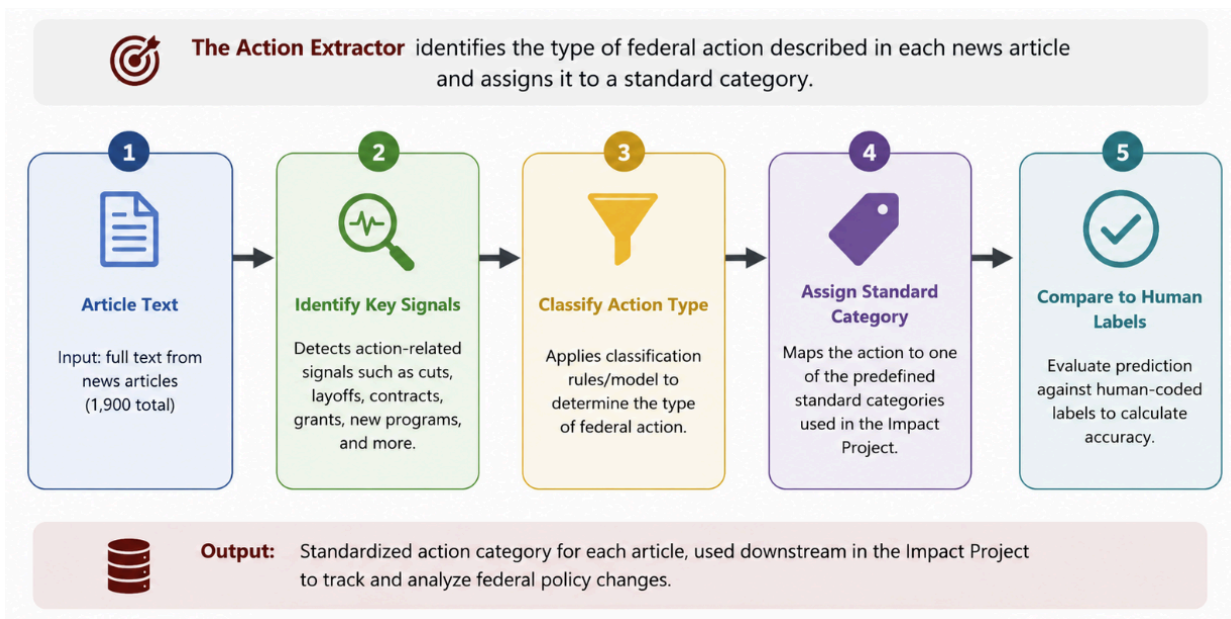


The Impact Project is a non-profit organization that tracks how federal policy, funding, and workforce changes affect communities across the United States. Its Federal Change Map is an interactive tool that helps journalists, policymakers, and community leaders connect government decisions to local consequences. Currently, key information from news articles is extracted manually, making the process time-intensive and prone to inconsistencies and errors.

Our team developed machine learning models to automate this workflow using approximately 1,900 scraped news articles. The models identified six key fields: federal action type, responsible authority and agency, date and location of the action, and funding amount (Fig. 1).



*Fig. 1* Example pipeline for extracting “Action” field from scraped article text

Compared against human-coded labels, the models achieved an average accuracy of 57%. While this demonstrates progress toward automation, the models are not yet sufficiently reliable for independent usage and may benefit from human review. We identified noisy scraped text, ambiguous policy language, and incomplete human-coded labels as the primary sources of misclassification. To improve performance, we recommend additional text-cleaning procedures and providing confidence scores alongside extracted outputs to support human validation. The modeling framework and extraction techniques developed in our project provide a foundation for future work, enabling more timely tracking and reporting of federal policy impacts.