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: Szolovits, Peter

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

POSITION TITLE: Professor of Computer Science and Engineering (Post-Tenure), Professor of Health Sciences and Technology

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
California Institute of Technology, Pasadena, CA	BS	05/1970	Physics
California Institute of Technology, Pasadena, CA	PhD	10/1974	Information Science

A. Personal Statement

I direct the Clinical Decision Making Group at the MIT Computer Science and Artificial Intelligence Laboratory. That group is engaged in a variety of research efforts, including: clinical decision support systems; building predictive models for patient outcomes by extrapolating from large patient cohorts; improving natural language processing methods to extract clinical data from doctors' and nurses' notes and de-identifying clinical narratives; combining lab, medication, notes, reports, discharge summary, imaging, and genetic/genomic data to better understand diseases and the state of individual patients; "big data" techniques to allow analysis of vast and growing data sets; modeling temporal evolution of patient states; and technologies to help protect the confidentiality of electronic medical records. We also focus on issues of explainability and fairness of machine learning models. I have supervised 42 doctoral theses, served on the committees of another 83 doctoral students, and supervised nearly a hundred master's theses.

B. Positions, Scientific Appointments and Honors

Positions and Employment

2025- Professor (Post-Tenure) of Computer Science and Engineering, MIT
2001- Professor of Health Science and Technology, Harvard / MIT Division of HST, MIT Institute for Medical and Engineering and Science
1997-98 Visiting Professor of Medicine, Stanford University
1993- Professor of Computer Science and Engineering, MIT
1989-90 Visiting Associate Professor of Medicine, Stanford University
1983 Visiting Scientist, University of Paris VI
1980-93 Associate Professor of Computer Science and Engineering, MIT
1974-80 Assistant Professor of Computer Science and Engineering, MIT

Other Experience and Professional Memberships

2023- United Healthcare Responsible Use of AI External Board
2018 NLM Blue Ribbon Panel on review of internal research
2013-17 NLM Biomedical Library and Informatics Review Committee
2010-11 Comm. on Future Information Architectures, Processes, and Strategies for the Centers for Medicare and Medicaid Services, NAS/NRC
2008-14 National Research Council, Computer Science and Telecommunications Board

2005- AIMBE
2005- National Academy of Medicine
2007-08 Comm. on Engaging the Comp. Sci. Research Community in Health Care Informatics, NAS/NRC
2000 Committee on the Role of Institutional Review Boards in Health Services Research Data Privacy Protection, IOM
1996-97 Committee on Maintaining Privacy and Security in Health Care Applications of the National Information Infrastructure, NAS/NRC
1986 IEEE task force on research needs for biomedical engineering systems
1980- AAAI
1977- AMIA (and predecessors)
1972- ACM

Honors

2019 Burgess and Elizabeth Jamieson Teaching Award, MIT-EECS
2019 Top 100 AI Leaders in Drug Discovery & Advanced Healthcare
2017 Founding Member of the International Academy of Health Sciences Informatics
2016 Intelligence in Medicine Award, from Pediatrics 2040
2013 Morris F. Collen Award of Excellence (American College of Medical Informatics)
2009 Thomas McMahon Mentoring Award, Harvard/MIT Division of Health Sciences and Technology
2006 Harvard Medical School Center for Biomedical Informatics Award for Innovation in Personally Controlled Health Record Infrastructure
2005 Elected Fellow of the American Institute for Medical and Biological Engineering
2005 Elected to the National Academy of Medicine (then known as the Institute of Medicine)
1992 Elected Fellow of the American Association for Artificial Intelligence
1984 Elected Fellow of the American College of Medical Informatics
1981, 1988 Whitaker Health Sciences Fund Awards
1970 NSF Fellowship (declined)
1970-74 Fannie and John Hertz Foundation Fellowship

C. Contributions to Science

1. I was one of the pioneers of the field of "Artificial Intelligence in Medicine," which attempts to model human clinical decision-making and to implement similar techniques in computer programs as a basis for decision support. Such programs may perform diagnosis, therapy selection and monitoring, prediction of the likelihood of future events and outcomes, critiques of planned interventions, and explanations of the basis of their reasoning. Technical challenges include how to represent clinical and pragmatic knowledge in such programs and how to make inferences from that knowledge and facts about individual cases. One of my earliest papers in this field reviewed the state of the art of medical diagnosis programs in *Artificial Intelligence* and was recognized as a most frequently cited article in that journal. I edited and contributed to an early book of influential articles on this topic, and have continued to synthesize ideas in later years.
   P. Szolovits and S. G. Pauker. Categorical and probabilistic reasoning in medical diagnosis. *Artificial Intelligence*, 11:115-144, 1978. Reprinted in W. J. Clancey and E. H. Shortliffe, *Readings in Medical Artificial Intelligence: The First Decade*, Addison-Wesley, 1984.
   R. S. Patil, P. Szolovits, and W. B. Schwartz. Causal understanding of patient illness in medical diagnosis. In *Proc. Seventh International Joint Conference on Artificial Intelligence*, 893-899, 1981. Reprinted in W. J. Clancey and E. H. Shortliffe, *Readings in Medical Artificial Intelligence: The First Decade*, Addison-Wesley, 1984.
   W. B. Schwartz, R. S. Patil, and P. Szolovits. Artificial intelligence in medicine. Where do we stand? N *Engl J Med*, 316(11):685–688, Mar 1987. PMID: 3821801
   Patel, V L, E H Shortliffe, M Stefanelli, P Szolovits, M Berthold, R Bellazzi, A Abu-Hanna. The Coming of Age of Artificial Intelligence in Medicine. Artif Intell Med 2009 May; 46(1):5-17. PMC2752210

2. Nearly thirty years ago, I advocated for the adoption of life-long personal health information systems based on the World Wide Web technologies that could integrate a lifetime of medically relevant data, educate patients about their conditions, provide decision support to allow their subjects and families to participate in their own care, and provide communication support with their providers and peers. Although this vision has not yet come to pass, my work was highly influential on my colleagues at Boston Children's Hospital, who developed the personal health record systems Ping, Indivo, Dossia and the more recent SMART on FHIR, commercial efforts such as Google Health and Microsoft Healthvault, and organizations such as the Health Record Banking Alliance. I also helped to pioneer Web-based institutional medical record systems.
   Peter Szolovits, Jon Doyle, William J. Long, Isaac S. Kohane, and Stephen G. Pauker. Guardian Angel: Patient-centered health information systems. Technical Report 604, MIT Laboratory for Computer Science, 1994.
   K. D. Mandl, P. Szolovits, and I. S. Kohane. Public standards and patients' control: how to keep electronic medical records accessible but private. *BMJ*, 322(7281):283–287, Feb 2001. PMC1119527
   D. M. Rind, I. S. Kohane, P. Szolovits, C. Safran, H. C. Chueh, and G. O. Barnett. Maintaining the confidentiality of medical records shared over the Internet and the World Wide Web. *Ann Intern Med*, 127(2):138–141, Jul 1997. PMID: 9230004
   H. S. F. Fraser, P. Biondich, D. Moodley, S. Choi, B. W. Mamlin, and P. Szolovits. Implementing electronic medical record systems in developing countries. *Inform Prim Care*, 13(2):83–95, 2005. PMID: 15992493

3. Since the early 2000's, I have worked on adoption of natural language processing techniques to extract from articles and narrative notes meaningful mentions of diseases, signs and symptoms, medications, procedures, test results, etc., to identify temporal, causal and associational relationships among these, and to identify and pseudonymize protected health information from clinical notes using various ML techniques, recently including large language models, with special concern for vulnerabilities to bias.
   Luo, Y., Xin, Y., Hochberg, E., Joshi, R., Uzuner, Ö., & Szolovits, P. (2015). Subgraph Augmented Non-Negative Tensor Factorization (SANTF) for Modeling Clinical Narrative Text. *J Am Med Inform Assoc*. 2015 Apr 9. PMCID: PMC4986663
   Yu, S., Liao, K. P., Shaw, S. Y., Gainer, V. S., Churchill, S. E., Szolovits, P., et al. (2015). Toward high-throughput phenotyping: unbiased automated feature extraction and selection from knowledge sources. *J Am Med Inform Assoc.* 2015 Apr 29. PMC4986664
   Jin, D, Jin, Z, Zhou, JT, Szolovits P. Is BERT really robust? A strong baseline for natural language attack on text classification and entailment. In: Proceedings of the AAAI Conference on Artificial Intelligence 2020. p. 8018–25.
   Szolovits P. Large Language Models Seem Miraculous, but Science Abhors Miracles. NEJM-AI 2024 May 23;1(6).

4. Current medical research depends on finding a fruitful combination of genetic, environmental and clinical factors that can explain the genesis and progression of disease. I have participated in various collaborations to identify highly specific patient cohorts and within these to find the novel relationships. The methods apply to a broad variety of diseases, and we have investigated rheumatoid arthritis, inflammatory bowel disease, psychiatric illness, autism, heart disease and type II diabetes.
   K. P. Liao, T. Cai, V. Gainer, S. Goryachev, Q. Zeng-treitler, S. Raychaudhuri, P. Szolovits, S. Churchill, S. Murphy, I. Kohane, E. W. Karlson, and R. M. Plenge. Electronic medical records for discovery research in rheumatoid arthritis. *Arthritis Care Res* (Hoboken), 62(8):1120–1127, Aug 2010. PMC3121049
   Sinnott JA, Dai W, Liao KP, Shaw SY, Ananthakrishnan AN, Gainer VS, Karlson EW, Churchill S, Szolovits P, Murphy S, Kohane I, Plenge R, Cai T. Improving the Power of Genetic Association Tests with Imperfect Phenotype Derived from Electronic Medical Records. *Human Genetics* 133:1369–1382 (2014). PMC4185241
   Luo, Y., Riedlinger, G., & Szolovits, P. (2014). Text Mining in Cancer Gene and Pathway Prioritization. *Cancer Informatics*, 69–79. PMC4216063
   Luo, Y., Eran, A., Palmer, N., Avillach, P,Levy-Moonshine, A., Szolovits, P., and Kohane, I. S. A multidimensional precision medicine approach identifies an autism subtype characterized by dyslipidemia. *Nat Med* **26** 1375-1379 (2020) PMID: 32778826

5. With the growing availability of vast clinical data sets, I have led and contributed to machine learning efforts to build predictive models that estimate the likely outcomes of various interventions and analytic models that provide accurate characterizations of patients from imaging and text. These can form the basis for clinical decision support.
   - Liao KP, Cai T, Gainer V, Goryachev S, Zeng-treitler Q, Raychaudhuri S, Szolovits P, Churchill S, Murphy S, Kohane I, Karlson EW, Plenge RM. Electronic medical records for discovery research in rheumatoid arthritis. Arthritis Care Res (Hoboken). 2010 Aug;62(8):1120-7. PMC3121049
   - Goehler A, Harry Hsu TM, Lacson R, Gujrathi I, Hashemi R, Chlebus G, Szolovits P, Khorasani R. Three-Dimensional Neural Network to Automatically Assess Liver Tumor Burden Change on Consecutive Liver MRIs. J Am Coll Radiol. 2020 Nov;17(11):1475-1484. doi: 10.1016/j.jacr.2020.06.033. Epub 2020 Jul 25. PubMed PMID: 32721409.
   - Boag W, Kovaleva O, McCoy TH Jr, Rumshisky A, Szolovits P, Perlis RH. Hard for humans, hard for machines: predicting readmission after psychiatric hospitalization using narrative notes. Transl Psychiatry. 2021 Jan 11;11(1):32. doi: 10.1038/s41398-020-01104-w. PubMed PMID: 33431794; PubMed Central PMCID: PMC7801508.
   - Lehman E, Hernandez E, Mahajan D, Wulff J, Smith MJ, Ziegler Z, et al. Do We Still Need Clinical Language Models? In: CHIL 2023. New York, NY, USA

Complete List of Published Work in MyBibliography:
https://www.ncbi.nlm.nih.gov/myncbi/peter.szolovits.1/bibliography/public/