15(6):2107–2115, jun 2018. ISSN 1543-8384. doi: 10.1021/acs.molpharmaceut.7b01077. URL http://pubs.acs.org/doi/10.1021/acs.molpharmaceut.7b01077.

A. Talattof, J. C. Price, and G. L. Amidon. Gastrointestinal Motility Variation and Implications for Plasma Level Variation: Oral Drug Products. *Molecular Pharmaceutics*, 13(2):557–567, feb 2016. ISSN 1543-8384. doi: 10.1021/acs.molpharmaceut.5b00774. URL http://pubs.acs.org/doi/10.1021/acs.molpharmaceut.5b00774.