

AWS メディアセミナー

朝日新聞社メディアラボの取り組み

2021/3/18 文責: 田森秀明

<https://cl.asahi.com>

Copyright 2021 The Asahi Shimbun Company.

朝日新聞社メディアラボ 研究チーム



田森 秀明

メディアラボ 主査、博士(情報科学)

2003年朝日新聞社入社。製作本部東京
発送部、同工程管理セクション、システム
セクションなどを経て現職。2015年7月
～2016年6月スタンフォード大学アジア
太平洋研究所客員研究員。



田口 雄哉

メディアラボ兼情報技術本部開発部

2017年奈良先端科学技術大学院大学
情報科学研究所修了。大学院では自然
言語処理の研究に従事。同年、朝日新聞
社に入社。



浦川 通

メディアラボ

2013年早稲田大学大学院基幹理工学
研究科 数学応用数理専攻修了。2019年
朝日新聞社入社。インタラクティブアー
ト・広告領域でのエンジニア・テクノロジ
ストなどを経て現職。

研究チーム メンバー

■ 5名



飯塚 荘太

メディアラボ

2016年慶應義塾大学理工学研究科修
了。同年、朝日新聞社入社。大学での画
像処理の知識をベースに、自然言語処
理にも研究分野を広げる。



杉野 かおり

メディアラボ

広報、情報誌編集者、編集デスク、編集
タイアップ担当等を経て、機械学習を学
び、2020年朝日新聞社入社。2児の母。

研究チームのミッションとモチベーション

将来の社内外のニーズに備え、基礎研究を通じてノウハウを先取りして蓄積していく

技術の社内保有

コアとなる技術は社内に保有したい
優秀な人材を集めるトリガー

社会課題の解決 マネタイズ

社内外の問題解決
膨大な記事データの再活用でマネタイズ

学術・社会への貢献

データの公開、特許出願、論文執筆 など

TSUNA

https://cl.asahi.com/api_data/headlinegeneration.html



TSUNA

Text SUMmarization Application

自動要約API TSUNA

九州電力玄海原発3、4号機(佐賀県玄海町)の運転差し止めを住民らが求めた仮処分申し立ての即時抗告審で、福岡高裁(山之内紀行裁判長)は10日、住民側の抗告を棄却した。主な争点は、耐震設計の基になる基準地震動(想定される最大の揺れ)の合理性、原発周辺の火山の噴火リスク、配管の安全性の3点。住民側は「基準地震動が過小評価されている」と主張。原子力規制委員会の安全審査の…

記事本文

13文字で出力
新幹線要約風…

スタイル
文字数



要約文

- 玄海原発差し止め抗告を棄却(13文字)
- 九州電力玄海原発3、4号機の運転差し止め申請を却下。福岡高裁、再稼働差し止めの即時抗告を棄却(新幹線)

- スタイルや文字数(8~100文字)をコントロール
- いくつも見出しや要約をつけないといけない、編集者が少ないメディアで活用を期待

自動要約生成API

入力文

九州電力玄海原発3、4号機(佐賀県玄海町)の運転差し止めを住民らが求めた仮処分申し立ての即時抗告審で、福岡高裁(山之内紀行裁判長)は10日、住民側の抗告を棄却した。主な争点は、耐震設計の基になる基準地震動(想定される最大の揺れ)の合理性、原発周辺の火山の噴火リスク、配管の安全性の3点。住民側は「基準地震動が過小評価されている」と主張。原子力規制委員会の安全審査の内規(火山影響評価ガイド)は、破局的噴火が予測できることを前提としている点是不合理と指摘。阿蘇カルデラ火山の噴火による火砕流が原発の敷地に到達する可能性も、十分小さいとは言えないと訴えていた。配管については「九電の検査方法では損傷が見逃されることがありうる」としていた。九電側は、基準地震動の評価について「各種調査で地域的な特性を把握した上、過小にならないようにしており合理的だ」と反論。「原発の運用期間中に破局的噴火が起きる可能性は極めて低い」と主張していた。配管については「健全性の確保に向け万全を期している」としていた。佐賀地裁は2017年6月、「安全性に欠けるところがあるとは認められない」などとして、仮処分の申し立てを却下。住民側が即時抗告していた。

生成例

- 玄海原発差し止め抗告を棄却(13文字)
- 玄海原発の運転差し止め、住民側の抗告棄却 福岡高裁(26文字)
- 九州電力玄海原発3、4号機の運転差し止め申請を却下。福岡高裁、再稼働差し止めの即時抗告を棄却(新幹線)
- 九州電力玄海原発3、4号機の運転差し止めを求めた仮処分申し立ての即時抗告審で、福岡高裁は住民側の抗告を棄却した。(題字下インデックス)

実用例(社内)

新幹線ニューステロップ惜しむ声 AI開発の矢先に悲報

有料会員記事 新幹線のいま

細沢礼輝、佐藤恵子 2020年3月5日 17時00分

シェア ツイート ブックマーク スクラップ メール 印刷



■ ANDES用短文ニュース生成

- 短文の作成時間が90分の1に



■ おすすめ記事ツイート

- ツイート文の作成補助で3割効率UP(ツイート数)
- 詳細は次ページ...



朝日新聞デジタル @asahicom

「デジタル分野の核兵器」急ぐ中国 桁違いの計算能力

見出し

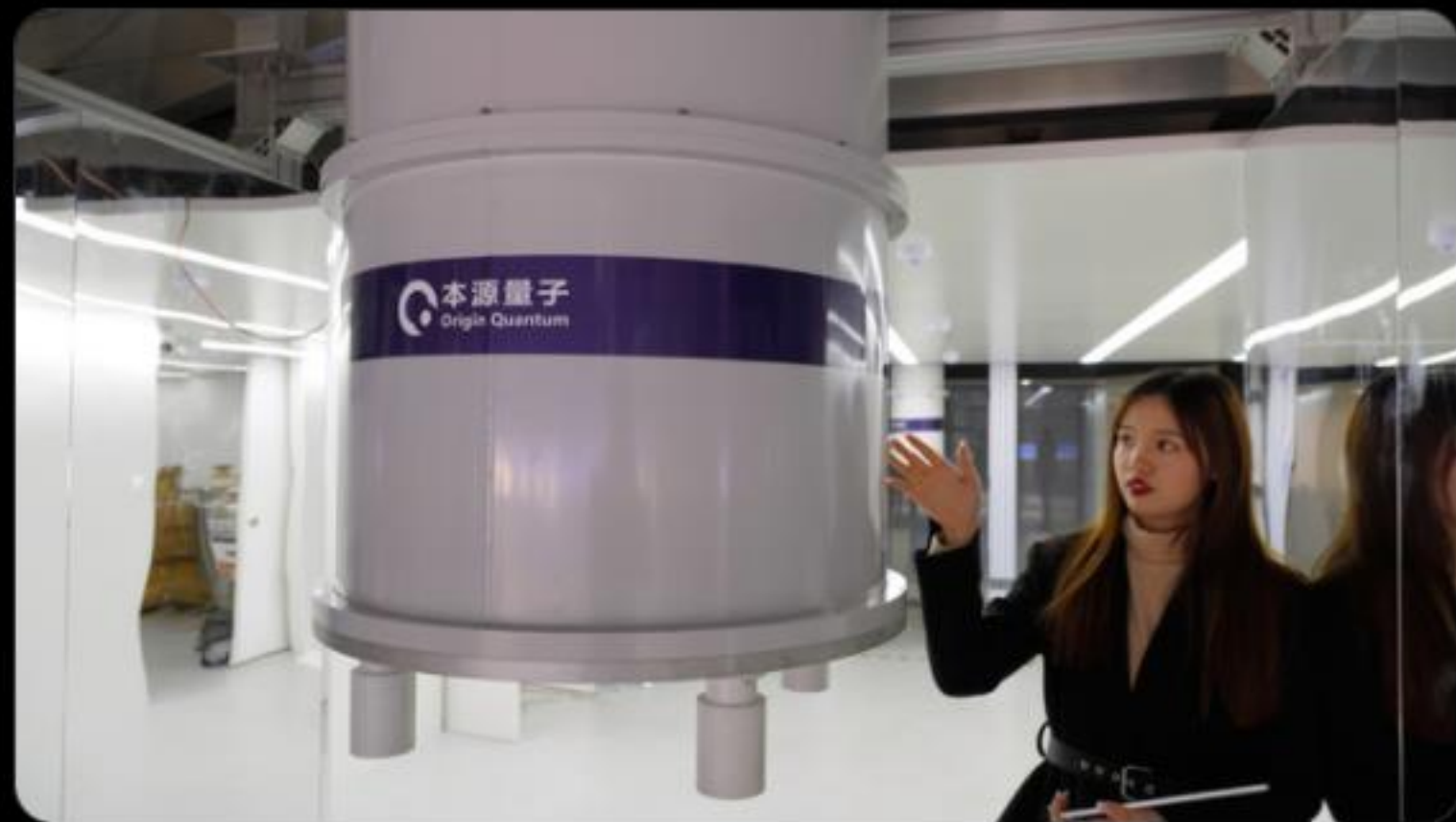
asahi.com/articles/ASP32...

URL

中国政府が「科学の都」と位置づける都市・合肥市では、約1100億円超を投じた量子技術の研究所の建設も進んでいます。中国で激化する量子コンピューターの開発競争を取材しました。

記事の要約
作成効率を3割UP

Translate Tweet



7:25 AM · Mar 3, 2021 · asahi_twico



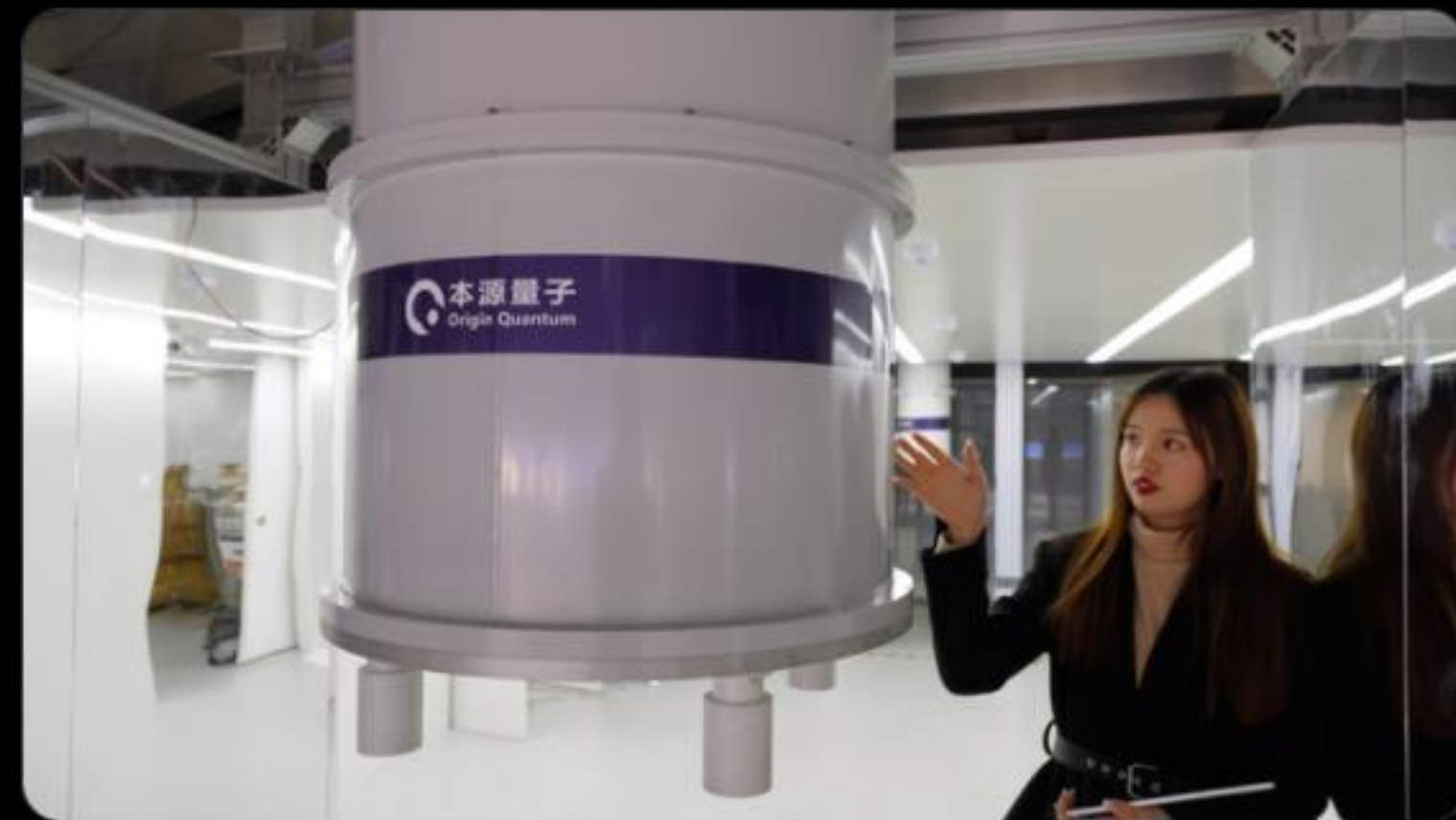
朝日新聞デジタル @asahicom

「デジタル分野の核兵器」急ぐ中国 桁違いの計算能力

asahi.com/articles/ASP32...

中国政府が「科学の都」と位置づける都市・合肥市では、約1100億円超を投じた量子技術の研究所の建設も進んでいます。中国で激化する量子コンピューターの開発競争を取材しました。

Translate Tweet



7:25 AM · Mar 3, 2021 · asahi_twico

言語モデルによる文書生成

言語モデル

言語モデル

「日本語の構造」を学んだ人工知能モデル

穴埋め問題を学習

ご都合 の よい お [MASK] を 教えて ください 。

文章の繋がりを学習

きょうは良いお天気です。 ・ 先日はありがとうございました。

言語モデルがあると...

- 日本語が正しくかけているかどうか分かる
 - 自動校正
- 文章が自動で書ける
 - 自動記事生成
- 文章の内容や特徴を数値化して計算できる
 - 文書分類

次頁以降のデモでご紹介

言語モデルを利用した自動文書生成

https://cl.asahi.com/api_data/gpt2-demo.html

#この記事は実在しません

GPT-2 Text Generation Demo

by Asahi Shimbun Media Lab

この記事の内容は、AIが自動生成したフィクションです_

言語モデルを利用した文書分類

https://cl.asahi.com/api_data/covid_19.html

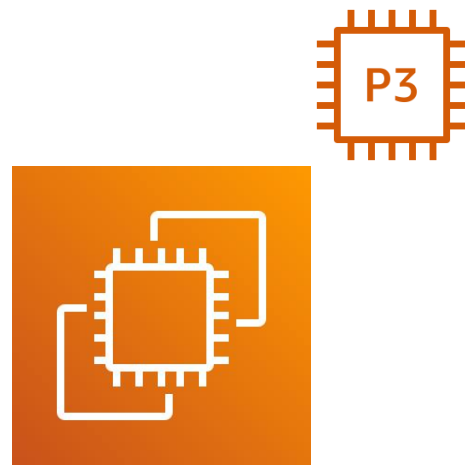
#COVID_19WordMap

powered by メディアラボ × Stroly

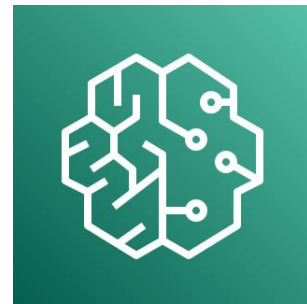
AWSの利用状況

AWS環境

学習

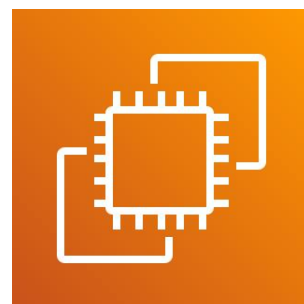


主にGPU インスタンス

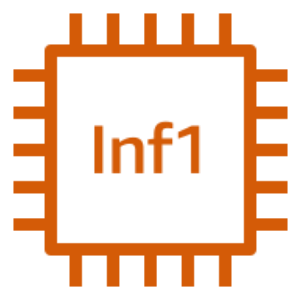


+ 社内オンプレサーバー

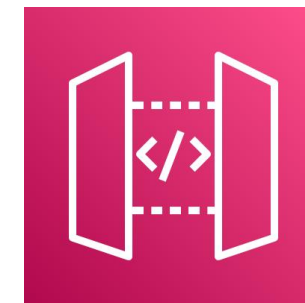
推論・API提供



主にCPU インスタンス



(テスト中)



最後に…

■ 4月より「メディア研究開発センター」へ改組

- 「メディアラボ・研究チーム」と「情報技術本部・研究開発チーム (ICTRAD)」が合体
- 12名の組織へレベルアップ
- AI技術を専門に扱うR&D専門部隊

ありがとうございました

朝日新聞社メディアラボ 研究チーム

<https://cl.asahi.com>

tamori-h@asahi.com