





# Renan Souza

✉ [rfsouza@br.ibm.com](mailto:rfsouza@br.ibm.com) •  [RenanSouza.org](https://github.com/RenanSouza.org) •  [in renansouza1](https://www.linkedin.com/in/renansouza1)  
 [Renan Souza](#) •  [renan-souza](#) • Generated on May 12, 2020

## Bio

---

Renan Francisco Santos Souza holds a Ph.D. (2019) and a M.Sc. (2015) in Computer Science from [COPPE/Federal University of Rio de Janeiro \(UFRJ\)](#), and a [B.Sc. in Computer Science](#) from UFRJ (2009-2013). Since 2015, he works at [IBM Research Brazil](#), where he is a Research Scientist in the [Industrial Cloud Technologies](#) group. He has been working both as a software engineer and a researcher in several projects since 2010 and has been actively publishing scientific papers in international conferences and referred journals since 2014. During his B.Sc., he spent a school year at [Missouri State University](#) and did a summer internship at [Stanford University](#) in the [SLAC](#) National Laboratory. During his Ph.D., he was a visiting researcher at [Inria/Univ. Montpellier](#) in France in 2019. In 2017, he won the best M.Sc. thesis award from SBBD, the main conference on data management in Latin America. He does research on large-scale data science and engineering for the support of Artificial Intelligence systems.

## Research Interests

---

Large-scale Data Science and Engineering • Parallel Workflows • Data Provenance • Big Data Analytics • High Performance Computing in Clusters and Clouds • Machine Learning •

## Education

---

- Ph.D. in Computer Science, Federal Univ. of Rio de Janeiro, Brazil Sep 2015 – Dec 2019  
Supervised by [Marta Mattoso](#) and [Patrick Valduriez](#)  
Title: [Supporting User Steering in Large-scale Workflows with Provenance Data](#)
- Visiting Ph.D. Student, Inria/Univ. Montpellier, France Jan 2019 – Mar 2019  
Supervised by [Patrick Valduriez](#)
- M.Sc. in Computer Science, Federal Univ. of Rio de Janeiro, Brazil Jan 2013 – Jul 2015  
Supervised by [Marta Mattoso](#)  
Title: [Controlling the Parallel Execution of Workflows Relying on a Distributed Database](#)
- Computer Science exchange student, Missouri State University, U.S. Jun 2011 – Jun 2012
- B.Sc. in Computer Science, Federal Univ. of Rio de Janeiro, Brazil Jan 2009 – Dec 2012  
Supervised by [Maria Luiza Machado Campos](#)  
Title: [Linked Open Data Publication Strategies: An Application in Network Performance Data \(in pt\)](#)
- Technical Degree in Information Systems, Lemos de Castro Jan 2005 – Dec 2007

## Experience

---

- IBM Research, Research Scientist Aug 2019 – present
- IBM Research, Research Engineer Sep 2015 – Aug 2019
- IBM Research, Software Engineer Intern Apr 2015 – Sep 2015
- Stanford University, SLAC, Research Collaborator Aug 2013 – Dec 2014
- Stanford University, SLAC, Research Intern May 2013 – Aug 2013
- CAPGov - Government Technologies, Leading Software Engineer Dec 2013 – Sep 2014

- CAPGov - Government Technologies, Software Engineer Sep 2013 – Jan 2014
- CAPGov - Government Technologies, Software Engineer Intern Jan 2011 – Jun 2012
- Federal Univ. of Rio de Janeiro, Software Engineer Intern Jan 2010 – Jul 2011

## Selected Publications

For complete list, visit: [RenanSouza.org/publications](https://RenanSouza.org/publications)

- [1] R. Souza, L. Azevedo, V. Lourenço, E. Soares, R. Thiago, R. Brandão, D. Civitarese, E. Vital Brazil, M. Moreno, P. Valduriez, M. Mattoso, R. Cerqueira, M. A. S. Netto, "Provenance data in the machine learning lifecycle in computational science and engineering," in *Workflows in Support of Large-Scale Science (WORKS) co-located with the ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis (SC)*, 2019, pp. 1–10. DOI: [10.1109/WORKS49585.2019.00006](https://doi.org/10.1109/WORKS49585.2019.00006). [Online]. Available: <https://arxiv.org/pdf/1910.04223>.
- [2] R. Souza, L. Azevedo, R. Thiago, E. Soares, M. Nery, M. Netto, E. V. Brazil, R. Cerqueira, P. Valduriez, M. Mattoso, "Efficient runtime capture of multiworkflow data using provenance," in *IEEE International Conference on e-Science (eScience)*, 2019, pp. 1–10. DOI: [10.1109/eScience.2019.00047](https://doi.org/10.1109/eScience.2019.00047). [Online]. Available: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-02265932>.
- [3] R. Souza, V. Silva, J. J. Camata, A. L. G. A. Coutinho, P. Valduriez, M. Mattoso, "Keeping track of user steering actions in dynamic workflows," *Future Generation Computer Systems*, vol. 99, pp. 624–643, 2019, ISSN: 0167-739X. DOI: [10.1016/j.future.2019.05.011](https://doi.org/10.1016/j.future.2019.05.011). [Online]. Available: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-02127456>.
- [4] R. Souza, V. Silva, A. L. G. A. Coutinho, P. Valduriez, M. Mattoso, "Data reduction in scientific workflows using provenance monitoring and user steering," *Future Generation Computer Systems*, vol. online, pp. 1–34, 2017, ISSN: 0167-739X. DOI: [10.1016/j.future.2017.11.028](https://doi.org/10.1016/j.future.2017.11.028). [Online]. Available: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-01679967/document>.