

# **SOUMYA BANERJEE**

## **PUBLICATIONS**

The top five papers are presented first.

### **Key Refereed Journal Papers**

- 1) A class-contrastive human-interpretable machine learning approach to predict mortality in severe mental illness, S. Banerjee, P. Lio, P. Jones, R. Cardinal, *Nature Partner Journal Schizophrenia*, 7, 60, 2021 \* (*Nature publishing group*, Impact Factor = 6.3)

*I developed an explainable machine learning technique to predict mortality. This was applied to data from hospital electronic healthcare records. I worked closely with clinicians to develop this model.*

- 2) Optogenetic tuning reveals Rho amplification-dependent dynamics of a cell contraction signal network, D. Kamps, J. Koch, V. Juma, E. Campillo, M. Graessl, S. Banerjee, T. Mazel, X. Chen, Y. Wu, S. Portet, A. Madzvamuse, P. Nalbant, L. Dehmelt, *Cell Reports*, 33(9):108467, 2020 (*Cell Press publishing group*, Impact Factor = 8.1)

- 3) Deconvolution of monocyte responses in inflammatory bowel disease reveals an IL-1 cytokine network that regulates IL-23 in genetic and acquired IL-10 resistance, D. Aschenbrenner, M. Quaranta, S. Banerjee, et al. *Gut*, 2020 (*British Medical Journal publishing group*, Impact Factor = 19.8)

*The work integrates clinical data with genomic information to predict if patients would respond to therapy. My role was leading the computational effort and liaising with clinicians.*

- 4) Hydroxychloroquine: balancing the needs of LMICs during the COVID-19 pandemic, S. Banerjee, *Lancet Rheumatology*, 2(7):385-386, 2020

*The work analyses the demand surge for hydroxychloroquine and the effect it may have on patients in developing nations.*

- 4) Influence of correlated antigen presentation on T cell negative selection in the thymus, S. Banerjee, S.J. Chapman, *Journal of the Royal Society Interface*, 15(148), 20180311, 2018 (Impact Factor = 4.3) \*

*I built a computer model that integrates genomic data on how the immune system is trained to recognize pathogens. I liaised with biologists and led the computational work.*

- 5) Modelling the effects of phylogeny and body size on within-host pathogen replication and immune response, S. Banerjee, A. Perelson, M. Moses, *Journal of the Royal Society Interface*, 14(136), 20170479, 2017 (Impact Factor = 4.3) \*

- 6) Towards a Quantitative Understanding of Within Host Dynamics of West Nile Virus Infection, S. Banerjee, J. Guedj, R. Ribeiro, M. Moses & A. Perelson, *Journal of the Royal Society Interface*, 13(117), 20160130, 2016 (Impact Factor = 4.3) \*

- 7) Predictive metabolomic profiling of microbial communities using amplicon or metagenomic sequences, H. Mallick, E. Franzosa, L. McIver, S. Banerjee, et al., *Nature Communications*, 10(1):3136, 2019 (Impact Factor = 11.8) \*

*This paper integrates clinical data with genomic data on the gut microbiome. I liaised with clinicians, experimentalists and industry experts and built the first draft of the machine learning model.*

- 8) Patient and public involvement to build trust in artificial intelligence: a framework, tools and case studies, S.Banerjee, P. Alsop, L. Jones, R. Cardinal, *Patterns*, 3(6):100506, 2022  
(Cell Press publishing group)
- 9) An excitable Rho GTPase signaling network generates dynamic subcellular contraction patterns, M. Graessl, J. Koch, A. Calderon, S. Banerjee, T. Mazel, N. Schulze, J. Jungkunth, A. Koseska, L. Dehmelt, P. Nalbant, *Journal of Cell Biology*, 216(12), 4271-4285, 2017 (Impact Factor = 9.7)
- 10) Involving patients in artificial intelligence research to build trustworthy systems, S. Banerjee, S. Griffiths, *AI & Society*, 2023
- 11) Software Application Profile: ShinyDataSHIELD—an R Shiny application to perform federated non-disclosive data analysis in multicohort studies, X. Escribà-Montagut, Y. Marcon, D. Avraam, S. Banerjee, T. Bishop, P. Burton, J. González, *International Journal of Epidemiology*, dyac201, 2022  
(Oxford University Publishing Group, Impact Factor 9.8)
- 12) dsSynthetic: Synthetic data generation for the DataSHIELD federated analysis system, S.Banerjee, T. Bishop, *BMC Research Notes*, 15(1):230, 2022
- 13) dsSurvival: Privacy preserving survival models for federated meta-analysis of clinical data, S.Banerjee, et al, *BMC Research Notes*, 15(1):197, 2022
- 14) The early impact of COVID-19 on mental health and community physical health services and their patients' mortality in Cambridgeshire and Peterborough, UK, S. Chen, P. Jones, B. Underwood, A. Moore, E. Bullmore, S. Banerjee, et al., *Journal of Psychiatric Research*, 131, 244-254, 2020 (Impact factor = 4.4)
- 15) A spatial model of the efficiency of T cell search in the influenza infected lung, D. Levin, S. Forrest, S. Banerjee, M. Moses & F. Koster, *Journal of Theoretical Biology*, 398(7), 52-63, 2016 (Impact Factor = 2.2)
- 16) Competitive dynamics between criminals and law enforcement explains the super-linear scaling of crime in societies, S. Banerjee, Hentenryck, P.V. & Cebrian, M. *Palgrave Communications*, 1(1), 15022, 2015 (Nature Publishing Group)  
*I worked with a social-scientist to understand why per-capita crime is higher in bigger cities. I collected the data, interacted with the social scientist and built the model.*
- 17) A bioorthogonal small-molecule switch system for controlling protein function in cells, P. Liu, A. Calderon, G. Konstantinidis, J. Hou, S. Voss, X. Chen, F. Li, S. Banerjee, J. Hoffmann, C. Theiss, L. Dehmelt & Y. Wu, *Angewandte Chemie*, 53(38), 10049-10055, 2014 (Impact Factor = 13.7)
- 18) Science and technology consortia in US biomedical research: A paradigm shift in response to unsustainable academic growth, Curt Balch, Hugo Arias-Pulido, S. Banerjee, Alex K. Lancaster, Kevin B. Clark, Michael Perilstein, Brian Hawkins, John Rhodes, Piotr Sliz, Jon Wilkins and Thomas W. Chittenden, *BioEssays*, 37(2), 119-122, 2014 (Impact Factor = 5.4)

- 19) Scale Invariance of Immune System Response Rates and Times: Perspectives on Immune System Architecture and Implications for Artificial Immune Systems, S. Banerjee & M. Moses, *Swarm Intelligence*, 4(4), 301-308, 2010 (Impact Factor = 2.1)

## Key Refereed Conference Papers

NOTE: In my parent discipline (Computer Science), people submit to conferences and conferences are peer-reviewed

- 1) Analysis of a Planetary Scale Scientific Collaboration Dataset Reveals Novel Patterns, S. Banerjee, *Proceedings of the Complex Systems Digital Campus 2015 – World e-Conference, Conference on Complex Systems*, 2016 (peer-reviewed conference)
- 2) A Multi-Agent System Approach to Load-Balancing and Resource Allocation for Distributed Computing, S. Banerjee and J. Hecker, *Proceedings of the Complex Systems Digital Campus 2015 – World e-Conference, Conference on Complex Systems*, 2016 (peer-reviewed conference)
- 3) Computationally Simulating Intermodal Terminal Attractiveness and Demand, S. Banerjee et al., *Proceedings of the 23rd World Congress on Intelligent Transport Systems*, 2016  
*I worked with companies and built a model that predicts demand for logistics and supply chains.*
- 4) Analysis of Demand and Operations of Intermodal Terminals, R. Garcia-Flores, S. Banerjee et al., *Proceedings of the 24th National Conference of the Australian Society for Operations Research*, 2016
- 5) Forecasting in the era of Big Data: Lessons and Pitfalls, Y. Tyshetskiy, S. Banerjee et al., *Proceedings of the Annual Conference of the International Association of Maritime Economists*, 2016
- 6) The Value of Inflammatory Signals in Adaptive Immune Responses, S. Banerjee, D. Levin, M. Moses, F. Koster & S. Forrest, *The 10th International Conference on Artificial Immune Systems (ICARIS)*, Lecture Notes in Computer Science, Volume 6825/2011, 1-14, 2011
- 7) Biologically Inspired Design Principles for Scalable, Robust, Adaptive, Decentralized Search and Automated Response (RADAR), M. Moses & S. Banerjee, *Proceedings of the 2011 IEEE Conference on Artificial Life*, 30-37, 2011
- 8) Modular RADAR: An Immune System Inspired Search and Response Strategy for Distributed Systems, S. Banerjee & M. Moses, *The 9<sup>th</sup> International Conference on Artificial Immune Systems (ICARIS)*, Lecture Notes in Computer Science, Volume 6209/2010, 116-129, 2010
- 9) A Hybrid Agent Based and Differential Equation Model of Body Size Effects on Pathogen Replication and Immune System Response, S. Banerjee & M. Moses, *The 8th International Conference on Artificial Immune Systems (ICARIS)*, Volume 5666-014, 2009

## FULL PUBLICATION LIST

### Refereed Journal Papers

- 1) Balancing Objects and Processes: Advocating Pluralism in Biology: Comment on “Thoughts and thinkers: On the Complementarity between Objects and Processes” by Chris Fields and Michael Levin, S. Banerjee, *Physics of Life Reviews*, 2025  
(Impact Factor 13.7)

- 2) DataSHIELD: Mitigating disclosure risk in a multi-site federated analysis platform, D. Avraam, R. Wilson, N. Chan, S. Banerjee et al. *Bioinformatics Advances*, vba046, 2025  
 (Oxford University Press publishing group)
- 3) Neural networks for abstraction and reasoning, M. Bober-Irizar, S. Banerjee, *Scientific Reports*, 14, 27823, 2024  
 (Nature Publishing Group, Impact Factor 3.8)
- 4) Generating Complex Explanations for Artificial Intelligence Models: An Application to Clinical Data on Severe Mental Illness, S. Banerjee, *Life*, 14(7), 807, 2024
- 5) Intelligence in animals, humans and machines: a heliocentric view of intelligence? H. Holm, S. Banerjee, *AI & Society*, 2024
- 6) Involving patients in artificial intelligence research to build trustworthy systems, S. Banerjee, S. Griffiths, *AI & Society*, 2023
- 7) dsSurvival 2.0: Privacy enhancing survival curves for survival models in the federated DataSHIELD analysis system, S. Banerjee, T. Bishop, *BMC Research Notes* 16, 98, 2023
- 8) Patient and public involvement to build trust in artificial intelligence: a framework, tools and case studies, S. Banerjee, P. Alsop, L. Jones, R. Cardinal, *Patterns*, 3(6):100506, 2022  
 (Cell Press publishing group)
- 9) Software Application Profile: ShinyDataSHIELD—an R Shiny application to perform federated non-disclosive data analysis in multicohort studies, X. Escribà-Montagut, Y. Marcon, D. Avraam, S. Banerjee, T. Bishop, P. Burton, J. González, *International Journal of Epidemiology*, dyac201, 2022  
 (Oxford University Publishing Group, Impact Factor 9.8)
- 10) dsSynthetic: Synthetic data generation for the DataSHIELD federated analysis system, S. Banerjee, T. Bishop, *BMC Research Notes*, 15(1):230, 2022
- 11) dsSurvival: Privacy preserving survival models for federated meta-analysis of clinical data, S. Banerjee, et al, *BMC Research Notes*, 15(1):197, 2022
- 12) A class-contrastive human-interpretable machine learning approach to predict mortality in severe mental illness, S. Banerjee, P. Lio, P. Jones, R. Cardinal, *Nature Partner Journal Schizophrenia*, 7, 60, 2021  
 (Nature publishing group, Impact Factor = 6.3)
- 13) Optogenetic tuning reveals Rho amplification-dependent dynamics of a cell contraction signal network, D. Kamps, J. Koch, V. Juma, E. Campillo, M. Graessl, S. Banerjee, T. Mazel, X. Chen, Y. Wu, S. Portet, A. Madzvamuse, P. Nalbant, L. Dehmelt, *Cell Reports*, 33(9):108467, 2020  
 (Cell Press publishing group, Impact Factor = 8.1)
- 14) Deconvolution of monocyte responses in inflammatory bowel disease reveals an IL-1 cytokine network that regulates IL-23 in genetic and acquired IL-10 resistance, D. Aschenbrenner, M. Quaranta, S. Banerjee, et al., *Gut*, 2020  
 (British Medical Journal publishing group, Impact Factor = 19.8)
- 15) Hydroxychloroquine: balancing the needs of LMICs during the COVID-19 pandemic, S. Banerjee, *Lancet Rheumatology*, 2(7):385-386, 2020  
 (Lancet publishing group)
- 16) The early impact of COVID-19 on mental health and community physical health services and their patients' mortality in Cambridgeshire and Peterborough, UK, S. Chen, P. Jones, B. Underwood, A. Moore, E. Bullmore, S. Banerjee, et al., *Journal of Psychiatric Research*, 131, 244-254, 2020  
 (Impact factor = 4.4)
- 17) Predictive metabolomic profiling of microbial communities using amplicon or metagenomic sequences, H. Mallick, E. Franzosa, L. McIver, S. Banerjee, A. Sirota-Madi, A. Kostic, C. Clish, H. Vlamakis, R. Xavier, C. Huttenhower, *Nature Communications*, 10(1):3136, 2019  
 (Nature publishing group, Impact Factor = 12.1)
- 18) Influence of correlated antigen presentation on T cell negative selection in the thymus, S. Banerjee, S.J. Chapman, *Journal of the Royal Society Interface*, 15(148), 20180311, 2018

(Royal Society publishing group, Impact Factor = 4.3)

19) Modelling the effects of phylogeny and body size on within-host pathogen replication and immune response, S. Banerjee, A. Perelson, M. Moses, *Journal of the Royal Society Interface*, 14(136), 20170479, 2017

(Royal Society publishing group, Impact Factor = 4.3)

20) An excitable Rho GTPase signaling network generates dynamic subcellular contraction patterns, M. Graessl, J. Koch, A. Calderon, S. Banerjee, T. Mazel, N. Schulze, J. Jungkunth, A. Koseska, L. Dehmelt, P. Nalbant, *Journal of Cell Biology*, 216(12), 4271-4285, 2017

(Cell publishing group, Impact Factor = 9.7)

21) Towards a Quantitative Understanding of Within Host Dynamics of West Nile Virus Infection, S. Banerjee, J. Guedj, R. Ribeiro, M. Moses & A. Perelson, *Journal of the Royal Society Interface*, 13(117), 20160130, 2016

(Royal Society publishing group, Impact Factor = 4.3)

22) A spatial model of the efficiency of T cell search in the influenza infected lung, D. Levin, S. Forrest, S. Banerjee, M. Moses & F. Koster, *Journal of Theoretical Biology*, 398(7), 52-63, 2016  
(Impact Factor = 2.2)

23) Competitive dynamics between criminals and law enforcement explains the super-linear scaling of crime in societies, S. Banerjee, Hentenryck, P.V. & Cebrian, M., *Palgrave Communications*, 1(1), 15022, 2015

(Nature publishing group)

24) A bioorthogonal small-molecule switch system for controlling protein function in cells, P. Liu, A. Calderon, G. Konstantinidis, J. Hou, S. Voss, X. Chen, F. Li, S. Banerjee, J. Hoffmann, C. Theiss, L. Dehmelt & Y. Wu, *Angewandte Chemie*, 53(38), 10049-10055, 2014

(Impact Factor = 13.7)

25) Science and technology consortia in US biomedical research: A paradigm shift in response to unsustainable academic growth, C. Balch, H. Arias-Pulido, S. Banerjee, A. Lancaster, K. Clark, M. Perilstein, B. Hawkins, J. Rhodes, P. Sliz, J. Wilkins and T. Chittenden, *BioEssays*, 37(2), 119-122, 2014

(Impact Factor = 5.4)

26) Scale Invariance of Immune System Response Rates and Times: Perspectives on Immune System Architecture and Implications for Artificial Immune Systems, S. Banerjee & M. Moses, *Swarm Intelligence*, 4(4), 301-308, 2010

(Impact Factor = 2.1)

## Refereed Workshop Papers

1) Bridging the Gap with Retrieval-Augmented Generation: Making Prosthetic Device User Manuals Available in Marginalised Languages Ogbonna Ikechukwu, Lesley Davidson, Soumya Banerjee, Abhishek Dasgupta, Laurence Kenney, Vikranth Harthikote Nagaraja *Data Science Africa 2025 Workshop*, 2025

2) Prime Directives for Responsible AI for Africa: A Manifesto for Inclusive Technology Soumya Banerjee, Vikranth Harthikote Nagaraja, Innocent Nyalala, Nirav Pravinbhai Bhatt *Data Science Africa 2025 Workshop*, 2025

3) Understanding Intelligence Across Humans, Machines, and Animals S. Banerjee, *Workshop on Exploring Interdisciplinary Frontiers: Cognitive Science, Computational Modeling, and Artificial Intelligence*, 2025

4) Towards a theory of life-like systems S. Banerjee, *IEEE Symposium on CI in Artificial Life and Cooperative Intelligent Systems (Late Breaking Abstract): Trondheim, Norway*, 2025

5) Augmenting human intelligence on abstraction and reasoning tasks: a path synergistic AI-human interaction for the Abstraction and Reasoning Corpus, S. Banerjee, *IEEE Symposium on CI in Artificial Life and Cooperative Intelligent Systems (Late Breaking Abstract): Trondheim, Norway*, 2025

- 6) "I apologize for my actions": Emergent Properties of Generative Agents and Implications for a Theory of Mind, N. Diamond, S. Banerjee *AAAI workshop on Advancing Artificial Intelligence through Theory of Mind (ToM4AI)*, 2025
- 7) The BACON system for equation discovery from scientific data: Reconciling classical artificial intelligence with modern machine learning approaches, Jonah Miller, S. Banerjee *4th Annual AAAI Workshop on AI to Accelerate Science and Engineering (AI2ASE)*, 2025
- 8) Enhancing AI Capabilities on the Abstraction and Reasoning Corpus: A Path Toward Broad Generalization in Intelligence, S. Banerjee, *4th Annual AAAI Workshop on AI to Accelerate Science and Engineering (AI2ASE)*, 2025
- 9) AugARC: Augmented Abstraction and Reasoning Benchmark for Large Language Models, K. Bikov, M. Bober-Irizar, S. Banerjee, *AAAI Workshop on Preparing Good Data for Generative AI: Challenges and Approaches*, 2025
- 10) AI and the Future of the Technosphere: A Path Towards Co-Existence with Superintelligence, S. Banerjee, *AAAI Workshop on Post-Singularity Symbiosis*, 2025
- 11) MASR: Multi-Agent System with Reflection for the Abstraction and Reasoning Corpus, K. Bikov, M. Bober-Irizar, S. Banerjee, *AAAI 2025 Workshop on Advancing LLM-Based Multi-Agent Collaboration*, 2025
- 12) What can machines teach us in our journey of reproducing human scientific creativity? A. Mazumdar, N. Karim, S. Banerjee, *IEEE Symposium on CI in Artificial Life and Cooperative Intelligent Systems (Late Breaking Abstract): Trondheim, Norway*, 2025
- 13) Reflection System for the Abstraction and Reasoning Corpus, K. Bikov, M. Bober-Irizar, S. Banerjee, *AAAI 2nd AI4Research Workshop: Towards a Knowledge-grounded Scientific Research Lifecycle*, 2025
- 14) On the Ethical Considerations of Generative Agents, N. Diamond, S. Banerjee, *NeurIPS Workshop, Socially Responsible Language Modelling Research (SoLaR)*, 2024
- 15) Neural networks for abstraction and reasoning, M. Bober-Irizar, S. Banerjee, *NeurIPS Workshop, Multimodal Algorithmic Reasoning (MAR)*, 2024  
(selected for oral presentation)
- 16) Enhancing patient stratification and interpretability through class-contrastive and feature attribution techniques, S. Olowu, N. Lawrence, S. Banerjee, *NeurIPS Workshop, Interpretable AI: Past, Present and Future*, 2024
- 17) Building Trustworthy AI: The Role of Patient and Public Involvement in Healthcare AI Development, S. Banerjee *Proceedings of the AAAI Symposium Series 4 (1)*, 329-331, 2024
- 18) The BACON system for equation discovery from scientific data: Transforming classical artificial intelligence with modern machine learning approaches, J. Miller, S. Banerjee, *Association for the Advancement of Artificial Intelligence (AAAI) Fall Symposium Series, Integrated Approaches to Computational Scientific Discovery*, 2024

## Refereed Conference Papers

- 1) Evaluating Machine Learning Models for Prediction of Attention Deficit Hyperactivity Disorder among Autistic Individuals Using Genetic Data, *2025 World Congress of Psychiatric Genetics*, 2025
- 2) Cosmicism and Artificial Intelligence: Beyond Human-Centric AI S. Banerjee *The 1st International Online Conference of the Journal Philosophies*, 2025
- 3) When Planes Fly Better Than Birds: Should AIs Think Like Humans? S. Banerjee *The 1st International Online Conference of the Journal Philosophies*, 2025
- 4) Enhancing patient stratification and interpretability through class-contrastive and feature attribution techniques, S. Olowu, N. Lawrence, S. Banerjee, *Proceedings of the 2025 IEEE Symposium Series on Computational Intelligence, Trondheim, Norway*, 2025

- 5) "I apologize for my actions": Emergent Properties of Generative Agents and Implications for a Theory of Mind, N. Diamond, S. Banerjee *Proceedings of the 2025 IEEE Symposium Series on Computational Intelligence, Trondheim, Norway*, 2025
- 6) Analysis of a Planetary Scale Scientific Collaboration Dataset Reveals Novel Patterns, S. Banerjee, *Proceedings of the Complex Systems Digital Campus 2015 – World e-Conference, Conference on Complex Systems*, 2016 (peer-reviewed conference)
- 7) A Multi-Agent System Approach to Load-Balancing and Resource Allocation for Distributed Computing, S. Banerjee and J. Hecker, *Proceedings of the Complex Systems Digital Campus 2015 – World e-Conference, Conference on Complex Systems*, 2016 (peer-reviewed conference)
- 8) Computationally Simulating Intermodal Terminal Attractiveness and Demand, S. Banerjee et al., *Proceedings of the 23rd World Congress on Intelligent Transport Systems*, 2016
- 9) Analysis of Demand and Operations of Intermodal Terminals, R. Garcia-Flores, S. Banerjee et al., *Proceedings of the 24th National Conference of the Australian Society for Operations Research*, 2016
- 10) Forecasting in the era of Big Data: Lessons and Pitfalls, Y. Tyshetkiy, S. Banerjee et al., *Proceedings of the Annual Conference of the International Association of Maritime Economists*, 2016
- 11) The Value of Inflammatory Signals in Adaptive Immune Responses, S. Banerjee, D. Levin, M. Moses, F. Koster & S. Forrest, *The 10th International Conference on Artificial Immune Systems (ICARIS)*, Lecture Notes in Computer Science, Volume 6825/2011, 1-14, 2011
- 12) Biologically Inspired Design Principles for Scalable, Robust, Adaptive, Decentralized Search and Automated Response (RADAR), M. Moses & S. Banerjee, *Proceedings of the 2011 IEEE Conference on Artificial Life*, 30-37, 2011
- 13) Modular RADAR: An Immune System Inspired Search and Response Strategy for Distributed Systems, S. Banerjee & M. Moses, *The 9<sup>th</sup> International Conference on Artificial Immune Systems (ICARIS)*, Lecture Notes in Computer Science, Volume 6209/2010, 116-129, 2010
- 14) A Hybrid Agent Based and Differential Equation Model of Body Size Effects on Pathogen Replication and Immune System Response, S. Banerjee & M. Moses, *The 8th International Conference on Artificial Immune Systems (ICARIS)*, Volume 5666-014, 2009

## **Refereed Book Chapters**

Using Optimisation and Machine Learning to Validate the Value of Infrastructure Investments, R. Garcia-Flores, S. Banerjee, G. Mathews (2016) Book chapter in *Infrastructure Investments: Politics, Barriers and Economic Consequences*