

Noah M. Daniels, Ph.D

CONTACT INFORMATION

The University of Rhode Island
Department of Computer Science and Statistics
9 Greenhouse Rd.
Kingston, RI 02881

Email: noah_daniels@uri.edu
GitHub: <https://github.com/URI-ABD/>
Homepage: <http://homepage.cs.uri.edu/~ndaniels/>
Last Updated: January 15, 2025

INTERESTS

Algorithms for “big data,” theoretical data science, computational biology, manifold learning, functional programming, FAIR metadata

EDUCATION

Ph.D. Computer Science, Tufts University, 2013

Dissertation: *Remote Homology Detection in Proteins Using Graphical Models*

Advisor: Lenore J. Cowen

M.S. Computer Science, Tufts University, 2006

B.S. *cum laude* Computer Science, Tufts University, 2001

PROFESSIONAL EXPERIENCE

Associate Professor Department of Computer Science and Statistics The University of Rhode Island Since 2023
Assistant Professor Department of Computer Science and Statistics The University of Rhode Island 2016-2023
Postdoctoral Associate Computation and Biology group (Berger lab), Department of Mathematics & CSAIL, Massachusetts Institute of Technology 2013-2016
Course Head & Lecturer *Machine Structure & Assembly-Language Programming*, Tufts University Fall 2012
Course Head & Lecturer *Exploring Computer Science*, Tufts University Spring 2011
Director of Data Analysis Panjiva, Inc. New York, NY USA 2006-2008
Vice President - Research IntrinsicQ Financial, LLC. Waltham, Massachusetts USA 2004-2007
Systems Engineer Analog Devices, Inc. Wilmington, Massachusetts USA 2001-2004
Network Engineer Net Daemons Associates, Woburn, Massachusetts USA 1999-2001

PUBLICATIONS

1. Hrytsenko Y, **Daniels NM**, Schwartz R 2025. Determining population structure from k-mer frequencies *PeerJ* (to appear).
2. Prior ME, Howard TJ, Light E, Ishaq N, **Daniels NM** 2024. Generalized compression and compressive search of large datasets *2024 IEEE International Conference on Big Data* (to appear).
3. Stubbs B, Shpilker P, Sayers M, Wallace S, **Daniels NM**, Couch A. 2024. AnVILMEDFORD: Documenting Cloud-Based Analysis Projects Using the MEDFORD Metadata Language *2024 IEEE International Conference on Big Data* (to appear).
4. Shpilker P, Stubbs B, Sayers M, Cowen LJ, Wallace S, Couch A, **Daniels NM**. 2024. Cross-referencing Metadata through an extension of the MEDFORD language *Research Conference on Metadata and Semantics Research (MTSR) 2024* (to appear).
5. Murthy MHS, Jasbi P, Lowe W, Kumar L, Olaosebikan M, Roger L, Yang J, Lewinski N, **Daniels NM**, Cowen L, Klein-Seetharaman J. 2024. Insulin signaling and pharmacology in humans and in corals. *PeerJ* 12:e16804 <https://doi.org/10.7717/peerj.16804>.
6. Strand L, Powers A, Shpilker P, Cowen L, Couch A, **Daniels NM** (2023). Context-Sensitive Editing for the MEDFORD Metadata Language *Research Conference on Metadata and Semantics Research (MTSR) 2023* (pp. 278-283). Cham: Springer Nature Switzerland, 2023. https://link.springer.com/chapter/10.1007/978-3-031-65990-4_25
7. Jonas T, **Daniels NM**, Macht G (2023). Electric Vehicle User Behavior: An Analysis of Charging Station Utilization in Canada *Energies* (16(4), 1592) <https://www.mdpi.com/1996-1073/16/4/1592>
8. Kumar L, Brenner N, Sledzieski S, Olaosebikan M, Lynn-Goin M, Klein-Seetharaman R, Berger B, Putnam H, Yang JK, Lewinski N, Singh R, **Daniels NM**, Cowen L, Klein-Seetharaman JK (2022). Transfer of

- Knowledge from Model Organisms to Evolutionarily Distant Non-Model Organisms: The Coral *Pocillopora damicornis* Membrane Signaling Receptome *PLoS ONE* (8(2): e0270965). <https://doi.org/10.1371/journal.pone.0270965>
9. Shpilker P, Freeman J, McKelvie H, Ashley J, Fonticella JM, Putnam H, Greenberg J, Cowen LJ, Couch A, **Daniels NM** (2022). MEDFORD: A human and machine readable metadata markup language *DATABASE - The Journal of Biological Databases and Curation*. (vol. 2022) Oxford University Press. <https://academic.oup.com/database/article/doi/10.1093/database/baac065/6670690>
 10. Howard T, Ishaq N, **Daniels NM** (2021). Clustered Hierarchical Anomaly and Outlier Detection Algorithms, *2021 IEEE International Conference on Big Data* (pp. 5163-5174). <https://arxiv.org/pdf/1908.08551.pdf>
 11. Shpilker P, Freeman J, McKelvie H, Ashley J, Fonticella JM, Putnam H, Greenberg J, Cowen LJ, Couch A, **Daniels NM** (2021). MEtaData Format for Open Reef Data (MEDFORD), *International Conference on Metadata and Semantics Research (MTSR) 2021* (pp. 206-211). Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-030-98876-0_18
 12. Ishaq N, Student G, **Daniels NM** (2019). Clustered Hierarchical Entropy-Scaling Search of Astronomical and Biological Data, *2019 IEEE International Conference on Big Data* (pp. 780-789). <https://arxiv.org/pdf/2103.11774.pdf>
 13. Hrytsenko Y, **Daniels NM**, Schwartz RS (2018). Efficient Distance Calculations Between Genomes Using Mathematical Approximation, *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. (pp. 546-546, abstract). <https://dl.acm.org/doi/abs/10.1145/3233547.3233654> (Note: here, an 'abstract' is a peer-reviewed 1-page manuscript accompanied by a short talk)
 14. Leonard T, Hamel L, **Daniels NM**, Katenka N (2017) Assortative Mixture of English Parts of Speech, *Proc. International Conference on Complex Networks and Their Applications*, (pp. 463-475). https://link.springer.com/chapter/10.1007/978-3-319-72150-7_38 Springer, Cham.
 15. Yuan L, Stevick R, Ahern O, **Daniels NM** (2017). Analysis of 16S Genomic Data using Graphical Databases, *Proceedings of the 2017 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. (pp. 599-599, abstract). <https://dl.acm.org/doi/abs/10.1145/3107411.3108208> (Note: here, an 'abstract' is a peer-reviewed 1-page manuscript accompanied by a short talk)
 16. Berger B, **Daniels NM**, Yu YW (2016). Computational biology in the 21st century: Scaling with compressive algorithms, *Communications of the ACM*, 59 (8): 72–80. <https://dspace.mit.edu/handle/1721.1/116419>
 17. Charles M, Das S, Daniels R, Kirkman L, Delva GG, Destine R, Escalante A, Villegas L, **Daniels NM**, Shigyo K, Volkman SK, Pape JW, Golightly LM (2016). *Plasmodium falciparum* K76T pfcr gene mutations and parasite population structure, Haiti, 2006–2009. *Emerging Infectious Diseases*, 22 (5).
 18. Alberti C, **Daniels NM**, Arrazola MH, Voges J, Lopez AAH, Mattavelli M, Berger B (2016). An Evaluation Framework for Lossy Compression of Genome Sequencing Quality Values, *Proc. IEEE Data Compression Conference* (pp. 221-230) <https://ieeexplore.ieee.org/abstract/document/7786166>
 19. Yu YW, **Daniels NM**, Danko DC, Berger B (2015). Entropy-scaling search of massive biological data, *Cell Systems*, 1 (2) (pp. 130-140). **Cover article (note: Yu and Daniels were co-first authors on this manuscript)**
 20. Daniels R, Rice BL, **Daniels NM**, Volkman S, Hartl DL (2015). The utility of genomic data for Plasmodium vivax population surveillance, *Pathogens and Global Health*, 109.3 (pp. 153-161).
 21. **Daniels NM**, Gallant A, Ramsey N, Cowen L (2015). MRFy: Protein Remote Homology Detection Using Markov Random Fields and Stochastic Search, **IEEE/ACM transactions on computational biology and bioinformatics* 12, no. 1 (2014): (pp. 4-16) **Best Student Paper at ACM-BCB 2013**
 22. Cao M, Zhang H, Park J, **Daniels NM**, Crovella M, Cowen L, Hescott B (2013). Going the Distance for Protein Function Prediction: A New Distance Metric for Protein Interaction Networks, *PLoS ONE* 8(10): e76339. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0076339>
 23. **Daniels NM**, Gallant A, Peng J, Baym M, Cowen L, Berger B (2013). Compressive Genomics for Protein Databases, *Bioinformatics* 29 (13): i283-i290. <https://academic.oup.com/bioinformatics/article/29/13/>

i283/189542

24. **Daniels NM**, Nadimpalli S, Cowen L (2012). Formatt: Correcting Protein Multiple Structural Alignments by Incorporating Sequence Alignment, *BMC Bioinformatics* 13:259 (pp. 1-8) <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-13-259>
25. **Daniels NM**, Gallant A, Ramsey N (2012) Experience Report: Haskell in Computational Biology, *Proceedings of the International Conference on Functional Programming*, Sep 2012 (pp 227-234). <https://dl.acm.org/doi/abs/10.1145/2364527.2364560> **Selected as one of the best papers of ICFP 2012.**
26. **Daniels NM**, Hosur R, Berger B, Cowen L (2012). SMURFLite: combining simplified Markov random fields with simulated evolution improves remote homology detection for beta-structural proteins into the twilight zone, *Bioinformatics* 28 (9): (pp. 1216-1222). <https://academic.oup.com/bioinformatics/article/28/9/1216/311484>
27. **Daniels NM**, Kumar A, Cowen L, Menke M (2012). Touring Protein Space with Matt, *IEEE/ACM Transactions on Bioinformatics and Computational Biology* 9: (pp. 286–293). <https://ieeexplore.ieee.org/abstract/document/6078456>
28. Nadimpalli S, **Daniels NM**, Cowen L (2011). Formatt: Correcting Protein Multiple Structural Alignments by Sequence Peeking, *Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine, Aug 2011*
29. Brady A, Maxwell K, **Daniels NM**, Cowen L (2009). Fault Tolerance in Protein Interaction Networks: Stable Bipartite Subgraphs and Redundant Pathways, *PLoS One* 4(4): e5364. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0005364>
30. Zhou W, Wu W, Palmer N, Mower E, **Daniels NM**, Cowen L, Blumer A (2003). Microarray Data Analysis of Survival Times of Patients with Lung Adenocarcinomas Using ADC and K-Medians Clustering, *Methods of Microarray Data Analysis* (pp. 175-190). Springer, Boston, MA. https://link.springer.com/chapter/10.1007/0-387-23077-7_14
31. Couch A, **Daniels NM** (2001). The Maelstrom: Network Service Debugging via “Ineffective Procedures”, *Proceedings of LISA 2001* (pp. 63-78), San Diego, CA. https://www.usenix.org/legacy/event/lisa2001/tech/full_papers/couch/couch_html/

HONORS, AWARDS, & FELLOWSHIPS

Best Student Paper Award, ACM-BCB 2013
NSF Travel Grant, ISMB 2013
Outstanding Contributor to Engineering Education, Tufts University 2013
SIGPLAN PAC Travel Grant, ICFP 2012
NSF Travel Grant, RECOMB 2012
Fellow, Graduate Institute for Teaching, Tufts University 2010
Best Poster Award, RECOMB 2009

GRANTS FUNDED

Undersea Object Recognition and Anomaly Detection by Manifold Mapping, London Tech Bridge Pathfinder Challenge Series- APEX Undersea Challenge (subaward from Rite-Solutions), (Oct 2021 – Feb 2022) 70K
Sea Floor Object Recognition System, Rhode Island Commerce Corporation, (July 2019 - June 2020) 50K
2019 Big Data Solicitation (Co-I with Dr. Gretchen Macht), University of Rhode Island, (Jun 2019) 3K
Failure Prediction of Networked Equipment, Rhode Island Commerce Corporation, (Jul-Dec 2017) 50K

INVITED TALKS

EXTERNAL

Great Lakes Bioinformatics Conference, Taxonomy and Metadata Panelist, virtual, May 2021
Allerton Conference, University of Illinois Urbana-Champaign, *The Shape of Genomic Data*, Monticello, IL, October 2018
Simons Institute for the Theory of Computing, *Algorithmic Challenges in Genomics*, Berkeley, CA, January 2016
INFORMS 2015, *Big Biological Data: Computational and Analytical Challenges*, Philadelphia, PA, November 2015
113th MPEG Meeting, *Panel on Genome Compression*, Geneva, Switzerland, October 2015

UCLA Institute for Pure & Applied Math, *Workshop on Multiple Sequence Alignment*, January 2015

INTERNAL

University of Rhode Island Computer Science & Statistics Colloquium Series, *Manifold Mapping Enables Fast Search, Anomaly Detection, and More*, Kingston, RI, October 2022

University of Rhode Island Big Data Collaborative, *Clustered Hierarchical Entropy-Scaling Search*, Kingston, RI, December 2019

University of Rhode Island Department of Mathematics, *Entropy-Scaling Search*, Kingston, RI, October 2016

Tufts University, Graduate School of Arts & Sciences and School of Engineering, *Plenary Session of the 16th Annual Graduate Research Symposium*, April 2012

CONFERENCE PRESENTATIONS

Cross-referencing Metadata through an extension of the MEDFORD language **Research Conference on Metadata and Semantics Research (MTSR) 2024** Athens, Greece, November 20, 2024.

MEDFORD: A human- and machine-readable markup language to facilitate FAIR coral metadata **IEEE Pacific Symposium on Biocomputing** virtual, January 8, 2021.

Metadata for Coral RNASeq **International FAIR Convergence Symposium 2020** virtual, November 27, 2020.

Clustered Hierarchical Entropy-Scaling Search of Astronomical and Biological Data **IEEE Big Data 2019** Los Angeles, California, December 12, 2019.

Analysis of 16S Genomic Data using Graphical Databases **ACM-BCB 2017** (poster session) Boston, MA, August 21st, 2017.

Entropy-scaling search of massive biological data **RECOMB 2016** (highlights talk) Santa Monica, California, April 18, 2016.

Entropy-scaling search of massive biological data **ISMB 2015** (late-breaking research) Dublin, Ireland, July 14, 2015.

Compressive Metagenomics **BioData 2014** (poster session) Cold Spring Harbor Laboratory, New York, November 6, 2014.

MRfY: Protein Remote Homology Detection Using Markov Random Fields and Stochastic Search **ACM-BCB 2013 (Best Student Paper Award)** Washington, D.C., September 23, 2013.

Compressive Genomics for Protein Databases **ISMB 2013** Berlin, Germany, July 22, 2013.

Experience Report: Haskell in Computational Biology **ICFP 2012** Copenhagen, Denmark, September 11, 2012.

SMURFLite: combining simplified Markov random fields with simulated evolution improves remote homology detection for beta-structural proteins **RECOMB 2012** (poster session) Barcelona, Spain, April 21, 2012.

Format: Correcting Protein Multiple Structural Alignments by Sequence Peeking **ACM-BCB 2011** Chicago, IL, August 2, 2011.

Touring Protein Space with Matt **ISMB 2010**, 3D SIG (poster session) Boston, MA, July 9, 2010.

Touring Protein Space with Matt **ISBRA 2010** Storrs, CT, May 24, 2010.

Fault Tolerance in Protein Interaction Networks: Stable Bipartite Subgraphs and Redundant Pathways **RECOMB 2009** (poster session, **Best Poster Award**). Tucson, AZ, May 18, 2009.

COURSES TAUGHT AND EDUCATIONAL ACTIVITIES

Exploring Computer Science, an introductory course for non-majors, Tufts University, Spring 2011

Machine Structure & Assembly-Language Programming, a mid-level undergraduate course required for majors, Tufts University, Fall 2010 (with Norman Ramsey) & Fall 2012

Advanced Functional Programming, advanced topics in the Haskell language, MIT Splash, Fall 2015

Design and Analysis of Algorithms, URI, Fall 2016, Fall 2017, Spring 2018, Spring 2019, Spring 2020, Spring 2021, Spring 2022, Spring 2023

Object-Oriented Programming, URI, Spring 2017, Fall 2017, Spring 2018, Fall 2018

Machine Organization, URI, Fall 2018, Fall 2019, Spring 2020, Fall 2020, Spring 2021, Fall 2021, Spring 2022, Fall 2022, Spring 2023

Algorithms for Big Data, URI, Spring 2017, Fall 2019, Fall 2021, Fall 2023

Introduction to Bioinformatics, URI, Spring 2019

Theory of Computation, URI, Fall 2020, Fall 2022

Machine Learning for Data Science, URI (online course), Fall 2021, Fall 2022
Supervisor for Undergraduate Research and Innovation scholars (URI²), URI, Summer 2020
Supervisor for Arts & Sciences Fellow, URI, Summer 2022
Supervisor for URI Honors Student, URI, 2022-2023
Supervisor for Computer Science Research Capstone Projects, URI, 2016-present (24 students up to now)

GRADUATE STUDENT
THESIS SUPERVISION

DOCTORAL

Martin Hellwig (2019) Dissertation Topic: *Computational Gene Association Based on Relative Species Displacement in Gene and Morphology Trees*

Yana Hrytsenko (2022) Dissertation Topic: *Clustering of Cancerous Genomes Using Efficient Distance Calculations*

MASTERS

Imam Kamran (2018) Thesis Topic: *Personal Health Card: Use of QR Code to Access Medical Data*

Martha Bellows (2018) Thesis Topic: *Exploration of Classifying Sentence Bias in News Articles with Machine Learning Models*

Ali Amani (2021) Thesis Topic: *Data Dimensionality Reduction Through Cluster Trees and Manifold Learning*

PROFESSIONAL
ACTIVITIES

EXTERNAL

Reviewer, Expert Systems With Applications, 2023
Reviewer, NIH/NLM Study Section, July 2021
Reviewer, *Bioinformatics*, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020
Reviewer, *PLoS ONE*, 2021
Reviewer, *ACM Transactions in Computational Biology and Bioinformatics*, 2013, 2014, 2018
Reviewer, *Journal of the Royal Statistical Society*, 2018, 2019
Reviewer, *Computers in Biology and Medicine*, 2018
Reviewer, *IEEE Spectrum*, 2018
Reviewer, *Nucleic Acids Research*, 2012, 2014, 2018, 2019
Reviewer, *PLoS Computational Biology*, 2017
Program Committee Member, ACM-BCB 2017
Scholarship Judge, Davidson Institute for Talent Development, 2017
Delegate, Motion Picture Expert Group, 2015 & 2016
Reviewer, RECOMB, RECOMB, 2013, 2014, 2015, 2016, 2017
Program Committee Member, ISMB, 2014, 2015, 2020, 2021, 2022, 2023
Program Committee Member, ACM-BCB, 2013
Reviewer, Workshop on Algorithms for Bioinformatics, 2012

INTERNAL

Senator, URI Faculty Senate, 2023-2026 term
Chair, URI Faculty Senate Committee on Technology and Infrastructure, 2023-2025 term
Chair, Tenure-Track Computer Science Hiring Committee, Dept. of Computer Science & Statistics, URI, 2022-2023
Chair, System Staff Hiring Committee, Dept. of Computer Science & Statistics, URI, 2022
Chair, Systems Committee, Dept. of Computer Science & Statistics, URI, since 2020
Member, Lecturer Hiring Committee, Dept. of Computer Science & Statistics, URI, 2021-2022
Member, Lecturer Hiring Committee, Dept. of Computer Science & Statistics, URI, 2019-2020
Member, Tenure-Track Computer Science Hiring Committee, Dept. of Computer Science & Statistics, URI, 2018-2019
Member, Tenure-Track Data Science Hiring Committee, Dept. of Computer Science & Statistics, URI, 2018-2019
Member, Graduate Committee, Dept. of Computer Science & Statistics, URI, 2018-2023
Member, Undergraduate Committee, Dept. of Computer Science & Statistics, URI, since 2016
Member, SPIN+ Executive Committee, URI, 2016-2018

PROFESSIONAL SOCIETY MEMBERSHIPS

Member, [International Society for Computational Biology](#), since 2010
Member, [Association for Computing Machinery](#), since 2011
Member, [Institute of Electrical and Electronics Engineers](#), since 2019

SOFTWARE PRODUCTS

CHAODA (Software implementation in Python for fast anomaly detection in high-dimensional spaces) <https://github.com/URI-ABD/chaoda>
CLAM (Software implementations in Rust and Python for entropy-scaling search of high-dimensional data, including a Rust implementation of CHAODA) <https://github.com/URI-ABD/clam>
medford (Software implementation Python for of the MEDFORD metadata description language) <https://github.com/TuftsBCB/clam>