



# 研究データの活用・共有・ 公開・長期保存に向けた AWSの活用

[aws-jpps-er@amazon.com](mailto:aws-jpps-er@amazon.com)

アマゾン ウェブ サービス ジャパン合同会社 パブリックセクター  
2022/12

# 内容についての注意点

本資料では2022年12月6日時点のサービス内容および価格についてご説明しています。最新の情報はAWS公式ウェブサイト(<http://aws.amazon.com>)にてご確認ください。

資料作成には十分注意しておりますが、資料内の価格とAWS公式ウェブサイト記載の価格に相違があった場合、AWS公式ウェブサイトの価格を優先とさせていただきます。

価格は税抜表記となっております。日本居住者のお客様が東京リージョンを使用する場合、別途消費税をご請求させていただきます。

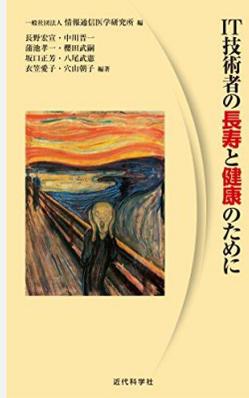
AWS does not offer binding price quotes. AWS pricing is publicly available and is subject to change in accordance with the AWS Customer Agreement available at <http://aws.amazon.com/agreement/>. Any pricing information included in this document is provided only as an estimate of usage charges for AWS services based on certain information that you have provided. Monthly charges will be based on your actual use of AWS services, and may vary from the estimates provided.



# 自己紹介

櫻田武嗣 (Takeshi SAKURADA) ,Ph.D

パブリックセクター シニアソリューションアーキテクト  
主に教育・研究機関のお客様を担当



情報処理学会出版委員会監修

経歴等 2000～ 郵政省(総務省)：通信総合研究所にて主に広帯域ネットワークに関わる研究に従事

2003～ 東京農工大学にてキャンパスITに関わる広範囲な研究・業務に従事

キャンパス内ITシステムその他、国立大学の1/5が利用する遠隔講義システム等の大学に関わるITシステムの立案から運用まで幅広く経験。学会、研究機関各種委員。企画した複数のシステムは製品化され他大学でも利用されるなど幅広く利用されている。

2018～ アマゾン ウェブ サービス ジャパン

シニアソリューションアーキテクト。高等教育機関や研究所等向けのシステム設計支援の他、クラウドコンピューティング活用のための啓蒙活動や講義・講演等実施。

好きなAWSサービス Direct Connect / Amazon S3



## Our Mission

地球上でもっともお客様を  
大切にする企業であること



# 最大規模のコミュニティ

AWS は世界中で何百万ものアクティブなお客様によって支えられています



## ● AWS User Group – Japan (JAWS-UG)

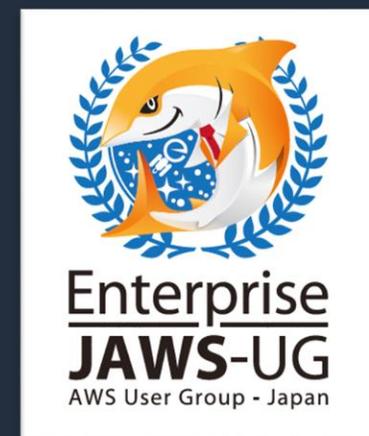
JAWS-UG は AWS が提供するクラウドコンピューティングを利用する人々の集まり（コミュニティ）です

一人ではできない学びや交流を目的として  
ボランティアによる勉強会の開催や交流イベントなどが行われています

## ● Enterprise-JAWS

エンドユーザー企業を対象としたクローズドな場として設立された  
AWS をお使いの企業のコミュニティです

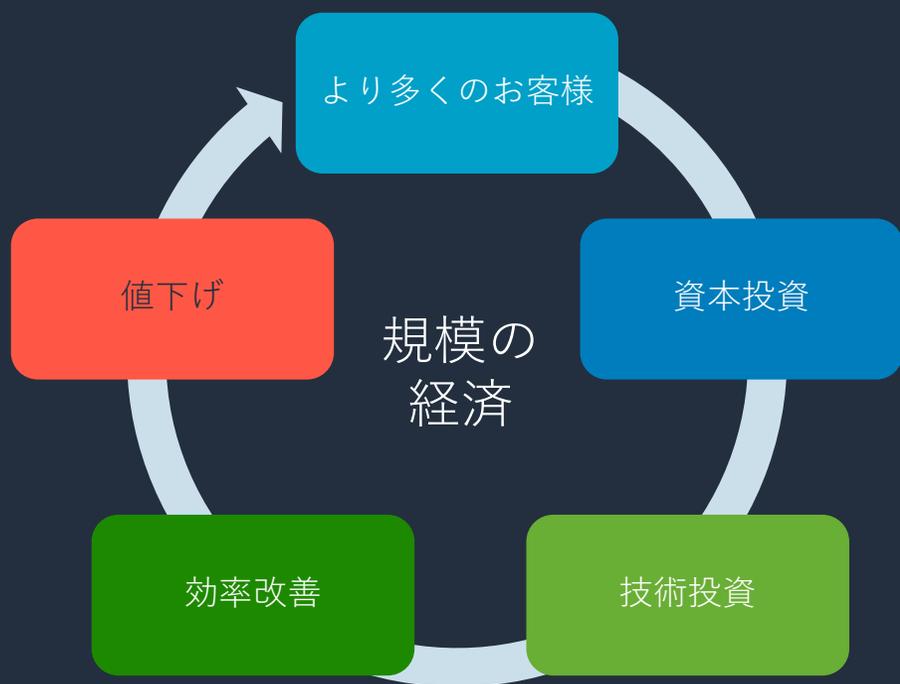
非公開を前提とした場で AWS のクラウド活用における課題や取り組みに  
関するさまざまなディスカッションが進んでいます



# スケールメリットをお客様に還元

AWSは2006年のサービス開始以降、非常に早いスピードで提供規模を拡大し、実質的に制限のないITリソースを提供しています  
このスケールメリットを生かし、コスト低下分をお客様に還元することで111回を超える値下げを実現しています

## 規模の拡大とイノベーション



## コストダウンを推進

継続的な値下げによる  
お客様への還元

115 回以上

\*2022年4月7日時点の情報



# データを安全に保管・長期保管・大容量データ保管

- 研究・業務データを保管したい
  - ✓ 長期保存が求められる (長期継続保管...、ストレージの入れ替えの対応を減らしたい...)
  - ✓ 将来どれだけのデータになるか分からない (研究データ、ログ...)
  - ✓ オンプレミスストレージのバックアップ、災害対策をしたい , etc ...

## 例) Amazon S3の活用

- ✓ 高耐久、3ヶ所以上に自動で隔地保存される
- ✓ 2006年からサービススタート、長期稼働の実績
- ✓ 利用した分だけの課金、無制限の容量 → スモールスタート可能
- ✓ データレイクとして様々なデータを集約、後から様々な方法で活用、S3へ直接クエリ等
- ✓ 利用頻度の低いデータは自動でアーカイブ階層としてコスト削減可能
- ✓ オンプレミスストレージのバックアップ先としてからはじめる、S3プラグインなどで対応しているストレージも多数
- ✓ 機関内と閉じた環境で接続、共同研究機関とのデータ共有での利

# 研究データの管理



<https://aws.amazon.com/jp/blogs/news/researchdatastore/>

- 研究データを10年以上保存の必要性
- 研究データ管理の仕組みが必要
  - 国立情報学研究所(NII)にて、研究データ管理基盤GakuNin RDMの提供を開始



下記の点を考慮する必要

1. メンバー管理・アクセスコントロール
2. ファイル等のバージョン管理
3. 研究証跡の記録
4. ファイル保管
5. 高い耐久性でかつ大容量のストレージ
6. 長期保管

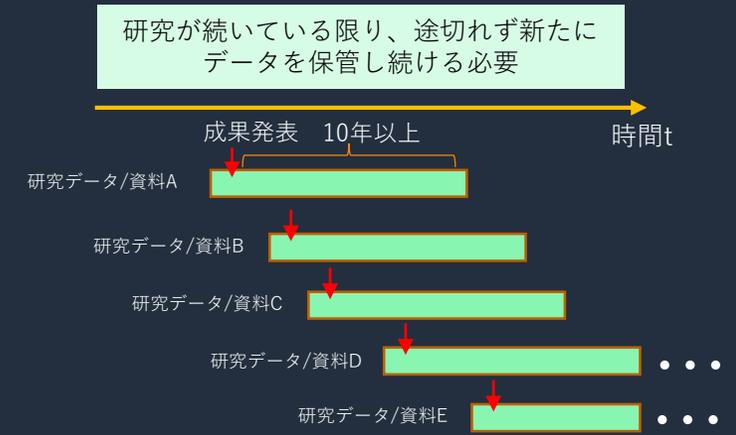
オープンソースソフトウェア  
等が開発され利用されている

ストレージの永続性が重要  
Amazon S3等の利用



Amazon S3は2006年からサービスを提供

- ✓ AWS上に構築されたシステムと閉じた環境で接続も可能
- ✓ 使用しただけのコスト、スモールスタートが可能
- ✓ 標準で99.99999999%の高い耐久性、地理的に離れた場所に3つの複製



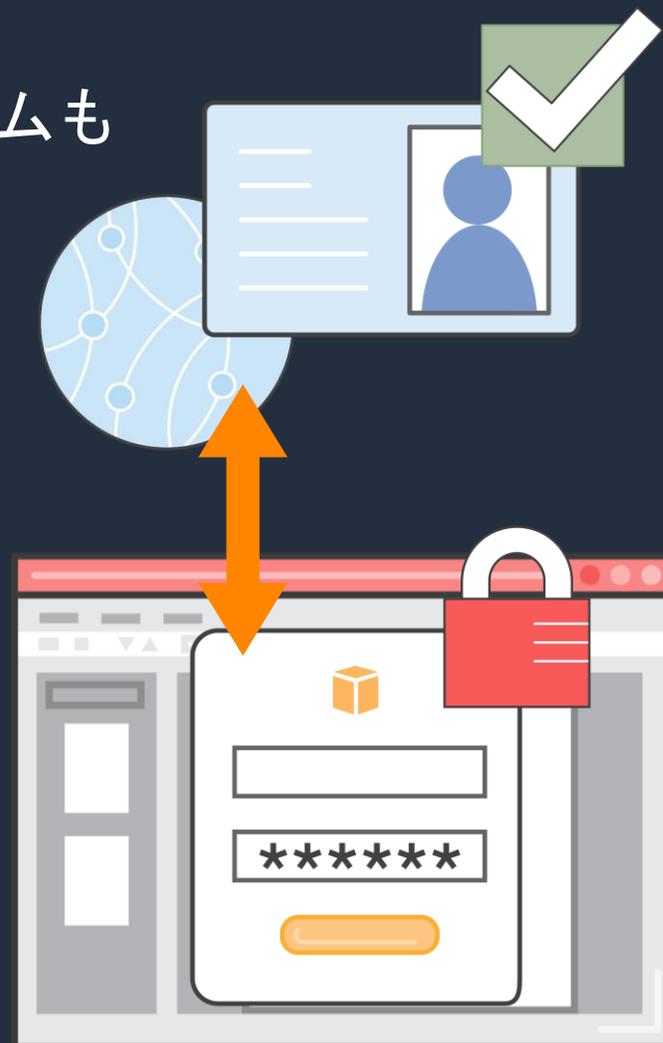
スモールスタートが可能



# 認証システムも安定した AWS の基盤上に

- 様々なシステムやサービスと連携する認証系のシステムも安定した AWS 基盤上で利用
- 学認(shibboleth) などの連携
  - IDaaS や on AWS上に環境を構築しての利用
    - Extic, okta, Cisco Duo, HENGE
    - AXEOLE, Account@Adapter+
    - KeyCloak

など



# 学認IdPとして機能するIDaaS「Extic」

- 学外から学認SP利用、多要素認証やシングルサインオンなど必要なセキュリティ機能をサービスとして提供しているIDaaSが「Extic」です。



Extic については、展示会場F-01 エクスジェン・ネットワークス社の  
展示ブースがございます。

# AWSのストレージ関連サービス

## データの格納 (ストレージリソースの提供)



Amazon S3 and  
Amazon S3 Glacier

Object (API)



Amazon EBS

Block



Amazon EFS

NFS



Amazon FSx for  
Windows File Server

SMB



Amazon FSx for  
NetApp ONTAP

SMB / NFS / iSCSI



Amazon FSx for  
Lustre

Lustre



Amazon FSx for  
openZFS

NFS

## データ管理

## エッジサービス

## データの移動 (物理デバイス)



AWS Backup

AWSサービス全体の  
バックアップとコン  
プライアンス



AWS Storage  
Gateway

オンプレミスからク  
ラウドストレージへ  
のゲートウェイ



AWS DataSync

AWSストレージサー  
ビスへのデータ転送



AWS Transfer  
Family

ファイル転送プロト  
コル (SFTP / FTPS /  
FTP) の提供



AWS Snowmobile

AWSストレージサービスと、クラウドのエッジやネットワークのエッジ間で  
データの移動を行うための物理デバイス

国内では、AWS Snowball EdgeとAWS Snowconeを提供中



AWS Snowball  
Edge

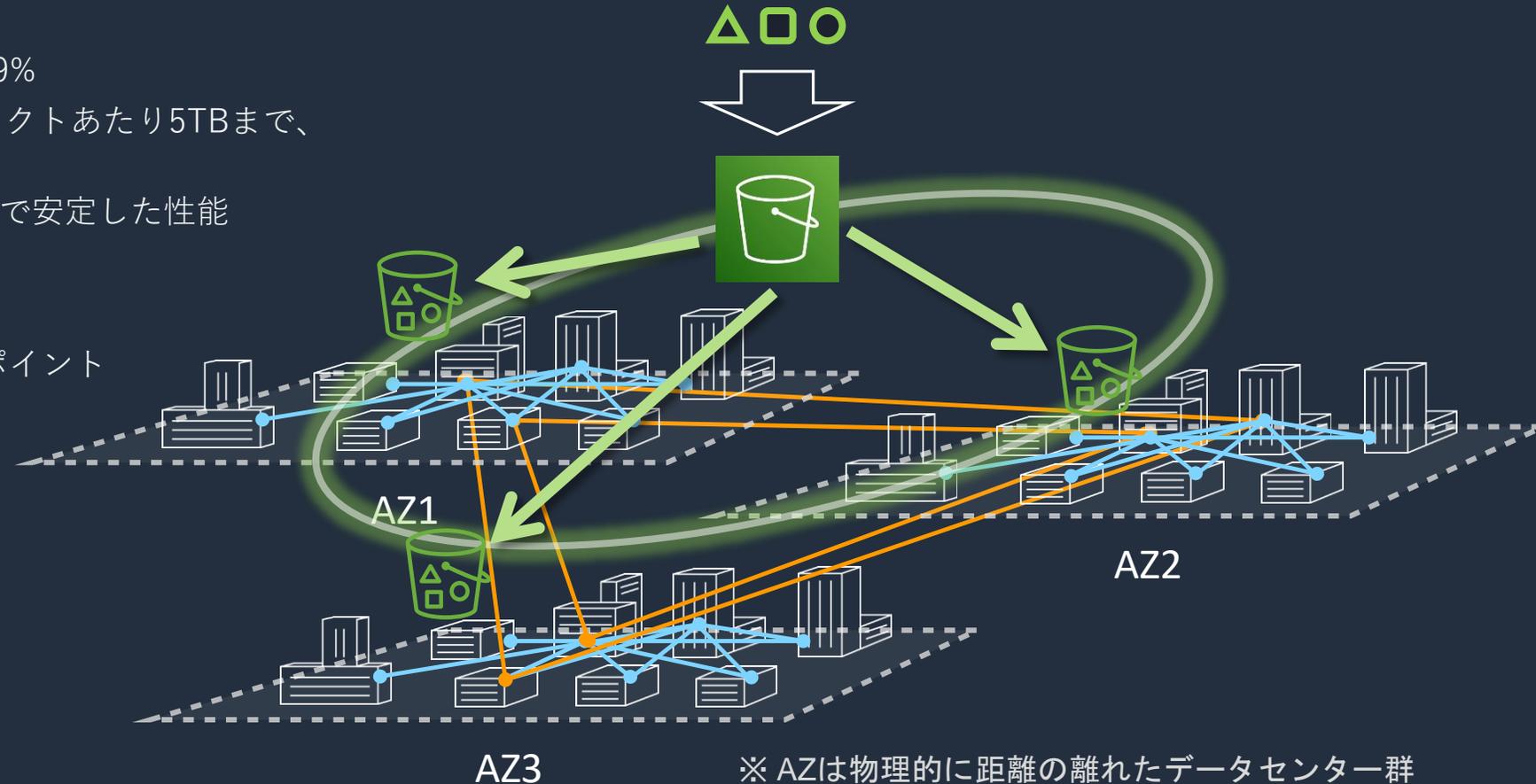


AWS Snowcone

# Amazon S3とAvailability Zone (AZ)

データ保存・バックアップ用途に向くオブジェクトストレージ Amazon S3

- http(s)でアクセス可能
- 自動的に3箇所以上のAZ (意味のある距離離れて独立したデータセンター群)へ隔地保管 (S3 OneZone-IA は除く)
- AZ間、データセンター間は、低遅延のプライベートネットワークで接続
- 設計上のデータ耐久性は、99.999999999%
- 容量無制限、サイジング不要 (1オブジェクトあたり5TBまで、オブジェクトの数は無制限)
- データ容量に依存しない、スケーラブルで安定した性能
- 暗号化をサポート
- WEBの静的コンテンツ配信機能
- IAM、バケットポリシー、S3 アクセスポイントなどによる細かなアクセス制御が可能
- PrivateLinkを利用しDirectConnect経由でアクセス可能



# Amazon S3をデータレイクとして活用

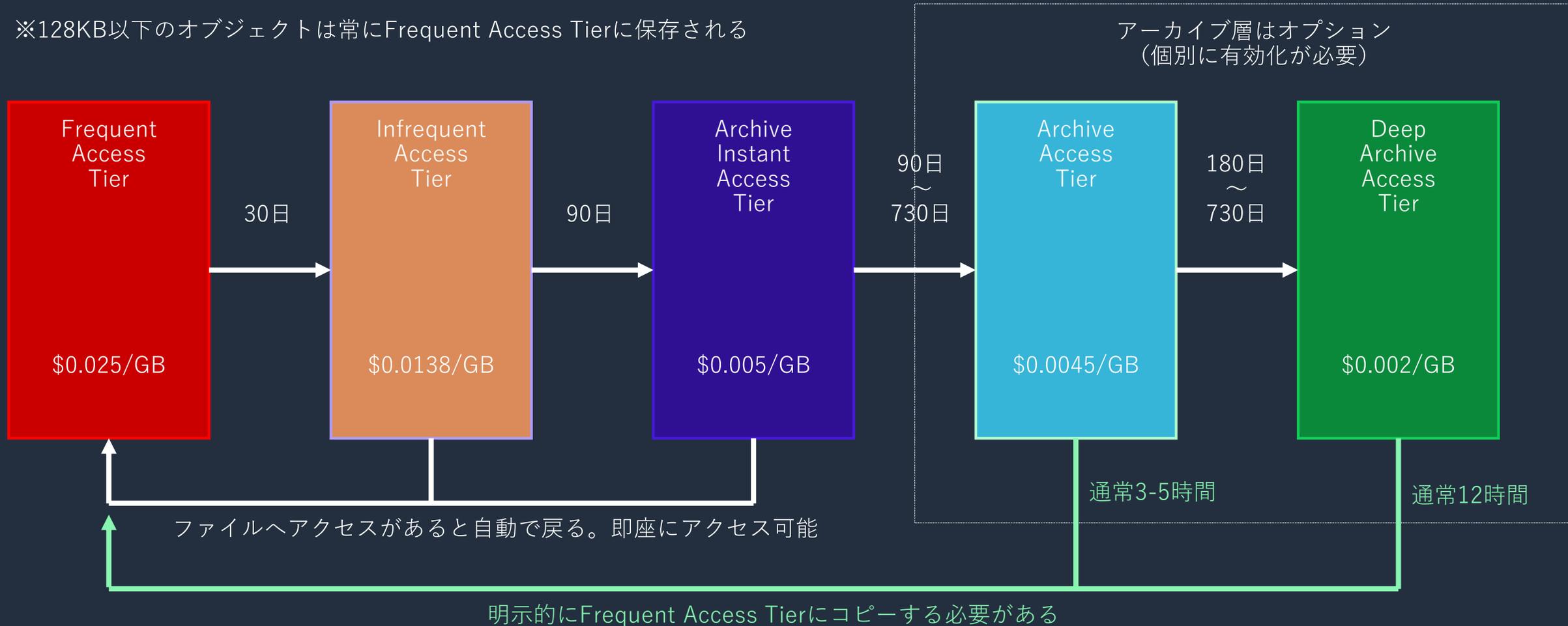
Amazon S3 を中心に様々なデータ活用のサービスと連携



# ストレージ階層を利用したデータ保護とコスト低減

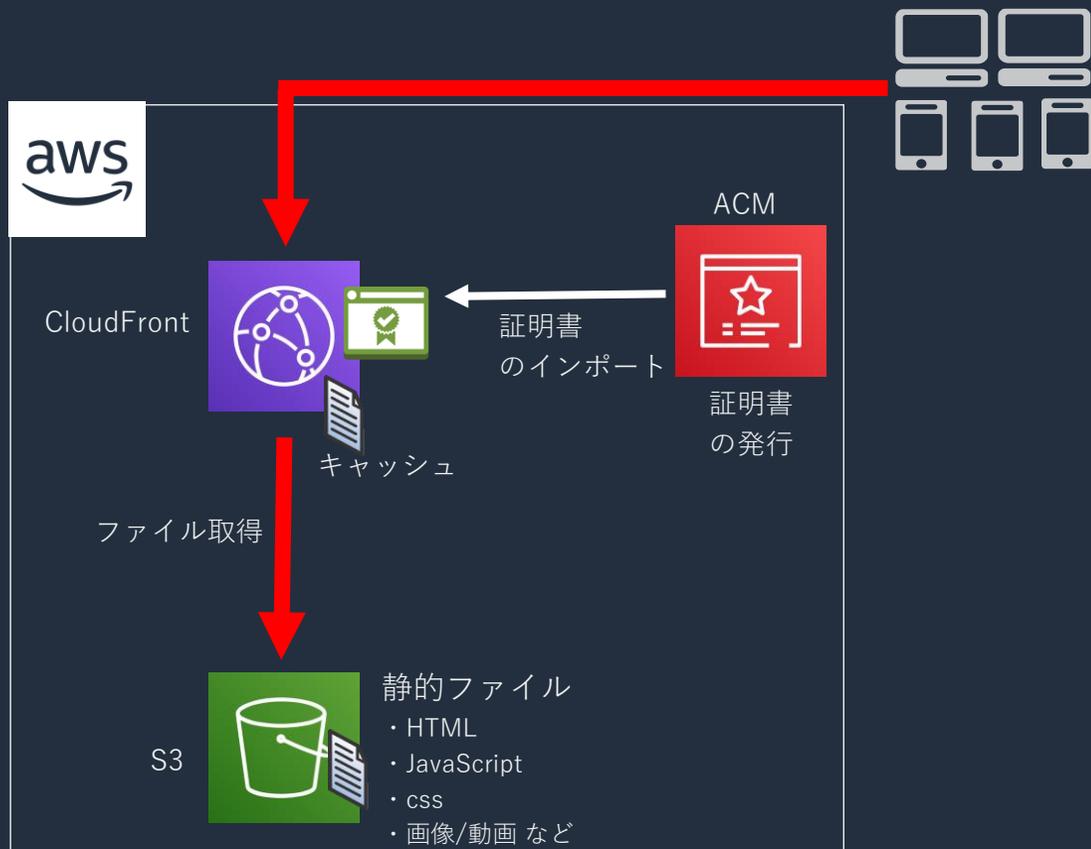
## Amazon S3 Intelligent Tiering の例

※128KB以下のオブジェクトは常にFrequent Access Tierに保存される



## 構成パターン例：

# CloudFront + S3で静的ウェブサイトを仮想サーバを使用せず構築



- S3の静的コンテンツ配信機能を利用し、CloudFrontより配信
- ウェブサイトが静的コンテンツのみ配信している場合に有効な構成
- フルマネージドサービスを利用しており、低コストかつサーバ管理は不要
- ACMを利用することで独自ドメインを簡単にHTTPS化可能

(アクセス例)

- `https://www.example.co.jp/index.html`
- CloudFrontはindex.htmlをS3から取得し、クライアントへ返す
- 2回目以降のアクセスは、CloudFrontがキャッシュから返す

サーバメンテナンス時や非常時の広報で利用などにも





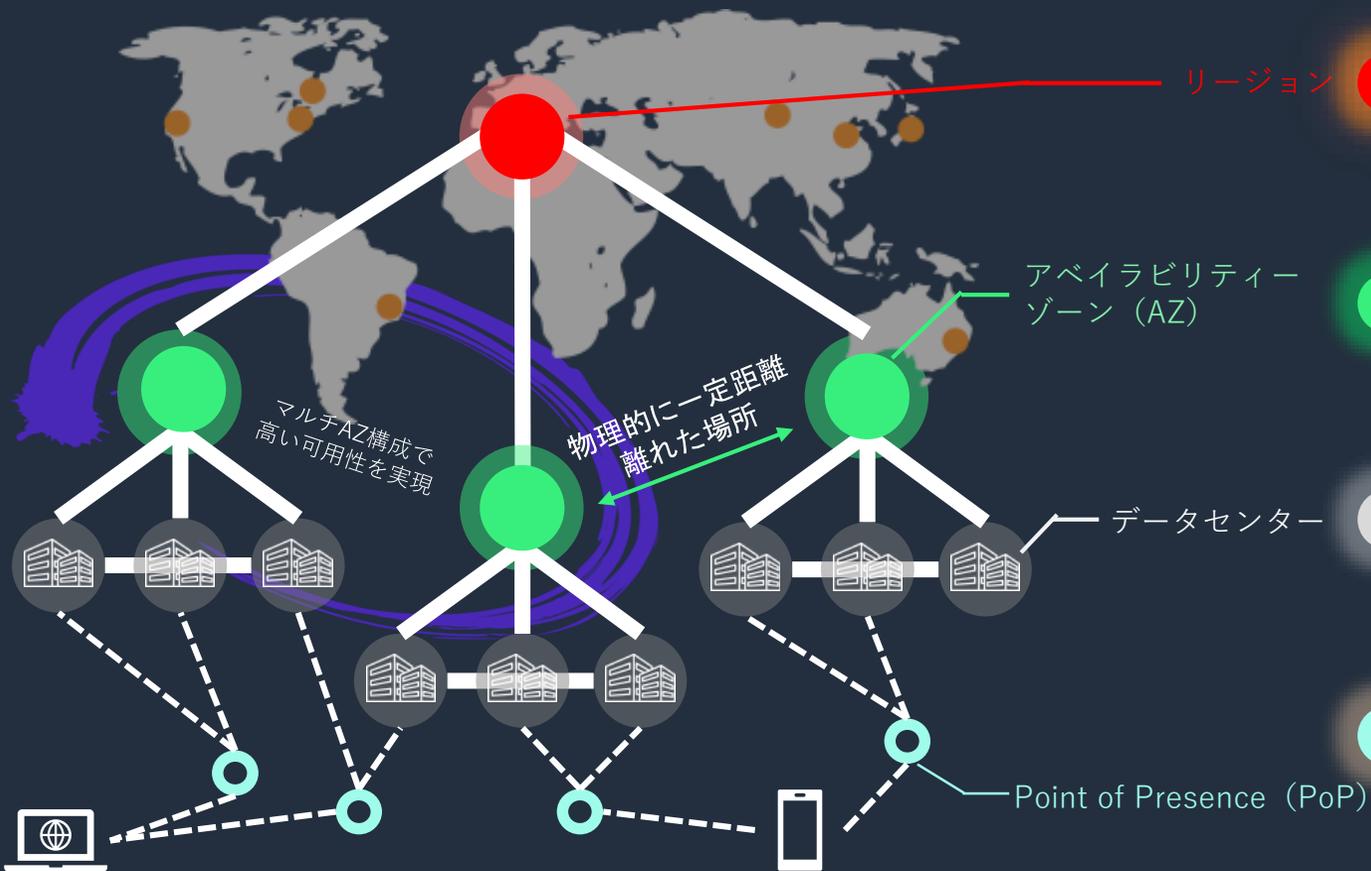
2021年  
大阪リージョン  
3 Availability zones

2011年  
東京リージョン  
4 Availability zones



# 耐障害性と高可用性を実現するインフラストラクチャ

AWSが世界規模で展開するグローバルインフラストラクチャは柔軟性、信頼性、拡張性、そしてセキュリティを兼ね備えています



リージョン



すべてのリージョンは独立し、物理的に離れた場所にある複数のアベイラビリティゾーンで構成されています

アベイラビリティゾーン (AZ)



複数のアベイラビリティゾーンでシステムを構成することで、高い耐障害性を実現できます

データセンター



アベイラビリティゾーンは1つ以上のデータセンターで構成され、互いに低遅延な専用線で接続されています

Point of Presence (PoP)



世界中に配置されたPoint of Presence からの低レイテンシーな高速転送で世界中のユーザーにコンテンツを安全に配信できます

# AWS グローバルクラウドインフラストラクチャ

96 のアベイラビリティゾーン

30 のリージョン

さらに 15+ のアベイラビリティゾーン・5+ リージョンを追加予定

410+ の Point of Presence

21 のローカルゾーン / 29 の AWS Wavelength Zone

\*2022年11月時点の情報



# AWS クラウドセキュリティ

AWSはクラウドコンピューティングの先駆者として、**セキュリティを最優先事項**としてお客様のイノベーションに迅速に対応可能なクラウドインフラストラクチャーを創造してきました。セキュリティ機能の実装や厳格なコンプライアンス要件へ対応でお客様は最も柔軟かつセキュアなクラウドコンピューティング環境を実現可能です

## AWS コンプライアンスプログラム

セキュリティとコンプライアンスのためにAWSに導入されている堅牢な管理は、独立した監査人によって評価されています。これにより、AWSはお客様のコンプライアンス要件への準拠をサポートします

### コンプライアンスプログラムの例



AWS コンプライアンスプログラム  
<https://aws.amazon.com/jp/compliance/programs/>

AWSのデータセンター <https://aws.amazon.com/jp/compliance/data-center/data-centers/>



## クラウドセキュリティのためのサービス

AWSの提供するセキュリティ、ID、コンプライアンスのための包括的なサービスと機能を活用いただくことでセキュリティとコンプライアンスの要件を満たす能力を向上させることができます



アイデンティティ & アクセス管理



脅威の検出と継続的なモニタリング



インフラストラクチャとデータの保護



インシデントへの対応



コンプライアンス

# データプライバシー

AWS はお客様のデータにアクセスできず、**データの保存、管理、保護方法はお客様が選択**します。

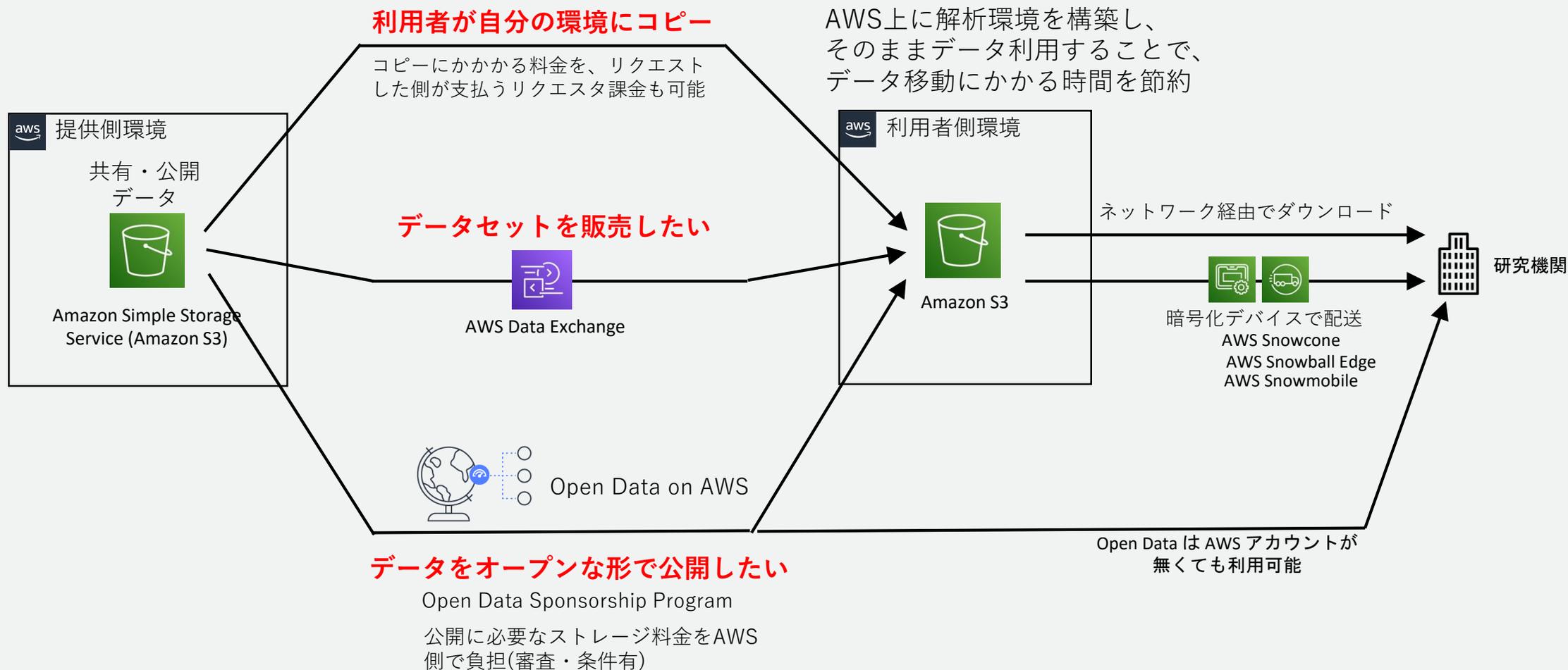
責任共有モデルにおけるデータの所有権と管理に関して、4つの重要な基本事項があります。

1. お客様は継続してデータを所有します。
2. お客様は、データを保存する地理的な場所を選択します。お客様がデータの移動を選択しない限り、データが移動されることはありません。
3. お客様はいつでもデータをダウンロードまたは削除できます。
4. お客様はデータの機密性を考慮し、転送中や保管時にデータを暗号化するかどうか、および暗号化方法を決定します。
  - 東京リージョンや大阪リージョンを選択すれば**データは日本国内**に置かれます
  - 住所が日本の場合、基本的に第一審裁判所は**東京地方裁判所**となります



# 大規模データの共有・公開

## Amazon S3を利用した共有・公開の例



## Open Data on AWS

- 様々なデータが公開されている
- 自由にダウンロード、利用可能
- AWS上で利用することにより、データ移動にかかる時間を減らして利用可能

The screenshot shows the 'Registry of Open Data on AWS' website. The header includes the AWS logo. The main content is divided into two columns. The left column has an 'About' section explaining the registry's purpose and a search bar for datasets (currently 320 matching). The right column features 'The Cancer Genome Atlas' dataset, with tags for 'cancer', 'genomic', 'life sciences', 'STRIDES', and 'whole genome sequencing'. It includes a detailed description of the TCGA project and a list of usage examples.

Open Data Sponsorship Program により、オープンデータのストレージコストをカバー可能です。(審査有)

<https://aws.amazon.com/jp/opendata/open-data-sponsorship-program/>

# AWS Data Exchange (ADX)

- 数百万のAWSユーザー向けに自社が持つデータを提供できるプラットフォーム
- 2019年のローンチ以降プロバイダー数は急増、200以上の**企業による**3,600以上の**データ**
- S3に保存できるテキスト、動画、画像、音声などフォーマットは様々
- COVID19、ESG関係の**データ**も多数
- サンプルデータ提供も可能（その場合費用はかからない）
- Private Products機能により**特定の企業・個人だけにデータ**を提供することも可能

The screenshot shows the AWS Data Exchange website. At the top, there is the AWS logo and navigation links for Products, Solutions, Pricing, Documentation, Learn, Partner Network, AWS Marketplace, Customer Enablement, Events, and Explore More. A 'Sign In to the Console' button is visible in the top right. Below the navigation, there is a search bar and a list of categories: AWS Data Exchange, Overview, For Providers, Features, Pricing, FAQs, and COVID-19. The main heading is 'AWS Data Exchange' with the subtext 'Easily find and subscribe to third-party data in the cloud'. There are two buttons: 'Explore available data products' and 'Learn about becoming a data provider'. The main content area contains text explaining that AWS Data Exchange makes it easy to find, subscribe to, and use third-party data in the cloud. It lists examples of data providers like Reuters, Change Healthcare, Dun & Bradstreet, and Foursquare. A video player is embedded on the right side, titled 'Introduction to AWS Data Exchange (3:02)'. The video player shows a play button and a question mark icon.

# Global Data Egress Waiver について

- 研究機関や学術機関のお客様を対象に、**データ out 転送料金**についてAWS利用料総額の最大15%まで免除する**割引プログラム**です。
- Payer Accountに対して適用されます。
- 中国リージョンを除く、すべてのAWSリージョンでの利用が適用対象です。
- AWS Region間（e.g 東京リージョンから大阪リージョン）の通信費用、CloudFront, Direct Connect (per-Gbps) port speed fees, VPNの費用は本プログラム適用対象外です。
- データ転送料金が月額利用総額15%を超えた場合、超えた分のみ通常のデータ転送費用が適用されます。（一般的なデータ転送費用の割合は月額費用の3~5%）

## 適用条件

1. 教育や国立研究所などの学術研究機関に所属していること。
2. SINETもしくはAWS Direct Connectを使用していること。
3. すべてのメンバーアカウントは、お客様またはお客様の従業員によって開設された、お客様またはお客様の従業員が使用するための機関用電子メールアドレスを使用していること。（例：個人用のgmailアドレスではなく、ac.jpアドレスなど）。AWSパブリックセクターが認定した Reseller 経由で発行したアカウントも適用対象になります。

<https://aws.amazon.com/jp/blogs/publicsector/data-egress-waiver-available-for-eligible-researchers-and-institutions/>



# AWS Cloud Credit for Research

aws

Contact Us Support English My Account

Products Solutions Pricing Documentation Learn Partner Network AWS Marketplace Customer Enablement Events Explore More Q

## AWS Cloud Credit for Research

Access to technology that accelerates innovation

Apply now

[AWS Cloud Credit for Research](#) · [FAQs](#) · [Terms and Conditions](#) · [Previous Recipients](#) · [Resources](#)

The AWS Public Sector Cloud Credit for Research Program supports researchers who seek to:

- Build cloud-hosted publicly available science-as-a-service applications, software, or tools to facilitate their future research and the research of their community
- Perform proof of concept or benchmark tests evaluating the efficacy of moving research workloads or open data sets to the cloud
- Train a broader community on the use of cloud for research workloads via workshops or tutorials

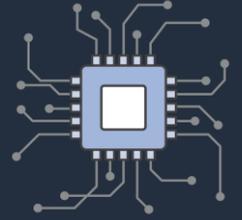
aws cloud credit for research

<https://aws.amazon.com/jp/government-education/research-and-technical-computing/cloud-credit-for-research/>

# 研究助成プログラム

- ✓クラウドでホストされるサービス、ソフトウェア、ツールの構築、研究プロセスのクラウド移行等の新しいプロジェクトをサポートするプログラム。
- ✓申請は四半期ごとに審査されます。
- ✓申請金額に上限はありません。
- ✓研究環境のオンプレミスからクラウドへの移行の検証
- ✓研究プロジェクトのメンバーのトレーニング
- ✓研究の公開のための基盤構築などの目的で応募可能。

# 研究でAWSの活用



スモールスタート、常に新しいリソースの利用、繁忙期にリソース増で研究を加速

- ✓ ITリソースは必要なときに起動、不要になったら削除
  - 研究立ち上げにスモールスタート、論文投稿や卒・修士論文研究時等の繁忙期にAWSの潤沢なリソースを活用
  - HPC環境はキュー待ちを無くした、自分だけのHPC環境
  - 潤沢なリソースを利用し、並列実行で短時間に。コストは同じ。
- ✓ 常に新しいGPU/アクセラレータ/CPU、FPGA等を利用可能して計算、高価なリソースも時間で利用
- ✓ マネージドサービス活用で、研究環境立ち上げの効率化
  - 手元のPC環境を壊すこと無く研究、研究環境の保全、再現
- ✓ 可視化もクラウドで行い、データを持ち出さず、データ移動にかかる時間も節約
- ✓ 機関外からリソースにアクセスして研究を継続
- ✓ 研究成果の公開、研究証跡管理での利用 (Amazon S3、CloudFrontの活用等)

AWSを活用し、リソース準備にかかる手間、時間を減らし、研究の本質に集中

# まとめ

- AWS の各種ストレージサービスを適材適所で活用
  - データを安全に保管・長期保管・大容量データ保管
  - 高耐久、3ヶ所以上に自動で隔地保存され、長期サービス提供し続けている **Amazon S3 の活用**、**学認RDMに連携可能**
  - 各ストレージ階層の活用してコスト効率良く利用
- 認証基盤等も **安定したAWS基盤上で**
- AWS上でのデータ共有や解析を行うことで、移動にかかる手間や時間を短縮
  - データセットの販売、データの公開のサービスの活用
  - オープンデータ公開には **AWS Open Data Sponsorship Program** の活用等
- 研究助成やデータ転送料低減プログラムなど、各種制度を活用



学術・研究機関向け問い合わせ先  
[aws-jpps-er@amazon.com](mailto:aws-jpps-er@amazon.com)



