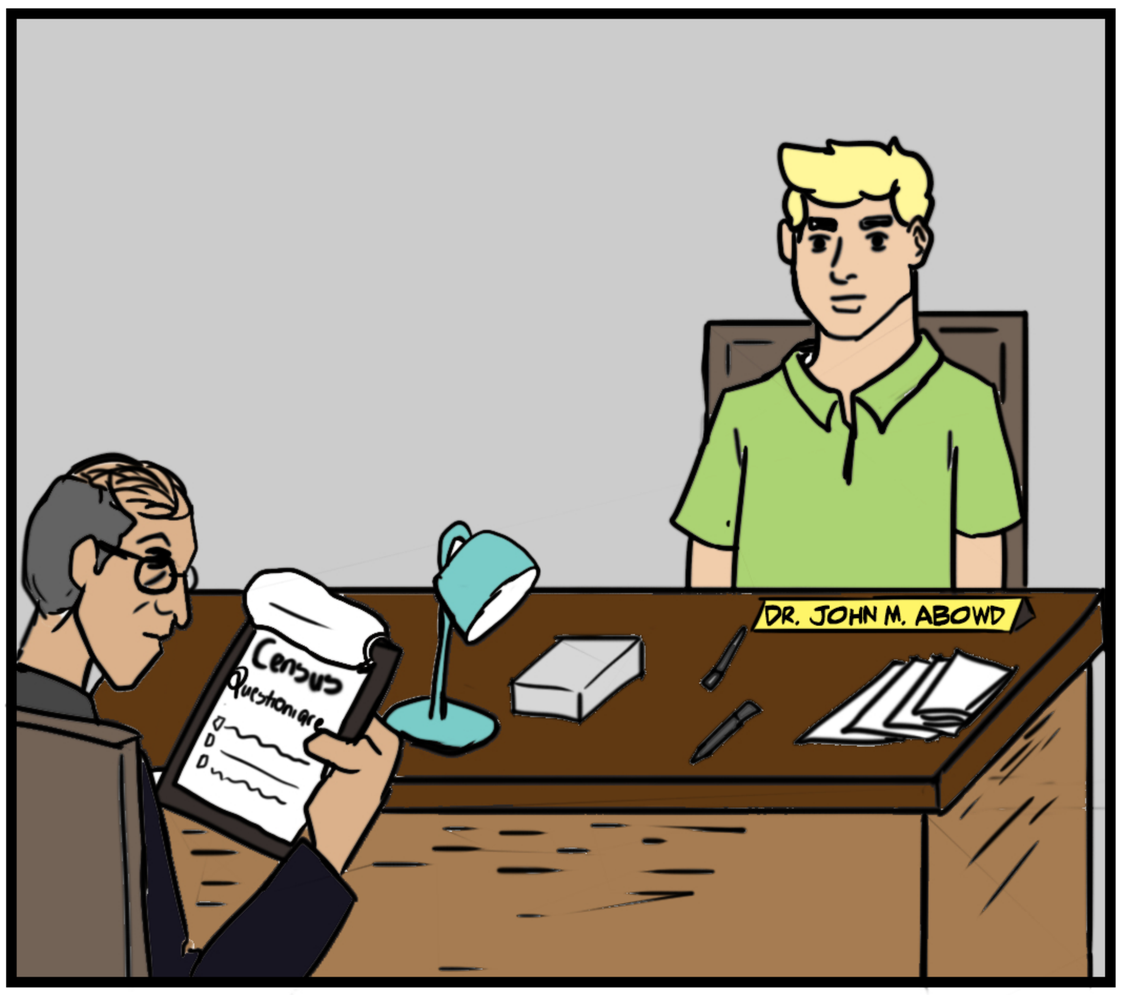
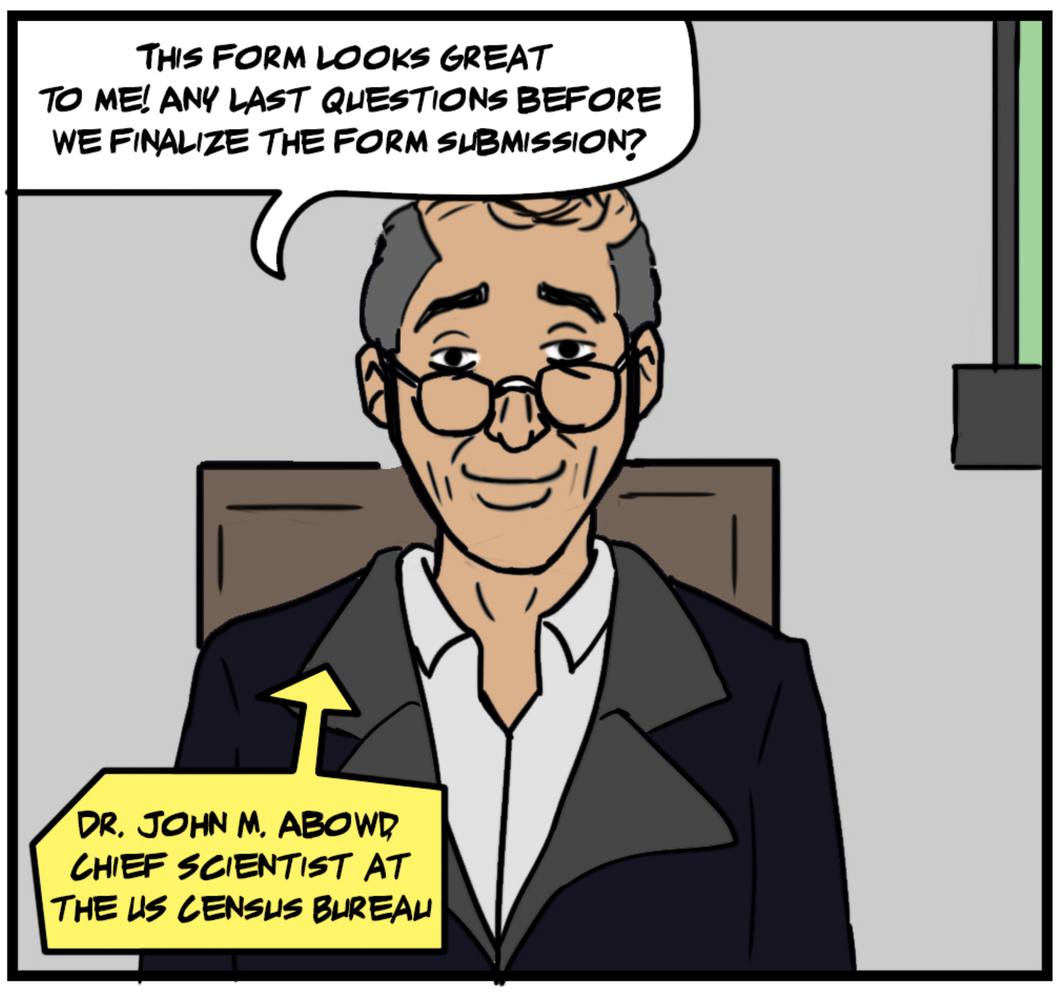
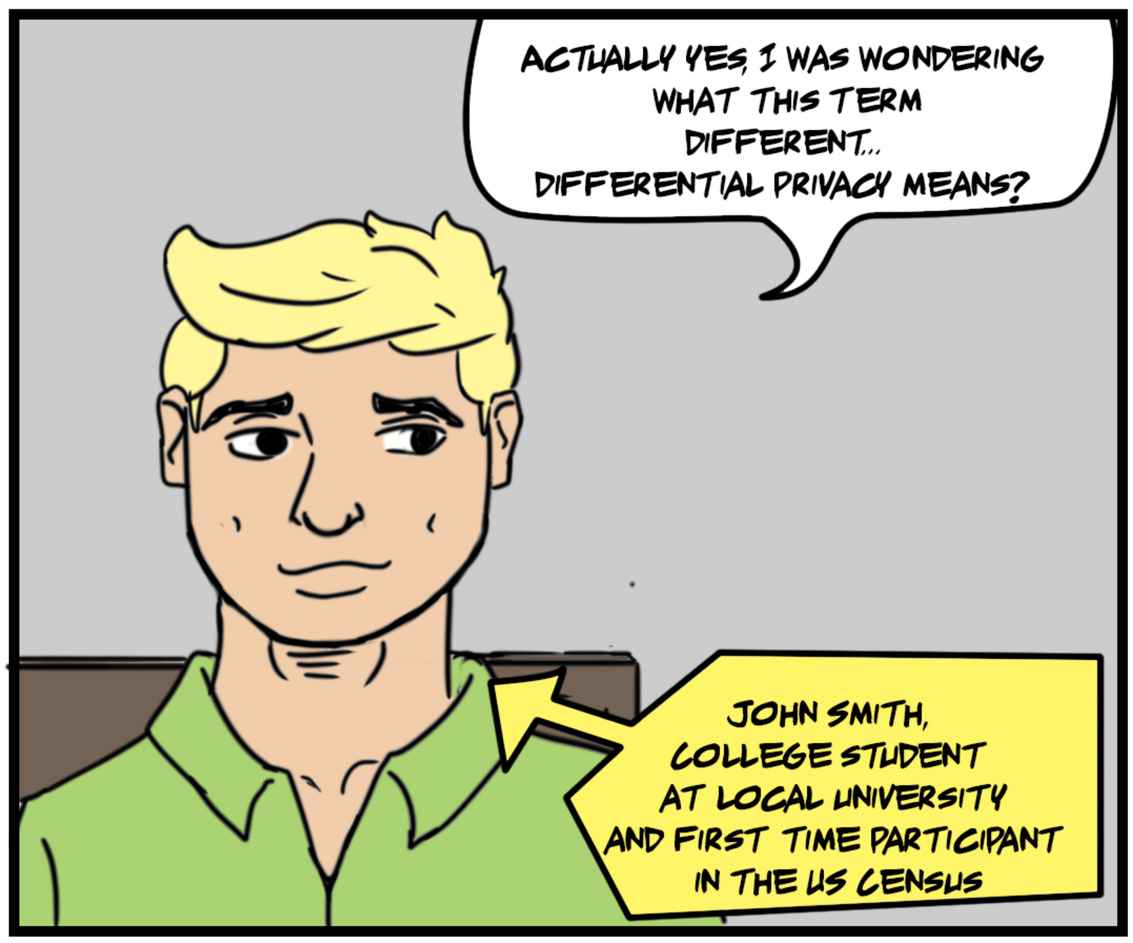
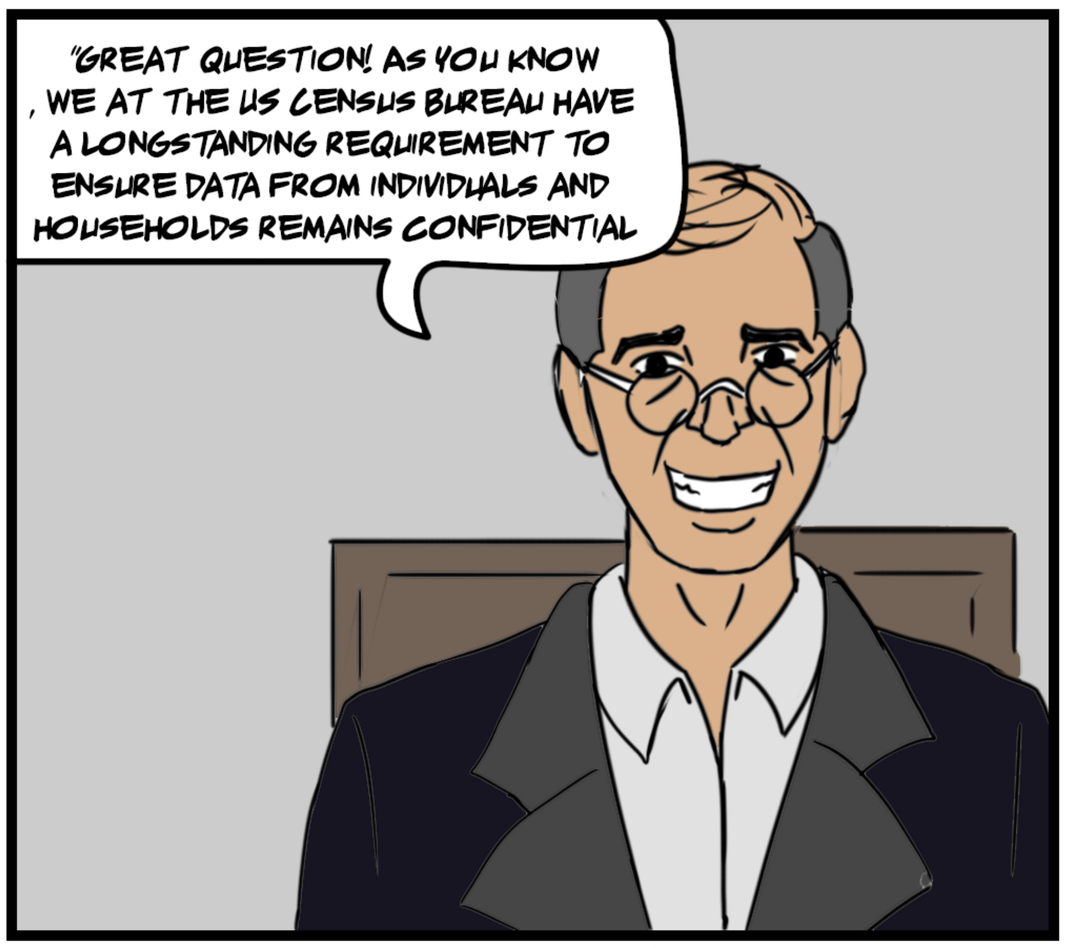
EXPLAINING DIFFERENTIAL PRIVACY









US CENSUS BUREAU

"DIFFRENTIAL PRIVACY", THE NEW GOLD STANDARD IN DATA PRIVACY PROTECTION.

DIFFERENTIAL PRIVACY IS OUR METHOD TO ENSURE THIS CONFIDENTIALITY AND PROTECT INDIVIDUAL PRIVACY.

GOOGLE CHROME DATA COLLECTION CONSENT

PROTECTED BY DIFFRENTIAL PRIVACY

APPLE DATA COLLECTION CONSENT

PROTECTED BY DIFFRENTIAL PRIVACY

YES NO

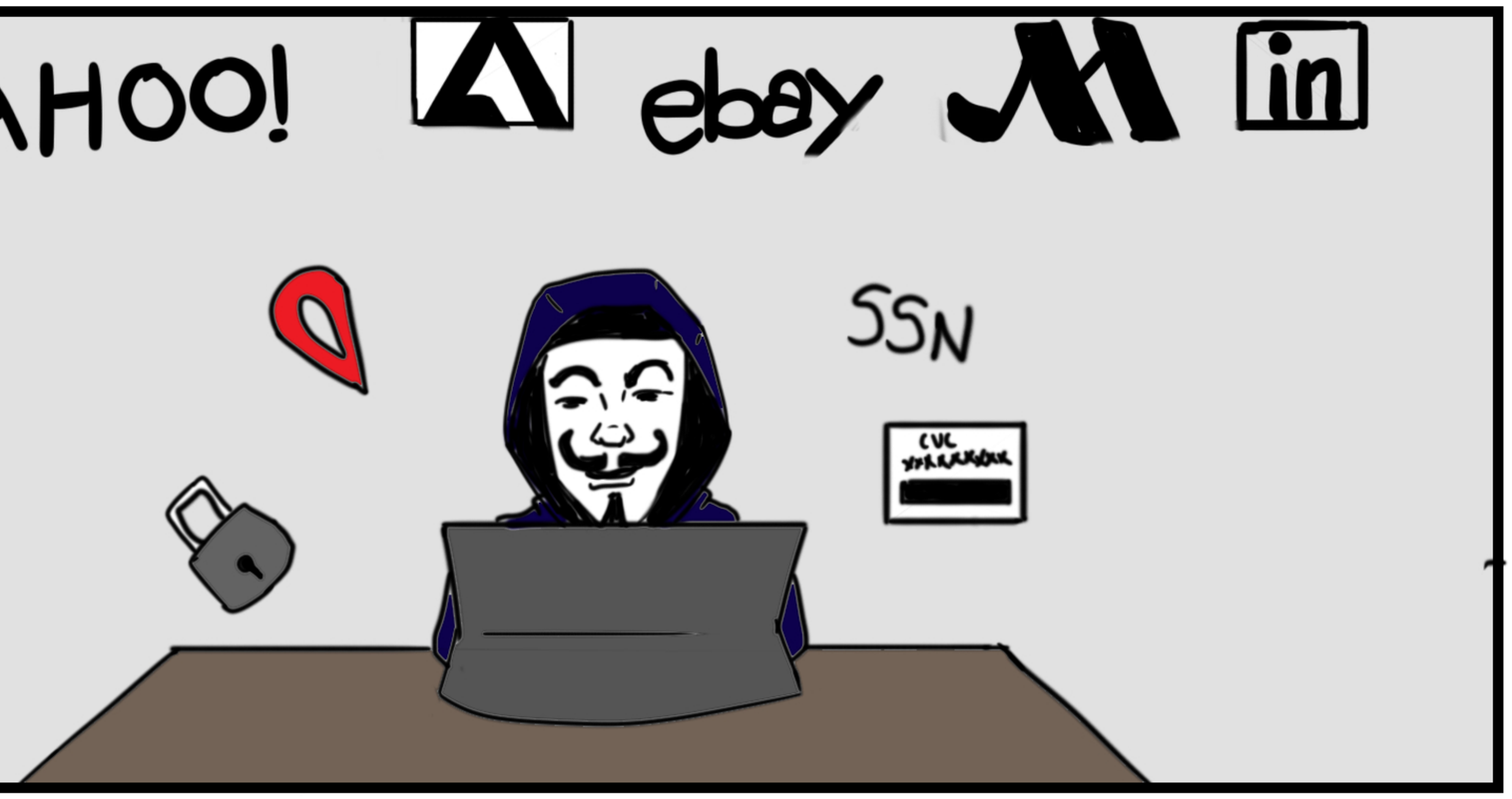
AWESOME! I HAD ACTUALLY SEEN THE PHRASE EARLIER AND WAS WONDERING WHAT IT WAS.

ABSOLUTELY, MANY COMPANIES ARE ADOPTING DIFFERENTIAL PRIVACY AS A METHOD OF PROTECTING YOUR PERSONAL DATA FROM EXTERNAL INDIVIDUALS OR ORGANIZATIONS WITH MALINTENT

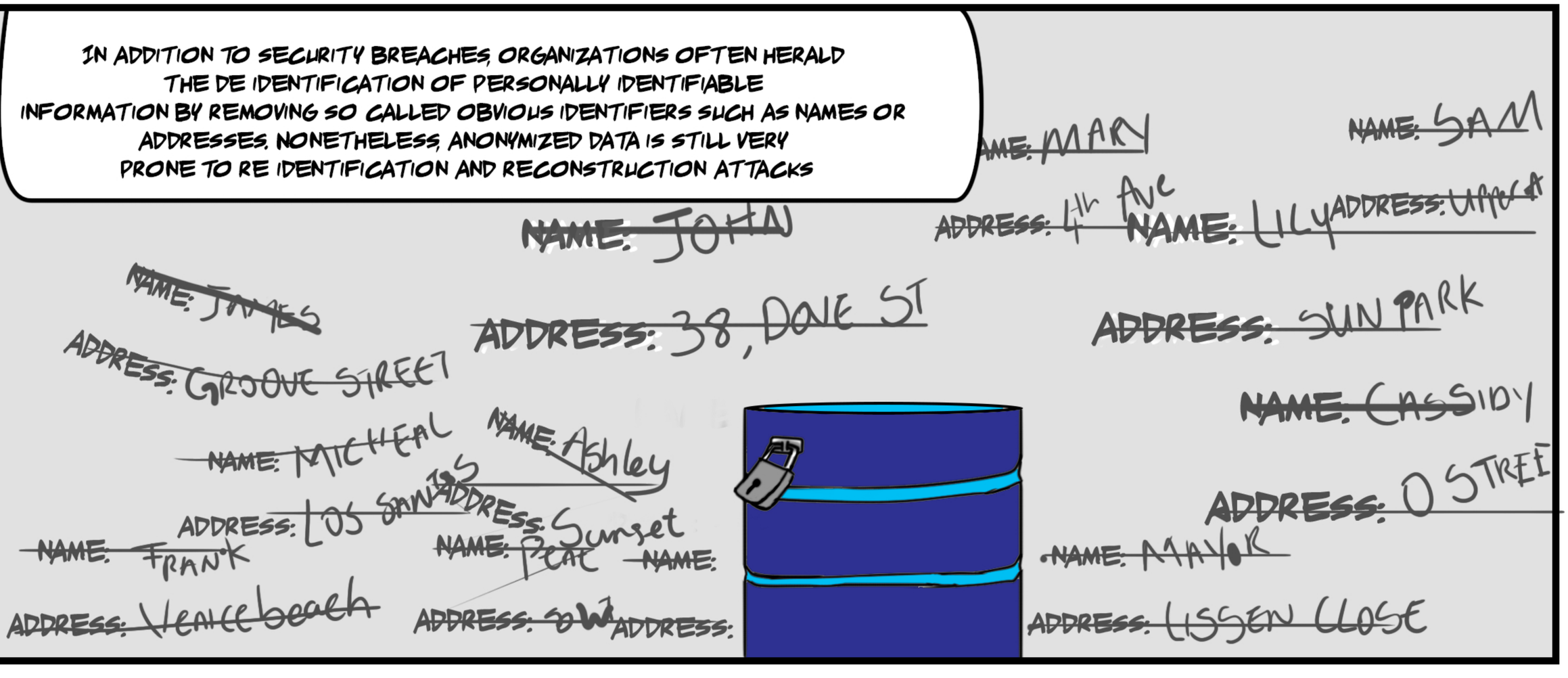
MY DATA? DON'T YOU ALL ALREADY ANONYMIZE OR ENCRYPT IT ANYWAYS SO IT'S SAFE? I THOUGHT PRIVACY WAS MORE ABOUT THE NSA TRACKING OUR CALLS OR MY SETTINGS ON FACEBOOK?"

UNFORTUNATELY, DESPITE ANONYMIZATION OR ENCRYPTION, YOUR DATA ISN'T AS SAFE AS YOU MIGHT THINK, AND THE PRIVACY RISKS ARE MUCH MORE MENACING THAN YOU'D EXPECT.

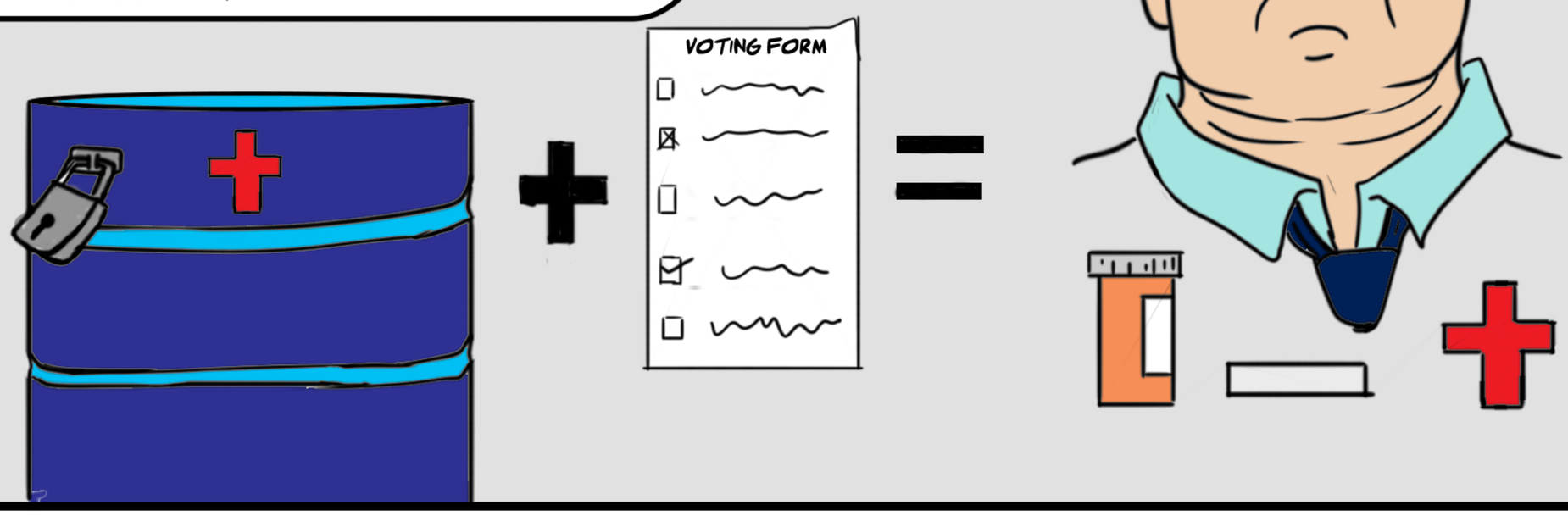
FOR INSTANCE, IN THE PAST TEN YEARS, BILLIONS OF PERSONAL USERS HAVE HAD THEIR DATA SUCH AS CREDIT CARD NUMBERS, PASSWORDS, LOCATIONS, SOCIAL SECURITY NU MBERS, AND MORE LEAKED BY 'TRUSTED' CORPORATIONS LIKE YAHOO, ADOBE, EBAY, MARRIOTT, LINKEDIN, AND OTHERS. EVEN WORSE, THESE ATTACKS ARE ONLY BECOMING MORE SOPHISTICATED AND FREQUENT IN NATURE.



IN ADDITION TO SECURITY BREACHES, ORGANIZATIONS OFTEN HERALD THE DE IDENTIFICATION OF PERSONALLY IDENTIFIABLE INFORMATION BY REMOVING SO CALLED OBVIOUS IDENTIFIERS SUCH AS NAMES OR ADDRESSES. NONETHELESS, ANONYMIZED DATA IS STILL VERY PRONE TO RE IDENTIFICATION AND RECONSTRUCTION ATTACKS



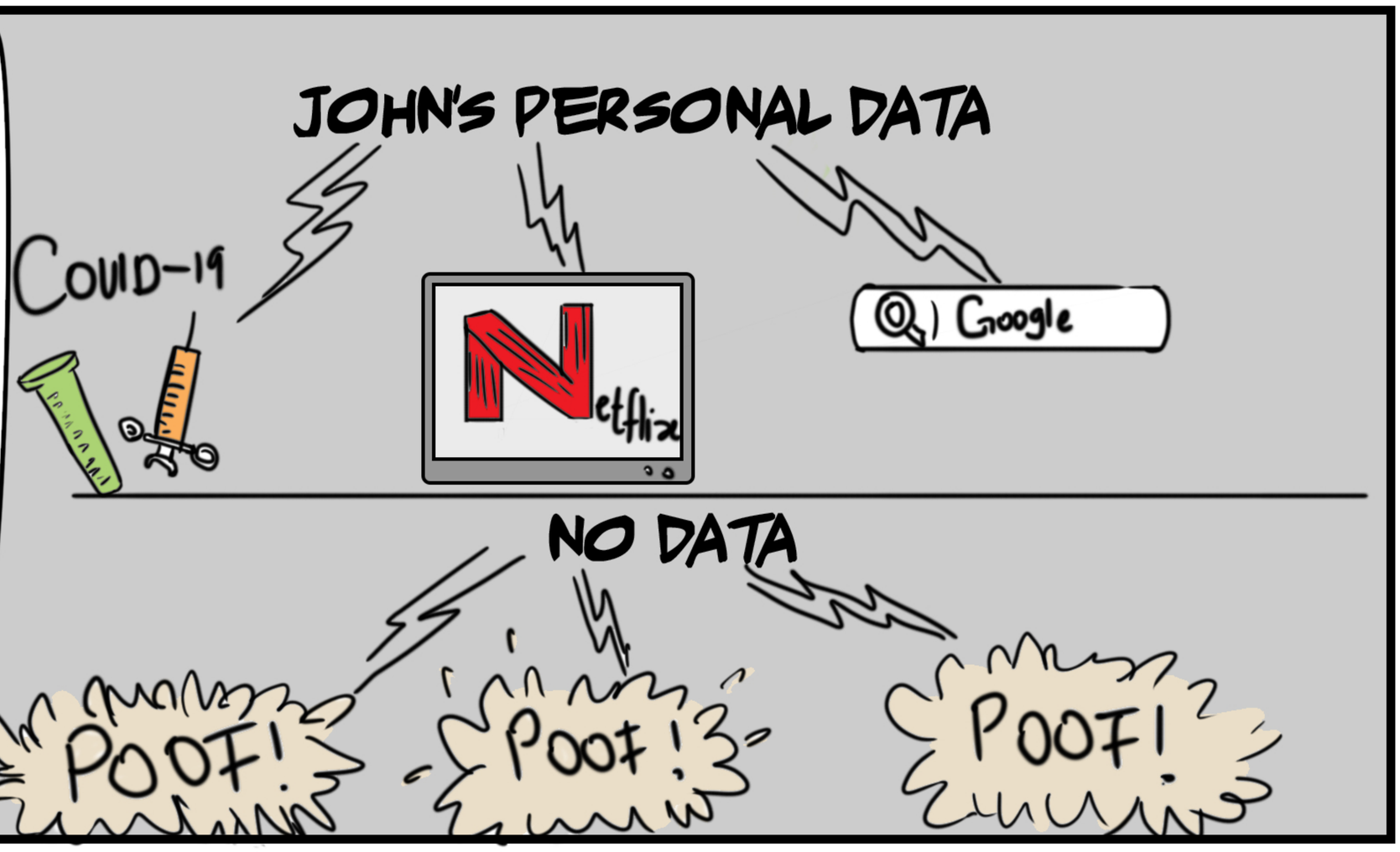
FOR EXAMPLE, IN HER 1997 GROUNDBREAKING RESEARCH, MIT PROF. LATANYA SWEENEY WAS ABLE TO LINK PUBLICLY AVAILABLE VOTER REGISTRATION DATA TO AN ANONYMIZED MEDICAL DATASET AND EASILY RE IDENTIFY PEOPLE FROM THE ANONYMIZED DATASET. MOST NOTABLY, SHE SENT THEN MASSACHUSETTS GOVERNOR BILL WELD HIS FORMERLY ANONYMOUS HEALTH RECORDS INCLUDING PRESCRIPTIONS AND DIAGNOSES TO HIS PERSONAL OFFICE



THAT'S HORRIBLE! I DON'T WANT MY PRIVATE HEALTH INFORMATION MADE PUBLIC! SHOULD I JUST NEVER CONSENT TO MY INFORMATION BEING SHARED WITH ANYONE?

WHILE IT'S DEFINITELY IN YOUR PERSONAL RIGHTS, I WOULDN'T RECOMMEND SO.

TECHNOLOGY TODAY RELIES ON DATA. WITHOUT YOUR PART, RESEARCHERS WOULD BE WORSE AT DEVELOPING VACCINES AND DRUGS, NETFLIX WOULDN'T HAVE HIGH QUALITY RECOMMENDATIONS, AND YOUR NEXT GOOGLE SEARCH WOULDN'T BE AS ACCURATE. IF EVERYONE DECIDED SUDDENLY TO NOT SHARE THEIR DATA WITH COMPANIES, WE'D LOSE ACCESS TO ALL THE BENEFITS TECHNOLOGY CURRENTLY GIVES US!"



SO MY TIKTOKS WOULDN'T BE AS POPULAR??

YES,, OR, HERE AT THE US CENSUS BUREAU LOCAL AND STATE GOVERNMENTS RELY ON THE DATA YOU PROVIDE TO DEVELOP IMPORTANT PROGRAMS AND BUDGETS ALONGSIDE ASSESSING ECONOMIC AND FINANCIAL CONDITIONS . WE WANT A WORLD WHERE YOU CAN HAVE BOTH THE BENEFITS OF DATA SHARING WITHOUT HAVING TO WORRY ABOUT YOUR PRIVACY!

Poof!!

AH, I SEE. SO THEN IF I SHARE, HOW DO I KNOW THAT MY DATA WILL BE PROTECTED FOR SURE? IS THIS WHERE DIFFERENTIAL PRIVACY COMES IN,, WAIT WHAT EXACTLY EVEN IS DIFFERENTIAL PRIVACY??

GREAT QUESTION! LET'S USE AN EXAMPLE TO HELP ILLUSTRATE

MEET ALICE! ALICE IS OUR HEAD DATABASE MANAGER HERE AT THE US CENSUS BUREAU AND SHE'S THE ONLY ONE WHO HAS ACCESS TO OUR PRIVATE INFORMATION

ALICE
HEAD DATABASE MANAGER

MANY LOCAL AND FEDERAL GOVERNMENTS WANT TO USE THIS DATA IN THEIR POLICYMAKING PROCESSES AND UNDERSTANDING OF THEIR AREA'S CONDITION . TO HELP THEM, ALICE CAN ANSWER THEIR INDIVIDUAL QUESTIONS, OTHERWISE KNOWN AS QUERIES, WITHOUT GIVING AWAY ALL OF THE PRIVATET INFORMATION IN THE DATABASE.

UNFORTUNATELY, THERE ARE ATTACKERS WHO WANT TO USE ALICE'S GENEROSITY IN ANSWERING QUESTIONS FOR ILL USE BY RE IDENTIFYING PEOPLE WITHIN THE DATABASE THEY DISGUISE THEMSELVES UNDER PSEUDO NAMES SUCH AS BOB

BOB
HACKER

IN ORDER TO PREVENT BOB FROM SIMPLY ASKING ENOUGH QUESTIONS TO ACCESS ALL THE CITIZEN DATA, ALICE DECIDES TO IMPLEMENT A DIFFERENTIALLY PRIVATE APPROACH. WHEN BOB ASKS A QUESTION , ALICE WILL GIVE A RANDOMIZED ANSWER THAT'S PROBABILISTICALLY CLOSE TO THE REAL ANSWER BUT NOT GUARANTEED

?

TRUE OR RANDOM
(95%) (5%)

THUS, IF BOB IS AN INNOCENT GOVERNMENT OFFICIAL HE'LL STILL GET A USEFUL ANSWER THAT'S APPROXIMATELY CORRECT. OTHERWISE, ALICE IS STILL CONFIDENT THAT NO ONE'S PRIVACY IS COMPROMISED. EVERYONE WINS!

TO PUT IT IN OTHER TERMS, DIFFERENTIAL PRIVACY GUARANTEES THAT ALICE'S ANSWER WOULD ALMOST ALWAYS BE THE SAME REGARDLESS OF WHETHER A SPECIFIC INDIVIDUAL WAS OR WAS NOT INCLUDED IN THE DATASET. THEREBY, BOB CAN'T USE HIS QUESTIONS TO DETERMINE IF A SPECIFIC PERSON'S DATA IS PRESENT OR NOT REGARDLESS OF WHAT HE KNOWS

HOLD UP! I DON'T KNOW IF I MISHEARD BUT DIDN'T YOU SAY ALMOST ALWAYS? DOES THIS MEAN THAT BOB COULD STILL EVENTUALLY FIGURE OUT IF AN INDIVIDUAL WAS IN THE DATABASE??

GREAT OBSERVATION! BOB DOES INDEED STILL LEARN A LITTLE BIT ABOUT THE DATABASE EACH TIME HE ASKS A QUESTION, BUT THAT'S WHERE THE CONCEPT OF A PRIVACY BUDGET COMES INTO PLAY.

WITH DIFFERENTIAL PRIVACY WE CAN PRECISELY MATHEMATICALLY DEFINE HOW MUCH PRIVACY BUDGET TO ASSIGN TO A SYSTEM WHERE EACH QUESTION WILL EXHAUST A SMALL PART OF THAT BUDGET.

Budget left
399

SO WE KNOW PRECISELY WHAT PRIVACY RISK WE HAVE?

YES, EXACTLY! THAT'S THE BEAUTY OF DIFFERENTIAL PRIVACY IN IT'S STRONG DEFINITION OF PRIVACY.

SO IF BOB HAD ASKED ENOUGH QUESTIONS SO THAT HE FULLY USED UP THE BUDGET, HE WOULD THEN BE ABLE TO IDENTIFY INDIVIDUALS IN THE DATASET. BUT FORTUNATELY WE HAVE ALICE TO MONITOR THIS BUDGET! SHE CAN SEE THAT WE'RE CLOSE TO EXHAUSTING THE PRIVACY BUDGET THEN STOP ANSWERING ANY QUESTIONS UNTIL WE MAKE FURTHER CHANGES AND CAN SAFELY ANSWER MORE QUESTIONS.

STOP

THAT'S AWESOME! SO NOW THAT DIFFERENTIAL PRIVACY IS IMPLEMENTED, ALL OF MY DATA IS SAFE THEN RIGHT? I CAN TRUST ALL SYSTEMS THAT ARE DIFFERENTIALLY PRIVATE!

WHOA, NOT SO FAST THERE! JUST BECAUSE A COMPANY MAKES DIFFERENTIAL PRIVACY CLAIMS FOR THEIR SYSTEM DOESN'T MEAN YOU CAN JUST BLINDLY TRUST THEM.

YOU WOULDN'T JUST BLINDLY ENTER A CAR AND IMMEDIATELY TRUST THE DRIVER WITH YOUR LIFE, ESPECIALLY GIVEN THIS CAR HAS FAR MORE DIRE CONSEQUENCES IF IT CRASHES. AS A PASSENGER, YOU'D STILL WATCH THE ROAD AND CHECK THE DRIVER'S OVERALL COMPETENCY BEST YOU CAN.

Google

Google

HERE, THE CAR IS THE DIFFERENTIALLY PRIVATE SYSTEM. JUST LIKE CARS, THESE SYSTEMS HAVE TO BE CAREFULLY DESIGNED, BUILT, AND THEN DRIVEN AS THERE'S TONS OF PLACES TO SCREW UP AND NEED TO BE ASSESSED. HOW MUCH OF THE PRIVACY BUDGET IS BEING ALLOCATED? WHO IS MONITORING THE BUDGET? HOW USEFUL ARE THE SLIGHTLY RANDOMIZED ANSWERS? WHY BUILD A CAR WHEN A BIKE IS SUFFICIENT?

?

Tool Box

IF ANYTHING, COMPANIES WHO MAKE EXTRAORDINARY CLAIMS ABOUT THEIR PRIVACY GUARANTEES THROUGH DIFFERENTIAL PRIVACY NEED TO BE HELD TO SIMILARLY EXTRAORDINARY STANDARDS IN ENSURING PROPER IMPLEMENTATION AND EXECUTION BY THE PUBLIC. ONE SHOULD NEVER ASSESS PRIVACY SOLELY BASED ON DIFFERENTIAL PRIVACY."

AH, I THINK I UNDERSTAND NOW! I SHOULD VALUE WHAT THESE COMPANIES AND ORGANIZATIONS ARE DOING WITH MY DATA AND DO MY BEST TO ENSURE THAT THEY ARE PROTECTING IT CORRECTLY.

YES, PRECISELY! YOU AS THE CUSTOMER AND CITIZEN, SHOULD ALWAYS BE CRITICAL AND IMPLORE THESE COMPANIES TO ADAPT PROPER, TRANSPARENT PRIVACY PRACTICES AS THE NORM. ABOVE ALL, YOU SHOULD BE GUARANTEED BOTH UTILITY AND PRIVACY ALONGSIDE BEING INFORMED ABOUT WHAT EXACTLY YOU'RE CONSENTING TO WHEN SHARING YOUR DATA REGARDLESS OF THE CIRCUMSTANCES. FOR US AT THE US CENSUS BUREAU, DIFFERENTIAL PRIVACY IS JUST THE BEGINNING OF THIS GUARANTEE!"

THANK YOU SO MUCH DR. ABOWD! I'M GOING TO GO MAKE A TIKTOK ABOUT THIS NOW; I'M SURE MY VIEWERS WILL LOVE THIS!