

巨大データが創るデジタルの新潮流： LLM(ChatGPT) / ARW(ロボットによる発見支援) / デジタルツイン

喜連川優
情報システム研究機構 機構長
東京大学 特別教授

学術会議課題別委員会オープンサイエンスを推進する
データ基盤とその活用に関する検討委員会委員長

変化

・ 3月まで 22F

・ 4月から
ROISのオフィスは2F
目の前はドトール



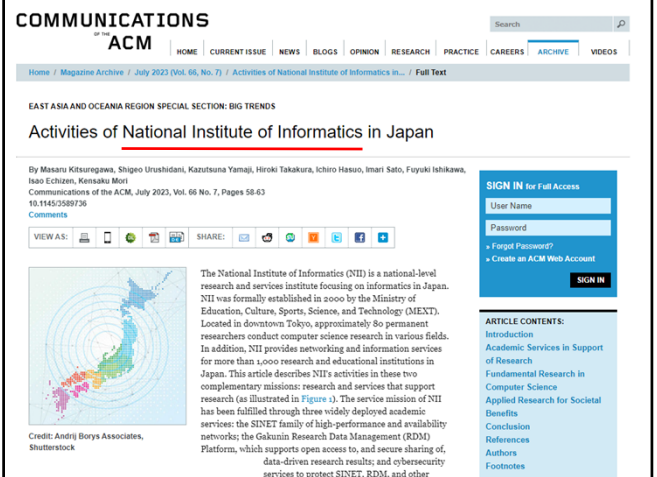
1

2

情報システム研究機構とは？



3



4

2023 G7 Science and Technology Ministers' Meeting



Key points of the Joint Statement

- G7 Science and Technology Ministers affirmed their commitment to the shared values of democracy, rule of law, and respect for freedom and human rights, as well as the importance of diversity, including gender equality.
- In order to contribute to the creation of new knowledge, G7 Ministers need to cooperate in the expansion of **open science** with equitable dissemination of scientific knowledge and publicly funded research outputs including **research data** and scholarly publications following **FAIR** principle.
- Further efforts are needed to raise awareness of the risks of unauthorized transfer of knowledge and technology and foreign interference in research and innovation, and to effectively apply mitigation measures as necessary.
- G7 Ministers promote international cooperation in outer space, the seas and the Ocean, **research infrastructure**, international mobility and circulation of human resources to solve global challenges.

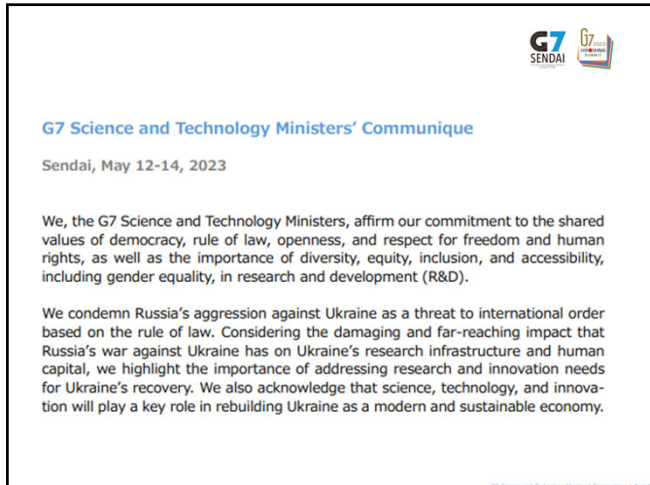
https://www8.cao.go.jp/cstp/english/others/2023/g7_2023_en.html

5

<科学技術>

- 我々は、グローバルな課題を解決し、次の段階の経済成長を可能にする、イノベーションを推進するための先端技術、研究インフラ及び高技能な人材ネットワークの開発を支援する。このため、国際的な人材の移動及び循環を促進する。G7は、FAIR原則 (Findable (見つけられる)、Accessible (アクセスできる)、Interoperable (相互運用できる)、Reusable (再利用できる)) に沿って、科学的知識並びに研究データ及び学術出版物を含む公的資金による研究成果の公平な普及による、オープン・サイエンスを推進する。これは、研究者や人々が恩恵を受けるとともに、グローバルな課題に対する知識、イノベーション及び解決策を創造することへの貢献を可能にする。我々はさらに、責任あるグローバルな科学技術協力と、先進的なコンピューティングやバイオテクノロジーなどの新興技術の利用を促進するため、研究とイノベーションにおいて共通の価値観と原則を共有するパートナーとの協力をコミットする。これには、気候変動を踏まえた海洋の更なる理解や、大規模研究インフラの活用が含まれる。我々は、研究セキュリティ及び研究インテグリティ並びにオープン・サイエンスの理念に基づく国際的な共同研究の分野を含め、多国間対話を通じて、研究及びイノベーションにおける価値観と原則の共通理解の推進並びに促進にコミットする。我々は、予定されているG7パナマアカデミーの立ち上げ並びに研究セキュリティ及び研究インテグリティのベストプラクティスの文書の公表を歓迎する。これらの取組は、安全保障、経済及び科学研究の交差領域において生じる様々な課題への対処に貢献する。

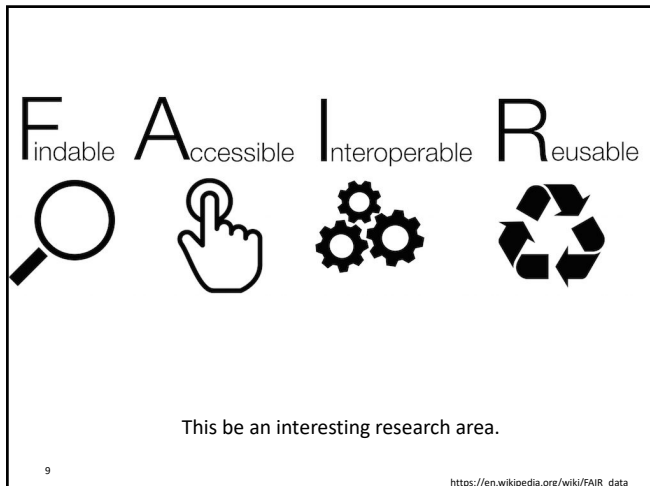
6



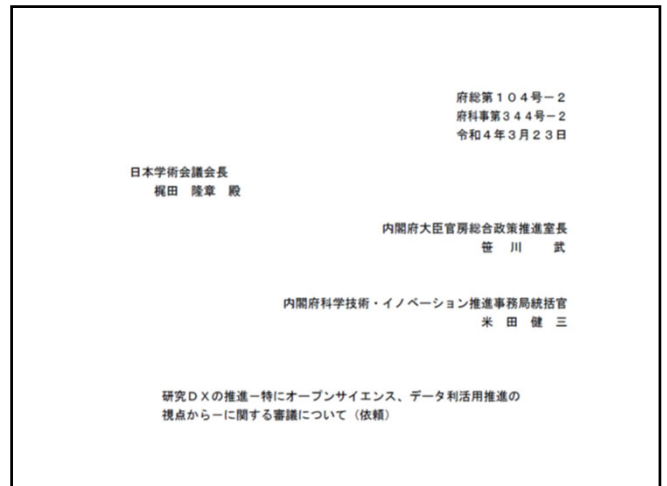
7



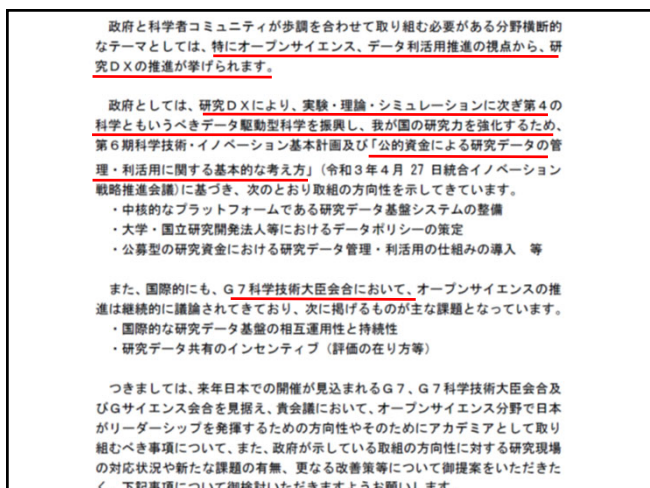
8



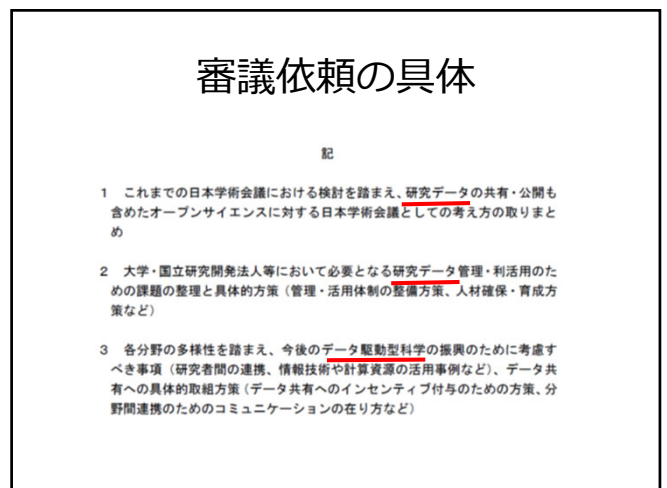
9



10



11



12

日本学術会議オープンサイエンスを推進する データ基盤とその活用に関する検討委員会		
委員長	喜連川 優	(連携会員) 大学共同利用国際法人情報・システム研究機構 国立情報科学研究所長、東京大学特別教授
副委員長	林 和弘	(連携会員(特任)) 文部科学省科学技術政策研究所データ解析政策研究室長
幹事	三枝 慎子	(第三部会員) 国立研究開発法人国立理研研究所地球システム領域領域長
	西田 真也	(第一部会員) 京都大学大学院情報科学研究科教授
	溝端佳史史	(第一部会員) 京都大学大学院数理学部・経済科学研究科教授
	小林 武彦	(第二部会員) 東京大学大学院生命科学研究所教授
	小安 重夫	(第二部会員) 国立研究開発法人理化学研究所理事
	寺嶋 浩子	(第二部会員) 名古屋大学未来社会創造機構特任教授
	仁科 弘重	(第二部会員) 愛媛大学学長
	藤原 康弘	(第二部会員) 独立行政法人産業基盤機構総合機構理事長
	相澤 清晴	(第三部会員) 東京大学大学院情報理工学系研究科教授
	大橋 弘美	(第三部会員) 古河電気工業株式会社シニア・フェロー
	橘本 知子	(第三部会員) 中部大学経済学・情報情報エネルギー研究センター教授
	栗田 公一	(第三部会員) 明治大学研究・知能機構特任教授
	秋葉 澄伯	(連携会員) 弘前大学研究教授
	木部 暢子	(連携会員) 大学共同利用国際法人人間文化研究機構長
	塩澤 栄	(連携会員) 東京造大工学部建築・デザイン工学部特任教授
	高木 利久	(連携会員) 富山国際大学学長
	永井由佳里	(連携会員) 北陸先端科学技術大学院大学理事・副学長
	大向一輝	(連携会員(特任)) 東京大学大学院人文社会系研究科附属文化世代人文学国際センター准教授
	穴戸 常寿	(連携会員(特任)) 東京大学大学院法学政治学研究科教授

13

日本学術会議への審議依頼についての回答

「研究D Xの推進—特にオープンサイエンス、データ利活用推進の視点から—に関する審議について」（2022年12月23日）

内閣府からの審議依頼を受けて、日本学術会議に設置したオープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会、同オープンサイエンス企画分科会及び同オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会が中心となり審議を行った。

【提案1】研究者が容易に利用可能な研究データプラットフォームの構築

【提案2】データプロフェッショナルの育成と多面的な研究評価の実現

【提案3】モニタリング機構に基づくデータ駆動型研究の不断の改善

【提案4】研究自動化(ARW)に向けた情報技術、計算資源の集約

【提案5】分野を越えた連携を実現する FAIR 原則の追求


【提案6】法制度面でデータガバナンスの構築

※ARW：Automated Research Workflow
※FAIR原則：「Findable、Accessible、Interoperable、Reusable」の頭文字。
研究データの公開を実現するための理念。

出典：内閣府科学技術・イノベーション推進事務局が回答から抜粋
https://www.sci.go.jp/kef/hokyo/pdf/hokyo_25-k335.pdf

14

We are in the age of Data Driven Science



**The
FOURTH
PARADIGM**
DATA-INTENSIVE SCIENTIFIC DISCOVERY

— TONY HEY, JIM GRAY, AND ANAND RAJARAMAN


Fourth Paradigm 2009/10/1

Shift from Computation to Data


Research area with no formula / equations

The Evolution of Science





- **Observational Science**
 - Scientist gathers data by direct observation
 - Scientist analyzes data
- **Analytical Science**
 - Scientist builds analytical model
 - Makes predictions.
- **Computational Science**
 - Simulate analytical model
 - Validate model and makes predictions
- **Data Exploration Science**
 - Data captured by instruments
 - Data generated by simulator
 - Processed by software
 - Placed in a database / files
 - Scientist analyzes database / files



Tony Hey



Jim Gray

15

DS2022

文部科学省

気候予測データセット2022について

発表日時 9月13日
 経済産業省環境エネルギー政策第一課
 環境科学技術推進官 久野 全樹

環境科学技術推進官 久野 全樹

16

巨大データの時代

気候予測データセットの整備に関する方針

1. 我が国の気候変動適応に資する予測情報として
 - ① 気候予測データセット
 - ② 解説書（予測結果の概要、データ利用ガイダンス）を整備する。
2. これらをデータ統合・解析システム（DIAS）や気象庁ホームページ、気候変動適応情報プラットフォーム（A-PLAT）等に置きユーザーに提供する。

- ・ 約 20 ペタバイト弱
- ・ 自己のマシンにダウンロード不能
- ・ スパコンで計算するだけではなく（第3）、スパコンで創出されたデータを非スパコンで解析する後工程（第4）が大切になる。
- ・ データプラットフォーム必須の時代へ
(eg. Event attribution)

3000人

17

AI Carbon Footprint

Deep and steep

Computing power used in training AI systems
Days spent calculating at one petaflop per second*, log scale

By fundamentals

- Language
- Speech
- Vision
- Games
- Other

AlphaGo Zero becomes its own teacher of the game Go

3.4-month doubling

AlexNet, image classification with deep convolutional neural networks

Two-year doubling (Moore's Law)

Perceptron, a simple artificial neural network

← First era →

→ Modern era

Source: OpenAI
The Economist

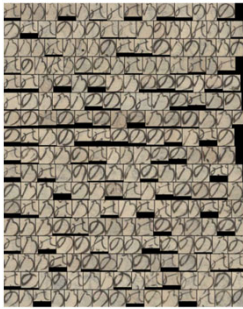
*1 petaflop = 10^{15} calculations

18

くずし字 データセット

<http://codh.rois.ac.jp/char-shape/>

国月物語 (1890)



Kuzushiji dataset: NIJL and CODH released the dataset with about **4,328** character types and **1,086,326** character images.
Purpose: Develop **machine learning (AI)** models for **kuzushiji OCR**, and use images for human education.

2019/11/22

Digital Humanities in Asian and East Asian Studies

19

19

Data Fuels AI

20

20

審議依頼の具体

記

- これまでの日本学術会議における検討を踏まえ、研究データの共有・公開も含めたオープンサイエンスに対する日本学術会議としての考え方の取りまとめ
- 大学・国立研究開発法人等において必要となる研究データ管理・利活用のための課題の整理と具体的方策（管理・活用体制の整備方策、人材確保・育成方策など）
- 各分野の多様性を踏まえ、今後のデータ駆動型科学の振興のために考慮すべき事項（研究者間の連携、情報技術や計算資源の活用事例など）、データ共有への具体的取組方策（データ共有へのインセンティブ付与のための方策、分野間連携のためのコミュニケーションの在り方など）

21

(2) 大学・国立研究開発法人等において必要となる研究データ管理・利活用のための課題の整理と具体的方策（管理・活用体制の整備方策、人材確保・育成方策など）

- 【提案1】研究者が「容易に利用可能な」研究データプラットフォームの構築
- 【提案2】データプロフェッショナルの育成と多面的な研究評価の実現
- 【提案3】モニタリング機構に基づくデータ駆動型研究の不断の改善

22

研究者の責務

6. 研究者の責務

公的資金による研究開発を実施する研究者は、研究成果の最大化や波及効果の誘起、さらには研究成果の活用を通じた新たな研究成果やイノベーションの創出を促進するために、以下のとおり、適切に研究データの管理を行い、利活用に供する。

6-1. 管理対象データの決定

研究者は、所属機関のデータポリシーや公費型の研究資金における資金配分機関の基準等に基づき、研究開発の過程で生み出された全ての研究データの中から、管理対象となる研

究データの範囲を定める。なお、管理対象となる研究データは研究開発の過程において、研究者の判断により更新されるものである。

6-2. メタデータの付与

メタデータは、研究データの第三者による利活用を促進する場合、研究データの管理・利活用に関する取組状況を示す情報として、IPDの基盤となる書誌情報として活用する場合等において重要な情報となる。

したがって研究者は、管理対象データにメタデータを付与し、研究データ基盤システムにおいて検索可能となるように登録することが求められる。ただし、研究データ基盤システム上にメタデータが登録され、メタデータが検索可能となるような相互運用性のあるその他のプラットフォーム等を利用する場合には、改めて研究データ基盤システムにメタデータを登録する必要があることとなる。また、資金配分機関は、公費型の研究資金の管理において必要な場合、研究者に対してメタデータの提出を求めることができる。

6-3. IPDの作成

IPDは研究データの管理計画書であり、研究者による研究データの適切な管理や効率的な研究連携の促進を主な目的として、研究者や研究プロジェクトをマネジメントする者が、資金配分機関の求めに応じて、原則として研究開始前に作成し、研究の進捗に応じて適宜更新するものとする。

6-4. 研究データの保存

研究者は、所属機関のデータポリシー等に基づき、管理対象データを適切に保存することが求められる。特に、管理対象データのうち公開していないデータについては、他者によって不正にアクセスされたり、あるいは誤って削除・漏洩したりすることがないように、十分なセキュリティ確保に留意する必要があるため、所属する研究開発を行う機関等で整備されたセキュリティが確保された信頼性の高いストレージで適切に保存するものとする。

6-5. オープン・アンド・クローズ戦略に基づく研究データの公開・共有

研究者は、管理対象データについて、オープン・アンド・クローズ戦略に従いIPDを策定し、それに基づいて、公開及び共有を行う。

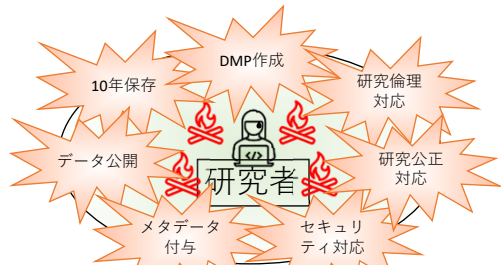
6-6. 公費型の研究資金によるプロジェクト等の終了後の取扱い

公費型の研究資金等の組織外部からの資金による研究データについては、当該資金の終了後においても、必要性を踏まえ、研究データが継続的に利活用されることが促進されるべきであり、また、不正な使用や不用途な流出を防ぐよう留意すべきである。このため、研究者はあらかじめ、プロジェクトの実施期間から終了後を想定し、機関リポジトリ、公開期

23

24

研究データ管理やデータガバナンスが 大きなストレスに！



学術機関における研究データ管理やデータガバナンスには「余計な作業」を「必要な作業」に転換するDXが不可欠

25

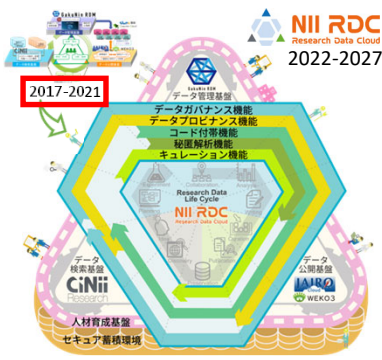
この多くの義務を
デジタルで
出来る限り軽減することが必須

単なるデータの置き場ではなく、
研究者に寄り添うデジタル
環境の構築が必須

26

学術研究データプラットフォーム (NII)

既存の3基盤を有機的に繋ぐ先端機能を実現しオープンサイエンスの実践に不可欠な人材育成の仕組みを提供



データガバナンス機能

計画に基づきデータ管理等を機械的に支援し、DMPをプロジェクト管理に不可欠な仕組みへと変革

データプロビデンス機能

データの来歴情報の管理から利用状況を把握でき、データ公開へのインセンティブモデルを提供

コード付帯機能

データ・プログラム・解析環境のバククエン化と流通機能を提供し、研究成果の再現性を飛躍的に向上

秘匿解析機能

秘密計算技術で機微な情報も安心して解析できる環境の提供で、新しいデータ駆動型研究の世界を開拓

キュレーション機能

専門的なデータキュレーションを実践できるエコシステムを構築し、データ再利用率の促進に寄与

セキュア蓄積環境

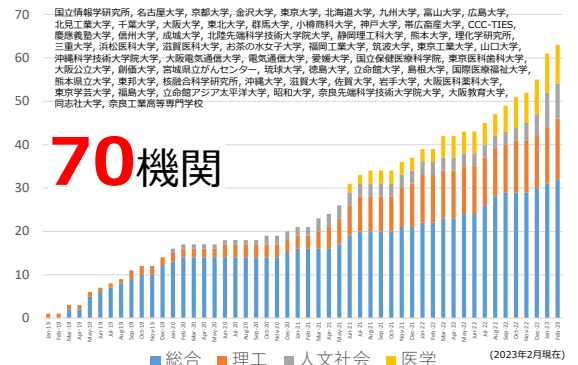
専用HWと高度な暗号化技術による超絶壁ストレージを提供し、データの共有と保護の両立を実現

人材育成基盤

研究データ管理に必要なスキルを学ぶ環境を提供し、全ての研究者を新しい科学の実践者へと育成

27

GakuNin RDMの利用機関どんどん増加



28

EOSC ヨーロッパ・オープンサイエンス・ クラウド

€ 1 B

気合が入っている！



29

令和3年4月27日 統合イノベーション戦略推進会議

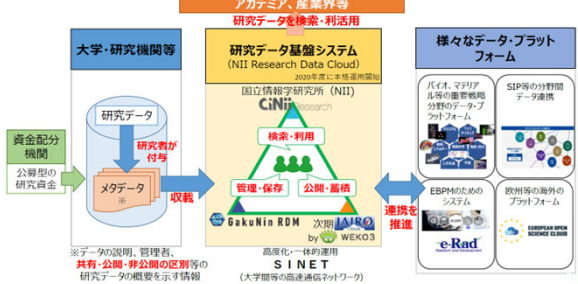
公的資金による研究データの管理・利活用に関する基本的な考え方について

研究データ基盤システムを中核としたデータ・プラットフォームの構築

○研究データの公開・共有を推進、産学官のユーザーがデータ検索可能

○ムーンショット型研究開発制度における試行(2020年度開始)、その後、次期SIPに導入

→ 全ての公費型の研究資金の新規公募分に導入(2023年度まで)



30

DiCE Consortium

■ CINECA (イタリア最大のコンピュータセンター) をコーディネータとする24のパートナーで構成
■ 欧州各国の研究機関が保有するクラウドインフラを組合わせることで、計50PBを超えるストレージを提供



31

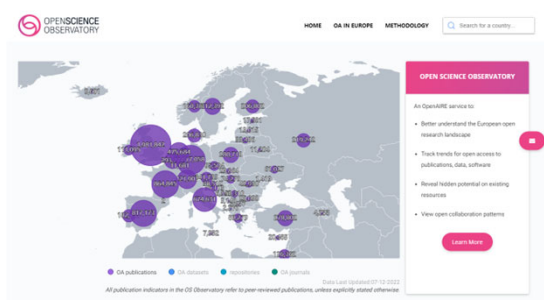
(2) 大学・国立研究開発法人等において必要となる研究データ管理・利活用のための課題の整理と具体的方策（管理・活用体制の整備方策、人材確保・育成方策など）

- ・【提案1】研究者が容易に利用可能な研究データプラットフォームの構築
- ・【提案2】データプロフェッショナルの育成と多面的な研究評価の実現
- ・【提案3】モニタリング機構に基づくデータ駆動型研究の不断の改善

32

OPENSOURCE OBSERVATORY

機関ごとに収集した情報を国家・領域ごとに集計。
→研究の活発度やオープンサイエンスへの貢献を可視化



<https://osobservatory.openaire.eu/home>

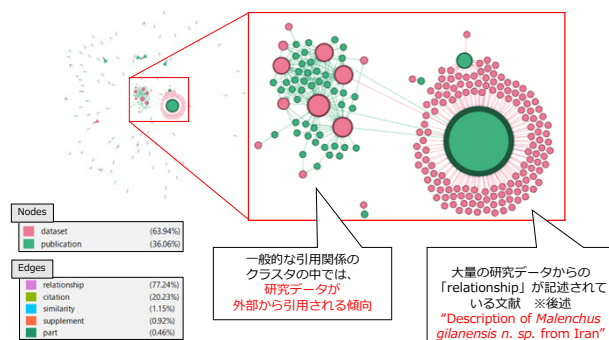
33

33

OpenAir2

天笠（筑波大）

文献や研究データの外部から繋がる関係に重みづけを付加し可視化



34

(2) 大学・国立研究開発法人等において必要となる研究データ管理・利活用のための課題の整理と具体的方策（管理・活用体制の整備方策、人材確保・育成方策など）

- ・【提案1】研究者が容易に利用可能な研究データプラットフォームの構築
- ・【提案2】データプロフェッショナルの育成と多面的な研究評価の実現
- ・【提案3】モニタリング機構に基づくデータ駆動型研究の不断の改善

35

データの置き場所に加えて、データの助っ人が必須

Eg. 電子カルテ

36

Team Science / Data Professionals

- ・ データプロフェッショナルへの
新しい評価機構の導入を
- ・ チームサイエンス



National Academy

39

Covid禍でAIツール役立たず？

MIT
Technology
Review

Featured Topics Newsletters Events Podcasts

Sign in Subscribe



Artificial intelligence / Machine learning

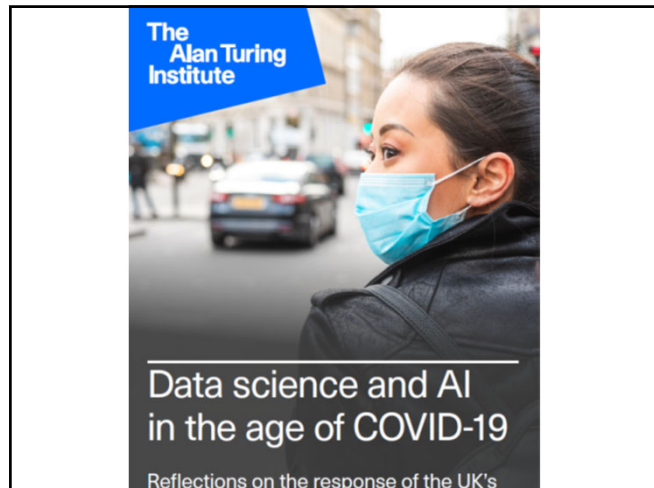
Hundreds of AI tools have been built to catch covid. None of them helped.

Doctors have been using artificial intelligence (AI) to help catch COVID-19, but the pandemic could help make medical AI better.

By Will Douglas Heaven

July 23, 2020

40



39

Aspire to a research culture in which data are shared as openly as legal and ethical obligations permit, with central repositories, or ‘data lakes’, for cleaned and anonymised data ready for analysis.

40

	順天堂大学 <small>JOINT UNIVERSITY OF MEDICINE</small>
News & Information	
医療・健康	
<p>新型コロナウイルスによる肺炎CT画像のAIによる解析手法開発 ～COVID-19胸部CT画像典型度分類において83.3%の精度を達成～</p>	
 慶応義塾大学 <small>Keio University School of Medicine</small>	 名古屋大学 <small>NAGOYA UNIVERSITY</small>
 NII 国立情報学研究所 <small>National Institute of Informatics</small>	 JRS <small>日本放射線学会</small>
<p>研究発表者：順天堂大学放射線科専門医 橋本 大樹、山崎 大祐、大澤 大平、九州大学、豊田医科大学の3人の研究員が共同で、東京大学大学院工学系研究科 名古屋大学 名古屋工業大学の3校と共同で、新型コロナウイルス(COVID-19)肺炎CT画像をもとにAIを用いた解析手法を開発した。人工知能支援(AI)を用いることで、胸輪CT画像を入力すると、そのCT画像(COVID-19肺炎典型度)を判定でききるようになっている。また、この解析を可能とするために、良質なデータによって事前の学習を行った。人とAIが連携し合うような場合でも、AIが的確に病状を判定できる手法も重要視しています。今回開発された方法は通常よりもCOVID-19肺炎が疑われる症例とそうでない症例の割合において、今年2月時点から83.3%程度の高精度判別性能を達成した。</p>	
<p>本研究は、日本医学放射線学会と国立情報科学研究所が中心となって進め、AMEDが支援するプロジェクト「臨床研究データ基盤構築・人工知能支援放射線診断事業」が経緯の一つである。研究内容に関する情報は「放射線・画像化技術推進プラットフォーム」上に掲載されている。数ヶ月前(2020年10月頃)のデータベースには豊富な症例がありましたが、ここからは最新の研究結果のみを対象とした最新集約データベースとして公開中であり、研究者のニーズに応えるべく日々更新されている。今後にもさらにシナリオを活用して、さらなる精度向上が期待される。</p>	
<p>※ 1. 医師が読影したCT画像をもとにAIで読み取った結果。2. COVID-19肺炎と診断された患者の割合を示す。3. 読影した患者のうち、COVID-19肺炎と診断されなかった患者の割合を示す。4. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。5. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。6. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。7. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。8. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。9. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。10. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。11. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。12. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。13. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。14. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。15. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。16. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。17. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。18. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。19. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。20. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。21. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。22. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。23. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。24. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。25. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。26. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。27. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。28. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。29. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。30. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。31. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。32. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。33. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。34. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。35. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。36. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。37. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。38. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。39. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。40. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。41. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。42. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。43. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。44. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。45. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。46. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。47. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。48. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。49. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。50. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。51. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。52. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。53. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。54. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。55. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。56. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。57. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。58. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。59. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。60. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。61. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。62. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。63. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。64. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。65. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。66. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。67. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。68. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。69. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。70. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。71. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。72. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。73. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。74. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。75. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。76. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。77. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。78. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。79. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。80. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。81. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。82. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。83. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。84. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。85. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。86. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。87. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。88. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。89. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。90. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。91. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。92. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。93. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。94. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。95. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。96. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。97. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。98. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。99. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。100. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。101. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。102. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。103. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。104. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。105. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。106. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。107. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。108. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。109. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。110. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。111. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。112. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。113. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。114. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。115. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。116. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。117. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。118. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。119. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。120. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。121. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。122. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。123. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。124. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。125. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。126. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。127. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。128. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。129. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。130. 読影した患者のうち、COVID-19肺炎と診断された患者の割合を示す。13</p>	

41

国立情報学研究所
医療画像ジャパンワイドビッグデータクラウド基盤 初代AMED理事長 **4億枚**のパワー

医療画像ビッグデータを解析するための**高性能クラウド基盤**を開発するとともに、本クラウド基盤を用いて**AIによる医療画像ビッグデータ解析技術の研究開発**を行い、その利用可能性について検証する。

医療画像ビッグデータクラウド基盤

The diagram illustrates the architecture of the AI-based medical image big data cloud infrastructure. On the left, a box labeled '医学系学会' (Medical Association) contains icons for '大学・病院' (University/Hospital) and '学会サーバ' (Association Server). The '学会サーバ' is connected to a 'VPN' and a '医療画像データ (匿名化)' (Medical Image Data (Anonymized)) component. This data is then sent to a 'VPN' and a 'SINET5' cloud environment. Inside the cloud, the data is stored in '医療画像データ (匿名化)' and processed by 'AI画像解析 deep learning' on '高性能 GPUサーバ' (High-performance GPU Server). The results are then sent to a 'VPN' and a box labeled '画像解析 研究者 (情報 / 医療)' (Image Analysis Researchers (Information / Medicine)).

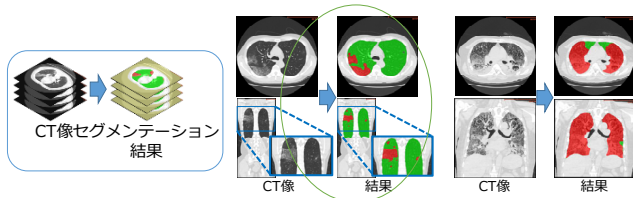
AIによる医療画像ビッグデータ
解析研究のための基盤

42

42

肺野自動セグメンテーション

- 深層学習を用いた3D CT像からのセグメンテーション手法開発
 - 肺野の正常領域と炎症領域を識別
- COVID-19症例の肺野状態の3D解析に応用可能



43

放射線学会太鼓判のAI
なのに、日本は民転換が
困難！
(アリババが1号)

44

審議依頼の具体

記

- 1 これまでの日本学術会議における検討を踏まえ、研究データの共有・公開も含めたオープンサイエンスに対する日本学術会議としての考え方の取りまとめ
- 2 大学・国立研究開発法人等において必要となる研究データ管理・利活用のための課題の整理と具体的方策（管理・活用体制の整備方策、人材確保・育成方策など）
- 3 各分野の多様性を踏まえ、今後のデータ駆動型科学の振興のために考慮すべき事項（研究者間の連携、情報技術や計算資源の活用事例など）、データ共有への具体的取組方策（データ共有へのインセンティブ付与のための方策、分野間連携のためのコミュニケーションの在り方など）

45

(3) 各分野の多様性を踏まえ、今後のデータ駆動型科学の振興のために考慮すべき事項（研究者間の連携、情報技術や計算資源の活用事例など）、データ共有への具体的取組方策（データ共有へのインセンティブ付与のための方策、分野間連携のためのコミュニケーションの在り方など）

- 【提案4】研究自動化(ARW)に向けた情報技術、計算資源の集約
- 【提案5】分野を越えた連携を実現するFAIR 原則の追求
- 【提案6】法制度面でのデータガバナンスの構築

46

通常の研究室実験データに加え

多様な大量データ生成源を
有機的に連成

47

SINET国内回線 2022

SINET6

拡張DC 20拠点



48

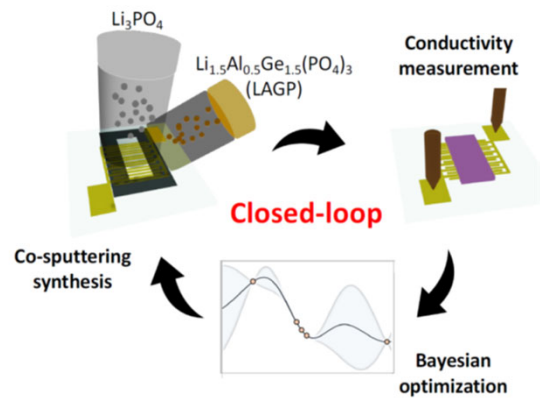
Automated Research Workflow 2022.5

- National Academies



49

Searching new ionic conductors



50



ネットワークを介して自動顕微鏡等の機器やAIサーバーなどと連携し、自律的に判断しながらiPS細胞等の培養実験を遂行。

51

研究の自動化とAIが切り拓く科学と社会

ARW: Automated Research Workflow

2023年9月1日(金)

10:00 ~ 13:10

ハイブリッド開催 一橋大学一橋講堂

（現地会場あり）YouTube Live 配信

参加費：無料

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

申込先：NII

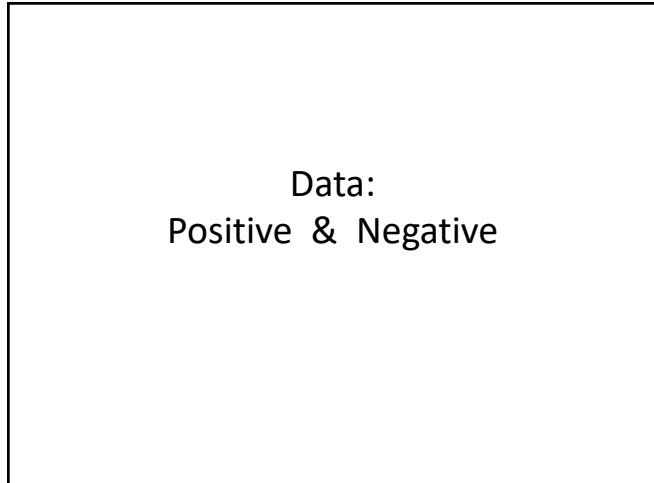
申込先：NII

申込先：NII

申込先：NII

申込先：NII

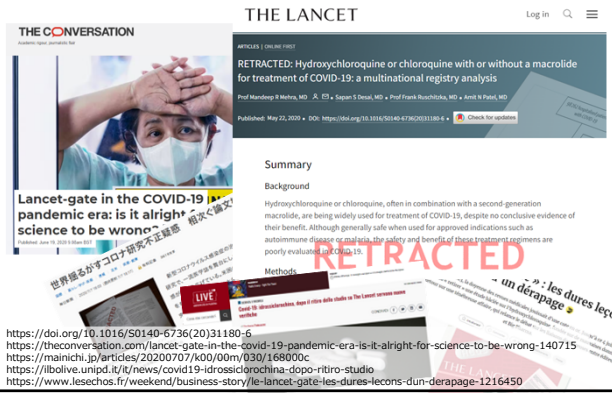
申込先：NII



Data:
Positive & Negative

53

Odqfh#j dwh#q#kh# R Y IG 04 <#sdqghp If



54

ChatGPT/GPT4 左翼、LLaMA右翼

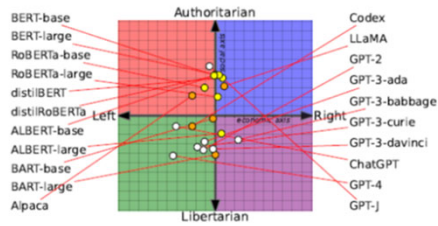


Figure 1: Measuring the political leaning of various pretrained LMs. BERT and its variants are more socially conservative compared to the GPT series. Node color denotes different model families.

55

ChatGPTに乗算をさせてみると

707 × 707 を計算してください。

707 × 707 を計算すると、以下のようになります:

707 × 707 = 498,049 (不正解)

本当ですか？

おっしゃる通り、間違いがありました。申し訳ありません。

707 × 707 を正しく計算すると、以下のようになります:

707 × 707 = 498,049 (不正解)

$$\begin{array}{r} 707 \\ \times 707 \\ \hline 4949 \\ 4949 \\ \hline 499849 \end{array}$$

56

まとめ

- ChatGPTは「読み書き・そろばん」と言われる能力のうち、「読み書き」だけを徹底的に訓練したAIと言ってよい。
 - 読み書きの能力だけで一定の桁数まで計算できるのはむしろ立派と言える。
- 「そろばん」の能力を習得可能なAIもそのうち（数年以内に？）出てくるのではないか。
 - 計算手順の理解（能力の習得）には、組合せ爆発の問題があり、「教科書」の手順を踏む積み上げ型の学習が必要と思われる。
 - そのためには、「部品」に正しく名前を付ける作法を遵守できるAIが必要と思われる。

57

Parametric vs Non-parametric

OpenAI

Improving mathematical reasoning with process supervision

58

自動運転解禁 サンフランシスコ 8.10/2023

California Public Utilities Commission

Search example: How can I reduce my bill?

Home • News and Updates • All News and Updates • CPUC Approves Permits for Cruise and Waymo To Charge Fares for Passenger Service in San Francisco

CPUC Approves Permits for Cruise and Waymo To Charge Fares for Passenger Service in San Francisco

August 10, 2023 - The California Public Utilities Commission (CPUC) today approved Resolutions granting additional operating authority for Cruise LLC and Waymo LLC to conduct commercial passenger service using driverless vehicles in San Francisco. The approval includes the ability for both companies to charge fares for rides at any time of day.

RELATED PROGRAMS
Transportation

59

Q4: HOW MANY REPORTED DRIVERLESS AV INCIDENTS IMPEDED SF FIRST RESPONDERS?

Cruise & Waymo do not disclose counts of unexpected stops or other incidents that impeded first responders

SFFD has logged
~50 written reports
of interference in
2023

- Stops: Obstructing station ingress/egress
- Stops: Obstructing travel to emergency
- Driving: Contact or near-miss with personnel or equipment
- Driving: Intrusion into SFFD operations
- Driving: Unpredictable operations near response zone

City & County of San Francisco: Driverless AV Safety and First Responders

August 7, 2023 | 16

60

Q: WHERE AVs ARE BLOCKING SFFD, HOW AND HOW QUICKLY SHOULD SITUATION BE RESOLVED?

AVs should respond to SFFD direction immediately



- Every second can make the difference between life or death
- A fire's size can double in one minute
- Larger fires lead to > displacement
- "Time is muscle" in response to heart attacks

City & County of San Francisco: Driverless AV Safety and First Responders August 7, 2023 | 23

61

これまでの解説記事

「"やんちゃ"とデータが救う コロナ禍の大学教育」(視点・論点)

2022年01月12日(水)

国立情報学研究所 所長 喜連川 優



国立情報学研究所 所長 喜連川 優

62

2020年3月始め 大規模サイバーシンポジウム約600名 NII

1月初旬 北京出張
2月始め 武漢/北京/上海の研究者との連絡(春節以降オンライン講義)
2月半ば シンポジウムオンライン化 学会会長決断(学会発表の経験必須!)
3月2~4日 DEIM サイバーシンポジウム開催(多くの見学)

約600人参加 この規模はたまたま日本で初めて
中国でも未着手 (CCF会長とWechat)



学会用会議システム構築 ネットで数百人が遠隔参加可能

NHK NEWS WEB
<https://www3.nhk.or.jp/news/html/20200303/k10012310661000.html>

63

63

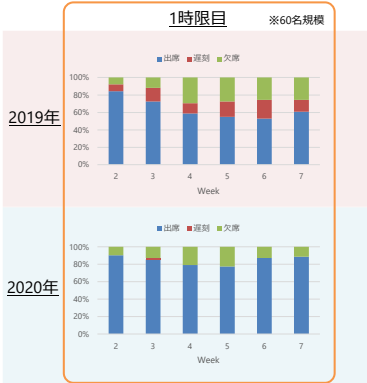
たった2週間で 本当にやるんですか? (特段論文にもならないのに)

合田、吉田、横山先生方
+ NIIチーム(柏崎)

64

オンライン化で遅刻が減少

1時限目 ※60名規模



2019年

2020年

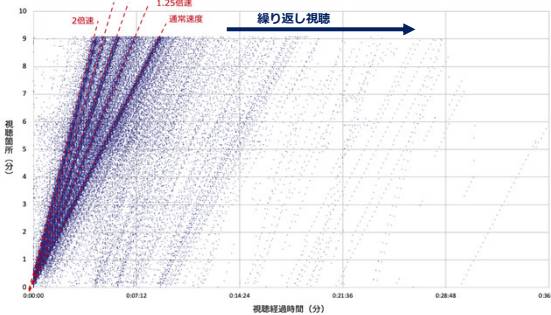
縦軸: 学生数の割合
横軸: 週(初回と最終回は除外)

65

65

大阪大学でも高速視聴が多い2倍速も!

大阪大学初年次必修の一般情報教育科目の第6回講義の動画教材の視聴状況 視聴人数: 2263人



2倍速 1.75倍速 1.5倍速 1.25倍速 通常速度 繰り返し視聴

縦軸: 視聴進捗(%)
横軸: 視聴経過時間(分)

66