



Indian Institute of Science
Bangalore, India
भारतीय विज्ञान संस्थान
बंगलौर, भारत

Department of Computational and Data Sciences

DS 294 – Data Analysis & Visualization

Project Proposal

Shriram R



Department of Computational and Data Sciences

Problem Statement

- Perform Analytics and Visualization of Knowledge Graphs
- Project Deliverable – Interactive Website
 - ▶ Complete Searching of Vertices/Edges
 - ▶ Traversal from one Vertex to another
 - ▶ Analytics [1]
 - Link Prediction
 - Link based Clustering
 - Autocorrelation

Resources

- Microsoft Azure/Google Cloud platform

- Hosting Website and Data



- D3.js (Data-Driven Documents)

- Graph Visualization

- Neo4j

- Graph database (backend)
- Analytics on graph data





Dataset

- Primary Dataset – YAGO [2]
 - ▶ Huge Semantic Knowledge Base
 - ▶ Sourced From Wikipedia, WordNet and GeoNames
 - ▶ Approx. 10 million entities & 120 million facts
 - ▶ Total size: 170 GB
 - ▶ Confirmed accuracy of 95%
- Secondary Datasets
 - ▶ DBPedia
 - ▶ NELL