

Similar States, Similar Solutions? LLM-based Identification of Similar Capital Funding Policies in K-12 Schools

Isabella Antonuccio-Amato

Department of Data Science and Business Analytics
Florida Polytechnic University
Lakeland, FL, USA
iantonuccioamato@floridapoly.edu

Abdulaziz Alhamadani

Assistant Professor, Department of Data Science and Business Analytics
Florida Polytechnic University
Lakeland, FL, USA
aalhamadani@floridapoly.edu

James Dewey

Associate Professor, Department of Data Science and Business Analytics
Florida Polytechnic University
Lakeland, FL, USA
jdewey@floridapoly.edu

Abstract

Capital funding is a critical component of school operations, yet there is limited publicly available information on how states define and implement their capital funding policies. Moreover, no studies systematically categorize states based on these policies to enable comparative analysis. This study addresses this gap by leveraging large language models (LLMs) to extract and synthesize state-level capital funding policies from websites, reports, and educational publications. The result is the National Capital Funding Mechanisms and Expenditures (NCFME) dataset—a novel, unified dataset that captures detailed funding mechanisms, allocation criteria, and legal statutes. Using the NCFME dataset, we propose a novel approach to categorize states with similar capital funding policies. Unlike traditional clustering techniques that rely on numerical data, our method employs LLMs to assess funding similarity based on shared qualitative and quantitative characteristics. This method was validated using expert-labelled operational funding data to both assess the LLM’s accuracy and guide prompt development. This approach enables a more nuanced understanding of policy alignment across states. The goal of this work is to conduct a comprehensive analysis of state-level capital funding policies, using LLMs to extract and categorize funding mechanisms. The analysis will focus on identifying states with similar funding policies and assessing what state characteristics impact funding structure. Key aspects include LLM-based text extraction and prompt engineering, identifying patterns in resource allocation, and predictive modeling of policy classifications. This work advances the understanding of capital funding mechanisms and introduces a methodology for analyzing complex policy documents, providing a foundation for future research.

Keywords

Schools, Facilities, Funding, Policy, LLM

Biographies

Isabella Antonuccio-Amato is a master's in Data Science student at Florida Polytechnic University. She received her bachelor's in Mathematics at Florida Southern College. Her research interests include applications of data science in criminology, public education, and public health.

Abdulaziz Alhamadani is an Assistant Professor of Data Science and Business Analytics at Florida Polytechnic University. His major research interests span machine learning, data mining, and deep learning, and minoring in machine learning applications, text classification, event detection, pandemic forecasting, and ethical AI.

Jim Dewey is a professor in the Department of Data Science and Business Analytics and Director of Economic Analysis at Florida Polytechnic University. He has a Ph.D. in Economics from the University of Florida. His master's degree is in Economics from the University of South Florida, where he also received his bachelor's degree with a double major in Economics and Political Science. His research interests are in the areas of applied microeconomics, policy analysis, education finance, labor economics, and urban and regional economics. He is a Research and Practice Fellow of the National Education Finance Academy, serves on the editorial advisory board of the Journal of Education Finance and Law, and is a member of Regional Science Association International and the Industrial Engineering and Operations Management Society International.