Sanchit Alekh

Sonnengasse 8, 74072 Heilbronn

☎ +49 1523.627.6687 ✉ sanchit.alekh@rwth-aachen.de · linkedin.com/in/sanchitalekh

Education

MSc. in Software Systems Engineering (Pursuing) ∘ RWTH Aachen University ∘ GPA: 1,4 ∘ 2016-2018

· Selected Coursework:

Theoretical Computer Science: Satisfiability Checking · Compiler Design

Data and Information Management: Implementation of Databases · Big Data in Medical Informatics · Privacy Enhancing Technologies in Data Science · Soft Computing · Artificial Intelligence · Information Retrieval

Applied Computer Science: Computer Vision · Virtual Reality · Graphics and Visual Computing

B.Tech (Honours) in Information Technology Indian Institute of Information Technology ∘ GPA 9.09/10 ∘ 2012-2016

Abitur/Higher Secondary School Exam Delhi Public School, Patna ∘ 93.6% ∘ 2010-2012

Work and Internship Experience

Audi AG ∘ Data Scientist Intern — since Jan 2018

· Currently working at the TechHub: Data Driven Production group, which aims to leverage big data technologies to make automotive production more automated and efficient

Technologies and Concepts: Data Lake, MQTT, RabbitMQ, Big Data, Smart Factories

Informatik 5, RWTH Aachen University ∘ Student Assistant — Jan 2016-Jan 2018

· Worked under PD Dr. Christoph Quix as a part of the mi-Mappa project (`dbis.rwth-aachen.de/cms/projects/mi-mappa`), which aimed to identify suitable actors for complex innovation in medical science

Technologies and Concepts: Medical Text Mining and Analysis, Information Extraction, Database Optimisation

Knowledge Mining and Assessment Group, TU Darmstadt ∘ Research Intern — May 2015-Jul 2015

· Under Prof. Ulf Brefeld, worked on computing a confidence measure on reliability of data on crowd-sourced and commercial encyclopediae

Technologies and Concepts: Python, MySQL, Information Extraction, Text Mining, Conditional Random Fields, Wikipedia API

Centre for Development of Advanced Computing Pune ∘ Project Trainee — Dec 2014-Jan 2015

· Built an anonymizer service for Medical Imaging standards: DICOM and HL7 using C-DAC's Medical Standards Toolkit.

Technologies and Concepts: Java, DICOM, HL7, DIMSE Services, HIPAA

University of Wisconsin Milwaukee ∘ Research Intern — May 2014-Jul 2014

· Analyzed and optimized present algorithms and toolkits for anonymizing personal information in Medical Health Data. Focused on Mu-Argus for structured and MIST for unstructured data.

Technologies and Concepts: Java, HIPAA, Text Mining

Skills

Programming Languages:	Java, Python, Scala
Operating Systems:	MacOS, Linux, Windows
Big Data Stack:	HDFS, MapReduce, Kafka, RabbitMQ, Apache Spark, Apache Hive
Machine Learning:	Supervised and Unsupervised learning, Privacy-Enhancing Technologies
Automotive IT:	MQTT, AMQP, OPC-UA
Medical Informatics:	DICOM, HL7, SNOMED-CT, LOINC, ICD, HL7-CDA
Databases:	Relational Databases, MongoDB, Cassandra

Selected Publications

An Integrated Ontology-based Approach for Patent Classification in Medical Engineering 12th International Conference on Data Integration in Life Science ∘ November 2017 ∘ Luxembourg

Ontology Matching for Patent Classification The Twelfth International Workshop on Ontology Matching ∘ October 2017 ∘ Vienna, Austria

References

PD Dr. Christoph Quix RWTH Aachen ∘ Aachen ,Germany ∘ quix@dbis.rwth-aachen.de

Mr. Christoph Kreibich Audi AG ∘ Ingolstadt ,Germany ∘ christoph.kreibich@audi.de