|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SBMI BIOGRAPHICAL SKETCH | | | | |
|  | | | | |
| NAME (Last, First Middle)  Occam, Bill | | CURRENT EMPLOYMENT TITLE AND EMPLOYER  Graduate Research Assistant School of Biomedical Informatics, UTHealth | | |
| STUDENT ID (7-digit number)  1010101 | |
| EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)* | | | | |
| INSTITUTION AND LOCATION | DEGREE  *(if applicable)* | | MM/YY | FIELD OF STUDY |
| University of Oxford | B.A. | | 05/03 | Philosophy |
| University of Oxford | M.A. | | 05/06 | Philosophy |
| Georgia Institute of Technology | M.S. | | 05/09 | Computer Science |
|  |  | |  |  |

*This Biosketch template is designed to provide a uniform resume/curriculum vitae for students at the UTHealth School of Biomedical Informatics for a variety of purposes (e.g., scholarship or GRA applications). If you find a problem with this template, please contact Kirk Roberts (*[*Kirk.Roberts@uth.tmc.edu*](mailto:Kirk.Roberts@uth.tmc.edu)*). For all other questions, please contact the Office of Academic Affairs (*[*SBMIAcademics@uth.tmc.edu*](mailto:SBMIAcademics@uth.tmc.edu)*).*

*Please name this Biosketch file using the following convention:***SBMI\_Biosketch\_Lastname(s)\_Firstname(s).pdf**

*There should be no gray text when you submit this Biosketch. Remove all the instruction text.*

1. Personal Statement

*In this section, the student is allowed to make a personal statement regarding his/her interests in biomedical informatics. If student is conducting research, this should focus on research interests (i.e., what informatics areas have you published in or plan to publish in). If the student is not conducting research, this should focus on areas of informatics the student is passionate about, has experience in, or plans to use in future employment.*

*It is preferred that this statement be around 300-500 words, and commensurate with the student’s experience in informatics (i.e., short for new students, long for experienced students). The statement cannot go beyond this page.*

*Mr. Occam, for instance, might describe his journey from philosophy and computer science to biomedical informatics. After that, he would describe how he hopes to use informatics as tool to cut through the entangled healthcare system.*

*Other notes about this Biosketch template:*

* *Do not change the font, margin, spacing, etc. unless necessary for clarity.*
* *Colors should be avoided, as should figures.*
* *Follow the structured elements below as close as possible to the template.*
* *Employment history should focus on post-undergraduate employment, unless centrally important to one’s career. Part time and summer positions during high school, college, and temporary situations should not be included. The exception is university positions related to teaching or research work.*
* *Attempt to place page-breaks between sections and/or sub-sections to improve readability for Sections B and C. E.g., lists (of courses, scholarships, etc.) should all be on a single page.*

*Part A above MUST end on this page. Part B below MUST start on next page (leave remaining space on this page blank).*

1. **Positions and Honors**

**Positions and Employment**

2003-2004 Teaching Assistant, Dept. of Philosophy, University of Oxford, Oxford, UK

2005-2006 Research Assistant, Dept. of Medicine, University of Oxford, Oxford, UK  
2008-2009 Teaching Assistant, Dept. of Computer Science, Georgia Institute of Technology, Atlanta, GA

2009-2012 Software Engineer, Electronic Health Corporation, Houston, TX

2012-2014 Vice President of Engineering, Electronic Health Corporation, Houston, TX

2014- Founder and CEO, Minimal Assumptions Technologies Inc., Houston, TX

*Note: Include any Graduate Research Assistant (GRA), Teaching Assistant (TA), and Student Assistant (SA) positions at UTHealth. State “N/A” if no previous employment.*

**Honors**

2006 Apostolic Book Prize, St. Francis Philosophical Society, Oxford, UK

2008-09 Louis Bavaria Scholarship ($5,000/year)

2015 Promising 40 Under 40 in Health, Top Houston Magazine, Houston, TX

*Note: Official awards from recognized organizations only. Remove section if no previous honors. Include substantial merit-based scholarships (minimum $1,000) prior for education prior to enrolling at SBMI. SBMI-related awards belong in the next section.*

1. **SBMI/UTHealth Academic Achievement**

**Planned Degree**:Master of Science in Biomedical Informatics

*Note: The only valid options to the above are:*

* *Certificate in [Biomedical Informatics/Applied Informatics/Public Health Informatics/Health Data Science/Pharmacy Informatics],*
* *Master of Science in Biomedical Informatics,*
* *Doctor of Philosophy (PhD) in Biomedical Informatics,*
* *Doctorate in Health Informatics (DHI)*

*Dual Degree and 4+1 students should just state the informatics degree they are seeking.*

**Semester Grade Course Course Name (Instructor)**

Fall 2015 A BMI 5300 Introduction to Biomedical Informatics (Smith)

Spr. 2016 A BMI 5310 Foundations of Health Information Sciences I (Smith)

Spr. 2016 A BMI 5007 Data Structures and Algorithms in Biomedicine (Rogith)

Sum. 2016 B BMI 5354 Cognitive Science in Biomedical Informatics (Franklin)

Fall 2016 A BMI 5311 Foundations of Health Information Sciences II (Cohen)

Fall 2016 A BMI 5352 Statistical Methods in Biomedical Informatics (Johnson)

Fall 2016 A PH 1690 Foundations of Biostatistics (Wu)

Spr. 2017 BMI 6323 Data Mining in Biomedical Informatics (Roberts)

Spr. 2017 BMI 6340 Health Information Visualization and Visual Analytics (Johnson)

*Notes: (1) Include current courses in the list, as above, and (2) Abbreviate course names where necessary to maintain one line per course.*

**Current Credit Hours**: 21

**Current GPA**: 3.86

**Scholarship Awards**

2015 Doris Ross Scholarship

2016 SBMI Scholarship Excellence Fund

2016 Travel Award (AMIA Annual Symposium)

2017 Travel Award (HIMSS)

*Note: Include all scholarships (no minimum) since enrolling at SBMI.*

**Service to SBMI and UTHealth**

2016-2017 Vice President, Student Governance Organization

**Honors**

2016 Best Student Paper, AMIA Annual Symposium

1. **Contribution to Science**

*Students not pursuing research at SBMI, and without a prior history of published research, should leave this section blank. Notably, at least for scholarship applications, this section is rarely used to judge non-doctoral students and thus not worth putting a lot of time into if you are just a Certificate/Masters student using this Biosketch to apply for a scholarship.*

*This section is designed to act like the NIH Biosketch Contribution to Science section. It allows the student to express research contributions made both prior to and after enrolling at SBMI. Break your research interests into 1-5 areas, as wide or narrow as appropriate. For each research area, provide a name, an overall description of your work in that area (maximum 300 words), and all peer-reviewed publications (journal articles, conference papers, posters, and peer-reviewed abstracts) in that area. A good organization of the description could contain two paragraphs: (1) a high-level description of the research area from the student’s perspective, and (2) the student’s actual or planned contributions to the area.*

*In the publication citation, place your name in bold. Place an asterisk (\*) at the front of publications prior to enrollment at SBMI. If the student has a substantial publication history prior to enrollment, and those publications have little relation to biomedical informatics, then all such publications may be placed in a single contribution. The publication list should be sorted chronologically, oldest first. For papers submitted but not yet accepted, place “In Submission” at the end of the citation.*

*The pseudonymous Mr. Occam is not engaged in research, so what follows is two example contributions from an older NIH Biosketch from Kirk Roberts with truncated publication lists:*

**Clinical Information Extraction**

A significant amount of important clinical information from electronic health records (EHRs) is in unstructured free text. For instance, much of the important information about a patient’s history and care plan are only available in free text. In order to perform the types of complex analysis necessary to improve EHR usability and medical research, natural language processing (NLP) methods are necessary to perform information extraction (IE) from unstructured text and convert it to a structured form.

Dr. Roberts’s contribution to this is the development of state-of-the-art clinical IE methods. He focuses on NLP techniques based on machine learning (ML), both traditional ML and deep learning methods [Roberts, 2016]. He has consistently placed among the top system submissions to the i2b2 NLP challenge. This includes tasks such as concept recognition and assertion classification [Roberts & Harabagiu, 2011], emotion classification, event and temporal expression extraction [Roberts et al., 2013], and identification of risk factors for heart disease [Roberts et al., 2015]. In this last task, the team led by Dr. Roberts finished first overall out of 20 teams. He demonstrated that fine-grained, high-quality training data combined with a simple IE system can outperform complex IE systems trained on coarser-grained data.

1. **Roberts, K.** & Harabagiu, S. A flexible framework for deriving assertions from electronic medical records. *Journal of the American Medical Informatics Association*, 18, 2011.
2. **Roberts, K.**, Rink, B., & Harabagiu, S. A flexible framework for recognizing events, temporal expressions, and temporal relations in clinical text. *Journal of the American Medical Informatics Association*, 20, 2013.
3. **Roberts, K.**, Shooshan, S., Rodriguez, et al. The role of fine-grained annotations in supervised recognition of risk factors for heart disease from EHRs. *Journal of Biomedical Informatics*, 58, 2015.
4. **Roberts, K.** Assessing the Corpus Size vs. Similarity Trade-off for Word Embeddings in Clinical NLP. *Proceedings of Clinical NLP*, 2016.

**Clinical Question Answering**

EHRs offer great promise to improve direct patient care as well as dramatically increase the availability of clinical data for researchers. However, the promises of improvement to patient outcomes is largely unfulfilled. One of the major reasons for this is that EHRs have focused mostly on electronic documentation, forcing physicians to become data entry workers, with little work on how the EHR can improve outcomes by interacting with clinicians at the point-of-care. A major barrier is the time and complexity of looking up information related to a patient’s history (since EHR design focuses more on initial documentation than retrieval). A promising solution is natural language question answering (QA). QA systems allow users to state their information needs in an intuitive format (questions) while the EHR does the heavy lifting of retrieving the answers to meet those information needs.

Dr. Roberts’s contribution to medical QA is the realization that this kind of QA requires deeply understanding the user’s question. He published the first proposal for a clinical question representation based on logical forms [Roberts & Demner-Fushman, 2015] and a corpus of such logical forms [Roberts & Demner-Fushman, 2016b] to enable natural language processing (NLP) systems to learn how to automatically convert EHR questions to logical forms. He has submitted a paper demonstrating an automatic method for a significant step in clinical question understanding, known as semantic parsing [Roberts & Patra, 2017]. Additionally, since such a QA approach might be useful for patients as well as clinicians, he conducted the largest study to date comparing how questions asked by medical professionals (representing clinicians) differ from those asked by consumers (representing patients) [Roberts & Demner-Fushman, 2016a]. Dr. Roberts continues to make progress in this area, exploring how crowdsourcing and various NLP techniques can improve EHR question understanding.

* + - 1. **Roberts, K.** & Demner-Fushman, D. Toward a Natural Language Interface for EHR Questions. *AMIA Summit on Clinical Research Informatics*, 2015.
      2. **Roberts, K.** & Demner-Fushman, D. Interactive use of online health resources: A comparison of consumer and professional questions. *Journal of the American Medical Informatics Association*, 2016a.
      3. **Roberts, K.** & Demner-Fushman, D. Annotating Logical Forms for EHR Questions. *Language Resources and Evaluation Conference*, 2016b.
      4. **Roberts, K.** & Patra, B.G. A Semantic Parsing Method for Mapping Clinical Questions to Logical Forms. *AMIA Annual Symposium*, 2017. In Submission.

**Complete List of Publications**

<https://www.ncbi.nlm.nih.gov/myncbi/browse/collection/51369299/?sort=date&direction=descending>

*Note: this can be a link to My NCBI, Google Scholar, or a personal website.*

E. Other Statement

*This section enables the student to advocate for any other relevant activities, including service to the biomedical informatics community, impactful volunteer efforts, or notable success in overcoming adversity. Responses are limited to 600 words.*