TRAK 1: Digital transformation enabling personalization of medicine



What is the real ethical question in Healtcare?

Bruno Schroder
National Technology Officer



Al is already everywhere in Healthcare

Germany:

- Project "Heart-Hero": Azure AI services and therapy configurator help treat heart failure. The customer story is available here in German
- Medical Research: SANA Clinics and HMS Analytical Software bet on Artificial Intelligence for early diagnosis of sepsis. <u>video.</u>

Mexico

Azure Genomics Services www.codigo46.com.mx

Malaysia

 Ministry of Health research on using retina images for early detection of diseases

Finland

HUS Health Village: https://www.terveyskyla.fi/

Czech

 Al-assisted recognition of prostate cancer in radiology https://news.microsoft.com/europe/features/doctors-and-ai-a-powerful-elixir-improving-patient-outcomes-in-europe/

Canada

- Automated diagnosis using images/photos
- Automated transcription of Physician conversations to help patients remember what was said

Poland

 Carna Life Analytical telemedicine identifying over 20 irregularities in EKG http://en.medapp.pl/

France

- Artificial Insight pour aider les médecins dans leur diagnostic https://customers.microsoft.com/en-us/story/724118-artificial-insight-healthcare-azure-fr-france
- Helpicto: troubles du langage et reconnaissance d'image

Sweden

 Optolexia : early detection of yslexia www.youtube.com/watch?v=rpQKf6vxm8U

South Africa

 Phulukisa's Specialist in a backpack bring medicine in the bush https://www.youtube.com/watch?v=Nyw8KZKURVY

Microsoft Research

- <u>Project InnerEye</u> ML for automatic, quantitative analysis of three-dimensional radiological images, turning radiological images into measuring devices.
- Project Premonition: trying to predict the spread of diseases by using mosquitoes to collect blood samples from animals in the wild and identify the diseases they're carrying

https://www.youtube.com/watch?v=tUnH5K8Dyts&feature=youtu.be

Project

Premonition

Problem

When trying to predict the spread of a disease, every second counts. About 60%–75% of all emerging infectious diseases originate in animals, but it's difficult to pinpoint how, when, and where.

Solution

Microsoft Researchers are using mosquitos to collect blood samples from animals in the wild and identify the diseases they're carrying.

Project Premonition uses drones to find mosquito breeding grounds, robotic traps to gather specimens, and cloud-scale genomics powered by machine learning to search the specimens' DNA for pathogens.

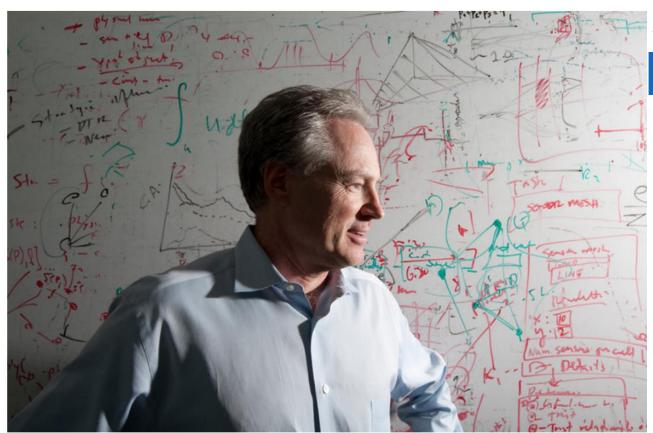
"Using the Microsoft Cloud, we can analyze more than 100 million pieces of DNA in every sample," says Microsoft Researcher Ethan Jackson.



The Microsoft Cloud is fighting disease by turning mosquitos into data-gathering devices, and analyzing pathogen data ... so we may one day stop outbreaks before they begin.

Microsoft Finds Cancer Clues in Search Queries

By JOHN MARKOFF JUNE 7, 2016



Dr. Eric Horvitz is one of the Microsoft researchers who conducted the study.

http://www.nytimes.com/2016/06/08/technology/online-searches-can-identifycancer-victims-study-finds.html?smid=tw-share& r=1

Journal of oncology practice

The Authoritative Resource for Practicing Oncology

This Journal V Enter search words / phrases / DOI / authors / keywords / etc.

Newest Content

Issues

Browse By Topic

Special Content

Authors

Screening for Pancreatic Adenocarcinoma Using Signals From Web Search Logs: Feasibility Study and Results

John Paparrizos, MSc, Ryen W. White, PhD 🔄, Eric Horvitz, MDPhD Columbia University, New York, NY; and Microsoft Research, Redmond, WA

Abstract Full Text Supplemental Material

Choose

Introduction:

Abstract

People's online activities can yield clues about their emerging health conditions. We performed an intensive study to explore the feasibility of using anonymized Web query logs to screen for the emergence of pancreatic adenocarcinoma. The methods used statistical analyses of large-scale anonymized search logs considering the symptom queries from millions of people, with the potential application of warning individual searchers about the value of seeking attention from health care professionals.

http://ascopubs.org/doi/full/10.1200/JOP.2015.010504

Define Al

FOURTH PARADIGM

DATA-INTENSIVE SCIENTIFIC DISCOVERY

EDITED BY TONY HEY, STEWART TANSLEY, AND KRISTIN TOLLE

No theory, no model, no experience
Data driven
The quest for signal
Probabilistic interpretation of the world

Medicine is already data driven

Would AI raise any (new) issue in collecting patient data?

- Randomized studies
- Ethical committees
- Patient informed consent

Al raises new issues in handling and sharing patient data!

The data issue

- Efficient research in AI requires data aggregation
- Minimization of data does not work in healthcare
- Re-identification will always be possible
- Unknown level of information contained could/will lead to discoveries not covered by consent
- It opens up use cases impossible to forecast

The traditional way of dealing with patient consent must evolve.

The DATA ethical questions

- Should we use data to advance care?
 - Can work with unlinked pseudonimized data
- Should we inform patients about new findings?
 - Requires link to patient
- Should we decide not to cure?
 - This is what we do if we do not share, aggregate, research
- Would basic medical ethic allow a decision not to care?

The USE CASE issue

- Al is a probabilistic interpretation of the world. Is this the right model?
 - Probability versus possibility
 - Discrimination risk is real
- Using AI for selection is a very risky business
 - Humans are better at understanding possibilities than probabilities
- Data and model bias are real risks
- Understanding of the outcome is frequently an issue
 - The model always gives an answer, and a probability
 - Although probably less a problem in health care
- Ethics is relative
 - Pre-existing conditions

The USE CASE ethical question: control

Is it really a solution to make it physically impossible by restricting data access?

What are then the socially (un)acceptable use cases?

Which Use Case control framework? permissive or restrictive? legal or deontological?

The way forward

Finding the signal is fundamental research requires as many data as possible uncertain outcome

Tuning the model is applied research requires targeted data generates economic value

The way forward

- Get as much data as possible
 - Informed consent
 - Aggregation
 - Accessible for research
 - Real world data
- De-biased it
 - What is the Shannon-Nyquist constraint of your data set?
- Incentivize data contributors
- Data sharing for the public good
 - Posthumous data donation
 - Data philanthropy

Or wait for other countries to come with an algorithm and use it under licence

Seriously, don't!

Risk mitigation: Technology

• ENABLE:

- The technology exists already in the cloud
- Use standards like ISO 19944 and certified infrastructure
- TEE:
 - Microsoft joins partners and the Linux Foundation to create Confidential Computing Consortium https://cloudblogs.microsoft.com/opensource/2019/08/21/microsoft-partners-linux-foundation-announce-confidential-computing-consortium/

• CREATE:

- Invest heavily in algorithmic research
- ADOPT:
 - Educate massively to the understanding of probabilistic models

The model does not remember the data. The physician does!

ISO/IEC FDIS 19944: data categories

Identified data

• Data that can unambiguously be associated with a specific person because PII is observable in the information. Guidance on what can be considered as identifiers can be found in 4.4.1 of ISO/IEC 29100:2011[04].

Pseudonymized data

Data for which all identifiers are substituted by aliases for which the alias assignment is such that it cannot be reversed by reasonable efforts of anyone other than the party that performed them.
 This corresponds to data defined as "pseudonymization" in 2.24 and described as "pseudonymous data" in 4.4.4; both in ISO/IEC 29100:2011.

Unlinked pseudonymized data

• Data for which all identifiers are erased or substituted by aliases for which the assignment function is erased or irreversible, such that the linkage cannot be re-established by reasonable efforts of anyone including the party that performed them.

Anonymized data

• Data that is unlinked and which attributes are altered (e.g., attributes' values are randomized or generalized) in such a way that there is a reasonable level of confidence that a person cannot be identified, directly or indirectly, by the data alone or in combination with other data.

This corresponds to data defined as "anonymized data" in 2.3 and process defined as "anonymization" in 2.2; both in ISO/IEC 29100:2011.

Aggregated data

• Statistical data that does not contain individual-level entries and is combined from information about enough different persons that individual-level attributes are not identifiable.



NEWS: Introducing the **Confidential Computing Consortium**, a new community effort to advance computational trust, security for cloud **computing**:



Risk mitigation: legal

- We need a strict legal framework
 - banning re-identification
 - defining forbidden use-cases
 - Framing data sharing and data aggregation
 - INHA project
 - Art 89 GDPR
- Heavy fines in case of breach

Conclusion

 Which ethical rule could make us refuse a data driven new therapeutic domain?

- There are no new ethical issues with AI in healthcare
 - Exclusion already happens through questionnaires and nullity of contract of the answers are not correct.
 - Issues are legal and practical

Introduction

Creating a trusted, responsible, and inclusive cloud

Ingo N.

Satya Nadella Chief Executive Officer Microsoft Corporation

Dent Sint

Brad Smith
President and Chief Legal Officer
Microsoft Corporation

Cloud Global

A policy roadmap for a trusted, responsible, and inclusive cloud

https://news.microsoft.com/cloudforgood/resources.html



Future is the only way forward! Microsoft