

RSAC[®]Conference2020

San Francisco | February 24 – 28 | Moscone Center

HUMAN
ELEMENT

SESSION ID: MLAI-W01

Disinformation Defense:

How AI inference threats might influence the outcome of 2020 election



KAREL BALOUN, KEN CHANG, MATT HOLMES

School of Information
UNIVERSITY OF CALIFORNIA – BERKELEY

#RSAC

Los Angeles Daily Times
SATURDAY MORNING, JULY 29, 1905.
California del Sur.
CITY AND COUNTRY.
IN ALL NEWS STATIONS
TRADE AND VENTURE 5 CENTS

EDITORIAL SECTION
PART II—LOCAL NEWS: 12 PAGES
CXIVTH YEAR.

TITANIC PROJECT TO GIVE CITY A RIVER.

Thirty Thousand Inches of Water to be Brought to Los Angeles.

Options Secured on Forty Miles of River Frontage in Inyo County—Magnificent Stream to be Conveyed Down to the Southland in Conduit Two Hundred and Forty Miles Long—Stupendous Deal Closed.

INDEPENDENCE (Cal.) July 28.—[Exclusive Dispatch.] Agents representing Los Angeles city have secured options on about forty miles of frontage on the Owens River north of Owens Lake. Fred Eaton, ex-Mayor of Los Angeles, and the superintendent of the Los Angeles water works were in the valley in an automobile the early part of this week. Two days ago they closed the last outstanding options. The price paid for many of the ranches is three or four times what the owners ever expected to sell them for. Everybody in the valley has money, and everyone is happy.

Three months ago, Eaton bought the holdings of the Rickey Cattle Company, comprising about 50,000 acres of water-bearing land. It was then thought that Eaton was going into the stock-raising business here, but it has since been learned that he was securing options for Los Angeles city. Eaton has made every option solid and secured all the land the city wanted. The deal is riveted.

THE cable that has held the San Owens River Valley between Lone Fernando Valley and for ten Pine and the northern edge of Owens centuries to the arid demon is Lake. In this territory are thousands of acres of government land. The water served during a number of trips to his son's ranch near Independence, the peculiar formation of the land along the route of the wagon road from Mojave to the Owens River Valley.

The engineers now all agree on what he first suggested, that the waters of the Owens River centuries ago flowed down through the arid valley from what is now Owens Lake, passing near the present site of Mojave and finally emptying into the Los Angeles River in the San Fernando Valley.

A series of mighty upheavals dislocated the ribs of a number of the lesser Sierras, throwing mountains across the path of the stream and for ten centuries at least the river has emptied into Owens Lake.

This lake is a great stretch of water ten miles wide and thirty miles long. Its waters are so permeated with soda that they contain no living thing.

The engineers contemplate cutting through the mountains that block the path of the river and bringing a canal from Charley's Butte, a foothill midway between Independence and Lone Pine, by way of Mojave to Los Angeles.

All the plans have been approved by the government engineers.

By expending about \$150,000 in cash for options and by guaranteeing the payment of over \$200,000 more, the Water Commissioners have pledged the city to build this conduit.

Ex-Mayor Eaton has acted as the city's agent in all the negotiations. The farmer folk in the Owens River Valley think that he has gone daffy on stock-raising. To them he is a millionaire with a fad.

INDEPENDENCE SACRIFICED.
It is the village of Independence that will probably be hardest hit. The town is kept alive partly by the trade of the ranchmen in the valley, and partly by the travel to and from the gold fields of Western Nevada. The trail to Goldfield and Inclineburg passes through Independence.

Already the engineers are recalling a canal from Lone Pine to Mojave. A great deal of the territory lies in a gold-bearing district. The sands in the bed of the Owens River are in many places rich with placer gold. Some of the tunnels will be run through buttes in which are promising ledges of gold and copper ore.

LAST SPIKE IN. DEAL CLOSED.
SUPT. MULHOLLAND BRINGS THE GLAD TIDINGS.

Says Options are Fixed and Los Angeles Becomes Owner of Thirty Thousand Inches of Purest Snow Water—Would Give All the Credit to Others.

Scorched and browned by the almost intolerable desert wind and sun Superintendent Mulholland returned yesterday afternoon from a daring nine days' automobile trip into the heart of the Owens River country, bearing the glad tidings that "The last spike has been driven, the options are all secured, the deal by which Los Angeles city becomes the owner of thirty thousand inches of the purest snow water has been nailed."

In the excited gratification born of a knowledge that the vexed water question has at last been solved, Mulholland laughed like a schoolboy.

"Fred Eaton did it. He has been working on it for thirteen years. He is the greatest natural engineer that the West has ever known. He has made it possible for us to accomplish the greatest scheme of water development ever dreamed of."

SUPT. MULHOLLAND AT WATERWORKS.



- 1928 St Francis Dam collapse killed 450
- 1926 St. Francis Dam completed
- 1907 Los Angeles voter approved construction bond
- 1905 Los Angeles started aqueduct project from the Owens Valley with initial \$1.5 million bond

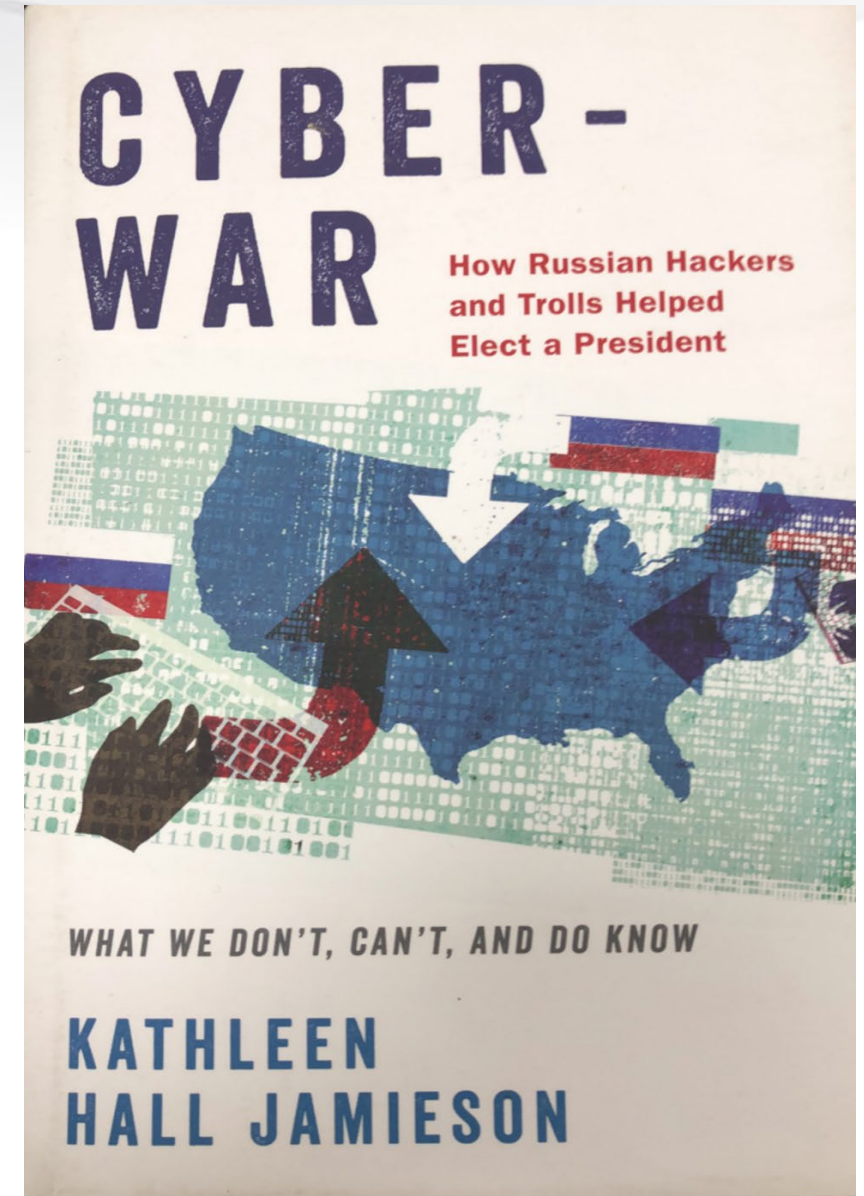


Freedom is on the line

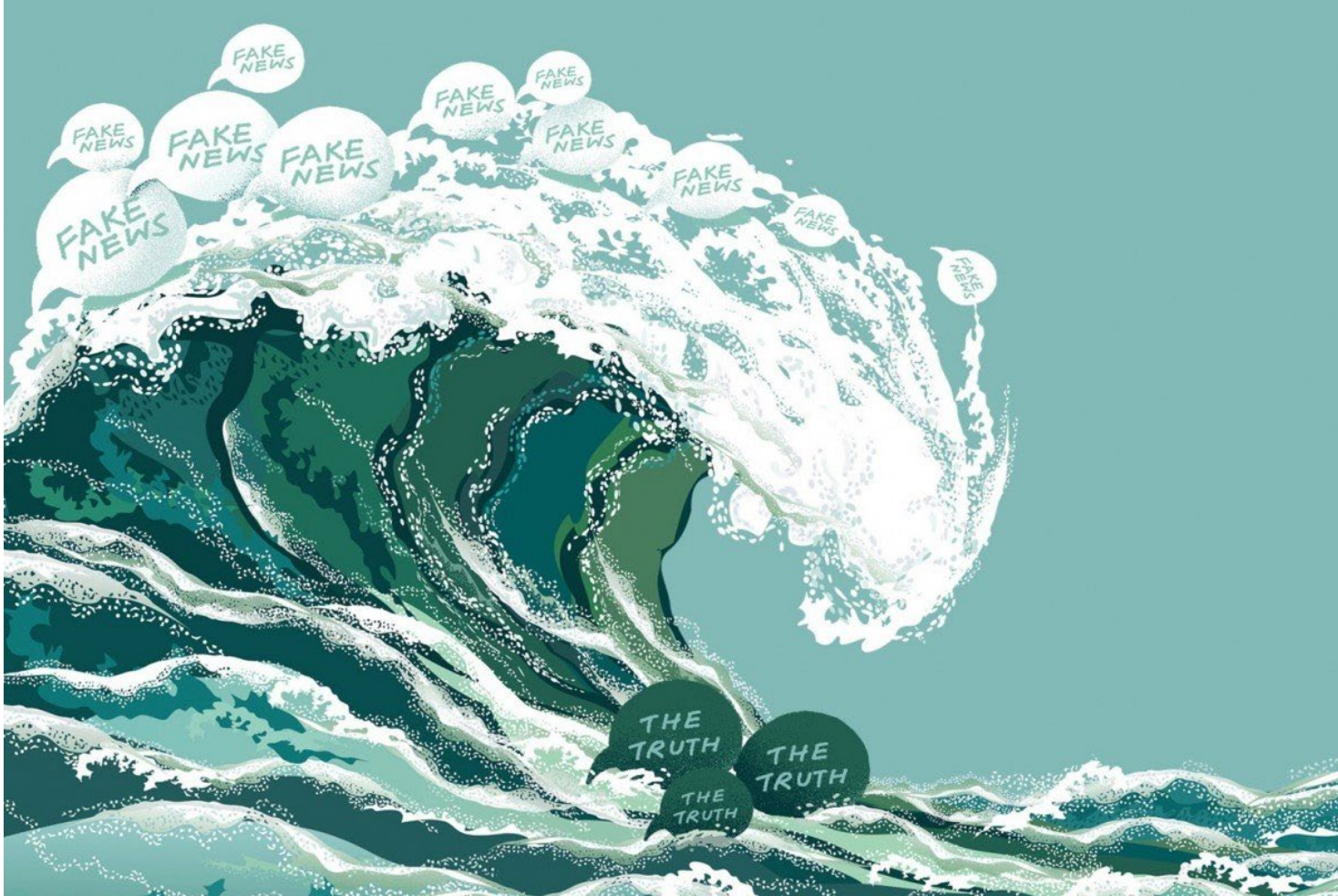
- Freedom is America's seminal and central value. AI can take it away.
- "A threat to free will is a threat to freedom, the imposition of a dangerous worldview without public awareness. **When free will itself is threatened, that is the ultimate threat to freedom.**"
 - George Lakoff on page 62 of "Whose Freedom?" (2006)

2016 U.S. Presidential Election

- National Security is never a partisan issue. Facts only.
- Jamieson is a trusted political historian.
- GRU hacked DNC
- Wikileaks sanitized the theft
- Media reported. Trolls inflamed.
- Key voters stayed home.



Disinformation Attack



- 2016 Ukraine Election
- 2016, 2019 UK Election on Brexit
- 2019 Hong Kong - Anti Extradition Law Protests
- 2020 Taiwan Presidential Election

RSA®Conference2020

AI Inference on Voter Preference

A matter of national security

Have you registered to Vote?

Voter Registration Application
Before completing this form, review the General, Application, and State specific instructions.

Are you a citizen of the United States of America?
Will you be 18 years old on or before election day?
If you checked "No" in response to either of these questions, do not complete form.
(Please see state-specific instructions for rules regarding eligibility to register prior to age 18.)

☐ Yes ☐ No ☐ Yes ☐ No

This space for [unclear] ly.

☐ Jr ☐ II
☐ Sr ☐ III
Code ☐ IV

1 ☐ Mr. ☐ Miss
☐ Mrs. ☐ Ms.

2 Home Address

3 Address Where You Get Your Mail If Different From Above

4 Date of Birth
Month Day Year

5 Telephone Number (optional)

6 ID Number - (See item 6 in the instructions for your state)






7 Choice of Party
(See item 7 in the instructions for your state)

8 Race or Ethnic Group
(See item 8 in the instructions for your state)

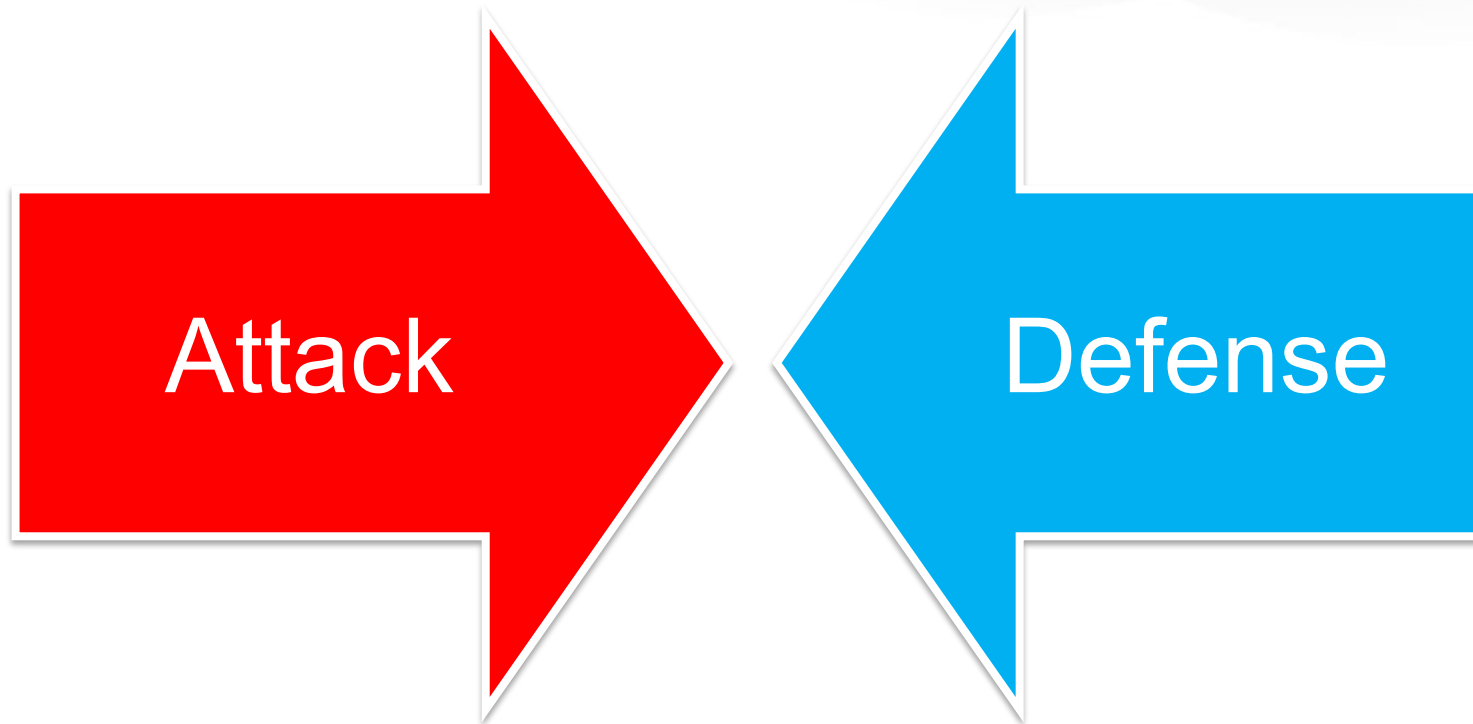
Please sign full name for put n

Day

AI/ML Threat Modeling on Personal Data

Assets	Confidentiality	Integrity	Availability
	data breach, linkage Inference	data poisoning	
	data breach Inference	margin of error	
	data breach, linkage Inference	disinformation, misinformation, malformation	
	data breach, linkage Inference	data poisoning adversarial perturbation	
	Inference {identity, attribute, membership}	algorithmic bias	deny of service

Hypothesis: Voter Preferences Disclosure

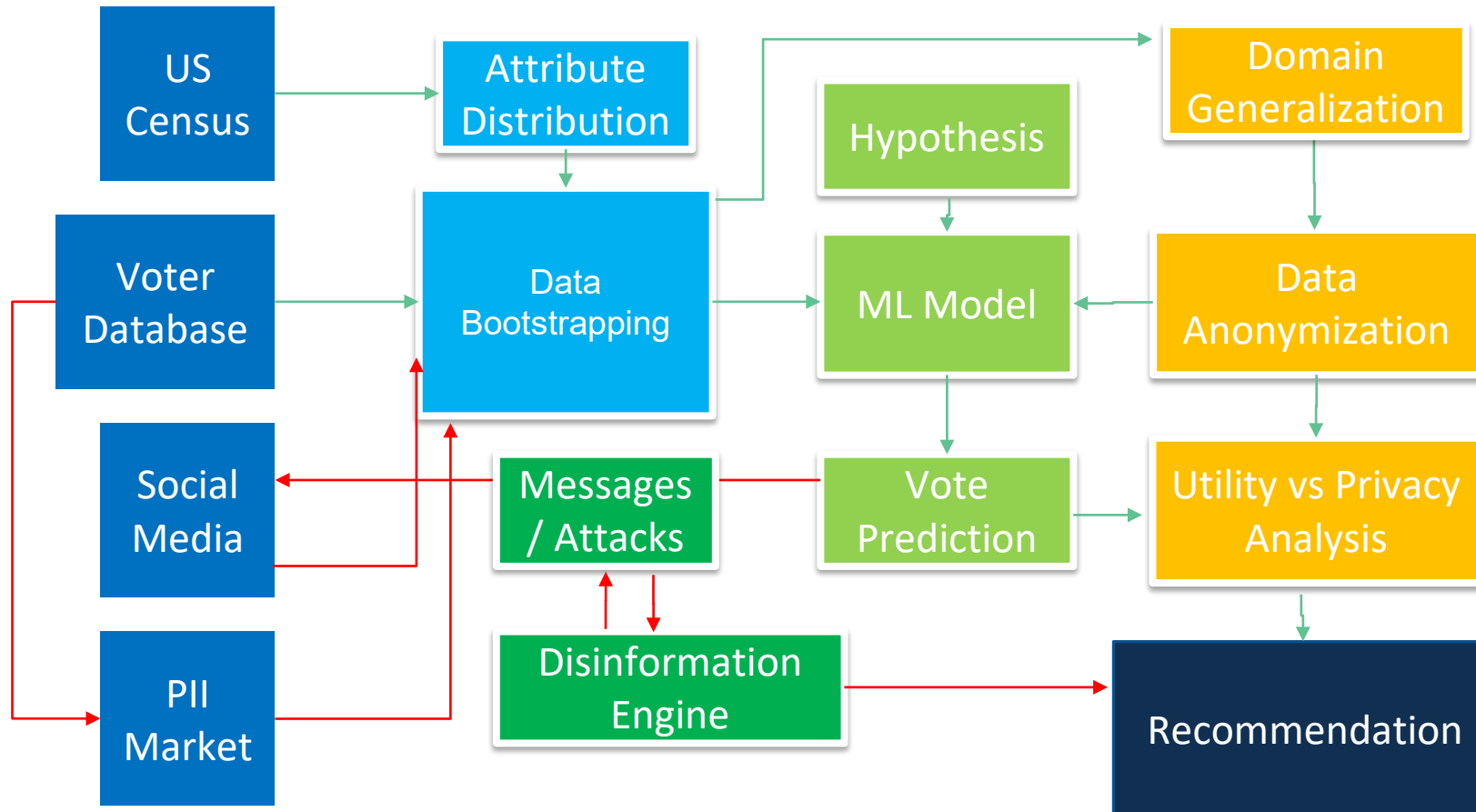


Use of Machine Learning algorithm with vast amount of personal data to infer secret personal attribute on vote preference.

Advocate data privacy laws and regulations with a holistic approach to anonymizing the personal data in the public market.

Design of Experiment

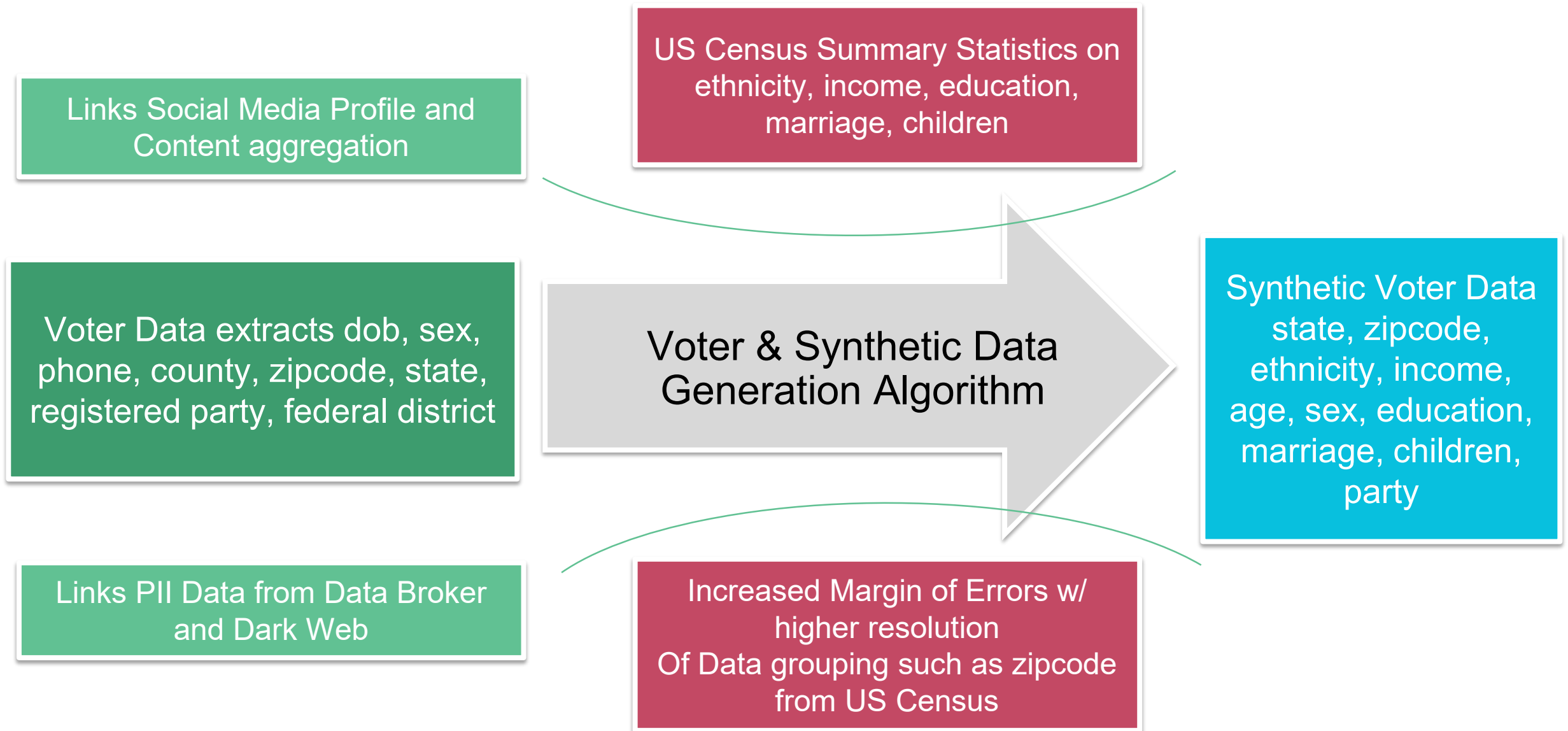
- Data Source
- Data Bootstrap
- Machine Learning
- Privacy Engineering
- Vote Influence



RSA®Conference2020

Defense: Privacy Engineering

Voter Data Bootstrapping



A Crash Course on k-Anonymity*

Lemma.

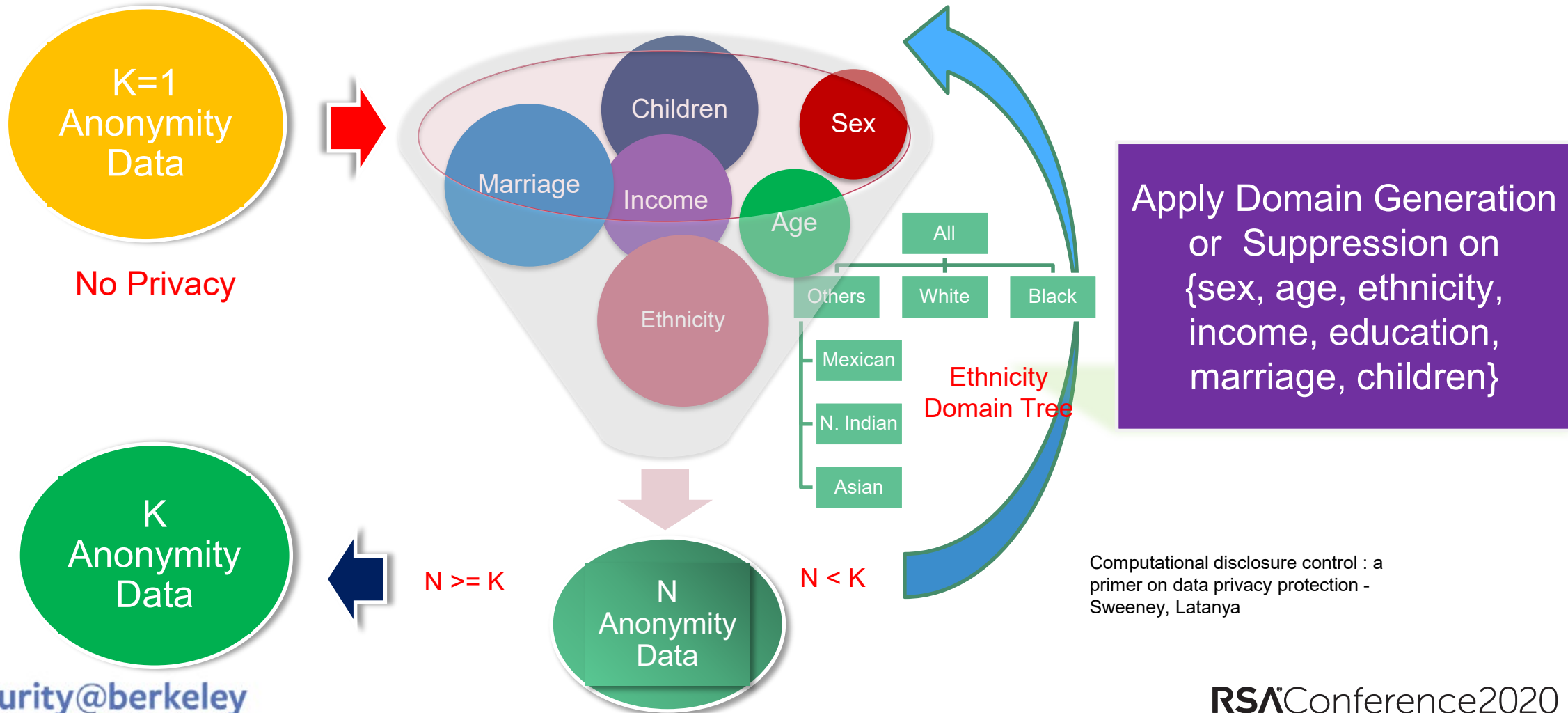
Let $RT(A_1, \dots, A_n)$ be a table, $QI_{RT} = (A_i, \dots, A_j)$ be the quasi-identifier associated with RT , $A_i, \dots, A_j \subseteq A_1, \dots, A_n$, and RT satisfy k -anonymity. Then, each sequence of values in $RT[A_x]$ appears with at least k occurrences in $RT[QI_{RT}]$ for $x=i, \dots, j$.

- Quasi-identifiers refer to attributes that are not considered explicit identifiers but could be linked with external information to re-identify personal records
- Given $QI = \{\text{Gender, Age, Marital Status, Race}\}$, $k=1$

Gender	Age	Marital Status	Country	Race	Alcohol/Day	Hard Drugs
M	32	Married	US	Non-Hispanic White	1	No
M	32	Married	US	Non-Hispanic White	2	Yes
M	32	Divorced	US	Non-Hispanic White	3	Yes
M	32	Never Married	US	Non-Hispanic White	1	No

* L. Sweeney,
2002

Voter Data Anonymization – Datafly Algorithm with desired K-Anonymity



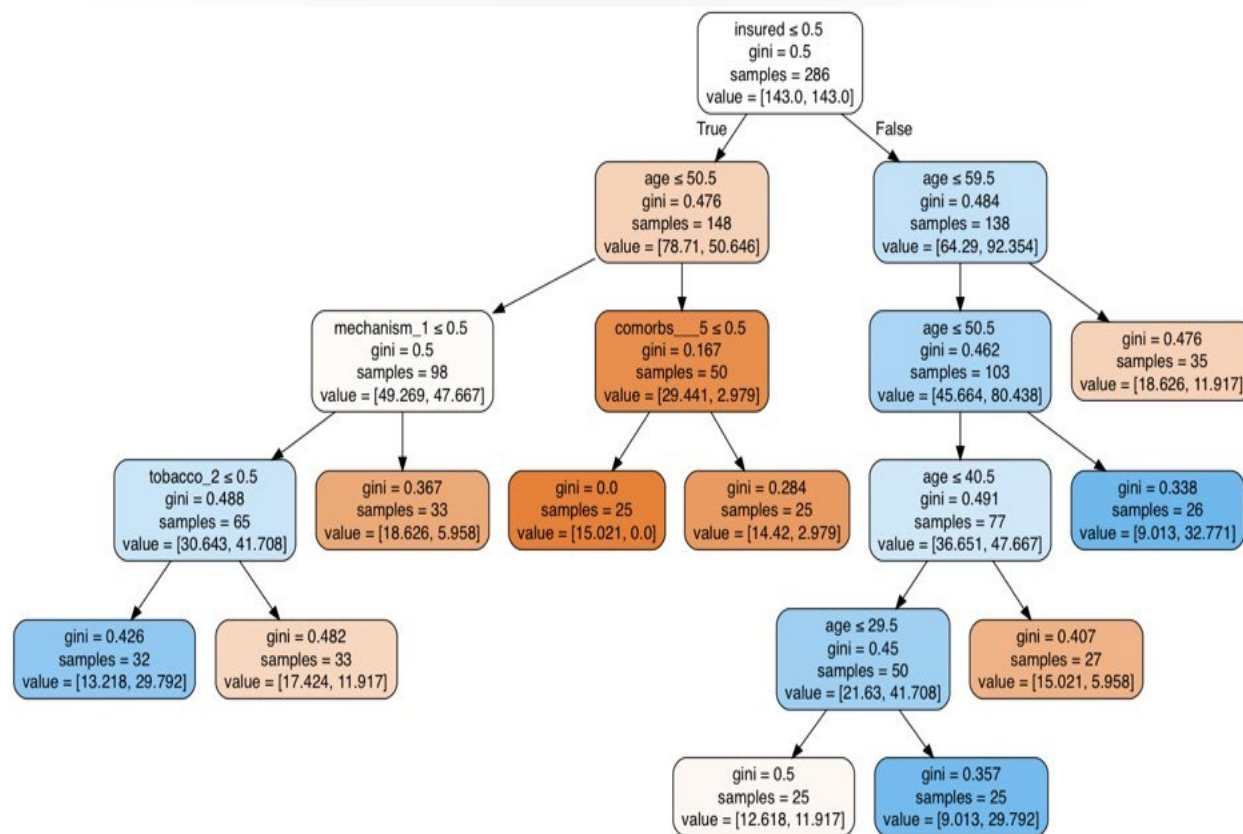
RSA®Conference2020

Attack: Machine Learning

Machine Learning Model - Decision Tree

Supervised Model - Party as Labels
(as a proxy for vote preference)

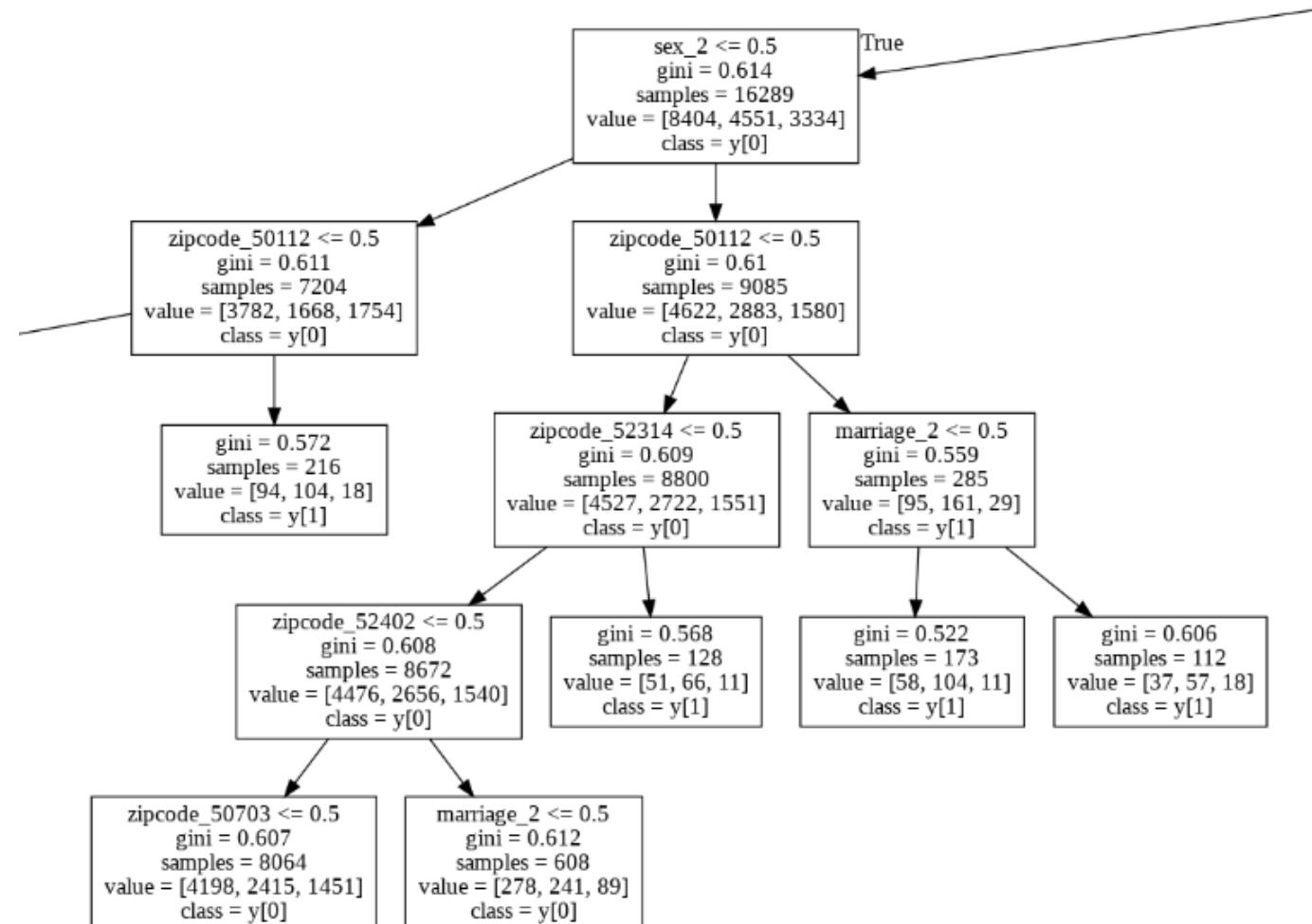
- Train & Predict raw data
 - Raw Data: 40% accuracy
- Train & Predict anonymized data
 - Anonymized Data: 47% accuracy
- Baseline: 3 labels ~ 33%
- Summary of Results



Machine Learning Model - Decision Tree Example

Record = [Age , Sex, Marital Status, Zip]

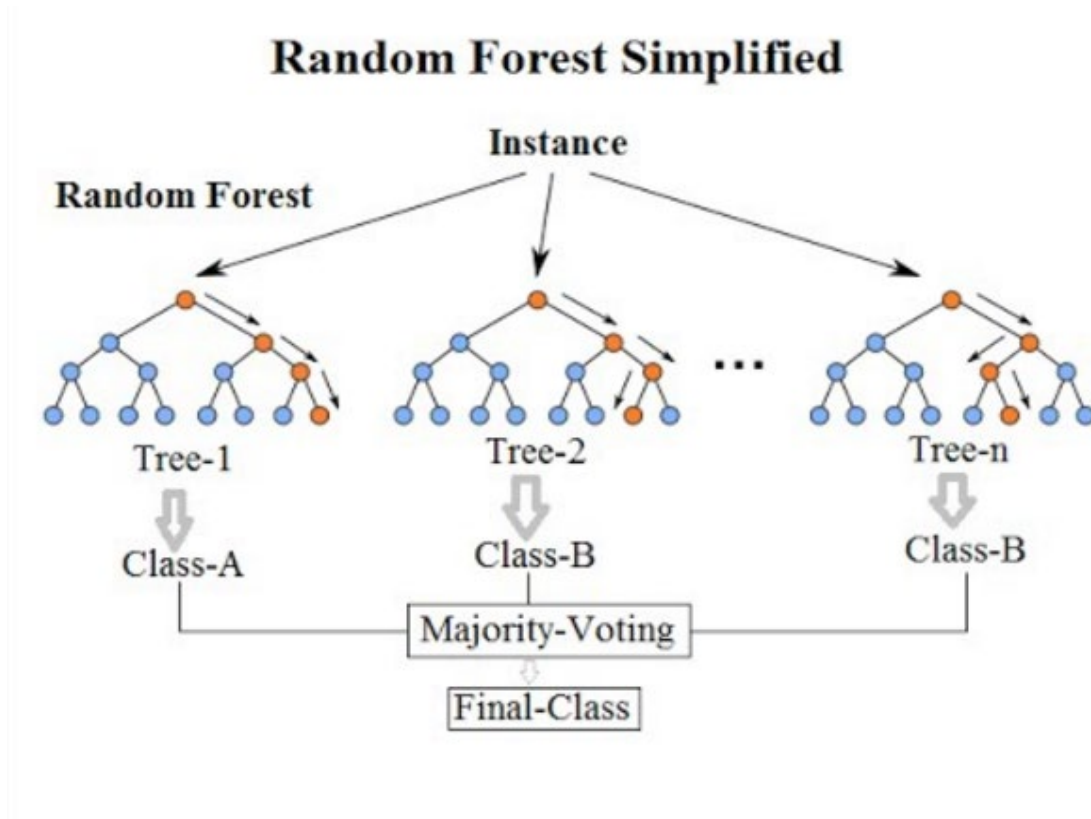
Bob = [55, M, Married, 50112]



Machine Learning Model - Privacy Aware Decision Forest

Decision Forest

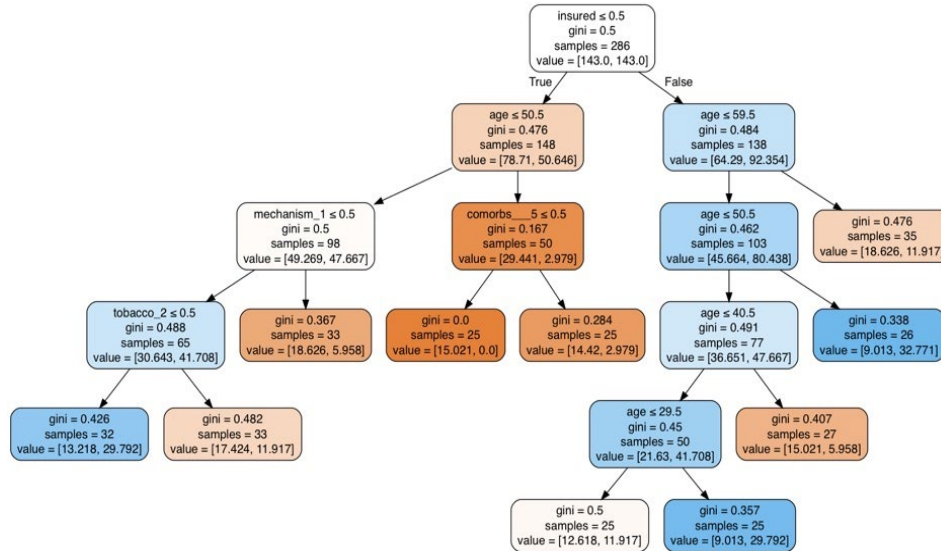
Privacy Aware



- Each decision tree will be tuned with Privacy Engineering in mind
- Increased accuracy when combined with real aggregated data linked from personal data sources
- More robust than a single tree; will help with capturing more structure in the data

“Privacy Aware”

Example:
Max Leaf Count



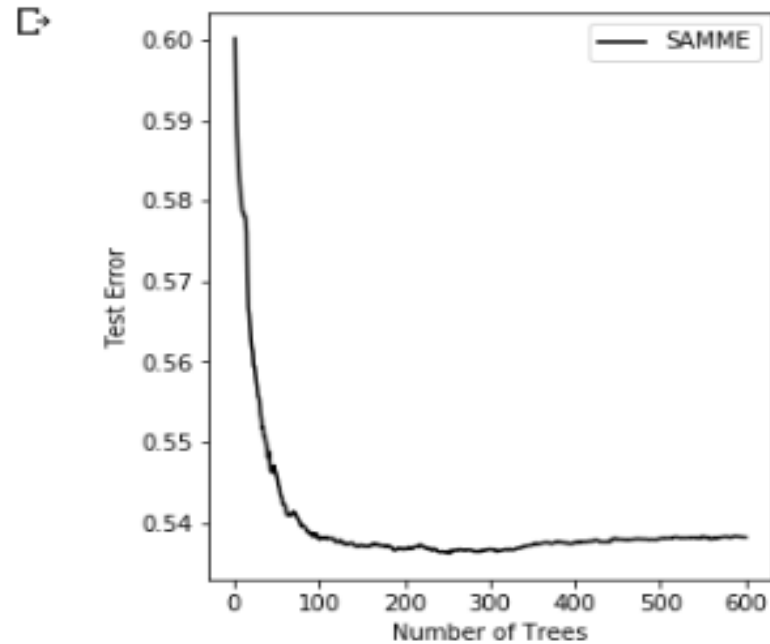
Leaf Count \leq Number of Equivalence Classes

Example: Carefully synthesized data and purchase personal data

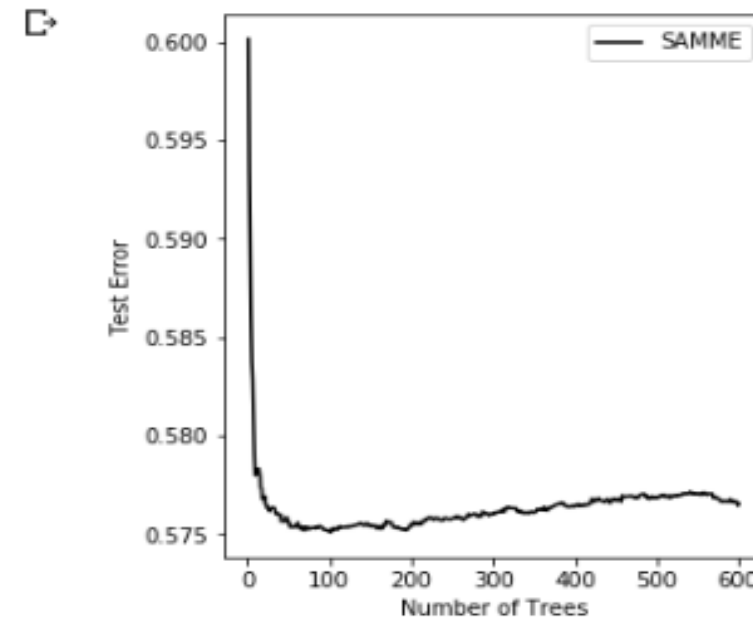
1. Use Census to collect probabilities of attributes for each EC
2. Find datasets that tie attributes to party affiliation (even in population level)

Decision Forest Results

K=1 Data: 46% accuracy



K=17 Data: 43% accuracy



RSA®Conference2020

Disinformation Engine

Disinformation Lifecycle



SOCIAL ENGINEERING PUBLIC RELATIONS
INFORMATION OPERATIONS ADS MISINFORMATION
ACTIVE MEASURES DISTORTION LIES
HALF TRUTHS
MEDDLING DECEPTION BOTNETS **DEEP FAKES**
MISREPRESENTATION
TROLL FARMS **PROPAGANDA** FAKE NEWS
MANIPULATION INTERFERENCE INFLUENCE ACTIVITIES
STRATEGIC COMMUNICATIONS REFLEXIVE CONTROL BRAINWASHING
REALITY JAMMING HACKING **FALSE** PSYCHOLOGICAL OPERATIONS
ARTIFICIAL INTELLIGENCE PUBLIC DIPLOMACY **MISO**
INFORMATION WARFARE DISINFORMATION

<https://images.app.goo.gl/xPswVgEs4q2EzFZL6>

Messaging & AI Threats

- Messaging Bots & Email Spam
 - Robot calls, including for push surveys
 - Media blogs, videos and posts
 - Advertising by PACs, Issue, and Campaigns
-
- AI impersonation, lie creation.
 - Overloading the media and messaging, especially locally
 - Not big data challenge, rather persuasion management + intensity
 - “Her”. Speed and interactivity: of natural language responses
 - Tuning of persuasive messaging (Cambridge Analytica at max scale)

RSA®Conference2020

Final Thoughts

What you can do!

- ❖ Advocate for each of the political and organizational recommendations on the next slide.
- ❖ Collect data, Build tools, Create video and text stories which document that this reality is happening.
- ❖ Stop disinformation around you!

Identify it, flag and report it, educate others proactively, and block the viral spread of this disease.

Recommendations

- ❖ Further study of **Privacy by Design** principles on collection, processing, and disclosure of personal voter and public census data.
- ❖ Regulate the **data broker** industry with consumer privacy law, so everyone can view and opt-out any sensitive information.
- ❖ Every Secretary of State should mandate **disclosure requirements** for access and use of **voter record files**.
- ❖ FEC must **ban use voter personal profiles** in voter messaging, and ban use completely by campaigns and PACs.
- ❖ Large **audience minimum** size for advertising



RSA[®]Conference2020

Thank You