

## CURRICULUM VITAE

1. **Name:** ANIMESH MUKHERJEE  

First
Middle
Last/surname (in UPPER CASE)
2. **Date of Birth:** 21/12/1979
3. **Field of Specialization:** a) Network Science, b) Social Media Analysis, c) Human Language Dynamics, d) Natural Language Processing and e) Information Retrieval
4. **Present Designation and Organization:** Associate Professor, Dept. of CSE, IIT Kharagpur
5. **Address** (for correspondence)  
 (a) Official: DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, IIT KHARAGPUR, WEST BENGAL, INDIA - 721302  
 (b) Residential: C1-159, IIT KHARAGPUR CAMPUS, WEST BENGAL, INDIA - 721302

Phone: 03222283472

Fax:

E-Mail address: animeshm@cse.iitkgp.ernet.in

Phone: 03222283473

Mobile: 9433124004

E-mail address: animeshm@gmail.com

6. **Career Profile in terms of (a) Academic Qualifications (Diploma/Bachelor onwards), b) Professional Experience (last five), (c) Awards/Distinctions:**

a)	Degree	University/Institution	Year	Branch/Stream	Class/CGPA/Remarks
	BTech (Hons.)	HIT, VU	2003	CSE	89.8% (Inst. Rank 1)
	Mtech	IIT Kharagpur	2005	CSE	9.43, (Dept. Rank 4)
	PhD	IIT Kharagpur	2009	CSE	

(b)	Designation	Organization, Place	Period (From – To)	Job/Responsibility
	Junior Researcher	ISI Foundation, Italy	2009 – 2011	Research
	Assistant Professor	IIT Kharagpur	2011 – 2016	Teaching/Research
	Associate Professor	IIT Kharagpur	2016 –	Teaching/Research

- c) **National/International Awards, Prizes, Certificates, Fellowships (Name, Year, Agency)**

a) Humboldt Fellowship for experienced researchers (2017 onward)

e) INSA Medal for Young Scientists (2014)

b) ACM Distinguished Speaker (2017 onward)

f) Simons Associate, ICTP, Trieste, Italy (2013 onward)

c) Gandhian Young Technological Innovation Award (GYTI) (2017)

g) INAE Young Engineer Award (2012)

**d) IBM Faculty Award  
(2015)**

**h) ISCA Young Scientist Award (2006)**

---

**8. Number of Major Sponsored/Consultancy Projects completed or in-progress:**

- Leveraging Bipartite Networks to Investigate the Dynamical Properties of Socio-technical Systems, Samsung GRO. (This grant was received through a world-wide highly competitive competition that we participated.)
- New perspectives for computational social science, Exchange of Researchers within the frame of the Executive Program of Scientific and Technological Cooperation, between the Republic of India and the Italian Republic for the years 2012 – 2014, Department of Science & Technology, International Division. (This project was in continuation of my post-doctoral research with Prof. Vittorio Loreto. One of the remarkable achievements of the work is our joint paper in the Proceedings of the National Academy of Sciences, USA).
- ICT for computational social science, Fast Track Scheme for Young Scientists, Department of Science & Technology, Government of India. (As a part of this project we developed various web experiments to study the dynamics of human languages.)
- Post-disaster situation analysis and resource management using delay-tolerant peer-to-peer wireless networks, ITRA, DIT. (This project was obtained through a very competitive selection process. It was a collaborative effort of six institutions including IIT Kharagpur, IIM Kolkata, IEST Shibpur, NIT Durgapur, Kalyani Government Engg. College and Heritage Institute of Technology. This project was massive appreciated because of its direct relevance to various government initiatives. Recently, there has been media coverage by Business Standard ([http://www.business-standard.com/article/current-affairs/human-mobility-powers-wireless-communication-to-aid-disaster-relief-117071500396\\_1.html](http://www.business-standard.com/article/current-affairs/human-mobility-powers-wireless-communication-to-aid-disaster-relief-117071500396_1.html)) and Financial Express (<http://www.financialexpress.com/industry/technology/human-mobility-powers-wireless-communication-for-disaster-relief/764520/>)).
- Won the IMPRINT Challenge Grant for the project Processing the language of social media for post-disaster situation analysis (funding awaited).
- Data driven approaches for inferring opinion dynamics on social networks, Indo-Polish bilateral exchange, DST, India. This is an outcome of the collaborations with Tyll Krueger.
- DST-DAAD PPP 2017 on Large-scale text analytics to identify temporal and media specific scopes of words. This project is an outcome of the continuous successful collaboration with Chris Biemann.
- Unrestricted research grant from MSR India for the KDD 2014 paper.
- Unrestricted research grant from Yahoo! Labs India for the KDD 2014 paper.
- Unrestricted research grant from Flipkart Internet Private Ltd., India to pursue cutting edge research in social recommendation in e-commerce (two times).

## List of publications

### Books/Edited Volumes

1. **Mukherjee, A.**, Choudhury, M., Peruani, F., Ganguly, N., and Mitra, B. (eds.) (2013). Dynamics on and of Complex Networks Volume 2: Applications to Time-varying Dynamical Systems, Birkhauser, Springer, Boston, ISBN: 978-1-4614-6728-1.
2. **Mukherjee, A.**, Choudhury, M., Hassan, S., and Muresan, S. (eds.) (2011). Editorial: Network Models for Cognitive and Social Dynamics of Language, Computer Speech and Language, 25(3), 635--638. The articles in this issue can be found here: Computer Speech and Language, Volume 25, Issue 3 under the head "special issue papers".
3. Ganguly, N., Deutsch, A., and **Mukherjee, A.** (eds.) (2009). Dynamics on and of Complex Networks: Applications to Biology, Computer Science, Economics, and the Social Sciences, Birkhauser, Springer, Boston, ISBN: 978-0-8176-4750-6 <http://www.springer.com/birkhauser/computer+science/book/978-0-8176-4750-6>

### Journals and book chapters (significant publications in bold face)

4. **Chakraborty, T., Dalmia, A., Mukherjee, A. and Ganguly, N. (2017). Metrics for community analysis: A survey. ACM Computing Surveys, 50(4), 54:1-54:37.**
5. Bodych, M., Ganguly, N., Krueger, T., **Mukherjee, A.**, Siegmund-Schultze, R. and Sikdar S. (2017). Threshold-based epidemic dynamics in systems with memory. *Europhysics Letters* 116(4), 48004-p1-7.
6. **Chakraborty, T., Srinivasan, S., Ganguly, N., Mukherjee, A. and Bhowmick, S. (2016). Permanence and Community Structure in Complex Networks. ACM TKDD, 11(2), 14:1--14:34. (ACM's 21<sup>st</sup> Best in Computing).**
7. **Chakraborty, T., Kumar, S., Ganguly, N., Mukherjee, A. and Bhowmick, S. (2016). GenPerm: A Unified Method for Detecting Non-overlapping and Overlapping Communities. IEEE TKDE, 28(8): 2101--2114.**
8. Singh, M., Chakraborty, T., **Mukherjee, A.** and Goyal, P. (2016). Is this conference a top-tier? ConfAssist: An assistive conflict resolution framework for conference categorization. *Journal of Informetrics*, 10(4), 1005--1022.
9. Sikdar, S., Ganguly, N., **Mukherjee, A.** (2016) Time series analysis of temporal networks. *EPJB Topical Issue on Temporal Network Theory and Applications*, 89(1), 1--11.
10. Gupta, A. K., Bhattacharya, I., Banerjee, P. S., Mandal, J. K. and **Mukherjee, A.** (2015) DirMove: direction of movement based routing in DTN architecture for post-disaster scenario. *Wireless Network Journal*, (accepted).
11. Chakraborty, T., Ganguly, N., **Mukherjee, A.** (2015) An author is known by the context she keeps: significance of network motifs in scientific collaborations. *Journal of Social Network Analysis and Mining*, 5:16.
12. **Mitra, S., Mitra, R., Maity, S. K., Riedl, M., Biemann, C., Goyal, P. and Mukherjee, A. (2015). An automatic approach to identify word sense changes in text media across timescales. Journal of Natural Language Engineering, Special Issue on Graphs for NLP, 21(5), 773--798.**
13. **Chakraborty, T., Kumar, S., Goyal, P., Ganguly, N. and Mukherjee, A. (2015). On the categorization of scientific citation profiles in computer sciences. Communications of the ACM, 58(9), 82--90.**
14. Chakraborty, T., Tammana, V., Ganguly, N. and **Mukherjee, A.** (2015). Understanding and Modeling Diverse Scientific Careers of Researchers. *Journal of Informetrics*, 9(1), 69--78.
15. Sur, S., Ganguly, N. and **Mukherjee, A.** (2015). Attack tolerance of correlated time-varying social networks with well-defined communities. *Physica A: Statistical Mechanics and its Applications*, 420, 98--107.
16. **Ganguly, N., Krueger, T., Mukherjee, A. and Saha, S. (2014). Epidemic spreading through direct and indirect interactions. Phys. Rev. E., 90, 032808**
17. Maity, R. R., Mallya, A., **Mukherjee, A.** and Ganguly, N. (2014). Understanding the Correlation of the Properties of Human Movement Patterns. *Advances in Complex Systems*, 17(6), 1450019.
18. **Saha, S., Ganguly, N., Mukherjee, A. and Krueger, T. (2014). Intergroup networks as random threshold graphs. Phys. Rev. E., 89, 042812.**

19. Chakraborty, T., Sikdar, S., Ganguly, N. and **Mukherjee, A.** (2014). A Quantitative Route to the Rise and Fall of Scientific Research. *Journal of Social Network Analysis and Mining*, 4:187.
20. Bhattacharjee, B., Manna, S. S. and **Mukherjee, A.** (2013). Information sharing and sorting in a community. *Phys. Rev. E.*, 87(6), 062808.
21. **Chakraborty, T., Srinivasan, S., Ganguly, N., Bhowmick, S. and Mukherjee, A. (2013). Constant communities in complex networks. Nature Scientific Reports, 3, 1825.**
22. Ganguly, N., Saha, S., Maiti, A., Agarwal, S., Peruani, F. and **Mukherjee, A.** (2013). Effect of attachment strategies on bipartite networks. *European Phys. Jour. B.*, 86, 287.
23. **Maity, S. K., Mukherjee, A., Tria, F. and Loreto, V. (2013). Emergence of fast agreement in an overhearing population: The case of the naming game. Euro. Phys. Lett., 101, 68004.**
24. **Ghosh, S., Saha, S., Srivastava, A., Kuerger, T., Ganguly, N. and Mukherjee, A. (2013). Understanding Evolution of Inter-Group Relationships using Bipartite Networks. IEEE Journal on Selected Areas in Communications, Special Issue on Emerging Technologies in Communications, 31(8), 1--11.**
25. **Mukherjee, A.**, Choudhury, M., Ganguly, N. and Basu, A. (2013). Language Dynamics in the Framework of Complex Networks: A Case Study on Self-organization of the Consonant Inventories. In *Cognitive Aspects of Computational Language Acquisition*, Poibeau, T., Villavicencio, A., Korhonen, A. and Alishahi, A. (eds.), Springer, 51--78, ISBN 978-3-642-31863-4.
26. **Maity, S. K., Venkat, M. T. and Mukherjee, A. (2012). Opinion formation in time-varying social networks: The case of the naming game. Phys. Rev. E, 86, 036110.**
27. **Loreto, V., Mukherjee, A. and Tria, F. (2012). On the origin of the hierarchy of color names. Proceedings of the National Academy of Sciences (PNAS), 109(18), 6819--6824. Full version.** Media Coverage: [PNAS Press Highlights](#), [Scientific American](#), [Live Science](#), [MedicalXpress](#), [The Hindu](#), [Business Line](#), [Times of News](#), [BBC Future](#), [Prospect Magazine \(UK\)](#), [Yahoo! News](#), [Deccan Herald](#), [MSN](#), [Homunculus](#), [Doctordisruption](#), [Galileo Net](#), [Scienza e Tecnica](#), [Chinese EurekAlert](#), [De Standaard](#), [Design by Coffee](#), [Blitz quotidiano](#), [Wired \(Italian\)](#). This work now has a Wikipedia entry in the page for [Color term](#).
28. **Mukherjee, A.**, Loreto, V. and Tria, F. (2012). Why are basic color names "basic"?, *Topical Issue on Language Dynamics, Advances in Complex Systems*, 15(3,4), 1150016 (1--13).
29. Tria, F., **Mukherjee, A.**, Baronchelli, A., Puglisi, A. and Loreto, V. (2011). A fast no-rejection algorithm for the Category Game, *Journal of Computational Science*, 2(4), 316--323, <http://arxiv.org/abs/1012.3583>
30. **Mukherjee, A.**, Choudhury, M. and Ganguly, N. (2011). Understanding how both the partitions of a bipartite network affect its one-mode projection, *Physica A: Statistical Mechanics and its Applications*, 390(20), 3602--3607 <http://arxiv.org/abs/1105.3902>.
31. Loreto, V., Baronchelli, A., **Mukherjee, A.**, Puglisi, A. and Tria, F. (2011). Statistical physics of language dynamics, *Journal of Statistical Mechanics: Theory and Experiment*, P04006.
32. **Mukherjee, A.**, Tria, F., Baronchelli, A., Puglisi, A. and Loreto, V. (2011). Aging in language dynamics, *PLoS ONE* 6(2): e16677.
33. Pal, R., **Mukherjee, A.**, Mitra, P., and Mukherjee, J. (2010). Modelling Visual Saliency Using Degree Centrality, *IET Computer Vision* 4(3), 218--229.
34. **Mukherjee, A.**, Choudhury, M., Basu, A., and Ganguly, N. (2010). Modeling the Redundancy of Human Speech Sound Inventories: An Information Theoretic Approach, *Journal of Quantitative Linguistics*, 17(4), 317--343.
35. **Choudhury, M., Ganguly, N., Maiti, A., Mukherjee, A., Bruschi, L., Deutsch, A., and Peruani, F. (2010). Modeling discrete combinatorial systems as Alphabetic Bipartite Networks: Theory and Applications Phys. Rev. E, 81 036103 <http://arxiv.org/abs/0811.0499>**
36. **Mukherjee, A.**, Choudhury, M., Basu, A., and Ganguly, N. (2009). Self-organization of the Sound Inventories: Analysis and Synthesis of the Occurrence and Co-occurrence Networks of Consonants, *Journal of Quantitative Linguistics*, 16(2), 157--184 <http://arXiv.org/physics/0610120>
37. Choudhury, M., and **Mukherjee, A.** (2009). The Structure and Dynamics of Linguistic Networks. In *Dynamics on and of Complex Networks: Applications to Biology, Computer Science, Economics, and the Social Sciences*, Ganguly, N., Deutsch, A., and Mukherjee, A., (eds.), Birkhauser, Springer, Boston, 145--166, ISBN: 978-0-8176-4750-6 <http://www.springerlink.com/content/x0g78760w753p781/>

38. **Mukherjee, A.**, Choudhury, M., RoyChowdhury, S., Basu, A., and Ganguly, N. (2008). Rediscovering the Co-occurrence Principles of the Vowel Inventories: A Complex Network Approach, *Advances in Complex Systems*, 11(3), 371--392 <http://arxiv.org/abs/physics/0702056>
39. **Mukherjee, A.**, Chakraborty, K., and Basu, A.(2008). SweepSticks - An Adaptive Virtual Mouse for People with Neuro-Motor Disorders, *Assistive Technology Journal of the Rehabilitation Engineering Society of North America*, 20(2), 111—124.
40. **Choudhury, M., Saraf, R., Jain, V., Mukherjee, A., Sarkar, S., and Basu, A. (2007).** Investigation and Modeling of the Structure of Texting Language, *Special Issue of the International Journal of Document Analysis and Recognition on Noisy Text Analytics, Springer*, 10(3-4), 157—174.
41. **Peruani, F., Choudhury, M., Mukherjee, A., and Ganguly, N. (2007).** Emergence of a non-scaling degree distribution in bipartite networks: a numerical and analytical study, *Euro. Phys. Letters*,79(2), 28001 <http://arXiv.org/abs/cond-mat/0703634>
42. **Mukherjee, A.**, Choudhury, M., Basu, A., and Ganguly, N. (2007). Modeling the Co-occurrence Principles of the Consonant Inventories: A Complex Network Approach, *Int. Jour. of Modern Phy. C*, 18(2), 281--295 <http://arXiv.org/physics/0606132>

#### Refereed Conferences/Workshops (including companion volumes)

43. **Sikdar, S., Chakraborty T., Sarkar, S., Ganguly, N. and Mukherjee, A. (2018).** CompAS: Community Preserving Sampling for Streaming Graphs. In *AAMAS (18)*, Stockholm, Sweden.
44. Maity, S. K., Gajula, R. and **Mukherjee, A.** (2018). Why Did They #Unfollow Me? Early Detection of Follower Loss on Twitter. In *ACM GROUP (18)*, (poster highlight), Florida, USA.
45. Maity, S. K., Kumar, A., Choudhary, V., Mallick, A., and **Mukherjee, A.** (2018). Understanding Book Popularity on Goodreads. In *ACM GROUP (18)*, (poster highlight), Florida, USA.
46. **Patro, J., Samanta, B., Singh, S., Basu, A., Mukherjee, P., Choudhury, M., Mukherjee, A. (2017).** All that is English may be Hindi: Enhancing language identification through automatic ranking of the likeliness of word borrowing in social media. In *EMNLP (17)*, Copenhagen, Denmark.
47. Maity, S. K., Panigrahi, A., and **Mukherjee, A.** (2017). Book Reading Behavior on Goodreads Can Predict the Amazon Best Sellers. In *ACM/IEEE ASONAM (17)*, (short paper), Sydney, Australia.
48. **Singh, M., Sarkar, R., Goyal, P., Mukherjee, A. and Chakrabarti, S. (2017).** Relay-Linking Models for Prominence and Obsolescence in Evolving Networks. In *ACM SIGKDD (17)*, Halifax, Canada.
49. **Sikdar, S., Marsili, M., Ganguly, N., Mukherjee, A. (2017)** Influence of Reviewer Interaction Network on Long-term Citations: A Case Study of the Scientific Peer-Review System of the Journal of High Energy Physics. In *ACM/IEEE JCDL (17)*, Toronto, Canada.
50. **Jana, A., Mooriyath, S., Mukherjee, A., and Goyal, P. (2017)** WikiM: Metapaths based Wikification of Scientific Abstracts. In *ACM/IEEE JCDL (17)*, Toronto, Canada.
51. **Singh, M., Niranjana, A., Gupta, D., Bakshi, N. A., Mukherjee, A., and Goyal, P. (2017)** Citation sentence reuse behavior of scientists: A case study on massive bibliographic text dataset of computer science. In *ACM/IEEE JCDL (17)*, Toronto, Canada (short paper).
52. **Singh, M., Jaiswal, A., Priya Shree, Pal, A., Mukherjee, A., and Goyal, P. (2017)** Understanding the Impact of Early Citers on Long-Term Scientific Impact. In *ACM/IEEE JCDL (17)*, Toronto, Canada.
53. **Maity, S. K., Kharb, A. and Mukherjee, A. (2017).** Language Use Matters: Analysis of the Linguistic Structure of Question Texts can Characterize Answerability in Quora In *ICWSM (17)*, (poster paper), Montreal, Canada.
54. **Maity, S. K., Chakraborty, A., Goyal, P. and Mukherjee, A. (2017).** Detection of sockpuppets in social media. In *CSCW (17) poster highlight*, Portland, OR.
55. **Maity, S. K., Jha, C., Kumar, A., Sengupta, A., Modi, M. and Mukherjee, A. (2017).** A Large-scale Analysis of the Marketplace Characteristics in Fiverr, In *HICSS-50*, Hilton Waikoloa Village, Hawaii.
56. **Singh, M., Barua, B., Palod, P., Garg, M., Satapathy, S., Bushi, S., Ayush, K., Rohith, K. S., Gamidi, T., Goyal, P. and Mukherjee, A. (2016)** OCR++: A Robust Framework For Information Extraction from Scholarly Articles. In *COLING (16)*, Osaka, Japan. Live System (This paper received the GYTI 2017 award).

57. Ufimtsev, V., Sarkar, S., **Mukherjee, A.** and Bhowmick, S. (2016). Understanding Stability of Noisy Networks through Centrality Measures and Local Connections, In *CIKM (16)*, (extended short paper), Indianapolis, IN.
58. Sikdar, S., Marsili, M, Ganguly, N. and **Mukherjee, A.** (2016). Anomalies in the peer-review system: A case study of the journal of High Energy Physics, In *CIKM (16)*, (extended short paper), Indianapolis, IN.
59. Sarkar, S., Kumar, S., Bhowmick, S., and **Mukherjee, A.** (2016). Sensitivity and Reliability in Incomplete Networks: Centrality Metrics to Community Scoring Functions, In *ASONAM (16)*, (short paper), San Francisco, CA.
60. **Maity, S. K., Saraf, R. and Mukherjee, A.** (2016). #Bieber + #Blast = #BieberBlast: Early prediction of popular hashtag compounds. In *CSCW (16)*, San Francisco, CA. (CSCW Honorable Mention, MIT Tech Review [best of the rest of arXiv](#). [Synopsis](#). [Other Comments](#).)
61. Chakraborty, T., Krishna, A., Singh, M., Ganguly, N., Goyal, P. and **Mukherjee, A.** (2016). FeRoSA: A Faceted Recommendation System for Scientific Articles. In *PAKDD (16)*, Auckland, New Zealand. ([Working Prototype](#))
62. **Maity, S. K., Ghuku, B., Upmaynu, A. and Mukherjee, A.** (2016). Out of vocabulary words decrease, running texts prevail and hashtags coalesce: Twitter as an evolving sociolinguistic system. In *HICSS (16)*, Kauai, Hawaii. (This paper featured as MIT Tech Review [best of the rest of arXiv](#))
63. Singh, M., Patidar, V., Kumar, S., Chakraborty, T., **Mukherjee, A.** and Goyal, P. (2015). The role of citation context in predicting long-term citation profiles: an experimental study based on a massive bibliographic text dataset. In *CIKM (15)*, Melbourne, Australia.
64. Chakraborty, T., Patranabis, S., Goyal, P. and **Mukherjee, A.** (2015). On the formation of circles in co-authorship networks. In *KDD(15)*, Sydney, Australia.
65. Singh, M., Chakraborty, T., Goyal, P. and **Mukherjee, A.** (2015). ConfAssist: A Conflict resolution framework for assisting the categorization of Computer Science conferences. In *JCDL (15)*, Tennessee, USA (poster).
66. **Maity, S. K., Singh, J. S. and Mukherjee, A.** (2015). Analysis and prediction of question topic popularity in community QA sites: A case study of Quora. In *ICWSM (15)*, Oxford, UK.
67. **Maity, S. K., Gupta, A., Goyal, P. and Mukherjee, A.** (2015). A stratified learning approach for predicting the popularity of Twitter Idioms. In *ICWSM (15)*, Oxford, UK (poster).
68. Sikdar, S., Bodych, M., Maiti, R. R., Paria, B., Ganguly, N., Krueger, T. and **Mukherjee, A.** (2015). On the broadcast of segmented messages in dynamic networks. In *IEEE NetSciCom(15)* (co-located with IEEE INFOCOM), Hong Kong, China.
69. Chakraborty, T., Ganguly, N. and **Mukherjee, A.** (2014). Automatic Classification of Scientific Groups as Productive: An Approach based on Motif Analysis. In *ASONAM(14)*, China.
70. Chakraborty, T., Kumar, S., Goyal, P., Ganguly, N. and **Mukherjee, A.** (2014). Towards a Stratified Learning Approach to Predict Future Citation Counts. In *JCDL(14)*, UK.
71. Chakraborty, T., Srinivasan, S., Ganguly, N., **Mukherjee, A.** and Bhowmick, S. (2014). On the permanence of vertices in network communities. In *KDD(14)*, New York, USA.
72. Chakraborty, T., Tammana, V., Ganguly, N. and **Mukherjee, A.** (2014). Analysis and modeling of lowest unique bid auctions. In *ASE Socialcom(14)*, Stanford, USA.
73. Maity, S. K., Manoj, T. V. and **Mukherjee, A.** (2014). Opinion dynamics in correlated time-varying social networks. In *ASE Socialcom(14)*, Stanford, USA.
74. **Mitra, S., Mitra, R., Riedl, M., Biemann, C., Mukherjee, A. and Goyal, P.** (2014). That's sick dude!: Automatic identification of word sense change across different timescales. In *ACL(14)*, Baltimore, USA.
75. Quasthoff, U., Mitra, R., Mitra, S., Eckart, T., Goldhahn, D., Goyal, P. and **Mukherjee, A.** (2014). Large Web Corpora of High Quality for Indian Languages. *LREC Workshop on Indian Language Data: Resources and Evaluation*, Reykjavik, Iceland (poster).
76. Maity, S. K., Porwal, A. and **Mukherjee, A.** (2013). Understanding how learning affects agreement process in social networks, In *ASE/IEEE Socialcom(13)*, Washington DC, USA.
77. Chakraborty, T., Kumar, S., Reddy, M. D., Kumar, S., Ganguly, N. and **Mukherjee, A.** (2013). Automatic Classification and Analysis of Interdisciplinary Fields in Computer Sciences, In *ASE/IEEE Socialcom(13)*, Washington DC, USA.

78. Chakraborty, T., Sikdar, S., Tammana, V., Ganguly, N. and **Mukherjee, A.** (2013). Computer Science Fields as Ground-truth Communities: Their Impact, Rise and Fall, In IEEE/ACM ASONAM(13), Niagra Falls, Canada.
79. Ghosh, S., Srivastava, H., Agarwal, S. and **Mukherjee, A.** (2013). Runtime delays in Indian railways: Is traffic the cause?, In ACM DEV(13) (poster), Bangalore, India.
80. Maity, S. K. and **Mukherjee, A.** (2012). Understanding how dominance affects the emergence of agreement in a social network: The case of Naming Game, In ASE/IEEE Socialcom(12), Amsterdam, The Netherlands.
81. Saha, S., Ganguly, N., and **Mukherjee, A.** (2012). Information Dissemination Dynamics in Delay Tolerant Network: A Bipartite Network Approach, In ACM MobiOpp(12), Zurich, Switzerland.
82. **Choudhury, M., Chatterjee, D., and Mukherjee, A.** (2010). **Global topology of word co-occurrence networks: Beyond the two-regime power-law**, In the proceedings of COLING(10), Beijing, China
83. **Biemann, C., Choudhury, M., and Mukherjee, A.** (2009). **Syntax is from Mars while Semantics from Venus! Insights from Spectral Analysis of Distributional Similarity Networks**, In the proceedings of ACL(09), (short paper), Suntec, Singapore
84. **Mukherjee, A.**, Choudhury, M., and Kannan, R. (2009). Discovering Global Patterns in Linguistic Networks through Spectral Analysis: A Case Study of the Consonant Inventories, In the proceedings of EACL(09), Athens, Greece <http://arxiv.org/abs/0901.2216>
85. Choudhury, M., **Mukherjee, A.**, Garg, A., Jalan, V., Basu, A., and Ganguly, N. (2009). Language Diversity across the Consonant Inventories: A Study in the Framework of Complex Networks, In the proceedings of EACL(09) workshop on Cognitive Aspects of Computational Language Acquisition, Athens, Greece <http://arxiv.org/abs/0904.1289>
86. **Mukherjee, A.**, Choudhury, M., Basu, A., and Ganguly, N. (2008). **Modeling the Structure and Dynamics of the Consonant Inventories: A Complex Network Approach** , In the proceedings of COLING(08), Manchester, UK <http://arxiv.org/abs/0806.0923>
87. Nath, J., Choudhury, M., **Mukherjee, A.**, Biemann, C., and Ganguly, N. (2008) Unsupervised Parts-of-Speech Induction for Bengali, In the proceedings of LREC, Morocco, Marrakech.
88. **Mukherjee, A.**, Choudhury, M., Basu, A., and Ganguly, N. (2007). **Redundancy Ratio: An Invariant Property of the Consonant Inventories of the World's Languages**, In the proceedings of ACL(07), Prague, Czech Republic.
89. **Mukherjee, A.**, Choudhury, M., Basu, A., and Ganguly, N. (2007). Emergence of Community Structures in Vowel Inventories: An Analysis based on Complex Networks, In the proceedings of ACL-SIGMORPHON9(07), Prague, Czech Republic.
90. Choudhury, M., Thomas, M., **Mukherjee, A.**, Basu, A., and Ganguly, N. (2007). How Difficult is it to Develop a Perfect Spell-checker? A Cross-linguistic Analysis through Complex Network Approach , Proceedings of TEXTGRAPHS-2, HLT/NAACL(07), Rochester, New York <http://arXiv.org/abs/physics/0703198>
91. Choudhury, M., **Mukherjee, A.**, Basu, A., and Ganguly, N. (2006). **Analysis and Synthesis of the Distribution of Consonants over Languages: A Complex Network Approach**, Proceedings of COLING-ACL(06), Sydney, Australia <http://arXiv.org/physics/0606131>
92. **Mukherjee, A.**, Bhattacharya, S., Halder, P.K., and Basu, A. (2005). A Virtual Predictive Keyboard as a Learning Aid for People with Neuro-motor Disorders, Proceedings of the 5th IEEE International Conference on Advanced Learning Technologies, (ICALT - 2005), Kaohsiung, Taiwan.



## Ten most significant publications by Animesh Mukherjee

1. Chakraborty, T., Srinivasan, S., Ganguly, N., Mukherjee, A. and Bhowmick, S. (2016). Permanence and Community Structure in Complex Networks. *ACM TKDD*, 11(2), 14:1--14:34. (ACM's 21<sup>st</sup> Best in Computing). [IF: 1.14, Citations: 3]
2. Chakraborty, T., Kumar, S., Ganguly, N., Mukherjee, A. and Bhowmick, S. (2016). GenPerm: A Unified Method for Detecting Non-overlapping and Overlapping Communities. *IEEE TKDE*, 28(8): 2101--2114. [IF: 2.47, Citations: 9]
3. Chakraborty, T., Srinivasan, S., Ganguly, N., Bhowmick, S. and Mukherjee, A. (2013). Constant communities in complex networks. *Nature Scientific Reports*, 3, 1825. [IF: 4.259, Citations: 22]
4. Ghosh, S., Saha, S., Srivastava, A., Kuerger, T., Ganguly, N. and Mukherjee, A. (2013). Understanding Evolution of Inter-Group Relationships using Bipartite Networks. *IEEE Journal on Selected Areas in Communications, Special Issue on Emerging Technologies in Communications*, 31(8), 1--11. [IF: 4.138, Citations: 3]
5. Loreto, V., Mukherjee, A. and Tria, F. (2012). On the origin of the hierarchy of color names. *Proceedings of the National Academy of Sciences (PNAS)*, 109(18), 6819--6824. Full version. Media Coverage: [PNAS Press Highlights](#), [Scientific American](#), [Live Science](#), [MedicalXpress](#), [The Hindu](#), [Business Line](#), [Times of News](#), [BBC Future](#), [Prospect Magazine \(UK\)](#), [Yahoo! News](#), [Deccan Herald](#), [MSN](#), [Homunculus](#), [Doctordisruption](#), [Galileo Net](#), [Scienza e Tecnica](#), [Chinese EurekaAlert](#), [De Standaard](#), [Design by Coffee](#), [Blitz quotidiano](#), [Wired \(Italian\)](#). This work now has a Wikipedia entry in the page for [Color term](#). [IF: 9.66, Citations: 37]
6. Singh, M., Sarkar, R., Goyal, P., Mukherjee, A. and Chakrabarti, S. (2017). Relay-Linking Models for Prominence and Obsolescence in Evolving Networks. In *ACM SIGKDD (17)*, Halifax, Canada. [Ranking: A\*, Citations: just published]
7. Maity, S. K., Saraf, R. and Mukherjee, A. (2016). #Bieber + #Blast = #BieberBlast: Early prediction of popular hashtag compounds. In *CSCW (16)*, San Francisco, CA. (CSCW Honorable Mention, MIT Tech Review [best of the rest of arXiv](#). [Synopsis](#). [Other Comments](#).) [Ranking: A, Citations: 8]
8. Chakraborty, T., Patranabis, S., Goyal, P. and Mukherjee, A. (2015). On the formation of circles in co-authorship networks. In *KDD(15)*, Sydney, Australia. [Ranking: A\*, Citations: 7]
9. Mitra, S., Mitra, R., Riedl, M., Biemann, C., Mukherjee, A. and Goyal, P. (2014). That's sick dude!: Automatic identification of word sense change across different timescales. In *ACL(14)*, Baltimore, USA. [Ranking: A\*, Citations: 24]
10. Chakraborty, T., Srinivasan, S., Ganguly, N., Mukherjee, A. and Bhowmick, S. (2014). On the permanence of vertices in network communities. In *KDD(14)*, New York, USA. [Ranking: A\*, Citations: 48]

## Details of collaborations

- **Matteo Marsili, ICTP, Trieste, Italy**

This collaboration started with the award of Simons Associateship to Dr. Mukherjee in 2014. Since then Dr. Mukherjee has been actively collaborating with Dr. Marsili on various research problems pertaining to (i) evaluation of data clustering algorithms for high dimensional data (ii) large-scale analysis of peer review systems. The collaboration has also got further extended to the Media Lab researchers at SISSA who have generously provided various datasets to run the experiments. The PhD student involved in this project received the prestigious STEP fellowship to visit ICTP in 2015 and 2016 to continue this research. The collaboration has resulted in two top-tier conference papers – ACM CIKM 2016 and JCDL 2017 (the work received a best paper nomination). Two more works are in submission.

- **Chris Biemann, University of Hamburg, Germany**

This collaboration started with a DAAD-IIT faculty exchange program jointly awarded to Dr. Mukherjee and Dr. Biemann in 2013. The two have been closely collaborating on automatic detection of word sense changes from time-stamped corpora. The collaboration has led to the success of a DST-DAAD joint project proposal and two top-tier publications – ACL 2014 and JNLE 2015. Recently, Dr. Mukherjee and Dr. Biemann have started work on co-hyponymy detection in natural language text and their evolution over time. Dr. Mukherjee has been awarded the prestigious Humboldt Fellowship for Experienced Researchers to continue this new thread of collaboration with Dr. Biemann at the University of Hamburg.

- **Sanjukta Bhowmick, University of Nebraska, Omaha**

Dr. Bhowmick and Dr. Mukherjee have been consistently collaborating for more than five years now. This collaboration has, to its credit, top-notch publications like Nature Scientific Reports, ACM SIGKDD, ACM CIKM, ACM TKDD, IEEE TKDE, ASONAM. The collaboration has also resulted in the production of an excellent quality Ph.D. thesis of Dr. Tanmoy Chakraborty. The thesis has attracted many notable awards like the INAE Best Dissertation Award, XRCI Best Dissertation Award etc. Both the collaborators have been actively working on noisy incomplete network analysis and stability of networks for the past one year with a new Ph.D. student. The current proposal is based on the problems to be solved in this thesis and the work has already started receiving appreciation in the community resulting in two notable publications in ASONAM 2016 and CIKM 2016.

- **Tyll Krueger, Wroclaw University of Science and Technology, Poland**

This collaboration started with the visit of Dr. Krueger to IIT Kharagpur in 2012. Dr. Mukherjee and Dr. Krueger have jointly studied various theoretical aspects pertaining to (i) evolution and (ii) diffusion in large networks. The two researchers have jointly come up with new ideas of segmented message diffusion and growth of threshold random graphs. This work resulted in a successful Indo-Polish bilateral researcher exchange project supported by the governments of both the countries. Various PhD students from both the groups have been working on these problems resulting in publications in leading physics and computer science journals including Physical Review E, Europhysics Letters, IEEE JSAC.

- **Vittorio Loreto, Sapienza University, Rome**

Dr. Mukherjee have been collaborating with Dr. Loreto since 2009 while he was in the group of Dr. Loreto as a post-doctoral researcher in ISI Foundation, Italy. Dr. Mukherjee has been working on various issues related to human language dynamics with a special focus on emergence of naming and categorization in linguistic communities. This research has resulted in the development of a computational basis to the emergence of hierarchy in color naming patterns. A successful Indo-Italian bilateral researcher exchange project supported by both the governments has resulted from this project. Many significant publications including PNAS, PLoS One, Europhysics Letters, Journal of Statistical Mechanics have come out of this collaboration.

- **Soumen Chakrabarti, IIT Bombay**

This collaboration started in 2016 to jointly study the dynamics of scientific citations. Dr. Chakrabarti and Dr. Mukherjee have together designed a new model of citation growth that is able to explain both entrenchment as well as obsolescence in a single model of evolution of citation network. Earlier models would mostly explain obsolescence by fitting exponential functions with either global parameters or per-node parameters. However, they could not explain where the diverted attention would go in the network when a certain paper is undergoing obsolescence. In this work, Dr. Chakrabarti and Dr. Mukherjee show the evidence of “relaying” of

citations from older papers to newer papers, i.e., how younger papers would “steal” citations from their older neighbors undergoing obsolescence. This work has been accepted to appear in AC SIGKDD 2017, which is the premium conference in knowledge discovery and data mining.

- **Monojit Choudhury, Microsoft Research Lab, India**

Dr. Choudhury and Dr. Mukherjee have been collaborating since 2006 while both of them were still PhD students. They have together significantly contributed to the literature of language evolution and change with publications in top-tier CS conferences like ACL, COLING, EACL and top physics journals like Physical Review E, Advances in Complex Systems, Europhysics Letters and Physica A.

- **Lipika Dey, Indrajit Bhattacharya, and Arindam Pal, TCS Innovation Labs, Kolkata**

Recently, Dr. Mukherjee’s continuous efforts in citation analysis got noted by one of the major Indian tech giant – Tata Consultancy Services. Consequently the company has engaged into active research collaboration with Dr. Mukherjee and his scholars. Three research scientists (Dr. Dey, Dr. Bhattacharya and Dr. Pal) from TCS innovation labs Kolkata and Delhi are actively collaborating on (i) analysis of patent citations and the growth of industry leaders, (ii) concept mining from citation networks for enhanced future co-authorship prediction and (iii) link dynamics in Wikipedia articles.

## **Media attention**

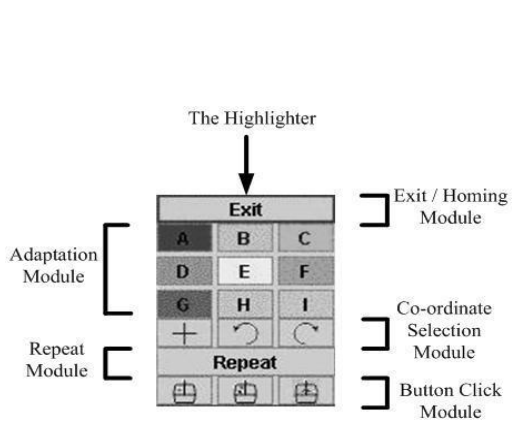
- **The 2012 PNAS paper: PNAS Press Highlights, Scientific American, Live Science, MedicalXpress, The Hindu, Business Line, Times of News, BBC Future, Prospect Magazine (UK), Yahoo! News, Deccan Herald, MSN, Homunculus, Doctordisruption, Galileo Net, Scienza e Tecnica, Chinese EurekAlert, De Standaard, Design by Coffee, Blitz quotidiano, Wired (Italian). This work now has a Wikipedia entry in the page for Color term.**
- **The TKDD 2016 paper: Included in ACM's 21<sup>st</sup> Annual Best of Computing**
- **The CSCW 2016 paper: Received the best paper honorable mention, was included in MIT Tech Review best of arXiv.**
- **The HICSS 2016 paper: Included in the MIT Tech Review best of arXiv.**
- **The work on disaster management has been recently covered by Business Standard and Financial Express.**
- **The JCDL 2017 work analysis of scientific peer review system received a best paper nomination.**
- **The ASONAM 2013 paper on dynamics of popularity of CS fields received a best paper nomination.**

## Invited Talks

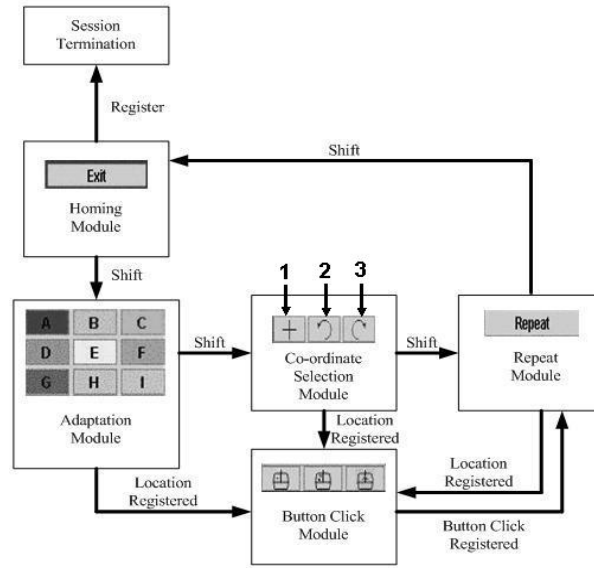
- Invited for the 2017 Microsoft Research India event on “AI for Social Good”.
- Invited talk at Microsoft Research India Lab, 2017.
- Invited speaker at the 2017 IISc workshop on User Generated Content Analysis.
- Invited talk at the workshop on Complex Networks, University of Kerala, 2017.
- Invited to attend and collaborate at the strategy workshop 2017 on “Computation, Sciences and Society” held at Mysore Park Infosys campus. This workshop was organised by K Vijay Raghavan, Kris Gopalakrishnan, Manindra Agrawal, Nisheeth Vishnoi and Somenath Biswas.
- Invited for the early career experience talk at the India-KDD event of ACM SIGKDD 2016 at San Francisco. This is a flagship program of the Indian data mining and data science community organized as a session of the premium data mining conference ACM SIGKDD. (<https://ikdd.acm.org/kdd2016/>).
- Invited speaker for the Machine Learning Data Science event of Microsoft (MLDS 2016) (<http://indiamlmeet.azurewebsites.net/#agenda>).
- Invited speaker at the IWML 2016 workshop at IIT Kanpur. This has been the flagship ML conference in India.
- Invited panelist at the Academic Research Summit 2016 jointly organized by ACM India and Microsoft Research. (<https://www.microsoft.com/en-us/research/event/academic-research-summit/?from=http%3A%2F%2Fresearch.microsoft.com%2Fen-us%2Fevents%2Facadsummit%2F>).
- Invited to participate and talk on the event of Language Day 2016 jointly organized by ICTP, Trieste, Italy and SISSA, Trieste, Italy.
- Invited talk at the workshop on Big Data Analytics and Management, IEST Shibpur, 2016.
- Invited talk at the workshop on Big Data Analytics, Heritage Institute of Technology, 2016.
- Invited talk at Microsoft Research India, Bangalore, 2015.
- Invited talk at MCKVIE, Liluah, 2014.
- Invited talk at ICTP, Trieste, Italy, 2014.
- Invited talk at RCCIT, Kolkata, 2014.
- Invited talk at TU Darmstadt, Germany, 2013.
- Invited talk at Indian Statistical Institute, Kolkata, 2013.
- Invited talk at workshop on Social Media Analysis and Data Mining, Dept. of CST, Bengal Engineering and Science University, Howrah, 2013.
- Invited talk at workshop on Recent Trends in Computing, Dept. of IT, Bengal Engineering and Science University, Howrah, 2013.
- Invited talk at workshop on Complex and Social Networks, Heritage Institute of Technology, Kolkata, 2013.
- Invited talk at Goethe University, Frankfurt, Germany on 11th December, 2012.
- Invited talk at IMPECS Workshop on Social Computing, IIT Kharagpur, 2012.
- Invited talk at S. N. Bose National Center for Basic Sciences, Kolkata, 2012.
- Invited talk at Indian Statistical Institute, Kolkata, 2012.
- Invited talk at Goethe University, Frankfurt, Germany, 2012.
- Invited talk at Technical University of Darmstadt, Germany, 2012.
- Invited talk at the workshop on Mathematical Physics of Complex Networks: From Graph Theory to Biological Physics, MPI PKS, Dresden, Germany, 2012.
- Invited talk at workshop on Social Networks, IMSc, Chennai, 2012.
- Invited talk at the workshop on Dynamics on and of Complex Networks, ECCS, Vienna, Austria, 2011.
- Invited talk at Max Planck Institute of Evolutionary Anthropology, Leipzig, Germany, 2007.

## Working Systems Developed

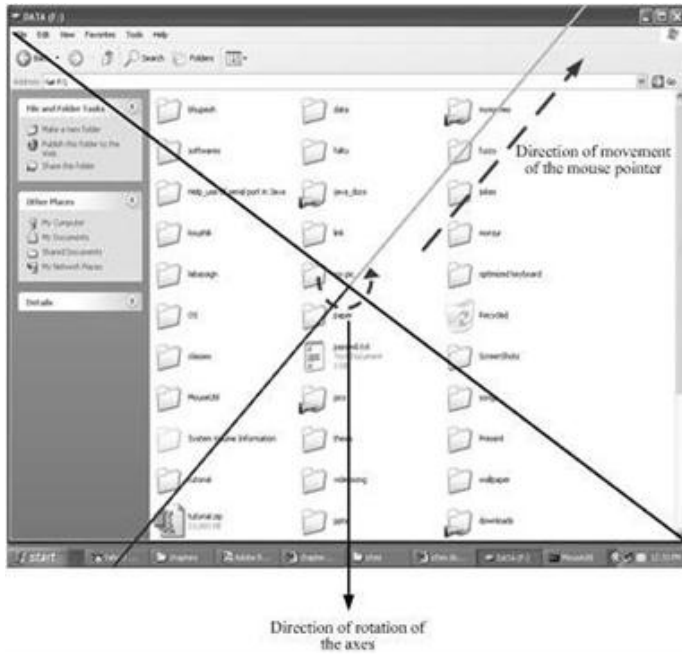
### A) Sweepsticks – An Adaptive Virtual Mouse for users with Neuro Motor Disorder



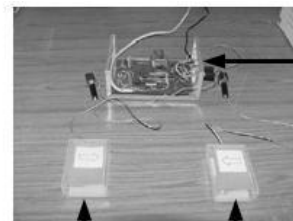
The main on-screen interface of SweepSticks.



Process flow of SweepSticks.



The 360 degree screen scan.



The Interface with the Computer  
 The Switch Emulating the Shift Operation  
 The Switch Emulating the Register Operation

The mechanical switches to operate SweepSticks.

## B) FeRoSA – A faceted recommendation system for scientific articles

The screenshot displays the FeRoSA web application interface. At the top left, the logo "FeRoSA" is shown in teal. To the right, there are navigation buttons for "HOME" and "RECOMMENDER ENGINE". Below the header is a horizontal menu with faceted navigation options: "All", "Background", "Alternative Approaches", "Method", "Comparison", and "Common". The main content area is a grid of nine article cards, each with a category label and a title. A scroll bar is visible on the right side, and a small square button with an upward arrow is located at the bottom right of the grid.

Category	Title
Method.	Mining WordNet For A Fuzzy Sentiment: Sentiment Tag Extraction From WordNet Glosses
Method. Comparison.	Effects Of Adjective Orientation And Gradability On Sentence Subjectivity
Comparison.	Converting Dependency Structures To Phrase Structures
Common	A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based On Minimum Cuts
Background.	The Sentimental Factor: Improving Review Classification Via Human-Provided Information
Common	Determining The Sentiment Of Opinions
Alternate Approaches. Background.	Extracting Product Features And Opinions From Reviews
Common	Co-Training for Cross-Lingual Sentiment Classification
Alternate Approaches. Method.	Recognizing Contextual Polarity In Phrase-Level Sentiment Analysis

## C) OCR++ – A robust framework for information extraction from scholarly articles

The screenshot displays the OCR++ web application interface. On the left is a dark blue sidebar with the OCR++ logo and a navigation menu. The main content area is titled "Upload file" and includes a "Choose File" button, a status indicator "No file chosen", and an "Upload" button. Below this is a preview of a scholarly article page from the Journal of Biomedical Informatics. The article title is "PDF text classification to leverage information extraction from publication reports" by Duy Duc An Bui, Guilherme Del Fiol, and Siddhartha Jonnalagadda. The interface also shows a search icon in the top right and a footer with navigation icons.

**OCR++**

Welcome :)

**FEATURES**

- Home
- Metadata
- Sections
- Tables and Figures
- Footnotes and URLs
- Citations & References

**OTHER OPTIONS**

- Team
- Source Code
- Feedback

**Upload file**

Select a file:

No file chosen

Journal of Biomedical Informatics 61 (2016) 141–148

Contents lists available at ScienceDirect

**Journal of Biomedical Informatics**

journal homepage: [www.elsevier.com/locate/yjbin](http://www.elsevier.com/locate/yjbin)

**PDF text classification to leverage information extraction from publication reports**

Duy Duc An Bui<sup>a,b,\*</sup>, Guilherme Del Fiol<sup>a</sup>, Siddhartha Jonnalagadda<sup>b</sup>

<sup>a</sup>Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, USA  
<sup>b</sup>Department of Preventive Medicine-Health and Biomedical Informatics, Northwestern University, Chicago, IL, USA

CrossMark