UBER_JCCP_MDL_002618039

Metadata

#Author	cescallon@uber.com	SEMANTIC
#Date Modified	05/12/2020	SEMANTIC
#DateCreated	10/30/2019	SEMANTIC
#Title	ACP - SRAD External socialization with experts	SEMANTIC
Account	brooke.anderson@uber.com;	SEMANTIC
All Custodians	Anderson, Brooke;Fuldner, Gus;Kansal, Sachin;Silver, Nick;	SEMANTIC
All Paths	Anderson, Brooke: \20240905\Anderson, Brooke\Drive\EDISCO-25640_DR_21.zip; Anderson, Brooke: \20240905\Anderson, Brooke\Drive\EDISCO-25640_DR_21.zip; Anderson, Brooke: \MassTort_Category4_DRIVE\MassTort_Category4_DRIVE_3.zip; Anderson, Brooke: \MassTort_Category4_DRIVE\MassTort_Category4_DRIVE_3.zip; Fuldner, Gus: \JCCP-EDISCO-23800_2019_to_2021\JCCP-EDISCO-23800_2019_to_2021_60.zip; Fuldner, Gus: \JCCP-EDISCO-23800_2019_to_2021\JCCP-EDISCO-23800_2019_to_2021_60.zip; Kansal, Sachin: \EDISCO-25937_sachin.kansal@uber.com\EDISCO-25937_sachin.kansal@uber.com2_56.zip; Kansal, Sachin: \EDISCO-25937_sachin.kansal@uber.com\EDISCO-25937_sachin.kansal@uber.com2_56.zip; Silver, Nick: \EDISCO-24394_Drive\EDISCO-24394_Drive_70.zip; Silver, Nick: \EDISCO-24394_Drive\EDISCO-24394_Drive_70.zip; Silver, Nick: \MassTort_Category2_DRIVE\MassTort_Category2_DRIVE_5.zip; Silver, Nick: \MassTort_Category2_DRIVE\MassTort_Category2_DRIVE_5.zip	SEMANTIC
Application	Microsoft 2007 Word Document	SEMANTIC
Begin Family	UBER_JCCP_MDL_002618039	SEMANTIC
Collaborators	hhydepage@uber.com; abyrne@uber.com; robert@igarape.org.br; akankshu@uber.com; usivaram@uber.com; gus@uber.com; gklinzman@uber.com; dsvirsky@uber.com; chelseamk@uber.com; brooke.anderson@uber.com; dkolta@uber.com; kallman@uber.com; sbryson@uber.com; binnings@uber.com; eboman@uber.com; sachin.kansal@uber.com; zrodriguez@uber.com	SEMANTIC
Confidentiality	Confidential	SEMANTIC
Date Created	10/30/2019 9:21 pm	SEMANTIC
Date Modified	05/12/2020 9:30 pm	SEMANTIC
Document Type	5/12/2020	SEMANTIC
End Family	UBER_JCCP_MDL_002618045	SEMANTIC
File Path	\20240905\Anderson, Brooke\Drive\EDISCO-25640_DR_21.zip	SEMANTIC
File Size	22591	SEMANTIC
Filename	ACP - SRAD External socialization with expert_1Qp3Ze9LuZpcVdVDkP5ZKvg_kQjp_X-VVYt_JjGo1VLc.docx	SEMANTIC
GoogleDocumentType	DOCUMENT	SEMANTIC
Hidden Content	Yes;	SEMANTIC
ILS All Bates	UBER_JCCP_MDL_002618039;UBER_JCCP_MDL_002618040;UBER_JCCP_MDL_002618041;UBER_JCCP_MDL_002618042;UBER_JCCP_MDL_002618043;UBER_JCCP_MDL_002618044;UBER_JCCP_MDL_002618045	SEMANTIC
ILS Prod Date	01/31/2025	SEMANTIC
ILS Prod Vol	JCCP_MDL083	SEMANTIC
LINKSOURCEBEGBATES	UBER_JCCP_MDL_001728645; UBER_JCCP_MDL_001728649	SEMANTIC
Other Custodians	Fuldner, Gus;Silver, Nick;Anderson, Brooke;Kansal, Sachin;	SEMANTIC
Primary Date	10/30/2019 9:21 pm	DOC_TYPE_ALIAS
Production Volume	JCCP_MDL083;JCCP_MDL157;	SEMANTIC
Redacted	No	SEMANTIC
Sort Date	05/12/2020 9:30 pm	SEMANTIC
SourceHash	1de147fb01eafc8344ee3e966a094dfd	SEMANTIC

Attorney Client Privileged / Attorney Work Product / Confidential

S-RAD External socialization with experts

Team:	In the doc:
Camila Escallon, Safety Public Policy	What's Happening
Dan Svirsky, Data Science	Key Messages
Daniel Kolta, Safety Legal	FAQ
Grant Klinzman, Safety Comms	Events with experts
Akankshu Dhawan, Safety Product	

WHAT'S HAPPENING

Uber will be holding sessions with experts in machine learning, civil liberties, and women's empowerment the week of November 19th. Based on the results of the work with our consultants ORCAA and Instituto Igarape, the sessions will focus on the safety impact. We will also cover the company's measures in terms of privacy and fairness as a demonstration of our responsible approach when it comes to machine learning for safety.

KEY MESSAGES

- **Safety -** Safety is at the heart of everything we do at Uber, and we are developing a new safety tool that builds on that commitment. The tool leverages data from billions of trips on the Uber platform to identify and prevent combinations of drivers and riders that may be more likely to result in a bad experience with the potential to escalate to an interpersonal conflict. While it only alters a very small percentage of rider and driver matches, our testing has shown promising results for improving both the user experience and platform safety.

- **Fairness -** Fairness is core to the Uber platform and something we take seriously. Working with top experts in the field we are conducting fairness testing to confirm that the technology does not lead to disparate outcomes by ethnicity or gender. This is a critical issue and we plan to test continuously to ensure the tool is fairly treating all users.

- **Privacy -** When you use Uber, you trust us with your personal data. We're committed to keeping that trust. We have taken steps to ensure we will be compliant with all applicable privacy and security laws, if the model is launched.

FAQs [Experts event Nov 19-21]

IMPORTANT NOTE: S-RAD is an internal name only. When referring to it externally please use the name "Safety Matching Tool" or "RideSense Matching" (NOT FINAL).

General

What is S-RAD?
S-RAD is a technology developed with rider and driver safety at its core. Drawing upon the billions of trips completed on the Uber platform, S-RAD seeks to identify and prevent combinations of drivers and riders that under certain circumstances may be more likely to result in a bad experience with the potential to escalate to an interpersonal conflict. S-RAD specifically looks at historical trip data, locational factors, and other driver and rider information. S-RAD is seamlessly integrated into the platform without changing the reliable Uber experience our users expect.

What is the safety effect?

For the past year and a half, Uber's safety team has been carefully exploring the potential of this technology. Pilots in Brazil and the US show promising results for improving both the user experience and platform safety, and we are increasingly confident the technology could be a critical part of our overall safety approach.

What type of personal data does S-RAD use?
S-RAD is narrowly tailored to only use the data necessary to deliver the greatest possible safety benefit, and is customized by country to optimize for effectiveness. Variable types include trip request location factors, historical trip data, and driver and rider information. Through extensive analysis, Uber has determined that use of gender data can meaningfully increase S-RAD's ability to prevent interpersonal conflicts. The model thus uses gender on the driver side, and inferred-gender on the rider side, as we do not collect rider gender. We understand the importance of dealing with gender information responsibly, and that is why we will allow our users to correct the information, opt out from its use, and we store the data responsibly. We commit to inferred rider gender being used exclusively for safety purposes.

What are other examples of variables used by S-RAD? How did you choose these variables?
S-RAD considers attributes of the location of the trip request, prior user feedback, prior user trip data, time of day, day of the week, driver partner gender as provided on their driver's license, and rider inferred gender based on name. Variables were selected based on their power to deliver safety and user experience benefits to the Uber platform, with significant consideration to privacy and fairness.

Do users have the ability to have S-RAD accurately reflect their gender?
At Uber we value diversity and respect every person's right to identify with the gender that best represents them. At this time, the model provides for the following gender alternatives: female, male, nonbinary/other, and prefer not to specify (used when users opt out). This mirrors the driver license system in 14 jurisdictions (examples include Oregon, California, New York City, and Washington DC) that provide for three gender categories.

We use the gender specified on the driver license for drivers, and we infer gender for riders. All users have the ability to find out what gender Uber is using for them in the model, to correct it as needed, and to request Uber no longer use their gender information.

How will S-RAD impact my user experience?
S-RAD will operate in the background and should not affect user experience. No rider or driver should ever be left without a trip due to S-RAD.

How does it help prevent sexual assault?
S-RAD seeks to identify combinations of drivers and riders that under certain circumstances may be more likely to result in a bad experience with the potential to escalate to an interpersonal conflict. Sexual and gender-based violence is a damaging and under-reported crime. While reported incidents to Uber are rare, they are devastating to victims and survivors. Because sexual and gender-based violence is associated with specific environmental and individual risk factors (e.g. time of day, exposure to alcohol), it may also be preventable. S-RAD is effective in helping to prevent sexual assault and sexual misconduct because it takes into account these contextual factors and optimizes matching to help prevent the types of scenarios that could give rise to an incident.

When are you launching this in the US?
We are doing some final product tests, assessing best practices in terms of privacy, and doing fairness testing to monitor for disparate impact. If the results of these are positive, we hope to be live in some US markets in the beginning of 2020.

How will you communicate this to riders and drivers?
This specific technology will be part of the RideSense portfolio, which we will introduce to the wider public prior to its launch. RideSense leverages information from Uber's safety platform to sense and detect potential risks through historical trip data and billions of trips, signals, and even technology that's built right into the device, such as GPS and telematics. All of this information combined is used to help predict, prevent and reduce safety incidents. Uber's RideSense technology is always on and constantly optimizing to inform its predictions and prevention capabilities in order to help reduce safety incidents.

Will you allow users to opt out of RideSense?
Because we have a high degree of confidence that RideSense makes our platform safer for its users, we do not think it would be appropriate to allow some users to opt out. Because S-RAD works at the dispatch level, the decision by one user to opt out could have an impact upon the people with which that user would be matched. We do understand that some users may prefer we not use their gender information in RideSense, so we allow users to opt out of our using that data.

If these users are dangerous, why are they still in your platform?
Uber maintains strict deactivation policies for those found to represent a safety risk to the platform. These policies represent an important piece of our overall safety strategy.

Through providing a platform for billions of trips, we have also found that certain combinations of drivers and riders under certain circumstances may be more likely to result in a bad experience or, in extremely rare situations, a safety incident. Leveraging Uber's historical data, S-RAD identifies these types of circumstances, and intervenes by blocking a small number of pairings. For example, if a driver tends to dislike and lowly rate trips leaving from bar districts at 2 AM, S-RAD may deprioritize such a trip request, in favor of one involving a different pickup address. Serious safety incidents are one in a million type events on the platform, and are incredibly difficult to anticipate. It is our hope that S-RAD can improve the overall user experience in enough situations where we can also help prevent some rare, serious safety incidents, from sexual assaults to physical altercations, from occurring all together.

Who has been supporting you in this process?
We brought in two external consultants to help us review our processes behind the algorithm, better understand the state of algorithmic transparency, and to figure out the best ways to communicate the project to experts and the general public. Internally, Uber has worked cross-functionally between product, ops, legal, public policy, and communications to ensure a responsible approach to this safety innovation. Uber is placing safety first, but we intend to do so in a way that respects privacy and is fair to all of our users.

If it reduces safety incidents, do you have an obligation to share this technology with your competitors?
We are always looking to lead on safety, and would be open to talking to others about our work in this space.

Privacy

What type of user data is being used in S-RAD?
S-RAD relies upon various types of data to assess what driver-rider pairings should be deprioritized, including attributes of the location of the trip request, prior user feedback, prior user trip data, time of day, day of week, driver's gender as provided on their driver's license, and rider inferred gender based on name.

S-RAD is also routinely subjected to testing to ensure the algorithm is fairly treating all users. In addition to using gender information for this purpose, Uber also infers ethnicity solely to help monitor and ensure the fairness of the algorithm. Users may access gender and ethnicity inferences made for them, and if desired, request Uber cease their use or change the inferences to better reflect their self-declared gender or ethnicity.

How does S-RAD use my data?
Uber's use of user data is guided by its Privacy Policy. This policy describes the information we collect, how it is used and shared, and your choices regarding this information. S-RAD specifically is an algorithm that uses data to help ensure you have a positive user experience when using Uber.

Will I be able to correct my gender or ethnicity?
Yes. To do so, please contact Uber at X@uber.com.

Will I be able to opt-out from the use of this data?
Yes. To do so, please email Uber at X@uber.com. Provide your name and phone number associated with your account, and specify whether you'd like Uber to cease use of the gender inference, ethnicity inference, or both.

Commented [1]: we should update this answer with the new information. +gklinzman@uber.com

Commented [2]: same

Where will this information be available?
This information will be available by contacting Uber at X@uber.com or through the Download Your Data tool.

What is Download Your Data?
Download Your Data (DYD) is a tool that allows users to understand what information Uber collects regarding them. The DYD website provides more information about the tool

Where can I learn more about your matching algorithms?
As part of our transparency plan for S-RAD, we intend to provide users with information about the safety feature.

How is data stored and managed?
Our safety team carefully manages and secures user data. We also take steps to ensure we are compliant with all applicable privacy and security laws.

Why isn't there the opportunity to opt-in? (by sharing gender with us)
When you use Uber, you trust us with your information. We are committed to keeping that trust. We believe in the power of using data in technology to help ensure the safety of the Uber platform. Use of gender data in S-RAD in particular has been shown to help enhance the user experience and user safety. However, if you do not wish Uber to use your gender information in this manner or if you wish to correct the data, please contact us at X@uber.com.

Fairness

Is this technology discriminatory? What type of fairness testing are you doing?
As of now, our assessment is that the technology does not lead to disparate outcomes by ethnicity or gender, but our testing is on-going, and the question is important enough that we plan to monitor it continuously.

We assess this question with a three-pronged approach, each prong answering a specific research question.

1. Does the data used to train the algorithm have ethnic disparities?
2. Does S-RAD lead to differences in earnings (or other market outcomes) for drivers across different genders or ethnicities?
3. Looking at the group of drivers most affected by S-RAD, does this group's ethnic composition look different from the ethnic composition of drivers as a whole?

Prong 1 is complete, with reassuring results. Among United States drivers in cities where the feature is being tested, roughly 10% of all drivers are African-American. Looking just at drivers who have been accused of a safety incident, or a sexual assault, the percentage of these drivers that are African-American is 9%. Hence, in contrast to stop-and-frisk data, or data on arrests, we do not see African-Americans being over-represented among the group of drivers accused of an incident.

Prongs 2 and 3 are underway now, but not yet complete. In the next month, we will conduct additional tests that will give us a more complete sense of whether the technology leads to ethnic or gender disparities in outcomes, like driver earnings or rider ETAs. We will test the feature using what we call a switchback experiment — turning the feature on and off at regular intervals — to assess its impact on the marketplace. This will let us test whether the feature affects driver earnings or rider wait times, and whether this impacts differ by ethnicity or gender. It will also let us hone in on drivers who are most affected by the feature, to see whether the ethnic composition of this group differs from drivers as a whole. In addition, there is an existing A/B experiment that assessed gender disparities in outcomes, which showed no disparate impact from the feature.

What are you doing to monitor and prevent disparate impact?
The research approaches described above are how we will monitor disparate impact. If we do find disparities, we will re-design the algorithm — dropping features that we suspect are driving any disparate impact — and then re-test the algorithm.

Why is Uber using ethnicity and gender data?

We use ethnicity data for only one purpose: to test for discrimination. While inferring or collecting ethnicity data is a sensitive matter, we felt that it was important to be able to test our compliance with antidiscrimination laws and make sure that we are doing the right thing.

We use gender data for only two purposes: to test for discrimination and for safety. For example, in some countries, like Saudi Arabia, we allow female drivers to choose to only have female riders. And in the context of S-RAD, we use gender to assess the risk of a sexual assault because academic research suggests that gender is an important determinant of risk.

Our internal policy is that gender and ethnicity data cannot be used for any other purposes without express approval from the Chief Privacy Officer.

Do you collect ethnicity data? How do you infer ethnicity?
Uber does not collect ethnicity data, so we infer ethnicity using name and location. The technique is widely used by academics in contexts from health care research (Eicheldinger & Bonito, 2008) to sociology (Gaddis, 2017) to labor economics (Bertrand & Mullainathan, 2004), as well as by the Department of Justice Civil Rights Division. The technique isn't perfect -- it gets many individual cases wrong. But it works well on aggregate populations when sample sizes are large enough. While we understand that such inferences are sensitive, and that they can make people uncomfortable, we ultimately felt that conducting antidiscrimination testing was important enough that it was the right thing to do.

What are other companies doing in this space?
Our conversations with external consultants, who work in this space and help companies design their fairness testing, is that there is very little work being done by our competitors.

How does Uber compare to other companies when it comes to algorithmic transparency?
Our approach to transparency for this algorithm will put Uber at the forefront of companies when it comes to algorithmic transparency. We plan to describe the algorithm, the process by which we trained and developed the algorithm, as well as results from fairness testing. This will give users a level of insight into how this feature works that is rare among tech companies. Some companies give broad descriptions of algorithms in an attempt to meet algorithmic transparency requirements, but what we propose here goes well beyond that.

Attorney Client Privileged / Attorney Work Product / Confidential

EVENTS WITH EXPERTS

Experts attending (full list of experts) *green = attended*

Boston (final):
- Ryan Budish, Harvard Berkman Klein [thank you note sent -camila]
- Diana Mancera, Jane Doe (Massachusetts Coalition Against SA and DV) [thank you note sent -camila]
- Sophie Hilgard, Harvard Data Science Initiative [thank you note sent -dan]
- Chelsea Barabas, MIT [thank you note sent -camila]
- Berk Ustun, Center for Research on Computation and Society, Harvard University

New York (final):
- Vincent Southerland, Center for Race Inequality and the Law (NYU) [thank you note sent -dan]
- Karen Levy, Cornell University [thank you note sent -camila]
- Elizabeth Anne Watkins, Data and Society Institute
- Esha Bhandari, ACLU (Speech, Privacy, and Technology Program)
- Amreeta Mathai, ACLU (Racial Justice Program)
- Roel Dobbe, AI Now Institute
- Amy Barasch, Her Justice [thank you note sent -dan]

DC (final):
- John Verdi, Future of Privacy Forum [thank you note sent -dan]
- Chad Sniffen, National Alliance to End Sexual Violence
- Erica Olsen, National Network to End Domestic Violence (Safety Net Initiative) [thank you note sent -dan]
- Spandana Singh, New America - Open Technology Institute
- Michelle Richardson, Center for Democracy and Technology
- Liz Woolery, Center for Democracy and Technology

Civil Liberties focused (NYC:
- Dariely Rodriguez, Lawyers' Committee for Civil Rights [thank you note sent -dan]
- Judge Shira Scheindlin [thank you note sent -dan]
- Faiza Patel, Brennan Center for Justice [thank you note sent -dan]
- Ames Grewart, Brennan Center for Justice [thank you note sent -dan]

Commented [3]: +zrodriguez@uber.com, Privacy groups (FPF, Open Technology Institute, Center for Democracy and Technology)

+ **Boston** (November 19th, 9:00 a.m. - 12:00 p.m.)
  Address: 239 Causeway Street, Boston, MA, 02114
  Room: 1st Floor, Boston Common (through door that says "Sanborn, Head & Associates")

+ **New York** (November 20th, 9:00 a.m. - 12:00 p.m.)
  Address: 1400 Broadway New York, NY
  Room: Jambalaya (32nd floor)

+ **DC** (November 21st, 9:00 a.m. - 12:00 p.m.)
  Address: 1717 Rhode Island Ave. NW
  Room: Air Force One (4th Floor)

REACTIVE COMMS STRATEGY

Attorney Client Privileged / Attorney Work Product / Confidential

+ Glassbreak comms plan