BRINGING PRECISION MEDICINE TO THE BRAIN:
UNLEASHING BIG DATA TO ANSWER ALS

At the Leading Edge
Big Data for Research

Elizabeth J. Bruce
Microsoft
Washington, DC – AnswerALS
October 24, 2017
All In on ALS

After the Ice Bucket

By Bill Gates | September 15, 2014

Like a lot of people, I dumped a bucket of ice water on my head a few weeks ago to call attention to the disease ALS. I soon got the chance to learn more about it—in a very inspiring way.

Shortly after I made the ice-bucket video, I heard that Steve Gleason, a former pro football player who was diagnosed with ALS in 2011, was visiting the Seattle area. He stopped by my office recently to talk about the disease, and to show me how he uses technology to break through some of the barriers it creates. It was really impressive to see how he overcomes the challenges of having ALS.

You may know about Steve from the Super Bowl ad he made with Microsoft this year, or from the
The world is full of complex problems...

What is the big deal about big data?
Financial Loans and Loan Defaults

Challenge: Predicting consumer credit risk?

Machine Learning algorithm detects potential loan defaults more accurately than FICO scores

Andrew W. Lo et al., MIT Sloan Sch
2.5B people do not have credit scores

Big data can enable new insights that can change people’s lives
Data Science

SCALE $\rightarrow$ Volume, Velocity, Variety
Data should be **Findable**

Data should be **Accessible**

Data should be **Interoperable**

Data should be **Re-usable**.

“The science system is in landslide transition from data-sparse to data-saturated. Meanwhile, scholarly communication, data management methodologies, reward systems and training curricula do not adapt quickly enough if at all to this revolution.”

Barend Mons, Chair Of The High-Level Expert Group, European Open Science Cloud
AnswerALS compute and storage needs

- Manage user accounts centrally
- Track and version computation
- Parallel computing
- Link data and metadata
- Rapid transfer of large files
- Encryption at rest
- Store and compute on HIPAA protected data
AnswerALS
- Latanya Sweeny, 2014 – Data map of data and data flows in personal health care
AI and Cognitive Services APIs

Language:
- Entity linking
- Bing news search
- Customer feedback analysis
- Recommendation API
- Text analytics
- Thumbnail generation
- Web language model
- Anomaly detection
- Sentiment scoring
- Bing autosuggest
- Computer vision

Speech:
- Bing image search
- Speech API
- Text to speech
- Spell check

Machine Learning:
- Bing image search
- Speech API
- Text to speech
- Academic knowledge
- Spell check
- Emotion

Knowledge:
- Bing news search
- Custom recognition (CRIS)
- Bing image search
- Text to speech
- Academic knowledge

Search:
- Bing web search
- OCR, tagging, captioning
- Bing news search
- Bing autosuggest

Vision:
- Bing image search
- Speech API
- Text to speech
- Academic knowledge
- Spell check
- Emotion

AI and Cognitive Services APIs
A legal standard of privacy

A mathematical equation of privacy

\[ M: X^n \to T \text{ satisfies } \epsilon\text{-differential privacy if} \]
\[ \forall x, x' \in X^n \text{ s.t. } \text{dist}_H(x, x') = 1 \forall S \subseteq T, \]
\[ \Pr[M(x) \in S] \leq e^\epsilon \Pr[M(x') \in S]. \]
to realize promise of big data:

1) enabling big data for research
2) public-private partnerships to break down data silos
Thank you!

elizabeth.bruce@microsoft.com