MICHAEL HAHSLER

Associate Professor
Computer Science
Lyle School of Engineering
Southern Methodist University
Dallas, TX 75205, USA

mhahsler@lyle.smu.edu
https://michael.hahsler.net

EDUCATION

2006: **Habilitation in Applied Computer Science** (business informatics), postdoctoral university degree with lecture qualification, Vienna University of Economics and Business (WU Wien), Austria.

2001: **Ph.D. in Social and Economic Sciences** with honors (major: information engineering and management; advisors: Wolfgang Janko, Andreas Geyer-Schulz), WU Wien, Austria. Thesis: *"Software Development Process Using Analysis Patterns with Examples for Information Management"* (in German). WU Wien is ranked 28 in the 2011 Financial Times European Business School Ranking, and is accredited by AACSB, AMBA, and EQUIS.

1998: **M.S. in Business Administration** (majors: information systems, applied computer science), WU Wien, Austria. Thesis: *"Software Patterns for Pinboards"* (in German).

1992: **Associate degree in Electrical Engineering** (communications engineering) with honors (first in class), College of Technology - HTBLA Wien I, Vienna, Austria.

APPOINTMENTS

2020–present: **Associate Professor of Computer Science (non-tenure track)**, Southern Methodist University (SMU), Dallas, TX, USA.

2015–2021: **Adjunct Assistant Professor of Clinical Sciences,** UT Southwestern Medical Center, Dallas, TX, USA.

2019–2020: **Data Science Researcher in Residence,** Office of Information Technology, SMU, Dallas, TX, USA.

2012–2019: **Assistant Professor of Engineering Management, Information, and Systems**, and (by courtesy) **Computer Science and Engineering**. Director of the Intelligent Data Analysis Lab (IDA@SMU), Lyle School of Engineering, SMU, Dallas, TX, USA.

2009-2012: **Visiting Assistant Professor of Computer Science and Engineering**, and Co-Director of IDA@SMU, Lyle School of Engineering, SMU, Dallas, TX, USA.

2001–2007: **Assistant Professor of Information Engineering and Management** (Universitätsassistent 2001–2006, Privatdozent 2006–2007), Department of Information Systems and Operations, and Core Researcher, Research Institute for Computational Methods, Vienna University of Economics and Business (WU Wien), Austria.

2002–2003: **Adjunct Professor of Computer Science**, Webster University, Vienna Campus, Austria.

1998–2001: **Lecturer of Applied Computer Science** (Universitätsassistent), Department of Applied Computer Science, WU Wien, Austria.

RESEARCH INTERESTS

My current research interests are focused on methods used in the interdisciplinary field of **Data Science** including:

Artificial Intelligence/Machine Learning/Data Mining: Data stream mining, recommender systems, reinforcement learning, association rule mining, data visualization.

Combinatorial Optimization: Clustering, sequential decision-making, seriation and optimal ordering problems.

Application areas: bioinformatics, healthcare analytics, quantitative marketing, earth sciences, and engineering problems.

AWARDS AND HONOR SOCIETIES

2021-2022 HOPE Professor of the Year Award nominee, Honoring Our Professors' Excellence, Residence Life and Student Housing, SMU, 2021.

Upsilon Pi Epsilon Honor Society, Elected Member, SMU Chapter, UPE Honor Society for the Computing and Information Disciplines, 2021.

2012 HOPE Award honoree, Honoring Our Professors' Excellence, Residence Life and Student Housing, SMU, 2012.

Graduate Student Council Outstanding Faculty Award, Computer Science and Engineering, Bobby B. Lyle School of Engineering, SMU, 2011.

WU Top Publication Award for "Data Mining and Marketing: Exploratory Market Basket Analysis" (in German: "Data Mining und Marketing am Beispiel der explorativen Warenkorbanalyse") in *Marketing ZFP - Journal of Research and Management*, WU Wien, 2007.

Finalist of the Global Bangemann Award (now Stockholm Challenge) with the Virtual University Project, Stockholm, Sweden, 1999.

Winner of the WU Innovation Award, WU Wien, 1997.

ADDITIONAL EXPERIENCE

Lead developer of several very popular R-extension packages for data mining. R is the leading free software environment for statistical computing, graphics, and data mining. The packages were downloaded more than 100,000 times in January 2025.

E-Business Adviser, Hall Financial Group, Frisco, TX, USA, 2007–2008.

Head of engineering, ePub-WU project. Development of an open-access digital library for working papers and Ph.D. theses, WU Wien, 2001–2003.

Designer, Assistant Project Manager and later **Project Manager,** Virtual University Project, WU Wien, 1997–2004.

PROFESSIONAL MEMBERSHIPS

ACM, GfKl Data Science Society, IEEE Computer Society, R Foundation.

Research

My research interests lie in the intersection of machine learning, statistical methods, and combinatorial optimization with applications in artificial intelligence, data mining, and data science. To support reproducible research, my students and I create and maintain well-documented open-source software as part of my research output. The developed software consists of several R-extension packages[1] and is published via the Comprehensive R Archive Network (http://CRAN.R-project.org) or as part of the Bioconductor project for bioinformatics (http://www.bioconductor.org/). Development versions and prototypes are maintained on GitHub (http://github.com/mhahsler) and r-universe (https://mhahsler.r-universe.dev/).

The software is very popular to support data science and analytics in research, educational and industrial settings. The most popular package (arules) was installed more than 15,000 times in January 2017. All packages together exceed 50,000 installs per month.

PUBLISHED RESEARCH SOFTWARE PACKAGES

AI / Machine Learning / Data Mining

arules: Infrastructure for analyzing transaction data with association rules.

arulesViz: A package for visualizing association rules based on package arules.

arulesSequences: Add-on package to handle and mine frequent sequences.

arulesCBA: Add-on package for classification based on association rules (CBA).

dbscan: A fast reimplementation of several density-based algorithms of the DBSCAN family for spatial data clustering.

pomdp, pomdpSolve: Define, solve and analyze partially observable Markov decision processes.

recommenderlab: Development and test environment for recommender algorithms.

rEMM: Temporal modeling for massive data streams using data stream clustering and Markov models.

stream, streamMOA, streamConnect: Infrastructure and algorithms for data stream mining.

Combinatorial Optimization

qap: Provides heuristics for the Quadratic Assignment Problem (QAP).

seriation: Seriation/sequencing techniques to reorder matrices, dissimilarity matrices, and dendrograms.

TSP: Infrastructure and algorithms for the traveling salesperson problem.

Bioinformatics

QuasiAlign: Efficient alignment-free methods for approximate sequence alignment.

rBLAST: Interfaces the Basic Local Alignment Search Tool (BLAST) to search genetic sequence databases from within the Bioconductor infrastructure.

rRDP: Interface to the Ribosomal Database Project (RDP) naive Bayes classifier for 16S rRNA.

rMSA: Interface for popular multiple sequence alignment (MSA) tools like ClustalW, MAFFT, MUSCLE, and Kalign.

1 R is the leading free software environment for statistical computing, graphics, and data mining.

RESEARCH FUNDING

2021-2022, Data science supplement for grant Evaluation of Liquefaction Potential of Saturated Granular Soils under Partial Drainage Conditions (NSF-CMMI 1728612), Senior Personnel with Usama El Shamy (PI or existing grant), **NSF**, $45,773.

2017–2021, SAFE-NET: An Integrated Connected Vehicle and Computing Platform for Public Safety Applications, Co-PI with Khaled Abdelghany (PI) and Barbara Minsker (Co-PI), **NIST** Public Safety Innovation Accelerator Program, $484,802.

2018, Deep Convolutional Networks for Forgery Classification and Anomaly Detection, Co-PI with Eric Larsen (PI), Paul Krueger (Co-PI) and Eli Olinick (Co-PI), Interdisciplinary Seed Funding, Lyle School of Engineering, **SMU**, $21,120.

2015, Robust and Efficient Flow Field Comparison Based on Point Properties, Co-PI with Paul Krueger (PI) and Eli Olinick (Co-PI), Interdisciplinary Seed Funding, Lyle School of Engineering, **SMU**, $22,230.

2013, Data-Enabled Characterization of the Seismic Response of Geotechnical Systems, Co-PI with Usama El Shamy (PI), Interdisciplinary Seed Funding, Lyle School of Engineering, **SMU**, $6,300.

2011–2014, Position Sensitive P-Mer Frequency Clustering with Applications to Classification, Co-PI with Margaret Dunham (PI) and Monnie McGee, NIH R21HG005912, National Human Genome Research Institute, **NIH,** $385,000.

2011–2012, Mobile Communication Innovation Lab at SMU, Co-PI with Mark Fontenot (PI), **Samsung**, $25,000 (equipment).

2009–2013, REU Supplement for NSF-IIS 0948893, Co-PI with Margaret Dunham (PI), **NSF,** $32,000.

2009–2013, III/EAGER: Temporal Relationships Among Clusters in Data Streams (TRACDS), Co-PI with Margaret Dunham (PI), NSF-IIS 0948893, **NSF**, Division of Information & Intelligent Systems. $180,000.

2009, An Experimentation Environment for Generating Top-N Recommendations from Binary Data, PI, **NSF I/UCRC**: Net-Centric Software & Systems Consortium, $60,000.

2005–2008, Infrastructure for interdisciplinary research focusing on machine learning and simulation, Co-PI with Kurt Hornik (PI), **Austrian Federal Ministry of Science and Education**, €179,000 ($230,000).

2001–2009, Digital Library – WU online publications, PI, University Library of the **Vienna University of Economics and Business**. €31,000 ($40,000).

2001, Supplementary funds for the virtual university project, PI, **Vienna Chamber of Commerce**, Austria, €11,000 ($14,000).

PATENTS

1. Jake Drew, Michael Hahsler and Tyler Moore: System and method for machine learning and classifying data, US20140344195A1, Application granted on 03/03/2020.

PUBLICATIONS

Underlined names indicate graduate students of Dr. Hahsler.

Names in bold indicate undergraduate students of Dr. Hahsler.

Manuscripts under review, under revision and in preparation

1. **Jerry Ma,** Michael Hahsler and Peter Moore, A Recommender Systems Architecture for University Curriculum Advising, 2025 AAAI Spring Symposium, Accepted.
2. Antônio Junior Alves Caiado and Michael Hahsler. AI content self-detection for transformer-based large language models. In: Intelligent Systems Conference (IntelliSys), 28 & 29 August, 2025. Preprint: arXiv:2312.17289 [cs.CL] (December 2023). Accepted.
3. Olga Bountali, Sila Cetinkaya, Michael Hahsler, Farnaz Nourbakhsh, Zhenghang Xu, Henry Quinones, Leveraging Advanced Analytics to Streamline the Emergent Dialysis Process at Parkland Hospital, under revision, *Healthcare Analytics*, 2024. Preprint: SSRN, DOI: 10.2139/ssrn.5066305
4. Farbod Tavakkoli Khomeini, Janille Smith-Colin, Michael Hahsler, Xinlei Wang, Barbara Minsker, Infrastructure Deficiency Correlation with Pedestrian/Cyclist Crashes: A Data-Driven Approach, *Cities*, Submitted, 2024.
5. Zahra Gharibi, Michael Hahsler, Mehmet Ayvaci. Optimal Quality Oversight in Kidney Transplantation and Its Impact on Transplant Centers' Waitlist Management, resubmission in preparation, 2024.
6. Michael Hahsler, Kurt Hornik, Confidence Intervals Estimation for Association Analysis, *Intelligent Data Analysis*, in preparation, 2024.

Papers in peer-reviewed journals (index in parentheses after journal name)

1. Michael Hahsler and Anthony R. Cassandra. Pomdp: A computational infrastructure for partially observable Markov decision processes. *R Journal* (ISI, SCI expanded), 16(2):1-18, 2025.
2. Michael Hahsler. An R companion for Introduction to Data Mining. *Journal of Open Source Education* (PubMed), 7(82):223, 2024.
3. M. Ledmi, C. Kara-Mohamed, M.E.H. Souidi, M. Hahsler, and A. Ledmi. Mining association rules for classification using frequent generator itemsets in arules package. *International Journal of Data Mining, Modelling and Management* (Scopus), 15(2):302–221, 2023.
4. Farzad Kamalzadeh, Vishal Ahuja, Michael Hahsler, Michael Bowen. An Analytics-Driven Approach For Optimal Individualized Diabetes Screening, *Production and Operations Management* (Scopus), 30(9):3161–3191, September 2021.
5. Zahra Gharibi and Michael Hahsler. A simulation-based optimization model to study the impact of multiple-region listing and information sharing on kidney transplant outcomes. *International Journal of Environmental Research and Public Health* (Scopus, SSCI), 18(3):873, January 2021.
6. Xinyi Ding, Zohreh Raziei, Eric C. Larson, Eli V. Olinick, Paul Krueger, and Michael Hahsler. Swapped face detection using deep learning and subjective assessment. *EURASIP Journal on Information Security* (Scopus), 2020(6):1–12, 2020.
7. Michael Hahsler, **Ian Johnson,** Tomas Kliegr, Jaroslav Kuchar. Associative Classification in R: arc, arulesCBA, and rCBA, The *R Journal* (ISI, SCI expanded), 11(2):254–267, 2019.
8. Michael Hahsler, Matthew Piekenbrock and Derek Doran. dbscan: Fast Density-based Clustering Algorithms in R, *Journal of Statistical Software* (SCI expanded), 91(1):1–30, 2019.

9. Paul S. Krueger, Michael Hahsler, Eli V. Olinick, Sheila H. Williams, Mohammadreza Zharfa. Quantitative Classification of Vortical Flows Based on Topological Features using Graph Matching, *Proceedings of the Royal Society A* (Scopus), 475(2228):1–16, August 2019.
10. Joan B B Soriano, Michael Hahsler, Cecilia Soriano, Cristina Martinez, Juan P de Torres, Jose M Marin, Pilar de Lucas, Borja G Cosio, Antonia Fuster, and Ciro Casanova. Temporal transitions in COPD severity stages within the GOLD 2017 classification system. *Respiratory Medicine* (SCI), 142:81–85 September 2018.
11. Michael Hahsler and Anurag Nagar. Discovering patterns in gene ontology using association rule mining. *Biostatistics and Biometrics Open Access Journal*, 6(3):1–3, April 2018.
12. Michael Hahsler. arulesViz: Visualizing Association Rules with R, *The R Journal* (ISI, SCI expanded), 9(2):163–175, 2017.
13. Michael Hahsler, **Matthew Bolaños** and **John Forrest.** Introduction to stream: An Extensible Framework for Data Stream Clustering Research with R. *Journal of Statistical Software* (SCI expanded), 76(14):1–52, 2017.
14. Jake Drew, Michael Hahsler, and Tyler Moore. Polymorphic malware detection using sequence classification methods, *EURASIP Journal on Information Security* (Scopus), 2:1–12, 2017.
15. Michael Hahsler. An Experimental Comparison of Seriation Methods For One-Mode Two-Way Data, *European Journal of Operational Research* (SCI Expanded), 257:133–143, 2017.
16. Zahra Gharibi, Mehmet Ayvaci, Michael Hahsler, Tracy Giacoma, Robert S. Gaston and Bekir Tanriover. Cost-Effectiveness of Antibody-Based Induction Therapy in Deceased Donor Kidney Transplantation in the United States. *Transplantation* (SCI), 101(6):1234–1241, 2017.
17. Shaiba Hadil and Michael Hahsler. A Comparison of Machine Learning Methods for Predicting Tropical Cyclone Rapid Intensification Events, *Research Journal of Applied Sciences, Engineering and Technology* (SJR), 13(8):638–651, 2016.
18. Michael Hahsler and Radoslaw Karpienko. Visualizing Association Rules in Hierarchical Groups. *Journal of Business Economics* (EconLit, Scopus), pages 1–19, May 2016.
19. Michael Hahsler and **Matthew Bolaños.** Clustering Data Streams Based on Shared Density Between Micro-Clusters. *IEEE Transactions on Knowledge and Data Engineering* (SCI), 28(6):1449–1461, June 2016.
20. Anurag Nagar and Michael Hahsler. Fast discovery and visualization of conserved regions in DNA sequences using quasi-alignment. *BMC Bioinformatics* (ISI, SCI expanded), 14(Suppl. 11), 2013.
21. Michael Hahsler, Sudheer Chelluboina, Kurt Hornik, and Christian Buchta. The arules R-package ecosystem: Analyzing interesting patterns from large transaction datasets. *Journal of Machine Learning Research* (SCI, Scopus), 12:1977–1981, 2011.
22. Michael Hahsler and Kurt Hornik. Dissimilarity Plots: A Visual Exploration Tool for Partitional Clustering. *Journal of Computational and Graphical Statistics* (ISI), 20(2):335–354, 2011.
23. Rao M. Kotamarti, Michael Hahsler, Douglas Raiford, Monnie McGee, and Margaret H. Dunham. Analyzing Taxonomic Classification Using Extensible Markov Models. *Bioinformatics* (SCI), 26(18):2235–2241, 2010.
24. Michael Hahsler and Margaret H. Dunham. rEMM: Extensible Markov Model for data stream clustering in R. *Journal of Statistical Software* (SCI expanded), 35(5):1–31, 2010.
25. Michael Hahsler, Christian Buchta, and Kurt Hornik. Selective association rule generation. *Computational Statistics* (ISI, SCI expanded) , 12(2):303–315, April 2008.
26. Michael Hahsler, Kurt Hornik, and Christian Buchta. Getting things in order: An introduction to the R package seriation. *Journal of Statistical Software* (SCI expanded), 25(3):1–34, March 2008.

27. Michael Hahsler and Kurt Hornik. TSP - Infrastructure for the traveling salesperson problem. *Journal of Statistical Software* (SCI expanded), 23(2):1–21, December 2007.
28. Michael Hahsler and Kurt Hornik. New probabilistic interest measures for association rules. *Intelligent Data Analysis* (SCI expanded), 11(5):437–455, 2007.
29. Thomas Reutterer, Michael Hahsler, and Kurt Hornik. Data Mining und Marketing am Beispiel der explorativen Warenkorbanalyse (in German). *Marketing ZFP - Journal of Research and Management*, 29(3):165–181, 2007.
30. Michael Hahsler. A model-based frequency constraint for mining associations from transaction data. *Data Mining and Knowledge Discovery* (SCI, Scopus), 13(2):137–166, September 2006.
31. Christoph Breidert, Michael Hahsler, and Thomas Reutterer. A review of methods for measuring willingness-to-pay. *Innovative Marketing* (EconLit), 2(4):8–32, 2006.
32. Michael Hahsler, Bettina Grün, and Kurt Hornik. arules - A computational environment for mining association rules and frequent item sets. *Journal of Statistical Software* (SCI expanded), 14(15):1–25, October 2005.
33. Michael Hahsler. Integrating digital document acquisition into a university library: A case study of social and organizational challenges. *Journal of Digital Information Management* (ISI), 1(4):162–171, December 2003.
34. Wolfgang Gaul, Andreas Geyer-Schulz, Michael Hahsler, and Lars Schmidt-Thieme. eMarketing mittels Recommendersystemen (in German). *Marketing ZFP - Journal of Research and Management*, 24:47–55, 2002.
35. Andreas Geyer-Schulz, Michael Hahsler, and Maximillian Jahn. Educational and scientific recommender systems: Designing the information channels of the virtual university. *International Journal of Engineering Education* (SCI), 17(2):153–163, 2001.
36. Andreas Geyer-Schulz, Michael Hahsler, and Georg Schneider. The virtual university and its embedded agents. *ÖGAI Journal* (Journal of the Austrian Society for Artificial Intelligence), 18(1):14–19, 1999.
37. Peter Bruhn, Andreas Geyer-Schulz, Michael Hahsler, and Markus Mottel. Genetic machine learning and intelligent Internet agents. *ÖGAI Journal* (Journal of the Austrian Society for Artificial Intelligence), 17(1):18–25, 1998.

Journal articles (not peer-reviewed)

1. Margaret H. Dunham, Michael Hahsler, and Myra Spiliopoulou. Novel data stream pattern mining, Report on the StreamKDD'10 workshop. *SIGKDD Explorations*, 12(2):54–55, 2010.

Book chapters

1. Michael Hahsler, Kurt Hornik, and Thomas Reutterer. Warenkorbanalyse mit Hilfe der Statistik-Software R. In Peter Schnedlitz, Renate Buber, Thomas Reutterer, Arnold Schuh, and Christoph Teller, editors, *Innovationen in Marketing*, pages 144–163. Linde-Verlag, 2006.
2. Michael Hahsler. A quantitative study of the adoption of design patterns by open source software developers. In S. Koch, editor, *Free/Open Source Software Development*, pages 103–123. Idea Group Publishing, 2005.
3. Andreas Geyer-Schulz, Michael Hahsler, Andreas Neumann, and Anke Thede. Behavior-based recommender systems as value-added services for scientific libraries. In Hamparsum Bozdogan, editor, *Statistical Data Mining & Knowledge Discovery*, pages 433–454. Chapman & Hall / CRC, July 2003.
4. Andreas Geyer-Schulz and Michael Hahsler. Comparing two recommender algorithms with the help of recommendations by peers. In O.R. Zaiane, J. Srivastava, M. Spiliopoulou, and B. Masand,

editors, *WEBKDD 2002 - Mining Web Data for Discovering Usage Patterns and Profiles 4th International Workshop, Edmonton, Canada, July 2002, Revised Papers*, Lecture Notes in Computer Science LNAI 2703, pages 137–158. Springer-Verlag, 2003.

5. Andreas Geyer-Schulz, Michael Hahsler, and Maximillian Jahn. A customer purchase incidence model applied to recommender systems. In R. Kohavi, B.M. Masand, M. Spiliopoulou, and J. Srivastava, editors, *WEBKDD 2001 - Mining Log Data Across All Customer Touch Points, Third International Workshop, San Francisco, CA, USA, August 26, 2001, Revised Papers*, Lecture Notes in Computer Science LNAI 2356, pages 25–47. Springer-Verlag, July 2002.
6. Andreas Geyer-Schulz, Michael Hahsler, and Maximillian Jahn. myvu: A next generation recommender system based on observed consumer behavior and interactive evolutionary algorithms. In Wolfgang Gaul, Otto Opitz, and Martin Schader, editors, *Data Analysis: Scientific Modeling and Practical Applications*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 447–457. Springer-Verlag, Heidelberg, Germany, 2000.

Edited books/proceedings

1. Margaret H. Dunham, Michael Hahsler, and Myra Spiliopoulou, editors. *Proceedings of the First International Workshop on Novel Data Stream Pattern Mining Techniques (StreamKDD'10)*. ACM Press, New York, NY, USA, 2010.

Books

1. Michael Hahsler. An R Companion for Introduction to Data Mining. Online textbook, Figshare, 2024.

Papers in refereed conference proceedings

1. Usama El Shamy and Michael Hahsler. Data analytics applied to a microscale simulation model of soil liquefaction. In *Geotechnical Earthquake Engineering and Soil Dynamics V*. ASCE, June 2018.
2. Michael Hahsler. Grouping association rules using lift. In C. Iyigun, R. Moghaddess, and A. Oztekin, editors, *11th INFORMS Workshop on Data Mining and Decision Analytics (DM-DA 2016)*, November 2016.
3. Jake Drew, Michael Hahsler, and Tyler Moore. Polymorphic malware detection using sequence classification methods. In *International Workshop on Bio-inspired Security, Trust, Assurance and Resilience (BioSTAR 2016)*, May 2016.
4. Becca Mokhtarpour, Jerrell T. Stracener, and Michael Hahsler. A data-analysis approach for improved decision-making in selecting the preferred SoS capability solution. In *2016 Conference on Systems Engineering Research*, March 2016.
5. Sudheer Chelluboina and Michael Hahsler. Trajectory segmentation using oblique envelopes. In *2015 IEEE International Conference on Information Reuse and Integration (IRI)*, pages 470–475. IEEE, August 2015.
6. Anurag Nagar, Michael Hahsler, and Hisham Al-Mubaid. Association rule mining of gene ontology annotation terms for SGD. In *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE, August 2015.
7. Jörg Lässig and Michael Hahsler. Cooperative data analysis in supply chains using selective information disclosure. In Brian Borchers, J. Paul Brooks, and Laura McLay, editors, *Operations Research and Computing: Algorithms and Software for Analytics, 14th INFORMS Computing Society Conference (ICS2015)*. INFORMS, January 2015.
8. Jake Drew and Michael Hahsler. Strand: Fast sequence comparison using MapReduce and locality sensitive hashing. In *Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Health Informatics (BCB 2014)*. ACM, September 2014.

9. Hadil Shaiba and Michael Hahsler. An experimental comparison of different classifiers for predicting tropical cyclone rapid intensification events. In *Proceedings of the International Conference on Machine Learning, Electrical and Mechanical Engineering (ICMLEME'2014*), Dubai, UAE, January 2014.

10. **Matthew Bolaños**, **John Forrest,** and Michael Hahsler. Clustering large datasets using data stream clustering techniques. In Myra Spiliopoulou, Lars Schmidt-Thieme, and Ruth Janning, editors, *Data Analysis, Machine Learning and Knowledge Discovery, Studies in Classification, Data Analysis, and Knowledge Organization*, pages 135–143. Springer-Verlag, 2014.

11. Hadil Shaiba and Michael Hahsler. Intensity prediction model for tropical cyclone rapid intensification events. In *Proceedings of the IADIS Applied Computing 2013 (AC 2013) Conference*, Fort Worth, TX, October 2013.

12. Anurag Nagar and Michael Hahsler. Genomic sequence fragment identification using quasi-alignment. In *Proceedings of the ACM BCB Conference 2013*, Washington D.C., September 2013.

13. Anurag Nagar and Michael Hahsler. A novel quasi-alignment-based method for discovering conserved regions in genetic sequences. In *Proceedings of the IEEE BIBM 2012 Workshop on Data-Mining of Next-Generation Sequencing*, October 2012.

14. Anurag Nagar and Michael Hahsler. Using text and data mining techniques to extract stock market sentiment from live news streams. In *2012 International Conference on Computer Technology and Science (ICCTS 2012)*, August 2012.

15. Charlie Isaksson, Margaret H. Dunham, and Michael Hahsler. SOStream: Self-organizing density-based clustering over data streams. In *International Conference on Machine Learning and Data Mining (MLDM'2012)*. Springer, July 2012.

16. Maya Eldayeh and Michael Hahsler. Biological pathway completion using network motifs and random walks on graphs. In *IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB 2012), pages 229–236*. IEEE, May 2012.

17. Maya Eldayeh and Michael Hahsler. Analyzing incomplete biological pathways using network motifs. In *27th Symposium On Applied Computing (SAC 2012), volume 2, pages 1355-1360*. ACM, 2012

18. **Vladimir Jovanovic**, Margaret H. Dunham, Michael Hahsler, and Yu Su. Evaluating hurricane intensity prediction techniques in real time. In *Third IEEE ICDM Workshop on Knowledge Discovery from Climate Data, Proceedings of the of the 2011 IEEE International Conference on Data Mining Workshops (ICDMW 2011)*. IEEE, 2011.

19. Michael Hahsler and Sudheer Chelluboina. Visualizing association rules in hierarchical groups. In *42nd Symposium on the Interface: Statistical, Machine Learning, and Visualization Algorithms (Interface 2011)*. The Interface Foundation of North America, 2011.

20. Michael Hahsler and Margaret H. Dunham. Temporal structure learning for clustering massive data streams in real-time. In *SIAM Conference on Data Mining (SDM11)*. SIAM, 2011.

21. Yu Su, Sudheer Chelluboina, Michael Hahsler, and Margaret H. Dunham. A new data mining model for hurricane intensity prediction. In *Second IEEE ICDM Workshop on Knowledge Discovery from Climate Data: Prediction, Extremes and Impacts, Proceedings of the of the 2010 IEEE International Conference on Data Mining Workshops (ICDMW 2010)*. IEEE, 2010.

22. Rao M Kotamarti, Michael Hahsler, Douglas W Raiford, and Margaret H Dunham. Sequence transformation to a complex signature form for consistent phylogenetic tree using extensible Markov model. In *Proceedings of the 2010 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (IEEE CIBCB 2010)*. IEEE, 2010.

23. Christoph Breidert and Michael Hahsler. Adaptive conjoint analysis for pricing music downloads. In R. Decker and H.-J. Lenz, editors, *Advances in Data Analysis, Proceedings of the 30th Annual Conference of the Gesellschaft für Klassifikation e.V., Freie Universität Berlin, March 8–10, 2006*,

Studies in Classification, Data Analysis, and Knowledge Organization, pages 409–416. Springer-Verlag, 2007.

24. Michael Hahsler and Kurt Hornik. Building on the arules infrastructure for analyzing transaction data with R. In R. Decker and H.-J. Lenz, editors, *Advances in Data Analysis, Proceedings of the 30th Annual Conference of the Gesellschaft für Klassifikation e.V., Freie Universität Berlin, March 8–10, 2006*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 449–456. Springer-Verlag, 2007.

25. Michael Hahsler, Kurt Hornik, and Thomas Reutterer. Implications of probabilistic data modeling for mining association rules. In M. Spiliopoulou, R. Kruse, C. Borgelt, A. Nürnberger, and W. Gaul, editors, *From Data and Information Analysis to Knowledge Engineering, Proceedings of the 29th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Magdeburg, March 9–11, 2005*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 598–605. Springer-Verlag, 2006.

26. Christoph Breidert, Michael Hahsler, and Lars Schmidt-Thieme. Reservation price estimation by adaptive conjoint analysis. In Claus Weihs and Wolfgang Gaul, editors, *Classification - the Ubiquitous Challenge, Proceedings of the 28th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Dortmund, March 9–11, 2004*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 577–584. Springer-Verlag, 2005.

27. Georg Fessler, Michael Hahsler, and Michaela Putz. ePubWU - Erfahrungen mit einer Volltext an der Wirtschaftsuniversität Wien. In Christian Enichlmayr, editor, *Bibliotheken - Fundament der Bildung, 28. Österreichischer Bibliothekartag 2004*, Schriftenreihe der OÖ. Landesbibliothek, pages 190–193, 2005.

28. Michael Hahsler. Optimizing websites for customer retention. In Bing Liu, Myra Spiliopoulou, Jaideep Srivastava, and Alex Tuzhilin, editors, *Proceedings of the 2005 International Workshop on Customer Relationship Management: Data Mining Meets Marketing, November 18–19, 2005, New York City, USA*, 2005.

29. Michael Hahsler and Stefan Koch. Discussion of a large-scale open source data collection methodology. In *38th Annual Hawaii International Conference on System Sciences (HICSS'05), January 3–6, 2005 Hilton Waikoloa Village, Big Island, Hawaii*. IEEE Computer Society Press, 2005.

30. Michael Hahsler and Stefan Koch. Cooperation and disruptive behaviour - learning from a multi-player internet gaming community. In Piet Kommers, Pedro Isaias, and Miguel Baptista Nunes, editors, *IADIS International Conference Web Based Communities 2004, Lisbon, Portugal, 24–26 March 2004*, pages 35–42. International Association for Development of the Information Society (IADIS), 2004.

31. Andreas Geyer-Schulz, Michael Hahsler, Andreas Neumann, and Anke Thede. An integration strategy for distributed recommender services in legacy library systems. In M. Schader, W. Gaul, and M. Vichi, editors, *Between Data Science and Applied Data Analysis, Proceedings of the 26th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Mannheim, July 22–24, 2002*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 412–420. Springer-Verlag, July 2003.

32. Andreas Geyer-Schulz, Michael Hahsler, and Anke Thede. Comparing association rules and repeat-buying based recommender systems in a B2B environment. In M. Schader, W. Gaul, and M. Vichi, editors, *Between Data Science and Applied Data Analysis, Proceedings of the 26th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Mannheim, July 22–24, 2002*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 421–429. Springer-Verlag, July 2003.

33. Edward Bernroider, Michael Hahsler, Stefan Koch, and Volker Stix. Data Envelopment Analysis zur Unterstützung der Auswahl und Einführung von ERP-Systemen. In Andreas Geyer-Schulz and Alfred Taudes, editors, *Informationswirtschaft: Ein Sektor mit Zukunft, Symposium 4.–5. September*

2003, Wien, Österreich, Lecture Notes in Informatics (LNI) P-33, pages 11–26. Gesellschaft für Informatik, 2003.

34. Andreas Geyer-Schulz, Michael Hahsler, Andreas Neumann, and Anke Thede. Recommenderdienste für wissenschaftliche Bibliotheken und Bibliotheksverbünde. In Andreas Geyer-Schulz and Alfred Taudes, editors, *Informationswirtschaft: Ein Sektor mit Zukunft, Symposium 4.–5. September 2003, Wien, Österreich*, Lecture Notes in Informatics (LNI) P-33, pages 43–58. Gesellschaft für Informatik, 2003.
35. Andreas Geyer-Schulz and Michael Hahsler. Software reuse with analysis patterns. In *Proceedings of the 8th AMCIS*, pages 1156–1165, Dallas, TX, August 2002. Association for Information Systems.
36. Andreas Geyer-Schulz and Michael Hahsler. Evaluation of recommender algorithms for an internet information broker based on simple association rules and on the repeat-buying theory. In Brij Masand, Myra Spiliopoulou, Jaideep Srivastava, and Osmar R. Zaiane, editors, *Fourth WEBKDD Workshop: Web Mining for Usage Patterns & User Profiles*, pages 100–114, Edmonton, Canada, July 2002.
37. Walter Böhm, Andreas Geyer-Schulz, Michael Hahsler, and Maximillian Jahn. Repeat buying theory and its application for recommender services. In O. Opitz and M. Schwaiger, editors, *Exploratory Data Analysis in Empirical Research, Proceedings of the 25th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Munich, March 14–16, 2001*, pages 229–239. Springer-Verlag, 2002.
38. Andreas Geyer-Schulz, Michael Hahsler, and Maximillian Jahn. Recommendations for virtual universities from observed user behavior. In W. Gaul and G. Ritter, editors, *Classification, Automation, and New Media, Proceedings of the 24th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Passau, March 15–17, 2000*, pages 273–280. Springer-Verlag, 2002.
39. Andreas Geyer-Schulz, Michael Hahsler, and Maximillian Jahn. Wissenschaftliche Recommendersysteme in Virtuellen Universitäten. In H.-J. Appelrath, R. Beyer, U. Marquardt, H.C. Mayr, and C. Steinberger, editors, *Unternehmen Hochschule*, Wien, Österreich, September 2001. Symposium UH2001, GI Lecture Notes in Informatics (LNI).
40. Andreas Geyer-Schulz, Michael Hahsler, and Maximillian Jahn. A customer purchase incidence model applied to recommender systems. In *WEBKDD2001 Workshop: Mining Log Data Across All Customer TouchPoints*, pages 35–45, San Francisco, CA, August 2001.
41. Andreas Geyer-Schulz and Michael Hahsler. Automatic labelling of references for information systems. In Reinhold Decker and Wolfgang Gaul, editors, *Classification and Information Processing at the Turn of the Millennium, Proceedings of the 23rd Annual Conference of the Gesellschaft für Klassifikation e.V., University of Bielefeld, March 10–12, 1999*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 451–459. Springer-Verlag, 2000.
42. Andreas Geyer-Schulz and Michael Hahsler. Lebenslanges virtuelles Lernen. In Franciszek Grucza, editor, *Europas Arbeitswelt von Morgen*, pages 51–54, Wien, 2000. Wiener Zentrum der Polnischen Akademie der Wissenschaften.
43. Michael Hahsler and Bernd Simon. User-centered navigation re-design for web-based information systems. In H. Michael Chung, editor, *Proceedings of the Sixth Americas Conference on Information Systems (AMCIS 2000)*, pages 192–198, Long Beach, CA, 2000. Association for Information Systems.
44. Andreas Geyer-Schulz, Michael Hahsler, and Georg Schneider. The virtual university as a network economy. In Heinrich C. Mayr, Claudia Steinberger, Hans-Jürgen Appelrath, and Uwe Marquardt, editors, *Informatik '99, Unternehmen Hochschule '99, Workshop-Unterlagen*, pages 75–86, Bielefeld, Germany, October 1999.

INVITED TALKS

Society, industry, and university sponsored talks (presenting author listed first)

1. Michael Hahsler, Common curriculum reading kick-off lecture: Introduction to AI, Common Curriculum Reading: Artificial Unintelligence, Office of General Education, SMU, September 10, 2024 and January 28&29, 2025.
2. Michael Hahsler, Moderation of SMU Tower Center Forum on Ethical AI, September 21, 2023.
3. Michael Hahsler, What is AI? How did we get here? Where will it lead us?, Invited talk: CBRE – Texas Multifamily Exclusive Client Forum, April 19, 2023.
4. Michael Hahsler, Probabilistic approaches to mine association rules, October 2018. Department Seminar, Department of Statistics & Actuarial Science, University of Waterloo, Waterloo, Canada, October 2018.
5. Michael Hahsler, Farzad Kamalzadeh, Vishal Ahuja, and Michael Bowen, Electronic health record analytics: The case of optimal diabetes screening, Artificial Intelligence in Medicine Seminar Series Division of Medical Physics and Engineering, UT Southwestern, April 13, 2018
6. Michael Hahsler and Kurt Hornik, Dissimilarity Plots: A Visual Exploration Tool for Partitional Clustering, invited seminar, IE Department Seminar Department of Industrial, Manufacturing, & Systems Engineering, University of Texas at Arlington, February 2017.
7. Michael Hahsler, Recommender systems: Harnessing the power of personalization, Curricular Recommender System Working Group, SMU, February 17, 2017.
8. Michael Hahsler, Data mining tutorial: Methods and tools, November 2016. DCII - Operations Research and Statistics Towards Integrated Analytics Research Cluster, SMU, November 30, 2016.
9. Michael Hahsler, Probabilistic approach to association rule mining, invited seminar, Center for Marketing Analytics, IESEG School of Management, Lille, France, May 2016.
10. Michael Hahsler, Recommender systems: Harnessing the power of personalization, Invited talks at the Southwest Airlines EDGe Analyst Community Meeting, Dallas, TX, November 18, 2015.
11. Michael Hahsler, Association Rule Mining: Introduction to the R package arules, Dallas R Users Group, Dallas, TX, March 2015.
12. Michael Hahsler and Kurt Hornik, Dissimilarity Plots: A visual exploration tool for partitional clustering, Graduate Seminar, School of Industrial and Systems Engineering, University of Oklahoma, Norman, OK, November 2013.
13. Michael Hahsler, Introduction to the predictive model markup language, Orange County R User Group, CA, Webinar together with Ray DiGiacomo, Alex Guazzelli and Rajarshi Guha, January 24, 2012.
14. Michael Hahsler, Recommender systems: User-facing decision support systems, Invited talk for EMIS 7357 - Decision Support Systems, Southern Methodist University, Dallas, Texas, February 22, 2012.
15. Michael Hahsler and Maya Eldayeh, Analyzing incomplete biological pathways using network motifs, Division of Biomedical Informatics Retreat, UT Southwestern Medical Center, Dallas, TX, May 6 and 12, 2011.
16. Michael Hahsler and Kurt Hornik, Dissimilarity plots: A visual exploration tool for partitional clustering, Invited talk at "2011 Best of JCGS (Journal of Computational and Graphical Statistics) Session," 42nd Symposium on the Interface, Cary, NC, June 1–3, 2011.

17. Michael Hahsler, Recommender systems: From content to latent factor analysis, CSE Colloquium, Department of Computer Science and Engineering, Southern Methodist University, Dallas, Texas, September 7, 2011.
18. Michael Hahsler, Dissimilarity plots: A visual exploration tool for partitional clustering, CSE Colloquium, Department of Computer Science and Engineering, Southern Methodist University, Dallas, TX, April 3, 2009.
19. Michael Hahsler, A probabilistic approach to association rule mining. CSE Colloquium, Department of Computer Science and Engineering, Southern Methodist University, Dallas, Texas, October 10, 2008.
20. Michael Hahsler, Generating top-N recommendations from binary profile data. Berufungsvortrag Wirtschaftsinformatik, WU Wien, Vienna, Austria, July 16, 2008.
21. Michael Hahsler, Warenkorbanalyse mit Hilfe der Statistiksoftware R. WU Competence Day, Wirtschaftsuniversität Wien, Vienna, Austria, October 19, 2006.
22. Michael Hahsler, Patterns im Softwareentwicklungsprozeß (in German), ADV - Arbeitsgemeinschaft für Datenverarbeitung, Vienna, Austria, September 2001.

CONFERENCE PRESENTATIONS

Presentations with paper

All papers in the refereed conference and workshop proceeding section above were either presented by me or one of my coauthors (see Publications Section).

Presentations without paper (presenting author listed first)

1. Michael Hahsler , Farzad Kamalzadeh, Vishal Ahuja and Michael Bowen, Electronic health record analytics: The case of optimal diabetes screening, December 2018. EMIS Industry Advisory Board and Outreach Meeting, December 3, 2018.
2. Zahra Gharibi, Michael Hahsler, and Mehmet Ayvaci, The impact of information sharing on organ transplantation: A simulation model, November 2018. 2018 INFORMS Computing Society Conference, Phoenix, AZ, November 3, 2018.
3. Farzad Kamalzadeh, Michael Hahsler, Vishal Ahuja and Michael Bowen, An Analytics-driven Approach for Optimal Diabetes Screening Decisions, 2017 INFORMS Annual Meeting, November 2017.
4. Zahra Gharibi, Michael Hahsler, and Mehmet Ayvaci, Modeling Kidney Transplantation Decisions: Optimal Regulatory Oversight, 2017 INFORMS Annual Meeting, November 2017.
5. Paul S. Krueger, Mohammadreza Zharfa, Michael Hahsler, Eli V. Olinick and Sheila Williams, Quantitative Classification of Vortical Flows Based on Topological Features, IUTAM Symposium on Dynamics and Topology of Vorticity and Vortices, June 2017
6. Michael Hahsler, Vishal Ahuja, Michael Bowen and Farzad Kamalzadeh, Predictive models for making patient screening decisions, 2016 INFORMS Annual Meeting, November 2016.
7. Michael Hahsler and Young Woong Park, Sequential aggregation-disaggregation optimization methods for data stream mining, 2016 INFORMS Annual Meeting, November 2016.
8. Paul S. Krueger, Sheila Williams, Michael Hahsler and Eli V. Olinick, Flow Field Classification Using Critical Point Matching, 69th Annual Meeting of the APS Division of Fluid Dynamics, November 2016.
9. Michael Hahsler, Ordering objects: What heuristic should we use? 2015 INFORMS Annual Meeting, Philadelphia, PA, November 1–4, 2015.
10. Michael Hahsler and Jörg Lässig, Cooperative data analysis in supply chains using selective information disclosure, 2014 INFORMS Annual Meeting, San Francisco, CA, November 9–12, 2014.
11. Michael Hahsler and Matthew Bolaños, A study of the efficiency and accuracy of data stream clustering for large data sets, 2013 INFORMS Annual Meeting, Minneapolis Convention Center, October 6–9, 2013.
12. Michael Hahsler and Sudheer Chelluboina, Visualizing association rules in hierarchical groups, 42th Symposium on the Interface, Cary, NC, June 1–3, 2011.
13. Michael Hahsler, Two applications of the TSP for data analysis. 31st Annual Conference of the German Classification Society (GfKl 2007), Freiburg, March 7–9, 2007.
14. Michael Hahsler, Probabilistische Ansätze in der Assoziationsanalyse. Habilitationsvortrag, Wirtschaftsuniversität Wien, May 19, 2006.
15. Michael Hahsler and Kurt Hornik, An association rule mining infrastructure for the R data analysis toolbox, 30th Annual Conference of the German Classification Society (GfKl 2006), Berlin, March 8–10, 2006.

16. Michael Hahsler and Andreas Geyer-Schulz, ePubWU - Erfahrungen mit einer Volltextplattform an der Wirtschaftsuniversität Wien, 28. Österreichischer Bibliothekartag 2004, Linz, Austria.
17. Michael Hahsler, Generating synthetic transaction data for tuning usage mining algorithms, March 2003. 27th Annual GfKl-Conference, Cottbus, Germany.
18. Michael Hahsler and Andreas Geyer-Schulz, Living Lectures - WU Virtual Library: Ein Lernportal, March 2000. in Vortragsreihe "Lernen per Internet", Technische Universität Wien.
19. Michael Hahsler and Andreas Geyer-Schulz, Living Lectures - Virtual University Projekt: Informationstechnologie im universitären Bildungsbereich, June 1999. Global Village 99.

SUPERVISED GRADUATE STUDENT RESEARCH

Antonio Caiada, *Security threats in Transformer models with Reinforcement Learning (working title), Adviser,* Ph.D. student in CS. Expected graduation year 2025.

Farzad Kamalzadeh, *A Data-Driven Framework for Decision Making Under Uncertainty,* Adviser, Ph.D. in OR, SMU, May 2020. Job: Data Scientist at Citibank.

Zahra Gharibi, *Modeling Kidney Transplantation Decisions: Regulatory Oversight, Information Sharing, and Post-Transplant Drug Choice*, Adviser & Committee Chair, Ph.D. in OR, SMU, May 2018. Job: Faculty member at SUNY Plattsburgh.

Pimprapai Thainiam, *Local Search Strategies for the Seriation Problem,* Adviser & Committee Chair, DE in Engineering Management, SMU, May 2017. Job: Faculty member at King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand.

William Spurgin, *Instability in Principal Component Analysis,* Adviser & Committee Chair, MS in CS, SMU, December 2016. Job: Principal Engineer at Stitch Fix.

Charlie Isaksson, *New Outlier Detection Techniques for Data Streams,* (Adviser & Committee Chair), Ph.D. in CS, SMU, December 2016. Job: Data Science Manager at Oportun.

Hadil Shaiba, *Machine Learning Methods for Tropical Cyclone Intensity Prediction*, Adviser & Committee Chair, Ph.D. in CS, SMU, May 2016. Job: Faculty member, CS, College of Computer and Information Sciences at Princess Nora University, Saudi Arabia.

Jake Drew, *Scalable Machine Learning for Big Data Applications in Bioinformatics and Cybercrime,* Ph.D. in CS, SMU, December 2015. Co-adviser with T. Moore & Committee Chair. Job: Adjunct professor, SMU.

Anurag Nagar, *A Quasi-Alignment based Framework for Discovery of Conserved Regions and Classification of DNA Fragments*, Ph.D. in CS, SMU, 2013. Adviser & Committee Chair. Job: Associate Professor of Instruction in CS at the University of Texas at Dallas.

Xiaodian Xie, Agent-based Simulation of Chinese Urban Demographics, MS in CS, SMU, 2013. Adviser & Committee Chair. Job: Developer at Dow Jones & Company.

Maya El Dayeh, Biological Pathway Completion using Network Motifs, Ph.D. in CS, SMU, 2012. Adviser. Job: Clinical Assistant Professor, CS at SMU.

Christoph Breidert, Estimation of Willingness-to-Pay: Theory, Measurement and Application, Ph.D. in Business Informatics, WU Wien, 2005. Adviser. Job: CEO of 1xInternet.

Strahil Ivanov, Bernhard Beran, Martin Kersch, Martin Groblschegg, Klaus Brosche, Danijela Mitrovic, Harald Lenz, Maria Shustitskiy, Parmis Parham, Karin Wurm, Lukas Kotoulek-Steiner, Norbert Fellinger, Manfred Friscic, Christian Grübl, Markus Fraisl, Patrick Gerdenits, Jürgen Haller, Andrea Ziegler-Skopecek, Jasna Tusek, Walter Schlögl, Alois Geith, Susanne Hafner, Ferdinand Nest, Dietmar Wessely, Marian Formanko, Thomas Wehling, Karin Ernsthofer, Erich Brenner, Isabelle Seidl, Martin Tuma, Gert Vasak, Thomas Teufer, Michael Linhart, Sabine Kuzdas, Martin Vodenicharov, Gerold Hämmerle, Emina Mehic, Matthias Redl, Lukas Helm, MS in Business Administration (Master's Thesis), WU Wien, 2001-2007. Note: A thesis was required for all master students.

SUPERVISED UNDERGRADUATE RESEARCH

Zerui Ma: Recommender Systems and Set-Cover Algorithms for University Curriculum Advising, Robert Mayer Undergraduate Research Fellowship, Dedman College Interdisciplinary Institute (DCII), SMU, 2024-25.

Alex Shockley: Deep Q learning for Connect-4, Independent Study, SMU 2024.

Elias Mann: Context-aware music recommendation systems, Bachelor Distinction Thesis (CS), SMU, 2023. Job: Founder and Director of ML at Rowan.

Ian Johnson: *arulesCBA: arules add-on for classification based on association rules (CBA),* Barry Goldwater Scholar, CS, SMU, 2019. Job: Software engineer at Google.

Oscar Vallner: *Social Network Analysis - Twitter as a Learning Technology for Mathematics Teachers* (Co-supervised with Annie Garrison Wilhelm, School of Education, SMU), CS, SMU, 2018.

Alexander Saladna: *BaoBao: Web-based System that Helps Individuals to Improve Their Foreign Language Skills (Intelligent User Interface),* Engaged Learning (CSE), SMU, 2016.

Kyle Nakatsuka: *A Bioinformatics Approach to Understand Aging Mechanisms,* Mayer Undergraduate Fellow Program (Biology/CS), SMU, 2015. Master student in Public Health Student at the University of Hawaii at Manoa.

Derek Phanekham: *Mining Frequent Patterns in Data Streams,* Bachelor Distinction Thesis (CS), SMU, 2015. He received a Ph.D. in CS from SMU and is now a network researcher at Google.

Matt Bolaños: *Data Stream Clustering in R,* Bachelor Distinction Thesis (CS), SMU, 2014. Completed an MS program at Carnegie Mellon University and works as a Principal Product Manager at Microsoft.

John Forrest: *Stream: A Framework for Data Stream Modeling in R,* Bachelor Distinction Thesis (CS), SMU, 2011. Job: Software engineer at Microsoft.

Teaching

UNDERGRADUATE COURSES AT SMU

"DS 1300: A Practical Introduction to Data Science," Lyle School of Engineering, SMU, Spring 2019, Fall 2019, Fall 2020, Spring 2021, Fall 2021, Spring 2022, Fall 2022, Spring 2023.

"CS 1341: Principles of Computer Science," Lyle School of Engineering, SMU, Fall 2009, Fall 2010.

"CS 1342: Programming Concepts," Lyle School of Engineering, SMU, Spring 2010, Spring 2011, Fall 2011.

"CS 2341: Data Structures," Lyle School of Engineering, SMU, Fall 2022, Spring 2023, Fall 2023, Spring 2024.

"OREM 2360: Engineering Economy," Lyle School of Engineering, SMU, Fall 2012, Spring 2013, Fall 2013, Spring 2014, Fall 2014, Spring 2015, Fall 2015, Spring 2016, Spring 2020.

"OREM 3309: Information Engineering," Lyle School of Engineering, SMU, Fall 2016, Spring 2017, Spring 2020, Spring 2021, Spring 2023.

"OREM 3363: Discrete Event Simulation," Lyle School of Engineering, SMU, Spring 2021, Spring 2022.

Course development:

- Development of the new course DS/OREM 1300 in 2019 as the introductory course for the new data science minor.
- OREM 3309 was completely redesigned in 2016 as an introduction to data sciences covering databases and descriptive analytics with a major hands-on project experience.
- The curriculum of OREM 2360 was expanded in 2012 by introducing a major project component that ties all techniques studied in the course together. The project using a real-world decision problem aims at preparing the students for their summer internships.

UNDERGRADUATE COURSES AT OTHER UNIVERSITIES

"Introduction to Information Engineering and Management" (in German "Grundlagen der Informationswirtschaft"), WU Wien, Austria, Fall 2008, Spring 2009.

"Introduction to Programming with Java" (in German "Grundzüge der Programmierung mit Java"), WU Wien, Austria, Spring 2002, Fall 2002, Spring 2003, Fall 2003, Spring 2004, Fall 2004, Spring 2005, Fall 2005, Spring 2006, Fall 2006, Spring 2007.

"Information Management for Businesses" (in German "Informationsmanagement in Organisationen I / Informationswirtschaft 2"), WU Wien, Austria, Spring 2003, Spring 2004, Spring 2005, Fall 2005, Spring 2006, Fall 2006, Spring 2007.

"Programming Lab (Java, C++, Perl, Databases)" (in German "Rechnerpraktikum aus Programmierung"), WU Wien, Austria, Spring 2001, Fall 2001, Fall 2002, Fall 2003, Fall 2004, Fall 2005, Fall 2006.

"IT Internship with Thesis" (in German "IT-Praktikum mit Bakkalaureatsarbeit,"), WU Wien, Austria, Spring 2005, Spring 2006, Spring 2007, Fall 2008, Spring 2009.

"COAP 2120: Data Handling on the Web," Webster University (Vienna Campus), Austria, Spring II 2002.

"COAP 3110: Interactive Web Site Development," Webster University (Vienna Campus), Austria, Fall II 2002.

"Introduction to Electronic Data Processing" (in German "Elektronische Datenverarbeitung: Markup-Konzepte"), WU Wien, Austria, Fall 1998.

GRADUATE COURSES AT SMU

"CS 5/7320: Artificial Intelligence," Lyle School of Engineering, SMU, Fall 2020, Spring 2021, Fall 2021, Spring 2022, Fall 2022, Fall 2023, Spring 2024, Fall 2024.

"CS/OREM 5/7331: Data Mining," Lyle School of Engineering, SMU, Spring 2013, Spring 2014, Fall 2014, Fall 2015, Spring 2016, Fall 2017, Spring 2018, Fall 2018, Summer 2020, Fall 2020, Spring 2021, Fall 2024.

"CS 7343: Operating Systems and System Software," Lyle School of Engineering, SMU, Spring 2009.

"OREM 5/7361: Computer Simulation Techniques," Lyle School of Engineering, SMU, Fall 2019.

"CS 5/7337: Information Retrieval and Web Search," Lyle School of Engineering, SMU. Spring 2012.

"CS 8331: Advanced Data Mining," Lyle School of Engineering, SMU. Spring 2012, Spring 2015, Spring 2017, Spring 2019, Fall 2021.

"CS 8091: Advanced Scientific Computing with R," Lyle School of Engineering, SMU, Fall 2011.

"CS 8098: Computer Science Seminar," Lyle School of Engineering, SMU, Fall 2009, Spring 2010, Fall 2010, Spring 2011, Fall 2011, Spring 2012.

Course development:

- Creation of an online version of CS 7320 – Introduction to Artificial Intelligence in 2025.
- Complete redesign of CS 5/7320 in 2020 to focus on problem solving, uncertainty, and machine learning.
- Creation of the new CS/OREM 5/7331 course in 2013 with a project-driven curriculum. Four major projects cover the main topics in data mining: data understanding, classification, association analysis and clustering.
- Creation of the new CS 5/7337 course in 2012 with several exercises focusing on modern Big Data methods including Hadoop and MapReduce for distributed index creation.
- Development of a new, research-centric curriculum for CS/OREM 8331 in 2012. The focus is now on reading and understanding current research articles in the fast-changing field of data mining. Students learn how to replicate key results in the papers, how to prepare scientific reviews and how to prepare a conference tutorial to cover a new research topic.
- CS 8091 was developed in 2011 as a special topics course to introduce engineering students to the R programming language used for data science. The course content is now available online as a self-study guide with video lectures and exercises. The developed content is now used as a blended learning component in CS/OREM 5/7331.

GRADUATE COURSES AT OTHER UNIVERSITIES

"Recommendation Tools," IESEG School of Management, Lille, France, May 2016.

"Process-Oriented Information Engineering and Management" (in German "Prozessorientierte Informationswirtschaft"), WU Wien, Austria, Fall 2006, Spring 2007.

"Current Topics in Information Engineering and Management" (in German "Seminar aus Informationswirtschaft"), WU Wien, Austria, Spring 2000, Fall 2000, Fall 2001, Spring 2002 Fall 2002, Spring 2003, Spring 2004, Spring 2005, Spring 2006, Spring 2007.

"Introduction to Object Oriented Programming" (in German "Einführung in das objektorientierte Programmieren"), WU Wien, Austria, Spring 1999, Fall 1999, Spring 2000, Fall 2000, Spring 2001.

EXECUTIVE PROGRAMS AND PROFESSIONAL TRAINING

"CS 7343: Operating Systems and System Software," Executive Master's Program in Security Engineering, Lyle School of Engineering, SMU, Spring 2009.

"UML Basics: Introduction to Object Oriented Modeling" (in German "UML-Basics: Einführung in Objekt-Orientierte Modellierung mit der Unified Modeling Language"), ADV (Arbeitsgemeinschaft für Datenverarbeitung), Vienna, Austria, 2000 to 2001.

"Introduction to Object Oriented Programming with C++" (in German "Einführung in den Einsatz von Objekt-Orientierung mit C++"), ADV (Arbeitsgemeinschaft für Datenverarbeitung), Vienna, Austria, 2000.

EXTRACURRICULAR STUDENT MENTORING

Matthew Piekenbrock, Mentor for the project Estimating the Empirical Cluster Tree sponsored by the **Google Summer of Code,** 2017.

Olivia Buerkle, Connor Gracie, Ian Johnson, Kelsey O'Leary: **INFORMS O.R. & Analytics Student Team Competition,** 2016/17.

Taghreed Alghamdi, Ali Almadan, and students from Economic: **SAS Data Mining Competition,** Co-mentored with Prof. Thomas Fomby (Economics), 2015.

Zizhen Chen, Ren Peng and three students from Economics : **Capital One Data Mining Cup**. As **finalists,** the students presented their solution for "Optimal Search Engine Adword Pricing" at the Capital One Corporate Headquarters in McLean, VA. Co-mentored with Prof. Thomas Fomby (Economics), 2014.

Gabriel Ayala, Matthew Bolaños, Jake Drew, and Derek Phanekham : **IBM's The Great Mind Challenge – Watson Edition. 1st place** in a data mining competition with 62 participating teams, 2014.

K12 MENTORING

Senior thesis project judge at TAG Magnet School for the Talented and Gifted, Dallas ISD, TX: Beckham Myers, Asynchronous Computer Processor Design (TTL prototype and compiler), 2017.

Mentor for the Independent Study and Mentorship program, Liberty High School, Frisco ISD, TX: Devanshi Padsala, Teaching girls in rural India CS basics, 2016–17.

Mentor and judge for the Network for Teaching Entrepreneurship competition, NFTE Dallas Chapter, 2014.

Service

EDITORIAL LEADERSHIP

INFORMS Journal on Computing, *Associate Editor* (2021–)

Journal of Statistical Software, *Associate Editor* (2014–).

International Journal of Open Source Software and Processes (IJOSSP), *Editorial Review Board Member* (2008–2016).

JSM Computer Science & Engineering, *Editorial Board Member* (2013–2014).

ORGANIZATION LEADERSHIP

INFORMS Data Mining Section, *Secretary and Treasurer*, 2016–2017.

FEDERAL GRANT REVIEWER

NSF CISE, *Review Panel Member*, June 2017.

NSF CISE, *Review Panel Member*, November 2015.

NSF CISE, *Review Panel Member*, March 2015.

CONFERENCE ORGANIZATION

ICS 2015 - INFORMS Computing Society Conference 2015, Data Mining *Stream Co-Chair*, January 2015.

StreamKDD'10 - Novel Data Stream Pattern Mining Techniques, a workshop held in conjunction with the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2010), *Organizer*, July 2010.

GfKl 2007 - 31st Annual Conference of the German Classification Society, *Session Organizer*, "Tools for Intelligent Data Analysis," March 2007.

GfKl 2006 - 30th Annual Conference of the German Classification Society, *Session Organizer*, "Tools for Intelligent Data Analysis," March 2006.

PROGRAM COMMITTEE MEMBER

PAKDD2019 - 23rd Pacific Asia Conference on Knowledge Discovery and Data Mining, *Program Committee*, April 14–17, 2019, Macau, China.

PAKDD 2018 - The 22nd Pacific-Asia Conference on Knowledge Discovery and Data Mining, *Program Committee*, June 3–6, 2018, Melbourne, Australia.

BigDaCI 2017 - International Conference on Big Data Analytics, Data Mining and Computational Intelligence, *Program Committee*, July 21–23, 2017, Lisbon, Portugal.

PAKDD 2017 - The 21st Pacific-Asia Conference on Knowledge Discovery and Data Mining, *Program Committee*, 2017.

BigDaCI 2016 - International Conference on Big Data Analytics, Data Mining and Computational Intelligence, *Program Committee*, July 2–4, 2016, Madeira, Portugal.

DATA ANALYTICS 2016 - The Fifth International Conference on Data Analytics, *Program Committee*, October 9–13, 2016, Venice, Italy.

PAKDD 2016 - The 20th Pacific-Asia Conference on Knowledge Discovery and Data Mining, *Program Committee*, April 19–22, 2016, Auckland, New Zealand.

BICoB 2016 - 8th International Conference on Bioinformatics and Computational Biology, *Program Committee*, April 4–6, 2016, Las Vegas, Nevada, USA.

QIMIE'15 - Quality issues, measures of interestingness and evaluation of data mining models Workshop organized in association with the PAKDD'15 conference, *Program Committee*, May 2015.

ECDM'15 - 9th European Conference on Data Mining 2015, Multi Conference on Computer Science and Information Systems (MCCSIS 2014), *Program Committee*, July 2015.

DATA ANALYTICS 2015 - The Fourth International Conference on Data Analytics, *Program Committee*, July 19–24, 2015.

PAKDD 2015 - The 19th Pacific-Asia Conference on Knowledge Discovery and Data Mining, *Program Committee*, May 2015.

BICOB 2015 - 7th International Conference on Bioinformatics and Computational Biology, *Program Committee*, March 2015.

ECDM'14 - 8th European Conference on Data Mining 2014, Multi Conference on Computer Science and Information Systems (MCCSIS 2014), *Program Committee*, July 2014.

DATA ANALYTICS 2014 - The Third International Conference on Data Analytics, *Program Committee*, August 2014.

PAKDD 2014 - The 18th Pacific-Asia Conference on Knowledge Discovery and Data Mining, *Program Committee*, May 2014.

BICOB 2014 - 6th International Conference on Bioinformatics and Computational Biology, *Program Committee*, March 2014.

DATA ANALYTICS 2013 - The Second International Conference on Data Analytics, *Program Committee*, September/October 2013

ECDM'13 - IADIS European Conference on Data Mining, IADIS Multi Conference on Computer Science and Information Systems (MCCSIS 2013), *Program Committee*, July 2013

PRIB 2013 - 8th IAPR International Conference on Pattern Recognition in Bioinformatics, *Program Committee*, June 2013.

QIMIE'13 - Quality issues, measures of interestingness and evaluation of data mining models Workshop organized in association with the PAKDD'13 conference, *Program Committee*, April 2013.

GfKl 2012 - 36th Annual Conference of the German Classification Society, *Program Committee* (Machine Learning and Knowledge Discovery), August 2012.

DM 2012 - Data Mining, IADIS Multi Conference on Computer Science and Information Systems (MCCIS 2012), *Scientific Committee*, July 2012.

PAKDD 2012 - The 16th Pacific-Asia Conference on Knowledge Discovery and Data Mining, *Program Committee*, May 2012.

KDD 2011 - 17th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, *Program Committee*, August 2011.

QIMIE'11 - Quality Issues, Measures of Interestingness and Evaluation of Data Mining Models, workshop organized in association with the PAKDD'11 conference, *Program Committee*, May 2011.

QIMIE'09 - Quality Issues, Measures of Interestingness and Evaluation of Data Mining Models, workshop organized in association with the PAKDD'09 conference, *Program Committee*, April 2009.

WebKDD 2008 - Knowledge Discovery on the Web, held in conjunction with the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2008), *Program Committee*, August 2008.

WebKDD 2006 - Workshop on Web Mining and Web Usage Analysis, held in conjunction with the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2006), *Program Committee*, August 2006.

REVIEWER FOR INTERNATIONAL JOURNALS

Annals of Operations Research (ANOR)
Archives of Data Science, Series A (AoDSA)
Computational Statistics (CompStat)
Computational Statistics & Data Analysis (CSDA)
Data & Knowledge Engineering (DKE)
Data Mining and Knowledge Discovery (DAMI)
Electronic Commerce Research
European Journal of Operational Research (EJOR)
IEEE Transactions on Network and Service Management (TNSM)
IEEE Transactions on Knowledge and Data Engineering (TKDE)
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
IEEE Transactions on Systems, Man and Cybernetics (SMC)
IEEE Transactions on Visualization and Computer Graphics (TVCG)
Information Sciences
International Transactions in Operational Research
Journal of Applied Statistics
Journal of Computational and Graphical Statistics (JCGS)
Journal of Intelligent Information Systems (JIIS)
Journal of Machine Learning Research (JMLR)
Journal of Modelling in Management (JM2)
Journal of Retailing and Consumer Service
Journal of Statistical Computation and Simulation
Journal of Statistical Software (JSS)
Knowledge and Information Systems: An International Journal (KAIS)
Management Science
Pattern Analysis and Applications (PAAA)
Pattern Recognition Letters
PLOS ONE
Psychometrica
R Journal
SIGKDD Explorations
The VLDB Journal

UNIVERSITY AND DEPARTMENT SERVICE

Current

Undergraduate Program Committee (UPC), CS, 2022–present.

TA Selection Committee, CS, 2022–present.

Faculty Adviser for the SMU Artificial Intelligence Club, CS, 2022–present.

Faculty Adviser of the SMU Chapter of the Upsilon Pi Epsilon Honor Society for the Computing and Information Disciplines, CS, 2020-present.

Member of the Guild of Marshals, SMU, 2011-present.

Member of the Ph.D. committees for Peter Furseth (OR), Ishna Satyarth (CS), Junhao Shen (CS), Beichen Hu (CS).

Member of the M.S. committees for Zhongdi Wu (CS).

Past

Member of the Data Science Minor/Major Committee, SMU, 2019–2023.

Member of the Course Recommender System Task Force, SMU, 2017–2018.

Member of the Center for Global Health Impact, SMU, 2014–2018.

Member of the Departmental Preliminary Counseling Exam Committee, EMIS, SMU, 2014–2016.

Member of the Senior Design Task Force, Lyle School of Engineering, SMU, 2016.

Program director for Master of Science in Information Engineering and Management, SMU, 2013–2015.

Liaison to the CIO executive development program committee of the EMIS Distinguished Advisory Council, EMIS, SMU 2012–2015.

Chair of the department's Undergraduate Program Committee, CSE, SMU, 2010–2012.

Department colloquium coordinator, CSE, SMU, 2009–2012.

Teaching assistant coordinator and member of the department's teaching assistant selection committee, CSE, SMU, 2009–2012.

Member of the Ph.D. committees for Christoph Breidert (Adviser, WU Wien, 2005), Mallik Kotamarti (CS, SMU, 2010), Yu Su, (CS, SMU, 2011), Maya El Dayeh (CS, Adviser, SMU, 2012), Anurag Nagar (CS, Adviser and Chair, SMU, 2013), Richard Goodrum (CS, SMU, 2013), Azi Sharif (Applied Sciences, SMU, 2014), Xiujun Zhu (Statistics, SMU, 2014), John Howard (CS, SMU, 2014), Jake Drew (CS, Chair and Co-adviser, SMU, 2015), Hadil Shaiba (CS, Chair and Co-adviser, SMU, 2016), Adel K. Alblawi (Systems Engineering, SMU, 2016), Rodman P. Abbott (Systems Engineering, SMU, 2016), Charlie Isaksson (CS, Adviser and Chair, SMU, 2016), Chatchai Wangwiwattana (CS, SMU, 2017), Eman A. Ababtain (CS, SMU, 2018), Zahra Gharibi (OR, Adviser and Chair, SMU, 2018),i Fernando Vilas (CS, SMU, 2020), Derek Phanekham (CS, SMU, 2022), Adam Colley (OR, SMU, 2022), John M. Green (Systems Engineering, SMU, 2022), Zizhen Chen (CS, SMU, 2023), Hao Jiang (OR, SMU, 2023), Xihao Xie (CS, SMU, 2024), Adreana Julander (OR, SMU, 2024), Clay Harper (CS, SMU, 2024).

Member of Doctor of Engineering committees for Laith Abuhilal (Engineering Management, SMU, 2014), Lewis A. Sykalski (Software Engineering, SMU, 2015), Pimprapai Thainiam (Engineering Management, Adviser and Chair, SMU, 2017).

Member of the master supervisory committees for Austin Hodges (CS, SMU, 2013), Xiaodian Xie (Adviser and Chair, CS, SMU, 2013) , William O'Connor (CS, SMU, 2016), Andrew Blanchard

(Electrical Engineering, SMU, 2016), William Spurgin (CS, Adviser and Chair, SMU, 2016). Farbod Tavakkoli (Environmental and Civil Engineering, SMU, 2021)

Member of the committee to implement a new Business Informatics degree program, Vienna University of Economics and Business (WU Wien), 2004–2006.

Member of the habilitation committee for Christopher Casey, WU Wien, 2004.

Department research evaluation coordinator, WU Wien, 2002.

Undergraduate EDP exam coordinator, WU Wien, 1999–2002.

REFERENCES (IN ALPHABETICAL ORDER)

Sila Çetinkaya
Cecil H. Green Professor of Engineering
Department Chair, Department of Operations Research and Engineering Management
Southern Methodist University
Dallas, TX 75275, USA
Phone: +1 214-768-1100
Email: sila@smu.edu

Kurt Hornik
Department Chair, Department of Finance, Accounting and Statistics
Head of the Research Institute for Computational Methods
Vienna University of Economics and Business
Welthandelsplatz 1, 1020 Wien, Austria
Phone: +43 1 31336-4756
Email: kurt.hornik@wu.ac.at

Suku Nair
University Distinguished Professor
Vice Provost for Research and Chief Innovation Officer
Southern Methodist University
Dallas, TX 75275, USA
Phone: +1 214-768-2856
Email: nair@smu.edu

Jia Zhang
Cruse C. and Marjorie F. Calahan Centennial Chair in Engineering
Interim Department Chair, Department of Computer Science
Southern Methodist University
Dallas, TX 75275, USA
Phone: +1 214-768-3083
Email: jiazhang@smu.edu