



日本におけるヒトゲノム情報や臨床情報の意義と 安全・安心な利用にむけた量子技術の活用について

京都大学 学際融合教育研究推進センター
スーパーグローバルコース 医学生命系ユニット
特定教授 長崎 正朗

量子技術の実用化推進ワーキンググループ（第二回）
令和4年（2022）/11/24 10:00-12:00 ハイブリッド開催



略歴



1998年 東京大学大学院理学系研究科 情報科学専攻

東京大学 医科学研究所 ヒトゲノム解析センター DNA情報解析分野 (宮野 悟 研究室)

2004年 東京大学大学院理学系研究科情報科学専攻 博士課程修了 学位 博士 (理学)

専門 バイオインフォマティクス、メディカルインフォマティクス、システム生物学、大規模情報解析

職歴・研究歴

2004年- 東京大学 医科学研究所ヒトゲノム解析センター 科学技術振興調整費特任研究員(2004-) 助手(2005-)、助教(2007-)、准教授(2011-)

システム生物学、ヒトゲノム情報解析とがんゲノム解析

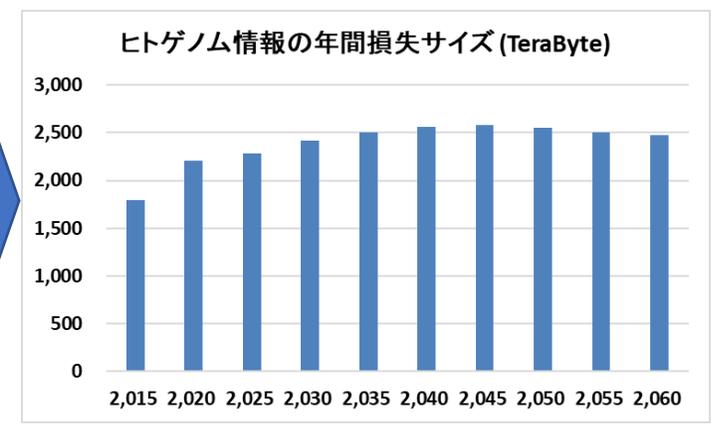
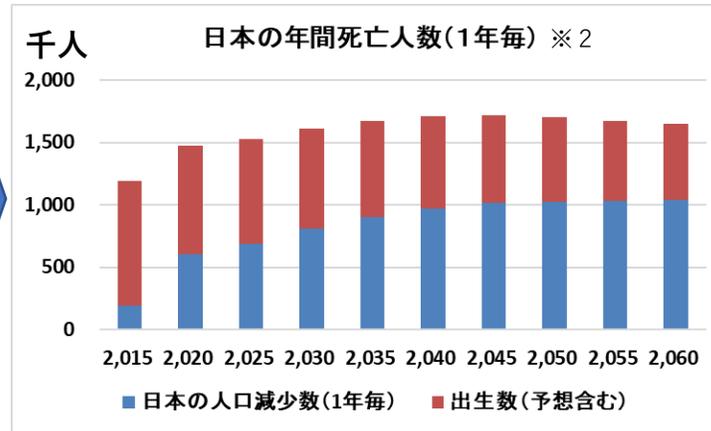
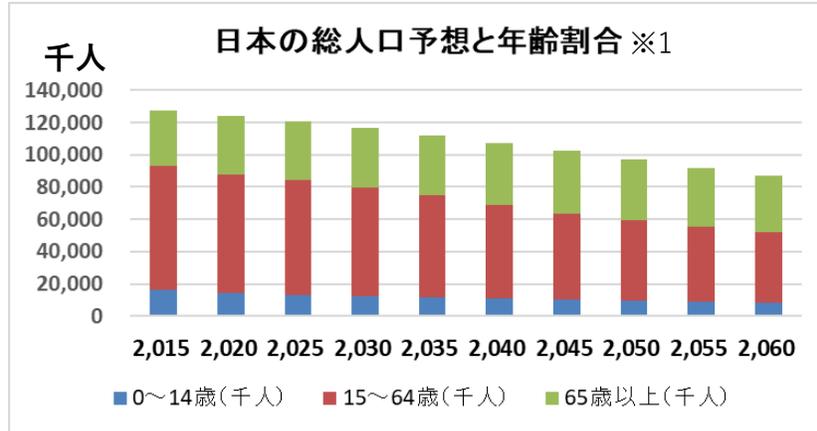
2012年- 東北大学 東北メディカル・カンパニオン ゲノム解析部門 バイオメディカル情報解析分野 教授

ヒトゲノム大規模情報解析 ヒトゲノムの数千検体の情報解析と疾患遺伝要因探索

2019年- 京都大学 学際融合教育研究推進センター スーパーグローバルコース 医学生命系ユニット (ゲノム医学センター併任) 特定教授

ヒトゲノム大規模情報解析 ヒトゲノムの1万検体以上の情報解析と疾患遺伝要因探索

はじめに 日本とヒトゲノムを取り巻く環境



年間 数百万人の日本人の**全ゲノム情報が焼失**※3 `rm -rf *`

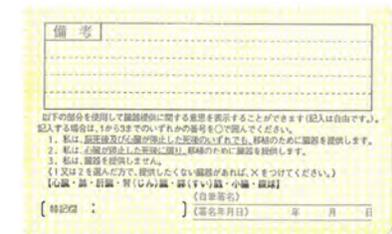
全ゲノム情報・医療情報とは？

「**国家において将来の医療に役立てることができる情報資源**」

どのような疾患にかかったか？どのくらい生存することができたか？検査値情報の履歴は貴重な情報

ヒト1名のゲノム情報
約30億塩基対 (A/T/G/C)
= 1.5GigaByte※4

外国に先んじて、超高齢化社会を迎える日本は、ゲノム情報や医療情報を活用することで、**疾患の予防や治療のための戦略立案を先行し、輸出産業を育成できるチャンスが到来**



どのようなセキュリティ技術によって、ゲノム情報を安心・安全を担保しつつ、情報解析資源として解析を行い実社会に活用するか？

臓器提供の意思の他にゲノム情報や臨床情報の提供などの意思表示を**本人の生前の同意**ができる環境を整備することで、**健康福祉に役立つ** (個人的意見)

ヒトゲノムの計測機器・情報解析の進展

計測機器の進展

2003

国際的な1つの
ヒト参照配列の確定

数十か国の国際プロ
ジェクト 数千億



2017-22

1台で6000人
1人あたり **\$800**



2023

Illumina社
NovaSeq X Plus

1台で年間 2万人
1人あたり **\$200**

参考：UK Biobank 2022 **15万人の解析完了済** 2024 **50万人完了予定**

1人あたり30Gbyte
ぐらい。30Pb^{※1}?

50台あれば100万人の
情報取得を/年で可能

情報解析の進展

CPUでの実装と改良



1人あたり約1日^{※2}

GPUでの実装と改良



1人あたり約1時間^{※2}

FPGAでの実装と改良



1人あたり約30分^{※2}

解析手法や実装は経
時的に改良されてい
る（要定期的再解析）

今までは、シーケンスが国産ではないので輸入的な要素が多かったが
2022年末で情報取得コストがかなり下がった。 **今後は全ゲノム情報活用が主戦場**



どのようなセキュリティ技術によって、ゲノム情報を安心・安全を
担保しつつ、情報解析資源として解析を行い実社会に活用するか？

[1] 計測・保管 → [2] 情報解析 → [3] 診断

ヒトゲノム情報を安全に管理・解析するための量子技術・秘密計算関係の長崎の主な活動

① 量子技術を用いたヒトゲノム配列情報の安全な転送技術の実証 [1]計測・保管

株式会社 東芝との共同研究

- | | | |
|---|-----------------------------|--------------------------|
| 2016 ヒトSNPアレイの情報の転送試験 | (7kmの既設光ファイバーでの実証) | (東北大学の長崎との共同研究・プレスリリース他) |
| 2018 ヒト全ゲノム情報の転送試験 | (平均10Mbpsを超える鍵配信速度での量子暗号通信) | (東北大学の長崎との共同研究・プレスリリース他) |
| 2020 ワンタイムパッドを用いた全ゲノム配列データのリアルタイム伝送 (※1) | | ※1 東北大学の長崎異動後の成果 |
| 2021 ワンタイムパッド/秘密分散によるゲノム情報によるバックアップの試験実装 (※2) | | ※2 東北大学の長崎異動後の成果、NICTも参画 |



詳述

② 大規模なヒトゲノム情報を秘匿したまま安全に統計解析を行う実時間で処理可能な秘密計算によるアルゴリズムと実装 [2]解析

NTTセキュアプラットフォーム研究所 (高橋克巳先生、千田浩司先生他) との共同研究

- 2017 1000人規模の秘密計算に基づく fisher検体手法の提案と実装 (Hamada *et al*/GenoPri 2017 国際会議 (査読付き))

③ 国内における研究機関におけるヒトゲノム解析の情報管理、再現性を担保した情報解析の在り方、国内外のヒトゲノム情報の倫理

国立遺伝学研究所、国立国際医療研究センターとの共同研究

- 2021 Practical guide for managing large-scale human genome data in research, Tanjo *et al* *Journal of Human Genetics* (review)

④ 量子技術を組み合わせたヒトゲノム配列の超高速情報解析 [1]保管 [2]解析 及び

医療機関における診療・診断を想定した最小限の個人ゲノム情報提供技術の枠組みの理論構築と試験実装 [3]診断

国立研究開発法人情報通信研究機構 (NICT) 藤原幹生先生、佐々木雅英先生、株式会社 東芝、株式会社 ZenmuTechとの共同研究

- 2022 Secure secondary utilization system of genomic data using quantum secure cloud, Fujiwara *et al* *Scientific Reports* 12:18530.

量子セキュアクラウド、one time pad、trusted node



詳述

[1] ヒトゲノム計測・保管量子基盤

① 量子技術を用いたヒトゲノム配列情報の安全な転送技術の実証 [1]計測・保管

株式会社 東芝との共同研究

2016 ヒトSNPアレイの情報の転送試験

(7kmの既設光ファイバーでの実証)

2018 ヒト全ゲノム情報の転送試験

(平均10Mbpsを超える鍵配信速度での量子暗号通信)

2020 ワンタイムパッドを用いた全ゲノム配列データのリアルタイム伝送 (※1)

2021 ワンタイムパッド/秘密分散によるゲノム情報によるバックアップの試験実装 (※2)

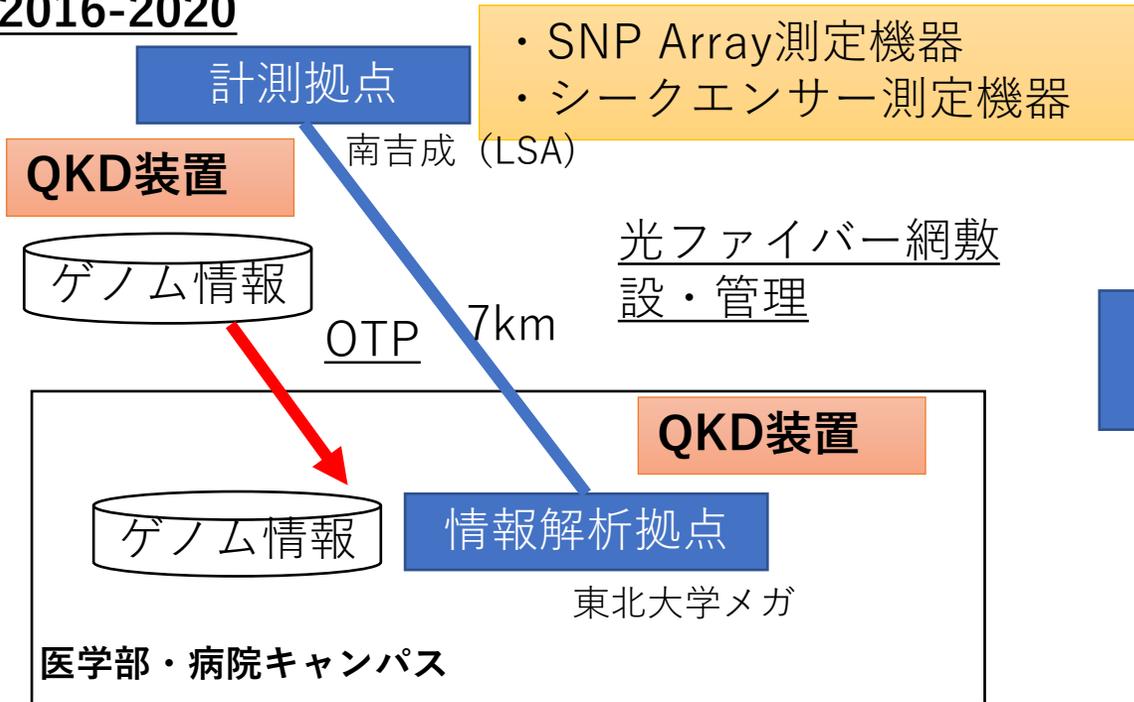
(東北大学の長崎との共同研究・プレスリリース他)

(東北大学の長崎との共同研究・プレスリリース他)

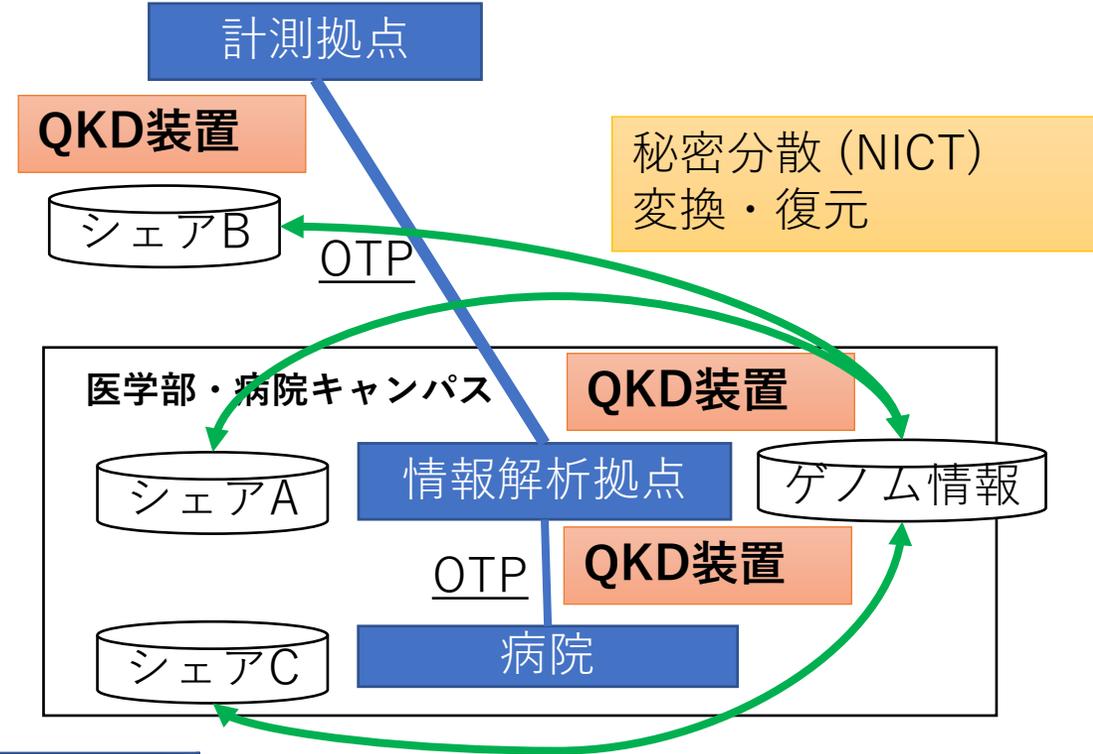
※1 東北大学の長崎異動後の成果

※2 東北大学の長崎異動後の成果、NICTも参画

2016-2020



2021-



QKD量子技術に基づく、ヒトゲノム計測・保管についての理論実装と試験運用は完了

課題：他拠点での実装・運用の展開、多拠点での運用など

[2] ヒトゲノム情報解析量子基盤

④ 量子技術を組み合わせたヒトゲノム配列の超高速情報解析 [1]保管 [2]解析 及び

医療機関における診療・診断を想定した最小限の個人ゲノム情報提供技術の枠組みの理論構築と試験実装 [3]活用

国立研究開発法人情報通信研究機構 (NICT) 藤原幹生先生、佐々木雅英先生、株式会社 東芝、株式会社 ZenmuTechとの共同研究

2022 Secure secondary utilization system of genomic data using quantum secure cloud, Fujiwara *et al* *Scientific Reports* 12:18530.

ヒト全ゲノム情報※

量子 + 秘密分散

シェアA

QKD装置

シェアB

QKD装置

シェアC

QKD装置

NEC/東芝/NTT等

東京QKDネットワークの活用

秘密分散情報復元

シェアC

シェアA

シェアB

データオーナー

trusted node内でのヒトゲノム情報解析

FPGAが最速 (1検体30分) なのでその最速解析速度を律速としない量子クラウドからの**復元と情報解析を達成**

※約30Gb (gzip圧縮時)



QKD量子技術に基づく、ヒトゲノム[1]保管・[2]情報解析についての理論実装と試験運用は完了

課題：ゲノム情報クラウドストレージサイズの拡大、試験以外での拠点実装、実運用ポリシーの策定

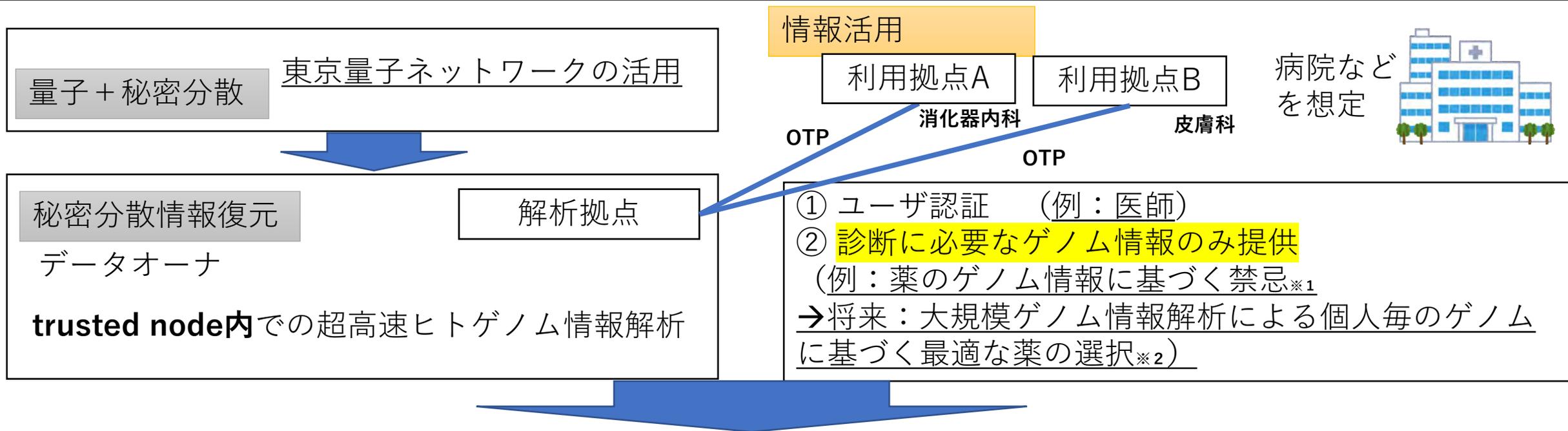
[3] ヒトゲノム診断量子基盤

④ 量子技術を組み合わせたヒトゲノム配列の超高速情報解析 [1]保管 [2]解析 及び

医療機関における診療・診断を想定した最小限の個人ゲノム情報提供技術の枠組みの理論構築と試験実装 [3]診断

国立研究開発法人情報通信研究機構 (NICT) 藤原幹生先生、佐々木雅英先生、株式会社 東芝、株式会社 ZenmuTechとの共同研究

2022 Secure secondary utilization system of genomic data using quantum secure cloud, Fujiwara *et al* *Scientific Reports* 12:18530.



QKD量子技術に基づく、ヒトゲノムの[3]診断などの利活用についての理論実装と試験運用は完了

課題：実環境での導入、実運用ポリシーの策定、臨床での実活用

※1 消化器内科：チオプリン製剤のNUDT15遺伝子の多型に基づく診断など

※2 なぜゲノム情報と医療情報を集積するかの理由

国内のゲノム情報の現在の公的データベースの保管・解析の状況と提案

NBDCヒトデータベースについて

ヒトに関するデータは、次世代シーケンサーをはじめとした解析技術の発達に伴って膨大な量が産生されつつあり、それらを整理・格納して、生命科学の進展のために有効に活用するためのルールや仕組みが必要です。

国立研究開発法人科学技術振興機構(JST)NBDC事業推進部(NBDC)では、個人情報の保護に配慮しつつヒトに関するデータの共有や利用を推進するために、ヒトに関する様々なデータを共有するためのプラットフォーム『NBDCヒトデータベース』を設立するとともに、国立遺伝学研究所 DNA Data Bank of Japan (DDBJ)と協力して、ヒトに関するデータを公開しています。

本Webサイトを通じて、ヒトに関するデータの利用及びヒトに関するデータの提供を行なうことができます。

なお、本データベースの目的・意義、扱うデータの種類、データ利用者の範囲、責任者については[こちら](#)をご覧ください。

NBDCヒトデータ取扱いセキュリティガイドライン (データ利用者向け) ver. 6.0

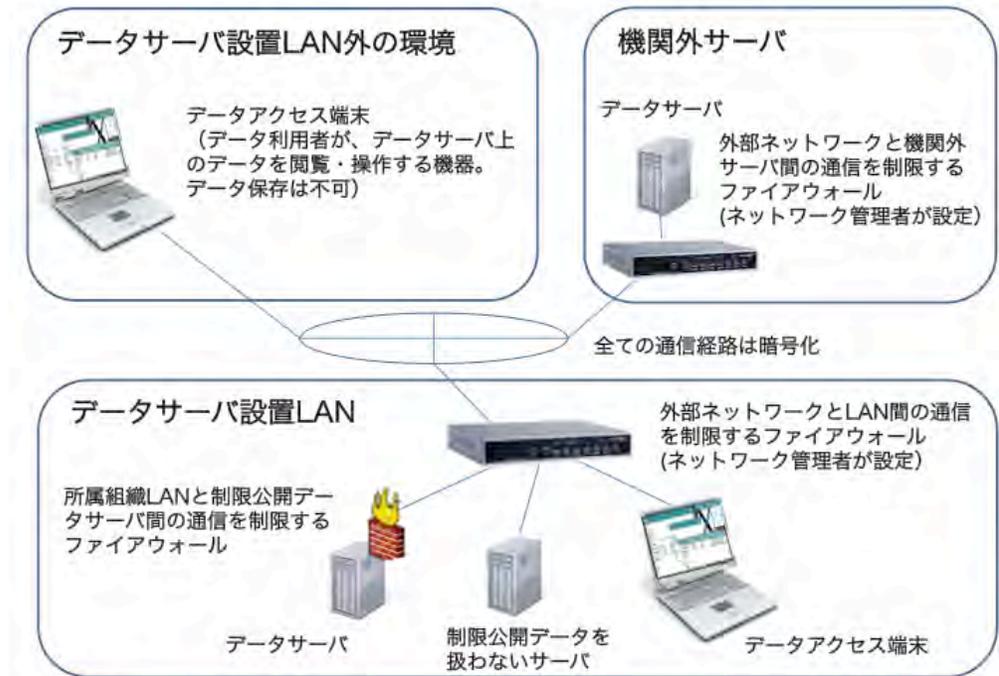


図1 データサーバ設置LAN、機関外サーバ、データサーバ、データアクセス端末

提案

量子系の技術を使うことで現在の研究者の解析アクティビティをさげないで、より安全に【情報管理】
【情報解析】を行える次世代のシステム(trusted serverの概念導入・量子を活用した認証等)のリファレンス構築やガイドラインを設定できないか？

国内の臨床情報の公的データベースの情報登録と保管の状況と提案

RADDAR-J <https://www.raddarj.org/>

統括的情報共有と連携。
難病の研究と治療方法の開発を推進します。

難病プラットフォームは、日本医療研究開発機構（AMED）および厚生労働省の難病研究班が収集した臨床情報や生体試料から得られた情報を集約する情報統合基盤で、AMEDの公的データベースです。集約したデータのシェアリングや二次利用を通じて、わが国の難病研究を推進させることを目指します。
難病プラットフォームは、AMEDの難治性疾患実用化研究事業の研究費により運営されています。

[日本医療研究開発機構（AMED）の「難治性疾患実用化研究事業」のページはこちら](#)



— 企業関係者・患者さんや研究者の方へ

- [レジストリ・バイオレポジトリ](#)一覧年度毎の研究課題・レジストリ・バイオレポジトリ等の情報を検索・閲覧できます。
- [企業マッチング](#)について研究班と企業（製薬企業など）との間で共同研究促進のための場を提供します。

— 難病研究班の方へ

- [各種支援](#)レジストリの新規構築や運営について、難病プラットフォームが行っている支援のご案内です。
- [個別相談](#)難病プラットフォームの活用方法に関する相談会を定期的実施しています。詳細は以下の連絡先へお問い合わせ下さい。
info@raddarj.org

AMEDの公的データベースの1つ
日本の難病研究の情報臨床情報の集約統合基盤システム

SSL/TLSサーバ証明書・クライアント証明書・分散暗号化データベース

医療情報システムの安全管理に関する
ガイドライン 第5.1版 令和3年1月
厚生労働省

また、IPsec を用いた VPN 接続等によるセキュリティの担保を行わず、インターネット等のオープンなネットワークを介し、他の医療機関や患者等が医療情報システムへ接続する場合（図 B-2-④）は、少なくとも TLS による暗号化を用いた HTTPS の利用が求められる。

さらに、オープンネットワークで接続する場合には、IPsec や TLS によるセッションが安全でも、他セッションが同居できるため、ネットワークに接続している機器やシステムが標的型メール等の攻撃にさらされるリスクがある。仮に、このような攻撃によってネットワークに接続する端末等がコンピュータウイルスに感染し、遠隔操作が可能になると、IPsec や TLS1.2 以上によるセッションへの正規のアクセスが発生し得る。

よって、IPsec や TLS を採用する場合でも、その端末にオープンネットワークに対する開放されたポートがある場合には、セッション間の回り込み（正規のルートではないクローズドセッションへのアクセス）等による攻撃からの防護について、適切な対策を実施する必要がある。

提案

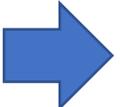
現在の臨床研究の先生方のアクセシビリティ※1を維持しつつ、量子系の技術を使うことでより、強固な次世代の情報集約のシステムを構築できないか？

個別もつことは困難なので、**認証の公共基盤として提供するような提供が必要？**

※1 臨床の傍ら自助努力で時間を確保して登録

論点

1. 伝送距離
2. 扱えるデータサイズ
3. 量子技術のプロトコルの規格化
4. 量子の運用技術担当者の育成
5. 導入・維持管理・教育コスト

 **学術系、6NCに関連した提案**

民間クラウドの進出 → 計算資源の維持管理の意義



量子技術の導入による、学術機関における研究開発情報資源のより**安全・安心な運用・管理の実践運用としての差別化**

- 1. 伝送距離
(県内)
- 2. 扱えるデータサイズ
(数TB-10TB)
- 3. 量子技術のプロトコルの規格化
(複数)
- 4. 量子の運用技術担当者の育成
(未成熟)

① 同一大学の場合には**各都道府県での距離の範囲に収まる**

② **同一組織なので運用が簡便**

③ 通常は何等かの大型計算機の技術に明るいメンバでの運用が行われているので**量子などの新規技術の親和性が高い**

④ 大型計算機を**Trusted Node**として扱いログインの認証に量子暗号の技術を使うことで学術機関におけるより安全な情報管理を担保

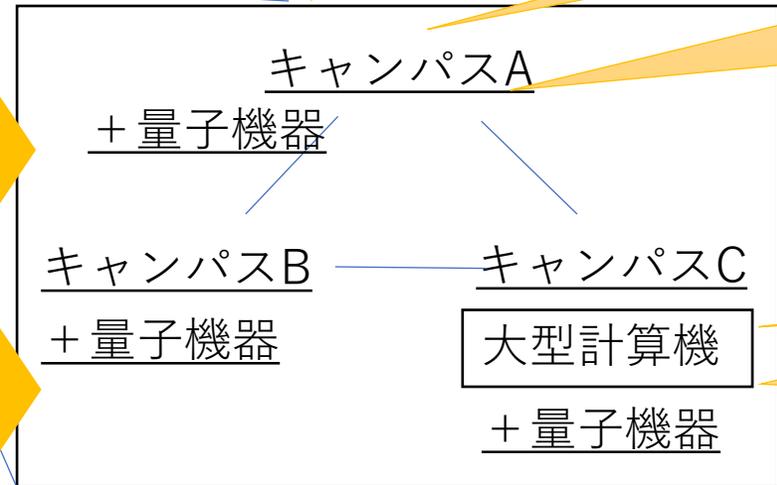
⑤ ログイン認証だけではなく、一部の**量子セキュアクラウドによる情報保全**



- ★ 産業界からの開発機器導入
- 運用支援
- セキュリティ投資20%ルール
- インセンティブ

産業界

政府



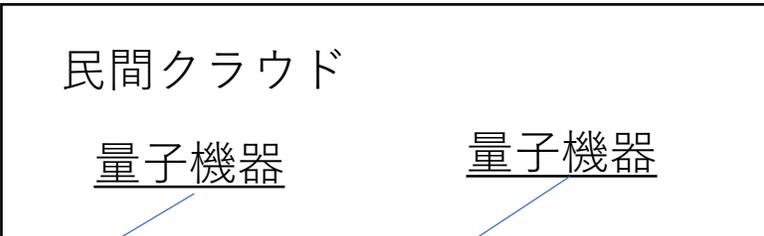
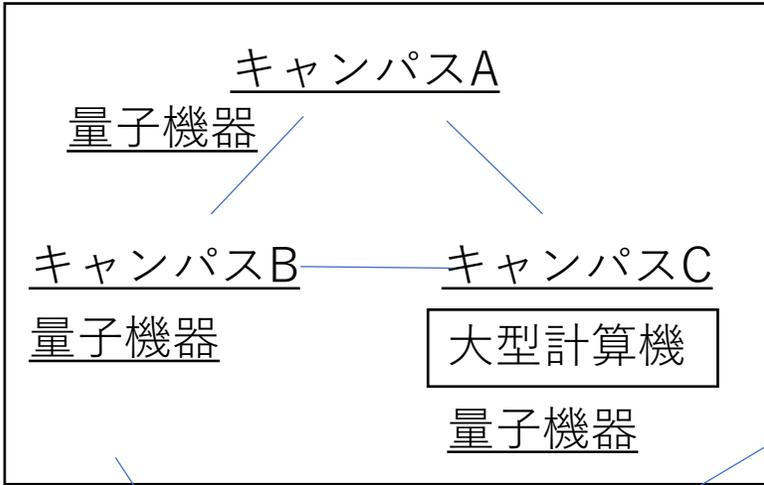
【達成目標】 量子技術担当者の育成、量子クラウドの運用ポリシーの策定、通信プロトコルの規格化

参考

提案

phase2

- 1. 伝送距離 (100km-400km)
- 2. 扱えるデータサイズ (数百TB-数PB)
- 3. 量子技術のプロトコルの規格化 (第一世代プロトコル確定・次世代プロトコル策定中)
- 4. 量子の運用技術担当者の育成 (基礎人材育成済→高度人材へ)



① 技術の進展に伴う **近隣都道府県間での実装と運用**

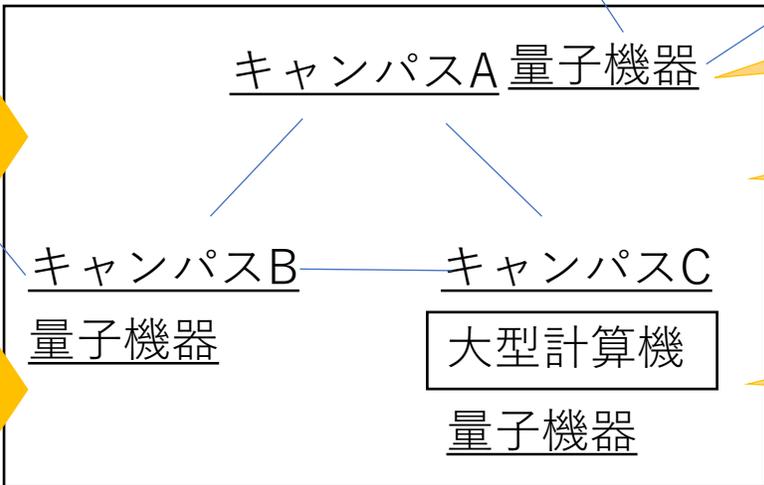
② 単一組織から **複数組織での運用ポリシーの策定と実運用**

③ 通常は何等かの大型計算機の技術に明るいメンバー間での運用実績の共有による **技術レベルの均てん化**

④ 冗長構成による認証における多重化による **より安定な運用**の担保できる。

⑤ ログイン認証だけではなく、**より大規模な量子セキュアクラウドによる情報保全**

SINET併設網など



産業界からの
開発機器導入
運用支援

セキュリティ投資20%ルール

インセンティブ

【達成目標】 高度量子技術担当者育成、量子クラウドの運用ポリシーの策定、次世代通信プロトコルの検証と規格化

phase3 長期

1. 伝送距離
(1000km以上)
2. 扱えるデータサイズ
(100PB以上)
3. 量子技術のプロトコルの規格化
(第2世代プロトコル確定・
第3世代プロトコル策定中)
4. 量子の運用技術担当者の育成
(基礎・高度人材育成済)



① 技術の進展に伴う **全道府県間での実装と運用**

② 地域レベルから **全国レベルでの運用ポリシーの策定と実運用**

③ 確立した教育プログラムによる **技術レベルの完全均てん化**

④ 単純な冗長化ではない効率のよい **より安定な運用の確立**

⑤ **量子セキュアクラウドによる非秘匿要件以外の全情報保全**