The Health Data Science Landscape at Indiana University

Brian E. Dixon, MPA, PhD, FHIMSS
Associate Professor and Research Scientist
December 6, 2017
Session Agenda

• Interested Parties Affiliated with IU
  – Schools, Departments, Institutes

• Programs and Curricula Related to Health Data Science
  – Core Components for the T15 Program
  – Our Supplemental Activities

• Challenges to Our Alignment of the IU Ecosystem
Data Science at Indiana University

School of Informatics and Computing

Regenstrief Institute / Center Biomedical Informatics

School of Public Health

Purdue School of Science / Computer Science Department

School of Informatics, Computing and Engineering

Health Data Science
Data Science Programs / Curricula

• B.S. Health Data Science  
  – FSPH & SOIC

• M.S. Data Science / Certificate Data Science  
  – School of Computing, Informatics, and Engineering (B)  
  – School of Informatics and Computing (Indy)

• Ph.D. Data Science  
  – School of Informatics and Computing (Indy)
Related Degrees and Programs

• M.S. and Ph.D. in Computer Science
  – IU Bloomington, Purdue
  – Concentration in Machine Learning

• M.S. and Ph.D. in Health Informatics
  – School of Informatics and Computing (Indy)

• M.S. and Ph.D. in Biostatistics
  – Fairbanks School of Public Health (Indy)
Our NLM T15 Program

• Core Curriculum drawn from Epidemiology / HPM
  – Foundations in the quantitative sciences
  – Layer with PH Informatics theories, methods, courses

• Ph.D. Minor in Population Health Analytics
  – Intro to Pop Health Analytics
  – Choose 2 Electives from ML, Viz, NLP, Text Mining, etc.
  – Applied Pop Health Analytics

• Applied PH Informatics Research @ Regenstrief CBMI
Example T15 Projects at Indiana University

• Build a dashboard for providers that monitors opioid usage in a health system population

• Analysis of HIE (health info exchange) usage and its impact on emergency department outcomes as well as costs

• Examine the rates and factors associated with syphilis testing in women with a stillbirth delivery

• Machine learning models to automate notifiable disease reporting as well as other forms of public health surveillance
T15 Supplemental Activities

- Creation of 2+ population health datasets
  - Syndromic surveillance*
  - Spectrum of diabetes care*

- Course materials to accompany datasets
  - Data dictionaries and exercise ideas
  - Assignments designed with real-world messiness of data
  - Jupyter notebooks with annotations to enable deployment
Challenges for Discussion

• Alignment (and negotiation) with other units
  – Many units have an interest; expertise is distributed
  – We do not have a Department of Biomedical Informatics
  – Separation of academic and research “homes”

• Symptoms of the broader challenge
  – Helping students navigate a sea of electives in disparate units, each with their own prerequisites
  – Recruitment of faculty with specialized knowledge for research committees, advising, mentoring
  – Keeping up with the changing programs, faculty, curricula, and courses
Brian E. Dixon, MPA, PhD, FHIMSS
Associate Professor, IU Fairbanks School of Public Health;
Research Scientist, Regenstrief Institute;
Health Research Scientist, Department of Veterans Affairs

Twitter: @dpugrad01
Email: bedixon@regenstrief.org