Big Data Architect

Summary

Creative Architect who started his adventure working High-Availability and High-Scalability solutions on which he learned 2 basic principles: "It has to be simple", "It needs to scale easily". I've designed and implemented a data streaming system based on Lambda Architecture before it became a Buzzword. I joined Teradata International Architecture CoE, where I was creating design patterns for Data Lake implementation in Teradata Corp. In the current role I'm using all my skill and experience, to help companies designing tailored Analytical ecosystems.

Career Summary

Lead Architect, Polish Telco   2018 – 2020

- I was responsible for creating design of integrated analytical platform that is using for
- creating and training data models used for optimization of x-sell and upsell products
- of the company. In this project I have to design overall platform, data layers, data
- ingestion processes, data curation processes, Semantic layer for accessing the data.
- Used technologies: HDP, HDF, Kylo, Nifi, Apache Spark, Semantic Layer,
- Kerberos, Ansible, Jupyter, Apache Hive, Python, FreeIPA.

Solution Architect, Global Pharma   2019 – 2020
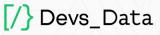
- I was responsible for designing some of components of the Data Hub created with
- Open Source technologies. Key achievement is design of Holistic backup and DR
- solution for Hadoop platform that covers customer business use cases. I was a
- working as a member of Excellence Community that was responsible for making key
- architecture decisions about the data ecosystem.
- Used Technologies: Cloudera CDH, Apache Hive, Apache Nifi, Teradata.

Ecosystem Architect, Austrian Telco   2018

- I was responsible for working with the customer on building an overall capability
- analysis for implementing a Data Virtualization layer and Unified cross platform ETL
- solution. I delivered a conceptual level architecture.
- Used technologies: HDP, ETL, Semantic layer, Data Virtualization.

Solution Architect, Austrian Bank   2016 – 2017

- Project was about creating an analytical environment in the Public Cloud that can process the data for optimising customer satisfaction and supporting risk scores. My responsibility was to create HLD, LLD, define data security policies and processes that were aligned with GDPR. Key achievements were: One of the first implementations of Data Lake in Public Cloud for the financial Sector. Dynamical scalable environment working on secure data repository.
- Used technologies: HDInsights, Apache Nifi, Kylo, Apache Ranger, Blob Storage, Apache Spark, Azure.

Education

Warsaw School of Economics, Post Graduate PM study   2014 – 2015

Krakow University of Technology, WIEiK, Teleinformatics
Finished with MSc thesis

Certifications: Hortonworks Certified Instructor 2016 | Hortonworks Certified Developer 2016 | Professional Scrum Master

Skills

Programming Languages: Python, Scala, Java SE, JavaScript, PHP7, R, Bash, Ansi C
Databases: Postgresql, MySQL, Teradata, HBase;
Big Data Ecosystem Components: Apache Hadoop, Apache Spark, Apache Hive, Apache Flume, Apache Nifi, Presto, Jupyter, Kylo, ELK Stack, Apache Atlas, Apache Ranger
Cloud Vendors: Azure, AWS,
DevOps Tools: Jenkins, Ansible, Git;
Project Management: Agile Scrum, Agile Kanban.

Languages

English: C1
Polish: Native
Spanish: A1