

# Iowa Initiative for Artificial Intelligence

## Final Report

Project title:	Distance Reading Patent Law		
Principal Investigator:	Principal Investigator: Jason Rantanen, College of Law Co PI: Kang Zhao		
Prepared by (IIAI):	Avinash Mudireddy		
Other investigators:	N/A		
Date(mm/dd/yyyy):	07/03/2022		
Were specific aims fulfilled:	Yes		
Readiness for extramural proposal?	Yes		
If yes ... Planned submission date	01/18/2022		
Funding agency	NSF		
Grant mechanism			
If no ... Why not? What went wrong?			

### **Brief summary of accomplished results:**

#### **Research report:**

#### **Aims (provided by PI):**

The United States Court of Appeals for the Federal Circuit is a paradox of legal decision making. The Federal Circuit is the federal appellate court that hears virtually all appeals involving patents. Congress created the Federal Circuit in 1982 for the purpose of bringing uniformity and predictability to patent law (Dreyfuss 1989). The uniformity of patent law jurisprudence and the predictability of outcomes is generally recognized as critically important in incentivizing investment in patentable research and development. Yet, it is also widely accepted that uniformity risks creating stagnancy and entrenchment in patent law, rendering it unable to improve and making it difficult to reverse misguided choices about the law. (Nard and Duffy, 2007; Schwartz 2014). For decades, legal scholars have assessed these theories by studying the rates of patent validity and infringement outcomes, reversals, and dissents. (Vacca 2019). These tools, however, provide only a limited and indirect way of assessing the uniformity, predictability, stagnancy and entrenchment hypotheses. Consequently, the degree to which these theories actually justify continuing the existence of the Federal Circuit or abolishing it remains unknown. There is thus a *critical need to develop and use an analytical tool that more directly captures the court's decision making*. Without use of more advanced methodologies than are presently available, future discussions about the continued existence of the Federal Circuit and the patent law will stagnate and remain blocked by a relatively low level of empirical data.

Our *long-term goal* is to understand the Federal Circuit's decision making, particularly as it shapes patent law. The overall objective for this project is to assess the uniformity and predictability of Federal Circuit opinions using computational text analysis techniques. Our *central hypothesis* is that the uniformity and predictability of the Federal Circuit's opinions have increased over time, but are subject to disruptions that can temporarily decrease uniformity and predictability. While the existing empirical evidence relying on reversal and dissent rates is mixed, Congress's intent was to bring uniformity and predictability to patent law. However, in recent years the Supreme Court has become more active in patent law, potentially acting as a disruptive element. (Golden 2009). In addition, the periodic replacement of judges on the court provides a logical discontinuity. (Rantanen & Petherbridge 2014). If the hypothesis of increased uniformity and predictability is correct, the text of the court's opinions should increase in similarity and become more predictable over time while also reflecting the effect of the Supreme Court's increased activity and the appointment of new judges. By using computational text analysis, we will be able to

directly observe changes in judicial decision making beyond conventional indicators, thus allowing us to more clearly assess the uniformity and predictability hypotheses.

Note- Important terms:

Main opinion = opinion of majority of judges in a case

Separate opinion = opinion of individual judges in case they differ/confer from the main opinion

### **Specific aims.**

We plan to attain the overall objective by pursuing the following specific aims:

- 1. Develop a functional dataset of the text for all precedential Federal Circuit opinions from 1982-2020.**  
While datasets of Federal Circuit opinions exist, they are not currently in a form that readily permits computational text and network analyses, nor do many of them have key metadata. This project will apply human-coded and automated techniques to existing resources to produce a functional dataset of all Federal Circuit opinions since the creation of the court in 1982, including the development of an internal taxonomy of key components of judicial opinions to permit more granular analysis.
- 2. Construct a network to analyze changes in the citation to legal precedent in Federal Circuit opinions.**  
By analyzing structural properties of the citation network, we will identify influential opinions over time, assess whether the judicial precedents that the Federal Circuit cites in its opinions are changing or remaining stable, and whether there are key network characteristics of tunnel vision and jurisprudential ossification.
- 3. Apply text mining methodologies to analyze the similarity of Federal Circuit opinions within and between judges, over time, and across different subnetworks in the citation network.**  
Our *working hypothesis* is that textual similarity between judicial opinions should demonstrate issue, temporal, and ideological effects, and should be greater in areas where the issues addressed by the court involve judicial specialization as opposed to areas in which the Federal Circuit is exposed to other courts' opinions such as on procedural issues.
- 4. Construct a proof-of-concept model that predicts the outcome of the court's decisions.** The components of the model will include components of judicial decision making such as the textual content of the parties' briefs, the procedural posture of the appeal, issues being appealed, and the identity of the panel. We *hypothesize* that predictability should increase over time, but decrease in proximity to Supreme Court decisions and the appointment of a new judge to the court.
- 5. Disseminate the dataset and computational tools for other researchers to use.**  
Once the dataset is created, we will make the functional dataset and research tools publicly available by archiving them on a data repository. Our goal is to foster further computational studies by creating a separate website with an easy-to-use interface for the data and publicizing the availability of these resources. In addition, we will host a workshop for researchers interested in using the dataset or research tools.

However, in place of aim 4, we addressed a different aim : Sentiment analysis of the opinions.

### **Data:**

The dataset contains all Federal Circuit written opinions 2004-2021 arising from the district courts. The team hand coded dataset of case characteristics about these decisions, including author, separate opinion, issues, nature of parties (Federal Circuit Dataset Project/Compendium of Federal Circuit Decisions (empirical.law.uiowa.edu))

The preprocessing steps include:

- Convert PDFs of Federal Circuit opinions to text fields
- Use key language to automatically identify the start & end of majority & separate opinions (verified against human coding)
- Match text & text analysis to existing data in the Federal Circuit Dataset Project
- Limited to non-en banc opinions authored by judges who have authored at least 20 opinions.
- Dataset consists of 1,763 unanimous opinions, 402 majority opinions, 244 dissenting opinions, 82 concurring opinions, and 94 opinions that are both concurring and dissenting

## AI/ML Approach:

While AIM 1&2 are addressed by PI and Co-PI, IIAI team addressed 3,4,&5.

### The methodology for AIM 3 is:

#### Data processing:

Once the data is prepared for text analysis, with the help of Natural Language Processing, we further process data to find features that are used to calculate a textual similarity score between every pair of judicial opinions.

To analyze text in its original form has some issues. In the English language, the usage of articles, prepositions, and few words are frequent though the words do not relate to the context of opinion. However, due to more frequency, they might influence the overall outcome statistically. These high-frequency words and non-useful words are termed as stopwords. As a part of this preprocessing step, we identified and eliminated a few stop-words in the context of vocabulary used in opinions. We also observed that usage of numbers and dates is quite often in opinion text. However, we believe the presence of the numbers is not useful for the overall analysis and hence, eliminated them.

After the stop-word removal, all the text in an opinion is treated as a single input sequence. In that case, the text represented in each opinion varies depending on case to case. The difference can be in the choice of words or the word length of the opinion. However, to mathematically compare two opinions, it is important to represent them in the same dimensional space. Hence, it is important to convert this variable-length input vector to a fixed-length input vector. Le & Mikolov, 2014 in their paper “Distributed Representations of Sentences and Documents” proposed a new algorithm doc2vec for this purpose. This model is based upon earlier proposed word2vec algorithm by Tomas Mikolov et al., 2013 in their paper “Distributed Representations of Words and Phrases and their Compositionality”.

In word2vec, a sequence of words is used to predict the next word given an entire corpus of various sentences. Such a representation encodes various relations between words like synonyms, antonyms, analogies, or frequent use together. These relations are encoded in terms of a fixed-length vector. If you consider this vector as a point in a multi-dimensional space, then words closer to each other in context will have similar encoding in this multi-dimensional space. In other words, they are spatially closer to each other.

Doc2vec is very similar to word2vec with an additional paragraph-id/document-id is given as an input along with a sequence of words to predict the next word. Logically, adding the paragraph id/document id will introduce the context of certain words occurring together to express a particular concept. Using this algorithm, we encode each document to a fixed-length vector. We call these vectors as document embeddings. In our case, each opinion is considered a document. We train the doc2vec algorithm to encode the opinions in the form of fixed-length vectors.

#### Methodology:

Cosine similarity measures the similarity between two vectors. It is measured by computing the cosine of the angle between two vectors in a multi-dimensional space. The higher score indicates that both the vectors are roughly pointing to the same direction in the space indicating a higher similarity. This measure of similarity can be used to compare documents or, say, give a ranking of all the opinions with respect to a given vector of a particular opinion. Let  $x$  and  $y$  be two opinion vectors for comparison. Then we have cosine similarity as

$$\text{similarity}(x, y) = \cos(\theta) = \frac{x \cdot y}{\|x\| \|y\|}$$

where  $\|x\|$  is the Euclidean norm of  $n$ -dimensional vector  $x = (x_1, x_2, x_3, \dots, x_n)$  also known as the length of the vector given by

$$\|x\| = \sqrt{(x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2)}$$

Similarly,  $y$  is the Euclidean norm of vector  $y$ . A similarity value of 0 means the vectors are orthogonal to each other indicating no match. Likewise, a perfect score of 1 indicates an exact match.

Using these similarity scores, we are able to analyze the similarity of different subsets of the data. With respect to the baseline assumptions about the nature of judicial opinions, the first thing we can test is the similarity of each judge’s opinions to that judge’s opinions versus other judges’ opinions. This is done by calculating an average

similarity score for each judge with respect to their own opinions and an average similarity score with respect to all other judges. We can also identify which judges are textually similar by calculating an average similarity score for each judge with respect to every other judge’s opinions. For example, Judge A’s opinion writing may be more similar to Judge B’s opinions writing than to Judge C’s opinion writing.

We can also analyze the degree to which the text of opinions involving different issues are more similar to one another than to opinions involving other issues by again calculating an average similarity score for each coded type of opinion.

These analyses can be extended by the incorporation of time to test our core hypotheses. To do so, it is necessary to compare opinions that are proximate in time to examine whether textual similarity changes over time. This can be done by calculating the average similarity score for each opinion to every other opinion within +/- 365 days. This average similarity score reflects the similarity of each opinion to all other opinions within a year of that opinion. By analyzing changes in these average similarity scores over time, we can (1) assess the degree to which the text of opinions has generally become more similar over time, and (2) assess whether predicted disruptions, such as the replacement of a judge on the court or a Supreme Court decision.

**The methodology for AIM 4 :** (The original aim is modified to sentiment analysis)

Data processing:

- The opinions are first tokenized
- Then the tokenized data can be of varied length. Hence we limited each text to a length of 512 and the remaining length is wrapped as a next vector.
- Hence for an opinion if there are multiple vectors, sentiment of each vector is voted and averaged for scores

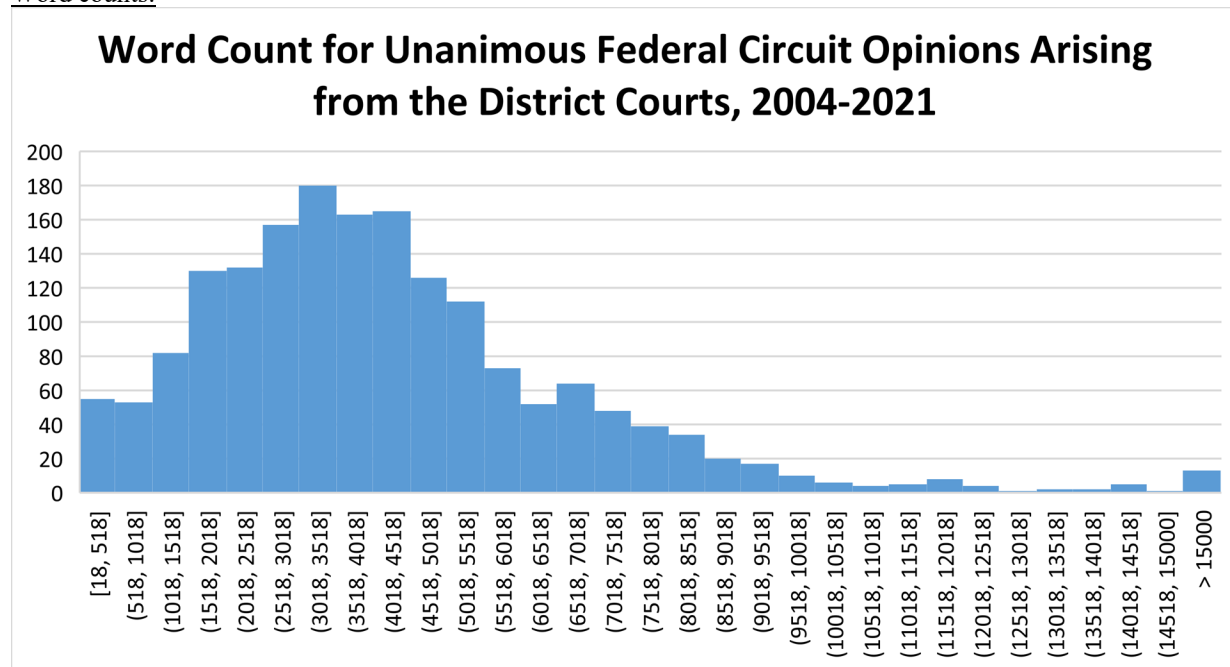
Methodology:

We used nltk’s setiment analysis pipline to compute the sentiment of each vector.

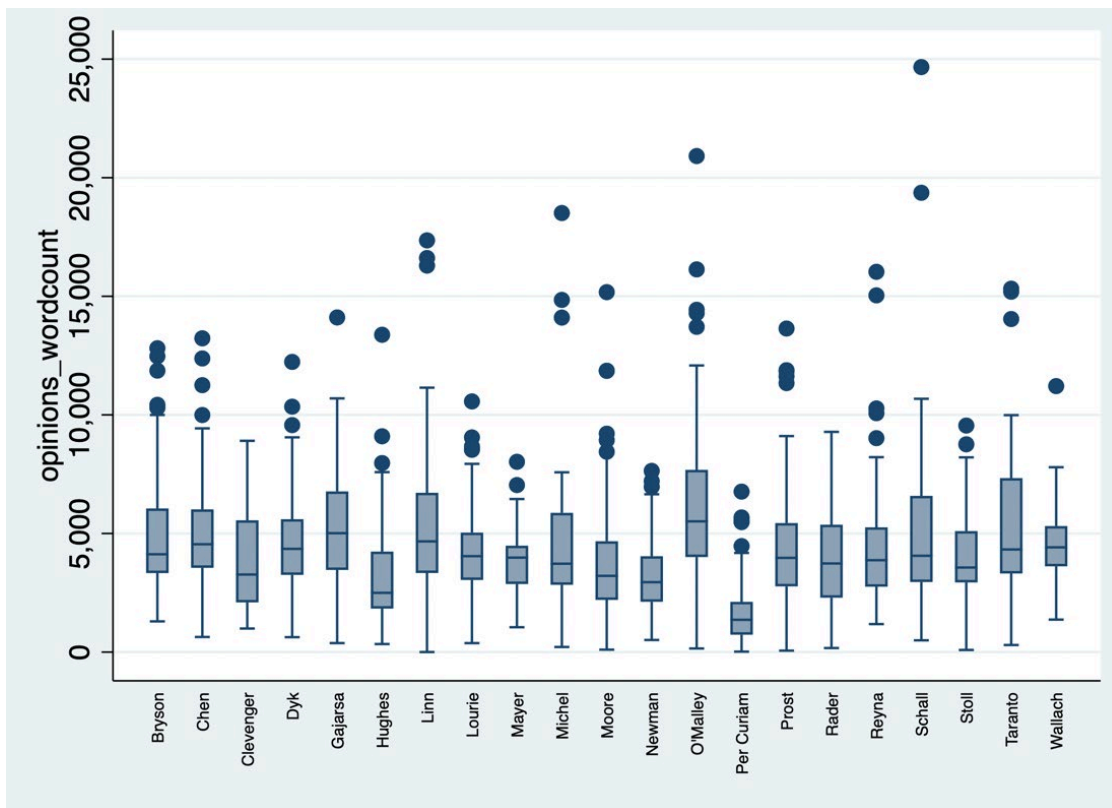
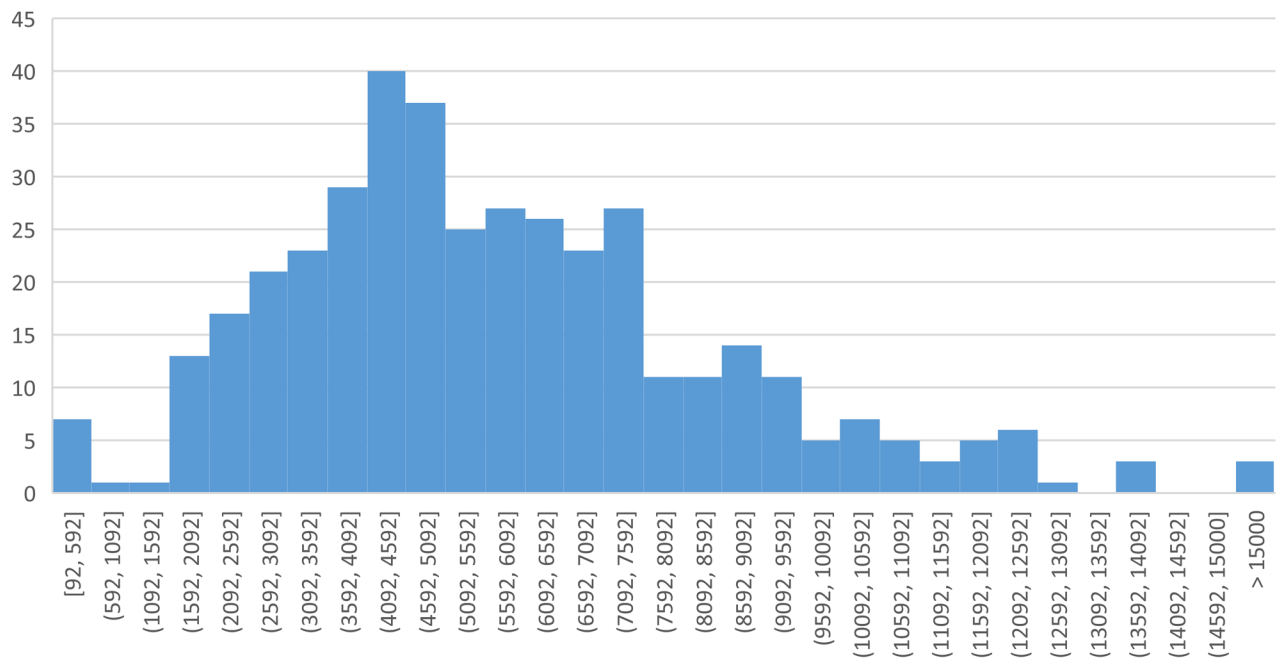
**Results:**

The preliminary results are as follows:

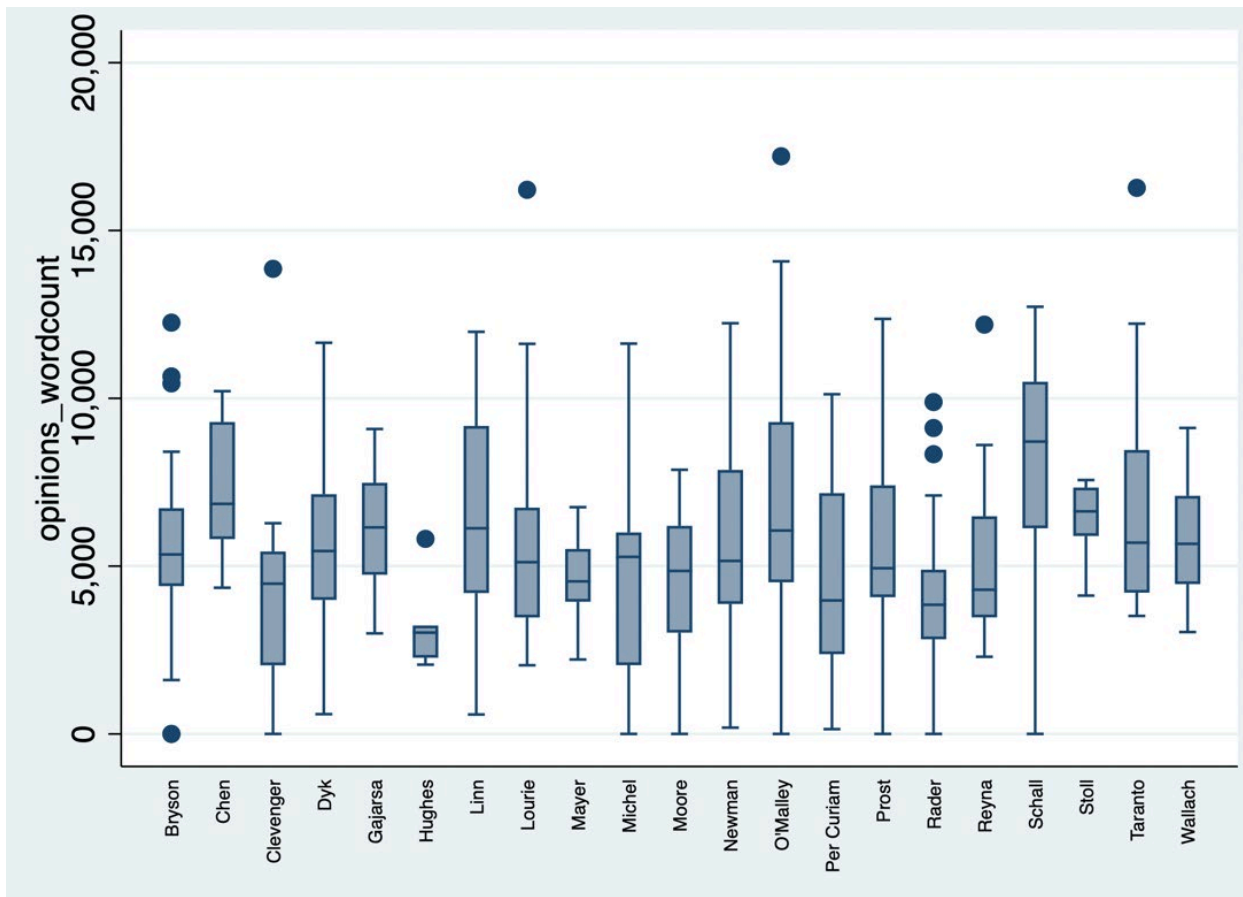
Word counts:



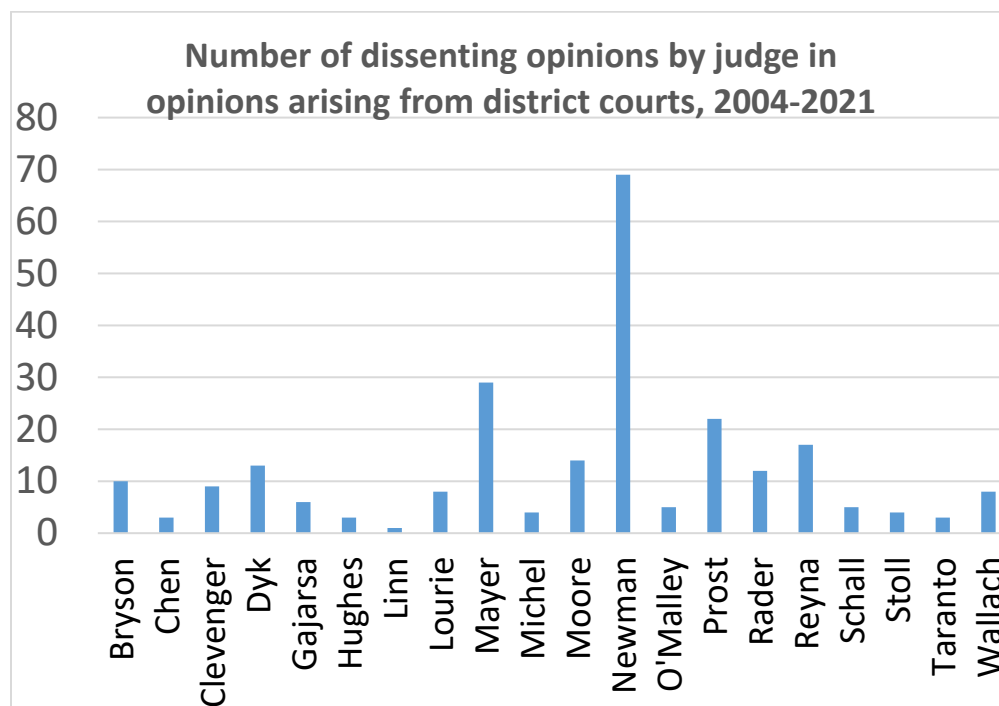
## Word Count for Majority Federal Circuit Opinions Arising from the District Courts, 2004-2021

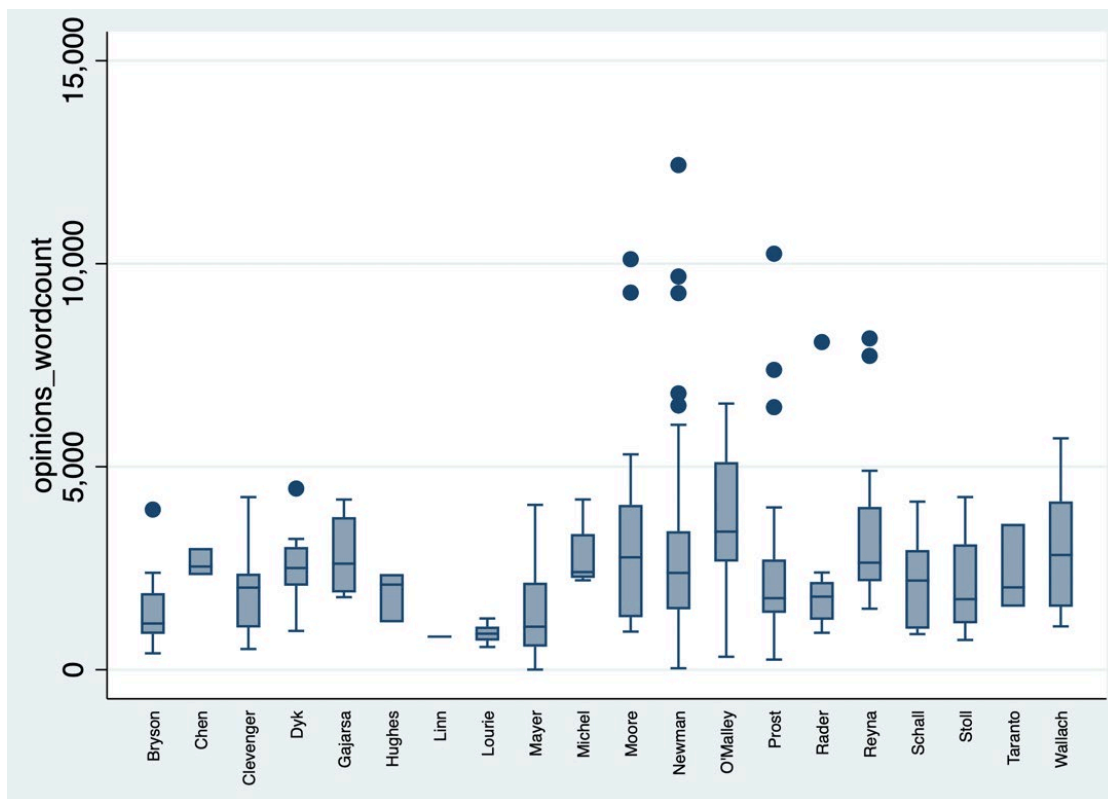


Unanimous Opinions, wordcount by author



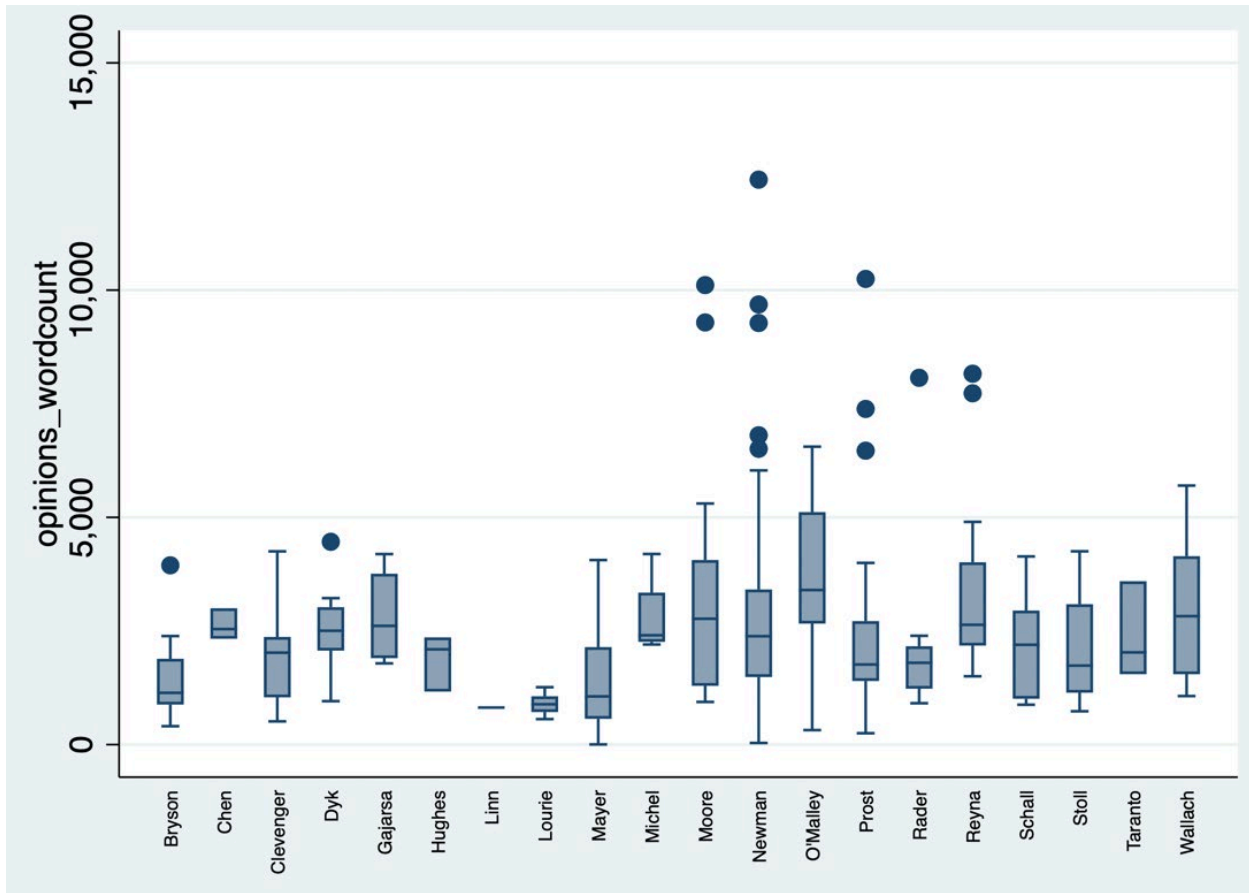
Majority Opinions, wordcount by author





**Dissenting Opinions, wordcount by author**

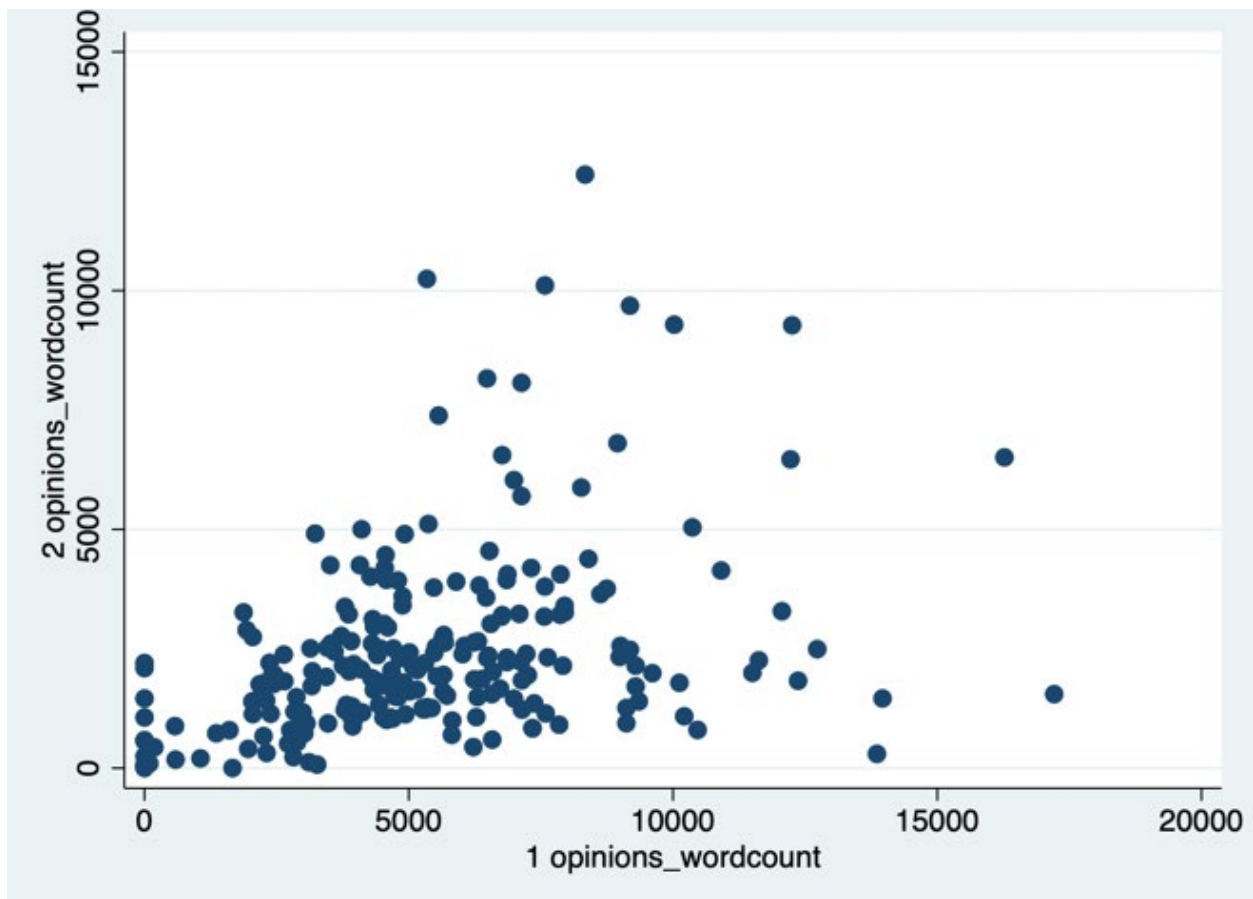
**Note: some of these n's are extremely small (e.g.: Linn (1); Chen, Hughes (3); O'Malley (5)).**



Note: some of these n's are extremely small (e.g.: Linn (1); Chen, Hughes (3); O'Malley (5)).

**Relationship between Majority & Dissenting opinion length**





**Length of majority opinions to corresponding dissenting opinions in CAFC opinions arising from the district courts**

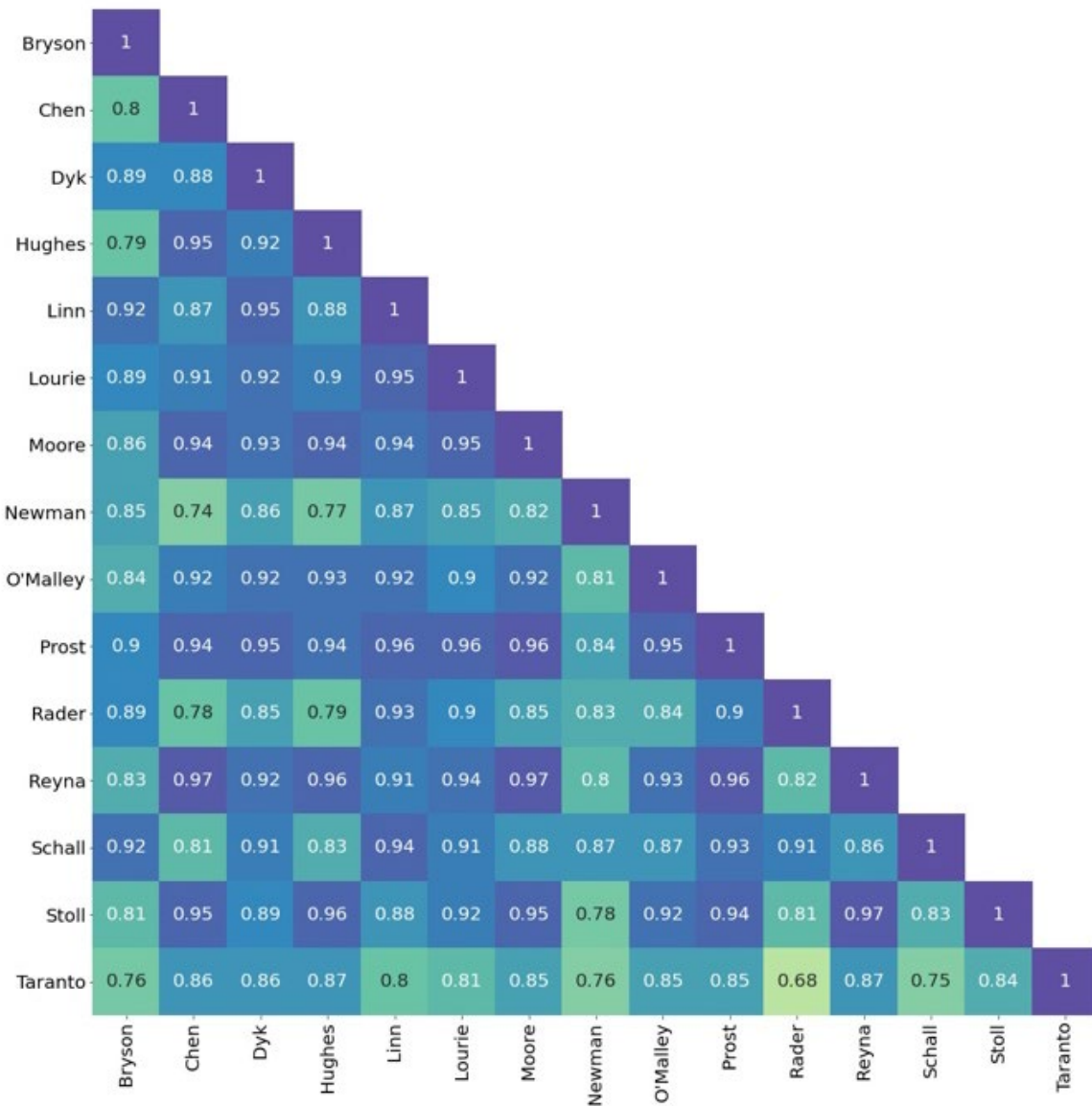
```
. regress opinions_wordcount2 opinions_wordcount1 if Opinion_type2=="Dissenting"
```

Source	SS	df	MS	Number of obs	=	227
Model	124763781	1	124763781	F(1, 225)	=	37.28
Residual	753061745	225	3346941.09	Prob > F	=	0.0000
Total	877825527	226	3884183.75	R-squared	=	0.1421
				Adj R-squared	=	0.1383
				Root MSE	=	1829.5

opinions_wordcount2	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
opinions_wordcount1	.2470272	.0404599	6.11	0.000	.1672985	.326756
_cons	1115.197	249.6692	4.47	0.000	623.2081	1607.186

**Length of majority opinions to corresponding dissenting opinions in CAFC opinions arising from the district courts**

**Similarity between judges' opinions:**



**Textual similarity between judges based on written unanimous & majority opinions arising from district courts, 2004-2021**

**Sentiment Analysis**

Preliminary Results:

- Opinions overwhelmingly negative with high scores (i.e.: most opinions have a negative sentiment with a score of 0.94-0.96).
- Benchmark comparison:
  - Ch. 1-6, "Pride and Prejudice": POSITIVE, score 0.96
  - Ch. 1-2, "Crime and Punishment": NEGATIVE, score 0.95

**Ideas/aims for future extramural project:**

- Construct initial dataset of text of all precedential Federal Circuit opinions from 1982 – 2021, including automated coding, manual document-level coding, and case citation coding.
- Manual coding of sub-opinion level attributes.
- Construction of document taxonomy.
- Extraction of intra-document citations and construction of citation network
- Conduct textual similarity analysis
- Conduct citation network analysis