

Portrait of an Indexer—Computing Pointers Into Instructional Videos

Andrew Lamb
Stanford University
andrew.lamb@stanford.edu

Jose Hernandez
Stanford University
josehdz@stanford.edu

Jeffrey Ullman
Stanford University
ullman@cs.stanford.edu

Andreas Paepcke
Stanford University
paepcke@cs.stanford.edu

ABSTRACT

1. INTRODUCTION

2. RELATED WORK

3. PREPARATION OF GOLD INDEX

Describe how we had the index prepared: Paid workers, the course, instructions they received. Some stats: average number of index terms they came up with, inter-rater reliability.

4. EXPERIMENTS

Introduce what we are trying to do: explore different methods for building the index, and comparing against gold. Foreshadow the methods: naive/pedagogy-inclusive/external-knowledge.

List whatever is common to the three experiments: stop word removal, method used for stemming...

4.1 Naive Approach: TF-IDF

Simply discard sequencing information of terms occurrences. For each video put the top-n tf/idf terms into the index [1].

4.2 Leverage Instructor Sequencing Decisions

Do we have any experiments that do take the positions of terms in the sequence of videos into account? Such an experiment would nicely tell a story from naive TF/IDF to including pedagogy to including external wikipedia.

4.3 Adding External Knowledge

All of the wikipedia inclusion work.

4.4 Summary of Results

Here is where the comparisons between the two/three experimental results happens.

5. DISCUSSION

Any observations we have about the results.

6. CONCLUSION

Overall summary, open questions, next steps.

7. REFERENCES

- [1] C. D. Manning, P. Raghavan, and H. Schutze. *Introduction to Information Retrieval*. Cambridge University Press, 32 Avenue of the Americas, New York, NY 10013-2473, USA, 2008.