

大学の研究データガバナンスを支えるポリシー・規程類とツール & 大学における研究データ管理お悩み相談

共同研究データ管理: 所内外の教員、学生、企業との試行錯誤

国立情報学研究所情報社会相関研究系

水野貴之

研究スタイル

- 民間等が保有する社会/経済のビッグデータから，データ解析によって，社会課題や人間活動に関する特徴量を抽出．データにもとづく現象のモデル化により，社会課題を解決するための戦略的な制御方法を探る．
- 他大学の研究者，学生（日本の修士→海外の博士課程で空白期間が発生する場合も），民間企業，官庁との共同研究
- 論文など，フィジビリティスタディ，技術指導



- ビッグデータ：大容量
- 産官学連携：アクセス権の異なる複数ユーザー
- リアルタイム性：ナウキャスト，定期的な更新

ハード面：共同研究データ管理

最近多い100GB程度のファイルを共有するときは、同期処理 or レジュームが必要

- ファイルサーバ + バックアップサーバ + コールドストレージ
 - オリジナルの社会/経済のビッグデータの管理（100TB単位）
 - 消さない，書き換ええない
 - 強いアクセス制限と，ファイル自身もパスワードによる保護
- 法人向け Google Workspace
 - 研究室全員分のアカウントを作成
 - Google Drive = アカウント数 × 2TB（実質，100GB～500GB単位）
 - データの受け渡し，バージョン管理，資料等の共同編集
 - （クラウド管理可能な）契約書等の書類，研究資料，ビデオ，議事録など
- Dropbox Plus
 - 共同研究先がGoogle driveが利用不可である場合に利用
- Mattermost（オンプレミス）
 - 日々の発言の記録

共同研究プロセスとデータ管理目的

1. 出会い
2. 顔合わせ
3. 共同研究計画の作成
4. 共同研究
5. 論文や知財化
6. 社会実装・技術指導

時間コストをかけずに共同研究データ管理がしたい

- ~~性善説にもとづいて、口約束で進める~~
 - プロジェクト解散後に問題発生
 - ちょっと応用した成果の帰属
 - オーサーシップ
 - 知的財産
- 当時のことが、改変なく分かる
- 同意事項として、いつ、誰が、何を言ったか。
- データの流出を防ぐ
- データの利用規約の確認
 - アクセス範囲
 - 利用範囲

2. 顔合わせ, 3. 共同研究計画の作成

- どのような協力関係がありえるのか意見交換するフェーズ
→ゴールは, NDA
- 共同研究契約の作成に向けたフェーズ
→ゴールは, 共同研究契約

必要なデータ管理項目

1. アジェンダ等, 議事次第
2. ミーティングの動画 (2 GB程度 × 数回)
3. ミーティング資料
4. 議事録 (特に, 発言の修正などがあれば記録する)
5. NDA/共同研究契約書の雛形の作成
6. 締結したNDA/共同研究契約書

4. 共同研究

- 共同研究フェーズ

必要なデータ管理項目

1. 各データの利用規約
2. データの共有（～200GB(オンライン)×数個）
3. データ保有者リスト（私→共同研究者→学生・ポスドク）
4. コードの共有
5. 解析結果（～5GB×数十個）
6. アジェンダ等，議事次第
7. ミーティングの動画（2GB程度×数十回）
8. ミーティング資料
9. 議事録（特に，発言の修正などがあれば記録する）

5. 論文や知財化

- 成果をまとめるフェーズ

必要なデータ管理項目

1. データの更新（社会課題は生もの）によるバージョン管理
2. データの入手先(担当部局等)や入手方法，入手条件
 - オープンデータに出来ない代替りの措置
3. 学会等での発表資料，アブストラクト
4. 論文，紀要，解説記事
5. 成果物の作成に利用したデータ
6. 特許明細
7. 離脱者（主に学生）からの引き継ぎ

6. 社会実装・技術指導

- 公知，知財，ノウハウの切り分け

必要なデータ管理項目

1. NDA/共同研究契約/受託研究契約/ライセンス契約の作成，締結
2. データや技術の利用範囲の管理
3. 技術を搭載したAPI等のアクセスログの管理
4. データやAPI等のバージョンの管理
5. アジェンダ等，議事次第
6. ミーティングの動画（2 GB程度×数回）
7. ミーティング資料
8. 議事録（特に，発言の修正などがあれば記録する）
9. アイデアや成果物の管理
 - ブレインストーミングでの雑多なアイデアの管理
10. 終了後のフォローアップ（報告義務，削除義務など）

まとめ：重視している点

1 プロジェクトに必要なスペック

- 共同研究のベースとなる1次整理済みのデータ（100GB）
 - データに紐づく利用規約
- 解析者が加工した2次整理済みのデータ（200GB）
- ミーティングの動画（2GB×数10回）
- アップロード/ダウンロードの容易さ
- 契約書，議事録，成果物，利用者の管理に関する書類
- バージョン管理

管理者の責任で，制限付きのアクセスを許可したいユーザー

- 卒業した学生
- 企業の共同研究者，協力者
- 海外の共同研究者（教員・ポスドク）
- （引き続き共同研究がしたい）海外の大学に進学した学生
- フィジビリティスタディの参加者
- URAやコーディネーター，契約担当部局