Knowspace: Integration with Elastic Search

*Author: Amir S. Husain*

*Email: ah3548@nyu.edu*

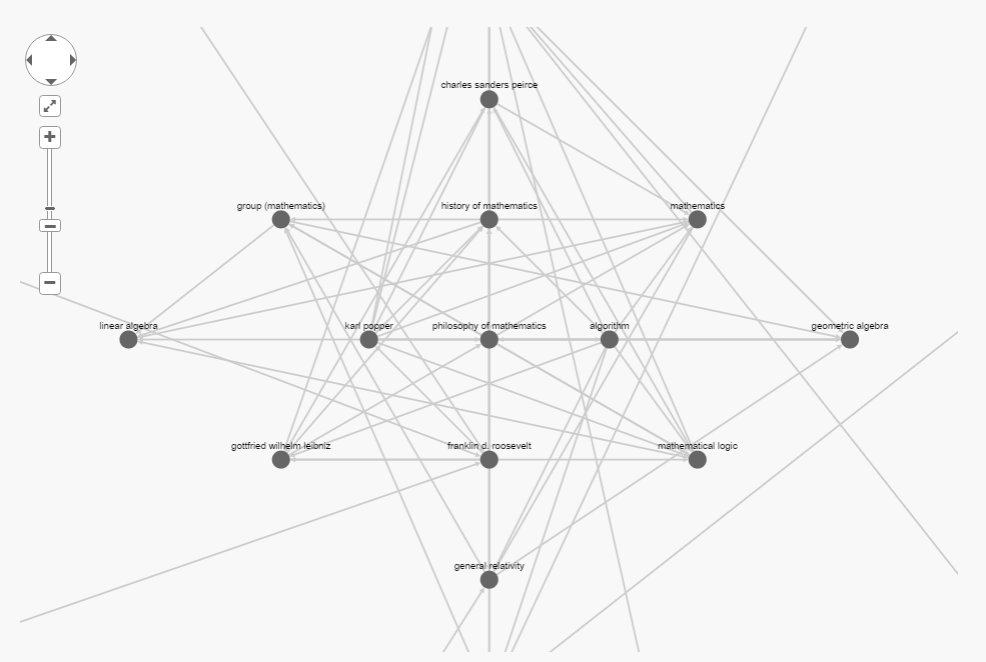
**ABSTRACT**

We aim to change the way adults (>14) approach education. Many learning styles exist and our approach is to let the user figure that out themselves. We do this by treating learning as a "Choose Your Own Adventure" game. By presenting the "Knowledge Space" through a node graph visualization we give the student a map. To present the associations and material we leverage Wikipedia to give clarification around complex topics.

**Keywords**

**1 INTRODUCTION**

The purpose of this project is to take an existing knowledge base (Linear Algebra) and present it in a new user-friendly manner. A node based graph approach is intuitive when discussing a subject but the methodology in generating this graph is the key piece of insight. Integrating with Elastic Search provided much benefit in its ability to store and index documents. The graph is generated off the ability to run a “more like this query” against a document. [1]



**2 SCOPE**

As a proof of concept we are going to present a use case to learn linear algebra using our tool. We are going to use elastic search to determine topic relationships. Elastic search will by default use tf-idf but this is configurable and can be run using other methods.

**3 MOTIVATION**

By trying to find documents that are highly related to each other we are trying to identify pre-requisites to knowledge. For example, if matrices are frequently mentioned in the document about linear algebra and/or the document about matrices frequently discusses linear algebra we can identify a potential dependence relationship. If we identify enough of these relationships we can find a “common core” to learn a subject (i.e. the node in figure 1 of history of mathematics) and potential fringe topics that may spark our interest (i.e. the node in figure 1 of franklin d. roosevelt).

**4. ARCHITECTURE AND TECHNOLOGIES**

**4 RESULTS AND COMPARISON TO EXISITING SOLUTIONS**

**4.0 Experiment Parameters**

When finding documents more like “Linear Algebra” we looked for the 7 most similar and for each result found the 7 most similar to that to generate the graph for a total depth of 3 include the root node. This is highly configurable and can be adjusted under different circumstances.

Some level of indexing is done pre user interaction and more data is indexed as user interact with the application. The below results are prior to any user interaction and ideally would further improve as more documents get indexed from intelligent user interaction.

**4.1 Standard Textbook**

|  |  |
| --- | --- |
| Introduction to Linear Algebra **[2]** *(<Chapter #> <Chapter Title #>)* | Wikipedia and Elastic Search*(<Distance from Linear Algebra Node> <Title>)* |
| **1 Introduction to Vectors** |  |
| **2 Solving Linear Equations** |  |
| **3 Vector Spaces and Subspaces** | **1 vector space** |
| **4 Orthogonality** |  |
| **5 Determinants** |  |
| **6 Eigenvalues and Eigenvectors** | **2 eigenvalues and eigenvectors** |
| **7 The Singular Value Decomposition (SVD)** | **1 singular value decomposition** |
| **8 Linear Transformations** |  |
| **9 Complex Vectors and Matrices** | **2 complex vectors and matrices** |
| **10 Applications** |  |
| **11 Numerical Linear Algebra** |  |
| **12 Linear Algebra in Probability & Statistics** |  |

**5 FUTURE WORKS**

* A key improvement would be the ability to read and write data back into stack overflow so that interaction is live and interactive.
* Stack overflow is not the be-all and end-all of subject forums, i.e. integration with reddit could improve our information repository.
* The tool can be applied to a classroom setting where teachers are experts guiding students along their journey.
* Gamification of tool can further improve motivation, i.e. awarding badges.
* Further research can be done on stack overflow question and answer owners.
* Non-text based information (i.e. images) can be cataloged as well.

**6 REFERENCES**

[1] Anon. More Like This Query | Elasticsearch Reference [5.3] | Elastic. Retrieved April 30, 2017 from <https://www.elastic.co/guide/en/elasticsearch/reference/current/query-dsl-mlt-query.html>

[2] Anon. Introduction to Linear Algebra, Fifth Edition (2016). Retrieved April 30, 2017 from <http://math.mit.edu/~gs/linearalgebra/>