1st. Civ. No. A150625

IN THE COURT OF APPEAL FOR THE STATE OF CALIFORNIA
FIRST APPELLATE DISTRICT, DIVISION 3

RICHARD SANDER and THE FIRST AMENDMENT COALITION,
a California non-profit corporation

Petitioners and Appellants,

v.

STATE BAR OF CALIFORNIA and
BOARD OF GOVERNORS OF THE STATE BAR,

Defendants and Respondents.

DWIGHT AARONS, et al.

Intervenors.

Appeal from the San Francisco Superior Court,
The Honorable Mary E. Wiss
Case No. CPF-08-508880

BRIEF *AMICUS CURIAE* OF ELECTRONIC FRONTIER FOUNDATION IN SUPPORT OF APPELLANTS

JENNIFER LYNCH (SBN 240701)
AARON MACKEY (SBN 286647)
ADAM SCHWARTZ (SBN 309491)
jlynch@eff.org
ELECTRONIC FRONTIER
FOUNDATION
815 Eddy Street
San Francisco, CA 94109
Telephone: (415) 436-9333
Facsimile: (415) 436-9993

Counsel for Amicus Curiae

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Cases

Statutes

Constitutional Provisions

Legislative Materials

Other Authorities

INTRODUCTION

Modern governments generate and consume vast amounts of digital data about members of the public. This implicates two fundamental rights, both explicitly recognized by the California State Constitution. First, the public has a right to access this data. This right enables the public to understand what its government is doing with this data and to use this data to expose government inefficiency or malfeasance. Second, the people described by this sensitive government data have a right to privacy. When these oceans of digital data are made public without protecting people's privacy, intimate details of their lives could be exposed to unwanted scrutiny by others. These two fundamental rights must be carefully balanced.

Across the country, data custodians are innovating ways to share digital data in a manner that advances access to information without invading privacy. A critical method is known as anonymization (or de-identification). This means that before data custodians release a dataset, they strip it of information that could be used to identify the people whose lives are described by the data. A growing field of scholarship is creating best practices to de-identify data, and government agencies in California are already using anonymization techniques.

In the case at bar, this Court must decide whether the California Public Records Act (CPRA) will advance or impede the use of anonymization and other sophisticated privacy-protecting techniques to ensure the proper balance between the competing interests of government transparency and individual privacy. After all, the "CPRA and the Constitution strike a careful balance between public access and personal privacy." (*City of San Jose v. Super. Ct.* (2017) 2 Cal.5th 608, 616.)

Unfortunately, the superior court made several erroneous legal rulings that stacked the deck against the proper use of anonymization to advance the fundamental rights at issue in this case. *Amicus curiae* Electronic Frontier Foundation respectfully urges this Court to correct these errors.

First, the court failed to properly apply the balancing tests under the CPRA's privacy exemption in Government Code § 6254(c),[1] and its catch-all exemption in § 6255. It ignored the significant public interest in the disclosure of state bar data, which the California Supreme Court recognized in an earlier stage of this case. (*Sander v. State Bar of Cal.* (2013) 58 Cal.4th 300.) Also, it improperly placed the burden of proof on the requester to show that disclosure would *not* harm privacy, as opposed to where that burden belongs: on the government, to show that disclosure clearly *would* harm privacy.

Second, the court improperly interpreted Business & Professions (B&P) Code § 6060.25 to create an absolute bar on the disclosure of state bar data. This is incorrect as a matter of law, given the California constitutional mandate to construe statutes narrowly if they limit the right of access. (Cal. Const., art. I, § 3, subd. (b), par. (2).) If this Court holds otherwise, it should be clear that this absolute bar applies only to § 6060.25, and not to other CPRA exemptions.

Third, the court improperly held that anonymizing data to protect privacy creates new records. Not so. Anonymization is simply a modern way of redacting exempt information from otherwise non-exempt records that the CPRA requires government to release.

---

[1] All statutory citations are to the Government Code unless otherwise specifically stated.

The California Supreme Court recently ruled: "Our case law recognizes that the CPRA should be interpreted in light of modern technological realities." (*ACLU Found. of S. Cal. v. Super. Ct.* (2017) 3 Cal.5th 1032, 1041.) But the lower court's cramped interpretation of the CPRA goes against this mandate. It is likely to have far-reaching consequences that frustrate access to vast amounts of government digital data in which the public has a legitimate interest. This Court must correct these errors.

ARGUMENT

I. With the Explosive Growth of Government Data, the CPRA Must Be Interpreted to Provide Broad Access to Data-Rich Records

With the growth of Internet-enabled technologies, the amount of data created worldwide each year continues to surpass the years before.[2] Government data has followed this trend as state agencies and local governments have invested in data collection, utilization, and management and are finding ways to digitize older records to make them more useful to the public. As data and data-driven algorithms increasingly become an integral part of how government agencies function, public access to that data is essential to government accountability and oversight. Although the CPRA was enacted before electronic records were as prevalent as they are today, the legislature has continued to update the law to ensure it allows access to electronic records and data sets. Local agencies in California have also established policies that promote access to the massive amount of data

[2] *See* Åse Dragland, *Big Data, for better or worse: 90% of world's data generated over last two years*, Science Daily (May 22, 2013) https://www.sciencedaily.com/releases/2013/05/130522085217.htm. All websites last visited on January 29, 2018.

that governments generate. Further, some agencies have developed anonymizing practices that promote transparency while protecting privacy.

These policies should inform how agencies process individual requests under the CPRA. In this age of explosive data creation and utilization, it would be antithetical to the original purpose of the CPRA to allow government agencies to shield their datasets and data practices from the public, merely by refusing to grapple with the consequences of creating sensitive data about identifiable people that are also public records. Instead, agencies must meaningfully evaluate and adopt new technologies that will allow them to release records in ways that properly balance privacy and transparency.

A. As Modern Data Capabilities Have Grown, Agencies Are Finding New Ways to Open Up Access to Their Records, Including Records Containing Highly Sensitive and Private Data

1. *State and Local Governments Generate Vast Amounts of Data Each Year*

State and local agencies are producing and collecting data at an unprecedented rate. As part of an effort to make the increased data collection more transparent, the California Legislature passed Senate Bill 272 in 2016. The law requires local agencies, with the exception of school districts, to post catalogs of their databases on their websites. (§ 6270.5.) These catalogs are meant to disclose all the data systems an agency uses as a primary source of records or to collect information about the public. The catalogs not only are important government records in their own right, they can serve as a menu of records that members of the public may request under the CPRA.

As cities, counties, and individual agencies have put their database catalogs online, the public has learned just how many data systems they each maintain.[3] Each agency's catalog may include a handful to several hundred databases, which in turn include countless records. For example, the City and County of San Francisco lists 463 individual data systems, with approximately half of the systems updated with new data either daily or continuously.[4]

It is not surprising that cities and counties maintain so many datasets, because so much of the information agencies rely on to do their work has either been digitized or is now digital throughout its lifecycle. For example, many counties and cities have digitized older vital and official records, such as birth and death certificates and property records, to make them searchable, both for the benefit of agency employees and for the general public.[5] The state has also taken steps to offer more services to the public online, such as the Department of Motor Vehicles' website that allows Californians to renew their vehicle registration or driver's license

---

[3] EFF lists available government digital datasets in a single linkable document that now includes 443 California local agencies. *(See California Database Catalogs 2016*, EFF, https://www.eff.org/pages/california-database-catalogs-2016.)

[4] *Inventory of citywide enterprise systems of record*, DataSF, https://data.sfgov.org/City-Management-and-Ethics/Inventory-of-citywide-enterprise-systems-of-record/ebux-gcnq/data.

[5] *See, e.g.*, *2017 was a busy year for the Recorder-Clerk*, The ARC Blog (Jan 10, 2018) https://sbcountyarcblog.org/2018/01/11/2017-busy-year-recorder-clerk/ (noting Recorder's Office digitized vital records, such as birth and death certificates, dating back to 1910 and property records dating back to 1958); Emily Alpert Reyes, *Many L.A. building records now just a few clicks away*, L.A. Times (June 18, 2015) http://www.latimes.com/local/lanow/la-me-ln-online-building-records-20150618-story.html.

and update their address information online.[6] Similarly, the Secretary of State has an online portal that allows businesses to file their required financial statements electronically.[7] These services not only allow the public to provide information to the state and conduct state business from the comfort of their own home, they also allow the public to search through the information provided to the state by others. Digital services, and transparency around those services, create increased efficiencies for connecting the public with the day-to-day operations of their government.

2. *State Agencies Are Finding Ways to Make Sensitive Data Available to the Public While Still Protecting Privacy*

As more and more data is collected and stored digitally, California agencies are finding ways to make even highly sensitive information available to the public without disclosing the identity of the people described by that information. For example, the California Health and Human Services Agency (CHHS) has established de-identification guidelines to allow public access to certain datasets while still complying with all laws and patient protections, including the de-identification requirements in the federal Health Insurance Portability and Accountability Act.[8] The CHHS open data portal provides de-identified records on the rates of certain diseases, such as the number of cases of antibiotic-resistant staph infections for each hospital in a given year;[9] the demographics of

---

[6] California Department of Motor Vehicles, Online Portal, https://www.dmv.ca.gov/portal/dmv/detail/online/onlinesvcs.
[7] Bizfile California, http://www.sos.ca.gov/business-programs/bizfile/.
[8] *Data De-identification Guidelines (DDG)*, CHHS (Nov. 22, 2016) http://www.dhcs.ca.gov/dataandstats/Documents/DHCS-DDG-V2.0-120116.pdf.
[9] *Methicillin-resistant Staphylococcus aureus bloodstream Infections (MRSA BSI) in Healthcare*, CHHS Open Data,

individuals who access healthcare services, including applicants for insurance affordability programs by country of origin;[10] and how transportation and other environmental factors may impact a person's health, including a comprehensive database of motor vehicle accidents leading to death or serious injury, with the individual incidents and races of the victims given a numerical value to prevent victim re-identification.[11]

The de-identification guidelines published by CHHS advocate for a case-by-case approach when creating an open dataset. In doing so, the agency recognizes that individual record-level data has a greater probability of including sensitive information than anonymized summary-level data.[12] This case-by-case approach increases the possibility for appropriately balancing access and privacy, as exemplified by the motor vehicle accident database described above that substitutes sensitive personal information with an anonymous value.

The de-identification guidelines, along with CHHS's initiative to make healthcare datasets publically available in an open-source format, show that government agencies are able, at every stage of the data management process, to balance the interests of individual privacy with the community's interest for access.

Similarly, the City of San Francisco has created a privacy-protective balancing test to determine how city data should be released to the public.

---

https://data.chhs.ca.gov/dataset/methicillin-resistant-staphylococcus-aureus-bloodstream-infections-mrsa-bsi-in-healthcare.

[10] *Ethnicity of Applicants for Insurance Affordability Programs*, CHHS Open Data, https://data.chhs.ca.gov/dataset/dhcs_ethnicity-of-applicants-for-insurance-affordability-programs.

[11] *Road Traffic Injuries 2002-2010*, CHHS Open Data, https://data.chhs.ca.gov/dataset/road-traffic-injuries-2002-2010.

[12] *Id*.

Following a 2009 directive,[13] all agencies must publish the data they collect and process on a single website for the public to access. Also, the city created a rubric to prioritize the publication of highly requested data, even where some privacy concerns may exist, before less publicly interesting data.[14] Further the city's Open Data Release Toolkit, which provides steps on how to release data, suggests masking identifying variables, like name or ID number; obscuring quasi-identifying information, like race and age; and undertaking other methods of de-identification when appropriate.[15]

While datasets that are published through the programs discussed above have been prioritized for public accessibility, many other datasets, like the state bar's database at issue here, contain records that are unquestionably of interest to the public but are not maintained in a way that would allow for automatic release without risking severe privacy harm. The public should not have to wait for government agencies to voluntarily modify and publish those datasets online. Instead, state agencies must find ways to disclose the data now in response to individual public records requests.

---

[13] S.F. Mayor Gavin Newsom, *San Francisco Government and Technology: How We're Innovating*, Mashable (Oct. 21, 2009). http://mashable.com/2009/10/21/san-francisco-government/#kX_VXug3cOqp.) San Francisco's Open Data Portal can be found at https://datasf.org/.

[14] S.F. Chief Data Officer Joy Bonaguro, *How to Unstick Your Open Data Publishing*, SFData (last updated July 7, 2015), https://datasf.org/blog/how-to-unstick-data-publishing/.

[15] *Open Data Release Toolkit: Privacy Edition v1.2*, DataSF, https://docs.google.com/document/d/1MhvEuGKFuGY2vLcNqiXBsPjCzxYebe4dJicRWe6gf_s/edit#heading=h.v77s0yo7ojk7.

B. Releasing Public Data Can Increase Government Oversight and Make Agencies More Efficient

The California Supreme Court squarely recognized the strong public interest in access to state bar data, holding "it seems beyond dispute that the public has a legitimate interest in whether different groups of applicants, based on race, sex or ethnicity, perform differently on the bar examination and whether any disparities in performance are the result of the admissions process or of other factors." (*Sander,* 58 Cal.4th at 324.) Access to the granular data contained in the state bar's databases—even in anonymized form—would allow the public to evaluate government activity on a much deeper level than mere access to summary data about the bar's programs. More broadly, public access to government data holds the potential to provide more accountability and to identify ways state and local agencies can better serve the public.

For example, researchers at Stanford University recently accessed Oakland Police Department (OPD) reports and body-worn camera data from thousands of routine police traffic stops in 2013 and 2014.[16] Their goal was to analyze interactions with the community and suggest points in need of reform.[17] In reviewing the records, they were able to find significant statistical variations between how OPD officers treat white and

---

[16] Stanford researchers reviewed the body camera data as part of their work to assist OPD in complying with court-mandated monitoring imposed as a result of a 2003 case alleging widespread police misconduct within the department. (*See* Clifton Parker, *Stanford big data study finds racial disparities in Oakland, Calif., police behavior, offers solutions*, Stanford News (June 15, 2016) https://news.stanford.edu/2016/06/15/stanford-big-data-study-finds-racial-disparities-oakland-calif-police-behavior-offers-solutions/.)

[17] *Id.*

African American suspects.[18] Although the data was not released pursuant to the CPRA, it shows how reviewing public data can serve to inform the public and lead to greater oversight.[19]

Similar to the bar records at issue in this case, the footage and other data provided to Stanford researchers likely contained tens of thousands of instances where personal information was revealed, such as a persons' face during a stop and the name that they provide when requested by a police officer. But the insights researchers gained from these records show the value of access to granular data. The hope is that the analysis of this data will help lead to reform efforts, including reducing police-community tensions and enhancing public safety.[20]

In myriad other contexts, the release of granular municipal data has the potential to increase community well-being and the ability to petition local government for services. For example, the City of Los Angeles has released more than 1,000 data sets to the public, including historic and

---

[18] Alex Shashkevich, *Police officers speak less respectfully to black residents than to white residents, Stanford researchers find*, Stanford News (June 5, 2017) https://news.stanford.edu/press-releases/2017/06/05/cops-speak-less-ommunity-members/

[19] The researchers had access to unredacted body camera video and audio as a result of the federal litigation. But if the same records were sought through a public records request, the CPRA's balancing test would apply, and would likely require redaction of at least some personally identifying information to protect privacy.

[20] Shashkevich, *Police officers speak less respectfully to black residents than to white residents, Stanford researchers find*, Stanford News; Clifton Parker, *Data can help rebuild police-community relationships, Stanford expert says*, Stanford News (July 19, 2016) https://news.stanford.edu/2016/07/19/data-can-help-rebuild-police-community-relationships-stanford-expert-says/.

current data on affordable housing construction projects,[21] the response rate for every incident in which the Los Angeles Fire Department is dispatched,[22] and a list of parking tickets showing the address or even longitude and latitude of the citation.[23] By making data available, residents not only have visibility into government decisions and spending projects, but now have reliable and authoritative statistics that they can bring to city planning meetings to advocate for better resources in their area. This data may help public officials respond more quickly to problems and make better informed decisions.

In light of the government oversight that public access to this data allows, agencies have an obligation to the public to make the data available in some form. Agencies can do this in ways that minimize harms to individual privacy by following the guidelines set out in the state's open data programs or by adopting anonymization protocols.

C. The California Legislature Demonstrated its Commitment to Greater Public Access to Government Data in its CPRA Amendments

The vast increase in data, along with the policy determination that this data must remain accessible to the public, has prompted the state legislature to increase access to digital records by amending the CPRA.

---

[21] *HCIDLA Affordable Housing Projects Catalog and Listing (2003 To Present)*, DataLA, (last updated Dec. 5, 2017) https://data.lacity.org/A-Livable-and-Sustainable-City/HCIDLA-Affordable-Housing-Projects-Catalog-And-Lis/u4mj-cwbz.

[22] *LAFD_ResponseMetrics_RawData*, DataLA (last updated Jan. 13, 2017) https://data.lacity.org/A-Safe-City/LAFD_ResponseMetrics_RawData/cthf-nngn.

[23] *Parking Citations*, DataLA (last updated Jan. 5, 2018), https://data.lacity.org/A-Well-Run-City/Parking-Citations/wjz9-h9np.

For example, in passing Assembly Bill No. 2799 in 2000, the state legislature made clear its intention to update the CPRA to ensure robust access to digital data. The bill amended § 6253 and added § 6253.9 to ensure that the public had a clear right to request electronic data, while also giving agencies the tools they need to accommodate those requests. (*See* Legislative Counsel's Digest, AB 2799 (Shelley).)[24] Agencies must make government data available "in any electronic format in which it holds the information." (§ 6253.9(a)(1).) Further, to the extent that an agency must construct a record or compile or extract data to fulfill a request, the CPRA states that the agency may have additional time to respond to the request and the requester "shall bear the costs of producing a copy of the record." (§§ 6253(c)(4); 6253.9(b).) The legislature thus recognized that the public is entitled to access the growing amounts of data being amassed by public agencies, while still giving those agencies the ability to recoup costs for responding to CPRA requests that require more complex redaction and production.

II. The Superior Court Misapplied the Balancing Tests Under Sections 6254(c) and 6255 and Misinterpreted Business & Professions Code Section 6060.25

By ignoring the recognized and substantial public interests in disclosure of state bar data, and instead determining that the privacy interests in this case predominate over all other issues, the lower court failed to properly apply the CPRA's balancing tests under §§ 6254(c) and 6255. (Super. Ct. Opn. at pp. 13-14, 17.) It also improperly interpreted B&P Code § 6060.25 to create an absolute bar against the disclosure of

[24] Available at http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=199920000AB2799.

state bar data. (Super. Ct. Opn. at p. 10.) This Court should reverse these legal errors.

A. Sections 6254(c) and 6255 Require Courts to Identify and Weigh Competing Interests Before Determining Whether the Government Proved that Privacy Clearly Outweighs Disclosure

The CPRA's privacy and catch-all exemptions require more careful analysis than the superior court conducted. The Supreme Court held it is "beyond dispute that the public has a legitimate interest in whether different groups of applicants, based on race, sex or ethnicity, perform differently on the bar examination and whether any disparities in performance are the result of the admissions process or of other factors." (*Sander*, 58 Cal.4th at p. 324.) This recognition of the strong public interest in the disclosure of state bar data creates a firm presumption in favor of disclosure. And under the CPRA, the burden is on the government to overcome this presumption. (*See Long Beach Police Officers Assn. v. City of Long Beach* (2014) 59 Cal.4th 59, 67 (rejecting a blanket prohibition on disclosing police officer's names in light of the public interest in disclosure).) However, instead of applying the Supreme Court's ruling and analyzing the competing interests as required by CPRA, the lower court improperly placed the burden on Petitioners to show an absence of harm. In so doing, the court undercut the CPRA's presumption of disclosure.

The balancing test courts must use under §§ 6254(c) and 6255 are the same. (*Los Angeles Unified School District v. Super. Ct.* (2014) 228 Cal.App.4th 222, 240.) With respect to § 6254(c), disclosure is required unless the information would compromise substantial privacy interests that would amount to a clearly unwarranted invasion of personal privacy. (*Versaci v. Super. Ct.* (2005) 127 Cal.App.4th 805, 818.) Under § 6255,

disclosure is required unless those seeking to withhold the records demonstrate a "clear overbalance on the side of confidentiality." (*City of San Jose v. Super. Ct.* (1999) 74 Cal.App.4th 1008, 1022 (internal quotation marks omitted).)

1. *The Superior Court Failed to Properly Analyze the Records Sought Under Section 6254(c)*

By not explaining how the privacy interests at issue in this case sufficiently override the legitimate public interest in disclosure, the superior court failed to properly apply the balancing test under § 6254(c). When examining withholdings under § 6254(c), courts must determine the precise nature of the privacy interests at stake in the particular records and categorize them as either *de minimus* or substantial (sometimes referred to as significant). (*Versaci*, 127 Cal.App.4th at p. 818.) This determination is a fact-specific "assessment of the extent and gravity of the privacy invasion under consideration," making it difficult to draw bright lines around particular categories of information. (*Teamster Local 856 v. Priceless, LLC* (2003) 112 Cal.App.4th 1500, 1519.)[25]

After assessing the degree of privacy invasion present in any records, courts next identify the public interest in disclosure of the information. Courts must ignore the specific requester's motives and focus instead on whether disclosure of the information would shed light on government activity or the conduct of public officials. (*BRV, Inc. v. Super. Ct.* (2006) 143 Cal.App.4th 742, 755-57; *ACLU of N. Cal. v. Super. Ct.* (2011) 202 Cal.App.4th 55, 69-70.) This analysis focuses on the relative

[25] Criticized on other grounds by *Int'l Fed'n of Prof'l & Technical Eng'rs v. Super. Ct.* (2007) 42 Cal.4th 319, 335 (distinguishing *Teamsters Local* based on the limited record in that case).

weight of the public's interest in the sought records, including whether it is significant. (*Long Beach Police Officers Assn.*, 59 Cal.4th at p. 74 (categorizing the public's interest in police officer's conduct as "significant" and "particularly great").)

With both the public and privacy interests categorized, courts then balance the competing interests to determine whether the interest in protecting privacy clearly outweighs the public's interest in disclosure. (*BRV, Inc.*, 143 Cal.App.4th at p. 755.) Assuming there are legitimate public interests in disclosure, the CPRA requires disclosure unless the privacy concerns clearly override them. (*Id.* at p. 756.) The party seeking to withhold records—the government agency—bears the burden to demonstrate this clear overbalance. (*Id.*)

In this case, the superior court failed to follow this test in two respects. First, it did not give appropriate weight to the public interest in disclosure, even though the Supreme Court had recently held that this interest is substantial. (Super. Ct. Opn. at p. 12; *Sander*, 58 Cal.4th at pp. 324-25.) The superior court acknowledged the Supreme Court's ruling, but it did not recognize that the legitimate public interest established a presumption that the records at issue should be disclosed under the CPRA. (*BRV, Inc.*, 143 Cal.App.4th at p. 755.) The court failed to explain how disclosure of the records in one of Petitioners' proposed privacy-protective formats constituted a clearly unwarranted invasion of privacy that overrode the public interest in disclosure. (Super. Ct. Opn. at pp. 12-13.) The court thus legally erred by failing to properly apply § 6254(c)'s balancing test.

Second, the court improperly placed the burden on Petitioners by requiring them to prove an absence, or at least a severe mitigation, of the privacy interests in non-disclosure. (Super. Ct. Opn. at pp. 13-14.) In

essence, the court required Petitioners not only to show there was zero risk that a bar applicant's identity would be disclosed directly through release of the bar records, but also that there was zero risk applicants could be re-identified by comparing the anonymized bar data at issue with other publicly-available non-bar data. This was incorrect under the CPRA. This legal error infected the superior court's entire analysis under § 6254(c) and foreclosed Petitioners' opportunity to attempt to show that their proposed protocols undercut the potential privacy harm and tilted the balance toward disclosure.

2. *The Court's Analysis Under Section 6255 Contains Similar Errors*

The court also failed to properly apply the balancing test under § 6255. The Supreme Court held that there is a substantial public interest in disclosure of the records at stake in this case. (*Sander*, 58 Cal.4th at pp. 324-25.). However, the superior court below did not analyze whether that interest was clearly outbalanced by the interests in nondisclosure. (*See* § 6255(a); *City of San Jose*, 2 Cal.5th at p. 616 (The "CPRA and the Constitution strike a careful balance between public access and personal privacy.").) In particular, although the court below stated that "the public interest in nondisclosure clearly outweighs the public interest in disclosure of the records in this case," (Super. Ct. Opn. at p. 17.), it failed to support this statement by conducting the analysis necessary to show how Petitioners' proposed anonymization protocols would not address the privacy interests presented by the State Bar. This is insufficient, given the well-recognized constitutional requirement that a statute, court rule, or other authority must be "broadly construed if it furthers the people's right of access, and narrowly construed if it limits the right of access." (Cal.

Const., art I, § 3, subd. (b), par. (2).)

For example, although the court found that the release of information "presents significant risk of identification" (Super. Ct. Opn. at p. 17.), it failed to explain how that risk clearly outweighs the public's ability to "independently ascertain and evaluate" the State Bar's examination and admissions process as a result of public disclosure. (*Sander*, 58 Cal.4th at p. 324.) The court below thus similarly diminished the public's interest in disclosure under § 6255 just at it had during its analysis of § 6254(c).

Further, the court failed to give proper weight to Respondent's ability to redact or make use of Petitioner's proposed protocols to address Respondent's interest in nondisclosure. Section 6255 does not bar disclosure when there is a significant interest in disclosure of records and the countervailing concerns can be addressed by redactions or by deleting portions of the record. (*CBS, Inc. v. Block* (1986) 42 Cal.3d 646, 655.) Because § 6253(a) requires release of any segregable portions of public records, if Petitioners demonstrate they can address the public interest in nondisclosure through their proposed protocols, the catch-all exemption would not prevent disclosure here. (*See L.A. Cty. Bd. of Supervisors v. Super. Ct.* (2016) 2 Cal.5th 282, 292 (holding CPRA "requires public agencies to use the equivalent of a surgical scalpel to separate those portions of a record subject to disclosure from privileged portions").)

Finally, in separately ruling that the data disclosed under Petitioner's proposed Protocol 3 would provide limited utility, the superior court repeated the error of its analysis under § 6254(c). First, the court's decision failed to account for whether Protocol 3 minimized the countervailing interests in nondisclosure. Second, its misjudgment about the supposedly low value of the data requested under Protocol 3 runs fundamentally

counter to both the Supreme Court's previous decision in *Sander* that there is a substantial public interest in disclosure of the disputed records, and also the legislature's clear policy choice to allow CPRA requesters to determine for themselves the usefulness of government data. As multiple courts have held, the motives of requesters and the purposes for which they would use records are irrelevant under the CPRA; the "question instead is whether disclosure serves the public interest." (*County of Santa Clara v. Super. Ct.* (2009) 170 Cal.App.4th 1301, 1324; *see also Caldecott v. Super. Ct.* (2015) 243 Cal.App.4th 212, 219; *Los Angeles Unified School District*, 228 Cal.App.4th at p. 242.) Here, Petitioners proposed Protocol 3 because they believe it provides useful records while sufficiently addresses the countervailing interests in nondisclosure. The superior court's insertion of its own assessment of the requested data's utility was thus reversible error.

B. The Superior Court Improperly Held that Business & Professions Code Section 6060.25 Absolutely Bars Disclosure of State Bar Data

The superior court's interpretation of B&P Code § 6060.25 appears to foreclose release of records even if there is only the most minimal risk that, after applying privacy-protective measures, such records might identify individual applicants. In effect, this interpretation creates an absolute bar to the disclosure of state bar data,[26] which would violate Californians' constitutional right of access to government records.

The court's interpretation of B&P Code § 6060.25 must be guided by the "constitutional imperative" to construe CPRA exemptions in a manner that furthers disclosure. (*ACLU Found. of S. Cal.*, 3 Cal.5th at p.

[26] Respondents argued B&P Code § 6060.25 "absolutely prohibits" the disclosure of state bar records. (Super. Ct. Opn. at 10.)

1039 (internal quotation marks omitted).) The constitution requires all statutes that purport to limit access to government records—including those that do so to protect countervailing privacy interests—to be narrowly construed. (Cal. Const., art. I, § 3, subd. (b), par. (2); *see also ACLU Found. of S. Cal.,* 3 Cal.5th at pp. 1039, 1042 (noting that even exemptions designed to protect privacy and ensure safety must be construed narrowly).) In failing both to acknowledge the possibility that bar data could be disclosed in such a way as to protect the identity of applicants, and to analyze whether Petitioners' proposed protocols would accomplish that, the superior court failed its constitutional duty to narrowly construe and apply the statute. (*Id*.)

The California Legislature, in direct response to this case, drafted B&P Code § 6060.25. This statute states that state bar records "*that may identify* an individual applicant, shall be confidential and shall not be disclosed pursuant to any state law, including, but not limited to, the California Public Records Act." (*Id.* (emphasis added).) As required by the California Constitution,[27] the legislature introduced findings that this exemption to the CPRA was necessary to "protect the privacy interests of those persons submitting information to the State Bar" and "to ensure that any personal or sensitive information is protected as confidential information." (Senate Bill No. 387, Sec. 25.)[28] However, neither the plain language of the statute nor the legislature's findings justify the absolute withholding of data. Rather, if privacy interests can be protected by anonymizing the data, then the statute cannot mandate nondisclosure.

---

[27] *See* Cal. Const., art. I, § 3, subd. (b), par. (2).

[28] Available at https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160SB387.

Moreover, the superior court's interpretation of B&P Code § 6060.25 runs counter to the well-recognized CPRA principle that promises of confidentiality cannot completely foreclose public access. The Supreme Court in *Sander* reaffirmed that promises of confidentiality recognized under Evidence Code § 1040 do not automatically require withholding of such records under the CPRA. (58 Cal.4th at p. 325.) Rather, the CPRA may still require disclosure of confidential records where portions of the records do not identify individuals—and thus do not violate the confidentiality promise—or where the CPRA's balancing test tilts in favor of public access. (*Id.*) As discussed above, both reasons apply in this case; there are ways to de-identify the data to satisfy B&P Code § 6060.25, and the interests in privacy must be balanced against the public interest in disclosure.

Were this Court to adopt the superior court's interpretation of B&P Code § 6060.25, it could set a precedent that all the legislature needs to do to exempt any public records that contain private, identifying information from disclosure is to pass a statute with similar language and findings. This would contradict the Constitution's requirement that CPRA exemptions must be construed narrowly in favor of public access. Also, it would fail to account for the ability of anonymization and other privacy-protecting techniques to prevent the release of the private and sensitive information at issue.

Even if this Court agrees with the superior court that B&P Code § 6060.25 bars the release of the specific data at issue here, it should make clear that the standard for withholding under this statute is distinct from the standards under Government Code §§ 6254(c) and 6255. The lower court's decision, on the other hand, appears to extend a restrictive legal standard

for withholding records under B&P Code § 6060.25 to the balancing tests under the CPRA's privacy and catch-all exemptions. This was incorrect.

B&P Code § 6060.25's language is narrower than that of either the CPRA's privacy or catch-all exemptions. (*See ACLU of N. Cal.*, 202 Cal.App.4th at pp. 69-70.) Therefore, applying these exemptions' balancing tests to the disputed records should not necessarily produce the same result. To the extent that the superior court applied the same test for all of the CPRA exemptions, this Court should not allow that legal error to stand. Left unchecked, that aspect of the superior court's ruling could be applied in future cases that involve the release of personally-identifying data to which B&P Code § 6060.25 does not apply, prompting courts to allow agencies to withhold personal data if there is any risk of re-identification at all. This could frustrate public access to records in which there is a significant public interest in disclosure, based solely on unsubstantiated fears that disclosure may somehow identify individuals.

In sum, the superior court's interpretation of B&P Code § 6060.25 is incorrect as a matter of law. But if this Court chooses to uphold the lower court's interpretation of this statute, it should avoid improper expansion of other CPRA exemptions by cabining the lower court's decision on B&P Code § 6060.25 to just the narrow category of records at issue in this case, *i.e.*, state bar records that may identify an individual applicant.

III. Requiring Agencies to Anonymize Records Does Not Result in the Creation of New Records

The superior court denied the petition because, among other reasons, the CPRA supposedly "does not require public agencies to create new records in order to respond to a records request." (Super. Ct. Opn. at p. 8.) This justification fails for two reasons. First, neither the plain language of

the CPRA nor the case law cited by the superior court supports its position. Second, the protocols proposed by Petitioners do not require the State Bar to create new records.

The text of the CPRA contains no express or implied limitation on agencies' duties to extract information from public records, or even to manipulate information within records, to withhold exempt information while making non-exempt information available to the public. Instead, the CPRA expressly anticipates that such extraction and manipulation will be necessary in some cases and that agencies may pass along to requesters the corresponding costs "to construct a record, and the cost of programming and computer services necessary to produce a copy of the record." (§ 6253.9.) Likewise, such extraction and manipulation is contemplated by the CPRA's rule that agencies may delay their responses based on the "need to compile data, to write programming language or a computer program, or to construct a computer report to extract data." (§ 6253.) Further, as all parties recognize, California statutory law and the California constitution provide broad access rights to the public and presume that all documents maintained by a public entity are subject to disclosure unless an exemption applies.

The superior court cited two cases to justify its position that agencies cannot be required under the CPRA to create new records from their existing data and information: *Haynie v. Super. Ct.* (2001) 26 Cal.4th 1061, and *Fredericks v. Super. Ct.* (2015) 233 Cal.App.4th 209. Neither case supports this proposition.

Haynie is factually distinguishable. In that case, the Supreme Court held only that an agency was not required to produce a list of potentially responsive records as part of its initial response to Mr. Haynie's request.

The court did not rule on anything beyond that or even address its own power to order the agency to prepare such a list once the petition for writ of mandate was filed. (*Haynie*, 26 Cal. 4th at pp. 1061, 1073.) In *Sander*, on the other hand, Petitioners have never asked the State Bar for a list of potentially responsive records but instead have asked it to apply an anonymization protocol to existing records.

Fredericks, which relies on *Haynie*, is similarly inapplicable because its discussion about agencies' duties under the CPRA was dicta. *Fredericks* addressed the scope of an agency's duties to extract information from law enforcement investigative files pursuant to § 6254(f)(2), a statute that specifically requires agencies to extract and produce certain delineated categories of information from police complaints or requests for assistance. (*Id.*) *Fredericks'* statement—that requiring agencies to create new records in response to a public records request would exceed the agencies' statutory duties under the CPRA—was dicta because the court never had to address whether the petitioner's request would require the agency to create a new record. Also, *Sander* does not involve extracting information from investigative files, so *Fredericks* is not on point.

Even if this Court were to agree with the superior court that the CPRA does not require agencies to create new records, that point is moot because the privacy protocols proposed by Petitioners do not require the State Bar to create new records. Instead, they require the Bar to manipulate existing public records to produce data in a format that serves the public interest in government transparency (by disclosing non-exempt information) while at the same time protecting the privacy interests of state bar applicants (by withholding exempt information). Again, the CPRA explicitly contemplates such data manipulation by allowing agencies to

charge requesters reasonable costs for producing records that require "data compilation, extraction, or programming." (§ 6253.9(b)(2).)

Other cases, involving both the CPRA and the federal Freedom of Information Act, are more closely aligned with Petitioner's request here than either *Haynie* or *Fredericks*. These courts have held that agencies may be required to manipulate existing records and databases to extract information sought by a public records requestor and that doing so does not create a new record. For example, in discussing possible anonymization protocols, the California Supreme Court recently recognized that agencies may be required to manipulate their data so that they can release records while still protecting privacy. The Court stated, "While real parties may not have designed their system to facilitate CPRA disclosure as a 'native function,' randomizing [data] or deleting columns from a spreadsheet, for example, would seem to impose little burden." (*ACLU Found. of S. Cal.*, 3 Cal.5th at p. 1047.)

Other cases are also on point. In *CBS Broadcasting Inc. v. Super. Ct.* (2001) 91 Cal.App.4th 892, the Court of Appeal ordered the California Department of Social Services to go through its records and compile and produce accurate lists of individuals granted a criminal conviction exemption to work in licensed child day care facilities and the identity of each facility employing such individuals.

In *May v. Department of Air Force* (5th Cir. 1986) 800 F.2d 1402, a federal appellate court addressed records maintained as handwritten forms that, if released, could reveal the identity of the author based on the distinctive style of the handwriting. The court held the Air Force could create a typewritten copy of the records or recreate them in a third-party's hand to protect the identity of the author. The court held, "such disclosure

would . . . ensure maximum disclosure under the [FOIA], and not unreasonably burden the agency." (*Id.* at p. 1403.)

In *Schladetsch v. H.U.D.* (D.D.C. 2000, No. 99-0175) [2000 WL 33372125], , the court held that neither the programming necessary to instruct a computer to conduct a search, nor the process of extracting and compiling the data resulting from such a search, constitute the creation of a new record. (*Id.* at p. *3.)

And in *Disabled Officer's Association v. Rumsfeld* (D.D.C. 1977) 428 F.Supp. 454, the court held that the fact that the agencies "may have to search numerous records to comply with the request and that the net result of complying with the request will be a document the agency did not previously possess is not unusual in FOIA cases nor does it preclude the applicability" of FOIA. (*Id.* at p. 456.)

The protocols proposed by Petitioners are more appropriately compared to requirements to redact exempted information from otherwise non-exempt records, which are well-established under both the CPRA and FOIA. (§ 6253(a); 5 U.S.C. § 552(a)(8)(A).) Redaction within records promotes the goals of the CPRA because it allows for the maximum disclosure of information while still protecting other interests. For example, in *CBS, Inc. v. Block*, the court held that releasing applications for concealed weapon licenses but deleting certain confidential information from those applications protected the privacy of applicants while still ensuring the public was provided with enough information to determine whether public officials were acting properly in issuing licenses for legitimate reasons. (*CBS, Inc.*, 42 Cal.3d at p. 655.)

Techniques for protecting exempt information while still releasing otherwise non-exempt government records that are of great interest to the

public must evolve as the government's means of collecting, compiling, and maintaining such records has evolved. Protocols that propose to anonymize data, such as those presented by Petitioners, represent one such technique. California courts should not avoid a determination of whether anonymization can protect privacy by dismissing it out of hand as the creation of a "new record."

CONCLUSION

For the foregoing reasons, *amicus* respectfully requests that this Court reverse the superior court's legal rulings and require it to properly apply the CPRA.

Dated: January 31, 2018

Respectfully submitted,

/s/ Jennifer Lynch
Jennifer Lynch

Aaron Mackey
Adam Schwartz
ELECTRONIC FRONTIER FOUNDATION

Counsel for Amicus Curiae

CERTIFICATE OF WORD COUNT

I certify pursuant to California Rules of Court 8.204 and 8.504(d) that this Amicus Brief of Electronic Frontier Foundation is proportionally spaced, has a typeface of 13 points or more, contains 6,913 words, excluding the cover, the tables, the signature block, verification, and this certificate, which is less than the total number of words permitted by the Rules of Court. Counsel relies on the word count of the Microsoft Word word-processing program used to prepare this brief.

Dated: January 31, 2018

/s/ Jennifer Lynch
Jennifer Lynch

ELECTRONIC FRONTIER FOUNDATION

Counsel for Amicus Curiae

CERTIFICATE OF SERVICE

STATE OF CALIFORNIA, COUNTY OF SAN FRANCISCO

I am over the age of 18 years and not a party to the within action. My business address is 815 Eddy Street, San Francisco, California 94109.

On January 31, 2018, I served the foregoing document entitled:

BRIEF *AMICUS CURIAE* OF ELECTRONIC FRONTIER FOUNDATION IN SUPPORT OF APPELLANTS

on the parties in this action on the attached Service List:

X BY ELECTRONIC TRANMISTION VIA TRUEFILING: I caused a copy of the foregoing documents to be sent via TrueFiling to the persons at the e-mail addresses listed in the Service List. The following parties and/or counsel of record are designated for electronic service in this matter on the TrueFiling website. I did not receive, within a reasonable time after the transmission, any electronic message or other indication that the transmission was unsuccessful.

X BY FIRST CLASS MAIL: I placed the envelope for collection and mailing following our ordinary business practices. I am readily familiar with this firm's practice for collecting and processing correspondence for mailing. On the same day that correspondence is placed for collection and mailing, it is deposited in the ordinary course of business with the United States Postal Service, in a sealed envelope with postage fully prepaid.

I declare under penalty of perjury under the laws of the State of Calif ornia that the foregoing is true and correct.

Executed on January 31, 2018 at San Francisco, California.

Madeleine Mulkern

SERVICE LIST

Via E-File Service

Jean-Paul Jassy
Kevin Lester Viek
Jassy Vick Carolan LLP
800 Wilshire Boulevard, Suite 800
Los Angeles, CA 90017

Attorneys for Richard Sander: Plaintiff and Appellant

James M. Chadwick *Via E-File Service*
Sheppard, Mullin, Richter & Hampton
Four Embarcadero Center, 17th Floor
San Francisco, CA 94111-4109

Andrea Nicole Feathers
Sheppard Mullin Richter & Hampton LLP
333 S Hope Street 43rd Floor
Los Angeles, CA 90071

Guylyn Remmenga Cummins
Sheppard Mullin Richter & Hampton LLP
501 W Broadway, 19th Floor
San Diego, CA 92101

David Edward Snyder
California First Amendment Coalition
534 4th Street, Suite B
San Rafael, CA 94901

Attorneys for First Amendment Coalition: Plaintiff and Appellant

James M. Wagstaffe *Via E-File Service*
Michael John VonLoewenfeldt
Melissa Perry
Kerr & Wagstaffe LLP
101 Mission Street, 18th Floor
San Francisco, CA 94105

Vanessa Lynne Holton
Destie Lee Overpeck
Office of General Counsel
State Bar of California
180 Howard Street
San Francisco, CA 94105

Attorneys for State Bar of California: Defendant and Respondent

Via E-File Service

William Frederick Abrams
David Howard Kwasniewski
Steptoe & Johnson LLP
1891 Page Mill Road, Suite 200
Palo Alto, CA 94304

Attorneys for Dwight Aarons: Intervener and Respondent

Via E-File Service

Robert E. Steinbuch
University of Arkansas
at Little Rock Law School
1201 McMath Ave.,
Little Rock, AR 72202

Richard J. Peltz-Steele
University of Massachusetts Law School
333 Faunce Corner Road
North Dartmouth, MA 02747

Attorneys for Amicus Curiae National Association of Scholars

Via E-File Service

Supreme Court of California
350 McAllister Street
San Francisco, CA 94102

Via First Class Mail

San Francisco Superior Court
Civic Center Courthouse
400 McAllister Street
San Francisco, CA 94102