IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF VIRGINIA
Richmond Division

UNITED STATES OF AMERICA	)	
	)	
v.	)	**Case No. 3:19cr130**
	)	
OKELLO T. CHATRIE,	)	
Defendant	)	

DEFENDANT'S MOTION TO SEAL RAW DATA RETURNS PROVIDED BY GOOGLE IN RESPONSE TO THE GEOFENCE WARRANT

COMES NOW the defendant, Okello Chatrie, by counsel, and for the following reasons, requests this Court to seal the attached raw data returns (Exhibit A) provided by Google in response to the geofence warrant obtained by government, ECF. No. 54.

BACKGROUND

The geofence warrant in this case presents an issue of first impression in this district, and in the country. At the Court's status conference on December 12, 2019, the parties and the Court discussed that the raw data returns relating to the geofence warrant should be entered on the record in order to advise the Court of the nature of the information at the heart of the upcoming discovery and suppression hearings, scheduled for January 21, 2020, and February 20, 2020, respectively. On December 17, 2019, the Court conducted a phone conference and the parties and the Court again discussed how best to provide the raw data returns to the Court on the record. At that time, the government indicated that it might be amenable to filing them under seal, subject to supervisory approval. The Court instructed the parties to confer and report back to the Court on December 19, 2019.

On December 19, 2019, the government filed a status update with the Court, *see* ECF No. 55, indicating that it would not file the raw data returns at this time, but would instead file them as part of a Notice to Introduce Evidence prior to the February 20, 2020, hearing. The same day, the defense filed a response stating that this data is critical to the Court's understanding of the nature of the search in this case, both for the discovery hearing on January 21, 2020, and the suppression hearing on February 20, 2020. ˆ*See* ECF No. 56. On December 23, 2019, the government replied, disagreeing that the Court already had an off-the-record copy of the raw return data for the geofence warrant. *See* ECF No. 61.

On December 30, 2019, the Court issued an order requiring counsel for both parties, and their supervisors, to appear for an in-person status conference on January 8, 2020, to address the disagreement expressed in the parties' filings. *See* ECF No. 65. On December 31, 2019, the parties conferred once again, at which point the government indicated that it would not oppose a defense motion to file the raw return data under seal, but without conceding the privacy issues that constitute the grounds for sealing. Consequently, and for the reasons described below, the defense now moves to seal the raw data returns relating to the geofence warrant.

ARGUMENT

The Fourth Circuit requires courts to do three things prior to sealing records: (1) provide public notice of the request to seal and allow interested parties a reasonable opportunity to object, (2) consider less drastic alternatives to sealing the documents, and (3) provide specific reasons and factual findings supporting its decision to seal the documents and for rejecting the alternatives. *Ashcraft v. Conoco, Inc.*, 218 F.3d 288, 302 (4th Cir. 2000); *see also Stone v. Univ. of Maryland Med. Sys. Corp.*, 855 F.2d 178, 181 (4th Cir. 1988); *In re Knight Pub. Co.*, 743 F.2d 231, 235 (4th Cir. 1984). All three factors are satisfied here.

First, the proposed order herein will provide notice and serve all interested persons with "an opportunity to object." *In re Knight Pub. Co.*, 743 F.2d at 235. Moreover, the government has already indicated that it will not oppose this motion. Second, as discussed below, good cause exists to seal the raw data returns at issue because they contain only pseudo-anonymous information that is both private and easily re-identifiable. Such information must be sealed to protect the privacy of bystanders caught up in the geofence warrant and no less drastic alternative will suffice. Third, publicizing the historic location information of bystanders or potential witnesses is injurious to their privacy; it is inherently identifiable and should be sealed.

The raw data returns at issue consist of three Excel spreadsheets and surrounding communications produced by Google pursuant to the geofence warrant. The warrant describes this data as "anonymized information," *see* ECF No. 54, because it does not include the names of Google users swept up in the search. Likewise, Google states that the "production version" of the data it provided to police does not contain any "account-identifying information." *See* ECF No. 59-1 at 18. But the geolocation coordinates for each device are real, and it is trivial to plot them on a map, connect the dots, and show the comings and goings of 19 people. In many cases, the location information alone is sufficient to deanonymize the data and identify individual users by name. The raw data here belongs to 19 people who went about their days and do not expect their private location information to become accessible to the general public.

The fact that Google masks the true "Device ID" with a pseudonym does not make the data anonymous. Internal consistency of the identifiers makes it possible to map the data and reveal each person's unique path through life. For instance, one individual travels from the bank to a high school down the road, and then back to a single-family residence, the owners of which are readily identifiable with a quick Google search. Another person drives by the church on their way home

from a hospital. And a third person registers in the bank, the nearby church, and then at a different bank before returning home. In each instance, it is trivial to match a name to the data. In short, this is pseudo-anonymous data, capable of being re-identified with ease.

Merely redacting the Device IDs would be insufficient to prevent re-identification given the unique nature of the geolocation data. Computer scientists have repeatedly shown that it is possible to "reidentify" or "deanonymize" individuals from ostensibly anonymous data. *See* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. Rev. 1701, 1716 (2010) (marshalling computer science research on the "science of reidentification" and arguing that it has taken the "robust anonymization assumption" and "essentially blown it up."). And as the *New York Times* recently demonstrated in spectacular fashion, "precise, longitudinal geolocation information is absolutely impossible to anonymize." Stuart A. Thompson & Charlie Warzel, *Twelve Million Phones, One Dataset, Zero Privacy*, N.Y. Times (Dec. 19, 2019) (quoting Paul Ohm).[1] The *Times* obtained a large dataset of mobile geolocation information and was able to identify and track celebrities, law enforcement officers, "high-powered lawyers (and their guests)," and even a Secret Service agent assigned to President Trump. *See id.*; Stuart A. Thompson & Charlie Warzel, *How to Track President Trump* (Dec. 20, 2019).[2] Indeed, the ability to map and analyze such supposedly "anonymous" geolocation information is precisely why law enforcement requested it from Google in this case. As a result, no lesser alternative to sealing is reasonable.[3]

[1] *Available at* https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html.
[2] *Available at* https://www.nytimes.com/interactive/2019/12/20/opinion/location-data-national-security.html.
[3] The third spreadsheet also contains the "Gaia ID" for three devices, *i.e.*, the "Google Accounts and ID Administration" ID, which is a number directly tied to specific Google users.

The raw data returns in this case constitute the real geolocation data of 19 people who happened to be near a crime. Recent privacy legislation, such as the California Consumer Privacy Act (CCPA), recognizes the inherent concerns in disclosing such information. The CCPA defines "personal information" not only as something that "identifies" a person, but also that which is "reasonably capable of being associated with" a person. Cal. Civ. Code § 1798.140(o)(1). Such information includes not just "a real name" or "postal address" but also "[g]eolocation data." Cal. Civ. Code § 1798.140(o)(1)(A)-(1)(G). Similarly, the European Union's (EU) General Data Protection Regulation (GDPR) defines "personal data" as "any information *relating to* an identified or *identifiable* natural person." Europ. Parl. and Coun. Reg. 2016/679, 2016 O.J. (L. 119) 1, Art. 4(1) (emphasis added). A person can be identified "directly or indirectly" not just by "name" but also "location data." *Id*. Even Google required law enforcement to obtain a warrant because of its statutory obligations to protect this data from disclosure. *See* ECF No. 59-1 at 14-17. The Court should therefore ensure it is not further disclosed publicly.

For these reasons, the defense moves to file the raw data returns from the geofence warrant under seal.

Respectfully submitted,

OKELLO T. CHATRIE

By: _____/s/_____

Michael W. Price
NY Bar No. 4771697 (pro hac vice)
Counsel for Defendant
National Association of Criminal Defense Lawyers
Fourth Amendment Center
1660 L St. NW, 12th Floor
Washington, D.C. 20036
Ph. (202) 465-7615
Fax (202) 872-8690
mprice@nacdl.org

/s/
Laura Koenig
Va. Bar No. 86840
Counsel for Defendant
Office of the Federal Public Defender
701 E Broad Street, Suite 3600
Richmond, VA 23219-1884
Ph. (804) 565-0881
Fax (804) 648-5033
laura_koenig@fd.org

CERTIFICATE OF SERVICE

I hereby certify that on January 3, 2020, I filed the foregoing with the Clerk of Court using the CM/ECF system, which will send a notification of such filing (NEF) to all counsel of record.

/s/
Laura Koenig
Va. Bar No. 86840
Counsel for Defendant
Office of the Federal Public Defender
701 E Broad Street, Suite 3600
Richmond, VA 23219-1884
Ph. (804) 565-0881
Fax (804) 648-5033
laura_koenig@fd.org