You Don't Say: An FTC Workshop on Voice Cloning Technologies
January 28, 2020
Transcript

HELLO, GOOD AFTERNOON.
THANK YOU VERY MUCH FOR
ATTENDING OUR WORKSHOP AND THANK
YOU TO THOSE OF US JOINING US ON
THE WEBCAST. MY NAME IS MIN
HEE KIM, I AM WITH THE FTC'S
OFFICE OF TECHNOLOGY THE,
RESEARCH AND INVESTIGATION.
BEFORE WE START I HAVE MANDATORY
ADMINISTRATIVE ANNOUNCEMENTS TO
GO OVER.
PLEASE SILENCE YOUR MOBILE
PHONES AND ANY OTHER ELECTRONIC
DEVICES.
IF YOU MUST USE THEM DURING THE
WORKSHOP, PLEASE BE RESPECTFUL
OF SPEAKERS AND FELLOW AUDIENCE
MEMBERS.
BE AWARE IF YOU LEAVE THE
BUILDING FOR ANY REASON DURING
THE WORKSHOP YOU HAVE TO GO BACK
THROUGH SECURITY SCREENING.
BEAR THIS IN MIND AND PLAN AHEAD
ESPECIALLY IF YOU'RE
PARTICIPATING ON A PANEL SO WE
CAN REMAIN ON SCHEDULE.
MOST OF YOU RECEIVED A LAN YARD
WITH A SECURITY BADGE, WHEN YOU
LEAVE FOR THE DAY, PLEASE RETURN
THE BADGE TO SECURITY ON YOUR
WAY OUT.
IF AN EMERGENCY OCCURS THAT
WOULD CAUSE YOU TO LEAVE THE
CONFERENCE, REMAIN IN THE
BUILDING, FOLLOW THE
INSTRUCTIONS PROVIDED OVER THE
BUILDING SYSTEM.
IF AN EMERGENCY OCCURS THAT
REQUIRES EVACUATION, ALARM WILL
SOUND, EVERYONE SHOULD LEAVE THE
BUILDING IN AN ORDERLY MANNER
THROUGH MEANS INSTRUCTED.

AFTER LEAVING THE BUILDING, PROCEED LEFT, PROCEED ACROSS E STREET TO THE EMERGENCY ASSEMBLY AREA.

REMAIN IN THE ASSEMBLY AREA UNTIL INSTRUCTED TO RETURN TO THE BUILDING.

IF YOU NOTICE ANY SUSPICIOUS ACTIVITY, PLEASE ALERT BUILDING SECURITY.

PLEASE BE ADVISED THIS EVENT WILL BE PHOTOGRAPHED, WEBCAST, RECORDED.

BY PARTICIPATING IN THIS EVENT YOU'RE AGREEING TO YOUR IMAGE AND ANYTHING YOU SAY OR SUBMIT MAY BE POSTED INDEFINITELY AT FTC.GOV OR ON ONE OF THE COMMISSION'S PUBLICLY AVAILABLE SOCIAL MEDIA SITES.

TAKE SEATS RATHER THAN STANDING AS IT'S AGAINST THE FIRE CODE.

RESTROOMS LOCATED IN THE HALLWAY OUTSIDE THE CONFERENCE ROOM.

THE CAFETERIA WILL BE OPEN UNTIL 2:00 P.M. AND THEN WILL OFFER LIMITED SERVICE UNTIL 3:00.

QUESTION CARDS ARE AVAILABLE IN THE HALLWAY ON THE INFORMATION TABLE IMMEDIATELY OUTSIDE THE CONFERENCE ROOM.

THERE ARE ALSO SEVERAL FTC STAFF THROUGHOUT THE ROOM WHO HAVE QUESTION CARDS.

IF YOU HAVE A QUESTION, PLEASE RAISE YOUR HAND, FILL OUT THE CARD AND FTC STAFF WILL COME TO RETRIEVE IT.

FOR THOSE OF YOU PARTICIPATING BY WEBCAST, TWEET YOUR QUESTIONS AT FTC, USING THE HASHTAG VOICE CLONING FTC.

PLEASE UNDERSTAND THAT WE MAY NOT GET TO ALL QUESTIONS.

NOW I AM HONORED TO INTRODUCE FTC COMMISSIONER ROHIT CHOPRA

WHO WILL PROVIDE OPENING
REMARKS.
[APPLAUSE]
>> ALL RIGHT.
THANKS A, MIN HEE.
THANK YOU ALL AND WELCOME.
ACCORDING TO RECENT SURVEYS,
AMERICANS ARE LOSING TRUST IN
TECHNOLOGY AND TECHNOLOGY
COMPANIES.
AND MANY ARE QUESTIONING THEIR
ROLE IN IMPACTING SOCIETY.
WE ALL KNOW THE BENEFITS OF NEW
TECHNOLOGIES AND THE IMMENSE
FORTUNES THAT THEY CAN SHOWER ON
US AND THE COMPANIES THAT CREATE
THEM.
BUT MANY OF US ARE CONCERNED
ABOUT HOW THESE TECHNOLOGIES ARE
MISUSED AND ABUSED.
TODAY TECHNOLOGY AND DATA ARE
WEAPONIZED BY THOSE WHO WISH TO
DO OUR COUNTRY AND OUR SOCIETY
HARM.
WE HAVE SEEN HOW THE HARVESTING
OF PERSONAL DATA CAN THREATEN
OUR PERSONAL SAFETY.
MASSIVE DATA BREACHES FROM THE
OFFICE OF PERSONNEL MANAGEMENT,
MARRIOTT, AND EQUIFAX, WE HEAR
ABOUT THESE EVERY DAY AND MANY
HAVE BEEN LINKED TO ACTORS IN
CHINA.
A CHINESE GAMING COMPANY WAS
RECENTLY ORDERED BY THE
COMMITTEE ON FOREIGN INVESTMENT
OF THE UNITED STATES TO DIVEST
RECENT ACQUISITION OF GRINDR,
POPULAR GAY DATING APP AS MANY
EXPECTED DATA USED FOR BLACK
MAIL, OTHERS TARGET US WITH
VAGUE INFORMATION AND CREATE
SOCIAL AND POLITICAL DIVISIONS
WHICH CAN DISRUPT OUR ECONOMY,
DEMOCRACY AND SOCIETY.
PRIVACY IS CLEARLY NOW A

NATIONAL SECURITY ISSUE AND A
PERSONAL SECURITY ISSUE.
IT'S NOT JUST ABOUT SURVEILLANCE
OF OUR MOVEMENTS AND OUR SOCIAL
INTERACTIONS.
NEW TECHNOLOGY CAN NOW ALLOW US
TO CLONE WHAT WE THOUGHT WAS
UNIQUELY OURS, OUR BIOMETRICS,
FROM OUR FINGERPRINTS TO OUR
FACES, LOSING CONTROL OF OUR OWN
BIOMETRICS POSES ANOTHER LEVEL
OF PERIL.
WHEN THIS HAPPENS, DEEP FAKES,
DISINFORMATION AND DISTRUST WILL
ACCELERATE.
WE'LL NEED TO DETERMINE HOW TO
CONTROL THIS TECHNOLOGY AND KEEP
IT OUT OF THE WRONG HANDS.
AND TO PROTECT OUR SAFETY AND
SECURITY FROM THE DANGERS OF
BIO
METRIC THEFT, WE NEED TO BE
FORWARD-THINKING WHICH IS SO
GOOD FTC IS HOSTING TODAY.
WE CAN'T WAIT FOR MASSIVE HARM
TO OCCUR.
THIS MAY INVOLVE NEW LAWS,
RULES, LICENSING, EXPORT
CONTROLS AND OVERSIGHT.
TODAY WE'RE GOING TO EXAMINE ONE
SPECIFIC ELEMENT OF BIOMETRICS:
VOICE CLONING.
A GOAL OF TODAY'S WORKSHOP IS TO
LEARN ABOUT THIS TECHNOLOGY, ITS
IMPLICATIONS AND WHAT CAN BE
DONE TO CONFRONT THE DANGER IT
POSES.
TO THAT END, WE'RE GOING TO
BEGIN THE DAY WITH A KEYNOTE
PRESENTATION FROM PATRICK
TRAYNOR, THE JOHN AND MARY LOU
DASBURG PRE-EMINENCE CHAIR OF
FLORIDA'S COLLEGE OF ENGINEERING
WHO WILL DISCUSS THE STATE OF
VOICE CLONING TECHNOLOGY.
FOLLOWING PROFESSOR TRAIN NOR'S

PRESENTATION, OUR FIRST PANEL WILL DISCUSS HOW VOICE CLONING TECHNOLOGY CAN BE HARNESSED TO ENABLE PEOPLE WHO LOST THEIR VOICES DUE TO INJURY OR DISEASE TO SPEAK AGAIN USING REALISTIC VERSIONS OF THEIR PRE-INJURY AND PRE-DISEASE VOICES.
BUT WE'RE ALSO GOING TO HEAR ABOUT THE RISKS VOICE CLONING TECHNOLOGIES POSE AND HOW THEY'RE LIKELY TO BE USED IN SOCIAL ENGINEERING ATTACKS.
THIS IS GOING TO GIVE US A CLEARER SENSE OF WHAT ARE THE RANGE OF HARMS THAT THIS TECHNOLOGY CAN POSE?
THE SECOND PANEL WILL FOCUS ON HOW TECHNOLOGISTS, CORPORATIONS AND SOCIETY SHOULD EVALUATE THESE NEW TECHNOLOGIES AND WHETHER WE SHOULD IMPOSE CONTROLS ON THE TECHNOLOGIES INTRODUCTION AND AVAILABILITY.
THE FIRST TWO PANELS WILL TEE UP THE LAST PANEL OF TODAY'S WORKSHOP, WHICH WILL EXAMINE WHETHER TECHNOLOGIES EXIST OR ARE LIKELY TO BE CREATED THAT CAN MITIGATE THE POTENTIAL HARMS OF VOICE CLONING TECHNOLOGY.
WE ALL LOOK FORWARD TO A LIVELY AND INFORMATIVE DISCUSSION AND WITH THAT, I JUST WANT TO THANK ALL OF THE FTC STAFF WHO HAVE WORKED TO BRING PANELISTS HERE TO DESIGN LOGOS, TO ESCORT YOU TO PROVIDE SECURITY, AND EVERYONE WHO MADE THIS DAY POSSIBLE.
SO I LOOK FORWARD TO EVERYONE'S PARTICIPATION AND IT'S MY PLEASURE TO TURN THE DAIS OVER TO OUR KEYNOTE SPEAKER, PROFESSOR PATRICK TRAYNOR.
[APPLAUSE]

GOOD AFTERNOON, EVERYBODY.
MY NAME IS PATRICK TRAYNOR AND
AS WAS SAID JUST A MOMENT AGO,
I'M A PROFESSOR IN COMPUTER
SCIENCE AT THE UNIVERSITY OF
FLORIDA.
NOW, THERE MY RESEARCH FOCUSES
ON THE INTERSECTION BETWEEN
CYBERSECURITY AND VOICE
INTERFACE TECHNOLOGY.
MY TASK TODAY IS NOT TO GIVE AN
OVERLY TECHNICAL TALK.
SO FOR THOSE OF YOU CONCERNED A
COMPUTER SCIENTIST IS SPEAKING
AND THAT YOU'VE GONE TO YOUR
E-MAIL ALREADY, MY GOAL IS IN
FACT TO BRING US TOGETHER AND TO
TALK AT A HIGH LEVEL ABOUT HOW
WE HAVE COME SO FAR SO FAST AND
HOPEFULLY SET UP THE OTHER
PANELS.
WE'LL GET IN TO MUCH MORE
TECHNICAL DISCUSSIONS AS THE DAY
GOES ON.
LET ME BEGIN BY ASKING A LARGER
AND MORE IMPORTANT QUESTION AND
THAT'S WHAT IS IT THAT MAKES US
UNIQUE?
IN PARTICULAR, WHAT IS IT THAT
SEPARATES HUMAN BEINGS FROM THE
REST OF THE ANIMAL KINGDOM?
IF I TOOK A POLL IN THIS ROOM
YOU MIGHT POINT TO ARCHITECTURE
OR ABILITY TO DO SCIENCE.
BUT SCIENTISTS WOULD ACTUALLY
TELL THAT YOU IT'S QUITE LIKELY
LANGUAGE AND SPEECH SEPARATES
US.
YOU MAY WANT TO INTERRUPT AND
SAY PATRICK, MY DOG SPEAKS, MY
PET BIRD SPEAKS, ANIMALS
CERTAINLY COMMUNICATE.
AND THAT'S RIGHT.
BUT WHAT MAKES HUMAN LANGUAGE
AND SPEECH SO INTERESTING IS
THAT WE TAKE A VERY SMALL SET OF

UTTERANCES, SOUNDS, AND COMBINE THEM EFFECTIVELY IN TO INFINITE OUTPUTS.
THAT IS THINK OF ALL THE SENTENCES YOU CAN POSSIBLY SAY WITH ONLY A FEW VERY SIMPLE SOUNDS.
IT'S REALLY AMAZING BECAUSE IT ALLOWS US TO DO SOMETHING THAT OTHER ANIMALS CAN'T.
WE DON'T JUST SAY THERE'S A PREDATOR OVER THERE BEHIND THE BUSH OR THERE'S GOOD FOOD BY THE RIVER.
IT ALSO ALLOWS US TO EXPLAIN CONCEPTS, ALLOWS US TO THINK ABOUT THE THEORETICAL, IT ALLOWS US TO LIE TO EACH OTHER, TO BE TRULY SOCIAL BEINGS.
SO SPEECH IS INCREDIBLY IMPORTANT.
IT IS IN FACT POWER.
NOW, ONCE HUMANS FIGURED THIS OUT, WE SET ABOUT THE TASK OF TRYING TO FIGURE OUT HOW WE COULD GIVE THAT POWER TO EVERYTHING ELSE AND THERE'S AT LEAST A THOUSAND YEARS OF RECORDED HISTORY OF HUMANS TRYING TO GET EVERYTHING FROM ROCKS TO STATUES TO OTHER ANIMALS TO SPEAK.
IF THEY COULD JUST SPEAK THEY WOULD REVEAL THEIR TRUTHS AND TRUTHS OF THE UNIVERSE TO US.
OKAY.
AS YOU CAN IMAGINE, ROCKS, ANIMALS, WE HAVEN'T DONE SO WELL WITH.
BUT LET'S JUMP FORWARD ABOUT 900 YEARS SINCE RECORDED HISTORY CAUGHT US TRYING TO DO THIS AND TALK VERY BRIEFLY ABOUT SOME OF THE TECHNOLOGIES WE HAVE HAD ALONG THE WAY AND WHAT'S REALLY ACCELERATED US.

I WANT TO BRING YOU IN TO THE
'60s AND '70s, WE HAD VERY EARLY
HARDWARE/ELECTRONIC MODELS FOR
SPEECH.
THESE ARE KIND OF HOW YOU DESIGN
SPEECH IF YOU DIDN'T HAVE ANY
BACKGROUND KNOWLEDGE.
SO THINK ABOUT IT IN THE
FOLLOWING WAY: YOU COULD RECORD
A DICTIONARY OF A LIMITED SET OF
TERMS AND MATCH THOSE TERMS
TOGETHER.
IF YOU WERE MORE SOPHISTICATED,
TAKE SOUNDS AND THEN GLUE THOSE
TOGETHER.
THE PROBLEM WITH THAT OF COURSE
IS THAT IMAGINE PUTTING A PUZZLE
TOGETHER THAT THE PIECES DON'T
QUITE FIT.
YOU COULD SMASH THEM OR GLUE
THEM AS MY KINDERGARTNER AT HOME
DOES.
BUT YOU WOULDN'T HAVE TRULY
HUMAN SOUNDING SPEECH.
LET ME GIVE YOU AN EXAMPLE NOW
OF WHAT I'M TALKING ABOUT.
EVERYONE I ASSUME WHO IS OLDER
THAN I AM KNOWS WHAT THIS IS.
FOR YOUNGER FOLKS, THIS IS
CALLED A SPEAK AND SPELL.
I THINK THE VOLUME IS BEING
CONTROLLED IN THE BACK.
>> NOW SPELL VST.
>> THAT IS DONE BY USING
SYNTHESIS.
WE TAKE A SMALL NUMBER OF SOUND
WAVES, PUT THEM TOGETHER AND WE
GET SOMETHING THAT'S CLOSE
ENOUGH TO HUMAN SPEECH.
ANOTHER EXAMPLE OF THIS, QUITE
PROMINENT, IS THE SCIENTIST
STEVEN HAWKING ORIGINALLY ABLE
TO SPEAK IN THE '70s BECAUSE OF
SYNTHESIS STYLE MODELS.
INTERESTINGLY ENOUGH WHILE
MODELS PROGRESSED SIGNIFICANTLY

SINCE THE '70s, HAWKING DECIDED
HE WAS GOING TO KEEP THAT VOICE
BECAUSE IT'S PART OF WHAT MADE
HIM UNIQUE.
IN FACT IT WAS HIS VOICE.
AS EVERYONE NOTICED OF COURSE,
THE THING THAT I JUST PLAYED TO
YOU WAS INCREDIBLY ROBOTIC.
THAT'S BECAUSE HUMAN SPEECH IS
FAR MORE COMPLEX THAN TAKING TWO
WAVES AND PUTTING IT TOGETHER.
THERE'S A LOT MORE THAT GOES IN
TO US MAKING THE SOUND.
TO PROGRESS FURTHER IN HISTORY
OR FURTHER IN TO THE PRESENT,
LET'S FIGURE OUT WHAT PIECES
COME NEXT.
THAT'S ACTUALLY THE MODELLING OF
THE HUMAN VOCAL TRACT.
THIS SHOULD LOOK FAMILIAR BUT
MOST OF YOU DON'T GET TO SEE
YOURSELF IN CROSS-SECTION ALL
THAT OFTEN.
THIS IS MOST OF THE HUMAN VOCAL
TRACT.
WHAT YOU SEE HERE ARE ROUGHLY
20, ALTHOUGH THERE ARE MORE,
IT'S MISSING MOST OF THE NOSE,
DIFFERENT PHYSICAL FEATURES THAT
ALLOW US TO MAKE DIFFERENT
SOUNDS, TO CHANGE THE SOUNDS
THAT ARE COMING OUT OF OUR BODY.
AS A SIMPLE EXAMPLE FOR ALL OF
YOU, THINK ABOUT THE DIFFERENCE
IF YOU SAY AN A SOUND AND A O
SOUND.
PUT HEY POOL TOGETHER.
THE A AND O ARE QUITE DIFFERENT.
ALL THAT CHANGES IS ONE
DIMENSION.
YOUR LIPS MOVE FROM WIDE OPEN
TO, HEY, TO POOL TO, A CIRCULAR
PATTERN.
NOW UNDERSTAND THAT TO MAKE
SOUND FROM THIS APPARATUS, YOU
CAN MAKE CHANGES IN ALL OF THESE

DIMENSIONS, ADDS COLOR,
VARIATION, DIMENSION TO THE KIND
OF NOISES THAT WE CAN MAKE.
WHEREAS THOSE EARLIER SYNTHESIS
BASED MODELS ARE GOING TO BE
SIMPLE WAVE OR TWO, WE CAN MAKE
COMBINATIONS OF LOTS OF WAVES TO
FORM MUCH MORE COMPLEX SOUNDS.
WHEN WE VIEW THEM TOGETHER, THEY
LOOK LIKE THIS.
NOW, THE FIGURE ON THE LEFT IS
JUST WHITE NOISE.
I'M SHOWING A SPEC TOE GRAM,
LOOKING AT POWER, MORE POWER
SIGNIFIED BY YELLOW AND LESS
POWER SIGNIFIED BY RED.
WHITE NOISE, THIS WILL BE AN AGE
TEST FOR THE AUDIENCE, REMEMBER
TURNING TO A TV STATION THAT
WASN'T THERE AND YOU GET STATIC?
THAT'S CLASSICAL WHITE NOISE,
ALSO GREAT FOR INFANTS TO SLEEP.
WHAT YOU SEE ON THE RIGHT IS THE
PHRASE TURN ON THE COMPUTER,
BEING SPOKEN.
AS WE MOVE THROUGH TIME, MOVE
LEFT TO RIGHT, THAT AT DIFFERENT
TIMES, DIFFERENT PARTS OF THE
SPECTRUM HAVE DIFFERENT POWER.
BUT IT'S NEVER JUST ONE SMALL
PART.
THAT YELLOW LIGHTS UP IN HIGHER
FREQUENCIES NOT JUST FROM TONES
I CREATE BUT FROM THE HARMONICS
FROM MY VOICE.
SO IT TURNS OUT THAT THIS
INSTRUMENT WITH MANY DIFFERENT
PLACES THAT I CAN TUNE AND
CHANGE CAN CREATE COMPLEX SOUNDS
THAT ARE REALLY VERY DIFFICULT
TO REPLICATE.
PEOPLE HAVE BEEN TRYING THIS FOR
A LONG TIME, BUT TO GET A
PHYSICAL MACHINE TO DO THIS IS
QUITE DIFFICULT.
SO WHAT HAVE WE DONE?

WELL, IT WOULD BE NICE IF WE HAD
SOME WAY TO AUTOMATICALLY
OPTIMIZE AND GUESS AND TIE IN
THINGS TO WHERE WE THINK HUMANS
ARE WHEN THEY SAY CERTAIN
THINGS.
AND IF YOU REMEMBER NOTHING ELSE
FROM THIS TALK, I WANT YOU TO
REMEMBER THIS SLIDE.
THIS SLIDE IS REALLY HOW WE
TRANSFORM FROM SAY HUNDRED OF
YEARS OF PROGRESS TO
RIDICULOUSLY FAST PROGRESS OVER
THE COURSE OF THE LAST FIVE
YEARS.
THE STATE OF THE ART IS
SOMETHING CALLED DEEP LEARNING.
YOU MAY HAVE HEARD OF IT THROUGH
OTHER SIMILAR NAMES, MACHINE
LEARNING, ARTIFICIAL
INTELLIGENCE.
THE VERY SIMPLE WAY TO VIEW THIS
IS FIRST, IT'S A VERY, VERY
COMPLICATED STATISTICAL MODEL OF
A DATA SET.
AND I LIKE TO VIEW THEM IN BLACK
BOXES, EFFECTIVELY I TAKE THIS
BOX AND PUSH SOME KIND OF INPUT
IN.
WHEN IT COMES OUT THE OTHER
SIDE, I SAY YES, YOU DID A GOOD
JOB OR NO, THAT'S NOT WHAT I
LOOKED FORMENT I WILL KEEP
PUSHING LOTS AND LOTS OF INPUT
THROUGH IT UNTIL IT REGULARLY IS
ABLE TO CLASSIFY THINGS THE WAY
THAT I WANT THEM TO.
PLEASE UNDERSTAND BY THE WAY,
THIS IS A GROSS
OVERSIMPLIFICATION OF THE STATE
OF THE ART, BUT THIS ALSO ISN'T
A MATH CLASS AT THE GRADUAL
LEVEL.
I'M CERTAINLY HAPPY TO TALK
ABOUT DETAILS OF THIS A LITTLE
LATER.

OKAY.
SO I HAVE THIS BLACK BOX AND I
SEND ALL KIND OF THINGS THROUGH
IT.
THE THING THAT I WANT YOU TO
UNDERSTAND IS THAT MACHINE
LEARNING IS NOT BY ANY MEANS A
PERFECT TOOL.
AND SO IT WILL ALSO LEARN THINGS
THAT ARE NOT CORRECT.
REMEMBER, BECAUSE IT'S A BLACK
BOX IT'S JUST ABOUT INPUTS AND
OUTPUTS.
IF IT LEARNS A CRAZY PATH
BETWEEN TWO THINGS, THAT PATH
MAY NOT ACTUALLY HAVE ANY
RELEVANCE IN REALITY.
AND IN THE LAST PANEL TODAY,
CERTAINLY WE'LL TALK ABOUT WHY
THAT MAKES THESE SYSTEMS
VULNERABLE THEMSELVES TO ATTACK
AND NOT JUST THE SOURCE OF IT
WHICH WE'LL TALK ABOUT LATER.
OKAY.
SO WE HAVE VERY COMPLEX
STATISTICAL MODELS.
WE'RE DOING A LOT OF MATH.
WHAT'S INTERESTING?
WELL, IT'S THIS HARDWARE
REVELATION.
THIS IS CALLED A GPU.
SOME OF YOU WHO ARE MORE
FAMILIAR WITH CRYPTOCURRENCIES
WILL DEAL WITH THIS EXACT SAME
DEVICE.
WHAT HAPPENED IS STARTING IN
2012 THESE DEVICES WERE
AVAILABLE TO THE MASS MARKET.
PEOPLE HAVE BEEN PURCHASING
THEM, IF YOU'RE A STOCKHOLDER
FOR EXAMPLE IN MANY COMPANIES
THAT MAKE THESE, YOU'RE PROBABLY
DOING VERY WELL.
THESE HAVE ADVANCED COMPUTING
POWER, ALLOW US TO THE THIS
SPECIFIC KIND OF MATH,

VERY EFFICIENTLY.
WE TAKE THESE ALGORITHMS THAT
EFFECTIVE.
CAN LEARN AND BE TAUGHT IN A
DIRECTION AND WE THROW
TREMENDOUS COMPUTING POWER AT
THEM AND GET VERY, VERY POWERFUL
MODELS.
WE DON'T HAVE TO KNOW
NECESSARILY THE POSITION OF
EVERY SINGLE ONE OF THE
BIOLOGICAL FEATURES IF WE THROW
ENOUGH DATA AT IT AND COMPUING
POWER THE MODEL WILL FIGURE IT
OUT OR CLOSE ENOUGH TO.
WRAP UP HERE, THE OLDER RESULTS
THAT I'M TALKING ABOUT HERE ARE
TWO TO THREE YEARS OLD.
THAT'S REALLY TO GIVE YOU AN
IDEA OF HOW RAPIDLY THIS FIELD
IS PROGRESSING.
SO I'M TALKING IN PARTICULAR
ABOUT DEGENERATIVE MODELS.
THE IDEA OF GENERATIVE MODELS,
WITH ENOUGH INPUTS THEY CAN HAVE
DEEP LEARNING, KNOWN AS DANS.
TRADITIONAL MEANS TWO TO THREE
YEARS AGO, TO THE THIS WAS YOU
TOOK AS MUCH AUDIO AS YOU COULD
GET AND YOU SHOVED IT IN TO THIS
BLACK BOX AND IT WOULD BE ABLE
TO DEVELOP A MODEL HOPEFULLY
THAT WOULD SOUND CLOSE TO WHAT
YOU WANTED IT TO SOUND TO.
I WILL PRESS THIS AGAIN.
HOPEFULLY GET --
>> THE BUSES AREN'T THE PROBLEM.
THEY PROVIDE A SOLUTION.
>> THIS AUDIO BY THE WAY IS
GENERATED BY A SYSTEM CALLED
TACOTRON BY GOOGLE AND THIS
MODEL IS FAIRLY ADVANCED.
IT CAN BE INSTRUCTED TO GIVE
EMPHASIS TO CERTAIN WORDS.
I'LL PLAY IT AGAIN SO YOU CAN
HEAR IT.

NOTICE THAT IT EMPHASIZES THE WORD PROBLEM AND SOLUTION. MAKING IT FAR MORE CONVINCING THAT IT IS ORGANIC OR HUMAN SPEECH.

>> THE BUSES AREN'T THE PROBLEM. THEY ACTUALLY PROVIDE A SOLUTION.

>> OKAY?

THE CHALLENGE WITH THESE KIND OF MODELS IS THE FOLLOWING: FIRST, TACOTRON IS ABOUT CREATING SYNTHETIC VOICES, NOT NECESSARILY TRYING TO MATCH VOICES.

THIS ISN'T GETTING EXACTLY AT THE PROBLEM THAT THE WORKSHOP IS ABOUT NECESSARILY.

BUT THE OTHER CHALLENGE IS THAT IT TAKES A LONG TIME IN ORDER TO TRAIN THESE MODELS.

YOU HAVE TO HAVE SIGNIFICANT COMPUTING POWER TO THE THIS CONVINCINGLY.

THAT THE REALISM REALLY IS RELATIVE.

I GAVE YOU A VERY SHORT PIECE OF AUDIO AND WHEN THE PROGRAMMERS BEHIND THIS FED THIS TEXT TO THE SYSTEM, THEY CAPITALIZED THE WORDS PROBLEM AND SOLUTION.

SO THEY GAVE THE MODEL HINTS WHERE IT SHOULD HAVE EMPHASIS.

IN GENERAL, ONE OF THE WAYS TO KNOW THAT YOU'RE TALKING TO THESE MODELS IS THEY HAVE A FLAT DELIVERY, THE EMOTION IS GOING TO BE ONE OF THE LIMITATIONS.

ALL RIGHT.

THE LAST MAIN SLIDE I WILL TALK ABOUT IS HOW FAR THINGS HAVE COME IN THAT TWO TO THREE-YEAR PERIOD.

WHAT WE HAVE NOW IS VERY, VERY FAST SYNTHESIS.

I'M TALKING LARGELY ABOUT A

PAPER THAT WAS PUBLISHED IN 2008, THE TOP VENUE IN COMPUTER SCIENCE FOR DEEP LEARNING, MACHINE LEARNING.
WHAT THESE FOLKS REALIZED IS THAT I DON'T ACTUALLY HAVE TO CAPTURE EVERYTHING TO CREATE A CONVINCING MODEL OF YOUR SPEECH.
WE ALL REUSE VERY SIMILAR PARTS OF SPEECH.
PARTICULAR THINGS LIKE Ps AND Ts, THOSE THINGS ALL SOUND THE SAME.
SO IF EVERYONE IN THIS ROOM WENT P-T, IT WOULD SOUND ARTICLE.
WE WILL LEARN FROM A LARGE CROWD, PRE-TRAIN BASED ON THINGS THAT ARE THE SAME AND LISTEN FOR PLACES LIKE THE VOWELS, WHERE WE'RE GOING TO HAVE VARIATION.
SO I CAN CREATE A GIANT POPULATION MODEL AND THEN PLUG IN SHORT BITS OF YOUR AUDIO IN TO THAT AND USE THAT TO CUSTOMIZE WHAT IS OUTPUT TO YOU, THE SPEAKER.
WHAT I WILL PLAY FOR YOU HERE FIRST IS A HUMAN SPEAKER AND THEN THE SECOND IS GOING TO BE SYNTHETIC SPEECH TRYING TO CONVINCE YOU THAT IT'S THE HUMAN SPEAKER.
THE HUMAN SPEAKER DID NOT EVER SAY THESE WORD.
AND THE SECOND SAMPLE WAS CREATED ONLY FROM FIVE, FROM ABOUT FIVE SECONDS AND FIVE SAMPLES OF THAT PERSON'S VOICE.
REAL PERSON FIRST.
>> WE ALSO NEED A SMALL PLASTIC SNAKE AND A BIG TOY FROG FOR THE KIDS.
>> OKAY.
ORGANIC HUMAN SPEECH.
>> LEARN ABOUT SETTING A NETWORK

CONFIGURATION.
>> FIVE FIVE-SECOND SAMPLES
ALLOWED US, NOT US, ALLOWED
SCIENTISTS TO CREATE THAT
SAMPLE.
NOW LISTEN.
FROM THE PERSPECTIVE OF
LIMITATIONS THERE WERE A LOT
FUTURE.
LISTENED CAREFULLY YOU HEARD
THINGS ABOUT THE SYNTHETIC
SPEECH YOU DIDN'T LIKE.
SMALL SETS OF SAMPLES CONVINCE
FOLKS IN LIMITED CONDITIONS, ALL
RIGHT?
YOU ONLY HAVE A SMALL SET OF
SAMPLES IT'S GOING TO SOUND
IMPERFECT, A LITTLE ROBOTIC.
HEARING THAT OVER THE THING
TODAY, ANYBODY, ANYBODY HAVE ANY
DIFFICULT HEARING THE DIFFERENCE
BETWEEN THE FIRST AND THE SECOND
SPEAKER?
IF I ASKED YOU TO CLASSIFY WHICH
WHICH WAS LIKELY ORGANIC AND
WHICH WAS GENERATED, COULD YOU
ALL HAVE SAID LIKELY THE SECOND?
YES, BUT IMAGINE THE SECOND
AUDIO CAME OVER THE PHONE, A
ROBO CALL, AUDIO AUTOMATICALLY
DEGRADED ANYWAY BY THE NATURE OF
HOW FAR PHONE NETWORKS WORK.
YOU MAY NOT HAVE BEEN ABLE TO DO
IT.
CERTAINLY REMOVING A, WE'RE
MOVING IN A DIRECTION WHERE THAT
IS THE LACK, A LARGE NUMBER OF
SAMPLES IS GOING TO BECOME LESS
PROBLEMATIC.
BUT IT DOES ALSO HOLD THAT AS
YOU ADD MORE AND MORE SAMPLES
YOU CAN OVERCOME THIS.
PLAY A PIECE OF AUDIO IN A
SECOND THAT SHOWS YOU THAT SAME
MODEL THAT'S, WITH A SINGLE
FIVE-SECOND UTTERANCE AND YOU'LL

REALLY HEAR THE DIFFERENCES.
>> MAKE IT LOOK HARMLESS.
>> HOW MANY COULD EVEN
UNDERSTAND THAT?
>> BUT IF YOU GIVE AT FEW MORE
PIECES OF SAMPLE AUDIO, IT
BECOMES MUCH CLEARER, MUCH MORE
QUICKLY.
AGAIN, COMPARE THAT FIRST
GENERATED ONE, FIVE SAMPLES TO,
THE SINGLE ONE YOU CAN SEE
QUICKLY HOW THAT CURVE GOES UP.
THE OTHER THING TO POINT OUT IS
THAT ORGANIC SPEECH VERSUS
SYNTHETIC SPEECH, SYNTHETIC
SPEECH TEND TO BE LESS DYNAMIC.
HUMAN BEINGS WHEN WE TALK TO
EACH OTHER TEND TO HAVE
DISFLUENCIES, SO I MIGHT SAY AND
PAUSE FOR SEEMINGLY NO REASON, A
MACHINE DOESN'T KNOW TO DO THAT
-- YET -- OR UNLESS WE TELL IT.
HOW LONG WILL THAT LAST?
I'M NOT SURE IF WE'RE PLANNING
TO HAVE ANOTHER EVENT LIKE THIS
NEXT YEAR BUT WE'LL SEE.
THESE DYNAMIC ELEMENTS OF SPEECH
ARE REALLY CRITICAL AT THE
MOMENT FOR DETERMINING WHETHER
OR NOT YOU'RE SPEAKING TO AN
ORGANIC OR SYNTHETIC INDIVIDUAL,
HOW LONG THEY WILL LAST IS
ENTIRELY UNCLEAR.
OKAY.
SO JUST TO WRAP UP, NOW WE'RE
ALL AT THEML
WE ALL UNDERSTAND WHAT'S
HAPPENED.
HUMANS HAVE WANTED TO GIVE
VOICE, SHARE OUR GIFT, THE THING
THAT MAKES US UNIQUE, REALLY FOR
AS LONG AS WE FIGURED OUT THAT
IT'S A USEFUL THING THAT ONLY WE
HAVE.
TURNS OUT THAT NOW THAT WE'RE
GIVING IT TO COMPUTERS, THEY TOO

WILL BE ABLE TO COMMUNICATE AND
LIE JUST LIKE WE DO.
SO THIS WORKSHOP IS DESIGNED TO
DEAL WITH THAT VERY REALITY.
LISTEN, WE HAVE HAD PEOPLE WHO
CAN IMITATE EACH OTHER FOR AS
LONG AS WE HAVE HAD VOICE.
BUT WHAT THIS TECHNOLOGY DOES IS
IT ALLOWS US TO DO THIS AT
SCALE, ESSENTIALLY FOR ANY
INDIVIDUAL WITH INTEREST
SURPRISINGLY SMALL NUMBER OF
SAMPLES.
WHEN YOU WALK OUT OF THIS, I
WANT YOU TO UNDERSTAND THE
THINGS THAT MAKE THIS POSSIBLE
ARE DEEP LEARNING;GPUs AND
REALLY GOOD MODELS OF WHAT HUMAN
SPEECH IS.
WITH THAT, I WOULD LIKE TO TURN
THIS OVER TO THE CONVERSATION OF
THIS NEXUS BETWEEN TECHNOLOGY,
OF THE FUTURE OF WORK, AND
SECURITY, CAN KICK OFF IN
EARNEST.
THANK YOU VERY MUCH.
[APPLAUSE]
I CAN TAKE A QUESTION OR TWO.
OR NONE?
[INAUDIBLE]
THEY ARE INVOLVED IN EQUALLY
COMPLEX EXCHANGES USING AND
OFTEN WE SIMPLY DON'T KNOW WHAT
WE DON'T KNOW, RIGHT?
>> ABSOLUTELY.
>> SO THERE'S AN ACADEMIC PIECE
TO THIS BUT THERE'S ALSO A GROUP
OF US WAS STANDING AROUND BEFORE
THE EVENT CHATTING AND SOMEONE
SAID WELL, DOGS ACTUALLY CAN
SPOT THE DIFFERENCE.
I DON'T KNOW IF, SOMEONE, A DOG
CAN SPOT THE DIFFERENCE BETWEEN
A REAL VOICE AND I SAID
JOKINGLY, MAYBE WE NEED TO PUT A
DOG FILTER ON OUR EAR SO THAT WE

CAN IN FACT SPOT THE FAKES.
THIS IS NOT IDLE SPECULATION.
IT JUST TAKES US TO THE END OF
THIS AFTERNOON'S EXCHANGES.
>> I DON'T MEAN TO SAY THAT WE
BY OURSELVES OWN LANGUAGE IN
PARTICULAR.
THERE ARE LOTS OF DOCUMENTED
CASE WHERE'S SCIENTISTS WILL
RECORD AUDIO AND REPLAY IT AND
THE ANIMALS IN QUESTION WILL DO
THE EXACT BEHAVIOR AS IT IF HAD
BEEN LIVE TRANSCRIPTION.
FOOD OVER THERE, THAT'S NOT IN
QUESTION.
WE DON'T OWN ALL OF THE SPEECH.
BUT WHAT WE DO OWN AT LEAST WHAT
WE SEEM TO BE DOING BETTER IS
THAT WE ARE ABLE TO PUT THESE
THINGS TOGETHER TO AT LEAST AS
WE CAN UNDERSTAND IT DESCRIBE
THINGS THAT MAY NOT EXIST YET TO
DESCRIBE CONCEPTS THAT ARE
THEORETICAL IDEAS, BELIEF
SYSTEMS, AND THAT, YOU KNOW,
WHILE WHALES FOR EXAMPLE SING
VERY INTERESTING SONGS AND
LIKELY HAVE VERY UNIQUE PATTERNS
TO THEM, WE DON'T KNOW OF ANY
COMPLEX BELIEF SYSTEMS THAT HAVE
ARISEN FROM WHALE SONG, FOR
EXAMPLE.
SO I DON'T THINK WE OWN IT
ENTIRELY.
BUT I THINK THAT THE CURVE LOOKS
A LOT LIKE THIS IN TERMS OF WHO
BENEFITS THE MOST FROM LANGUAGE.
BEES WAGGLE, RIGHT?
DOGS BARK, ABSOLUTELY.
>> THANK YOU.
>> SURE.
I LOOK FORWARD TO MORE QUESTIONS
ON THE FINAL PANEL BUT I'LL TURN
THINGS BACK OVER.
THANK YOU.
[APPLAUSE]

THANK YOU, PROFESSOR TRAYNOR, FOR LAYING THE FOUNDATION FOR OUR WORKSHOP THIS AFTERNOON. YOU HAVE GIVEN US A LOT TO THINK ABOUT AS WE EMBARK ON OUR DISCUSSIONS.

THANK YOU, EVERYONE, IN THE AUDIENCE FOR JOINING US FOR OUR FIRST PANEL OF THE WORKSHOP. DURING THIS PANEL WE WILL BE EXPLORING THE PROS AND CONS OF VOICE CLONING TECHNOLOGY.

I HAVE A VERY ESTEEMED PANEL JOINING ME TODAY. EACH PANELIST REPRESENTS A UNIQUE PERSPECTIVE ON THE USES OF VOICE CLONING TECHNOLOGY.

TO MY IMMEDIATE LEFT IS JOHN COSTELLO, THE DIRECTOR OF BOSTON CHILDREN'S HOSPITAL'S AUGUST MEANTIVE COMMUNICATION PROGRAM.

NEXT TO JOHN IS REBECCA DAMON, EXECUTIVE VP OF SAG AFTRA.

TO HER LEFT IS RUPAL PATEL, CEO AND FOUNDER OF VOCALID.

AND LAST BUT NOT LEAST IS MONA SEDKY, PROSECUTOR WITH THE DEPARTMENT OF JUSTICE COMPUTER CRIME AND INTELLECTUAL PROPERTY SECTION.

EACH PANELIST'S BIO IS POSTED ON THE WORKSHOP EVENT PAGE. PLEASE REFER TO THOSE BIOS TO SEE A FULLSOME DESCRIPTION OF EACH PANELIST'S ACCOMPLISHMENTS.

TO START THE PANEL DISCUSSION, I HAVE ASKED EACH PANEL TOYS GIVE A FEW REMARKS ABOUT THEIR EXPERIENCE WITH VOICE CLONING TECHNOLOGY.

FOLLOWING THESE REMARKS, THE PANEL WILL EXPAND ON THE ISSUES RAISED THROUGH A QUESTION AND ANSWER DISCUSSION.

WITH THAT, I'LL TURN OVER THE MIC TO OUR FIRST SPEAKER, JOHN

COSTELLO.
>> THANK YOU.
I WOULD LIKE YOU TO IMAGINE FOR
A MOMENT THAT YOU HAVE BEEN
STRIPPED OF THE ABILITY TO
SPEAK.
AND THEN YOU HAVE BEEN STRIPPED
OF THE ABILITY TO HAVE
VOLITIONAL MOVEMENT.
AND THEN YOU HAVE BEEN STRIPPED
OF THE ABILITY TO USE FACIAL
EXPRESSIONS.
NOW I WANT YOU TO TELL A BELOVED
PERSON HOW DEEPLY YOU CARE FOR
THEM AND LOVE THEM.
I WANT YOU TO COMFORT A
DISTRAUGHT AND SELF-DOUBTING
CHILD.
I WANT YOU TO DELIVER A
SARCASTIC QUIP TO YOUR FRIENDS.
HOW ARE YOU GOING TO REPRESENT
THAT AUTHENTIC SELF?
IT'S THROUGH YOUR VOICE.
SO THE BENEFITS OF CUSTOM VOICE
TECHNOLOGY, FOR PEOPLE WHO HAVE
ALS, THERE ARE 5,000 NEWLY
DIAGNOSED PEOPLE IN THE UNITED
STATES AND APPROXIMATELY 25,000
PEOPLE LIVING WITH ALS.
80% OF THOSE HAVE BULBAR
SYMPTOMS MEANING THEY WILL LOSE
THE ABILITY TO SPEAK.
THIS ARE MYRIAD OF OTHER
DEGENERATIVE CONDITIONS FOR
CHILDREN AND ADULTS THAT WILL
RESULT IN LOSING THE ABILITY TO
SPEAK.
THERE ARE ALSO PEOPLE WHO,
BECAUSE OF SURGICAL PROCEDURES,
THEY WILL HAVE EITHER A
TEMPORARY OR PERMANENT LOSS OF
SPEECH.
SO WHAT IS THE POWER OF PERSONAL
VOICE?
FAMILIAR PARTNERS ONLY NEED TO

HEAR "IT'S ME" ON THE PHONE AND
IMMEDIATELY THERE IS A FLOOD OF
ATTITUDE AND UNDERSTANDING AND
EMOTION FOR WHO THAT PERSON IS
AND WHAT YOUR RELATIONSHIP IS
WITH THAT PERSON.
OUR VOICE IS OUR ACCUSAL
FINGERPRINT.
THE ROLE OF, AKUFTIAL
FINGERPRINT.
IT'S A MARKER OF OUR PERSONALITY
AND IMPORTANT NOT ONLY FOR THE
SPEAKER, BUT FOR THE
COMMUNICATION PARTNERS AND AS
HAS BEEN REFERENCED AND AS
YOU'LL SEE IN VIDEO, THOSE
PARTNERS HUMAN AND ANIMAL.
SO IN THE VIDEO THAT YOU'RE
GOING TO SEE HERE, THE FIRST
VIDEO IS A 2 1/2 MINUTE
COMBINATION OF VOICE
DEMONSTRATION, VOICE CREATION,
THAT HAS BEEN DONE FOR PEOPLE
WHO ARE AT RISK OR WHO HAVE NOW
LOST THEIR ABILITY TO SPEAK.
AND JUST TO GIVE YOU A LAY OF
THE LAND HERE, WHAT I HAVE TRIED
TO DO IS TAKE SOMEONE'S
AUTHENTIC VOICE, AUTHENTIC
MESSAGES THAT THEY HAVE BANKED
AND WHAT WE DO IN BOSTON AND
MANY OTHER DOES NOW, IS MEET
WITH PEOPLE BEFORE THEY LOSE
THEIR ABILITY TO SPEAK AND HAVE
THEM BANK MESSAGES THAT ARE
IMPORTANT TO THEM, THAT THEY
WANT TO MAKE SURE THAT IF EVER
THEY CANNOT SPEAK, THEY HAVE
READILY AVAILABLE TO BE EMBEDDED
IN TO TECHNOLOGY.
SO THEY CAN AUTHENTICALLY
REPRESENT TO THEIR LOVED ONES,
TO PEOPLE AROUND THEM, EXACTLY
WHAT IS THEIR EMOTION AND
EXACTLY WHAT'S ON THEIR MIND.
WE NOW, BECAUSE MANY PEOPLE ALSO

WANT TO MAKE SURE THAT THEY HAVE
AN AUTHENTIC-SOUNDING SYNTHETIC
VOICE, SO FOR THOSE THINGS THAT
THEY NEVER THOUGHT TO BANK THEY
WANT TO BE ABLE TO SPEAK THOSE
AND HAVE IT SOUND LIKE
THEMSELVES.
SO WHAT WE HAVE NOW DONE IS WHAT
I CALL DOUBLE DIPPING.
TAKING HUNDREDS AND MOST
TYPICALLY UPWARD OF A THOUSAND
BANKED MESSAGES AND WORKING WITH
COMPANIES AND THE EXAMPLES
YOU'RE GOING TO SEE HERE ARE
WITH THE COMPANY ACAPELLA OUT OF
FRANCE, HAVE CREATED A SYNTHETIC
VOICE THAT SOUNDS CLOSE TO THAT
PERSON'S BIOLOGICAL VOICE.
IN THIS VIDEO WHAT YOU'RE GOING
TO SEE IS FIRST, IF YOU READ THE
TEXT THAT GOES ALONG, YOU'LL SEE
THE AUTHENTIC VOICE AND THEN THE
ACAPELLA VOICE.
YOU'LL ALSO SEE ONE
DEMONSTRATION OF A VOICE THAT'S
BEEN CREATED BY GOOGLE FOR TIM
SHAW, SO THIS IS A PUBLIC VIDEO
THAT IS OUT OF A VOICE THAT THEY
HAVE CREATED FOR HIM.
AND MY PURPOSE OF INCLUDING ALL
OF THESE IS SO THAT IT'S VERY
CLEAR THE IMPORTANCE TO THE
PERSON AT RISK OF LOSING THE
ABILITY TO SPEAK AND THE PEOPLE
AROUND THEM TO HAVE AUTHENTIC
VOICE.
IF YOU COULD START THE VIDEO,
PLEASE.
>> THIS IS SYNTHESIS.
WOULD YOU LIKE TO SEE A SHOW?
>> WHERE ARE THE SONGS?
I WANT TO WISH YOU A HAPPY
BIRTHDAY
§ I LOVE YOU SO MUCH.
>> THIS IS TIM SHAW, SPECIAL
TEAMS ANIMAL.

CHRISTMAS COMING UP.
WE NEED TO FIND OUT WHAT'S
HAPPENING WITH TITANS PLAYER F
YOU HAVE TO HESITATE, THAT'S
PROBABLY A NO.
>> I DECIDED TO WRITE YOU THIS
LETTER.
I HAVE SO MUCH TO TELL YOU.
>> IT'S HIS VOICE.
I'VE FORGOTTEN.
>> I REMEMBER THAT ONE.
>> WE DO.
>> HI, MY NAME IS JOHN.
>> HI, MY NAME IS JOHN.
>> THE FIRST IS SYNTHETIC.
>> THIS IS MY ACAPELLA VOICE.
>> THIS IS MY ACAPELLA VOICE.
>> AND IT'S BASED ON A THOUSAND
RECORDINGS.
>> DO YOU KNOW HOW MUCH I LOVE
YOU?
THAT'S MY NORMAL SPEECH PATTERN.
HOW LONG WILL YOU BE GONE?
>> NOW LET'S HEAR THE ORIGINAL
MESSAGE.
>> HOW LONG WILL YOU BE GONE?
>> CAN I HAVE A GLASS OF WATER?
>> NOW THE SYNTHETIC.
>> CAN I HAVE A GLASS OF WATER?
>>

[INAUDIBLE]
>> IT WAS A COLD DAY IN APRIL
AND AND
[INDISCERNIBLE]
>> CAN I HAVE SOME MAYONNAISE ON
MY SANDWICH?
>> CAN I HAVE SOME MAYONNAISE ON
MY SANDWICH?
>> IT IS SO AMAZING THAT I WILL
BE ABLE TO HAVE MY VOICE IN BANK
MESSAGES HERE.
>> I'M REALLY TIRED.
>> AND IN THE INTEREST OF TIME.
>> IT IS A NICE DAY.
>> LET'S GO TO THE MOVIES.

LET'S PLAY.
>> AGAIN, THAT'S A COMBINATION.
SO IN THIS VIDEO WHAT I WOULD
LIKE YOU TO SEE AGAIN IS THE
POWER OF AUTHENTIC VOICE AND
HOLLY WAS RECORDING TO, SHE WAS
ACTUALLY PRACTICING TO MAKE SURE
THE RECORDINGS IN THIS DEVICE
WERE WORKING.
WE WERE UNAWARE THAT HER DOG WAS
IN THE ROOM.
PLEASE PLAY THE VIDEO.
>> WALKER, COME HERE, WALKER.
WALKER, COME ON.
WALKER, COME HERE, WALKER.
>> THERE WERE FIVE WOMEN IN THE
ROOM.
>> WALKER, COME ON, YOU WANT A
COOK SNI.
>> WALKER ONLY WENT TO HOLLY.
>> AWESOME!
>> AWESOME!
>> WHERE'S THAT SILLY DOG?
WALKER?
COME HERE, WALKER, YOU WANT A
COOKIE?
WHERE'S THAT SILLY DOG.
>> HOLLY WOULD BE VERY HAPPY.
HOLLY PASSED AWAY A FEW YEARS
AGO THAT YOU WERE ABLE TO
WITNESS THAT.
PRIVACY AND SAFEGUARDS,
COMPANIES CURRENTLY WHO ARE
CREATING THESE SYNTHETIC VOICES,
ACAPELLA FOR EXAMPLE, ARE
REQUIRING AUTHENTICATION AND
CREDENTIALING BEFORE VOICE IS
INSTALLED.
ONE OF THE THINGS THAT I WOULD
LIKE TO PROPOSE FRANKLY IS THAT
THERE IS FURTHER PROTECTION PUT
IN PLACE BY WHITE LISTING, MUCH
LIKE THE DOG WHISTLE YOU WERE
DESCRIBING, HAVING A NOISE IN
THE BACKGROUND OF VOICE THAT IS
PLACED ON ALL VOICE, ALL THE

CODES EXCEPT FOR THE ONE THAT IS
GIVEN TO THE AUTHENTIC PERSON.
SO YOU KNOW ALSO FROM PATIENTS
AND FAMILIES, THERE ARE REQUESTS
FROM PATIENTS NOT FOR THE ET
THEY'RIAL VOICE, WHEN SOMEONE
PASSED AWAY, BUT FOR SPECIFIC
MESSAGES.
AND SO OUR REQUESTS SO FAR IN
CLINIC HAVE BEEN PEOPLE AND
PATIENTS WANTING TO MAKE SURE
SPECIFIC MESSAGES ARE MADE
AVAILABLE.
YOU SHOULD ALSO UNDERSTAND
CURRENTLY IN SPEECH GENERATING
DEVICES TO BE FULLY FUNCTIONAL
TOOL THAT VOICE NEED TO BE
INTEGRATED IN TO THAT
COMMUNICATION DEVICE, SPEECH
API, AND IN ORDER FOR SOMEONE
WHO USE THAT AS AUTHENTICALLY ASWKkI
POSSIBLE, AND TAKE ADVANTAGE OF
THE POWER OF THEIR SOFTWARE,
THEY CANNOT BE RELYING ON
SOMETHING BEING IN THE CLOUD,
WHICH SOME OF THE HIGHEST
QUALITY VOICES RIGHT NOW ARE
CLOUD-BASED.
AND REALLY SADLY FOR THE END
USER IT'S NOT REAL TIME
COMMUNICATION.
EVEN SOMEONE WHO IS AN EFFECTIVE
COMMUNICATOR WITH A SPEECH
DEVICE, THEY'RE ENCODING THE
MESSAGE AND THEN SPEAKING IT OUT
SO THERE IS NOT AT THIS POINT
REAL TIME COMMUNICATION.
I THINK I'M OUT OF TIME.
BUT MY CONTACT IS HERE AS WELL.
[APPLAUSE]
REBECCA DAMON WILL BE OUR NEXT
SPEAKER.
>> HELLO THERE.
HI, EVERYBODY.
MY NAMES REBECCA DAMON AND I AM
THE EXECUTIVE VICE PRESIDENT OF

SAL-AFTRA, A LABOR UNION, SCREEN
ACTORS GUILD, AMERICAN
FEDERATION AND TELEVISION AND
RADIO ARTISTS.
SAG AFTRA FOUNDED, SCREEN ACTORS
GUILD IN 1933, AFTRA SHORTLY
THEREAFTER.
WE MERGED IN 2012 TO BECOME ONE
UNION.
WE ARE THE FACES AND VOICES THAT
ENTERTAIN AND INFORM AMERICA AND
THE WORLD.
SO OUR UNION IS 160,000 MEMBERS
ACROSS THE UNITED STATES.
WE ARE WHAT PEOPLE THINK OF US
TRADITIONALLY AS ACTORS, BUT WE
ARE ALSO BROADCASTERS, BOTH IN
THE NEWS AND INFORMATION SPACE,
BUT ALSO IN THE ENTERTAINMENT
SPACE.
WHEN YOU THINK OF A DISK JOCKEY
OR THINK OF A SPORTS CASTER,Q
THOSE PEOPLE ARE SAG AFTRA
MEMBERS.
WE'RE RECORDING ARTISTS,
SINGERS, DANCERS, STUNT PEOPLE
AND VOICE OVER ARTISTS.
WHEN PEOPLE THINK ABOUT VOICE
OVER ARTISTS, A LOT OF TIMES
THEY THINK ABOUT WHERE DO THEY
WORK?
HOW DO THEY WORK?
WHAT DOES THAT LOOK LIKE?
AND I THINK IT'S KIND OF
IMPORTANT WHEN YOU'RE THINKING
ABOUT THAT TO THINK ABOUT IT IN
MYRIAD DIFFERENT WAYS.
YOU MIGHT THINK OF IT IN TERMS
OF ANIMATION, HOW MANY OF YOU
GUYS HAVE WATCHED AN ANIMATED
MOVIE LATELY WITH EITHER, WITH
YOUR KIDS OR YOU CAN CONFESS YOU
WATCHED IT BY YOURSELF BECAUSE
THERE'S A LOT OF GREAT ANIMATION
OUT THERE THESE DAYS.
VIDEO GAMES, WHICH MANY OF OUR

PERFORMERS ARE KNOWN FOR.
COMMERCIALS, BOTH OUR HIGH
PROFILE MEMBERS AS WELL AS OUR
RANK AND FILE WORKING MEMBERS,
WORK IN ALL THESE DIFFERENT
SPACES.
PROMOS, FILM TRAILERS,
DOCUMENTARIES, AUDIO BOOKS, MANY
OTHER AREAS, RADIO.
SO OUR MEMBERS ARE THE PEOPLE
THAT WHEN YOU'RE LISTENING TO
SOMETHING THAT IS PROFESSIONALLY
RECORDED, THOSE PEOPLE ARE SAG
AFTRA MEMBERS.
SO YOU REALLY DO HEAR US
EVERYWHERE.
NOW, WHAT SAG-AFTRA DOES, WE'RE
THE COLLECTIVE BARGAINING AGENT
FOR OUR MEMBERS FOR WAGES,
WORKING CONDITIONS LIKE HEALTH
AND SAFETY PROTECTS, HEALTH AND
REQUIREMENT CONTRIBUTIONS,
RESIDUALS, ALL OF THOSE THINGS
FLOW FROM OUR COLLECTIVE
BARGAINING AGREEMENTS IN ALL
THOSE DIFFERENT SPACES.
AND WHY IT'S IMPORTANT TO THINK
ABOUT THAT IS WHO WE ARE AND
WHAT WE THINK ABOUT WHEN WE'RE
TALKING ABOUT VOICE IS WE
ACTUALLY ARE OUR VOICES.
THAT IS A QUANTITY IN TERMS OF
WHAT WE HAVE, IT IS ALSO OUR
ART.
SO IT'S SOMETHING THAT WHEN
WE'RE HAVING THESE CONVERSATIONS
IT'S ACTUALLY REALLY EXCITING TO
BE UP HERE AFTER JOHN AND RIGHT
AHEAD OF RUPAL BECAUSE OF
INTERESTING ETHICAL APPROACHES
TO HOW VOICE TECHNOLOGY CAN WORK
BUT A LOT OF WAYS WE THINK ABOUT
IT WHEN WE SAABING ABOUT WHAT
SAG-AFTRA IS, WHAT WE DO, WHAT
WE THINK SOCIETY AT LARGE NEEDS
TO BE THINKING ABOUT.

AT SAG-AFTRA, WE BELIEVE YOUR
IMAGE IS VALUABLE, YOUR VOICE IS
VALUABLE AND YOUR VOCAL
PERFORMANCE IS ACTUALLY
VALUABLE.
IT IS SOMETHING THAT OUR MEMBERS
ACTUALLY MONETIZE AND MAKE THEIR
LIVINGS USING.
SO BECAUSE OF THAT WE ARE THE
STRONGEST ADVOCATES OUT THERE
FOR STATE RIGHT OF PUBLICITY
LAWS THAT REQUIRE CONSENT TO THE
USE OF IMAGES AND PERFORMERS'
VOICES.
THAT IS SOMETHING THAT'S VERY
IMPORTANT AND IT'S VERY
IMPORTANT ALSO THAT THERE IS A
CONSENT PIECE TO IT, OF COURSE,
WHERE PEOPLE ACTUALLY GET TO
MAKE INFORMED DECISIONS ABOUT
HOW THEIR VOICES ARE GOING TO BE
USED.
IT IS ALSO IMPORTANT BECAUSE
THERE NEED TO BE COMPENSATION
THAT IS TIED TO THAT USE OF A
PERFORMER'S VOICE.
SO WHEN WE ARE THINKING ABOUT
IT, IT'S NOT THE KIND OF THING
WHERE WE WANT TO HAVE OUR
MEMBERS' VOICES, LET'S SAY WE
HAVE MADE A CHOICE FOR WHATEVER
REASON TO APPEAR IN A CERTAIN
KIND OF FILM OR CERTAIN KIND OF
TELEVISION.
IT'S NOT ACCEPTABLE FOR OUR
PERFORMERS TO BE USED IN A WAY
THAT WOULD BE INCONSISTENT WITH
THEIR BELIEFS, THAT MIGHT BE
INCONSISTENT WITH OTHER
AGREEMENTS THAT THEY HAVE.
SO AS PERFORMERS WE OFTEN HAVE
AN EXCLUSIVITY ARRANGEMENT WITH
A CERTAIN PRODUCT, SERVICE.
SO IT'S VERY IMPORTANT AS WE'RE
THINKING ABOUT ALL THIS THAT
THERE ARE RIGHTS THAT ATTACH TO

THAT, BOTH THROUGH COLLECTIVE
BARGAINING AGREEMENTS AS WELL AS
LAWS ON A STATE TO STATE BASIS.
SO SOME OF THE THINGS THAT WE
ADVOCATE FOR IS AN ETHICAL
APPLICATION OF THESE VOICE
CLONING TECHNOLOGIES.
IT'S EXCITING.
WE ALWAYS KIND OF WHEN WE'RE
TALKING ABOUT IT IN OUR BOARD
ROOM AND WITH OUR MEMBERS, THINK
ABOUT ALL THE WONDERFUL USES.
I WAS SO COMPELLED AS JOHN WAS
TALKING ABOUT THAT INCREDIBLE
ABILITY TO HAVE SOMEBODY THAT
WILL NO LONGER HAVE A VOICE TO,
ACTUALLY HAVE A VOICE.
THAT IS EXCITING AND WONDERFUL
AND TREMENDOUS.
I'M ALSO EXCITED WHEN I HEAR
ABOUT DEFENSES THAT YOU'RE
SUGGESTING GO UP BECAUSE A LOT
OF TIMES PEOPLE GET EXCITED AND
RUSH IN WITH THE NEW TECHNOLOGY
AND THEN THEY DON'T NECESSARILY
THINK THROUGH ALL THE
APPLICATIONS, SO I ALSO WANT TO
CONGRATULATE CHRISTINE AND THE
FTC FOR PUTTING THIS
CONVERSATION TOGETHER.
THERE ARE NOT ENOUGH OF THESE
CONVERSATIONS.
I KNOW, BECAUSE I FEEL LIKE I GO
TO EVERY SINGLE ONE OF THEM ON
IT.
WE ALSO THINK ABOUT THIS FROM A
PERSONAL AND NATIONAL SECURITY
POINT OF VIEW, AS WELL AS
PRESERVING TRUST IN BROADCAST
JOURNALISM.
THIS IS INCREDIBLY IMPORTANT TO
US BECAUSE AS I MENTIONED OUR
BROADCASTERS ARE THOSE FACES AND
VOICES THAT DO INFORM AMERICA.
SO IMAGINE YOU'RE ON THE EVE OF
BREAKING A HUGE STORY, I DON'T

KNOW, ANYTHING COULD BE GOING
ON, HUGE STORY, YOU'RE ABOUT TO
BREAK THAT STORY AND YOU FIND AN
AUDIO OF YOUR VOICE SAYING
SOMETHING OR A VISUAL DOING
SOMETHING YOU DIDN'T DO.
AND IT TOTALLY RUINS YOUR
CREDIBILITY.
SO WHEN WE ARE THINKING ABOUT IN
TERMS OF OF THEIR PERSONAL
BRAND, IN TERMS OF THE TRUST
THAT THEY MIGHT HAVE SPENT A
LIFETIME IN BUILDING, IT DOES
HAVE AN IMPACT IN WHAT IS
BELIEVED TODAY.
SO AS WE LOOK THROUGH THE
IMPLICATIONS FOR THIS KIND OF
TECHNOLOGY, WE REALLY BELIEVE IT
HAS TO BE DONE IN A WAY WITH AS
MANY SAFEGUARD AS POSSIBLE.
SO THAT IS WHY WE'RE EXCITED TO
BE HERE TODAY.
LOOKING FORWARD TO YOUR
QUESTIONS ONCE WE GET TO THE
PANEL.
[APPLAUSE]
>> OUR NEXT SPEAKER WILL BE
RUPAL PATEL.
>> THANK YOU.
MY NAME IS RUPAL PATEL.
TIM I AM THE FOUNDER OF VOCALID,
CREATE SYNTHETIC VOICES THAT
SOUND AS REALISTIC AS POSSIBLE.
I'M A PROFESSOR AT NORTHEASTERN,
SPEECH LANGUAGE PATHOLOGY AND
AUDIOLOGY AS WELL AS THE COLLEGE
OF COMPUTER SCIENCES.
THE RESEARCH IN MY LAB BACK FROM
2007 THROUGH 2014 WAS ON TRYING
TO CREATE UNIQUE SYNTHETIC
VOICES FOR INDIVIDUALS WHO WERE
UNABLE TO SPEAK, MUCH LIKE WHAT
JOHN WAS TALKING ABOUT THROUGH
VOICE BANKING RATHER THAN
MESSAGE BANKING.
WHAT WE WERE INTERESTED IN

ACTUALLY DOING, THOUGH, WAS
CREATE A VOICE FOR SOMEONE DID
NOT HAVE THE ABILITY TO SPEAK
FROM BIRTHMENT WE TOOK A
CROWDSOURCE APPROACH.
WE HAD PEOPLE LIKE YOU AND I
CONTRIBUTE THEIR VOICE TO OUR
VOICE BANK AND THEN WHAT WE
WOULD DO IS TAKE SAMPLES OF THAT
PERSON WHO COULDN'T SPEAK, BUT
CAN STILL VOCALIZE.
THOSE VOCAL STATIONS WOULD THEN
BE USED TO LOOK FOR A MATCH
WITHIN OUR VOICE BANK OF SOMEONE
SIMILAR, AND THEN WHAT WE WOULD
DO IS THROUGH THESE TECHNOLOGIES
BASICALLY INVERSE FILTER AND
REMAP OR MORPH A VOICE SO IT
COULD SOUND LIKE IT COULD
POTENTIALLY COME FROM THEM.
STEVEN HAWKING WOULD HAVE BEEN A
GREAT CANDIDATE IF HE THOUGHT
HIS DESKTOP VOICE WAS HIS ONLY
VOICE.
BUT THERE ARE OTHER PEOPLE WHO
BENEFITTED FROM THE TECHNOLOGY.
THAT'S HOW VOCALID STARTED.
2015-2018 FOCUSED ON INDIVIDUALS
NON-SPEAKING, FROM BIRTH AND
INDIVIDUAL WHO WERE ABOUT TO
LOSE THEIR VOICE.
WE DON'T FOCUS ON MUCH ON ALS
BUT PEOPLE WITH HEAD AND NECK
CANCER.
PRIOR TO THEM LOSING THEIR VOICE
TO OR BECAUSE EITHER THEIR VOICE
BOX WILL BE REMOVED OR PART OF
THEIR TONGUE WILL BE REMOVED
THEY BANK THEIR VOICE AHEAD OF
TIME AND REMEMBER W. HE RECREATE
IT.
WE'RE ON GENERATION FOUR OF OUR
VOICE SYNTHESIS ENGINE, ALLOWS
US TO MAKE THE MOST HIGHLY

REPFULLY REPLICATABLE VOICE.

THESE ARE DEEP-LEARNING
METHODOLOGIES BUT THOSE ARE THE
ONLY ONES THAT WORK ON ASSISTIVE
TECHNOLOGY DEVICES LIKE JOHN WAS
MENTIONING EARLIER.
THOSE ARE ALL THE GOODS, RIGHT?
THERE'S ALWAYS A PUSH, BECAUSE
NOW THAT WE'RE ALSO MAKING
VOICES FOR COMPANIES AND
ORGANIZATIONS, SO A COMPANY
OFTENTIMES WILL HAVE A BRAND
VOICE, LIKE REBECCA WAS TALKING
ABOUT.
AND THAT VOICE TYPICALLY HAS
BEEN AN INDIVIDUAL.
BUT MANY TIMES TOO THAT
INDIVIDUAL MIGHT BE RECORDING
THEIR PHONE TREE FOR THEIR IBR
SYSTEM THAT.
INDIVIDUAL MAY BE DOING
E-LEARNING SCRIPTS FOR THEIR
CORPORATE TRAINING VIDEOS, SO
ON, SO FORTH.
SO THIS IS ALL, THE PERSON WHO
KNOWS THEY'RE HAVING THEIR VOICE
SYNTHESIZED KNOWS THIS IS A WAY
TO INCREASE THEIR CAPABILITIES
IN TERMS OF BEING ABLE TO SERVE
MORE THAN ONE CLIENT AND THEY
COULD GET RESIDUALS BACK AND SO
ON.
I THINK THERE ARE WAYS IN WHICH
THIS TECHNOLOGY CAN BENEFIT THE
VOICE ACTORS THEMSELVES, THAT
THEY DON'T ALWAYS HAVE TO COME
BACK AND DO THOSE PICKUPS AND
SMALL THINGS THAT CAUSE THEM
PAIN AND ANNOYANCES, THEY CAN DO
MORE CREATIVE THINGS.
BUT IT IS A WORKFORCE ISSUE AS
WELL IN TERMS OF DOES IT MEAN IT
WILL DISPLACE SOME?
YES.
ANY NEW TECHNOLOGY WILL HAVE A
DISPLACEMENT FACTOR AS WELL.
THEN THERE'S THIS UNCANNY VALLEY

PART OF IT WHERE ANY TIME YOU
MAKE A TECHNOLOGY THE, EVERYONE
WANTS TO PUSH TO THE EXTREME OF
WHEN IT STARTS SOUNDING REAL,
CAN YOU MAKE IT SOUND MORE REAL?
MAKE IT WITH MORE MELODY,
EMPHASIS?
EVERYONE WANTS THAT.
THINK ABOUT COMPUTER GRAPHICS.
WE'RE NOT OKAY WITH IT JUST
LOOKING LIKE AN ANIMATION.
WE WANT IT TO BE REALISTIC, THAT
VIDEO GAME TO FEEL LIKE YOU'RE
IN IT, RIGHT?
WELL, WE'RE PUSHING THAT
TECHNOLOGY BUT WHEN WE PUSH THAT
TECHNOLOGY WE'RE ALSO PUSHING
THE ANTE ON WHAT CAN HAPPEN IN
THE REVERSE.
I THINK THAT ONE OF THE
INTERESTING THINGS IS THAT WE
ARE BY VIRTUE OF THIS TECHNOLOGY
ASKING FOR MORE IN THE
TECHNOLOGY ITSELF.
BUT THE GOOD PART OF IT IS THE
POTENTIALLY CAN INCLUDE MORE
DIVERSE POOL OF INDIVIDUALS THAN
WE HAVE SEEN BEFORE.
MIGHT INCLUDE THAT VOICES THAT
ARE BEING HEARD ON THESE NEW
TYPES OF MEDIA, MAY HAVE MORE
DIVERSITY AND ACCENT WHICH WE
DON'T HEAR ON THE RADIO, DON'T
SEE ON TV THAT MUCH.
WE'RE STARTING TO SEE MORE BUT
MIGHT SEE IN THIS EMBEDDED
WORLD.
THERE IS A BENEFIT THAT WAY.
ON THE OTHER END, NEFARIOUS
USES, WITH ANY NEW TECHNOLOGY,
THERE WILL ALWAYS BE NEFARIOUS
USES.
SOME TECHNOLOGIES RACED AHEAD
AND NOT THOUGHT ABOUT NEGATIVE
CONSEQUENCES.
BECAUSE VIDEO CAME FIRST.

WITH AUDIO THERE ARE COMPANIES, INCLUDING OURS, THAT ARE STARTING TO THINK ABOUT WHAT'S OUR ROLE AS THE ORIGINATOR OF THE TECHNOLOGY TO ALSO PUT IN THE SAFEGUARD?

YOU KNOW, IF YOU KNOW HOW THE TECHNOLOGY WORKS, YOU'RE ALSO BEST ABLE TO KNOW HOW TO DECODE IT OR HOW TO FIGURE OUT A WAY TO PUT IN SOMETHING SO THAT IT CANNOT BE MISUSED.

I THINK THAT REGULATION, SOME OF IT WILL COME FROM US AS WELL.

THE OTHER I THINK BIGGEST SAFEGUARD IS AWARENESS.

AWARENESS, PEOPLE DON'T EVEN KNOW THAT THIS TECHNOLOGY EXISTS.

PEOPLE DON'T KNOW HOW GOOD IT CAN BE.

I THINK WE NEED TO START EDUCATING PEOPLE ABOUT THAT AND UNDERSTANDING THIS TECHNOLOGY AND ITS MECHANISMS, HOW CAN IT SPREAD IS AN IMPORTANT PIECE OF PUBLIC AWARENESS AS WELL.

IN ADDITION TO THAT, I CAN'T SPEAK FOR THE -- I DO THINK THAT IT IS IMPORTANT TO THINK ABOUT WHAT ALREADY EXISTS IN TERMS OF COMMON LAW TOO, TO PROTECT US.

IT'S NOT ONLY LIKE YOU MAKE A NEW LAW FOR THIS NEW TYPE OF TECHNOLOGY WILL BE VERY NARROW.

WE HAVE TRIED TO DO THAT SOME OF THE PRIVACY ISSUES.

WE HAVE TO THINK ABOUT SOMETHING MORE GENERALLY IN TERMS OF THERE'S WAYS TO PROTECT AGAINST IMPERSONATION FOR EXAMPLE.

SO HOW DO WE TAKE THOSE AND APPLY THAT TO THIS NEW TECHNOLOGY?

MUCH OF THE THINGS WE'RE TALKING ABOUT HERE ARE NOT NEW, RIGHT?

IMPERSONTATION HAS BEEN THERE
FOR AWHILE.
NOW IT'S MAKING IT EASIER AS
PATRICK SAID, ALLOWING US TO
SCALE MORE.
THE OTHER THING I WOULD SAY IS
THERE'S SORT OF AN AGREEMENT
AMONG THOSE OF US CREATING NEW
SYNTHETIC MEDIAS THAT OUR GOAL
IS FOR THIS TO REALLY BE A
TECHNOLOGY THAT BRINGS PEOPLE
TOGETHER RATHER THAN BEING
DIVISIVE.
WE¦ KNOW THAT THERE WILL BE BAD
ACTORS.
HOW DO WE MAKE SURE THOSE BAD
ACTORS DO NOT PREVAIL OVER THAT?
THAT'S BECAUSE WE'LL PUT IN
THESE SAFEGUARD ALONG THE WAY,
ALONG WITH ORGANIZATIONS LIKE
WORKING WITH THE ACTORS UNION TO
MAKE SURE THERE IS FAIR
COMPENSATION AND SO ON.
BUT THIS IS A GROUP EFFORT.
IT CAN'T JUST BE ANY ONE SEGMENT
THAT TAKE IT ON.
I THINK WHAT IT WILL HAPPEN, IT
WILL BE BACKDOORS, WILL BE
BAND-AIDS RATHER THAN ACTUAL
FIXES OF THE ENTIRE ECOSYSTEM.
THANK YOU AND I LOOK FORWARD TO
YOUR QUESTIONS.
[APPLAUSE]
>> MONA SEDKY.
>> THANK YOU FOR INVITING ME
HERE.
SO I AM THE VOICE OF DOOM.
I DON'T WANT TO -- I'M AN IN THE
TRENCHES FEDERAL PROSECUTOR IN
THE COMPUTER CRIME AND
INTELLECTUAL PROPERTY SECTION AT
DOJ.
I DEAL WITH CYBERCRIMINALS ALL
DAY LONG AND HAVE FOR ABOUT 15
YEARS.
MY SPECIALTY AT DOJ ARE SORT OF

TWO DIFFERENT KINDS OF CYBER
CRIME, BOTH OF WHICH ARE FERTILE
GROUNDS FOR PEOPLE ENGAGED IN
DEEP FAKE AUDIO FOR LACK OF A
BETTER WORD.
I DO A LOT OF CASES INVOLVING
PRETTY PERNICIOUS CYBER
STALKING, NON-CONSENSUAL
PORNOGRAPHY, EXTORTION, SEX
TORSION AND THINGS LIKE THAT.
I DO A FAIR AMOUNT OF ONLINE
FRAUD THAT INVOLVES SOCIAL
ENGINEERING.
GENERALLY SPEAKING, THE CRIMINAL
BUCKETS THAT I DEAL WITH ARE
PEOPLE, THERE ARE CRIMES THAT
ARE DEPENDENT UPON COMMUNICATION
AND THEN THERE ARE CRIMES THAT
INVOLVE ALMOST NO COMMUNICATION.
AND COMMUNICATION DEPENDENT
CRIMES ARE YOU HAVE THE FOR
CRIMINALS BECAUSE FIRST OF ALL,
IF CRIMINALS DON'T WANT TO
COMMUNICATE WITH THE PUBLIC,
THEY DON'T WANT TO COMMUNICATE
WITH LAW ENFORCEMENT, THEY DON'T FOR CRIMINALS.
FIRST OF ALL, IF CRIMINALS DON'T
WANT TO COMMUNICATE WITH THE
PUBLIC, THEY DON'T WANT TO
COMMUNICATE WITH LAW
ENFORCEMENT, THEY DON'T WANT TO
COMMUNICATE WITH VICTIMS.
THEY LEAVE FINGERPRINTS ALL THE
TIME.
IT'S BAD.
THERE ARE CRIMES YOU HAVE TO
COMMUNICATE.
IF YOU'RE THREATENING OR HE CAN
TOGETHERRING, A HOEBG THREAT OR
A REAL THREAT, P YOU'RE TRYING
TO DEFRAUD SOMEONE YOU HAVE TO
COMMUNICATE WITH THEM AND YOU
LEAVE FINGERPRINTS EVERY TIME
YOU DO.
THEY DON'T LIKE TO COMMUNICATE
BECAUSE IT'S DANGEROUS FOR THEM.

THEY ALSO DON'T WANT TO
COMMUNICATE BECAUSE IT'S COSTLY.
THOSE FRAUD SCHEMES TAKE A HOT
OF TIME.
TIME IS MONEY FOR CRIMINALS.
THEY DON'T WANT TO SPEND A BUNCH
OF TIME ON THE PHONE LOOKING TO
STEAL YOUR MONEY.
HE DON'T WANT TO COMMUNICATE
WITH YOU, IT COST THEM TILE AND
MONEY.
THIRDLY, THEY'RE BAD AT IT.
IT'S DIFFICULT TO CONVINCINGLY
IMPOSE AS SOMEONE ELSE.
ESPECIALLY IF YOU'RE A FOREIGN
NATIONAL AND HAVE AN ACCENT.
IT'S VERY DIFFICULT TO KHAOUPB
KATE WITH A CRIME VIC TUPL AND
CON SRUPBS THEM YOU ARE SOMEONE
ELSE.
FOR THESE THOUGH REASONS REALLY
COMMUNICATION FOCUSED CRIME IS,
HAS HISTORICALLY BEEN LOWER ON
THE TOTUM POLL.
NOW WITH DEEP FAKE AUDIO, AMNON
AMNONMIZIG TOOLS.
YOU CAN COMMUNICATE ANONYMOUSLY
WITH PEOPLE ANYWHERE FROM THE
WORLD.
THERE IS A UPTICK ON
COMMUNICATION FOCUSED CRIME.
TO BE THE VOICE OF DOOM.
I'M A IN-THE-TRENCHES
PROSECUTOR.
I DON'T DO THE SOCIETY COST
BENEFIT A MAL SIS.
IT'S FANTASTIC TO HAVE THIS.
IT'S EASY TO BE IN THE DARK
SPACE I WORK TO ENVISION ALMOST
GUARANTEED CRIMINAL USE OF THIS
TECHNOLOGY GO.
LET ME GIVE YOU A FEW EXAMPLES.
THERE WILL BE THREATS AND
HARASSMENT FOCUSED CRIMINALS
THAT WILL LOVE THIS TECHNOLOGY
GO AND FRAUD BASED CRIMINALS

THAT WILL LOVE THIS TECHNOLOGY
GO.
FOR THE THREATS AND HARASSMENT
FOCUSED CRIMINALS HERE ARE A
COUPLE SCENARIOS.
I PROSECUTE CYBER STALKERS.
THEY POSE AS THEIR VICTIMS AND
PUT UP BASICALLY RAPE FANTASY
ADVERTISEMENTS OR COME HITHER,
COME RAPE ME, BDS SOLICITATIONS
ON BE HALF OF YOUR VICTIM.
NOW YOU POST HER PHOTOGRAPH,
VIDEO, ADDRESS AND WORK PLACE.
YOU KNOWS AS HER AND CALL PEOPLE
TO ESSENTIALLY IT'S RAPE BY
PROXY.
COMMON PATTERN.
IF YOU HAVE VIDEO.
IF YOU CAN ALSO MIMIC HER VOICE
TO CONVINCINGLY INVITE PEOPLE TO
HER HOME FOR THIS GANG BANG,
EXCUSE MY FRENCH OR WHATEVER
IT'S NOT FRENCH, IT'S A ENHANCED
TOOL FOR CYBER STALKERS.
I ALSO PROSECUTE WHAT WE CALL
SEXTORTION.
IT'S ANOTHER FORM OF CYBER
STALKING.
THE OFFENDER TRIES TO CREATE
FAKE EXPLICIT SEXUAL IMAGES OR
STAET SEXUAL EXPLICIT IMAGES
THAT LOOK LIKING THE VICTIM.
THEY THEN COMMUNICATE WITH THE
VICTIM AND SAY IF YOU DON'T SEND
ME NAKED PHOTOS EVERYONE IN THE
WORLD WILL SEE THIS IMAGE.
DEEP FAKE AUDIO COULD BE USED IN
SORT OF TANDEM WITH DEEP FAKE
VIDEO TO CREATE A REALISTIC
SEXUALLY EXPLICIT VIDEO TO BE
USED BACK TO THE VICTIM AND SAY
IF YOU DON'T GIVE ME REAL
CONTENT BY MIDNIGHT TONIGHT THIS
DEEP FAKE IS GOING ALL OVER THE
INTERNET.
CONVERSELY IMAGINE IF YOU COULD

POSE AS YOUR VICTIM'SENT MATE PARTNER.
YOU COULD CALL YOUR VICTIM AND EASILY SPOOF THE TELEPHONE NUMBER OR SPOOF THE E-MAIL ADDRESS AND CALL AS A FOLLOW-UP AND SAY MAYS SEND ME SEXUALLY EXPLICIT IMAGES OF YOU, THEY WOULD USE DIFFERENT LANGUAGE FROM WHAT I SAID JUST.
THEY WOULD CONVINCE THEIR VICTIM TO TURN OVER SEXUALLY EXPLICIT IMAGES, OTHER PERSONAL A DEPARTMENT FIABLE INFORMATION, PERHAPS PASSWORDS TO THEIR ACCOUNTS.
THAT WOULD BE USED TO FURTHER EXPLOIT THE VICTIM.
WHAT THEY'RE HOPING IS THAT THERE WILL BE SOME KIND OF CON PROPB STATION WITH LAW ENFORCEMENT AND THE VICTIM.
IF YOU HAVE A CONVINCING DEEP FAKE AUDIO YOU CAN EASILY SWAT YOUR VICTIM IN A CONVINCING WAY.
PERHAPS THEY KNOW, PERHAPS IT'S A SMALL TOWN AND THEY KNOW YOUR VOICE.
IT CAN BE WEAPONIZED IN THE CYBER WORLD AGAINST VICTIMS.
THE BEAUTY OF SOMETHING LIKE DEEP FAKE AUDIO IS IT'S EMINENTLY SCALABLE.
BACK TO THE COST OF ONE-ON-ONE COMMUNICATION WITH SOMEONE.
YOU CAN CLICK A BUTTON AND DO A MASSIVE AMOUNT OF IMAGE.
THE FRAUD CONTENT SUP.
I WILL GIVE YOU THE TEN SECOND -- THE FRAUD SCENARIOS ARE PRETTY OBVIOUS.
THESE ARE YOUR GRANDPARENT SCAMS.
WE CALL THEM BUSINESS E-MAIL COMPROMISED.
BECs PROBABLY THE SINGLE

LARGEST FINANCIAL INJURY OF ALL
COMPUTER CRIME RIGHT NOW.
CLASSIC SCENARIO YOU'RE A
UNIVERSITY.
I SEND YOU A FAKE PHISHING
E-MAIL PRETENDING TO BE THE
CONSTRUCTION COMPANY.
I SAY THE BANK ACCOUNT HAS
CHANGED PLEASE WIRE FUNDS TO THE
NEW BANK.
IF I CAN FOLLOW-UP A FAKE SPEAR
TPHISHING E-MAIL WITH A PHONE
CALL IT'S HUGE DOLLAR LOSSES
INSTANTLY.
VOICE ENHANCED BECs
COMPROMISES WILL BE VERY VERY
FRUIT FULL.
FUL.
SO MY VOICE OF DOOM HERE IS,
THIS TECHNOLOGY IS FANTASTIC.
JUST LIKE THE INTERNET CAN BE
WEAPONIZED AGAINST PEOPLE.
DOESN'T MEAN WE SHOULDN'T HAVE
THE INTERNET BUT THESE ARE
THINGS TO THINK ABOUT.
EXACTLY WHAT RUPAL WAS SAYING
THERE ARE THINGS TO BAKE IN ON
THE FRONT END TO MAKE IT HARDER
TO WEAPONIZE AGAINST PEOPLE.
THANK YOU AGAIN FOR INVITING ME.
>> THANK YOU, VERY MUCH.
I WANT TO ADDRESS THE FIRST
QUESTION TO RUPAL.
WE HAVE TALKED ABOUT A COUPLE OF
DIFFERENT TYPES OF TECHNOLOGY.
MESSAGE BANKING.
.SYNTHETIC VOICE AND VOICE
CLONING.
CAN YOU GIVE US A BRIEF OVER
VIEW OF EACH TECHNOLOGY.
>> SURE MESSAGE BANKING IS WHEN
YOU PRESTORE MESSAGES LIKE VOICE
MEMOS ON YOUR PHONE OR WHATEVER.
YOU CAN REPLAY THOSE MESSAGES.
WHAT JOHN IS TALKING ABOUT IS A
WAY FOR PEOPLE WITH DISABILITIES

TO ACTUALLY HAVE THESE IN ANARCH
ANARRAY TO FIND THEM AND MAY THEM
BACK.
SYNTHETIC VOICE IS WHEN YOU HAVE
A BUNCH OF RECORDINGS AND USE
THOSE TO MACHINE WISE RECREATE A
VOICE.
THERE ARE MANY TECHNIQUES OF
SYNTHETIC SPEECH INCLUDING WHAT
WAS SAID EARLIER IN TERMS OF
TRYING TO CREATE VOICES WHERE
YOU GLUE TOGETHER SEGMENTS OF
SPEECH.
THE NEWEST METHOD IS USING
MACHINE LEARNING TECHNIQUES.
AI IS THE MOST ADVANCE OF
SYNTHETIC VOICE.
AI IS A METHOD USED, DEEP
LEARNING IS USED TO CREATE THESE
VOICES.
IT'S A SYNTHETIC VOICE AT THE
END OF THE DAY.
WE CALL IT AI IT'S USED WITH
DEEP LEARNING RATHER THAN
GLUEING TOGETHER PIECES OF
SPEECH.
>> ARE THERE PARTICULAR BENEFITS
OF CLONED VOICES VERSUS
SYNTHETIC VOICES AND MESSAGE
BANKING?
>> VOICE CLONING.
SO THINK ABOUT SIRI RIGHT.
A VOICE ACTOR, SUSAN BENNETT,
THE VOICE BEHIND SIRI.
SHE RECORDED FOR HOURS AND
HOURS.
THOSE WERE CUT INTO SNIPPETS OF
SPEECH AND COMBINATIONS OF
HOSES.
IT ALLOWS FOR US TO TYPE
ANYTHING IN AND SAY THINGS SHE
DIDN'T NECESSARILY SAY.
SYNTHETIC VOICE, THE POINT IS
IT'S GENERATED.
YOU CAN TYPE IN ANY NOVEL PHRASE
AND CREATE, GENERATE AUDIO FROM

IT.
RIGHT.
SO THE PROBLEM WITH OLDER
METHODOLOGIES IS IT'S SO
EXPENSIVE TO BUILD AND YOU NEED
SO MUCH TIME OF THE VOICE ACTOR
OR SOMEONE IN CLEAN STUDIO
RECORDINGS TO CREATE THAT KIND
OF A VOICE.
IT'S VERY EXPENSIVE.
BANK OF AMERICA MADE A VOICE
RECENTLY, MILLIONS OF THRARZ.
YOU KNOW IT'S OBVIOUSLY MADE
WITH LOTS AND LOTS OF TIME.
THE NEWER METHODOLOGIES THAT PAT
YOU CAN TRAYNOR DESCRIBED YOU
DON'T NEED THIS WITH THE SEN
THET I CAN METHOD OL TPWAOES YOU
TAKE A ENTIRE RECORDING BASE
ABOUT A THOUSAND SENTENCES.
ABOUT AN HOUR OF SPEECH.
YOU RECREATE A VOICE.
I GUESS YOU CALL THAT VOICE
CLONING.
WE CALL IT A SYNTHETIC VOICE
BECAUSE YOU WANT TO SOUND LIKE
THAT PERSON.
NOW YOU CAN USE THAT FOR A
MULTITUDE OF DIFFERENT OPTIONS.
THE POINT IS YOU CAN TYPE IN
THINGS AND GENERATE THE AUDIO
WHICH WASN'T SAID BEFORE.
WASN'T PRERECORDED.
>> OKAY.
JOHN, DO YOU HAVE ANYTHING TO
ADD ON THAT FROM THE PERSPECTIVE
OF THE PATIENTS THAT YOU WORK
WITH?
>> I THINK TO YOUR QUESTIONS
ABOUT THE VALUE OF THE MESSAGE
BANKING AS YOU SAW IN THE VIDEO,
THE POWER OF HAVING THAT
AUTHENTIC REPRESENTATION OF
EXACTLY WHAT YOU WANT TO SAY AND
HOW YOU WANT TO SAY IT IS
IMPORTANT.

FROM THE VOICE BANKING SIDE THE
PEOPLE WHO YOU HEARD AND SOME OF
WHO YOU MET IN VIDEO RECORDED UP
TO A THOUSAND MESSAGES.
THIS TOOK A BIT OF TIME.
THEY DID THAT FUNCTIONALLY
OVERTIME.
THEN THAT BEING INTEGRATED WITH
OUR SITE AND WITH ACAPELLO IT
WAS HELPFUL FOR MANY PEOPLE.
NOW PEOPLE LOSING THE ABILITY TO
SPEAK HAVING THAT CUSTOM
SYNTHETIC VOICE IT'S ATTAINABLE
FOR SOMEONE SO VULNERABLE.
>> HOW WIDE SPREAD WOULD YOU SAY
THE TECHNOLOGY IS BEING USED NOW
BY PEOPLE WHO HAVE LOST THE
ABILITY TO SPEAK?
>> I WOULD SAY THAT THE IDEA OF
VOICE BANKING HAS REALLY COME IN
TO IT'S OWN EVEN IN THE PAST
COUPLE OF YEARS AS SOME OF THE
TECHNOLOGY HAS BECOME BETTER AND
BETTER.
CERTAINLY IN THE WORLD OF ALS
THERE IS A GREAT DEAL OF
INTEREST IN AWARENESS OF THIS
NATIONALLY AND INTERNATIONALLY.
WHERE AS IF YOU SPOKE WITH
SOMEONE FROM A CLINIC THREE OR
FOUR YEARS AGO THEY WOULD
SUGGEST IT WAS THE A GOOD OPTION
FOR MOST PEOPLE BECAUSE THEY
DIDN'T UNDERSTAND HOW EASY OR
AFFORDABLE IT WAS.
I THINK NOW IT'S STANDARDLY
RECOGNIZED, CERTAINLY IN THE
DISABILITY WORLD AS AN OPTION TO
CONSIDER.
>> AND REBECCA FROM YOUR
PERSPECTIVE HAVE YOU SEEN
INSTANCES WHERE THIS IS USED
WITHOUT CONSENT?
>> WE ACTUALLY HAVE.
A LOT OF TIMES IT'S ALSO
COMBINED IN VIDEO.

THERE WAS A PARTICULAR PERSON
WHO WAS INSERTED INTO A VIDEO, A
BROADCASTER WE WERE TALKING
ABOUT THIS EARLIER.
THEY MADE THEIR LIVING AS A
SPORTSCASTER AT A PARTICULAR
KIND OF EVENT DOING THE SAME
KIND OF THING.
THEY WERE INSERTED INTO A VIDEO
GAME WITHOUT THEIR KNOWLEDGE AND
WITHOUT THEIR CONSENT.
THAT WAS A WAY THAT, THAT WAS
THAT PERSON'S LIVELIHOOD AND
CAREER, THEIR WHOLE TRAINING,
THEIR WHOLE ABILITY -- IT WAS
THEIR BRANDIE SENSUALLY.
IT WAS INSERTED.
THOSE ARE THE KINDS OF THINGS AS
WE GO FORWARD, OBVIOUSLY THERE
ARE THINGS WE CAN CONTRACTUALLY
DO IN CERTAIN SITUATIONS.
THAT'S WHY THE LEGISLATION PIECE
IS SO IMPORTANT AND CRITICAL.
THERE ARE NOT ALWAYS ETHICAL
PLAYERS IN THE MIX.
>> AT THE TIME THAT HAPPENED WAS
THIS TECHNOLOGY NOT AS WIDELY
KNOWN?
>> I THINK FOR EVERYBODY THIS IS
MOVING SO QUICKLY PEOPLE ARE NOT
AS FAMILIAR WITH HOW IT HAPPENS.
I DO TALK TO PEOPLE IN THE
GOVERNMENT BOTH ELECTED AND
OTHERWISE WHO ARE NOT AS
FAMILIAR WITH WHAT THIS MEANS
AND THE USES FOR IT.
I AM OFTEN IN THE MIDDLE OF A
MEETING.
I'M LIKE I IMMEDIATE TO SHOW YOU
OR MAKE YOU HEAR WHAT I'M
TALKING ABOUT TO MAKE THAT
KNOWN.
AS THIS TECHNOLOGY IS GETTING
BETTER AND BETTER AND THE
DIFFERENCES ILLUSTRATED EARLIER
BY PATRICK ARE PERFECT EXAMPLES

HOW QUICKLY THIS REALLY MOVES.
WE'RE INCREASINGLY MORE CONCERNS
ABOUT IT.
>> MONA, HAVE YOU SEEN ANY
INSTANCES WHERE THIS TECHNOLOGY
HAS BEEN USED?
>> NO, I HAVEN'T.
I HAVE SEEN REAL LIFE
APPLICATIONS OF DEEP FAKE VIDEO.
IT DIDN'T HAVE DEEP FAKE AUDIO
IN THE CONTEXT OF A PRETTY
PERNICIOUS SEXTORTION CAMPAIGN.
I WANT TO HIGHLIGHT THE DAMAGE
THAT CAN BE DONE.
PUT AGO SIDE FRAUD AND FINANCE
LOSS.
I CAN'T BEGIN TO DESCRIBE THE
LIFE-LONG DAMAGE THAT IS DONE
WHEN A EITHER REAL OR DEEP FAKE
VIDEO WITH OR WITH OUT AUDIO
GETS OUT FOR WHAT ARE OFTEN
YOUNG PEOPLE IN VERY VULNERABLE
TIMES OF THEIR LIVES ALREADY.
EVEN IF THREE NANOSECONDS LATER
IT'S DEBUNK AS A FAKE IT DOESN'T
MATTER THOSE PEOPLE.
THOSE IMAGES CAN'T BE CLAWED
BACK.
NINE OUST TIMES I INTERVIEW
VICTIMS AND WE'RE GETTING READY
FOR A SENTENCING MY VICTIMS ARE
MORE CONCERNED ABOUT GETTING
THEIR CONTENT OFF THE INTERNET
THEN THEY ARE ABOUT HOW LONG
THIS GUY WILL BE IN JAIL.
YOU KNOW THE, THERE ARE REPORTS
OF AUDIO, DOPE FAKE AUDIO BEING
USED IN A BEC TYPE BUSINESS
E-MAIL TYPE SCAM.
I DON'T KNOW FOR SURE.
I HAVE NO FIRST HAND KNOWLEDGE
WITH THAT.
WHEN IT'S FINANCIAL ARGUABLY YOU
GET YOUR MONEY BACK.
IN THIS OTHER HARASSMENT CONTEXT
YOU'RE FIRED FROM YOUR JOB, YOUR

PERSONAL RELATIONS WILL OVER.
YOUR PRIVACY IS FOREVER
VIOLATED.
THAT GENIE CAN'T BE PUT BACK IN
THE BOTTLE EVEN IF IT'S EXPOSED
AS A DEEP FAKE IN FAIRLY SHORT
ORDER.
>> OKAY.
>> RUPAL, I THINK YOU MENTIONED
THE TECHNOLOGY CAN BE COSTLY IN
CERTAIN INSTANCES.
DO YOU THINK COST IS A DETERRENT
FOR BAD ACTORS TO USE THE
TECHNOLOGY?
>> NO, THE COST IS A FRACTION OF
THE COST IT USE TO BE.
I ACTUALLY THINK PART OF THE
ISSUE TOO IS THERE IS SO MUCH
OPEN SOURCE TECHNOLOGY AVAILABLE
THESE DAYS.
IT'S GOOD.
THERE IS AN ADVANCEMENT OF
TECHNOLOGY WITH THE OPEN SOURCE
MOVEMENT.
IT ALSO MEANS YOU DON'T KNOW WHO
ACT SEXER EVERYONE HAS ACCESS.
PEOPLE PICK UP FRAGMENTS AND
START TO USE THESE CODE BASES.
OPEN SOURCE ISN'T A RECIPE ON
HOW TO MAKE A SYNTHETIC VOICE.
YOU HAVE TO KNOW SOMETHING ABOUT
THE TECHNOLOGY ITSELF.
IT DOESN'T TAKE ALL THAT MUCH
TIME.
IF YOU'RE DETERMINED, LIKE YOU
SAID FRAUDSTERS ARE DETERMINED
TO FIGURE THIS OUT.
THEY WILL FIGURE OUT HOW TO DO
THIS.
I DON'T THINK IT'S COST OR
ACCESS.
FRANKLY IT'S NOT EVEN AUDIO.
THERE IS SO MANY AUDIO ON THE
INTERNET NOW.
ANYONE OF US.
IN FACT YOU'RE TAPING THIS

RIGHT.
THERE IS SO MUCH AUDIO FOR MANY
PEOPLE WHO, WHO OFTEN GIVE
TALKS.
IT'S CLEAN AUDIO TOO.
THAT AUDIO IS THEN AVAILABLE.
LIKE WE HAVE BEEN TALKING ABOUT,
YOU ONLY NEED A LITTLE BIT TO
THEN GET SOMETHING CLOSE ENOUGH.
THEN AS YOU BUILD THE MODEL YOU
CAN ADD MORE DATA TO IT AS WELL.
RIGHT.
I DON'T, I THINK THERE ISN'T --
I DON'T THINK THIS ARE BARRIER
COSTS OR ACCESS ANYMORE.
>> ARE COMPANIES THINKING ABOUT
WHETHER OR NOT TO RELEASE THE
TECHNOLOGY THROUGH OPEN SOURCE
CODE BECAUSE OF THE CONCERNS
THAT COSTS MAY NOT BE A BARRIER?
>> MEANING IF YOU BUILD THE
TECHNOLOGY -- SO, I THINK THERE
IS DEFINITELY A BIG OPEN SOURCE
MOVEMENT.
THERE ARE SOME THINGS THAT MOST
COMPANIES WILL STILL KEEP PRO
PROPRIETARY BECAUSE YOU DON'T
HAVE EVERYTHING OPEN SOURCED
UNLESS YOU ARE A OPEN SOURCE
COMPANY COMPLETELY.
THIS ARE THINGS YOU DO,
OTHERWISE SOMEONE CAN DO WHAT
YOUR ART IS.
I THINK THAT ONE OF THE THINGS
IS YOU DON'T ALSO WANT TO PATENT
EVERYTHING.
AGAIN YOU TELL PEOPLE HOW TO DO
SOMETHING.
SO THERE IS, IT'S A TRICKY
INTELLECTUAL PROPERTY ISSUE.
YOU HAVE TO WORK AROUND, WELL
WHO DO YOU WANT TO HAVE ACCESS
TO THIS?
HOW DO YOU WANT TO PROTECT IT?
HOW DO YOU WANT TO PROTECT YOUR
OWN INTELLECTUAL PROPERTY?

HOW DO YOU MACH SURE IT'S NOT
REVERSED ENGINEERED.
IT'S COMPLICATED AND NOT A
SIMPLE SOLUTION TO WORK.
>> OKAY.
JOHN, DO YOU FIND THAT COST IS A
FACTOR IN YOUR PATIENT'S
DECISION WHORPB TO USE THE
TECHNOLOGY?
>> CURRENTLY COST HAS BEEN A
MINIMAL DECISION.
IN PART BECAUSE AS THE
IMPORTANCE OF THE TECHNOLOGY HAS
BECOME SO EVIDENT THERE ARE A
THUB OF FILL AN THOP I CAN
ORGANIZATIONS THAT HAVE STEPPED
UP TO COVER THE COST FOR PEOPLE
WHO ARE AT RISK OF LOSING THEIR
ABILITY TO SPEAK.
THE COST HAS CONSIDERABLY
DROPPED.
WHERE THE VOYAGESES HEARD FROM
ACAPELA COSTS UPWARDS OF A
THOUSAND DOLLARS THEY'RE NOW
CHARGING A HUNDRED DOLLARS A
YEAR TO CREATE THE VOICE OVER
TEN YEARS.
AT THE END OF TEN YEARS IT'S
DONE.
WITH THAT THERE ARE
PHILANTHROPIC ORGANIZATIONS.
FOR EXAMPLE TEAM GLEASON OUST
NEW ORLEANS FOUNDED BY STEVE
GLEASON A STORMER NFL PLAYER,
THEY COVER THE COST OF VOICES
FOR PEOPLE WITH ALS.
>> I WANT TO SWITCH THE
CONVERSATION TO THE PRIVACY
IMPLICATIONS OF THE TECHNOLOGY.
JOHN, YOU MENTIONED SOME OF THE
PATIENTS USING THE TECHNOLOGY
FEE ARE NOT WORRIED ABOUT THEIR
VOICE LIVING IN PER I STEW IT
BUT THEY WANT TO MAKE SURE
CERTAIN MESSAGES ARE MAINTAINED
FOR FAMILY MEMBERS.

IS PRIVACY DISCUSSED WITH
PEOPLE?
>> YES PRIVACY IS DISCUSSED
WHERE THE RECORDINGS ARE STORED,
WHO HAS ACCESS, THE SIGN-ON
CREDENTIALS FOR THAT.
THEN WHEN A VOICE IS CREATED
AGAIN HAVING CREDENTIALS
CONTROLLED BY THE PATIENT AND
MAKING SURE THAT THEY'RE FULLY
IN CONTROL OF WHO HAS ACCESS TO
THAT.
TO YOUR QUESTION AS WELL MANY
PATIENTS HAVE INDICATED
SOMETIMES AT THE FIRST TIME
MOATING THEM AS THEY THINK ABOUT
THE POTENTIAL OF HAOZING THE
ABILITY TO SPEAK THAT THEY WOULD
LIKE TO MAKE SURE IF OR WHEN
THEY PASS AWAY THAT MESSAGES
THEY HAVE BANKED ARE ACTUALLY
AVAILABLE TO THEIR LOVED ONES.
BECAUSE OF COURSE THEY'RE OFTEN
THINKING ABOUT ALSO LEGACY
MESSAGES PEOPLE HAVE UP FRONT IN
PERPETUITY.
THERE HAS BEEN LESS INTEREST.
TO DATE I DON'T BELIEVE ANYONE
HAS ASKED ME TO MAKE SURE THEIR
SYNTHETIC, CUSTOM SYNTHETIC
VOICE IS AVAILABLE.
INSTEAD THEY WANT SPECIFIC
DIRECTIVE MESSAGES ARE PRESERVED
FOR LOVED ONES.
>> WE HAVE A NUMBER OF CASES
WHERE THE SPOUSES AFTER A
PATIENT HAVE PASSED AWAY HAVE
WANTED THEIR VOICES.
THEY, THEY REALLY WANT TO USE
THAT VOICE AFTER AS A WAY TO
REMEMBER THAT PERSON.
IT DOES MAKE ME WONDER ABOUT
THINGS LIKE OTHER TYPE OF
CRIMES.
INSURANCE CRIMES.
YOU KNOW COULD YOU KNOWS -- AS

THAT PERSON.
I KNOW THAT MANY OF THESE
PEOPLE.
YOU THINK THEY WOULDN'T DO
THINGS LIKE THAT.
YOU HAVE TO START THINKING ABOUT
HOW DO YOU, THIS PERSON IS
GRIEVING.
HOW DO YOU MAKE SURE, YOU'RE NOT
CREATING A BACK DOOR FOR OTHER
ISSUES.
THEY'RE EMBEDDED TECHNOLOGY GOS
AS WELL.
THE VOICES ARE OLDER AND MAYBE
WON'T FOOL ANYONE.
BUT THE MARCH OF TECHNOLOGY IS
MOVING FORWARD AND TWO YEARS
FROM NOW I MAY NOT FEEL THAT
WAY.
WE HAVE TO THINK ABOUT THE
NEGATIVE CONSEQUENCES DOWN THE
ROAD EVEN THOUGH PEOPLE ARE
ASKING OUT OF SOMETHING THAT
SEEMS SO BENIGN AND SO WORTH
WHILE TO SUPPORT.
HOW DO WE KNOW THAT SOMEONE IN
THE FAMILY WON'T GET ACCESS TO
THAT EUFPLT PHONE OR HOW MANY
DEVICES CAN YOU DOWNLOAD IT ON,
RIGHT.
THEN WHAT WOULD HAPPEN WITH
THAT.
IT'S SOMETHING I'M SCRAPPING MY
HEAD ABOUT THESE DAYS.
IT'S SOMETHING WE DON'T
ANTICIPATE.
>> I SUSPECT THE DIFFERENCE IN
OUR EXPERIENCE IS BECAUSE MANY
OF OUR PATIENTS UP FRONT HAVE
CREATED MESSAGES THAT ARE
ALREADY KIND OF PRECONSTRUCTED
AND THAT'S WHAT IS IMPORTANT TO
PEOPLE.
SO INTERESTINGLY BECAUSE
FAMILIES HAVE THAT.
THEY DON'T ASK FOR THE VOICE

BECAUSE THEY HAVE THE HEARTFELT
MESSAGES IN HAND.
THAT MAKES PERFECT SENSE.
>> I'M THE VOICE OF DOOM HERE.
THINK OF -- IF YOU THINK I WON'T
TALK TO PEOPLE ON THE PHONE I
DON'T KNOW.
PEOPLE CAN'T CALL ME AS A PRANK
CALL TO RECORD ME.
JUST HAVING SIR A TIPPOUS
RECORDING DEVICES.
THEY'RE HIGH QUALITY.
THEY'RE EASY TO HIDE.
YOU CAN STAND NEAR SOMEONE WHILE
THEY'RE HAVING A CONVERSATION
WITH SOMEONE ELSE AND GET YOUR
SEGMENT THAT NEED FOR A HIGH
QUALITY DEEP FAKE AUDIO OF THEM. .
>> YES.
I THINK ABOUT THE ACTS ARE, THE
TRUSTED FACES THAT YOU HOOK AT.
YOU LOOK AT A PERSON, MAKING A
CHOICE WHAT YOU DO OR NOT DO,
THE BROADCASTERS, PUBLIC
POLITICIANS, PUBLIC OFFICIALS
OUT THERE.
PUT THAT IN THE POSSIBILITIES OF
SOMEONE AND WHAT THEY CAN DO.
IF ONE OF THOSE PEOPLE ARE
GIVING ADVISOR DIRECTIVE ON
SOMETHING.
IT MAY SEEM BENIGN.
I THOUGHT WHEN MONA TALKED
EARLIER WHAT IF ONE OF OUR
MEMBERS ASKED FOR DONATIONS TO X
CAUSE IN A DISASTER.
NOW YOU HAVE THE VOICE AND MAYBE
A DEEP FAKE VIDEO OF SOMEONE
ASKING YOU TO DONATE TO THE
CAUSE.
THESE ARE NOT THE THINGS PEOPLE
ARE OFTEN THINKING ABOUT.
WE THINK ABOUT IT IN THOSE WAYS
OF SECURITY.
IF YOU'RE THE PERP MAKING YOUR
LIVING DOING THAT THEN YOU'RE

NOT A TRUSTED FIGURE THE NEXT
TIME YOU DO IT.
PEOPLE SPEND LIVES BUILDING UP
THOSE CAREERS AND TRUST WITH THE
PUBLIC.
THEN IT'S GONE AND YOU CAN'T
CLAW IT BACK.
>> CAN I PLAY DEVIL'S ADVOCATE
ON THE OTHER SIDE.
THERE ARE ARTISTIC AND
ENTERTAINMENT APPLICATIONS AS
WELL.
THINK ABOUT TAKING A
MOTIVATIONAL -- TAKE EINSTEIN.
IF YOU HAVE GREAT AUDIO OF
EINSTEIN AND YOU CAN CREATE A
PHYSICS, YOU KNOW, TUTORIAL OR
EDUCATIONAL APPLICATION AROUND
EINSTEIN, NOBODY BELIEVES HE'S
STILL HERE.
YOU'RE NOT TRYING TO FOOL
ANYONE.
IT'S INSPIRATIONAL.
THERE ARE ARTISTIC ELEMENTS
WHERE YOU THINK ABOUT THE USE OF
VOICE AND TECHNOLOGY IN
INTERESTING WAYS.
WE HAVE PROJECTS NOW WE'RE BEING
ASKED TO BRING BACK TO LIFE
CERTAIN CHARACTERS.
IT'S JUST THE WAY WE THINK AND
MEMORIALIZE SOME OF AS WELL.
THERE IS DEFINITELY THE DOOM
STUFF.
THIS IS ALSO SOME OF THE THINGS
THAT CAN BE INSPIRATIONAL.
IF YOU'RE LOOKING AT EDUCATIONAL
APPLICATIONS WE DON'T HEAR
DEMOGRAPHIC DIFFERENT VOICES IN
THOSE.
WHAT DOES IT MEAN TO THOSE
COMMUNITIES?
THAT THEY'RE NOT REALLY HEARD.
IF YOU HAVE VOICES THAT
REPRESENT PEOPLE THAT THE
TECHNOLOGY SOUNDS LIKE, THAT'S

GREAT.
MAYBE THERE IS AN ADOPTION AND
TRUST IN A DIFFERENT WAY.
THIS ARE TWO SIDES TO LOOKING AT
THIS.
LET'S NOT GET TOO DOWN THE
RABBIT HOLE THAT IT'S ALL GOING
TO POT.
>> I DON'T THINK -- I DON'T KNOW
MAYBE MONA SAYS IT IS.
I THINK THIS ARE EXAMPLES OF
PEOPLE WHO HAVE MADE DIRECTIVES
AND THE LAWS DIDN'T CATCH UP
WITH IT.
PRINCE DIDN'T WANT TO BE A
HOLOGRAPHIC PERFORMER.
NOBODY WOULD DOUBT THE ARTISTIC
EXPRESSION OF PRINCE AND THE
BEAUTY OF THAT.
IT TOOK AN OUTCRY FROM PEOPLE TO
SAY THAT'S NOT ACCEPTABLE.
THE QUESTION IS, SHOULDN'T IT BE
PEOPLE'S ESTATES.
SHOULDN'T PEOPLE BE ABLE TO MAKE
THOSE CHOICES FROM A POSTMORTEM
RIGHT OF PUBLICITY TO MAKE THOSE
DECISIONS AND OTHERS ARE NOT
MAKING THE DECISIONS FOR PEOPLE.
I THINK THAT'S PART OF A UPON
ANSWERRIZATION.
IT CAN BE WONDERFUL WHEN, HEY, I
LOVED TO GET SEEING CARRY FISHER
ONE MORE TIME AS A STAR A WARS
FAN.
THAT WAS DONE WITH CONSENT.
WHAT ARE THE APPROPRIATE TENS
AND CAN WE WORK WITH ETHICAL
PLAYERS TO DO IT.
IF THE ANSWER TO THAT IS YES,
THEN THAT'S A FANTASTIC THING.
>> I THINK JUST TO FIND THE
COMMON GROUND BETWEEN WHAT YOU
ARE BOTH SAYING.
I THINK YOU HAVE BOTH DISCUSSED
THE ISSUE OF CONSENT.
WHICH LEADS ME TO MY NEXT

QUESTION DEALING WITH NOT
DORISMENTS.
SO YOU TALKED EARLY ABOUT EX EX
CLUE CYST AND CON TRAPBTS.
--
>> YES.
IF YOU'RE THE VOICE OF A FAST
FOOD CHAIN AND THEN THE
COMPETITOR THE NEXT DAY CAN TAKE
YOUR VOICE AND MAKE YOU THEIR
VOICE AS WELL.
YOU HAVE DIMINISHED THE VALUE OF
THE PERFORMER BUT ALSO THE
BUSINESS WHO HAS INVESTED THE
TIME AND MONEY.
THERE ARE MANY SITUATIONS WE
HAVE TO HAVE AWARENESS OF THAT.
THE WONDER OF BEING ABLE TO --
MAYBE ARE YOU THE VOICE AND
THROUGH TECHNOLOGY AND CONSENT
YOU CAN EXPANDED.
PAEB THE HIGH PROFILE PERFORMER
ISN'T GOING TO DO THE PHONE
TREE.
MAYBE NOW THEY WILL.
THERE IS AN EXCITING SPACE IN
THIS WHOLE WORLD WE'RE EXCITED
ABOUT THE VOICE FIRST TECHNOLOGY
IN TERMS OF APPLICATIONS AND
OTHERWISE.
THAT TO ME IS AMAZING.
THERE IS WONDERFUL POTENTIAL
THERE.
IT HAS TO BE SOMETHING WE'RE
DOING THE DOPE DIVE TO MAKE SURE
THIS IS THE PRECAUTIONS.
WHOLE BUSINESS MODELS NEED TO BE
PROTECTED UNDER THAT.
>> MOAN A I CAN SEE FROM A LAW
ENFORCEMENT PERSPECTIVE TOO.
IF A PARTICULAR INDIVIDUAL'S
VOICE IS USED TO ENDORSE A
PRODUCT OR SERVICE OR SOMETHING
THAT COULD CREATE LAW
ENFORCEMENT CHALLENGES.
ALSO CONCEPTIONS FROM CONSUMERS

PERSPECTIVE AS TO WHO IS
SUPPORTING.
>> ABSOLUTELY.
YOU CAN COP COME UP WITH A STOCK
MARKET MANIPULATION SCENARIO.
YOU HAVE SOMEONE POSING AS THE
CEO OF A COMPANY MAKING A BOGUS
EARNINGS CALL TO MANIPULATE THE
STOCK PRICE.
THAT WOULD BE WIRE FRAUD AND A
NUMBER OF SECURITY REGULATIONS.
CONVERSELY YOU COULD HAVE A
COMPETITOR BASICALLY SABOTAGE
ANOTHER COMPETITOR SIMILAR TO
WHAT REBECCA SAID.
POSE AS THE COMPETITOR CEO AND
MAKE FALSE NEGATIVE CLAIMS ABOUT
THE COMPANY.
THE MARKET WOULD RESPOND
INSTANTLY.
MAYBE THE INFORMATION WOULD
FINALLY COME TO LIGHT IT WAS A
FAKE CALL.
MAYBE THE MARKETS ULTIMATELY
RECOVER.
IN THE MEANTIME MASSIVE AMOUNTS
OF SHAKING THE CONFIDENCE OF THE
MARKETPLACE HAS HAPPENED.
NOT TO MENTION PRESUMABLY WHO IS
SHORT OF THE STOCK HAS MADE A
FORTUNE.
SO YES THERE ARE, THE CRIMES,
THE LAW WOULD COVER ABUSIVE USED
OF THE, OF THE VOICE.
IT'S LESS CLEAR TO ME QUITE
FRANKLY WHETHER MERELY TAKING
SOMEONE'S VOICE WITHOUT
AUTHORIZATION, YOU KNOW RIGHT
NOW THE STATUTES THAT COME TO
MIND.
I'M PROBABLY WAY INTO THE WEEDS
NOW.
WE HAVE 18 USC TURN 29 THE
UNAUTHORIZED USE OF A ACCESS
DEVICE AND THEN WE HAVE ONE FOR
A MEANS OF IDENTIFICATION.

QUERY WHETHER SOME ONES VOICE
WOULD BE ACCESS DEVICE OR MEANS
OF IDENTIFICATION PARTICULARLY
IF IT WAS USED IN A BIO MET I
CAN UNLOCK CONTEXT.
IF YOU HAVE AN ACCOUNT AND ONE
PRIMARY WAY, A BIOMETRIC LOCK IS
YOUR VOICE.
YOU COULD HAVE SOMEONE POSE AS
YOU AND STEAL YOUR MONEY.
THOSE SCENARIOS ARE COVERED ONCE
THEY ACTUALLY MISUSE THE VOICE
EXEMPLAR.
THE QUESTIONING IS IT A CRIME TO
ACQUIRE WITHOUT AUTHORIZATION
SOMEONE'S VOICE EXEMPLAR IN THE
FIRST INSTANCE.
I THINK THAT'S A MORE OPEN
QUESTION.
>> HOW HAVE YOU ALL SEEN THE
TECHNOLOGY CHANGE.
WE HAVE TALKED ABOUT HOW RAPIDLY
THE TECHNOLOGY HAS CHANGED THE
LAST COUPLE OF YEARS.
HOW HAS IT AFFECTED YOUR
EXPERIENCE WITH THE TECHNOLOGY
AND PEOPLE USING IT?
>> SO, I CAN SAY THAT WITH THE
AVAILABILITY OF DNN TECHNOLOGY
FOR CREATING VOICE, THE QUALITY
OF THE VOICE CREATED HAS BEEN
SIGNIFICANTLY IMPROVED.
WITH THAT OBVIOUSLY IS MUCH MORE
ATTRACTIVE TO SOMEONE WHO IS AT
RISK OF LOSING THEIR ABILITY TO
SPEAK.
I THINK THAT SPEAKS TO WHY IN
THE PAST COUPLE OF YEARS THERE
IS SO MUCH MORE INTEREST AND
AWARENESS OF THE ABILITY TO
CREATE A CUSTOM SYNTHETIC VOICE.
IT'S SO MUCH MORE ATTRACTIVE
BECAUSE OF THE QUALITY.
THAT'S DIRECTLY RELATED TO THE
SUBSTANTIAL RAPID CHANGE IN THE
TECHNOLOGY.

>> REBECCA.
>> YES FOR US IT HAS CAUSED A
COMPREHENSIVE REVIEW OF ALL OF
OUR CONTRACTS.
AT THE END OF THE DAY BEING ABLE
TO NEGOTIATE AND ENFORCE THE
CONTRACTS DIRECTLY IS AT THE FOR
FRONT OF WHAT WE DO.
IT CERTAINLY HAS EXPANDED OUR
LEGISLATIVE EFFORTS AROUND THE
COUNTRY, FOR SURE.
I THINK IT'S DONE A THIRD THING,
FOR US TO BE LOOKING AT AND
WORKING WITH PEOPLE THAT ARE
ETHICAL PARERS IN IN THIS SPACE.
MAKING SURE PEOPLE USING THIS
TECHNOLOGY AND UNDERSTANDING
FARE COMPENSATION ARE ABLE TO
WORK WITH OUR PERFORMERS.
>> WHEN YOU SAY LEGISLATIVE
EFFORTS YOU ARE TALKING ABOUT
THE STATE RIGHT TO PUBLICITY
LAWS YOU REFERENCED EARLIER
DEALING SPECIFICALLY WITH AN
ACTOR'S RIGHT TO THEIR OWN
VOICE.
NOT NECESSARILY --
>> IMAGE AND VOICE RIGHTS WITH
REGARD TO THEIR VOICE.
WE TALKED EARLIER ABOUT STATES,
YOU KNOW GREAT STATES LIKE, YOU
KNOW CALIFORNIA AND TENNESSEE
THAT HAVE DIFFERENT PROVISIONS
ALREADY.
THIS ARE A LOT OF STATES.
I'M A NEW YORKER.
NOT A GREAT RIGHT OF PUBLICITY
IN LAW, THEY DON'T HAVE A POST
MORTEM RIGHT THINGS LIKE.
THAT WE HAVE GOVERNMENT AFFAIRS
AND PUBLIC POLICY DEPARTMENT
THAT SPECIALIZES WITH ANALYZING
THOSE LAWS FOR STATE
LEGISLATURES FOR OUR MEMBERS AND
THE PUBLIC.
>> AT THIS POINT SAG HASN'T BEEN

WORKING ON LEGISLATION
SPECIFICALLY FOR VOICE CLONING.
TO BE CLEAR.
>> NO, SO IN THE RIGHT OF
PUBLICITY IT IS A BROADER BASED
RIGHT OF PUBLICITY.
WE'RE NOT LOOKING AT IN TERMS OF
ONE ASPECT OF IT.
>> IN TERMS OF HOW I THINK THE
TECHNOLOGY HAS CHANGED.
CERTAINLY BEFORE WE HAVE HAD SO
MANY CHANGES IN THE TECHNOLOGY
WE'RE OFFERING AND BUILDING.
I THINK IT'S CHANGED THE MARKETS
WE CAN SERVE.
INITIALLY PEOPLE WITH DISABILITY
WE CONTINUE TO SERVE THAT
POPULATION.
IT'S A SMALL FRAGMENT OF THE
POPULATION.
YOU CAN'T NECESSARILY BUILD A
BUSINESS AROUND THAT.
NOW THE VOYAGESES SOUND BETTER
THERE ARE OTHER INDUSTRIES YOU
CAN HAVE AS A RESULT OF THAT.
IT'S ALSO BECAUSE THE MELODY IS
MORE APPROPRIATE AND MORE HUMAN
LIKE THAT IT ACTUALLY BECOMES
MORE REALISTIC SOUNDING FOR THE
APPLICATIONS.
THE OTHER THING IS TOO WE ENDED
UP HAVING SOME WORK THAT WE DID
WITH A FINANCIAL INSTITUTION.
WE WERE ASKED TO BUILD THESE
VOICES AND SEE IF WE COULD BREAK
THROUGH THE AUTHENTICATION
SYSTEM.
THAT IS WHERE WE STARTED TO
THINK MORE ABOUT HOW TO
SAFEGUARD THIS.
AS WE SAW THE MARCH OF THE
TECHNOLOGY MOVING FORWARD WE
REALIZED WE CREATED THIS THING
THAT INITIALLY WAS BENEVOLENT
AND HAD A MISSION DRIVEN ASPECT
TO IT WE NEEDED TO ALSO BE

THINKING ABOUT THE PROTECTION
PART OF IT AS WE EVOLVE IN THIS
WE HAVE KIND OF GONE FROM OKAY
IT SOUNDS LIKE THIS PERSON.
HOW DO WE SCALE IT.
HOW DO WE MAKE THE COSTS MORE
ACCESSIBLE.
THEN IT WENT FROM THAT TO IT'S
NOW SOUNDING BETTER AND WE HAVE
DIFFERENT TECHNOLOGIES TO NOW WE
HAVE TO PROTECT IT.
IT'S ACROSS.
>> I WOULD SAY THE ANCILLARY
CRIMINAL TOOLS THAT WOULD
AMPLIFY, NO PUN INTENDED, THE
VOICE TECHNOLOGY HAVE
EXPONENTIALLY IMPROVED THE LAST
TWO TO FIVE YEARS.
FIVE YEARS AGO WHEN I GOT A CASE
I COULD PRETTY MUCH CUT FIVE
SUBPOENAS AND CATCH THE GUY IN
THREE MONTHS, SIX MONTHS.
NOW WITH AMNON HIGHING

TECHNIQUES -- THE ANCILLARY
TOOLS HAVE GOTTEN VERY CHEAP,
VERY READILY AVAILABLE AND VERY
GOOD.
THE CRIMINALS ARE LOVING IT.
SO IT'S THAT IN TANDEM WITH THE
VOICE ISSUES THAT ARE A LAWENCE
FORCEMENT CHALLENGE.
>> WE TALKING ABOUT TECHNOLOGIYS
WITH A SLIGHT DELAY WHAT HAD THE
PERSON USING THE TECHNOLOGY
WANTS TO SAY AND THE
TECHNOLOGY'S OUTPUT.
ARE WE SEEING THIS TECHNOLOGY
BEING USED IN REALTIME?
URGES YES.
THERE ARE METHODS THIS IS REAL
TIME USE.
SYNTHETIC USE HAS BEEN USED REAL
TIME FOR A LONG TIME.
IT JUST DOESN'T SOUND VERY HUMAN
LIKE.

>> I THINK THE QUESTION BETWEEN EX TESTIMONY POOR AIRIOUS SPEECH AND PREPARED SPEECH IS IMPORTANT THERE.
THE SUNG WILL MESSAGE CAN BE REAL TIME.
IF SOMEONE IS FORTUNATE TO HAVE COME UP WITH ALL OF THE RIGHT SCRIPT ITEMS FOR A CONVERSATION AND PRESTORED THOSE, POTENTIALLY A CONVERSATION COULD SEEM TO BE REAL TIME.
>> AS THE REAL TIME TECHNOLOGY GO GETS BETTER I WOULD EXPECT ANY CRIMINAL USING THE TECHNOLOGY WOULD BE LIKELY SUCCESSFUL IMPLICATED WHAT MONA WAS DISCUSSING.
>> I REALLY DO WANT TO SAY I'M SELF CONSCIENCE ABOUT BEING THE VOICE OF DOOM.
I THINK THE TECHNOLOGY GO IS FANTASTIC.
I DON'T WANT THE TAKEAWAY TO BE WE SHOULDN'T DO THIS.
I JUST WANT TO EDUCATE CONSUMERS, ASKING QUESTIONS BACK ABOUT SECRET INFORMATION THAT PERHAPS ONLY THE SPEAKER WOULD KNOW, A TRUSTED PERSON, DOING FRONT END TECHNOLOGY, I DON'T KNOW IF THEY CAN WATER MARK THE AUDIO SO THAT IT'S READILY, THE TECHNOLOGY IS READILY AVAILABLE TO CONSUMERS TO DO A LITMUS TEST ON A REAL VOICE OR NOT.
I DON'T WANT THE TAKE AWAY TO BE THIS IS BAD WE SHOULDN'T DO IT.
IT'S MORE LIKE, WE NEED TO THINK UP FRONT ABOUT PROTECTING CONSUMERS WHO WILL DEFINITELY BE VICTIMIZED IN WAYS BY CRIMINALS.
IN MY DARK WORLD EXPERIENCE THE CRIMINALS ARE USUALLY AHEAD OF US IN TERMS OF LAW ENFORCEMENT.
I THINK EDUCATING CONSUMERS IS

MORE IMPORTANT.
I'M NOT SURE THE TECHNICAL FIX
WILL BE READILY AVAILABLE OR UP
TO DATE ENOUGH FOR THE AVERAGE
CONSUMER TO USE AND DETECT
WHETHER IT'S AUTHENTIC OR DEEP
FAKE AUDIO.
WHAT THAT CONSUMER EDUCATION
PLATFORM LOOKS LIKING.
I GUESS THAT'S PARTLY WHY WE'RE
HERE TODAY.
THINKING ABOUT EDUCATING
CONSUMERS.
YOU'RE PROBABLY IN THE BEST
POSITION TO TELL US WHAT TO LOOK
FOR TO DETECT A DEEP FAKE.
>> YES.
>> I THINK YOU TOUCHED ON THAT
IN THE OPENING REMARKS.
YOU TALKED ABOUT AWARENESS BEING
ONE OF THE MOST IMPORTANT
THINGS.
WHAT ARE COMPANIES DOING?
OR ARE COMPANIES NOT THE PROPER
MECHANISM OF MAKING PEOPLE AWARE
OF THE TECHNOLOGY THAT EXISTS?
>> I THINK COMPANIES ARE
STARTING TO DO THAT.
RECENTLY THERE WAS A PRODUCTION
PUT ON BY THE MIT, IT'S A LONG
NAME LIKE CITIZEN AWARENESS OR
SOMETHING LIKE THAT.
THE PEACE WAS CALLED "IN THE
EVENT OF MOON LANDING."
IT WAS TAKING NIXON'S SPEECH
THAT HAD THEY NOT LANDED ON THE
MOON WHAT A DISASTER IT WOULD
HAVE BEEN.
RECREATED NICK ON DOING THAT
WITH VIDEO AND AUDIO.
IT WAS PLAYED AT A FILM FESTIVAL
IN THE NETHERLANDS TPHURBLLY.
THERE WILL BE FOOTAGE RELATED TO
THAT AFTERWARDS.
PART OF THE GOAL OF THE PROJECT
WAS TO EDUCATE PEOPLE ABOUT WHAT

IS AVAILABLE WITH THE TECHNOLOGY
THESE DAYS.
AS THEY SHOW THE VIDEO THERE ARE
MARKINGS ON THE SIDE AND
ANNOTATIONS.
THAT'S FAKE.
THE LIPS MOVED IN THIS WAY.
THE LOWER LIP IS SLOWER THAN THE
UPPER LIP.
THINGS LIKE THAT TO EDUCATE
PEOPLE WHAT TO LOOK FOR.
IF YOU BRING PEOPLE UP TO SPEED
ON SIMPLE THINGS.
YOU KNOW YOUR UPPER LIP IS
ATTACHED TO THE PART OF THE FACE
THAT DOESN'T MOVE.
YOUR LOWER LIP MOVES.
LOCK FOR THAT, RIGHT.
IN VIDEO YOU CAN'T ALWAYS, IN
ANIMATED VIDEOS SOME THINGS LIKE
THAT CHANGE A LITTLE BIT.
THERE MIGHT BE WAYS.
THESE ARE WAYS TO EDUCATE THE
PUBLIC.
IT WAS DONE THROUGH ART.
IT WAS DONE IN A WAY NOT TO
SCARE PEOPLE BUT TO EXPLAIN TO
THEM WHAT THE TECHNOLOGY CAN DO.
WE'RE WORKING WITH AN ARTIST
NOW, AN OPERA SINGER.
WE'RE MAKING A SHOW WHERE THERE
WILL BE A COMBINATION OF AUDIO
AND HER SYNTHETIC, HER SINGING
AND HER SYNTHETIC VOICE, NOT
SINGING.
WE CAN'T DO THAT YET.
PRODUCING SOUND.
THIS COMBINATION WILL BE ANOTHER
WAY TO BRING PEOPLE UP TO
THINKING ABOUT THE
POSSIBILITIES AND WELL AS
NEGATIVE CONSEQUENCES.
>> REBECCA FROM YOUR PERSPECTIVE
FOR ACTORS NOT PART OF THE
UNION.
ARE ANY EFFORTS BEING MADE TO

MAKE THEM AWARE OF THE ISSUES AS
WELL.
>> SO SAG-AFTRA TAKES THE VIEW
THAT THOSE WHO ARE NOT MEMBERS
YET ARE MEMBERS THAT HAVEN'T
JOINED YET.
WE DO THAT BECAUSE WE REALIZE
ALL PEOPLE KNOWN IN THE INDUSTRY
START OUT NOT BEING A MEMBER.
WHEN THEY RISE TO A CERTAIN
LEVEL OF PROFESSIONALISM BECOME
PART OF OUR UNION AND WE'RE
THANKFUL TO HAVE THEM.
WHEN WE THINK ABOUT THIS SPACE
WE THINK ABOUT EDUCATING ALL
PERFORMERS, ALL PEOPLE IN THE
EARLY STAGES.
WE DO EXTENSIVE OUTREACH TO
COLLEGES, UNIVERSITIES,
BROADCAST PROGRAMS TO MAKE SURE
THOSE PEOPLE UNDERSTAND THIS.
IT'S IMPORTANT TO US IN A LARGER
SCHEME FROM THE ARTISTIC AND THE
NEWS COMMUNITIES.
SO WE ARE VERY MUCH THINKING
ABOUT IT AS PART OF OUR HOLISTIC
APPROACH TO EDUCATING EVERYBODY:
>> YOU MENTIONED, REBECCA,
EARLIER YOU WERE VERY HAPPY THAT
THIS WORKSHOP WAS TAKING PLACE.
YOU HAVE BEEN AT THE TABLE WITH
A LOT OF DISCUSSIONS HAPPENING.
ARE YOU HAVING THE DISCUSSIONS
WITH COMPANIES OR THE
DISCUSSIONS WITH PROFESSORS,
INTELLECT YOU'LLS?
WHO IS AT THE TABLE NOW
DISCUSSING THESE ISSUES YOU'RE
PART OF?
>> WE'RE HAVING DISCUSSION BOTH
WITH COMPANIES, WITH PROFESSORS,
I THINK THIS IS PART OF A LARGER
QUESTION IN TERMS OF THE FUTURE
OF WORK IN GENERAL.
YOU KNOW WE HAVE ALWAYS TAKEN AN
APPROACH IN SAG-AFTRA IN TERMS

OF EMBRACING INNOVATION AND
HELPING OUR MEMBERS ADAPT TO IT.
OUR PRESIDENT IS THE SECTOR
CHAIR THROUGH THE AFLCIO.
PART OF WHAT WE HAVE DONE IN THE
U.S. AND FIA THE ORGANIZATION
FOR ACTORS AROUND THE WORLD IS
WORKING AND HAVING CONVERSATIONS
IN ALL OF THOSE PLACES THAT WANT
TO HAVE IT.
WE HAVE DONE CONVERSATIONS WITH
FOLKS ACROSS.
NOT JUST TO HAVE A UNDERSTANDING
WHAT IS THERE NOW AND BEING USED
BUT HOW IT WILL ADAPT.
IT COMES DOWN TO ARE ALL THE
COMPANIES MOVING INTO THE SPACE
THINKING ABOUT IT FROM A
ETHICAL POINT OF VIEW.
THERE ARE A LOT OF GREAT
INDUSTRY PLAYERS THAT ARE.
THEY VIEW IT FROM THE POINT OF
VIEW OF PEOPLES VOICES BEING
UNIQUE AND THEY BELONG TO THEM.
THEY WANT TO DO THIS IN A WAY
THAT MAKES SENSE.
THERE ARE OTHER, MAYBE SOME OF
THE TECH COMPANIES THAT DON'T
EVEN ACKNOWLEDGE THAT PIECE IN
THE PUZZLE THAT VOICES ARE
SOMETHING THAT IS UNIQUELY
YOURS.
THAT IS PART OF THE CONVERSATION
WE HAVE BEEN ENGAGED WITH.
>> ASIDE FROM THE POSITIVE USE
CASES WE SRAO DISCUSSED TODAY
ARE THERE OTHER POSITIVE USES OF
THE TECHNOLOGY DISCUSSED IN THE
MEETINGS.
>> YES, I MEAN OBVIOUSLY AS WE
GO TO THE VOICE-FIRST UNIVERSE.
THAT'S HOW WE INTERACT WITH THE
PHONES.
WE'RE USE TO TAPPING ON THE
SCREEN AND FINDING WHAT WE WANT
TO DO.

THERE ARE MULTIPLE USES, AS
LEAST FOR OUR MEMBERS AND THE
PUBLIC AT LARGE USING THAT
SOMEWHERE I THINK IT'S
INCREDIBLE.
IT PROVIDES A HUGE SPACE WHERE I
THINK WE CAN ALL AGREE THERE ARE
HUGE IMPROVEMENTS NOT
INTERACTING PHYSICALLY WITH YOUR
PHONE.
WE SEE THE BASICS OF THE
TECHNOLOGY ALREADY.
IMAGINE IT BEING DONE IN A WAY
SO MUCH MORE ROBUST.
IT'S EXCITING AND I KNOW OUR
MEMBERS ARE EXCITED ABOUT THAT
PIECE OF IT.
>> I THINK THERE ARE
APPLICATIONS IN HEALTHCARE, FOR
EXAMPLE.
NOT NECESSARILY PEOPLE WITH
DISABILITIES.
IF YOU THINK ABOUT PHYSICIANS.
THINK ABOUT PHARMACISTS AND
SOCIAL WORKERS.
WAYS THEY CAN ACTUALLY TOUCH
MORE LIVES AND ACTUALLY HAVE
BETTER CARE.
IT MIGHT BE THROUGH A VOICE AH
AVATAR OR VOICE CLONE, WHATEVER
YOU WANT TO CALL IT THESE ARE
NOT VOICE ACTORS.
THEY'RE PROFESSIONALS.
THE VOICE IS IMPORTANT IT'S
TRUSTED WITH THEIR INTERACTION
WITH THEIR PATIENTS.
THINKING ABOUT THE WAYS YOU
EXERCISE, PHYSICAL THERAPY, IT
MIGHT BE ASSOCIATED WITH AN APP
AS WELL.
THE COACH IS YOUR PT, RIGHT.
THERE ARE VARIOUS THINGS YOU CAN
THINK ABOUT HOW THE TECHNOLOGY
CAN BE USED IN ORDER TO CONTINUE
THE RELATIONSHIP THAT YOU HAVE
WITH A KNOWN INDIVIDUAL AS WELL.

\>> I THINK ALSO JUST THINKING ABOUT, YOU KNOW TECHNOLOGIES THAT WE'RE ALL RELYING ON IN OUR HOME, HOME AUTOMATION TECHNOLOGIES.

WE USE OUR VOICE TO CONTROL THAT.

A HIGH QUALITY EFFECTIVE VOICE FOR A PERSON WHO IS UNABLE TO SPEAK OR WHOSE SPEECH IS NOT CLEAR, IT'S REALLY ESSENTIAL IN TERMS OF GIVING THAT LEVEL OF I INDEPENDENTENCE THAT OTHERWISE IS NOT AVAILABLE TO THEM.

\>> OKAY.

IF THERE WAS ONE THING THAT YOU WANT THE PUBLIC KNOW KNOW AND WE HAVEN'T TALKED ABOUT ON THE PANEL, WHAT WOULD IT BE?

\>> I'M A BIG GOOSE EGG ON THIS ONE.

I -- I REALLY HAVE TO PASS.

\>> JOHN IS GOING TO GO.

\>> SO I WILL JUMP IN THEN.

I STARTED MY BRIEF DISCUSSION ABOUT BY HIGHLIGHTING OUR VOICE IS TRULY HOW WE CONNECT WITH THE WORLD AROUND US.

IT REPRESENTS US.

IT BEST ALLOWS US TO, TO DEVELOP SOCIAL RELATIONS AND REALLY REFLECT OURSELVES.

SO THIS TECHNOLOGY IS ALLOWING PEOPLE EITHER WHO ARE AT-RISK OR HAVE LOST THEIR ABILITY TO SPEAK TO MAINTAIN THAT SOCIAL CONNECTEDNESS.

THAT THREAD THAT ALLOWS US ALL TO RELATE TOGETHER AS HUMAN BEINGS.

ALSO WITH THE WORK THAT RUPAL IS DOING FOR PEOPLE WHO HAVE NEVER BEEN ABLE TO SPEAK.

IT GIVES THEM AN AUTHENTIC CONNECTION TO THE PEOPLE AND WORLD AROUND THEM.

THAT ESSENTIAL IMPORTANT PIECE.
AS MONA INDICATED WHAT IS
IMPORTANT IS THAT WE RECOGNIZE
THAT AND THEN FIGURE OUT HOW TO
MANAGE THIS SO IT'S SAFE.
ABSOLUTELY RECOGNIZE THE, THE
CRITICAL NATURE OF PERSONAL
VOICE.
>> I THINK I WOULD SAY THAT, YOU
KNOW, VOICES ARE VALUABLE TO
THE PEOPLE THAT DON'T HAVE THEM
AND NEED THEM.
TO PEOPLE WHOSE LIVES AND
CAREERS ARE MADE BY THEM, BUT
ALSO BY US WHO ARE THE
BENEFICIARIES OF HEARING THOSE
THINGS KNOW THEM, TRUST THEM,
AND UNDERSTAND THAT THOSE ARE
SOURCES OF GREAT POWER AND
INFLUENCE.
SO WHEN I'M THINKING ABOUT A
TAKE AWAY FOR ME IT'S FINDING
THE BALANCE BETWEEN WHAT IS
UNDOUBTEDLY TECHNOLOGY EXCITING
AND NOT MAKING SURE WE DON'T END
UP IN A BAD EPISODE OF BATTLE
STAR GALACTICA.
>> ANYTHING ELSE FROM OUR --
>> I THINK WE HAVE COVERED
EVERYTHING.
>> OKAY.
THANK YOU EVERYBODY FOR TAKING
THE TIME TO SPEAK WITH US ON THE
FIRST PANEL.
I THINK IT WAS VERY INFORMATIVE
AND VERY ENGAGING.
THANK YOU.
[APPLAUSE]
>> THANK YOU.
>> THANK YOU FOR THE
OPPORTUNITY.
THANK YOU.
>> GOOD AFTERNOON, WELCOME BACK.
I'M DAN SALSBURG FROM THE FTC.
BEFORE THE BREAK WE HEARD HOW
VOICE CLONING TECHNOLOGY WHO

HAVE LOST THEIR ABILITY TO
SPEECH TO SOUND LIKE FORMER
SELVES.
WE ALSO LEARNED THAT THIS
TECHNOLOGY CAN HAVE SERIOUS DARK
SIDE THAT ENABLES CRIMINALS TO
ENGAGE IN SCHEMES OF SCALE,
EXTORTION, OTHER CRIMES.
SO HOW DO WE BALANCE THE
POTENTIAL UPSIDES TO HARMS OF
NEW ARTIFICIAL INTELLIGENCE
TECHNOLOGY?
WHO SHOULD BE DOING THIS
BALANCING?
THESE ARE AMONG THE QUESTIONS
THAT WE'LL BE EXPLORING THE NEXT
HOUR WITH FANTASTIC GROUP OF
PANELISTS, FULL BEOS ARE
INCLUDED I'LL JUST BRIEFLY
INTRODUCE THEM NOW.
ON MY IMMEDIATE LEFT IS ROSALYN
BERNE A COMPUTER SCIENCE
PROFESSOR AT UNIVERSITY OF
VIRGINIA.
ALTHOUGH SHE'S NOT SPEAKING ON
BEHALF NATIONAL ACADEMY OF
ENGINEERING.
ON HER LEFT IS LEIGH HAFREY.
LECTURER AND BEHAVIORAL AND
POLICY SCIENCE ON SLOAN SCHOOL
OF MANAGEMENT AUTHOR OF NUMBER
OF BOOKS ON LEADER SHIPS AND
ETHICS INCLUDING FIVE STEPS TO
MASTERING ETHICS IN BUSINESS.
ASHISH JAIMAN IS DIRECTOR OF
MICROSOFT DEFENDING DEMOCRACY
PROGRAM WHERE HE'S WORKING ON
DISINFORMATION DEFENSE AND DEEP
FAKE INTERVENTION STRATEGIES.
AND FINALLY AT THE END OF THE
TABLE KIRSTEN MARTIN.
SHE IS CHAIR OF THE DEPARTMENT
OF STRATEGIC MANAGEMENT AND
PUBLIC POLICY AT GEORGE
WASHINGTON UNIVERSITY'S BUSINESS
SCHOOL AND AUTHOR SEVERAL

ARTICLES ARE ETHNICS AND
TECHNOLOGY.
IF YOU HAVE QUESTIONS FOR THE
PANEL, PLEASE WRITE THEM DOWN ON
THE QUESTION CARDS AND HAND THEM
TO FTC STAFF WHO WILL BE
CIRCULATING.
HOPEFULLY WE'LL HAVE TIME TO ASK
SOME OF THEM.
I'M GOING TO BEE GENERAL WITH
AQUAPHOR ASHISH.
YOUR PROGRAM AREA AT MICROSOFT
DEFENDING DEMOCRACY.
AT THE SAME TIME MICROSOFT
OFFERS A CUSTOM VOICE SERVICE
THAT CAN CREATE A VOICE CLONE.
CAN YOU DESCRIBE KIND OF WHAT
THE CUSTOM VOICE PROGRAM IS AND
THEN HOW YOU WENT A BOUT
CONSIDERING THE ETHICAL
IMPLICATIONS OF THIS TECHNOLOGY
AND HOW TO PREVENT ITS MISUSE.
>> WELL, THANK YOU, DAN.
I'LL START WITH MICROSOFT AS A
PLATFORM AND TOOLS COMPANY WE
HAD CUSTOM VOICE TECHNOLOGY WAY
BACK IN THE DAY ESSENTIALLY.
THE WINDOWS, WE DID HAVE THE
NARRATOR WHICH WAS MORE LIKE
MECHANICAL ROBOTIC VOICE.
BUT RECENTLY SINCE LAST MANY
YEARS WE HAVE, AS WE HEARD IN
THE EARLIER PANEL WE HAVE
FINESSED HOW REAL THESE VOICES
HAVE BECOME.
SYNTHETIC VOICES HOW REAL THEY
ARE.
LATELY FROM OUR TEXT TO SPEECH
SERVICE, WE HAVE BEEN THINKING
ABOUT HOW DO WE ACTUALLY CREATE
SOMETHING, BECAUSE AGAIN WE
LEARNED ABOUT A LOT OF GOOD USE
CASES AND THERE ARE, I'LL SAY,
95% OF THESE SYNTHETIC VOICE
DEFECTS ARE USED -- THEY ARE
POSITIVE USE CASES.

5% OF FAILURE, BUT AS WE STARTED
THINKING ABOUT WE THINK IN TWO
TERMS.
WHAT ARE THE POTENTIAL BENEFIT
OF ANY TECHNOLOGY THAT WE BRING
TO MARKET.
AND THEN BALANCE IT OUT WITH
RISK -- I'LL SAY RISK MODEL OF
POTENTIAL HARMS OF THAT
TECHNOLOGY CAN BRING.
BUT SOME MECHANISM TO REDUCE
RISK OF THE HARM.
SO, BENEFITS ON ONE END, WHICH
WE THINK THERE ARE HUGE NUMBER
OF THEM.
RISKS, ESSENTIALLY CREATE A
THREAT MODEL OR RISK THEN
GOVERNANCE MODEL WHERE WE SAY,
OKAY, WHAT CAN WE DO TO BRING
THE RISK DOWN.
WE LAUNCHED AS A SERVICE.
WHERE DOES THIS SERVICE LIVE?
IT LIVES IN AN ORGANIZATION IN
MICROSOFT CALLED COGNITIVE
SERVICES.
THERE'S A GROUP CALLED ETHICS
AND AI AND THAT'S ACTUALLY THE
GOVERNANCE TEAM OR ANY OF THOSE
COGNITIVE SERVICES.
SOME OF OUR FACE RECOGNITION
TECHNOLOGY ALSO SITS THERE,
VOICE TECHNOLOGY, KNOWLEDGE
TECHNOLOGY.
ESSENTIALLY MOSTLY AI
TECHNOLOGIES THAT CAN ENABLE SO
MANY SCENARIOS SIT WITH
GOVERNANCE MODEL WHICH WE CALL
ETHICS AND AI TEAM IN MICROSOFT.
WHEN WE LAUNCHED CUSTOM VOICE,
THAT WAS A TEAM THAT CAME UP
WITH WHAT WE CALL INTERNALLY A
HARMS FRAMEWORK, SAYING, HEY,
THESE ARE THE POTENTIAL MISUSES
OF THIS TECHNOLOGY.
GOVERNANCE MODEL IF THESE
POTENTIAL MISUSERS IF WE SEE ANY

OF THESE USED WE CAN CUT ACCESS
TO THE SERVICE.
AND THAT'S HOW WE STARTED
LOOKING AT THIS.
WE CREATED AN ENFORCEMENT
MECHANISM HOW DO WE CUT ACCESS
TO THIS.
COME UP WITH IF IT IS USED FOR
X, Y PURPOSES YOU CAN THINK OF
AS A CHECKLIST, AUTO SYSTEM
SHELLY.
WHICH A CUSTOMER CAN USE
TECHNOLOGY FOR ALL GOOD CASES
BUT IF THEY FALL IN ANY OF THOSE
HARMS FRAMEWORK CHECKLIST THEN
WE'LL CUT ACCESS.
AND THAT'S HOW WE ARE THINKING
ABOUT IT AS NOT ONLY THE
BALANCING OF POTENTIAL GOOD USE
CASES, BUT MANAGING IT WITH WHAT
KIND OF RISK FRAMEWORK THEN HAVE
A GOVERNANCE MODEL AROUND IT.
>> JUST SO I UNDERSTAND LET'S
SAY I WANTED TO TAKE ROSALYN'S
VOICE AND UPLOAD IT TO
MICROSOFT.
WOULD I BE ABLE TO CREATE A
VOICE CLONE OF HER VOICE?
>> SO ESSENTIALLY ONE THING THE
GOVERNANCE MODEL REQUIRES A
FRAMEWORK THAT WE CREATED, NOT
ONLY -- YOU CAN, IF YOU CAN SHOW
US THAT YOU ACTUALLY WOULD BE
USING THAT CUSTOM VOICE WITH
CONSENT OF ROSALYN IN THE RIGHT
WAY.
WITHIN THOSE PER VIEW OF THE
BENEFIT AND HARMS FRAMEWORK.
>> YOU HAVE METHOD FOR
DETERMINING THAT, I'M NOT REALLY
ROSALYN WHEN I SUBMIT THE VOICE?
>> SO AGAIN, YOU ARE A CUSTOMER
OF MICROSOFT, YOU ARE IN THE
BUSINESS OF CREATING SYNTHETIC
VOICE FOR WHATEVER REASONS, IT
COULD BE I WANT TO CREATE A BOT

SERVICE OR CREATING FOR ART AND
EXPRESSION OR INDIVIDUAL, GOOD
USE CASES SO YOU'RE A CUSTOMER.
AND ROSALYN IN THIS CASE IS THE
VOICE ACTOR THAT YOU WILL
ACTUALLY USE THAT SERVICE FOR.
YOU'LL UPLOAD IT SO YOU SIGN UP
LIKE WHAT WE CALL TERMS OF
SERVICE, A FRAMEWORK TERMS OF
SERVICE, UPLOAD THIS WITH THE
CONSENT FRAMEWORK F. WE SEE THAT
VOICE THAT WE GAVE YOU IS USED
FOR ANY NEFARIOUS PURPOSES WE'LL
CUT ACCESS TO IT.
>> DOES HA MAKE YOU SATISFIED?
>> NO.
HERE'S WHY.
I'M GOING TO TELL TWO QUICK
STORIES.
I'M A CUSTOMER OF SPRINT AND
LAST WEEK I WAS HAVING A TEXT
EXCHANGE WITH SOMEONE I REFER TO
AS MY WASBAND, WE WERE TALKING
ABOUT FINANCIAL MATTERS, SORT OF
PERSONAL.
AFTER FOUR DAYS, WHAT IS WRONG,
HE'S NOT RESPONDING, THIS IS
IMPORTANT.
A TEXT CAME IN THAT SAID, WHO
ARE YOU WRITING TO?
I DON'T THINK I'M THIS PERSON.
I THOUGHT, WHOA, WHO IS THIS?
WELL IT TURNS OUT WHO'S THIS WAS
JUST A THIRD PARTY THAT HAD
ACCESS TO OUR TEXTS.
SO AS SPRINT CUSTOMER I GOT ON
THE PHONE I SAID, HELLO, I'M
SCARED!
WHAT'S GOING ON, GET THIS PERSON
OFF MY PHONE.
THREE DAYS LATER THEY HAD NOT
BEEN ABLE TO RESOLVE THE
PROBLEM.
FINALLY, FINALLY I REALIZED
BECAUSE THE PERSON WAS NOT
INTENDING ANY HARM THEY JUST

HAPPENED TO BE SEEING ALL MY
TEXT MESSAGES BUT MY WAS --
WUSBAND TURNS OUT THAT PERSON'S
NUMBER, ENDED UP IN CONTACTS,
NOBODY KNOWS HOW.
THE STORY IS TOLD, BECAUSE
SPRINT INTENDED TO PROTECT ME.
I NEEDED PROTECTION, NOBODY
COULD FIGURE OUT HOW TO DO IT.
SO WHEN WE TALK ABOUT ETHICS
WE'RE LOOKING AT MINIMIZING
HARM, MAXIMIZING GOOD, GENERALLY
SPEAKING.
I'M NOT CONVINCED THAT WE HAVE
BEEN ABLE TO DESIGN INTO THE
SYSTEMS THEMSELVES THE CAPACITY
TO PROTECT THE INDIVIDUAL.
I THINK MICROSOFT IS A GREAT
COMPANY, DON'T GET ME WRONG.
BUT SOMETIMES I DON'T THINK WE
HAVE THE CONTROL WE BELIEVE WE
MIGHT.
OR THAT WE MIGHT WANT.
>> DO YOU THINK PART OF THE
ISSUE IS THAT IT'S HARD TO KNOW
WHAT THE HARMS ARE, SOMETIMES
THEY'RE UNANTICIPATED.
MAYBE THE HARM THAT -- THINKING
OF SECURITY OF THEIR TEXTING
SYSTEM THIS NEVER CROSSED THEIR
MIND THAT THIS WOULD HAPPEN.
>> WE ARE EYE INVOLVING,
TECHNOLOGY EVOLVES AS WELL AS
THE BAD ACTORS THINKING IS
INVOLVING.
I WOULD SAY EVEN THOUGH YOU CAN
CREATE THE CHECKLIST THAT WE
HAVE TODAY, BUT IT HAS TO KEEP
PACE WITH THE WAY WE SEE OTHER
THREATS COMING IN.
I THINK I CAN GO BACK TO WHAT
THEY WERE SAYING, THERE WILL BE
THING THAT WE DON'T ANTICIPATE
RIGHT NOW IN TERMS OF HARMS
CUSTOM VOICE CAN CREATE WE HAVE
TO MAKE SURE THAT WE ARE KEEPING

PACE WITH SOCIETY, WITH
TECHNOLOGIES AND KEEP ON
EVOLVING OUR FRAME WORKS OF
GOVERNANCE AS WELL AS HARMS THAT
WE HAVE.
>> CAN I JUST --
SURE.
>> I'M CURIOUS YOU MENTIONED
ASHISH THE PROGRAM IS CALLED
DEFENDING DEMOCRACY?
YOU HAVE MADE A POLITICAL CHOICE
AS A FRAMEWORK FOR YOUR
TECHNOLOGY, YOUR INVESTIGATIONS.
WHY DEMOCRACY?
>> AGAIN, I AM PART OF DEFENDING
DEMOCRACY WE THINK DEFECTS ARE
ONE OF THE EMERGING THREATS TO
DEMOCRACY BUT TALKING ABOUT
VOICE CLONING WHICH FOR
MICROSOFT AS I SAID THE CREATION
OF IT SERVES UNDER COGNITION
SERVICES, GOFF THAN TEAMS CALLED
ETHICS AND AI I'M JUST
REPRESENTING TALKING ABOUT THAT
PIECE.
NOW FROM THAT PERSPECTIVE FROM
DEFENDING DEMOCRACY PERSPECTIVE,
WE HAVE BEEN LOOKING INTO THIS
ISSUE AREA AS I SAID, WE THINK
THIS WHOLE IDEA OF DEFECT BOTH
FROM PERSPECTIVE OF VOICE OR
AUDIO OR AUDIO VOICE AS WELL AS
VIDEO, ONE OF THE GOOD USE CASES
FOR DISRUPTION ACTUALLY IS
DEMOCRACY, WHERE YOU CAN HAVE
SOCIETAL DISCORDS AND ALL KINDS
OF ARMS THAT CAN BE CREATED BY
DISRUPTING, USING DEFECTS AS
BIGGEST TECHNIQUE FOR THIS
INFORMATION.
>> SO YOU HAVE THIS ETHICS AND
AI GROUP.
ARE THEY INDEPENDENT FROM THE
DEVELOPERS OF THE PROGRAM SO
THAT THEY CAN ASSESS THE HARMS
AND --

>> YOU CAN THINK OF THEM AS GOVERNANCE BODY.
WE HAVE THIS RESPONSIBLE AI, IF YOU LOOK INTO THE ORGANIZATION STRUCTURE OF MICROSOFT, OUR LEGAL HEAD IS PRESIDENT BRAD SMITH.
SO ALL THESE RESPONSIBLE AI ORGANIZATION AND OUR ORGANIZATION SITS UNDER BRAD VERSUS ENGINEERING LEADERS, WE HAVE AI ENGINEERING, WE HAVE CLOUD ENGINEERING LEADER AND THEY SIT -- THEY ALL BUT WE ACTUALLY, YOU CAN THINK OF AS A GOVERNANCE BODY IN OUR ORGANIZATION NOT IN ENGINEERING ORGANIZATIONS.
>> LET'S SAY THE ETHICS AND AI GROUP IS DOING THIS INDEPENDENT REVIEW OF POTENTIAL PROGRAM, WHEN DOES THAT OCCUR?
IS IT WHEN THERE'S A GLIMMER OF HOPE ABOUT A PROGRAM OR DURING --
>> WE HAVE LEARNED FROM -- BACK IN THE DAY, IT WAS I THINK BACK IN EARLY 2000S IN FACT WHEN WE ANNOUNCED THIS PROGRAM CALLED TRUSTWORTHY COMPUTING.
SECURITY IS NOT AN ADD ON, IT HAS TO START AS YOU START ENVISIONING A SOLUTION.
WE ARE TAKING SIMILAR KIND OF 'MILOSEVIC ON RESPONSIBLE AI WHERE WE DON'T THINK THAT RESPONSIBLE AI ACTUALLY IS ONCE THE PRODUCT IS DONE WE HAVE CHECKLIST WE GO THROUGH IT.
WE WANT TO MAKE SURE THAT IT IS PART OF THE CULTURE AS WE START ENVISIONING A SOLUTION, ESPECIALLY AROUND AI THINK ABOUT WHAT THE POTENTIAL IMPLICATIONS, WHAT OUR CHOICES WE MAKE AT DEIS IN WOULD BE.

>> KIRSTEN, DOES THAT TIMING
MAKE SENSE THAT YOU START
THINKING ABOUT ETHICS FROM DAY
ONE OR EARLIER?
>> I THINK TWO THINGS ABOUT WHAT
YOU SAID THAT MAKES SENSE.
ONE IS THE IDEA THAT THESE
VALUE-LADEN DECISIONS ARE MADE
IN DESIGN.
THEY'RE ACTUALLY BEING MADE, ONE
THEORY THEY'RE ACTUALLY BEING
MADE WHETHER YOU TALK ABOUT THEM
OR NOT.
SO YOU CAN DO THEM WELL OR HAVE
THEM BE COMPLETELY UNGOVERNED SO
THE IDEA IS THAT A COMPUTER
SCIENTIST, SOMEONE WHO IS
DEVELOPING AI OR ANY OTHER TYPE
OF COMPUTER PROGRAM, IS MAKING
DECISIONS ABOUT WHAT'S
EFFECTIVE, HOW IT WILL WORK,
WHAT'S THE RIGHT FACTORS, WHAT
SHOULD WE BE INCLUDED IN THIS,
WHO SHOULD I BE CONSIDERING.
AND MAKING THOSE DECISIONS IN
THEIR DESIGN THROUGHOUT DESIGN
AND DEVELOPMENT.
THIS WAY MICROSOFT IS SAYING,
LET ME HELP WITH YOU THAT AND
LIEUTENANTS THINK ABOUT THIS IN
A MORE SYSTEMATIC WAY.
THE IDEA THE COMPUTER PROGRAMS
ARE DOING IT, MICROSOFT IS
ACTUALLY SAYING I'D LIKE TO HELP
WITH YOU THAT BECAUSE WE THINK
IF WE DO -- DON'T TALK ABOUT IT
YOU'RE GOING TO DO IT IN
UNGOVERNABLE WAY.
THE FACT THAT IT'S EARLY ONLY
TIME TO DO IT BECAUSE ONCE
TECHNOLOGY IS OUT IN THE WILD
IT'S MUCH MORE DIFFICULT TO
RECONSTRUCT AND CHANGE ONCE IT'S
OUT BEING USED.
FOR BOTH OF THOSE REASONS, I
THINK THE OTHER THING THAT WAS

INTERESTING THAT WAS MENTIONED
IS THAT THERE'S THIS IDEA OF
WHAT TYPE OF RELATIONSHIP DO YOU
HAVE WITH THE CONSUMER OR USER
AND TECHNOLOGY WHEN IT'S OUT AND
WHEN YOU'VE SOLD IT.
SO A LOT OF TIMES COMPANIES WILL
SEND SOMETHING OUT THEN KIND OF
GO LIKE THIS SAY YOU USE IT
HOWEVER YOU WANT.
I DON'T WANT TO BE INVOLVED.
IT WAS INTERESTING WHAT I PICKED
UP ON WAS THIS IS -- IT'S
ACTUALLY HOUSED AT MICROSOFT AND
A SUBTLE DIFFERENCE OF SAYING
YOU HAVE TO COME TO ME EVERY
TIME YOU WANT TO USE IT AND I
DECIDE USE IT CORRECTLY OR NOT.
IT'S ACTUALLY MUCH MORE TIME
INTENSIVE FOR MICROSOFT TO DO IT
THAT WAY.
THEY ARE CONSTANTLY INVOLVED IN
THE ONGOING USE, THEY CAN KEEP
TRACK OF WHEN IN PRACTICE IS
BEING USED IN A NOVEL WAY THAT'S
NOT APPROPRIATE.
WHERE AS IF THEY JUST SOLD IT
THEY WOULD JUST HAVE ALMOST LIKE
PRODUCT LIABILITY TYPE APPROACH
THEY CAN SAY, IT WASN'T MY
FAULT.
THEY ARE TAKING MORE OWNERSHIP.
>> JUST ADD TO THAT, THAT IS
POSSIBLE, BY THE WAY, BECAUSE OF
CLOUD SERVICES.
BECAUSE WE ARE THE CLOUD WINDOW,
WE HOST SOME OF OUR SERVICES, WE
ALSO ENABLE PARTNERS TO OUR
SERVICES SO EVEN THEY HAVE
CAPABILITY TO CREATE -- AS THEY
START CREATING THE SERVICES AND
SELL TO CUSTOMERS.
THEY CAN TAKE LEARNING FROM US
AS WE DO, IT IS ALL POSSIBLE
BECAUSE OF THE CLOUD.
BUT A CUSTOMER CAN ALSO NOW

START THINKING ABOUT RESPONSIBLE
AI AND GOVERNANCE MODEL BECAUSE
THEY ALSO ACTED ALSO IN THE
CLOUD.
>> PART OF THE REASONS, THE
COMPUTING POWER IS NEEDED FOR
THE CLOUD IN ORDER TO ENGAGE IN
THIS SORT OF VOICE CLONING IS
THAT RIGHT?
>> AGAIN, WE CAN GO BACK TO
THIS, IT'S LIKE THE WHOLE IDEA
OF -- FOR FIVE YEARS WHY
EVERYTHING IS SO ON AN
EXPONENTIAL GROWTH IT'S A
PERFECT STORM OF TECHNOLOGY WITH
COMPUTING POWER, YOU HAVE
RESEARCH AND THEN YOU HAVE
COMPUTE, YOU CAN TALK ABOUT THE
TECHNOLOGY ACCESS AND THE
DEMOCRATIZATION OF TECHNOLOGY
AND GETTING VERY QUICKLY WITHOUT
PUTTING A LOT OF RESEARCH INTO
BUYING HARDWARE AT ONCE.
SO ALL OF THAT CONFLUENCE
ACTUALLY ENDED UP WHERE WE
ARE -- THE PACE OF INNOVATION
WHICH ALSO GIVES US NEW BUSINESS
MODELS BUT ALSO SOME ADVANTAGE
IN TERMS OF GOVERNANCE MODELS IF
YOU WANT TO START THINKING THAT
WAY.
>> WHEN YOU AI FOLKS ARE LOOKING
AT THE HARMS AND BENEFITS, IS
THERE LIKE A LIST MADE OF EACH?
IS IT LIKE, SHOULD I GO OUT WITH
HER, LIKE THAT, PROS AND CONS OR
AND HOW IS VALUE PUT ON EACH
HARM.
>> YOU'RE RIGHT.
AGAIN, I DON'T KNOW THE DETAILS
AT THIS POINT.
WHAT I KNOW IS THAT ESPECIALLY
FOR CUSTOM VOICE, WE MADE SURE
THAT BEFORE WE EVEN TALK ABOUT
RELEASING IT, WE ACTUALLY GO
THROUGH THIS HARMS FRAMEWORK

THAT WAS CREATED.
I DON'T KNOW HOW VALUE IS PUT IN
ON ALL THAT I'M NOT THE RIGHT
PERSON TO ANSWER THAT.
BUT I AM THE RIGHT PERSON TO SAY
THAT WE ARE THINKING ABOUT IT
FROM THE GET GO OF HOW WE
DESIGN.
HOW WE DEVELOP AND DEPLOY
TECHNOLOGY.
>> I GUESS SOMEBODY HAS THE
AUTHORITY TO SAY WE'RE NOT GOING
TO ACTUALLY RELEASE THIS OR
WE'RE GOING TO PUT CONTROLS ON
IT.
>> IN FACT, IF YOU READ THE BLOG
FROM ALEX SHIPMAN WHO WAS
TECHNICAL FELLOW WHO OWNS THIS
CUSTOM VOICE TECHNOLOGY, HIS
BLOG WAS ALL ABOUT THAT.
HOW WE WOULD EVEN ABOUT
ANNOUNCING IT.
YES, A CONSIDERATION FROM
MICROSOFT'S POINT OF VIEW.
THERE IS ABILITY TO SAY, NO, IF
THINK HARM FRAMEWORK OR
POTENTIAL HARMS ACTUALLY
OUTWEIGH THE BENEFITS OF
TECHNOLOGY.
>> I WAS WATCHING THE FIRST
PANEL THINKING ABOUT THE VALUE
OF THE BENEFITS OF VOICE
TECHNOLOGY.
AND SEEING THE VIDEO OF PERSON
WITH ALS AND THEN HEARING THE
DOOM AND GLOOM ABOUT ALL THE
POSSIBLE MISUSES.
IS IT ACTUALLY POSSIBLE TO PUT
DOLLAR VALUES ON SOME OF THIS?
CERTAINLY LIKE THE AMOUNT OF
FRAUD YOU CAN TALLY SOMETHING
UP.
HOW DO YOU ENGAGE IN THIS SORT
OF BALANCE.
>> PRICE EMOTIONS?
>> WELL, EMOTIONS, YEAH, THAT --

THAT IS, RIGHT.
THIS IS FAMILY LOVE THAT YOU
WANT TO -- CONTINUE INTO THE
INDEFINITE FUTURE AND CAN YOU
PRICE THAT.
>> CAN YOU?
>> I'M CONFUSED ABOUT THE
QUESTION.
FOR WHAT PURPOSE WOULD YOU WANT
TO PUT A MONETARY VALUE ON ANY
OF THIS?
>> THAT IS A VERY GOOD QUESTION.
BUT I GUESS THE IDEA IS THAT IF
A BUSINESS IS COMING OUT WITH
VOICE CLONING ENGAGING ANY SORT
OF BALANCING OF HARMS AND
BENEFITS THAT'S GOING TO
DETERMINE WHETHER YOU KNOCK THEM
OUT WITH TECHNOLOGY OR PUT
CONTROLS ON TECHNOLOGY THAT
THOSE CONTROLS HAVE CERTAIN
PRICE, HOW DO YOU ENGAGE IN THAT
BALANCE?
>> MY GUESS WOULD BE, SO WHERE
IS JOHN AT THE MOMENT?
SO, MY GUESS WOULD BE THAT THE
PEOPLE WITH WHOM YOU WORK FEEL
THAT ANY PRICE IS THE
APPROPRIATE PRICE FOR THE
BENEFIT THAT YOUR TECHNOLOGIES
WOULD GIVE THEM.
WHETHER THEY COULD ACTUALLY PAY
THAT PRICE IS ANOTHER MATTER.
BUT IT IS BEYOND PRICE, IS IT
NOT?
TO HAVE THAT LOVED ONE CONTINUE
TO EXIST AT LEAST IN AUDIBLE
FORM?
>> HOLD ON ONE SECOND WE'LL GET
YOU THE MICROPHONE.
>> I'M HAPPY TO SAY THAT A
DOLLAR AMOUNT HAS NOT EVER BEEN
PART OF A DISCUSSION WITH ANY OF
THE PATIENTS OR FAMILIES.
THE EMOTION, AS YOU POINTED OUT,
IS KEY.

BUT WE HAVE NOT EVER NEEDED TO
OR NOR HAS ANY PATIENT TO MY
KNOWLEDGE MONETIZED WHAT THAT
VALUE IS TO THEM.
IT'S MORE THE EMOTIONAL
CONNECTEDNESS.
>> RIGHT.
I THINK THAT IS A BEGINNING OF
AN ANSWER.
BUT YOU'RE ALSO ASKING, THIS IS
TO THIS SITUATION, HOW DO YOU
PRICE THE TECHNOLOGY THAT YOU
PUT ON THE MARKET GIVEN THIS
REALITY.
THAT IS FOR SOME SEGMENT OF THE
POPULATION YOU DON'T EVEN WANT
TO TALK ABOUT PRICE.
AND YET YOU KNOW AS YOU'RE
PRODUCING THE ITEM THAT IT DOES
COST.
AND THAT YOU CAN LEGITIMATELY AS
A BUSINESS FACTOR COST INTO YOUR
CALCULATIONS FOR THE WAYS IN
WHICH YOU PROVIDE THAT
TECHNOLOGY, RIGHT?
THE QUESTION DOESN'T GO TO JOHN
BECAUSE HE'S DEALING WITH THE
END USER, BUT TO YOU WHO PRODUCE
THE ITEM WITH A VERY CLEAR
INTENT TO MAKE A GOING CONCERN
OF YOUR COMPANY.
SO COST DOES MATTER.
>> DAN, YOU WERE ALSO TALKING
ABOUT WHAT COST WOULD YOU PUT TO
THE MONETIZATION OF THE HARMS AS
WELL, RIGHT?
LIKE THAT WAS --
>> SURE.
>> I THINK SO THAT IS -- I DON'T
KNOW WHAT THE FRAMEWORK IS BUT
LIKE DOING COST BENEFIT ANALYSIS
IN UNDERSTANDING ETHICS, I MEAN
THIS IS JUST THE QUINTESSENTIAL
PROBLEM OF WHOSE HARMS COUNT,
HOW DO YOU QUANTIFY THOSE HARMS
AND COST BENEFIT ANALYSIS IS

ALWAYS WILLING GO TO SACK IF I
CRIES MINORITY OR MAJORITY.
THAT IS JUST HOW IT WORKS.
YOU'RE SUMMING SOMETHING UP.
SO SOMETIMES -- BUT YOU CAN ALSO
DO OTHER WAYS AS LONG AS YOU'RE
NOT LOOKING FOR A PURELY
NUMERICAL VALUE.
I WOULD SAY THIS IS NOT -- FOR
EXAMPLE, I MEAN, THERE ARE
CLASSIC AM COMPELS WHERE THE
COST BENEFIT ANALYSIS WENT AWRY
SO ORANGE O RINGS ON THE
SHUTTLE, THE PINTO CASE, WE PUT
NUMBER VALUE, PERHAPS SAY IT
WASN'T THE CORRECT NUMBER VALUE
WE DID A COST BENEFIT ANALYSIS
AS CORPORATION, WE CAME OUT
INCORRECTLY.
THAT WASN'T IN THE LINE DID NOT
SURVIVE.
BUT THIS IS NOT UNIQUE TO
TECHNOLOGY.
YOU CAN SAY THE SAME THING ABOUT
LIKE, DO YOU AMY LOU A
PHARMACEUTICAL COMPANY TO CREATE
PILLS WHERE WE KNOW THAT A
CERTAIN PORTION OF THEM ARE
GOING TO BE USED ILLICIT TEE.
THAT THEY THEN ARE SHODDY ON
GOVERNANCE MECHANISMS TO
IDENTIFY HARMFUL USES THEY SAVE
MONEY BY DOING THAT BUT CREATED
SOME SORT OF SYSTEMATIC HARM.
THAT'S NOT TECHNOLOGY.
THAT'S JUST A PAIN PILL.
WE HAVE THE SAME -- THIS -- THE
IDEA OF LIKE A NEW PRODUCT
HAVING HUGE BENEFITS UPSIDE
COULD BE USED IN NEFARIOUS WAYS.
THE SECONDARY MISUSE OF PRODUCTS
IS NOT UNIQUE TO TECHNOLOGY.
NOT JUST -- WE HAD THIS THIS
WITH OPIOID.
>> THAT'S A GREAT POINT.
MAYBE TALK A BIT ABOUT -- HOW

HAVE OTHER TECHNOLOGIES WHEN
THEY HAVE BEEN INTRODUCED HOW
HAVE PEOPLE DEALT WITH THE
ETHICS ISSUES THERE?
>> I'M THINKING EARLY ON IN THE
DEVELOPMENT OF THE MECHANISMS
THAT MADE POSSIBLE, KIDNEY
TRANSPLANTS, VERY EARLY ON THEY
ACTUALLY CALCULATED THE VALUE OF
INDIVIDUAL LIVES TO DETERMINE
WHO SHOULD GET THE TRANSPLANTED.
THERE WAS A FORMULA TO FIGURE
OUT WHO WARRANTS THIS, WHO IS A
GOOD PERSON.
AND THEN THEY DETERMINED
ULTIMATELY WHEN THE PRINCIPLES
OF BIOETHICS CAME FORWARD THAT
YOU CAN'T DO THAT.
YOU GET A NUMBER YOU GET IN
LINE.
THAT'S THE END OF THE STORY.
WE DON'T HAVE THE CAPACITY TO
DETERMINE WHO IS MORE VALUABLE
THAN ANYBODY ELSE.
WITH OTHER MORE RECENT
TECHNOLOGIES I'M THINKING ABOUT
HOW THE MARKET DRIVERS FOR
COMPETITION SEEM TO SOMETIMES
PUSH THEM OUT QUICKER THAN WE
HAVE THE CAPACITY TO THINK ABOUT
THE ETHICAL IMPLICATIONS OF
THEM.
WE'RE TRYING TO COMPETE,
INNOVATE, MEANWHILE WE'RE HAVING
THESE CONVERSATIONS BUT IT'S A
LITTLE TOO LATE TO PUT IT INTO
THE ACTUAL DESIGN.
I DON'T KNOW IF IT ANSWERS YOUR
QUESTION BUT I'M TRYING TO THINK
OF CASE WHERE WE'RE ACTUALLY
ABLE TO METHODICALLY, CAREFULLY
BUILD IN CONCERN R CONCERN FOR
ETHICS BEFORE THESE THINGS GET
RELEASED.
THE OTHER PROBLEM IS WE'RE
TALKING ABOUT MICRO ETHICS,

MACRO ETHICS, TALKING ABOUT
INDIVIDUAL HARM WITHOUT THINKING
ABOUT BROADER, LARGER SCALE
QUESTIONS, FOR EXAMPLE, WHAT
DOES IT MEAN TO BE HUMAN.
WHAT DOES VOICE HAVE TO DO WITH
THAT SENSE OF IDENTITY.
>> IF WE TOOK IT TO THAT MACRO
LEVEL, A LITTLE BIT IN FROM THAT
MACRO LEVEL WE'D SAY, IF YOU
WERE DOING SOME SORT OF COST
BENEFIT ANALYSIS HERE, WHAT IS
THE VALUE OF LOSING THE ABILITY
TO KNOW THAT THE PERSON ON THE
OTHER END OF THE PHONE IS REALLY
WHO THEY SAY THEY ARE.
>> OR WHO THEY ARE OR WHAT IT
MEANS TO BE THAT PERSON.
A DISEMBODIED VOICE LEADS ME TO
HAVE TO RECONSIDER WHO AM I AND
WHO AM I TO YOU AND WHAT IS THIS
OBJECT IN OUR RELATIONSHIP, WHAT
DOES IT MEAN.
HOW DOES IT CHANGE THAT
RELATIONSHIP IN MY SENSE OF
SELF.
WHAT HAPPENS IF IT BREAKS, DO I
DISAPPEAR WHEN IT NO LONGER
FUNCTIONS?
THOSE KINDS OF QUESTIONS.
>> I THINK WE'RE VERY USED TO
BALANCING PROS AND CONS HOW WE
OFTEN THINK ABOUT THINGS.
WHAT IS THE ALTERNATIVE MODEL
THAT A BUSINESS SHOULD BE USING
WHEN IT'S THINKING HOW IT SHOULD
ACT ETHICALLY?
>> I JUST WANT TO -- THIS IS A
GREAT QUESTION.
AS ORGANIZATION, THIS IS
SOMETHING THAT WE TAKE PRIDE IN
AS WELL AS WE CAME UP AS A
COMPANY, AS WE START BUILDING
TECHNOLOGY, AS WE DEVELOP AS
DISTRIBUTE IT, THE CORE AI
PRINCIPLES THAT WILL WORK WITH,

WITHIN THE FRAMEWORK.
IT INCLUDES INCLUSIVENESS, THE
TECHNOLOGY THAT WE WILL BUILD
HAS TO BE INCLUSIVE IT CANNOT
EXCLUDE, WHATEVER POPULATION.
THAT IS FAIRNESS.
TRANSPARENCY, EVEN IN CUSTOM
NEURAL VOICE, IN THIS CASE,
RIGHT, YOUR EXAMPLE YOU HAVE TO
BE TRANSPARENT THAT IT'S
ROSALYN'S VOICE.
THOSE KIND OF THINGS.
HAVE TO BE BUILT INTO THE MODEL,
ACCOUNTABILITY, PRIVACY AND
SECURITY AND RELIABILITY AND
SAFETY.
IF YOU THINK ABOUT ANY
TECHNOLOGY THAT WE BUILD IF
THERE ARE CORE BRONZE THAT THEY
HAVE TO ADHERE TO THE SIX
PRINCIPLES OF INCLUSIVE,
NEFARIOUS.
RELIABILITY AND SAFETY.
THEN THAT IS ALL -- IDENTIFY ANY
OF THOSE BREACHES, WE HAVE THE
CAPABILITY TO SAY NO.
THAT IS HOW WE THINK ABOUTF A
BIGGER SOCIETAL VALUE THAN
INDIVIDUAL HARM WHATEVER WE ARE
DOING, AS A GOOD CUSTODIAN, AS A
BIG TECHNOLOGY COMPANY, WHAT IS
OUR PLACE.
AND HOW DO WE ACTUALLY USE THAT
TO CREATE HARMONY.
THAT'S WHAT IT IS.
>> SHOULD BE SAID, TOO, THAT THE
WAY ASHISH DESCRIBES WHAT'S
HAPPENING AT MICROSOFT AT LEAST
FOR THIS PROJECT RESPONSE TO
WHAT YOU WERE SAYING ROSALYN,
THIS IS NOT LAGGING.
THIS IS BUILDING.
THEY ARE, YOU ARE BUILDING
ETHICAL CONCERNS INTO THE
PROCESS OF DEVELOPMENT FROM THE
GET GO.

>> IT'S TRUE, I HEARD THAT.
THERE'S ALSO GOVERNANCE WHICH IS
SORT OF MORE THEORETICAL.
BUT THAT DEEPEN GAUGEMENT OF THE
PUBLIC FROM THE BEGINNING, FROM
THE EARLY DEVELOPMENT OF SOME OF
THESE TECHNOLOGIES, PRACTICALLY
SPEAKING I DON'T KNOW HOW THAT
PLAYS OUT BUT CERTAINLY SEE IT
IN THE LITERATURE.
>> FOR PURPOSES OF THIS PANEL
WE'RE GOING TO HOLD MICROSOFT
OUT AS THE MODEL IN HOW YOU'RE
SUPPOSED TO DO THINGS WHEN IT
COMES TO ACTING EAT I CAN --
ETHICALLY DEALING WITH AI AND
VOICE CLONING AND RELATED
ISSUES.
IS IT REASONABLE TO EXPECT
SMALLER COMPANIES WHO HAVE ONLY
ONE PRODUCT IN DEVELOPMENT,
VOICE CLONING, TO ENGAGE IN THIS
SORT OF REALLY THOUGHTFUL
ANALYSIS OF THEIR PRODUCT
PIPELINE?
>> YES.
I THINK IT IS.
I WOULD EXPECT WHOEVER RAN THAT
BUSINESS TO FEEL THAT HE, SHE,
THEY IN FACT NEEDED TO HOLD
THEIR OPERATIONS AND THEIR
ENTITY TO A HIGH STANDARD.
ABSOLUTELY.
>> ASHISH, DO YOU KNOW WHETHER
OR NOT THERE IS ANYTHING OUTSIDE
OF MICROSOFT, LIKE A BROADER
CONSORTIUM OF COMPANIES THAT HAS
SIMILAR POLICIES ABOUT WHAT
CONSTITUTES RESPONSIBLE AI?
>> I DON'T KNOW IF THERE IS LIKE
AN INDUSTRY GROUP WHICH IS
ACTUALLY -- THERE IS A LOT OF
THINKING AND I'M NOT -- YES, WE
ARE ACTUALLY VERY CONSCIOUSLY
TALKING ABOUT AI AND ETHICS
PRINCIPLES.

THE OTHER BIG THING, I'VE
ACTUALLY HEARD AND I DON'T KNOW
HOW IT IS IMPLEMENTED, GOOGLE
ALSO ACTUALLY TRIED TO OR HAS AN
AI PRINCIPLES THERE.
SO, I DON'T KNOW ABOUT THE
CONSORTIUM BUT I KNOW THAT OTHER
BIG TECH COMPANIES ARE ALSO
THINKING -- AGAIN THE THING THAT
I WANT TO EMPHASIZE HERE IS
THAT, TRUST.
IS A BIG CURRENCY.
PURPOSE, AS THE NEW GENERATION
MILLENNIALS COME NOT ONLY THE
WORKFORCE BUT IN THE CONSUMER
BASE, THOSE BECOME THE HIGH
VALUE CURRENCY THAT WE AS
BUSINESSES HAVE TO SUPPORT.
IT'S THE RIGHT THING TO DO ALSO
THE MARKET DEMANDING AT THIS
POINT IN TIME.
TO YOUR PREVIOUS QUESTION IS,
EVEN IF IT'S SMALL COMPANY, I
THINK THE CUSTOMERS OF THAT
SMALL COMPANY ACTUALLY WANT THAT
COMPANY.
FOR A LONG TERM SUCCESS OF THAT
COMPANY.
SHORT TERM, FLY BY NIGHT
OPERATORS, YES, IT MAY WORK.
ANY SMALL COMPANY DO NOT THINK
ABOUT THE CONSCIENCE OF A
CUSTOMER AS CUSTOMER DEMANDED
TODAY, IT'S VERY SHORT.
>> ROSALYN, YOU TEACH ETHICS TO
COMPUTER SCIENCE UNDERGRADUATE
STUDENTS.
>> ALL ENGINEERS.
>> HOW DO YOU GET THEM TO
RECOGNIZE WHEN THEY LEAVE IVA
AND THEY GO WORK FOR A BUSINESS
THAT THEY HAVE KIND OF A
HIGHER -- HIGH RESPONSIBILITY
THAN JUST THE BOTTOM LINE OF THE
COMPANY?
>> WELL, THEY MIGHT LEAVE WITH

THAT UNDERSTANDING BUT DEPENDING ON THE CULTURE OF THE COMPANY THEY MAY NOT OR MAY BE ABLE TO EXERCISE THAT UNDERSTANDING.

I HAVE AN EXAM PER OF STUDENT WHO WROTE LAST YEAR AND SAID, I WORK FOR THIS CONSULTING FIRM THEY PUT ME ON X, Y, Z, CLIENT I HAD TO DECLINE, WHEN WE TALKED ABOUT THIS, I DON'T FEEL LIKE I CAN DO THIS WITH FULL CONSCIOUSNESS.

SHE PAID A PRICE BECAUSE SHE GOT REASSIGNED TO A SCHOOL, WHICH WASN'T IN HER AREA OR WHERE -- ASPIRED.

SHE HAD TO DECIDE, OKAY, THAT'S ALL RIGHT.

I'M JUST GOING TO GO WHERE I'M TOLD BECAUSE I'M REALLY NOT GOING TO BE ABLE TO GO HERE ETHICALLY.

DEPENDING ON THE ORGANIZATION, DEPEND CAN ON THE CULTURE AND COMMITMENT, SOMETIMES CERTAINLY NOT IN THE CASE OF MICROSOFT, BUT SOMETIMES THE ETHICS CLAIMS ARE REALLY MORE ABOUT HOW DO YOU SAY IT, IMAGE.

THAN INTERNAL TO THE ORGANIZATION WHERE THEY ARE ACTUALLY ARE NOT REAL TRUE CHECKS AND BALANCES AND METRICS TO KNOW THAT'S THE CASE.

IT DEPENDS ON WHERE THEY GO.

THEY GET THEIR ASSIGNMENT AND GET THEIR CHECK AND THEY'RE HAPPY BECAUSE THEY'RE WELL PAID AND GOING TO GET A NEW HOUSE THEN ETHICS SORT OF FADES.

>> DO ANY OF YOU KNOW WHETHER OR NOT BUSINESSES THAT HAVE LOFTY ETHICAL STATEMENTS ACTUALLY PROVIDE PROTECTIONS FOR THEIR EMPLOYEES WHO MIGHT HAVE AN OBJECTION LIKE THAT?

>> SO EMPLOYEES CURRENT
EMPLOYEES.
>> LET'S TAKE EXAMPLE OF ROSALYN -- WHISTLE-BLOWER.
>> THERE WAS NO WHISTLEBLOWING
HERE SHE DIDN'T HAVE THE -- IT
WAS ASSAULT TO HER MORAL SENSE
OF MORAL COMMITMENTS.
>> I WONDER WHETHER SHE COULDN'T
HAVE CONVERTED THAT INTO -- I
DON'T KNOW WHAT SHE WAS BEING
ASKED TO DO.
>> IT WAS ABOUT -- A CLIENT
THAT'S BRINGING THE PIPELINE IN
THROUGH IMMINENT DOMAIN, SOME
NATURAL WILDERNESS IN PEOPLE'S
FARMS.
IN MY CLASS THEY WENT OUT AND
TALKED TO THOSE FARMERS AND THEY
LEARNED ABOUT WHAT WAS AT STAKE.
SHE SAID, I KNOW ABOUT THE
PEOPLE CONNECTED TO THIS PROJECT
AND THEIR LIVES, I DON'T WANT TO
WORK ON IT.
SO THERE WAS NOTHING ILLEGAL,
NOTHING EGREGIOUS OR WRONG SHE
JUST SAID, I KNOW TOO MUCH ABOUT
SOCIAL IMPLICATIONS OF THIS
PROJECT TO FEEL GOOD ABOUT IT.
WHAT SHOULD SHE DO?
SOCIAL IMPLICATIONS TO VOICE
CLONING TECHNOLOGY.
>> ASHISH, DOES MICROSOFT HAVE
ANY WHISTLE-BLOWER TYPE
PROTECTIONS?
>> WE HAVE ACTUALLY NOT ONLY
HAVE ALL PROCESSES IN PLACE, WE
ARE VERY ROBUST MECHANISMS
INTERNALLY AS WELL.
GO THROUGH ALL KIND OF TRAININGS
AS WELL.
FOR PEOPLE WHO MAY NOT KNOW
WHERE TO GO TO, IF THEY FIND
THEMSELVES IN THESE KIND OF
SITUATIONS.
WE MAKE IT VERY ACCESSIBLE
INTERNALLY IN MICROSOFT BOTH

FROM MANAGEMENT PERSPECTIVE BUT
HR AS WELL.
I THINK WE ARE DOING PRETTY
GOOD.
>> ENGINEERING STUDENT, UNDER
GRAD GOES TO YOUR ETHICS CLASS
THEY GO OFF INTO THE BUSINESS
WORLD.
HOW OFTEN DO BUSINESSES BRING IN
ETHICISTS TO REFRESH THE
TRAINING FOR THEIR EMPLOYEES.
>> THERE ARE COMPLIANCE
OFFICERS, WHICH IS A LITTLE
DIFFERENT.
THAT'S MORE ABOUT THINGS LIKE
BREAKING THE LAW,
MISREPRESENTING THE COMPANY,
MISAPPROPRIATING FUNDS, ET
CETERA.
I CERTAINLY HAVE COMMUNICATION
WITH PEOPLE WHO ARE IN
COMPLIANCE WHO ARE INTERESTED IN
THAT.
THAT'S GOOD.
I I IT MAY NOT GET TO SOME OF
WHAT WE'RE CONCERNED ABOUT HERE
EBB TERMS OF A COMMITMENT TO A
HIGHER ORDER OF VALUES AROUND
GREATER PROTECTIONS FOR
INDIVIDUALS AND SOCIETY.
>> SOME OF THIS I THINK IS ALSO
TERMINOLOGY.
THAT IS A LOT OF WHAT HAPPENS
NOW UNDER THE RUBRIC OF
SUSTAINABILITY, I THINK WITHIN
LIMITS SOME OF WHAT HAPPENS IN
THE RUBRIC OF LEADERSHIP IS IN
FACT AN ATTEMPT TO GET AT
ETHICAL BEHAVIOR.
BOTH OF THOSE TERMS AT THIS
POINT ANYWAY HAVE THE VALUE THAT
WE DON'T THINK OF THEM AS
PUNITIVE.
WHERE AS ETHICS ALWAYS COMES
WITH THE CONNOTATION OF, YOU DID
SOMETHING WRONG, NOT NECESSARILY

ILLEGAL BUT SOMETHING THAT FEELS
INAPPROPRIATE.
GIVEN THE CULTURE, WHETHER IT'S
THE CULTURE WITHIN ORGANIZATION
OR IN THE SOCIETY AT LARGE.
AND SO IT'S HARD TO SELL ETHICS
IN THE FACE OF THAT SENSE THAT
THERE'S A VIOLATION FOR WHICH
YOU WILL BE PUNISHED.
AND WHEN WE TALK ABOUT
LEADERSHIP WE TALK ABOUT
SUSTAINABILITY.
WE CAN TALK ABOUT MILLENNIAL
VALUES ALL THAT HAVE FEELS
SOMEHOW MORE EXPANSIVE AND --
>> SOCIAL RESPONSIBILITIES AND
OTHERS.
>> ABSOLUTELY.
BUT ALL OF THAT TO YOUR
QUESTION, DAN, IN FACT ALLOWS
COMPANIES TO ADDRESS ETHICAL
ISSUES.
>> LET'S SAY I OWN A VOICE
CLONING BUSINESS AND I'VE HIRED
YOU TO COME IN AND GIVE ME MY
LEADERSHIP OR ETHICS TRAINING ON
HOW I SHOULD ASK RESPONSIBLY AND
MY COMPANY'S ETHICS PROGRAM
SHOULD LOOK LIKE.
THE 30-SECOND SPEECH THAT YOU
GIVE TO COMPANIES.
>> WHY THEY "HAVE ETHICS OR WHAT
--
>> WHAT SHOULD MY ETHICS PROGRAM
LOOK LIKE?
UNFAIR IN 30 SECONDS?
>> I WOULD ACTUALLY -- I THINK
MICROSOFT'S AI AND ETHICS COVERS
MOST OF THE BASES, PRIVACY AND
SECURITY, TRUSTWORTHINESS, THE
IDEA OF ACCOUNTABILITY AND
TRANSPARENCY.
LOOKING AT HARMS AND DIGNITY,
YOU KNOW, THOSE TYPES OF
QUESTIONS THAT YOU WOULD ASK
ABOUT WHO ARE THE MAJOR

STAKEHOLDERS, HOW IS VALUE
CREATED OR DESTROYED.
WHETHER MAJOR NORMS AND
PRECEDENTS NEED TO BE AWARE OF
HOW ARE RIGHTS DIMINISHED OR
ENABLED.
THOSE ARE THE BIG THREE.
>> IF I'M A SMALL COMPANY AND I
DON'T HAVE THE LUXURY OF HAVING
SEPARATE AI GROUP IN MY ETHICS
GROUP IN MY ORGANIZATION, WHO AM
I GOING TO TURN TO HELP ME MAKE
DECISIONS TAKE IT OUT OF MY
PROFIT MOTIVE.
>> BUT THAT'S MAKING THE
ASSUMPTION THAT SOMEONE THAT IS
TO COME FROM OUTSIDE AND LIKE
CURE THE ILLS OF THE EVIL DOERS
IN BUSINESS.
I JUST DON'T THINK THAT MOST
BUSINESSES -- I THINK THAT
THEY -- I'M OF THE BELIEF THAT
MOST OF THE COMPUTER SCIENTISTS
OR CEOs THE LEADERSHIP TEAM IS
ACTUALLY TRYING TO DO THE RIGHT
THING THEY JUST DON'T HAVE THE
FRAMEWORK TO SAY LIKE, HOW WOULD
I THINK ABOUT WHAT'S RIGHT.
I DON'T KNOW WHAT TO THINK
ABOUT.
THAT'S WHEN YOU START TALKING
ABOUT STAKEHOLDER AND VALUE OR
RIGHTS ENABLING LANGUAGE VERSUS
LIKE AN EXTERNAL GROUP THAT
NEEDS TO COME IN AND CORRECT
THEM IN SOME WAYS.
DOES THAT MAKE SENSE?
JUST A DIFFERENT MODEL OF WHAT
YOU THINK THAT THEY ACTUALLY
WANT TO DO NATURALLY VERSUS WHAT
YOU NEED TO PREVENT THEM FROM
NOT DOING.
>> I'M THINKING MORE THAT --
MICROSOFT IS VERY ADVANCED
COMPANY.
IT HAS GIVEN TONS OF THOUGHT TO

DEVELOPING AN ETHICS PROGRAM.
PROBABLY BEYOND WHAT AN
INDIVIDUAL STARTING A COMPANY
HAS REALLY THOUGHT OF.
>> I PERSONALLY THINK THAT YOU
HAVE TO HAVE A COST FOR MAKING
CERTAIN TYPES OF CREATING AI.
MY ARGUMENT, I'VE WRITTEN ABOUT
THIS, IS THAT RIGHT NOW THERE'S
ACTUALLY NO INTERNALIZATION OF
THE COST OF DEVELOPING THESE
TYPES OF ALMOST INFRASTRUCTURE
COMPUTER PROGRAMS THAT PEOPLE
RELY UPON IN THE MARKETPLACE OR
IN SOCIETY.
SO JUST LIKE WE HAVE THE
PROFESSIONALIZATION OF CIVIL
ENGINEERS THAT CREATE BRIDGES,
WE NEED TO HAVE THE
PROFESSIONALIZATION OF COMPUTER
SCIENTISTS BE REQUIRED TO BE ON
CERTAIN PROG HOAX WOULD BRING IN
THAT TYPE OF KNOWLEDGE.
EVEN IF YOU JUST HIRE TWO HOOD
HE'S IN A GARAGE OFF THE STREET
HAVE NO DEGREES WHATSOEVER BUT
ARE GOOD CODERS HE CAN STILL
HAVE SOMEONE WHO IS PROFESSIONAL
THAT HAS THAT TYPE OF TRAINING
THAT IS THEN RESPONSIBLE FOR THE
ETHICAL IMPLICATIONS.
WE DO THIS WITH CPA, ENGINEERS.
THERE'S LOT OF AREAS WHERE WE AS
SOCIETY HAVE DECIDED THAT THESE
TYPES OF MARKET DECISIONS ARE
TOO CONSEQUENTIAL FOR SOCIETY
THAT WE NEED TO HAVE A
PROFESSIONAL PLACED AND
RESPONSIBLE.
STATISTICS HAS THIS.
THESE TYPE OF AREAS.
MY ANSWERS YOU HAVE ACTUALLY
HAVE PROFESSIONALIZATION CLASS
OF COMPUTER SCIENTISTS IN HIGH
DEMAND.
BUT THAT'S DIFFERENT FROM WHAT

WE HAVE RIGHT NOW.
BUT THAT'S MY VISION FROM TEN
YEARS FROM NOW.
>> THIS IS JUST.
-- NOW MY SON WHO IS NOW IN
COLLEGE STARTED CODING IN LIKE
7th, 8th GRADE.
I AM DOING A MASTER'S FIRST TIME
I TOOK AN ETHICS COURSE IS LIKE
ETHICS IN ICT IN A GRADUATE
COURSE RIGHT NOW.
I THINK YOU HAVE TO PUSH THAT
WHOLE COURSE ACTUALLY TO A
MIDDLE SCHOOL STUDENT BECAUSE
RIGHT NOW THEY ARE CREATING CO
CODE.
THEY ARE THINKING ABOUT
TECHNOLOGY.
WE HAVE TO BRING THAT COURSE
CURRICULUM TO A LEVEL OF MIDDLE
SCHOOL AND ABOVE.
BECAUSE RIGHT NOW IF IT'S -- IN
GRADUATE SCHOOL IT'S ALREADY TOO
LATE.
>> WE WANT TO MAKE SURE THAT OUR
7th GRADERS CAN REFUSE TO DO
THE HOMEWORK.
>> THIS IS UNETHICAL.
>> HAVE ANY THOUGHTS ON THIS
BEFORE I GO ON TO THE NEXT
QUESTION?
>> NO.
I AGREE.
I THINK WE NEED TO PUSH THE
EDUCATION FURTHER DOWN THE
LADDER AND IF IN FACT STUDENTS
ARE CODING IN 7th GRADE,
ETHICS COMPONENT SHOULD BE A
PART OF THE CURRICULUM, THAT'S
WHAT'S HAPPENING IN THE THE
TERTIARY LEVEL NOW SO ETHICS, I
THINK FACTORS IN AND SERIOUS
REQUIRED WAYS TO WORK THAT
COMPUTER SCIENTISTS ARE DOING,
COMPUTER SCIENCE STUDENTS ARE
DOING IN THIS PROGRAM, YOU WOULD

KNOW THIS, TOO, RIGHT?
>> YES, I HAVE SOMETHING TO ADD.
WHEN I THINK ABOUT MY STUDENTS,
THEY OFTEN EXCITED ABOUT WHAT
THEY CAN DO.
THEIR MASTERY OVER THEIR
CAPACITY TO CODE OR BUILD OR
DESIGN SOMETHING.
SAME WITH ENTREPRENEURS.
THOSE WHO I KNOW ARE EXCITED
ABOUT THE NEW IDEA AND HOW IT
CAN DO WELL.
I WOULD LOVE TO HAVE A
CONVERSATION WITH THESE
ENTREPRENEURS ABOUT, WHY ARE YOU
IN THIS?
AND WHY YOU'RE IN IT CAN YOU
POSSIBLY BE THINKING ABOUT THE
OTHER SIDE OF IT FROM THE USER'S
PERSPECTIVE AND DO THE VALUES IN
ANY WAY CONFLICT.
I THINK A LOT OF TIMES WHEN WE
RUN INTO ETHICS PROBLEMS IS
BECAUSE WE HAVE CONFLICKING
VALUES.
SO, THE VALUE OF DISCOVERY,
CREATION, MASTERY.
MAY BE WOULD CONFLICT WITH THE
THE VALUE OF SELF PROTECTION,
INDEPENDENCE, AUTONOMY, THE
ABILITY TO HAVE SOME CONTROL
OVER YOUR PERSONAL, WHAT'S
YOURS.
AS A CONSUMER, I OFTEN FEEL THAT
I HAVE LOST CONTROL AND I CAN'T
GET IT BACK.
DOES THAT MATTER TO THE PERSON
WHO IS MASTERED THE CODE.
DOES THAT MATTER TO THE
ENTREPRENEUR WHO IS EXCITED
ABOUT THE INVESTMENT
OPPORTUNITIES.
TO MAKE THIS PRODUCT GET INTO
THE MARKETPLACE.
THOSE COULD BE INTERESTING
CONVERSATIONS IF PEOPLE ARE

WILLING TO HAVE THEM.
>> HOW DO YOU FORCE THOSE
CONVERSATIONS?
>> YOU ASK THE FTC -- TO ENGAGE
THE PUBLIC AT THE BEGINNING OF
THE DEVELOPMENT OF THE NEW
TECHNOLOGY, WHAT ABOUT THAT?
IS THAT FAIR?
>> BUT IT'S ALSO PUBLIC
DISCOURSE.
MICROSOFT IN FACT BECOMES A
PLAYER IN THIS CONVERSATION WHAT
YOU JUST PHRASED, ROSALYN, I
THINK IS A RECOGNITION THAT
SPECIFICALLY IN THE DOMAIN OF
DATA, WE ARE NO LONGER JUST
CUSTOMERS.
>> ABSOLUTELY.
>> WE ARE ALSO -- THE PHRASE
SURVEILLANCE CAPITALISM SUDDENLY
IS EVERYWHERE AND THERE'S A GOOD
REASON FOR THAT.
>> YES.
>> RIGHT?
>> WE'RE ACTUALLY CONSEWELLERS,
LIKE WE PAID MICROSOFT.
IN OTHER AREAS WE'RE NOT A
CONSUMER.
IF YOU LOOK AT THEIR 10K YOU DO
SEARCH ON CUSTOMER, YOU FIND
THREE INSTANCES BECAUSE THEIR
CUSTOMER IS THE ADVERTISING
AGENCIES, WE'RE NOT CUSTOMERS TO
FACEBOOK OR ANYONE ELSE.
IT'S NOT -- WE USE THAT TERM
LIGHTLY BUT THEY DON'T REFER TO
US AS THAT AS THEY REPORT TO THE
SEC OWE MICROSOFT IS A LITTLE
BIT DIFFERENT.
THE MODEL IS A LITTLE BIT
DIFFERENT.
>> WE HAVE A QUESTION FROM THE
AUDIENCE ASKING, WHAT DOES
MICROSOFT POLICE WHETHER USERS
TEMPORARILY USING SOMEONE'S
VOICE WITHOUT THEIR CONSENT OR

WITH THEIR CONSENT EVEN?
HOW DO YOU ENSURE THAT CONSENT
IS THEREAFTER THE FACT?
>> I'LL DEFER THAT QUESTION,
I'LL SHARE MY E-MAIL
THERE'S A LOT OF TALK ABOUT WHAT
CHANGES TO A CORPORATION'S
RESPONSIBILITY IS AND THERE WAS
A NEW DEFINITION OF WHAT A
CORPORATION IS SUPPOSED TO BE
OTHER THAN MAKING MONEY FOR
SHAREHOLDERS IS SUPPOSED TO HAVE
AN ETHICS RESPONSIBILITY.
RESPONSIBILITY TO SOCIETY.
DO YOU SEE THIS AS HAVING ANY
EFFECT ON COMPANIES THAT MIGHT
BE DEVELOPING VOICE CLEANING AND
RELATED TECHNOLOGIES?
SHOULD IT HAVE SOME SORT OF
EFFECT?
>> IT CERTAINLY SHUT.
AND I DON'T -- SHOULD.
I DON'T KNOW, WE HAVE SOMEBODY
IN THE ROOM ENGAGED IN THE
PRACTICE.
WHAT DOES SO THE BUSINESS
ROUNDTABLE AND LARRY FINK'S
LETTER WHERE HE FOCUSES ON
RESPONSIBILITY AND THERE'S BEEN
A SERIES OF TOPICS HE AND OTHER
BUSINESS LEADERS HAVE TOUCHED
ON.
THE ASPEN PRINCIPLES, THE
FRAMEWORK THAT JUST CAME OUT OF
DAVOS WHICH IS THE SECOND
EDITION THAT LOOKS AT PRIVACY
ISSUES.
ALL THESE QUESTIONS ARE BEING
RAISED BY HIGH-PROFILE
INDIVIDUALS AND ORGANIZATIONS.
ONE WOULD LIKE TO BELIEVE IF YOU
HAVE ENOUGH CONVERSATIONS WITH
ENOUGH AUTHORITY FIGURES OR JUST
THE GENERAL PUBLIC ONCE THEY
START TAKING ABOUT STAKEHOLDER
VALUE RATHER THAN SHAREHOLDER IT

WILL HELP PEOPLE, PEOPLE BEING
CORPORATIONS AND BUSINESSES
PAYING MORE ATTENTION THAN PRIOR
WAS THE CASE THAT NOTION OF
SOCIAL RESPONSIBILITY.
YOU CAN ARGUE SOME CORPORATIONS
HAVE ALWAYS HAD THAT COMMITMENT
AND EXERCISED THAT COMMITMENT
FOR A WHOLE RANGE OF REASONS.
SOMETIMES IT COULD BE RELIGION
SOMETIMES IT'S JUST EDUCATION.
YOU HAVE A SENIOR STAFF
INCLUDING A CEO THAT THOUGHT
ABOUT THE ISSUES AND WANTS TO
PRACTICE THEM.
I THINK THE PUBLIC CONVERSATION
AND AGAIN THE WORKSHOP WOULD BE
AN EXPRESSION OF THAT INTENT.
I THINK THAT COULD HELP.
I WOULD ALSO LIKE TO BELIEVE AND
I THINK THE HISTORICAL EVIDENCE
WOULD SUPPORT THIS, THAT WE ARE
NOT SUDDENLY IN A NEW AGE WHERE
WE'RE AWARE CORPORATIONS HAVE
SOME RESPONSIBILITY TO THE REST
OF SOCIETY.
PEOPLE HAVE TALKED ABOUT THAT
FOR A LONG TIME AND THEY HAVE
EXERCISED THAT TYPE OF
RESPONSIBILITY FOR A LONG TIME.
>> IS IS THERE ANY LIMIT TO THE
CORPORATE RESPONSIBILITY FOR
VOICE CLOTH -- CLONING
TECHNOLOGIES?
IS THERE THERE SOMETHING THAT
PRESSES UP THE POSSIBILITY THAT
CONTROL COULD BE LOST?
I USE A LOT OF FILM IN TEACHING.
I THINK ABOUT THE FILM WAR GAMES
AND THE WHOLE PURPOSE OF THE
FILM WAS TO SUGGEST ULTIMATELY
IT WAS MACHINE LEARNING THAT HAD
TO SUPERCEDE HUMAN ERROR IN
ORDER TO PROTECT US IN THE END.ua
WE HEARD A TALK EARLIER ABOUT
MACHINE LEARNING AND THIS IS

EVOLVING IN PART BECAUSE OF THE
CAPACITY OF MACHINE TO LEARN.
AT SOME POINT IS THE CORPORATION
NOT GOING TO BE ABLE TO BE
RESPONSIBLE FOR THIS GIVEN THE
TECHNOLOGY?
I DON'T KNOW.
WE NEED TO THINK EARLY IN ABOUT
DESIGN IN THE FACE OF THE
POSSIBILITY WE WON'T HAVE
CONTROL.
LIKE THE NUMBER THAT SHOWED UP
IN MY PHONE.
>> IT'S A GOOD QUESTION.
>> KIRSTEN.
>> THE RESPONSIBILITIES
DISTRIBUTED THE USERS, FIRM AND
TECHNOLOGY DOESN'T JUST
DISAPPEAR LIKE THE MISSING
MASSES OF LATEUR IDEA AND IF THE
DEVELOPER OF AI AND MACHINE
LEARNING AND THE PROGRAM IN THE
WILD DOESN'T CREATE A MECHANISM
TO GIVE GOVERNANCE AND JUDGE AND
CORRECT MISTAKES DONE IN THE
WILD IF IT'S DESIGNED THANE WAY
TO BLACK BOX IT.
IT'S A DESIGN DECISION, IS MY
ARGUMENT AND IF THIY MAKE THAT
DECISION IT CAN'T BE LATER USED
TO IDENTIFY JUDGE AND FIX
MISTAKES THE DEVELOPER'S THEN
RESPONSIBLE FOR EVERYTHING THAT
GOES UPON IT'S ACTUALLY THE
OPPOSITE OF THE CONVERSATION
GOING ON WHICH IS IT'S BLACK BOX
IT'S INSCRUTABLE THEREFORE
NOBODY'S RESPONSIBLE.
MY ARGUMENT IS IF YOU DESIGN
SOMETHING TO BE INDISCRUTABLE
AND IT WILL FIGURE ITSELF OUT
BUT WHEN WE SAY NO ONE'S
RESPONSIBLE THAT'S THE WORSE
THING YOU CAN DO.
>> THE WAY IT PLACE OUT INVOICE
CLOTH -- CLONING IF I WERE TO

UPLOAD HER VOICE SAMPLE TO
MICROSOFT AND GO TO A FINANCIAL
INSTITUTION SHE HAD AN ACCOUNT
ON USING THE VOICE PRINT OF HER
WOULD YOUR COMPANY BE LIABLE?
AS LONG AS THERE'S INTERVENTION
AS WE ARE BUILDING WE NEED TO
MAKE SURE THE STAKEHOLDERS AND
SOFTWARE AND THE USER WHO MAY
USING THE SOFTWARE BOTH HAVE
MECHANISMS TO INTERVENE IN TO
WHAT'S GOING ON IN THE SYSTEM IF
THERE'S MECHANISMS FOR HER TO
INTERVENE IN TO WHAT HAPPENED IS
THE WAY WE SHOULD APPROACH THIS.
>> OR THE WAY YOU'RE DOING IT
WHERE I'M NOT SURE IT'S USED IN
A DIRECT FASHION OR COME TO ME.
SO I'M NOT SURE WHAT'S GOING ON
IN SOCIETY SO IN MICROSOFT I'LL
HOLD IT CLOSER IN AND PUT IN THE
CLOUD.
>> AS AN ANALYZING ENGINEER IT'S
VERY HARD THE THINGS WE'RE
TALKING ABOUT AND YOU'LL ATTEST
TO THAT.
IT'S VERY HARD TO ONE THE WAY AI
ESPECIALLY DEEP LEARNING MODELS
ACTUALLY ARE BUILT AND WHY THE
OUTPUT IS THE OUTPUT BECAUSE
IT'S HARD TO EXPLAIN EVEN BY THE
ENGINEERS AND RESEARCHERS
BUILDING IT.
AND FORCING THEM IS HARD AS WELL
BUT WE'RE HERE TO SOLVE HARD
PROBLEMS SO EVENTUALLY WE'LL GET
TO A PROBLEM OF SOLVING IT.
>> YOU DON'T THINK WE'LL SOLVE
IT IN THE NEXT FIVE MINUTES OF
THIS PANEL?
>>ONE THING THAT MAY HAPPEN IS
WE WANT MORE HUMAN INVOLVE MANY
IN THE USE OF THE MACHINE
LEARNING SO IT'S NOT ROBUST AND
NOT USE NEURAL NETWORKS THAT'S
NOT EXPLAINABLE BUT USE A

REGRESSION EASILY IDENTIFIABLE
AND THE SENTENCING ALGORITHM
THAT WAS A BLACK BOX AND THERE'S
WAYS TO MAKE IT SIMPLE SO
THERE'S DUE PROCESS.
ONE PLACE THAT'S INTERESTING IS
INSURANCE AND HOW YOU ADJUDICATE
A CLAIM AND THEY'RE TRYING TO
USE MACHINE LEARNING IN WHO IS A
GOOD CANDIDATE AND ADJUDICATE A
CLAIM AND WHO YOU INSURANCE WITH
BUT INSURANCE IS HIGHLY
REGULATED BY THE STATE.
THEY'RE ACTUALLY VERY GOOD AT
EXPLAINING THE DECISIONS THEY
MADE BECAUSE THEY HAVE TO.
I WAS JUST AT SOMETHING AT
BENTLEY RECENTLY WHERE IN ALL
THE EXAMPLES THEY USED INSURANCE
AND WERE SHOWING ALMOST A
DASHBOARD OF HOW DECISIONS WERE
MADE WHETHER IT WAS SKEWING ONE
WAY OR ANOTHER.
HOW IT COULD BE FIXED.
THEY WERE CREATING A WAY THE
USER COULD BE INVOLVED THOUGH
THEY WERE USING PRETTY
SOPHISTICATED STUFF.
YOU CAN HAVE SOMETHING THAT'S
REALLY SOPHISTICATED BUT THERE'S
ALSO WAYS TO FIGURE OUT A
MECHANISM SO THE HUMAN COULD BE
MORE INVOLVED.
I'D LOOK AT THE PLACES THAT ARE
AT THE FOREFRONT OF THAT HIGHLY
REGULATED, INSURANCE IS ONE OF
THEM.
YOU HAVE EMPLOYMENT WHICH WILL
BE ANOTHER.
AT SOME POINT SOMEONE'S GOING TO
GET SUED AND THEY'LL HAVE TO GO
THROUGH DISCOVERY AND FIGURE OUT
A WAY WHERE WE'RE MAKING AI
DECISIONS AND HOW TO EXPLAIN IT.
>> INVOLVEMENT OF THE USER SEEM
IMPORTANT PRINCIPLES.

IN PREPARATION FOR THIS WORKSHOP
WE FOUND ON GITHUB A CODE
SOMEONE HAD CREATE FROM AN
ARTICLE ON VOICE CLONING THAT IS
NOW OUT THERE FOR SOMEBODY BAD
TO USE FOR ONE OF THE HORRIBLE
CRIMES THAT WAS MENTIONED.
HOW ARE WAYS OF DEALING WITH
THIS?
RIGHT NOW MOST RESEARCH IS
PUBLISHED WITH THE IDEA OF BEING
OPEN ABOUT YOUR MATERIALS AND
METHODOLOGY.
IS THAT SOMETHING THAT SHOULD BE
RECONSIDERED TO PREVENT DEEP
LEARNING PROGRAMS GETTING IN THE
WRONG HANDS?
>> PERSONALLY, AGAIN, I THINK
THE CODE THAT WE PUBLISHED
BECAUSE THERE'S SO MANY GOOD USE
CASES I'LL SAY WAY MORE THAN IF
YOU JUST DO THE BENEFITS AND
AGAIN, NOT ONLY WE DON'T KNOW
THE HARMS, RIGHT, AT THIS POINT
HOW THE TECHNOLOGY COULD BE USED
IN A HARMFUL WAY.
WE DON'T KNOW HOW THE TECHNOLOGY
CAN BE USED IN A GOOD WAY AND IT
COULD EVOLVE TO BE USED FOR
SOMETHING WHICH CAN CHANGE THE
WORLD.
PERSONALLY I THINK OPENNESS IS
VERY GOOD AND NOW THE MODEL OF
SAYING IF IT WERE USED AND
GITHUB IS ONE OF THE MICROSOFT
THINGS WE HAVE AN ACCEPTABLE USE
POLICY THERE WHERE THE CODE EVEN
IF IT IS USED IF IT'S FULL
LEARNING AND TRAINING PURPOSES
WHICH SHOWS IT CAN BE USED FOR
THOSE PURPOSES IT'S FINE BECAUSE
IT'S MORE FOR LEARNING AND
TRAINING AND IF IT'S USED FOR
NEFARIOUS PURPOSES WE HAVE TO
TAKE IT DOWN.
>> AGAIN IT'S AFTER THE FACT.

>> YES, BECAUSE AGAIN, IT'S HARD
TO FIGURE OUT THE INTENT OF THE
PUBLISHER AT THIS POINT.
THE PUBLISHER HASN'T ANTICIPATED
USE BY BAD HANDS.
>> IN OUR REMAINING MINUTE AND A
HALF, LET'S GO BACK TO THE 30
SECONDS COULD EACH OF YOU SAY
WHAT IF ANYTHING SHOULD BE DONE
BY BUSINESSES, GOVERNMENT,
INDEPENDENT INDUSTRY GROUPS,
WHATEVER ACT YOU THINK IS
APPROPRIATE TO ENSURE SOCIETY
BENEFITS FROM VOICE CLONING
TECHNOLOGIES BUT REDUCE THE
LIKELY HARMS.
>> DO YOU CARE THE ORDER?
>> I DON'T CARE AT ALL.
>> SEEMS WE WANT ALL THE PLAYERS
YOU JUST NAMED TO WORK TOGETHER
TO MINIMIZE THE LIKELIHOOD OF
HARM EVEN AS YOU RECOGNIZE THE
POTENTIAL UPSIDE AND THE FIRST
PLAN MADE IT CLEAR WHAT IT WOULD
BE FOR A SMALL SEGMENT OF
SOCIETY AND MORE BUT IF YOU WANT
PUBLIC SECTOR AND PRIVATE SECTOR
TO WORK TOGETHER TO MAKE THAT
HAPPEN.
I DON'T THINK YOU WANT TO
OVERLOAD THE RESPONSIBILITY ON
ONE PLAYER OR ONE PARTICULAR
GROUP OF PLAYERS.
AS A TECHNOLOGY COMPANY THEY
SHOULD HAVE NOT JUST ENGINEERS
BUT PEOPLE TAKING ABOUT THAT
PARTICULAR PROBLEM THEY'RE
TRYING TO SOLVE.
>> THE TWO THINGS THAT I THOUGHT
OF IS INTERESTING DETECTION
MODELS WHEN WE HAVEN'T CRACKED
SINGING I THOUGHT THAT'S
INTERESTING BECAUSE THAT'S AN
AUTHENTIC WAY TO MAKE SURE IT
WAS HUMAN.
ARE PEOPLE GOING TO START

SINGING LITTLE BO PEEP AND WE
SAW THE DOG WAS FOOLED BY THE
FAKE VOICE.
CERTAIN DETECTION MECHANISM CAN
THAN BE PROLIFIC THROUGH PHONES
OR OTHER SPEAKER SYSTEMS TO BE
ABLE TO IDENTIFY WHEN SOMETHING
IS FAKE OR NOT IS MY FIRST
THOUGHT.
PUT MORE EMPHASIS ON THAT.
THE SECOND IS THE GOVERNMENT AND
OTHER LARGE ORGANIZATIONS HAVE
AN ABILITY TO CREATE THE DEMAND
FOR GOOD TECHNOLOGY NOT JUST
THROUGH REGULATION BUT TO SAY I
WILL NOT DO BUSINESS WITH YOU
UNLESS YOU HAVE THESE THINGS IN
PLACE.
THAT'S A STRONG LIKE GOVERNMENT
AS A GOVERNMENT ACT OR BIG
CORPORATIONS SIMILAR TO WHAT WE
TALKED ABOUT AT LUNCH WHEN
MICROSOFT MAKES A STAND ON
SOMETHING THEY SAY BY THE WAY
THAT'S THROUGHOUT MY SUPPLY
CHAIN AND WE'LL MAKE THAT A
DECISION EVERYBODY HAS TO ABIDE
BY.
THERE'S WAYS TO MAKE DEMAND FOR
VOICE CLONING.
THE MORE THE GOVERNMENT CAN MAKE
BIG DEMAND SAYING YOU HAVE TO
ABIDE BY THESE PRINCIPLES IT
WON'T STOP THE FRAUDSTERS ON
GETHUB AND PULL DOWN BAD CODE
BUT IT WILL STOP SOME PEOPLE WHO
WANT TO GET THE GOVERNMENT'S
BUSINESS.
THAT'S WHAT CAME TO MIND.
>> WHAT CONCERNS ME THE MOST
ABOUT TECHNOLOGY ETHICS IS THE
HUBRIS THAT CAN SOMETIMES COME
ALONG WITH NEW TECHNOLOGICAL
ADVANCES AND I WAS HEARTENED
WHEN I HEARD SOMEONE FROM A TECH
COMPANY SAYING THERE'S SOME

THINGS WE DON'T UNDERSTAND OR
CAN'T CONTROL.
THAT LEVEL OF HUMILITY TO ME IS
WHAT OPEN THE DOOR FOR5ú
AUTHENTIC, OPEN CONVERSATIONS TO
KNOW WHERE WE ARE AND GET TO
WHERE WE WANT TO GO.
>> WELL, THANK YOU ALL.
IT'S BEEN THE MOST ETHICAL PANEL
I'VE EVER BEEN ON.
THANK YOU VERY MUCH.
THANK YOU.
[APPLAUSE]
>> MY NAME IS WILL MAXIM AND WE
HEARD ABOUT THE SCAMS AND

EXTORTIONS AND THINGS THAT CAN
HAPPEN WITH VOICE CLOTH AND WE
TALKED ABOUT THE ETHICAL ISSUES
AND WHAT BOUNDS COMPANIES AND
WHAT THEIR OBLIGATIONS ARE FOR
BALANCING THE PROS AND CONS.
THE THIRD PANEL WILL HOPEFULLY
TALK ABOUT HOPEFUL ASPECTS AND
MITIGATION, DETECTION AND OUR
ABILITY TO HOPEFULLY LET PEOPLE
KNOW WHEN THEY'RE DEALING WITH A
VOICE CLONE A FAKE SO THEY'RE
LESS LIKELY TO BE SCAMMED OR A
VOICE TO FOOL THEM.
I ENCOURAGE EVERYONE TO GO
ONLIBE TO LOOK AT OUR -- ONLINE
TO LOOK AT OUR PANEL'S FULL
BIOS.
I'LL DISTILL THEM AS I
INTRODUCE.
WE HAVE JOHN AMEIN VICE
PRESIDENT AT ID&AND TO HIS LEFT
NEIL JOHNSON A SCIENCE AND
ENGINEERING TECHNICAL ADVISER
CONTRACTOR FOR THE DEFENSE
ADVANCED RESEARCH PROJECTS
AGENCY ALSO KNOWN AS DARPA.
TO HIS LEFT WE HAVE PATRICK
TRAYNOR AS YOU KNOW A PROFESSOR
IN THE DEPARTMENT OF COMPUTER

AND ENGINEERING AT THE
UNIVERSITY OF FLORIDA AND
CO-DIRECTS THE CYBER INSTITUTE
AND JOINING US VIA SKYPE IS
PROFESSOR SIWEI LYU AND DIRECTOR
OF MACHINE LEARNING LAB AT
ALBANY AND WE'LL HAVE A SHORT
PRESENTATION FROM PROFESSOR LYU
AND THEN HAVE A PRESENTATION ON
DARPA'S PROGRAM AND JOHN WILL DO
A SHORT DEMO OF DETECTION
TECHNOLOGY AND THEN WE'LL HAVE
TIME FOR PANEL Q&A.
SO FIRST I'LL TURN TO PROFESSOR
LYU.
YOU CAN GET US STARTED.
>> ALL RIGHT.
CAN YOU HEAR ME CLEARLY?
>> YES, WE CAN HEAR YOU.
>> ALL RIGHT.
GREAT.
THANK YOU FOR THE INTRODUCTION.
TODAY I'M GOING TO TALK ABOUT
OUR RECENT WORK COLLABORATED
WITH U.C. BERKELEY ON USING
ANALYSIS TO RECOGNIZE FAKE
VOICES.
SO AS WE HEARD A LOT ABOUT
SYNTHESIZED VIDEOS WE'RE SEEING
A GROWING NUMBER OF FAKES.
TO GIVE AN EXAMPLE, LAST SUMMER
A COMPANY IN CANADA RELEASED
SOME EXAMPLES OF AMAZINGLY HIGH
QUALITY SYNTHETIC VOICE OF JOE
ROGGAN.
I'LL PLAY TWO VIDEOS.
>> SOME NEED TO IMPROVE THE
QUALITY OF YOUR EXISTENCE ON
EARTH.
YOU HAVE TO DO THE RIGHT THINGS.
>> HEALTHY, LOCAL INGREDIENTS,
FARM FRESH INGREDIENTS.
>> ALL RIGHT.
I DON'T KNOW.
ANYONE WANT TO GUESS WHICH ONE'S
REAL AND FAKE.

NEXT, PLEASE.
IT TURNS OUT THAT THE FIRST ONE
IS THE SYNTHETIC VOICE AND THE
SECOND IS THE REAL ONE.
THE QUALITY OF THE SYNTHETIC
VOICE IT'S VERY DIFFICULT TO
TELL THEM APART.
IT SHOWS WE'RE AT THE VERGE OF
AI SYNTHESIZED VOICES BEING
INDISTINGUISHABLE TO OUR EARS.
LATER LAST YEAR WE SAW THE FIRST
REPORTS OF A FRAUD WITH
SYNTHESIZED VOICES AND THIS IS
CAUSING DAMAGES TO CONSUMERS AND
COMPANIES.
BEFORE I TALK ABOUT HOW WE CAN
ACTUALLY DEAL WITH THIS PROBLEM
LET ME JUST GIVE YOU A ROUGH
IDEA HOW SYNTHETIC VOICE
CREATED.
I USED THE OPEN SOURCE GOOGLE
WAVE LENGTH AS A BASIC MODEL.
THIS IS ONE OF THE MOST DOMINANT
VOICE SYNTHESIZERS.
THE THERE'S TTS AND THE NETWORK
TOGETHER WITH THE IMPULSE THAT
HAS BEEN GENERATED.
THE SYNTHESIS IS SOLVED AS A

REDICTION -- PREDICTION AND
PREDICTS WHAT THE SAMPLE WILL
SOUND LIKE AND ALL THE PREVIOUS
GENERATED SAMPLES.
SO CURRENTLY THE NEW NETWORK
USES HUNDREDS OF VOICES IN THE
CASE OF JOE ROGAN'S CASE AND
TAKES WEEKS ON GPU SERVERS.
THE SYNTHESIS TAKES TIME.
NEVERTHELESS THE TECHNOLOGY
IMPROVES AND THE EFFICIENCY OF
AI VOICE SYNTHESIS WILL IMPROVE
AND THEY'RE TRYING TO GIVE UP TO
PROVIDE TECHNOLOGY TO RESIST THE
PROBLEM.
I'LL BRAVELY DISCUSS OUR WORK IN
THIS RESPECT WHICH IS CURRENTLY

SUPPORT ED BY THE PRECURSOR
DESCRIBED AFTER ME.
SO WE USE DIFFERENCES TO
DIFFERENTIATE AI SYNTHESIZED AND
REAL VOICES AND THIS IS WELL
KNOWN IN THE NEURAL SYSTEMS.
OUR EARS ARE SENSITIVE TO FACES
OF SIGNALS.
LET ME EXPLAIN AT THIS POINT IN
ANOTHER DETAIL.
LET'S CONSIDER A CONVERSATION
LIKE IN THE PICTURE SHOWED HERE
BETWEEN TWO SPEAKERS, TWO
PEOPLE.
NEXT, PLEASE.
NOW THE VOICE OF THE SPEAKER
PROP -- HOWEVER THE DIRECT PATH
BETWEEN THE SPEAKER'S MOUTH AND
EARS IS NOT THE ONLY WAY SOUND
PROPAGATES WE HAVE SOME IN OTHER
DIRECTIONS AND THEN THEY BOUNCE
FROM THE WALLS OR CEILINGS TO
REACH THE LISTENER'S EARS.
THE END RESULT IS THE LISTENER'S
EARS RECEIVED SOUND WAVE FROM
DIFFERENT SOURCES AT DIFFERENT
TIMES BUT THROUGH DIFFERENT
TRAJECTORIES.
THIS IS DIFFERENCES IN THE SOUND
WAVE.
SO WE CAN DETRACT USEFUL
MESSAGES FROM THE SOUND WE HEARD
INSTEAD OF SOUND WAVES ARRIVING.
HUMAN EARS ARE MORE EVOLVED IN
THE ARRIVING TIMES AND ONLY
FOCUSSING ON THE MAGNITUDE OR
LOCAL ENERGY THE SOUND WAVES
CARRY TO HEAR SOMEONE TALKING
CLEARLY THOUGH THE SPEAKER IS
MOVING AROUND IN A ROOM.
SO THIS IS A VERY INTERESTING
OBSERVATION.
HOWEVER, AI SYNTHESIZED VOICES
ARE GENERATED WITH A SPECIFIC
NETWORK MODEL TO HAVE A
DIFFERENT SIGNATURE THAN

DIFFERENT VOICES.
BECAUSE THEY SIMULATE THE SAME
MAGNITUDE AS HUMAN VOICES OUR
BRAIN HEARS THE SOUND SIMILARLY
BUT IF WE HAVE SOME WAY TO
REVEAL THE DIFFERENCE IN FACES
WE MAY BE ABLE TO IDENTIFY THE
AI SYNTHETIC VOICES.
SO OUR METHOD USES THE
BISPECTRAL ANALYSIS AND WE USE
VOICES TO EXPOSE THE DIFFERENCES
IN THEIR FACES WE REPRESENT THE
SOUND WAVES AS A COMBINATION OF
PURE HARMONICS AND WE LOOK AT
CORRELATIONS AMONG FREQUENCY
COMPONENTS AND THEY'RE SENSITIVE
THROUGH THE FACE DIFFERENCES
THROUGH DIFFERENT SOUND WAVES.
THE RESULT OF THE BISPECTRAL
ANALYSIS ARE REALIZED AT THE
BOTTOM OF THE SLIDES AND I WILL
SKIP OVER THE DETAILS BUT WE CAN
OBSERVE THAT THEY'RE VERY
DIFFERENT CHARACTERISTICS AMONG
HUMAN VOICES ON THE RIGHT COLUMN
THAN ALL THE AI SYNTHESIZED
VOICES TO GIVE UP A HOPE TO GIVE
UP SOME WAY TO DIFFERENTIATE
THEM USING FEATURES CAPTURED BY
BISPECTRAL.
WE WE USED THE BISPECTRUM
ANALYSIS AND WE CAN SEE THE
HUMAN VOICES WHICH CORRESPOND TO
THE REGION HIGHLIGHTED CAN BE
SEPARATED FROM THE SYNTHESIZED
VOICES.
THE CLASSIFICATION ACCURACY OF
THIS METHOD IS AROUND 90%
MEASURED.
SO THIS SAY VERY PROMISING
INITIAL RESULT BUT AS WE STUDY A
LITTLE FURTHER IT'S HILTE --
HIGHLY AFFECTED BY HIGH
FREQUENCY VOICES AN
DEGRADATIONS.
IT'S MY BELIEF THERE'S MORE TO

BE DONE IN THIS AREA THAN WE
HAVE DONE SO FAR.
SO THAT'S SOMETHING WE'RE
WORKING ON ACTIVELY ON THESE
DAYS.
NOW SOME CONCLUSIONS AND
DISCUSSIONS.
FIRST OF ALL, I HOPE HAVE YOU
CONVINCED AI ALGORITHMS ARE MAKE
THE GENERATION OF FAKE VOICES
EASIER THAN BEFORE THOUGH IT
STILL TAKES SOME TIME AND THESE
KIND OF FAKE VOICES CAN CAUSE
REAL DAMAGE TO CONSUMERS.
DEVELOPMENTS IN DIGITAL MEDIA
FOREENCE SICKS ARE PROVIDING
TECHNICAL TOOLS AND PROVIDE A
COUNTERMEASURE TO THE PROBLEM.
THE COMPETITION BETWEEN THE
MAKING AND DETACHING OF AI
SYNTHESIZED VOICE WILL BE GOING
ON FOR QUITE A LONG TIME AS A
CAT AND MOUSE GAME.
AS WE IMPROVE THE TECHNOLOGY THE
SYNTHESIS WILL ALSO IMPROVE.
BUT I THINK OVERALL LIKE
FIGHTING FAKE MEDIA IN GENERAL
AND DETECTING FAKE VOICES IS A
COMMUNITY EFFORT.
RESEARCHERS LIKE ME AND MY
COLLEAGUES WORK ON THIS BUT ALSO
REQUIRES INVOLVEMENT OF
COMMERCIAL COMPANIES DOMINANCE
AND USERS.
EVERYBODY TAKES THEIR DUE
DILIGENCE AND WE CAN HAVE THIS
PROBLEM UNDER CONTROL.
THAT'S MY HOPE.
THANK YOU VERY MUCH.
>> THANK YOU, PROFESSOR LYU AND
NOW WE'LL TURN TO JOHN ABOUT THE
DARPA PROGRAM.
>> INSIDE THE TECHNICAL ADVISER
FOR THE FORENSIC AND SEMANTICS
PROGRAM AT DARPA AND EVIDENCE
PRESENTED ARE BOTH PERFORMERS IN

OUR METAPHOR PROGRAM AND WHAT
METAPHOR IS ABOUT AROUND
DETECTING WHAT WE CALL
MANIPULATIONS IN IMAGE AND
VIDEO.
MANIPULATIONS COULD BE
COMPLETELY SYNTHETIC OR
PARTIALLY OR CHEAP FAKES,
PHOTOSHOP.
ADDING FRAMES, REDUCING FRAMES
TO SPEED UP VIDEO, ANYTHING IN
BETWEEN, ADDING, REMOVING
OBJECTS.
SO THERE'S DIFFERENT LEVEL OF
COMPUTE AND SKILL THAT IS
REQUIRED IN THESE DIFFERENT
AREAS.
HOPEFULLY YOU GOT THERE'S SKILL
INVOLVED IN BUILDING THESE FROM
SCRATCH.
IF WE'RE GOING TO BUILDING MORE
AND WE HEARD SNIPPETS OF
PRE-FORMED VERBIAGE OR A GROUP
OF SAMPLES FROM FIVE-SECONDS
EACH PRODUCING SOME KIND OF
SOUND.
ARE THEY COMPELLING?
ARE THEY GIVING CADENCE?
THERE'S OTHER FEATURES BEYOND
PITCH AND VOLUME THAT ARE
DYNAMIC TO OUR VOICE.
WE ARE LOOKING AT THE NEXT
PROGRAM COMING OUT OF DARPA THIS
SUMMER.
WE'RE LOOKING ACROSS ALL
MODALITIES.
YOU CAN'T LOOK AT AUDIO WITHOUT
OTHER CONTEXT.
WHAT STORY'S BEING TOLD.
WHAT'S THE MESSAGE BEING PLAYED.
WE'RE LOOKING AT VISUAL CONTENT
FROM THE SAME TYPE PERSPECTIVE.
WE CAN'T LOOK AT A VIDEO AND
ANALYZE THAT SUFFICIENTLY
WITHOUT CONSIDERING THE AUDIO
UNDERNEATH THE HOOD AND AUDIO'S

MUCH MORE THAN SPEED.
IT'S THE ENVIRONMENTALS AND
ATMOSPHERICS AND THE INFORMATION
CAPTURED THEREIN THAT TELLS US
ABOUT THE ENVIRONMENT.
I ALLUDED TO THIS AS THE
REVERBER ECHO THAT ENTERS OUR
EARS, OUR SENSORS THAT COLLECT
THIS INFORMATION THAT GIVES US
SOME KIND OF UNDERSTANDING.
WHAT I'M SHOWING YOU ON THE
SCREEN RIGHT NOW ARE SOME MANUAL
MANIPULATIONS THAT HAVE FOOLED
PEOPLE.
WE HAVE THE SHOOT DOWN OF MA-17.
THERE WAS A SMALL GROUP LOOKING
AT MEDIA AND THE RUSSIANS
REPORTED UKRAINIANS SHOT DOWN AN
AIRCRAFT.
THIS IS THE IMAGE THEY SHOWED.
IT WAS PROVEN THE AIRCRAFT THE
JUMP JUMP -- JUMBO JET ISN'T
THE SAME AND THE AIRCRAFT THAT
CLAIMED TO HAVE SHOT DOWN THE
JUMBO JET ISN'T THE SAME ONE
THAT WAS CLAIMED AND THIS IS AN
IMAGE FOUND ON GOGGLE EARTH AND
TURNED TO GRAY SCALE AND CLOUD
WERE ADDED USING PHOTOSHOP.
THE IMAGE A PROTEST AND THE FLAG
WAS ADDED AGAIN USING PHOTOSHOP.
IT'S A PROTEST POSSIBLY
SOMEWHERE ELSE IN THE WORLD BUT
THE FLAG WAS ADDED TO ELICIT AN
EMOTIONAL RESPONSE AND WE'RE
STARTING TO SEE THE SAME THING
WITH VIDEO AND AUDIO COMBINED
WITH THE IMPERSONATION OF OBAMA
AND JORDAN PEELE, FAKE SOCIAL
MEDIA AND HERE WE ARE LOOKING AT
AUDIO.
VOICE, CLONING VOICE.
CLONING IS ALONG THE LINE OF
SAMPLING.
I WANT TO TAKE SAMPLES OF
SOMEONE'S VOICE AND BUILD

SOMETHING UP FROM AND CLONE
THEIR INFORMATION OPPOSED TO
SYNTHETICALLY GENERATE I POINTED
TO A COMPANY AND IF I GO TO FAKE
JOE ROGAN.com YOU'LL GET
INTERESTING SAMPLES AND IT'S
HARD TO TELL THE REAL VOICE AND
SYNTHETIC RULE AND THEY'LL BOTH
FOOL SPEAKER I.D. SYSTEMS AND
THEY'LL BOTH SAYING THESE JOE

ROGGAN AND WE HAVE BERKELEY AND
OTHERS WORKING ON
DIFFERENTIATING SPEAKER AND
SYNTHETIC.
THIS IS AN EXAMPLE WHERE A U.K.
COMPANY WAS EXTORTED OVER
$200,000 EQUIVALENT BECAUSE
SOMEONE WAS ABLE TO SPOOF THE
CEO'S VOICE FOR THE WIRE
TRANSFER.
WE'RE LOOKING ACROSS MODALITY.
THE INTENT IS WE'RE EXPECTING
DISINFORMATION AND FAKE NEWS TO
BE AROUND US FOREVER.
IT HAS BEEN SINCE WE CAN
COMMUNICATE WHEN PEOPLE LIED TO
ONE ANOTHER.
IT'S NOT GOING CHANGE.
WE'RE GOING LIVE WITH IT.
WHAT CAN WE DO TO ENSURE WE HAVE
TRUST IN WHAT WE SEE AND HEAR?
WE STARTED THAT WITH METAPHOR
PROGRAM IN IDENTIFYING
MANIPULATED OR REAL IMAGE AND
VIDEO BUT WE NEED TO GO BEYOND
THAT BECAUSE A LOT OF THE
DETECTERS OF ANALYSES THEY'RE IN
RELY ON STATISTICAL FEATURES
FRAGILE DUE TO RECOMPRESSION.
HOW MANY TIMES CAN A VOICE GET
EXPRESSED OR VIDEO.
YOU GO FROM HIGH QUALITY YOUTUBE
TO CELL PHONE AND RE-ENCODE IT
SEVERAL TIMES THROUGH SOCIAL
MEDIA.

WILL WE HAVE DEGRADATION?
YES, I BET WE WILL BUT THERE'S
OTHER CHARACTERISTICS AND
FEATURES THAT SURVIVED AND WE
REFER TO THEM AS SEMANTIC
FEATURES.
IT TELLS US ABOUT THE CONTACTS
UNDER THE HOOD AND WHY IS THE
INFORMATION BEING CONVEYED FROM
SENSOR TO CONSUMER.
I'M SHOWING YOU ON THE BOARD A
TIME LINE OVER A FEW MONTHS OR
THE LAST TWO YEARS WHERE THERE'S
BEEN A HUGE ADVANCEMENT IN
GENERATIVE MEDIA.
GOOGLE DUPLEX HAS COME OUT WITH
A SYNTHETIC VOICE.
IT'S NOT INTENDED TO REPLACE A
SPEAKER'S VOICE BUT THERE'S
PAUSES, UMS AND AHS INTRODUCED,

CADENCE CHANGES.
THE INTENT IS TO USE THE
SYNTHETIC CHAT AS A WAY TO
CORRESPOND WITH OTHERS TO SET UP
YOUR AN APPOINTMENT AND GOOGLE,
MAKE ME AN APPOINTMENT AND THIS
PROXY WILL CALL ON YOUR BEHALF.
WE'VE SEEN FACES BEING DRIVEN BY
PUPPET MASTERS IN FACE TO FACE.
IT'S A TYPE OF DEEP FAKES.
DEEP FAKES THE WORD'S BEEN
HIJACKED TO MEAN ANYTHING
SYNTHETIC TO PARTIALLY SYNTHETIC
TO WHERE IT LOOKS REALISTIC BUT
ISN'T AND IS BECOME TERM SIMILAR
TO HOW WE ALL UNDERSTAND HATE
NENE PHOTO -- WHAT IT MEANS TO
PHOTOSHOP SOMETHING.
IT'S NOW A VERB.
DEEP FAKE IS BECOMING MEDIA,
VISUAL, AUDIBLE THAT RESEMBLES
SOMETHING ELSE THAT WE HOLD TO
BE TRUE BUT IT ISN'T.
WHAT'S INTERESTING ABOUT THIS IS
ANYTHING YOU SEE ON THE SCREEN

IS GENERATED.
THIS PERSON DOES NOT EXIST.
THAT AIRBNB DOES NOT EXIST.
THE TEXT DOES NOT EXIST.
BUT THERE ARE SEMANTIC ERRORS
THAT COMPUTERS GENERATE EITHER
IN THE CONSTRUCTION OF THE MEDIA
THAT WE CAN USE AS FEATURES FOR
DETECTION OR IN THE CONTEXT OF
THE TEXT FOR OTHER SEMANTIC
RELATIONSHIPS OF THE CONTEXT
THAT CAN BE TRIGGERED ON TO TELL
US WHETHER IT'S E REAL EVENT OR
REAL SITUATION OR NOT.
IT'S THE OBJECTIVE THE SEMANTICS
FORENSICS PROGRAM GOES TO ALL
MODALITIES, VIDEO, AUDIO, TEXT
BUT IN CONCERT IN TO WHAT
STORIES ARE BEING TOLD.
WHAT ARE THE THREADS OF TRUTH
BEING CONVEYED IN MESSAGES THAT
ARE SENT.
THIS SAY PROBLEM NOT JUST FROM A
PERSPECTIVE, THIS SAY WORLD WIDE
ISSUE.
SOCIAL MEDIA PLATFORMS ARE THE
FRONTLINE.
THAT IS WHERE MESSAGES ARE
GETTING CONVEYED LATCHED ON TO
AND DISSEMINATE.
ALL WE WANT TO DO AT DARPA AND
COUNTER FAKE NEWS AND
INFORMATION AT SCALE.
AND PLACE TRUST VALUE IN THE
MEDIA BEING PRODUCED AND STORIES
BEING TOLD?
THANK YOU.
>> THANK YOU, NEAL.
NOW WE'LL HEAR FROM JOHN WITH A
DEMO OF DETECTION TECHNOLOGY.
>> WELL, MY NAME IS JOHN AMEIN
WITH IR R&D.
IT MAY NOT BE THE SEXIEST NAME
IN THE WORLD BUT IT'S AT LEAST
VERY EFFICIENT BECAUSE IN FOUR
LETTERS WE TELL YOU WHAT WE DO,

IDENTITY, RESEARCH AND
DEVELOPMENT.
THAT'S OUR FOCUS.
SPECIFICALLY BIOMETRICS.
SO THE CAPABILITIES THAT WE HAVE
IN THE AREA OF BIOMETRICS ARE
INVOICE BIOMETRICS.
FOR EXAMPLE, WE TALKED A NUMBER
OF TIMES HERE ALREADY HOW YOU
CAN MATCH A VOICE.
YOU HEAR A VOICE SIT THAT
PERSON'S VOICE.
I HAVE ON MY PHONE AS I OPEN AN
APP HERE OLD CHICAGO WHITE SOX.
IT'S SUPPOSED TO MATCH ME AND
YOU CAN SEE THAT IT DID MATCH
ME.
THAT'S ONE EXAMPLE OF WHAT WE
DO.
HOWEVER, THERE'S A MAJOR PROBLEM
WITH JUST DOING THE MATCHING.
THE OBVIOUS THING BEING THAT IT
COULD BE SPOOFED.
SOMEONE CAN BE RECORDED OR
SYNTHESIZING YOUR VOICE AND ONE
THING WE PUT EFFORT IN IS VOICE
ANTI-SPOOFING.
OTHER THINGS WE DO, WE ALSO DO
FACE ANTI-SPOOFING.
IF YOU'RE DOING FACE RISK ABLE
TO KNOW WILL OR NOT IT'S A
PHOTOGRAPH OF A PHOTOGRAPH OR IF
THE CAMERA'S TAKING A PICTURE OF
A LIVE PERSON.
THESE ARE SAME KINDS OF THINGS
WE NEED TO DO IN THE FACE WORLD
AND THIS IS OUR BUSINESS.
THIS IS WHAT WE DO.
WE SELL THIS SOFTWARE TO ENABLE
OTHER COMPANIES TO CREATE
AUTHENTICATION SYSTEMS.
SO I'M NOT SURE HOW WELL YOU CAN
SEE THIS PARTICULAR CHART BUT
THE POINT OF THE CHART IS TO
SHOW YOU HOW YOU MEASURE THE
EFFECTIVENESS OF ANTI-SPOOFING

AND FORTUNATELY THERE'S A GLOBAL
ORGANIZATION CALLED INTERSPEECH
THAT INVESTS IN FIGURING OUT HOW
TO TEST THESE KINDS OF
ANTI-SPOOFING SYSTEMS TO MEASURE
THE EFFECTIVE NESS OF THEM.
THEY'RE RUNNING THIS COMPETITION
THERE WERE 49 COMPANIES AND ONE
FROM 2019 IN MARCH AND THERE'S
TWO AXES BECAUSE IF YOU'RE
TRYING TO DETECT A SYNTHETIC
VOICE WHAT CAN HAPPEN?
THE SYSTEM CAN GET IT RIGHT OR
WRONG.
HOW CAN IT GET IT WRONG?
ONE OF TWO WAYS.
ONE AXIS HERE I GOT IT WRONG
BECAUSE IT WAS SYNTHETIC VOICE
THAT I MARKED AS A REAL VOICE.
THE OTHER AXIS IS THE OPPOSITE.
IT WAS A REAL VOICE BUT I CALLED
IT SYNTHETIC.
SO THAT'S WHAT THAT CHART IS.
THE IDEA IS THAT YOU WANT TO BE
AS CLOSE TO ZERO AS POSSIBLE.
THESE ARE THE TWO KINDS OF
ERRORS AND EVERY SYSTEM WILL
INTEREST A TRADE-OFF BETiS:THE
TWO ERRORS.
YOU MEASURE THE SYSTEMS AND YOU
GET SCORES BACK AND BECAUSE THE
DATA IS LABELLED ON INPUT SIDE
YOU KNOW WHETHER OR NOT THE
SYSTEM WAS ACCURATE.
THAT'S THE PURPOSE OF THE
CHALLENGES.
AND YOU CAN SEE THAT WE WERE
HERE AT CLOSE TO ZERO WHICH IS
WHY I THINK WE GOT INVITED TO
THIS EVENT.
WE ARE THE MARKET LEADER IN
TERMS OF TECHNOLOGY THROUGH THIS
GLOBAL INTERSPEECH EVENT.
SO WHAT I WANT TO TALK ABOUT NOW
IS WELL, OKAY, HOW CAN YOU
DETECT THIS SYNTHESIZED VOICE.

IT TURNS OUT ONE OF THE FIRST
WAYS TO DO IT IS THE FACT THAT A
SYNTHETIC VOICE IS NOT COMING
FROM A BODY AND WE HAVE
TECHNOLOGY TO DETERMINE WHETHER
OR NOT SPEECH IS COMING FROM A
REPRODUCTION DEVICE, FLAT,
SURFACE LOUD SPEAKER VERSUS THE
HUMAN VOCAL TRACT AND THAT'S AN
OVER SIMPLIFICATION OF WHAT
MAKES THESE THINGS DIFFERENT BUT
IN FACT WE ARE ABLE TO MAKE THAT
DISTINCTION BETWEEN A PLAYED
BACK VOICE WHETHER SYNTHETIC OR
REAL AND A REAL HUMAN VOICE.
SO WITH THAT I'M ACTUALLY GOING
TO ASK YOU TO PLAY THE VIDEO.
I HAD TO MAKE A QUICK AND DIRTY
VIDEO OF ME SHOWING A FEW
EXAMPLES.
>> THE WORSE JOB OF FLYING
AROUND THE WORLD BEING AWAY FROM
MY FAMILY RECORDING 10 MINUTES
OF AUDIO.
>> THAT WAS ELLEN, YOU MAY HAVE
RECOGNIZED.
THAT WAS TOTALLY SYNTHESIZED.
>> AND THE WORSE PART OF MY JOB
IS BEING AWAY FROM HOME AND
RECORDING VIDEO.
>> THAT WAS ME.
>> MY VOICE MAY BE GENERATE
COMPUTER BUT I THINK IT SOUNDS
PRETTY HUMAN.
I DON'T KNOW EXACTLY HOW THEY
MADE IT BUT I'M REALLY
IMPRESSED.
>> THAT'S AN APP ON MY iPHONE IF
ANYBODY WOULD LIKE IT HAVE THAT
APP I CAN ENABLE IT FOR YOU AS A
TEST FLIGHT.
THOSE ARE ONE OF THE WAYS WE CAN
DETECT SYNTHETIC VOICE.
HOWEVER, A LOT OF THE EXAMPLES I
THINK WE TALKED ABOUT ALREADY
LIKE YOU'RE DOWNLOADING

SOMETHING FROM THE INTERNET OR
WHATEVER, THEN YOU'RE NOT GOING
TO BE PLAYING THAT BACK THROUGH
ANOTHER DEVICE THAT IS THEN
GOING TO BE ABLE TO DETECT IT.
IT'S A DIRECT PLAY.
FOR THAT WE HAVE TO TAKE ANOTHER
APPROACH.
THAT APPROACH IS TO LOOK AT THE
ARTIFACTS IN SYNTHETIC SPEECH.
SO THE PREVIOUS SPEAKERS
ACTUALLY TALKED ABOUT THAT.
THERE ARE ARTIFACTS AND WE LOOK
AT THEM AND I WISH I COULD TELL
YOU MORE OF EXACTLY WHAT WE'RE
LOOKING AT BUT I DO NOT WANT TO
DO THAT ON A PUBLIC RECORD SO
I'M NOT GOING TO OTHER THAN
HERE'S A VISUAL TO GIVE YOU AN
IDEA.
THERE'S SOME FREQUENCY DOMAIN
KINDS OF THINGS THAT ARE
DIFFERENT BETWEEN THE
SYNTHESIZED VOICE AND HUMAN
VOICE AND THERE'S ALSO CERTAIN
OTHER THINGS I THINK YOU
MENTIONED PHASE INVERSIONS.
THAT'S ONE OF THE THINGS WE'RE
LOOKING AT IN OUR TECHNOLOGY AND
SOME OTHER THINGS THAT FRANKLY
SYNTHETIC SPEECH IS A LITTLE TOO
PERFECT.
THAT'S ONE OF THE THINGS THAT
COMES THROUGH WHEN YOU DO THIS
ANALYSIS.
ONE THING THAT WORRIES ME A
LITTLE BIT BECAUSE I'M SURE THE
CLONING TECHNOLOGY WILL MAKE IT
LESS PERFECT SO IT'S MORE HUMAN.
THIS IS THE OTHER THING WE DO
WITH THE TECHNOLOGY.
THE WAY OUR ENGINE WORKS AND I
THINK THIS IS WILL BE TRUE OF
ANY ENGINE COMBINED NEURAL
NETWORKS LOOKING AT SPEECH AND
FUSING A RESULT AND PROVIDING A

SCORE.
THAT'S ALL OUR SOFTWARE DID.
YOU GIVE IT INPUT AND IT COMES
OUT WITH A SCORE AND PROBABILITY
OF HOW LIVE IS IT VERSUS NOT
LIVE.
ONE OF THE KEY THINGS THAT IS
REALLY CRITICAL I THINK IT CAME
UP SEVERAL TIMES IS DATA.
HAVE TO HAVE A LOT OF DATA.
THIS IS ONE OF THE REASONS WHY
WE WERE SUCCESSFUL IN THESE
TESTS IS BECAUSE THE AMOUNT OF
PUBLICLY AVAILABLE DATA IS BACK
IN LATE 2018, WE DIDN'T HAVE
THAT MUCH MORE DATA THAN WAS
AVAILABLE PUBLICLY WHERE YOU CAN
BE EITHER BUY IT OR RESEARCH
UNIVERSITIES AND PROVIDE THIS
KIND OF DATA IT EACH OTHER, THAT
SORT OF THING.
WE INVESTED BECAUSE IT'S OUR
BUSINESS IN A CORE TEAM OF
PEOPLE THAT COLLECTS DATA OF ALL
KINDS OF BIOMETRICS DATA AND WE
NOW HAVE A VERY LARGE CORPUS WE
USE AS PART OF OUR WORK.
THIS IS AN ONGOING EFFORT FOR
SURE.
SO THE LAST THING I THINK I CAN
MAKE THIS PLAY.
I DO HAVE SOME EXAMPLES OF
SYNTHESIZED SPEECH WE RAN
THROUGH THE SOFTWARE DIRECTLY
LIKE IN OUR API CALL AND
DETECTED ALL OF THESE ARE
SYNTHESIZED AND YOU HEARD SOME
OF THE TECHNOLOGIES MENTIONED
EARLIER FROM GOOGLE AND THE WAVE
NET FROM GOOGLE AND THE OTHER
AND I DON'T KNOW HOW TO MAKE
THEM PLAY.
>> MY VOICE MAY BE GENERATED BY
A COMPUTER BUT I THINK IT SOUNDS
PRETTY HUMAN.
I DON'T KNOW EXACTLY HOW THEY

MADE IT BUT I'M REALLY
IMPRESSED.
>> A SINGLE WAVE NET CAN CAPTURE
THE CHARACTERISTICS OF DIFFERENT
SPEAKERS WITH EQUAL FIDELITY.
NOW IT'S FAST.
GEORGE WASHINGTON WAS THE FIRST
PRESIDENT OF THE UNITED STATES.
SO THAT'S IT.
THIS IS THE DEMO I ALREADY
SHOWED YOU.
WITH THAT I GUESS I WANT TO SAY
I TOTALLY AGREE WITH THE
PROFESSOR EARLIER THIS IS GOING
TO BE A CAT AND MOUSE GAME.
THIS IS NEVER GOING TO END
COLLECTING DATA, IMPROVING
ALGORITHMS AND WORKING TOGETHER
IS WHAT IT'S GOING TO TAKE TO
DETECT AND PREVENT VOICE CLONING
FROM BECOMING A HAZARD AND
HARMFUL TO ALL OF US.
THANK YOU.
[APPLAUSE]
>> THANK YOU, VERY MUCH, JOHN.
ALSO I'LL MENTION IF YOU HAVE
QUESTION FOR THE PANEL AND PRIOR
PANELS IF YOU CAN WRITE THEM
DOWN I'LL TRY TO WORK THEM IN A
LIMITED AMOUNT OF TIME.
MY FIRST AND PREVENT VOICE CLONING
FROM BECOMING A HAZARD AND
HARMFUL TO ALL OF US.
THANK YOU.
(APPLAUSE)
>> THANK YOU VERY MUCH, JOHN.
IF YOU HAVE QUESTIONS FOR THE
PANEL, LIKE THE PRIOR PANELS, IF
YOU CAN WRITE THEM DOWN AND GET
THEM DELIVERED AND I'LL TRY TO
WORK THEM OUT.Ñi
>> SO NO MATTER HOW THE
DETECTION METHOD WORKS ACCURATE,
THE CLEAN EXAMPLES, THEY MAY OR
MAY NOT GENERALIZE THROUGH THESE
KIND OF CASES.

SO THE SECOND POINT I WANT TO
MAKE IS, YOU KNOW, THIS
EVALUATION MAY OR MAY NOT BE THE
UPKIND OF EVALUATION WE'RE
INTERESTED IN, YOU KNOW, ARE
RELEVANT.
WHICH PART OF THIS VOICE IS
WRONG, WHY WOULD WE THINK THIS
IS NOT A REAL VOICE?
SO THE KIND OF ACROSS-THE-BOARD
EVALUATION MATCHUP WE'RE USING
NOW PROBABLY IS NOT ACCURATE TO
REFLECT THE P PERFORMANCE OF
THIS ALGORITHM OF INDIVIDUAL
VOICE SAMPLES.
AND THEN I THINK, YOU KNOW, WHAT
NEIL IS SAYING IS LOOK BEYOND
THE SIGNAL.
WE LOOK FOR SOMETHING THAT IS
MORE SEMANTIC, MORE ROBUST AND
HARDER TO FAKE.
TO THOSE ARE, I THINK, MY
COMMENTS ON THIS.
>> SO YOU'RE BOTH TALKING
SOMETHING ABOUT -- YOU KNOW, THE
BACK AND FORTH AND SEMANTIC PART
OF DETECTION, DOES THAT MEAN YOU
THINK IT'S MORE LIKELY THAT WE
WOULD HAVE TO HAVE BACK AND
FORTH TO IDENTIFY WHERE
SOMETHING IS A.I. OR IS IT
POSSIBLE WE'RE GOING TO BE ABLE
TO USE THE ANALYTICS WITHIN THE
AUDIO TO DETERMINE WHETHER IT'S
A CLONE VOICE OR NOT?
>> CAN I THROW IN A GREAT "IT
DEPENDS"?
(LAUGHTER)
>> THAT'S WHERE I'M GOING.
ONE TO HAVE THE REASONS WE
WANT TO DO THAT IS A LOT OF THIS
IS FOCUSED ON SOCIAL MEDIA.
I THINK WE'VE LEFT OUT WHAT WAS
JUST HIT UPON AND THAT IS
ROBOCALLS AND TARGETED PHISHING
OVER THE PHONE WHERE AUDIO IS

GOING TO BE INCREDIBLY
COMPRESSED.
IT WON'T SOUND AS GOOD AS SKYPE
USUALLY DOES, CERTAINLY ANYTHING
ON MY MOBILE NEVER SOUNDS THAT
GOOD.
YOU HAVE VERY LITTLE CONTEXT
THERE.
BUT THINGS THAT WOULD HELP THERE
WOULD NOT JUST --
(TECHNICAL DIFFICULTY)
-- IN THAT IS IT CONSISTENT WITH
THE SOURCE THAT WE ARE EXPECTING
IT TO COME FROM.
EITHER THEñr STORYTELLER, THE NEWS
ORGANIZATION, THE AUTHOR OR THE
SPEAKER.
AND THEN THERE'S
CHARACTERIZATION.
YES, WE ARE EXPECTING DATA THAT
COMES OFF THE INTERNET TO BE
MANIPULATED, COMPRESSED,
ENHANCED SOME WAY OR DEGRADED.
DO WE CARE?
IS THE MANIPULATION OR
AUGMENTATION DONE WITH MALICIOUS
INTENT?
IS IT CHANGING THE STORY THAT'S
BEING TOLD?
IF SO, WHAT'S BEING DONE AND HOW
DO WE DETECT THAT.
VERY HARD PROBLEMS.
IT'S ONE REASON IT'S A DARKER
PROBLEM.
(LAUGHTER)
THANK YOU.
SO THE CHARACTERIZATION PIECE IS
JUST AS IMPORTANT AS THE
DETECTION PIECE BECAUSE IT HELPS
US NARROW DOWN WHAT WE NEED TO
FOCUS ON AND THIS IS ANOTHER
REASON WE CAN'T LOOK AT ONE
MODALITY WITHOUT ANOTHER WHEN
WE'RE CONSIDERING EVENTS AND
STORIES THAT ARE BEING CONVEYED,
RELAYED NEWS OR POTENTIAL

ACTIONS THAT HAVE HAPPENED
INTERNATIONALLY.
>> COULD I ADD TO THAT?
BECAUSE I THINK THE QUESTION
WAS AROUND CAN WE DETECT, AND WE
CANNOT DETECT PERFECTLY, AND
THAT'S -- BUT THAT'S TRUE OF
MANY THINGS.
MANY THINGS AROUND SECURITY.
IF YOU TALK TO ENTERPRISE
SECURITY EXPERTS,
EXPECTING 100% PERFECTION,
DETECTION AUTHENTICATION.
SO WHAT DO THEY DO?
THEY COMBINE MULTIPLE METHODS
AND WE STRONGLY ENCOURAGE HAVING
VOICE AND FACE FOR EXAMPLE AS
AUTHENTICATION BECAUSE IT WILLñr
NEVER BE PERFECT.
YOU MAY NOT BE ABLE TO DETECT
CLONED VOICE SO YOU MAY BE ABLE
TO COMBINE FACTORS.
ENTERPRISE SECURITY PEOPLE KNOW
THIS.
THEY PUT IN STRATEGIES TO
IMPLEMENT EXACTLY THOSE THINGS,
MAYBE WE HAVE TO DO MORE TO HELP
CONSUMERS BE ABLE TO LEVERAGE
THAT KIND OF CAPABILITY, TOO.
>> PROFESSOR TRAYNOR MENTIONED
THE PROBLEMS WITH THE VOICE
CLONES WITH ROBO CALLS AND
AUTOMATED CALLS.
RIGHT NOW WE HAVE ON THE MARKET
ANALYTICS PRODUCTS THAT ARE OVER
THE TOP, USING ANALYTICS TO
COMPLAINTS OR WHAT IS THE CALLER
I.D. NUMBER, WHAT DO WE KNOW
ABOUT THE CALL, SIT POTENTIALLY
SCAM OR A FRAUD.
COULD THE KIND OF I.D.
TECHNOLOGY THAT YOU CAULKED
ABOUT HERE BE IMPLEMENTED WITH
THE OTHER ANALYTICS TO TRY TO
CREATE A BETTER PICTURE OF
WHETHER A CALL THAT'S COMING IN,

IFçó LISTENING TO THE AUDIO OF THE
MESSAGE THAT'S PLAYING, WHETHER
OR NOT THAT'S LIKELY A SCAM CALL
OR NOT?
>> WELL, I THINK THE SYSTEMS CAN
ABSOLUTELY DETECT IF IT'S TEXT
IF THAT'S HELPFUL, IT HAS TO
LISTEN TO SOME OF IT BEFORE IT
CAN MAKE THAT DECISION, THOUGH.
AND I'M FINDING MORE AND MORE --
I GET LEGITIMATE CALLS THAT ARE
ALSO TEXT TO SPEECH BECAUSE
COMPANIES ARE FINDING IT EASIER
TO SEND ME MESSAGES AND ALERT ME
THAT MY PACKAGE IS THERE,
WHATEVER, AND THEY'RE USING TEXT
TO SPEECH.
SO, YOU KNOW, I DON'T KNOW, IT
MIGHT BE QUESTIONABLE WHETHER OR
NOT YOU CAN ACTUALLY DEPLOY THAT
BECAUSE IT'S BEING USED FOR GOOD
REASONS AS WELL AS BAD ONES.
>> ANOTHER WAY TO LOOK AT THAT,
THOUGH, IS EMAIL.
WE HAVE GOOD EMAIL AND BAD
EMAIL.
WE HAVE SPAM EVERY SINGLE DAY.
THAT'S FAKE INFORMATION.
DEPENDING ON THE SERVERS
PROVIDERS ARE USING, THERE ARE
DIFFERENT WAYS OF HAVING SPAM
FELT REALS DEPLOYED EITHER
AUTOMATICALLY OR USER-SELECTED
FILTERING.
I SEE THE SAME TYPE OF APPROACH
CAN BE USED WITH DISINFORMATION
OF THIS TYPE AND PROVIDING
INDICATORS AND FLAGS AS A WAY TO
INFORM WHETHER WE BELIEVE THIS
IS REAL OR NOT.
IN SOME CASES, YOU POINTED OUT,
YES, I'VE GOTTEN THE AUTO
INSURANCE CLAIM THAT THEY WILL
BE SHOWING UP BETWEEN THE HOURS
OF 8:00 P.M. AND 6:00 P.M. TO DO
SOMETHING.

ON THE OTHER HAND, I ALSO HAVE
MY CELL PROVIDER HAS RECENTLY
IMPLEMENTED A FEATURE THAT
ALERTS ME TO POSSIBLE SPAM WHEN
NUMBERS ARE COMING IN.
>> SPINAL LIKELY?
AM LIKELY, YEAH.
I LOVE THAT.
I HAVE A CHOICE.
DO I ANSWER THE CALL AND PLAY
WITH THE ROBOVOICE?
OR DO I LET IT GO TO VOICEMAIL,
THEY'LL LEAVE ME A VOICEMAIL,
MAYBE, IF IT'S IMPORTANT, OR IT
WILL DROP.
GIVE THE CONSUMER SOME KIND OF
EDUCATION AND OPTION, I THINK IS
ONE WAY TO GO.
>> WE HEARD, YOU KNOW, SOME OF
THE EXAMPLES EARLIER ABOUT THE
JOE ROGEN RECORDINGS, FOR
INSTANCE, YOU KNOW, THAT I
LISTENED TO YESTERDAY AND I GOT
WRONG.
I -- IN FACT, I NOT ONLY GOT
THEM WRONG, I WAS CONFIDENT I
WAS RIGHT WHEN I MADE THE GUESS
AND I WAS COMPLETELY WRONG.
THE TECHNOLOGY CLEARLY EXISTS,
WHAT ARE THE ODDS -- OR IS IT
POSSIBLE THAT CONSUMERS COULD
NOW BE RECEIVING CALLS FROM
PEOPLE USING EFFECTIVELY CLONED
VOICES AND THE VICTIMS OF A
GRANDPARENT SCAM OR SOMETHING
>> LIKE TODAY.
ABSOLUTELY.
 LOT OF THESE CALLS COME
FROM PRE-RECORDED SCRIPTS
ANYWAY, BARRIER TO ENTRY.
>> RIGHT.
I WILL SAY THEY ARE -- THE
CURRENT TTS TECHNOLOGY MAY NOT
HAVE A LEVEL OF CONVERSATION.
SO IF PRE-RECORDED, THAT CAN BE
MADE CONVINCING, BUT WHEN YOU

TALK WITH THEM, THEY CANNOT RUN
AS REAL TIME AS -- AS FAST AS
REAL TIME.
THAT'S THE CURRENT PROBLEM.
>> WHERE DO YOU THINK THE
TECHNOLOGY TO IDENTIFY VOICE
CLONES AND FAKED VOICES WILL BE
IN FIVE YEARS?
IF THIS TECHNOLOGY IS DEVELOPING
AS FAST AS IT SOUNDS LIKE AND
IT'S GETTING BETTER AND BETTER
AND EVERY SIX MONTHS OR A YEAR
IT'S GOING TO BE MORE EFFECTIVE
AND HARDER TO DISTINGUISH, IF
YOUR PHONE IS RINGING IN FIVE
YEARS AND IT SOUNDS LIKE YOUR
GRANDMOTHER ON THE PHONE, IS IT
LIKELY THAT THERE'S GOING TO BE
SOMETHING THAT YOU CAN IMPLEMENT
EITHER ON YOUR PHONE OR AT THE
NETWORK LEVEL OR SOMEWHERE THAT
IS GOING TO TELL YOU WITH SOME
DEGREE OF CONFIDENCE WHETHER
THAT IS A CLONED VOICE OR A REAL
HUMAN VOICE?
>> WELL, IN FIVE YEARS, YOU WILL
BE COMING DOWN TO THE END OF THE
SEMANTIC FORENSICS PROGRAM, AND,
BY THAT TIME, WE'LL DEFINITELY
HAVE TRANSITIONABLE CAPABILITIES
JUST AS WE'RE HAVING WITH MEDIA
FORENSICS RIGHT NOW.
I BELIEVE IT'S VERY PROMISING.
I'M EXCITED IN THE REACHES THAT
THE RESEARCH TEAMS ARE DOING ON
THE MEDIA FORENSICS SIDE AND
JUST THE SCOPE AND BREADTH THAT
WE ARE GOING AFTER ON THE
SEMANTIC FORENSICS, I'M REALLY
EXCITED TO SEE WHAT THE NEXT
FOUR AND A HALF YEARS HOLD ON
THAT PROGRAM.
>> WHAT ABOUT YOU, PROFESSOR
TRAYNOR, JOHN, DO YOU HAVE ANY
THOUGHTS ON WHAT WE'RE GOING TO
SEE IN THE NEXT FIVE YEARS?

>> I THINK CERTAINLY ADVERSARIES
AREN'T GOING TO GET ANY WORSE.
SO WE'LL SEE THAT SOPHISTICATION
CONTINUE.
ULTIMATELY, IT WILL BE A
COMBINATION OF TECHNIQUES AND
CONTEXTINGS THAT WILL GET US
THERE.
WE HAVE THE SAME CONVERSATION.
THE DETAILS ARE DIFFERENT.
IF WE COULD WALK AWAY WITH A
PIECE OF ADVICE, IF SOMEBODY
CALLS YOU URGENTLY AND YOU THINK
IT'S IMPORTANT, HANG UP AND CALL
THEM BACK, UNLESS IT'S A STATE
ACTOR WHO CAN REROUTE PHONE
CALLS OR A VERY, VERY
SOPHISTICATED HACKING GROUP,
CHANCES ARE THAT'S THE BEST WAY
NOW TO FIGURE OUT IF YOU WERE
TALKING TO WHO YOU THOUGHT YOU
WERE.
>> MY KIDS AND I HAVE SAFETY
WORDS.
IF I GET AN URGENT CALL, AND
THIS HAPPENED WITH MY SON, I GOT
AN URGENT CALL, YOU KNOW, AND I
ASKED HIM A QUESTION THAT I KNEW
HE WOULD ONLY HAVE THE ANSWER
TO, AND IT WAS SPUR OF THE
MOMENT, AND HE ANSWERED IT AND
HE SAID, DAD, WHAT WOULD YOU DO
IF I DIDN'T ANSWER THAT QUESTION
RIGHT?
I SAID, I WOULD TURN ON THE
LOCATORRER AND COME GET WHOEVER
HAD YOUR PHONE.
(LAUGHTER)
>> YEAH, I THINK THAT'S -- SO WE
HAVE BEEN TALKING ABOUT HOW IT'S
FAKED.
THE OTHER SIDE IS
AUTHENTICATION, WHAT IS REAL.
SO YOU CAN DO SOMETHING, AS YOU
CAN IMAGINE, FOR ALL THE REAL
VOICES, WE CAN TRACK SOME SPHESK

FEATURES AND WHEN THERE'S NEVER
A VOICE COMING IN, ASSUMING TO
BE THAT IDENTITY, WE COMPARE
THAT FEATURE.
THIS IS IS TRADITIONAL BYO
METRIC FRAMEWORK FOR
AUTHENTICATION.
SO I STRONGLY AGREE WITH JOHN
THAT THE SOLUTION WILL NOT COME
BY ONE METHOD OR NOT EVEN BY
JUST DETECTION.
IT SHOULD BE CHARACTERIZATION,
AUTHENTICATION ALL PUT TOGETHER.
SO I THINK BY FIVE YEARS WE'LL
REACH AN EQUILIBRIUM WHERE THE
CONSUMER WILL ALSO BE
SUBJECTCATED.
THE MERE NUMBER OF FAKE AWED
OWES, BUT THEY ALSO WILL GIVE AN
IMMUNE SYSTEM TO THESE THINGS.
HUMANS ARE AMAZING TO ADAPT TO
THIS KIND OF SITUATIONS.
LIKE THINK ABOUT SPAM E-MAILS,
AT THE VERY BEGINNING WE WERE
VERY AFRAID THEY WERE GOING TO
DOMINATE THE E-MAILS, OR THIS
RECENT PHOTOSHOP THE IMAGES OR
EVEN THE DEEP FAKE VIDEO, PEOPLE
ARE BECOMING BETTER AT
IDENTIFYING THEM.
SO I'M ASSUMING SIMILAR
SITUATIONS WILL PLAY OUT FOR
FAKE VOICES, ALSO, IN THE COMING
YEARS.
>> GREAT, I SEE OUR RED LIGHT IS
BLINKING SO WE'VE COME TO THE
END OF OUR TIME.
PLEASE JOIN ME IN THANKING THIS
FANTASTIC PANEL.
(APPLAUSE)
AND I'M GOING TO INTRODUCE
LAURA, THE ASSOCIATE DIRECTOR IN
F.T.C.'S DIVISION OF LITIGATION
TECHNOLOGY AND ANALYSIS TO GIVE
A FEW BRIEF CLOSING REMARKS.
>> THANK YOU, WILL.

MY COLLEAGUES LOIS GREISMAN
PLANNED TO GIVE CLOSING REMARKS
TODAY AND I THINK IT'S SOMEWHAT
APROPOS THAT SHE LOST HER DAY
THE DAY BEFORE HER VOICE CLONING
TECHNOLOGY WORKSHOP, BUT
FORTUNATELY IT'S JUST LARYNGITIS
AND I WILL NOT BE USING ANY
TECHNOLOGY SOUND LIKE HER.
WHENEVER THERE'S CUTTING EDGE
TECHNOLOGY, I TEND TO MAKE A
SILLY REFERENCE TO 1970s
"STAR TREK."
MY STAFF, HOWEVER, REMINDED ME
THAT, FOR THIS, YOU CAN GO A
LITTLE BIT FURTHER BACK IN TIME.
GO ALL THE WAY BACK TO THE
ODYSSEY.
HOMER WRITES HELEN NOW THWARTED
THE INVASION OF TROY BY
BECKONING THE SOLDIERS HIDDEN
INSIDE THE TROJAN HORSE.
SHE BECKONED THE SOLDIERS TO
COME OUT MIMICKING THE VOICE OF
THE SOLDIERS' WIVES.
ODYSSEUS RECOGNIZED THAT THE
VOICES WERE BEING HIM BINGED AND
KEPT THE SOLDIERS FROM EXITING
THE HORSE.
SO FAST FORWARD A FEW YEARS, AND
HERE WE ARE TODAY, AND AS OUR
PANELISTS AND SPEAKERS HAVE
MENTIONED, PEOPLE HAVE BEEN
MIMICKING VOICES FOR YEARS, BUT
JUST IN THE LAST FEW YEARS,
TECHNOLOGY HAS ADVANCED TO THE
POINT WHERE WE CAN CLONE VOICES
AT SCALE USING A VERY SMALL
AUDIO SAMPLE.
I APPRECIATE OUR FLANNELLISTS
GIVING US REAL, CONCRETE
EXAMPLES OF THE BENEFITS AND THE
KERNS OF THE TECHNOLOGY.
WE HEARD A VERY POWERFUL EXAMPLE
OF THE USE OF TECHNOLOGY TO HELP
PEOPLE WHO HAVE LOST THE ABILITY

TO SPEAKERS COMMUNICATE WITH FAMILY MEMBERS, INCLUDING OUR FOUR-LEGGED FAMILY MEMBERS AND, AS WELL, USING THESE TYPES OF TECHNOLOGIES TO HELP PEOPLE WHO HAVE NEVER BEEN ABLE TO SPEAK AND, ALSO, POTENTIAL FOR USING TECHNOLOGIES LIKE THIS TO INCREASE THE DIVERSITY OF VOICES THAT WE HEAR IN A VARIETY OF SITUATIONS.

BUT THE CONCERNS ALSO ARE REAL. THE IMPACT FOR EXAMPLE ON VOICE ARTISTS WHO, IF THEIR VOICE IS USED, WILL THEY BE COMPENSATED? IS IT WITH THEIR CONSENT?

AND THEN, OF COURSE, THE USE OR POTENTIAL USE BY CRIMINALS, WHETHER IT'S FOR HARASSMENT OR FRAUD.

CRIMES DEPEND ON COMMUNICATION AND THE USE OF THESE TECHNOLOGIES, ESPECIALLY COMBINED WITH VOICE COMPRESSION PHONE CALLS, CAN MAKE IT MORE CONVINCING, MAKE A CRIMINAL SOUND MORE CONVINCING.

THIS COULD JUST RESULT IN A GREATER AMOUNT OF CONSUMER INJURE.

SO THE ETHICAL CONSIDERATIONS OF VOICE CLONING TECHNOLOGIES ARE PENNILESS, TALKING ABOUT THE IMPORTANCE OF BUILDING IN ETHICAL PROCESSES DURING TECHNOLOGY DEVELOPMENT.

AND WE HEARD EXAMPLES OF THAT HAPPENING TODAY.

ALSO, OUR PANELISTS TALKED ABOUT THE IMPORTANCE OF INVOLVING USERS, HAVING COMPANIES TAKE END USERS' VIEWS AND VALUES INTO ACCOUNT, INVOLVING THEM IN THE DISCUSSION, AND THE NEED FOR ALL PLAYERS -- CIVIL SOCIETY, PRIVATE SECTOR, GOVERNMENT -- TO

BE INVOLVED IN THE DISCUSSIONS
OF THE ETHICAL USES OF
TECHNOLOGY.
FINALLY, OUR PANEL WHICH WE JUST
HEARD FROM, I WON'T TRY TO
EXPLAIN THE TECHNOLOGY THAT YOU
DESCRIBED, BUT THE TECHNOLOGIES
THAT ARE BEING DEVELOPED TO HELP
DETECT CLONED OR A.I.
SYNTHESIZED VOICES.
THE IMPORTANT PART, OUR PANELS
TALKED ANT WHAT I EXPECT -- AND
THIS HAS BEEN THE CASE OF OTHER
EMERGING TECHNOLOGIES -- IS
THERE CAN BE A CAT AND MOUSE
GAME.
AS TECHNOLOGY GETS BETTER SO DO
CLONING TECHNOLOGIES AND SO
FORTH.
BUT WHAT I HEARD AT THE END IS
COMBINE AGO NUMBER OF FACTORS
INCLUDING DETECTION TECHNOLOGIES
TO HELP EDUCATE CONSUMERS OR HOW
THEY MIGHT DETECT CLONED VOICES.
SO, ALL IN ALL, I WOULD LIKE TO
THANK ALL OF OUR PANELISTS
TODAY, ESPECIALLY PATRICK
TRAYNOR FOR LEADING OFF OUR
SESSION.
WE APPRECIATE THE TIME AND YOUR
PARTICIPATION IN THIS EVENT.
THANK YOU TO THE F.T.C. STAFF,
MOST IMPORTANTLY, INCLUDING OUR
EVENTS PLANNING STAFF, OUR MEDIA
TEAM, OUR PARALEGALS, FOR
EVERYONE WHO MADE THIS EVENT
POSSIBLE.
IN THE COMING WEEKS OF VIDEO AND
A ROUGH TRANSCRIPT OF THE
WORKSHOP WILL BE AVAILABLE AT
F.T.C..GOV.
IF YOU DIDN'T GET ENOUGH TODAY,
YOU CAN REVISIT IT AND SEE IF
YOU CAN GET THE EXAMPLE CORRECT.
THANK YOU ALL FOR COMING.
(APPLAUSE)