



Promoting Data Equality by Improving Open Government Data Users' Data Literacy

Fanghui Xiao
fax2@pitt.edu

Daqing He
dah44@pitt.edu

David Walker
david.walker@pitt.edu

School of Computing and Information
University of Pittsburgh

Western Pennsylvania Regional Data Center
University of Pittsburgh

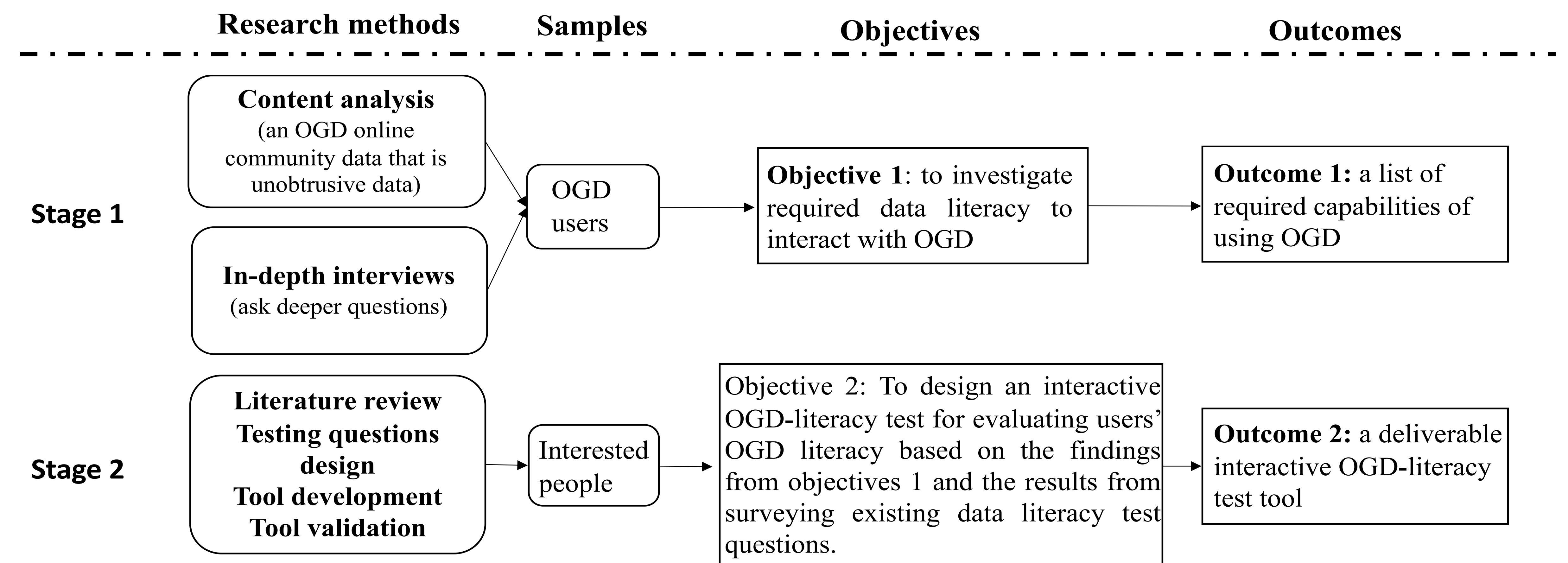


Project Description

This project aims to improve the data literacy of open government data (OGD) users, to increase awareness and use of OGD, and finally, to help address the issue of data inequality.

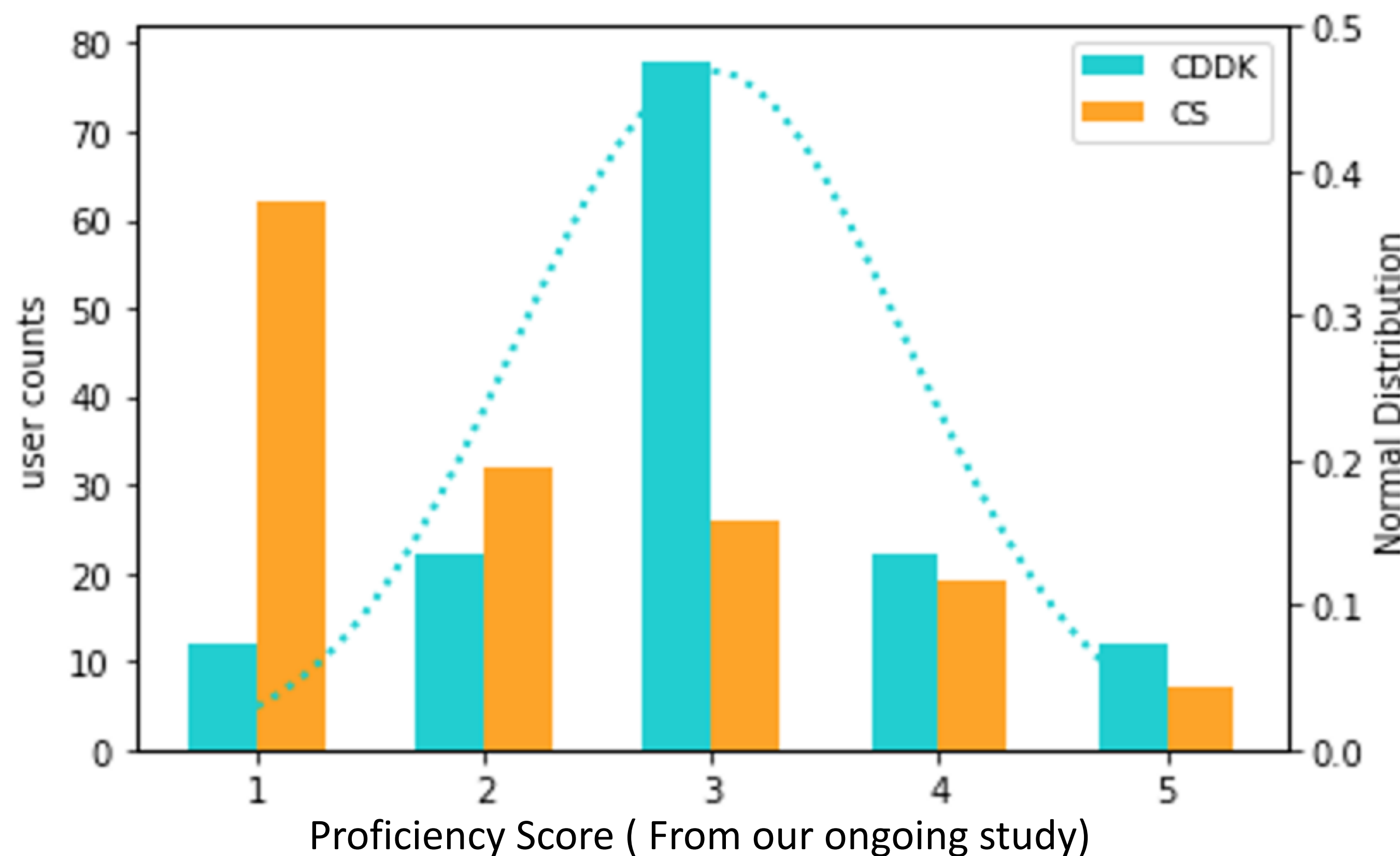
Contributions & Impact

1. Help OGD portals **design their systems** to be more accessible to novice users, which could broaden the fraction of the population capable of using open government data.
2. Assist local-level OGD portals to **develop more effective training workshops and online documentation**, which will directly contribute to the data literacy of users or potential users by helping them learn related knowledge and skills.
3. This project could be **disseminated and applied to other local-level OGD portals in the U.S.**



Why OGD Literacy is Needed?

- 1) the **inequality** between those who are proficient in data-domain knowledge, e.g., storage and collection, and those who are not (Andrejevic, 2014) and
- 2) the **inequality** between those who are proficient in the technical skills required to effectively work with data and those who are not (D'Ignazio, 2017).



Stage 1: How Can We Help People Improve their OGD Literacy?

User OGD behaviors	Required OGD literacy
Data needs	ability to frame the data needs
Finding data	ability to identify the data sources ability to formulate queries/keywords ability to develop seeking strategies
Evaluating data	ability to evaluate if the identified data is topic relevant
Scrutinizing data	ability to assess if the identified data is usable
Acquiring data	ability to acquire datasets
Understanding data	ability to perceive the context of the identified datasets ability to explain the meanings of the field names
Schematizing	ability to explain the logic reasoning of the columns inside of one table or among multiple tables. ability to make sense of the corresponding visualizations
Using data	ability to prepare data for further use, such as extracting, importing and cleaning data ability to transform data from different types for further merging or comparing data ability to merge datasets for further data analysis ability to reasonably analyze data ability to present quantitative data (specific data, tables, graphs) ability to draw meaning from analysis results or patterns ability to apply results to make decisions or solve problems ability to be aware of using data ethically
Sharing data	ability to publish data on various platforms

Stage 2: How Can We Test Someone's OGD Literacy?

Basic OGD knowledge

Category	Sub-category	Skills/Concepts
Basic OGD knowledge	Basic context of OGD	Existence of the civic data The concept of OGD Ethic use of OGD
	Basic knowledge for Data	The file formats of OGD The types of OGD
Information literacy	Find	Seeking strategies (keyword searching, browsing)
	Scrutinize	Needed dates, missing data,
	Understand	Metadata, data dictionary
	Schematize	
Computational thinking	Decomposition	Computational skills Statistics Programming Tool
	Pattern recognition	
	Abstraction	
	Algorithms	
Data communication	Report	Draw meaning from analysis results or patterns
	Decision Making	Apply results to make decisions or solve problems