

Assignment 4

Project Data Management Plan & Costing

Presented by:

Shreya Goel
Devan Goto
Chuheng Hu

Project Description and Goals

Goals:

1. To testify if the gaming detectors can effectively detect gaming.
2. To distinguish between gaming that hurts from gaming that does not hurt.
3. To generalize the gaming detector to different data set.



Key Words:

“Cognitive Tutor”

“Gaming the system”

“Generalizable detector”

Data Life Cycle

Data Stage	Data Description
Raw Data	Collect Data from schools coordinating with Cognitive tutors
Pre-Processed Data	Dataset – I and II will both include: Knowledge component model, Class observational data, Pretest & posttest scores for each lesson, Log files, Metadata
Processing Data	Data will be processed/cleaned and Gaming Detector will be build: Knowledge map, Class Observation Data, Cognitive tutor log Dataset, metadata, algorithms
Post processing Data - I	Test and Improve the Gaming Detector if needed using Dataset – I
Post processing Data - II	Test the generalizability of gaming detector using Dataset – II.
Analyzing Data	Analyze if the gaming detector can be used as Generalized Detector
Publish Data	Publish the results: Generalized gaming detector
Curate Data	Upload the data to Data Shop and Columbia Academic Commons, including metadata: Scripts, Codebooks; Logfiles, Algorithms, Datasets

Data to be produced

- Algorithms
- Datasets
- Metadata:
 - Scripts
 - Codebooks
- Anonymised Raw
Log files



Challenges and suggested solutions

- Metadata quality -> labour division and collaboration
- Potential for redistribution v.s economic case ->by request
- Credibility Issue ->share log files
- Potential risk in Data Shop ->Multiple repository
- Timeliness of the data -> metadata

Challenge: Vulnerable Population

Solutions

- TC IRB
- Participant Consent Forms
- Anonymization

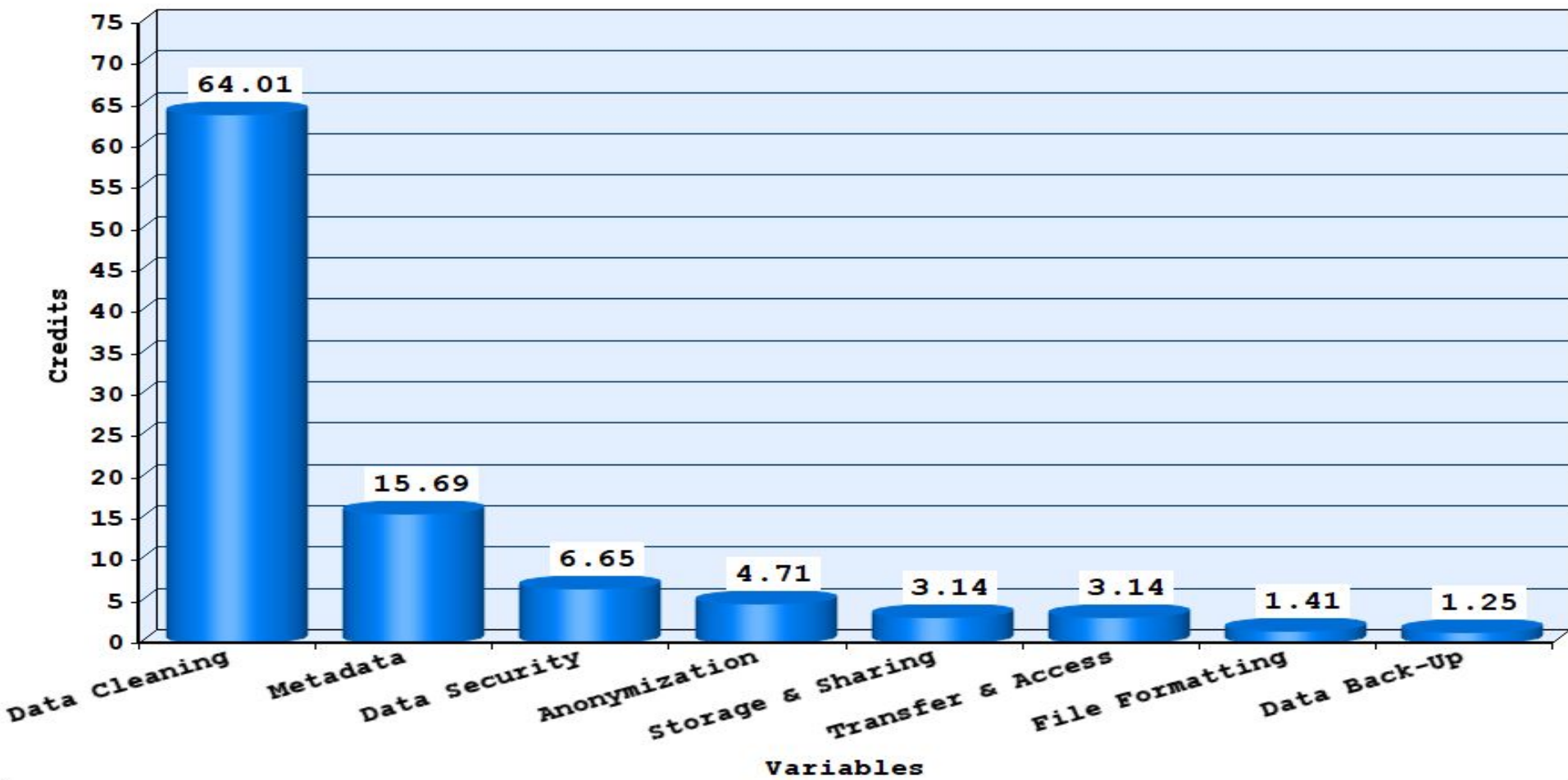
Instill Ethics

- Senate Bill 1177 Ch. 839
 - Do not advertise based on any student data.
 - Do not use student data to amass a profile about a student.
 - Do not sell student data to a third party

Why Impose High Ethical Standards

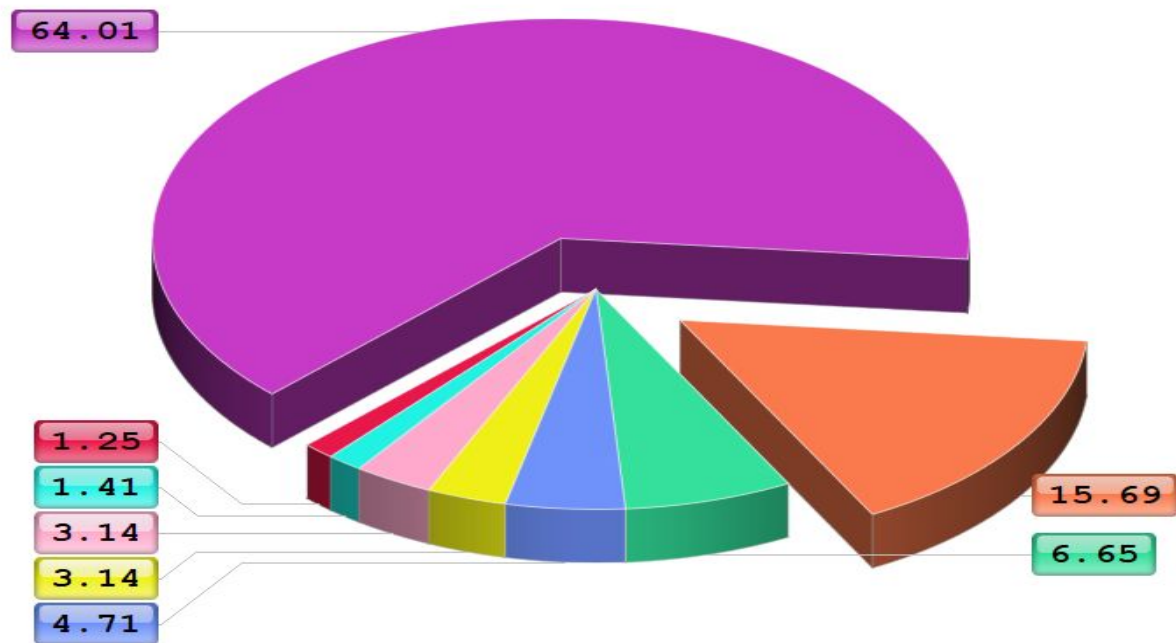
- Precedent
- Trust
- Transparency
- Accountability

Budget



Budget

■ Data Cleaning ■ Metadata ■ Security ■ Anonymization ■ Storage & Sharing
■ Transfer & Access ■ File Formatting ■ Data Back-Up



Questions