No. 18-1150

In The Supreme Court of the United States

STATE OF GEORGIA, et al.,

Petitioners,

v.

PUBLIC.RESOURCE.ORG, INC.,

Respondent.

On Writ Of Certiorari To The United States Court Of Appeals For The Eleventh Circuit

BRIEF OF 36 COMPUTATIONAL LAW SCHOLARS AS *AMICI CURIAE* IN SUPPORT OF RESPONDENT

MICHAEL A. LIVERMORE
Counsel of Record
Professor of Law
UNIVERSITY OF VIRGINIA SCHOOL OF LAW
580 Massie Road
Charlottesville, VA 22903
Tel: (434) 982-6224
mlivermore@virginia.edu

CHARLOTTE S. ALEXANDER
Associate Professor of Law and Analytics
J. MACK ROBINSON COLLEGE OF BUSINESS
GEORGIA STATE UNIVERSITY
35 Broad Street, NW
Atlanta, GA 30303
Tel: (404) 413-7468
calexander@gsu.edu

ANNE M. TUCKER
Professor of Law
GEORGIA STATE UNIVERSITY COLLEGE OF LAW
85 Park Place, NE
Atlanta, GA 30303
Tel: (404) 413-9179
amtucker@gsu.edu

COCKLE LEGAL BRIEFS (800) 225-6964
WWW.COCKLELEGALBRIEFS.COM

TABLE OF CONTENTS

TABLE OF CONTENTS—Continued

TABLE OF AUTHORITIES

Page

CASES

TABLE OF AUTHORITIES—Continued

TABLE OF AUTHORITIES—Continued

Page

TABLE OF AUTHORITIES—Continued

TABLE OF AUTHORITIES—Continued

Page

TABLE OF AUTHORITIES—Continued

TABLE OF AUTHORITIES—Continued

TABLE OF AUTHORITIES—Continued

TABLE OF AUTHORITIES—Continued

Page

INTEREST OF *AMICI CURIAE*[1]

Amici are a group of 36 scholars who study the law using computational methodologies and whose research requires access to digital versions of legal texts.[2] *Amici* scholars have a range of disciplinary backgrounds, including law, political science, history, economics, finance, computer science, and mathematics. Common to *amici*'s scholarly work is the need for unfettered, copyright-free access to legal texts, in order to synthesize, interpret, and study the law.

SUMMARY OF THE ARGUMENT

The People have interests both in knowing what the law says and in understanding what it means. The government edicts doctrine furthers these interests by granting unfettered, copyright-free access to legal texts. This access is available to legal subjects so that they can have notice of the rules that apply to their

[1] Counsel of record for all parties received notice at least 10 days prior to the due date of the *amici curiae*'s intention to file this brief. All parties have given consent. No counsel for a party authored this brief in whole or in part, and no counsel or party made a monetary contribution intended to fund the preparation or submission of this brief. No person other than *amici curiae* made a monetary contribution to this brief's preparation or submission. A list of all of the *amici* is set forth in the Appendix to this brief.

[2] The views expressed herein are those of the *amici* in their capacity as scholars. No part of this brief purports to express the views of any institution, including the University of Virginia School of Law and Georgia State University.

conduct. Access to legal texts is also foundational to law scholarship, which has informed shared understanding of the law and shaped legal development for centuries.

In a digital age, digital access to the law is the touchstone. Such access not only lowers barriers to the public; it also facilitates the application of new research tools, such as natural language processing, computational text analysis, and machine learning, that can help illuminate the law's meaning. Narrowing the scope of the government edicts doctrine will inhibit scholars' ability to access the law and apply these new tools and techniques.

Since at least the sixth century, when Justinian I ordered the organization and codification of Roman Law, the work of law scholars, jurists, and legal practitioners has been intertwined. As societies grew in scale and complexity, the law became a learned profession, requiring substantial study to come to grips with the rules and rulings issued by government bodies diffused across increasingly sprawling states. In the common law tradition, legal commentaries, treatises, and other works of law scholarship have played a particularly important role, aggregating and synthesizing what would otherwise be an unmanageable body of case law. Such scholarship has frequently been referenced by state and federal courts, including this Court, since the founding of the Republic.

From the printing press to the internet, law scholarship has evolved alongside information technology.

With the growing availability of digital versions of legal texts, scholars of the law have begun to take advantage of related advances in mathematics, computer science, statistics, and machine learning. Recent work applies these tools to a range of legal texts, and researchers are actively developing methodologies and techniques that can help improve both scholarly and public understanding of the law. Unfettered, copyright-free access to large bodies of legal texts in digital form is a precondition for future development in this area.

Official annotations to state statutory codes fall into the heartland of the types of texts that law scholars can usefully analyze with computational techniques. Such annotations are state-endorsed interpretations of and commentary on state statutes. Whatever their officially binding character, they are used by courts and other legal actors—including scholars of the law—to understand and apply the law. Other forms of legal commentary and scholarship created by private authors are informative and persuasive, as evidenced by the Court's long history of reliance on such works, but are not authoritative. Legislative endorsement of official annotations confer the legitimacy of the state on these interpretations, distinguishing them from other forms of legal commentary and raising their status to the level of a government edict.

ARGUMENT

I. Extending copyright protection to official annotations of state statutes will inhibit legal scholarship

New technologies such as natural language processing and other methods of computational text analysis have created new approaches to legal scholarship. This work has already borne early fruit as scholars have developed new insights and courts have looked to techniques such as corpus linguistics to aid legal interpretation. This type of scholarship requires access to large data sets of legal texts. Statutory annotations fall squarely within the types of data that can be usefully subjected to computational legal analysis. As a consequence, subjecting official annotations to copyright would unnecessarily hinder the growth of this new form of legal scholarship.

A. Digitized, publicly available legal texts facilitate the use of computational tools in legal scholarship

Law and legal scholarship have long been intertwined with information technology. Robert C. Berring, *Full-Text Databases and Legal Research: Backing into the Future*, 1 HIGH TECH L.J. 27 (1986). A technological change enabled the transition from exclusive reliance on cultural norms to formal, consistent legal rules: the advent of the written word. Mireille Hildebrandt, *The Force of Law and the Force of Technology*, *in* THE ROUTLEDGE HANDBOOK OF TECHNOLOGY, CRIME AND

JUSTICE 597, 599 (M.R. McGuire & Thomas J. Holt eds., 2017). The printing press, and later the creation of searchable legal databases, also profoundly influenced how law was distributed, understood, and studied. *See* Berring, *supra*.

More recently, two related trends are transforming the practice and study of law: the large-scale digitization of legal texts; and advances in information processing technology and theory. Michael A. Livermore & Daniel L. Rockmore, *Introduction: From Analogue to Digital Legal Scholarship*, *in* LAW AS DATA: COMPUTATION, TEXT, AND THE FUTURE OF LEGAL ANALYSIS xv (Michael A. Livermore & Daniel L. Rockmore eds., 2019). In the 1970s, commercial databases led the digitization of legal texts, which later spread through the burgeoning internet. Now, growing public data availability has intersected with developments in the fields of artificial intelligence, natural language processing, text mining, and machine learning to increase the role of computational methods in the professional lives of lawyers, law scholars, and courts.

Researchers engaged in the computational analysis of legal texts depend not only on access to digital legal texts, however, but also, and critically, on *open* access to them. This is because proprietary databases such as Lexis and Westlaw prevent researchers (even those with paid subscriptions) from downloading textual data in bulk using automated approaches. If data is available on an open-source site such as Public.Resource.Org, researchers can automate the data collection process—essentially programming their computers to collect the

data for them. When legal texts are not publicly available, but are instead locked away in proprietary databases, computational research is extremely costly and inefficient, at best, and may be entirely infeasible.

Once text is assembled in bulk, however, techniques in natural language processing can be used to extract quantitatively useful information from complex legal texts. For example, similarities between documents can be discovered and used to sort and classify documents into meaningful categories, which is the basis of e-discovery, a legal practice at issue in the Court's recent decision in *Rimini Street Inc. et al. v. Oracle USA, Inc.*, 139 S. Ct. 873 (2019). Further, algorithms originally designed to detect plagiarism can be used to measure similarity between texts, enabling researchers to explore, for example, how lower federal court opinions influence the content of Supreme Court opinions. Pamela C. Corley, Paul M. Collins, Jr. & Bryan Calvin, *Lower Court Influence on U.S. Supreme Court Opinion Content*, 73 J. POL. 31 (2011). Another technique, known as sentiment analysis, uses the presence of positive or negative words to estimate the emotional content of texts, and accordingly facilitates research into attitudes, feelings, and biases of authors and institutions. Bing Liu, SENTIMENT ANALYSIS: MINING OPINIONS, SENTIMENTS, AND EMOTIONS (2015). Topic modelling, which extracts semantic content from textual data, David M. Blei, *Probabilistic Topic Models*, 44 COMM. ACM 77 (2012); David M. Blei & John D. Lafferty, *A Correlated Topic Model of Science*, 1 ANN. APPL. STAT. 17 (2007), is another approach that legal scholars

employ to quantitatively represent legal texts, and study similarities and differences among them. Michael A. Livermore, Allen B. Riddell & Daniel N. Rockmore, *The Supreme Court and the Judicial Genre*, 59 ARIZ. L. REV. 837 (2017). Moreover, text mining tools can extract citation information from legal texts, which can then be coupled with various forms of network analysis to, for example, examine legal complexity or reveal patterns of influence among courts. J.B. Ruhl, Daniel Martin Katz & Michael J. Mommarito II, *Harnessing Legal Complexity: Bringing Tools of Complexity Science to Bear on Improving Law*, 355 SCIENCE 1377 (2017); James H. Fowler et al., *Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court*, 15 POL. ANALYSIS 324, 325 (2007). The information extracted through computational methods such as these can be analyzed using traditional statistical models as well as newer, machine-learning approaches, to generate both descriptive and predictive results of widespread interest—both to the legal community and broader society.

Technology-driven analytic methods applied to legal texts can inform long-standing inquiries in the law. One approach that has received considerable recent attention is the use of corpus linguistics in legal analysis. Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 YALE L.J. 788 (2018). Corpus linguistics is a computer-based method of collecting information regarding the use and context of a phrase or word by interrogating a large body, or *corpus*, of naturally occurring language. *Id.* This tool permits

scholars, parties to litigation, and judges to address ambiguity in a law by considering the ordinary meaning of a word or phrase in the historical context of the legislation's enactment.

Corpus linguistics has already been recognized as a valuable tool by courts, some of which have employed the analytic method to inform their legal interpretations. In recent years, courts including the Sixth Circuit Court of Appeals, *Wilson v. Safelite Grp., Inc.*, 930 F.3d 429, 438-39 (6th Cir. 2019), and the Supreme Courts of Utah, *Richards v. Cox*, No. 20180033, 2019 Utah LEXIS 157, at *10-14 (Utah Sept. 11, 2019); Michigan, *People v. Harris*, 885 N.W.2d 832, 838-39 (Mich. 2016); and Idaho, *State v. Lantis*, No. 46171, 2019 Ida. LEXIS 127, at *13-17 (Idaho Aug. 23, 2019), have independently conducted corpus linguistic inquiries and reported their results in published opinions. Other courts have noted positively the value of such analysis and have encouraged parties to present empirical support derived from the method. *See, e.g.*, *Muddy Boys, Inc. v. DOC*, 440 P.3d 741, 749 (Utah Ct. App. 2019) ("[O]ne of the chief benefits of a corpus-linguistics-style analysis is that it offers a systematic, nonrandom look at the way words are used across a large body of sources."); *Craig v. Provo City*, 389 P.3d 423, 428 (Utah 2016). The Utah Supreme Court encouraged lawyers to "provide courts with meaningful tools using the best available methods when the court is tasked with determining ordinary meaning," noting that there is a general shortcoming in human ability to select the most common meaning of language. *Fire*

Ins. Exch. v. Oltmanns, 416 P.3d 1148, 1163 n.9 (Utah 2018).

Computationally based legal scholarship promises to shed light on classic questions of legal interpretation and may open up entirely new avenues for future legal development. But the success of this research depends on open access to unbiased data—the very type of access that Public.Resource.Org as well as similar open-source sites, such as CourtListener and Justia, are working to facilitate. CourtListener, https://www.courtlistener.com/; Justia, https://www.justia.com/. Scholars in this area are taking advantage of open access to the law and legal materials to leverage increasingly sophisticated analytic techniques. This trend has already shown important potential to contribute to the long-standing and productive dialogue between jurists and legal scholarship.

B. Open access to large volumes of digital texts is required for many forms of computational legal analysis

Many of the computational techniques discussed above share an important characteristic: their ability to return useful results is a function of the quantity and quality of data available to them. Without a large amount of textual data, these forms of analysis are less effective and, in some cases, cannot be conducted at all. Limiting access to legal texts—and particularly those legal texts that have the most value for computational

legal analysis—directly interferes with the field's ability to grow.

There are two primary reasons why bulk data is needed for computational legal analysis. The first concerns complexity. Legal texts, when represented in a quantitative fashion, can be understood as complex, high-dimensional objects. A single judicial opinion, for example, might address multiple different legal claims, analyze multiple different strands of precedent relevant to each, and conclude with different rulings on different sub-parts of the parties' arguments. To render all of this information quantitatively and in high fidelity requires an extremely large number of variables (i.e., dimensions). In other words, many factors are necessary for computational tools to be precise. Even when dimensionality reduction tools are used to make the data more tractable, there are often limits to the simplicity with which legal texts can be accurately represented. As the number of dimensions in a data set grows, the number of observations (amount of text) needed to carry out meaningful analysis also grows. In the technical literature, this fact is sometimes referred to as the "curse of dimensionality." Eamonn Keogh & Abdullah Mueen, *Curse of Dimensionality*, *in* ENCYCLOPEDIA OF MACHINE LEARNING AND DATA MINING 314 (Claude Sammut & Geoffrey I. Webb eds., 2017). For analysis that is sensitive to fine distinctions between legal documents (implying a relatively large number of variables), a large number of observations is needed, in the form of large masses of legal text.

A second consideration that favors large volumes of data is the problem of bias. Nikhil Garg et al., *Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes*, 115 PROC. NAT'L ACAD. SCI. E3635 (2018). Datasets that are systematically limited create the risk that conclusions drawn from this data will be skewed in some unobservable fashion that makes analysis and interpretation difficult. To take one example, corpus linguistics examines how words are used in context. Were a corpus to systematically exclude texts produced by certain groups of language users, then the resulting analyses would be biased, in the sense that alternative usages that are common among the excluded community might not appear in the corpus at all. If some states take their annotated codes out of the public domain, then nation-wide surveys of statutory law will be compromised, as will efforts to compare and contrast statutory law across states—since the only data available for large-scale computational study will be biased by the omission of the excluded text.

The government edicts doctrine mitigates the problems of complexity and bias by enabling unfettered, copyright-free access to legal text in bulk. All parties to this case acknowledge that judicial opinions, statutes, and regulations fall within the doctrine. As a consequence, copyright cannot successfully be asserted over these materials. Governments can also intentionally release to the public materials that fall outside of the doctrine.

In the instant case, the respondent asserts that official legal materials developed under the direct

supervision of the legislature are covered by copyright. If supported by this Court, this assertion would render these materials effectively inaccessible to scholars engaged in computational analysis of legal texts. The consequence would be a reduction in the amount of data that can be subject to computational legal analysis, and the introduction of potential biases based on differences between jurisdictions. Both of these effects would limit the ability of researchers to study the law.

C. Official annotations are legal texts that can be usefully analyzed by scholars

Prior scholarship in computational legal analysis has taken a wide range of legal documents as objects of study, including judicial opinions, Douglas Rice, *The Impact of Supreme Court Activity on the Judicial Agenda*, 48 L. & SOC'Y REV. 63 (2014), constitutions, Daniel Rockmore et al., *The Cultural Evolution of National Constitutions*, 69 J. ASS'N INFO. SCI. & TECH. (2017), statutory texts, David S. Law & David Zaring, *Law Versus Ideology: The Supreme Court and the Use of Legislative History*, 51 WM. & MARY L. REV. 1653 (2010), administrative materials, Michael A. Livermore, Vladimir Eidelman & Brian Grom, *Computationally Assisted Regulatory Participation*, 93 NOTRE DAME L. REV. 977 (2018), and docket sheets, Charlotte S. Alexander, *Litigation Migrants*, 56 AM. BUS. L.J. 235 (2019); Charlotte S. Alexander, *#MeToo and the Litigation Funnel*, 22 EMPL. RTS. & EMPL. POL'Y J. 101 (2019). Official annotations of state statutes fall squarely in

the heartland of the types of documents that can usefully be analyzed using computational legal analysis.

For example, courts conducting corpus linguistic analyses have, to date, relied on the Corpus of Contemporary American English (COCA) and the Corpus of Historical American English (COHA), depending on the relevant time period of their inquiry. *Wilson*, 930 F.3d at 438-39; *Richards*, 2019 Utah LEXIS 157, at *10-14; *Harris*, 885 N.W.2d at 838-39; *Lantis*, 2019 Ida. LEXIS 127, at *13-17. These corpora are impressively sized, publicly available resources. COCA contains more than 560 million words, "equally divided among spoken, fiction, popular magazines, newspapers, and academic texts" from 1990 to 2017. Corpus of Contemporary American English, https://www.english-corpora.org/coca/. COHA, similarly, contains over 400 million words spanning the 1810s to 2000s. Corpus of Historical American English, https://www.english-corpora.org/coha/.

But neither the COCA nor the COHA focuses on legal texts. Rather, they serve as a window into how words or phrases are used in common language. The availability of more specialized sets of documents facilitates the application of corpus linguistics techniques to legal texts. For example, the *British Law Report Corpus* is a collection of British judicial opinions with 8.5 million words that allows for the study of specifically legal language. *See* María José Marín, *Legalese as Seen Through the Lens of Corpus Linguistics—An Introduction to Software Tools for Terminological Analysis*, 6 INT'L J. LANGUAGE & L. 18, 21 (2017). Just as judicial

opinions have certain characteristics that distinguish them from common speech, legislators use a special language when engaged in statutory drafting. In particular, legislators frequently use terms of art or other turns of phrase that are less pervasive in daily speech. Further, annotations could be especially helpful as additional texts that elaborate on the meaning of statutes, and in the case of official annotations, that have been given the imprimatur of the legislature.

II. Computational legal scholarship builds on a long tradition of scholarly synthesis of legal materials that courts have found useful

When tasked with giving meaning to legal text, courts rely on an array of interpretive aids, including commentaries, treatises, and other works of legal scholarship. These resources organize, synthesize, and distill vast amounts of raw legal material, broadly defined, to help identify patterns, tease out meaning, and explain historical usage. While not law themselves, these resources aid in the production of legal knowledge, and support courts' interpretive function. *Kernan v. Cuero*, 138 S. Ct. 4, 9 (2017), *reh'g denied*, 138 S. Ct. 724 (2018) (quoting *Glebe v. Frost*, 574 U.S. 21 (2014) (per curiam)) (observing that treatises and law review articles "do[] not constitute 'clearly established Federal law, as determined by the Supreme Court' ").

Beginning long before the advent of computational methods, there is a rich tradition of the scholarly creation of commentaries and treatises that summarize

and synthesize large bodies of law. Blackstone's Commentaries, for example, published between 1765 and 1769, were designed "to provide a systematic and explanatory presentation of the doctrines of all branches of English law." Stephen Skinner, *Blackstone's Support for the Militia*, 44 AM. J. LEGAL HIST. 1, 1 (Jan. 2000) (describing Blackstone's Commentaries); Albert W. Alschuler, *Rediscovering Blackstone*, 145 U. PA. L. REV. 1, 16 (1996) ("The United States Supreme Court still cites the *Commentaries* approximately ten times each year."). Even earlier are codification efforts that extend back to the Byzantine period. H.F. Jolowicz & Barry Nicholas, HISTORICAL INTRODUCTION TO THE STUDY OF ROMAN LAW 478-504 (3d ed. 1972) (describing creation of *Codex Justinianeus* in the sixth century). The value of scholarly synthesis arises at least in part from the natural tendency of law to proliferate through legislation and judicial decision making alongside the desire (and sometimes obligation) on the part of legislators and jurists to achieve coherence with an increasingly sprawling body of law.

The tools of computational legal analysis are particularly well-suited to continue in the long tradition of scholarly synthesis of legal materials because of their ability to deal with very large volumes of documents. Accordingly, broad and unfettered public access to raw legal material will facilitate the future production of research in the tradition of commentaries, treatises, and related forms of scholarship, which will in turn support courts' interpretive enterprise. Inhibiting that access, as respondents seek to do, threatens the

operation of this important feedback loop in the generation of legal knowledge and interpretation of the law.[3]

A. Commentaries, treatises, and related forms of legal scholarship have proven useful to courts

Courts frequently rely on scholarly work that synthesizes legal material. For example, a non-exhaustive review of U.S. Supreme Court opinions over the last decade identified 467 citations to Blackstone alone, along with 896 references to other commentaries, treatises, restatements of law, and additional field-building or field-interpretive works. Further, scholars have documented nearly 230 citations in Supreme Court opinions to law review articles during just the 2016-2017 term. Adam Feldman, *With A Little Help from Academic Scholarship*, Empirical SCOTUS, Oct. 31, 2018, https://empiricalscotus.com/2018/10/31/academic-scholarship/.

---

[3] An example of this feedback loop may be found in *Green v. Bock Laundry Mach. Co.*, 490 U.S. 504, 511-21 (1989), where this Court interpreted the *Federal Rules of Evidence* on the impeachment of witnesses, citing three treatises for their distillation of the common law on impeaching evidence, along with scholarly articles that surveyed the landscape of court decisions on the admission or exclusion of such evidence. *Id.* (citing Jack Weinstein & Margaret Berger, WEINSTEIN'S EVIDENCE 609-58 (rev. ed. 1988) (citing James Wigmore, EVIDENCE § 519 (3d ed. 1940)); James Moore & Helen Bendix, MOORE'S FEDERAL PRACTICE VI-134 (2d ed. 1988); Edward Cleary et al., MCCORMICK ON EVIDENCE 93 (3d ed. 1984); Mason Ladd, *Credibility Tests—Current Trends*, 89 U. PA. L. REV. 166, 176, 191 (1940); Carl McGowan, *Impeachment of Criminal Defendants by Prior Conviction*, 1970 L. & SOC. ORDER 1 (1970).

These interpretive aids draw not only on case law, as Blackstone did, but also on regulatory and statutory text, legislative history, and—importantly for the present case—on statutory annotations. *See infra* pp. 21-23.

The following discussion briefly illustrates judicial reliance on commentaries, treatises, and other legal scholarship in Supreme Court opinions from 1790 to the present to highlight the importance of these resources and, by extension, of access to the full array of raw legal material upon which they draw.

Commentaries. William Blackstone's Commentaries is one of several early works that sought to survey the state of English law, extract it "from the rubbish in which it was buried and . . . show[] it to the public in a clear, concise, and intelligible form." Alschuler, *supra* at 9 n.36 (quoting 1767 ANNUAL REGISTER 286, 287 (8th ed. 1809)) (a journal edited by Burke), quoted in A.V. Dicey, *Blackstone's Commentaries*, 4 CAMBRIDGE L.J. 286, 286 (1932) (originally published in 54 NAT'L REV. 645, 653 (1909)). As noted above, the Court cites Blackstone extensively, as many as ten times per year as of the late 1990s, by one count. *Id.* at 16. In *District of Columbia v. Heller*, for example, the Court relied heavily on Blackstone himself, as well as on notes on Blackstone by St. George Tucker, an early American legal scholar, and on additional commentaries on the interpretation of the Constitution. *D.C. v. Heller*, 554 U.S. 570, 605-07 (2008).

Similarly, the Court has relied on Kent's Commentaries on American Law and Sir Edward Coke's Institutes of the Lawes of England in almost 300 cases. *See, e.g.*, *Sun Oil Co. v. Wortman*, 486 U.S. 717, 726 (1988); *Betterman v. Montana*, 136 S. Ct. 1609, 1614 (2016); *Stoneridge Inv. Partners, LLC v. Scientific-Atlanta, Inc.*, 552 U.S. 148, 176 n.12 (2008); *Washington v. Glucksberg*, 521 U.S. 702, 711 n.10 (1997); *U.S. v. Maine*, 475 U.S. 89, 104 (1986); *Surplus Trading Co. v. Cook*, 281 U.S. 647, 655 (1930); *U.S. v. Wong Kim Ark*, 169 U.S. 649, 654 (1898); *The William Bageley*, 72 U.S. 377, 410 (1866); *Bradlie v. Md. Ins. Co.*, 37 U.S. 378, 393 (1838). The Court's reliance on commentaries is relevant here because of the function that these works played at the time they were written, sweeping together the relevant set of raw legal texts in order to distill their meaning. In *Heller*, for instance, Justice Scalia's description of Blackstone's works as "the preeminent authority on English law for the founding generation." *Heller*, 554 U.S. at 593-94. That opinion's extensive reliance on the commentaries and treatises of the day implicitly also relied on the authors' *access* to the underlying legal materials on which they commented.

Treatises. As legal material proliferated, single, field-wide commentaries like Blackstone's began to be replaced by subject-matter-specific treatises in the production and organization of legal knowledge. Alschuler, *supra* at 9. Treatise authors have remained active throughout the development of American law, and have been cited heavily by the Court.

Examining only six exemplar treatises, for instance, we found over 500 Supreme Court citations in the years 1905 through 2018.[4] These include reliance on John Henry Wigmore's *Treatise on the Anglo-American System of Evidence in Trials at Common Law*, widely known as Wigmore on Evidence, over 320 times, including in a 2019 concurrence by Chief Justice Roberts, who commented that a court "would want to know what John Henry Wigmore said about an issue of evidence law." *Kisor v. Wilkie*, 139 S. Ct. 2400, 2442 (2019) (citation and internal quotation marks omitted). Similarly, contract law interpreters Arthur Corbin and Samuel Williston remain important sources in the Court's interpretation and application of the common law of contracts. Arthur L. Corbin et al., CORBIN ON CONTRACTS (rev. ed. 2019); Samuel Williston & Richard A. Lord, A TREATISE ON THE LAW OF CONTRACTS (4th ed. 1993 & Supp. 1999).

Like Blackstone before him, Corbin's contribution has been described as "weaving together the vast body of our case law of contracts into understandable patterns." Friedrich Kessler, *Corbin on Contracts: Part I:*

---

[4] Treatises cited include William M. Fletcher et al., FLETCHER CYCLOPEDIA OF THE LAW OF PRIVATE CORPORATIONS (perm. ed., rev. vol. 1999) (32 citations); Henry W. Ballantine, BALLANTINE ON CORPORATIONS (rev. ed. 1946) (7 citations); Arthur L. Corbin et al., CORBIN ON CONTRACTS (rev. ed. 2019) (52 citations); Richard R. Powell, POWELL ON REAL PROPERTY (Patrick J. Rohan ed., 1995) (11 citations); John H. Wigmore, A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW (3d ed. 1940) (322 citations); Samuel Williston & Richard A. Lord, A TREATISE ON THE LAW OF CONTRACTS (4th ed. 1993 & Supp. 1999) (87 citations).

Formation of Contract, 61 YALE L.J. 1092, 1092 (1952). As above, these treatise authors provide interpretive aid to the Court by virtue of their access to masses of raw legal text and the meaning they extract therefrom.

Restatements. The American Law Institute's Restatements offer yet another example of synthesizing works of legal analysis, where scholars mine and distill the text of an array of raw legal material in a variety of subject areas. In the words of Justice Kennedy, the Restatements remedy the inaccessibility of the law, caused by the "great outpouring of case decisions and statutes," by performing a function similar to Blackstone hundreds of years before. Andrew Hamm, *Retired Justice Kennedy promises message of civility at American Law Institute's annual meeting*, SCOTUSBlog (May 20, 2019, 5:17 PM), https://www.scotusblog.com/2019/05/retired-justice-kennedy-promises-message-of-civility-at-american-law-institutes-annual-meeting/.

Though the neutrality of the more recent Restatements has sometimes been questioned, *Kansas v. Nebraska*, 135 S. Ct. 1042, 1064 (2015) (Scalia, J., concurring in part and dissenting in part), the Supreme Court and lower courts alike frequently cite these works as reliable sources for the governing rule of law. In *Norfolk & W. Ry. Co. v. Ayers*, 538 U.S. 135, 148 (2003), for example, both the majority and concurrences cite to the (Third) Apportionment of Liability (2000) twenty-eight times, including for the proposition that the Restatement "states the general rule." *Id.* at 148 ("Unlike stand-alone claims for negligently inflicted emotional distress, claims for both pain and

suffering associated with, or 'parasitic' on, a physical injury are traditionally compensable. The Restatement (Second) of Torts states the general rule. . . .") (internal citations omitted).

Law Review Articles. Finally, law review articles can serve the same field-gathering and distilling role as the other interpretive aids discussed above, and the Court cites them for this very function. Most recently, in *Kisor v. Wilkie*, legal scholarship played a prominent role in framing the Court's analysis of the parties' arguments in an administrative law case. 139 S. Ct. at 2421. The law review articles in question provided a historical summary of this Court's approach to judicial review of agency interpretations, *id.* (describing "a widely respected law review article," John Manning, *Constitutional Structure and Judicial Deference to Agency Interpretations of Agency Rules*, 96 COLUM. L. REV. 612 (1996)), and drew on the authors' experience reviewing "thousands" of agency rules for vagueness, *id.* (citing Cass R. Sunstein & Adrian Vermeule, *Constitutional Structure and Judicial Deference to Agency Interpretations of Agency Rules*, 96 COLUM. L. REV. 612, 635-36 (1996)).

The Value of Official Annotations. These legal scholars' insights, which have informed central parts of the Court's analysis, were possible only because of their broad and unfettered access to case law and regulations—the same family of raw legal material to which statutory annotations belong. In addition, annotations themselves have served as an important input for influential interpretive legal scholarship.

For example, Fletcher Cyclopedia of Corporations, a frequently cited treatise, relies upon official annotations and commentary throughout the treatise for such interpretive tasks as distilling the history of the Model Business Corporation Act, William M. Fletcher et al., FLETCHER CYCLOPEDIA OF THE LAW OF PRIVATE CORPORATIONS § 3439.20 (perm. ed., rev. vol. 1999) (citing to M.B.C.A. § 3.04 annotations on historical background), understanding the scope of the Clayton Act as it relates to corporations, *id.* at § 5062.30 (perm. ed., rev. vol. 1999) (citing to the American Law Reports, 23 A.L.R. Fed. 878 (1975)), and interpreting the Sentencing Reform Act to understand the rights of corporate criminal defendants, *id.* at § 4965.10 (citing extensively to commentary to the Act to interpret the meaning).

Similarly, Williston on Contracts cites to annotations to the Uniform Laws and the United States Code Annotated for specific interpretative points such as the regulation of insurance collateral. Samuel Williston & Richard A. Lord, A TREATISE ON THE LAW OF CONTRACTS § 20:27 (4th ed. 1993 & Supp. 1999) (citing to the Uniform Consumer Credit Code 4.104 annotations for the proposition that "statutes have been enacted regulating the nature and extent to which lenders may insist on insurance collateral to the loan transaction."). Williston also cites annotations in the official research references to whole chapters such as those on mutual assent, *id.* at § 8, rights and duties under contract, *id.* at § 36, and bargains and restraints on trade, *id.* at § 13. Importantly, Williston cites to state statutory

annotations for important interpretive and contextual information. For example, in evaluating the appropriate time for a creditor to make tender of payment, Williston cites to the New York annotations to the restatement to understand the exact parameters of New York law allowing for up to one year for payment as opposed to merely a "reasonable time." *Id.* at § 27:14.

B. Computational tools are well suited to continue the tradition of interpretive legal scholarship by synthesizing large volumes of texts

The convergence of the empirical legal studies movement with computational methodologies continues the Blackstonian tradition of law-oriented synthesizing and interpretive scholarship. And researchers today can perform this work more rigorously and efficiently than ever before, so long as the relevant textual data remains accessible. *See* Michael A. Livermore & Daniel N. Rockmore, *Distant Reading the Law, in* LAW AS DATA, *supra* at 3. In a recent important integration of empirical quantitative methods with the traditional goals of scholarly synthesis, the reporters for the Restatement of Consumer Contracts informed their deliberations based on quantitative analysis of patterns found in the universe of relevant cases. Oren Bar-Gill, Omri Ben-Shahar & Florencia Marotta-Wurgler, *Searching for the Common Law: The Quantitative Approach of the Restatement of Consumer Contracts*, 84 U. CHI. L. REV. 7 (2017).

The particular advantage of computational text analysis is its ability to survey a vast landscape of documents and examine them using quantitative tools. *See* Allen Riddell, *How to Read 22,198 Journal Articles: Studying the History of German Studies with Topic Models*, *in* DISTANT READINGS: TOPOLOGIES OF GERMAN CULTURE IN THE LONG NINETEENTH CENTURY 91 (Matt Erlin & Lynne Tatlock eds., 2014). For example, computational legal analysis can effectively process the entire corpus of Supreme Court decisions. Keith Carlson, Michael A. Livermore & Daniel L. Rockmore, *A Quantitative Analysis of Writing Style on the US Supreme Court*, 93 WASH. U. L. REV. 1461 (2016). The ability to extract information from such a large, complex body of text allows analysts to draw conclusions that would otherwise be elusive due to the limits of human cognition and resources. There are also economic gains associated with the ability to process large corpora of legal texts in a cost-efficient manner.

As mentioned above, corpus linguistics provides one set of techniques for extracting valuable information from large collections of documents totaling many millions of words. Another approach applies network analysis to a large number of cases to detect patterns in how judicial opinions cite relevant authority. Examples of work along these lines include efforts to develop network-based metrics that correspond to expert judgment about the highest impact Supreme Court precedents. James H. Fowler & Sangick Jeon, *The Authority of Supreme Court Precedent*, 30 SOC. NETWORKS 16, 17 (2008); James H. Fowler et al., *supra*.

Other scholars have introduced an alternative way of describing dependencies between opinions and determining the legal significance of cases, which builds on estimating a "genealogical model" from citation data. Tom S. Clark & Benjamin E. Lauderdale, *The Genealogy of Law*, 20 POL. ANALYSIS 329 (2012). Related work investigates the citation networks of courts such as the European Court of Justice. Mattias Derlén & Johan Lindholm, *Is It Good Law? Network Analysis and the CJEU's Internal Market Jurisprudence*, 20 J. INT'L ECON. L. 257 (2017).

Computational tools can also facilitate the study of conceptual categories in large groups of legal texts. Examples include examinations of constitutional preambles, David S. Law, *Constitutional Archetypes*, 95 TEX. L. REV. 153 (2016), executive pronouncements, J.B. Ruhl, John Nay & Jonathan M. Gilligan, *Topic Modeling the President: Conventional and Computational Methods*, 86 GEO. WASH. L. REV. 1243 (2018), successor liability cases, Frank Fagan, *Successor Liability from the Perspective of Big Data*, 9 VA. L. & BUS. REV. 391 (2014), and veil-piercing decisions, Jonathan Macey & Joshua Mitts, *Finding Order in the Morass: The Three Real Justifications for Piercing the Corporate Veil*, 100 CORNELL L. REV. 99 (2014). In each instance, the authors use computational tools to examine a large number of legal documents for the purpose of testing the extent to which existing doctrinal or substantive categorizations map onto the actual language used by legal decision makers.

Scholars engaged in historical research of legal phenomena have also taken advantage of the tools of computational text analysis to survey and extract information from large collections of legal documents in order to identify patterns that would be impossible to otherwise discern. For example, one study examined over one hundred thousand digitized trial records from the Central Criminal Court in London over a several-hundred-year period ending in the early twentieth century. Sara Klingenstein, Tim Hitchcock & Simon DeDeo, *The Civilizing Process in London's Old Bailey*, 111 PROC. NAT'L ACAD. SCI. 9419 (2014). The authors found that testimony in violent cases became increasingly distinctive from testimony in other cases from the late 1700s to the early 1900s. They attributed this long-term shift not to any specific legislative or bureaucratic policy change, but to "a gradual process driven by evolving social attitudes" and especially "the decreasing acceptability of interpersonal violence as part of normal social relations." *Id.* at 9423.

Other legal history-oriented projects using large volumes of documents and computational text analysis techniques include:

- Text-based representations of judicial opinions collected in the first eight volumes of the Hawaii Reports to examine how jurists' interpretation of the interaction of racial status and access to the writ of habeas changed over the course of the mid-to-late nineteenth century, Charles W. Romney, *Using Vector Space Models to Understand the Circulation*

of Habeas Corpus in Hawai'i, *1852-92*, 34 L. & HIST. REV. 999 (2016);

- Analysis of a dataset of state session laws to examine the adoption of laws affecting juvenile criminal defendants in the 1990s. Eric C. Nystrom & David S. Tanenhaus, *The Future of Digital Legal History: No Magic, No Silver Bullets*, 56 AM. J. LEGAL HIST. 150 (2016); Eric C. Nystrom & David S. Tanenhaus, *"Let's Change the Law": Arkansas and the Puzzle of Juvenile Justice Reform in the 1990s*, 34 L. & HIST. REV. 957 (2016).

Each of these projects integrates quantitative text analysis tools and data derived from digitized legal documents with the methods and questions familiar to legal historians. Collectively, the works summarized here demonstrate the broad applicability of computational tools to legal scholarship, which will inform legal understanding and interpretation. All of this research relies on bulk access to non-copyrighted legal materials facilitated by the government edicts doctrine.

III. Official annotations such as Georgia's are created by state action, and as such are government edicts rather than legal commentary

The official statutory annotations at issue in the instant case do not directly state rules of conduct, but instead explain and contextualize the language contained in statutory texts. As such, they share some

characteristics of legal commentary, in the sense that they help clarify the meaning, scope and application of legal texts. The official annotations are different in kind from legal commentary, however, because they are created by the state. This Court should interpret the government edicts doctrine to maintain a clear distinction in which official government statements about the content of the law remain outside the copyright regime, while other forms of commentary retain their traditional copyright status.

A. The Eleventh Circuit's "hallmarks" clearly identify official legal texts

The Eleventh Circuit "navigate[d] the ambiguities surrounding how to characterize" Georgia's official annotations based on three "hallmarks" that distinguish texts that are covered by the government edicts doctrine from those that are not. These hallmarks are: "the *identity* of the public officials who created the work; the *authoritativeness* of the work; and the *process* by which the work was created." *Code Revision Comm'n for General Assembly of Georgia v. Public.Resource.Org, Inc.*, 906 F.3d 1229, 1232 (11th Cir. 2018) (emphasis in original).

In applying this standard to Georgia's official annotations, the court below found that all three were satisfied. The annotations were "created by the duly constituted legislative authority of the State of Georgia," they "clearly have authoritative weight," and the process of their adoption includes core features of the

legislative process, "namely bicameralism and presentment." *Id.* at 1233.

The test used by the court below helps clarify the difference between Georgia's official annotations and standard works of legal commentary. The fundamental difference is that the latter are not created by the state. They are not produced through legislative or other legal channels, and are only statements about the law, rather than statements that either directly contain or authoritatively characterize the law.

The three hallmarks used by the court below to demarcate the difference between government edicts and works of legal commentary that may be subject to copyright provide a workable distinction that can be easily implemented by researchers engaged in computational law scholarship. When building a corpus for analysis, material that bears the imprimatur of the state and contains official interpretation of the law can generally be understood as falling within the government edicts doctrine, while material issued by a private publisher falls outside the doctrine. It would be difficult and costly for researchers to comb through each document in an attempt to differentiate between official and unofficial text. Such an exercise would substantially hinder the development of computational legal techniques because the preliminary task of corpus assembly alone would become resource- and time-prohibitive. Respondent's proposed interpretation of the government edicts doctrine would require just such a subdocument level review, and would pose a significant barrier to scholarship in this area.

Because of their integration into the legislative process and special significance in the state legal system, the official annotations are not appropriately thought of as resulting from the voice of the individuals charged with their drafting, but instead as emanating from the People and therefore outside the domain of the copyright system. To the extent that the case of official annotations represents a close case, the Court should err on the side of a clear rule based on institutional features (such as was used by the Eleventh Circuit) that grants unfettered access to legal documents to facilitate the study, interpretation, and synthesis of the law. Joseph Scott Miller, *Error Costs and IP Law*, 2014 U. ILL. L. REV. 175.

B. Copyright will continue to protect works of law scholarship, including annotations, that have not been officially endorsed by the state

Copyright plays a vital role in our economic system by providing incentives for creators, including, where applicable, creators of commentary on the law. Just as there is no controversy in this case concerning whether the government edicts doctrine covers statutes or judicial opinions, there is no controversy over whether copyright protection covers legal commentary created by private actors. The dispute in this case is over statutory annotations generated by state action with the three hallmarks of government edicts prescribed by the Eleventh Circuit. Drawing the line between state and private actors clearly demarcates

copyright boundaries that both serve the People and incentivize creative works by any individual or group.

CONCLUSION

For the foregoing reasons, the Court should uphold the Eleventh Circuit's decision in this case.

Respectfully submitted,

MICHAEL A. LIVERMORE
Counsel of Record
Professor of Law
UNIVERSITY OF VIRGINIA SCHOOL OF LAW
580 Massie Road
Charlottesville, VA 22903
Tel: (434) 982-6224
mlivermore@virginia.edu

CHARLOTTE S. ALEXANDER
Associate Professor of Law and Analytics
J. MACK ROBINSON COLLEGE OF BUSINESS
GEORGIA STATE UNIVERSITY
35 Broad Street, NW
Atlanta, GA 30303
Tel: (404) 413-7468
calexander@gsu.edu

ANNE M. TUCKER
Professor of Law
GEORGIA STATE UNIVERSITY COLLEGE OF LAW
85 Park Place, NE
Atlanta, GA 30303
Tel: (404) 413-9179
amtucker@gsu.edu