BEFORE THE PUBLIC UTILITIES COMMISSION OF THE

STATE OF CALIFORNIA

Order Instituting Rulemaking on Regulations Relating to Passenger Carriers, Ridesharing, And New On-Line-Enabled Transportation Services.	R.12-12-011 (Filed December 20, 2012)

JOINT OPENING COMMENTS OF THE SAN FRANCISCO MUNICIPAL TRANSPORTATION AGENCY AND THE SAN FRANCISCO COUNTY TRANSPORTATION AUTHORITY ON THE PROPOSED DECISION REQUIRING TRANSPORTATION NETWORK COMPANIES TO SUBMIT THEIR ANNUAL REPORTS FOR THE YEARS 2014-2019 TO THE COMMISSION WITH LIMITED REDACTIONS

TILLY CHANG
Executive Director
San Francisco County Transportation Authority
1455 Market Street, 22nd Floor
San Francisco, CA 94103
(415) 522-4832
tilly.chang@sfcta.org

JEFFREY P. TUMLIN
Director of Transportation
San Francisco Municipal Transportation Agency
One South Van Ness Avenue, 7th Floor
San Francisco, CA 94103
(415) 646-2522
jeffrey.tumlin@sfmta.com

Dated: November 29, 2023

SUBJECT INDEX OF RECOMMENDED CHANGES TO THE PROPOSED DECISION

The proposed decision ("PD") should be modified as follows:

- The Commission should revise the PD to require precise reporting of timestamp data in the Annual Reports as the decision to aggregate timestamp data to the nearest 30-minute interval is unsupported under the California Public Records Act and abrogates the people's right to access public records.

TABLE OF AUTHORITIES

Cases

Statutes

Other Authorities

.

Pursuant to California Public Utilities Commission Rule of Practice and Procedure 14.3(a), the San Francisco Municipal Transportation Agency and the San Francisco County Transportation Authority (together "San Francisco") submit these Joint Opening Comments on the Proposed Decision Requiring Transportation Network Companies to Submit Their Annual Reports for the Years 2014-2019 to the Commission with Limited Redactions (the "Proposed Decision" or "PD") filed by the California Public Utilities Commission (the "Commission" or "CPUC") on November 9, 2023.

I. INTRODUCTION

San Francisco supports the Proposed Decision's determination that the Transportation Network Company ("TNC") Annual Reports for reporting years 2014 to 2019 should no longer be afforded the presumption of confidentiality provided in footnote 42 in Decision ("D.")13-09-045[1] and the Proposed Decision's conclusion that the TNCs have failed to carry their burden of proving that the trip data at issue should be shielded from public disclosure on privacy or trade secret grounds.[2] However, San Francisco strongly opposes the Proposed Decision's determination to aggregate timestamp data to the nearest 30-minute interval. The decision to aggregate timestamp data constitutes a legal error under the California Public Records Act ("CPRA") as the Proposed Decision does not establish that any exemption applies to warrant abrogating the people's right of access.[3] Further, timestamp aggregation is inconsistent with, and contrary to, the Proposed Decision's conclusion that the trip data does not warrant protection on privacy grounds.

[1] Proposed Decision at 17.

[2] *Id.* at 53 ("we conclude that except for the information identified above in the table, the balance of the trip data in the Annual Reports from 2014-2019 is not protected from disclosure on privacy grounds"); *id.* at 56 (". . . we, again, reject the argument that trip data and other information in the Annual Reports for the years 2014-2019 is trade secret protected.")

[3] Cal. Const., Article I, § 3(b)(2).

II. DISCUSSION

A. The PD Commits Clear Legal Error by not Establishing That an Exemption Applies to Warrant Withholding Precise Timestamp Data

1. The Commission Must Justify its Decision to Aggregate Timestamp Data Under the CPRA.

Under the California Constitution, the public has a "fundamental and necessary" right to access public records.[4] The CPRA presumes that records are public, unless those records are "exempt from disclosure by express provisions of law."[5] An agency must "justify withholding any record by demonstrating that the record in question is exempt under express provisions of [the CPRA], or that on the facts of the particular case the public interest served by not disclosing the record clearly outweighs the public interest served by disclosure of the record."[6] The public agency claiming the right to withhold information carries the burden of demonstrating the need for nondisclosure of a public record.[7] Mere assertions of possible endangerment do not clearly outweigh the public interest in access to public records.[8] As the Proposed Decision accurately states:

> The California Constitution also states that statutes, court rules, and other authority limiting access to information must be broadly construed if they further the people's right of access, and narrowly construed if they limit the right of access. Rules that limit the right of access must be adopted with findings demonstrating the interest protected by the limitation and the need for protecting that interest.[9]

Contrary to these established principles, the Proposed Decision contains no analysis or findings supporting the decision to aggregate timestamp data – data which the Commission has held to be a

[4] Cal. Const. Article I, § 3; Cal. Gov't. Code § 7921.000.

[5] Cal. Gov't. Code § 7922.530(a).

[6] Cal. Gov't Code § 7922.000.

[7] *See*, *e.g.*, *Braun v. City of Taft*, 154 Cal.App.3d 332, 345 (Ct. App. 1984).

[8] *New York Times Co. v. Superior Ct.*, 218 Cal.App.3d 1579, 1585 (Ct. App. 1990) (finding that agency water district did not carry its burden of justifying the withholding of the names and addresses of water users as the mere assertion of possible endangerment did not "clearly outweigh" the public's interest in access to records) citing *CBS, Inc. v. Block*, 42 Cal.3d 646, 652 (1986).

[9] Proposed Decision at 18 citing Cal. Const., Article I, § 3(b)(2).

public record in its unredacted form.[10,11] The Proposed Decision simply asserts, without justification, that aggregating timestamp data "strike[s] a balance between promoting public use of trip data while protecting personal privacy." But these privacy risks are never explained. This is clear legal error.

2. The Decision to Aggregate Timestamp Data to Protect Privacy is Internally Inconsistent with the PD's Finding that the Trip Data *Does Not* Implicate Privacy Concerns.

In Section 3.2 of the PD, in deciding that the Annual Reports for years 2014-2019 should be released to the public, the PD considers whether trip data, which includes timestamp data, should be withheld from disclosure on privacy grounds. After engaging in extensive analysis,[12] the PD determines, consistent with the Commission's treatment of the 2020 and 2021 Annual Reports, that the trip data *does not* implicate privacy concerns.[13] The PD finds that neither Lyft nor Uber have met their burden of establishing that the trip data at issue includes a legally protected privacy interest, and explicitly rejects the TNC's claims that the trip data can be reidentified to reveal personal information about a rider as "speculative at best."[14] Moreover, in rejecting the TNC's privacy arguments, the PD itself admits that "the Commission has already taken steps to permit the redaction of trip data information likely to infringe on driver and passenger privacy interests."[15] These findings are incompatible with a decision to aggregate timestamp data on privacy grounds.

---

[10] *See Assigned Administrative Law Judge's Ruling on Uber Technologies, Inc.'s and Lyft's Motion for Confidential Treatment of Certain Information in Their 2020 Annual Reports* (December 21, 2020). CPUC Rulemaking 12-12-011. <https://docs.cpuc.ca.gov/PublishedDocs/Efile/G000/M355/K738/355738454.pdf> [as of Dec. 21, 2020].

[11] *Assigned Administrative Law Judge's Ruling Granting, in Part, the Motions of Uber Technologies, Inc., Lyft, Inc., HopSkipDrive, Inc., and Nomad Transit, LLC for Confidential Treatment of Portions of Their 2021 Annual Transportation Network Company Reports* (November 24, 2021). CPUC Rulemaking 12-12-11. <https://docs.cpuc.ca.gov/PublishedDocs/Efile/G000/M425/K517/425517150.pdf> [as of Nov. 24, 2021].

[12] Proposed Decision at 30-53.

[13] *Id.* at 52.

[14] *Id.* at 53 ("Finally, Lyft and Uber fail to establish that the disclosure of the trip data would be a serious invasion of privacy. As noted above, the claims that the trip data can be reidentified to reveal personal information about a rider's politics, religious beliefs, sexual orientation, or medical status are speculative at best.")

[15] *Id.* at 45.

Later in Section 7, the PD relies on the analysis of authors Ann Cavoukian and Daniel Castro to reject arguments from Lyft regarding data aggregation.[16] The PD notes that Castro and Cavoukian posit that data anonymization can be successful if it addresses three privacy risks – 1) "data aggregating must protect an individual's records from being uniquely identified in the dataset;" 2) "data aggregation must prevent an individual's records from being linked to other datasets;" and 3) "data aggregation must make it difficult to infer sensitive information about an individual."[17] The PD then explains how the public versions of the Annual Reports address these three risks:

> First, the public versions of Annual Reports do not contain any unique identifiers for each passenger. Neither names nor code names are used for a passenger's trips. Thus, someone reviewing the dataset would not be able to tell all the times that an individual passenger made use of the TNC passenger service. Second, no information is provided about an individual passenger trip that would allow that information to be linked to other datasets. The Annual Reports do not contain gender information, dates of birth, or other data that would permit such linkages. Third, nothing is required in the public version of the Annual Reports that would allow a third party to determine sensitive information about an individual. The usual examples that parties' offer in support of their objection to the public disclosure of trip data is that it can be manipulated to determine a passenger's sexual predisposition or political party affiliation, determine if a passenger is going to an abortion clinic, or if a passenger is going to conduct an illicit assignation. But as the Annual Reports do not contain latitude and longitude, one cannot tell by a zip code if a passenger is going to or coming from such a sensitive location.[18]

As can be seen from the above, by the PD's own analysis, aggregation of timestamp data is unnecessary to address potential data reidentification under the Castro and Cavoukian framework as the public versions of the Annual Reports do not contain unique identifiers, the data cannot be linked to other data sets, and granular location data is already withheld which prevents third parties from determining sensitive information. Similar to the findings in Section 3.2, this analysis is incompatible with a decision to withhold unaggregated timestamp data – and on the contrary, shows that aggregation is unnecessary to protect privacy.

[16] *Id.* at 109.

[17] *Id.* at 110.

[18] *Id.* at 110-111.

3. The PD's Reliance on the City of Chicago's Aggregation Practices is Unsupported.

The only clear basis that San Francisco can ascertain for the decision to aggregate timestamp data is the Commission's independent review of how the City of Chicago aggregates timestamp data for TNC trips.[19] The City of Chicago aggregates timestamp data to the nearest 15-minute interval.[20] The PD notes that Lyft and Uber have not, in seven years since Chicago adopted this practice, reported any breaches of passenger privacy to the Commission, and then summarily concludes that timestamp data should be aggregated to 30-minute intervals.[21] The decision to aggregate on this basis should be reversed for at least two reasons. First, the PD does not explain why the 15-minute Chicago standard needs to be doubled in order to afford California passengers the same level of privacy as Chicago passengers when the PD itself explains that no breaches of privacy have been reported at the shorter 15-minute interval. More importantly, a description of another jurisdiction's data aggregation practice, without more, does not carry the Commission's burden under the CPRA of justifying that an express exemption exists to allow withholding of precise timestamp data (a public record) or show that on the facts of this particular case the public interest served by not disclosing the record clearly outweighs the public interest served by disclosure of the record.[22] Therefore, the decision to aggregate based on Chicago's practice, without more, is insufficient under the CPRA.

Second, as discussed in San Francisco's Opening Comments on the Ruling Reopening the Record on Timestamp Data, New York City's Taxi and Limousine Commission ("NYC TLC") publishes precise timestamp data monthly with a three-month lag between reporting updates, and "[n]either Uber nor Lyft have cited any issues arising from NYC's requirement in this rulemaking, despite collectively reporting 780 million trips there."[23] Yet the PD does not explain why the approach utilized by NYC TLC is insufficient to protect privacy when the rationale for adopting the Chicago approach—a lack of reported breaches of privacy—is equally applicable to NYC TLC's practice.

[19] Proposed Decision at 107-108.

[20] *Id.* at 108.

[21] *Id.* at 109.

[22] Cal. Gov't Code § 7922.000.

[23] *See* Proposed Decision at 100 citing San Francisco Opening Comments on the Ruling Reopening the Record on Timestamp Data at 6.

Since the PD's stated rationale for the Chicago approach is as applicable to NYC TLC's non-aggregation approach, but NYC TLC's approach would better serve the CPRA's policy of "maximum disclosure,"[24] the decision to aggregate based on Chicago's approach is again inconsistent with the CPRA.

B. The Decision to Aggregate Timestamp Data at 30-Minute Intervals will Severely Impact the Public's Use of the TNC Trip Data.

As explained in San Francisco's Opening Comments on the Ruling Reopening the Record on Timestamp Aggregation, timestamp data at high-levels of precision enables a host of functions which serve the public interests highlighted by the PD,[25] including travel demand modeling, curb passenger loading capacity and other active curb management decisions, and traffic assignment models.[26] Of particular note, lack of precise timestamp data severely impacts the public's ability to ascertain data quality issues within the TNC Annual Reports, hindering quality control and quality assurance. Data whose precision has been artificially lowered may obscure errors in the underlying data and may make them impossible to detect. Precise timestamp data can reveal inconsistencies in the data that may not be apparent in less precise form.

Additionally, while the PD only deals with data from reporting years 2014-2019, an inability to verify data would present an especially acute problem in the TNC Access for All context if the decision were extended prospectively. As explained in our prior comments, the TNC Access for All program collects fees on all TNC trips in California to fund on-demand wheelchair accessible vehicle ("WAV") service, identifies response times (the time between when a trip is requested and the vehicle arrives to pick up the passenger) as a key performance metric, and requires TNC WAV service to meet response time standards in order to qualify for reimbursement of the cost of providing WAV service. Those standards are benchmarked against response times for non-WAV service. If timestamp data is produced with low precision, then response times cannot be accurately calculated. This limits the ability of the Commission and interested parties to validate TNCs' self-reported response times, which

[24] *See*, *e.g.*, *New York Times Co. v. Superior Ct.*, 218 Cal.App.3d 1579, 1585 (Ct. App. 1990).

[25] Proposed Decision at 74.

[26] San Francisco Opening Comments on the Ruling Reopening the Record on Timestamp Data at 3-4.

are the basis for administering reimbursements from the multimillion-dollar TNC Access for All program funds. It also limits the ability of the Commission and the public to gain insight into trends in this key metric over time, between companies, and between WAV and non-WAV service. Lack of access to high precision data would inhibit the public's ability to ascertain whether persons with disabilities have equal access to TNC rides – a public interest highlighted in the PD.[27]

III. CONCLUSION

San Francisco appreciates the opportunity to provide comments on the Proposed Decision and supports the Proposed Decision's determination that the TNC Annual Reports for reporting years 2014 to 2019 should no longer be afforded the presumption of confidentiality provided in footnote 42 in D.13-09-045.[28] However, San Francisco strongly opposes the Proposed Decision's determination to aggregate timestamp data to the nearest 30-minute interval. The decision to aggregate timestamp data rests on several legal errors under the CPRA, and improperly undermines the people's right of access. For this reason, San Francisco opposes adoption of the Proposed Decision unless the Commission amends the Proposed Decision to require precise timestamp reporting.

Dated: November 29, 2023

Respectfully submitted,

DAVID CHIU
City Attorney
LILLIAN A. LEVY
Deputy City Attorney
(415) 554-3876
Lillian.Levy@sfcityatty.org

By: /s/*Lillian A. Levy*
LILLIAN A. LEVY

On behalf of: THE SAN FRANCISCO MUNICIPAL TRANSPORTATION AGENCY AND SAN FRANCISCO COUNTY TRANSPORTATION AUTHORITY

[27] Proposed Decision at 88-90.

[28] *Id.* at 17.

APPENDIX

San Francisco recommends the following changes to the Proposed Decision's conclusions of law and ordering paragraphs. Additions are in red, and deletions are struck through.

Proposed Amendments to Conclusions of Law

16. It is reasonable to conclude that the Commission should require each TNC to submit its public Annual Report data with precise ~~all~~ timestamps as the TNC's have failed to carry their burden of proving that precise timestamp data is protected from disclosure on privacy grounds. ~~aggregated to the nearest 30-minute interval in order to strike a balance between promoting public use of trip data while protecting personal privacy.~~

Proposed Amendments to Ordering Paragraphs

3. The following categories of trip data shall be disclosed, for each ride provided, as part of each Transportation Network Company's public version of its Annual Reports for the years 2014-2019. All other data fields may be marked "Redacted" per Ordering Paragraph 1:

- Trip Requester Zip Code;
- Driver Zip Code;
- Trip Request Date/Time ~~(aggregated to the nearest 30-minute interval)~~;
- Miles Traveled (Period 1);
- Request Accepted Date/Time ~~(aggregated to the nearest 30-minute interval)~~;
- Request Accepted Zip Code;
- Passenger Pick Date/Time ~~(aggregated to the nearest 30-minute interval)~~;
- Miles Traveled (Period 2);
- Passenger Pick Up Zip Code;
- Passenger Drop Off Date/Time ~~(aggregated to the nearest 30-minute interval)~~;
- Passenger Drop Off Zip Code;
- Miles Traveled (Period 3); and
- Total Amount Paid.