# Sinchan Banerjee

## Summary

- Principal Data architect with 15 years of hands-on experience to lead and deliver enterprise grade data platforms, large-scale data migrations and digital transformations.
- Expert in architecting, implementing, and delivering enterprise grade data engineering solutions for both batch and streaming workloads with terabytes of data in cloud as well as in on-premises.
- Possess Organizational skills with the ability to coordinate with different internal and external teams under difficult situations and the ability to bring out the best in others while creating a healthy work environment.
- Provides Technical advisory to executives and top leadership with respect to data strategy, data governance and tools.
- Won prestigious national award - **Indian Achievers Award, 2024**.
- Awarded and recognized as one of the "**Top Men Leaders to look up to in 2024**" globally by Passion Vista Magazine.
- Invited to author a book on data engineering by packt publication- "Scalable Data Architecture with Java", published on 2022.
- Recipient of multiple awards across the industry for Innovation, Leadership and Technical excellence.
- Lead Inventor of a **Patent on Predictive Analytics** and author of multiple international publications.
- Judged multiple industry wide awards in his domain of expertise.
- Start-up and Big Data Enthusiast who helped and guided companies to adopt scalable data engineering solutions for their problem.
- Experienced in mentoring, motivating, and building highly capable team from scratch.

## Experience

**UST - Data Architect II (Principal Data Architect)**
Mason, OH
*01/2020 - Current*

- Serves as the principal data architect for all UST engagements and partnerships with Elevance Health across the business units.
- Architected, developed, and lead the solution delivery of Enterprise Data Platform (EDP) for Elevance Health as Principal Architect. EDP is the centralized real-time data repository for all engagement platforms of ELV directly delivering useful data and insights to over 80 million subscribers. Established multiple innovative framework, patterns, standard and best practices which enabled fast-paced solution delivery with lesser efforts and robust quality standard.

## Contact

- mr.sinchan.banerjee@gmail.com
- 2673045030
- West Chester, OH 45069
- https://www.linkedin.com/in/sinchan-tech/

## Education and Training

01/2024
**Massachusetts Institute of Technology**
Cambridge, MA
Artificial Intelligence in Healthcare: Fundamentals and Applications

07/2023
**Cornell University**
Ithaca, NY
Executive Leadership Program: Organizational Leadership

07/2018
**University of California - San Diego**
La Jolla, CA
Course: Machine Learning with Big Data
- Grade: 98.20%

2009
**SRM University**
Chennai, India
Bachelor of Technology: Computer Science Engineering
- Merit Scholarship Recipient
- 9.17/10 GPA

- Enabled first-of-its-kind real-time organized data layer for engagement portals for MEDICAID business.
- Key contributor and served as lead data architect in analyzing the whole footprint of Elevance Health in AWS for all its business tower and create plan, estimate and roadmap for transformation and migration to Azure.
- Analyzed the whole ecosystem of Platform Services and Data Engineering tower at enterprise level for Elevance Health to identify and document patterns, improvement areas, gaps, and opportunities. Worked closely with Client VPs and directors to establish a roadmap to mitigate the existing gaps and holistically improve the platforms, processes, and practices.
- Instrumental in solutioning, estimation, creation of delivery roadmap and presentation of multiple successful UST's RFPs worth millions of dollars for healthcare, telecom and Food & beverages sector.
- Served as Principal technical advisor for client executives across Elevance Health and its subsidiaries where UST is involved.
- Reviewed multiple architectures, established best practices, identified performance problems, and provided sustainable solutions, help troubleshoot problems and conducted training and mentorship for newer technologies.
- Leads team of architects and SMEs to create and establish reusable and scalable solutions and frameworks.
- Provides consultation to cross-functional teams and initiatives across the organization for technical strategies, technical solution reviews, suggestion, or cost optimization.
- Mentored and built multiple cross-functional teams from scratch in EDP area.
- Conducts training and presentation for other engineers in UST as well as for client for newer technologies and mentor them in as is need basis.
- Works closely with talent acquisition team and account managers to assess and interview both internal and external candidates
- Member of corporate wide Internal Advancement Review (IAR) council for UST
- Member of UST iMentor Council (mentorship program)
- Received **UST Star awards** in **2021,2022, 2023 and 2024** for Leadership, Innovation and Agility

**DataCore Systems - Big Data Architect**
Scottsdale, AZ
*01/2019 - 01/2020*

- Architected multiple real-time data pipelines to ingest data from Salesforce or other external sources to Hadoop and MemSQL/SingleStore clusters using Kafka Streams and Kafka Connect. Designed and developed realtime federated updates and search on MemSQL realtime data layer using GraphQL braids.
- Architected and developed batch based raw data aggregation and ingestion pipelines from Hive and HBase to MemSQL realtime database using Apache Spark and Apache NiFi.

- Created Multiple Micro-services on Openshift Cloud using Spring Boot, Spring Cloud , Slueth, Ishtio Service Mesh, Kubernetes and Docker to Publish, read and analyze Data from Kafka and MemSQL real-time Database
- Architected and implemented above mentioned framework components which increased ease of development by 50% and time to market by 40% for teams across the organization.
- Helped teams across organization to adapt to this new framework.
- Collaborated and coordinated with various teams to analyze business needs and their pain-points and propose newer features for the platform. Also, brainstorm such features with Staff architects to create feature blueprints for the platform.
- Coordinated and managed other engineers to successfully implement and develop the framework components.
- Mentor, monitor and guide engineers in the team to deliver efficiently.
- Train and Guide engineers from other teams to implement and use the platform components. Also, created solution architecture for their use-cases using the platform.
- Actively participated in multiple seminars, presentation, and demos to create broader awareness and acceptability of the platform across Amex.
- Helped, coordinated, worked and reviewed with external committers across the Amex for the platform towards the goal of making the platform open source.

**United Health Group - Senior Big Data Engineer**
*02/2017 - 12/2018*
- Redesigned data ingestion framework for Unified Data Lake (known as SDR) for HEDIS Data coming from various sources to give a 7x performance boost using Spark on MapR clusters. Presented the new framework POC and got funding to implement the solution.
- Lead delivery of the solution .
- Designed and lead the delivery of generic data provisioning framework to provision on-demand data to downstream HEIDIS sources from the HEDIS SDR Data lake .
- Architected and proposed a real-time ingestion framework for SDR, paving the way for next-generation streaming ETL for SDR data lake.
- Designed and developed reporting framework and workflows for SDR.
- Established CI/CD pipeline for big data jobs using Jenkins, GitHub, docker and Cucumber.
- Architected , implemented and delivered Common Analytic Framework, a Data Lake for Storing and searching Fraud Analytic Solutions. It was developed and designed using Elastic Search, Log Stash and Spring Boot.
- **Awarded 4 times by UHG** for Innovation, teamwork and Spectacular Performance in the year 2018
- Nominated as innovation champion to drive and improve innovation within the team.

## *Membership*

Senior IEEE member
Member of MIT CSAIL

## *Judging*

- Judge of Regeneron International Science and Engineering Fair (ISEF), 2025
- Judge for Big Innovations Award, 2025
- Judge for Stratus Award (cloud computing), 2024
- Judge for Big Awards for Business, 2024

## *Media Coverage*

- Feature Story by Passion Vista Magazine: - https://www.passionvista.com/sinchan-banerjee/
- Interview with Pack publication - https://partnerships.packt.com/interview-with-sinchan-banerjee/

## *Skills*

- **Leadership**: Collaborative Leadership, Technical Project Leadership, Mentoring, RFP,

- Involved in creation of automation test framework for an ingestion tool for Enterprise Data-Lake using Python, HBase, hive and shell script.
- Ingest huge volume of data into Optum Data-Lake using HBase, Hive, Sqoop, Spark and Talend.

**Impetus Technologies - Module Lead Software Engineer (Big Data)**
*12/2015 - 01/2017*

- Designed and developed a data validation tool for an ingestion framework using Spark, Drools and Hive for a huge data Lake for an US based Financial organization.
- Involved in enhancement and performance tuning of a Spark based archival reporting tool for a financial institution.
- Worked as a developer in Hadoop map-reduce based Marketing and offer recommendation engine for a US based financial firm.

**HPE - Software Engineer (R&D- Data Storage)**
*01/2012 - 11/2015*

- Owned the online database and the Datawarehouse for HPE StoreOnce Enterprise Manager.
- Owned Forecasting feature of HPE StoreOnce.
- **Patent on a Storage Capacity Forecasting** model which **outperformed in accuracy the industry-leading** capacity forecasting model at that time (Patent Id: US 15102997)
- **Awarded 4 times by HP** for thought leadership, exceptional customer support and amazing work delivery of multiple features and tools in **the year 2012,2013 and 2014**.

**Tata Consultancy Services Limited - Systems Engineer**
*10/2009 - 12/2011*

- JAVA developer & designer with Spring 3.0 , IBM Websphere / BEA Weblogic , XML/SAX parsing, Java Concurrency, performance tuning (Application level) experience
- Proposed technical feasibility solutions for new system designs and suggested options for performance improvement of technical components.
- Automated data collection and testing activities to increase productivity.

## *Patents*

"Storage capacity regression", Patent (Patent Id: US 15102997) on Storage

capacity forecasting-2016

## *Publications*

- "Scalable Data Architecture with Java", book on Data Engineering by Packt Publication-2022. This book is indexed in the catalog of top

---

- Planning, Strategic Decision making, Team building
- **Data Engineering:** Spark, Kafka, Kinesis, NIFI, Flink, Hadoop, Map Reduce, Hive, Sqoop, Yarn, Flume, Pig, Zookeeper, Informatica, Talend, DeltaLakes
- **Cloud:** AWS, Azure, Openshift, Confluent Cloud, Cloudera Cloud platform, Databricks
- **Languages:** Java, Scala, Python, Groovy
- **Databases:** Memsql, Postgresql, mysql, Oracle, AWS Aurora, Snowflake
- **NoSql :** HBASE, MaprDB, Cassandra, MongoDB, AWS DocumentDB, Redis, AWS ElasticCache, KsqlDB
- **Search Engines:** Elasticsearch, Apache Solr
- **Visualization & Monitoring:** Kibana, Graphana, Splunk, Tableu, Wiley, Data Dog
- **Machine Learning Tools:** Spark ML, KNIME, Fuzzify, AWS ML, TensorFlow
- **Workflows:** Oozie, CntrlM
- **OS:** Linux, Windows, Mac,Unix
- **CICD:** Kubernetes, Docker, Jenkins, Sonar, Prisma, Cucumber, Bamboo
- **Industry**: US Health Insurance, Finance

university libraries across the globe including MIT Library, Harvard Library, Berkley Library, Stanford Library, Oxford Library, CSU library to name a few.

- "Storage Capacity Forecasting for De-duplicating Systems using Segmented Linear Regression Algorithm and Intelligent Fuzzy Logic", HP TechCon-2014
- "Trust Enhanced Dynamic Source Routing Protocol For Adhoc Networks", WASET-2009
- "A Novel Routing Technique Against Packet Dropping Attack in Adhoc Networks", Journal of Computer Science-2008