



Strengthening
University-Industry
Partnerships

Data Part 1

Data Generation



What is “data” ?

- ❖ **Overbroad or unclear definitions, e.g.,**
 - “results”
 - “information”
 - lumped into “Intellectual Property”
 - Included in “Confidential Information”

- ❖ **Where did the data come from?**
 - Provenance of Data
 - Permissions and consents

- ❖ **Preexisting or New; Coming In or Going out**
 - Required to be generated per SOW
 - Required to be provided for use by other party to perform SOW

- ❖ **Project context**
 - Human subjects
 - Trade secrets
 - Export controlled
 - Unique to the SOW

Data Considerations Stem From 5 Areas

1. Laws and Regulations
2. Ethics, Privacy, Societal Concerns
3. Policies
4. Contracts and Agreements
5. Practicalities and Logistics



Strengthening
University-Industry
Partnerships

Data Part 2

Data Use

.

Who can do what with the defined “data”?

- **Ownership vs. use**
 - “commercial” use
 - “internal” use
- **Exclusive or nonexclusive**
 - Publication and sharing
 - Monetization and licensing
- **Expectations and limitations**
 - Training AI models
 - Deposit to registry
 - Augment
 - Regulatory use



Situation #1

A faculty member (no signature authority to bind University), executes an online agreement with a .org entity for access to their data

Uses the data to train Machine Learning /AI algorithm

Receives interest from a 3rd party for the trained data model (not the data)

University wants to license the trained model (but technically does not have access to the data that was used to create it)

TO DO?

Does the University need an agreement for data access?

Will the agreement need to include terms enabling the use to train and license the model?

What if the data used to train the model is “scraped” from the internet?

What if the data allows identification of the institution or individuals?



Data Registry

- A U.S. Tax exempt non-profit entity maintains a data registry with information about patients who have undergone a particular medical procedure. The entity is not a Covered Entity under HIPAA.
- The entity has Members that submit data about the patients which contains some variables like race and dates of service but not other personal identifiers (i.e., a HIP Limited Data Set.)
- The entity permits the Members to use data from the Registry for research and QI but the data provided by the Entity is deidentified and does not connect the data to the Member who submitted it.
- There are hundreds of Members some of which are in the EU and other foreign countries.
- Some Members are corporate device developers.

TO DO?

How can the Entity comply with foreign data laws like the GDPR subject's right to know who has their data and the right to have it deleted?

How can the Entity meet security requirements that may vary across the Member countries?

Can Entity policies cover these foreign regulatory requirements?



Challenges for Data Projects

- Personalized healthcare solutions require access and coordination of large amounts of data
- Other industry sectors use big data in product development and consumer products
- AI and ML need to “learn” from data
- Does social good result from data sharing?
 - Is data a “public utility”?
 - Should non-identifiable data be freely shared for the “right” purposes?
- Is privacy now the dominating factor?



Societal and Institutional Challenges

- What institutions/beliefs must change to accommodate the new paradigm?
- Can/should regulatory/legislative changes favor data use and sharing?
- What is needed to ensure consumers are comfortable with data sharing?
- What is the best way to incentivize data collection and data sharing by research participants, researchers, funders, archivists, users?
- How can divergent interests be aligned?
- How can regulatory schemes be harmonized across
 - U.S. and state agencies?
 - U.S. and other countries?
 - Corporate Sectors?
 - Universities?



Strengthening
University-Industry
Partnerships

Thank You!

•

Let's discuss