

材料科学分野における データ利用のライセンスの考え方

- 図書館からの視点

2017年2月14日 第3回SPARC Japanセミナー2016

小野寺 千栄（物質・材料研究機構 科学情報プラットフォーム）

物質・材料研究機構（NIMS）について

- 国内で唯一の物質・材料科学研究の専門機関
- ミッション
 - 物質・材料科学技術に関する基礎研究および基盤的研究開発
 - 研究開発成果の普及とその活用の促進
 - NIMSの施設および設備の共用
 - 研究者・技術者の育成およびその資質の向上
- 職員数：約1,500人
 - 研究に関わる職員は約半分

図書セミナー「オープンサイエンスと著作権」第1回

- **2016年1月8日開催**
- **テーマ：オープンサイエンスと著作権**
- **講師：内閣府、弁護士**
- **参加者：101名**
 - NIMS（研究系職員）：48名
 - 外部：53名（研究機関、図書館、政府機関、出版社・学会、大学）
- **概要：**
 - オープンサイエンス政策への取り組みと今後の展開。
 - 論文と研究データの著作権。現行の著作権法はデジタル開示された著作物の利用には不十分であり、契約で補う。データは著作物ではなく、著作権法の対象にならないため、利用者が利用する条件は契約で規定する。

図書セミナー「オープンサイエンスと著作権」第2回

- 2016年3月10日開催
- テーマ：（続）オープンサイエンスと著作権－データ共有
- 講師：出版社、研究者、弁護士
- 参加者：83名
 - NIMS（研究系職員）：51名
 - 外部：32名（企業、研究機関、政府機関、大学、図書館、出版社）
- 概要：
 - データキュレーターから見るデータ共有（データジャーナル）。
 - NIMSにおけるデータ収集と共有。
 - データベースが著作物として認められることは、過去の判例を見ても非常に稀。著作物ではないデータは自由に利用できるが、データ出典元との契約があれば、その契約に従う。

図書セミナー「オープンサイエンスと著作権」第3回

- 2016年11月24日開催
- テーマ：(続々)オープンサイエンスと著作権－データマイニングからAIへ
- 講師：弁護士、出版社
- 参加者：82名
 - NIMS（研究系職員）：61名
 - 外部：21名（研究機関、大学、企業、出版社、政府機関、図書館）
- 概要：
 - AIをめぐる知的財産権の論点について：AIの学習用データの権利をどこまで著作権で保護するか。アルゴリズム・学習モデルの保護範囲（あるいは特許、プログラムの著作権として保護）、AIによる自動生成物は著作物か否か。
 - 出版社におけるテキストデータマイニングの実例紹介。

NIMSにおけるデータ：case 1 作成者として

●研究データのメタデータ付与の効率化、共有化の取り組み

○計測データのメタデータの記録システム

課題	対応方法
手書きのラボノート	データ採取時にデジタル化して記録
装置の操作で手がふさがる	音声入力
全てを入力するのは面倒	可能な箇所はシステムが自動計算
実験手順と入力項目の順番の相違	入力したい項目を入力したいタイミングで入力
装置が異なると、語句が異なる	語句の違いを吸収するためのテンプレート
メタデータと実験データの対応付け	メタデータ記録システムに実験データも保存

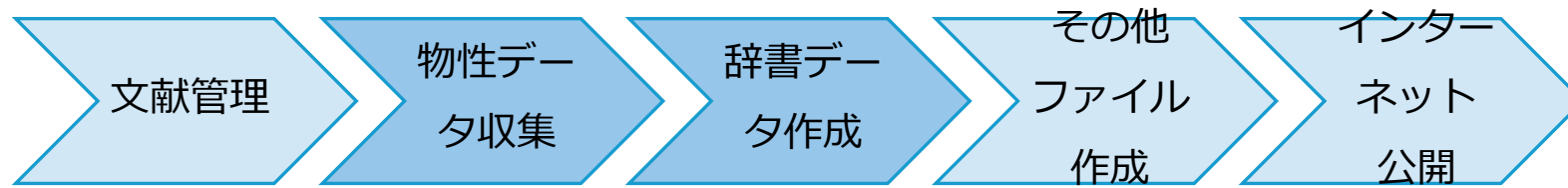
○学会でも議論を共有して、課題解決に取り組む

NIMSにおけるデータ：case 2 利用者として

～ データサイエンティストによるデータ収集 ～

●物質・材料データベースサイト（MatNavi） > 高分子データベース

- 高分子材料設計に必要とされるデータの提供
- データ収集元：出版論文（購読誌、OA誌）から**データのみ**を収集
- 作業工程



データ収集・辞書データ作成はデータサイエンティストとして研究OBが担当



<http://mits.nims.go.jp/>

NIMSにおけるデータ：case 3 利用者として

～ ソフトによるデータ収集 ～

●マテリアルキュレーションのためのデータ収集

- 背景：材料探索の前提となる多量のデータを取得する
- 方法：スキャンした論文PDFのグラフから、**数値データを抽出**するソフト開発
- 開発したソフトを特許取得し、希望者とライセンス契約
 - ソフト自体はデータを持たない
 - 著作権上の留意点については、研究者自身が弁護士に相談している（ソフトの配布は問題ないが、**第三者がソフトを利用してグラフから抽出したデータをグラフに描画することは、著作権侵害の責任を負う可能性がある**）

NIMSにおけるデータ：case 4 提供者として(1)

～ 段階的な提供範囲の限定 ～

●クリープデータシート

○ 公開方法

1. 無償

- 印刷物（配布申請機関のみ）
- Web（要ユーザー登録、閲覧のみ、印刷・保存不可）

2. 有償（英国 グランタ・デザイン社を通じて）

- データベースソフトウェアとのセット販売

3. 詳細データの個別契約（NIMSとの直接契約）



NIMSにおけるデータ：case 4 提供者として(2)

●利用を認める範囲

1. 無償配布

- OK：データをグラフにプロットする、データ解析結果を公表する
- NG：データの営利目的利用
 - ⇒ 配布/ユーザー申請時の理由によっては、拒否することもある

2. 有償販売

- 利用範囲は契約に基づく

3. 詳細データの個別契約

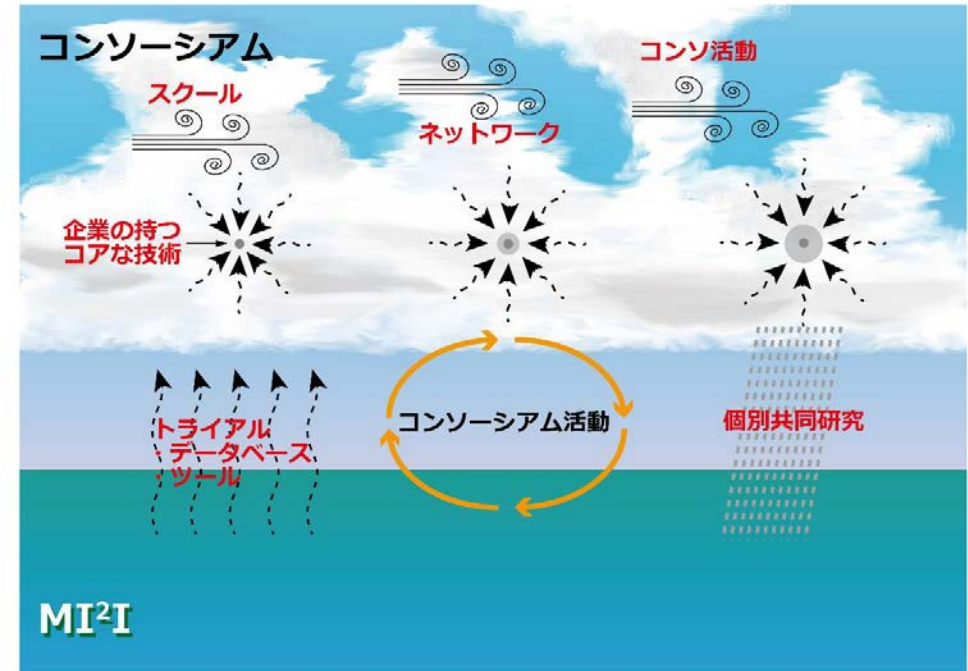
- データ開示禁止
- 第三者への提供禁止
- 解析結果の公表は可能

NIMSにおけるデータ：case 5 提供者として

～ 限定的な提供範囲 ～

●情報統合型物質・材料開発イニシアティブ (MI²I)

- JSTのイノベーションハブ構築支援事業の採択課題
- データ駆動型の物質・材料研究への取り組み
- コンソーシアム会員を募って活動
- 会員の利用規則に沿ったデータ利用
- データプラットフォームはNIMSが提供

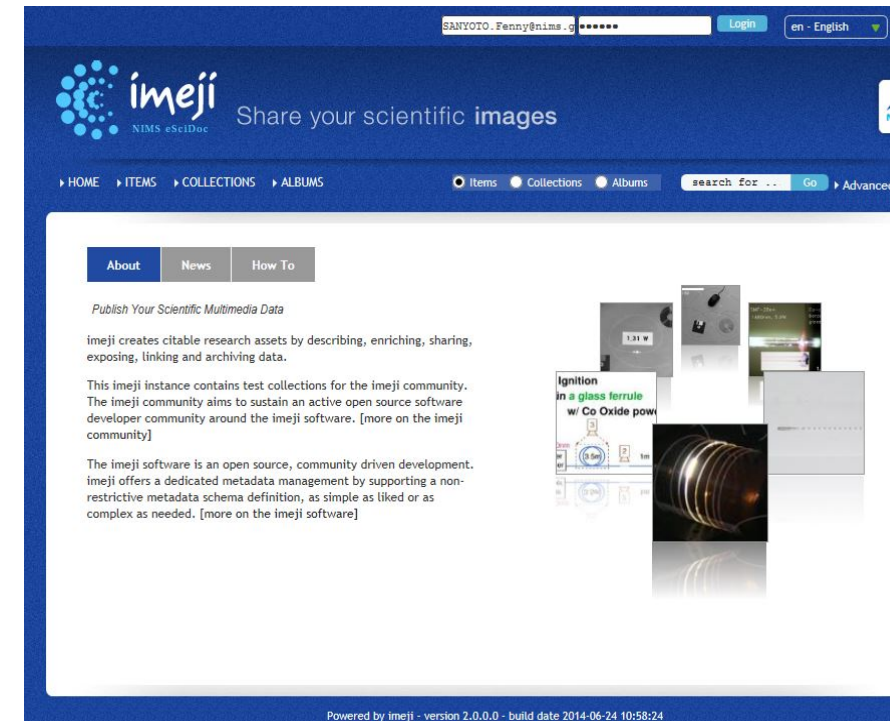


コンソーシアム活動の概念図

<http://www.nims.go.jp/MII-I/h6g4rf00000002oc-att/h6g4rf00000002pe.pdf>

図書館業務とデータ(1)

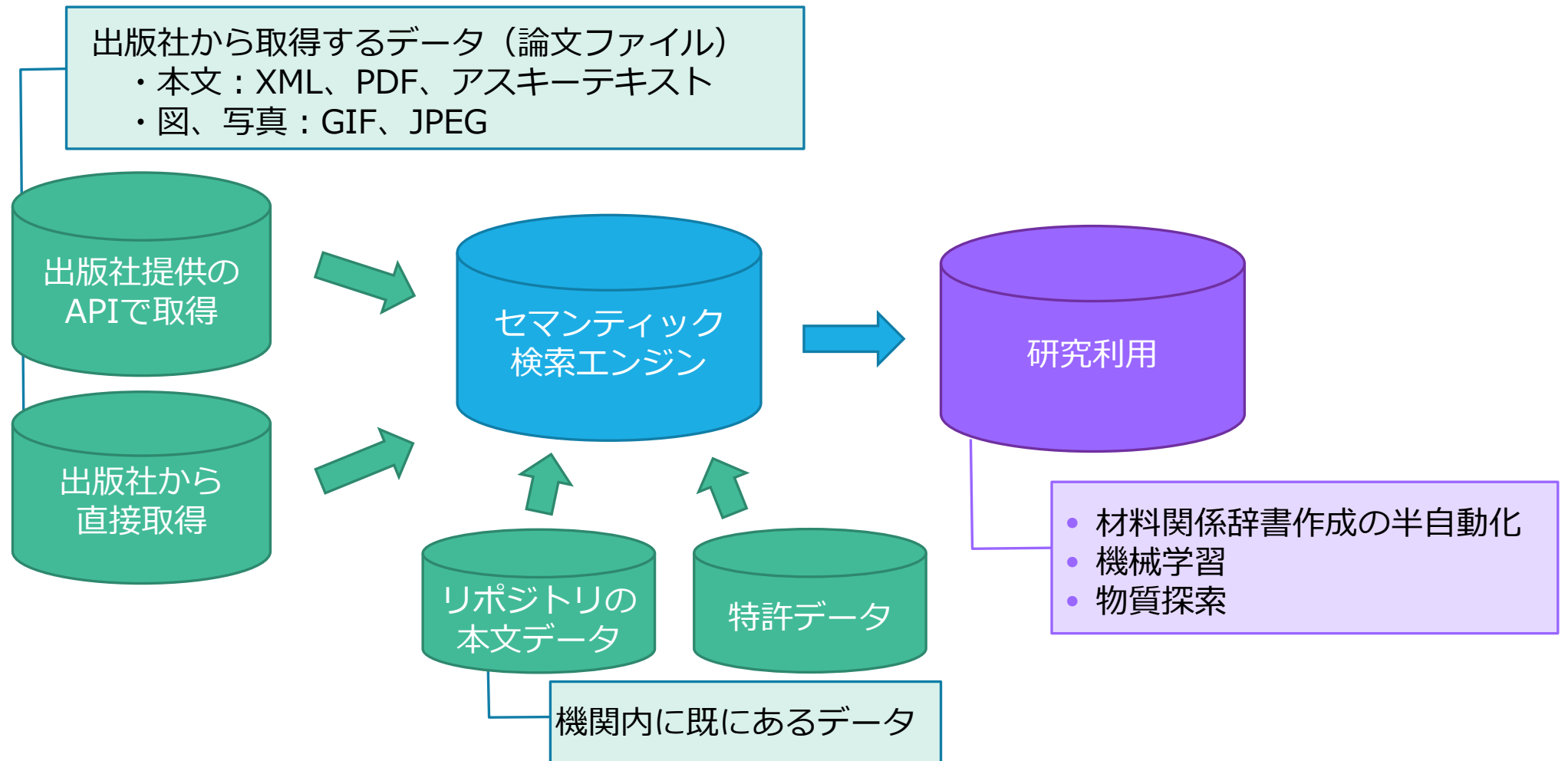
- データを扱うには、その分野の専門知識も必要
- 図書館員が今できることは？
 - 資料の保存と管理
 - ⇒ 研究者にデータの保存と管理に必要な情報を伝える
 - 機関リポジトリのノウハウ
 - ⇒ 研究者に保存と提供のプラットフォームを用意する
 - ジャーナル購読
 - ⇒ テキストデータマイニングのための
 - ・ ジャーナルアクセスライセンス
 - ・ 論文ファイルの収集



<http://imeji.nims.go.jp/imeji/>

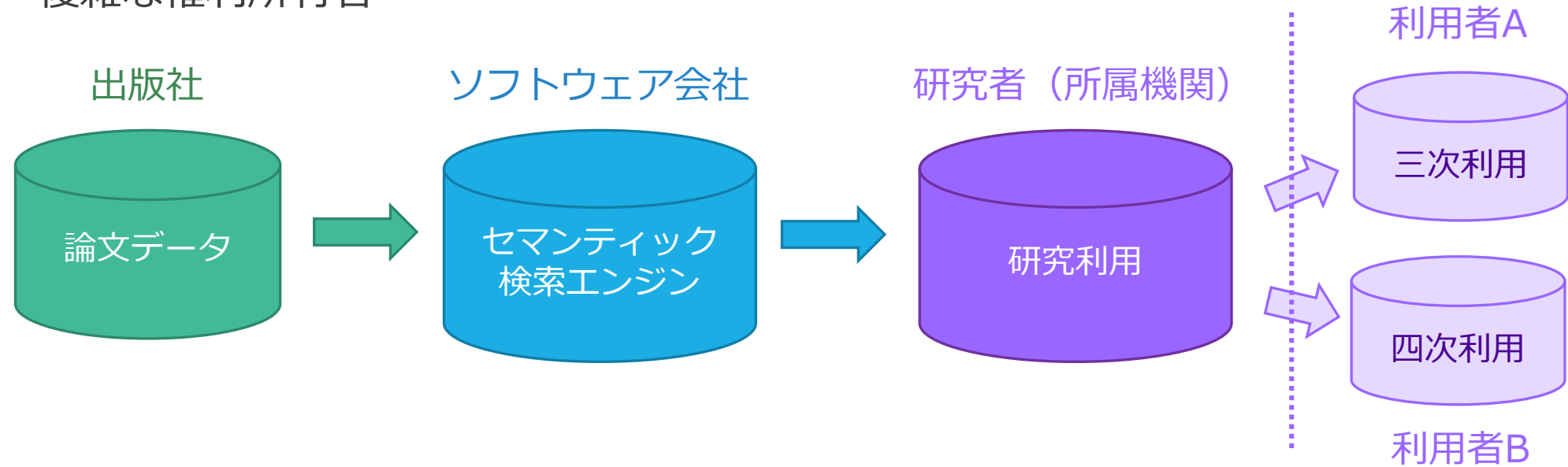
図書館業務とデータ(2)

●テキストデータマイニングのための論文ファイルの収集



- ライセンス契約の際に、どこまでを利用範囲とするか

- 複雑な権利所有者



- 利用をコントロールしたいデータをどのように提供するか

- 利用範囲の明確化