

**Open Data/Open Science
for Researchers, Stakeholders and Citizens:
From the viewpoint of researcher**

Haruki Nakamura

中村 春木

**Professor Emeritus, Institute for Protein Research,
The University of Osaka**

Editor-in-Chief of *Biophysics and Physicobiology*

(Official Journal of Biophysical Society of Japan:

一般社団法人日本生物物理学会)

Self-Introduction*

- 1999～2018: Professor at Inst. Protein Res., The Univ. of Osaka: Research in the field of Structural Biology and Biophysics (構造生命科学, 生物物理学)
- 2001～2017: Database management of Protein structures (Protein Data Bank: PDBj and wwPDB)
- 2012～2013: Editor-in-Chief of Japanese Journal, Seibutsu Butsuri: (生物物理誌) published by Biophysical Society of Japan (一般社団法人 日本生物物理学会) with 3,000 members founded in 1960)
- 2015～2017: President of Biophysical Society of Japan
- 2020～: Editor-in-Chief of English Journal, *Biophysics and Physicobiology*, which has been published since 2005 (Peer reviewed, open access, online journal)

*Hall, D., Basu, G. & Ito, N. Computational biophysics and structural biology of proteins -a Special Issue in honor of Prof. Haruki Nakamura's 70th birthday. Biophys Rev 14, 1211–1222 (2022). <https://doi.org/10.1007/s12551-022-01039-0>

自己紹介*

- 1999～2018: 大阪大学蛋白質研究所教授: 構造生命科学・生物物理学分野の研究
- 2001～2017: 蛋白質構造のデータベース管理 (Protein Data Bank: PDBj and wwPDB)
- 2012～2013: 日本生物物理学会が刊行する日本語雑誌『生物物理』誌編集委員長。同学会は1960年に設立され、3,000人の会員が在籍
- 2015～2017: 日本生物物理学会 会長
- 2020～: 2005年から刊行されている英文誌 "*Biophysics and Physicobiology*" (査読有り、オープンアクセス、オンラインジャーナル) 編集委員長

*Hall, D., Basu, G. & Ito, N. Computational biophysics and structural biology of proteins -a Special Issue in honor of Prof. Haruki Nakamura's 70th birthday. Biophys Rev 14, 1211–1222 (2022). <https://doi.org/10.1007/s12551-022-01039-0>

Open Data/Open Science

Benefits and Requirements

1) for Researchers

2) for Stakeholders

**(Funding Agency, University, Research Institute,
and Journal Publisher)**

3) for Citizens

オープンデータ/オープンサイエンス

利点と要件

- 1) 研究者に対して
- 2) ステークホルダーに対して
(助成機関、大学、研究機関、雑誌出版社)
- 3) 市民に対して

Recent Paradigm Shift in Science

First and Second Paradigm: Experiment and Theory



Third Paradigm (the 2000s) : Computational science,
Simulation **by Super-computers**



Fourth Paradigm (the 2010s) : Data Science,
Data-driven Science **by BigData**



Fifth Paradigm (the 2020s) : AI-driven Science
by Generative AI, LLM, etc.

科学における最近のパラダイムシフト

第1、第2のパラダイム: 実験と理論



第3のパラダイム (2000年代): 計算科学、スーパーコンピュータによるシミュレーション



第4のパラダイム (2010年代): ビッグデータによるデータ科学、データ主導科学



第5のパラダイム (2020年代): 生成AI、LLMなどによるAI主導科学

Open Data/Open Science: Benefits for Researchers

- **The advancement of data science and science in general is anticipated through Open Data/Open Science initiatives.**

オープンデータ/オープンサイエンス:研究者にとっての利点

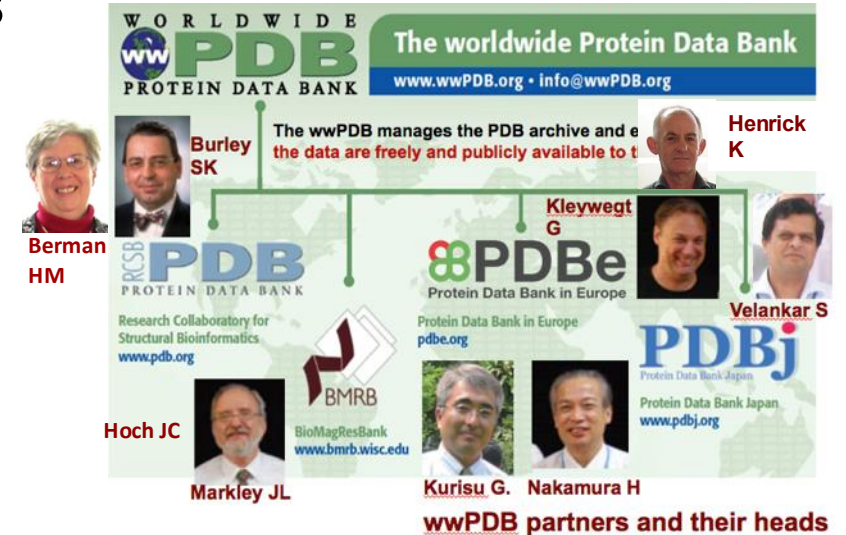
- オープンデータ/オープンサイエンスにより、
データ科学および科学一般の発展が期待される

Open Data/Open Science: Benefits for Researchers

Protein Data Bank: International Database of protein 3D structures

- All the structures are available under CC0 1.0 Universal Public Domain Dedication.
- The Data management is made by the research funds from the Governments of USA, EU, and Japan
- Publishers allow the authors to submit their manuscripts **only when their structures have been deposited to wwPDB with the validation report and the experimental data**

(Photos from wwPDB website)



Brief History of PDB

- Protein Crystallography started in 1960s
- PDB started with 7 structure data in 1971
- PDBj and wwPDB started in 2000s
- Now about 237,000 data were registered

<https://pdj.org/>

The screenshot shows the PDBj website interface. At the top, the 'PDBj' logo is displayed, along with the text 'Protein Data Bank Japan' and the website 'www.pdbj.org'. The interface includes a search bar, a navigation menu with links like 'Home', 'About PDBj', 'Data deposition', and 'Download', and a 'Latest news' section. The 'About PDBj' section provides information about the PDBj's role in maintaining the global PDB/BMRB/EMDB archives. The 'Latest news' section lists recent updates, including the release of 250 new PDB entries and the announcement of the PDBj Newsletter Vol. 25 release.

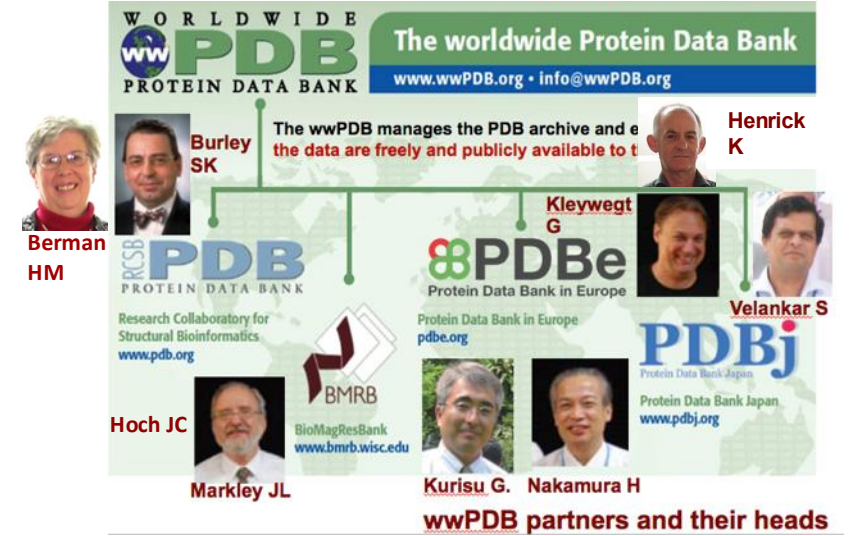
オープンデータ/オープンサイエンス:研究者にとっての利点

Protein Data Bank: タンパク質の3D構造の国際データベース

- すべての構造は、CC0 1.0 Universal Public Domain Dedicationの下で利用可能
- 米国、EU、日本グループがそれぞれの国と地域の研究資金によって協力して運営している
- 構造情報に加えてvalidation reportと実験データがwwPDBに登録された場合のみ、論文を発表することができる

PDBの簡単な歴史

- 1960年代に蛋白質の構造解析が開始
- PDBは、1971年に7つの構造データから開始
- PDBjとwwPDBは、2000年代に開始
- 現在、約237,000件のデータが登録されている



<https://pdbj.org/>

The screenshot shows the PDBj (Protein Data Bank Japan) website. The top header includes the PDBj logo, navigation links (Home, Top Page, Statistics, Help, FAQ, Contact Us, Cite Us / Terms and Conditions, Links, Settings), and a search bar. The main content area features:

- About PDBj**: A paragraph describing PDBj's role in maintaining the global PDB/BMRB/EMDB archives and its support by JST-NBDC (JPMND2205) and AMED-BINDS.
- Notice of temporary suspension of PDBj services**: A message stating that all PDBj services will be temporarily unavailable for about 15 minutes at most twice between 2:00 a.m. and 5:00 a.m. on Saturday, May 31 (JST) due to network maintenance.
- Latest news**: A list of recent updates, including the release of 250 new PDB entries on May 28, 2025, and the announcement of the PDBj Newsletter Vol. 25 release.

On the right side, there are links to 'Molecule of the Month', 'Article List', 'PDB-IHM', 'EM Navigator', and 'BMRBj'.

Science has been developed by Open Data/Open Science

The Nobel Prize in Chemistry 2024

To:
David Baker
for **protein design**

John Jumper and Demis Hassabis
for **“accurate” protein 3D structure prediction by AI** (AlphaFold 2)

➡ • **Most biologists get the benefits of those methods of designs and predictions**

• **Even structural biologists can determine the precise protein structures much more quickly with other information of functions and dynamics**



(Photo from wwPDB website)

**Comment by
Prof. Helen M. Berman:
One of the founders of
wwPDB**

“The huge protein database that spawned AlphaFold and biology’s AI revolution.”

“Other communities can, should and must do this. Otherwise, we’re not going to get the big break throughs.”

E. Callaway, *Nature* 634, 1028-1029 (2024)

科学はオープンデータ/オープンサイエンスによって発展してきた

2024年ノーベル化学賞

David Baker
人工蛋白質の設計

John Jumper and Demis Hassabis
AIによる”正確な” 蛋白質の3D構造予測
(AlphaFold 2)

- ほとんどの生物学者が、蛋白質の人工的なデザインと構造予測という恩恵を受けられるようになった。
- 新たな構造決定をしている研究者でさえ、その構造決定が迅速に行われ、分子機能やダイナミクスなどの別の情報を得る研究が進むという恩恵を受けている。



(写真はwwPDB websiteより)

Helen M. Berman教授の
コメント:
wwPDBの創設者の一人

「AlphaFoldと生物学のAI革命を生み出した巨大な蛋白質データベース」

「他のコミュニティもこれを行うことができ、そうすべきであり、そうしなければなりません。そうでなければ、大きなブレークスルーを得ることはできません。」

E. Callaway, *Nature* 634,
1028-1029 (2024)



Open Data/Open Science for Researchers

- The advancement of data science and science in general is anticipated through Open Data/Open Science initiatives.

→ However, some scientists do not perceive this as an urgent issue. Consequently, there is considerable reluctance to pay **high APCs** or to go through the cumbersome process of **depositing articles in Institutional Repositories**. In recent years, the Cabinet Office of Japan requires that researchers funded by public sources should publish their **articles along with the evidence data openly accessible without delays**, either via publishers or institutional repositories.

- Fair Evaluation of Researchers' Contributions

Currently, researchers' achievements are predominantly evaluated based on **Journal Impact Factor (JIF)**, **citation counts**, and the **perceived prestige** of the journals in which their work is published. As a result, evaluations often fail to reflect the true value of their contributions, hindering progress in **securing grants, academic appointments, promotions, and degree conferrals**.

→ If making all publications open access and providing the associated data—preferably raw data—in a machine-readable format enables **more accurate assessment of research contributions**, it is likely that **more researchers would willingly embrace such openness**

オープンデータ/オープンサイエンス:研究者にとっての利点

- オープンデータ/オープンサイエンスにより、データサイエンス及び科学一般の進歩が期待される

→ 科学者自身にとって喫緊の課題とは感じられにくく、高額なAPCの支払いや面倒な機関リポジトリへの掲載には消極的な姿勢が強い。
内閣府の近年の施策により、公的な研究資金で行われた研究結果は、根拠データとともに即時に論文を公開することが求められている。

- 研究者の業績が正当に評価される

現在、研究者の業績は、主にジャーナル・インパクトファクター(JIF)、論文引用数及び掲載ジャーナルのブランドに基づいて評価されている。その結果、正確な評価がなされにくい状況にあり、助成金の獲得、アカデミックな職の獲得、昇進、学位取得の進行を妨げている。

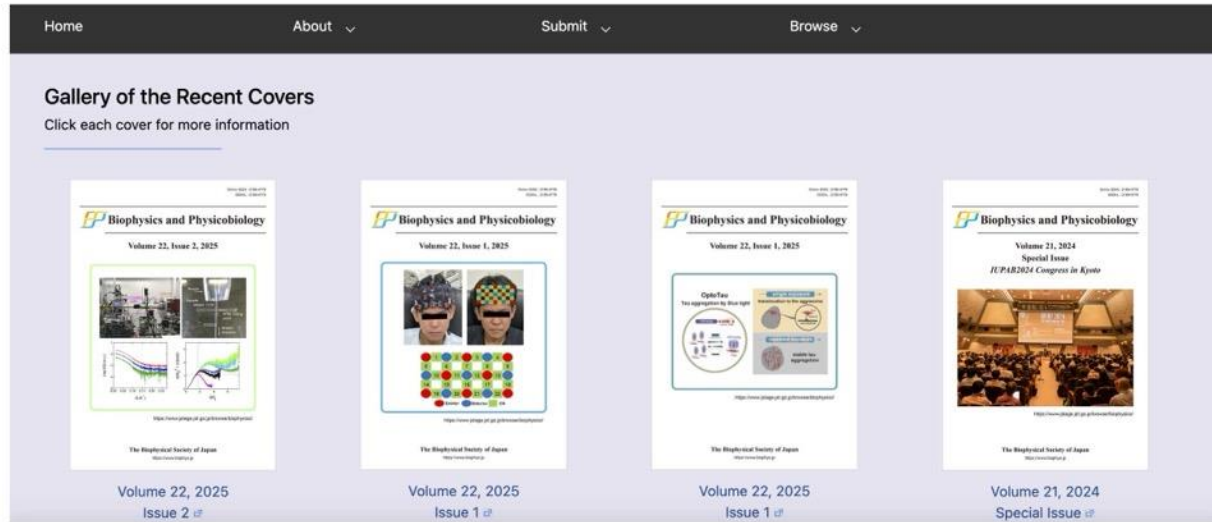
→ すべての出版物をオープンアクセスにし、根拠データ(できれば生データ)を機械可読形式で提供することにより研究業績をより正確に評価できるのであれば、論文と根拠データを公開する研究者は増えるはずである。

Open Data/Open Science for Stakeholders

- At funding agencies, the availability of open-access publications and datasets enables reviewers to conduct **fair and transparent evaluations** during various stages of the review process.
- To ensure fairness in evaluations, **only research outputs that are publicly available should be considered**, especially those most relevant to the project under review.
- In the planning, interim, and final evaluations by funding agencies, **inclusion of items related to the dissemination of data and research outcomes must be mandatory** (such as "Initiatives for disseminating information").
- Organizations responsible for the accumulation, curation, and maintenance of data and publications (e.g., **database administrators, journal publishers**) should **assign unique identifiers (such as DOIs)** to each dataset and publication, ensure systematic organization and public access with **minimum embargo periods**.

ステークホルダーにおけるオープンデータ/オープンサイエンス

- ・Funding Agencyにおいて、研究評価者が種々の評価時において、公開された論文やデータを参考にできるため、公正な審査が可能となる。
- ・公正な審査のためには、業績リストのうちで特に評価対象研究と関係が深いものは、公開されているもののみに限定する。
- Funding Agency における研究開発採択時の計画評価、中間評価、事後評価には、データや成果の公開に関する項目を必須とする（具体的には、「国内外への情報発信などの取組内容」などでの記載）
- ・データ及び論文の蓄積、キュレーション、維持を行う組織（Database Administrator, Journal Publisher）は、各データ及び論文にID（DOIなど）を付して整理・公開し、EMBARGO期間は可能な限り短いものとする。



**Peer reviewed, Open access, online Journal
 (JIF 1.6 in 2023)**

Year 2020: 25 article,

2021: 41

2022: 52

2023: 70

2024: 59

- BPPB is the official **peer reviewed open access journal** of the Biophysical Society of Japan (BSJ) **on the J-STAGE platform**, under either the **CC BY 4.0 license** or the **CC BY-NC-SA 4.0** license, depending on the authors' selection.
- BPPB applies **IT** (Information Technology) to reduce publication costs and time. In fact, the **APC** (Article Processing Charges) is now **44,000 JPY** including tax independent of the paper length.
- BPPB promotes the idea of **Open Data/Open Science** collaborating with **J-STAGE Data** (Data repository). **PMC** (PubMed Central) and **DOAJ** (Directory of Open Access Journals) certify us as the **Gold OA** (open access) journal.



査読有り、オープンアクセス、オンラインジャーナル
 (2023年のJIF 1.6)

2020年: 25記事

2021年: 41記事

2022年: 52記事

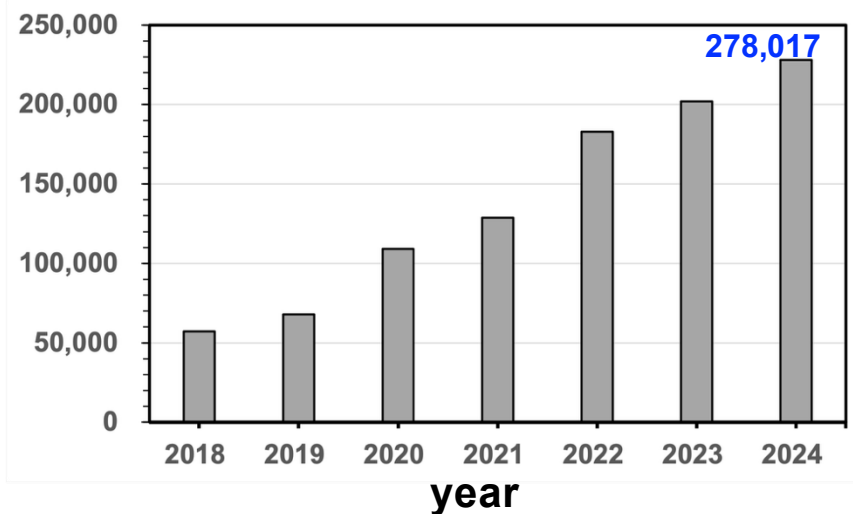
2023年: 70記事

2024年: 59記事

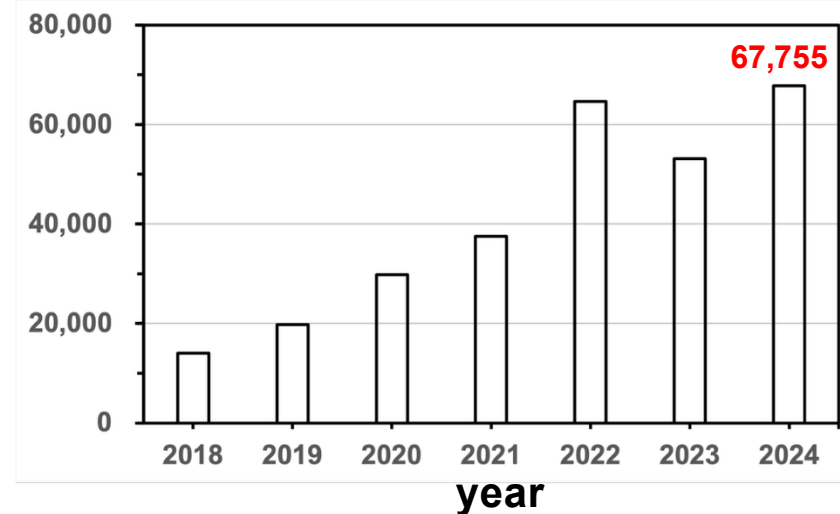
- BPPBは、**J-STAGE**登載の日本生物物理学会(BSJ)の公式**査読付きオープンアクセスジャーナル**であり、著者の選択に応じて**CC BY 4.0ライセンス**または**CC BY-NC-SA 4.0ライセンス**のいずれかで提供される
- BPPBは**IT**(Information Technology)を応用し、出版コストと時間を削減している。実際出版費、**APC**(Article Processing Charges)は、用紙の長さに関係なく、税込みで**44,000円**である
- BPPBでは、**J-STAGE Data**(データリポジトリ)と連携した**オープンデータ/オープンサイエンス**の考え方を推進しており、**PMC** (PubMed Central) と **DOAJ** (Directory of Open Access Journals)は、**Gold OA** (Open Access) ジャーナルとして認定している

Publication statistics

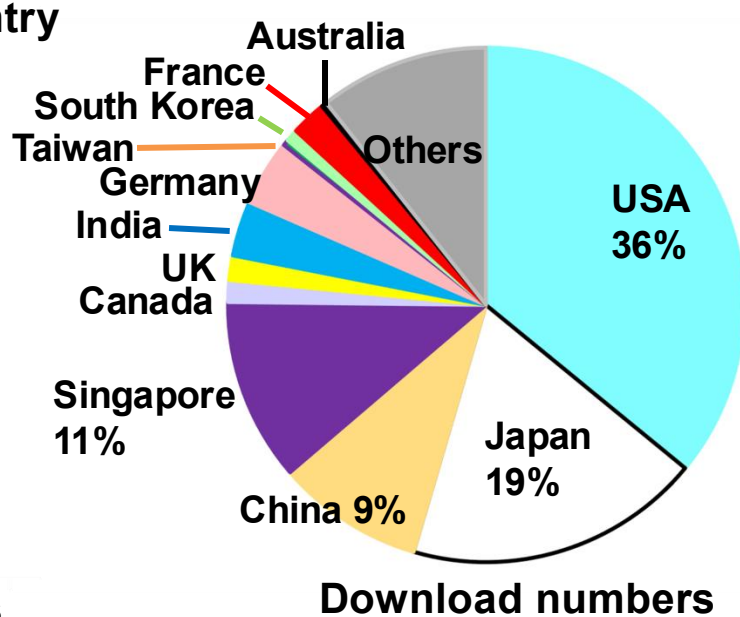
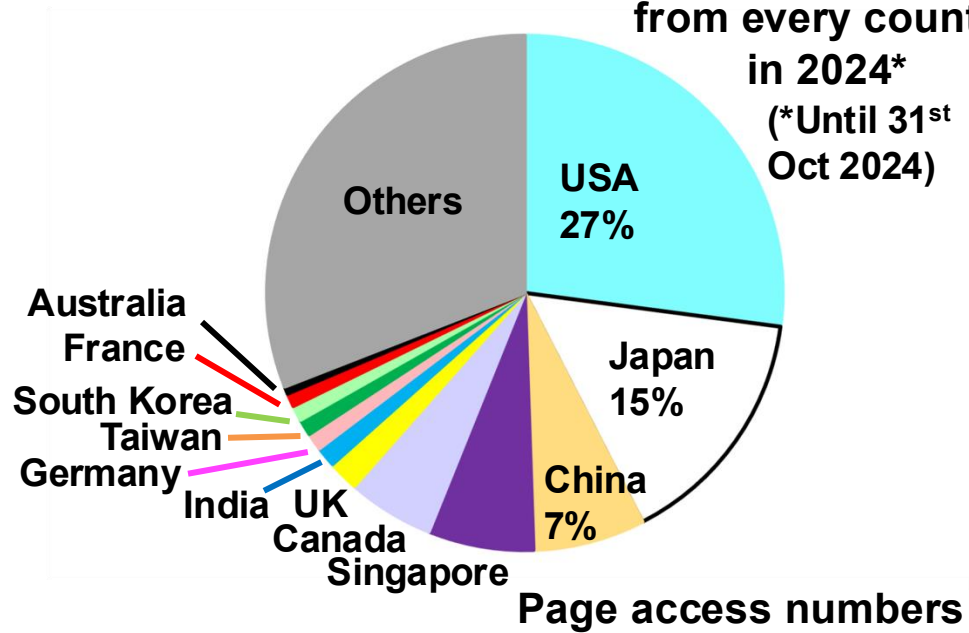
Access numbers to J-STAGE pages



Paper download numbers from J-STAGE



Access/DL numbers from every country in 2024* (*Until 31st Oct 2024)



Open Data/Open Science: Benefits for Citizens

Now, citizens are indirect stakeholders by paying taxes to Government.

However, they may become the largest stakeholders in the near future.

オープンデータ/オープンサイエンス:市民にとっての利点

現在、市民は政府に税金を支払うという形で、
間接的なステークホルダーである

しかし、今後、最大のステークホルダーになりうる

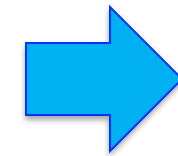
Open Data/Open Science: Benefits for Citizens

Popularization of Science: The current wave of the **science reformation** resembles **The Reformation (宗教改革)** in early 16th-century Europe.

The Reformation in 16th-Century Europe

New Testament of the Roman Catholic had been written **in Latin**.

- Martin Luther translated the Bible into **German (Luther Bible)**
- development of **printing technology** by Gutenberg.

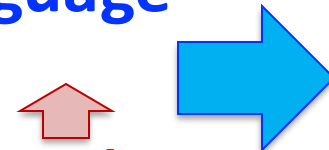


Collapse of the feudal system in Europe

The Science reformation in the 21st Century

Scientific publications have been written in highly technical language, primarily in English.

- Explanation of scientific content **in plain language across multiple languages**, by **Generative AI**
- Development of **Internet**



Open Data/Science

Popularization of science

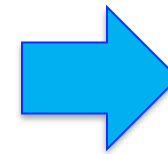
オープンデータ/オープンサイエンス:市民にとっての利点

科学の大衆化:現在の科学改革の波は、16世紀初頭のヨーロッパの宗教改革に似ている

16世紀ヨーロッパにおける宗教改革

ローマ・カトリックの新約聖書はラテン語で書かれていた

- マルチン・ルターによるドイツ語のルター聖書
- グーテンベルクによる印刷技術の発展

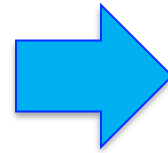


ヨーロッパにおける
封建制度の崩壊

21世紀の科学改革

科学の論文は、主に英語で難解な専門用語で出版されている

- 生成AIによる多言語での平易な解説
- インターネットの発達



科学の大衆化



オープンデータ/サイエンス

Conclusion: Open Data/Open Science Benefits for Researchers, Stakeholders and Citizens

Citizens can evaluate correctly individual researches and activities of researchers, and **will become the Largest, Richest and Direct Stakeholders of Research and Researchers**

Individual researchers and research may **obtain funds from lots of citizens in the world**, independent of prominent members of the society, Government Authority, etc. **by crowdfunding and/or Social Tipping** through Internet

**Open Data/Open Science may lead to a Paradigm shift:
The Science Reformation**

結論:オープンデータ/オープンサイエンス 研究者、ステークホルダー、市民にとっての利点

市民は、個々の研究者と業績を正確に評価することができ、研究と研究者にとって最大で最も裕福な直接的なステークホルダーになり得る

また、学会の重鎮や政権とは独立した形で、世界中の多くの市民から研究資金を得ることができる

例「クラウドファンディング」や「投げ銭」など

オープンデータ/オープンサイエンスは、
新しいパラダイムを生み出す可能性がある
「科学改革」

Acknowledgements

- Biophysical Society of Japan
(BSJ: 一般社団法人日本生物物理学会)
- Editorial Board members and Editorial Office
of Biophysics and Physicobiology (BPPB)
- Staffs of wwPDB and PDBj
- J-STAGE, J-STAGE Data, and JST

Thank you for your time.

謝辞

- 一般社団法人日本生物物理学会(BSJ)
- Biophysics and Physicobiology (BPPB)の編集委員および編集室
- wwPDBとPDBjのスタッフ
- J-STAGE、J-STAGE Data、JST

ご清聴、ありがとうございました