

**It's 3 a.m.....**

Do you know where your health data are?

# No Records Responsive to Your Request

Barriers to State Oversight of Health Data Sharing

Search Popular Latest

*The Atlantic*

**TECHNOLOGY**

# Google's Totally Creepy, Totally Legal Health-Data Harvesting

Google is an emerging health-care juggernaut, and privacy laws weren't written to keep up.

SIDNEY FUSSELL NOVEMBER 14, 2019

*The New York Times*

# Google to Store and Analyze Millions of Health Records

The tech company's deal with Ascension is part of a push to use artificial intelligence to aid health services.

**npr** SIGN IN NPR SHOP DONATE NOW

NEWS ARTS & LIFE MUSIC SHOWS & PODCASTS SEARCH

TECHNOLOGY

# Google Health Data Project Under Scrutiny

November 14, 2019 - 5:02 AM ET  
Heard on Morning Edition

SHANNON BOND

Home World U.S. Politics English Edition Print Edition Business Tech Video Podcasts Latest Headlines Markets Opinion Life & Arts Real Estate WSJ Magazine

# THE WALL STREET JOURNAL

## Inside Google's Quest for Millions of Medical Records

The company has struck deals that grant it access to troves of patient data; 'We want to be helpful'

By [Rob Copeland](#), [Dana Mattioli](#) and [Melanie Evans](#)

Jan. 11, 2020 12:15 am ET

# What does Google want health data for, anyway?

“[Google] was planning to build a search tool for medical professionals that would employ machine-learning algorithms to process data and make suggestions about prescriptions, diagnoses, and even which doctors to assign to, or remove from, a patient’s team.” (As per the Atlantic)

- Insight: a lot of problems in healthcare are actually well-suited for the application of data-driven / ML approaches
  - Diagnosis, treatment recommendations, admissions / discharges, bed assignments, ...

# Growing Demand for Health Data?

- “Health Data Analytics” sector
  - ~\$12 billion in 2018, projected to grow to \$80 bil by 2026
- Access to data is a huge competitive advantage when trying to train predictive models for care delivery
  - Top-tier algorithms are usually published in papers / available to the research community --- the “secret sauce” is usually in the parameters learned through extensive training.

# Thank God for HIPAA!

- Protected under the Health Insurance Portability and Accountability Act
- Key principle: Health providers can only share protected health data for reasons related to providing care
  - When a business wants to receive health data in order to provide a service (lab tests, EHR, diagnostic tools) they must enter into a Business Associate Agreement
  - BAAs specify the purposes for and the ways in which the business can use the data

# With One Little Problem...

- Anonymized health data is not protected under HIPAA
  - HIPAA has fairly robust standards for anonymization
    - (...Query whether anything can ever be anonymized)
- Certain BAAs allow the business to create anonymized health data from protected health data
  - Once anonymized, it can be used for anything (no protections!), inc. sold for profit

What can Google do with that data?

&

Are they the only one?

# The original research questions

- Investigate the BAAs entered into by different health care providers
  - Which companies are the agreements with?
  - What are the processing purposes for each?
  - What other data uses are allowed by the contracts?

# Methods

- Freedom of Information Act Requests!
  - FOIA laws for federal and state gov'ts
- Advantages of FOIA
  - Entity has to respond (or indicate that they cannot) within a certain timeframe
  - Entity has to turn over actual records (rather than just relating their impressions / PR line)
  - Should get a broader swath of records for each request
- Disadvantages of FOIA
  - Limited to gov't orgs and public hospitals

# Methods

- FOIA requests to:
  - State health agencies
  - County / municipal health agencies
  - Public hospitals & public hospital groups
- Submitted through different mediums
  - Online portal, email, phone call, request to use mail (?) or fax (???)

Maia Hamin  
15 Bank St  
Princeton, NJ 08542  
Phone: (413) 336-7757

Karen Onderdonk, Board Secretary  
St. Vincent General Hospital District  
822 W. 4th Street  
Leadville, CO 80461  
Phone: (719) 486-0230

Dear Ms. Onderdonk:

Under Colorado Open Records Law § 24-72-201 et seq., I am requesting an opportunity to inspect or obtain copies of public records that relate to data-sharing agreements between St. Vincent General Hospital District and private technology firms. I'm interested in records that pertain to business associate agreements that allow for the sharing of identifiable patient health data and partnerships which involve the transfer of de-identified patient data. These may take the form of documents related to contracts and procurement, budget and other fiscal information, internal guidelines around system usage or patient data sharing, or any other document which mentions or materially describes a patient health data recipient. I'm interested in records from within the last five years (e.g. those from after March 2015).

If there are any fees for searching or copying these records, please inform me if the cost will exceed \$15. However, I would also like to request a waiver of all fees, since the disclosure of the requested information is in the public interest. I am a student researcher seeking to compile and publish a study on patterns in health data sharing for the benefit of lawmakers and patients --- this information is not being sought for commercial purposes.

The statute requires a response to this request within three days following receipt of this letter. However, I understand that you are at the front lines of a developing crisis --- if access to the records is likely to impose a burden on your operations, just let me know and I'm happy to narrow the scope or to withdraw the request, as is most helpful.

Thank you for considering my request --- I hope you and your loved ones are in good health.

Sincerely,  
Maia Hamin

# Results

The Department is unable to identify any records responsive to your request.

- We submitted 28 public records requests in total
  - Six cities, seven hospitals, three counties, and nine state-level agencies
  - In 11 states (and the District of Columbia)
- Of the 22 requests that yielded a response, 20 were rejections, and only two yielded responsive documents:
  - Lemuel Shattuck Hospital (MA) And Cook County Health (IL)

Please be advised that there are no records responsive to your above-referenced request. With this response, your above-referenced request is now closed.

**you do not appear to be a citizen of Arkansas.**

# Findings (from rejections)

- Almost every state-level agency indicated that they did not hold the records *and* that they were not sure who would, other than the hospitals themselves
  - This indicates that there is little to no oversight of BAAs & their provisions
  - Hospitals are the entrants into these contracts *and* their oversight bodies
  - Makes it extremely challenging to document any kind of “state of the world” (would you need to FOIA every single hospital?)

# Findings (from successes)

- Both hospitals returned only a few BAAs, and indicated that returning any more would be onerous
  - Indicates that hospitals likely do not have a centralized or searchable database of their own agreements and the provisions / limitations of each
  - May actually violate some provisions of HIPAA (which requires that patients be able to request that their information be corrected or deleted)

# Findings (from successes)

- For both hospitals, the language of BAAs was extremely ambiguous --- even with a close reading, we struggled to determine whether the BAA would permit the BA to anonymize data for their own business functions
  - Use of jargonistic phrases like “health care operations”
  - Reliance on state law whose application may be ambiguous
  - Indicates that mere collection of records may not be enough for oversight, unless BAAs are more explicit about the specific uses of data that are allowed or prohibited

# Case Study: LabCorp

- One of the BAs from Lemuel Shattuck Hospital
- “Health care operations”
  - One of the seemingly-innocuous phrases that actually has a definition within HIPAA which encompasses business functions like anonymization
  - We’re still not sure whether LabCorp can anonymize data for its own use
  - However...
- Quest, LabCorp’s main competitor, makes millions of dollars per year from the sale of de-identified data
- LabCorp had also recently announced a partnership with Cioux Health to create a de-identified database of COVID-19 test results for use in epidemiological research and predictive modelling

# Some of the challenges

- Absolute lack of oversight seems... dangerous
  - However, there are reasons for the federated system of health oversight
- Do we want to impose limits on health data sharing?
  - Fine balance between protecting personal privacy and stifling innovation in an area where innovation is actually... really good.

What would you recommend?

# Recommendations

## Problem:

- Lack of any oversight mechanism, and general confusion about who is responsible for oversight

## Solutions:

- Require state Depts of Health to collect and maintain a database of these agreements as a part of their annual auditing work
- Require Depts of Health to have at least one employee whose responsibilities include collection, oversight, and auditing of BAAs

# Recommendations

## Problem:

- Lack of record-keeping at hospitals, making it difficult for them to answer questions about the use of patient data by third parties

## Solutions:

- Fed and state Depts of Health should clarify the reporting requirements for hospitals, including a requirement to standardize the process of querying the third parties to whom a specific patient's data has gone

# Recommendations

## Problem:

- Lack of transparency around what kinds of data use are permitted by BAAs, caused by patchwork of Fed and state laws as well as HIPAA jargon

## Solutions:

- Fed and state Depts of Health should introduce a taxonomy of health data uses and provide template language for allowing / restricting each
- Hospitals should store each BAA with information about the uses permitted and disallowed under this taxonomy,
  - Would make it possible to answer questions like “How many of your agreements allow for the anonymization and reuse of data?”

# Acknowledgements

HUGE thanks to my partner Maggie Baughman, who wrote this with me through fits of sleep-deprivation-induced delirium and got all the hard parts right

HUGE thanks to Prof Mayer, who helped us focus in on the most parts of the questions we were asking and the most-likely-to-succeed methods we had proposed

**No Records Responsive to Your Request:**  
Barriers to State Oversight of Hospital Data-Sharing Practices

**Eyes on the PHIs**  
Barriers to State Oversight of Public Hospitals' Data-Sharing Practices

**Curses, FOIA'd Again:**  
Record Availability Barriers to State Oversight of Hospital Data-Sharing Practices