

Steven Li

Shanghai, China

stevenlee77@126.com

+8618616569017

[linkedin.com/in/steven-li-65911833](https://www.linkedin.com/in/steven-li-65911833)

Summary

Head of Data & AI at Starbucks till now

Chief Data Officer | Head of Data at Sanofi 2 years

Global Head of Data Services & Solutions at eBay 17 years

Experience



Head of Data & AI

Starbucks

Sep 2023 - Present

- Lead an integrated data team to deliver high-quality data solutions and services, fostering a data literacy culture and unlocking business value through data and AI initiatives.
- Develop and implement data-driven strategies to optimize supply chain, inventory management, and sales forecasting.
- Enhance customer personalization and engagement through advanced AI and machine learning models.
- Build up Data Middle Layer for big data foundation across 8 business domains includes Transaction, Member, Online Behavior, Store OPS, Marketing Promotion, Store Development, Supply-Chain and LBS
- Establish the One Analytics Hub 1.0 to enable descriptive, diagnostic, and actionable insights for Sales, Marketing, Customer360, and Product Innovation.
- Collaborate with cross-functional teams, including Marketing, Digital Ventures, Supply Chain, Store Development & Operations, and Finance, to identify and prioritize data initiatives aligned with business goals.
- Lead a team of data scientists, analysts, and engineers, cultivating a culture of innovation and continuous improvement.
- Implement data governance policies to ensure data quality, security, and compliance.
- Drive the adoption of advanced analytics and AI technologies, staying current with industry trends and innovations.



Head of Data & AI, Greater China

Sanofi

Sep 2021 - Sep 2023 (2 years 1 month)

- Led the integrated data team at Sanofi, comprising Data Architecture, Data Science, Data Analytics, Data Engineering, and Data Governance, to build a robust data foundation and advance analytics and AI technologies, supporting healthcare and Pharmaceutical GBUs.
- Established a big data foundation and OneMesh on a cloud-native platform to drive commercial business initiatives.
- Directed the Data & Analytics team, overseeing data strategy, governance, and analytics across the organization.
- Led the AI Engineering team in creating a Generative AI internal platform and framework, pioneering AIGC use cases to drive business impact.

- Developed and implemented data-driven strategies to optimize drug discovery, clinical trials, and patient engagement.
- Collaborated with R&D, clinical, and commercial teams to identify and prioritize data initiatives aligned with business goals.
- Mentored and managed a team of data scientists, analysts, and engineers, fostering a culture of innovation and continuous improvement.

Implement data governance policies to ensure data quality, security, and compliance with regulatory standards.

Drive the adoption of advanced analytics and AI technologies, building OneAI and OneBI to enable business innovations

Global head of Data Services & Solutions

eBay

Sep 2019 - Jul 2021 (1 year 11 months)

Lead eBay Global (San Jose and Shanghai) Data Services and Solutions team in eBay.

Build Data Platform and Analytics Platform in eBay to enable data driven eBay e-commerce business including buyer, selling, marketing, advertising and risk control.

Build Graph Intelligence for Fraud Detection, Ads and Marketing

Build Data management ecosystem including data governance, metadata management and data quality automation platform for entire eBay big data foundation.

Director of Data Solutions & Services at eBay

eBay

May 2018 - Aug 2019 (1 year 4 months)

Leading team build Augmented Analytics product to deliver automated insights for eBay product and business by self-enrichment Knowledge Graph with nature language experience.

Leading key initiative to build open source Data Processing platform (Spark+Hadoop) to enable entire eBay enterprise DW migrating from Teradata, which includes 60PB daily processing data. Based upon that, data services and products enables eBay business decisions and site features, so it has to be always available and accurate. 1. Build a full automation framework from code to process with 90% of them migrated in automatic way 2. Scale Apache Spark capability and scalability for a such large data warehouse, we extend key features like Adaptive Execution, Indexed Buckets(PPI in TD)

Director of Data Solutions & Services eBay

eBay

Oct 2016 - May 2018 (1 year 8 months)

As the Director of eBay Data Services & Solutions CCOE team.

Have been built strong Platform engineering, Data engineering and Data Science team in CCOE.

Created the vision and strategy for Big data and analytics ecosystem for eBay.

Lead team to build big data service, enable self-service analytics capability and establish machine learning engineering solution for eBay global.

Built data driven products and services to support company key initiatives including Structured Data, B2C and C2C Selling, Trust, Risk and Offebay.

Sr. Manager of Data Service & Data Solutions

eBay

Dec 2004 - Sep 2016 (11 years 10 months)

As the leader for DSS China team, created strategy and team goal, and established tigher partnership with North America and Europe business units and stakeholders,

- # Lead Product Managers and Engineers to design and build Analytics Products in multiple domains like Seller Experience, Inventory and Trust.
- # These Data driven products are not only for providing insight but also giving actionable recommendation to improve eBay business.
- # Lead Data Architecture team to establish Real Time platform – Kafka/Storm and also Hadoop/ Teradata Big Data Platform to enable Data Services for Site, 3rd party applications and end users with real time and big data process capabilities.
- # Lead team to build eBay Hadoop Data Platform. The platform enables big data processing and global data access with unified Metadata. It also enables self-service of data landing, publish and 3rd party data integration for Business User, Analyst and Data Scientist. The new platform dramatically promotes eBay analytics capability on Hadoop.
- # Lead DW Dev team to build Tri-Load (2 Teradata + 1 Singularity) Data Warehouse system (which is largest DW system in the world). Including ETL infrastructure (Abinitio 2.5) implementation, Job Scheduling tool (UC4) design and ETL automation desing. Also responsible for effective communication and collaboration between 3 Dev Centers (Shanghai, San Jose, and Seattle)
- # As Dev Manager for Search Analytics & Mobile Analytics teams, built Web Analytics products & BI reports. Including data analysis on eBay Sojourner data and transactional data, building user behavior metrics and purchase metrics.
- # Lead the DW Operations, Unix team and the Business Units to build eBay Data Licensing Platform which providing eBay Core Data for 3rd party partners, In this program, I worked with US Distinguish Architect for designing the architecture, building the automation ETL process and analyzing the business requirement.

Senior DW Engineer

BearingPoint

May 2003 - Jan 2005 (1 year 9 months)

Business Area: Medicine manufactory

Worked as a Senior Consultant in large scale Data Warehouse on-site project for Pfizer UK, The Client is the one of biggest medicine manufactory in the world. The client also had a number of legacy systems (ERP, CRM), and with the company rapidly expanding, the sales, client and product data is increasing extremely. Considering the data volume is huge and data structure is extremely complex, the team decided to use Informatica 7.0 as ETL tools to consolidate data from various data source into Data Warehouse which is built on, and use Cogons as business intelligence tool to build BI report for trend analysis. As DW senior Consultant, I collected and analyzed client's requirement and worked with Bearingpoint US. Expert to design the Data Model, implement the ETL process and create Cubes (Data Mart) for BI reports.

Business Area: Retailing

Worked as a Senior Consultant in large scale Data Warehouse on-site project for a leading multination accouterments company (Kellwood) in USA, The Client is one of USA's largest accouterments manufactory, rapidly expanding its operations. The client also had a number of legacy systems(ERP,CRM), the data format and schema of these systems are different that needed to be consolidated for BI Report and Trend Analysis. Data structure in the legacy systems is extremely complex and data volume is very huge. The team decide to utilize the leading ETL tools, DataStage, to extract, transform and load data to Data Warehouse and Oracle Database to store and maintain the data, which involved all aspects of Data Warehouse implementation, from requirement collection, data model build, metadata management, ETL to OLAP and Reports. The solution allowed client to all

aspects data and history data in time from various systems for BI Report and Trend Analysis whenever data volume grew extremely fast.

Software Developer

Envocis (Shanghai) Co., Ltd

Jun 2001 - Mar 2003 (1 year 10 months)

Worked as core developer for PTP collaboration office software. The software is for business collaboration in Peer To Peer communication channel. This system is built on Microsoft .NET platform.

- Defined system architecture and communication model
- Developed low level communication mechanisms with socket and .Net Remoting
- Developed XML-base message encoder and decoder package
- Provided technical support to development team and troubleshooting

Technologies Used: MS SQLServer 2000, Windows 2000 Server, Internet Information Server 6.0, .Net Framework, C#, WebServices, XML

Software Engineer

GUYA

Jul 2000 - Jun 2001 (1 year)

Worked as senior software engineer on the HIS (Hospital Information System) for the local hospital, which included three different systems (DIC System, PACS System, RAS System) which consolidate and analyze separate data sources and build report. The DICOM interface for the MRI, DSA is high technology in medical equipment domain. The protocol of the DICOM is extremely complex. Additional, the imaging data from MRI, DSA is very huge. The team decided to utilize OPENGL API and C++ to develop the Imaging receiving and processing System, and utilize low level TCP/IP Socket API to encode and decode DICOM protocol. The solution allowed client to receive the imaging from digital medical equipment and load to PACS,RAS system automatically, and view diagnosis report in local network terminal.

Education



Shanghai University

Bachelor, Electronic & Communication Engineering

1996 - 2000

Licenses & Certifications



Neural Networks and Deep Learning - DeepLearning.AI

UMNQWNN7SS57



Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization - Coursera

NSWXAT6YB8VK



Data Science in Python - Coursera

BKKAZZF6SWMW



Applied Machine Learning in Python - University of Michigan

TSYYBR5FBGJW

Skills

Big Data Analytics • Generative ai • Big Data • Technical Leadership • Data Warehousing • Machine Learning • Data Governance • ETL • Hadoop • Business Intelligence