BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Luis M. Rocha

eRA COMMONS USER NAME (credential, e.g., agency login): lmrocha

POSITION TITLE: George J. Klir Professor of Systems Science

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Instituto Superior Técnico, Portugal	B.S.	1985-1988	Mechanical Engineering
Instituto Superior Técnico, Portugal	M.S.	1988-1990	Systems Engineering
State University of New York at Binghamton	Ph.D.	1992-1997	Systems Science
Los Alamos National Laboratory	Post Doc	1997-1998	Complex Systems

A. Personal Statement

My research and leadership are in complex networks and systems, biomedical informatics, computational and systems biology, and computational intelligence. Over decades of publishing in these areas, I have laid the groundwork to apply network science and machine learning methods to integrating and analyzing multiscale, heterogeneous biomedical datasets from multiomics, pharmacology, electronic health records, social media, and other sources. My lab has considerable experience in developing text mining and NLP pipelines for producing network models by extracting evidence from the scientific literature, social media, and electronic health records in biomedicine and public health. I also have an established record in complex systems and network science, with a focus on multivariate dynamics in biomedical problems. I am committed to interdisciplinary research and graduate training, having created and directed various interdisciplinary PhD programs. I was the lead PI of a NSF-NRT interdisciplinary PhD training program in complex networks & systems (NSF-1735095). I have graduated 14 PhD students and 9 postdocs in these topics and am currently training 7 PhD students and 4 postdocs.

Products

Products Closely Related to the Proposed Project

1. R.B. Correia, J.C. Rozum, L. Cross, J.Felag, M. Gallant, Z. Guo, B.W. Herr II, A. Min, J. Sanchez-Valle, D.S. Rocha, A. Valencia, X. Wang, K.Börner, W.R.Miller, and L.M. Rocha. [2026]. "myAURA: Personalized health library for epilepsy management via knowledge graph sparsification and visualization". *Journal of the American Medical Informatics Association*, **33**(1):167-181. PMCID: PMC12758476.
2. M. De Domenico, L. Allegri, G. Caldarelli, V. d'Andrea, B. Di Camillo, L.M. Rocha, J. Rozum, R. Sbarbati, F. Zambelli [2025]. "Challenges and opportunities for digital twins in precision medicine: a complex systems perspective". *npj Digital Medicine* **8**, 37. PMCID: PMC11742446.
3. M. Ginda, K. Börner, O. Scrivner, W. Trochim, L.M. Rocha, Selma Šabanović. [2026] "Preparing the next generation of complex networks and systems scientists: Evaluation results for the complex networks and systems NSF research training program at Indiana University". *PLOS One*. **21**(1): e0334779. PMCID: PMC12843521.

4. A.A. Hamed and L.M. Rocha [2026]. "Protocol for evaluating ChatGPT in biomedical association generation and verification using a RAG-enabled, cross-model majority voting workflow". *STAR protocols*, **7**(2): 104533. PMID: 42133493.
5. J. Sánchez-Valle, M. Flores-Rodero, F.X. Costa, J. Carbonell-Caballero, I. Núñez-Carpintero, R. Tabarés-Seisdedos, L.M. Rocha, D. Cirillo, A. Valencia [2026]. "Sex-specific transcriptome similarity networks elucidate comorbidity relationships." *Communications Medicine*, **6**: 61. PMCID: PMC12847782.

Other Significant Products Highlighting Contributions to Science

1. R.B. Correia, J.M. Almeida, Ma.J. Wyrwoll, I.Julca, D. Sobral, C.S. Misra, L.G. Guilgur, H.-C. Schuppe, N. Silva, P. Prudêncio, A. Nóvoa, A. S. Leocádio, J. Bom, M. Mallo, S. Kliesch, M. Mutwil, L.M. Rocha, F.Tüttelmann, J. D. Becker, P. Navarro-Costa [2024]. "The conserved transcriptional program of metazoan male germ cells uncovers ancient origins of human infertility". *eLife*. **13**: RP95774. PMCID: PMC11466473.
2. A. Gates, Correia, R.B., X. Wang, L.M. Rocha [2021]. “The effective graph reveals redundancy, canalization, and control pathways in biochemical regulation and signaling.” *Proceedings of the National Academy of Sciences*. **118**. PMCID: PMC8000424.
3. Z. Guo, J. Felag, J.C. Rozum, R.B. Correia, and L.M. Rocha. [2025]. " Focused digital cohort selection from social media using the metric backbone of biomedical knowledge graphs." *J. of Biomedical Informatics*. **168**: 104847. PMCID: PMC13227082.
4. J. Sánchez-Valle, R.B. Correia, M. Camacho-Artacho, R. Lepore, M.M. Mattos, L. M. Rocha, A. Valencia [2024]. “Prevalence and differences in the co-administration of drugs known to interact: an analysis of three distinct and large populations.” *BMC Medicine*. **22**: 166. PMCID: PMC11027217.
5. T. Simas, R.B. Correia, and L.M. Rocha. [2021] “Distance backbone of complex networks”. *Journal of Complex Networks*, **9** (6), cnab021. PMCID: PMC10859673.

B. Positions, Scientific Appointments, and Honors

Appointments and Positions:

2021-	George J. Klir Professor of Systems Science, Binghamton University, SUNY
2024-	Visiting Professor, Católica Medical School, Portugal
2019-	Co-director, Consortium for Social and Biomedical Complexity, Indiana University
2019-2021	Visiting Professor, NOVA School of Business and Economics, Portugal
2017-2023	Director, NSF-NRT Interdisciplinary Training in Complex Networks and Systems.
2016-2018	Visiting Professor, Center for Theoretical Physics, University of Aix-Marseille, France
2016-2018	Visiting Professor, Champalimaud Foundation, Portugal
2013-2021	Professor, Indiana University, Luddy School of Informatics, Computing & Engineering.
2008-2020	Director, Indiana University, Complex Networks & Systems track, Informatics PhD Program, Luddy School of Informatics, Computing & Engineering.
2004-2013	Associate Professor, Indiana University, Luddy School of Informatics, Computing & Engineering.
2002-2024	Principal Investigator, Instituto Gulbenkian de Ciencia, Portugal
2002-2012	Director, Computational Biology Collaboratorium, and co-director of PhD Program in Computational Biology, Instituto Gulbenkian de Ciencia, Portugal
1999-2004	Technical Staff Member, Los Alamos National Laboratory
1998-2002	Team Leader Complex Systems Modeling Team, Los Alamos National Laboratory
1997-1999	Postdoctoral Fellow, Los Alamos National Laboratory
1995-1997	Adjunct Professor, State University of New York, Binghamton, Systems Science Dep.
1990-1991	Graduate Research Assistant, Laboratório Nacional de Engenharia Civil, Portugal.

Other Experience and Professional Service (selected):

2026-2028 Elected member of the Binghamton University Faculty Senate.
2026- Advisory Board. Computational Sciences and Deep Tech Technologies PhD Program (Collegio di Dottorato), at the Università degli Studi del Piemonte Orientale (Italy)
2026 Program Chair, *2026 Conference on Complex Systems* (CCS), October 12-16, 2026, Binghamton, NY.
2025 General Co-Chair, Eight *Northeast Regional Conference on Complex Systems* (NERCCS25). April 9-11, 2025. Binghamton, NY.
2024- External Advisory Board, Environmental Health Institute (Instituto de Saúde Ambiental), University of Lisbon Medical School (ISAMB-FMUL), Portugal
2023- Section Editor, *PLoS Complex Systems*, Complex Systems in Nature section
2023 Reviewer, Clinical Data Management and Analysis (CDMA) study section at CSR NIH
2022- Editorial Board, Frontiers in Physics - Complex Systems section
2021- Editorial Board, Journal of Computational Science
2021, 2023 Reviewer, Clinical Informatics and Digital Health (CIDH) study section at the CSR NIH
2020 Reviewer, Biomedical Computing & Health Informatics (BCHI) study section at CSR NIH
2020-2027 Program Chair, Complex Networks Conference
2019 General Chair, Complex Networks 2019
2018 -2021 Council (elected member), Complex Systems Society (CCS)
2018 External Review Committee, *Center for the Study of Complex Systems*, U. Michigan.
2018 -2021 Advisory Council, *Indiana University Network Science Institute* (IUNI)
2017- Editorial Board, Complexity
2017- Editorial Board, BMC Bioinformatics
2014 NSF Site Visit Committee of the *Biocomputational Evolution in Action CONsortium* (BEACON) NSF Science and Technology Center at Michigan State University
2014- Editorial Board, Frontiers in Robotics and AI | Computational Intelligence
2011 Member, NSF, Networks, Synthetic Biology and Evolution review panel
2011- Editorial Board, PLoS ONE
2006 Chair, Organizing Committee, Artificial Life X

I regularly serve on the program committees of conferences in my field and referee articles for peer-review journals such as *Adaptive Behavior, Advances in Complex Systems, Artificial Intelligence, Artificial Life, Biosystems, BMC Bioinformatics, BMC Syst. Biology, Cognitive Science, Complex Systems, Data Mining & Knowledge Discovery, J. Biomedical Informatics, Genome Biology, IEEE/ACM Trans. Comp. Biology & Bioinformatics, J. of Theor. Biology, Molec. Biosystems, Nature Communications, Network Science, Nucleic Acids Research, PLOS Computational Biology, PLOS ONE, PNAS, Phys. Rev. E, Science Advances, Scientific Reports*.

Honors and Awards (selected):

2025 Lois B. DeFleur Prize for Academic Achievement, Binghamton University
2019 Research Achievement Award, Luddy School of Informatics, Computing, and Engineering, Indiana University
2015, 2006 Trustees Award for Teaching excellence Award, Indiana University
2016- Fulbright Scholar, United States Department of State.
2009 Professor António Xavier Award, Oeiras Valley Municipal Council, Portugal

Media Coverage (selected):

TUN, Lab Worldide, BingUNews . June 2, 2026. "AI without hallucinations: Binghamton University researchers develop new way to reduce troublesome fake info".
ICT&Health. "Seven-AI voting system reduces medical hallucinations". May 26, 2026.
Times Higher Education. "Using Complex Systems to Predict Disease Progression". Nov. 12, 2025.
Alto Adige, Italy. "The new generation of digital twins is ready". January 21, 2025.
PYMNTS (online). "Making Sense of Spatial Computing and Digital Twins for Industry". Jan. 17, 2025.
Complexity Thoughts. "AI, Cybernetics, and Complexity: unpacking the 2024 Nobel Prizes". Also available as podcast. October 14th, 2024.

Magazine Digest for eLife. "How to build a sperm cell". October 10th, 2024.
National Weather Desk: "Love is in the Air Pressure". Off the Radar Podcast. February 14th, 2024.
New Scientist, "Life may be less chaotic than we thought, say physicists", Dec. 15, 2023.
The Atlantic: "The Calendar of Human Fertility Is Changing", by Katherine Wu. Sep. 22nd, 2023.
National Geographic, CNN Portugal, etc., coverage of our *Epilepsy & Behavior* paper, Mar. 1, 2022.
Publico, Sabado, Jornal Economico, etc., *coverage of our PNAS cover paper "Control pathways in biochemical regulation and signaling"*, Mar. 16, 2021.
Medical News, Globo News, Futurity, Publico, Medical X Press, Diario de Noticias, SIC TV, etc., wide coverage of our work on "Drug Interactions and Gender and Age Biases.", Aug., 2019.
Time, Newsweek, The Independent, The Times, Mother Jones, The Atlantic, etc., Viral coverage of our paper "Human Sexual Cycles are Driven by Culture and Match Collective Moods.", Dec. 2017.
Reddit top trending topic: "After studying 129 countries, academics discovered that the majority of people become more interested in sex and pornography around religious festivals.", Dec 2017.
Nature, "Computer fact-checker and news reader grab attention online.", Jun. 25, 2015.
Various news pieces in *Wired, ACM Tech news, Deutschlandfunk, GCN, Pacific Standard, Phys.org, Scientific Computing, GizModo, Big Think*, etc., about Computational Fact-Checking, Jun. 2015.
EuroScientist, "When Real Science Falls Short In Hollywood.", Jan. 28, 2015.
Nautilus, USA, "Scientists Create Cybernetic Links Between People—by DJing", Apr. 16, 2014.
Público, Portugal, "O cérebro é uma máquina de Turing?", Dec. 27, 2012.
The Daily Telegraph, U.K., "Can we make software that comes to life?", Aug. 5, 2008.
The New York Times, "On Facebook, Scholars Link Up With Data." Dec. 17, 2007.
The Washington Post, "Web May Hold the Key to Achieving Artificial Intelligence", Sep. 6, 2002.

Recent Invited Lectures (selected):

Institute Pascal, Workshop "Building the Inflammation Digital Twin". Paris, France. October 19-23, 2026. Invited Speaker:
Artificial Life (ALIFE 2026) *and Applied Mathematics, Modeling and Computational Science* (AMMCS) *Convention*, Waterloo, Ontario, Canada, August 17-21, 2026. Keynote Speaker: "Canalizing Dynamics as a Network Mechanism for Collective Decision and Control, in Living Systems"
26th Int. Conference on Computational Science (ICCS), Hamburg, Germany, June 29 - July 1, 2026. Keynote Speaker: "Digital Twins for Precision Medicine from a Complex Systems Perspective."
National Institute for Theory and Mathematics in Biology (NITMB), Workshop on "Modularity in Biological Systems", Chicago, IL, USA, April 13-17, 2026. Invited Speaker: "Redundancy shapes the dynamical modularity of biochemical regulation"
Te Ao Mārama: Centre for Fundamental Inquiry, University of Auckland, Auckland, New Zealand, February 16, 2026. Invited Speaker: "canalizing dynamics as a network mechanism for collective decision and emergence in living systems."
Santa Fe Institute, Working Group: "Navigating the space of collective transitions in biology and society", Santa Fe, NM, USA, January 7–9, 2026. Invited Speaker: "canalizing dynamics as a network mechanism for transition control, collective decision, and emergence."
Ancient History Faculty Seminar: AI and Classics, University of Oxford, UK, October 21, 2025. Invited Speaker: "From cybernetics to the global brain: symbolic control in art and creativity."
NetBioMed 2025: Network Science, Biology, and Medicine, Symposium at the International School and Conference on Network Science, Maastricht, the Netherlands, June 2-6, 2025. Keynote Speaker: "Digital Twins for Understanding multiscale factors in disease."
Oeiras Valley Science Festival, TagusPark, Oeiras, Portugal May 24-25, 2025
Colloquium Speaker: "From cybernetics to the global brain: the challenge of complex systems."
Clinical Information Systems Working Group at the *American Medical Informatics Association* (AMIA). April 4, 2025. Colloquium Speaker: "myAURA: a personalized health library for epilepsy management via knowledge graph sparsification and visualization."

Center for Cell Analysis and Modeling, University of Connecticut School of Medicine, Farmington, CT, February 27, 2025. Colloquium Speaker: "Multilayer network inference via redundancy removal: linking the exposome to multiomics in human health."

Department of Mathematics and Computer Science, University of Catania, Italy, March 7, 2025. Colloquium Speaker: "Towards Extraction of Multiscale Factors via Redundancy Removal from the Structure and Dynamics of Complex Networks."

Purdue Systems Collaboratory Summit, Purdue University, West Lafayette, IN, USA, April 29, 2024 *Keynote Speaker*: "Redundancy in the Structure and Dynamics of Complex Networks."

Social Media Mining for Health 2023 (#SMM4H). Workshop at AMIA 2023 Annual Symposium, New Orleans, 11/11/2023. Keynote Speaker: "Network science to characterize health vulnerabilities."

Multiscale & Integrative Complex Networks (MIX-NEXT), Symposium at NetSci 2023, Vienna, Austria, 7/10/2023. Keynote Speaker: "Redundancy in the Structure and Dynamics of Complex Networks."

Network Science for The Sustainable Development Goals, Satellite Symposium NetSci 2023, Vienna, Austria, 7/11/2023. Keynote Speaker: "Network science to characterize health vulnerabilities."

Boolean networks and related modeling frameworks Symposium at the ECMTB 2022, Heidelberg, Germany, 9/19-23/2022. Invited Speaker: "The effective graph: nonlinear canalization in biochemical systems."

Northwestern Institute on Complex Systems Seminar Series, Northwestern University, 5/11/2022 Colloquium Speaker: "Redundancy in the Structure and Dynamics of Complex Systems."

Gender Bias in Artificial Intelligence and Health: Building a Future for Equality, Barcelona Supercomputing Center, Spain, April 14, 2021. Keynote Speaker: "Sex and Gender perspective in Personalized Medicine."

NERCCS 2021: 4th Northeast Reg. Conf. on Complex Systems, March 31-April 2. Keynote Speaker.

5th European Conference on Translational Bioinformatics: Biomedical Big Data Supporting Precision Medicine, Madrid, November 30th-December 1st, 2020. Keynote Speaker.

Universidad Nacional del Sur, Bahia Blanca, Argentina, November 20-22, 2018 Lecture Series Speaker: "Interdisciplinarity, Data Science and Complex Systems".

Network Medicine: Personalized Medicine in the Era of Big Data. NetSci 2018 Satellite. Paris, France. June 11. Keynote Speaker: "The multi-level complexity of human health".

2nd Week of Complexity Sciences, Universidad Nacional Autonoma de México, Mexico City, January 31st 2018. Keynote Speaker: "Multi-level complexity of human health: from drug-interaction to human reproduction".

C. Contributions to Science

1. **Multilevel Complex Systems Approach to Biomedical Data Integration**. My group has been involved in biomedical text mining and NLP from the inception of the field; we participated successfully in the first four *BioCreAtIvE* (Critical Assessment for Information Extraction) challenges between 2004 and 2012. Much biomedical research relies on the inference of correlations and interactions from data at multiple scales of biological organization from the molecular to the social. To understand the biochemical, functional, and behavioral roles of genes and proteins in organisms, there is a need to integrate and study multiscale associations reported in the literature and multiomics databases, as well as exposome data via electronic health records, social media, and digital phenotyping. My contributions toward this goal have been in the development of novel methods based on network science and complex systems. This data-driven approach has enabled the automatic discovery, classification, and annotation of disease factors, protein-protein and drug-drug interactions, biosocial health risks, pharmacokinetics in drug interaction and adverse reaction studies, protein sequence and structure prediction, functional annotation of transcription data, etc..

2. **Redundancy and Control in Complex Networks.** Network science has provided many insights into the organization of complex systems, capturing multilevel interactions as (multilayer) graphs or multivariate dynamical systems. As the field matures, however, there is a need to move from understanding to controlling the organization of complex systems. This is particularly true in

systems biology and medicine, where increasingly accurate models of biochemical regulation are being produced. I have contributed to this advancement with two mathematical concepts developed in my group that allow us to remove different forms of redundancy in networks: distance closures and dynamical canalization. The first concept allows us to extract sparse distance backbone subgraphs, which are sufficient to compute all shortest paths in a weighted graph, and preserve modularity, transmission dynamics, and centrality measures better than existing graph sparsification methods. This has demonstrated that there is massive redundancy in many networks in different domains, whereby most edges in a network are not necessary to compute shortest paths (e.g., 90% of conserved protein interactions in male germ cells). Removing redundant edges facilitates computation and discovery of important regulatory pathways, such as our recent study of genes involved in male infertility. The second concept is used to remove redundancy from the logical rules of biochemical regulation models in systems biology, revealing that most variables (e.g., chemical species) rely on a small subset of their inputs for regulation (canalization). The removal of this redundancy simplifies multivariate dynamics into scalable effective graphs, that reveal how to control large biochemical models, which are otherwise too large to study analytically, such as our study of a large model of pharmacological interventions in breast cancer.

3. **Computational Social Science for Public Health Monitoring**. My group is one of the first to use social media data to study collective social behavior in biomedical problems. For instance, we were the first to use *Instagram* for pharmacovigilance of drug interactions, adverse reactions, and behavior pathology, focusing on depression and epilepsy, and *Facebook* to study sudden death in epilepsy. The work demonstrates that social media provides a very promising source of large-scale data for monitoring the exposome in ways that have not been hitherto possible. Indeed, given the large number of users, social media data allows us to identify under-reported, population-level pathology, environment, and behavior. My group and collaborators have used other sources of large-scale data to tackle problems of biomedical and public-health relevance, such as: *electronic health records* for studying gender and age biases and comorbidity in medical care, *Twitter* and *Google Trend*s data to study human reproductive behavior on a global scale, and *Wikipedia* to automatically establish the veracity of online statements. Another important contribution is the design of human-centered, personalized applications and biomedical corpora.
4. **Multivariate timeseries data and Network Dynamics.** Modern social and biomedical analysis relies on large complex multivariate timeseries data. It is thus essential to develop methods to infer time-varying data associations such as pairwise variable interactions for reverse-engineering networks or discovering subsets of variables that mostly interact with one another (modularity). Our contributions to these problems have been in the areas of spectral methods, statistical inference, information theory, and multilayer network analysis. These are being used to uncover interactions and multiscale modularity in various domains, such as gene regulation, transcriptomics, multiomics and phylogenetics in disease, epidemiology, and brain activity.
5. **Automata Models of Evolutionary Systems**. I have contributed to theoretical biology with computational models of systems whose evolutionary role or adaptive capabilities are not well understood. This includes the first computational model of RNA Editing and an agent-based model of T-Cell cross-regulation. We have also studied how the dynamics of many existing systems biology models relate to their evolvability. More generally, I have contributed to the study of the interplay between self-organization and natural selection by introducing the concept of selected self-organization and developing bioinspired computing and agent-based models to study it.

Complete List of Published Work:
https://scholar.google.com/citations?user=SIouE70AAAAJ&hl=en,
https://orcid.org/0000-0001-9402-887X , or
https://casci.binghamton.edu/publications.php,