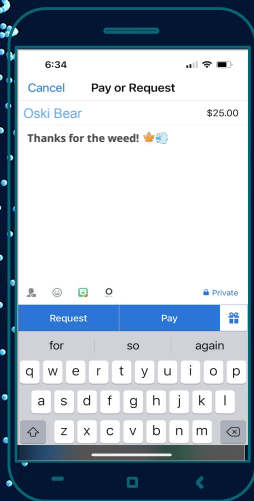


Applying Differential Privacy to Venmo Transactions

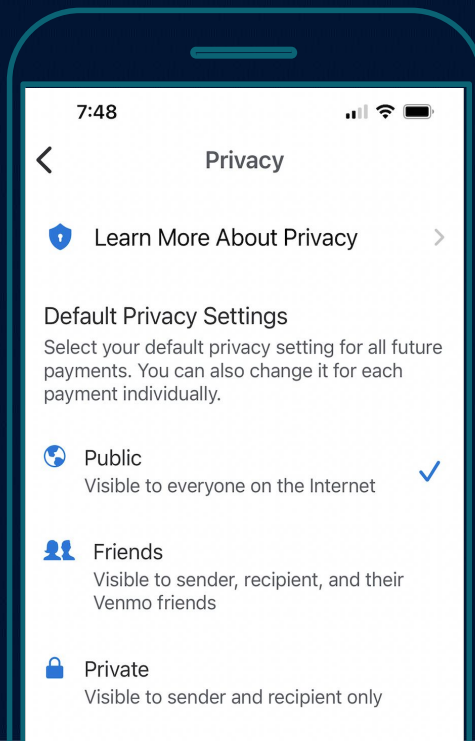
UC Berkeley MIDS and MICS
W233, Spring 2022

Ram Ben-David, Amanjot Samra,
Jordan Meyer, Kristy Edwards

BACKGROUND



Venmo UX with default
privacy settings:
“Public. Visible to everyone
on the Internet”



Scenario

Preparing data for Venmo's data
scientists

Find the best privacy/utility
trade-off

Link with another dataset
(Census data) which leads to
inference of sensitive attributes

	app.name	payment.note	payment.action	payment.date_created	id	payee	payee.last_name	recipient	recipient.last_name	et
20	Venmo for iPhone		charge	2018-08-07T02:11:13	2540404988111225470					
74	Venmo for iPhone		pay	2018-07-26T18:47:27	2532209097340617537					
79	Venmo for iPhone		pay	2018-07-26T18:47:28	2532209106912018852					
152	Venmo for iPhone		pay	2018-07-26T18:47:59	2532209366279389682					u
174	Venmo for iPhone		pay	2018-07-26T18:48:08	2532209440745062473					

RESEARCH

DAN SALMON SECURITY JUN 26, 2019 9:00 AM

I Scraped Millions of Venmo Payments. Your Data Is at Risk

Opinion: Venmo makes sending and receiving money a social affair. But those emoji-laden payment descriptions leave you exposed to cyberattacks.

WIRED

CNN BUSINESS

Markets Tech Media Success Video

A researcher studied a year of public Venmo transactions. Here's what she learned

by Kaya Yurieff @kyurieff
JULY 17, 2018, 5:31 AM ET

If you've used Venmo in the past year, there's a good chance Hang Do Thi Duc knows more about you than you might like.



FEDERAL TRADE COM
PROTECTING AMERICA'S

PayPal Settles FTC Charges that Venmo Failed to Disclose Information to Consumers About the Ability to Transfer Funds and Privacy Settings; Violated Gramm-Leach-Bliley Act

February 27, 2018



METHODOLOGY

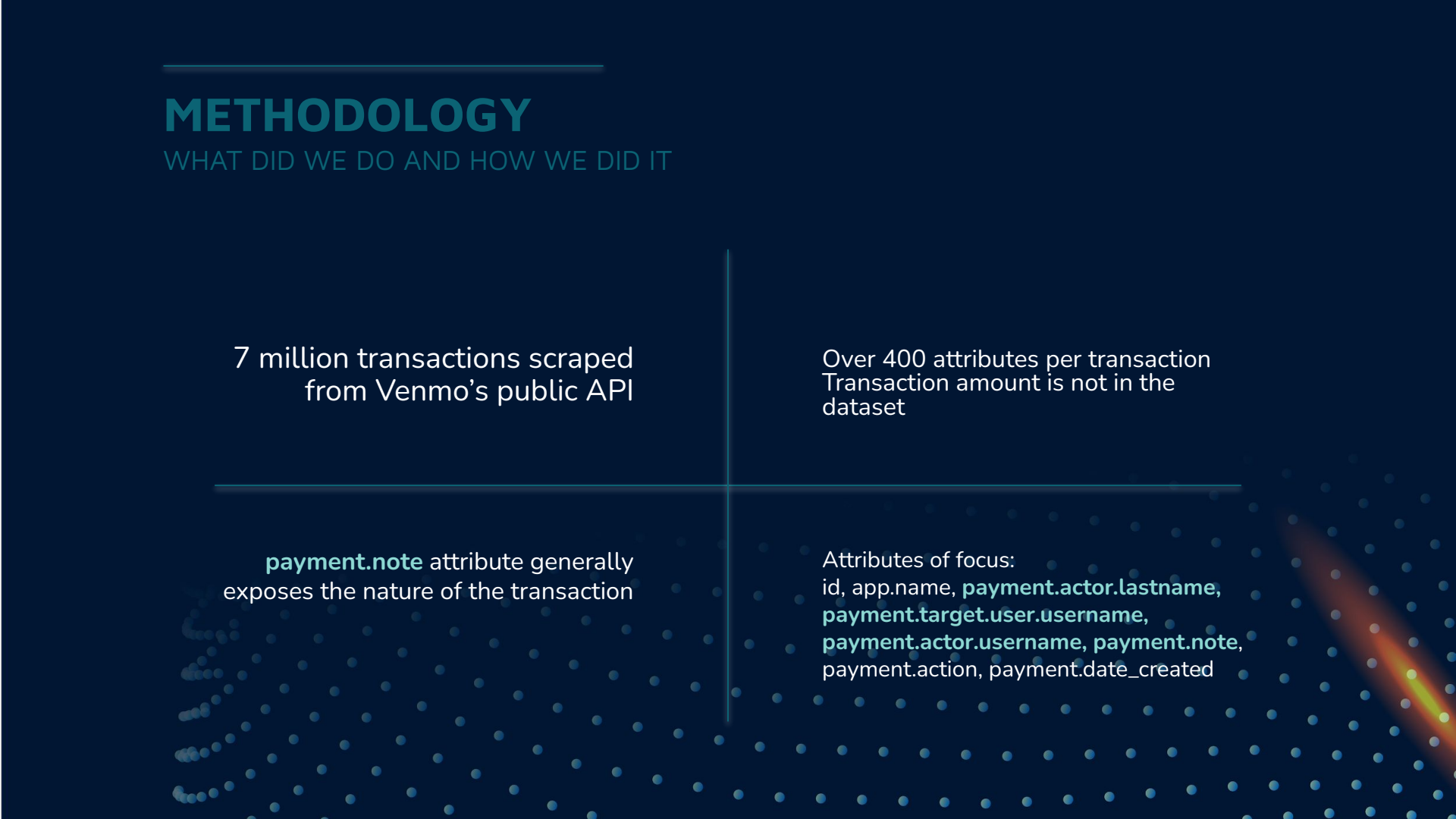
WHAT DID WE DO AND HOW WE DID IT

7 million transactions scraped
from Venmo's public API

Over 400 attributes per transaction
Transaction amount is not in the
dataset

payment.note attribute generally
exposes the nature of the transaction

Attributes of focus:
id, app.name, **payment.actor.lastname**,
payment.target.user.username,
payment.actor.username, **payment.note**,
payment.action, payment.date_created





WEED



ALCOHOL



COCAINE

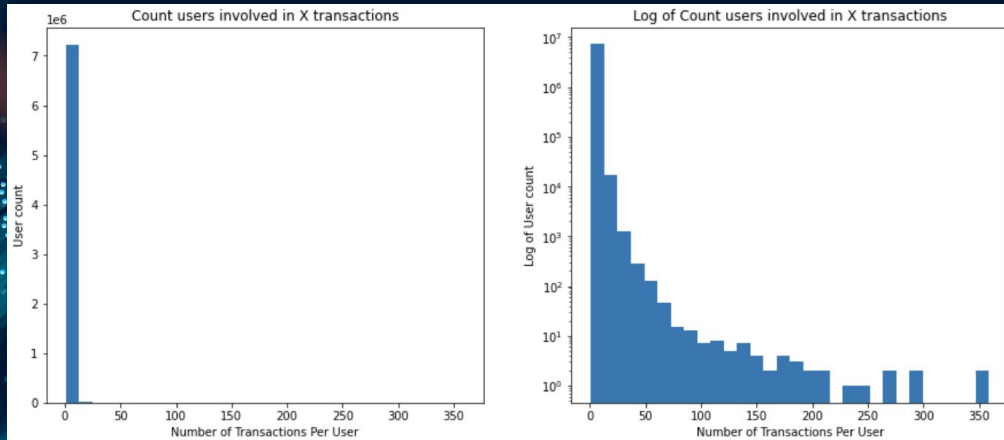
IMPLEMENTING PRIVACY

How much privacy is enough?

What mitigations?

Concerns

- Attribute Disclosure
- Membership Disclosure
- Identity Disclosure



DIFFERENTIAL PRIVACY

- Global Application
- Focus on counting queries

Laplace Mechanism

	Bin Label	Bin Count
0	Venmo for iPhone	6155051
1	Venmo for Android	884367
2	splitwise	25619
3	Venmo.com	9269
4	venmo payouts	484
5	tab	201
6	Venmo Developer	61
7	Workflow	6
8	drupe	6
9	BottleRocketUtility	6
10	Pay.mo	3
11	Kasisto KAI	3
12	Alexa for PayPal	2
13	Developer Settings	2
14	Georgetown University Alumni & Student FCU	1

No Differential Privacy

- Sensitivity
 $\Delta f = 1$
- Budget
 $\epsilon = 0.2$

	Bin Label	Bin Count
0	Venmo for iPhone	6155041.00063
1	Venmo for Android	884361.96104
2	splitwise	25617.35947
3	Venmo.com	9266.68018
4	venmo payouts	483.34252
5	tab	197.09750
6	Venmo Developer	66.29931
7	Alexa for PayPal	22.93999
8	Kasisto KAI	6.95960
9	BottleRocketUtility	5.02315
10	drupe	4.39540
11	Workflow	3.81721
12	Developer Settings	1.93746
13	Georgetown University Alumni & Student FCU	1.61341
14	Pay.mo	1.41615

0.2-Differentially Private

RESULTS

HOW DID IT ALL WORK OUT?

Racial and Ethnic Group	2010 Census (% of total population)	Venmo Transactions (counts)	Venmo Transactions (% of identified last names)
White	78.695	4,652,214.266	75.721
Black	3.120	120,741.423	1.965
Asian Pacific Islander	3.773	670,103.077	10.907
American Indian and Alaskan Native	0.089	2,039.672	0.033
Two or more races	0.008	656.268	0.011
Hispanic	14.313	698,152.832	11.363

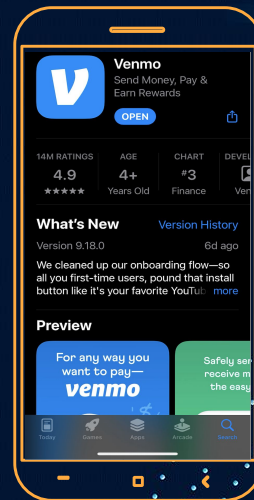
RECOMMENDATIONS

- Venmo
 - Do not include PII in API response
 - Consider deprecating the API in favor of an interactive, differentially private data set
 - Add UX friction for public transactions as an extra safety measure
- Users
 - CCPA
 - Request deletion from public data set
 - Change usernames moving forward

THANKS!

Questions?

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik.



**this will be replaced by
the next slide by Jot**

01

**WE OPTIMIZED THE
PRIVACY / UTILITY**

We were able to
maintain granularity
and integrity of the
data

02

**AVAILABLE TO
VENMO'S INTERNAL
ANALYSTS**

Here you can
describe the topic of
the section

03

**RACE AND
ETHNICITY OF USER
BASE**

Here you can
describe the topic of
the section

RESULTS (1)

HOW DID IT ALL WORK OUT?

SUBSTANCE	COUNT*	PERCENTAGE OF TRANSACTIONS
ALCOHOL	223,414.42	3.158
MARIJUANA	32,306.11	0.457
COCAINE	227.58	0.003

*Count of Transactions (not users) scraped from Venmo's public API

FUTURE WORK



USER LEVEL

Future work
could transform
the data from
transactional to
grouping by
users



USER LEVEL

Future work
could transform
the data from
transactional to
grouping by
users



TRACK