Diploma thesis Outline

1. **Introduction**
- Brief description of Similarity Searching field - why is it important?
- Motivation for Sketch approach.
- Intruduction to this thesis - What is it going to be about?

2. **Definitions**
- Proper definition of Metric space, Distance function, Similarity search query types, Partitioning, Filtering, Indexing and everything that is mentioned in this thesis and reader needs to understand.
- Proper definition of Sketches, how are they meant to be used - in memory, with hamming distance, candidate set, …
- Description of attributes which make a 'good' Sketch according to Vladimir's research.

3. **Problem definition**
- Description of problem that is going to be solved in this thesis - We have close to 'good' Sketchces and now we need to search among them in-memory.
- What are challenges?

4. **Exiting approaches description**
- Sequential Scan - in which cases it might be sufficient, what are the disadvantages
- Hash Index - why it appears as good approach, why it is not for larger r etc.
- Multi-index hashing - key idea, how does it affect number of buckets, how is m affecting number of visited buckets and number of candidates… The special case when r<m.

5. **Designed custom approach description**
- TO DO
- we will focus on special case when r<m
- design index and it's architecture - maybe multiple hash indices
- one simple application (1 hash index), one complex (multiple MHI)
- Idea is that indices for different value of r are explored in order
- System architecture

6. **Experiments definition**
- heavily depends on 5
- evaluate attributes of both single and multiple MHI
- in this section all questions that need to be asnwered will be defined

7. **Experiments implementation**
- Description of used data set
- Notes on architecture implementation, etc

8. **Experiments evaluation**
- Results of experiment - visalization, description
- Conclusions for each result

9. **Conclusion**
- Conclusion of results of this thesis - what have we shown

- Suggestions for future work