EPSTEIN
BECKER
GREEN

# Data Governance and Artificial Intelligence in the Health Care

*Alaap Shah, JD MPH*

- October 16, 2019

# Presented by

**Alaap B. Shah**

Member of the Firm

Epstein Becker Green P.C.

ashah@ebglaw.com

202.861.5320

# My Background

- Member of the Firm, Epstein Becker & Green P.C.
  - Partner in Heath Care and Life Sciences Division
  - Co-Lead of Data Privacy, Cybersecurity and Data Asset Management Team
- American Society of Clinical Oncology/CancerLinQ
  - Senior Counsel, Chief Privacy and Security Officer
  - Helped launch CancerLinQ – Big Data in Oncology
  - Helped manage enterprise-wide risk associated with privacy and security
- Certified by IAPP as a Privacy Professional
- Certified by HIMSS as a Health Information Systems Professional
- Certified by HITRUST on the Common Security Framework

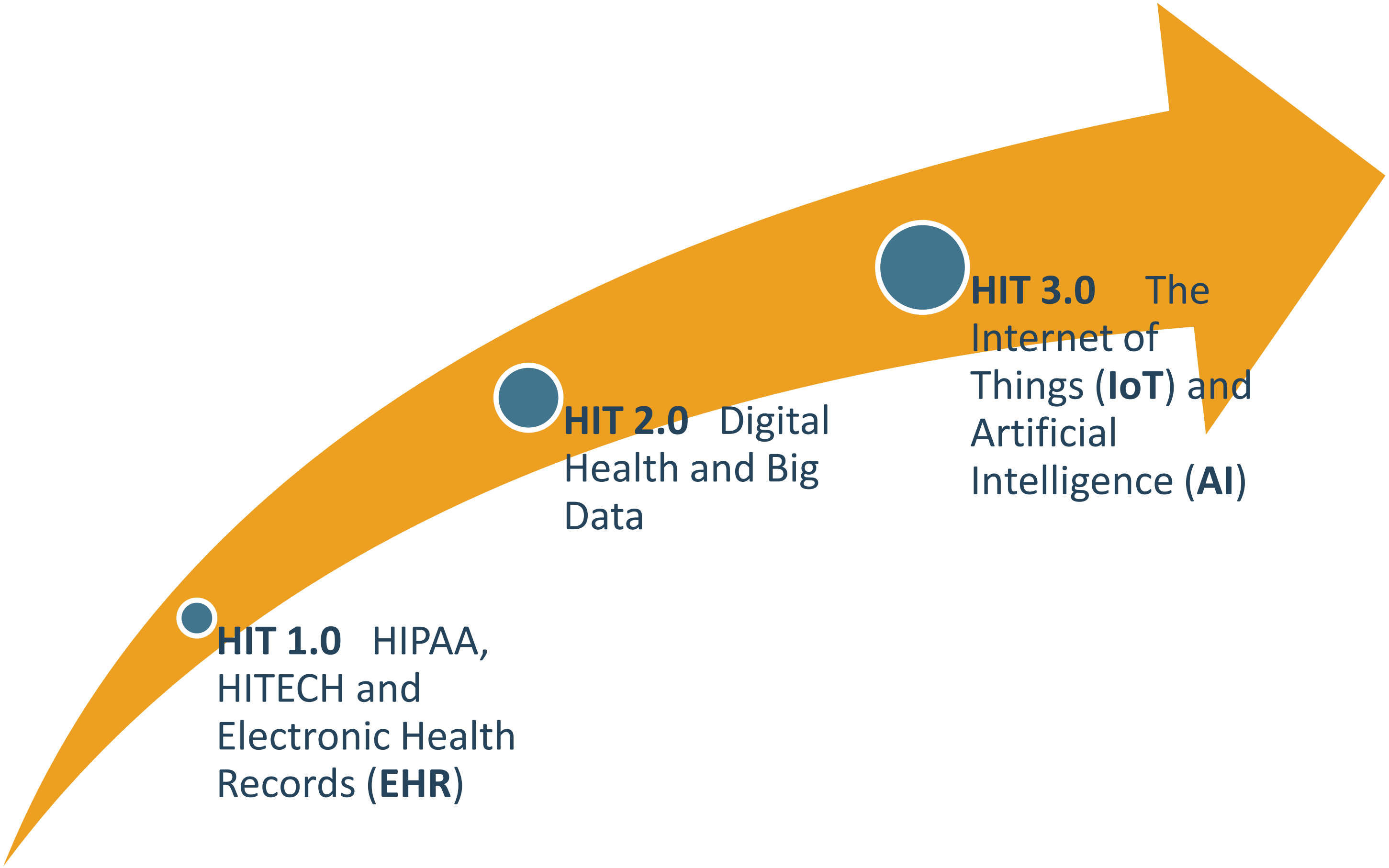
# Today's Agenda

- Data-Driven Healthcare
- Building Trust Networks
- Building Trust with a Data Governance Program
- Q&A

EPSTEIN
BECKER
GREEN

# Data-Driven Healthcare

# Health Information Technology Trajectory



**HIT 1.0** HIPAA, HITECH and Electronic Health Records (**EHR**)

**HIT 2.0** Digital Health and Big Data

**HIT 3.0** The Internet of Things (**IoT**) and Artificial Intelligence (**AI**)

# Healthcare is an Data Driven Enterprise

- Data drives learning, quality improvement and efficiency
- Interconnectivity and data sharing is increasing rapidly
- Establishing trust is critical for robust information exchange
- HIPAA is a starting point to manage data risk, but *only* a starting point
- Data may be shared in ways that HIPAA will no longer apply
- Establishing adequate levels of assurance often demands going beyond minimum HIPAA requirements

Image Credit: Shutterstock

# OCR’s Role in Individual-Directed Data

## Recent HIPAA FAQ guidance:

- Individuals have right to access PHI (including transmission to a third party app)
  - Cannot deny request based on concerns about app privacy or security
- Apps developed for or on behalf of a Covered Entity by a Business Associate will likely be covered by HIPAA
  - BAA required
  - Subsequent use and disclosure of PHI will be subject to HIPAA
- Covered Entities that transmit PHI to a non-HIPAA covered apps will not be liable for subsequent unlawful uses or disclosure of that data
  - Terms of use and privacy policy of third party app will govern
  - Transmission of PHI may be unsecured if requested by an individual as long as risks are explained to the individual

# ONC's Role in Individual-Directed Data

## *Proposed Rules on Interoperability*

- Proposed Rules issued on March 4, 2019 (pursuant to 21st Century Cures Act)
- Proposed Rules geared toward promoting patient access and consumer-directed sharing of data to spur digital health innovation
- ONC Proposed Rule Comment Period closed on June 3, 2019 (2013 comments received)

Image Credit: Shutterstock

# FTC’s Role in Individual-Directed Data

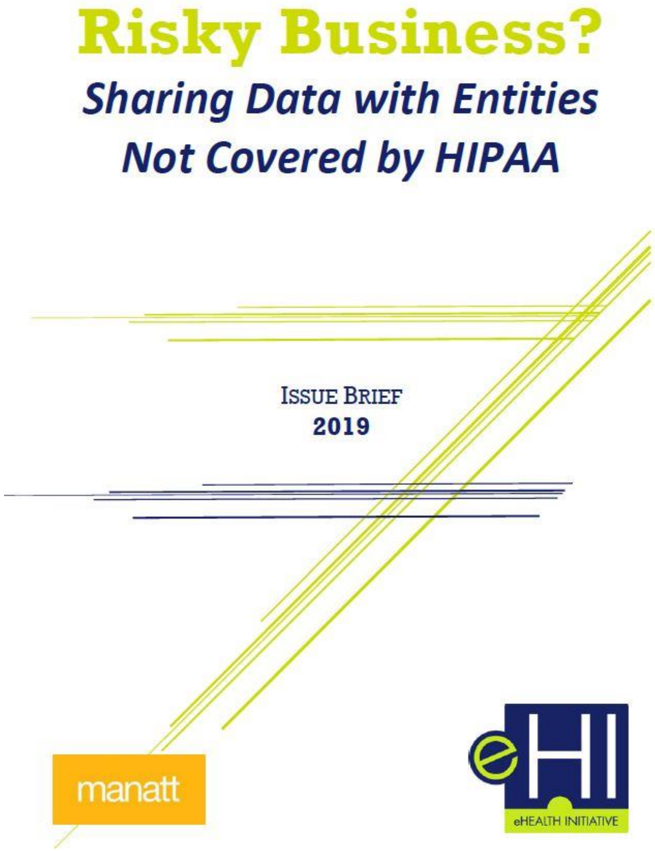
## Section 5 of the FTC Act

- Prohibits unfair methods of competition

- Section 5(n) provides the standard for “unfairness”
  - If an act “causes or is likely to cause substantial injury to consumers”; the injury to be caused “is not reasonably avoidable by consumers themselves”; and the injury is “not outweighed by countervailing benefits to consumers or competition.”

- FTC actions have been based on:
  - Failure to safeguard information;
  - Failure to adequately disclose how information will be used or disclosed;
  - Misrepresenting how information collected would be used

# Healthcare Internet of Things (IoT)



Risky Business?

*Sharing Data with Entities Not Covered by HIPAA*

ISSUE BRIEF
2019

manatt

eHI
eHEALTH INITIATIVE

Credit: Peerbits - https://www.peerbits.com/blog/internet-of-things-healthcare-applications-benefits-and-challenges.html

# Complex Health Networks

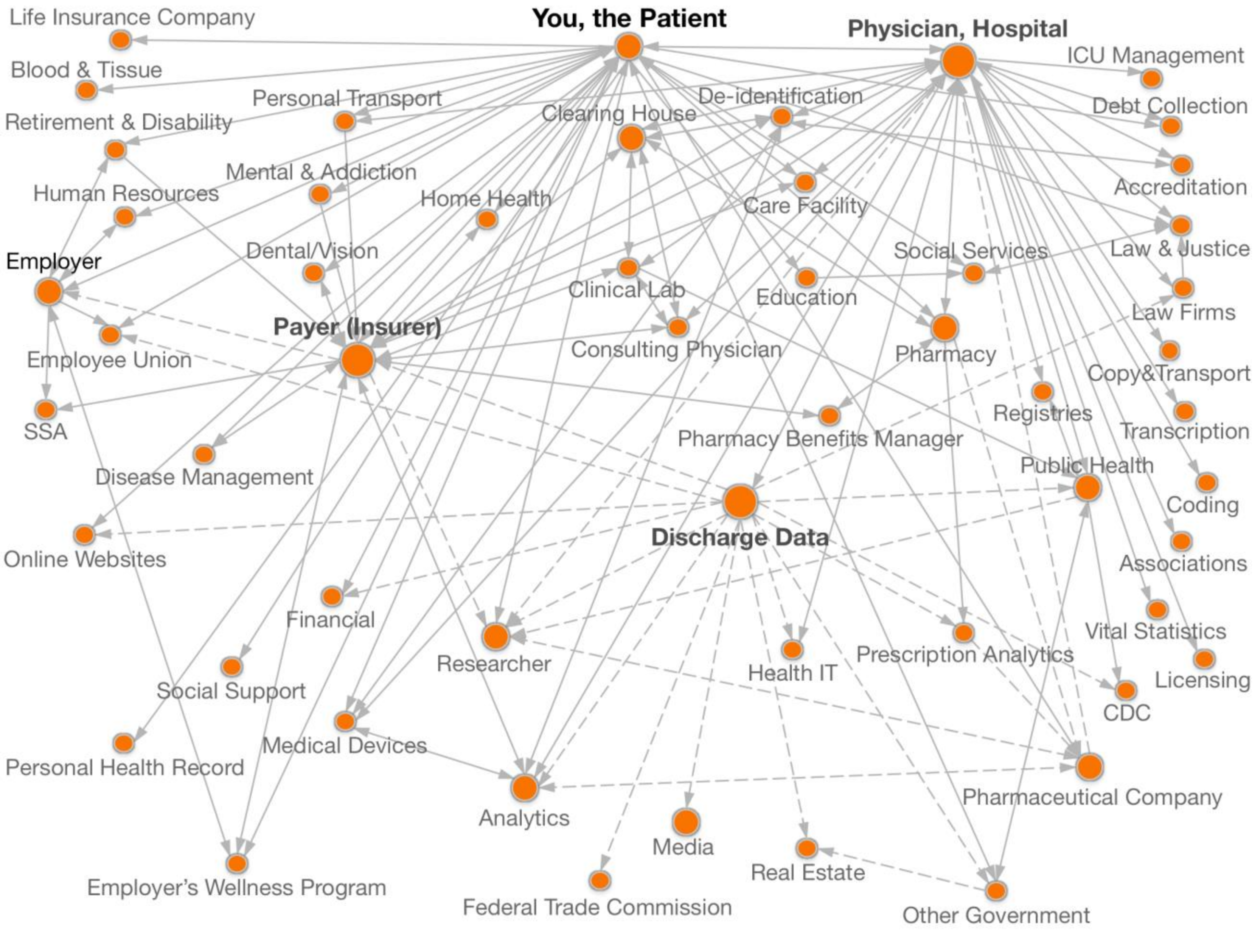


Image Credit: thedatamap.org

# Another Kind of Complex Network . . . A Neural Network

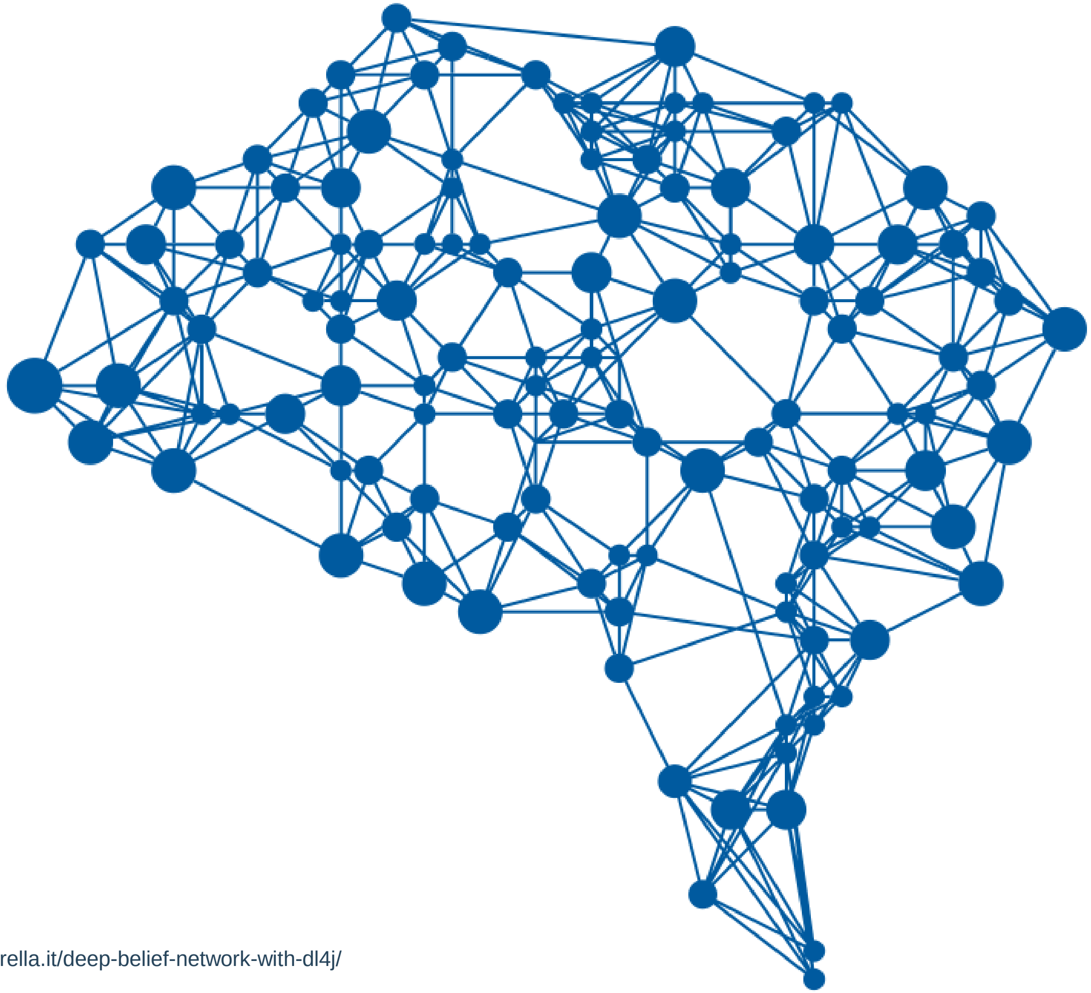


Image Credit: https://www.luigicardarella.it/deep-belief-network-with-dl4j/

EPSTEIN
BECKER
GREEN

# Artificial Intelligence

# What is Artificial Intelligence?

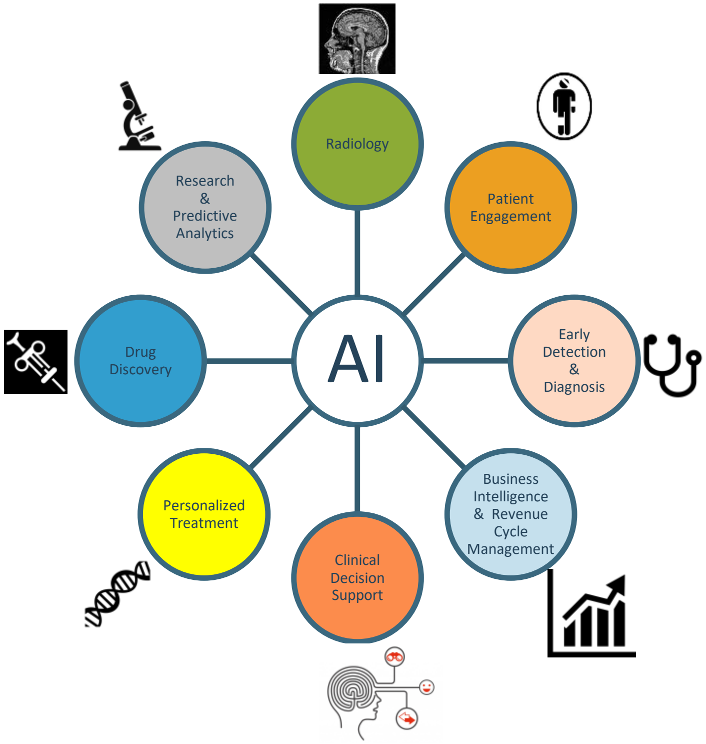
- Emerging field with various terminology:
  - Machine learning
  - Smart algorithms
  - Artificial neural networks
  - Deep learning
  - Data analytics
  - Big data
  - Data mining
  - Continuously learning system

*using computers to analyze data and make decisions by mimicking human "intelligence" but at a greater speed and scale than humanly possible*

# AI Market:  Rapid Innovation

- Rapid digitization coupled with technological advances accelerates development and implementation of AI

- AI value propositions:
  - ○ Generating efficiencies
  - ○ Reducing costs
  - ○ Improving quality and safety
  - ○ Bridging gaps in the continuity care
  - ○ Improving patient engagement



Radiology
Research & Predictive Analytics
Patient Engagement
Drug Discovery
AI
Early Detection & Diagnosis
Personalized Treatment
Clinical Decision Support
Business Intelligence & Revenue Cycle Management

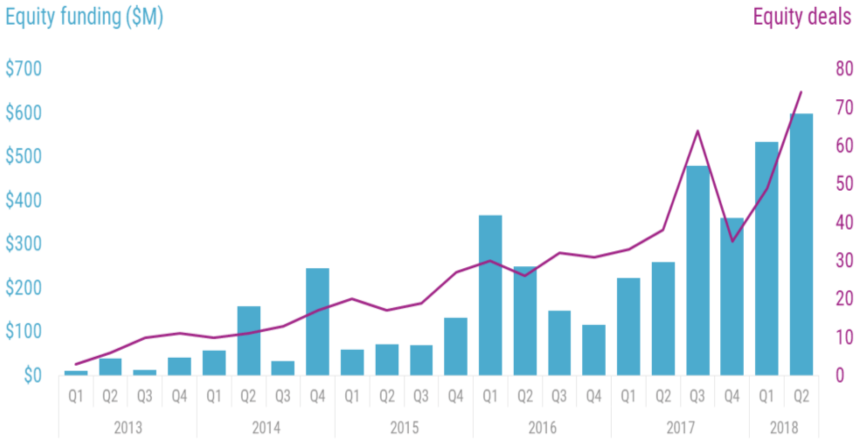
# AI Market:  Growing Investment

- Potential for improved patient care, scaling, and savings has spurred significant and increasing financial investment

- Health AI startups have over $4B since 2013, which tops all other industries in AI deal activity

AI in healthcare funding hit a historic high in Q2’18

Disclosed equity funding, Q1’13 – Q2’18



Source: cbinsights.com

CBINSIGHTS

# AI Market:  Data is King

*The world's most valuable resource is no longer oil, but data.*

- *"[Alphabet, Amazon, Apple, Facebook and Microsoft . . . are the five most valuable listed firms in the world."*
- *"With data there are extra network effects.  By collecting more data, a firm has more scope to improve its products, which attract more users, generating even more data, and so on."*

**The 'Data Economy' is at a fever pitch.  Enormous value may be realized as long as data continues to flow and trust is maintained.**

Credit: The Economist, May 6, 2017

# Key Legal and Ethical Issues with AI

## Privacy and Data Security Risks

- **Data Rights:** Ensure adequate authority exists to use data to train AI
- **AI Security:** Ensure secure collection, storage, processing, and manipulation
- **AI Lifecycle:** Ensure secure transfer and disposal of data
- **Cybersecurity Risk:** Unauthorized access and tampering with data integrity or AI functionality could negatively impact AI outputs
- **Garbage in-Garbage Out:** AI training hinges on quality inputs to produce reliable outputs
- **Bias:** Bias in AI training can lead to unreliable and potentially dangerous outputs

# AI Diligence Checklist

1. Who is responsible for overseeing the AI deployment?

2. What is the scope of intended use and required data for training?

3. How to investigate and diligence?
   – Vendor Privacy Compliance
   – Technology Security

4. What legal and contractual challenges exist?
   – Upstream data rights to use data for processing
   – Downstream data sharing rights
   – Allocation of Risk

5. What is the timeline and implementation plan?

6. How to conduct pre-deployment testing and ensure validation prior to go-live?

EPSTEIN
BECKER
GREEN

# Building Trust Networks

# Building a Trust Network

TRUST take time and effort to EARN and PRESERVE

TRUST takes seconds to LOSE

Image Credit: Shutterstock

# Building Trust with a Data Stewardship Mindset

- **Good Data Stewardship is about *TRUST*:**
  - Respect for persons is a fundamental principle undergirding health care
  - Doctor-patient relationship is based on maintaining confidence
  - Loss of trust compromises information exchange and patients suffer

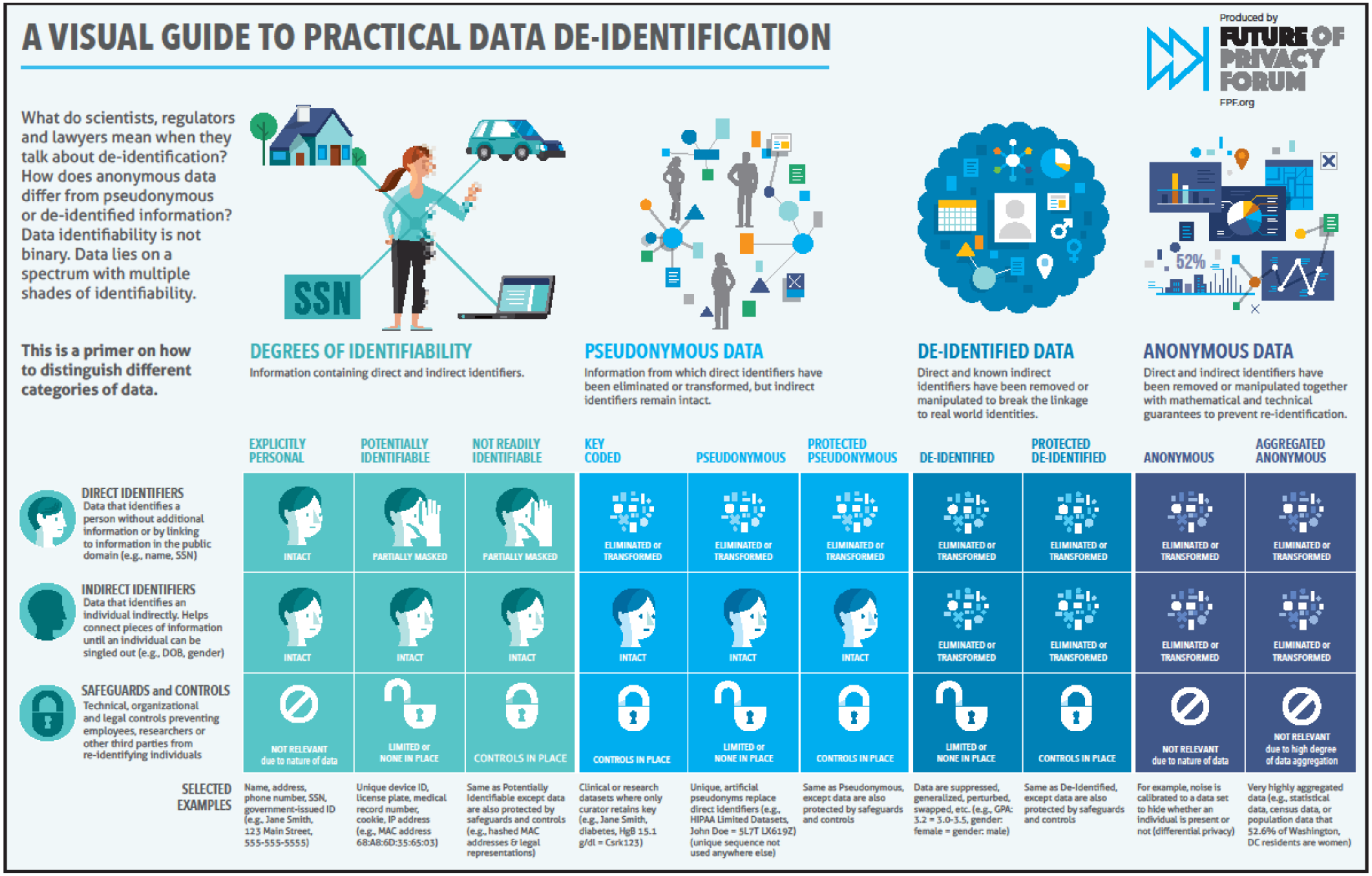
- **Reasons to Practice Good Data Stewardship:**
  - Organizations require data to drive learning, quality improvement and efficiency
  - Organizations must build trust with many partners with whom data is shared
  - Organizations must maintain the trust of patients for long-term success

- **Who Should Practice Good Data Stewardship?**
  - Everyone who collects, views, stores, exchanges, aggregates, analyzes, and/or uses patient data should practice data stewardship

# Principles of Good Data Stewardship

- **Transparency:** Provide notice regarding collecting, using, disclosing, and retaining data
- **Individual Participation:** Engage individuals, and to the extent practicable provide individual with a meaningful choice as to participation
- **Purpose Specification:** Articulate the purpose(s) for using the data
- **Data Minimization:** Only collect data that is directly relevant and necessary to accomplish the specified purpose(s) and only retain data for as long as is necessary to fulfill the specified purpose(s)
- **Use Limitation:** Use and disclose data solely for the specified purpose(s)
- **Data Quality and Integrity:** To the extent practicable, ensure that data is accurate, relevant, timely, and complete
- **Security:** Protect data through appropriate security safeguards
- **Accountability and Auditing:** Organizations should be accountable for complying with these principles, providing training to all who use data, and auditing the actual use of data

# Building Trust through De-identification and Anonymization



## A VISUAL GUIDE TO PRACTICAL DATA DE-IDENTIFICATION

Produced by FUTURE OF PRIVACY FORUM
FPF.org

What do scientists, regulators and lawyers mean when they talk about de-identification? How does anonymous data differ from pseudonymous or de-identified information? Data identifiability is not binary. Data lies on a spectrum with multiple shades of identifiability.

**This is a primer on how to distinguish different categories of data.**

### DEGREES OF IDENTIFIABILITY

Information containing direct and indirect identifiers.

### PSEUDONYMOUS DATA

Information from which direct identifiers have been eliminated or transformed, but indirect identifiers remain intact.

### DE-IDENTIFIED DATA

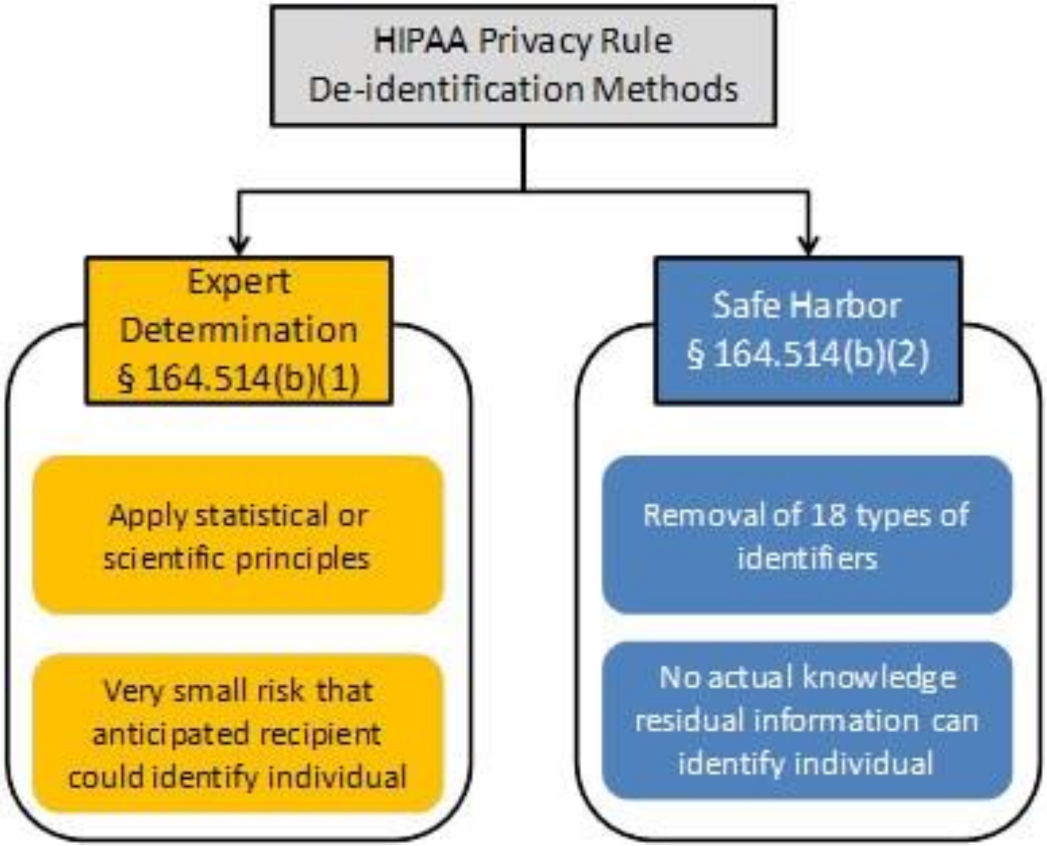
Direct and known indirect identifiers have been removed or manipulated to break the linkage to real world identities.

### ANONYMOUS DATA

Direct and indirect identifiers have been removed or manipulated together with mathematical and technical guarantees to prevent re-identification.

| | EXPLICITLY PERSONAL | POTENTIALLY IDENTIFIABLE | NOT READILY IDENTIFIABLE | KEY CODED | PSEUDONYMOUS | PROTECTED PSEUDONYMOUS | DE-IDENTIFIED | PROTECTED DE-IDENTIFIED | ANONYMOUS | AGGREGATED ANONYMOUS |
|---|---|---|---|---|---|---|---|---|---|---|
| **DIRECT IDENTIFIERS**<br>Data that identifies a person without additional information or by linking to information in the public domain (e.g., name, SSN) | INTACT | PARTIALLY MASKED | PARTIALLY MASKED | ELIMINATED or TRANSFORMED | ELIMINATED or TRANSFORMED | ELIMINATED or TRANSFORMED | ELIMINATED or TRANSFORMED | ELIMINATED or TRANSFORMED | ELIMINATED or TRANSFORMED | ELIMINATED or TRANSFORMED |
| **INDIRECT IDENTIFIERS**<br>Data that identifies an individual indirectly. Helps connect pieces of information until an individual can be singled out (e.g., DOB, gender) | INTACT | INTACT | INTACT | INTACT | INTACT | INTACT | ELIMINATED or TRANSFORMED | ELIMINATED or TRANSFORMED | ELIMINATED or TRANSFORMED | ELIMINATED or TRANSFORMED |
| **SAFEGUARDS and CONTROLS**<br>Technical, organizational and legal controls preventing employees, researchers or other third parties from re-identifying individuals | NOT RELEVANT due to nature of data | LIMITED or NONE IN PLACE | CONTROLS IN PLACE | CONTROLS IN PLACE | LIMITED or NONE IN PLACE | CONTROLS IN PLACE | LIMITED or NONE IN PLACE | CONTROLS IN PLACE | NOT RELEVANT due to nature of data | NOT RELEVANT due to high degree of data aggregation |
| **SELECTED EXAMPLES** | Name, address, phone number, SSN, government-issued ID (e.g., Jane Smith, 123 Main Street, 555-555-5555) | Unique device ID, license plate, medical record number, cookie, IP address (e.g., MAC address 68:A8:6D:35:65:03) | Same as Potentially Identifiable except data are also protected by safeguards and controls (e.g., hashed MAC addresses & legal representations) | Clinical or research datasets where only curator retains key (e.g., Jane Smith, diabetes, HgB 15.1 g/dl = Csrk123) | Unique, artificial pseudonyms replace direct identifiers (e.g., HIPAA Limited Datasets, John Doe = 5L7T LX619Z) (unique sequence not used anywhere else) | Same as Pseudonymous, except data are also protected by safeguards and controls | Data are suppressed, generalized, perturbed, swapped, etc. (e.g., GPA: 3.2 = 3.0-3.5, gender: female = gender: male) | Same as De-Identified, except data are also protected by safeguards and controls | For example, noise is calibrated to a data set to hide whether an individual is present or not (differential privacy) | Very highly aggregated data (e.g., statistical data, census data, or population data that 52.6% of Washington, DC residents are women) |

# HIPAA De-identification Methods



HIPAA Privacy Rule
De-identification Methods

Expert Detemination
§ 164.514(b)(1)

Apply statistical or scientific principles

Very small risk that anticipated recipient could identify individual

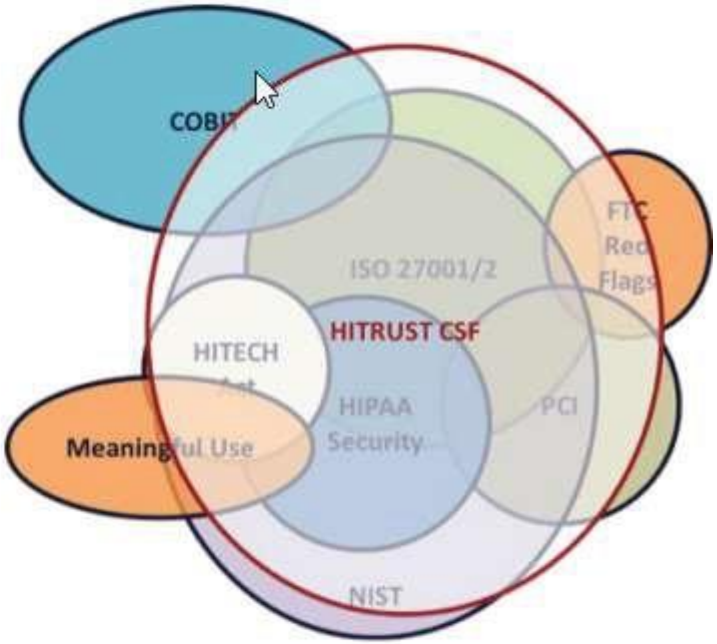
Safe Harbor
§ 164.514(b)(2)

Removal of 18 types of identifiers

No actual knowledge residual information can identify individual

# Building Trust with HITRUST Common Security Framework

- CSF provides a prescriptive and flexible privacy and security audit framework
- HITRUST is mapped to international and national cyber security standards
- Widely accepted in healthcare, and utilized to manage Business Associates
- Required by many healthcare payers
- An assessor reviews a series of controls covering multiple security domains



EPSTEIN
BECKER
GREEN

# Building Trust with a Data Governance Program

# Leadership Buy-In for a Data Governance Program

- Data governance programs should be top-down and bottom-up

- Directors should formally adopt guiding principles of data governance
  - Board-level policy
  - Charter of Board or Subcommittee thereof
  - Sets the tone for the data governance operations of an organization

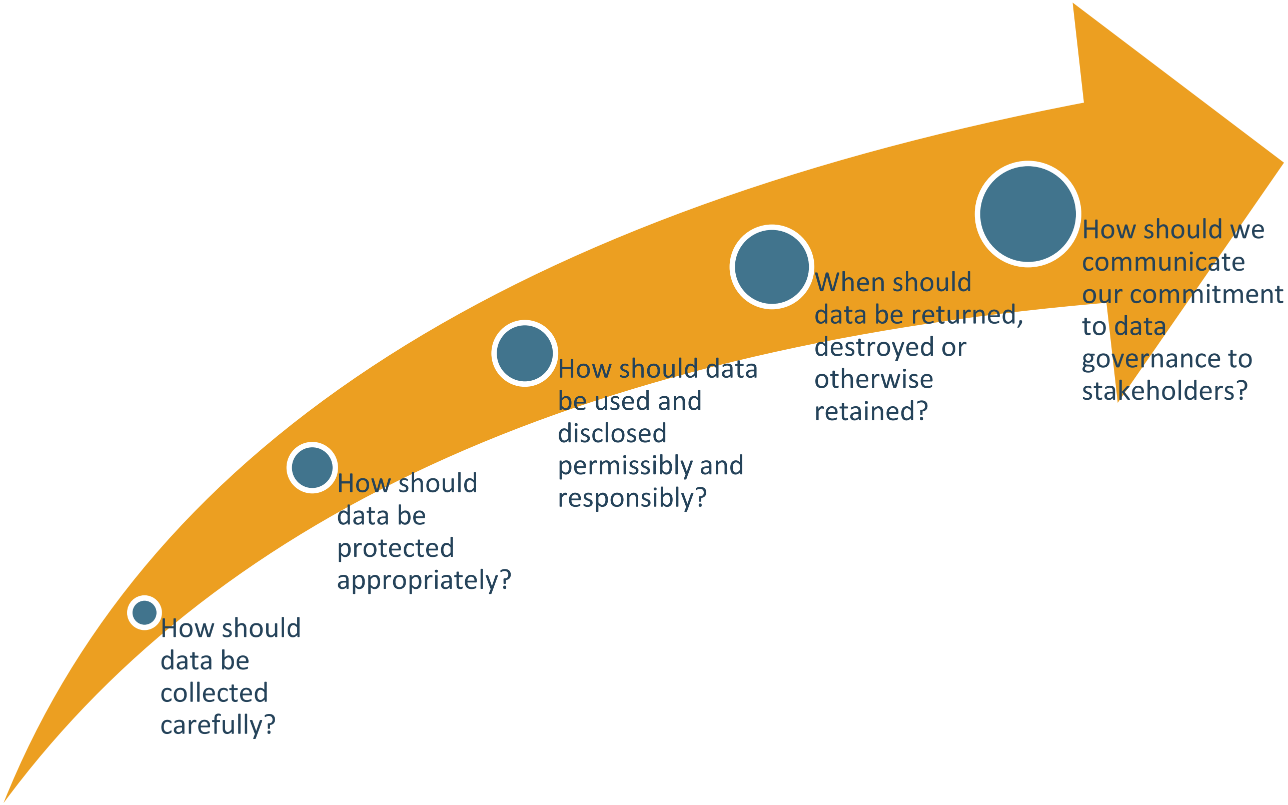
- Guiding principles may include:
  - An affirmative statements regarding commitment to good data stewardship
  - Statements evidencing the organization’s commitment to:
    - Adequate oversight and resourcing of data governance activities
    - Responsible collection, usage, and protection of health information
    - Transparency through open communication regarding data governance policies
    - Accountability by requiring the organization to report data governance issues

# Operationalizing a Data Governance Program

- Educate the C-Suite regarding the importance of data governance
- Adopt corporate-level data governance policy
- Establish a Data Governance Committee to develop policies and procedures in line with privacy and security requirements
- Appoint individuals to oversee the program as data stewards and data champions
- Establish reporting structures and metrics

Image credit: Shutterstock

# Key Data Governance Questions



How should data be collected carefully?

How should data be protected appropriately?

How should data be used and disclosed permissibly and responsibly?

When should data be returned, destroyed or otherwise retained?

How should we communicate our commitment to data governance to stakeholders?

# Data Governance Committees

Role of Data Governance Committees:

- Multi-stakeholder bodies
- Provide expertise on data governance issues
- Evaluate responsible internal and external uses and disclosures of data
- Distill legal, regulatory and business requirements into policy and guidance documents
- Promote transparency regarding an organization's data governance program

Image credit: Shutterstock

# Subcommittee Examples

## Data Governance Oversight Committee

- Examine policy, ethical issues, and legal and regulatory requirements
- Draft principal documents
- Draft policy statements
- Establish subcommittees as appropriate

## Data Access Committee

- Create processes to evaluate data sharing requests
- Evaluate requests before fulfillment
- Validate recipient of data has appropriate security safeguards in place to protect data
- Define contractual requirements governing use, disclosure and protection of disclosed data

## Data Quality Committee

- Contribute to development of data quality policies and procedures
- Create standard operating procedures for identifying, measuring, reporting, and resolving data quality issues
- Oversee routine data quality reviews
- Monitor and assess technology trends to identify potential strategies and solutions for enhancing data quality
- Identify and recruit subject matter experts to guide data quality improvement

# Data Access Policy

- Establish organization’s purpose of the data collection in line with its mission
- Consider developing a public-facing document explaining why and how your organization shares data
- Explain the process for requesting data including the review process
- Define key requirements and restrictions of the data access request process

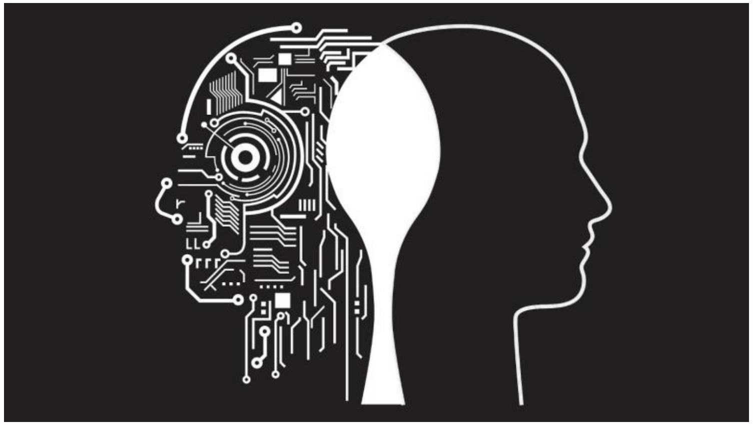
Image credit: World Meteorological Organization

# Data Access Review Process

- Sharing data can be beneficial to society, but it requires diligence
- Ensure your organization has sufficient authority to share the data (contractual/legal)
- Evaluate the propriety of requestors of data
- Evaluate the legitimacy of purposes for which data is requested
- Evaluate the feasibility of providing adequate date to support a request
- What is the minimum necessary data required to support a data request?
- Employ contracts to govern the sharing of data
  - Who owns the data?
  - What is the scope of the license to the data?
  - What minimum security safeguards must the recipient have in place?
  - If the data is de-identified, what restrictions exist relative to linking or re-identification?
  - What limits exist on re-disclosure of data to other third parties
  - What audit rights and notification obligations exist relative to compliance issues and security incidents?

# Looking Forward

- We are in a data-driven world
- AI, interoperability, and patient access rights will drive more data sharing
- HIPAA can serve as a great starting point to develop trust networks
- We ought to go beyond HIPAA's requirements to develop trust
- Robust data governance will be paramount in helping organizations navigate complexity in how to share data in the health ecosystem



Credit: Android Authority - https://www.androidauthority.com/complex-ai-ethics-833133/

# Questions?

**Alaap B. Shah**

**ashah@ebglaw.com**

Bedankt

谢谢您

# Thank you!

Grazie

Danke

Merci

謝謝您

Takk

Obrigado

Gracias