

ACM Task Force on
Data Science Education:
First Draft Curriculum Report, Plans,
Collaborations

Andrea Danyluk
Williams College & Northeastern University
adanyluk@williams.edu

Outline

- The Data Science Curriculum Task Force Effort
 - Committee
 - Background
- The First Draft Curriculum Report
 - Contents
 - Knowledge Areas and Competencies
- Community Engagement
 - Timeline
 - Discussion



Data Science Task Force

- Andrea Danyluk, Williams C. and Northeastern U., USA (co-chair)
- Paul Leidig, Grand Valley State University, USA (co-chair)
- Scott Buck, Intel, USA
- Boots Cassel, Villanova University, USA
- Andrew McGettrick, University of Strathclyde, UK
- Jian Pei, Simon Fraser University, Canada
- Christian Servin, El Paso Community College, USA
- Hongzhi Wang, Harbin Institute of Technology, China
- Weining Qian, East China Normal University, China
- Tin Kam Ho, IBM, USA

ACM Data Science Task Force Charter

To add to the broad, interdisciplinary conversation on data science, with an articulation of the role of computing discipline-specific contributions to this emerging field. The task force should seek to define what the computing contributions are to this new field, and should provide guidance for undergraduate data science programs of study.

To create a report, which may then be used to invite collaboration and coordination with other (non-computing) professional societies.

Background

- ACM Ed. Council summer meeting 2017
 - Build on the efforts of Boots Cassel & Heikki Topi, as well as other groups
 - Articulate importance of computing in the interdisciplinary data science space
 - Identify computing-based competencies for an undergraduate data science curriculum

Other Data Science Efforts

- EDISON Project (2017)
 - A competency-based framework to be used as guidance for educators, employers, etc.
 - Most similar to ACM effort; Europe focus.
 - Missing some key competencies
- Park City Report (2017)
 - Topics and learning outcomes for undergraduate data science curricula
 - Sample course outline
- National Academies report (2018)
 - Articulation of the importance of data science education
 - Synthesis of approaches
 - High level recommendations for this emerging field

Other ACM Efforts

- ACM/IEEE-CS Joint Task Force on Computing Curricula 2013
 - www.cs2013.org
- <https://www.acm.org/education/curricula-recommendations>

Draft Report Contents

Chapter 1 Introduction

- 1.1 Charter
- 1.2 Prior work on defining data science curricula
- 1.3 Committee work and processes
- 1.4 Survey of academic and industry representatives
- 1.5 Knowledge areas
- 1.6 Data Science in context
- 1.7 Competency framework
- 1.8 Motivating the study of data science
- 1.9 Overview of this report

Chapter 2 The Competency Framework

- 2.1 Competency in theory
- 2.2 Competencies and professional practice

References

Appendix A Draft of Competencies for Data Science

Appendix B Summary of Survey Responses

Draft Report Contents

Chapter 1 Introduction

Chapter 2 The Competency Framework

References

Appendix A Draft of Competencies for Data Science

Appendix B Summary of Survey Responses

Draft Report Contents

Chapter 1 Introduction

Chapter 2 The Competency Framework

Chapter 3 Toward an Interdisciplinary Data Science Curriculum

Chapter 4 Broadening Participation

References

Appendix A Body of Knowledge

Competencies for Data Science

Competency Details

Appendix B Summary of Survey Responses

Appendix C Exemplar Courses and Curricula

Knowledge Areas (Draft 1)

- Computing Fundamentals
 - Programming, Data Structures, Algorithms, Software Engineering
- Data Acquisition and Governance
- Data Management, Storage, and Retrieval
- Data Privacy, Security, and Integrity
- Machine Learning
- Data Mining
- Big Data
 - Complexity, Distributed Systems, Parallel Computing, and HPC
- Analysis and Presentation
 - HCI, Visualization
- Professionalism

Knowledge Areas (Draft 2)

- Computing Fundamentals
 - Programming, Data Structures, Algorithms, Software Engineering, Complexity
- Data Acquisition and Governance
- Data Management, Storage, and Retrieval
- Data Privacy, Security, and Integrity
- AI
 - Machine Learning, NLP/Text Processing, Vision/Image Processing
- Data Mining
- Big Data Systems
 - Cloud storage and computing, Parallel Computing, HPC
- Analysis and Presentation
 - HCI, Visualization
- Professionalism

Competency Framework

- Following ACM/IEEE-CS IT 2017; moving in the direction of CC 2020.
- Utilize a working definition of competency that connects knowledge, skills, and dispositions.
- Includes, but moves beyond articulation of topics and learning outcomes. [e.g., CS2013]

Competency = Knowledge + Skills + Dispositions

- Knowledge
 - Mastery of content
 - Transfer of learning
- Skills
 - Capabilities and strategies for higher-order thinking
 - Interactions with others and world around
- Dispositions
 - Personal qualities (socio-emotional skills, behaviors, attitudes) associated with success in college and career

From IT 2017; adapted from a publication by Council of Chief State School Officers (2013).

Timeline

- Early 2019:
 - Draft report out for comment
 - Outreach and gathering of feedback
 - Note: Initial comment period ends March 31
- Spring 2019 (f2f @ SIGCSE)
 - Begin work on next phase, including new KAs and competency details
- Spring 2019+: Outreach, presentations and information gathering
- Summer 2019:
 - Next draft to SIGCSE Education Advisory Committee (formerly Education Council)
 - Call for joint task force
- Fall 2019:
 - Draft report out for comment
- Early 2020
 - Release final report

Discussion

- <http://www.cs.williams.edu/~andrea/DSReportInitialFull.pdf>
- Comments?
- Collaboration?

