

# Shahab Helmi

[shahab.helmi@ucdenver.edu](mailto:shahab.helmi@ucdenver.edu) | <http://www.shelmi.com> | <https://www.linkedin.com/in/shelmi>

## EDUCATION

**Doctor of Philosophy in Computer Science**, Since January 2015

University of Colorado Denver, CO

**Master of Science in Computer Science**, December 2014

University of Southern California, Los Angeles, CA

**Bachelor of Science in Software Engineering**, June 2012

Azad University of Tehran, Tehran, Iran

## EXPERIENCE

**Teaching**, Since January 2014

**University of Colorado Denver**, Since January 2015

- Co-instructor/Course Designer: Big Data Systems [CSCI 5800]
- Instructor: Database System Concepts [CSCI 3287]
- Instructor: Fundamentals of Computing (C++ Programming) [CSCI 1410]
- Instructor: Fundamentals of Computing Lab (C++ Programming) [CSCI 1411]
- Teaching Assistant: Database System Concept [CSCI 5559], Introduction to Data Science [CSCI7800], Introduction to Data Mining [CSCI 7702]

**University of Southern California**, January 2014 to May 2014

- Teaching Assistant: Database System Concept [CSCI 585]

**Synthea-OMOP ETL Tool**, June 2017 to September 2017

**University of Colorado Denver, Anschutz Medical Campus, Denver, CO**

- This ETL tool extracts data from [Synthea FHIR](#) outputs, transforms it to the [OMOP](#) common data model, and loads it to an OMOP-based database hosted on PostgreSQL server. This application also handles missing data, missing vocabularies, etc.
  - Roles: design, implementation, and testing the application using agile methodology
  - Technologies: Java, JSON, PostgreSQL, IntelliJ

**COPD**, Since January 2017

**University of Colorado Denver, Anschutz Medical/Downtown Campus, National Jewish Health, Denver, CO**

- As a part of this ongoing project, we are developing a suite of novel tools for efficient data integration, storage, querying, and analysis, all as part of an end-to-end data management and mining system that enables effective data-driven study of COPD. As an example, this tool is going to be used as an underlying layer for applying machine learning techniques on COPDGene data. For more info please visit the [BDLab](#) website.
  - Technologies: C#, WPF, Microsoft SQL Server, LINQ, Microsoft Visual Studio, R

**RadOn**, September 2016 to September 2017

**University of Colorado Denver, Anschutz Medical Campus, Denver, CO**

- The RadOn project focuses on developing data integration and management solutions to automate verification of radiation oncology treatments. In particular, such solutions are expected to compare

the prescription provided by a physician, treatment plans created by dosimetrists, and the details of actual treatment carried out by the treatment instruments, in order to identify inconsistencies in the treatment accurately. For more info please visit the [BDLab](#) website.

- Technologies: C#, WPF, Microsoft SQL Server, LINQ, Eclipse Scripting API

**CiSoft**, January 2014 to September 2017

**University of Southern California, Los Angeles, CA**

- Implemented an application using Microsoft StreamInsight to receive streaming sensor data from Chevron, detect errors, and reconstruct them.

**iWatch**, May 2013 to December 2013

**University of Southern California, Los Angeles, CA**

- This project was implemented for the Department of Public Safety at University of Southern California to detect and announce crimes based on Geospatial data.
  - Technologies: Oracle Spatial Database, Microsoft SQL Server, Java, JDBC, and Microsoft C#.NET

## SKILLS

- **Programming Languages:** C#, Java, Python, C++, WPF, R, Scala
- **Scripting Languages:** SQL [Microsoft SQL Server, PostgreSQL, MySQL, Oracle], JavaScript, HTML
- **Libraries:** Pandas, scikit-learn
- **Tools:** git, QGIS, Tableau
- **Big Data:** Apache Spark, Apache Kafka

## CERTIFICATES

- **Microsoft Professional Program for Data Science:**
  - Data Science Orientation | Query Relational Databases | Introduction to Python for Data Science
- **Udemy Certificates:**
  - Complete Python Bootcamp | Tableau 9 For Data Science | R Programming A-Z™

## PUBLICATIONS

- S. Helmi, F. Banaei-Kashani, "Spatiotemporal Range Pattern Queries on Large-scale Co-Movement Pattern Datasets", in IEEE International Conference on Big Data, 2017
- S. Helmi, F. Banaei-Kashani, "Efficient Processing of Spatiotemporal Pattern Queries on Historical Frequent Co-Movement Pattern Datasets", in proceedings of the 43rd VLDB International Conference on Mobility Analytics for Spatio-temporal and Social Data, 2017
- S. Helmi, F. Banaei-Kashani, "Mining Frequent Episodes from Multivariate Spatiotemporal Event Sequences", in proceedings of the 7th ACM SIGSPATIAL International Workshop on GeoStreaming. ACM, 2016.

## AWARDS

[Student Travel Award](#), 2017 IEEE International Conference on Big Data, Boston, MA

## COMMUNITY SERVICES

- **Peer Review:** PMC 2016 | TSAS 2016 | ACMGIS 2015
- **Webmaster:** [BSD 2016 and 2017](#), [BDSE Program at UC Denver](#), [BDLab at UC Denver](#)