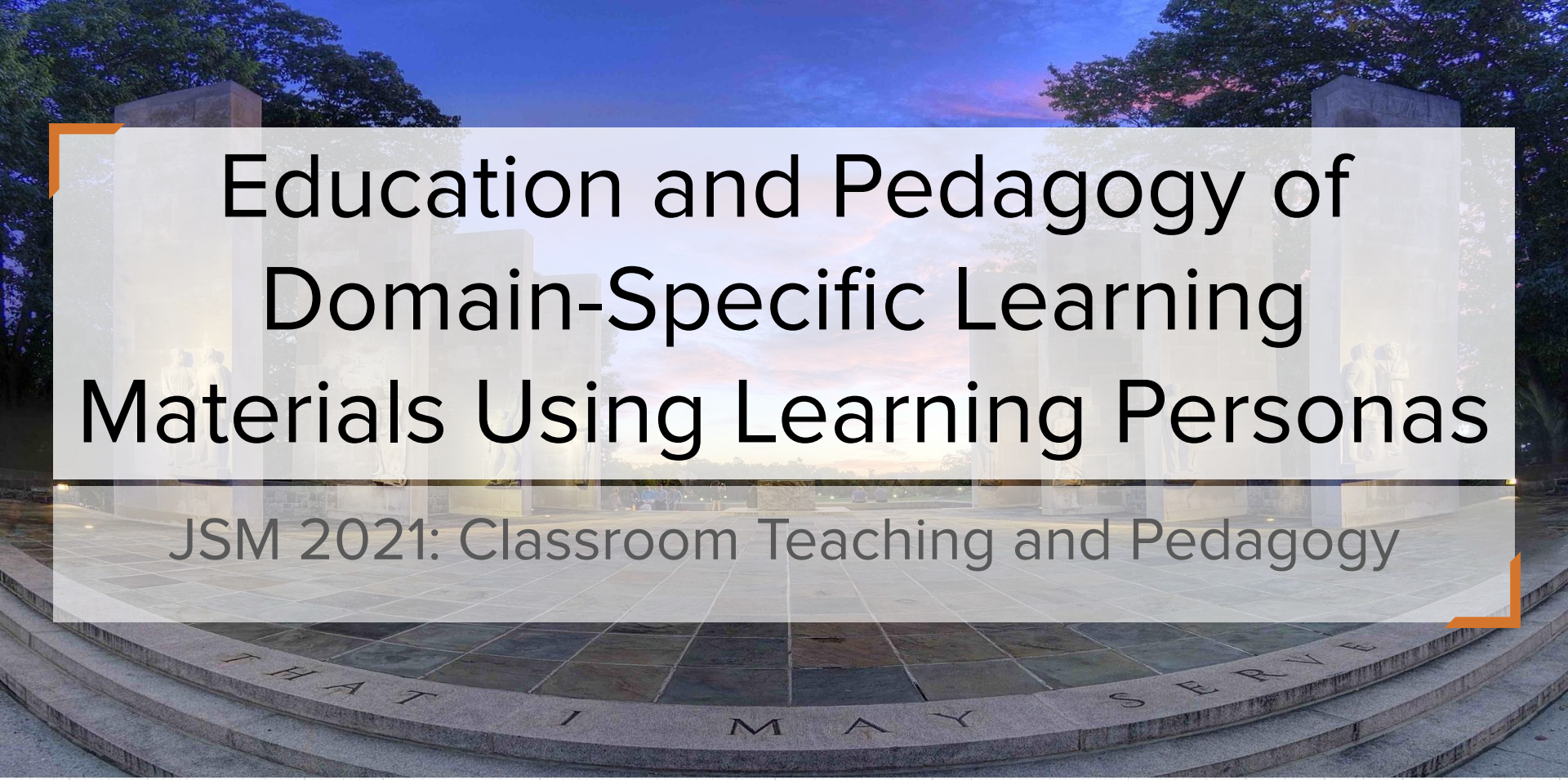


Education and Pedagogy of Domain-Specific Learning Materials Using Learning Personas

JSM 2021: Classroom Teaching and Pedagogy

Daniel Chen, MPH - GBCB, Virginia Tech
Anne Brown, PhD - University Libraries, Dept. of Biochemistry, Virginia Tech

VIRGINIA TECH™

PhD Candidate, Virginia Tech
Genetics, Bioinformatics, Computational Biology
@chendaniely
https://daniel.rbind.io/

Assistant Professor, Science Informatics Consultant and Health Analytics Coordinator, University Libraries

VIRGINIA TECH™

Data science programs are too general

- Data science programs target **single broad audiences**
- Opportunity to **branch out** to different disciplines
- Democratization of data science education enables more **domain specific** learning materials

- Statistics: statistical literacy vs statistical methods

- Consumers vs Producers

(Kross et al., 2020; GAISE College Report 2016)

Domain Needs: NIH Strategic Plan for Data Science

Data Infrastructure	Modernized Data Ecosystem	Data Management, Analytics, and Tools	Workforce Development	Stewardship and Sustainability
•Optimize data storage and security •Connect NIH data systems	•Modernize data repository ecosystem •Support storage and sharing of individual datasets •Better integrate clinical and observational data into biomedical data science	•Support useful, generalizable, and accessible tools and workflows •Broaden utility of and access to specialized tools •Improve discovery and cataloging resources	•Enhance the NIH data-science workforce •Expand the national research workforce •Engage a broader community	•Develop policies for a FAIR data ecosystem •Enhance stewardship

Figure 2. NIH Strategic Plan for Data Science: Overview of Goals and Objectives

(National Institutes of Health, 2020)

VIRGINIA TECH™

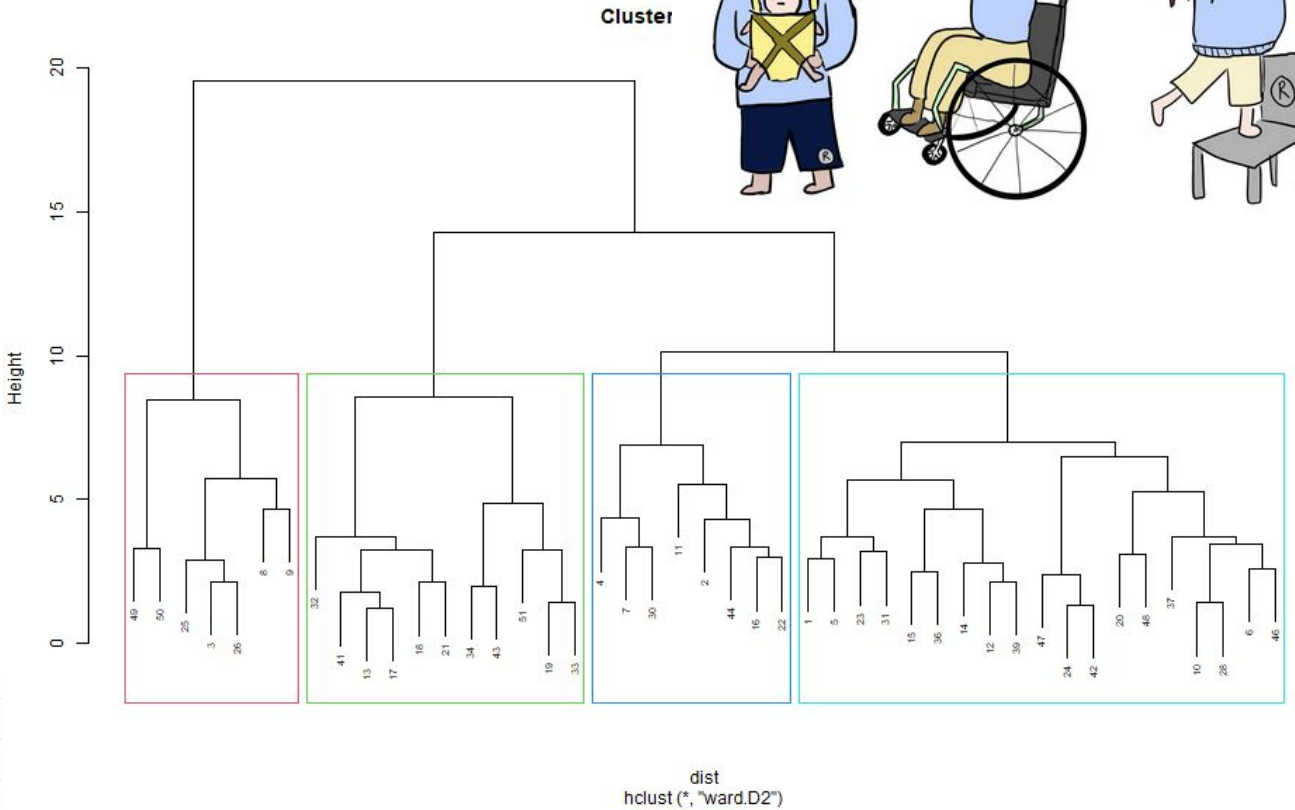
Personas: Learner Self-Assessment Survey

- Surveys:
  - https://github.com/chendaniely/dissertation-irb/
  - Combination of The Carpentries, “How Learning Works”, “Teaching Tech Together” (Ambrose et al., 2010; Koch and Wilson, 2016; Wilson, 2019)

1. Demographics (6)
2. Programs used in the past (1)
3. **Programming experience (6)**
4. **Data cleaning and processing experience (4)**
5. **Project and data management (2)**
6. **Statistics (4)**
7. Workshop framing and motivation (3)
8. Summary likert (7)

VIRGINIA TECH™

Learner Personas



Hierarchical cluster using Ward's method to identify personas (clusters)

VIRGINIA TECH™

Future Implications

- Data-driven approach to identifying learners
- Domain specific data science learning materials

Parts of a learner persona

1. Background
2. Relevant prior knowledge and experience
3. Perception of needs
4. Special considerations

Clare Clinician

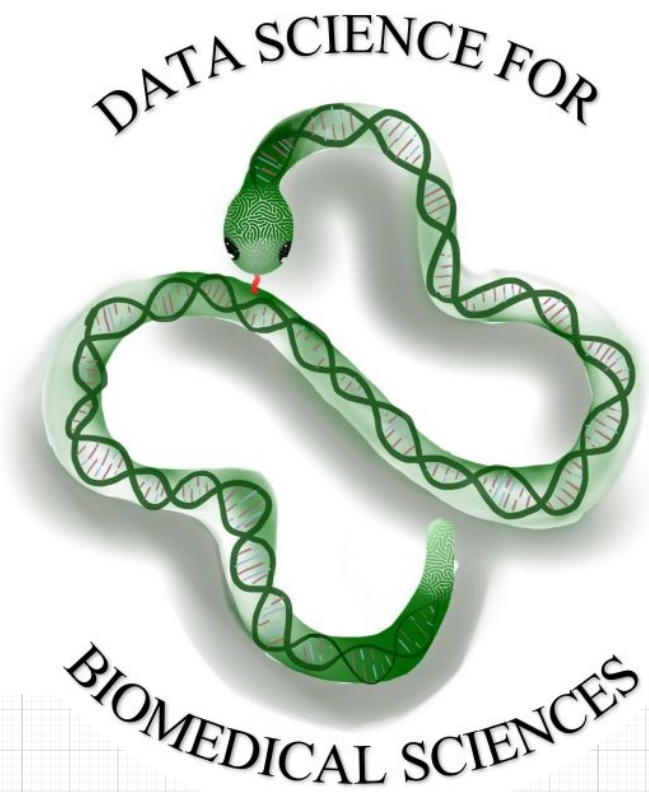
Figure 0.3: Drawn by Julia Chen

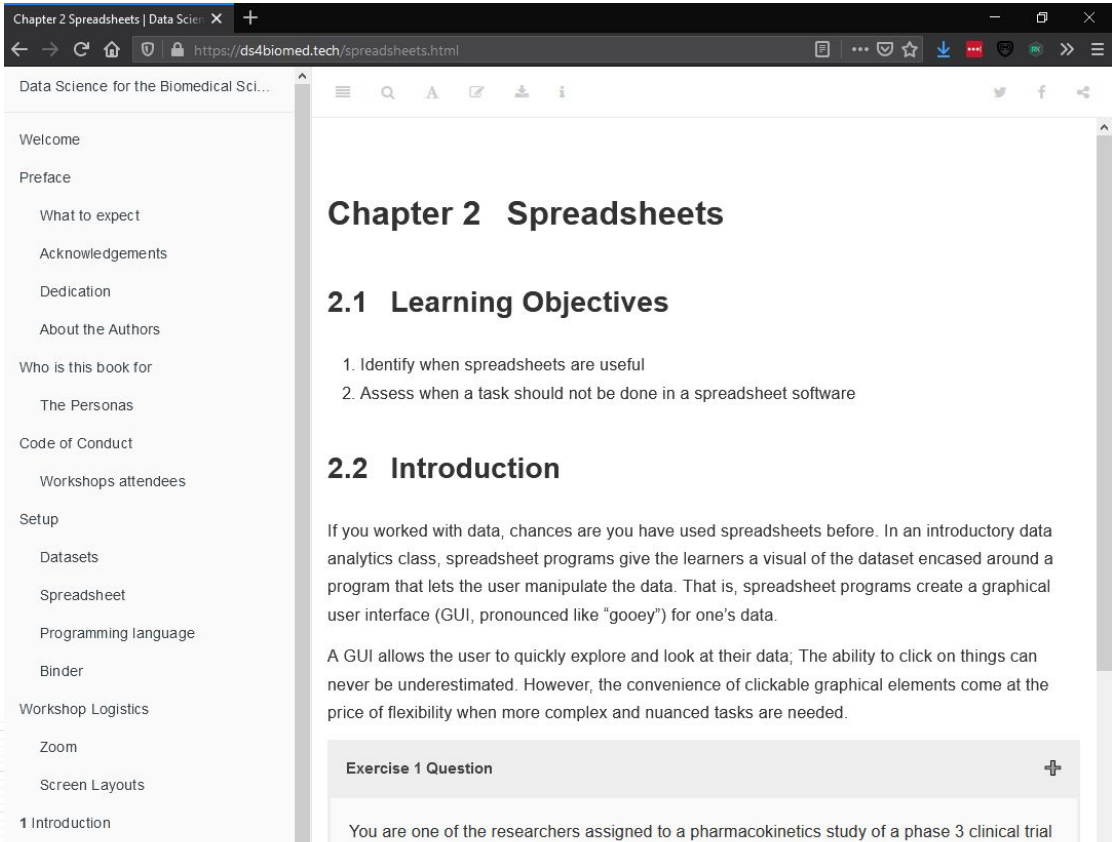
Background

Clare has spent the last 6 years working in the Cardiothroasic ICU in a large medical hospital system. They read lots of gushing articles about data science, and was excited by the prospect of learning how to do it, but nothing makes sense when trying to learn it on their own. Clare has always been a good student and always excelled at things they tried to learn; they are hard on themselves when struggling to learn a new skill and would rather place blame on the long hours at work than having their peers know they could use assistance.

Relevant prior knowledge or experience

Data Science for the Biomedical Sciences: ds4biomed.tech





Thank You!

ds4biomed learning materials: https://ds4biomed.tech/

Slides: https://github.com/chendaniely/jsm-2021-learner_personas