

LIS 2021 - Fall 2020

Interview Protocol Group 5 - DataTrackers

Introduction

Hello and thank you for taking the time to meet with us today. Our group is made up of four graduate students from the Master of Library and Information Science program at the University of Pittsburgh School of Computing and Information. This project is part of the Design Methods course, which will span the course of two terms, Fall and Spring 2021-2022. The Design Methods course requires that our group partner with an organization to identify a problem within the organization, to design and implement solutions, and to report the results.

We have been asked by the data services team of the City of Pittsburgh to develop a strategic plan for Data Governance Health Status Checks. We will be interviewing a number of stakeholders to gather different perspectives about the City of Pittsburgh's Open Data efforts.

We will take notes during this interview for our own use and anonymize the information you provide by assigning user IDs to keep your responses confidential. Responses will be accessible only to group members, our professor, teaching assistants, and other students within this course. We will share only summarized information with stakeholders of the City of Pittsburgh and the Data Services Team. However, because of the small sample size of no more than 10 interviews, it may be possible to re-identify summarized responses. You may stop the interview, decline to answer, or request to go off record at any time.

Interview Questions

1. Project Lead - History and Scope of Project
 - a. What can you tell us about the origins of the City of Pittsburgh data governance effort?
 - b. We've reviewed the Data Governance Engagement Plan. Describe your expectations for our group within your Data Governance process: how can we help you?
 - c. What are the target government departments for the data governance effort and the Data Governance Health Status Checks?
 - d. How was Data Governance Health Status monitoring developed?
 - e. What do you expect to monitor?

- i. Does this include number of datasets selected/uploaded, timeliness from data preparation to publication, data cleanliness and/or completeness, data documentation?
- f. Tell us about how the city manages Dog License data?
 - i. What department handles this?
- g. Describe any resources you have developed or researched and plan to use for the Training Camp in Phase 3.
 - i. What is your learning platform?

2. Role within organization

- a. What is the primary business of your department?
- b. How large is your department?
- c. What is your role within your department?
- d. What is your role on the Data Services Team?
- e. What is your role with respect to the City's Open Data efforts?
- f. What are your responsibilities with respect to the proposed Data Governance Engagement Plan and Health Status Checks?

3. Background with the Data Services Team and Pittsburgh's Open Data Ordinance

- a. Can you give us a brief description of your connection to the City of Pittsburgh Data Services Team?
- b. What led to your current work with the Data Services Team?
- c. What do you know about Open Data and the Open Data movement?
- d. What is your perspective on the Open Data movement, in general?
- e. What is your perspective on the City of Pittsburgh's Open Data Legislation?
- f. What previous project experience do you have that is related or can be applied to the Dog License data project?

4. Current Processes and Procedures

- a. What kinds of data do you use in your work?
 - i. Where do you get this data?
 - ii. How would you define the data you use on a daily basis?
- b. Describe the process you go through to use your data.
 - i. How do you collect or obtain your data? In what form is it when you get it, e.g., paper, computer?
 - ii. How do you prepare/clean your data?
 - iii. How do you store your data?
- c. What can you tell us about your department's current data governance practices?

- i. Are your current practices aligned with any digital data management and/or data preservation standards?
 - ii. Does your department define their own standards for data management?
 - iii. Do you have any documentation you could share?
- d. What criteria do you think must be met for data to be considered 'healthy'?
- e. What are your current practices for monitoring the health of your data?
- f. Who are the stakeholders and users who access the data?
 - i. What information do you think they would be interested in accessing?
- g. Describe how you share your data within your department?
- h. Describe how you share your data with other City of Pittsburgh departments?
- i. What bothers you about how data works in your department?
- j. What do you think works well about how your department handles data?
- k. What changes would you make to improve how your department handles data?

5. Information Systems

- a. What are the IT systems and tools currently at your disposal?
 - i. How long do you expect to be working with those systems?
 - ii. Which of these tools do you use on a daily basis?
 - iii. Which of these tools do you feel most comfortable using?
- b. What databases does your department use?
- c. What applications does your department use to capture data?
- d. Are any data kept in spreadsheets? Are these on shared or individual drives?
- e. What type of information is captured in the spreadsheet (what are the variables and values)?
- f. How is the data in the spreadsheet created?
 - i. If users create the data, where do they input the data?

6. Open Data Planning and Preparation

- a. For your Open Data, what do you know about the data portal that the City of Pittsburgh will use for publishing data?
- b. What do you know about publishing open data?
- c. Are any of your data currently available to the public online?
- d. What data sources and datasets have you identified as potential for publishing to the Open Data portal?

- i. Do you have an inventory of datasets and sources that you could share with us?
 - e. Do you currently anonymize data?
 - i. If so, What is your process for anonymizing data?
 - f. Describe any work you've done that involves data dictionaries?
 - g. Describe any metadata you have used to describe data.
 - i. What standards do you follow?
 - h. What file formats have you worked with?
 - i. Describe your data cleaning process.
 - i. What tools do you use for cleaning or preparing data for publication/sharing?
 - j. What do you know about licensing data, i.e., granting permissions for access and use?
 - k. How will you promote the use of your data?
 - l. How do you see the data management requirements fitting into your workflow?
7. Sustainability
- a. What would make it easy for you to add the data management requirements to your workflow?
 - b. Describe what might hinder your data management efforts?
 - c. What are your thoughts about the Data Governance Health Status Checks?
 - d. Who will be responsible for managing and maintaining your data going forward?
 - e. How do you propose to publish this data once organized?

Closing

We appreciate you taking the time to discuss this project with our group. If you have additional information you think will be helpful feel free to contact any team member. If you have questions or concerns about our work, contact Professor Chelsea Gunn at cmg100@pitt.edu. Likewise, feel free to contact us if you have further questions or concerns.

Thank you,
DataTrackers