

Sponsor Meeting 5

Date:	26 Jan 2017
Time:	3.30pm – 4.30pm, followed by 4.30pm-5.30pm Team Meeting
Venue:	SMU Library
Attendees:	Tan Jun Rong Wang Jing Xuan Aaron Tay (LKS Library Liaison) Nursyeha Binte Yahaya (Librarian) Salihin Mohd Ali (Digital Initiatives Manager) Ler Wee Hiong (Librarian) Don Danping (Research Data Librarian)
Absentees:	Lim Yu Xiang Bendexter
Agenda:	Obtaining full years' worth of Dataset & General updates

Background about this meeting

This meeting was originally supposed to be a mere obtaining of a full years' worth of data from Nursyeha, who had previously passed us only 1 months' worth of dataset to conduct our preliminary analysis.

As Jing Xuan had managed to separate the strings in the log files into search queries, Nursyeha decided to call for a meeting with Aaron Tay, who was impressed with the developments as none of the library team had managed to do such a thing before. Thus, Aaron called us to the conference room for an impromptu meeting with the other librarians.

No.	Details	Action By	Due Date
1	Obtaining full years' worth of Dataset from Nursyeha - This dataset was transferred to Jing Xuan's laptop	Nursyeha Jing Xuan	-
2	Impromptu Meeting with the Library Team - Jing Xuan presented how he extracted the search queries from strings of log data to the amazement of the Library Team as this is something that the Library Team have not managed to accomplish before - Jun Rong presented the proposal done thus far and explained why there is a delay in the confirmation of the proposal <ol style="list-style-type: none"> The delay in confirmation of the proposal is due to our group having to change Supervisor from Prof. Kam to Prakash as Prof. Kam's schedule is overloaded with other groups We have taken time to bring everything up to date with Prakash after our proposal was done according to Prof. Kam's instructions 	Aaron Tay Nursyeha Salihin Ler Wee Hiong Don Danping Jing Xuan Jun Rong	-

3. Prakash has looked through our proposal post-submission date and asked us to amend some details to make it more in line with what he expects of this project, hence the delay
- Jing Xuan then presented the anomaly whereby log data is logged every 2 seconds or with every key press

Example 1: Log data is logged every 2 second

12:55:02PM Search Query: **Re**

12:55:04PM Search Query: **Resol**

12:55:06PM Search Query: **Resoluti**

12:55:08PM Search Query: **Resolution**

Example 2: Log data is logged with every key press

Search Query: **T**

Search Query: **Tr**

Search Query: **Tru**

Search Query: **Trum**

Search Query: **Trump**

- This proves to be cumbersome when determining the final keyword to identify as the search query as we need to be able to pinpoint which is the exact completed term that a user is searching for. Example 3 below will further illustrate this problem:

Example 3: List of 3 keywords being logged with every keystroke by User A

[Start of Session for User A]

Re

Regu

Regula

Regulati

Regulation

Ja

Jack

	<p>Si Sing Singap Singapor Singapore</p> <p>[End of Session for User A]</p> <ul style="list-style-type: none"> - Therefore, we need to think of how do we extract the middle term & the term before that as a search query when it is not the end of the session for User A. 		
3.	<p>Sharing by the Library Team on possible directions this project could embark</p> <ul style="list-style-type: none"> - Searches done by law students on law database/resources may be based on assigned readings and hence not all that interesting to analyze - Euromonitor could potentially be a database to focus our analysis on as the searches are more open-ended and not limited to prescribed readings (Eg. This database is commonly used among Marketing students when finding out which companies to base their projects on) - Aaron will provide a list of databases categorized by Industry, Ebooks etc - Same searches done at the same time by many individuals may show that a librarian is teaching search querying skills in a class, best to identify these and omit from actual analysis - Researchguides provided by majors/faculty may be why students are searching in the same way - Boolean searches are boring to the Library Team, we can do this but they would not recommend us to do so - PWR's topic for this semester is Medical Sports - Could potentially narrow analysis to Masters/PhD candidates who are doing research for real research behavior, setback is that the volume of data may not be large enough to confirm any hypothesis 	<p>Aaron Tay Nursyeha Salihin Ler Wee Hiong Don Danping Jing Xuan Jun Rong</p>	-

The meeting was adjourned at 5.30pm after a tabulation of meeting minutes between [Jun Rong](#) and [Jing Xuan](#). These minutes will be circulated and adopted if there are no amendments reported in the next three days.

Agenda for next meeting

1. Address all the concerns with Prakash

Prepared by:

Tan Jun Rong

Vetted and edited by:

Jing Xuan

Read and acknowledged by:

Bendexter Lim