

## Profile

**Name:** Medari Janai Tham  
**Designation:** Associate Professor  
**Department:** Computer Science  
**Institution:** St. Anthony's College, Shillong  
**Email:** medaritham16@gmail.com

1. **Educational Qualification:** MCA; NET (2007); Ph.D.
2. **Research Interest:** Natural Language Processing
3. **Teaching Experience:**
  - i. Post Graduate (MCA) : 14 years (concurrent with UG)
  - ii. Under Graduate (BSc, BA) : 19 years
4. **Book:** Ka Grammar Khasi: Da ka Jingdro. Macmillan Education, India. 2018  
ISBN: 978-93-88296-32-8  
Companion Website: <https://grammarkhasi.in>
5. **Software Tools Developed:** Khasi Shallow Parser and Khasi POS Tagger  
Website: <https://medaritham.pythonanywhere.com>
6. **Papers Published/Presented:**
  - i. "NLP Tools for Khasi, a low resourced language" Proceedings of the 17th International Conference on Natural Language Processing (ICON): System Demonstrations, 2020 (ICON2020), NLP Association of India (NLPAI), <https://aclanthology.org/2020.icon-demos.10>, pp.26-27
  - ii. Tham, M. J. (2020): Bidirectional Gated Recurrent Unit For Shallow Parsing. Indian Journal of Computer Science and Engineering (IJCE), Vol. 11 No. 5 Sep-Oct 2020, DOI: 10.21817/indjce/2020/v11i5/201105167
  - iii. Tham, M. J. (2020): A Hybrid POS Tagger for Khasi, an Under Resourced Language. International Journal of Advanced Computer Science and Applications (IJACSA), 11(10), 2020. <http://dx.doi.org/10.14569/IJACSA.2020.0111042>
  - iv. Tham, M. J. (2018): **Khasi Shallow Parser. Proceedings of ICON2018:** 15<sup>th</sup> International Conference on Natural Language Processing. Punjabi University, Patiala India. 15-18 December, 2018. pp. 43-49. Organised by Natural Language Processing Association India (NLPAI), IIIT Hyderabad, Punjabi University, and Linguistic Data Consortium for Indian Languages (LDC-IL).
  - v. Tham, M. J. (2018): **Challenges and Issues in Developing an Annotated Corpus and HMM POS Tagger for Khasi. Proceedings of ICON2018:** 15<sup>th</sup> International Conference on Natural Language Processing. Punjabi University, Patiala India. pp. 10-19. 15-18 December, 2018. Organised by Natural Language Processing Association India (NLPAI), IIIT Hyderabad, Punjabi University, and Linguistic Data Consortium for Indian Languages (LDC-IL).

- vi. Tham, M. J. (2015) : ‘**Experimental Observations on applying BIS standard annotation to Khasi language**’ Proceedings of the International Conference on Computing and Communication Systems , I3CS’15, ISBN: 978-1-4799-5857-01, pp. 157-159; Information Technology Dept., NEHU, Shillong, 9-10 April, 2015
- vii. Tham, M. J. (2013): ‘**Preliminary Investigation Of a Morphological Analyzer and Generator for Khasi**’ Proceedings of the 1st International Conference on Emerging Trends and Applications in Computer Science (Technically sponsored by IEEE), ISBN: 978-1-4673-5249-9, pp. 256-259; Dept. of Computer Science, St. Anthony’s College, Shillong, 13-14 September, 2013
- viii. Tham, M. J. (2012): ‘**Proposed Finite State Morphological Analysis for Nouns and Verbs in Khasi**’, paper presented in the National Conference of the 34<sup>th</sup> All India Conference of Linguists, NEHU, Shillong. 1 -3 November, 2012
- ix. Tham, M. J. (2012): ‘**Design Considerations for developing a Part-of-Speech Tagset for Khasi**’, Proceedings of the 3<sup>rd</sup> National Conference on Emerging Trends and Applications in Computer Science (Technically sponsored by IEEE), ISBN: 978-1-4577-0747-6, pp. 277-279; Dept. of Computer Science, St. Anthony’s College, Shillong, 30-31 March, 2012
- x. Tham, M. J. (2011): ‘**Natural Language Processing (NLP): A means for human-computer interaction using the Khasi language**’, paper presented in National seminar on ICT in Education, NEHU, Shillong, 13-15 September, 2011

## 7. Chapter in Edited Book

- i. Natural Language Processing (NLP): A means for human-computer interaction using the Khasi language, ICT in Education, Volume 1, Edited by Brinda Bazely (2013), EBH Publishers, ISBN :978-93-80261-91-1; pp.45-50

## 8. Research Project:

- i. Minor Project titled “Part-of-Speech Tagger for Khasi”, No.F.5-123/2014-15/MRP/NERO/1860; **Duration-** 2015-17; **Amount-** Rs. 2,40,000/-; **Status-** Completed;

## 9. Workshop Organized/Coordinated:

- i. Organizing Committee member of the 17<sup>th</sup> International Conference on Natural Language Processing, ICON 2020, 18-21 December 2020, IIT Patna, India.
- ii. Coordinator National Seminar on “Towards Standardizing Khasi for Computational Purposes”, 28-29 October 2014, in association with Dept. of Arts and Culture, govt. of Meghalaya, St. Anthony’s College, Shillong.
- iii. Event Coordinator **MeghaBytes 2011**, Meghalaya IT summit, September 2011
- iv. Organizing Committee member ICETACS 2013
- v. Organizing Committee member NCETACS 2010, 2011, 2012, 2018

## 10. Additional Information

- i. Honorary Rosalind Member of London Journals Press, February 2021