**Workshop Proposal for BIBM 2018 Workshops**

**1.  Workshop Title and Acronym**

The Third International Workshop on Semantics-Powered Data Analytics (SEPDA 2018)

<http://semantics-powered.org/>

**2. Introduction to workshop**

Biomedical ontologies and controlled terminologies provide structured domain knowledge to a variety of health information systems. The rich thesaurus with concepts linked by semantic relationships has been widely used in natural language processing, data mining, machine learning, semantic annotation, and automated reasoning. The dramatically increasing amount of health-related data poses unprecedented opportunities for mining previously unknown knowledge with semantics-powered data mining and analytics methods. However, due to the heterogeneity of different data sources, it is a challenging problem to exploit multiple sources to solve real-world problems such as designing cost-effective treatment plan for patients, designing generalizable clinical trials, drug repurposing, and clinical phenotyping. The goal of this workshop is to bring people in the field of data mining, knowledge representation, knowledge management, and data analytics to discuss innovative semantic methods, applications, and data analytics to address problems in healthcare, biomedicine, public health, and clinical research with biomedical, clinical, behavioral, and social web data.

We are inviting original research submissions as well as work-in-progress. Selected full papers will be invited to submit an extended version to a special issue of a mainstream journal in biomedical informatics (TBA).

This is the third installment of the SEPDA workshop. The first SEPDA workshop (SEPDA 2016) was held in conjunction with IEEE BIBM 2016 in Shenzhen, China (<http://semantics-powered.org/sepda2016.html>). SPEDA 2016 was a half-day workshop. Dr. Hua Xu, a leading expert in the field of Biomedical Informatics, gave the keynote speech on the bioCADDIE project. Eight quality papers were presented. Selected papers were invited to publish the extended version to the special issue on “Semantics-Powered Healthcare Engineering and Data Analytics” in the Journal of Healthcare Engineering (<https://www.hindawi.com/journals/jhe/si/461480/>).

The second SEPDA Workshop (SEPDA 2017) was held in conjunction with IEEE BIBM 2017 in Kansas City, Missouri, USA (<http://semantics-powered.org/>). SPEDA 2017 was a full day workshop. 17 high quality papers were presented in the workshop. Extended journal version of the papers will be published in BMC Medical Informatics & Decision Making (IF: 1.6).

**3. Research Topics included in the workshop**

Topics of interest include but not limited to:

Semantics-based Data Mining and Analytics

* Ontology-based text mining and natural language processing
* Semantics-powered data mining and machine learning from biomedical, clinical, or social web data
* Information Extraction from biomedical, clinical, or social web data
* Semantic annotation on biomedical, clinical or social web data

Ontologies and Controlled Terminologies

* Ontology development and enrichment
* Quality assurance of ontologies and controlled terminologies
* Semantic harmonization and ontology alignment
* Knowledge representation and reasoning

Data Integration

* Linked open data
* Novel approaches for data integration of heterogeneous data sources
* Large scale data integration

Applications

* Novel tools and ontologies for data interpretation and visualization
* Pharmacovigilance and drug repurposing using ontologies
* Clinical trial generalizability assessment using ontologies
* Algorithmic phenotyping and cohort identification using ontologies
* Improving the literacy of health information consumers

**4. Important Dates**

Sept 30, 2018: Due date for full workshop papers submission

Oct 27, 2018: Notification of paper acceptance to authors

Nov 15, 2018: Camera-ready of accepted papers

Dec 3, 2018: Workshops

**5. Plan to Attract Quality Submissions**

We plan to attract quality submissions by sending the CFP to the mailing lists of the American Medical Informatics Association with over 5000 active members. We will also post our CFP in WikiCFP. The workshop co-chairs will leverage their professional networks to attract quality submissions. We will also invited selected conference papers to high quality journals in biomedical informatics such as BMC Medical Informatics & Decision Making (IF: 1.6), Health Informatics Journal (IF: 3.02), and Journal of Biomedical Semantics (IF: 1.85). Co-chairs Dr. Zhe He and Dr. Jiang Bian are Associate Editors of BMC Medical Informatics & Decision Making.

**6. Workshop Co-chairs and Bios:**

**Zhe He, PhD, School of Information, Florida State University, USA**

Bio: Dr. Zhe He is a tenure-track Assistant Professor in the School of Information and a Core Affiliate of the Institute for Successful Longevity at Florida State University. He has a strong background in computer science and biomedical informatics, with specific training and extensive interdisciplinary research experience in clinical research informatics, biomedical ontologies, natural language processing, and data analytics. On the foundational track, he is developing semantic and structural methods to enhance controlled terminologies/ontologies and improve their utility in natural language processing, clinical decision support, data analytics, and electronic health records. On the application track, he is developing novel data-driven methods and informatics tools to assess the generalizability of clinical studies using the electronic data in clinical trial registries, public patient databases, and clinical data warehouses. His research has produced over 40 articles in major biomedical informatics venues including Journal of Biomedical Informatics, Artificial Intelligence in Medicine, Methods of Information in Medicine, American Medical Informatics Association (AMIA) Annual Symposium, and MEDINFO. He received two Distinguished Paper Awards from AMIA Annual Symposiums in 2015 and 2017, respectively. His research has been funded by NIH, NVIDIA, Amazon, FSU Council on Research and Creativity, and Institute for Successful Longevity.

**Cui Tao, PhD, School of Biomedical Informatics, University of Texas Health Science Center at Houston, USA**

Bio: Dr. Cui Tao is a tenured associate professor at the School of Biomedical Informatics, University of Texas Health Science Center at Houston. She is a recipient of the Presidential Early Career Award for Scientists and Engineers (PECASE) of 2014 named by President Obama. Dr. Tao is an expert in medical informatics with focus on ontologies, semantic web, biomedical data and knowledge extraction, normalization, and integration, as well as applying artificial intelligent technologies on biomedical applications. She has been a principal investigator (PI) or a co-investigator (co-I) of projects supported by over $60 Million in federal or state funding. She is currently the PI for three NIH funded projects with over $6 million funding. Dr. Tao has published more than 100 peer-reviewed articles and has served as a reviewer for many different journals and conferences.

**Jiang Bian, PhD, Department of Health Outcomes and Policy, University of Florida, USA**

Bio: Dr. Bian is an Assistant Professor of Biomedical Informatics in Health Outcomes & Biomedical Informatics (HOBI) at the University of Florida (UF). He has strong expertise in developing informatics tools and systems as well as extensive experience in machine learning methods for the interpretation of biomedical and textural data. His research areas can be divided into four logical sections: (1) development of novel informatics methods, tools and systems to support clinical and clinical research activities; (2) data-driven medicine-applications of informatics techniques, including machine learning methods in medicine on solving big data problems; (3) mining the Internet, including the social web, to provide insights into health-related behavior and health outcomes of various populations and finding ways to develop interventions that promote public and consumer health; and (4) leveraging ontology and semantic web technologies for data management and data integration targeting biomedical researchers, as well as knowledge management and knowledge dissemination targeting health information consumers. Dr. Bian’s research has produced over 60 articles in major biomedical informatics venues such as Journal of the American Medical Informatics Association, Journal of Biomedical Informatics, Journal of Medical Internet Research, PLOS One.

**Rui Zhang, PhD, Institute for Health Informatics, University of Minnesota, USA**

Bio: Dr. Zhang is a tenure-track Assistant Professor in the Department of Pharmaceutical Care & Health Systems, Institute for Health Informatics, and Data Science at the University of Minnesota. He has extensive background and research experience in the field of health informatics, especially biomedical and clinical natural language processing (NLP), data and text mining of electronic health records (EHR), translational informatics and statistical analysis. His research interests include the secondly analysis of electronic health record (EHR) data for improving quality of patient care as well as discovery of pharmacovigilance knowledge through mining a large scale of biomedical literature. Dr. Zhang is the Principal Investigator on a NIH/NCCIH R01 grant investigating informatics framework to discover drug-supplement interactions in the biomedical literature and electronic health record. He also serves as a Co-I on a R01 project through Agency for Healthcare Research and Quality (AHRQ) on the development of automated methods to identify and navigate clinically relevant new information within the clinical notes, and leads a team to implement the new information identification algorithm to visualize clinical notes in the Epic EHR system. His work has been recognized on a national and international scale including *Journal of Biomedical Informatics* Editor’s Choice (2014), a nomination for *Marco Ramoni* Distinguished Paper Award for Translational Bioinformatics (2015), and the best student papers in American Medical Informatics Association (AMIA) Annual Symposium (2011), MEDINFO (2013), and AMIA Joint Summits (2014, 2017, 2018). In additional, his work on mining 23 million biomedical publications to discover drug-supplement interactions has been highlighted by *The Wall Street Journal* and *Fox News*.

**7. Tentative Program Committee Members**

* GQ Zhang, University of Kentucky, USA
* James Geller, New Jersey Institute of Technology, USA
* Guoqian Jiang, Mayo Clinic, USA
* Zhihui Luo, University of Wisconsin, Milwaukee, USA
* Tianyong Hao, The cisco School of Informatics, Guangdong University of Foreign Studies, China
* James Cimino, University of Alabama at Birmingham, USA
* Licong Cui, University of Kentucky, USA
* Xia Jing, Ohio University, USA
* Riccardo Miotto, Icahn School of Medicine at Mount Sinai, USA
* Fleur Mougin, University of Bordeaux, France
* Praveen Chandar, IBM Research, USA
* Yonghui Wu, University of Florida, USA
* Ronald Cornet, University of Amsterdam, Netherlands
* Zhengxing Huang, Zhejiang University, China
* Gang Feng, Northwestern University