



Our People...Your Edge

EDGEWATER

FEDERAL SOLUTIONS

AI Modeled Concept Tagging: Enhancing Breadth and Relevance of Search Results by Adapting Open-Source AI Tools

In the ever-evolving landscape of data management, the ability to quickly find and relate information is paramount. Edgewater's latest innovation stands at the forefront of this revolution. By harnessing the power of open-source AI tools, this system not only enhances the breadth of search results but also ensures their relevance, providing a seamless research experience that transcends traditional keyword searches.

Introduction to Concept Tagging

Concept tagging leverages AI models to analyze text and generate concept labels pertinent to the content. This method facilitates the grouping of documents through related or similar tags, enabling users to uncover relationships that simple keyword searches might miss. By tagging papers, authors, and institutions with related concepts, users can discover connections that enhance their research experience.

User Engagement through Concept Tags

The journey from a mere concept to a fully operational tagging system is a testament to the ingenuity of Edgewater's Data Engineering Team. The team's dedication to improving user engagement through concept tags has led to a more intuitive and enriching user experience. As users navigate through related records, they are encouraged to delve deeper into their research, uncovering layers of information that were once hidden. This enhancement not only enriches the user experience, but also increases the utility of research tools by keeping users engaged for longer periods.



Our People...Your Edge

EDGEWATER

FEDERAL SOLUTIONS

Implementation by our Data Engineering Team

The technical expertise of the Data Engineering Team shines through in their implementation of the concept tagging system. The team identified and implemented tools to introduce concept tagging for a collection of federal research documents. Adapting tools to fit the unique needs of these research documents, they have crafted a method that not only serves the immediate needs, but also creates an operational method for adding concept tags to records within the collection.

Project Impetus

At the heart of this project lies a simple yet profound goal: to augment the researcher's ability to discover relevant papers and unseen connections. The Data Engineering Team was tasked with improving researchers' ability to find relevant papers in this collection and uncover relationships that are not visible through keyword searches. The goal was to enhance 'findability,' enabling deeper, more efficient research and better discovery of new and relevant ideas.

Identifying Relationships for Capture

Our Data Engineering Team explored methods to assist researchers in finding items related to their keyword searches. An improved search system would address questions such as:



What works are similar to this work?

Which authors have conducted other work related to this topic?

Which authors have collaborated or worked on related topics?

Which institutions are involved in these subject areas?

Tools and Packages for Article Review and Concept Tag Extraction

The project identified an economically and operationally sustainable solution using an open-source AI model developed with Python and TensorFlow, provided by the OpenAlex-concept-tagging GitHub project. A PostgreSQL database running on a Linux server was chosen to store the article and tag references generated by the AI model.

Application Development and Implementation

Upon reviewing the package, it was determined that the original OpenAlex project contained unnecessary code and libraries for this implementation. The project was streamlined to its core components, and then extended specifically using Python and TensorFlow. The system now runs daily, assigning concept labels to new documents and storing the results in a PostgreSQL database hosted on a Linux server.

The system is currently utilized internally and under review for public deployment.



Our People...Your Edge

EDGEWATER

FEDERAL SOLUTIONS

Use Case Example

Scenario

A researcher searches the collection for "Radio frequency" and retrieves the following record. On the record page, the AI-generated concept tag "RF power amplifier" appears, along with other potential concepts of interest, linking to records in the "RF power amplifier" category or others, thereby expanding the user's access to related research topics.

Input: Title and Abstract

An ongoing process detects new documents and feeds their metadata into the AI model to generate scientific topics or "concept tags."



Final Technical Report - Back-gate Field Emission-based Cathode RF Electron Gun

Title

The objective was to complete the design of an electron gun utilizing a radio frequency (RF) power source to apply voltage to a field emission (FE) cathode, a so-called cold cathode, to produce an electron beam. The RF electron gun concept was originally conceived at Argonne National Laboratory but never realized. The research completed the design based on integrating the FE electron source. Compared to other electron guns, the RF gun is very compact, less than one-third the size of comparable guns, and produces a high-energy (several MeV), high-quality, high-power electron beam with a long focal length and high repetition rates. The resultant electron gun may be used in welding, materials processing, analytical equipment, and waste treatment.

Abstract

AI Concept Tagging

A user searching for "radio frequency" research might select this report from their results list. The pre-determined concept tags would be displayed and help the user to related documents.

- Cold cathode
- Electrical engineering
- Electron
- Electron gun
- Engineering
- Field electron emission
- Field emission gun
- Materials science
- Nuclear engineering
- Nuclear physics
- Optoelectronics
- Physics
- Power (physics)
- Quantum mechanics
- Radio frequency
- RF power amplifier

Concept Tags Generated using AI

This example illustrates how AI-generated concept tagging can significantly enhance the user experience by broadening the scope of accessible research topics and facilitating deeper exploration of related content.

To learn more about AI-generated concept tagging, contact our team via www.EdgewaterIT.com.