OMB No. 0925-0001 and 0925-0002 (Rev. 03/2020 Approved Through 02/28/2023)

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, Maryland, USA | PhD | 06/2003 | Computer Science |
| National Institute on Aging, Baltimore, MD, USA | Post-Doctoral | 07/2006 | Computational Biology |

**A. Personal Statement**

I am a biomedical informatician with an interest in understanding issues in human-computer interaction, usability, and the overall goal of making information tools more useful for researchers and clinicians.

To address these challenges, I draw upon nearly 20 years of experience in human-computer interaction, including numerous journal articles and conference papers in multiple domains and several years of teaching graduate-level courses in human-computer interaction. I am also co-author of Research Methods in Human-Computer Interaction, 2nd Edition (Morgan Kaufmann, 2017).

My research efforts draw upon a range of techniques and practices from human-computer interaction, including qualitative investigations of user needs, design and implementation of prototype systems, collection of data from systems in use, and qualitative and quantitative user studies. My current research projects include the development of natural language processing (NLP) tools for cancer registries (**UG3 CA243120), NLP in support of clinical research at the Veterans Administration, the design of data portals and user tools for curating and sharing COVID-19 related datasets (**3U24GM132013-02S1), and the development of explainable predictions for models of pediatric deterioration (proposal pending with NLM). I have previously worked with colleagues at the University of Pittsburgh 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), and on the Center for Causal Discovery (**U54HG008540)**. Other efforts an NLM-funded (1R01LM010964) project investigation of interactive natural language processing (Trivedi, et al. 2017), an AHRQ-funded (5R01HS021290) investigation of EMR usage patterns in physician-patient encounters, the development of user tools for exploring links between human phenotypes and animal models (R24OD011883), and an investigation into information needs for curation of drug-drug interaction information (R01 LM011838 ), among others.  These projects have involved a combination of ethnographic methods with design and empirical evaluations. My research has also covered a range of other topics, including information visualization, bioinformatics, universal usability, security, privacy, and public policy implications of computing systems. I am director of the Pittsburgh Biomedical Informatics Training Program (T15 LM007059).

1. Lazar J, Feng J, **Hochheiser H**. Research Methods in Human-Computer Interaction. Cambridge, MA: Morgan Kaufmann ; 2017.
2. Romagnoli, KM, Boyce R, Empey, P, Ning Y, Adams S, **Hochheiser H**. Design and evaluation of a pharmacogenomics information resource for pharmacists J. Am Med Inform Assoc 2017 DOI: 10.1093/jamia/ocx007. PMID: 28339805, PMCID: PMC6080676.
3. Goldberg IG, Allan C, Burel JM, Creager D, Falconi A, **Hochheiser H**, Johnston J, Mellen J, Sorger PK, Swedlow JR. The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging. Genome biology. 2005; 6 (5):R47. PMCID: PMC1175959.
4. **Hochheiser H**, Shneiderman B. Dynamic Query Tools for Time Series Data Sets: Timebox Widgets for Interactive Exploration. Information Visualization. 2004; 3:1-18.

**B. Positions and Honors  
Positions and Employment**

|  |  |
| --- | --- |
| 1990 - 1991 | Graduate Student, Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA |
| 1991 - 1993 | Research Staff, Tufts University School of Medicine, Boston, Massachusetts, USA |
| 1993 - 1995 | Software Developer, Biomedical Engineering, Massachusetts General Hospital, Boston, Massachusetts, USA |
| 1995 | Software Developer, AT&T Bell Labs, Murray Hill, New Jersey, USA |
| 1996 - 1998 | Consultant, H. Systems, Princeton, New Jersey, USA |
| 1998 - 2003 | Graduate Student, Computer Science, University of Maryland, College Park, Maryland, USA |
| Summer 1999 | Summer Intern, IBM T.J. Watson Labs, Hawthorne, New York, USA |
| 2003 - 2006 | Postdoctoral Researcher, Laboratory of Genetics, National Institute on Aging, Baltimore, Maryland USA |
| 6/1999 - 9/1999 | Summer Intern, IBM T.J. Watson Labs, Hawthorne, New York, USA |
| 2006 - 2009 | Assistant Professor, Computer and Information Sciences, Towson University, Towson, Maryland USA |
| 2009 – 2017 | Assistant Professor, Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania, USA |
| 2012 – 2017 | Assistant Professor, Intelligent Systems Program, University of Pittsburgh, Pittsburgh, Pennsylvania, USA |
| 2017-Present | Associate Professor, Biomedical Informatics and Intelligent Systems Program, University of Pittsburgh, Pittsburgh, Pennsylvania, USA |
| 2017- Present | Director, Biomedical Informatics Training Program, University of Pittsburgh, Pittsburgh, Pennsylvania, USA |

**Other Experience and Professional Memberships**

|  |  |
| --- | --- |
| 2004 - Present | Member, Association of Computing Machinery |
| 2009 - Present | Member, American Medical Informatics Association |

**Honors**

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

**C. Contributions to Science**

|  |  |  |  |
| --- | --- | --- | --- |
| 1. | | **Predictive Learning for EHRs**: My work in clinical informatics has explored the development of a predictive learning electronic medical record, capable of using information access patterns to identify data items likely to be of high value when reviewing patient records. This project has combined many aspects of biomedical informatics, including observational data collection, eye-tracking, machine learning, and user interface design (King, et al, 2015, 2017, 2018, 2019). | |
| a. | | King, AJ, Cooper, GF, Clermont, G, **Hochheiser H**, Hauskrecht, M, Sittig, DF, Visweswaran, S, Using Machine Learning to Selectively Highlight Patient Information, Journal of Biomedical Informatics (2019), doi: <https://doi.org/10.1016/j.jbi.2019.103327>. PMC Journal in-process. |
| b. | | King, AJ, Cooper C, **Hochheiser H,** Clermont C, Hauskrecht M, Visweswaran S Using Machine Learning to Predict the Information Seeking Behavior of Clinicians Using an Electronic Medical Record System 2018 AMIA Annual Symposium; 2018 Nov 3-7; San Francisco, California PMID: 301815109 PMCID: PMC 6371238. |
| c. | | King AJ, **Hochheiser H**, Visweswaran S, Clermont G, Cooper GF. Eye-tracking for clinical decision support: A method to capture automatically what physicians are viewing in the EMR. AMIA Translational Summits 2017, San Francisco, CA Winner, Clinical Research Informatics Best Student Paper. PMID: 28815151 PMCID: PMC5543363 |
| d. | | King A, Cooper GF, **Hochheiser H**, Clermont G, Visweswaran S. Development and Preliminary Evaluation of a Prototype of a Learning Electronic Medical Record System. 2015 AMIA Annual Symposium; 2015 Nov 14-18; San Francisco, California p.1967-1975 PMID: 26958296 PMCID: PMC4765593. |
|  | | | |
| 2. | | **Information visualization/Visual analytics for biomedical data:** The application of visual, dynamic tools to the understanding of complex biomedical data has been a recurring theme in my work. My dissertation work on time series visualization explored applications in temporal gene expression profiles and genetic sequences (Hochheiser and Shneiderman, 2004, above). I applied these techniques to digital microscopy workflows, the development of visualizations deployed for the FaceBase Data Management and Coordination Hub, a data portal for the Sarcoidosis and A1AT Genomics & Informatics Center (1U01HL112707), collaborator identification via research networking systems (Borromeo, et al. 2014), to interactive tools for Natural Language Processing (Trivedi, et al. 2017, 2019); and to longitudinal cancer patient histories (Zhou, et al. 2020). | |
| a. | | 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 in progress |
| b. | | 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: https://doi.org/10.1055/s-0039-1695791. PMCID: PMC6727024 |
| c. | | Trivedi G, Pham P, Chapman WW, Hwa R, Wiebe J, **Hochheiser H**. NLPReViz: an interactive tool for natural language processing on clinical text. J Am Med Inform Assoc 2017 DOI: 10.1093/jamia/ocx070. PMID: 29016825, PMCID: . PMCID: PMC6381768. |
| d. | | Borromeo CD, Schleyer TK, Becich MJ, **Hochheiser H**. Finding collaborators: toward interactive discovery tools for research network systems. Journal of Medical Internet Research. 2014; 16 (11):e244. doi: 10.2196/jmir.3444 PMID: 25370463, PMCID: PMC4376239. |
|  | | | |
| 3. | | **Biomedical data models and ontologies:** My work on information visualization exposed me to the importance of both defining and populating well-structured data models. I led the team responsible for the data models software infrastructure necessary to coordinate data from 10 projects collaborating in the FaceBase consortium. My work on FaceBase led to my participation in the development of the Ontology of Craniofacial Development and Malformation (Brinkley, et al. 2014). Related efforts have involved data modeling for pharmacogenomic data (Boyce, et al. 2013), cancer patient profile data models (PI Savova 1 U24 CA184407-01), and cancer treatment regimens (Warner, et al. 2019) | |
| a. | | 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 [Available on 2020-08-01]. |
| b. | | 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. |
| c. | | **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. |
| 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. |
|  | | | |
| 4. | | **Biomedical data integration, sharing, and analysis:** My interest in developing tools that help people share and exploit data has led me to projects at the intersection of team science, biomedical data, and data sharing and resuse. For the FaceBase Data Management and Coordination Hub, I coordinated the development of a data portal for 10 diverse projects generating data on the genetics of craniofacial development. For the Monarch Initiative, I worked with colleagues from several institutions to develop tools in support of the use of phenotype information for understanding human disease. | |
| a. | | Mungall CJ, McMurry JA, Köhler S, Balhoff JP, Borromeo C, Brush M, Carbon S, Conlin T, Dunn N, Engelstad M, Foster E, Gourdine JP, Jacobson JOB, Keith D, Laraway B, Lewis SE, NguyenXuan J, Shefchek K, Vasilevsky N, Yuan Z, Washington N, **Hochheiser H**, Groza T, Smedley D, Robinson PN, Haendel MA. The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species. Nucleic Acids Res (2017) 45 (D1): D712-D722 PMID: 27899636, PMCID: PMC5210586. |
| b. | | Smedley D, Schubach M, Jacobsen JOB, Köhler S, Zemojtel T, Spielmann M, Jäger M, **Hochheiser H**, Washington NL, McMurry JA, Haendel MA, Mungall CJ, Lewis SE, Groza T, Valentini G, Robinson PN A Whole-Genome Analysis Framework for Effective Identification of Pathogenic Regulatory Variants in Mendelian Disease. American Journal of Human Genetics 2016 99(3)505-606. doi 10.1016/j.ajhg.2016.07.005 PMID: 27569544, PMCID: PMC5011059. |
| c. | | Mungall CJ, Washington NL, Nguyen-Xuan J, Condit C, Smedley D, Köhler S, Groza T, Shefchek K, **Hochheiser H**, Robinson PN*,* Lewis SE, Haendel MA. Use of Model Organism and Disease Databases to Support Matchmaking for Human Disease Gene Discovery. Human Mutation 2015 36(10): 979-984. doi: 10.1002/humu.22857 PMID:26269093, PMCID: PMC5473253. |
| d. | | **Hochheiser H**, Aronow BJ, Artinger K, Beaty TH, Brinkley JF, Chai Y, Clouthier D, Cunningham ML, Dixon M, Donahue LR, Fraser SE, Hallgrimsson B, Iwata J, Klein O, Marazita ML, Murray JC, Murray S, de Villena FP, Postlethwait J, Potter S, Shapiro L, Spritz R, Visel A, Weinberg SM, Trainor PA. The FaceBase Consortium: a comprehensive program to facilitate craniofacial research. Dev Biol. 2011 Jul 15; 355 (2):175-82. PMCID: PMC3440302. |
|  | | | |
| 5. | | **Qualitative and quantitative investigations of human-comptuer interaction in biomedicine:** Building on my training in qualitative methods from human-computer interaction and related fields, I have applied qualitative and quantitative human-computer interaction techniques to understand informatics challenges facing clinicians. Among other efforts, I have participated in investigations of information needs for curation of drug-drug interaction data (Romagnoli, et al. 2017), measurements of physician activity during outpatient visits (Calvitti, et al. 2017), explorations of educational resources for pharamacists working on pharmacogenomics (Romagnoli, et al. 2016), and inquiries into barriers in documentation for nursing staff (Kohle-Ersher, et al. 2012). | |
| a. | | 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. |
| b. | | Calvitti A, **Hochheiser H**, Ashfaq S, Bell K, Chen Y, El Kareh R, Gabuzda MT, Liu L, Mortensen S, Pandey B, Rick S, Street RL, Weibel N, Weir C, Agha Z. Physician activity during outpatient visits and subjective workload, J. Biomed. Inf. 2017, May, 69:135-149. PMID: 28323114. DOI: 10.1016/j.jbi.2017.03.011. |
| 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. PMID: 26725696. PMCID: PMC4720137 doi: 10.1016/j.ijmedinf.2015.11.015. |
| 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 |

**D. Additional Information: Research Support and/or Scholastic Performance**

**Ongoing Research Support**

**T15 LM007059-33 (Hochheiser) 07/01/1987 – 06/30/2022**

**NIH/NLM/NIDR**

**Pittsburgh Biomedical Informatics Training Program**

**This grant provides funding support to graduate students in the Biomedical Informatics Training Program, Department of Biomedical Informatics, University of Pittsburgh.**

**UG3 CA243120 (Savova, Durbin, Hochheiser, Warner) 07/19/2019 - 06/30/2021**

**NIH/NCI**

**Natural Language Processing Platform for Cancer Surveillance**

**Development of cancer natural language processing tools in support of cancer registry data curation.**

**IPA (Weisbord) 04/01/2020 – 03/31/2021**

**U.S. Department of Veterans Affairs**

**IPA Harry Hochheiser**

**Natural language processing for extraction of details clinical scores regarding cardiogenic shock and EKG tests as needed for computation of clinical scores.**

3U24GM132013-02S1 (Van Pahhuis) 07/01/2020-06/30/2021

NIGMS

Accelerating COVID-19 Modeling Research by Improving the Discovery and New Use of Data: Leveraging Engagement and Automation of Curation Workflows

**Completed Research Support (end date within last 4 years)**

5U24CA184407-02 (Savova) 05/01/2014 – 04/30/2019 (NCE through 04/30/2020)

NIH/NCI

Cancer Deep Phenotype Extraction from Electronic Medical Records (DeepPhe)

Role: Co-Investigator

information models and visualization tools for longitudinal models of cancer patient histories,

**R01 LM012095 (Visweswaran) 9/15/2015 - 6/30/2019**

**NLM/NIH**

**Development and Evaluation of a Learning Electronic Medical Record System**

**Role: Co-Investigator**

**User studies and user interface design and evaluation for the Learning Electronic Medical Record System**

**U54HG008540 (Cooper) 09/15/14 - 08/31/18**

**NHGRI**

**Center for Causal Modeling and Discovery of Biomedical Knowledge from Big Data**

**Role: Co-Investigator**

**User needs investigations, user interface design, and data storage and management infrastructure for causal modeling tools and analysis results.**

**U2GGH00729 (Douglas) 02/01/14 - 02/01/18**

**Baobab Health Trust (CDC)**

**Role: Co-Investigator**

**Training in user-centered design methods for staff at Baobab Health Trust**

**1 R01 LM011838 (Boyce) 02/01/14 - 01/31/18**

**NIH/NLM**

**Addressing gaps in clinically useful evidence on drug-drug interactions**

**Role: Co-Investigator**

**Qualitative inquiries with drug-drug interaction experts; design of minimal information models for potential drug-drug interaction evidence.**