

シンポジウム・ワークショップ「大学における研究 データサービス」

Imker, Heidi
University of Illinois at Urbana-Champaign

Mischo, William H.
University of Illinois at Urbana-Champaign

Schlembach, Mary C.
University of Illinois at Urbana-Champaign

Miyazaki, Yuta
Kyushu University Library

他

<https://hdl.handle.net/2324/2547228>

出版情報 : 2019-12-05. Department of Library Science, Graduate School of Integrated Frontier Sciences, Kyushu University

バージョン :

権利関係 :

Workshop Outline

- Three sessions:
 - Session 1: What is Research Data?
 - Session 2: Data Policies in Academic Journals
 - Session 3: Curating Research Data
- Each session contains:
 - Introduction to the topic
 - Examples from the literature
 - An activity using example articles, policies, or data
 - Discussion questions

ワークショップのアウトライン

- 3つのセッション:
 - Session 1: 研究データとは何か?
 - Session 2: 学術ジャーナルにおけるデータポリシー
 - Session 3: 研究データのキュレーション
- 各セッションに含まれるもの:
 - トピックの紹介
 - 文献からの例示
 - サンプルの論文、ポリシー、データを使用した実習
 - 議論の課題

Session 1: What is Research Data?

Heidi Imker, PhD
Director, Research Data Service

Bill Mischo
Head, Grainger Engineering Library and Information Center

Mary Schlembach
Chemistry & Physical Sciences Librarian

2019年12月6日

九州大学統合新領域学府ライブラリーサイエンス専攻・附属図書館共同開催 (代表者 石田栄美)

ワークショップ「大学における研究データサービス」

※このスライドの日本語訳は、ワークショップの配布資料として作成したものです。

専門家による訳ではないため、誤りがある場合があります。

日本語訳：九州大学附属図書館eリソース課 泉 愛

Session 1: 研究データとは何か?

Heidi Imker, PhD

Director, Research Data Service

Bill Mischo

Head, Grainger Engineering Library and Information Center

Mary Schlembach

Chemistry & Physical Sciences Librarian

Session 1 Outline

1. Definitions of research data
2. Examples of research data
3. Activity: identifying research data in academic journal articles
4. Discussion questions

Session 1 アウトライン

1. 研究データの定義
2. 研究データの例
3. 実習: 学術雑誌論文における研究データの特定
4. 議論の課題

Session 1 Outline

1. Definitions of research data
2. Examples of research data
3. Activity: identifying research data in academic journal articles
4. Discussion questions

Session 1 アウトライン

1. 研究データの定義
2. 研究データの例
3. 実習: 学術雑誌論文における研究データの特定
4. 議論の課題

Definitions of Research Data

Australian National Data Service Disclaimer:

Providing an authoritative definition of research data is challenging, as any definition is likely to depend on the context in which the question is asked.

https://www.ands.org.au/data/assets/pdf_file/0006/731823/Whatis-research-data.pdf

研究データの定義

Australian National Data Serviceの免責事項:

質問された文脈に依存する場合があるため、研究データの正式な定義を示すことは困難である

https://www.ands.org.au/data/assets/pdf_file/0006/731823/Whatis-research-data.pdf

Definitions of Research Data

Nature Definition:

“Research data comprises research observations or findings, such as facts, images, measurements, records and files in various formats, and can be stored in databases.”

<https://www.nature.com/subjects/research-data>

研究データの定義

*Nature*の定義:

“研究データは、事実、画像、測定値、記録、さまざまな形式のファイルなど、研究観測または結果から構成され、データベースに保存できる。”

<https://www.nature.com/subjects/research-data>

Definitions of Research Data

US Government Definition:

“recorded factual material commonly accepted in the scientific community as necessary to validate research findings including data sets used to support scholarly publications”

<https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A110/2cfr215-o.pdf>

研究データの定義

アメリカ政府の定義:

“研究結果を必要に応じて検証するために、科学コミュニティで一般に受け入れられている、記録された事実資料。学術出版物を裏付けるために使用されるデータセットを含む。”

<https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A110/2cfr215-0.pdf>

Definitions of Research Data

CASRAI Definition:

“Data that are used as primary sources to support technical or scientific enquiry, research, scholarship, or artistic activity, and that are used as evidence in the research process and/or are commonly accepted in the research community as necessary to validate research findings and results. All other digital and non-digital content have the potential of becoming research data. Research data may be experimental data, observational data, operational data, third party data, public sector data, monitoring data, processed data, or repurposed data.”

https://dictionary.casrai.org/Research_data

研究データの定義

CASRAIの定義:

“技術的または科学的な調査、研究、学術、または芸術活動をサポートするための主要なソースとして使用されるデータ。および研究プロセスで根拠として使用されるデータ、もしくはは研究結果や結論を必要に応じて検証するための、研究コミュニティで一般的に受け入れられているデータ。すべてのデジタルおよび非デジタルコンテンツは、研究データになる可能性がある。研究データは、実験データ、観測データ、運用データ、第三者データ、公的機関のデータ、モニタリングデータ、処理データ、または別の目的に利用されたデータがある”

https://dictionary.casrai.org/Research_data

Session 1 Outline

1. Definitions of research data
2. Examples of research data
3. Activity: identifying research data in academic journal articles
4. Discussion questions

Session 1 アウトライン

1. 研究データの定義
2. 研究データの例
3. 実習: 学術雑誌論文における研究データの特定
4. 議論の課題

Example #1 – Library and Information Science

Wiley C.A. (2017) Assessing Research Data Deposits and Usage Statistics within IDEALS. *Journal of eScience Librarianship* 6(2): e1112. <https://doi.org/10.7191/jeslib.2017.1112>.

Author's research question:

What is the file composition of datasets ingested into the University of Illinois at Urbana-Champaign campus repository? Are datasets more likely to be single-file or multiple-file items?

Example #1 – 図書館情報学

Wiley C.A. (2017) Assessing Research Data Deposits and Usage Statistics within IDEALS. *Journal of eScience Librarianship* 6(2): e1112. <https://doi.org/10.7191/jeslib.2017.1112>.

著者のリサーチクエスチョン:

イリノイ大学アーバナシャンペーン校のリポジトリに採録されたデータセットのファイル構成は何か？

データセットは単一ファイルまたは複数ファイルのどちらのアイテムである可能性が高いか？

Example #1 – Library and Information Science

Method:

To answer this question, the author located deposits in Illinois' institutional repository, IDEALS, and gathered metadata and statistics for each deposit characterized as “data.”

Example #1 – 図書館情報学

方法:

この問いに答えるために、著者はイリノイ大学リポジトリ (IDEALS) のデポジットを検索し、“データ”として登録されたメタデータと統計を収集した。

Example #1 – Library and Information Science

The “raw” or “underlying” data is a CSV file.

IDEALS ID	Date Available in IDEALS	Date Available in IDEALS Rounded	Month Downloads Recorded	Time available (months)	Downloads/ Month	Discipline	Title	Dataset Composition
2142/45709	9/15/13	9/1/13	9/1/16	36	68.8	English	List of stop v	Single Text
2142/88828	1/15/16	1/1/16	9/1/16	8	63.9	LIS	Telecommur	Single PDF
2142/8775	6/30/08	6/1/08	9/1/16	99	57.9	Agriculture a	Hard times f	Single HTML
2142/3495	1/11/08	1/1/08	9/1/16	104	51.9	University Lil	Identifying th	Single PDF
2142/3483	12/14/07	12/1/07	9/1/16	105	82.7	University Lil	What Can I E	Multiple Mixe
2142/13183	7/27/09	7/1/09	9/1/16	86	28.2	Mechanical S	Modeling an	Single PDF
2142/78980	8/3/15	8/1/15	9/1/16	13	23.9	University Lil	ARL Academ	Single Excel
2142/54952	9/24/14	9/1/14	9/1/16	24	23.8	Special Colle	Michigan To	Single Excel
2142/88863	2/3/16	2/1/16	9/1/16	7	19.1	Prarie Resea	Urbana Irriga	Single PDF
2142/50053	8/7/14	8/1/14	9/1/16	25	18.4	Chemistry	Bibliography	Single PDF
2142/55672	11/5/14	11/1/14	9/1/16	22	17.3	Computer Sc	Single Core E	Single PDF



plus more columns...



plus more rows...

Example #1 – 図書館情報学

“raw” または “underlying” データはCSVファイル

IDEALS ID	Date Available in IDEALS	Date Available in IDEALS Rounded	Month Downloads Recorded	Time available (months)	Downloads/ Month	Discipline	Title	Dataset Composition
2142/45709	9/15/13	9/1/13	9/1/16	36	68.8	English	List of stop v	Single Text
2142/88828	1/15/16	1/1/16	9/1/16	8	63.9	LIS	Telecommur	Single PDF
2142/8775	6/30/08	6/1/08	9/1/16	99	57.9	Agriculture a	Hard times f	Single HTML
2142/3495	1/11/08	1/1/08	9/1/16	104	51.9	University Lil	Identifying th	Single PDF
2142/3483	12/14/07	12/1/07	9/1/16	105	82.7	University Lil	What Can I E	Multiple Mixe
2142/13183	7/27/09	7/1/09	9/1/16	86	28.2	Mechanical S	Modeling an	Single PDF
2142/78980	8/3/15	8/1/15	9/1/16	13	23.9	University Lil	ARL Academ	Single Excel
2142/54952	9/24/14	9/1/14	9/1/16	24	23.8	Special Colle	Michigan To	Single Excel
2142/88863	2/3/16	2/1/16	9/1/16	7	19.1	Prarie Resea	Urbana Irriga	Single PDF
2142/50053	8/7/14	8/1/14	9/1/16	25	18.4	Chemistry	Bibliography	Single PDF
2142/55672	11/5/14	11/1/14	9/1/16	22	17.3	Computer Sc	Single Core E	Single PDF



列はさらに続く...



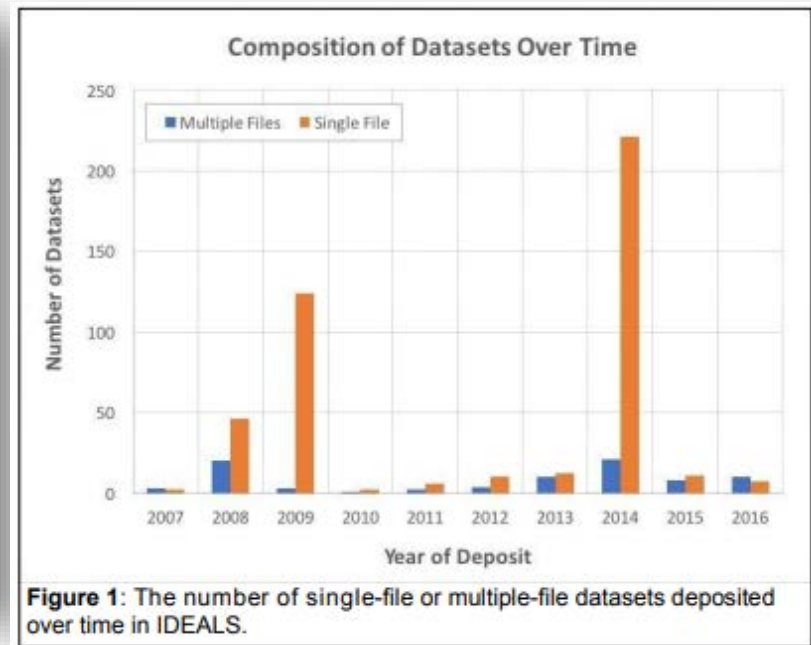
行はさらに続く...

Example #1 – Library and Information Science

In the article, the data are shown in graphs and tables.

Table 1: Composition of IDEALS Datasets

Multiple-File Datasets	Count	Single-File Datasets	Count
Multiple Mixed	73	Single Excel	208
Multiple CSVs	3	Single PDF	186
Multiple PDFs	3	Single QuickTime Video	11
Multiple TIFFs	1	Single Zip	10
Multiple XMLs	1	Single AVI Video	9
Multiple Zips	1	Single AIFF Audio	4
		Single CSV	3
		Single HTML	2
		Single JPEG	2
		Single Unknown	2
		Single Text	2
		Single PowerPoint	1
Total	82	Total	440

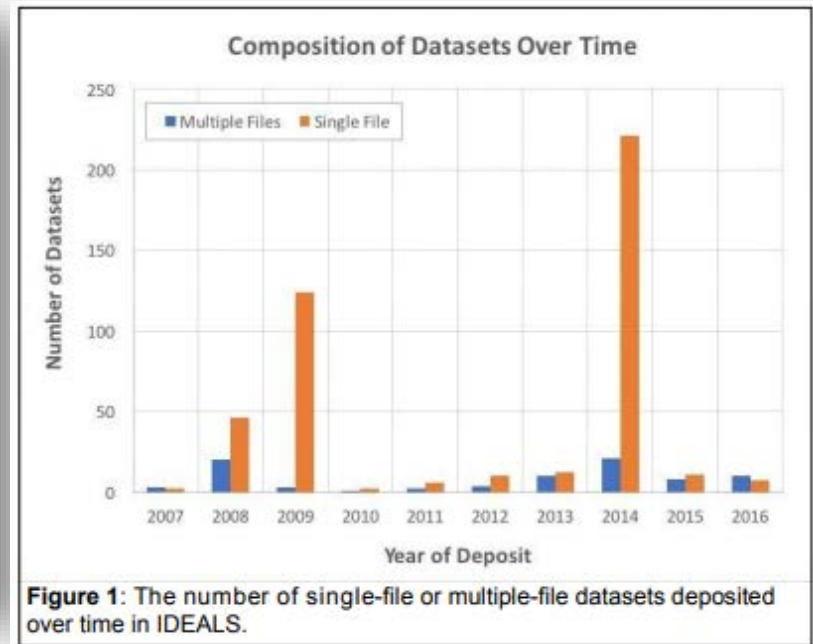


Example #1 – 図書館情報学

論文では、データはグラフや表で示されている

Table 1: Composition of IDEALS Datasets

Multiple-File Datasets	Count	Single-File Datasets	Count
Multiple Mixed	73	Single Excel	208
Multiple CSVs	3	Single PDF	186
Multiple PDFs	3	Single QuickTime Video	11
Multiple TIFFs	1	Single Zip	10
Multiple XMLs	1	Single AVI Video	9
Multiple Zips	1	Single AIFF Audio	4
		Single CSV	3
		Single HTML	2
		Single JPEG	2
		Single Unknown	2
		Single Text	2
		Single PowerPoint	1
Total	82	Total	440



Example #2 – Life Science

Wachi, N. , Abe, Y. , Inomata, N. , Szmidt, A. E. and Tachida, H. (2012), Speciation history of three closely related oak gall wasps, *Andricus mukaigawae*, *A. kashiwaphilus*, and *A. pseudoflos* (Hymenoptera: Cynipidae) inferred from nuclear and mitochondrial DNA sequences. Mol Ecol, 21: 4681-4694. doi:[10.1111/j.1365-294X.2012.05727.x](https://doi.org/10.1111/j.1365-294X.2012.05727.x)

Authors' research question:

During evolution of different species of wasps, which characteristic came first: the preference for host plants or how sex is determined during wasp development?

Example #2 – 生命科学

Wachi, N. , Abe, Y. , Inomata, N. , Szmidt, A. E. and Tachida, H. (2012), Speciation history of three closely related oak gall wasps, *Andricus mukaigawae*, *A. kashiwaphilus*, and *A. pseudoflos* (Hymenoptera: Cynipidae) inferred from nuclear and mitochondrial DNA sequences. *Mol Ecol*, 21: 4681-4694. doi:[10.1111/j.1365-294X.2012.05727.x](https://doi.org/10.1111/j.1365-294X.2012.05727.x)

著者のリサーチクエスチョン:

スズメバチの異なる種の進化の間に、寄生植物の好み、発育期間に性別がどのように決定されるか等、どの特徴が最初に起こったか。

Example #2 – Life Science

Method:

To answer this question, the author gathered DNA from 3 different wasp species and compared the gene sequences.

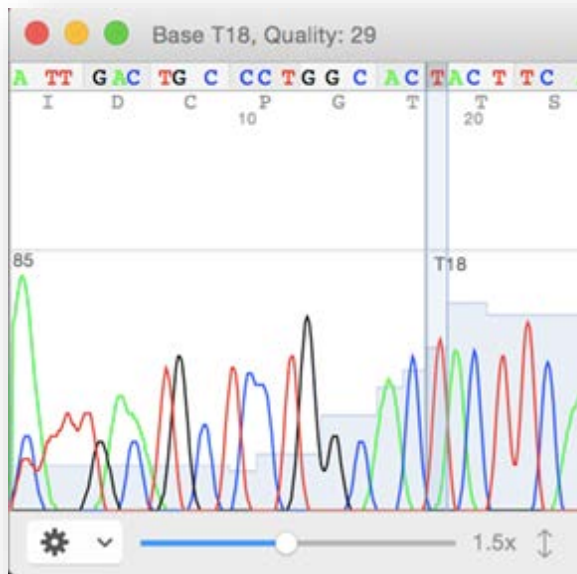
Example #2 – 生命科学

方法:

この問いに答えるために、著者は異なる3種類のスズメバチのDNAを収集し、遺伝子配列を比較した

Example #2 – Life Science

The “raw” or “underlying” data are the DNA sequences.



>F17_APMN-1

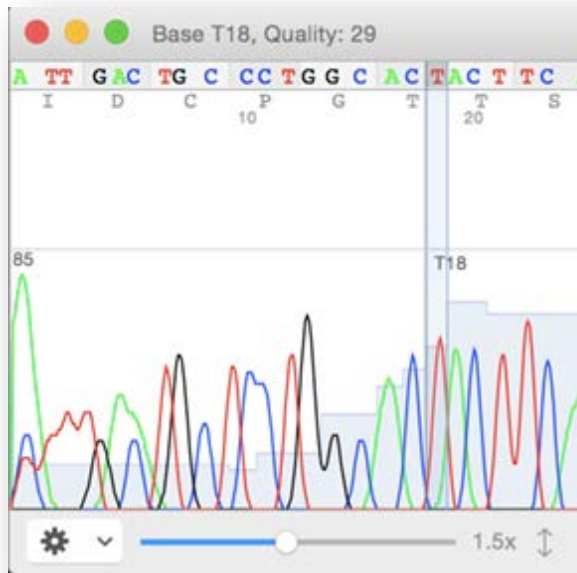
```
TTACTGTTTTTATTTCTTTTTTGGATCTTATTTTCATTCTGCTTTATCCCCAAATGTTGAAATTGGAAGATTATGACC
TCCTAAAAATATTAATCAATTTAATCCCTATGAAATTCCTTTATTAATACTATTATTTTATTATCTTCCGGGGTTACAA
TTACTTGATCACATCATTGTTTATTAATTGGTAATAAATTTAAAAAGATTAATTGGTTTATTTTAAACAATTTTATTAGGA
TTATTATTTTCTTAATCAATTATATGAATATAATCAATCATGTTTTTGTTAAGAGATTTCAGTTTATGGATCATTATT
TTTTATAACAACCTGGATTCATGGGTTACATGTAATTATTGGTAGATTATTTTATTAATTATATTATTCCGAATGTATT
TTAATCATACATCTTCATATCATCTTTGGGTTTGAAGCTTCATCTTGGTATTGACATTTTGTGATGTAATTTGAATT
ATTGATTTTACAATAATTTATTGATGAAATTATTTTATATAGTATATAAAA-TTACATTTTATTTCCAATAAAAAAA
ATTAATAAATTTTAAATAAAAAAATTTTATTTATTTTATAAATTTATTATTTTATAATTAGAAGAATTTTATTA
ATATTAATTTTAAAAAAGAAAAAATTTAATAAATCAAGGAATAAATTAACCTTCATTGAAATGTTGGGTTTGATAATAT
TTCATTATTACGATTACCATTTTCTGTTTATTTTATAAATTGGAAATTTTAAAAATTTTAAAAATTTTAAAAATTTCTT
TAATTTTCCAATTTAAATTTTAAATTTAATAAATTTTAAAAAATTTTAAAAAATTTTAAAAATTTTAAAAATTTTAA
ATTTTAGGCTTATATGAAATGAAACAAGTTTCAAAAATGATTATAAAAATAAATTTTGAATAATAGTTAAATTTTA
TAACATTTAAATTGCATTTAAAAATTTAATAAAAAATTTATTTTGAATAAAAACTAAAACCTTAGTATTTAATTTTCA
CTTAAATTTATAATT
```

>F17_APMN-2

```
TTACTGTTTTTATTTCTTTTTTGGATCTTATTTTCATTCTGCTTTATCCCCAAATGTTGAAATTGGAAGATTATGACC
TCCTAAAAATATTAATCAATTTAATCCCTATGAAATTCCTTTATTAATACTATTATTTTATTATCTTCCGGGGTTACAA
TTACTTGATCACATCATTGTTTATTAATTGGTAATAAATTTAAAAAGATTAATTGGTTTATTTTAAACAATTTTATTAGGA
TTATTATTTTCTTAATCAATTATATGAATATAATCAATCATGTTTTTGTTAAGAGATTTCAGTTTATGGATCATTATT
TTTTATAACAACCTGGATTCATGGGTTACATGTAATTATTGGTAGATTATTTTATTAATTATATTATTCCGAATGTATT
TTAATCATACATCTTCATATCATCTTTGGGTTTGAAGCTTCATCTTGGTATTGACATTTTGTGATGTAATTTGAATT
ATTGATTTTACAATAATTTATTGATGAAATTATTTTATATAGTATATAAAA-TTACATTTTATTTCCAATAAAAAAA
```


Example #2 – 生命科学

“raw” もしくは “underlying” データはDNA配列



>F17_APMN-1

```
TTACTGTTTTTATTTCTTTTTTGGATCTTATTTTCATTCTGCTTTATCCCCAAATGTTGAAATTGGAAGATTATGACC
TCCTAAAAATATTAATCAATTTAATCCCTATGAAATTCCTTTATTAATACTATTATTTTATTATCTTCCGGGGTTACAA
TTACTTGATCACATCATTGTTTATTAATTGGTAATAAATTTAAAAAGATTAATTGGTTTATTTTAAACAATTTTATTAGGA
TTATTATTTTCTTAATCAATTATATGAATATAATCAATCATGTTTTTGTTAAGAGATTTCAGTTTATGGATCATTATT
TTTTATAACAACCTGGATTTTCATGGGTTACATGTAATTATTGGTAGATTATTTTATTAATTATATTTATTCCGAATGTATT
TTAATCATACATCTTCATATCATCTTTGGGTTTGAAGCTTCATCTTGGTATTGACATTTTGTGATGTAATTTGAATT
ATTGATTTTACAATAATTTATTGATGAAATTATATTTTATATAGTATATAAAA-TTACATTTTATTTCCAATAAAAAAA
ATTAATAAATTTTAAATAAAAAAATTTTATTTATTTATAAATTTATTATTTTATAAATTAGAAGAATTTTATTA
ATATTAATTTTAAAGAAAAAATTTAATAAATCGGAATAAATTAACCTTCATTGAAATGTTGTTGATAATAT
TTCATTATTACGATTACCATTTTCTGTTTATTTTATAAATTGGAAATTTTAAATTTTATTTTATTTGATGTAGAAATTTCTT
TAATTTTCCAATTTAAATTTTAAATTTAATAAATTTTATAAAAAAATTTTAAATTTTATTTATTTAATTTTAA
ATTTTAGGCTTATATGAAATGAAACAAGTTTCAATAAATGATTATAAATAAATTTTGAATAATAGTTAAATTTTA
TAACATTTAAATTGCATTTAAAAATTTAATAAAAAATTTATTTTGAATAAAACTAAAACCTTAGTATTTAATTTTCA
CTTAAATTTATAATT
```

>F17_APMN-2

```
TTACTGTTTTTATTTCTTTTTTGGATCTTATTTTCATTCTGCTTTATCCCCAAATGTTGAAATTGGAAGATTATGACC
TCCTAAAAATATTAATCAATTTAATCCCTATGAAATTCCTTTATTAATACTATTATTTTATTATCTTCCGGGGTTACAA
TTACTTGATCACATCATTGTTTATTAATTGGTAATAAATTTAAAAAGATTAATTGGTTTATTTTAAACAATTTTATTAGGA
TTATTATTTTCTTAATCAATTATATGAATATAATCAATCATGTTTTTGTTAAGAGATTTCAGTTTATGGATCATTATT
TTTTATAACAACCTGGATTTTCATGGGTTACATGTAATTATTGGTAGATTATTTTATTAATTATATTTATTCCGAATGTATT
TTAATCATACATCTTCATATCATCTTTGGGTTTGAAGCTTCATCTTGGTATTGACATTTTGTGATGTAATTTGAATT
ATTGATTTTACAATAATTTATTGATGAAATTATATTTTATATAGTATATAAAA-TTACATTTTATTTCCAATAAAAAAA
```

Example #2 – Life Science

In the article, the data are shown in genetic trees and tables.

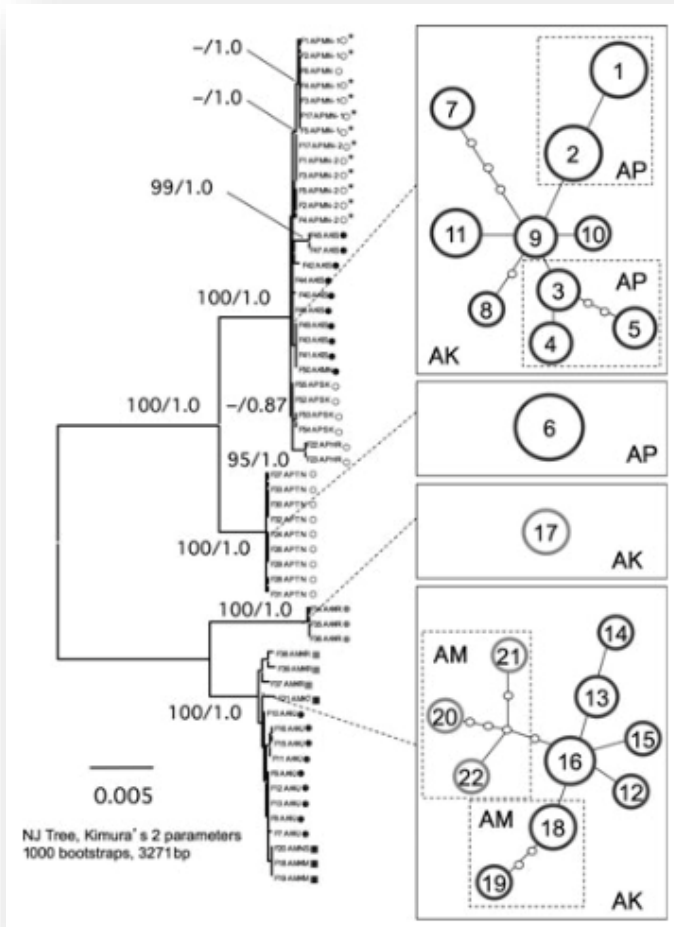


Table 2 Estimates of nucleotide variation

Species	Locus	n	S	K	π	θ_W	
<i>Andricus mukaigawae</i>	mtDNA	7	12	4.286	0.00131	0.0015	
	nDNA	<i>Ant_SesB</i>	14	5	2.297	0.00134	0.00092
		<i>RpL37</i>	14	4	0.956	0.00298	0.00392
		<i>RpS18</i>	14	3	1.56	0.00147	0.00089
		<i>Locus 2</i>	14	0	0	0	0
		<i>Locus 4</i>	14	22	5.275	0.00365	0.00479
		<i>COXlike</i>	14	2	0.67	0.00042	0.00039
		<i>Locus 5</i>	14	19	6.538	0.00401	0.00366
		<i>RpL27A</i>	14	23	5.527	0.0283	0.0037
		<i>LWRh</i>	14	0	0	0	0
Total/mean		78	2.536	0.00469	0.00203		
<i>Andricus kashiwaphilus</i>	mtDNA	22	140	64.069	0.01959	0.01174	
	nDNA	<i>Ant_SesB</i>	44	6	0.951	0.00055	0.0008
		<i>RpL37</i>	44	1	0.0089	0.00028	0.00072
		<i>RpS18</i>	44	13	5.834	0.00548	0.00281
		<i>Locus 2</i>	44	4	0.525	0.00038	0.00066
		<i>Locus 4</i>	44	2	0.091	0.00006	0.00032
		<i>COXlike</i>	44	5	0.542	0.00034	0.00072
		<i>Locus 5</i>	44	3	0.266	0.00016	0.00042
		<i>RpL27A</i>	44	3	0.416	0.00021	0.00035
		<i>LWRh</i>	44	0	0	0	0
Total/mean		37	0.959	0.000829	0.000756		
<i>Andricus pseudoflos</i>	mtDNA	28	37	15.376	0.0047	0.00291	
	nDNA	<i>Ant_SesB</i>	44	2	0.584	0.00034	0.00027
		<i>RpL37</i>	44	0	0	0	0
		<i>RpS18</i>	44	12	5.666	0.00533	0.00259
		<i>Locus 2</i>	44	1	0.089	0.00006	0.00016
		<i>Locus 4</i>	44	0	0	0	0
		<i>COXlike</i>	44	0	0	0	0
		<i>Locus 5</i>	44	0	0	0	0
		<i>RpL27A</i>	44	1	0.495	0.00025	0.00012
		<i>LWRh</i>	44	0	0	0	0
Total/mean		16	0.759	0.000664	0.000349		

Example #2 – 生命科学

この論文では、データは遺伝系譜図と表で示される

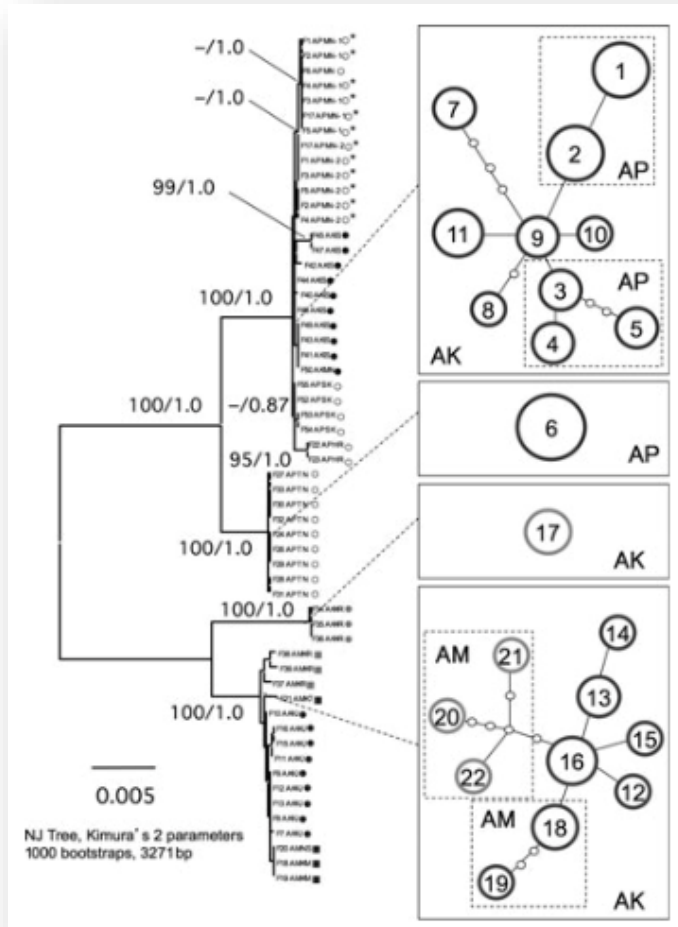


Table 2 Estimates of nucleotide variation

Species	Locus	n	S	K	π	θ_W	
<i>Andricus mukaigawae</i>	mtDNA	7	12	4.286	0.00131	0.0015	
	nDNA	<i>Ant_SesB</i>	14	5	2.297	0.00134	0.00092
		<i>RpL37</i>	14	4	0.956	0.00298	0.00392
		<i>RpS18</i>	14	3	1.56	0.00147	0.00089
		<i>Locus 2</i>	14	0	0	0	0
		<i>Locus 4</i>	14	22	5.275	0.00365	0.00479
		<i>COXlike</i>	14	2	0.67	0.00042	0.00039
		<i>Locus 5</i>	14	19	6.538	0.00401	0.00366
		<i>RpL27A</i>	14	23	5.527	0.0283	0.0037
		<i>LWRh</i>	14	0	0	0	0
Total/mean		78	2.536	0.00469	0.00203		
<i>Andricus kashiwaphilus</i>	mtDNA	22	140	64.069	0.01959	0.01174	
	nDNA	<i>Ant_SesB</i>	44	6	0.951	0.00055	0.0008
		<i>RpL37</i>	44	1	0.0089	0.00028	0.00072
		<i>RpS18</i>	44	13	5.834	0.00548	0.00281
		<i>Locus 2</i>	44	4	0.525	0.00038	0.00066
		<i>Locus 4</i>	44	2	0.091	0.00006	0.00032
		<i>COXlike</i>	44	5	0.542	0.00034	0.00072
		<i>Locus 5</i>	44	3	0.266	0.00016	0.00042
		<i>RpL27A</i>	44	3	0.416	0.00021	0.00035
		<i>LWRh</i>	44	0	0	0	0
Total/mean		37	0.959	0.000829	0.000756		
<i>Andricus pseudoflos</i>	mtDNA	28	37	15.376	0.0047	0.00291	
	nDNA	<i>Ant_SesB</i>	44	2	0.584	0.00034	0.00027
		<i>RpL37</i>	44	0	0	0	0
		<i>RpS18</i>	44	12	5.666	0.00533	0.00259
		<i>Locus 2</i>	44	1	0.089	0.00006	0.00016
		<i>Locus 4</i>	44	0	0	0	0
		<i>COXlike</i>	44	0	0	0	0
		<i>Locus 5</i>	44	0	0	0	0
		<i>RpL27A</i>	44	1	0.495	0.00025	0.00012
		<i>LWRh</i>	44	0	0	0	0
Total/mean		16	0.759	0.000664	0.000349		

Example #3 – Materials Science

DeCost, B.L. and Holm, E.A. (2017) Characterizing powder materials using keypoint-based computer vision methods. *Computational Materials Science*, 126: 438-445.

<https://doi.org/10.1016/j.commatsci.2016.08.038>

Authors' research question:

How can images of materials of irregular size and with textured surfaces be quantitatively characterized?

Example #3 – 物質科学

DeCost, B.L. and Holm, E.A. (2017) Characterizing powder materials using keypoint-based computer vision methods. *Computational Materials Science*, 126: 438-445.

<https://doi.org/10.1016/j.commatsci.2016.08.038>

著者のリサーチクエスチョン:

不規則な大きさの材料の画像とテクスチャ面は、どのように定量的に特徴付けられるか？

Example #3 – Materials Science

Method:

To answer this question, the authors used a computer program to synthetically generate a set of realistic scanning electron micrographs (SEM) for 8 different materials with different powdery textures.

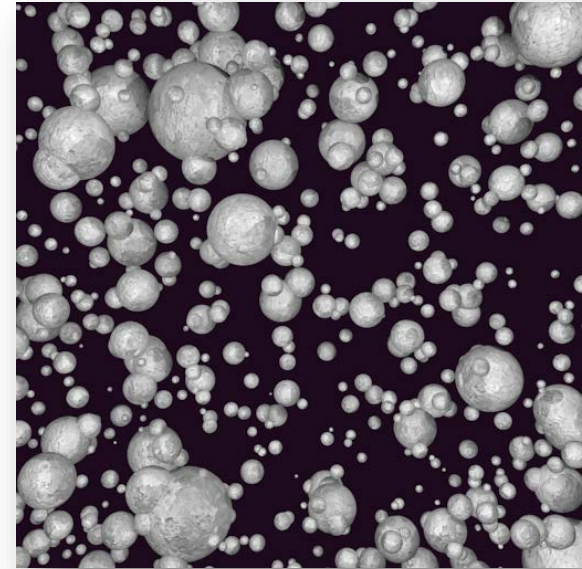
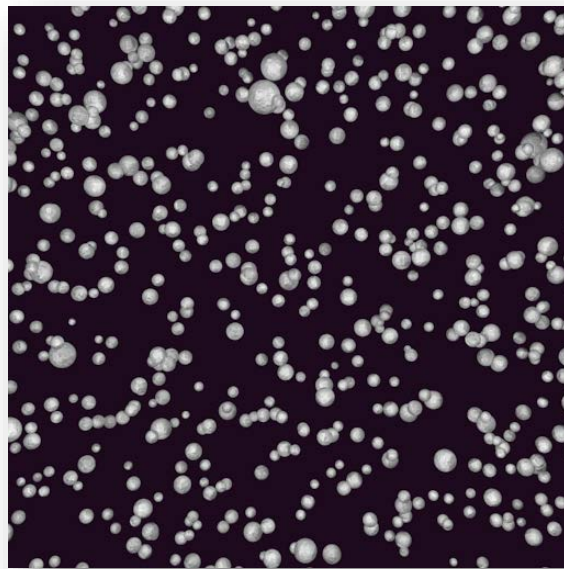
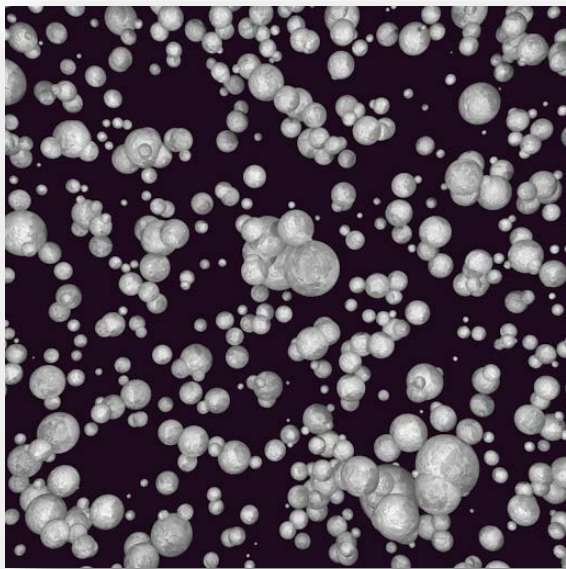
Example #3 – 物質科学

方法:

この問いに答えるために、著者はコンピュータープログラムを使用して、さまざまな粉末状のテクスチャーと8種類の異なる材料のリアリティのある走査型電子顕微鏡写真（SEM）を生成した

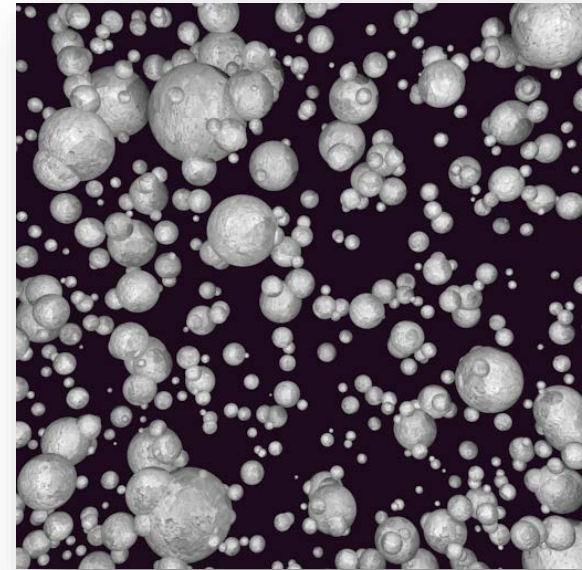
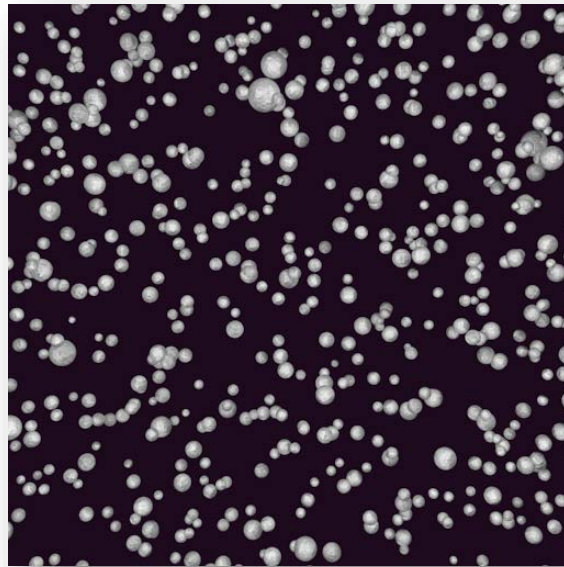
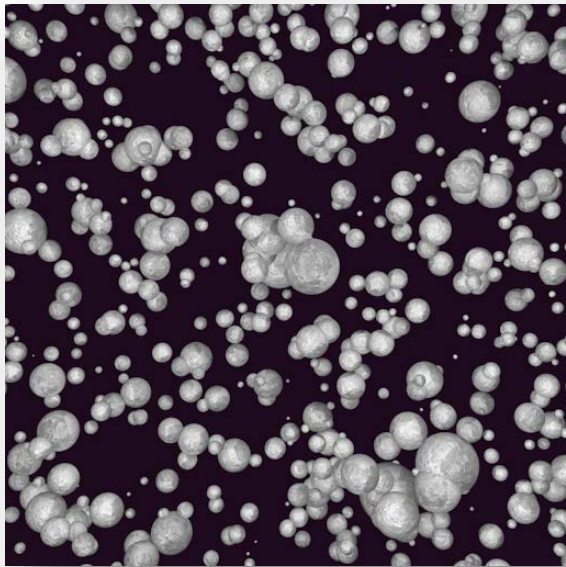
Example #3 – Materials Science

The “raw” or “underlying” data are 1000s of synthetic SEM images.



Example #3 – 物質科学

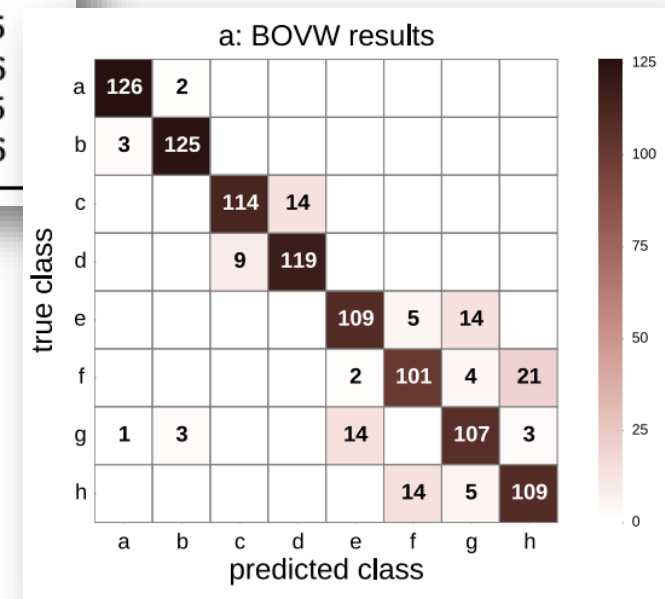
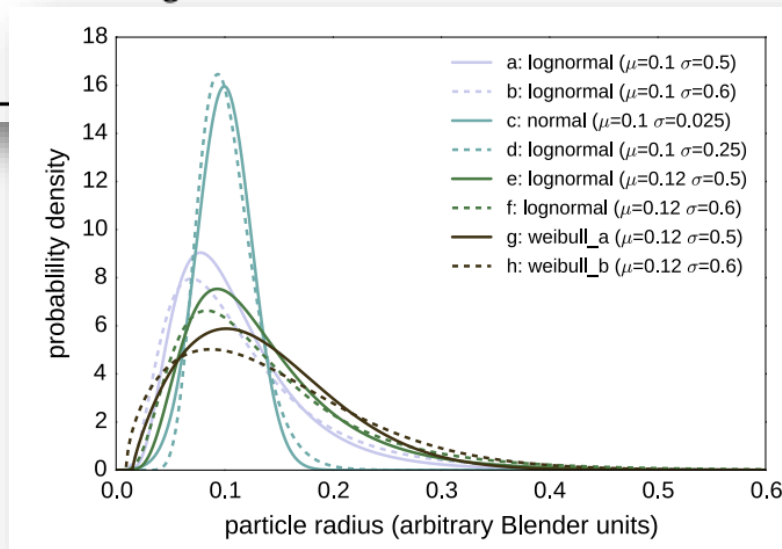
“raw” もしくは“underlying” データは、数千の合成SEM画像



Example #3 – Materials Science

In the article, the data are shown in graphs, tables, and other figures.

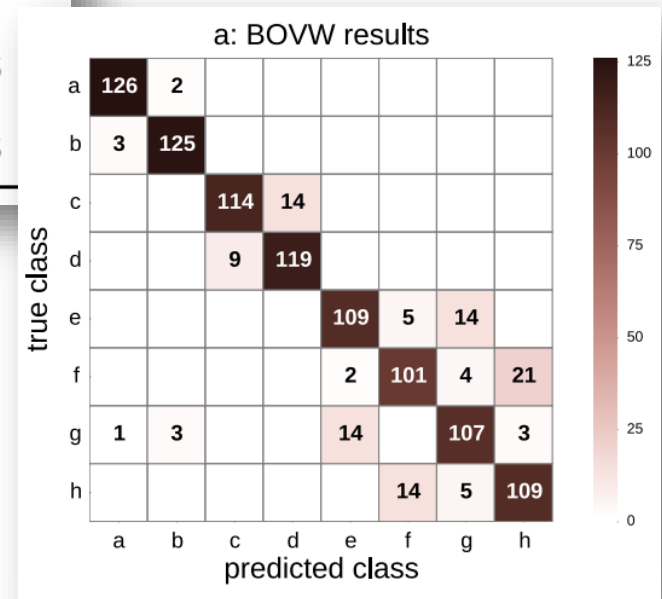
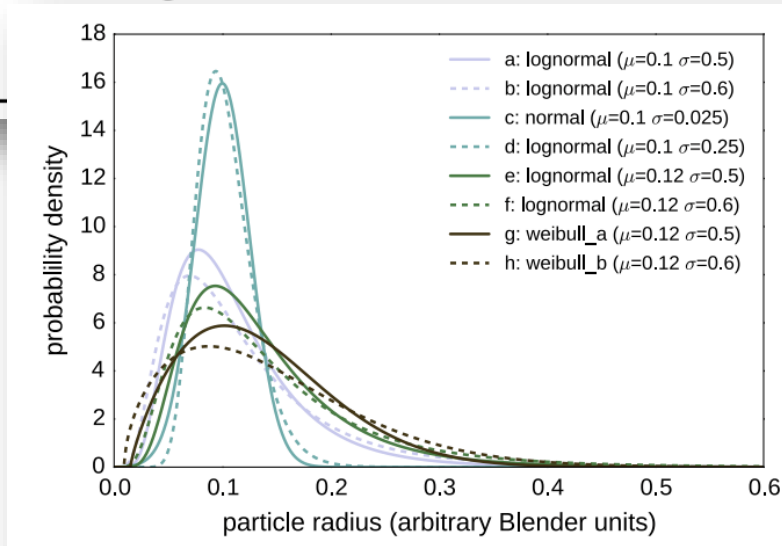
Class	Family	Mean (μ)	Shape (σ)
<i>a</i>	Lognormal	0.1	0.5
<i>b</i>	Lognormal	0.1	0.6
<i>c</i>	Normal	0.1	0.025
<i>d</i>	Lognormal	0.1	0.25
<i>e</i>	Lognormal	0.12	0.5
<i>f</i>	Lognormal	0.12	0.6
<i>g</i>			0.5
<i>h</i>			0.6



Example #3 – 物質科学

この論文では、データはグラフ、表、その他の図に示されている

Class	Family	Mean (μ)	Shape (σ)
<i>a</i>	Lognormal	0.1	0.5
<i>b</i>	Lognormal	0.1	0.6
<i>c</i>	Normal	0.1	0.025
<i>d</i>	Lognormal	0.1	0.25
<i>e</i>	Lognormal	0.12	0.5
<i>f</i>	Lognormal	0.12	0.6
<i>g</i>	weibull_a	0.12	0.5
<i>h</i>	weibull_b	0.12	0.6



Session 1 Outline

1. Definitions of research data
2. Examples of research data
3. **Activity: identifying research data in academic journal articles**
4. Discussion questions

Session 1 アウトライン

1. 研究データの定義
2. 研究データの例
3. 実習: 学術雑誌論文における研究データの特定
4. 議論の課題

Session 1 Activity – Identifying Data

Choose one of the following for Activity 1:

1. Library and Information Science

Ikeuchi, Y. and Ishimura, Y. (2016). Data Sharing Policy by Academic Journals: Inter-field Comparison and Feature Analysis. *Journal of Japan Society of Library and Information Science*, 62(1), 20-37. https://doi.org/10.20651/jslis.62.1_20

2. Psychology

Wakabayashi, A., Tojo, Y., Baron-Cohen, S. and Wheelwright, S. (2005) The Autism-Spectrum Quotient (AQ) Japanese version. *The Japanese Journal of Psychology*, 75(1), 78-84. <https://doi.org/10.4992/jjpsy.75.78>

3. Your Choice!

An article of your own choosing.

Take 15 minutes to familiarize yourself with the article by reading the abstract and methods. Then review the tables and figures to evaluate what the “underlying” data might be.

Session 1 実習 – データの識別

下記の中から1つ課題を選択してください:

1. 図書館情報学

Ikeuchi, Y. and Ishimura, Y. (2016). Data Sharing Policy by Academic Journals: Inter-field Comparison and Feature Analysis. *Journal of Japan Society of Library and Information Science*, 62(1), 20-37. https://doi.org/10.20651/jslis.62.1_20

2. 心理学

Wakabayashi, A., Tojo, Y., Baron-Cohen, S. and Wheelwright, S. (2005) The Autism-Spectrum Quotient (AQ) Japanese version. *The Japanese Journal of Psychology*, 75(1), 78-84. <https://doi.org/10.4992/jjpsy.75.78>

3. お好みで!

自身で選択した論文

15分間、論文のアブストラクトと方法を読み、論文に慣れましょう。
表と図を確認し“underlying”データは何か判断しましょう。

Discussion Questions

1. What questions did you have as you thought about what “underlying” data might be?

For example, Table 1 in Article 1 shows Repositories in DataBib by field, but you might wonder what the specific repositories are within each listed field.

2. For any of the tables or figures, can you think of a way the data may be reused in a future study?

For example, Table 1 in Article 1 shows Repositories in DataBib by field as of 2014, but you might wonder how this compares if the same analysis was done in 2019.

議論の課題

1. “underlying” データが何であるか考えたことを通し、何か感じたことはありますか？

例えば、論文「1」のTable 1はDataBibにおけるリポジトリ数を分野別に示しているが、それぞれの分野において特定のリポジトリが何であるかを疑問に思うかもしれません。

2. 表や図が、データとして今後の研究で再利用される方法を考えられますか？

例えば、論文「1」のTable 1はDataBibでは2014年のデータとして記録されていますが、もしも同じ分析が2019年に行われた場合どのように比較すべきか迷うのではないのでしょうか。