Notes:

Sophie: using smaller data in order to get the processes running, rather than trying to do it all.

Scheduling future hackathons – where are the dependencies? How do the groups need to interact and what do they need in advance to do it.

Vittorio: good to have various groups gather, even for an hour in advance; the further in

advance they can get access to preparatory material.

Frank: $100, Sophie, $100; Neha, $250.

Me: report: what we’ve done and what we plan to do. Roadmap for each of the areas.

Stephen – Write Road map, as well as reflective report: point out how we define minimal viable

projects at various points, so people have a way to orient themselves and their skill sets to the project. **Structuring engagement and giving people a place to start; hierarchy of projects; key accomplishments we want to complete along a timeline; goals that would require people with specific skill sets that we might want to recruit. Identify bottlenecks. List of people that we have and people we want.**

Sophie – keep in the loop;

**People**:

Metadata Enrichment: Christine and Sophie. Would have benefitted with more

connection to search and visualization groups in order to define direction. Having more coders generally, any coders could be good. Is there a better way to do this? Can someone take this thing we’re working on and run with it, once we’ve defined something. Python could be the glue to bring the group together,

Neha and Frank; Cole, Artie

Darian. Pretty crucial for keeping things together late at night. Great at corralling the

novice coders who were in the room. This is a really important skill for someone

to do, also just to get people on to github. Essentially we need sacrificial lambs in

terms of social cohesiveness and emotional labour and explains technical problems (getting on internet, github, instances) to people. Not sacrificial lamb, but rather **facilitator**. Some people just want to hang out.

Vittorio, Artie, Cole, Web Development people.

Do you have a set of tasks (Y/N); Will someone teach you and they have the bandwidth

(Y/N); Can you do something less demanding (Y/N); did you make a bad decision in coming here (Y/N). Printouts of the results: there could have been just a big printout of all the metadata; a big printout of a results set; 300 passages. Having something visual that people can gather around is very good.

Having a survey that you can give to people based on check

Next Meeting: Search Dev and Metadata Enrichment:

April 13:

Morning: Stephen, Ceilyn,

Aft: Stephen Neha, Ceilyn, Christine

April 20:

Morning: Stephen, Ceilyn, Neha, Christine

Aft:Everyone

April 27:

Morning: Stephen, Neha, Christine

Aft: Christine, Neha, Stephen

Others to invite: Claire, Cole, Artie, Elliott, Rashmi, Daina, Vernica

Preparatory Tasts: Run chunking script to retain tails; Run Neha’s script; Decide on and set up AWS instance or MongoDB account, Increase the size of the instance by 1GB;

Goal: with the relational database already complete for TF and passageID, enrich metadata and associate mongoDB with database for metadata; give everyone administrative access; complete reconciliation with wikiData (the full thing, not just the first chunk); partially address authornames in advance

1) Associate metadata with MongoDB

2) Write a search query for MongoDB + metadata

3) Get metadata into mongodb

4) Other statistical measures: cosine similarity, non-negative occurrence of

terms, etc.

Next meeting: Early May, after the first week of may.First Friday of May. Neha has a take-home exam that ends on that. Ceilyn has a presentation on the 8th, so the fourth isn’t great. Vittorio’s available in the afternoon: Rope in others from HUIT website development: Artie, Daniel Seaton. Christine might not be available, as well.