

# ROBSON LEONARDO FERREIRA CORDEIRO

## PERSONAL

University Address:

Computer Science Department  
University of São Paulo  
400 Trabalhador São-carlense Avenue  
São Carlos, SP, Brazil 13566-590

Phone: +55 16 3373-6715

Email: robson@icmc.usp.br

## EDUCATION

Year	Degree	Institution
1999-2002	<b>B.Sc.</b> Computer Science	University of Western São Paulo (UNOESTE), Brazil
2003-2005	<b>M.Sc.</b> Computer Science	Federal University of Rio Grande do Sul (UFRGS), Brazil
2007-2011	<b>Ph.D.</b> Computer Science and Computational Mathematics	University of São Paulo (USP), Brazil <b>Visiting researcher:</b> one year at the <b>Carnegie Mellon University, USA</b> , under the local supervision of the co-advisor Prof. Dr. Christos Faloutsos.
2011-2013	<b>Post-Doctorate</b> Computer Science and Computational Mathematics	University of São Paulo (USP), Brazil
2021	<b>Post-Doctoral Degree</b> "Livre Docência" Development of Complex Data Managers	University of São Paulo (USP), Brazil

## UNIVERSITY EXPERIENCE

**2004 - 2006** Assistant Professor, University of Western São Paulo (UNOESTE), Brazil.  
**2013 - 2021** Assistant Professor (Tenured), University of São Paulo (USP), Brazil.  
**2021 -** **Associate Professor (Tenured)**, University of São Paulo (USP), Brazil.

## RESEARCH EXPERTISE

**Computer Science**

**Main Areas:** Artificial Intelligence, Machine Learning, Data Mining, and Databases.

## PUBLICATIONS

### BOOK

1. R. L. F. Cordeiro, C. Faloutsos, and C. Traina Jr. Data Mining in Large Sets of Complex Data. **SpringerBriefs in Computer Science**. Springer, 2013. 116 pages. ISBN: 978-1-4471-4889-0. DOI: 10.1007/978-1-4471-4890-6. URL: <https://doi.org/10.1007/978-1-4471-4890-6>.

→ **ACM Computing Reviews' Notable Books and Articles 2013**<sup>1</sup>

→ **13 thousand chapter downloads**<sup>2</sup>

### EDITED BOOK

1. A. J. M. Traina, C. Traina Jr., and R. L. F. Cordeiro. Similarity Search and Applications. **Lecture Notes in Computer Science**. Springer, 2014. 301 pages. Volume 8821. ISBN: 978-3-319-11987-8. DOI: 10.1007/978-3-319-11988-5. URL: <https://doi.org/10.1007/978-3-319-11988-5>.

<sup>1</sup> [http://computingreviews.com/recommend/bestof/notableitems\\_2013.cfm](http://computingreviews.com/recommend/bestof/notableitems_2013.cfm)

<sup>2</sup> <https://link.springer.com/book/10.1007/978-1-4471-4890-6>

## BOOK CHAPTERS

1. A. J. M. Traina, C. Traina Jr., R. L. F. Cordeiro, M. X. Ribeiro, and P. M. Azevedo-Marques. Issues and Techniques to Mitigate the Performance Gap in Content-based Image Retrieval Systems. **New Technologies for Advancing Healthcare and Clinical Practices**. IGI Global, pages 60-83, 2011. ISBN: 978-1-60960-780-7. DOI: 10.4018/978-1-60960-780-7.ch004. URL <https://doi.org/10.4018/978-1-60960-780-7.ch004>.
2. I. R. V. Pola, R. L. F. Cordeiro, C. Traina Jr., and A. J. M. Traina. A New Concept of Sets to Handle Similarity in Databases: the SimSets. **Similarity Search and Applications**. Springer, pages 30-42, 2013, Lecture Notes in Computer Science, Volume 8199. ISBN: 978-3-642-41061-1. DOI: 10.1007/978-3-642-41062-8\_4. URL [https://doi.org/10.1007/978-3-642-41062-8\\_4](https://doi.org/10.1007/978-3-642-41062-8_4).
3. H. Gualdrón, R. L. F. Cordeiro, J. F. Rodrigues Jr., D. H. P. Chau, M. Kahng, and U. Kang. M-flash: Fast Billion-scale Graph Computation Using a Bimodal Block Processing Model. **Machine Learning and Knowledge Discovery in Databases**. Springer, pages 623-640, 2016, Lecture Notes in Computer Science, Volume 9852. ISBN: 978-3-319-46226-4. DOI: 10.1007/978-3-319-46227-1\_39. URL [https://doi.org/10.1007/978-3-319-46227-1\\_39](https://doi.org/10.1007/978-3-319-46227-1_39). **World-class Publication**<sup>3</sup>
4. G. Q. Vasconcelos, D. S. Kaster, and R. L. F. Cordeiro. On the Support of the Similarity-aware Division Operator in a Commercial RDBMS. **Advances in Databases and Information Systems**. Springer, pages 142-155, 2018, Lecture Notes in Computer Science, Volume 11019. ISBN: 978-3-319-98397-4. DOI: 10.1007/978-3-319-98398-1\_10. URL: [https://doi.org/10.1007/978-3-319-98398-1\\_10](https://doi.org/10.1007/978-3-319-98398-1_10).
5. A. G. Bispo Jr., and R. L. F. Cordeiro. Fast and Scalable Outlier Detection with Metric Access Methods. **Computational Science**. Springer, pages 189-203, 2019, Lecture Notes in Computer Science, Volume 11537. ISBN: 978-3-030-22740-1. DOI: 10.1007/978-3-030-22741-8\_14. URL: [https://doi.org/10.1007/978-3-030-22741-8\\_14](https://doi.org/10.1007/978-3-030-22741-8_14).
6. A. I. A. Lescano, and R. L. F. Cordeiro. Relational Conditional Set Operations. **New Trends in Database and Information Systems**. Springer, pages 38-49, 2021, Communications in Computer and Information Science, Volume 1450. ISBN: 978-3-030-85081-4. DOI: 10.1007/978-3-030-85082-1\_4. URL: [https://doi.org/10.1007/978-3-030-85082-1\\_4](https://doi.org/10.1007/978-3-030-85082-1_4).

## REFEREED JOURNALS

1. L. F. D. Santos, W. D. Oliveira, M. R. P. Ferreira, R. L. F. Cordeiro, A. J. M. Traina, and C. Traina Jr. Evaluating the Diversification of Similarity Query Results. **Journal of Information and Data Management**, Volume 4, Number 3, pages 188-203, 2013. URL: <http://seer.lcc.ufmg.br/index.php/jidm/article/view/205>.
2. R. L. F. Cordeiro, A. J. M. Traina, C. Faloutsos, and C. Traina Jr. Halite: Fast and Scalable Multi-resolution Local-correlation Clustering. **IEEE Transactions on Knowledge and Data Engineering**, Volume 25, Number 2, pages 387-401, 2013. ISSN 1041-4347. DOI: 10.1109/TKDE.2011.176. URL: <http://dx.doi.org/10.1109/TKDE.2011.176>. **World-class Publication**<sup>4</sup>
3. R. L. F. Cordeiro, F. Guo, D. S. Haverkamp, J. H. Horne, E. K. Hughes, G. Kim, L. A. S. Romani, P. P. Coltri, T. T. Souza, A. J. M. Traina, C. Traina Jr., and C. Faloutsos. QuMinS: Fast and Scalable Querying, Mining and Summarizing Multi-modal Databases. **Elsevier Information Sciences**, Volume 264, pages 211-229, 2014. DOI: 10.1016/j.ins.2013.11.013. URL: <https://doi.org/10.1016/j.ins.2013.11.013>. **World-class Publication**<sup>5</sup>
4. I. R. V. Pola, R. L. F. Cordeiro, C. Traina Jr., and A. J. M. Traina. Similarity Sets: A New Concept of Sets to Seamlessly Handle Similarity in Database Management Systems. **Elsevier Information Systems**, Volume 52, pages 130-148, 2015. DOI: 10.1016/j.is.2015.01.011. URL: <https://doi.org/10.1016/j.is.2015.01.011>. **World-class Publication**<sup>6</sup>
5. G. P. Gimenes, R. L. F. Cordeiro, and J. F. Rodrigues Jr. ORFEL: Efficient Detection of Defamation or Illegitimate Promotion in Online Recommendation. **Elsevier Information Sciences**, Volume 379, pages 274-287, 2017. DOI: 10.1016/j.ins.2016.09.006. URL: <https://doi.org/10.1016/j.ins.2016.09.006>. **World-class Publication**<sup>7</sup>

<sup>3</sup> CORE Rank: A. See <http://portal.core.edu.au/conf-ranks/491/>

<sup>4</sup> CORE Rank: A\*. See <http://portal.core.edu.au/jnl-ranks/415/>

<sup>5</sup> CORE Rank: A. See <http://portal.core.edu.au/jnl-ranks/680/>

<sup>6</sup> CORE Rank: A\*. See <http://portal.core.edu.au/jnl-ranks/520/>

<sup>7</sup> CORE Rank: A. See <http://portal.core.edu.au/jnl-ranks/680/>

6. A. S. Gonzaga, and R. L. F. Cordeiro. The Similarity-aware Relational Division Database Operator with Case Studies in Agriculture and Genetics. **Elsevier Information Systems**, Volume 82, pages 71-87, 2019. DOI: 10.1016/j.is.2019.01.002. URL: <https://doi.org/10.1016/j.is.2019.01.002>. **World-class Publication**<sup>6</sup>
7. A. Lima, A. Florez, A. Lescano, J. Novaes, N. Martins, C. Traina Jr., E. P. M. de Sousa, J. F. Rodrigues Jr., and R. L. F. Cordeiro. Analysis of ENEM's attendants between 2012 and 2017 using a clustering approach. **Journal of Information and Data Management**, Volume 11, Number 2, pages 115-130, 2020. URL: <https://periodicos.ufmg.br/index.php/jidm/article/view/24835>.
8. J. J. M. Oliveira, and R. L. F. Cordeiro. Unsupervised Dimensionality Reduction for Very Large Datasets: are we going to the right direction? **Elsevier Knowledge-Based Systems**, Volume 196:105777, 14 pages, 2020. DOI: 10.1016/j.knosys.2020.105777. URL: <https://doi.org/10.1016/j.knosys.2020.105777>.
9. M. A. C. Alves, and R. L. F. Cordeiro. Effective and Unburdensome Forecast of Highway Traffic Flow with Adaptive Computing. **Elsevier Knowledge-Based Systems**, Volume 212:106603, 2021. DOI: 10.1016/j.knosys.2020.106603. URL: <https://doi.org/10.1016/j.knosys.2020.106603>.
10. D. F. Milon-Flores, and R. L. F. Cordeiro. How to Take Advantage of Behavioral Features for the Early Detection of Grooming in Online Conversations. **Elsevier Knowledge-Based Systems**, Volume 240:108017, 2022. DOI: 10.1016/j.knosys.2021.108017. URL: <https://doi.org/10.1016/j.knosys.2021.108017>.

#### REFEREED CONFERENCES

1. R. L. F. Cordeiro, C. S. dos Santos, N. Edelweiss, and R. M. Galante. Classificação de Restrições de Integridade em Bancos de Dados Temporais de Versões. In **19<sup>th</sup> Brazilian Symposium on Databases - SBD 2004**, pages 363-377. Brasília, Brazil, 2004. ISBN 85-7669-009-8. (in Portuguese)
2. R. L. F. Cordeiro, N. Edelweiss, R. M. Galante, and C. S. dos Santos. TVCL - Temporal Versioned Constraint Language. In **20<sup>th</sup> Brazilian Symposium on Databases - SBD 2005**, pages 55-69. Uberlândia, Brazil, 2005. ISBN 85-7669-029-2.
3. D. S. Kaster, R. L. F. Cordeiro, and R. P. M. Fortes. Um Framework Web Colaborativo Baseado em Ontologias para Recuperação e Descoberta de Conhecimento em Imagens Médicas. In **13<sup>rd</sup> Brazilian Symposium on Multimedia and the Web - WebMedia 2007**, pages 246-253. Gramado, Brazil, 2007. ISBN 978-85-7669-150-1. (in Portuguese)
4. R. L. F. Cordeiro, R. M. Galante, N. Edelweiss, and C. S. dos Santos. A Deep Classification of Temporal Versioned Integrity Constraints for Designing Database Applications. In **19<sup>th</sup> International Conference on Software Engineering & Knowledge Engineering - SEKE 2007**, pages 416-421. Boston, USA, 2007. ISBN 1-891706-20-9.
5. R. L. F. Cordeiro, A. J. M. Traina, C. Faloutsos, and C. Traina Jr. Finding Clusters in Subspaces of Very Large, Multi-dimensional Datasets. In **26<sup>th</sup> IEEE International Conference on Data Engineering - ICDE 2010**, pages 625-636. Long Beach, USA, 2010. ISBN 978-1-4244-5444-0. DOI: 10.1109/ICDE.2010.5447924. URL: <https://doi.org/10.1109/ICDE.2010.5447924>. **World-class Publication**<sup>8</sup>
6. R. L. F. Cordeiro, F. Guo, D. S. Haverkamp, J. H. Horne, E. K. Hughes, G. Kim, A. J. M. Traina, C. Traina Jr., and C. Faloutsos. QMAS: Querying, Mining and Summarization of Multi-modal Databases. In **10<sup>th</sup> IEEE International Conference on Data Mining - ICDM 2010**, pages 785-790, Sydney, Australia, 2010. ISBN 978-0-7695-4256-0. DOI: 10.1109/ICDM.2010.150. URL: <https://doi.org/10.1109/ICDM.2010.150>. **World-class Publication**<sup>9</sup>
7. P. P. Coltri, R. L. F. Cordeiro, T. T. de Souza, L. A. S. Romani, J. Zullo Jr., C. Traina Jr., and A. J. M. Traina. Classificação de Áreas de Café em Minas Gerais por meio do Novo Algoritmo QMAS em Imagem Espectral Geoeye-1. In **15<sup>th</sup> Brazilian Symposium on Remote Sensing - SBSR 2011**, pages 539-546, Curitiba, Brazil 2011. ISBN 978-85-17-00057-7. (in Portuguese)
8. R. L. F. Cordeiro, C. Traina Jr., A. J. M. Traina, J. López, U. Kang, and C. Faloutsos. Clustering Very Large Multi-dimensional Datasets with Mapreduce. In **17<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 2011**, pages 690-698, San Diego, USA, 2011. ISBN 978-1-4503-0813-7. DOI: 10.1145/2020408.2020516. URL: <https://doi.org/10.1145/2020408.2020516>. **World-class Publication**<sup>10</sup>

→ **217 citations**<sup>11</sup>

<sup>8</sup> CORE Rank: A\*. See <http://portal.core.edu.au/conf-ranks/986/>

<sup>9</sup> CORE Rank: A\*. See <http://portal.core.edu.au/conf-ranks/642/>

<sup>10</sup> CORE Rank: A\*. See <http://portal.core.edu.au/conf-ranks/26/>

<sup>11</sup> According to Google Scholar. See <https://scholar.google.com/citations?user=bDCcqZ8AAAAJ&hl=en>

9. K. Fasolin, R. Fileto, M. Krüger, D. S. Kaster, M. R. P. Ferreira, R. L. F. Cordeiro, A. J. M. Traina, and C. Traina Jr. Efficient Execution of Conjunctive Complex Queries on Big Multimedia Databases. In **15<sup>th</sup> IEEE International Symposium on Multimedia - ISM 2013**, pages 536-543, Anaheim, USA, 2013. ISBN 978-1-4799-2171-3. DOI: 10.1109/ISM.2013.112. URL: <https://doi.org/10.1109/ISM.2013.112>.
10. A. E. D. Silva, and R. L. F. Cordeiro. Halite<sub>ds</sub>: Agrupamento de Dados em Subespaços de Séries Temporais Multidimensionais. In **29<sup>th</sup> Brazilian Symposium on Databases - SBBD 2014**, pages 97-106, Curitiba, Brazil, 2014. ISBN: 978-85-7669-290-4. URL: <https://www.inf.ufpr.br/sbbd-sbbsc2014/sbbd/proceedings/artigos/pdfs/74.pdf>. (in Portuguese)
11. A. S. Gonzaga, and R. L. F. Cordeiro. Fast and Scalable Relational Division on Database Systems. In **31<sup>st</sup> Brazilian Symposium on Databases - SBBD 2016**, pages 169-174, Salvador, Brazil, 2016. URL: [http://sbbd2016.fpc.ufba.br/sbbd2016/sbbd/sbbd\\_s012.pdf](http://sbbd2016.fpc.ufba.br/sbbd2016/sbbd/sbbd_s012.pdf).
12. P. H. Oliveira, A. C. Fraideinberze, N. A. Laverde, H. Gualdrón, A. S. Gonzaga, L. D. Ferreira, W. D. Oliveira, J. F. Rodrigues Jr., R. L. F. Cordeiro, C. Traina Jr., A. J. M. Traina, and E. P. M. de Sousa. On the Support of a Similarity-enabled Relational Database Management System in Civilian Crisis Situations. In **18<sup>th</sup> International Conference on Enterprise Information Systems - ICEIS 2016**, pages 119-126, Rome, Italy, 2016. ISBN: 978-989-758-187-8. DOI: 10.5220/0005816701190126. URL: <https://doi.org/10.5220/0005816701190126>.
13. A. E. Silva, L. L. Sanches, A. C. Fraideinberze, and R. L. F. Cordeiro. Halite<sub>ds</sub>: Fast and Scalable Subspace Clustering for Multidimensional Data Streams. In **16<sup>th</sup> SIAM International Conference on Data Mining - SDM 2016**, pages 351-359, Miami, USA, 2016. ISBN: 978-1-61197-434-8. DOI: 10.1137/1.9781611974348.40. URL: <https://doi.org/10.1137/1.9781611974348.40>. **World-class Publication**<sup>12</sup>
14. A. S. Gonzaga, and R. L. F. Cordeiro. The Similarity-aware Relational Division Database Operator. In **32<sup>nd</sup> ACM Symposium on Applied Computing - SAC 2017**, pages 913-914, Marrakech, Morocco, 2017. ISBN: 978-1-4503-4486-9. DOI: 10.1145/3019612.3019869. URL: <https://doi.org/10.1145/3019612.3019869>.
15. A. S. Gonzaga, and R. L. F. Cordeiro. A New Division Operator to Handle Complex Objects in Very Large Relational Datasets. In **20<sup>th</sup> International Conference on Extending Database Technology - EDBT 2017**, pages 474-477, Venice, Italy, 2017. ISBN: 978-3-89318-073-8. DOI: 10.5441/002/edbt.2017.48. URL: <https://doi.org/10.5441/002/edbt.2017.48>. **World-class Publication**<sup>13</sup>
16. G. Q. Vasconcelos, G. F. Zabet, D. M. Lima, J. F. Rodrigues Jr., C. Traina Jr., D. S. Kaster, and R. L. F. Cordeiro. TendeR-Sims - Similarity Retrieval System for Public Tenders. In **20<sup>th</sup> International Conference on Enterprise Information Systems - ICEIS 2018**, pages 143-150, Funchal, Portugal, 2018. ISBN: 978-989-758-298-1. DOI: 10.5220/0006697601430150. URL: <https://doi.org/10.5220/0006697601430150>.
17. L. F. Kunze, T. Amaral, L. M. P. Moraes, J. J. M. Oliveira, A. G. B. Junior, E. P. M. Sousa, and R. L. F. Cordeiro. Classification Analysis of NDVI Time Series in Metric Spaces for Sugarcane Identification. In **20<sup>th</sup> International Conference on Enterprise Information Systems - ICEIS 2018**, pages 162-169, Funchal, Portugal, 2018. ISBN: 978-989-758-298-1. DOI: 10.5220/0006709401620169. URL: <https://doi.org/10.5220/0006709401620169>.
18. M. T. Cazzolato, L. S. Rodrigues, L. C. Scabora, G. F. Zabet, G. Q. Vasconcelos, D. Y. T. Chino, A. E. S. Jorge, R. L. F. Cordeiro, C. Traina Jr., and A. J. M. Traina. A DBMS-based Framework for Content-based Retrieval and Analysis of Skin Ulcer Images in Medical Practice. In **34<sup>th</sup> Brazilian Symposium on Databases - SBBD 2019**, pages 109-120, Fortaleza, Brazil, 2019. ISSN: 2763-8979. DOI: 10.5753/sbbd.2019.8812. URL: <https://doi.org/10.5753/sbbd.2019.8812>.
19. C. Traina Jr., A. Moriyama, G. M. Rocha, R. L. F. Cordeiro, C. D. A. Ciferri, and A. J. M. Traina. The SimilarQL framework: Similarity Queries in Plain SQL. In **34<sup>th</sup> ACM Symposium on Applied Computing - SAC 2019**, pages 468-471, Limassol, Cyprus, 2019. ISBN: 978-1-4503-5933-7. DOI: 10.1145/3297280.3299736. URL: <https://doi.org/10.1145/3297280.3299736>.
20. L. M. P. Moraes, and R. L. F. Cordeiro. Detecting Influencers in Very Large Social Networks of Games. In **21<sup>st</sup> International Conference on Enterprise Information Systems - ICEIS 2019**, pages 93-103, Heraklion, Greece, 2019. ISBN: 978-989-758-372-8. DOI: 10.5220/0007728200930103. URL: <https://doi.org/10.5220/0007728200930103>.
21. E. F. Cabral, and R. L. F. Cordeiro. Fast and Scalable Outlier Detection with Sorted Hypercubes. In **29<sup>th</sup> ACM International Conference on Information and Knowledge Management - CIKM 2020**, pages 95-104, Virtual Event, Ireland, 2020. ISBN: 978-1-4503-6859-9. DOI: 10.1145/3340531.3412033. URL: <https://doi.org/10.1145/3340531.3412033>. **World-class Publication**<sup>14</sup>

<sup>12</sup> CORE Rank: A. See <http://portal.core.edu.au/conf-ranks/1727/>

<sup>13</sup> CORE Rank: A. See <http://portal.core.edu.au/conf-ranks/528/>

<sup>14</sup> CORE Rank: A. See <http://portal.core.edu.au/conf-ranks/25/>

22. S. Jiang, R. L. F. Cordeiro, and L. Akoglu. D.MCA: Outlier Detection with Explicit Micro-Cluster Assignments. In **22<sup>nd</sup> IEEE International Conference on Data Mining - ICDM 2022**, pages 987-992, Orlando, USA, 2022. ISBN 978-1-6654-5099-7. DOI: 10.1109/ICDM54844.2022.00119. URL: <https://doi.org/10.1109/ICDM54844.2022.00119>. **World-class Publication**<sup>9</sup>

#### REFEREED WORKSHOPS

**Other 8 articles** were published in Brazilian and International Workshops.

#### PRESENTATION OF MY WORK IN CONFERENCES / WORKSHOPS

- **ACM SIGKDD**; **IEEE ICDM**; **IEEE ICDE**; **SIAM SDM**; LSNA-WWW; ICEIS; SIAM AN; SBBD; CTD-CSBC.

#### GRANTS

##### COORDINATOR / BENEFICIARY

1. **1998 - 1998**; **Title**: Instrumentation Laboratory for Teaching Exact Sciences - LIENCE; **Funding Agency**: State University of Mato Grosso - UNEMAT.
2. **2007 - 2008**; **Title**: Data Mining in Large Sets of Complex Data; **Funding Agency**: Coordination for the Improvement of Higher Education Personnel - CAPES; **Values**: R\$ 16,728.00.
3. **2008 - 2010**; **Title**: Data Mining in Large Medical Image Databases; **Funding Agency**: The São Paulo Research Foundation - FAPESP; **Grant Number**: 2007/01639-4; **Values**: R\$ 84,963.84.
4. **2009 - 2010**; **Title**: Data Mining in Large Sets of Complex Data; **Funding Agency**: Coordination for the Improvement of Higher Education Personnel - CAPES; **Grant Number**: BEX 2454/09-2; **Values**: US\$ 15,240.00.
5. **2011 - 2013**; **Title**: Fractal Theory and Correlation Clustering Techniques to Improve Climate Change Forecast; **Funding Agency**: The São Paulo Research Foundation - FAPESP; **Grant Number**: 2011/06655-3; **Values**: R\$ 128,001.60.
6. **2015 - 2017**; **Title**: The Similarity-aware Relational Division Database Operator; **Funding Agency**: The São Paulo Research Foundation - FAPESP; **Grant Number**: 2014/21483-2; **Values**: R\$ 29,503.95.
7. **2015 - 2016**; **Title**: Relational Division by Similarity in Databases; **Funding Agency**: The São Paulo Research Foundation - FAPESP; **Grant Number**: 2015/05607-6; **Values**: R\$ 26,900.28.
8. **2018 - 2018**; **Title**: Evaluating Similarity Join Algorithms for Data Streams; **Funding Agency**: The São Paulo Research Foundation - FAPESP; **Grant Number**: 2017/21512-0; **Values**: R\$ 7,444.80.
9. **2018 - 2020**; **Title**: Mining Frequent Data Streams of High Dimensionality with a Case Study in Digital Games; **Funding Agency**: The São Paulo Research Foundation - FAPESP; **Grant Number**: 2018/05714-5; **Values**: R\$ 44,791.12 + US\$ 2,826.24.
10. **2022 - 2023**; **Title**: Exploratory Analysis of Microdata from the National Secondary Education Examination - ENEM; **Funding Agency**: The São Paulo Research Foundation - FAPESP; **Grant Number**: 2022/01210-8; **Values**: R\$ 10,561.32.

#### STUDENTS DIRECTED

##### Ph.D. ADVISING

1. **2019 -** — **Student**: Braulio Valentin Sánchez Vinces; **Title**: Anomaly Detection in High Frequency and Dimensionality Data Flows; **Funding Agency**: Coordination for the Improvement of Higher Education Personnel - CAPES; **Grant Number**: PROEX-10712476/D.
2. **2022 -** — **Student**: Felipe Pires Ferreira; **Title**: Correlation of Databases of Public and Private Institutions to Identify Anomalies in Gold Mining Activities in Brazil.

##### M.Sc. ADVISING

1. **2015 - 2017** — **Student**: André dos Santos Gonzaga; **Title**: The Similarity-aware Relational Division Database Operator; **Funding Agencies**: Coordination for the Improvement of Higher Education Personnel - CAPES (beginning), The São Paulo Research Foundation - FAPESP (posterior); **Grant Numbers (CAPES; FAPESP)**: PROEX-7152890/M; 2015/05607-6.

→ Ranked among the **9 best M.Sc. Theses in Computer Science** concluded in 2017 in Brazil<sup>15</sup>.

<sup>15</sup> <https://sol.sbc.org.br/index.php/ctd/issue/view/233> (in Portuguese)

2. **2015 - 2017** — **Student:** Antonio Canabrava Fraideinberze; **Title:** Effective and Unsupervised Fractal-based Feature Selection for Very Large Datasets: removing linear and non-linear attribute correlations; **Funding Agencies:** Coordination for the Improvement of Higher Education Personnel - CAPES (grant); **Microsoft Azure (US\$20,000.00); Amazon (US\$2,000.00); Grant Number (CAPES):** PROEX-7152292/M.
3. **2017 - 2019** — **Student:** Guilherme Queiroz Vasconcelos; **Title:** On the Support of the Similarity-aware Division Operator in a Relational Database Management System; **Funding Agency:** Coordination for the Improvement of Higher Education Personnel - CAPES; **Grant Number:** PROEX-9254542/M;
4. **2017 - 2019** — **Student:** Altamir Gomes Bispo Junior; **Title:** Fast and Scalable Outlier Detection with Metric Access Methods; **Funding Agency:** Brazilian National Council for Scientific and Technological Development - CNPq; **Grant Number:** 132861/2017-8.
5. **2017 - 2020** — **Student:** Leonardo Mauro Pereira Moraes; **Title:** Modeling and Analyzing Social Networks of Games; **Funding Agencies:** Brazilian National Council for Scientific and Technological Development - CNPq (grant); **Amazon (US\$8,000.00); Grant Number (CNPq):** 166889/2017-2.
6. **2017 - 2020** — **Student:** Jadson José Monteiro Oliveira; **Title:** Unsupervised Dimensionality Reduction in Big Data via Massive Parallel Processing with MapReduce and Resilient Distributed Datasets; **Funding Agencies:** Brazilian National Council for Scientific and Technological Development - CNPq (grant); **Microsoft Azure (US\$20,000.00); Amazon (US\$19,000.00); Grant Number (CNPq):** 166887/2017-0.
7. **2018 - 2021** — **Student:** Eugênio Ferreira Cabral; **Title:** Fast Outlier Detection Using Similarity Self-Join Techniques; **Funding Agency:** Brazilian National Council for Scientific and Technological Development - CNPq; **Grant Number:** 132788/2018-7.
8. **2019 - 2021** — **Student:** Alexis Iván Aspauza Lescano; **Title:** Relational Conditional Set Operations; **Funding Agency:** Coordination for the Improvement of Higher Education Personnel - CAPES; **Grant Number:** PROEX-9249525/M.
9. **2019 - 2022** — **Student:** Daniela Fernanda Milón Flores; **Title:** How to Take Advantage of Behavioral Features for the Early Detection of Grooming in Online Conversations; **Funding Agency:** Brazilian National Council for Scientific and Technological Development - CNPq; **Grant Number:** 130384/2020-8.
10. **2019 - 2022** — **Student:** Guilherme Domingos Faria Silva; **Title:** C-AllOut: Catching and Calling Outliers by Type; **Funding Agency:** Coordination for the Improvement of Higher Education Personnel - CAPES; **Grant Number:** PROEX-9361094/M.

#### ADVISING UNDERGRADUATE STUDENTS

- **Final Year Project Dissertation:** 8 students directed.
- **Undergraduate Research Work:** 7 students directed.

#### QUANTITATIVE INDICATORS

1) Books Published	1
2) Books Edited	1
3) Book Chapters Published	6
4) Referred Journal Publications	10
5) Referred Conference Publications	22
4) Concluded M.Sc. Theses Directed	10
5) Concluded Ph.D. Dissertations Directed	0
6) Number of Citations <sup>16</sup>	472

#### AWARDS

1. **Best undergrad student award:** first prize as the best student among graduates of 2002 in the Bachelor's Degree in Computer Science at University of Western São Paulo - UNOESTE, Brazil.

<sup>16</sup> According to Google Scholar. See <https://scholar.google.com/citations?user=bDCcqZ8AAAAJ&hl=en>

2. **Best CS Dissertation award:** Doctoral Dissertation awarded the first prize as the best in computing of 2011 by the Instituto de Ciências Matemáticas e de Computação - ICMC at University of São Paulo - USP. ICMC is one of the most important Brazilian institutions in Computer Science, Mathematics, Applied Mathematics and Statistics.
3. **Best CS Dissertation award:** first prize in the 2012 edition of the largest and most renowned Brazilian competition for theses and dissertations in computing. Through this competition, the **Brazilian Computer Society - SBC** annually analyzes and recognizes the best works generated in the country in all sub-areas of Computer Science. Link: <http://horizontes.sbc.org.br/old/doku.php?id=v05n02:9> (in Portuguese)
4. **Notable book award:** R. L. F. Cordeiro, C. Faloutsos, and C. Traina Jr. Data Mining in Large Sets of Complex Data. Springer. Book listed among the “Notable Computing Books and Articles of 2013” by the **Association for Computing Machinery - ACM**, which is the world’s largest computing society with nearly 100,000 members all across the world. Link: [http://computingreviews.com/recommend/bestof/notableitems\\_2013.cfm](http://computingreviews.com/recommend/bestof/notableitems_2013.cfm)
5. **Best paper award:** I. R. V. Pola, R. L. F. Cordeiro, C. Traina Jr., and A. J. M. Traina. A New Concept of Sets to Handle Similarity in Databases: the SimSets. In 6<sup>th</sup> International Conference on Similarity Search and Applications - SISAP 2013. Paper selected as one of the best papers of the conference, and invited for a journal publication.
6. **Best CS Thesis award (as advisor):** one M.Sc. work developed under my guidance and supervision by André dos Santos Gonzaga was selected among the nine best theses in the 2018 edition of the largest and most renowned Brazilian competition for theses and dissertations in computing. Through this competition, the **Brazilian Computer Society - SBC** annually analyzes and recognizes the best works generated in the country in all sub-areas of Computer Science. Link: [http://www2.sbc.org.br/csbc2018/?page\\_id=191](http://www2.sbc.org.br/csbc2018/?page_id=191) (in Portuguese)
7. **Best paper award:** L. M. P. Moraes, and R. L. F. Cordeiro. Detecting Influencers in Very Large Social Networks of Games. In 21<sup>st</sup> International Conference on Enterprise Information Systems - ICEIS 2019. Paper selected as one of the best papers of the conference, and invited for a journal publication.
8. **Award for teaching excellence:** São Carlos School of Engineering, University of São Paulo - USP, 2020.
9. **Award for teaching excellence:** São Carlos School of Engineering, University of São Paulo - USP, 2021.

## IN THE NEWS

1. **Aug. 2012** — News about the “*best CS Dissertation award*”:
  - a. News article in the website of University of Western São Paulo; Local audience; Link: <https://www.unoeste.br/noticias/2012/8/egresso-tem-melhor-doutorado-do-pais-na-area-computacional> (in Portuguese)
  - b. Article in newspaper O Imparcial; Local audience. (in Portuguese)
2. **Jul. 2013** — *Book review* published by ACM Computing Reviews. **Global audience**, in potential: nearly 100,000 members of ACM all across the world. Link: [https://www.computingreviews.com/review/review\\_review.cfm?review\\_id=141412](https://www.computingreviews.com/review/review_review.cfm?review_id=141412)
3. **May 2017** — ORFEL algorithm to *detect defamation and illegitimate promotion*:
  - a. News article in the website of University of São Paulo, campus of São Carlos; Local audience; Link: <http://www.saocarlos.usp.br/software-detecta-difamacao-fraudulenta-em-sites-de-comercio-eletronico/> (in Portuguese)
  - b. News article in magazine Exame; **National audience** (Brazil), in potential: Exame is one of the leading business magazines in Brazil; Link: <https://exame.com/tecnologia/software-detecta-difamacao-fraudulenta-no-comercio-eletronico/> (in Portuguese)
  - c. News article in portal Agência FAPESP; Regional audience, in potential: Agência FAPESP is a well-known research-related portal in the Brazilian state of São Paulo, which attracts millions of views per year; Link: <https://agencia.fapesp.br/software-detecta-difamacao-fraudulenta-em-sites-de-comercio-eletronico/25235/> (in Portuguese)
  - d. News article in portal Periódicos CAPES; **National audience** (Brazil), in potential: Periódicos CAPES is a well-known research-related portal in Brazil, which attracts millions of views per year; Link: [http://www-periodicos-capes.gov.br/ez67.periodicos.capes.gov.br/?option=com\\_pnews&component=NewsShow&view=pnewsnewsshows&cid=499&mn=71](http://www-periodicos-capes.gov.br/ez67.periodicos.capes.gov.br/?option=com_pnews&component=NewsShow&view=pnewsnewsshows&cid=499&mn=71) (in Portuguese)
4. **Aug. 2021** — Analysis of the *performance of students* from all regions of Brazil:
  - a. Article in newspaper Jornal da USP; Regional audience, in potential: Jornal da USP reaches all 8 campuses of USP spreading all over the Brazilian state of São Paulo, including nearly 100,000 students, 5,000 faculty members and 13,000 staff members; Link: <https://jornal.usp.br/ciencias/dados-do-enem-podem-contribuir-para-a-construcao-de-futuras-politicas-educacionais-aponta-estudo/> (in Portuguese)
  - b. Interview in radio CBN São Carlos; Local audience; Link: <https://www.cbnsaocarlos.com.br/noticias/som.0.1.88160.pesquisa-do-icmc-da-usp-de-sao-carlos-analisou-os-numeros-do-enem-e-o-que-eles-revelam-sobre-o-ensino-do-pais.aspx#> (in Portuguese)

5. **Feb./Mar. 2022** — BF-PSR *sexual grooming detection* algorithm:

- a. Article in newspaper Jornal da USP; Regional audience, in potential: Jornal da USP reaches all 8 campuses of USP spreading all over the Brazilian state of São Paulo, including nearly 100,000 students, 5,000 faculty members and 13,000 staff members; Link: <https://jornal.usp.br/ciencias/ algoritmo-monitora-conversas-on-line-de-criancas-e-adolescentes-e-detecta-assedio-sexual/> (in Portuguese)
- b. News article in the magazine Crescer; **National audience** (Brazil), in potential: Crescer is a magazine from Grupo Globo, which is the largest mass media group of Latin America; Link: <https://revistacrescer.globo.com/Educacao-Comportamento/noticia/2022/02/algoritmo-desenvolvido-pela-usp-pode-ajudar-detectar-casos-de-assedio-sexual-online-contra-criancas.html> (in Portuguese)
- c. Reportage in TV newscast Jornal da EPTV; Regional audience, in potential: Jornal da EPTV reaches the Brazilian cities of São Carlos, Araraquara, Campinas, and Piracicaba; it is a newscast from Grupo Globo, which is the largest mass media group of Latin America.
  - i. Link for Campinas/Piracicaba: <https://globoplay.globo.com/v/10332896/> (in Portuguese)
  - ii. Link for São Carlos/Araraquara: <https://g1.globo.com/sp/sao-carlos-regiao/noticia/2022/02/25/sistema-que-monitora-conversa-online-de-menores-e-alerta-pais-sobre-pedofilia-e-criado-na-usp.ghtml> (in Portuguese)
- d. Reportage in TV newscast Jornal da Gazeta; **National audience** (Brazil), in potential: Jornal da Gazeta is a well-known TV newscast in Brazil, which attracts nearly one million views every weekday; Link: [https://youtu.be/2xO\\_6-a6o6s](https://youtu.be/2xO_6-a6o6s) (in Portuguese)
- e. Interview in radio CBN Vale; Local audience; Link: <https://www.cbnvale.com.br/pesquisadores-da-usp-desenvolvem-algoritmo-que-combate-assedio-sexual-contra-criancas/> (in Portuguese)

**COURSES TAUGHT**

At UNOESTE - UNDERGRADUATE PROGRAMS (usually, one course has around 50 hours in class per semester)

Undergraduate Program	Course — Semester(s)
B.Sc. in Information Systems	Data Structures II — 2004/1 <sup>st</sup> ; 2005/1 <sup>st</sup> ; 2006/1 <sup>st</sup> ; 2006/2 <sup>nd</sup> Algorithms and Programming Techniques I — 2004/1 <sup>st</sup> ; 2006/2 <sup>nd</sup> Computational Tools I — 2004/1 <sup>st</sup> ; 2004/2 <sup>nd</sup> ; 2005/1 <sup>st</sup> ; 2005/2 <sup>nd</sup> ; 2006/1 <sup>st</sup> ; 2006/2 <sup>nd</sup> Computational Tools II — 2004/1 <sup>st</sup> ; 2004/2 <sup>nd</sup> ; 2005/1 <sup>st</sup> ; 2005/2 <sup>nd</sup> ; 2006/2 <sup>nd</sup> Programming Languages II — 2004/2 <sup>nd</sup> ; 2005/1 <sup>st</sup> ; 2005/2 <sup>nd</sup> ; 2006/1 <sup>st</sup> ; 2006/2 <sup>nd</sup> Databases I — 2005/1 <sup>st</sup> ; 2005/2 <sup>nd</sup> ; 2006/1 <sup>st</sup> ; 2006/2 <sup>nd</sup> Supervised Internship II — 2005/1 <sup>st</sup> ; 2006/1 <sup>st</sup> Supervised Internship I — 2005/1 <sup>st</sup> ; 2006/1 <sup>st</sup> Data Structures I — 2006/1 <sup>st</sup> Computer Laboratory II — 2006/1 <sup>st</sup>
Technology in Informatics	Introduction to Databases — 2004/1 <sup>st</sup>
B.Sc. in Computer Science	Databases I — 2004/2 <sup>nd</sup> Software Engineering II — 2004/2 <sup>nd</sup> Supervised Internship II — 2005/1 <sup>st</sup> ; 2005/2 <sup>nd</sup> ; 2006/2 <sup>nd</sup> Final Year Project Dissertation I — 2005/1 <sup>st</sup> Final Year Project Dissertation II — 2005/2 <sup>nd</sup>
Technology in Web Development	Databases — 2005/1 <sup>st</sup> ; 2005/2 <sup>nd</sup> Supervised Internship II — 2005/2 <sup>nd</sup>



At USP - UNDERGRADUATE PROGRAMS (usually, one course has around 50 hours in class per semester)

Undergraduate Program	Course — Semester(s)
B.Sc. in Computer Science	Theory of Computation and Formal Languages — 2013/2 <sup>nd</sup> Final Year Project Dissertation and Supervised Internship — 2013/2 <sup>nd</sup> ; 2016/1 <sup>st</sup> ; 2018/1 <sup>st</sup> ; 2019/1 <sup>st</sup> ; 2019/2 <sup>nd</sup> ; 2020/1 <sup>st</sup> Databases Laboratory — 2014/2 <sup>nd</sup> Databases — 2015/1 <sup>st</sup> ; 2017/1 <sup>st</sup> File Organization — 2016/1 <sup>st</sup> Algorithms and Data Structures I — 2018/2 <sup>nd</sup> ; 2019/2 <sup>nd</sup> Seminars in Computing — 2021/2 <sup>nd</sup> ; 2022/2 <sup>nd</sup>
B.Sc. in Information Systems	Topics in Databases — 2013/2 <sup>nd</sup> Algorithms and Data Structures II — 2014/1 <sup>st</sup> ; 2016/1 <sup>st</sup> Databases Laboratory — 2014/1 <sup>st</sup> Databases — 2015/2 <sup>nd</sup> ; 2016/2 <sup>nd</sup> ; 2017/2 <sup>nd</sup> ; 2020/2 <sup>nd</sup> Algorithms and Data Structures I — 2018/2 <sup>nd</sup>
Computer Engineering	Algorithms and Data Structures I — 2014/2 <sup>nd</sup> Databases — 2015/2 <sup>nd</sup> ; 2016/2 <sup>nd</sup> Data Structures II — 2019/1 <sup>st</sup> ; 2021/1 <sup>st</sup> ; 2022/1 <sup>st</sup> Object-oriented Programming — 2020/2 <sup>nd</sup> ; 2021/2 <sup>nd</sup> ; 2022/2 <sup>nd</sup>
Civil Engineering	Introduction to Programming for Engineering Courses — 2018/1 <sup>st</sup> ; 2020/1 <sup>st</sup>
B. Sc. in Applied Mathematics and Scientific Computing	Data Structures II — 2019/1 <sup>st</sup> ; 2021/1 <sup>st</sup> ; 2022/1 <sup>st</sup> Object-oriented Programming — 2020/1 <sup>st</sup>

At USP - GRADUATE PROGRAM (each course has 45 hours in class per semester)

Ph.D. and M.Sc. Graduate Program	Course — Semester(s)
Computer Science and Computational Mathematics	Theory and Techniques for the Construction of Relational Database Management Systems — 2017/2 <sup>nd</sup> Multimedia Databases and Image Retrieval — 2018/1 <sup>st</sup> Database Topics (Knowledge Discovery in Large Databases) — 2019/2 <sup>nd</sup> Development of Complex Data Managers — 2022/2 <sup>nd</sup>

## PROFESSIONAL SERVICE

### EDITORSHIP AND ADVISORY BOARD

- 2015 - 2017: **Associate Editor** of the Journal of Information and Data Management - JIDM, published by the Brazilian Computer Society - SBC. Link (first issue): <https://periodicos.ufmg.br/index.php/jidm/article/view/320/225>. Link (last issue): <https://periodicos.ufmg.br/index.php/jidm/article/view/394/276>
- 2018 - 2018: Member of the **PIPE Advisory Board**, The São Paulo Research Foundation - FAPESP, Brazil.

### EVALUATION OF RESEARCH PROJECT PROPOSALS

- 2017 - : Referee for **3 (three)** research project proposals submitted to the São Paulo Research Foundation - FAPESP, Brazil.

### DEFENSE COMMITTEES

- Member in **3 (three)** Ph.D. defense committees.

- Member in **20 (twenty)** M.Sc. defense committees.
- Member in **99 (ninety-nine)** defense committees of Final Year Project Dissertations regarding various Undergraduate Programs.

#### CONFERENCE ORGANIZATION

1. Publications Chair, International Conference on Similarity Search and Applications - SISAP, October 29-31, 2014. Los Cabos, Mexico. Link: <http://www.sisap.org/2014/organization.html>
2. Local Chair, IEEE International Symposium on Computer-Based Medical Systems - CBMS, June 22-25, 2015. São Carlos/Ribeirão Preto, Brazil. Link: <http://www.gbdi.icmc.usp.br/events/cbms2015/organization.html>
3. Demonstrations and Applications Chair, Brazilian Symposium on Databases - SBBD, October 7-10, 2019. Fortaleza, Brazil. Link: <https://sbbd.org.br/2019/organization-committee/>

#### CONFERENCE PROGRAM COMMITTEE MEMBER

- **ECML-PKDD**; **SIAM SDM**; ACM SAC; IEEE ISM; SISAP; SITIS; BRACIS; SBBD.

#### JOURNAL REVIEWS

- **VLDB**; **IEEE TKDE**; **ACM TKDD**; Elsevier Information Sciences; Elsevier Information Systems.

#### CONFERENCE REVIEWS

- **ACM SIGKDD**; **ACM SIGMOD**; **SIAM SDM**; **ECML-PKDD**; **CIKM**; SSDBM; SISAP; ACM SAC; ACM GIS; SITIS; BRACIS; SBBD.

#### UNIVERSITY SERVICES OTHER THAN TEACHING

##### **VICE COORDINATION** OF UNDERGRADUATE PROGRAM

1. 2022 - : B.Sc. in Data Science; USP.

#### MEMBER

1. 2015 - 2021 : Board of the Computer Science Department; USP.
2. 2016 - : Coordination Committee of Undergraduate Program; B.Sc. in Information Systems; USP.
3. 2016 - 2018: Coordination Committee of Graduate Program; Professional Master's Course in Applied Mathematics, Statistics and Computing Applied to Industry; USP.
4. 2017 - : Research Committee; USP.
5. 2022 - : Committee of International Relations; USP.
6. 2022 - : Coordination Committee of Undergraduate Program; B.Sc. in Data Science; USP.

#### ALTERNATE MEMBER

1. 2019 - 2021 : Coordination Committee of Undergraduate Program; B.Sc. in Computer Science; USP.
2. 2019 - 2021: Coordination Committee of Graduate Program; Professional Master's Course in Applied Mathematics, Statistics and Computing Applied to Industry; USP.
3. 2020 - 2022 : Committee of International Relations; USP.
4. 2020 - 2021 : Coordination Committee of Undergraduate Program; B.Sc. in Data Science; USP.
5. 2021 - : Coordination Committee of Graduate Program; Ph.D. and M.Sc. Graduate Program in Computer Science and Computational Mathematics; USP.

##### **MEMBERSHIP IN PROFESSIONAL ASSOCIATIONS**

1. Member of the Institute of Electrical and Electronic Engineers - IEEE.
2. Member of the Association for Computing Machinery - ACM.
3. Member of the Society for Industrial and Applied Mathematics - SIAM.
4. Member of the Brazilian Computer Society - SBC.