AMERICAN CRIME STORY
a visual investigation into a decade of US crime

VALUENEX • Published on January 18, 2021 • 6 min read

Electoral campaigns often promise to be tough on crime and strong on justice but with the White House about to pass to a president-elect who helped head two previous administrations, whose focus increasingly became criminal justice reform, it's worth looking at what the Department of Justice (DoJ) priorities were at the time and how they evolved over the intervening years to see whether the current stated policy objectives are merely a continuation of the same goals, or whether there has been a policy shift.

To do this we have chosen to analyze DoJ press releases as being a truer reflection of the focus of the DoJ's work than news media coverage or political speeches. The VALUNEX Radar visualizations of these press releases by keywords, dates, topics, and content clearly show where priorities were concentrated as well as emerging trends, painting an accurate and comprehensive picture of an evolving situation.

The Analysis & Dataset

VALUENEX Radar is a visualization tool that organizes and clusters text data based on semantic similarities quickly exposing areas of interest. These areas represent collections of documents pertaining to content-specific topics. In this case, the size and density of the area indicates the incidence of certain crimes.

MAJOR areas refer to crimes that are heavily occurring within the dataset. GROWING areas tell us the crimes in the United States that are developing rapidly. While SPARSE areas (or whitespace) could be interpreted as "underreported" crimes, for this specific radar the SPARSE area is in the middle of the radar and can be fundamentally connected to all other crimes.

The dataset used for this analysis is from Kaggle and includes 13,000 press releases from the Department of Justice archive dated between 2009 and 2018.

The Output

Looking at the Radar (fig.1) we can easily identify five Major areas, and one Growing area. Hovering over the areas displays content and related keywords, which can give a quick overview of the contents of the press releases within.
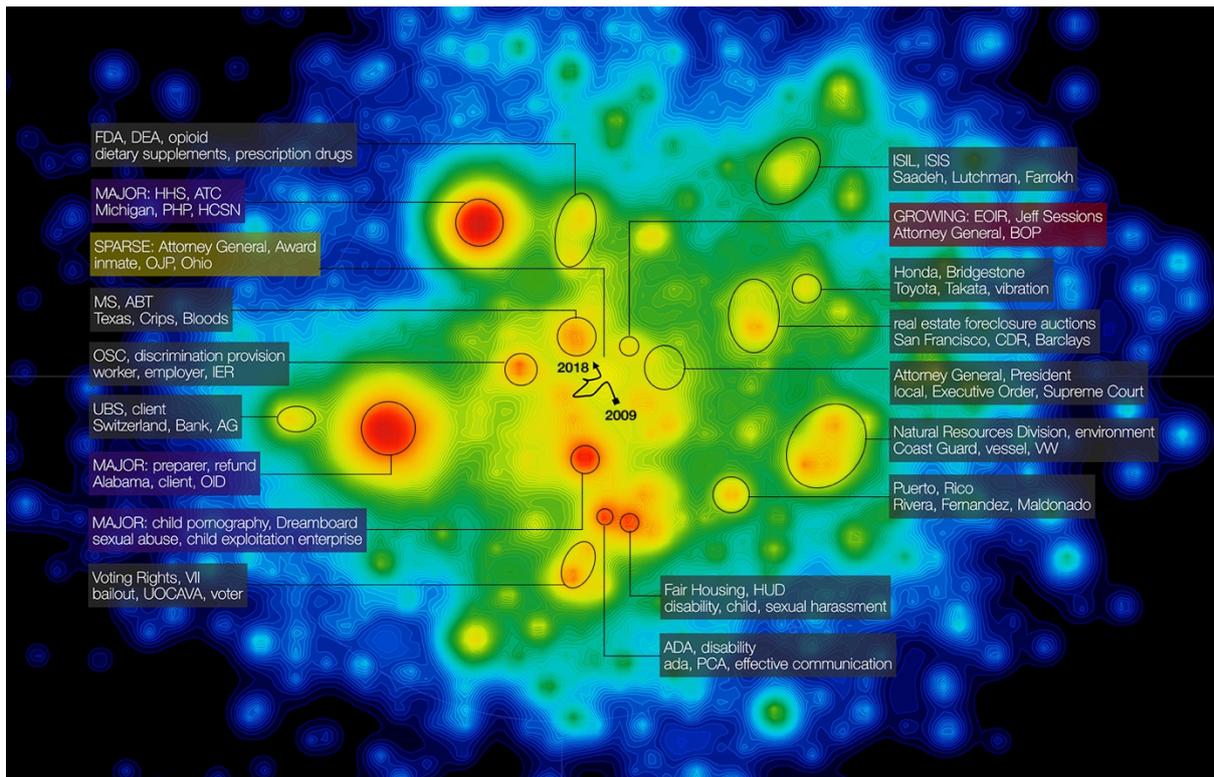
Fig. 1 • US Department of Justice Press Releases (2009 – 2018)

MAJOR AREA: Tax Fraud
Keywords: preparer, refund, Alabama, client, OID

The keywords "OID" and "Alabama" come up frequently in this area. A quick search onto the term "OID" reveals it to be related to bonds, an "Original Issue Discount (OID)" is the discount in price from a bond's face value at the time a bond or other debt instrument is first issued". Investigating the area more closely, most documents are related to tax fraud.

A similar search for "Alabama" reveals documents on identity fraud and tax theft, a statistic on identity fraud in Alabama reveals the number of identity theft complaints in Alabama grew by 32% from 2010 to 2014. From further examining the documents and viewing individual clusters, we find varied crimes related to taxes and identity fraud. The density of this area, and the frequency by which area keywords show up across the radar, tells us that tax and identity crimes are frequent and complex.
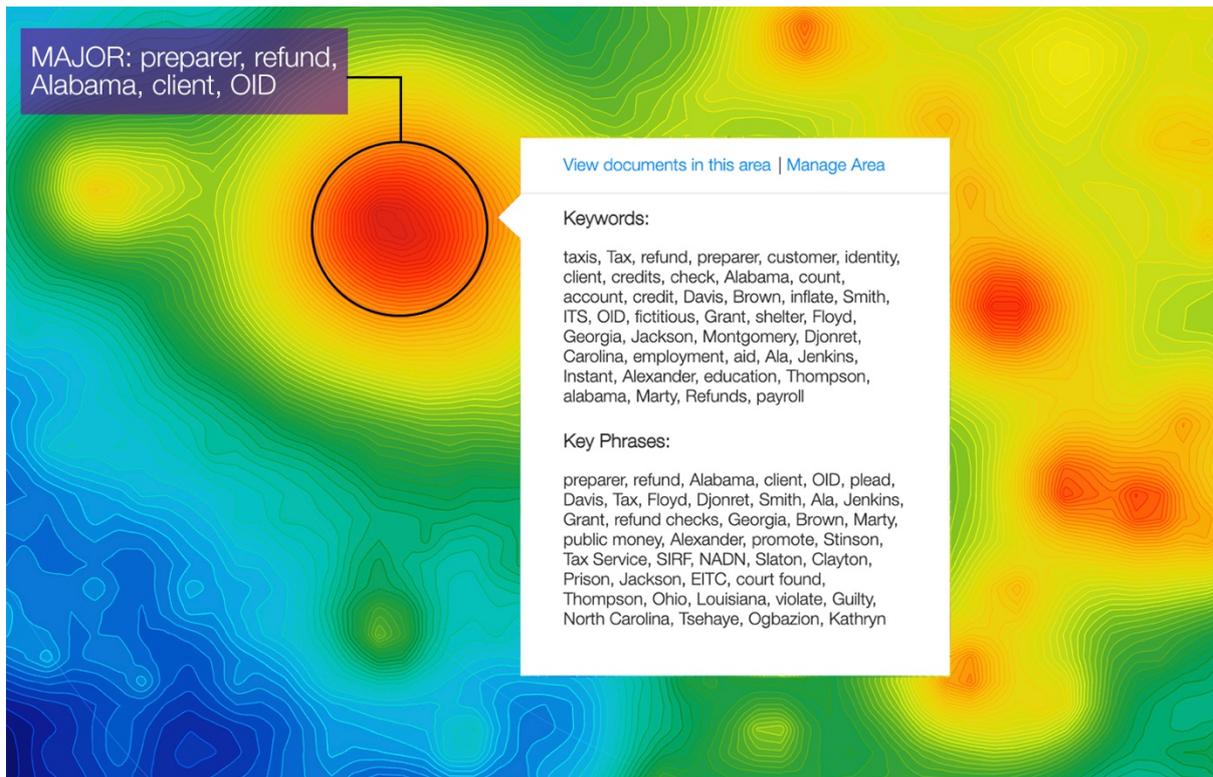
Fig. 2 • Keywords and Key Phrases for Major Area #1

Cluster #339

Keywords: Jenkins, OID, Heber, Utah, proof at trial, HHS, barring, Michigan, refund, preparer, ...

Years: 2010, 2013, 2014

Document sampling (of 7)
- Former Utah Certified Public Accountant Convicted of Filing False Claims for Tax Refunds Totaling More Than $8 Million and Presenting a $300 Million Fictitious Financial Instrument
- Utah Man Previously Charged with Filing False Claims for Tax Refunds Indicted for Additional Charge of Passing a Fictitious Financial Instrument
- Utah Man Charged with Filing False Claims for Tax Refunds
- Federal Court in California Shuts Down Tax Preparer
- Utah Accountant Sentenced for Filing Over $9 Million in False Tax Refund Claims and $300 Million Fictitious Financial Instrument
- Salt Lake Federal Court Bars CPA from Preparing Tax Returns for Others
- Owner of Detroit Adult Day Care Centers Pleads Guilty in Connection with Medicare Psychotherapy Fraud Scheme

Fig.3 • Keywords and Key Phrases for Tax Fraud area

MAJOR AREA: Medical Fraud
Keywords: HHS, ATC, Michigan, PHP, HCSN

The main keywords of this area are HHS (Health and Human Services, a cabinet department), ATC (American Therapeutic Corporation), PHP which is an abbreviation for a partial hospitalization program, and HCSN (Health Care Solutions Network). Keywords in the area are not frequently mentioned, so we generated a sub-radar to analyze it more closely by isolating the documents from all other data points, allowing us to detect keywords and trends at a more granular level. Once this was done, custom areas were outlined.
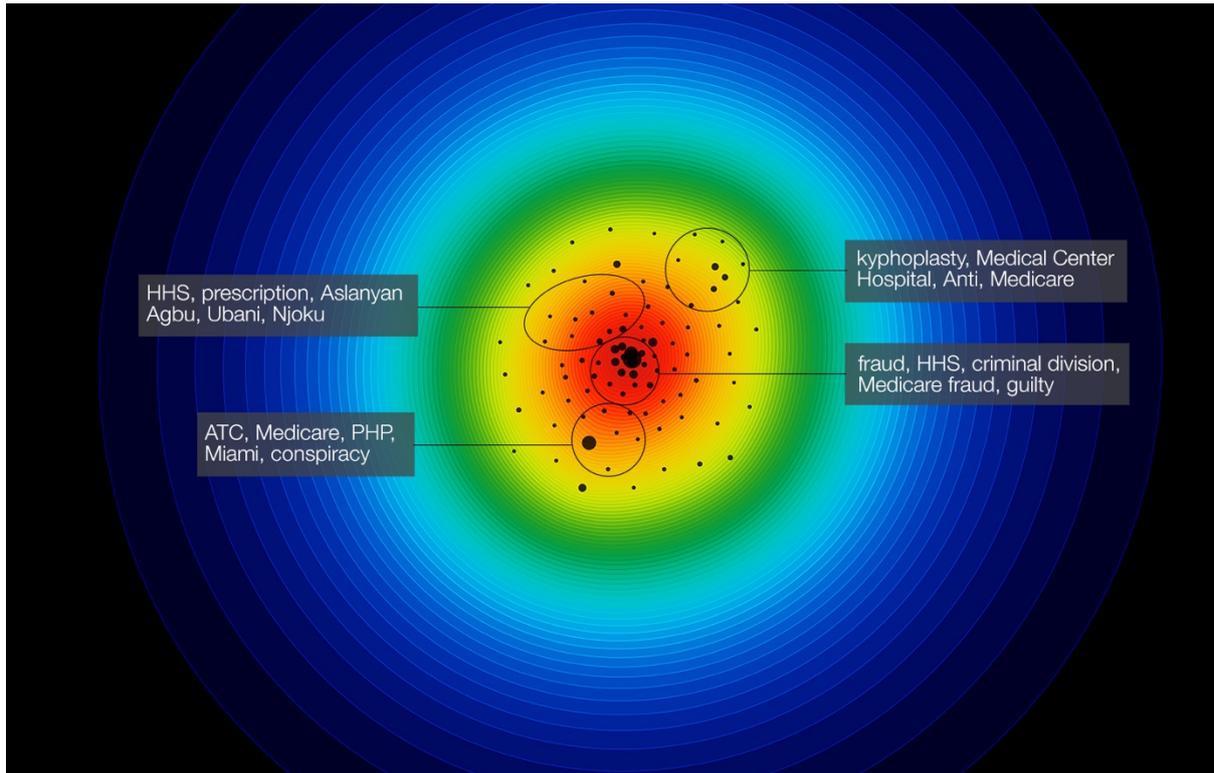

Fig.4 • Sub-radar for the Medical Fraud area

From the sub-radar (Fig. 4), we observe various crimes:
- Areas in the sub-radar with "HHS" as a keyword cover Medicare fraud and cases across the United States. A quick search shows that The DoJ received roughly $2.1 billion in 2017 and $2.5 billion in 2018 from lawsuits involving healthcare fraud and false claims.
- Other crimes discussed in the areas include falsifying diagnoses, other forms of healthcare fraud, and unreasonable or unnecessary rehabilitation.

MAJOR AREA: Civil Rights Abuses

Keywords: child pornography, Dreamboard, sexual abuse, child exploitation enterprise | fair housing, HUD, disability, child, sexual harassment | ADA, disability, PCA, effective communication

Using the center of gravity feature, we can gauge the space on the radar occupied by releases from the Civil Rights Division, the DoJ office on discrimination cases. Noticeably, the bottom half of the radar (Fig. 5) includes the remaining 3 Major areas encompassing offenses related to a plethora of civil rights abuses. The most salient area in this center of gravity includes a major area on child sexual abuse and child pornography.
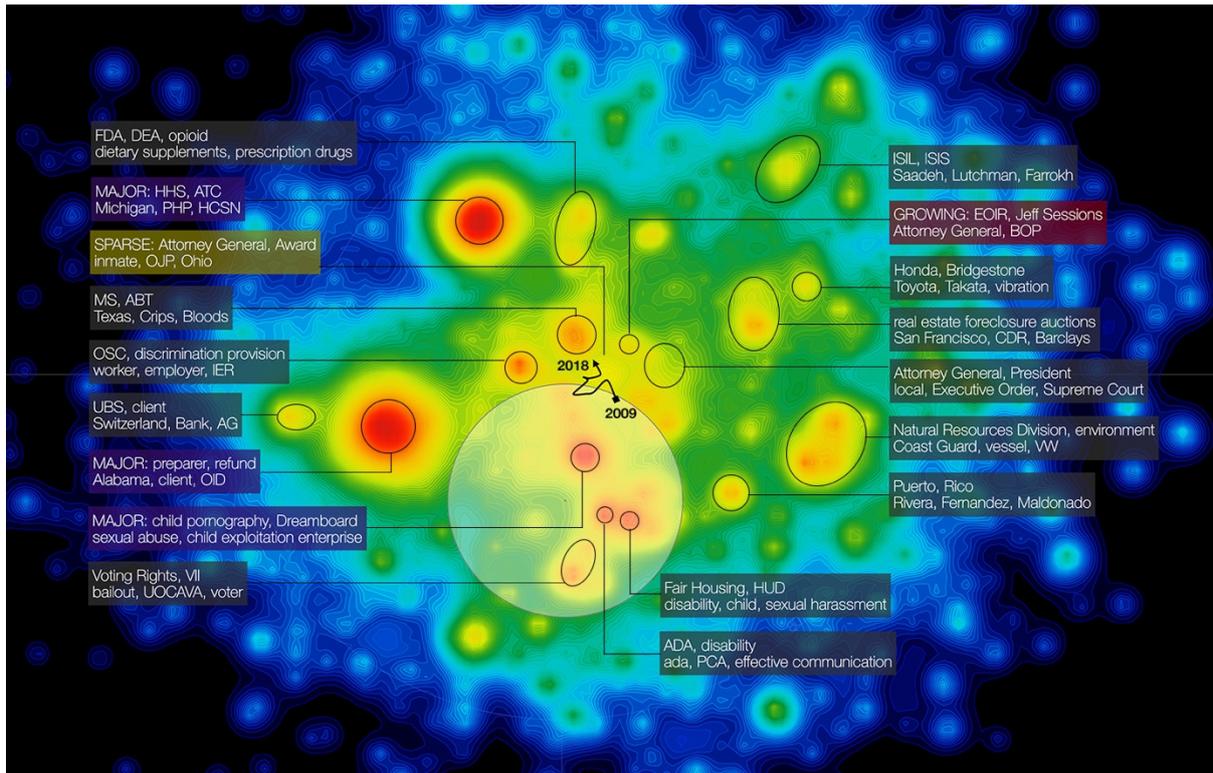

Fig.5 • Center of Gravity representing Civil Rights Abuses

Many releases in this area come from the Department of Justice's Project Safe Childhood, the DoJ's initiative to combat child exploitation and abuse. Various statistics confirm the scope of crimes across this area:
- From 2010 to 2014, there was a 31% increase in the sexual exploitation of minors
- The total lifetime economic burden of child sexual abuse (CSA) in the United States in 2015 was estimated to be at least $9.3 billion, though this statistic is likely an underestimate.

GROWING AREA: Jeff Sessions

The one growing area on the radar covers news related to former Attorney General Jeff Sessions (who was the Attorney General under the Trump administration) and assorted controversies from the administration. To analyze the area, we chose to highlight clusters containing "Attorney General" (in purple) and "President" (in red) (Fig.6).
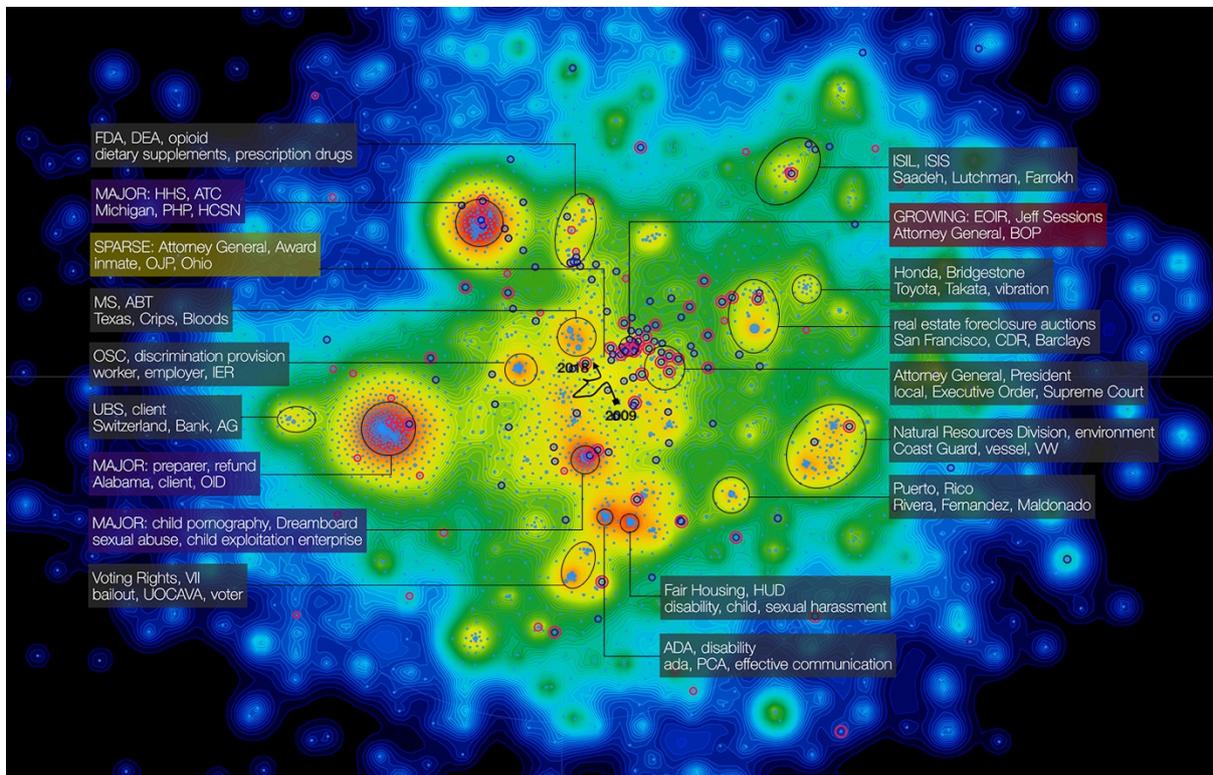

Fig.6 • Quick Highlights feature activated to show documents
containing the words 'President' and 'Attorney General' between 2017 – 2018

A variety of documents related to immigration reform and crimes, Guantanamo Bay, and same-sex marriage are contained in this area. All of these topics were heavily discussed during the Trump presidency.

The Insight

This macroscopic view of how policy focus has shifted can inform decision-making on various aspects of the department's functioning, such as reviews into its policy objectives, assessments of budget allocations for these objectives, reviews into the underlying causes of these shifts (Has the issue been consciously deprioritized or has it been neglected?). The sharp increase in press releases relating to the Attorney General, and the President are the sign of a DoJ on the defensive rather than offensively tackling its brief — did this only affect the press office or was it to the detriment of the wider department and its ability to tackle policy objectives?

The spatial aspect of the visualization makes it possible to identify links between issues that could be leveraged for efficiency and effect. The Center of Gravity feature, for example, reveals a close proximity between clusters on child exploitation, fair housing and disability, indicating that a more encompassing approach to these issues might be more beneficial than tackling each in isolation.

There are many ways to take these insights further. Once an area of interest has been identified, one can run a more targeted analysis to draw out more detail. It is also possible to combine analyses. For instance, one could analyze DoJ press releases alongside news media reports to visualize the divide between government policy priorities and with what is considered public interest.