

Yakub Sebastian  
Research Active Lecturer  
College of Engineering, IT and Environment



## Research interests

Dr Yakub Sebastian's main research interest is in data science, information retrieval, text mining, and network theory. He is an active advocate of the field Literature-Based Discovery, a specific branch of information retrieval whose goal is to improve computational approaches to hypothesis generation and novel knowledge discovery from multi-modal data sources. Aside from these, he shows interest in scientometrics and science mapping. Dr Sebastian was awarded Doctor of Philosophy in Computer Science by Monash University and is currently a Guest Associate Editor for journal *Frontiers in Research Metrics and Analytics* (Text Mining and Literature-Based Discovery section).

## Employment

### Lecturer - Information Technology

Research Active Lecturer  
College of Engineering, IT and Environment  
30 Mar 2021 → present

### Adjunct Research Fellow

Swinburne University of Technology Sarawak Campus  
Sarawak, Malaysia  
23 Mar 2020 → 22 Mar 2022

## Research outputs

### The boundary-spanning mechanisms of Nobel Prize winning papers

Sebastian, Y. & Chen, C., Aug 2021, In: *PLoS One*. 16, 8, p. 1-31 31 p., e0254744.

### Evaluation of the Neural-network-based Method to Discover Sets and Representatives of Nonlinearly Dependent Variables

Ohsaki, M., Sasaki, H., Kishimoto, N., Katagiri, S., Ohnishi, K., Sebastian, Y. & Then, P. H. H., 8 Jun 2021, *2021 5th IEEE International Conference on Cybernetics (CYBCONF)*. Sendai, Japan: IEEE, Institute of Electrical and Electronics Engineers, p. 101-106 6 p. 9464151. (2021 5th IEEE International Conference on Cybernetics, CYBCONF 2021).

### Accelerating innovations through literature-based discovery: Qualitative analysis of an early usability evaluation

Phang, C. S. J., Vong, W. T., Sebastian, Y., Raman, V., Tiong, X. T. & Then, P. H. H., 23 Feb 2021, *Proceedings of 2020 IEEE 8th R10 Humanitarian Technology Conference, R10-HTC 2020*. IEEE, Institute of Electrical and Electronics Engineers, p. 1-6 6 p. 9356968. (IEEE Region 10 Humanitarian Technology Conference, R10-HTC; vol. 2020-December).

### A graph-theoretic approach for the detection of phishing webpages

Tan, C. L., Chiew, K. L., Yong, K. S. C., Sze, S. N., Abdullah, J. & Sebastian, Y., Aug 2020, In: *Computers and Security*. 95, p. 1-14 14 p., 101793.

### Design of the synthetic multivariate coefficient of variation chart based on the median run length

Lee, M. H., Lim, V. Y. C., Chew, X. Y., Lau, M. F., Yakub, S. & Then, P. H. H., 2020, In: *Advances in Mathematics: Scientific Journal*. 9, 9, p. 7397-7406 10 p., 9.

### Design of IoT monitoring architecture for smart organic farming in Sarawak state of Malaysian Borneo

Jo, R. S., Raman, V., Lu, M., Sebastian, Y., Cameron, S. J., Ten, J. Y. H., Chong, T. K. S., Yap, F. L., Chiew, V., Yeo, D., Koay, R. K. M., Bong, B. S. C., Chew, K. T., Lim, V. Y. C. & Then, P. H., 2019, *SDPS 2019 Proceedings - 24th International Conference: Changing the World via Innovative Design and Applications of Automation and Artificial*

*Intelligence*. Jololian, L., Guldal, S. & Lipscomb, M. M. (eds.). p. 120-124 5 p.

**GLADIATER: An algorithm to support anomaly detection from a longitudinal cardiovascular and diabetes dataset**

Chew, J. T., Sebastian, Y., Raman, V., Tiong, X. T., Fong, A. Y. Y. & Then, P. H. H., Oct 2018, In: *Journal of the American College of Cardiology*. 72, 16 (Supplement), p. C244-C245 2 p., GW29-e0830.

**Anomaly detection from diabetes similarity graphs using community detection and Bayesian techniques**

Sebastian, Y., Chew, J. T., Tiong, X. T., Raman, V., Fong, A. Y. Y. & Then, P. H. H., 5 Jan 2018, *Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication, IMCOM 2018*. Association for Computing Machinery (ACM), 9 p. 32. (ACM International Conference Proceeding Series).

**Advances in Diabetes Analytics from Clinical and Machine Learning Perspectives**

Sebastian, Y., Tiong, X. T., Raman, V., Fong, A. Y. Y. & Then, P. H. H., Oct 2017, In: *International Journal of Design, Analysis and Tools for Integrated Circuits and Systems*. 6, 1, p. 32-37 6 p.

**Literature-Based Discovery by Learning Heterogeneous Bibliographic Information Networks**

Sebastian, Y., Jun 2017, In: *SIGIR Forum*. 51, 1, p. 75-76 2 p.

**Emerging approaches in literature-based discovery: techniques and performance review**

Sebastian, Y., Siew, E. G. & Orimaye, S. O., 16 May 2017, In: *Knowledge Engineering Review*. 32, p. 1-35 35 p., e12.

**Learning the heterogeneous bibliographic information network for literature-based discovery**

Sebastian, Y., Siew, E. G. & Orimaye, S. O., 1 Jan 2017, In: *Knowledge-Based Systems*. 115, p. 66-79 14 p.

**Predicting future links between disjoint research areas using heterogeneous bibliographic information network**

Sebastian, Y., Siew, E. G. & Orimaye, S. O., 1 Jan 2015, *Advances in Knowledge Discovery and Data Mining - 19th Pacific-Asia Conference, PAKDD 2015, Proceedings*. Cao, T., Lim, E-P., Ho, T-B., Zhou, Z-H., Motoda, H. & Cheung, D. (eds.). Springer-Verlag London Ltd., p. 610-621 12 p. (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); vol. 9078).

**Domain-driven KDD for mining functionally novel rules and linking disjoint medical hypotheses**

Sebastian, Y. & Then, P. H. H., Jul 2011, In: *Knowledge-Based Systems*. 24, 5, p. 609-620 12 p.

**A paradigm shift: Combined literature and ontology-driven data mining for discovering novel relations in biomedical domain**

Sebastian, Y., Loh, B. C. S. & Then, P. H. H., 1 Dec 2009, *ICDM Workshops 2009 - IEEE International Conference on Data Mining*. p. 51-57 7 p. 5360528. (ICDM Workshops 2009 - IEEE International Conference on Data Mining).

**Towards natural language interface framework for automated medical analysis**

Sebastian, Y., Loh, B. C. S. & Then, P. H. H., 12 Nov 2009, *Proceedings - 2009 2nd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2009*. p. 5-9 5 p. 5234782. (Proceedings - 2009 2nd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2009).

**Cardiovascular diagnostics mining on Borneo island: From labour-intensive to prototype**

Then, P. H. H., Fong, A. Y. Y., Sebastian, Y., Raman, V., Wong, Y. H. C. & Sim, K. H., 17 Aug 2009, In: *Informatics for Health and Social Care*. 34, 1, p. 1-9 9 p.

**Software engineer's approach towards cardiac knowledge discovery**

Sebastian, Y., Then, H. H. P., Raman, V., Wong, Y. H., Fong, A. Y. Y. & Sim, K. H., 16 Dec 2008, *Proceedings - International Symposium on Information Technology 2008, ITSIM*. 6 p. 4631644. (Proceedings - International Symposium on Information Technology 2008, ITSIM; vol. 2).

**Impact on student learning from traditional continuous assessment and an e-assessment proposal**

Jacob, S. M., Issac, B. & Sebastian, Y., 1 Dec 2006, *PACIS 2006 - 10th Pacific Asia Conference on Information Systems: ICT and Innovation Economy*. p. 1482-1496 15 p.

## **Activities**

### **International Conference on Advances in ICT for Emerging Regions (Event)**

Yakub Sebastian (Reviewer)

2 Dec 2021 → 3 Dec 2021

### **PLoS One (Journal)**

Yakub Sebastian (Reviewer)

19 Jul 2021

### **Journal of Information Science (Journal)**

Yakub Sebastian (Reviewer)

16 Jun 2021

### **The 12th International Conference on Information Technology in Asia 2021 (Event)**

Yakub Sebastian (Reviewer)

9 May 2021

### **Frontiers in Research Metrics and Analytics (Journal)**

Yakub Sebastian (Reviewer)

May 2021