

# A Picture Is Worth A Thousand Words

An Application Of Knowledge Graph To Electronic Records Systems

Shin-Chung Shao, Infodoc Technology Corporation  
Cheng-Wei Tsai, Infodoc Technology Corporation

# Contents

- Research Motivation
- Storyboard Like Display -- Our Approach
- Conclusion and Future Research

# Research Motivation

- The Google-Like search interface

The screenshot shows a search interface with a search bar containing "i have a dream". Below the search bar are navigation links for All, Images, Videos, News, Books, More, Settings, and Tools. The search results indicate about 1,350,000,000 results found in 0.82 seconds. The first result is a PDF link from the National Archives titled "'I Have a Dream' speech - National Archives" with a link to <https://www.archives.gov/files/press/exhibits/dream-speech.pdf>. The second result is a link to the King Institute at Stanford University titled "'I Have a Dream,' Address Delivered at the March on Washington for ..." with a link to <https://kinginstitute.stanford.edu/.../i-have-dream-address-delivered-march-washingto...>. Below these results, there is a section titled "Videos" showing three thumbnail images.

About 1,350,000,000 results (0.82 seconds)

[PDF] "I Have a Dream" speech - National Archives  
<https://www.archives.gov/files/press/exhibits/dream-speech.pdf> ▾  
I am happy to join with you today in what will go down in history as the greatest demonstration for freedom in the history of our nation. Five core years ago a ...

"I Have a Dream," Address Delivered at the March on Washington for ...  
<https://kinginstitute.stanford.edu/.../i-have-dream-address-delivered-march-washingto...> ▾  
In his iconic speech at the Lincoln Memorial for the 1963 March on Washington for Jobs and Freedom, King urged America to "make real the promises of democracy." King synthesized portions of his earlier speeches to capture both the necessity for change and the potential for hope in ...

Videos

I Have a Dream

Speech

"I Have a Dream" is a public speech delivered by American civil rights activist Martin Luther King Jr. during the March on Washington for Jobs and Freedom on August 28, 1963, in which he calls for an end to racism in the United States and called for civil and economic rights. [Wikipedia](#)

Originally published: August 28, 1963

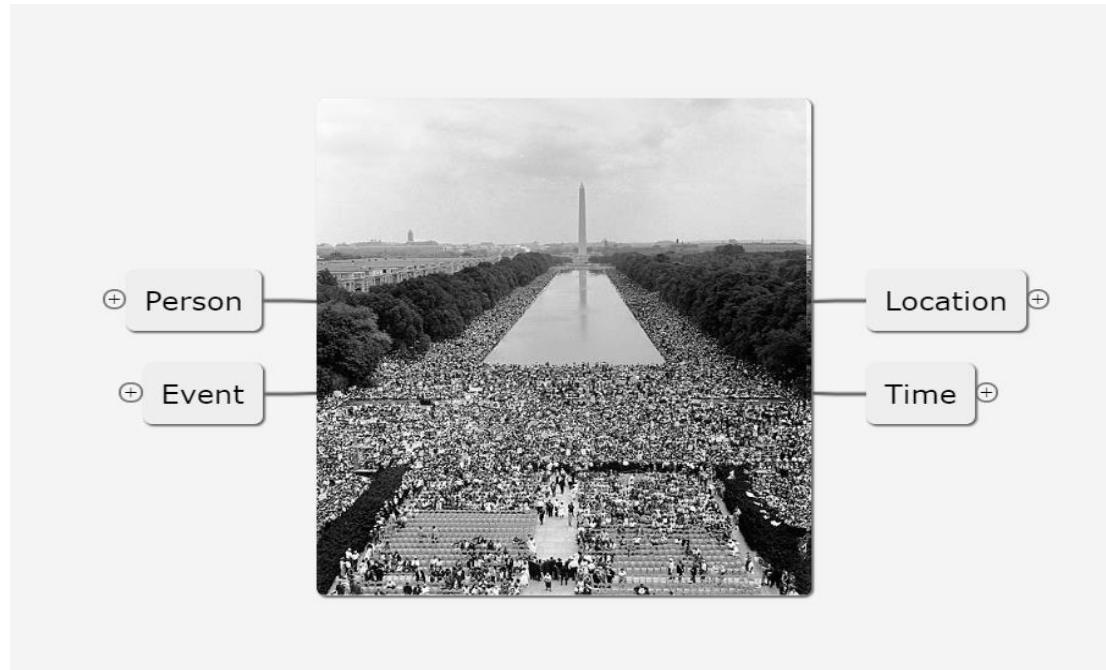
Author: Martin Luther King, Jr.

# Research Motivation

- The search results, a list of links directing to articles, audios, videos, images, may themselves related, in the sense that there may be ***classification rules***, ***association rules***, ***chronicle orders*** and ***semantic rules*** among them, but not appropriately presented.
- Indeed, those historical electronic assets are history, they have ***stories*** involved. Therefore, we ask ourselves, can we, insteadly, provide a ***Storyboard-like search results display?***

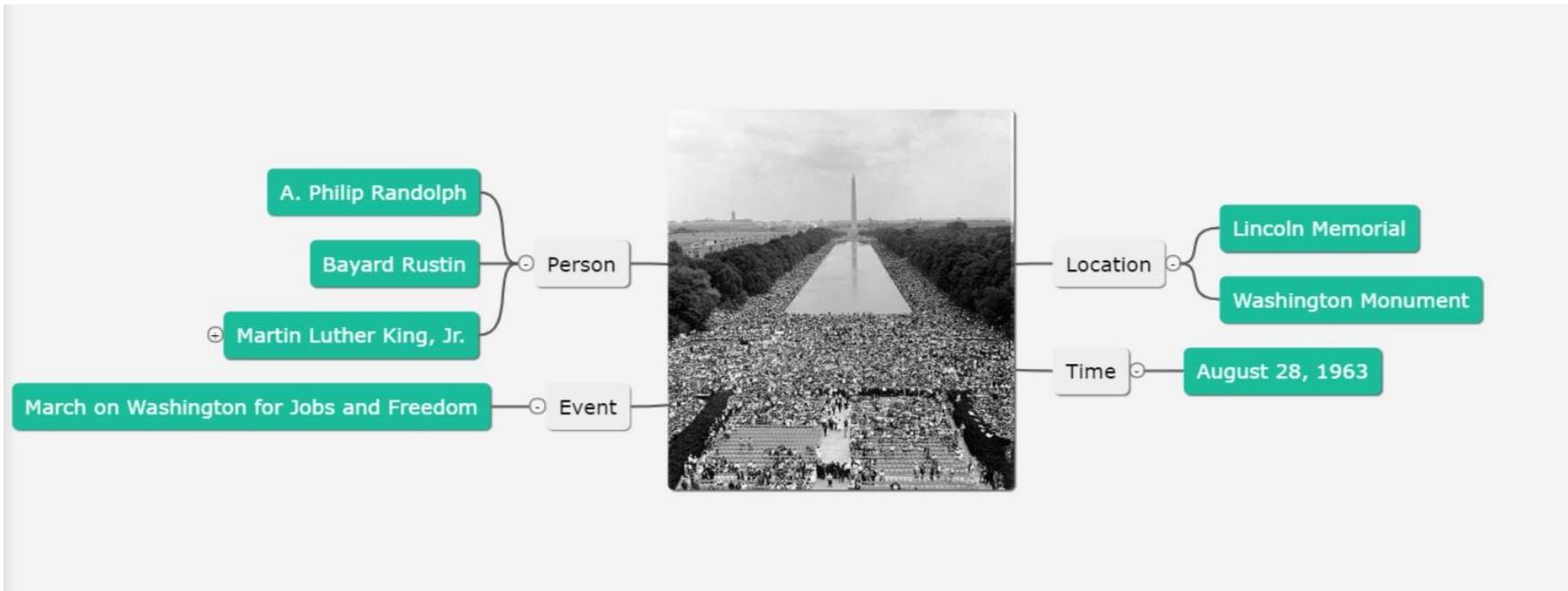
# Storyboard-Like Display

When a user inputs “I have a dream”, the system responds something like the following:



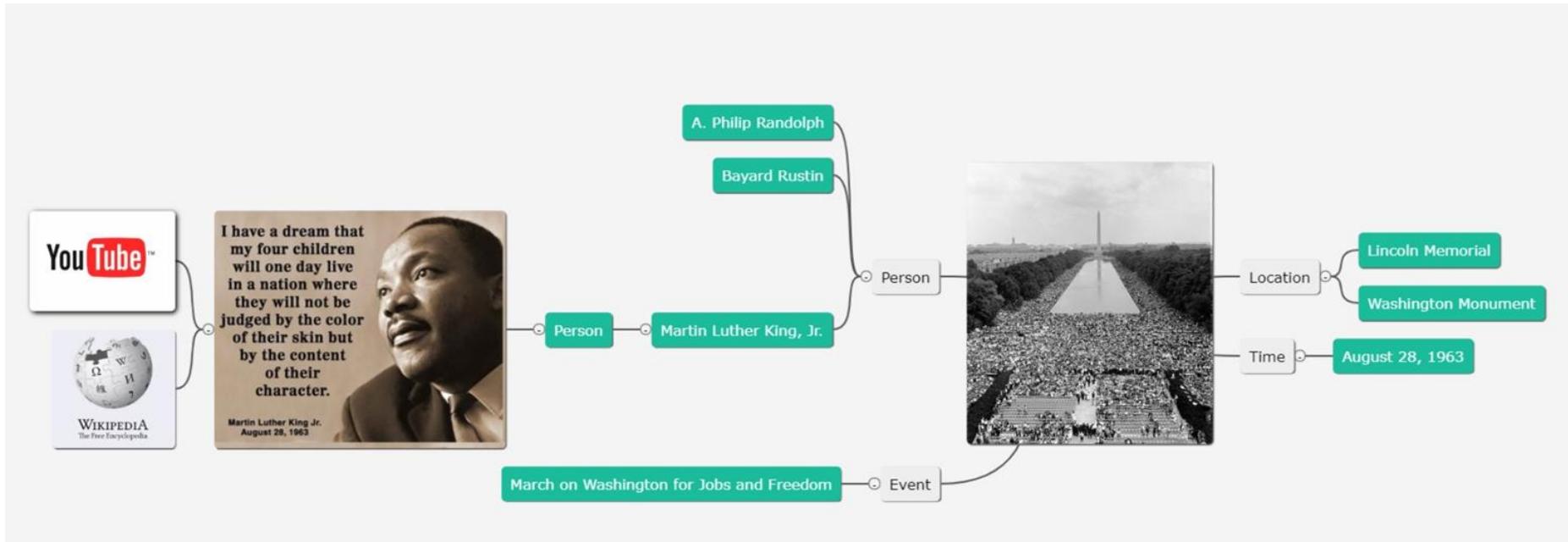
# Storyboard-Like Display

Clicking the + icon of each node, the node expands and displays more nodes:



# Storyboard-Like Display

The leaf nodes are actually links to web resources, such as Youtube, Wikipedia.



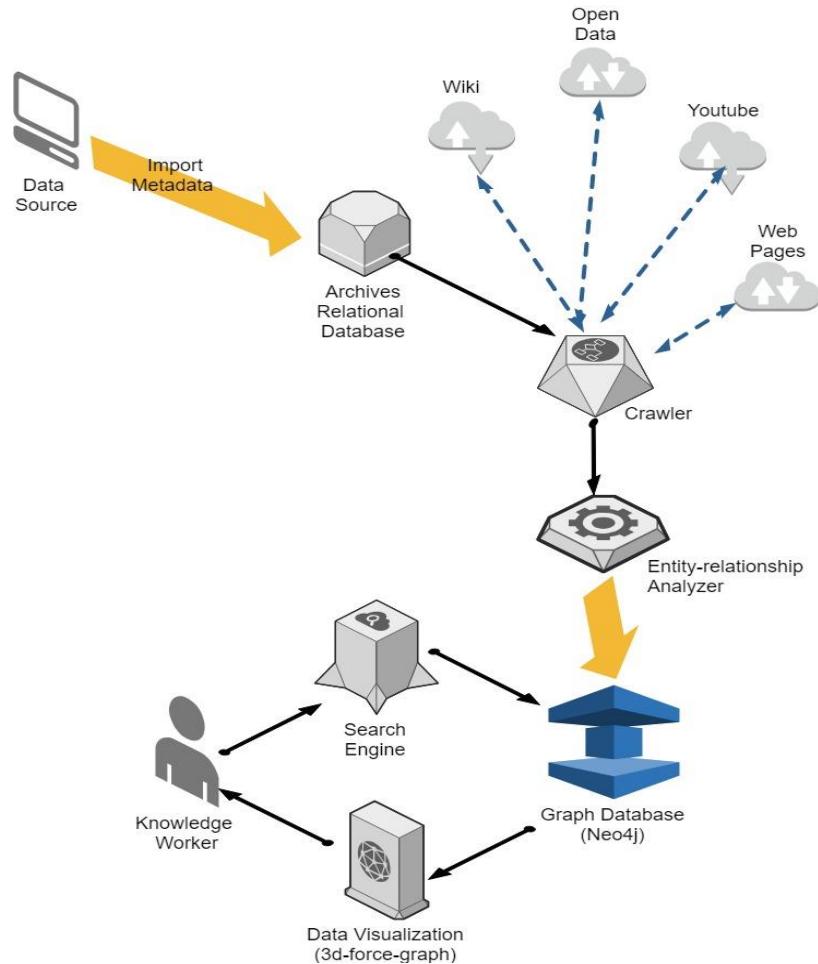
# Storyboard-Like Display

Clicking any leaf node will display the contents of the target web resource:



# Our Approach

- Each ER is associated with a set of metadata, featuring the Person, Events, Places, Time, and Facilities (人事時地物 )
- Metadata are first stored in a relational database, and then used as keywords to search internal and external resources using crawlers.
- Search results are then analyzed using Entity-Relationship Analyzer, or called Inference Engine.
- Analyzed results, nodes and links, are then stored in Graph database Neo4j (Open Source).
- When a user inputs a keyword, the search engine searches the graph database, and presents the results using data visualization tool, 3d-force-graph (Open Source)



# Our Approach

- The entity-relationship's are actually nodes and arcs. Each node represents a feature, or a web resource, and each arc represents a semantic rule, an association rule, a classification rule or a chronological order between two nodes. Therefore, they are best presented as networks.
- In implementing the Entity-Relationship Analyzer, we use “semantic web” or “semantic network”, machine learning, natural language processing (NLP), and data mining techniques, however, we are still working on turning the analyzer.
- Our approach is, by no means, a substitute of Google-Like search approach, but instead, we provide an alternative manner to display search results as storyboard, which, we think, is more appropriate for electronic records containing historical, cultural contents.

# Future Research

- This research is still in the infancy stage, from the development of the prototype system, we learned the following lessons and future research directions:
  - How to filter out irrelevant information, skip web pages containing repetitive, false, doubtful, unnecessary information?
  - How to enhance the functionality of the Entity-Relationship Analyzer, e.g., exploiting more advanced semantic network or data mining techniques?
  - How to enhance disambiguation?
  - How to determine the optimal number of degrees of a network, i.e., the size of the network?
  - How to deal with ad-hoc web pages, i.e., the page contents are dynamically generated on demand, therefore, crawler cannot grab their dynamic contents.

# Thank You!

- Comments?
- Suggestion?