

## Realizing the Promise of AI

*United States Telecommunications Training Institute* December 10, 2020

Please Do Not Reproduce Without Permission



#### Agenda

- 14:00-14:05 Introduction (Jamal Anderson, Carolyn Nguyen)
- 14:05-15:05 Trustworthy AI: principles, governance and enabling policy (Carolyn Nguyen)
- 15:05-15:15 Break
- 15:15-15:55 AI and COVID-19 (Geralyn Miller)
- 15:15-16:15 Privacy-preserving technologies (Geralyn Miller)
- 16:15-16:55 Al use cases in the real world (Ted Way)
- 16:55-17:10 Summing up and Q&A



# Trustworthy AI: principles, governance and enabling policy

M-H. Carolyn Nguyen Director, Technology Policy, Microsoft

- 1. Al overview
- 2. Principles for trustworthy AI
- 3. Responsible AI at Microsoft
- 4. Shaping AI policy development
- 5. AI policy considerations



### Artificial intelligence



#### Economic Impact of Artificial Intelligence



#### Computers Understanding the World



### Accelerating Al innovation



### Project Tokyo: inclusive design

Cecily Morrison, MBE, Microsoft Researcher

Visual agent technology that extends the capabilities of people who are blind and low-vision by dynamically lighting up the social landscape and enable more natural social interaction.



#### Advances That Make AI Real

#### Vast amounts of data



#### Huge computational power



#### Al as a Cloud Service



**Software as a Service (SaaS)** Cloud applications (e.g., AI, cognitive services)

**Platform as a Service (PaaS)** On-demand application-hosting environment (e.g., Azure)

Infrastructure as a Service (laaS)
Basic compute, network and storage resources

### The Promise of Artificial Intelligence

Al for Good



#### FarmBeats

FarmBeats provides farmers with access to Microsoft Cloud and Al technologies, enabling data-driven decisions to improve agricultural yield, lower overall costs, and reduce the environmental impact of agricultural production





.............

Local Farm Computer Windows 10 IoT-based Azure IoT Gateway performs computations and AI at the FarmBeats Edge in the farmer's house/office



.........

.............

\*\*\*\*\*\*\*\*\*\*\* \*\*\*\*\*\*\*\*\*\*\*\*

X

Low-cost Tethered Eye continuously stream images and video of the farm to the Azure IoT Gateway in places where drones are not permitted



Soil Sensor Solar-powered soil sensors gather moisture, soil chemistry, and temperature data

**Azure Cloud** 

Machine learning algorithms integrate sensor data with aerial imagery and crop predictions

Drones

Autopilot drones gather

aerial imagery (visible

and multi-spectral)

..... ............ .......... \*\*\*\*\*\*\*\*\*\*\*\*

X



Algorithms process data to create a detailed heat map of the field

............

\*\*\*\*\*\*\*\*\*\*

Farmer Heat map data is sent to a farmer who applies insight to increase efficiency and conserve resources

### Design AI to amplify human ingenuity

### Principles and Policies for Trustworthy Al



#### Design AI to Earn Trust



Al isn't just another piece of technology. It could be one of the world's most fundamental pieces of technology the human race has ever created.

#### — Satya



#### Fairness

Understand how bias can be introduced and affect recommendations

Attract diverse pool of AI talent

Develop analytical techniques to detect and eliminate bias

Human review and domain expertise

# The New York Times

#### Facial Recognition Is Accurate, if You're a White Guy

By STEVE LOHR FEB. 9, 2018



Gender was misidentified in up to 1 percent of lighter-skinned males in a set of 385 photos.



Gender was misidentified in up to 12 percent of darker-skinned males in a set of 318 photos.



Gender was misidentified in up to 7 percent of lighter-skinned females in a set of 296 photos.



Gender was misidentified in 35 percent of darker-skinned females in a set of 271 photos.

### Reliability & Safety

Evaluate training data, data provenance and integrity

Test extensively (and enable a user feedback loop)

Monitor ongoing performance

Design for unexpected circumstances including nefarious attacks

Human in the loop



### Privacy & Security

Comply with existing privacy laws (e.g. the General Data Protection Regulation)

Provide transparency about data collection and use

Use de-identification techniques to promote both privacy and security

Design systems to protect against bad actors



#### Inclusiveness

Inclusive design practices to address potential barriers that could unintentionally exclude people

Enhances opportunities for those with disabilities

Build trust through contextual interaction

EQ in addition to IQ



#### Transparency

People should understand how decisions were made

Provide contextual explanations

Make it easier to raise awareness of potential bias, errors and unintended outcomes



#### Accountability

People must be accountable for how their systems operate

Norms should be observed during system design and in an ongoing manner

Role for internal review boards



#### Today's Workplace Transformed

Changing nature of work and jobs

The on-demand economy

Focus on skills training



### Al Policy Development

- Responsible AI at Microsoft
- Shaping AI policy development



#### Our Approach to Responsible Al





Sensitive Uses of AI



AI Reliability and Safety

Human-Al

Collab and Interaction



Fairness and Bias Intelligibility & Explainability

Engineering Practices for Al



Human Attention & Cognition

### **AETHER Committee** AI Ethics and Effects in Engineering and Research

#### The Office of Responsible AI

#### Internal policy

Set company-wide policies and practices for responsible AI.

Ensure that roles and responsibilities are defined clearly.

#### **Enablement**

Ensure readiness to adopt responsible AI practices within Microsoft.

Support our customers and partners to do the same.

#### Case management

Operate intake and triage function for Sensitive Uses.

#### Public policy

Formulate and advocate for responsible AI policies externally.



#### Mission:

- Develop and share best practices
- Advance public understanding
- Provide an open & inclusive platform for discussion & engagement
- Foster AI for socially beneficial purposes



### Al Policy Framework

Enable broad deployment and continued innovation

- Practical principles for trustworthy and humancentered Al
- Sharing of best practices
- Voluntary, consensus-based, market-driven standards
- Sectoral, technology-neutral regulation if necessary
- Multi-stakeholder dialogues

A regulatory environment that fosters trustworthy Al innovation and adoption



#### Global AI Policy Development: Competitiveness and Leadership

- National AI strategies
  - Economic growth & innovation
  - Al research, infrastructure, platforms
  - Al governance commission & oversight
  - Education & awareness
  - Skills development
  - Strategic relationships
  - National security
- Multilateral organizations
- Bilateral and regional collaborations























Inter-American Development Bank



### UN Roadmap on Digital Cooperation

"Supporting Global Cooperation on AI: AI must be trustworthy, human rights based, safe, and sustainable and promote peace".

Create a multi-stakeholder advisory body on Global AI cooperation to:

- Build global capacity for the development and use of AI
- Address the lack of representation and inclusiveness in global discussions
- Harness AI to support the Sustainable Development Goals



### OECD Principles and Recommendations for Artificial Intelligence

- OECD countries (36) + Argentina, Brazil, Colombia, Costa Rica, Peru, Romania adopted 22 May 2019.
- G20 countries adopted Principles 9 June 2019.
- Principles for responsible stewardship of trustworthy AI:
  - 1. Inclusive growth, sustainable development and well-being
  - 2. Human-centred values and fairness
  - 3. Transparency and explainability
  - 4. Robustness, security and safety
  - 5. Accountability
- National policies and international co-operation for trustworthy AI:
  - 1. Investing in AI research and development
  - 2. Fostering a digital ecosystem for AI
  - 3. Shaping an enabling policy environment for AI
  - 4. Building human capacity and preparing for labour market transformation
  - 5. International co-operation for trustworthy AI





#### OECD: Moving from Principles to Practice



- Platform to share and shape public policies for responsible, trustworthy and beneficial AI
  - Multi-disciplinary
  - Evidence-based analysis
  - Global multi-stakeholder partnerships
  - Real-time data
- COVID-19 watch
- International cooperation
- OECD Network of Experts on AI (ONE AI)
  - 1) Classifying AI systems
  - 2) Implementing trustworthy AI
  - 3) National AI policies
  - 4) Al Compute



#### OECD.AI:

Evidence-based policy making with Microsoft Academic Graph & LinkedIn Economic Graph



#### **Between-country AI skills migration** Year: Resume # items: All 2015 2016 2017 2018 2019 -2,500 -2,000 -1,500 -1.000-500 1,000 1,500 2.000 2,500 Al skills migration (per 10 000 LinkedIn members) Luxembour United Arab Emirate Ireland Canada Singapore Netherland German Switzerland Australia Japar Saudi Arabia Sweder Norwa Finland Austria United Kingdon Oata Cyprus Thailand United States Belgiun Czech Republi New Zealand France Denmark Portuga Spair Hungary ndonesia Israel Korea Poland Hong Kong (China) Colombia China (People's Republic of) Italy Romania Mexico Malaysia Chinese Taipei Argentina Brazil South Africa Greece Ukraine India Viet Nam Turkey Egypt Pakistan Bangladesh Iran Tunisia Venezuela

Note: Linearly average from 2015 to 2019 for a selection of countries with 100 000 LinkedIn members or more. Migration flows are normalized according to LinkedIn membership in the country of interest. Please see methodological note for more information.

#### Source of data: LinkedIn Economic Graph.

Please cite as: OECD.AI(2020), visualisations powered by JSI using data from LinkedIn, accessed on 2/12/2020, www.oecd.ai

### Al Policy Considerations



### Facial Recognition

Government regulation is needed

Principles Microsoft adopted:

- Fairness
- Transparency
- Accountability
- Non-discrimination
- Notice and consent
- Lawful surveillance

#### U.S. State of Washington law passed Mar 2020



#### Privacy and General Data Protection Regulation

Privacy is fundamental to trust and adoption of AI

- Comprehensive data protection regulation enables responsible Al innovation
- General Data Protection Regulation (GDPR) principles
  - Transparency and accountability
  - Individual empowerment and control
  - Corporate responsibility
  - Enforcement and redress
- Additional considerations
  - Data minimization may lead to bias
  - Encourage use of pseudonymized data and privacypreserving technologies
  - Legitimate interest in data processing
  - Contextual use of data
- Global interoperability facilitates cross-border data flow



#### Trusted Data Sharing

Access to data is fundamental to AI development

- There will be lots of data\*
  - 2025: 75 billion IoT devices; ~463 exabytes of data generated per day, most from sensors (200M DVDs)
  - 2030: 90% of population older than six will be online
- Value of data
  - Use of public-sector data can add 0.1%-1.5% of GDP, and 1%-2.5% of GDP when private-sector data are included^
  - Informed policy-making
- Growing data divide
- Trusted data sharing requires holistic approach
  - Technology securing data at rest, in transit, and in use
  - Policy & regulatory framework
  - Tools to enable data sharing
  - Data-sharing culture


### Cross-border data flow

- Challenges to "data free-flow with trust" (DFFT)
  - Lack of interoperable privacy frameworks & Schrems II
  - Unlimited government access to data
  - Data localization measures
- People are concerned about privacy & security of their data and (foreign) government access
- Companies need regulatory certainty to protect customers data & fulfill potential of cloud
- Law enforcements are concerned about national security
- Some principles for consideration:
  - Internationally interoperable legal regimes that minimize conflicts of law, with mechanisms to resolve conflicts.
  - Data minimization in seeking access to data.
  - Transparency & accountability on government access to data.



### Ensuring Al Empowers Us All

### Technology + Policy

### **OECD Integrated Policy Framework**



### Key Learnings

#### Trust is foundational to realizing the potential of AI

- All stakeholders have roles in the development and deployment of trustworthy Al
- Organization-wide governance is essential to developing a responsible-AI culture
- The OECD *Principles for responsible stewardship of trustworthy AI* is the most widely accepted set of principles by governments around the world and is a good basis for any national/regional AI strategy
- Comprehensive privacy and data protection regulation enables responsible-AI innovation
- Trusted data sharing requires a holistic approach and a data sharing culture
- Cross-border data flow with trust enables full cloud and AI potential
- A whole-of-government approach can ensure that AI empowers us all



# Thank you!



© 2013 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

### Microsoft Resources

- The Future Computed: Artificial Intelligence and its Role in Society (<u>https://blogs.microsoft.com/blog/2018/01/17/future-computed-artificial-intelligence-role-society/</u>)
- The Future Computed: AI and Manufacturing (<u>https://news.microsoft.com/futurecomputed/</u>)
- Al resources: <u>www.microsoft.com/ai</u>
  - "Learn": AI Business School, AI School, AI Lab, Experience AI
  - "Approach": Responsible AI
- Digital skills: <u>www.microsoft.com/digital-skills</u>
- Al for good: <u>www.microsoft.com/ai/ai-for-good</u>
- Al research: <u>www.microsoft.com/research/research-area/artificial-intelligence</u>
- OECD AI Policy Observatory: <u>oecd.ai</u>



# AI for Health

Microsoft

Geralyn Miller Health Strategist Al for Good Research Lab Microsoft geramill@microsoft.com

### Al for Good

Empowering the people working to solve humanity's greatest issues

Al for Accessibility Al for Earth Al for Humanitarian Action Al for Cultural Heritage Health

# Health is a global issue

### Technology's role:

- Advancing research
- Democratizing data
- Protecting against health crises
- Improving access to care for underserved populations



### AI for Health

A \$60 million, five-year philanthropic program created to empower those working to tackle some of the toughest challenges in health.

## **Collaborating with grantees**





A foundation of privacy, security, and ethics

 $\bigotimes$ 

Developed in collaboration with leading health experts

ිරිද්

Open source models shared with research community

### Focus areas

#### **Quest for Discovery**

Accelerating medical research to advance the prevention, diagnoses, and treatment of diseases.

#### **Global Health Insights**

Increasing our shared understanding of mortality and longevity to protect against global health crises.

### **Health Equity**

Improving access to care for underserved populations.



Al for Health partners and grantees

### Grantee: Novartis Foundation

### Working to eliminate leprosy

- 200,000 new cases diagnosed yearly
- Early diagnosis helps improve prognosis and limit disease transmission
- Developing AI-enabled digital tool to classify images of lesions
- Tool enables referral to healthcare professionals for early diagnosis
- Through the Global Partnership for Zero Leprosy, the imagery database and AI code will be made publicly accessible to empower leprosy researchers around the world.



### **Novartis Foundation**

# AI for Health COVID-19 Response

### Impact areas

- Data and insights
- Treatment and diagnostics
- Allocation of resources
- Dissemination of accurate information
- Basic scientific research





# Accelerating research and keeping communities safe

### **Progress to date**

- Awarded 150+ grants since April 2020
- Research length sometimes reduced from months to days
- Expedited work helps policymakers make data-driven public health and safety decisions



**COVID-19** partners and grantees

### Grantee: Folding@home

# Working across borders to better understand COVID-19

- Distributed computing project at Washington University in St. Louis
- Organizing millions of volunteers to run COVID-19 protein simulations
- Accelerated by Azure to run simulations at speed and scale
- Using AI to better understand the relationship between proteins and diseases with a goal of accelerating new therapeutics, including for COVID-19



### **Grantee: UCB**

### The COVID-19 Moonshot Project

Using AI and computational design to accelerate drug discovery

- A grassroots employee volunteer effort to accelerate research to combat COVID-19
- Identified 150 new molecules that could potentially counteract replication of the SARS-CoV-2 virus and aid in drug therapies
- Utilized the massive compute power of the cloud to do in weeks what was not possible on internal infrastructure





## COVID-19 High Performance Computing Consortium

# Advancing the pace of scientific discovery to stop the virus

- Private-public effort spearheaded by White House Office of Science and Technology Policy
- Provides researchers access to world's most powerful high-performance computing resources
- Microsoft is a founding member and provides grants through AI for Health



https://covid19-hpc-consortium.org/

## Grantee: Duke University

#### Addressing short supply of ventilators

- COVID-19 High Performance Computing Consortium + AI for Health grantee
- Two-person ventilator splitting is possible and can help with supply issues, but only for patients with similar characteristics
- Conducted millions of simulations to determine optimal volumes and pressures expected for ventilator splitting between two patients with different characteristics
- Submitted for emergency FDA approval





# Understanding and tracking progress

### AI for Health COVID-19 Data Visualizations

- Interactive visualizations available to everyone
- Full transparency into scope of problem
- Unique metric, Progress to Zero, to track progress globally



https://www.microsoft.com/en-us/ai/ai-for-health-covid-data

### **Grantee: Washington State**

#### **Department of Health**

Providing the public with timely and accurate data

 Interactive dashboard gives local population the most up-to-date information on the COVID-19 outbreak in their area

CURRENT STATUS					Data as of October 14, 2020 11:5	
CORRENT STATUS					Data as 01 October 14, 2020 11.3	SFIVI
Select an Option Confirmed Cases Hospitalizations	COVID-19 in Washing Confirmed Cases, Ho COVID-19 has spread have had a positive mo	gton State spitalizations throughout mo olecular test for	and Deaths by County st of Washington State. Confirmed cases, COVID-19.	hospitalizations and	deaths happened to people who	Le M
Deaths	Confirmed Cases	96,185	Confirmed Cases by County	<b>4</b>	Legend	
Select County	Hospitalizations	7,956		<u>,</u>	4,920 +	
All	Deaths	2,232			3,691 to	4,92
Tabular View	Percent of Deaths (deaths/confirmed cases)	2.3%			2,461 tc 1,231 tc	3,69 2,46
	Total Tests	2,169,192			1 to 1,2	30
County-level information can be found on Local Health Jurisdiction					TTP:	
(LHJ) websites	Please click "Learn Mi information.	ore" for more	252 of 06 105 confirmed second a not	have an assigned sou	ntu	
La Love Instres			552 UI 90, 165 CUITITIMED Cases do not	nave an assigned cou	nty	

#### **Governor's Office**

#### Data-driven, county-level decision-making

• Dashboard provides data used to decide whether it's safe for a county to enter a new phase of reopening

Phase and Risk Assessment	COVID-19 Disease Activity	Testing Capacity		Healtho	are System	Case Investigations and Contact Tracing	Populations at Higher Risk
🔉 PHASE AND RISH	ASSESSMENT					Data as of Oct	ober 28, 2020 11:59PM
Phase by County	COVID-19 in Washington S County Phases and Risk As This map shows the current ph Select a county to see the statu	itate sessment: ases and the key me us of key metrics for	etrics used that partice	to determine Jar county, d	e county readiness to or select a key metri	o move between phases. c for additional detail.	Lea Moi
Select a County	Washington State key met	trics				Phases by County	
		Value	Goal	Meeting Goal			
Select a key metric Rate per 100K newly	Rate per 100K of newly diagn cases over two weeks	losed 112.2	<25	No			
diagnosed cases	Average daily COVID-19 testi per 100K over a week*	ng rate 221.0	-	-		FL S S	
Daily testing rate	Percent of positive tests over week*	a 3.7%	<2%	No		╘╧╾╡╵╍┾┍╡	7-52
Percent of positive tests	Percent of licensed beds occupatients	upied by 62.8%	<80%	Yes			
Percent of licensed beds occupied	Percent of licensed beds occu COVID-19 patients	upied by 3.1%	<10%	Yes		Phase 1 Phase 2 Phase 3	Phase 4
Percent of licensed beds occupied by	*Metrics updated on 8/25/20. See Learn More for details Number of Counti					ties 0 22 17	0
Chart View Tabular View	Sources: Washington State De	partment of Health					

https://coronavirus.wa.gov/what-you-need-know/covid-19-riskassessment-dashboard

#### https://www.doh.wa.gov/Emergencies/COVID19

### Grantee: Institute for Health Metrics and Evaluation (IHME)

#### **Forecasting COVID-19 impact**

- Forecasts show demand for hospital services, daily and cumulative deaths, rates of infection and testing, and the impact of social distancing, organized by country and state
- Working with AI for Health enabled rapid response to scale and share data globally



#### www.healthdata.org

### References

#### Microsoft AI for Health Announcement

Using AI to advance the health of people and communities around the world - Microsoft On the Issues

Using AI to advance the health of people and communities around the world Jan 29, 2020 | John Kahan - Chief Data Analytics Officer

#### f in У



Photograph courtesy of Novartis Foundation

The health of people and communities around the world has been improving over time. For example, the steep decline in child and maternal mortality is a key indicator of positive

#### Microsoft AI for Health Website

https://www.microsoft.com/enus/ai/ai-for-health





#### Understanding COVID-19 We want the world to better understand COVID-19. As such, we have developed a set of interactive visualizations so everyone has

We want the work to better unarts and CUUL 19. As such we have developed a set of interactive visualizations so everyone h full transparency into the scope of the problem and the progres we are making together to heal the world. We will continue to update and refine this visualization with new data and insights.

#### Our commitment

Al for Health is a new 560 million, five-year philamtropic program from Microsoft, created to empower nonprofits, researchers, and organizations tackling some of the toughtest challenges in global health. We are providing access to Al and expertise in these areas: Learn more about the program > \_ \_ b Play. Al for Health video

Quest for discovery Global health insights Accelerating medical research to advance the prevention, diagnoses, and treatment of disposes. Health equity Research capabilities Supporting fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities, including data collaboratives and differential removing fundamental rese capabilities collaboratives and differential removing fundamental rese capabilities collaboratives and differential removing fundamental rese capabilities collaboratives and differential removing fundamental rese collaboratives and differential removing fundamental collaboratives collabora

Programs making an impact

#### Microsoft AI for Health COVID-19 Announcement

Mobilizing AI for Health to fight against COVID-19 -Microsoft On the Issues

Mobilizing AI for Health to fight against COVID-19 Apr 9, 2020 | John Kahan - Chief Data Analytics Officer

f in ⊻



On January 29, 2020, <u>we announced the launch of AI for Health</u>, an initiative to advance the health of people and communities around the world. This five-year commitment was created to empower nonprofits, researchers and organizations with AI and data science tools.

# **Privacy Challenges**

Privacy is a major issue by which we are all impacted, but current methods for preserving privacy have limitations.

### Option 1: Opt out

Limitation: Blocks critical insights needed to unlock additional knowledge relevant for cancer research, rare disease, opiod crisis, etc

### Option 2: De-identify the data

Limitation: Eliminates critical elements needed to answer research questions.

Limitation: When combined with other data, personal information can often be re-identified

We need new techniques that preserve privacy yet enable robust data analysis

# **Case Study in Re-identification**

**Governor William Weld's Medical Information** 

William Weld was the governor of Massachusetts (1991-1997)

HMO's released hospital data to researchers for the purpose of improving healthcare and controlling costs Latanya Sweeney, an MIT grad student in computer science, bought for \$10 voter registration data for Cambridge, MA

She cross linked this voter data with the HMO data and was able to identify the medical records of the governor



### **Researchers rely on techiniques like de-identification**



Technique does not always work to protect privacy and keep the meaning of the data, limiting researchers' ability to analyze and find new insights.





### **Researchers rely on techiniques like de-identification**



Technique does not always work to protect privacy and keep the meaning of the data, limiting researchers' ability to analyze and find new insights.



< **□** ▷ ▷ ♡ \*

### **Researchers rely on techiniques like de-identification**



Technique does not always work to protect privacy and keep the meaning of the data, limiting researchers' ability to analyze and find new insights.



< ⊡ ▷ ዖ ♡ ,⊀

# Differential Privacy—What is it?

Developed in 2006 by a team of researchers led by Cynthia Dwork, Distinguished Scientist at Microsoft Research

A privacy preserving technique that provides a mathematically provable guarantee of privacy protection

Works by introducing relatively small inaccuracies, or statistical noise, into the calculations/queries

These inaccuracies are large enough to protect privacy, but small enough such that the answers provided to analysts and researchers are still useful

The amount of information revealed from each query is calculated and deducted from an overall privacy budget to halt additional queries when personal privacy may be compromised.

Protects an individual's information as if the information was not used in the analysis at all

Masks the contribution of any single individual, making it impossible to infer information specific to an individual

# Differental privacy ensures data is kept private while providing statistically equivalent outcomes.



Differential privacy



# Differental privacy ensures data is kept private while providing statistically equivalent outcomes.







Microsoft CELA Data Science and Analytics

# Cardiovascular disease, sudden unexpected infant death, and breast cancer research using differential privacy

The cardiovascular disease (CVD) prevalence rate is 118.0934 per 1,000, as compared to 118.1486 per 1,000 using differential privacy

Data sourced from OHSU included 446,000 records



Microsoft

#### Areas of research

Analysis

- Cardiovascular disease (CVD)
- Sudden unexpected infant death (SUID)
- Wisconsin breast cancer research

	-	
)	CVD	Prevalence rate
	CVD	Rate for age group (0-20)
	CVD	Rate for age group (20-40)
	CVD	Rate for age group (40-55)
	CVD	Rate for age group (55-70)
	CVD	Rate for age group (70-90)
	CVD	Rate for race (AmInd)
	CVD	Rate for race (Asian/PI)
	CVD	Rate for race (Black/AfAm)
	CVD	Rate for race (White)
	Mea	n of age of CVD patient:

Differential privacy introduces relatively small inaccuracies, or statistical noise, into the calculations. These inaccuracies are large enough to protect privacy, but small enough that the answers provided to analysts and researchers are still useful.
# Cardiovascular disease, sudden unexpected infant death, and breast cancer research using differential privacy

The infant mortality rate is 5.7400 per 1,000, as compared to 5.7399 per 1,000 using differential privacy

Data sourced from CDC included 3,700,000 records



Microsoft

#### Areas of research

• Infant Mortality rate

SUID Rate for gender (F)

Analysis

SUID Rate

- Cardiovascular disease (CVD)
- Sudden unexpected infant death (SUID)
- Wisconsin breast cancer research



# Cardiovascular disease, sudden unexpected infant death, and breast cancer research using differential privacy

The breast cancer malignant rate is 37.2583 per 100, as compared to 36.7669 per 100 using differential privacy

Data sourced from University of Wisconsin, Clinical Sciences Center included 569 records



Microsoft

#### Areas of research

- Cardiovascular disease (CVD)
- Sudden unexpected infant death (SUID)
- Wisconsin breast cancer research

#### Analysis

Malignant rate
Mean of benign tumor radius

- O Mean of malignant tumor perimeter
- O Mean of malignant tumor radius
- O Number of patients

Differential privacy introduces relatively small inaccuracies, or statistical noise, into the calculations. These inaccuracies are large enough to protect privacy, but small enough that the answers provided to analysts and researchers are still useful.

### Microsoft and Harvard Institute for Quantitative Social Science

# An open collaboration to build a platform for differential privacy

- Platform and algorithms open to all developers, researchers, companies worldwide to ensure transparency and build trust <u>https://github.com/opendifferentialprivacy</u>
- A royalty-free license under Microsoft's own differential privacy patents will be granted to the world
- Builds on Microsoft's work around homomorphic encryption and confidential computing, which looks to advance security in the cloud.



### **Differential Privacy References**

- Cynthia Dwork's brief <u>A Firm Foundation for Private Data Analysis</u>, written for a broad technical audience
- <u>Differential Privacy: A Primer for a Non-technical Audience</u>, the product of a working group of the Privacy Tools for Sharing Research Data project at Harvard University
- John Kahan's LinkedIn article on the collaboration between Microsoft and Harvard's Institute for Quantitative Social Science

## THANK YOU!

United States Telecommunications Training Institute December 10, 2020



### Al in the Real World

Ted Way, PhD Senior Program Manager, Azure Cloud AI Systems and Technology

#### tedway@microsoft.com

in /in/tedway



Please Do Not Reproduce Without Permission





Exit

### AI demands more and more

### $\rightarrow$ Flexible scale

Data explosion: 2013 4.4 ZB - 2025 175 ZB Machine Learning & AI proliferating

#### $\rightarrow$ Low latency

Real-time insights Interactive user experiences Autonomous decision making

### $\rightarrow$ High throughput

Cloud-scale services Converting data to intelligence



### Overview

- М-Кора
  - Predictive maintenance of solar batteries in Africa
- RXR
  - · Covid-19 safety: social distancing measurements
- Chevron
  - Automatically extracting data from reports
- Jabil
  - · Manufacturing defect analysis

### **AI Platform**



### Al transcription demo

### **Case Studies**





Using the Cloud and IoT to deliver safe, affordable energy in Africa

M-KOPA Solar Africa



. 0 0

s Airband technology

#### Sign in $(R_+)$

# Closing the rural broadband gap

Many rural areas lack broadband access, limiting their ability to thrive in a digital economy. The Airband Initiative brings affordable broadband access to everyone.

Play video on Airband Initiative

### My children will be learning without any shortages, as compared with the paraffin (kerosene).

### Case study 1: M-Kopa

- Customer overview
  - Kick-started pay-as-you-go solar market in Africa
- Problem to solve
  - Predicting failures of solar-charged batteries





#### M-KOPA 6

Clean, bright, affordable light and phone charging

#### M-KOPA 600

24" HD TV, bright lights, and phone charging to connect your home to the world



#### M-KOPA 6000

32" HD TV and more lights for super-charged viewing

#### M-KOPA 6500

Reliable refrigeration and lighting to make your business stand out

### **Anomaly detection**



### Safe spaces

### Case study 2: RXR

- Customer overview
  - $\cdot$  3<sup>rd</sup> largest real-estate owner in New York City with 230000 m<sup>2</sup> of space
- Problem to solve
  - $\cdot$  Helping tenants navigate the "new abnormal" of COVID-19 and beyond



### Computer Vision for Spatial Analysis

Enables building applications to analyze live video and understand people's movement in physical space.















### Oil and gas

### Case study 3: Chevron data mining

### Customer problem

- Numerous custom, unstructured drilling and completions reports in PDFs for entry into spreadsheets
  - Each well generates dozens of reports
  - Each report 70-100 pages
- Manually analyzing and copying and pasting critical data from drilling and completion reports into backend systems





### **OCR (Read) Overview**



### Form Recognizer

12345 Main St. Palo Alto, CA 98473 USA

Easily extract text, keys, and tables from your documents

Form <b>4506-T</b> (June 2019) Department of the Treasury Internal Revenue Service	Request for Transcript of Tax Return	OMB No. 1545-1872
	Do not sign this form unless all applicable lines have been completed.	
	Request may be rejected if the form is incomplete or illegible.	
	► For more information about Form 4506-T, visit www.irs.gov/form4506t.	
Tin Use Form 4506 T to order a transprint or other return information free of abore. See the product list below You can guiddly request transprints by using		

Tip. Use Form 4506-T to order a transcript or other return information free of charge. See the product list below. You can quickly request transcripts by using our automated self-help service tools. Please visit us at IRS.gov and click on "Get a Tax Transcript..." under "Tools" or call 1-800-908-9946. If you need a copy of your return, use Form 4506, Request for Copy of Tax Return. There is a fee to get a copy of your return.

entification		
3 Current name, address (including apt., room, or suite no.), city, state, and ZIP code (see instructions)		
12345 Main St. Palo Alto, CA 98473 USA		
4 Previous address shown on the last return filed if different from line 3 (see instructions)		

Name: Haas Jonathan Spouse Name: Haas Mary First SSN: 234-65-2134 Second SSN: 543-05-3243 Current Address: 12345 Main St. Palo Alto, CA 98473 USA Previous Address: 12345 Main St. Palo Alto

**Key-value pairs** 

### What kept us up at night

#### What we trained with



Licensed under the Creative Commons Attribution-Share Alike 4.0 International license, attributed to MCruz

#### What customers uploaded



Licensed under the Creative Commons Attribution-Share Alike 4.0 International license, attributed to Donald Trung

### Case study 3: Solution

- Solution
  - Use Form Recognizer to extract data
  - $\cdot\,$  Use robotic process automation to enter data into backend systems
- Value
  - Streamlined operations
  - Enables analysis with greater speed, accuracy, and depth



### Documents



### Mixed Languages...

C1阿里云 000 FOOD . LA NOURRITURE . DIE NAHRUNGSMITTEL . LOS ALIMENTOS . IL CIBO 奥运会全球和宝营赔务商 nuts and dried fruit . les noix et les fruits secs . die Nüsse und das Dörrobst · los frutos secos - le noci e la frutta secca City Brain: A City Intelligence Infrastructure pine nut . le pignon pistachio \_ la pistache cashewnut . la noix heanut la hazelnut \_ la noisette die Piniennuss die Pistazie \_ el de cajou . die cacahouète \_ die die Haselnuss 城市治理的前瞻性探索实践 pistacho . il pistacchio el piñón . il pinolo Cashewnuss \_ el Erdnuss \_ el cacahuete , la avellana anacardo, l'anacardio l'arachide la nocciola 全局实时分析城市运行状态 利用数据资源调配公共资源 prazilnut . la noix du almond \_ l'amande walnut . la noix . die chestnut \_ le marron necan la noix Brésil \_ die Paranuss pacane \_ die Pecannuss . die Mandel . la Walnuss - la nuez , die Esskastanie \_ la · la nuez de Brasil la pacana . la almendra \_ la mandorla la noce castaña - la castagna 修正城市运行过程中的缺陷 · la mandorla brasiliana noce pecan shell la coquille die Schale la cáscara il auscio date - la datte fig. la fique die nacadamia rune . le pruneau le macadamia Feige \_ el higo \_ il fico - die Dattel \_ el dátil die Backpflaume die Macadamianuss il dattero · la ciruela pasa 西梨橙西梨西梨橙 **Drive on the left** la macadamia 霸霸霸霸霸霸 la prugna secca FRUIT SUMMER 瓜 汗 汗 瓜 橙 瓜 柠 柠 la noce di macadamia TEA FRUIT 水芒柠橙红 梨 橙 檬 檬 la chair 霸煮水果茶 鲜榨果 Conduise das Fruchtfleisch 果果檬子柚 (1000cc) (400cc/ la pulpa 500cc) a polpa #12 12 15 12 13 13 13 16 a gauche 15 15 12 14 14 #15 15 18 15 16 16 16 20 raisin . le raisin sec currant . le raisin de coconut - la noix de sultana le raisin de Links fahren Smyrne \_ die Sultanine die Rosine \_ la pasa Corinthe \_ die Korinthe coco . die Kokosnuss CIASSIC 丝 红 血 珍 港 香 蜜 冻柠冻冻龙 la pasa de Corinto SUMMER l'uvetta · el coco . la noce 林林豆糯珠式港鸯 柠檬柠柠凤 TEA PORT 奶奶奶米奶咖鸳奶 茶水蜜干柠 夏日凉飲 經典灌式 茶盖茶奶茶啡鸯盖 (500cc/ (400cc) Tenere la sinistra (000cc) www.LEGOsurvey.com/pr #8 8 10 10 10 10 12 10 10 10 12 12 14 #12 10 12 12 14 Conduzca por la GAGNE ! 原咸巧葡黑桂红 助口 助口 助口 WIN! **GEWINNE!** CIASSIC CIASSIC a izquierda 哩 哩 哩 味蛋克萄芝花豆 FOOD FOOD a naar www.LEGOsurvev.com/product. 鸡黄力干麻鸡鸡 鱼鱼双 ondre à un questionnaire et 港式小吃 Il een enqueteformulier in en maak kans 潜式小吃 hance de gagner un produit LEGO 丸 蛋 拼 蛋蛋蛋蛋蛋蛋蛋 p een cool LEGO® product. Links rijden 仔仔仔仔仔仔仔 the LEGO logo are trademarks of the 10 15 12 12 12 12 12

### Manufacturing

### Case study 4: Jabil



- Customer overview
  - Vision: Be the most technologically advanced and trusted manufacturing solutions provider
- Customer problem
  - · Manufacturing defect analysis



Ξ

2

1|8¥Γ

#### Microsoft and Jabil Innovate with **Predictive** Analytics on the Plant Floor

FVI **Functional Verification Test** 

ICT - In Circuit Test

AXI - Xray Inspection

AOI - Automatic Optical Inspection

Reflow - Oven

and?

SMT - Pick and Place

SPI Paste





PASS



FAIL


## Low-code/no-code AI

# Low-code/no-code Al

- Azure Machine Learning Studio
- Automated machine learning



#### Flight Delays දරූ

96 assets in total	Ö	00	Autosave or	n 🖪	୬ ୯	$\Box$	ĥ	Ŵ	<b>1</b>	100% ~	Ð	Q	1:1	2ª	2
Applies a mathematical operation to colu- values.	mn 9/7/2020		ſ	🖯 Flig	ht Delays	Data									
Apply SQL Transformation Microsoft Runs a SQLite query on input datasets to transform the data.	9/7/2020														
Clean Missing Data A Microsoft Specifies how handle the values missin dataset.	g from a 9/7/2020			€ No	rmalize D	ata									
C Clip Values A Microsoft Detects outliers and clips or replaces their	r values. 9/7/2020				O			)		J					
Convert to CSV A Microsoft Converts data input to a comma-separate format.	ed values 9/7/2020														

Convert to Dataset

# Conclusions

# **Conclusions and Takeaways**

### Trustworthy Al

- · All stakeholders have roles and responsibilities
- $\cdot$  Enabling policy, and technology as a solution
- AI and COVID-19
- Privacy-preserving technologies
- AI in the real world
  - $\cdot\,$  AI technology is real and used all over the world
  - Al may move faster than policy can keep up



# Thank you!

# Resources

- Microsoft Airband
  - https://www.microsoft.com/en-us/corporate-responsibility/airband
- Case studies
  - · Social distancing
    - https://customers.microsoft.com/en-us/story/843823-rxr-realty-reopens-for-business-using-azure-iot
  - Form recognizer
    - https://customers.microsoft.com/en-us/story/chevron-mining-oil-gas-azure-cognitive-services
  - · Manufacturing
    - https://customers.microsoft.com/en-us/story/811123-jabil-discrete-manufacturing-teams
- Azure Cognitive Services
  - https://azure.microsoft.com/en-us/services/cognitive-services/
- Azure Machine Learning
  - https://azure.microsoft.com/en-us/services/machine-learning/