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: Hochheiser, Harry S

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

POSITION TITLE: Associate Professor, Biomedical Informatics, University of Pittsburgh

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
Massachusetts Institute of Technology, Cambridge, Massachusetts, USA	BS	06/1991	Computer Science and Engineering
Massachusetts Institute of Technology, Cambridge, Massachusetts, USA	MS	06/1991	Electrical Engineering and Computer Science
University of Maryland, College Park, MD, USA	PhD	06/2003	Computer Science
National Institute on Aging, Baltimore, MD, USA	Post-Doctoral	07/2006	Computational Biology

A. **Personal Statement**

A Computer Scientist by training, I have been active in biomedical informatics research for the past 15 years. My research takes a broad view of the field, focusing on the application of techniques ranging from human-computer interaction to data science and machine learning, all with the goal of making data and computing tools more useful and usable for researchers, clinicians, patients, policymakers, and the general public.

Much of this work has focused on clinical informatics and decision support. I am currently a co-PI on an NINDS project aimed at using clinical data to predict neuromorbidities in pediatric ICU patients (R01NS118716). This project includes both the development of machine learning models and user interfaces for displaying these predictions to clinicians. To design these interfaces, we will be conducting qualitative inquiries with clinicians. I have previously used similar techniques on an NLM-funded project R01 LM012095, (PI Visweswaran) project involving the development of a learning electronic medical record (King, et al. 2018, 2017, 2015). Other relevant efforts include my work on an NLM-funded (1R01LM010964, PI Chapman) project investigation of interactive natural language processing (Trivedi, et al. 2017), and an AHRQ-funded (5R01HS021290, PI AGHA) investigation of EMR usage patterns in physician-patient encounters. These projects have involved a combination of ethnographic and human-computer interaction methods with design and empirical evaluations.

For the past 10 years, I have been working on the development of natural language processing tools for cancer data, through the DeepPhe project (U24CA2480, co-PIs G. Savova and J. Warner; UH3CA243120, co-PIS G. Savova, J. Warner, E. Durbin). For this project we have developed tools for extracting patient treatment timelines from clinical notes, data models for representing those timelines, visualization tools for displaying extracted data, and web service architectures for integrating these tools into existing software stacks.

The third major focus of my current efforts is on team science support for infectious disease modeling research. Since 2021, I have been PI and director of the Models of Infectious Disease Agent Study (MIDAS) Coordination Center (R24GM153920), funded by the NIGMS to support a community of over 1000 active researchers. In addition to providing a website, running a webinar series, and coordinating community activities, the MIDAS Coordination Center supports the Scenario Modeling Hub, a consortium of researchers developing medium-term forecasts of infectious disease outbreaks under possible future scenarios. The

MIDAS Center also conducts novel research on the reproducibility of infectious disease modeling work and the application of data science principles to infectious disease modeling.

My research has covered a range of other topics, including information visualization, bioinformatics, universal usability, security, privacy, and public policy implications of computing systems. I have published more than 80 peer-reviewed journal and conference papers, two book chapters, and I am co-author of Research Methods in Human-Computer Interaction, 2nd Edition (Morgan Kaufmann, 2017). I am also director of the Pittsburgh Biomedical Informatics Training Program (T15 LM007059).

Citations most relevant to this proposal:

1. Lazar J, Feng JH, Hochheiser H Research Methods in Human-Computer Interaction, 2nd edition. 2017 Morgan Kaufmann (3005 citations as of May 27, 2023).
2. Bitterman DS*, Goldner E*, Finan S, Harris D, Durbin EB, **Hochheiser H**, Warner JL, Mak RH, Miller T, Savova GK. An end-to-end natural language processing system for automatically extracting radiotherapy events from clinical texts. Int J Radiat Oncol Biol Phys. 2023 Mar 26;117(1):262-273. doi: 10.1016/j.ijrobp.2023.03.055. PMID: 36990288; PMCID: PMC10522797
3. Barda A., Horvat, CM, **Hochheiser H.** A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. BMC Medical Informatics and Decision-Making (2020) 20: 257. DOI: 10.1186/S12911-020-01276-x PMID: 33032582 PMCID: PMC754557
4. Calzoni L, Clermont G, Cooper GF, Visweswaran S, **Hochheiser H.** Graphical Presentations of Clinical Data in a Learning Electronic Medical Record. Applied Clinical Informatics (2020) 11(04) DOI: 10.1055/s-0040-1709707. PMID: 33058103 PMCID: PMC7560537

Projects that I would like to highlight include:

R01NS118716-01A1
Au, Clark, Hochheiser, Horvat (MPI)
08/01/2021-07/31/2026
Bio-digital Rapid Alert to Identify Neuromorbidity

R24GM153920
Hochheiser (PI)
9/15/2024-06/30/2029
MIDAS Coordination Center

U24CA248010
Savova, Hochheiser, Warner (MPI)
09/24/2020-08/31/2025
Cancer Deep Phenotype Extraction from Electronic Medical Records

UH3CA243120
Savova, Durbin, Hochheiser, Warner (PI)
07/19/2021-06/30/2024
Natural Language Processing Platform for Cancer Surveillance

T15LM007059
Hochheiser (PI)
07/01/2017-06/30/2022
Pittsburgh Biomedical Informatics Training Program

B. Positions, Scientific Appointments, and Honors

Positions and Scientific Appointments

2020- Present	Associate Professor, Clinical and Translational Science, University of Pittsburgh, Pittsburgh PA
2017- Present	Director, Biomedical Informatics Training Program, University of Pittsburgh, Pittsburgh, PA
2017-Present	Associate Professor, Biomedical Informatics and Intelligent Systems Program, University of Pittsburgh, Pittsburgh, PA

2012 – 2017	Assistant Professor, Intelligent Systems Program, University of Pittsburgh, Pittsburgh, PA
2009 – 2017	Assistant Professor, Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania
2009 - Present	Member, American Medical Informatics Association
2006 - 2009	Assistant Professor, Computer and Information Sciences, Towson University, Towson, MD
2004 - Present	Member, Association of Computing Machinery
2003 – 2006	Postdoctoral Researcher, Laboratory of Genetics, National Institute on Aging, Baltimore, MD
2002	Doctoral Consortium, ACM Special Interest Group on Computer-Human Interaction
2002 – 2003	America Online Fellowship in Human-Computer Interaction
1999	Summer Intern (June-September 1999), IBM T.J. Watson Labs, Hawthorne, NY
1998 - 2003	Graduate Student, Computer Science, University of Maryland, College Park, MD
1996 - 1998	Consultant, H. Systems, Princeton, NJ
1993 - 1995	Software Developer, Biomedical Engineering, Massachusetts General Hospital, Boston, Massachusetts, USA 1995 Software Developer, AT&T Bell Labs, Murray Hill, NJ
1990 - 1991	Graduate Student, Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, MA
1991 - 1993	Research Staff, Tufts University School of Medicine, Boston, MA

Honors

2002 - 2003	America Online Fellowship in Human-Computer Interaction
2002	Doctoral Consortium, ACM Special Interest Group on Computer-Human Interaction

C. Contributions to Science

1. **Clinical and clinical research human-computer interaction**: I have used qualitative and quantitative methods to design and evaluate tools for a variety of domains, including interacting with results of NLP systems, information needs for curation of drug-drug interaction data, clinical pharmacogenomics, and and inquiries into barriers in documentation for nursing staff.

a. Trivedi G, Pham P, Chapman WW, Hwa R, Wiebe J, **Hochheiser H** NLPReViz: an interactive tool for natural language processing on clinical text. JAMIA 2017. DOI: 10.1093/jamia/ocx070 PMID: 29016825 PMCID: PMC6381768
b. Romagnoli, KM, Nelson SD, Hines L, Empey P, Boyce R, **Hochheiser H** Information needs for making clinical recommendations about potential drug-drug interactions: a synthesis of literature review and interviews BMC Medical Informatics and Decision Making 2017 1721 Doi: 10.1186/s12911-017-0419-3 PMID: 28228132, PMCID:PMC5322613.
c. Romagnoli KM, Boyce R, Empey PE, Adams S, **Hochheiser H.** Bringing clinical pharmacogenomics information to pharmacists: a qualitative study of information needs and resource requirements. International Journal of Medical Informatics. 2016 Feb; 86:54-61.DOI:10.1016/j.ijmedinf.2015.11.015 PMID: 26725696. PMCID: PMC4720137.
d. Kohle-Ersher A, Chatterjee P, Osmanbeyoglu HU, **Hochheiser H**, Bartos C. Evaluating the barriers to point-of-care documentation for nursing staff. Comput Inform Nurs doi: 10.1097/NCN.0b013e3182343f1. 2012 Mar; 30 (3):126-33. PMID: 22024972

2. **Cancer Informatics:** For the "DeepPhe" project (U24CA248010, UH3CA243120), we are developing natural language processing and visual analytics tools in support of extracting and analyzing longitudinal patient histories from cancer clinical notes, including a containerized implementation providing API-based access to the DeepPhe NLP tools, in support of integration of DeepPhe NLP functionality with cancer registrar workflows; a visual analytics tools; and a computable information model.

a. **Hochheiser H**, Finan S, Yuan Z, Durbin ED, Jeong JC, Hands I, Rust D, Kavuluru R, Wu XC, Warner JL, Savova G. DeepPhe-CR: Natural Language Processing Software Services for Cancer Registrar Case Abstraction medRxiv May 5 2023; DOI: 10.1101/2023.05.05.23289524
b. Yuan, Z, Finan, S, Warner, J, Savova, G, **Hochheiser, H**. Interactive Exploration of Longitudinal Cancer Patient Histories Extracted from Clinical Text JCO Cancer Clinical Informatics May 2020 DOI: 10.1200/CCI.19.00115 PMID: 32383981 PMCID: PMC7265796
c. Savova, GK, Tseytlin E, Finan E, Castine M, Miller T, Medvedeva O, Harris D, **Hochheiser H**, Lin

C, Chavan G, Jacobson RS. DeepPhe: A Natural Language Processing System for Extracting Cancer Phenotypes from Clinical Records, Cancer Research 77(21), November 2017 DOI: 10.1158/0008- 5472.CAN-17-0615. PMID: 29092954, PMCID: PMC5690492.

d. **Hochheiser H**, Castine M, Harris D, Savova G, Jacobson RS An information model for computable cancer phenotypes. BMC Medical Informatics and Decision Making 2016 16:121. doi: 10.1186/s12911-016-0358-4. PMID: 27629872 PMCID: PMC5024416.

3. **Natural Language Processing**: I have worked on a number of projects exploring the application of NLP to extract information from clinical text, including the development of a novel NLP pipeline for review of information regarding reproductive health counseling for young adults with epilepsy, a conceptual model of translational NLP, and tools for extracting incidental findings in radiology notes and exploring those findings in the context of the notes.

a. Harrison E, Kirkpatrick LA, **Hochheiser HS**, Sogawa Y, Kazmerski TM. A retrospective textual analysis of sexual and reproductive health counseling for adolescent and young adult people with epilepsy of gestational capacity. Epilepsy Behav. 2023 Jun 20;145:109321. doi:10.1016/j.yebeh.2023.109321. PMID: 37348408

b. Newman-Griffis D, Lehman JF, Rosé C, **Hochheiser H**. Translational NLP: A New Paradigm and General Principles for Natural Language Processing Research. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Jun; 2021:4125-4138 DOI: 10.18653/v1/2021.naacl-main.325 PMID: 34179899 PMCID: PMC8223521

c. Trivedi G, Dadashzadh ER, Handzel RM, Chapman WW, Visweswaran S, **Hochheiser H**. Interactive NLP in Clinical Care: Identifying Incidental Findings in Radiology Reports. Appl Clin Inform 2019;10:4, 655–69. DOI: 10.1055/s-0039-1695791. PMID: 31486057; PMCID: PMC6727024

d. Trivedi G, Hong C., Dadashzadeh E, Handzel RM, **Hochheiser H**, Visweswaran S. Identifying incidental findings from radiology reports of trauma patients: an evaluation of automated feature representation methods IJMI 2019 129: 81-87. DOI: 10.1016/j.ijmedinf.2019.05.021 PMID: 31445293 PMCID: PMC6717529

4. **Knowledge Representation**: To increase the utility and reusability of research data, I have been involved in several projects focused on ontologies and data models, including the development of information models for potential drug-drug interactions, tools for representing and interacting with drug-drug information evidence, vocabularies for cancer treatment regimens, and an ontology for craniofacial disorders.

a. **Hochheiser H**, Jing X, Garcia EA, Avvaz S, Sahay R, Dumontier M, Banda JM, Beyan O, Brochhausen M, Draper E, Habiel S, Hassanzadeh O, Herrero-Zazo M, Hokum B, Horn B, LeBaron B, Malone DC, Nytrø Ø, Reese T, Romagnoli K, Schneider J, Zhang LY, Boyce RD. Minimal Information Model for Potential Drug-Drug Interactions. Frontiers in Pharmacology 2021 08 March. DOI: 10.3389/fphar.2020.608068 PMID: 33762928 PMCID: PMC7982727

b. Grizzle AJ, Hines LE, Malone DC, Kravchenko O, **Hochheiser H**, Boyce RD. Testing the face validity and inter-rater agreement of a simple approach to drug-drug interaction evidence assessment. J Biomed Inform. 2020 Jan;101:103355. Epub 2019 Dec 12. DOI: 10.1016/j.jbi.2019.103355. PMID: 31838211; PMC7537787.

c. Warner JL, Dymshyts D, Reich CG, Gurley MJ, **Hochheiser H**, Moldwin ZH, Belenkaya R, Williams AE, Yang PC. HemOnc: A New Standard Vocabulary for Chemotherapy Regimen Representation in the OMOP Common Data Model. J Biomed Inform. 2019 Jun 22. doi: 10.1016/j.jbi.2019.103239 PMID:31238109; PMCID: PMC6697579.

d. Brinkley JF, Borromeo C, Clarkson M, Cox TC, Cunningham MJ, Detwiler LT, Heike CL, **Hochheiser H**, Mejino L, Travillian RS, Shapiro LG. The Ontology of Craniofacial Development and Malformation for translational craniofacial research. Seminars in Medical Genetics DOI: 10.1002/ajmg.c.31377. 2013 Oct 4; 1-14. PMCID: PMC4041627. PMID: 24124010

5. **Infectious Disease Modeling**: As Director of the MIDAS Coordination Center (3U24GM132013), I lead a time actively working on supporting infectious disease modeling effort. Research efforts have focused in two areas: 1) The development and evaluation of a framework for increasing reproducibility in infectious disease modeling. 2) Under my leadership, the MCC team has provided technical support for the COVID-19 Scenario

modeling Hub, a community effort aimed at developing long-term projections of pandemic trajectories under differing future conditions. We have validated data, developed visualizations, and supported the operation of this effort, which was cited by the White House as influencing decisions on vaccination policy.

a. Pokutnaya D, Childers B, Arcury-Quandt AE, **Hochheiser H,** Van Panhuis WG An Implementation Framework to Improve the Transparency and Reproducibility of Computational Models of Infectious Diseases. PLoS Computational Biology. March 2023 DOI: 10.1371/journal.pcbi.1010856 PMID: 36928042 PMCID: PMC10019712.
b. Pokutnaya D, Van Panhuis WG, Childers B, Hawkins MS, Arcury-Quandt AE, Matlack M, Carpio K, **Hochheiser H** Inter-rater reliability of the Infectious Disease Modeling Reproducibility Checklist (IDMRC) as applied to COVID-19 computational modeling research. March 2023 medRxiv DOI: 10.1101/2023.03.21.23287529 PMID: 36993426 PMCID: PMC10055605.
c. Truelove S, Smith CP, Qin M, Mullany LC, Borchering RK, Lessler J, Shea K, Howerton E, Contamin L, Levander J, Salerno J, **Hochheiser H**, Kinsey M, Tallaksen K, Wilson S, Shin L, Rainwater-Lovett K, Lemaitre JC, Dent J, Kaminsky J, Lee EC, Perez-Saez J, Hill A, Karlen D, Chinazzi M, Davis JT, Mu K, Xiong X, Y Piontti AP, Vespignani A, Srivastava A, Porebski P, Venkatramanan S, Adiga A, Lewis B, Klahn B, Outten J, Schlitt J, Corbett P, Telionis PA, Wang L, Peddireddy AS, Hurt B, Chen J, Vullikanti A, Marathe M, Hoops S, Bhattacharya P, Machi D, Chen S, Paul R, Janies D, Thill JC, Galanti M, Yamana T, Pei S, Shaman J, Reich NG, Healy JM, Slayton RB, Biggerstaff M, Johansson MA, Runge MC, Viboud C. Projected resurgence of COVID-19 in the United States in July-December 2021 resulting from the increased transmissibility of the Delta variant and faltering vaccination. eLife June 21,2022. DOI: 10.7554/eLife.73584 PMID: 34494030 PMCID: PMC8423228 (20 citations as of May 27, 2023)
d. Borchering RK, Mullany LC, Howerton E, Chinazzi M, Smith CP, Qin M, Reich NG, Contamin L, Levander J, Kerr J, Espino J, **Hochheiser H**, Lovett K, Kingsey M, Tallaksen K, Wilson S, Shin L, Lemaitre JC, Hulse JD, Kaminsky J, Lee EC, Hill AL, Davis JT, Mu K, Xiong X, Piontti AS, Vespignani A, Srivastava A, Porebski P, Venkatramanan S, Adiga A, Lewis B, Klahn B, Outten J, Hurt B, Cheng J, Morveit H, Wilson A, Marathe M, Hoops S, Bhattacharya P, Machi D, Chen S, Paul R, Janies D, Thill J-C, Galanti M, Yamana T, Pei S, Shaman J, Espana G, Cavany S, Moore S, Perkins A, Healy JM, Slaton RB, Johansson MA, Biggerstaff M, Shae K, Truelove SA, Runge MC, Viboud C, Lessler J. Impact of SARS-CoV-2 vaccination of children 5-11 years on COVID-19 disease burden and resilience to new variants in the United States, November 2021-March 2022: A multi-model study. The Lancet Regional Health – Americas. January 2023. Volume 17, DOI: 10.1016/j.lana.2022.100398 PMID: 36437905 PMCID: PMC9679449

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