

A Licensing Model and Ecosystem for Data Sharing

Jane Greenberg, Sam Grabus (Drexel Univ.); Famien Koko, Sam Madden, Tim Kraska (MIT)

Project Goals

1. Licensing Framework / Generator
2. Data-Sharing Platform (Enforce Licenses)
 - DataHub → ShareDB
3. Metadata (Search Licenses and Data)

Progress

- Cognitive walkthrough improved workflow support and ShareDB design.
- ShareDB testing, small-scale datasets.
- Differential Privacy model developed that can select the best differentially private algorithm, given an error tolerance.
- Rights and Licensing Landscape environmental scan.

Regulations

- HIPAA ⓘ
- FERPA ⓘ

Privacy ⓘ

- PII Anonymized or Removed
- PII Anonymized
- PII Removed

ShareDB

Guide to using ShareDB: Privacy Profiles

N E X T

- Investigate additional anonymization, PII detection, and watermarking techniques.
- First pilot study w/industry and academia.
- Continuing to collect agreements, revise ontology
- Organize stakeholders workshop.
- Explore secure hold-out datasets to support false-discovery control.



IIS/BD Spokes/Award
#1636788