



OKAYAMA
UNIVERSITY

米・コーネル大学出張報告

研究データ管理 etc.

2025/5/13 COAR Annual Conference 2025

岡山大学図書館 情報管理課
情報整備グループ 石丸 綾華

期間：2024年12月14日～20日
訪問先：コーネル大学（アメリカ）

- 1865年設立の私立大学
メインキャンパス：ニューヨーク州イサカ
- アイビーリーグ所属
THE世界大学ランキング20位
- 学部数：8
大学院数：7
- 学生数：26,793人
教職員数：11,674人



12月16日～17日：Olin Library, Mann Library (イサカ)

12月18日：Weill Cornell Medicine (マンハッタン)

メンバー：図書館長、電子ジャーナル契約担当 主査、
機関リポジトリ担当、医学系分館職員、URA (計5名)



Cornell Data Services

Data Management Plan の作成からストレージ選定まで至れり尽くせり

研究を始めようかな（と研究者が思ったら）

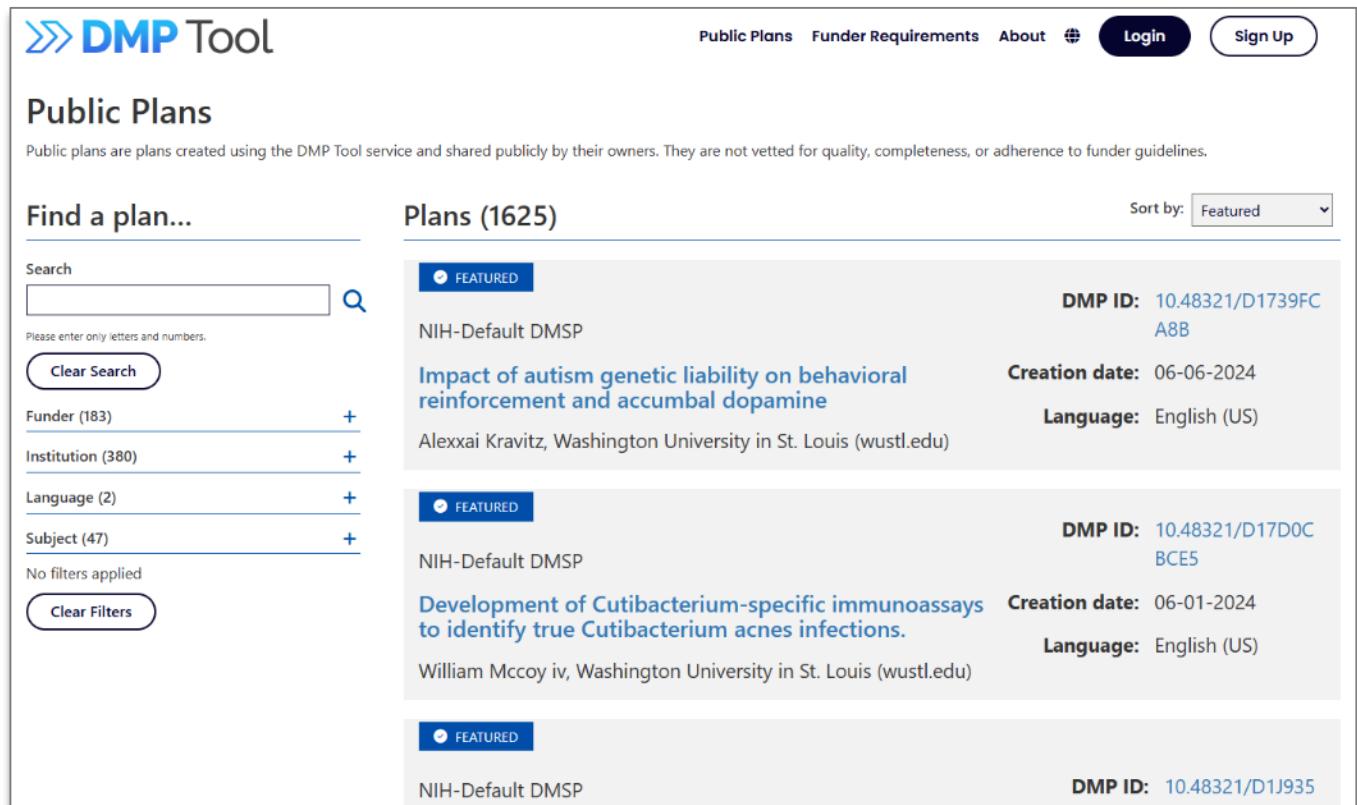
研究開始前に DMP Tool を参考にして DMP を作成

研究データを管理するストレージを Data Storage Finder で選定

研究スタート！

研究中も、研究データのデジタル化や分析までサポート

- 既存の DMP を参考にして作成を進めることができる
資金提供者・機関・言語・主題で絞り込み、用語検索
- 各資金提供者ごとのテンプレートもあり



The screenshot shows the 'Public Plans' section of the DMP Tool. On the left, there is a search bar with a placeholder 'Search' and a note 'Please enter only letters and numbers.', a 'Clear Search' button, and a 'Clear Filters' button. Below the search bar are filters for 'Funder (183)', 'Institution (380)', 'Language (2)', and 'Subject (47)'. The main area displays a list of 1625 plans, each with a 'FEATURED' badge, the plan title, the creator's name and institution, the DMP ID, the creation date, and the language. The first three plans are listed:

Plan Title	Creator	DMP ID	Creation Date	Language
Impact of autism genetic liability on behavioral reinforcement and accumbal dopamine	Alexxai Kravitz, Washington University in St. Louis (wustl.edu)	10.48321/D1739FC A8B	06-06-2024	English (US)
Development of Cutibacterium-specific immunoassays to identify true Cutibacterium acnes infections.	William Mccoy iv, Washington University in St. Louis (wustl.edu)	10.48321/D17D0C BCE5	06-01-2024	English (US)
NIH-Default DMSP		10.48321/D1J935		

Data Storage Finder



- 研究で扱うデータの特性を入力するだけで、利用できるストレージサービスを絞り込み比較できる
- コーネル大学図書館職員とコーネル大学IT部門が内製コードが [GitHub](#) で公開されている

Describe your data

Answer these questions to help identify data storage services that are suitable for your needs. Checking these boxes will change the list of available services. If you are uncertain how to answer, leave the question blank to maximize your resulting options.

[Clear Answers](#)

1. What is the classification of your data?

- Public / Low Risk
- Sensitive / Moderate Risk
- Confidential or Restricted / High Risk
- HIPAA-Regulated

2. Do you need backups, snapshots or replication of your data?

Select data storage services you would like to compare.

[Select All](#) [Clear Selections](#)

Amazon Web Services Elastic Block Store	Amazon Web Services Elastic File System	Amazon Web Services Glacier	Amazon Web Services S3
BioHPC Cloud	CAC Archival Storage	CAC Red Cloud Storage	CCSS Data & Reproduction Archive
CCSS Regulated Research Environment	CCSS Research Servers	CUGIR	CUL eCommons

Compare services that match your selected criteria

[Select All](#) [Clear Selections](#) [Amazon Web Services S3](#)

Amazon Web Services Elastic Block Store	Amazon Web Services S3
Brief Description	Amazon Elastic Block Store (EBS) provides persistent block storage volumes for use with Amazon EC2 instances in the AWS Cloud. EBS volumes are automatically replicated to provide high availability and consistency. EBS offers high performance and the ability to scale your storage up or down quickly while paying only for what you use.
Example Use	Storing data files to be accessed by a single Amazon EC2 server.
Cost	6 666 Cost dependent on storage class, replication and other services used.
Capacity	No maximum file size limit. No practical overall limit. No practical limit to number of files (limits imposed by EBS volume size).
Access and Collaboration	End user access control to the content of EBS volumes is implemented by the EC2 instance to which an EBS volume is attached. Access and management of EBS volumes themselves is controlled by AWS IAM (Identity and Access Management), which can be used to control Active Directory users for authentication and authorization.
Data Allowed	Allowed: Public data, sensitive / moderate risk data. Allowed with special configuration: HIPAA-protected data, confidential / restricted / high risk data. Contact Amazon Web Services Elastic Block Store for more information. Not Allowed: HIPAA-regulated data, export controlled information.
Durability	High

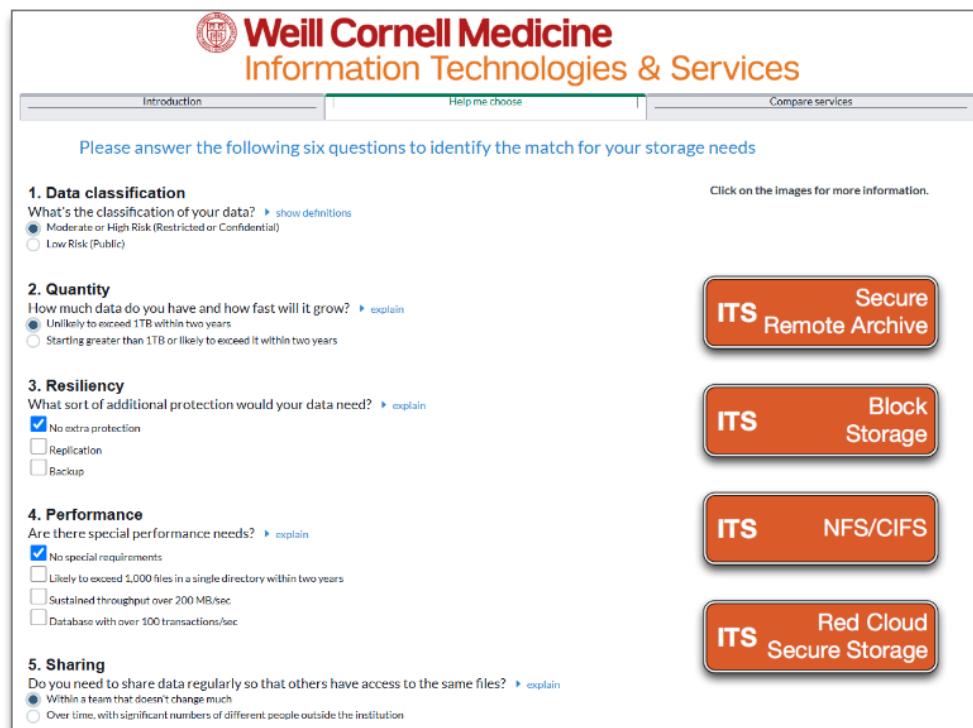
扱うデータの特性について
チェックを入れていくだけ

選択したストレージ
間で比較ができる

“Data Storage Finder”. Cornell Data Services. <https://finder.research.cornell.edu/> (最終閲覧日 : 2025/5/26)

5

- 限られた管理者のみアクセス可能なセキュアなシステム
- AWS 上に完全に閉じた計算資源を確保し、外部ネットワークとの接続ができないようになっている
- こちらも研究データの特性に合わせてストレージを選定可能



Weill Cornell Medicine
Information Technologies & Services

Please answer the following six questions to identify the match for your storage needs

1. Data classification

What's the classification of your data? [show definitions](#)

Moderate or High Risk (Restricted or Confidential)
 Low Risk (Public)

Click on the images for more information.

2. Quantity

How much data do you have and how fast will it grow? [explain](#)

Unlikely to exceed 1TB within two years
 Starting greater than 1TB or likely to exceed 1TB within two years

3. Resiliency

What sort of additional protection would your data need? [explain](#)

No extra protection
 Replication
 Backup

4. Performance

Are there special performance needs? [explain](#)

No special requirements
 Likely to exceed 1,000 files in a single directory within two years
 Sustained throughput over 200 MB/sec
 Database with over 100 transactions/sec

5. Sharing

Do you need to share data regularly so that others have access to the same files? [explain](#)

Within a team that doesn't change much
 Over time, with significant numbers of different people outside the institution

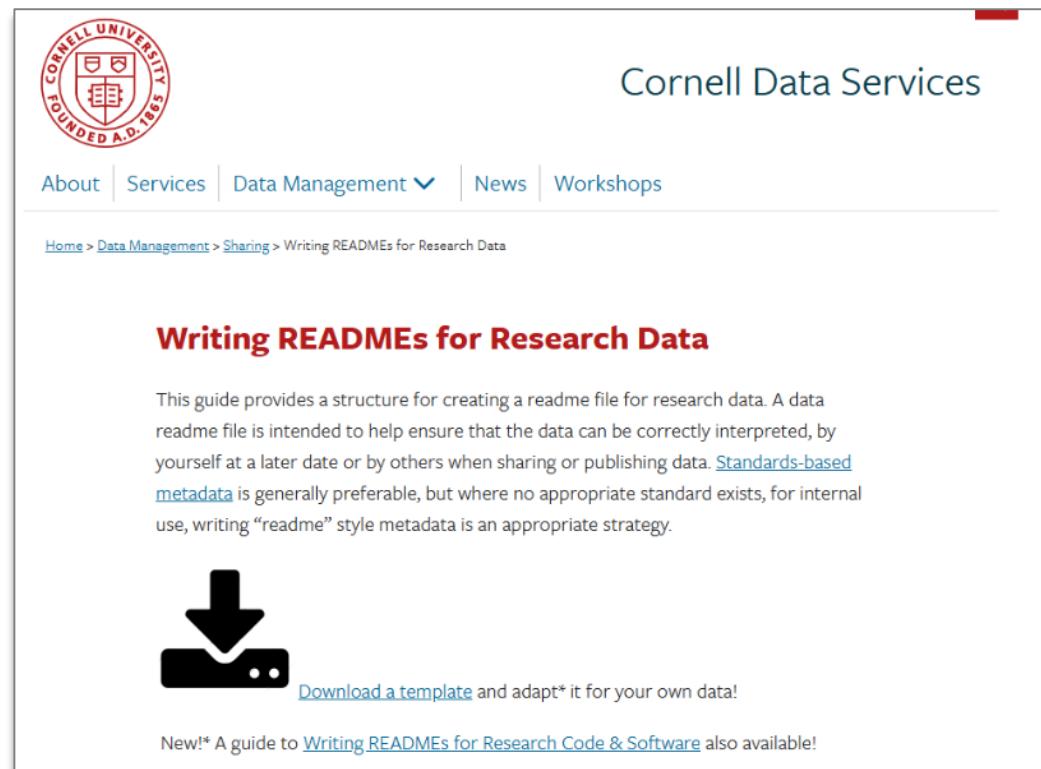
ITS Secure Remote Archive

ITS Block Storage

ITS NFS/CIFS

ITS Red Cloud Secure Storage

- 分野によってスタンダードな形式あり
- readme
連絡先、データに適用される制限やライセンス情報、引用文献、プロトコルや略語の説明、親フォルダ etc.

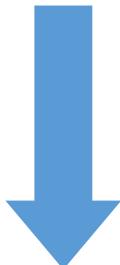


The screenshot shows the Cornell Data Services website. At the top, there is a red Cornell University seal with the text "CORNELL UNIVERSITY" and "FOUNDED A.D. 1865". To the right of the seal, the text "Cornell Data Services" is displayed. Below the seal, there is a navigation bar with links for "About", "Services", "Data Management", "News", and "Workshops". The "Data Management" link is currently selected, indicated by a dropdown arrow. Below the navigation bar, the URL "Home > Data Management > Sharing > Writing READMEs for Research Data" is shown. The main content area has a red header "Writing READMEs for Research Data". The text in this section explains that the guide provides a structure for creating a readme file for research data. It states that a data readme file is intended to help ensure that the data can be correctly interpreted, by yourself at a later date or by others when sharing or publishing data. It notes that "Standards-based metadata" is generally preferable, but where no appropriate standard exists, for internal use, writing "readme" style metadata is an appropriate strategy. At the bottom of this section, there is a large black download icon with a white arrow pointing down and three dots below it. To the right of the icon, the text "Download a template and adapt* it for your own data!" is displayed. Below this, a note states "New! A guide to [Writing READMEs for Research Code & Software](#) also available!".

"Writing READMEs for Research Data". Cornell Data Services.
<https://data.research.cornell.edu/data-management/sharing/readme/>
(最終閲覧日 : 2025/5/26)

データキュレーターチームが事前にデータ確認

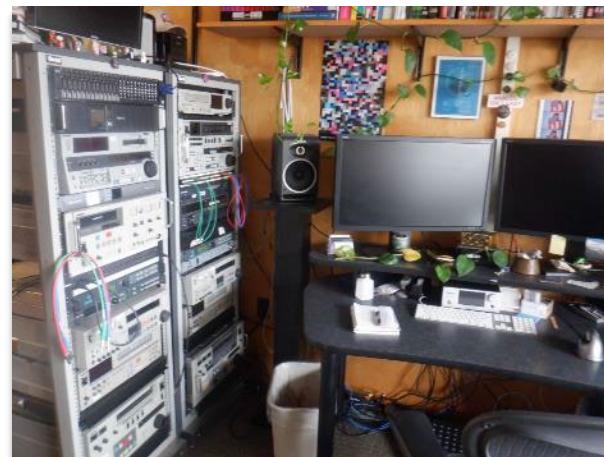
研究者がデータセット、ドキュメント、メタデータファイル、
readme ファイル等を提供



ファイルレベルでキュレーション、フィードバックとアドバイス
ファイルが揃っているか、開いて正常に動作するか

- DCN (Data Curation Network) に所属
データセットを DCN に送信 → 他機関のキュレーターへ

アナログデータからデジタルデータへの変換
様々な機材を用いて専門スタッフがサポート



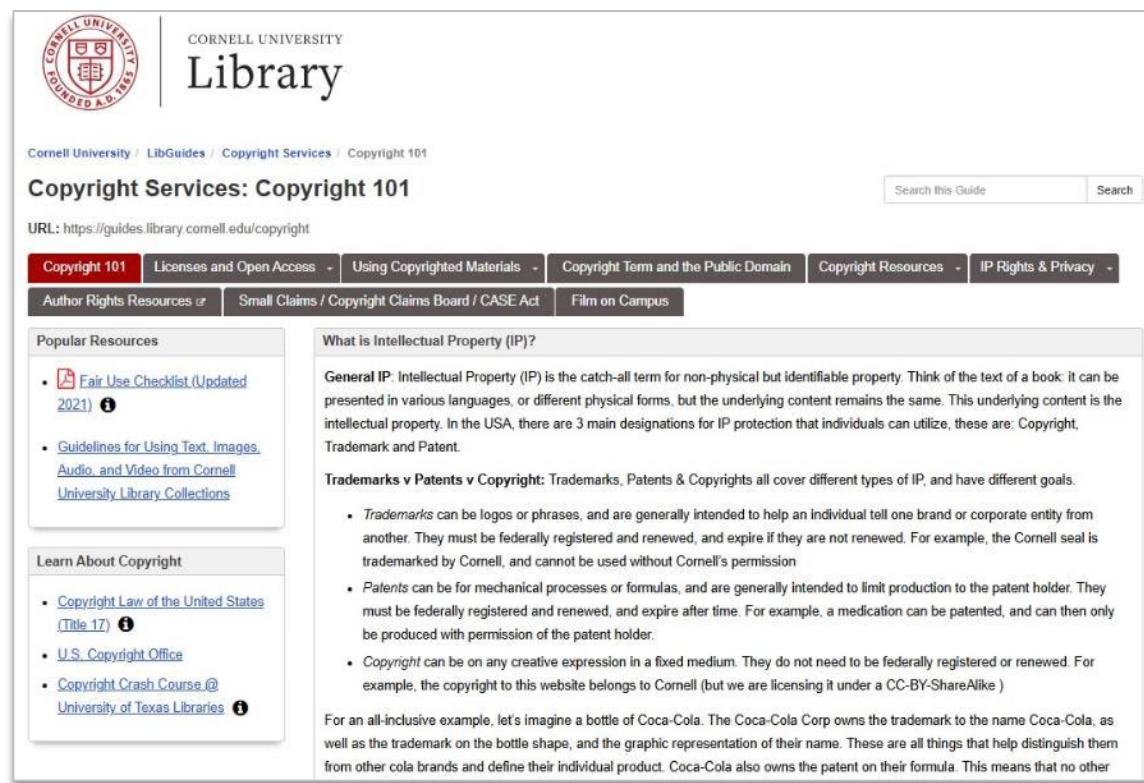
- デジタルデータを扱うための支援
ライブラリアン3名 (+図書館の他スタッフもサポート)

例えば.....

テキストマイニング、機械学習やアルゴリズムの使用、
データベースの作成、メタデータ付与、データの可視化
等を支援（かなり専門的）

- 様々な学部の学生や研究者の集まる場として、
関連する研究内容を持つ研究者を紹介するという支援も

- 包括的なウェブサイトを作成
パブリックドメインのチャートが特に人気
- ウェブサイト+アウトリーチで提供する教育活動



CORNELL UNIVERSITY
Library

Cornell University / LibGuides / Copyright Services / Copyright 101

Copyright Services: Copyright 101

URL: <https://guides.library.cornell.edu/copyright>

Search this Guide

Copyright 101 | Licenses and Open Access | Using Copyrighted Materials | Copyright Term and the Public Domain | Copyright Resources | IP Rights & Privacy

Author Rights Resources | Small Claims / Copyright Claims Board / CASE Act | Film on Campus

Popular Resources

- [Fair Use Checklist \(Updated 2021\)](#)
- [Guidelines for Using Text, Images, Audio, and Video from Cornell University Library Collections](#)

Learn About Copyright

- [Copyright Law of the United States \(Title 17\)](#)
- [U.S. Copyright Office](#)
- [Copyright Crash Course @ University of Texas Libraries](#)

What is Intellectual Property (IP)?

General IP: Intellectual Property (IP) is the catch-all term for non-physical but identifiable property. Think of the text of a book: it can be presented in various languages, or different physical forms, but the underlying content remains the same. This underlying content is the intellectual property. In the USA, there are 3 main designations for IP protection that individuals can utilize, these are: Copyright, Trademark and Patent.

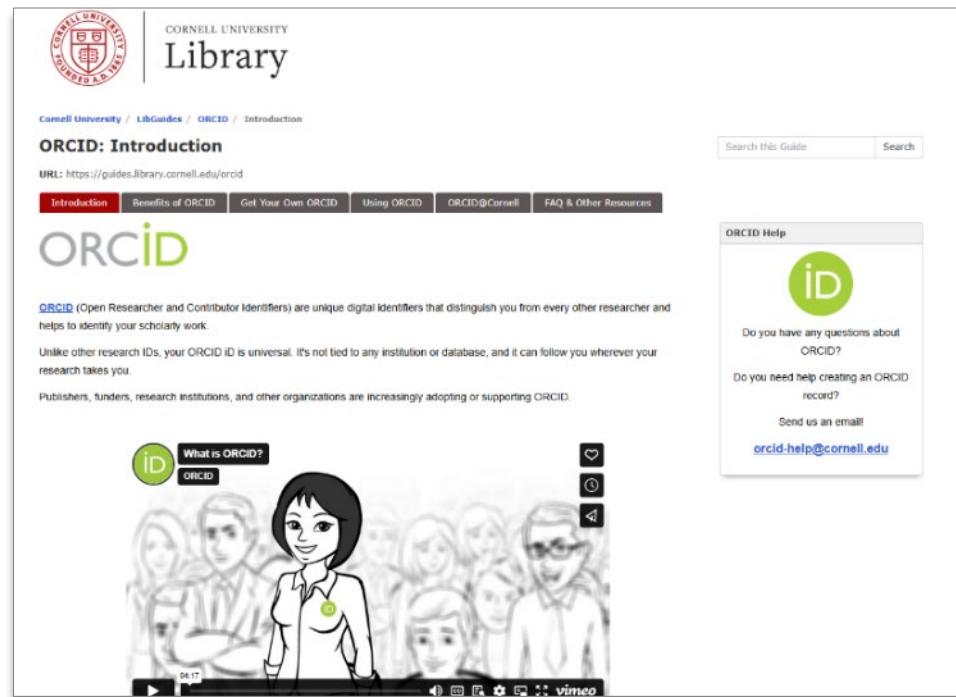
Trademarks v Patents v Copyright: Trademarks, Patents & Copyrights all cover different types of IP, and have different goals.

- Trademarks** can be logos or phrases, and are generally intended to help an individual tell one brand or corporate entity from another. They must be federally registered and renewed, and expire if they are not renewed. For example, the Cornell seal is trademarked by Cornell, and cannot be used without Cornell's permission.
- Patents** can be for mechanical processes or formulas, and are generally intended to limit production to the patent holder. They must be federally registered and renewed, and expire after time. For example, a medication can be patented, and can then only be produced with permission of the patent holder.
- Copyright** can be on any creative expression in a fixed medium. They do not need to be federally registered or renewed. For example, the copyright to this website belongs to Cornell (but we are licensing it under a CC-BY-ShareAlike.)

For an all-inclusive example, let's imagine a bottle of Coca-Cola. The Coca-Cola Corp owns the trademark to the name Coca-Cola, as well as the trademark on the bottle shape, and the graphic representation of their name. These are all things that help distinguish them from other cola brands and define their individual product. Coca-Cola also owns the patent on their formula. This means that no other

- 過去には著作権サービスチームが構成されており
図書館の著作権に関すること全般を担当
現在は教員の出版契約に重点を置く
- 教員は著作権やフェアユースに関する意識があまりない
出版契約に署名する前に質問してもらう、何に注意すべきか
知ってもらう（署名した後になつて権利を失うことになると気づく）
- 博士課程の学生のためのプログラムでも著作権を担当
出版前に注意すべき点について具体的に説明

- 2012年～ORCIDのOrganizational Memberに所属
(創設時からの初期メンバー)
- 新入学の大学院生は2学期の終わりまでにORCID IDの取得義務あり
→ Affiliation Managerを利用して新入生のリストをORCIDに
一括インポート・管理



The screenshot shows a web page titled "ORCID: Introduction" from Cornell University Library. The page features the Cornell University logo and the word "Library". The main content area is titled "ORCID" and includes a brief description of what ORCID is and how it works. A video player at the bottom shows a cartoon character with an ORCID logo. On the right side, there is a "ORCID Help" section with a green "ID" icon and contact information: "orcid-help@cornell.edu". The URL of the page is <https://guides.library.cornell.edu/orcid>.

"ORCID: Introduction". Cornell University Library.
<https://guides.library.cornell.edu/orcid> (最終閲覧日：2025/5/26)

- 図書館で ORCID ID を管理しているのは院生のみ
研究者はトラッキングしていない
- ORCID ID を使って研究者の出版物を追跡することは行っていない
学部ごとに独自ルール・ポリシーを所持
=大学全体で統一的な方法を使用することができない
- eCommons (機関リポジトリ、 DSpace) とは未連携
DSpace をバージョンアップすると著者 ID とORCID ID を
統合可能になる

- OSTP の研究成果公開方針において、パブリックアクセスを実現する方法とその資金については明確にされていない
- 資金調達における大学の役割と責任について、コーネル大学内でもはっきりとした合意は得られていない
- APC は高額すぎるので、Gold OA だけで即時OA義務に対応するのは非現実的
- 様々な分野の論文を、同じ大学の著者が書いたから全て同じリポジトリに保存するというのは意味がない
- 機関リポジトリでアーカイブするならば、それらを見つけられるようにする仕組みが必要



OKAYAMA
UNIVERSITY



～ご清聴ありがとうございました～