

Harish Chandra

🌐: Hyderabad, India ✉ : harry.rathor@gmail.com ☎: +91 8588863362 [LinkedIn](#)

About me:

I am a nocturnal developer, building innovative self-service-based platform, frameworks and solutions at world biggest tech giants. I have collected an array list of technical skills in my bag of tricks.

Professional Summary: 10 Years

#	Company	Designation	Start	End
1	🍏 Apple Inc.	Software Engineer (ICT 3)	February 2019	CURRENT
2	Intuit Inc.	Senior Software Engineer	June 2018	February 2019
3	Walmart Global	Development Engineer III	May 2016	June-2018
4	Snapdeal	Data Engineer II	September 2014	May 2016
5	InfoTrellis Pvt Ltd	Software Engineer	June 2012	September 2014

Academic Profile:

#	Board/University	Course	Year
1	University Grant Commission India	UGC- NET (Computer Science)	2012
2	Department of Computer Science, University of Delhi , New Delhi	Master of Computer Application (MCA)	2012
3	Jiwaji University Gwalior (M.P.)	Bachelor of Computer Application (BCA)	2009
4	M.S.V inter collage Kalpi (U.P.)	Intermediate (12th)	2006
5	Mahavir collage Neem Gaon (U.P.)	Higher Secondary (10th)	2004

Technical Proficiency:

Operating System	LINUX, UNIX, Windows, Mac
Development Editor	Eclipse / IntelliJ / STS
Programming Languages	Java, Python & Unix Shell scripting
Technologies	Hadoop (HDFS, Spark, Sqoop, Oozie, Hive), JSON APIs (GSON/Jackson), Logging (slf4j, Log4j), API, parquet MS - {IBM-MQ/ActiveMQ/ Kafka }, Hazelcast, Alluxio
Cloud Technologies	AWS, Azure, GCP

Web Technology	Spring- {boot, data, JDBC, JMS}, Tomcat server
Database	RDBMS (MySQL, Snowflake, Teradata), NoSQL (Mongo, Cassandra)
Others	Design Patterns, Data Structure, Networking, Operating System, Git

Key Project Details:

🍏 **ACD/Keystone Pipelines:-** I have written applications on **Keystone Data Platform**, where data comes from GBI/ACI Kafka and get persist into multiple Teradata & Snowflake simultaneously

- Create and Deploy pipeline for ACD team
- I have done troubleshooting of pipeline when it fails because issues like format, environment, Retry, Scheduling
- Performance analysis and tuning of pipelines do benchmarking based on requirement

🍏 **Evija™** - Build data lake software to cater data services where data get replicated and processed to/from **AWS S3/Kafka/Teradata**

- Design & Development of Framework where user can create pipeline to transfer data into **S3 Storage**
- Run periodic query and send notification when it complete
- Enable encryption of data on the fly and data at Rest
- Do Performance analysis based on format of source and find out optimal performance

🍏 **FalconX™** - Project is self service based tool to cater transfer data from **Teradata <-> Kafka/Hadoop/S3** and vice-versa.

- Design & Development of framework to serve Marketing data transfer request
- Make use of encryption/decryption of password when store locally
- Created DB Utility for audit jobs/pipelines
- Setup many pipeline for Marketing team

🍏 **Workload Simulator™** - **Workload simulation framework** is to mimic production load and generate performance reports to evaluate different warehousing 3rd party software and It help to decide optimal Cloud based on requirements. This project help in making decision between **BigQuery(GCP)** / **Snowflake** to replace Teradata Based on running Load in GBI

- Build Framework to run query parallel queries
- It can be configured based on number user and Threads
- Generated report of performance and conclude inferences of performance

🍏 **Walmart Customer 360™** -

Project is mainly focus on matching of Customer's demographic and transactional data on various domain and create a 360 view of customer using deterministic and probabilistic technique and tracking customer activity on real-time basis. It was deployed in **Azure**

Responsibilities –

- Development of matching deterministic and probabilistic on java
- Apply these on real time using Kafka and Spark Streaming
- Create 360 view of Customer demographic and transactional data on Janus Graph database.

Walmart Raven™ -

Raven Platform is based on Java Micro Service Architecture, composed of REST based services. It supports Assortment Team to run Analytical Requests and send back data in requested format.

Responsibilities –

- Development of Spring-boot micro services that REST asynchronous request and process into spark
- It built on cutting edge technologies like: Alluxio, ActiveMQ, Spark, Spring-boot, Cassandra
- Horizontally scalable & distributed using Spring eureka, Config, Zuul, admin

Snapdeal Pulse™ -

Snapdeal pulse is real time analytical engine build on top of Hadoop which do analytics based on social media data twitter, and customer User Question & Answer data.

Responsibilities –

- Fully responsible for End to End design and development of core engine.
- It is in production and is able to find product trend in twitter
- Collection data from twitter and internal data
- Do matching in internal product/service and generate report real time and periodic basis.

Snapdeal Replication Engine™ -

Framework is does replication of data from different data sources. It Design in such a way it moves data from source to sink in incremental load fashion and Framework is plug-able enough to accept new source and sink Vertica, MySQL, mongo, HBase, HDFS etc.

Responsibilities –

- Fully responsible for End to End design and development of framework.
- It is in production and doing replication about 30 databases and 1600 table with huge data from Vertica/ MySQL to HDFS in 3 time in day. Destination data size about 3 TB.

Snapdeal Log Processing Engine™ -

I innovated this module, it processes server log offline using Hadoop components and extract fact about users and product which is proving positive to the business.

InfoTrellis AllSight™ -

A Big Data product built using Hadoop technology like MapReduce 1, Hadoop, HBase, Hive, Oozie, Sqoop and Pig, to provide a platform for big data analytics and also includes a set of big data solutions built on top of it.

Responsibilities –

- Developed “**Social Media Master**” Framework is core framework in “**AllSight**” to extract hidden customer information in social media data like twitter and Facebook on top of (Oozie, pig, HDFS)
- Originated “**Social Media Toolkit**” to collect Facebook feed on top of HBase, H2 database.
- Worked on developing Twitter and Facebook data collection jobs for “**Social-Media-Master (SMM)**” module. SMM is implemented in java and invoked through pig UDF for collecting Facebook, Twitter data. The collected data is then transformed into JSON format. HBase and RDBMS (MySQL/H2) used for Data storage.

- Formulated a unique extractor pluggable feature in SMM Framework to allow users to plug in required Extractor among the set of OOTB SMM extractors.
- Developed major SMM Extractors like Keyword extractor, User Profile extractor, User Location extractor and Relevance, Sentiment extractor using various self-evolved algorithms.
- Created UNIX shell script “**AllSight Installer**” for installing AllSight in Production environment. The basic duty of the installer is to,
 - ✓ Install component in local and HDFS file system
 - ✓ Creation of database, tables, indexes in MySQL
 - ✓ Deploy war files in tomcat
 - ✓ Prepare dynamic launcher for launching job
 - ✓ Dynamic uninstaller for AllSight Uninstallation
- Developed “**REST Services**” using java, HBase, MySQL, Spring JDBC on top of Apache CXF framework and tomcat server. These rest services are utilised by “**AllSight Big Data Solutions**”.
- Hands on experience implementing **Machine learning** for **Sentiment Classifier** module. This classified used to find out whether a sentence is conveying positive or negative sentiment.
- Developed **Data Aggregator** and **Loader** framework that aggregates data stored in HDFS using HIVE loads the aggregated data using Sqoop.
- Worked on **Transformers Engine** which transforms data from another format to AllSight Compatible.
- Played a vital role in fine tuning the **Performance** of various modules of AllSight.

Awards and Achievements

- Awarded as Best Performer of the Year 2012 & 2013 in InfoTrellis Pvt Ltd
- Awarded Excellence award at Intuit Inc
- Awarded Spot Award at Walmart
- Cracked UGC-NET (Computer Application) June 2012 in very first attempt
- Shortlisted as Junior System Administrator @ Indian Institute of Science Bangalore (**IISc**)

Personal Profile

Father Name	: Kamlesh Kumar
Languages Known	: English, Hindi
Hobbies	: Swimming, cooking, driving, bike riding