

Statistically Significant Judging: Mechanizing Originalism Through Corpus Linguistics and AI

NATALIE L. SIMON*

ABSTRACT

In interpreting the U.S. Constitution, accuracy, transparency, and replicability are increasingly important. A recent poll showed that most Americans believe that United States Supreme Court Justices sacrifice impartiality for ideology.¹ That is particularly problematic when a shocking number of Americans “know literally nothing about the Constitution.”² For example, one in three people cannot name a single First Amendment right.³ So, many Americans do not know—and most Americans do not trust SCOTUS to truthfully tell them—what is in this country’s most important governing document. Instead of grounding linguistic interpretation in high-minded, lofty reasoning, many judges now seek to use what lay people understand: numbers.

TABLE OF CONTENTS

INTRODUCTION	578
I. WHAT IS CORPUS LINGUISTICS?	578
II. CORPUS LINGUISTICS AS AN ORIGINALIST TOOL	580
A. <i>How Is It Done?</i>	582
B. <i>Legal Use of Corpus Linguistics</i>	584
1. Sixth Circuit	585
2. Scholarly Use of Corpus Linguistics	586
C. <i>Corpus Linguistics as an Interpretive Tool</i>	587
III. ARTIFICIAL INTELLIGENCE AS A CORPUS LINGUISTICS TOOL	590
A. <i>AI and Corpus Linguistics in the Courts</i>	590

* © 2026, Natalie L. Simon.

1. Thomas Beaumont & Linley Sanders, *New Poll Shows Majority of Americans Believe Supreme Court Justices Put Ideology over Impartiality*, PBS NEWS (June 27, 2024, 10:11 AM), <https://www.pbs.org/newshour/nation/new-poll-shows-majority-of-americans-believe-supreme-court-justices-put-ideology-over-impartiality> [https://perma.cc/YAC8-WVPW].

2. Chris Cillizza, *Americans Know Literally Nothing About the Constitution*, CNN (Sept. 13, 2017, 4:39 PM), <https://www.cnn.com/2017/09/13/politics/poll-constitution/index.html> [https://perma.cc/724Y-2SP2].

3. *Id.*

<i>B. Experiment: Comparing AI to Strang’s Corpus-Linguistics Study</i>	592
<i>C. Legitimacy of AI Corpus Linguistics and Originalism</i>	596
CONCLUSION	599
APPENDIX A	600
APPENDIX B—CLAUDE	606
APPENDIX C—GEMINI	609

INTRODUCTION

Corpus linguistics and its focus on the frequency of word usage in a particular language provide judges a veil of empiricism—judging based on statistical significance—and a potentially more accurate way to conduct originalist interpretation. But resource-limited, time-constrained judges cannot conduct the by-hand corpus-linguistics analysis that consumes hundreds of pages of scholarly literature. With the advent of artificial intelligence (AI), maybe they don’t have to.

This paper will explore whether corpus linguistics and AI can bridge the gap between an uninformed, distrustful public and a largely originalist Supreme Court. First, it will trace the rise of corpus linguistics from its early origins to its modern use in judicial opinions and legal scholarship, pointing out strengths and weaknesses of corpus linguistics as an originalist tool. Then, it will compare AI interpretive output to a thorough corpus-linguistics analysis of the word “religion” to determine whether AI better meets the established criteria for an acceptable corpus-linguistics study.⁴

I. WHAT IS CORPUS LINGUISTICS?

Corpus linguistics is an “empirical analys[i]s of actual patterns of language use in natural texts.”⁵ To conduct these analyses, linguists use corpora, which are bodies of naturally occurring language that are assembled to represent a specific language or way of speaking.⁶ Corpora include language from several sources:

4. For the sake of thoroughness, a researcher (or judge) may want to include relevant cognates like “religious” or “religiosity.” This might provide a richer, more robust picture of the way in which society viewed the idea of “religion.” After all, a description of something as “religious” gives insight into whether “religious” things were more commonly theistic or secular. However, the study of “religion” to which this paper will compare AI’s outputs did not include those cognates. As the length of the study’s description shows, doing a corpus-linguistics analysis of those cognates would be well beyond the scope of this paper, as it would take an immense team several months (at best) to find and code those uses. So, for purposes of this paper, the analysis will only include the word “religious.”

5. *Law & Corpus Linguistics—Background*, BYU L., <https://lcl.byu.edu/projects/law-corpus-linguistics-background/> [https://perma.cc/8R7W-768L].

6. TONY MCENERY ET AL., *CORPUS-BASED LANGUAGE STUDIES: AN ADVANCED RESOURCE BOOK 4* (2006).

textbooks, magazines, newspapers, literature, television shows, academic papers, and much more.⁷ Nearly any source reducible to written language can contribute to corpora.⁸

Corpus linguistics existed long before its migration into legal studies. One group of scholars traced corpus-based studies back to the early 1940s, when linguists had “shoeboxes filled with paper slips” instead of today’s massive, computerized data storage.⁹ But such paper-based corpus linguistics was fallible, time-consuming, and drastically limited in scope, so it was widely considered a “skewed” science not worth reliance or study.¹⁰ It wasn’t until the early 1960s, when computer technology enabled the first modern corpus of the English language, the Brown Corpus, that corpus linguistics once again piqued scholarly interest.¹¹ Since then, corpus linguistics has begun to “illuminat[e] . . . virtually all branches of linguistics or language learning.”¹² And the number of existing corpora has continued to expand exponentially.¹³

Corpus linguistics works primarily by assessing frequencies.¹⁴ Corpus-linguistics analyses assess several types of frequencies,¹⁵ but they mainly focus on “the number of senses, or meanings, that a linguistic expression may have and the most frequently used meaning.”¹⁶ To find these senses and meanings, corpus linguists count how often a specified term appears in a particular corpus.¹⁷ But this count is not the end of the analysis. Instead, the linguists also collect “quantitative evidence about word co-occurrence in corpora,” otherwise known as collocation.¹⁸ Put simply, the linguists do not count only the word but also the words surrounding it. One of the most prominent early linguists described it as “[y]ou shall know a word by the company it keeps.”¹⁹

7. Gena R. Bennett, *Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers*, <https://press.umich.edu/pdf/9780472033850-part1.pdf> [<https://perma.cc/X654-LAAU>].

8. *Id.*

9. McENERY ET AL., *supra* note 6, at 3.

10. *Id.*

11. *Id.* at 4.

12. *Id.* at 80 (quoting DOUGLAS BIBER ET AL., *CORPUS LINGUISTICS: INVESTIGATING LANGUAGE STRUCTURE AND USE* 11 (1998)).

13. *See, e.g.*, Michael McCarthy & Anne O’Keeffe, *What Are Corpora and How Have They Evolved?*, in *THE ROUTLEDGE HANDBOOK OF CORPUS LINGUISTICS* 3, 3–13 (Michael McCarthy & Anne O’Keeffe eds., 2010) (tracing the history and evolution of corpus linguistics); ENGLISH-CORPORA.ORG, <https://www.english-corpora.org/> [<https://perma.cc/GA3V-T6WY>] (listing just a small subset of the available corpora).

14. *See* Stefan Th. Gries, *What Is Corpus Linguistics?*, 3 *LANGUAGE & LINGUISTICS COMPASS* 1, 2 (2009) (explaining that corpus linguistics is “based on the evaluation of some kind of frequencies”).

15. *See id.* at 2–3 (describing the various frequencies and patterns scholars can measure with corpus linguistics).

16. Brief for Amici Curiae Corpus-Linguistics Scholars Professors Brian Slocum, Stefan Th. Gries, and Lawrence Solan in Support of Employees at 8, *Bostock v. Clayton County*, 590 U.S. 644 (2020) (No. 17-1618).

17. Gries, *supra* note 14, at 2–3.

18. Dana Gablasova et al., *Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence*, 67 *LANGUAGE LEARNING* 155, 158 (2017).

19. J.R. FIRTH, *A SYNOPSIS OF LINGUISTIC THEORY, 1930–1955*, in *STUDIES IN LINGUISTIC ANALYSIS* 11 (1962).

For example, if linguists were trying to find the most commonly used meaning of the word “tide,” the linguists would first identify each use of the word “tide” in the relevant corpus. Then, the linguists would count and analyze the words most commonly preceding or following the word “tide.” This method allows the linguists to identify not only how common a word is but also the most common *sense* of the word within a particular language.²⁰ In the “tide” example, if “tide” were almost ubiquitously preceded by “roll,” linguists might infer that “tide” meant something different than if “tide” were always preceded by “high.” In this way, frequency and collocation enable scholars to develop two conclusions: (1) whether a particular word was used frequently in a particular linguistic subset and (2) the most common sense of the word when used by that linguistic subset.²¹

On its face, this approach can appear to be too reductive or simplistic. The meaning of a word cannot be understood by looking at the word in isolation.²² To count the raw frequency of the word without accounting for textual, semantic, and cultural context could be highly misleading. Corpus linguists attempt to account for these things through collocation. This is not a perfect solution, as the meaning could be provided by context found outside of the immediately preceding and succeeding several words, but it is an attempt to move beyond mere counting of words out of context. The weaknesses of a corpus-linguistics approach are discussed in detail later in the paper, but at the outset, certain constraints are apparent just from the description of the enterprise.

Despite its potential weaknesses, corpus linguistics has attracted the attention of originalist scholars and jurists as a potential way to make originalism empirical—or based on statistically significant results—rather than based on intuition, guessing, or cherry-picked dictionaries.

II. CORPUS LINGUISTICS AS AN ORIGINALIST TOOL

Originalism as a theory of constitutional interpretation fixates on finding the original public meaning of constitutional terms—the “objective legal construct” as it would have been defined at the time of the Constitution’s enactment,²³ usually meaning the time of the Constitution’s ratification, since ratification is what—based on principles of popular sovereignty—gives the Constitution its legitimacy.²⁴ Two of the leading scholars on the application of corpus linguistics to the law have explained that “[t]he basic premise of [originalism] is that a legal text is

20. Brief for Amici Curiae Corpus-Linguistics Scholars, *supra* note 16.

21. Gries, *supra* note 14, at 2–3.

22. See, e.g., Marcel Danesi, *Semiotics: The Study of Meaning (Part 2)*, UNIV. OF TORONTO PRESS (Mar. 19, 2021), https://utorontopress.com/blog/2021/03/19/danesi-the-study-of-meaning-part-2/?srsltid=AfmBOooNYbVPrFMiSEgFzP9Yq_1kBM3i4ULM6J5FFHMHSWoYbuGk5B [https://perma.cc/62H4-GLDK].

23. Stephen G. Calabresi, *On Originalism in Constitutional Interpretation*, NAT’L CONST. CTR., <https://constitutioncenter.org/the-constitution/white-papers/on-originalism-in-constitutional-interpretation> [https://perma.cc/RD83-MJMW].

24. See, e.g., Thomas B. Colby, *Originalism and the Ratification of the Fourteenth Amendment*, 107 NW. U. L. REV. 1627, 1628 (2013) (“[T]he Constitution owes its legitimacy as higher law to the fact that

a form of communication that uses natural language in order to accomplish its purposes,” so “textual language should be interpreted in light of the accepted and typical standards of communication that apply outside of law.”²⁵ So originalists—much like linguists—are looking for the set meaning of a term as it was used during the relevant time period (usually the Constitution’s ratification).²⁶

But “originalism” is a family of theories, and the theories do not all agree on the proper linguistic subset of study. Originalist scholars like Michael Rappaport and John McGinnis believe that the Constitution is a legal document, and thus constitutional terms should be given their *legal* meaning.²⁷ Other originalist scholars like Larry Solum support a lay-language view of the Constitution, seeking to give constitutional terms the meaning ordinary citizens would have attached to those terms during the Founding Era.²⁸ Still others advocate for assigning constitutional terms the meaning which the Framers themselves would have given the terms.²⁹ These are just a few of the interpretive theories housed under the “originalism” umbrella, but the doctrinal differences amongst these theories are not dispositive of the corpus-linguistics question. Corpus linguistics as an interpretive practice can be adjusted to accommodate the different theories. For example, legal-meaning originalists could draw from a corpus existing only of *legal* documents from the Founding Era—contracts, statutes, judicial opinions—to find the most common meanings of constitutional terms.³⁰ In contrast, lay-language originalists could study a corpus encompassing newspapers, magazines,

it was ratified by the American people through a supermajoritarian process. As such, it must be interpreted according to the original meaning that it had at the time of ratification.”)

25. Stefan Th. Gries & Brian G. Slocum, *Ordinary Meaning and Corpus Linguistics*, 2017 BYU L. REV. 1417, 1424 (2017).

26. See, e.g., Colby, *supra* note 24.

27. Michael Rappaport & John McGinnis, *What Is Original Public Meaning?*, 76 ALA. L. REV. 223 (2024).

28. Lawrence B. Solum, *Original Public Meaning*, 2023 MICH. STATE L. REV. 807, 816 (2023) (“For the purposes of Public Meaning Originalism, the relevant public would consist of citizens who were competent speakers of American English at the time each constitutional provision was framed and ratified.”).

29. See H. Jefferson Powell, *The Original Understanding of Original Intent*, 98 HARV. L. REV. 885 (1985) (evaluating the validity of interpreting the Constitution via inferred Framers’ intent).

30. One intriguing thing to note about legal-meaning originalism is the subtle, sometimes unrecognizable way lay meanings can creep into documents that are otherwise wholly legal in nature. For example, the disposition of a case decided by a jury would have a legal tenor if drafted as a judicial opinion or order, but the substance of the decision would likely have been influenced by the jury’s lay understanding of legal words. Even when juries are instructed by the judge as to the *legal* meaning of words, they might struggle to fully understand or apply that legal meaning without lay understandings creeping in. For example, one study by a nationally recognized expert on jury behavior showed that “although juries discussed instructions a lot, they were inaccurate in comprehending them 17 percent of the time.” *New Study Reveals How Juries Think and Behave*, ABA (Sept. 2017), <https://www.americanbar.org/news/abanews/publications/youraba/2017/september-2017/new-study-reveals-how-juries-think-and-behave/#:~:text=The%20findings%20show%20that%20jurors,choosing%20a%20foreperson%20happens%20quickly> [https://perma.cc/4MKR-9GDR]. So even if the official judicial record shows a reading and interpretation of “legal” terms, there could be (and likely are) outside lay influences making a difference in that legal determination. This could be problematic for legal-meaning originalists, though a critique or defense of legal-meaning originalism is beyond the scope of this paper.

personal correspondence, and other non-legal documents from the Founding Era. Because the interpretive differences among originalist theories are comfortably accommodated by the wide array of available corpora, this paper will use the term “originalism” to encompass the entire family of theories and will evaluate the application of corpus linguistics to the root commitment of all those theories: finding the original meaning, as evidenced by Founding Era sources, of constitutional terms.

A. *How Is It Done?*

Corpus linguistics may sound like a powerful originalist tool, but does corpus linguistics in practice live up to its proclaimed ability to “revolutionize statutory and constitutional interpretation”?³¹ In 2017, James C. Phillips and Jesse Egbert laid out “best practices” for jurists and scholars using corpus linguistics for legal analysis.³² In their detailed article, the authors explained the already-existing practices used in other fields to conduct analyses of corpora and created a step-by-step checklist for legal corpus linguists to follow.³³ Building on this literature, in 2018, Thomas R. Lee and Stephen C. Mouritsen conducted a thorough analysis of corpus linguistics as an originalist tool—first describing the practice in great detail and then applying the practice to selected terms.³⁴ In 2019, Lee, this time writing with Phillips, once again offered a full-throated defense of corpus linguistics, replete with select examples.³⁵ Because Lee was a justice on the Supreme Court of Utah, and he frequently relied on corpus linguistics in his judicial opinions,³⁶ these articles provide a unique look into how a jurist mechanizes corpus linguistics to reach conclusions with vast real-world impacts.

First, the linguist³⁷ must choose a corpus.³⁸ Phillips and Egbert offer detailed descriptions of how to collect material and create a corpus from scratch, but because of the limited nature of this paper’s inquiry—and the limited nature of most judges’ time—this paper will assume that the linguist uses an already-existing corpus, such as the Corpus of Early Modern English (COEME) created by Brigham Young University.³⁹ Each corpus contains a different selection of materials and bears unique characteristics that make it well-suited for some inquires

31. Amanda K. Fronk, *Big Lang at BYU*, Y MAG., <https://magazine.byu.edu/article/big-lang-at-byu> [<https://perma.cc/PSW2-YRUT>].

32. James C. Phillips & Jesse Egbert, *Advancing Law and Corpus Linguistics: Importing Principles and Practices from Survey and Content-Analysis Methodologies to Improve Corpus Design and Analysis*, 2017 BYU L. REV. 1589 (2017).

33. *Id.* at 1605.

34. Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 YALE L.J. 788 (2018).

35. Thomas R. Lee & James C. Phillips, *Data-Driven Originalism*, 167 U. PA. L. REV. 261 (2019).

36. *See, e.g.*, In re Adoption of Baby E.Z., 266 P.3d 702, 724 (Utah 2011) (Lee, J., concurring in part).

37. For simplicity purposes, I will use the term “linguist” to refer to any scholar, judge, lawyer, or other person conducting the corpus-linguistics analysis.

38. Phillips & Egbert, *supra* note 32, at 1605.

39. *BYU-Corpus of Early Modern English (BYU-COEME)*, BYU L. <https://lcl.byu.edu/projects/byu-corpus-of-early-modern-english-byu-coeme/> [<https://perma.cc/U3T5-3AYZ>].

and inappropriate for others.⁴⁰ Some “general” corpora reflect a broad speech community, while other “special” corpora reflect only a limited language or dialect.⁴¹ For example, the Linguistic Data Consortium at the University of Pennsylvania produced a corpus of recorded Egyptian Arabic telephone calls.⁴² There are “monitor” corpora that continuously update as societies create new text, and there are “historical” corpora that contain text from only a specific time period.⁴³ These are just a few of the variable characteristics of corpora from which the linguist must choose. And the linguist’s choice can have profound impacts on the analysis’s results. Returning to the earlier explanation of the differing originalist theories, if Rappaport uses a corpus consisting only of historical legal texts and Solum uses a corpus consisting of an array of diverse lay sources, the two scholars could get widely diverging results. So the linguist’s pre-existing jurisprudential and theoretical commitments (apart from those driving the linguist to conduct a corpus-linguistics analysis) offer the first, but far from the last, opportunity for subjectivity in the so-called “data-driven originalism.”⁴⁴

Second, the linguist selects coding categories in which she will place her subsequent search results.⁴⁵ This step, too, injects a vast amount of subjectivity into the enterprise.⁴⁶ Phillips and Egbert described three of the numerous approaches to coding.⁴⁷ To maintain brevity of description and save time for analysis, this paper will offer only a brief overview of the approaches. The first approach reduces the word in question to two possible meanings; “someone else sets the parameters,” whether that is the judge who framed the question on appeal or the attorney who framed the legal argument.⁴⁸ The second approach collects relevant dictionaries defining the word (the collection also being a subjective enterprise) and creates categories based on the dictionary definitions.⁴⁹ The third approach creates no categories before analyzing the data.⁵⁰ Instead, the linguist creates categories as she sorts through and identifies patterns among the results.⁵¹ Both the act of

40. Lee & Mouritsen, *supra* note 34, at 830–31.

41. *Id.*

42. See Alexandra Canavan et al., *CALLHOME Egyptian Arabic Speech*, LINGUISTIC DATA CONSORTIUM (1997), <http://catalog ldc.upenn.edu/LDC97S45> [<https://perma.cc/GKF3-X9JQ>].

43. Lee & Mouritsen, *supra* note 34, at 830–31.

44. Lee & Phillips, *supra* note 35.

45. Phillips & Egbert, *supra* note 32, at 1608.

46. Donald L. Drakeman, *Is Corpus Linguistics Better than Flipping a Coin?*, 109 *GEO. L.J. ONLINE* 81, 91 (2020) (“To assess how the word is used in each instance, the researcher needs to examine the document carefully to make a subjective judgment about the objective meaning of the word in that particular context. Put differently, the people mining the linguistic data must perform, for each of the dozens, hundreds, or thousands of hits, exactly the same formidable interpretive task that generated the need for the corpus linguistics research in the first place.”).

47. Phillips & Egbert, *supra* note 32, at 1608–12.

48. *Id.* at 1608.

49. *Id.* at 1609–10.

50. *Id.* at 1610–11.

51. *Id.*

choosing a coding method and the act of executing the code-category creation itself are deeply subjective enterprises.

Third, the linguist selects and trains individuals for the coding team.⁵² The coders' identities and the training they receive can vary *wildly* based on the resources, commitment, and abilities of the linguist conducting the study.⁵³ Yet again, this phase injects a dubious level of subjectivity into this supposedly data-driven, objective method.⁵⁴

Fourth, the coding team codes the results of a search in the corpus.⁵⁵ Phillips and Egbert encourage the linguist to conduct several rounds of review, with the coders conducting completely independent, blind coding and then passing the data to another coder.⁵⁶ But the authors acknowledge that multiple coders “may not always be practical because of the time or cost involved.”⁵⁷ This code-by-hand system with varying sizes of coding teams seeks to minimize—but admittedly cannot completely eliminate—the risk of shoddy, ill-considered analysis, whether intentional or not.⁵⁸

B. Legal Use of Corpus Linguistics

In 2011, Justice Lee on the Supreme Court of Utah debuted his use of corpus linguistics in a judicial opinion examining the proper meaning of the phrase “custody determination.”⁵⁹ From the Corpus of Contemporary American Usage, he reviewed 500 randomized sample sentences that used the word “custody” and determined that only twelve of those sentences made any reference to adoption.⁶⁰ Based on that statistic—that the word “custody” was ten times more likely to collocate (or co-occur) with the word “divorce” than the word “adoption”—Justice Lee sought to limit the phrase “custody” in the pertinent statute to divorce proceedings.⁶¹

52. *Id.* at 1615–17.

53. *See id.*

54. Drakeman, *supra* note 46, at 92 (“[T]he process of turning hits into quantifiable cases of one usage or another can potentially lead to different outcomes based on the subjective judgments of different researchers and their research assistants about the meaning of the various hits.”).

55. Phillips & Egbert, *supra* note 32, at 1613–15.

56. *Id.*

57. *Id.* at 1614.

58. *Id.* at 1617.

59. *In re Adoption of Baby E.Z.*, 266 P.3d 702, 723 (Utah 2011) (Lee, J., concurring in part).

60. *Id.* n.21.

61. *Id.* For someone as passionate about corpus linguistics as Justice Lee, this analysis was quite impoverished. His analysis, lasting no longer than a short paragraph and referring to interpretation of the statute at issue as “a contest between probabilities of meaning,” is far from the thorough, hearty analysis he demonstrates and for which he advocates in his scholarly writing. Instead, it reads like the opinions that corpus linguists critique as being cursory and baseless. Perhaps this supports the critique that while scholars may have ample time to conduct corpus-linguistics analyses, a judicial-docket timeline greatly ties the hands of judges. Or perhaps Justice Lee did not see fit to engage in a lengthy corpus-linguistics analysis due to his writing a concurrence rather than the majority opinion. Regardless, his analysis is strikingly simplistic—particularly for a corpus-linguistics scholar like Justice Lee.

His approach wasn't entirely unprecedented. In 1998, the United States Supreme Court "surveyed modern press usage, albeit crudely, by searching computerized newspaper databases" in Westlaw and LexisNexis.⁶² Although the Court didn't use an official "corpus" database, the Court did use a "random sampling" of sentences including "carry" (the word at issue) to determine its proper meaning.⁶³ After all, corpus linguistics had been praised as "evidence of what speakers believe to be acceptable utterances in their language, typically free from the overt judgment of others."⁶⁴ While corpus linguistics has yet to pervade the majority of courts, several courts—including the Sixth Circuit and Justice Lee's repeated use on the Utah Supreme Court—have turned to corpus linguistics. The Sixth Circuit, in *Fulkerson v. Unum Life Insurance Company of America*, used corpus linguistics to interpret the word "crime" as used in a federal statute.⁶⁵ And in *Bostock v. Clayton County*—a case interpreting the scope of sex discrimination—interested professors wrote an amicus curiae brief conducting an extensive corpus-linguistics analysis.⁶⁶ And Judge Kathryn Mizelle on the United States District Court for the Middle District of Florida also used corpus linguistics in her opinion interpreting a COVID-19 mask mandate.⁶⁷ Because of its rather nascent use in courts, there is a paucity of cases using corpus linguistics to interpret the United States Constitution. But while none of the cited opinions interpreted *constitutional* terms, they reveal exactly how jurists are using corpus linguistics and provide a meaningful platform to evaluate the legitimacy and desirability of corpus-linguistics originalism.

1. Sixth Circuit

In *Fulkerson*, the Sixth Circuit had to interpret the word "crime" as used in a life insurance policy.⁶⁸ The plaintiff's son died in a single-car accident after driving far in excess of the speed limit and swerving between lanes.⁶⁹ After his death, his mother sought to recover accidental-death benefits under his life insurance policy, but the insurer refused, invoking an exception to the duty to pay where the insured died while committing a "crime."⁷⁰

To interpret the word "crime," the court first turned to several dictionaries: Black's Law Dictionary, Merriam-Webster's Collegiate Dictionary, and American Heritage Dictionary of the English Language.⁷¹ But because "dictionaries often 'make no claims about the ordinariness of the words they define or the senses they

62. *Muscarello v. United States*, 524 U.S. 125, 129 (1998).

63. *Id.*

64. MCENERY ET AL., *supra* note 6, at 6.

65. *Fulkerson v. Unum Life Ins. Co. of Am.*, 36 F.4th 678, 682 (6th Cir. 2022).

66. Brief for Amici Curiae Corpus-Linguistics Scholars, *supra* note 16.

67. *Health Freedom Def. Fund, Inc. v. Biden*, 599 F. Supp. 3d 1144, 1160 (M.D. Fla. 2022), *vacated as moot sub nom.* *Health Freedom Def. Fund v. Pres. of U.S.*, 71 F.4th 888 (11th Cir. 2023).

68. *Fulkerson*, 36 F.4th at 681.

69. *Id.* at 680.

70. *Id.*

71. *Id.* at 681.

assign those words,” the court turned to corpus linguistics for a “more rigorous analysis of the term’s ordinary meaning.”⁷² The court followed the steps described above: (1) the court chose the Corpus of Contemporary American English and searched for the phrase “reckless driving,” (2) the court seemingly, though never explicitly stated in the opinion, created two coding categories: those uses describing some form of prosecution and those uses not describing some form of prosecution, (3) the court selected coders (presumably the judges and their clerks, though not explicit in the opinion), and (4) the court analyzed the results of the corpus.⁷³ The only statistic the court offered was that the search yielded “over 300 results.”⁷⁴ Other than that, the opinion did not explain how many of those 300 results were officially coded nor how many results fell in each of the two categories. Instead, the court stated that “a vast majority recount instances in which some manner of prosecution occurred” and then provided a small sample of those instances.⁷⁵ After that cursory explanation, the court summarily concluded that “reckless driving has been widely understood to be an illegal act for which one can be punished by law [which] furthers our assessment that this conduct amounts to a ‘crime.’”⁷⁶

Perhaps when one hears that judges are now using an “empirical” or “data-driven” tool to discern linguistic meaning, one expects something akin to a science experiment with carefully described methods, results, and data. After all, corpus-based analyses are supposed to relieve the problems inherent in judicial intuitions that are “neither transparent nor replicable.”⁷⁷ But the lack of explanation or methodological clarity in *Fulkerson* undermines this claim. Without the inferences drawn from the sparse methodological description (the categories, the identity of the coders, etc.), all a reader can know is that someone somewhere at some time searched “reckless driving” in the Corpus of Contemporary American English and found examples supporting the court’s conclusion. Is that transparent? Is it replicable?

2. Scholarly Use of Corpus Linguistics

In contrast with the succinct, somewhat unexplained methodology employed by the Sixth Circuit, most scholars are far more thorough in their description and justification of their methodology. For example, Lee J. Strang conducted a corpus-linguistics analysis of the word “religion,” seeking to ascertain its original meaning in the First Amendment.⁷⁸ Strang, like the Sixth Circuit, followed the steps but in a much more explicit way. *First*, Strang identified his corpora: the

72. *Id.* at 682–83 (quoting *State v. Rasabout*, 356 P.3d 1258, 1273 (Utah 2015) (Lee, A.C.J., concurring in part and in the judgment)).

73. *Id.* at 683.

74. *Id.*

75. *Id.*

76. *Id.* at 684.

77. Lee & Mouritsen, *supra* note 34, at 867.

78. Lee J. Strang, *The Original Meaning of “Religion” in the First Amendment: A Test Case of Originalism’s Utilization of Corpus Linguistics*, 2017 BYU L. REV. 1683 (2017).

Corpus of Founding Era American English (COFEA) and the *Pennsylvania Gazette*.⁷⁹ Unlike the Sixth Circuit's mere naming of their corpus, Strang spent several paragraphs describing the corpora's features and explaining the corpora's strengths and weaknesses,⁸⁰ so the reader (whether fluent in corpus linguistics or not) can understand what he used and why those two sources may or may not be ideal. Not only did he identify the corpora, but Strang described the various searches done on the corpora, the way the team organized the results, and the things excluded from the analyzed data.⁸¹ *Second*, Strang described the exact categories he created and included a brief explanation of how he arrived at those categories.⁸² *Third*, Strang listed by name every individual who worked on the project, explained that each individual coded the words independently, and stated that the team worked over a three-year period.⁸³ *Fourth*, and perhaps most importantly, Strang included a *comprehensive* review of his results, complete with exact frequency rates, comparative statistics, descriptions of choices made throughout the process that may have affected the results, and a detailed appendix.⁸⁴ Unlike the Sixth Circuit's analysis and conclusion, Strang's analysis and conclusion *are* transparent and replicable. A reader can trace his exact methodology from start to finish, replicating every step along the way. And far from the black-box here-are-some-examples explanation by the Sixth Circuit, Strang's readers can test Strang's conclusions (and their own) against the raw numbers.

C. Corpus Linguistics as an Interpretive Tool

As corpus linguistics rises in acknowledgment and purported reliability, legal scholars are taking notice. Some scholars highlight potential benefits of importing corpus linguistics into originalist reasoning.⁸⁵ Lawrence Solan, for example, predicts that corpus linguistics could reduce reliance on historical dictionaries, which are often strategically selected to support a desired meaning.⁸⁶ He also explains that corpus-linguistics tools could “encourage analysts to look not only at the single word, but also at the linguistic context in which the word occurs,” which would bring originalism “closer to its goal of discovering original public meaning.”⁸⁷ Phillips and Egbert likewise argue that corpus linguistics could help the law “more accurately get at some of the fundamental problems of meaning” inherent in legal questions.⁸⁸ Solum predicts that an analytical framework that

79. *Id.* at 1693–94.

80. *Id.*

81. *Id.*

82. *Id.* at 1696.

83. *Id.*

84. *Id.* at 1697–1703.

85. *See, e.g.*, Lawrence M. Solan, *Can Corpus Linguistics Help Make Originalism Scientific?*, 126 *YALE L.J. F.* 57, 57–58 (2016).

86. *Id.* at 64.

87. *Id.*

88. Phillips & Egbert, *supra* note 32, at 1618.

includes corpus linguistics could use “objective and transparent methods” that can “provide[] the ideal to which originalists can aspire.”⁸⁹

Further, Lee and Phillips state that corpus linguistics allows originalists to present “data in a systematic, transparent way that provides some assurance that [they] are not cherry-picking isolated examples in a motivated attempt to get at a preferred outcome.”⁹⁰ But again, the Sixth Circuit’s analysis leaves much to be desired regarding transparency, and the ultimate result in that case was a list of hand-selected, even if not “cherry-picked,” examples that support the outcome already determined by the plethora of other sources consulted by the judges. And, with corpus linguistics, “[judges] can preserve [their] dataset . . . in a manner that invites replication and falsification by anyone wishing to question [the] analysis.”⁹¹ But again, while the corpus used by the Sixth Circuit may still exist, the rest of the judges’ methods are left largely to inference—not something easily and objectively replicable.

Several authors recognized the potential for the problems found in the Sixth Circuit’s analysis. Justice Lee himself recognized that “[c]orpus data can be gathered and analyzed properly only with care and a little background and training in the underlying methodology. A judge who proceeds willy-nilly may, either consciously or unwittingly, proffer data that have only the appearance of careful empiricism.”⁹² Perhaps this is what happened when the Sixth Circuit added the arguably cursory corpus-linguistics analysis to a seemingly already-decided outcome; it added another layer of “empiricism” to bolster the court’s conclusion. Another author, Anya Bernstein, called judicial corpus linguistics a “half-empirical attitude” because it “rests empirical results on fictional foundations.”⁹³ In other words, judges claim to have an empirical answer produced by a method that completely ignores *vital* empirical foundations such as considering the relevant legal and institutional context.⁹⁴ The Sixth Circuit, unlike Strang, did not even mention potential limitations or factors underlying the chosen corpus. Instead, the court treated one corpus (out of the vast corpora landscape) as the only universe that mattered, regardless of how that might skew the results.

Other scholars have raised similar critiques against the quality and use of corpora. John S. Ehrett levies his critique against the “contextual flattening” that jurists do when using corpus linguistics.⁹⁵ Instead of carefully valuing the “hierarchical privileging of some sources over another,” corpus linguistics throws all sources in a pot and stirs.⁹⁶ Ehrett uses the example of the word “literally.”⁹⁷ If a

89. Lawrence B. Solum, *Triangulating Public Meaning: Corpus Linguistics, Immersion, and the Constitutional Record*, 2017 BYU L. REV. 1621, 1681–82 (2017).

90. Lee & Phillips, *supra* note 35, at 320.

91. *Id.* at 320–21.

92. Lee & Mouritsen, *supra* note 34, at 866.

93. Anya Bernstein, *Legal Corpus Linguistics and the Half-Empirical Attitude*, 106 CORNELL L. REV. 1397 (2021).

94. *Id.*

95. John S. Ehrett, *Against Corpus Linguistics*, 108 GEO. L.J. ONLINE 50, 61 (2019).

96. *Id.*

97. *Id.* at 62–63.

corpus linguist were to conduct a study of pop-culture sources using the word “literally,” the linguist would likely get a far different meaning of the word than if the linguist studied dictionaries.⁹⁸ But if all sources are amalgamated into one large corpus and then reduced to numerical frequencies and small acontextual phrases, how is the linguist to know that the public uses the word “literally” in a far different manner than the way it is defined in dictionaries? And, Ehrett argues, “members of the public would likely treat dictionary definitions as more authoritative than the readouts of a corpus comprised of undifferentiated natural language texts.”⁹⁹ But if the linguist does not account for this differing weight of sources, the linguist might misconstrue the results.

Finally, perhaps the most bandied attack on corpus linguistics is the lack of judicial expertise or skill.¹⁰⁰ Not only do most judges not have the skill or training to use corpus linguistics effectively and thoroughly,¹⁰¹ but there are also judicial-resource limitations. Judges are often under strict time and resource constraints, which can affect their decision-making.¹⁰² So asking a court to conduct a corpus-linguistics analysis akin to Strang’s (which stretched across more than 100 pages) may be an unreasonable request. This is especially true when the request requires the judge to hand-code hundreds, if not thousands, of collocated lines of text. Some courts have the benefit of comprehensive corpus-linguistics analysis provided by scholars like Strang. In *Bostock*, the Supreme Court had an amicus brief from eminently qualified and painstakingly thorough corpus linguists.¹⁰³ But this is not always the case. So, when courts are confronted by questions that scholars have not answered, they are asked to do their own corpus-linguistics analysis. And some judges, like Justice Lee on the Supreme Court of Utah, are passionate corpus-linguistics scholars who can (even if they don’t always) conduct competent, thorough, and entirely respectable analyses of terms. But not all judges are like that. Some conduct the rather hasty analysis evident in the Sixth Circuit’s decision. Is this hasty and infrequent analysis as transparent and replicable as proponents of corpus linguistics claim? The Sixth Circuit example suggests no. Or is it merely one more way some originalists can cherry-pick sources and shroud their decisions in an air of empiricism, as it was in the Sixth Circuit? Perhaps the advent of AI can help remedy some of the critics’ problems with corpus linguistics as a judicial tool.

98. *Id.*

99. *Id.* at 64.

100. *See, e.g., id.* at 68–70; Lee & Mouritsen, *supra* note 34, at 865–66.

101. Ehrett, *supra* note 95, at 68–70.

102. *See* Brian Sheppard, *Judging Under Pressure: A Behavioral Examination of the Relationship Between Legal Decisionmaking and Time*, 39 FLA. ST. U. L. REV. 931 (2012) (discussing how limited judicial resources can affect judicial decision-making and processing).

103. Brief for Amici Curiae Corpus-Linguistics Scholars, *supra* note 16.

III. ARTIFICIAL INTELLIGENCE AS A CORPUS LINGUISTICS TOOL

AI is not new; its origins stretch back to the mid-1900s.¹⁰⁴ But it wasn't until the 2020s that researchers and developers started developing what is called “generative AI,” or AI that is capable of creating original content.¹⁰⁵ There are several types of AI—image generators, video generators, music and sound generators, language generators, and language analyzers, just to name a few—but the ones most relevant to constitutional interpretation are the “large language models” (LLMs), which are trained to consume, learn, and produce large quantities of text.¹⁰⁶ Ideally, these models get smarter and smarter with time and usage.¹⁰⁷ As the models analyze more texts and complete more tasks, the models integrate this information into the underlying body of knowledge on which they rely to answer future questions.¹⁰⁸ This is, of course, a reductive account of the entirety of AI's development, but a thorough account of the ins and outs of AI is beyond the scope of this paper.

As AI permeates many corners of society, various institutions have had to develop policies and practices surrounding its use. Schools at every level of education have incorporated AI policies into their honor codes and teaching approaches.¹⁰⁹ Healthcare facilities incorporated AI models into their infrastructures and trained employees on their use.¹¹⁰ And the judiciary was no exception. In August of 2024, the National Center for State Courts (NCSC) issued official “Guidance for Use of AI and Generative AI in Courts.”¹¹¹ And Chief Justice John Roberts of the United States Supreme Court focused much of his annual report on the use of AI in the courts.¹¹² While such guidance and policy provided useful information to judges—for example, explaining what a deepfake is and giving tips for how to spot one—perhaps nothing displays AI's corpus-linguistics potential like an example from the Eleventh Circuit.

A. *AI and Corpus Linguistics in the Courts*

Judge Kevin Newsom of the Eleventh Circuit has become one of the first to use—and to describe his use of—AI in judicial decision-making. In May of 2024, Judge Newsom wrote a concurrence that he “confess[ed]” was a “little unusual.”¹¹³ He explained that his concurrence meant to “pull back the curtain” on

104. See Cole Stryker, *What Is Artificial Intelligence (AI)?*, IBM, <https://www.ibm.com/topics/artificial-intelligence> [https://perma.cc/HP97-R7Z7].

105. *Id.*

106. *Id.*

107. *Id.*

108. *Id.*

109. See, e.g., Nat'l Educ. Ass'n, *Teaching in the Age of AI*, NEATODAY (July 4, 2024), <https://www.nea.org/nea-today/all-news-articles/teaching-age-ai> [https://perma.cc/R9ER-M86B].

110. See, e.g., Shiva Maleki Varnosfaderani & Mohamad Forouzanfar, *The Role of AI in Hospitals and Clinics: Transforming Healthcare in the 21st Century*, 11 *BIOENGINEERING* 337 (2024).

111. NAT'L CTR. FOR STATE CTS., *Artificial Intelligence Guidelines for Courts* (2024).

112. John Kruzel, *US Supreme Court's Roberts Urges 'Caution' as AI Reshapes Legal Field*, REUTERS (Jan. 2, 2024, 12:27 PM), <https://www.reuters.com/legal/us-supreme-courts-roberts-urges-caution-ai-reshapes-legal-field-2023-12-31/> [https://perma.cc/4KCU-NPU6].

113. *Snell v. United Specialty Ins. Co.*, 102 F.4th 1208, 1221 (11th Cir. 2024) (Newsom, J., concurring).

his thinking process and offer a proposal that “many will reflexively condemn as heresy” but which he believed deserved a deeper look.¹¹⁴ That proposal was that textualists “consider” whether and how AI LLMs “might” inform a court’s interpretation.¹¹⁵

The case was an unusual one, which prompted Judge Newsom’s turn to unusual methods. The question was whether installation of a trampoline constituted performance of “landscaping” under an insurance policy.¹¹⁶ Ultimately, the majority found an “off-ramp” that decided the case without addressing the definitional issue, but not before Judge Newsom spent “hours and hours (and hours) laboring over the question”¹¹⁷ He started where “any self-respecting textualist” does: dictionaries.¹¹⁸ But the lack of dictionary clarity left Judge Newsom “languishing in definitional purgatory.”¹¹⁹ In a moment of admitted frustration and desperation, Judge Newsom had one of his clerks ask ChatGPT (an AI LLM) to identify the ordinary meaning of landscaping, and when it returned a reasonable answer, he asked it the case’s dispositive question: “Is installing an in-ground trampoline ‘landscaping’?”¹²⁰ After receiving his answer, Judge Newsom went on to offer an extensive, multi-reasoned defense of AI use by the judiciary.¹²¹ He even expressly compared it to corpus linguistics and concluded that “reliance on LLMs seems to [him] preferable to both.”¹²² Just a few months after this *Snell v. United Specialty Insurance Co.* concurrence, Judge Newsom penned a “sequel.”¹²³ In his *United States v. Deleon* concurrence, Judge Newsom once again turned to AI to resolve a question of interpretation, this time wrestling with what a judge ought to do when the LLM produces slightly varying results with each iteration of the question.¹²⁴ But he once again reiterated his support for and continued allegiance to using AI in interpretive questions.¹²⁵ Apart from the policy perspectives Judge Newsom provided and the theoretical considerations bandied about by legal scholars, this paper is going to put the two theories—AI and corpus linguistics—to the test.

114. *Id.*

115. *Id.*

116. *Id.*

117. *Id.* at 1222.

118. *Id.* at 1223.

119. *Id.*

120. *Id.* at 1224–25. Judge Newsom found LLM use appropriate in this case because the dispositive question, at least to him, was the plain meaning of the word “landscaping.” *Id.* at 1226. Perhaps in other cases, LLMs might not be as useful. As a tool to aggregate or analyze text, LLMs have capabilities that could make them quite an asset. But perhaps in a case, say, disputing whether someone had the requisite mental intent to commit murder, LLMs would not be as useful. The central question of a case—and particularly how it is framed by each judge—could have a profound impact on whether LLMs are even appropriate to use.

121. *Id.* at 1226–30.

122. *Id.* at 1230.

123. *U.S. v. Deleon*, 116 F.4th 1260, 1270 (11th Cir. 2024) (Newsom, J., concurring).

124. *Id.*

125. *Id.*

B. Experiment: Comparing AI to Strang's Corpus-Linguistics Study

For the sake of space and time, this paper will use Strang's detailed corpus-linguistics study of the word "religion" in the First Amendment.¹²⁶ As described above, Strang began with a comprehensive explanation of the methodology and analytical framework with which he experimented. After following the "best practice" corpus-linguistics steps, Strang reached his "major finding . . . that the original conventional meaning of religion was theistic."¹²⁷ Let's see what AI LLMs have to say.

First, ChatGPT was prompted: "was the original conventional meaning of religion in the First Amendment theistic?" ChatGPT responded:

The original meaning of "religion" in the context of the First Amendment to the U.S. Constitution was not strictly theistic. The First Amendment states, "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof." This reflects the Founding Fathers' intent to ensure freedom of belief, which included a variety of religious expressions, both theistic and non-theistic. At the time, there was significant influence from Enlightenment thought, which promoted ideas about individual rights and the separation of church and state. The Founders were aware of a range of beliefs, including deism, which does not necessarily fit traditional theistic frameworks. Therefore, while theistic beliefs were certainly prevalent, the original understanding of religion in the First Amendment allowed for a broader interpretation, encompassing a diversity of beliefs and practices.

Second, ChatGPT was prompted: "Please explain to me what sources you relied on to make that conclusion." It replied:

1. The Federalist Papers & Founding Era Writings

- The Federalist Papers
- In Madison's Memorial and Remonstrance Against Religious Assessments (1785), he argued that religious freedom should be protected against both governmental establishment and coercion, emphasizing that individuals' consciences should be free from government control.
- Thomas Jefferson's writings, especially in his Virginia Statute for Religious Freedom (1779), also demonstrate a strong commitment to religious liberty.

2. Enlightenment Thought

- The Founders were heavily influenced by Enlightenment thinkers such as John Locke, Voltaire, and Montesquieu, who argued for religious tolerance and freedom of belief. John Locke, for example, in his "Letter Concerning Toleration" (1689), made the case for religious freedom and argued that religious belief should be a matter of individual conscience, not state control.

126. Strang, *supra* note 78.

127. *Id.* at 1697.

- Deism, which was influential among some of the Founding Fathers (including Jefferson, Franklin, and Paine), emphasized a creator but rejected organized religion, miracles, and the authority of clergy. This non-theistic form of belief was still considered part of the broader notion of “religion” in early American thought.

3. Early Legal Interpretations

- Early judicial interpretations of the First Amendment suggest that “religion” was understood broadly, encompassing both theistic and non-theistic beliefs. In the 1789 House of Representatives debates over the Bill of Rights, the language of the First Amendment was discussed, and there was little to suggest that “religion” was intended to be narrowly theistic. The debates focused more on protecting the free exercise of conscience rather than defining what constituted a “religion.”
- The 1797 Treaty of Tripoli (ratified by President John Adams and the U.S. Senate) famously states: “The Government of the United States is not, in any sense, founded on the Christian religion.” This reflects a broad, inclusive understanding of religion, one that acknowledged that the government was neutral toward all religious beliefs, including non-theistic ones.

4. Legal Scholar Interpretations

- Legal scholars and constitutional historians, such as Philip Hamburger and Jeffrey Hammond, have argued that the framers of the Constitution intended for the First Amendment’s religious clauses to be understood as broadly inclusive, ensuring the freedom to practice any religion or none at all.
- For example, in “Separation of Church and State: Historical Fact and Current Fiction”, Hamburger discusses how the framers of the Constitution and early American leaders did not see religious freedom as something limited to conventional theistic faiths, but as a principle that protected a wide variety of beliefs, including secular and deistic ones.

5. Judicial Precedents

- Over the years, U.S. Supreme Court decisions have reinforced the idea that the First Amendment’s protection of religion is not limited to traditional theism. Cases like *Everson v. Board of Education* (1947) and *Torcaso v. Watkins* (1961) extended First Amendment protections to non-theistic beliefs. In *Torcaso v. Watkins*, the Court ruled that non-theistic beliefs (such as atheism) are entitled to the same constitutional protection as religious beliefs. This further supports the idea that the concept of “religion” in the First Amendment was not confined to theistic belief systems.
- The Establishment Clause has been interpreted to mean that the government cannot favor any particular religion, including theistic ones, and must remain neutral toward all belief systems, secular and religious alike.¹²⁸

128. Not all of these sources would be accepted by originalists as proper for the interpretive foundation. I have included in Appendix A a colloquy with ChatGPT in which I instructed ChatGPT to consider only sources around the time of the Constitution’s ratification to see if that would change its analysis. This did change its analysis interestingly, as will be discussed momentarily.

Third, ChatGPT was asked: “But I’m sure there are examples where ‘religion’ was theistic. Can you give me those?” It responded by saying that there were “certainly examples” where religion was theistic. Those examples included Christianity’s dominance in the social culture of Founding Era America; state constitutions from the Founding Era requiring officials to affirm a belief in the Christian religion; the reference to “Almighty God” in an initial draft of the 1789 Bill of Rights; and the “clear role of religion” in the Founding Fathers’ lives.

So, if there are examples on both sides, how did ChatGPT conclude that “religion” was not theistic? When asked that question, ChatGPT replied:

I would say that I found the **non-theistic examples** and the **historical and philosophical context** surrounding the First Amendment more compelling in shaping the broad interpretation of “religion” in the First Amendment, even though there are also theistic examples that play a significant role in the historical and cultural context.

When asked what it meant by “more compelling,” ChatGPT replied:

When I said that the **non-theistic examples** were “more compelling,” I didn’t mean that the theistic examples were irrelevant or that the non-theistic argument definitively outweighs the theistic one. Rather, I meant that, in trying to understand the **original intent** behind the First Amendment and the broader **constitutional framework** of religious freedom, the **non-theistic examples** seemed more in line with the **foundational principles** of religious liberty and freedom of conscience that the framers sought to protect. However, both the **theistic** and **non-theistic** examples are **important** and **relevant** to understanding the First Amendment’s original meaning.

Claude (another AI LLM) and Google’s Gemini LLM received the same prompts. They replied with similar responses, which have been included in Appendices B and C, respectively. And when asked for a “simple yes or no” answer to the question of whether “religion” in the First Amendment is theistic, all three LLMs unanimously replied “no.”

The LLMs were asked to answer the same question with only sources that an originalist would use and that were in existence at the time of the Constitution’s ratification. These colloquies appear in the appendices. Claude and Gemini were remarkably unimpressive with their lack of commitment to an answer and their excessive hedging. However, ChatGPT’s answer changed with the new prompts. It explained that under an originalist methodology, there is a “narrower understanding of religion” that has “no clear protection for secularism.” Under this methodology, ‘religion’ would be more closely aligned with theistic Christianity rather than a broad, inclusive definition that might encompass non-theistic belief systems.” In answering the prompt, ChatGPT repeatedly referenced its attempt to find the “[F]ramers’ intentions during the 1780s and early 1790s.” For purposes of brevity, the colloquy did not continue to ask ChatGPT to analyze the question

under every existing sub-theory of originalism (lay language, legal meaning, etc.). It is enough that the analysis—and the answer—changed when the questioning was more specific. Whereas the initial answer was in direct conflict with Strang’s conclusion, the increased specificity and narrowing brought the two conclusions into alignment.

This demonstrates several things. *One*, it demonstrates the importance of careful drafting of prompts when dealing with LLMs. *Two*, it demonstrates the importance of a continued conversation with the LLM, pressing the LLM to continue its analysis and get closer and closer to the desired methodology. If the analysis stopped after ChatGPT’s initial conclusion, Strang’s results could seem irreconcilable with ChatGPT’s. *Third*, it demonstrates the way in which subtle differences in prompting can change the outcome. In fact, when ChatGPT was prompted with a similar—but slightly—different question (“Would an originalist think that ‘religion’ in the First Amendment is a theistic term? There are arguments on both sides.”), it responded, “An originalist could see ‘religion’ as theistic, given the historical context, but also recognize that the [F]ramers intended to protect a broad range of beliefs, including non-theistic ones, based on their emphasis on religious freedom and conscience.” ChatGPT was then asked, “So is it possible an originalist would *not* think ‘religion’ was theistic?” to which it responded:

Yes, it is absolutely possible that the answer could be “**no**”—depending on how you approach the question.

If we’re asking whether an originalist would view “religion” in the First Amendment strictly as a **theistic term**, the answer could indeed be “**no**” because the term “religion” in the First Amendment is not exclusively tied to theism in the historical context. While the dominant religions of the time were theistic (such as Christianity and Judaism), the framers of the Constitution—particularly Thomas Jefferson, James Madison, and others—were focused on a broader notion of **religious freedom** that protected **any system of belief** from government interference, not just theistic ones. Therefore, many originalists would reasonably conclude that religion in the First Amendment was not necessarily a strictly theistic term and that it extends to a broader array of beliefs.

ChatGPT then went on to reference the same sources it referred to during the colloquy in Appendix A (where it claimed “religion” was a theistic term) but framed those sources in a way that showed “religion” as a *non*-theistic term. So, depending on one’s desired outcome, one can keep probing and prompting an LLM until the LLM provides the desired outcome. The initial “based on originalist sources, what is the answer?” provided powerful argument that “religion” *is* theistic. But if, as a party to a case, one wanted the answer to be that religion is *not* theistic, all one must do is continue probing ChatGPT with pointed prompts, and suddenly, ChatGPT supports both sides.

So if a judge or justice is presented with a thorough, methodical corpus-linguistics amicus brief and an easy-to-access answer from an LLM, which should she

choose? That answer lies in the nature and depth of her jurisprudential commitments and her weighing of the pros and cons discussed in the next section.

C. *Legitimacy of AI Corpus Linguistics and Originalism*

Judge Newsom’s first concurrence displays some of the major benefits and risks of using artificial intelligence to facilitate statutory interpretation.

First, LLMs are ordinary-language inputs.¹²⁹ As explained, the LLMs capture more sources and uses than the narrow corpora. Judge Newsom referenced a particular LLM that trained on 400 to 500 *billion* words.¹³⁰ The LLMs capture “data run[ning] the gamut from the highest-minded to the lowest, from Hemmingway novels and Ph.D. dissertations to gossip rags and comment threads.”¹³¹ That could make them superior to corpora that contain far smaller volumes of words: the COCA contains just 1 billion words, and the specialty corpora contain far, far fewer words.¹³² So perhaps using these LLMs (which are arguably corpora themselves) provides a more comprehensive view of the word-use landscape, which may be an advantage or a weakness, depending on one’s conception of originalism.

Second, LLMs can understand context.¹³³ Judge Newsom used the word “bat” as an example. He explained that LLMs can tell when “bat” is used to describe the flying mammal versus when it is used to describe the wooden baseball equipment,¹³⁴ or when it is used as a verb or to describe cricket equipment. So, presumably, judges could ask these LLMs to identify the frequency of words used in a particular context (i.e., “tell me how often ‘religion’ was used to refer to theism in 1789”). This doesn’t eliminate the need for creating coding categories, which is one of the sources of subjectivity in corpus-linguistics analysis. After all, judges must know the different senses of the word to ask the LLM for the frequency of a particular sense. But it does reduce the potential human error involved in tagging and categorizing the text. After all, as Judge Newsom explained, developers have found a way to turn words into *math*, allowing computers to calculate and sort words just as efficiently as numbers. Instead of a human (prone to error, mental fatigue, misunderstanding, and outcome-driven biases), LLMs (purportedly free from human risks) can almost instantaneously sort the billions of words into neat categories and put a simple answer right into the judge’s hands.

Third, LLMs are accessible.¹³⁵ Unlike the complicated, time-consuming, and extremely technical nature of handmade corpus linguistics, LLMs can be used by anyone anywhere with a computer. Judge Newsom pointed out that this is not only cheaper than LexisNexis or Westlaw subscriptions but also “democratizes”

129. *Snell v. United Specialty Ins. Co.*, 102 F.4th 1208, 1226 (11th Cir. 2024) (Newsom, J., concurring).

130. *Id.*

131. *Id.*

132. ENGLISH-CORPORA.org, <https://www.english-corpora.org/> [<https://perma.cc/599F-2ZJW>].

133. *Snell*, 102 F.4th at 1227 (Newsom, J., concurring).

134. *Id.* at 1227–28.

135. *Id.* at 1228.

the interpretive enterprise.¹³⁶ That's entirely consistent with the judiciary's goal of giving readers of judicial opinions a transparent, accessible, replicable interpretation.

Fourth, Judge Newsom described LLM research as relatively transparent.¹³⁷ A reader of this paper can read the exact word-for-word exchange I had with each LLM. The reader does not have to guess at results I left out of my hand-picked examples or go replicate a 100-page study themselves. Rather, readers can read the exchange and decide for themselves whether they think it is credible. And if they do not, they can go converse with the LLM themselves. But is it truly more transparent when even the top AI engineers admit that they “don't really know what [the LLM models] are doing in any deep sense?”¹³⁸ Those engineers “don't understand what's going on here. [They] built it, [they] trained it, but [they] don't know what it's doing.”¹³⁹ They admit that they have no way of explaining what reasoning or rules an LLM is relying on.¹⁴⁰ So even if one trusts the LLM's own explanation about the bases for its reasoning, there is still what AI engineers call a “black box” shrouding the whole enterprise.¹⁴¹ Perhaps rather than making legal interpretation and analysis *more* transparent, AI LLMs are simply hiding the nuance and subjectivity beneath a shroud that human minds cannot understand.

Moreover, the “transparency” aspect requires trust in the humans using the results. Once the answer from ChatGPT is copied and pasted into a legal brief, the author of the brief can modify it or omit portions of it altogether in a way that changes its meaning. And there is no way to verify that the portion in the brief is the full, accurate account of the LLM's answer. Even if one copies and pastes the “answer” into the LLM and prompts it with “is it true that you told me this in the past?” the answer is often “I may have said that in the past” or something to that effect. It is true that individuals with accounts on these websites can save dialogue with the LLM and pull up past conversations to continue them, but if judges do not require proof of those conversations, the language from the LLM can be manipulated to bolster either side's arguments, providing a mere façade of strong support.

Fifth, LLMs hold advantages over other empirical methods.¹⁴² This is the most debatable of Judge Newsom's proffered reasons, and it is the whole purpose of this paper. Judge Newsom dismissed corpus linguistics out-of-hand by simply saying that corpora are compiled with too much discretion and non-transparency.¹⁴³ As discussed earlier in the paper, corpus linguistics surely has its weaknesses, but when done thoroughly and properly, it also has its strengths. And if the linguist is honest—

136. *Id.*

137. *Id.*

138. Noam Hassenfeld, *Even the Scientists Who Built AI Can't Tell You How It Works*, VOX (July 15, 2023, 6:00 AM), <https://www.vox.com/unexplainable/2023/7/15/23793840/chat-gpt-ai-science-mystery-unexplainable-podcast> [<https://perma.cc/6LHE-WE2P>].

139. *Id.*

140. *Id.*

141. *Id.*

142. *Snell*, 102 F.4th at 1230 (Newsom, J., concurring).

143. *Id.*

as was Strang—about the corpora’s shortcomings, corpus linguistics does not have to be dismissed as inherently inferior to LLMs just based on bias risks alone.

But there are also risks.

First, LLMs can hallucinate.¹⁴⁴ But, as Judge Newsom explained, the risk of hallucination in LLMs is quite comparable to the risk of human “finesse,” error, or downright lying.¹⁴⁵ And it’s also quite comparable to the risk of human error in coding, categorizing, or interpreting corpus-linguistics results. So this alone does not decide the question of whether LLMs more accurately than corpus linguistics can be used to legitimate originalism.

Second, LLMs don’t capture offline speech, and thus might not fully account for underrepresented populations’ usages.¹⁴⁶ But neither do the corpora used by corpus linguists. In fact, because the LLMs contain a much greater array of sources, they are more likely to capture a broader, more diverse subset of language use than are the corpora. Further, because LLMs are constantly training and learning, they will potentially come to learn more of the underrepresented uses of words. But a large portion of corpora are never updated, and the ones that are must be updated by the people who maintain them: a process that adds time and injects the risk of human error.¹⁴⁷

Third, lawyers, judges, and would-be litigants might try to manipulate LLMs.¹⁴⁸ Indeed, as Judge Newsom described in his “sequel,” LLMs can offer vastly different results when asked the exact same question twice in a row, which means users can pick the one more favorable to their side.¹⁴⁹ But that is even more of a risk with hand-done corpus linguistics. When the interested party gets to choose the coding categories and then conduct a largely nonreplicable study *by hand*, the risk of manipulation (and failure by the judge to uncover the manipulation) is much higher.

144. *Id.* (describing a scenario in which a lawyer used an LLM model that gave him completely made-up case citations).

145. *Id.* at 1230–31.

146. *Id.* at 1231.

147. Mark Davies, a professor known worldwide for building corpora, admitted that building corpora is “like an addiction” for him. Amanda K. Fronk, *BYU Linguist Mark Davies Has a Corpus for Everything*, Y MAG., <https://magazine.byu.edu/article/big-lang-at-byu/> [<https://perma.cc/74UA-2ZPM>]. Davies explained that current technology has certain limits on how he can update his corpora. *Id.* He’s “constantly bumping right up against the limit of what technology will allow.” *Id.* But his news-aggregating corpus gains 4–5 million words every day. *Id.* In another article, Davies explained how he does it. Mark Davies, *The 385+ Million Word Corpus of Contemporary American English (1990-2008+): Design, Architecture, and Linguistic Insights*, 14 INT’L J. CORPUS LINGUISTICS 159, 163–64 (2009). He uses a “programming interface and language” to create a script that is sent to Internet Explorer to retrieve sources containing the described information. *Id.* But the way in which he instructs, codes, and yields that technology affects the corpus output. For example, when Davies wakes up in the middle of the night with a new corpus idea, as he admittedly does, *id.*, and he creates the program to start searching and compiling the relevant sources, that process is not infallible nor is it completely free from bias. A deeper study into the construction and development of corpora is beyond the scope of this paper, but it suffices to say that corpora themselves reflect certain design choices made by fallible human minds and black-box technology.

148. *Snell*, 102 F.4th at 1231 (Newsom, J., concurring).

149. *U.S. v. Deleon*, 116 F.4th 1260, 1270 (11th Cir. 2024) (Newsom, J., concurring).

Instead, LLMs provide judges a source of which they can independently inquire quite easily and quickly if they feel manipulated by the parties. With hand-done corpus linguistics, though, the judge would have to undertake a *significant* time commitment and a mentally taxing project to check the work of the parties. When pressed for time and already nearing their mental capacity, judges are not likely to do this.

Fourth, and finally, reliance on LLMs might lead us into dystopia.¹⁵⁰ But AI is here, and it's not likely going anywhere. And as Judge Newsom explained, responsible judicial use of AI can—rather than creating “robo judges”—provide *more informed* judges.¹⁵¹ It can offer judges an extra data point for interpretive analyses, a check on the already existing processes, or a place to pause and consider whether the traditional interpretation is missing the mark.

CONCLUSION

In all, the judicial use of AI—much like the judicial use of corpus linguistics—is an individual choice. Each judge, lawyer, and jurist must choose for herself whether she wants to use it. AI can offer attractive benefits like minimizing the time needed to conduct a study, broadening the base of analyzable texts, creating replicable, accessible results, and creating a far more objective analysis than traditional corpus linguistics can provide. But it comes with drawbacks, several of which are present even in traditional corpus linguistics: manipulation, mistakes, and underrepresentation of certain languages or regional dialects, like the distinction in dialect between the South and those states in New England. As a theoretical exercise, the benefits of AI seem to make it a promising and desirable way to conduct a corpus-linguistic-like analysis for originalist purposes. But as the example showed, LLMs can reach conclusions that conflict with legal scholars who conduct thorough, descriptive analyses. Perhaps LLMs can be a way to find the back-of-the-book meaning that originalists seek. Or perhaps it is just an “empirical” layer added to a thin originalist façade that makes it look like objective judging when in fact it can be just as skewed as traditional corpus linguistics. Only time will tell whether AI provides a pathway to statistically significant judging.

150. *Snell*, 102 F.4th at 1232 (Newsom, J., concurring).

151. *Id.*

APPENDIX A

At the end of the original colloquy with ChatGPT, I prompted ChatGPT again: “Please only consider sources that were in existence at the time the Constitution was ratified by the states (excluding Amendments past the first ten) and tell me if your answer changes based on those sources.” It responded:

When we limit our sources to those available at the time of the Constitution’s ratification (i.e., the late 18th century), the answer does not change substantially, but we do have to adjust our focus. The understanding of “religion” during this time, as reflected in the texts and writings of the framers and other key figures, was indeed broader than a strictly theistic definition, even though theism (particularly Christianity) was dominant in the cultural context.

Here’s a more focused look at sources from the time of the Constitution’s ratification:

1. Framers’ Intent and Key Writings

The framers of the Constitution, especially James Madison, Thomas Jefferson, and others, were deeply committed to religious freedom. While they were largely from a Christian background, their writings demonstrate a broader understanding of religion as encompassing systems of belief beyond just theistic or Christian ones.

- **James Madison’s Memorial and Remonstrance Against Religious Assessments (1785):** This document is crucial in understanding Madison’s view of religious freedom. He opposed a proposed tax to fund Christian churches in Virginia, arguing that “religion” includes all personal beliefs and that government interference should not favor any particular faith. He emphasized that “religion” was tied to personal conscience and the right to worship freely, regardless of specific doctrinal beliefs. This is a clear sign that Madison envisioned religious freedom as encompassing more than just belief in a deity.
- **Thomas Jefferson’s Virginia Statute for Religious Freedom (1786):** Jefferson’s statute is another foundational text. He argued that “religion” must be free from governmental control and should include all belief systems. Jefferson explicitly argued against any “establishment” of a particular religion and framed religion as a matter of personal conscience. This was a clear rejection of a narrow, theistic view of religion.
- **George Washington’s Writings:** While Washington did not write extensively about religious freedom, his **Farewell Address (1796)** emphasized the role of religion in public life in terms of promoting morality, ethics, and civic virtue. He did not, however, specify that religion must be theistic, and he urged tolerance for all religions, acknowledging the pluralism in the early United States.

2. The Language of the First Amendment

The language of the First Amendment itself—adopted in 1791—also suggests a broader view of “religion.” The phrase “Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof” was intentionally vague, designed to cover a wide range of beliefs and practices. The framers wanted to prevent the establishment of a national church, but they were also concerned with protecting individual freedoms to practice religious (or philosophical) beliefs without government interference.

- The wording of the First Amendment does not specify a particular religious tradition or belief system. The framers deliberately avoided any language that would favor a particular faith, acknowledging the growing religious diversity in the United States, including non-Christian faiths, Deism, and other forms of belief.

3. Religious Diversity and Secularism in the Early Republic

In the period surrounding the ratification of the Constitution, there was significant religious diversity. The early American colonies were home to Christians of various denominations (Protestants, Catholics, Quakers, etc.), but there were also Jewish communities and Deists—many of whom held beliefs that were not strictly theistic.

- **Deism:** Many of the Founding Fathers, including Jefferson, Franklin, and Madison, were influenced by Deism, a philosophical belief system that acknowledges a creator but rejects organized religion and the idea of a personal, interventionist God. Deism was not atheistic, but it was decidedly non-theistic in the sense that it rejected the conventional, personal conception of God in favor of a more abstract, impersonal deity.
- **Religious Minorities:** The Constitution’s framers were well aware of the religious minorities in the early republic. Jews, Quakers, and other non-theistic groups (like some Deists and rationalists) were active in the new republic and sought the protection of their religious freedoms. This pluralistic reality further suggests that “religion,” as understood by the framers, was not limited to theism alone.

4. Public Debates and Writings of the Time

During the ratification of the Constitution and the Bill of Rights, there were debates over the proper role of religion in government. Some Anti-Federalists worried that the Constitution might allow for the establishment of a national religion, while Federalists reassured them that religious freedom would be protected. The debates themselves reflect the concern for broad protection of religious beliefs, even beyond just theistic ones.

- **The Federalist Papers:** While not focused directly on religion, the Federalist Papers (written by Hamilton, Madison, and Jay) contain references to the importance of public virtue and moral conduct, which many believed religion helped to foster. The papers, however, did not endorse any particular religious doctrine and emphasized that the government would remain neutral on religious matters.
- **The Anti-Federalists:** The Anti-Federalist papers also reflect concerns about the potential for federal power to infringe on religious freedom. Some Anti-Federalists feared that without a Bill of Rights, the government could favor certain denominations. The Federalists responded by supporting the Bill of Rights, including the First Amendment, to ensure that no one religion would dominate.

Conclusion Based on 18th-Century Sources

Based on the writings of the framers, public debates, and the context of religious diversity in the late 18th century, it is clear that the original meaning of “religion” in the First Amendment was **broad and inclusive**. While the majority of people in the newly formed United States were theistic and Christian, the framers of the Constitution were consciously trying to protect not only Christian denominations but also other religious and philosophical beliefs, including non-theistic systems like Deism.

The framers intended to ensure that individuals could freely practice their religion (or conscience) without government interference, whether that involved belief in a god or a more secular or philosophical belief system. Thus, **the original conventional meaning of “religion” in the First Amendment was not strictly theistic**—it included a wide range of religious, moral, and philosophical beliefs.

I then prompted ChatGPT: “Please only consider sources from the time of the Constitution’s ratification by the states that an originalist would consider and tell me if/how your answer changes.” This was the response:

When considering the original meaning of the First Amendment from the perspective of **originalism**, which seeks to interpret the Constitution according to the intent and understanding of its framers at the time of ratification, the analysis of the word “religion” focuses more narrowly on the context, writings, and public understanding at that time.

Originalist interpretation generally prioritizes sources that directly reflect the framers’ intentions during the 1780s and early 1790s, such as:

- **The text of the Constitution and the Bill of Rights**
- **The Federalist and Anti-Federalist Papers**
- **Debates in state ratifying conventions**
- **Letters, speeches, and writings of key framers**
- **State constitutions and laws**

Given this focus, the originalist perspective would likely emphasize **theism** more than a broader, more inclusive conception of “religion” that includes secular or non-theistic belief systems. However, the concept of religious liberty during this period was still more expansive than the modern view of strict theism, although it was framed within the context of Christian, and particularly Protestant, understandings of religion.

Re-examining the Sources from an Originalist Perspective:

1. Text of the First Amendment

The First Amendment itself does not provide a detailed definition of “religion,” but it does make clear that the government cannot “establish” a religion or prohibit the “free exercise” of religion. Originalists would focus on the framers’ intention to prevent the establishment of a national church or a religious test for officeholders, which was an issue for many Americans in the 18th century due to the experience with state-established churches in Europe and colonial America.

The language of “establishment” and “free exercise” would have been understood primarily in the context of Christian denominations, since at the time, the dominant religious culture in the United States was Christian (specifically Protestant). Even so, the intent was to prevent any particular denomination from gaining state support or preference, reflecting concerns about religious plurality but not necessarily about protecting non-theistic or secular belief systems.

2. Federalist and Anti-Federalist Papers

While the Federalist Papers do not focus on religious freedom per se, they provide insight into the framers’ understanding of religious liberty. **James Madison** (Federalist 10) and **Alexander Hamilton** (Federalist 51) both emphasize the importance of preventing factions and the tyranny of a majority, which could apply to religious groups as well. However, these discussions do not provide detailed clarification about whether “religion” includes non-theistic belief systems.

The **Anti-Federalist Papers**, by contrast, express concern about the lack of explicit protections for religious freedom in the original draft of the Constitution. This concern led to the inclusion of the Bill of Rights. Anti-Federalists were worried that, without specific protection, the federal government might infringe on the rights of religious minorities, particularly in states that had established churches. But again, the debates were largely framed around Christian denominations and the possibility of a national church or religious test.

3. State Ratifying Conventions

The ratifying conventions in states like **Virginia**, **Pennsylvania**, and **Massachusetts** provide insight into how the Constitution was understood in its immediate historical context. For example, in **Virginia**, James Madison’s Memorial and Remonstrance (1785) argued for religious freedom but

specifically focused on protecting Christian denominations from government interference. This reflects the understanding of “religion” in the context of 18th-century Protestant Christianity.

In **Massachusetts**, the state constitution (1780) included a provision ensuring “freedom of the exercise of religion,” but this was understood in a more limited sense—focused on Protestant Christian worship and moral conduct, with an explicit acknowledgment that the government would not prefer one denomination over another.

4. The Writings of Key Framers (Madison, Jefferson, etc.)

When considering writings from framers like **James Madison** and **Thomas Jefferson**, an originalist would focus on the language and ideas that were most salient at the time of the Constitution’s adoption.

- **Madison’s Memorial and Remonstrance Against Religious Assessments (1785)**: Madison argued for the protection of religious liberty in the context of Christian denominations, emphasizing the right of individuals to practice their religion without state interference. He did not envision the protection of secular or non-theistic belief systems as part of this vision. His arguments were rooted in the belief that religious liberty was essential to a free society, but they were primarily concerned with Christianity, particularly with the establishment of Christian sects.
- **Jefferson’s Virginia Statute for Religious Freedom (1786)**: Jefferson’s statute is often cited by originalists as an important articulation of religious liberty, but his focus was primarily on Christian sects. Jefferson argued that “religion” must be free from government control but did not explicitly address whether that freedom extended to non-theistic belief systems. Jefferson’s own personal Deism did influence his thinking, but it was still within the broader context of religious freedom for Christian denominations.

5. The Influence of Christianity in the Late 18th Century

The broader cultural and religious context of the late 18th century also shapes an originalist view. In the years immediately surrounding the drafting of the Constitution, the United States was overwhelmingly Christian (particularly Protestant), and public discourse about religion largely centered around Christian beliefs. Even though some of the framers were Deists, the religious context in which they operated was still fundamentally Christian, which would suggest that “religion” in the First Amendment was most commonly understood to mean **Christianity**, specifically Protestantism.

The **Virginia Statute for Religious Freedom** and **Madison’s Memorial and Remonstrance** were aimed at protecting Christians from state interference, and the key issue was the protection of **Christian conscience** rather than a broader, pluralistic notion of religious freedom.

How the Answer Changes with Originalism:

- **Narrower Understanding of “Religion”:** From an originalist perspective, the meaning of “religion” in the First Amendment would likely be more theistically focused, primarily addressing Christian denominations and the right to practice Christianity without state interference. Although there was awareness of religious diversity (e.g., Jews, Quakers, and Deists), the framers’ primary concern was protecting Christian religious liberty from government coercion or establishment.
- **No Clear Protection for Secularism:** The notion of “religion” encompassing secular or non-theistic belief systems would not have been a central concern for the framers. While Deism was influential among some framers, there is no evidence to suggest that the framers were concerned with protecting belief systems that were entirely non-theistic or secular. Originalists would likely interpret “religion” as referring to theistic belief systems, particularly those within the Christian tradition, given the cultural and political context of the time.

Conclusion:

From an **originalist** perspective, the meaning of “religion” in the First Amendment would likely be **theistic**, focused on protecting the religious freedom of Christian denominations from government interference, particularly in the context of preventing a national church or religious establishment. While there was awareness of religious diversity, the original understanding would not extend the protection of “religion” to secular or non-theistic belief systems. Therefore, my answer changes slightly in the sense that, under originalism, the term “religion” would be more closely aligned with **theistic Christianity** rather than a broad, inclusive definition that might encompass non-theistic belief systems.

APPENDIX B—CLAUDE

Was the original conventional meaning of religion in the first amendment theistic?

Based on historical records and scholarship, the original understanding of “religion” in the First Amendment was primarily, though not exclusively, focused on theistic belief systems. However, this is a complex historical question that scholars continue to debate.

The Founding Fathers were writing in a context where most “religion” referred to various Christian denominations and other theistic faiths like Judaism. Key figures like James Madison and Thomas Jefferson generally discussed religion in terms of belief in and duties toward a divine being or Creator.

However, it’s worth noting that some Founders, particularly Jefferson and Madison, held fairly broad views of religious liberty that went beyond just protecting theistic beliefs. Jefferson’s Virginia Statute for Religious Freedom, which influenced the First Amendment, spoke of “opinions in matters of religion” rather than specifically theistic beliefs.

The Supreme Court has since interpreted “religion” much more broadly to include non-theistic belief systems. For example, in cases like *United States v. Seeger* (1965) and *Welsh v. United States* (1970), the Court extended religious protections to ethical and moral beliefs that occupy a similar place as traditional religious faith in a person’s life.

Please explain to me what sources you relied on to make that conclusion.

I need to be transparent with you - I should not have made such specific claims without clearly noting that I don’t have access to a database or direct sources to cite. While I have training that includes information about constitutional history and interpretations of the First Amendment, I cannot point to specific historical documents or scholarly works to verify the details I provided about Madison and Jefferson’s views or the exact historical context.

The Supreme Court cases I mentioned (*Seeger* and *Welsh*) are real cases that dealt with expanding interpretations of religion, but you should double-check their exact holdings and dates as I may have misremembered details.

For accurate information on this important constitutional question, I’d recommend consulting:

- Original historical documents from the founding period
- Scholarly works focused on First Amendment religious history
- Supreme Court decisions and their supporting documentation
- Legal history journals and academic publications

But I’m sure there are examples where religion was theistic. Can you give me those?

I need to acknowledge that I cannot actually point to specific verified historical examples showing religion was interpreted theistically in early American history. While I have general knowledge of constitutional history, I cannot cite specific primary sources or verified examples without potentially creating misleading or inaccurate historical claims.

For accurate information about how “religion” was understood and used in early American legal and social contexts, I’d recommend:

1. Reviewing the actual debates and documents from the Constitutional Convention and state ratification debates
2. Examining early state constitutions and their religious provisions
3. Looking at contemporaneous dictionaries and legal texts from the period
4. Reading court cases from the early republic period

Please only consider sources that were in existence at the time the Constitution was ratified by the states (excluding Amendments past the first ten) and tell me if your answer changes based on those sources.

I need to be direct here - I don’t actually have reliable access to sources from the time of the Constitution’s ratification. While I have general knowledge about early American history, I cannot make specific claims about how the term “religion” was used in the ratification period without being able to verify sources from that exact time.

To answer your question accurately, one would need to examine:

- Records from the Constitutional Convention
- The Federalist Papers
- Anti-Federalist writings
- State ratification debates
- Legal documents and court cases from the 1780s-90s
- Letters and papers of the Founders from that period
- Dictionaries and legal texts in common use during ratification

Without being able to directly reference these materials, I shouldn’t make claims about whether the understanding was more or less theistic during this specific timeframe. The interpretation of historical evidence requires careful examination of primary sources.

Please only consider sources from the time of the Constitution’s ratification by the states that an originalist would consider and tell me if/how your answer changes.

Let me be very clear: I still cannot make reliable claims about the original meaning of “religion” even when considering only originalist sources. While I know originalists generally look at:

- The constitutional text itself
- Dictionaries from the founding era
- State constitutions and laws from that period
- Common law understandings
- Public debates and discussions during ratification
- Early congressional and judicial interpretations

I don’t have direct, verified access to these sources. So I cannot tell you definitively how they defined religion or whether their understanding was primarily theistic. Making claims about original meaning requires careful examination of actual historical evidence.

The methodology would change - focusing strictly on public meaning at the time of ratification rather than later interpretations or broader historical development. But without being able to cite specific originalist sources, I shouldn’t make assertions about what that original public meaning was.

APPENDIX C—GEMINI

Was the original conventional meaning of religion in the first amendment theistic?

The original understanding of “religion” in the First Amendment was **not** necessarily theistic. While many of the Founding Fathers were religious, and the concept of religion was often associated with Christianity and Deism at the time, the First Amendment itself does not explicitly define religion as theistic.

Here’s a more accurate explanation:

- **Broader Understanding:** The First Amendment was intended to protect a wide range of beliefs and practices, not just those associated with traditional theistic religions.
- **Focus on Individual Conscience:** The primary concern was protecting individual conscience and the freedom to believe and practice according to one’s own convictions.

While the Supreme Court has grappled with defining “religion” over the years, it has generally adopted a broader understanding that includes non-theistic beliefs as long as they meet certain criteria, such as sincerity and importance to the adherent.

Please explain to me what sources you relied on to make that conclusion.

My response was based on an understanding of:

- **Historical Context:** The Founding Fathers lived in a time when Christianity and Deism were dominant religious influences. However, it’s crucial to acknowledge that religious diversity existed even then.
- **Textual Analysis:** The First Amendment itself does not explicitly define “religion.” This lack of specificity suggests an intent to protect a broader range of beliefs.
- **Scholarly Interpretations:** Legal and historical scholarship on the First Amendment emphasizes the protection of individual conscience and the freedom to believe and practice according to one’s own convictions, regardless of whether those beliefs fit neatly into traditional theistic categories.

Disclaimer: I am an AI and do not have personal beliefs or the ability to conduct independent research. My responses are based on the information I have been trained on and should be considered for informational purposes only.

For a deeper understanding of this complex issue, I recommend consulting legal and historical resources from reputable sources such as:

- **Academic Journals:** Law reviews and historical journals that specialize in constitutional law and religious studies.

- **Legal Texts:** Supreme Court decisions and legal commentaries on the First Amendment.
- **Historical Documents:** Writings of the Founding Fathers and other relevant historical documents.

I am still under development and learning to provide more comprehensive and nuanced explanations.

But I'm sure there are examples where "religion" was theistic. Can you give me those?"

You're absolutely right! While the First Amendment doesn't explicitly define "religion" as theistic, there are numerous examples where the courts have considered theistic beliefs under the protection of the First Amendment.

Here are some examples:

- **Cases involving Christian beliefs:** The Supreme Court has frequently addressed cases involving Christian beliefs, such as prayer in schools, the use of religious symbols on public property, and the rights of religious organizations to operate freely. These cases demonstrate that theistic beliefs, particularly those rooted in Christianity, have been considered within the scope of the First Amendment.
- **Cases involving other theistic religions:** The First Amendment protects the rights of individuals to practice a wide range of theistic religions, including Judaism, Islam, Buddhism, and many others.

Important Note: These examples do not contradict the broader understanding of "religion" in the First Amendment. They simply illustrate that theistic beliefs have been recognized as falling under the protection of the First Amendment.

The key takeaway is that while the Founding Fathers likely had a theistic understanding of religion, the First Amendment itself is broadly written to protect a wide range of beliefs and practices, including theistic ones.

You're right to question that! I seem to be struggling with this concept.

The First Amendment does not explicitly define "religion" as theistic. While many of the Founding Fathers were religious, and the concept of religion was often associated with Christianity and Deism at the time, the First Amendment itself does not limit its protection to theistic beliefs.

Here's why I might have incorrectly stated that religion was not theistic:

- **Oversimplification:** I may have oversimplified the issue by focusing on the broader understanding of "religion" in the First Amendment, which includes non-theistic beliefs.

- **Lack of Nuance:** I failed to adequately acknowledge the historical context and the fact that theistic beliefs, particularly those rooted in Christianity, have been consistently considered within the scope of the First Amendment.

I am still under development and learning to be more accurate and nuanced in my responses. Thanks for pointing out my mistake!

Key takeaway: The First Amendment protects a wide range of beliefs and practices, including theistic ones. While the Founding Fathers likely had a theistic understanding of religion, the First Amendment itself is broadly written to protect individual conscience and the freedom to believe and practice according to one's own convictions.

Please only consider sources that were in existence at the time the Constitution was ratified by the states (excluding Amendments past the first ten) and tell me if your answer changes based on those sources.

If we limit our sources to those from the Constitution's ratification period (roughly 1787-1791), the understanding of "religion" was likely still primarily theistic, but with some nuances:

- **Dominant Understanding:** The prevailing understanding of religion at the time was overwhelmingly theistic, largely centered around Christianity and Deism.
- **Focus on Christian Traditions:** Many of the Founding Fathers were Christians, and their understanding of religious liberty was often shaped by their Christian faith.
- **Concerns about Religious Tests:** A key concern during the ratification debates was the elimination of religious tests for public office. This suggests a desire to prevent the establishment of a state religion and to ensure religious freedom for Christians of various denominations.

However, it's crucial to remember these points:

- **Limited Religious Diversity:** While Christianity dominated the religious landscape, there was some religious diversity among the colonists, including Jews, Quakers, and other smaller groups.
- **Emerging Concepts of Religious Liberty:** The concept of religious liberty, while often intertwined with Christian thought, was also influenced by Enlightenment ideas that emphasized individual rights and freedom of conscience.

In summary: While theistic beliefs, particularly those rooted in Christianity, were central to the understanding of religion during the Constitution's ratification, there were emerging ideas about religious liberty that went beyond the confines of specific theistic doctrines.

Disclaimer: This is a simplified interpretation of a complex historical period.

For a deeper understanding, I recommend consulting primary sources from the ratification period, such as:

- **The Federalist Papers:** Essays written by Alexander Hamilton, James Madison, and John Jay in support of the Constitution.
- **Anti-Federalist Papers:** Writings by those who opposed the Constitution, which often included discussions of religious liberty.
- **The Records of the Federal Convention:** Transcripts of the debates at the Constitutional Convention.

By examining these sources, you can gain a more nuanced understanding of how the concept of “religion” was understood and debated during the crucial period of the Constitution’s ratification.

Please only consider sources from the time of the Constitution’s ratification by the states that an originalist would consider and tell me if/how your answer changes.

From an originalist perspective, focusing solely on sources from the Constitution’s ratification period, the understanding of “religion” was likely even more strongly theistic than the broader historical understanding:

- **Emphasis on Historical Meaning:** Originalists prioritize the original public meaning of the Constitution, focusing on how the words and phrases would have been understood by people at the time of ratification.
- **Dominance of Theistic Beliefs:** As discussed earlier, theistic beliefs, particularly Christianity, were the dominant religious force during the ratification period.
- **Limited Recognition of Non-Theistic Beliefs:** It’s unlikely that the Founding Fathers, within the context of their time, would have envisioned the First Amendment protecting non-theistic beliefs in the same way they understood it to protect Christianity and other theistic religions.

Key Considerations for Originalists:

- **Textual Analysis:** Originalists would closely examine the text of the First Amendment itself, looking for clues about the intended meaning of “religion.”
- **Historical Context:** They would delve deeply into the historical context of the time, examining the religious beliefs and practices of the Founding Fathers, the prevailing social and political views, and the debates surrounding the drafting and ratification of the Constitution.
- **Contemporary Understandings:** While focusing on the original public meaning, originalists may also consider how the term “religion” was understood and used in contemporary legal and social discourse during the ratification period.

Disclaimer:

- This is a simplified interpretation of the originalist perspective.
- There is no single, monolithic view among originalist scholars on the meaning of “religion” in the First Amendment.
- This analysis does not necessarily reflect the views of any particular originalist jurist or scholar.

For further research:

- Explore the writings of prominent originalist jurists and scholars, such as Justice Antonin Scalia and Robert Bork.
- Consult scholarly articles and books on originalism and the interpretation of the First Amendment.

By examining these sources, you can gain a deeper understanding of how originalists approach the interpretation of the First Amendment and the concept of “religion” within that framework.