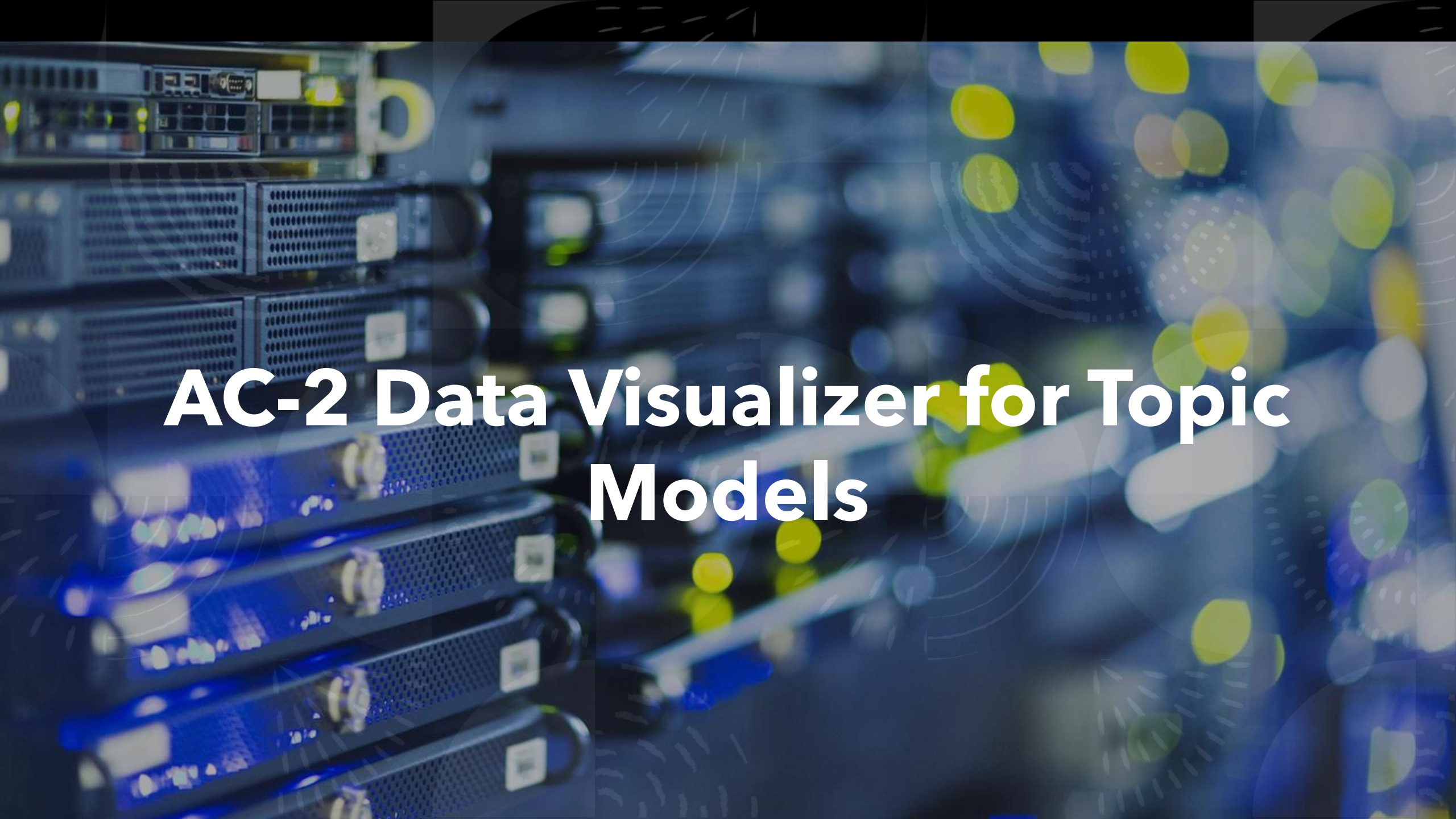
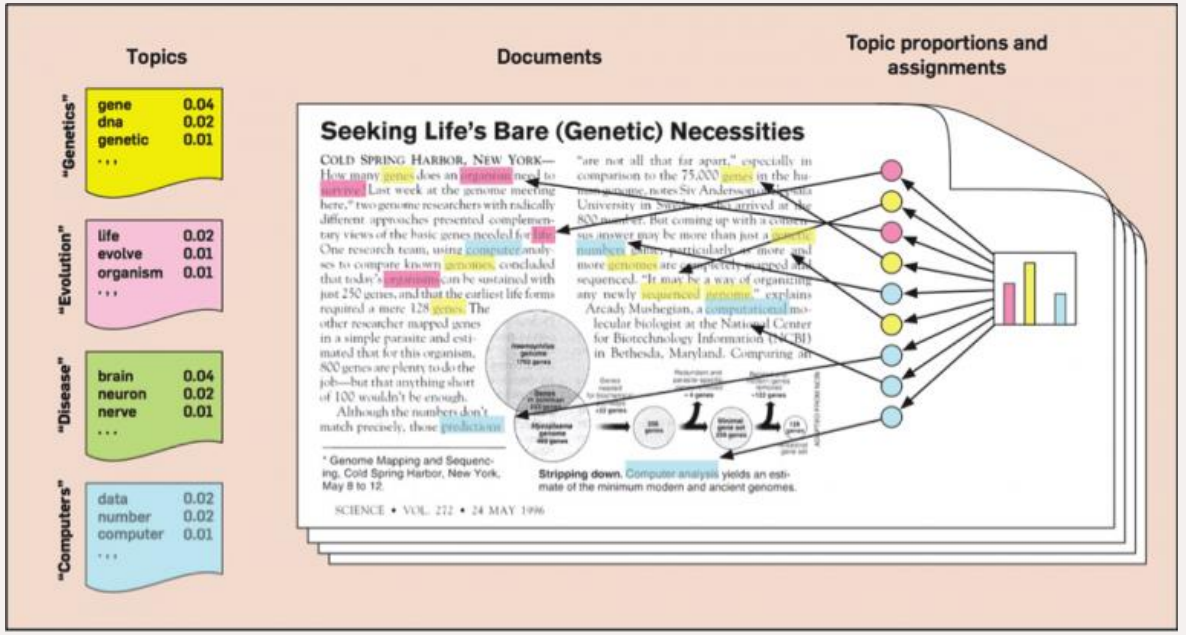


AC-2 Data Visualizer for Topic Models

What is Topic Modeling?

- Topic modeling is a statistical technique that helps to identify the hidden topics in a collection of documents.
- It is an unsupervised machine learning technique that analyzes text data and discovers the underlying themes that are present within it.
- The algorithm works by analyzing word frequency patterns and grouping words that appear together frequently, forming topics.



How does Topic Modeling work?

- Topic modeling is based on the assumption that documents are created from a mixture of topics.
- The algorithm tries to discover the underlying topics by identifying the co-occurring words in the documents.
- The algorithm scans through the entire document corpus, looking for groups of words that are frequently used together. These groups of words form the basis for the topics.

What is our Topic Model doing?

- Our topic model is analyzing a vast amount of text data in the Wikipedia dataset to identify similar articles based on their content.
- By using topic modeling, we are trying to discover the underlying themes and patterns present in the dataset.
- These themes and patterns can help to identify similar articles and group them together, making it easier for users to find the information they need.

How does our Topic Model Work?

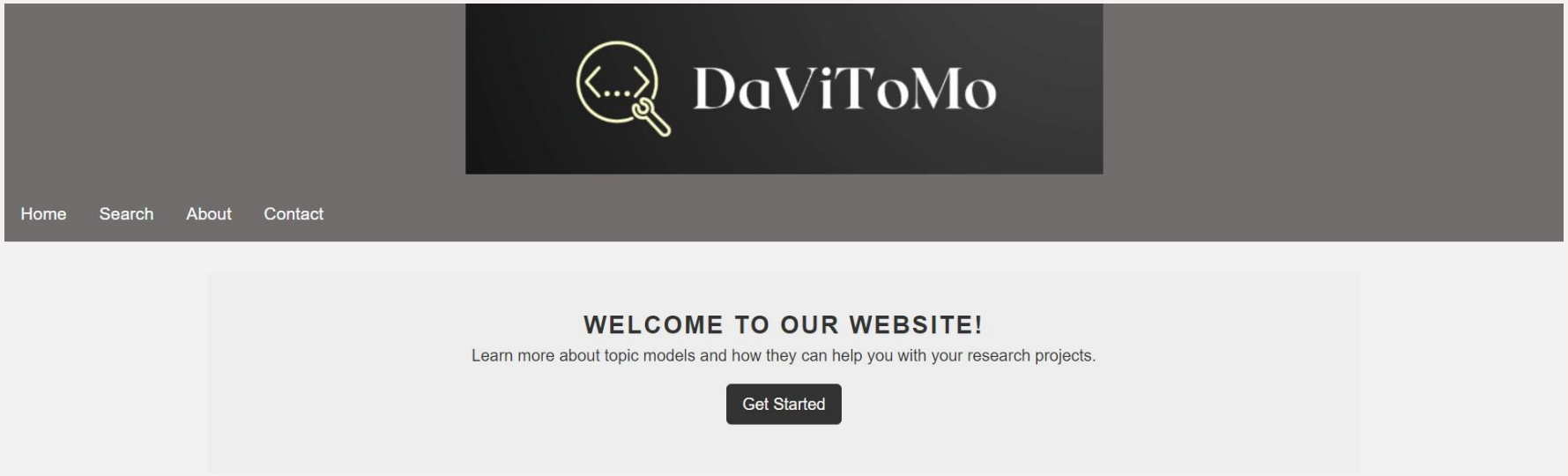
- Our topic model algorithm scans through the entire Wikipedia dataset, looking for groups of words that frequently appear together, forming topics.
- It scores each article based on how strongly it relates to each topic, creating a topic distribution for each article.
- This distribution is used to group similar articles together and recommend them to the users based on the content present in the article.
- For example, if a user is reading an article about Artificial Intelligence, our model will recommend similar articles related to AI, such as Machine Learning, Computer Vision, and Natural Language Processing.

Benefits of Topic Modeling

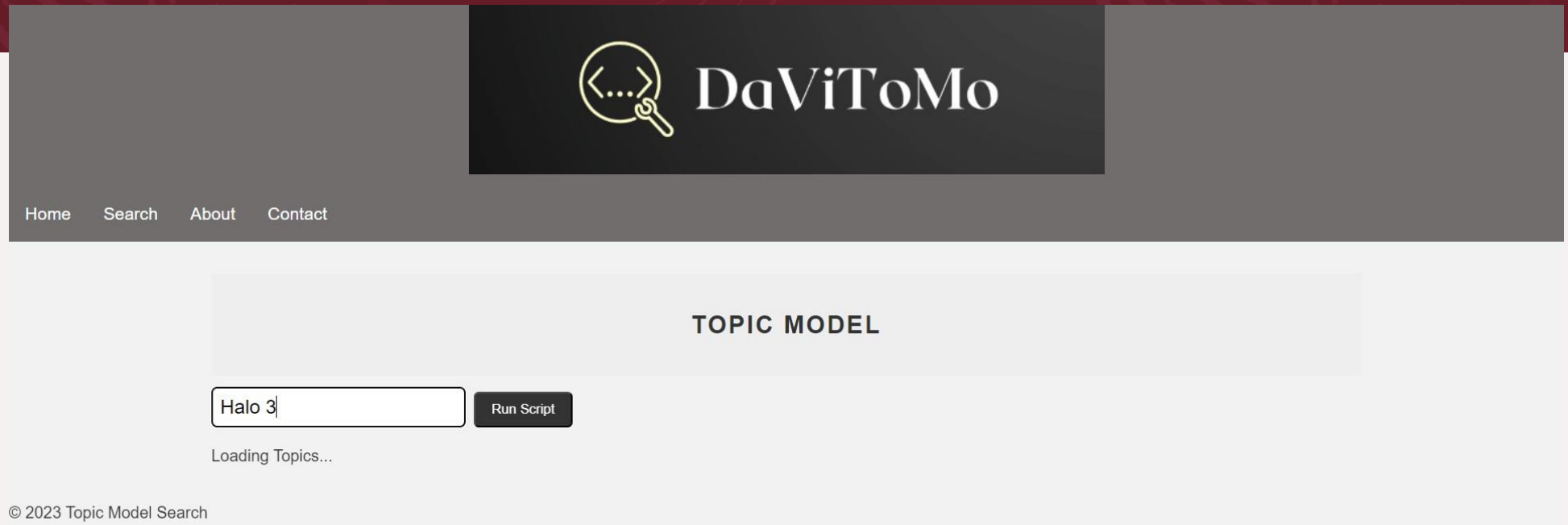
Topic modeling has several benefits. It can help to:

- Organize large volumes of unstructured data
- Identify hidden themes and patterns within the data
- Gain insights into complex issues
- Improve decision-making
- Sentiment analysis

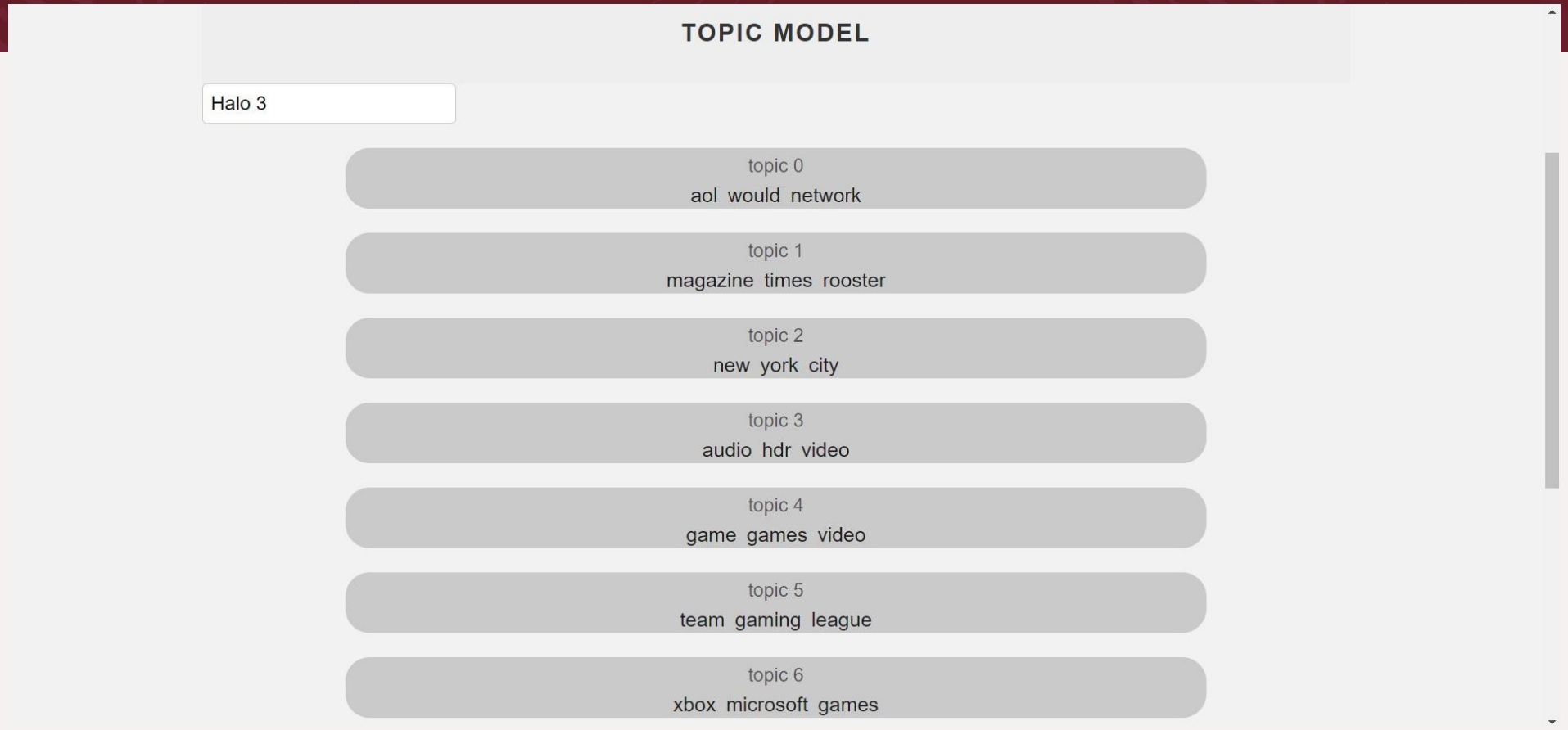
Demo Screenshots



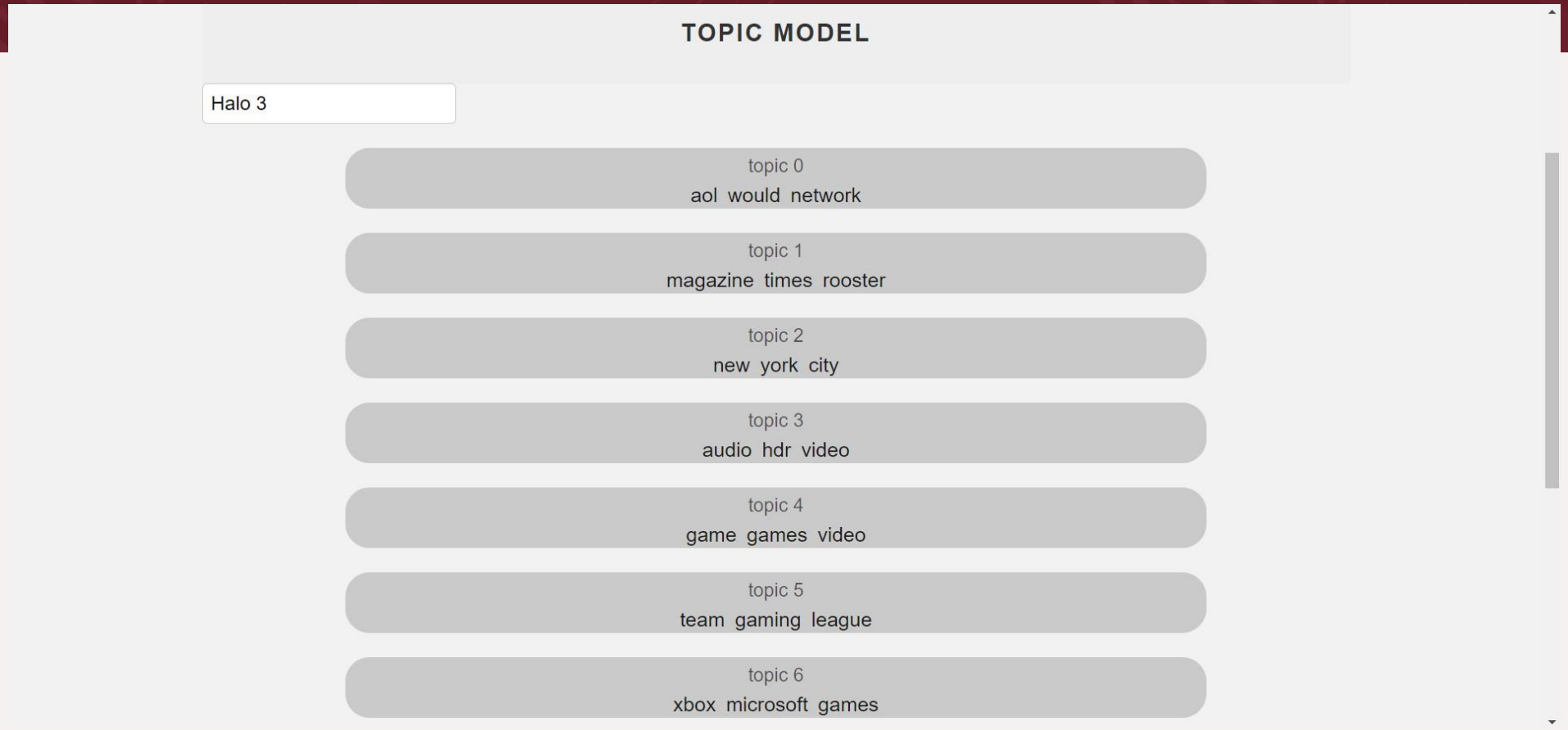
Demo Screenshots



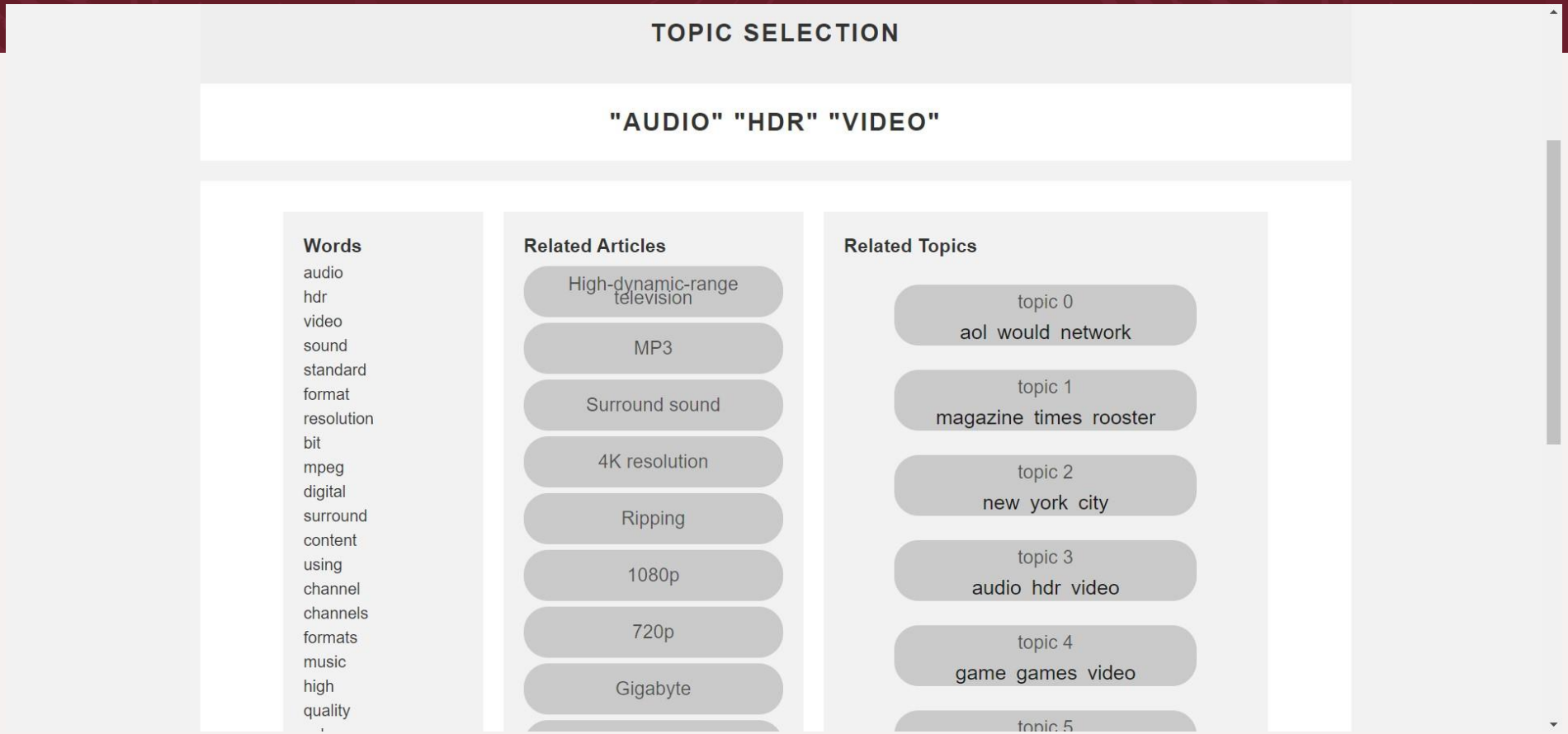
Demo Screenshots



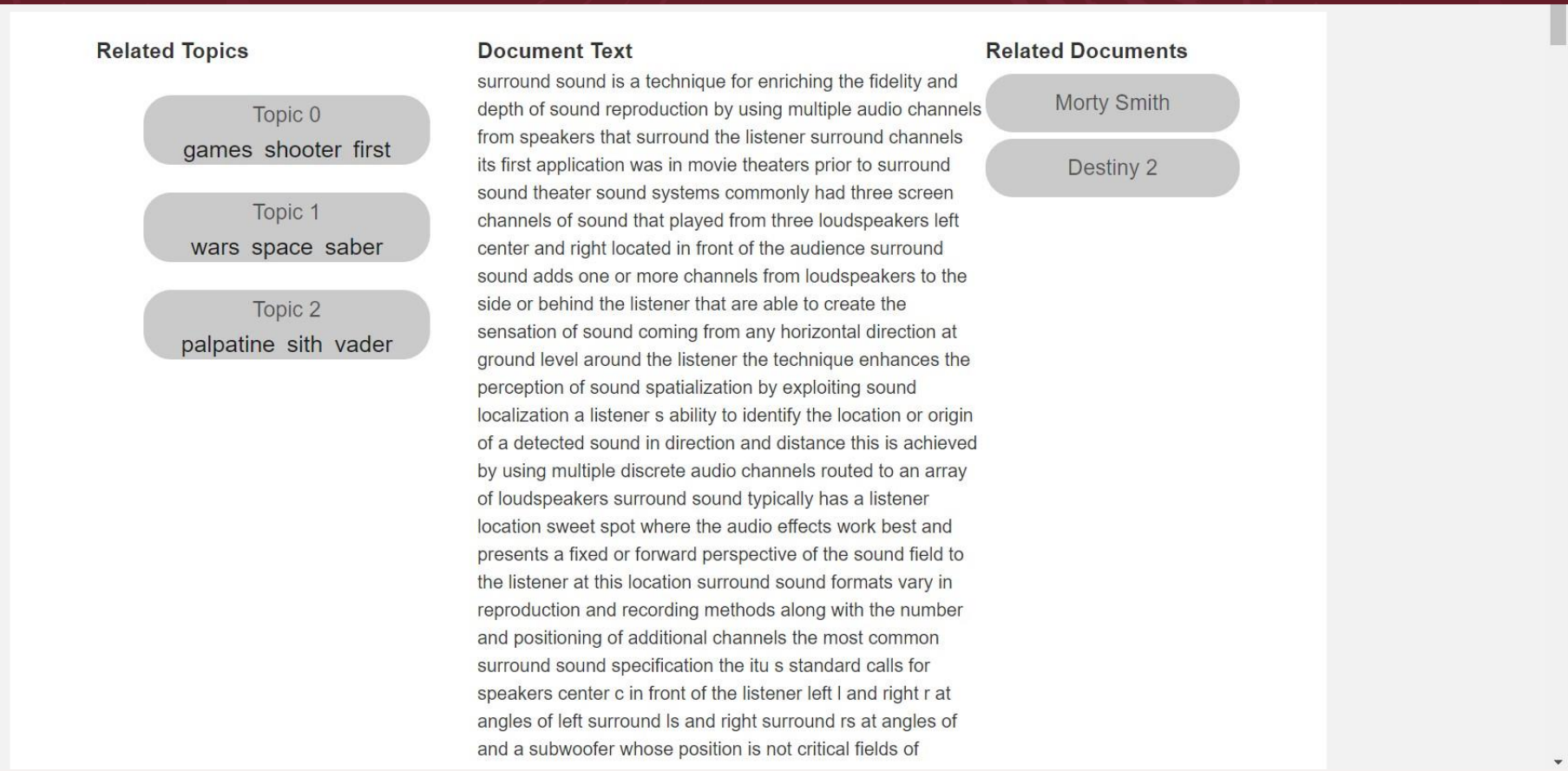
Demo Screenshots



Demo Screenshots

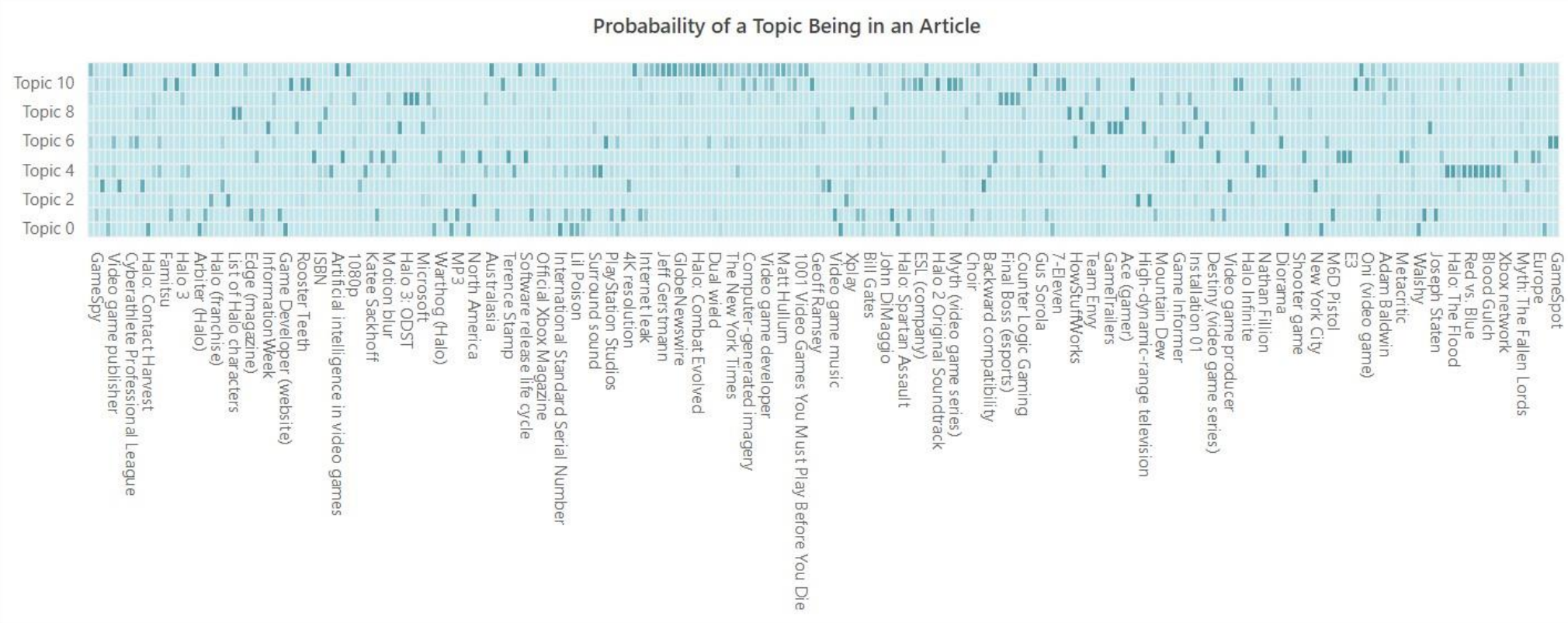


Demo Screenshots

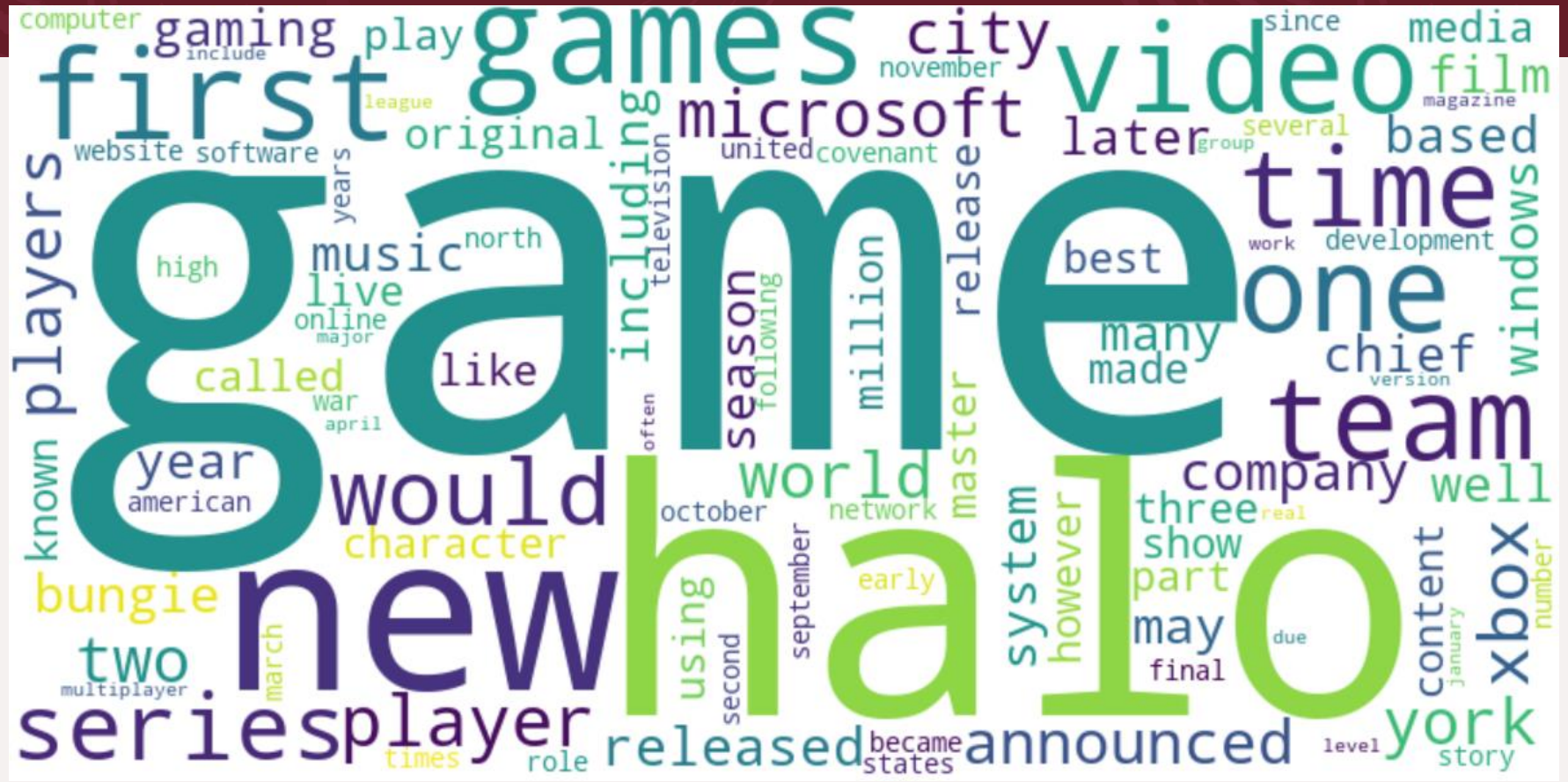


Demo Screenshots

HEATMAP



Demo Screenshots



References

- Wikipedia Contributors. “Topic Model.” *Wikipedia*, Wikimedia Foundation, 27 May 2019, en.wikipedia.org/wiki/Topic_model.
- Chaney, Allison. “Visualizing Topic Models - Cs.columbia.edu.” Visualizing Topic Midels, https://www.cs.columbia.edu/~blei/papers/ChaneyBlei2012.pdf.
- Sharma, Deepak, et al. “Intelligent Systems and Applications.” *Intelligent Systems and Applications*, vol. 2, 2019, pp. 70–82, www.mecs-press.org/ijisa/ijisa-v11-n2/IJISA-V11-N2-8.pdf, https://doi.org/10.5815/ijisa.2019.02.08. Accessed 20 Oct. 2019.