Sareh Sadeghianasl

Curriculum Vitae

+61 7 3138 1395
s.sadeghianasl@qut.edu.au
ssdn70@gmail.com
http://www.sarehsadeghianasl.com
September, 1991
https://github.com/sarehsadeghianasl
My LinkedIn

EDUCATION

2018 – 2022 **Doctor of Philosophy**
Information Systems
Queensland University of Technology, Australia

2017 – 2014 **Master of Science**
SECOND TOP STUDENT (OUT OF APPROX. 100)
Computer Software Engineering
Ferdowsi University of Mashhad, Iran

2010 – 2014 **Bachelor of Science**
SECOND TOP STUDENT
Computer Software Engineering
Ferdowsi University of Mashhad, Iran

RESEARCH EXPERIENCE

CURRENT, FROM JUNE 2023 (FT)

Queensland University of Technology
Lecturer in Information Systems

In my role as a Lecturer, I coordinate various units, such as Business Process Modelling and Modern Data Management. Furthermore, I continue my research on process-data quality, and work on various industry projects, including data governance for an Australian Health and Safety organisation. I currently manage the open-source software framework PraeclarusPDQ at QUT which involves international collaboration (with UNIST in South Korea and University of Bayreuth in Germany).

2021-2023 (FT)

Queensland University of Technology
Research Fellow in Process Science

In my role as a Research Fellow, I recognised that data quality problems have to evolve with the emergence of the new Object-Centric log format which has given rise to a new collection of event log imperfection patterns. I also developed a repair algebra in order to be able to specify repairs at a high(er) level of abstraction. In addition, I worked on various industry projects, including data governance for an equipment supplier, and quality assessment and root cause identification of data quality problems in electronic medical records of patients presenting at the Emergency Department of a Queensland hospital.

2018 – 2021 (FT)

Queensland University of Technology
Researcher

My PhD research focused on improving the quality of process data, in particular activity labels, which capture the tasks performed in an organisation. I looked at labels with the same meaning but different syntax. I developed computational as well as human-based approaches to improve the quality of such activity labels. I was the first to apply crowdsourcing and gamification techniques for data cleaning in the field of process mining and I implemented a number of solutions as well.

2014 – 2017 (FT)

Ferdowsi University of Mashhad
Researcher

My research focused on semantic web and process-oriented data mining. More specifically, I was looking at semantic process similarity assessment using graph-based similarity measures, semantic process fragmentation, and developing ontologies from event log data. I implemented a tool for building an ontology from a log.

TEACHING EXPERIENCE

Queensland University of Technology, Brisbane, Australia

- *IAB203: Business Process Modelling (2024–unit coordinator and lecturer).*

  This unit introduces the principles of business process modelling. The main language that we focus on is BPMN and students learn different notations, such as gateways, loops, events, and exceptions.

- *IAB206: Modern Data Management (2023/4 unit coordinator and lecturer, 2021–super tutor, 2019/20–tutor).*

  This unit introduces different means to address challenges related to managing Big Data, i.e. volume, variety, velocity, and veracity, which is being relied on in today's decision making in organisations. It covers different types of NoSQL databases, e.g. column family, document, key-value, and in-memory databases along with some of their most famous tools, e.g. MongoDB, Hadoop, MapReduce, and Cassandra. I received a score of 4.5/5 in the Spot survey 2019, and an average agreement of 84.6% in the Student Voice survey 2021.

- *IAB321: Business Process Technologies (2023–lecturer, 2021–tutor).*

  This unit introduces the principles and techniques of business process automation and process mining. It covers different well-known business process technologies and tools such as Bonita, YAWL, ProM, Disco, and Celonis. In the Student Voice survey, I received an average agreement of 96.8% for the unit items measuring student learning and teaching quality.

- *IAB201: Modelling Techniques for Information Systems (2020/1–tutor).*

  This unit provides knowledge and skills related to conceptual modelling, designing artefacts, and analysing different aspects in regard to information systems. It covers most modern systems modelling languages such as ER, UML, ORM, and Petri net.

Jinling Institute of Technology, China

- *IAB206: Modern Data Management (2024–unit coordinator, 2021–super tutor, 2020–tutor)*

Ferdowsi University of Mashhad, Iran

- *Database Management (2015–tutor)*

INDUSTRY PROJECTS

Process Mining and Data Governance Strategy for RSHQ
2023-Current

I have led the industry project on process mining and data governance in RSHQ. A number of different types of data and process analyses and a new concept of learning mine analysis was introduced. Furthermore, we conducted interview with 14 staff at RSHQ to identify data governance strengths and road blocks.

Data Governance Strategy for Hastings Deering
2022-2023

I have led the data quality assessment research activities in this project with Hastings Deering and funded by them through QUT Centre for Data Science. I analysed data sets from this organisation with more than 100,000 records and 200 attributes. I developed a data quality framework, used statistical techniques to measure the quality of the data sets, was involved in the root-cause analysis and prevention and mitigation strategies, and presented the results of my research to the senior decision-makers within the organisation. Additional funding is under discussion to continue this work.

Data Quality Analysis for an Australian Hospital
2021-Current

This project focuses on quality assessment and root cause identification of data quality problems in electronic medical records of patients presenting at the Emergency Department of a Queensland hospital. The project is funded through QUT Centre for Data Science. My tasks include (i) pre-processing of more than 130,000 patient records, (ii) analysing the quality of data, and more specifically, timestamps, (iii) data-driven compliance analysis of patients' Length of Stay (LOS), and Seen In Time (SIT), and (iv) bottleneck analysis for patients with long LOS and SIT. I presented the results of the quality analysis to the senior stakeholders at the ED and am a co-author of four research publications as the outcomes of this project.

AWARDS & RECOGNITIONS

2023 **Runner-Up of the Best Process Mining PhD Thesis**
Princess Alexandra Hospital Health Symposium, Brisbane, Australia

2023 **Early Career Researcher Best Presentation Award**
International Conference in Process Mining, Rome, Italy

2020 **Associate Fellow of Higher Education Academy**
Advanced HE, UK

2020 **Best Student Reviewer Award**
School of Information Systems, QUT

2020 **Medical Datathon Finalist**
Queensland AI Hub

2017 **HDR Tuition Fee Sponsorship**
Queensland University of Technology

2016 **Top Master Student Award**
Ferdowsi University of Mashhad, Iran

2013 **Top Bachelor Student Award**
Ferdowsi University of Mashhad, Iran

2013 **Master of Science National Entry Exam Waiver**
Ferdowsi University of Mashhad, Iran

2009 **Top 1% Candidate in National University Entry Exam**
Iran Ministry of Science, Research, and Technology

MEMBERSHIPS

2020 **Data for Discovery Theme**
QUT Centre for Data Science

2020 **Women in Technology**

2020 **Business Process Management Association**

2019 **IEEE Task Force in Process Mining**

SELECTED PRESENTATIONS

Process Mining Day Webinar Series
2022

I delivered a talk on the topic of Data Quality in Process Mining at the 6th Process Mining Day webinar hosted by Behfalab, the largest process mining company in Iran. I talked about common types of data quality issues in event logs that hinder organisations from deriving meaningful insights from their data. I also discussed the current state of the art on detection and repair approaches for event log quality issues, including my own gamification solutions to the problem. This presentation was well-received by several process mining academics and practitioners.

Data Science Workshop for Researchers & Industry
2020

I was an Academic Speaker in this workshop, which was co-hosted by ACEMS, the Australian Data Science Network, the QUT Centre for Data Science and Integrity Systems Company (ISC). The goal of this workshop was to explore the latest research in data quality, management, and integration, as well as challenges and opportunities in industry. My talk outlined my research on data quality improvement through gamification and crowdsourcing techniques, which received positive feedback from both industry and academic guests.

ICPM Online Conference – Research Track 3: Data Quality and Preparation
2020

I presented our paper, "Collaborative and Interactive Detection and Repair of Activity Labels in Process Event Logs", at the online ICPM conference event. My talk was well-received by more than 130 live attendees, followed by an insightful discussion involving leading and internationally recognised researchers in data and process mining afterwards. The recorded video of the session is available at: https://www.youtube.com/watch?v=_Q1exV-o53E

CoopIS Conference, Rhodes, Greece – Research Track: Machine Learning and Knowledge Discovery III
2019

I have presented our paper "A Contextual Approach to Detecting Synonymous and Polluted Activity Labels in Process Event Logs" in the CoopIS conference in Greece. This talk was well-appreciated followed by an insightful discussion involving internationally recognised researchers in data and process mining afterwards.

SOFTWARE DEVELOPMENT

- *PraeclarusPDQ*

  An open-source framework for process data quality management. This framework will be the central focus of software development for our research in data quality. This software is written as a Java Spring Boot application and is available on GitHub.

- *The Quality Guardian*

  A web-based gamified system for detecting and repairing imperfect activity labels in even logs. This software was the first tool for gamified data cleaning in process mining. The software is written as a Java web application hosted on a Tomcat server and is available on GitHub.

- *The Quality Guardian Redux*

  A card matching game where each card contains an imperfect activity label in event logs. This tool was a continuation to the Quality Guardian game desiged specifically for domain experts. The software is written as a Java web application hosted on a Tomcat server and is available on GitHub.

- *The Quality Guardian Rosebud*

  A web-based gamified system for creating an ontology of activity labels in event logs. This was the first tool for gamified ontology creation in process mining. The software is written as a Java web application hosted on a Tomcat server and is available on GitHub.

- *Synonymous Label Repair*

  A ProM plugin for detecting and repairing imperfect activity labels in even logs. It uses data mining and statistical techniques, e.g. probability distribution functions, or clustering algorithms to detect synonym labels. The plugin is written in Java and is available on the ProM Server.

RESEARCH INTERESTS

Process mining, Data quality, Gamification, Visualisation, User experience, Data mining, and Data bases.

PUBLICATIONS

I have 76 citations in Google Scholar and an h-index of 5.

International Journal Publications

1. Sareh Sadeghianasl, Arthur H.M. ter Hofstede, Moe T. Wynn, and Selen Turkay. *Humans-in-the-loop: Gamifying activity label repair in process event logs*. Engineering Applications of Artificial Intelligence, vol. 132, pp. 107875, 2024 [SCImago D1]

   https://doi.org/10.1016/j.engappai.2024.107875

2. Arthur H.M. ter Hofstede, Agnes Koschmider, Andrea Marrella, Robert Andrews, Dominik A. Fischer, Sareh Sadeghianasl, Moe T. Wynn, Marco Comuzzi, Jochen De Weerdt, Kanika Goel, Niels Martin, Pnina Soffer. *Process-data quality: the true frontier of process mining*, vol. 15, No. 3, pp. 1-21, 2023 [SCImago Q2]

   https://dl.acm.org/doi/full/10.1145/3613247

3. Rehan Syed, Rebekah Eden, Tendai Makasi, Ignatius Chukwudi, Azumah Mamudu, Mostafa Kamalpour, Dakshi Kapugama Geeganage, Sareh Sadeghianasl, Sander JJ Leemans, Kanika Goel, Robert Andrews, Moe T. Wynn, Arthur H.M. ter Hofstede, and Trina Myers. *Digital Health Data Quality Issues: Systematic Review*. Journal of Medical Internet Research (JMIR), vol. 25, pp. e42615, 2023 [SCImago D1]

   https://www.jmir.org/2023/1/e42615/

4. Sareh Sadeghianasl, Arthur H.M. ter Hofstede, Moe T. Wynn, Selen Turkay, and Trina Myers. *Process activity ontology learning from event logs through gamification*. IEEE Access, vol. 9, pp. 165865-165880, 2021. [SCImago Q1]

   https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9648191

International Conference Papers

5. Kanika Goel, Sareh Sadeghianasl, Robert Andrews, Arthur Ter Hofstede, Moe Thandar Wynn, Dakshi Kapugama Geeganage, Sander JJ Leemans, James McGree, Rebekah Eden, Andrew Staib, Rob Eley, and Raelene Donovan. *Digital Health Data Imperfection Patterns and Their Manifestations in an Australian Digital Hospital*. Hawaii International Conference on System Sciences, pp. 3235-3244, 2022. [HICSS; acceptance rate 47%]

   https://scholarspace.manoa.hawaii.edu/items/5a8bbd79-a0d0-4620-8325-ca539e641220

6. Sareh Sadeghianasl, Arthur H.M. ter Hofstede, Suriadi Suriadi, and Selen Turkay. *Collaborative and Interactive Detection and Repair of Activity Labels in Process Event Logs*. International Conference in Process Mining, IEEE, 2020. [ICPM; acceptance rate 32%]

   https://ieeexplore.ieee.org/abstract/document/9230305/

7. Sareh Sadeghianasl, Arthur H.M. ter Hofstede, Moe T. Wynn, and Suriadi Suriadi. *A Contextual Approach to Detecting Synonymous and Polluted Activity Labels in Process Event Logs*. International Conference on Cooperative Information Systems, LNCS, vol. 11877, pp. 76-94, Springer, 2019. [CoopIS; ranked as an A conference by CORE2020, acceptance rate 22% for full papers]

   https://link.springer.com/chapter/10.1007/978-3-030-33246-4_5

Papers in National or Regional Conferences

8. Sareh Sadeghianasl and Mohsen Kahani. *Process Similarity Assessment Using Ontology Mapping (in Persian)*. Computer Society of Iran Computer Conference (CSICC), Tehran, Iran, 2016.

   https://csi.org.ir/en/paper/view/id/1932

Theses

9. Sareh Sadeghianasl. *The Quality Guardian: Improving Activity Label Quality in Event Logs through Gamification*. PhD thesis, Queensland University of Technology, Brisbane, Australia, 2022.

10. Sareh Sadeghianasl. *The Quality Guardian: Improving Activity Label Quality in Event Logs Through Gamification (Extended Abstract)*. Proceedings of the Best Dissertation Award, Doctoral Consortium, and Demonstration & Resources Track at the BPM Conference, CEUR Workshop Proceedings, vol. 3216, pp. 1–5, Münster, Germany, 2022.

    https://ceur-ws.org/Vol-3216/paper_122.pdf

11. Sareh Sadeghianasl. *Process Similarity Assessment Using Ontology Mapping (in Persian)*. Master's thesis, Ferdowsi University of Mashhad, Mashhad, Iran, 2017.

Submitted Papers

12. Sareh Sadeghianasl, Arthur H.M. ter Hofstede, Moe T. Wynn, Robert Andrews, and Wil van der Aalst. *Object-Centric Data Imperfection Patterns*. Submitted.

13. Sareh Sadeghianasl, Dominik A. Fischer, Arthur H.M. ter Hofstede, Michael Adams, Robert Andrews, Marco Comuzzi, Jonghyeon Ko, Agnes Koschmider, Moe T. Wynn, and Tobias Ziolkowski. *PraeclarusPDQ: A Reference Architecture for Process Data Quality Management*. Submitted to *Big Data Research*.

14. Rehan Syed, Rebekah Eden, Ignatius Chukwudi, Sareh Sadeghianasl, Sander JJ Leemans, Kanika Goel, Moe T. Wynn, and Robert Andrews. *Data quality and Influencing Factors on Decision Making*. Submitted.

15. Rehan Syed, Rebekah Eden, Ignatius Chukwudi, Sareh Sadeghianasl, Kanika Goel, Moe T. Wynn, Robert Andrews, and Dakshi Kapugama Geeganage. *Tergeo: A Framework for Data Quality Root Cause Analysis*. Submitted.

Papers Under Preparation

16. Sareh Sadeghianasl, Arthur H.M. ter Hofstede, Alistair Barros, Robert Andrews, and Remco Dijkman. *Repair Algebra for Event Log Imperfection Patterns*. Under preparation.

ACADEMIC SERVICE

- *Deputy Lead* of International & Engagement portfolio, School of Information Systems, QUT, 2024.

- *PC Member* of International Conference on Advanced Information Systems Engineering (CAiSE), 2024.

- *Co-Chair* of Data Quality and Transformation in Process Mining Workshop in ICPM 2022/3.

- *Associate Editor* of the BPM track at the 18th International Conference on Wirtschaftsinformatik (WI), 2023.

- *Reviewer* of ICPM 2021 and 2022, CoopIS 2022, ECIS 2022 and 2023, and CAiSE 2023 (all conferences) and the IEEE Access journal (2021).