Tracking Medical Marijuana Policies over Time Using Tf-idf

Arka Gupta
JURIST Digital Scholars Program

Medical marijuana transitioned from being heavily criminalized to legalized in some form in 47 states. The movement for legalization has strengthened in the past two decades, especially as the opioid crisis has sparked a growing need for alternative pain therapies. This project aims to leverage natural language processing techniques to identify policy developments in medical marijuana and the legal questions associated with it.

Methodology

This paper analyzes the legal evolution of medical marijuana policy in the United States from 2004 to 2020 using JURIST's coverage, totaling 145 articles. These articles were selected by parsing through JURIST's news articles and selecting only those with "medical marijuana" in the title. Then, articles with ▨ tags such as "international", "Canada" and "Mexico" were removed, so only articles regarding medical marijuana within the United States remained.

These articles were imported into a pandas data frame and pre-processed by removing hyperlinks, punctuation, and stop words (words such as "a", "an", and "the" that add little semantic meaning to a sentence). Then, each article is tokenized - a technique that splits text into smaller subunits known as tokens. Each word constitutes a token, which is then "stemmed" - the process of reducing a word to its stem. For example, a sentence ("the dog runs") that is tokenized could be broken into three tokens ("the" "dog" "runs"). By removing stop words and stemming, we would be left with "dog run." After stemming, the tokens are recombined into a text format.

These articles were separated across four time periods: 2004-2007 (13 articles), 2008-2011 (18 articles), 2012-2015 (66 articles), and 2016-2020 (48 articles). In order to tease out distinct policy developments in the medical marijuana policy, an n-gram term frequency-inverse document frequency (tf-idf) was employed. A tf-idf is a common natural language processing tool that uses statistical weighting to determine a word's importance relative to the corpus. A word's importance is increased by being used frequently but discounted by the number of documents in which it is used (a.k.a. its document frequency). This technique is generally used in order to identify terms that occur frequently, but not include "filler" words such as "a", "the", or "and".

In addition to tf-idf, this analysis also included n-grams, a technique that divides text into all combinations of consecutive "n" words. This is meant to identify not only important single terms, but also phrases. The top 50 words/phrases were generated for each time period and distilled into a graph.

First Period: 2004-2007
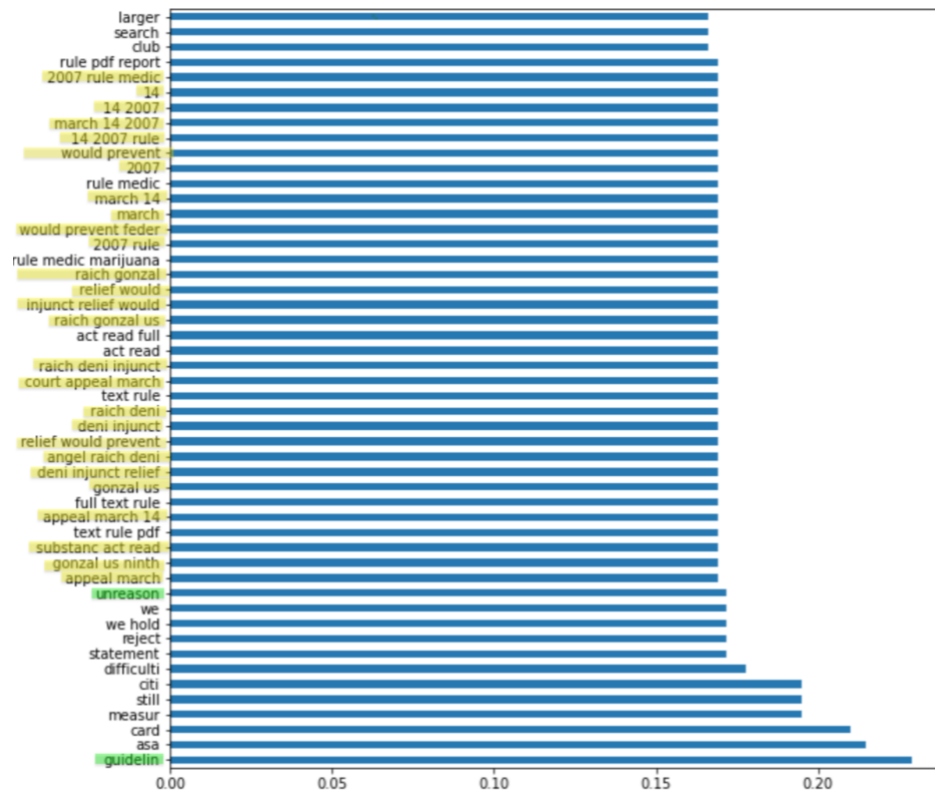
One defining policy development during this time period was the Supreme Court decision Gonzales v. Raich ("gonzal us ninth", "gonzal us", "angel raich deni", "raich gonzal us", "raich gonzal", "substance act read", "denied injunct relief", "deni injunct", 'injunct relief', "would prevent", "march 14 2007", "appeal march 14"). In 1996, California legalized medical marijuana despite a federal prohibition under the Controlled Substances Act. Raich, a medical marijuana patient, relied on medical marijuana to alleviate her pain when alternative medications had failed. Raich and two other anonymous medical marijuana users filed for an injunctive relief against the federal government to not criminalize their use of medical marijuana. They argued that the Controlled Substances Act was an unconstitutional overreach of Congress' commerce clause power, while the government argued that medical marijuana use anywhere could affect interstate commerce. The federal district court ruled against Raich and her fellow respondent, but the Ninth Circuit Court of Appeals affirmed that the Controlled Substances Act was unconstitutional. The United States Supreme Court heard the case, and ruled in 2005 that the Controlled Substances Act is constitutional, denying injunctive relief for Raich but remanded the case to the federal district court on the doctrine of medical necessity, referring to medical treatments that are deemed necessary for care. After the district court ruled against her, Raich later appealed to the Ninth Circuit Court of Appeals, which again ruled against her in March of 2007.

This time period was also defined by the First Circuit Court's decision on Ridley et al v. Massachusetts Bay Transportation Authority (MBTA) (<mark>"guidelin", "unreason"</mark>). The First Circuit ruled that the MBTA violated First Amendment rights of Change the Climate, an advocacy group that placed an ad challenging state prohibition against marijuana at the time. The MBTA argued that these ads were rejected out of concern of promoting illicit marijuana use; however, the court ruled this denial constituted First Amendment violations due to viewpoint discrimination.

The unhighlighted terms are less distinct since they may refer to basic search terms ("search", "text", etc.) or very broad terms ("statement", "we hold", "larger") that cannot identify a single case.

Second Period: 2008-2011
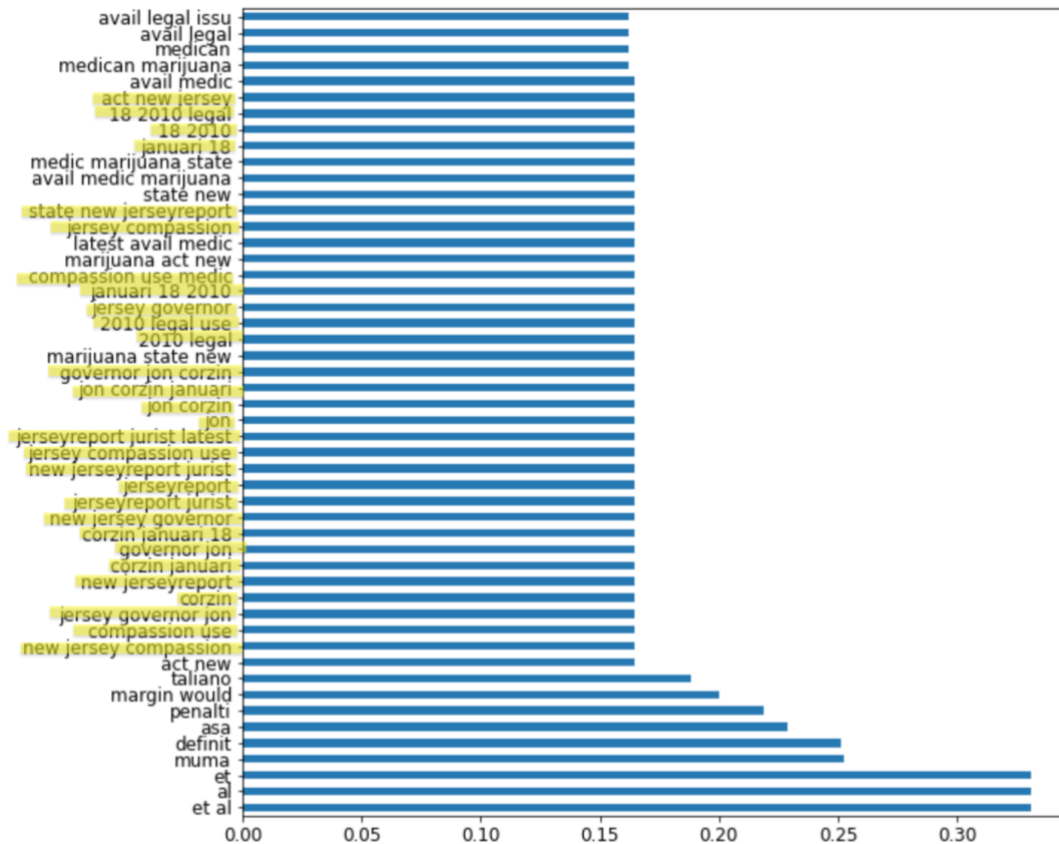
From 2008-2011, there were a total of 18 articles. A prominent policy development occurred in New Jersey (==“act new jersey”, “state new jerseyreport”, “jersey governor”, “governor jon corzin”, “compassion use”, “18 2010 legal”==). Despite the 2005 decision in Raich v. Gonzales that affirmed the constitutionality of the Controlled Substances Act, many states proceeded with legalization of medical marijuana. New Jersey enacted a compassionate medical marijuana bill after approval by the state legislature and Governor Jon Corzine. This bill permitted only those with chronic illnesses to use marijuana obtained from a certified dispensary, notably prohibiting public growing of marijuana or use in public. For unhighlighted terms, they were generally so broad they could not be tied to a distinct policy development.

Third Period: 2012-2015
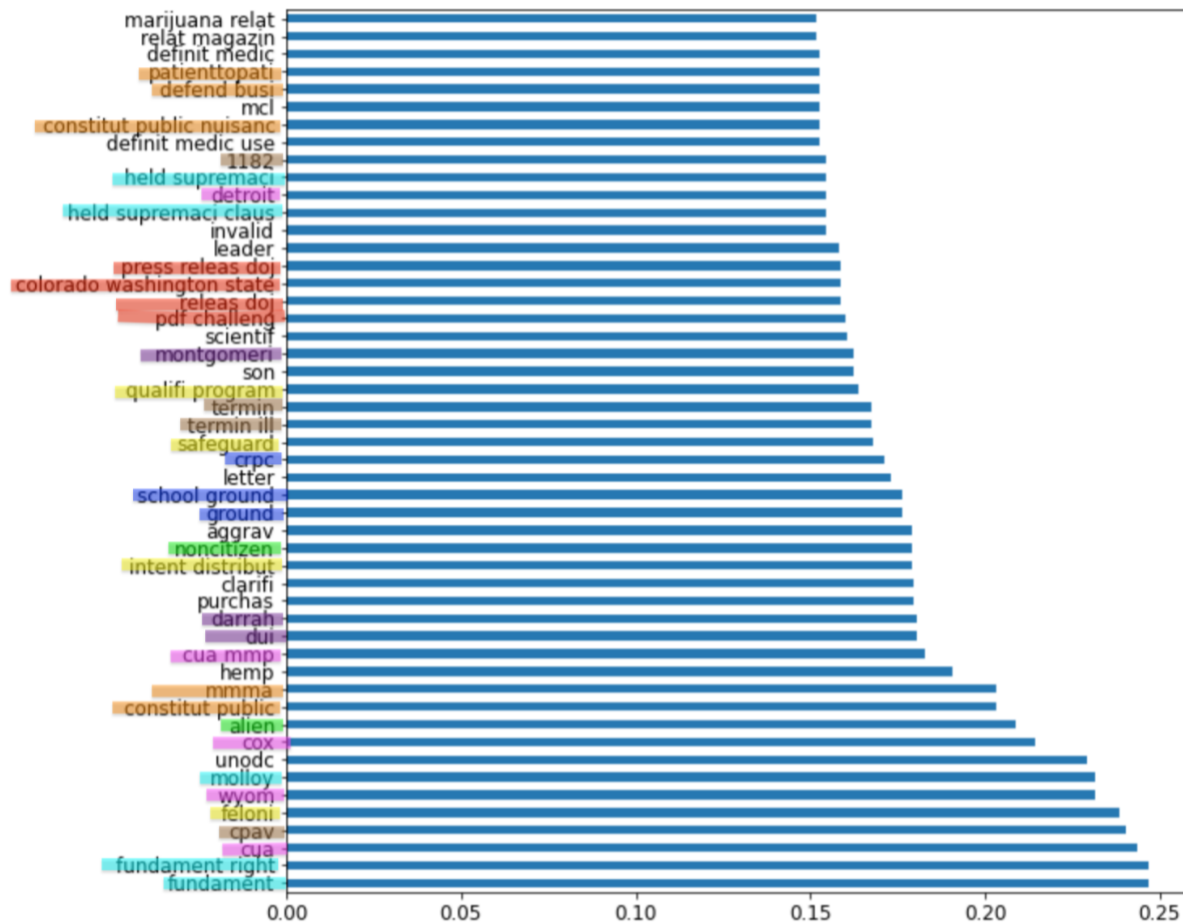
During this time period, a focal point of medical marijuana policy was distinguishing between medical users versus illicit distributors (==“feloni”, “intent distribut”, “safeguard”, “qualifi program”==). State medical marijuana policies from 2012 to 2015 focused on gradual decriminalization. For example, [Connecticut](#) in 2011, the city of [Chicago](#) in 2012, [Vermont](#) and [Maryland](#) in 2013, [Washington D.C.](#) in 2014, and [Delaware](#) in 2015 passed legislation to decriminalize only small amounts of marijuana possession. In November of 2012, US Congressional Representatives [sent](#) a letter to the Department of Justice asking it to respect state laws permitting small usage of medical marijuana. Building on nationwide trends, the United States Supreme Court (==“alien”, “noncitizen”, “feloni”==) even [ruled](#) in a 7-2 decision that state misdemeanors for possession of small amounts of marijuana would not constitute grounds for deportation. The distinguishing factor between misdemeanors and felonies across many states is the amount of marijuana possession, underscoring the importance of amount. A similar question was the distribution of medical marijuana, through shops or individual sales. In Michigan (==“constitut public nuisanc”, “defend busin”, “patienttopati”, “mmma”==), the Supreme Court ruled in 2013 that the private sale of medical marijuana was illegal because it constituted a public

nuisance. The implications of this decision would defend medical marijuana businesses from competition by limiting patient to patient sales, requiring patients to only purchase medical marijuana from licensed dispensaries.

Montana saw multiple developments in medical marijauana policy ("fundament right", "molloy", "held supremaci"). In January of 2012, Judge Molloy of the Montana federal district court ruled that the state's medical marijuana law does not protect against federal enforcement, citing the Supremacy clause and the ruling held in *Gonzales v. Raich*. That same year, a bill was introduced to repeal Montana's existing medical marijuana law. Plaintiffs challenged this law, but the Montana Supreme Court held that marijuana possession is not a fundamental right.

While Montana withdrew from medical marijuana, other states continued to increasingly embrace it. The Michigan ("wyom", "detroit") Supreme Court invalidated a local ordinance that prohibited medical marijuana, directly conflicting with state law, and Georgia ("cox") enacted a new medical marijuana law. Interestingly, the California Supreme Court ("cua", "mmp") ruled in an opposite manner to the Michigan court when presented the same situation. The California high court upheld a local prohibition on medical marijuana despite a conflicting state law permitting medical marijuana, citing scope of the legislation does not preclude prohibitions.

This period saw additional states legalize medical marijuana ("1182", "termin ill, "cpav") including Pennsylvania legalizing for several forms, New York legalizing medical marijuana for terminally-ill medical patients, and Arkansas' Supreme Court permitting a medical marijuana ballot initiative to proceed, rejecting a challenge by the Coalition to Preserve Arkansas Values (CPAV). Some states even expanded access and reduced barriers to medical marijuana ("school ground", "ground", "crpc") In New Jersey, Governor Chris Christie signed a bill permitting medical marijuana for disabled students on school grounds while Colorado permitted attorneys to advise medical marijuana companies without fear of violating state ethics laws.

Meanwhile, states such as Arizona enacted restrictions on medical marijuana ("darrah", "dui", "montgomeri"). In 2013, William Montgomery, a county prosecutor, filed charges against parents seeking medical marijuana as a treatment for their son suffering from epilepsy despite their clearance by the Arizona Department of Health. In 2014, an Arizona appeals court ruled that licensed holder Travis Darrah could not use medical marijuana as a valid defense against a DUI. At the federal level, the Department of Justice was initially hostile towards legalization in 2012, particularly towards Colorado and Washington ("colorado washington state", "doj", "pdf challeng"). However, by 2014 the U.S. Justice Department acquiesced.

Fourth Period: 2016-2020
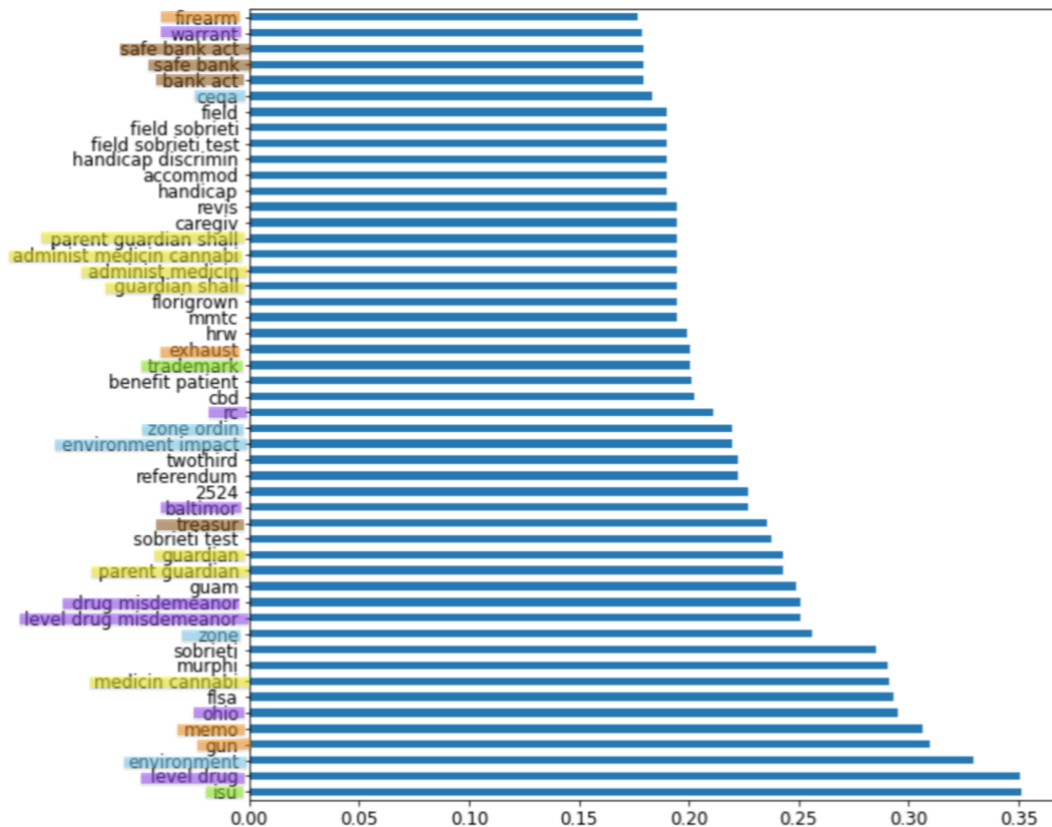
A front for changes in medical marijuana policy during this time period was students. California (<mark>"guardian", "medicin cannabi"</mark>) enacted legislation allowing medical marijuana use on school grounds. Students from Iowa State University (<mark>"isu", "trademark"</mark>) sued the administration for prohibiting the use of its official logo in a fundraiser for a medical marijuana student advocacy group. A federal district court and appeals court ruled in favor of the students, citing right to free speech in public forums, a similar logic to the aforementioned Ridley v. MBTA case.

Another noteworthy case occurred in 2019, when the California Supreme Court ruled that the city of San Diego should have conducted an environmental impact study, under the California Environmental Quality Act (CEQA), for a new dispensary zoning ordinance (<mark>"environment", "zone", "ceqa"</mark>).

Decisions at the federal level during this time period moderated the expansion of medical marijuana (<mark>"exhaust", "gun", "firearm", "memo"</mark>). For example, Attorney General Sessions rescinded an Obama-era memo that limited federal prosecutions against medical marijuana patients. In addition, the federal courts took action with the District Court for the Southern District of New York dismissing a challenge to the federal prohibition on medical marijuana,

citing the plaintiff's inability to exhaust all other remedies before suing. In addition, the Ninth Circuit Court of Appeals upheld a ban on firearm sales to medical marijuana patients.

A noteworthy trend juxtaposing the Attorney General's efforts to increase marijuana prosecutions is local prosecutors withdrawing from such prosecutions ("baltimor", "drug misdemeanor", "warrant", "ohio"). For example, the Manhattan District Attorney dropped over 3,000 medical marijuana charges for low-level misdemeanors, explaining that issuing warrants ruined the lives of New Yorkers, particularly those in communities of color. In a similar vein, Baltimore's state prosecutor also announced that she would not continue prosecuting marijuana cases, citing concerns for public safety and unjust burdens on communities of color. Notably, Ohio attempted but failed to pass legislation that would form a medical marijuana quota for minority owners.

Another avenue of policy development was finance ("bank", "treasur"). Given that medical marijuana remains illegal at the federal level, many banks were hesitant to hold accounts for medical marijuana companies out of fear of federal crackdown. To address this problem, 38 state and territory Attorneys General signed a letter supporting passage of the federal SAFE (Secure And Fair Enforcement) Banking Act, which would loosen banking regulations for medical marijuana companies. In addition, the state of West Virginia enacted legislation that established a separate treasurer's fund for collected fees until it can be reimbursed into the state's general account.

The unhighlighted terms belonged to individual cases at state level, including a Massachusetts ruling that field sobriety tests could not determine marijuana impairment, and Florida medical marijuana legislation enacted in March of 2019, only to be ruled unconstitutional by a state appeals court only a few months later.

Conclusion

This project aimed to leverage tools of natural language processing to better understand the avenues of medical marijauna policy development. Tf-idf provided an objective avenue to identify distinct cases and events. However, there are limitations in that the entities recognized are limited to the corpus used, which will inevitably contain some sort of bias. A larger corpus would reduce the probability or possibly the strength of such a bias occurring. Another challenge faced was distinguishing between words that signal important policy developments particular to a time period, and words with a high frequency due to being commonly used, and therefore should be omitted.

The tf-idf algorithm identified the Gonzales v. Raich Supreme Court case, which established precedent for later decisions denying protection from federal enforcement of the controlled substances act. In addition, the algorithm identified common policy topics such as amount of possession, schools, and financing regulations.

This algorithm will be helpful for future social scientists seeking to identify policy trends across time. Tf-idf is an exploratory tool that allows a researcher to identify distinct developments within a specified field as a starting point for additional research. While a search

engine will search for a key term or phrase in a comprehensive list of sources, a tf-idf analysis will identify the most salient terms within a given corpus of documents. However, a prerequisite for tf-idf is an existing corpus that can be converted into a machine-readable format and then preprocessed. Therefore, a user entirely new to a field may lean on a search engine in order to "scratch the surface" of a field, while a tf-idf can summarize major developments and provide a roadmap for additional analysis.