

Anjan Goswami

San Francisco Bay Area, CA, 614-429-7689

Objective: *Search, Data, and AI Leadership*

goswami.anjan@gmail.com

<https://www.linkedin.com/in/goswamianjan/>

Executive Summary:

Versatile leader with a proven track record in building, managing and scaling high-performing engineering and applied science teams across startups and established large corporations. Deep expertise in developing impactful AI, search and recommendation products from inception targeted toward significantly improving business metrics. Skilled at providing technical leadership to small, agile teams as well as large, diverse engineering units successfully driving results within varied corporate structures.

Areas of Expertise:

- AI Domains: Computer Vision, Natural Language Processing, Generative AI, Large Language Models, Reinforcement Learning.
- Responsible AI: Fairness, Bias, Privacy, Legal Consideration, Model Security.
- Information Retrieval: Search, Recommendation, Personalization.
- Data Science: Product Analytics, Controlled Experiments, Evaluation of Models and AI Products, Statistical Aspects.
- Enterprise AI: Ideation, Development, Deployment, Customer Success, Solution Engineering.
- ML Platforms: Training and Inference Systems, MLOps.
- Edge ML: Model Miniaturization, Energy Efficiency, Platforms for Mobiles and Desktops.
- Data Platforms: Ingestion, Storage, Governance, Quality, Retrieval.
- Software Engineering: Software Systems Design, Development Strategies, Agile Methodologies.
- Organization Design: Building and Scaling Applied Science and Software Engineering Teams.

Professional Experience

Director of Machine Learning Science and Engineering, Adobe (July 2019 - May 2023):

Adobe Document Cloud (December 2021 - May 2023):

- Directed an organization of over 50 machine learning scientists and engineers, fostered the creation of AI-driven experiences through collaborative efforts with research, design and product teams.
- Standardized development processes and introduced fast, efficient and intuitive methods for AI model development, validation and improvement.
- Enhanced model development by implementing parallel models, distributed training, optimal inference and MLOps practices. Substantially improved model efficiency through applications of data-centric AI techniques, shadow testing, fast acquisition of high-quality training data, weak-labeled and synthetic data, feedback loop, hyper-parameter optimization, neural architecture search, model distillation.
- Developed essential tools for debugging models, privacy-preserving visualization, monitoring model performance and data quality.
- Amplified user experience in Adobe Sign by tripling the efficiency of form field detection.
- Improved PDF views in liquid-mode for Acrobat iOS app with state-of-the-art vision models significantly enhancing document layout detection and table structure extraction.
- Released a new fast and comprehensive multilingual document layout detection (compared to its competitor) service for enhancing view and edit experience of Adobe Express in 15+ languages.
- Led the prototype development for Q&A, Summarization and Highlights functionalities in Acrobat using large language models. Established an effective pipeline for document processing and prompt engineering.

Adobe Search (January 2021 - December, 2021):

- Led the design of machine-learned ranking for Adobe Stock (image search) and image recommender systems for the Creative Cloud improving key search quality and business metrics with a focused team.
- Architected indexing low-dimensional embedding-based tokens using a multi-modal (CLIP) model and k-sparse autoencoder to enhance long-tail of image recommendation results.

Digital Experience Cloud (July 2019 - December 2020):

- Led 50+ people teams to develop enterprise AI for Adobe's Experience Platform, including AI-driven audience discovery and machine-learned cross-device identity stitching frameworks.
- Achieved 4x cost efficiency with Beta release of an AI-driven audience discovery platform.
- Doubled the recall of a cross-device identity stitching framework while retaining its precision.
- Introduced a no-code recommendation model for retail and data quality assessment for Adobe's Digital Experience platform.

Director - Data Science and Engineering, Salesforce (July 2016 - July 2019): Service Cloud, Salesforce (August 2018 - July 2019):

- Led multiple machine learning teams to create a roadmap for Salesforce Service Cloud's AI-powered customer service.
- Built and launched an article recommendation system for customer agents, improved service cloud chatbot's adoption across multiple verticals such as retail, banking and hospitality and expanded chatbot functionality to multiple languages.
- Introduced a semi-automated case completion tool for customer agents and prototyped a dialogue recommendation system for Salesforce's service cloud chatbot.

Community Cloud, Salesforce (July 2016 - July 2018):

- Formed and led the Community Cloud (an enterprise social network) AI division comprising 40+ people across science and engineering teams in EU and US. Built first dedicated data science modeling team. Team in EU under my leadership received best team award for fast and successful delivery.
- Released four recommender systems, a feed relevance system and a question-answering engine which led to 20% increase in user engagement.
- Strategically onboarded 5 large enterprise customers adding significant annual contractual value and played a key role in M&A efforts.
- Led several product analytics initiatives to track trends in engagement, daily and monthly active users (DAU and MAU) to understand the impact of machine learning models over time.
- Designed key quality metrics and visualizations for explainability of the machine learning-driven functionalities for developers and customers.
- Led the development of a counterfactual simulation framework to evaluate feed ranking models.

Founder and Managing Consultant, SmartInfer LLC (November 2015 - August 2019):

- Founded and ran a successful revenue positive consulting firm providing high-level consultation services for clients like Etsy.com and Neurotrack.
- **Client Neurotrack:** Devised a plan to enhance the algorithmic performance of Neurotrack's product – a diagnostic software designed to detect early-onset Alzheimer's. Provided technical leadership for the development of a fast and accurate pupil-tracking algorithm from video data and managed Neurotrack's first successful product release.
- **Client Etsy:** Conceived a comprehensive roadmap for augmenting Etsy's search technology encompassing both organizational design and technological improvements. I designed a statistical framework to accurately interpret key business metrics such as clicks, add-to-cart, sales and revenue with respect to queries and items aiding in the detection of high-priority customer search issues. Advised on the UI design to enhance user interaction with data visualization. Architected and oversaw the development of a machine-learned spell engine for Etsy.

Director of Search Science, Engineering and Product, Walmart E-commerce (June 2014 - October 2015):

- Oversaw engineering and product teams of 80+ personnel, working on areas including relevance, discovery, query understanding, ranking infrastructure, autocompletion, spell correction, browse relevance, mobile and grocery search and product analytics.
- Pioneered a statistical framework for identifying key challenges with existing search system and crafted an engineering roadmap resulting in a 23% increase in revenue and 17% sales growth within a year.

- Spearheaded the development of critical AI-driven commerce search software components such as machine-learned ranking, query understanding, price search, spell correction and auto-completion frameworks significantly boosting search optimization efforts.
- Launched a new search engine for Sam's Club enhancing sales and engagement metrics. Developed new commerce searches for ASDA, Walmart Grocery and Stores.
- Established a university partnership program to promote academic collaboration and expand R&D efforts.
- Boosted productivity of the e-commerce R&D team in Bangalore by fostering a culture of autonomy and accountability optimizing workflows and outcomes.
- Streamlined hiring processes successfully recruiting over 10+ skilled machine learning engineers within a year and significantly enhancing productivity.

Director (Principal Architect) Data Science and Recommendation Systems, Elance/UpWork (June 2013 - June 2014):

- Built and led the data science and recommendation engineering at Elance, a startup that became public through an IPO as UpWork. Successfully scaled the data science organization and crafted the roadmap for its AI technologies.
- Architected and developed a contractor recommender system from scratch leading to a 4% increase in hiring rate.
- Architected and led the development of a controlled experiment platform and wrote the statistical package to analyze the results.
- Fostered academic partnerships with management and economics departments at UC Berkeley and Harvard for research on labor markets.

Lead Software Engineer, eBay (April 2010 - June 2013):

- Played a pivotal role in architecting ranking and query serving components of eBay's new commerce search engine (Cassini). Led design and development efforts on multiple search relevance enhancements to transition eBay's rule-based ranking to machine-learned ranking resulting in a revenue increase of over 6%.
- Led development of eBay's image search for fashion via the addition of the more-like-this feature and designed algorithms for assessing product image quality leading to a significant improvement in user experience and search efficiency.

Lead Software Engineer, Amazon Product Search (A9) (December 2007 - April 2010):

- Led the development and architectural design of key search components including federated search ranking, related search and spell correction systems. Reengineered text processing algorithms and feature engineering using Hadoop MapReduce improving the relevance and customer experience.
- Enhanced search functionality by introducing new ranking features such as item availability, query and item's title phrase match score, left-hand-side relevance.
- Contributed to global growth initiatives by aiding in the launch of Amazon China and enhancing search relevance for Amazon Japan and Germany.

Senior Software Engineer, Microsoft (May 2006 - November 2007):

- Developed components of indexing, ranking and crawling for Microsoft's search engine. Key contributor to Microsoft's search codebase in C++.

Lead Search Software Engineer, MyFamily (Ancestry) (April 2005 - April 2006):

- Led the development of search ranking and indexing for Ancestry's historical newspaper data digitized using optical character recognition (OCR).
- Implemented an Optical Character Recognition Engine and a Name Extraction Service for historical newspaper search.

Education

- **Ph.D., Computer Science, University of California, Davis (2018)**
 - Specialization: Artificial Intelligence, Distributed and Parallel computing.
- **Executive Education, Haas School of Business, University of California, Berkeley (2018)**
 - Certificates: Negotiation and Influence, High Impact Leadership.
- **M.S., Computer Science, Ohio State University (2005)**
 - Specialization: Distributed Data Mining and Big Data.
- **M.S., Computer Science, University of South Florida (2002)**
 - Specialization: Computer Vision and Graphics.
- **M.Tech, Mechanical Engineering, Indian Institute of Technology, Kanpur (1999)**
 - Specialization: Mechanical Design and Robot Vision.
- **B.E., Mechanical Engineering, Government Engineering College, West Bengal (1996)**

Selected Publications

- Machine-Learned Ranking Algorithms for E-commerce Search and Recommendation Applications. Doctoral Dissertation, University of California, Davis, 2018.
- Efficient Clustering Algorithms for Out of Core, Distributed and Streaming Data. Masters thesis, Ohio State University, 2005.
- Quantifying and Visualizing the Demand and Supply Gap from E-commerce Search Data using Topic Models. WWW 2019 workshop for NLP in E-commerce.
- Learning to Diversify for E-commerce Search with Multi-Armed Bandit. Submitted SIGIR ECOM 2019.
- Towards Optimization of E-Commerce Search and Discovery. SIGIR ECOM 2018.
- Controlled Experiments for Decision-making in E-commerce Search. IEEE Big Data, 2015.
- A Bayesian Framework for Modeling Price Preference in Product Search. NIPS 2015 workshop on ranking.
- Recommendation Systems for Markets with Two-Sided Preferences. ICMLA 2014.
- Assessing product image quality for online shopping. IS&T/SPIE Electronic Imaging, 2012.

Selected Patents

- Extraction of Keywords for Generating Multiple Queries. US 16140443
- Routing of Cases Using Unstructured Input and Natural Language Processing. US 16140418
- Service Agent Personal Recommender System. US 102.0369US
- An Omni-platform Question Answering System. US Non-provisional
- Method and System for Classifying Queries to Improve Relevance of Search Results. US 20120221557
- Behavior-driven multilingual stemming. US 8793120
- Image Quality Assessment to Merchandise an Item. US 13300305
- Availability-Aware Item Ranking. US 550731

Awards

- Graduate Fellowship at UC Davis, Research and Teaching Assistantships at University of South Florida, Ohio State University and IIT Kanpur.
- Ranked in Several All India Tests.