

.... and its relevance to Data Science

Peter Foltz

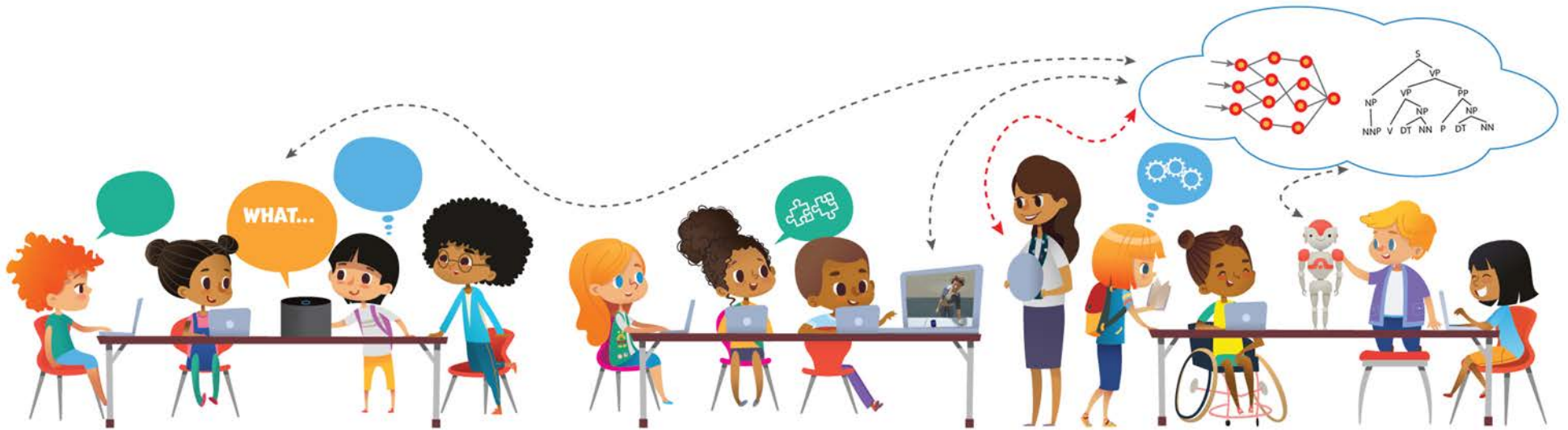
Research Professor, Institute of Cognitive Science

iSAT Overview

The Problem:

**How to promote deep
conceptual learning via rich
socio-collaborative learning
experiences for all
students?**

In our vision, AI is viewed as a **social, collaborative partner** that helps both students and teachers work and learn more effectively, engagingly, and equitably



AI Partners: A bold and transformative approach

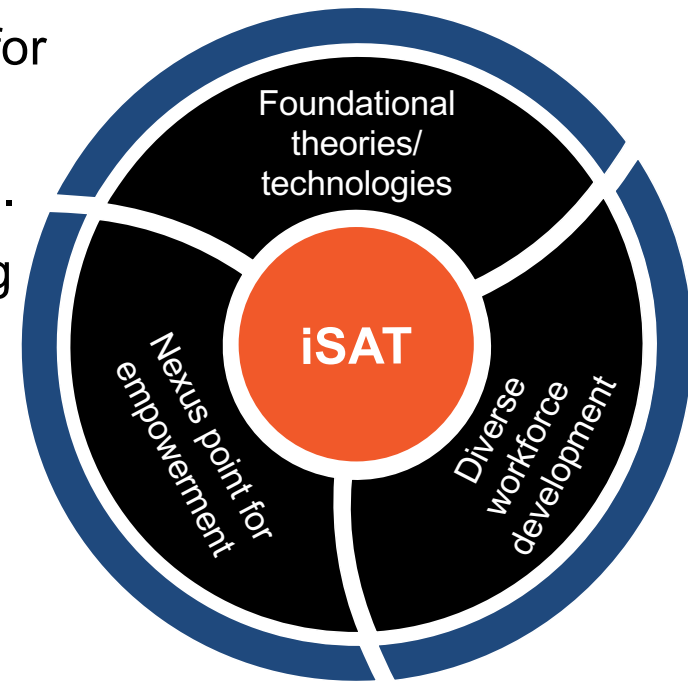
- Most existing educational AI and orchestration systems are **unimodal**, interact with **individual** students, or monitor **typed chats** among students*
- Our AI Partner will observe, participate in, and facilitate collaborative learning conversations by interacting naturally through speech, gesture, gaze, and facial expression in a real-world classrooms (i.e., **multimodal, multiparty**)
- It will help **teachers orchestrate** more effective collaborative learning



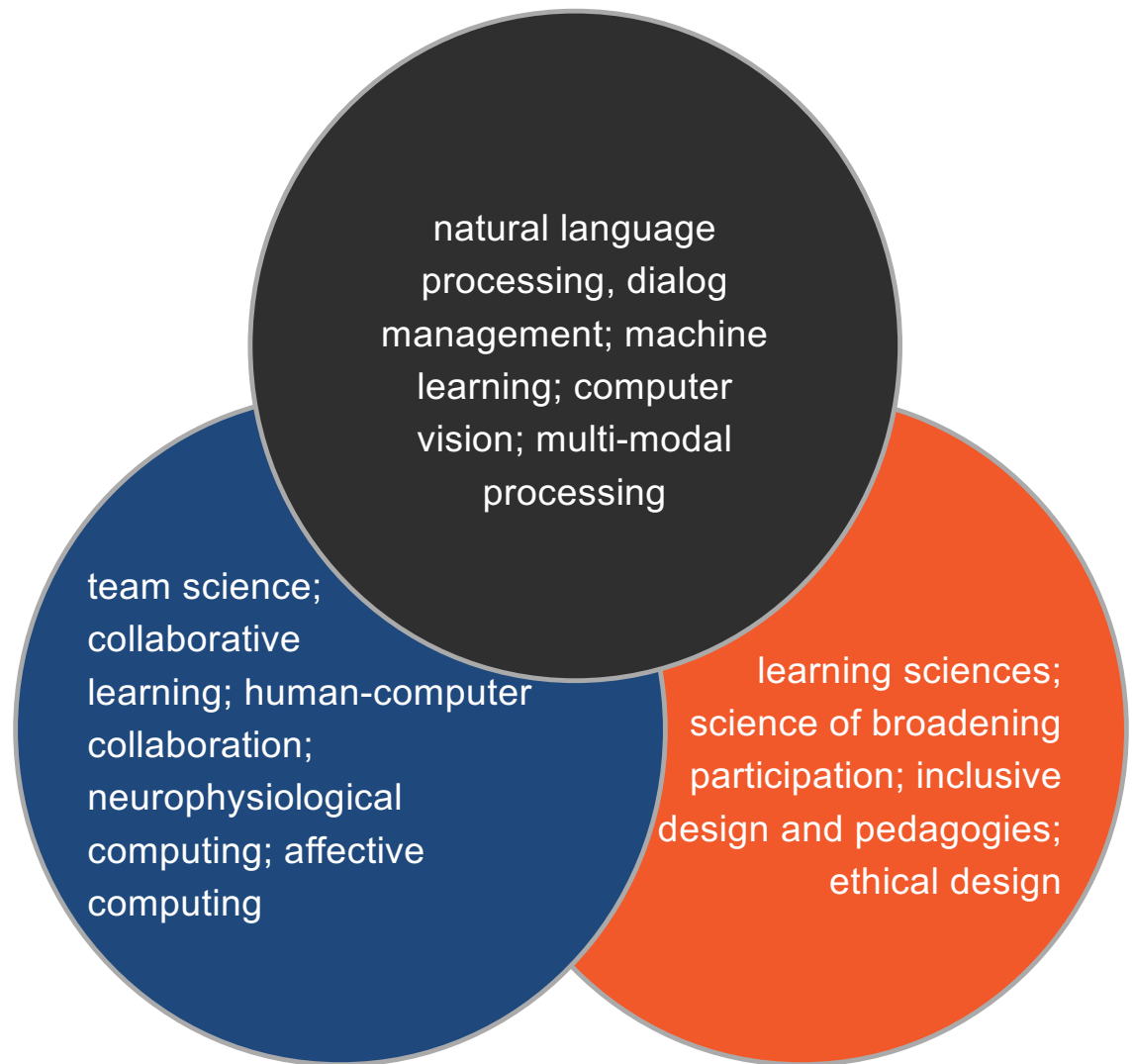
*[Arroyo et al. 2014; Nojavanasghari et al. 2017; VanLehn et al. 2016; Rose et al.]

The iSAT mission

- Develop **foundational theories & AI technologies** for creating next-generation collaborative learning environments composed of diverse student-AI teams.
- **Grow a diverse workforce** of the future by engaging 5,000 middle/high school students in innovative AI education through AI-enabled pedagogies.
- Serve as a **national nexus point for empowering** diverse stakeholders to envision, co-create, critique, and apply student-AI teaming in their communities.



iSAT integrates
researchers from **14**
research areas



iSAT partnerships

Unites 45 researchers from 14 research areas with partners from academia, K-12, and industry network affiliates

University Partners

- University of Colorado Boulder
- Colorado State University
- University of California, Santa Cruz
- University of California, Berkeley
- Brandeis University
- Worcester Polytechnic Institute
- Georgia Institute of Technology
- U. of Illinois at Urbana Champaign
- University of Wisconsin-Madison

K12 Partners

- Denver Public Schools (DPS)
- St. Vrain Valley School District (SVVSD)
- Project VOYCE
- Rigorous Love
- Youth Empowerment Broadcasting

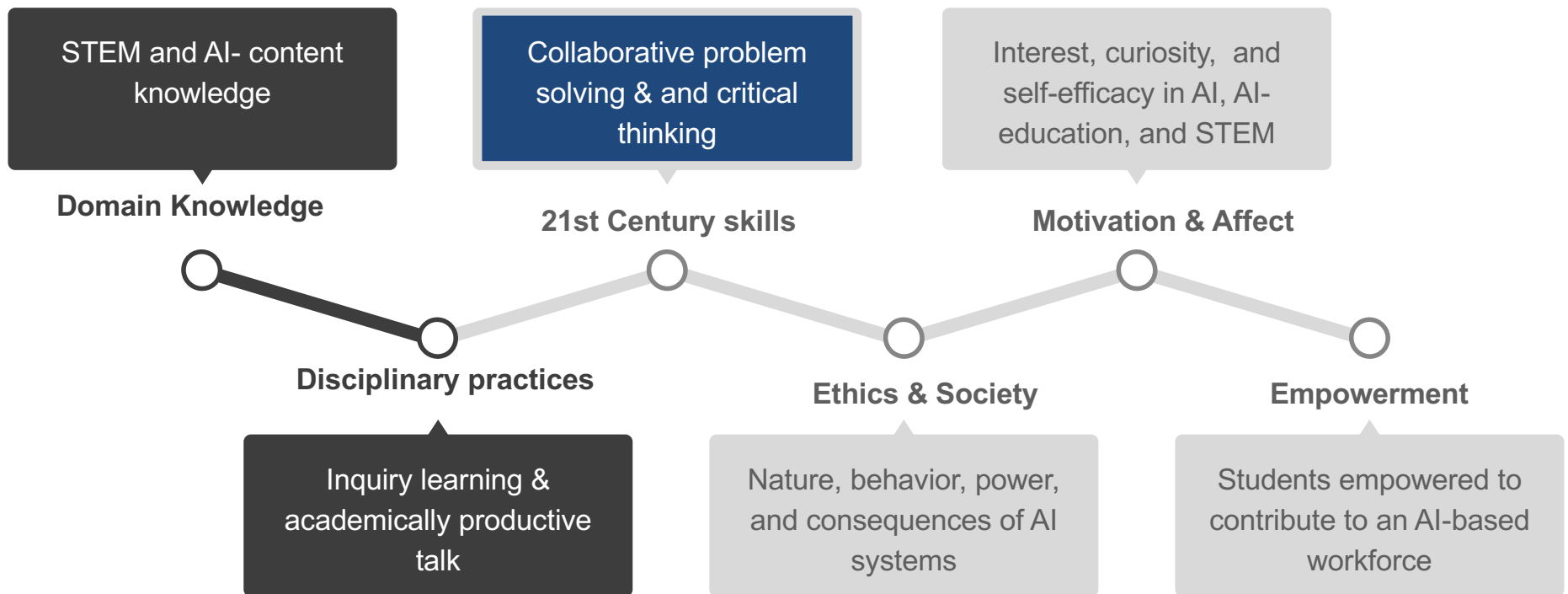
Development Partners

- SparkFun Electronics
- Curve 10

Industry Network Affiliates

- Front Range Technology Network
- Center on Pervasive Personalized Intelligence

iSAT will integrate **AI-education** in science & tech courses to provide measurable learning outcomes



A multi-faceted AI Partner



- We envision an interactive AI Partner that listens to, analyzes and **facilitates** student conversations, **promoting problem solving** and keeping the conversation focused
- Our AI Partner must apply a **control policy** to make **Dialog Management** choices that best promote the group's learning, e. g., correct **Talk Move**
 - **Talk Moves*** are established ways of moving an academic conversation along (REVOICE, PRESS FOR ACCURACY, etc.)
- Good Dialog Management choices will depend on **rich content analysis** and **multimodal perception (speech, vision)** of the learning context

*[Tegos et al, 2015,16; Dyke et al, 2014; Kumar & Rose, 2010, Adamson et al, 2013]

Levels for AI Partner Interaction



Sheepdog

Only “Keeping Everyone Together”
Talk Moves



Shepherd

More guidance, but no true communication with students



Guide

Follows the gist of most content-based conversation and responds encouragingly

Progressive levels of functioning for the AI partner

Level 1

Y1-2

Proof-of-concept prototype for rapid data collection

Level 2

Y2-4

First concrete rendition of the entire system but interaction/orchestration based on speech data only.

First major field trial with middle school students.

Level 3

Y4-5

Will embody the entire vision and will include both speech and video data and more sophistication throughout.

Second major classroom field trial with middle & high school students.

How does Data Science impact iSAT?

How does iSAT impact Data Science?

The Digital Ocean*

Our ability to record, store and process information from learners suggests a paradigm shift

- Focus on a broad range of attributes versus measuring narrowly defined knowledge and skills
- Assessment from in vivo naturalistic tasks vs. via pre-made tests
- Integration of data over activity and time as opposed to over singular events
- Detailed tracking of context outside of testing situations
- Remove current distinctions such as “informal” vs. “formal” learning
- Collection and permanence of learner profile data to make ongoing, intelligent recommendations.

*Dicerbo & Behrens, 2014

https://assets.pearsoncanadaschool.com/asset_mgr/current/20194/Impacts-of-the-digital-ocean-on-education.pdf

Data Collection and modeling

AI Partner

Dialogue
Management

Content Analysis

Automatic
Speech
Recognition

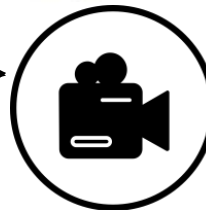
Speaker
Diarization

Multimodal
Student
Modeling

Multimodal SCA
Modeling

Multimodal
Group Modeling

Reinforcement
Learning



AI Partner

Teacher
Dashboard



Data collection in live environments

Data:

- Multi-party
- Multi-modal
- Multi-curricular
- Real-time interaction
- Noisy, Incomplete

Data Collection is Evidence Accumulation:

- Frameworks of what skills are being assessed
- How do we operationally define the skills so they can be measured?
- How do we surface feedback to students and teachers?
- How do we validate performance?

Ethics:

- Avoiding algorithmic bias and its consequences
- Accounting for the role of privacy and security
- Responsible Innovation framework: Promote creativity and opportunities for science and innovation that are socially desirable and undertaken in the public interest

The screenshot displays the goReact interface for an "AI Retreat Annotation Activity". The top section shows a video of a classroom where students are seated at tables. Below the video, a list of annotations is visible, each with a timestamp and a description of the observed behavior. The interface includes a search bar, a "BACK" button, and a "COMMENTS" section. At the bottom, there is a text input field for adding comments and a character count.

goReact COURSES

AI Retreat Annotation Activity INSTRUCTIONS

William Penuel: "You want to do diameter and height" - noticing this is directive speaking, but also turning her body toward the other girl in the group.

00:14

00:12 James Pustejovsky: B turns toward A. B gestures towards A. Clarifying "you want to do diameter and height."

00:14 William Penuel: "You want to do diameter and height" - noticing this is directive speaking, but also turning her body toward the other girl in the group.

00:18 James Pustejovsky: A looks toward C, says "write this down guys"

00:18 Eddie Dombrower: One thing that occurs to me is that this task could easily still be done in a Zoom/Google Meets room... what role would the AI play in that situation... it might even be more easily integrated into the workflow as a 5th member of the room rather than a physical object fighting for attention in a noisy room.

00:18 Terri Dunbar: This could be a good role for the AI--note taker--starting about here

Text Video Audio Upload YouTube Library

Add a text comment...

Characters: 0/500 Timeline: Start Typing... Hit "Enter" to save your comment

Example: Using Educational Data

Building on TalkBack: Accountable Talk Moves DataSet NSF 1600325 & 1837986 (Big Data) – PI, Sumner

Accountability to the learning community

- Keeping everyone together (e.g., "What did she just say?")
- Getting students to relate to another's ideas (e.g., "Who agrees and who disagrees?")
- Restating (e.g., "Let me say back what I heard.")

Ensuring purposeful coherent and productive group discussion

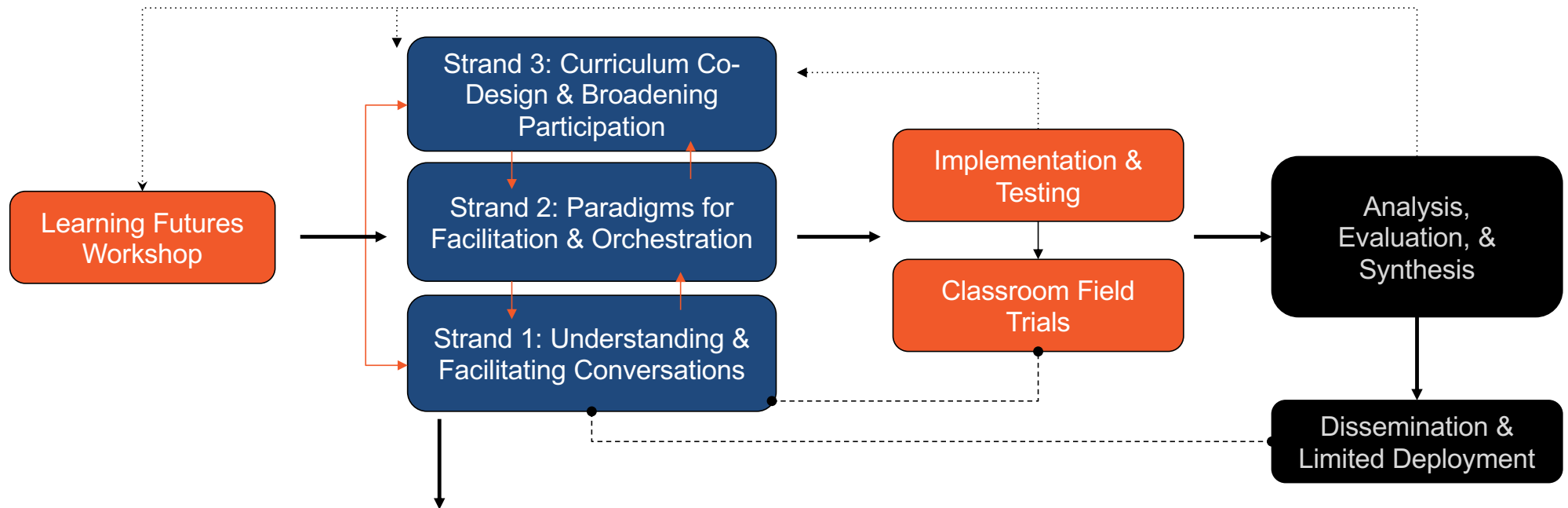
Accountability to rigorous thinking

Dataset & Performance

- ~250K sentences, recorded, transcribed annotated with Talk Moves
- 66% Teacher sentences/ 34% Student sentences
- >60% No Move/ < 40% Talk Moves
- Classifier Performance: Teacher Moves: ~79%; Student Moves: ~77% [Suresh, et. al., 2019, 2021]

Data science project life cycle

iSAT Research Process/Methodology



Foundational Research occurring continually: Curriculum Development; Design & Rapid Prototyping; Data Curation & Annotation; Algorithm Development; Experimentation & Evaluation; Community Resource Development

Designing with the End User in Mind: Co-Design

In what ways can inclusive co-design processes empower stakeholders with diverse identities to envision, co-create, critique, and apply AI learning technologies for their schools and communities?*



Select “Anchor Problem”

- Begin with Futures Vision
- Examine standards
- Brainstorm possibilities with educators and researchers
- Solicit input from students & educators

Build a Unit

Develop a “storyline” for unit driven by student questions
Write lessons with partners
Test and revise lessons based on teacher and students experience & evidence

*[Penuel & Reiser, 2018; Severance, Penuel, Sumner & Leary, 2016; Penuel & Watkins, 2019]

Summary: Data Science in Educational Technology (and other fields as well)

Think of Digital as a Process



From:
Intrusive off-line data
instrumentation & collection
with slow feedback cycles.



To:
Continuous data for
continuous customer insight
and product improvement.

Apply Intelligent Emergent Services



From:
Repetitive human activity that
varies in quality and
consistency.

PEARSON

Writing Assignment: Essay: Gender and Sexuality (Autofeedback)

Due Date: 06/27
Status: Viewing Assignment
Recent Activity
Alerts: 07/17/2015 Assignment instructions updated

Start Assignment

Instructions

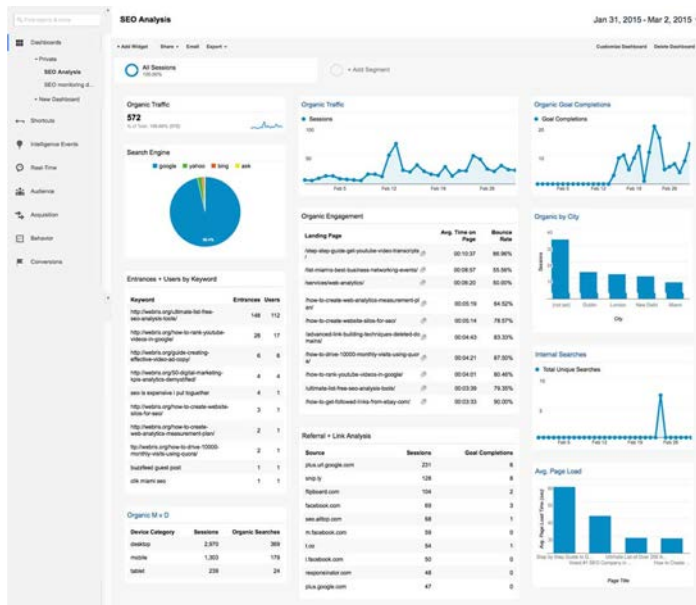
For the past century, part of what sociologists have studied is how social forces affect sexuality, something that is often considered to be largely natural, taken for granted and in the biological arena. While biology no doubt has an effect on sexual identities and behaviors, sociologists have shown that our sexualities are also social. For example many societies put restrictions on sexual behaviors, and other studies have shown that what is considered attractive also varies across societies. Sexual orientation refers to whether individuals are attracted to the other sex, the same sex, or both. While we have categories for each of these states research has shown that most people do not fit neatly into any one category and that sexuality can be thought of as a continuum rather than separate categories.

Importantly, many societies today are what sociologists call 'hetero-normative.' This is one possible reason that sexual minorities

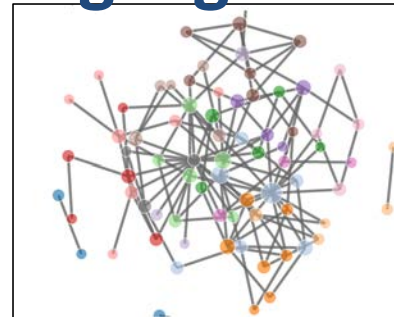
Tools: Sample Documents, Assignment Instructions, Assignment Resources, Checklist, Dictionaries & Thesaurus

To:
Capturing and replicating human
intelligence to scale value and free
people for high-touch needs.

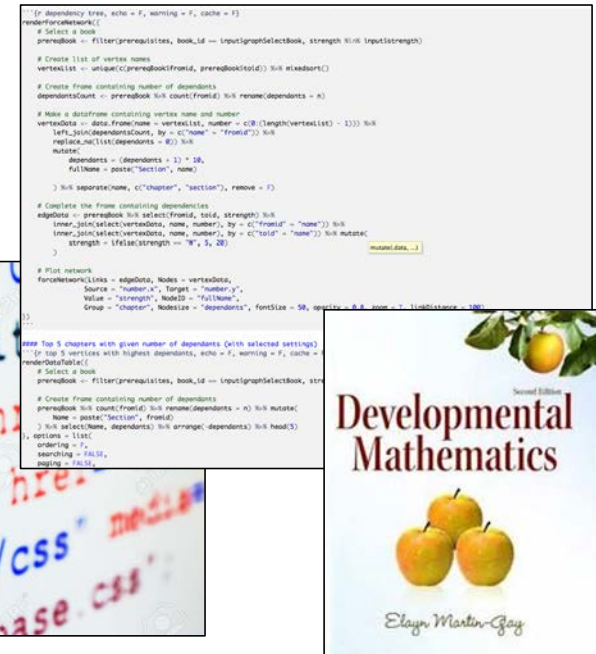
Exploit the fact that “Everything digital is data”



From:
Analytics focused only on
human behavior via
“instrumentation”.



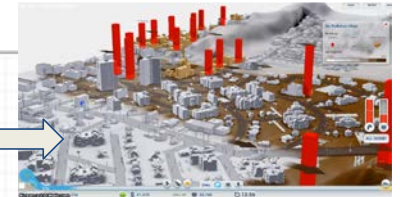
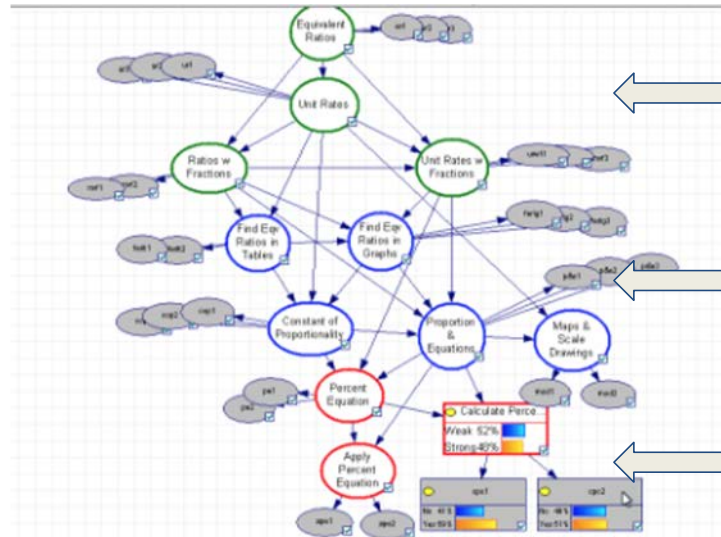
To:
Value creation from all existing
digital assets, whether created by
your teams and customers.



Leverage Assessment as Knowledge Creation



From:
Isolated and non-persistent
descriptions lost in time.



To:
Continually updated and persistent
profiles that support customers along
the journey.

Questions/Discussion

Peter.Foltz@Colorado.edu