\Orchestrating a brighter world NEC

NECと大阪大学による
秘密計算のゲノム解析システムへの適用実証

2019年7月

NEC セキュリティ研究所

Orchestrating a brighter world

未来に向かい、人が生きる、豊かに生きるために欠かせないもの。
それは「安全」「安心」「効率」「公平」という価値が実現された社会です。

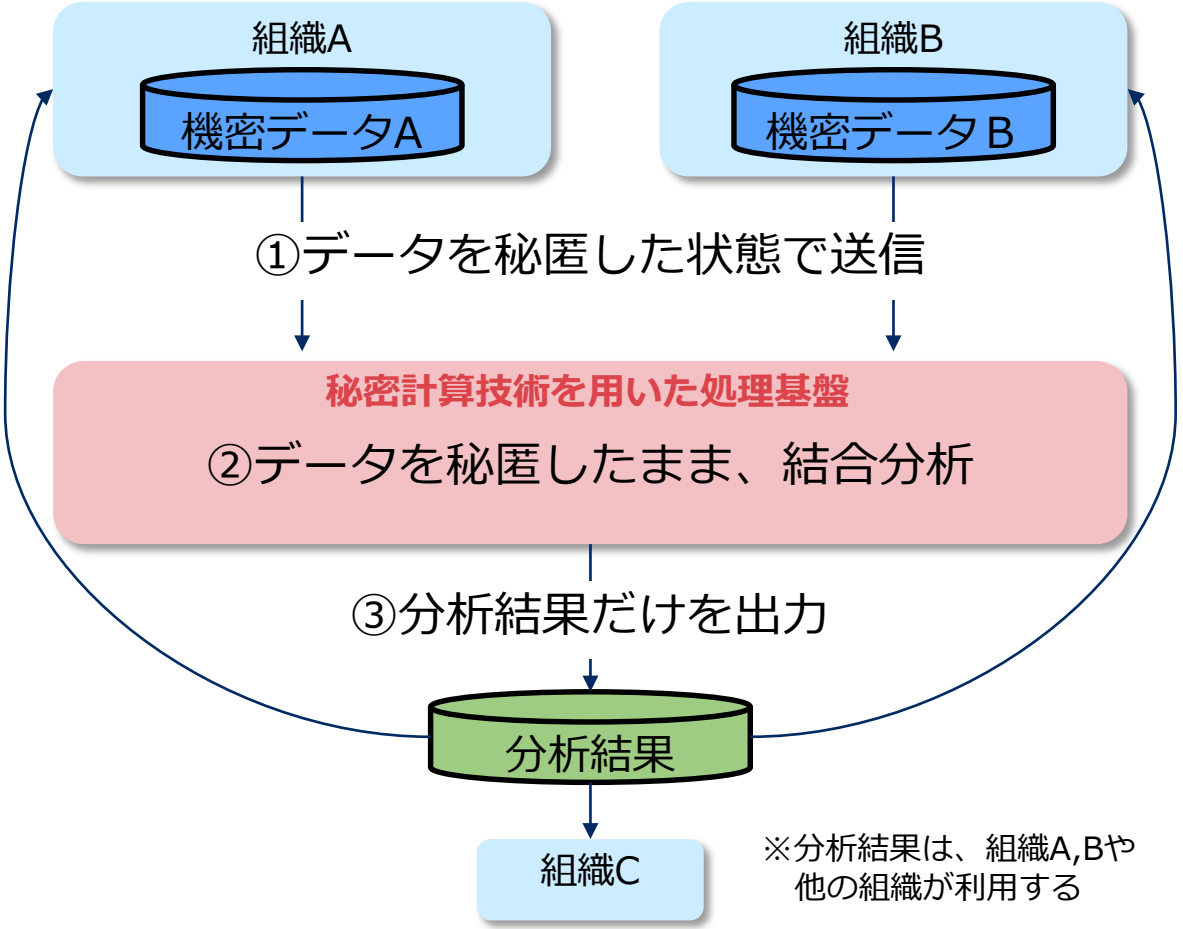
NECは、ネットワーク技術とコンピューティング技術をあわせ持つ
類のないインテグレーターとしてリーダーシップを発揮し、
卓越した技術とさまざまな知見やアイデアを融合することで、
世界の国々や地域の人々と協奏しながら、
明るく希望に満ちた暮らしと社会を実現し、未来につなげていきます。

目次

１．秘密計算とは

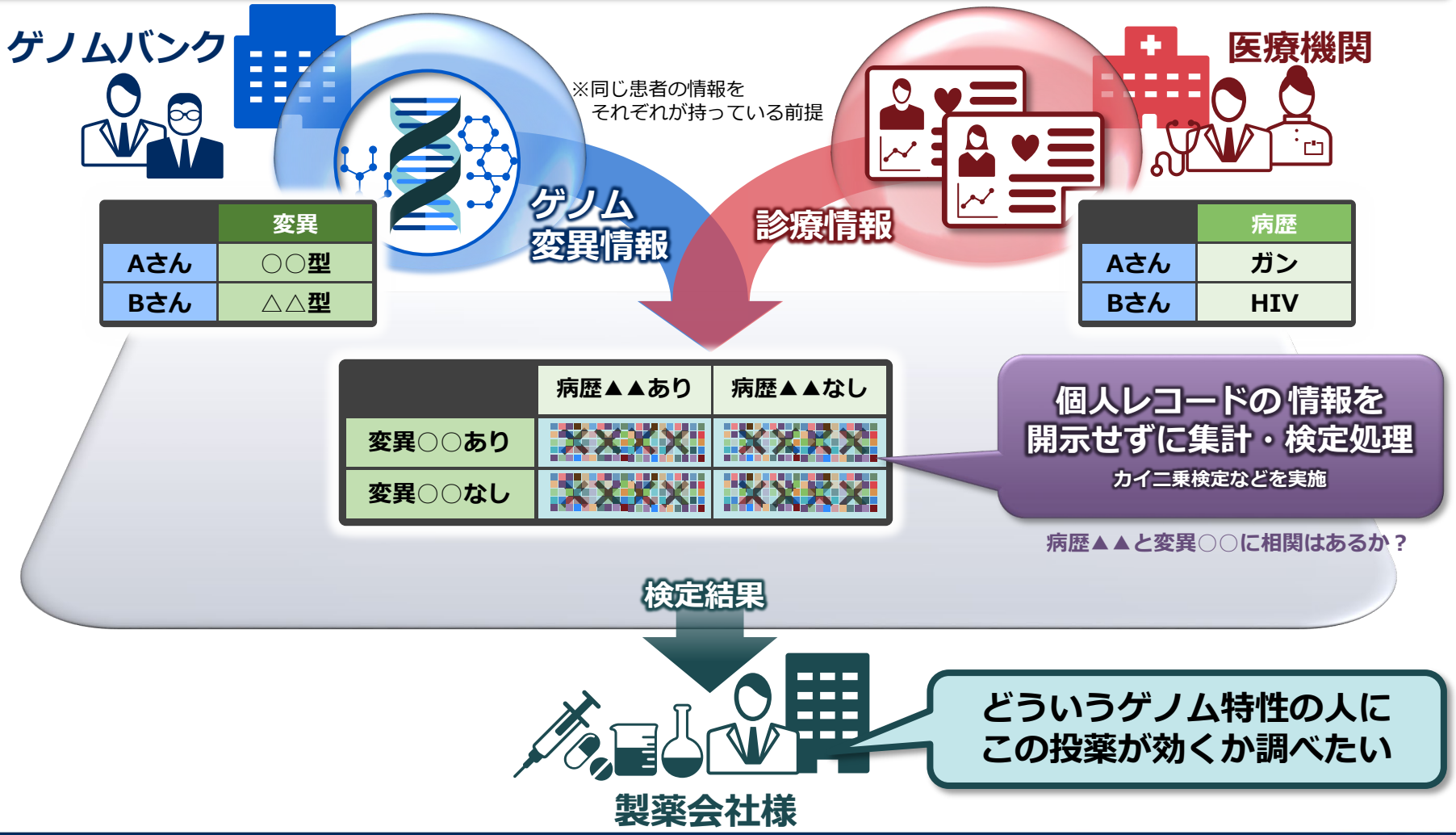
秘密計算の概要

- 秘密計算とは、データを秘匿したまま処理できる技術
- 異なる組織のデータを、組織外に元データを一切開示せずに、データを結合した分析が可能



ユースケース例：ゲノム解析

秘密計算によって、ゲノム情報や診療情報を安全に結合分析し、ゲノム特性に応じた薬の開発やオーダーメイド医療を促進



2．実証実験内容

実証目的と検証項目

実証目的

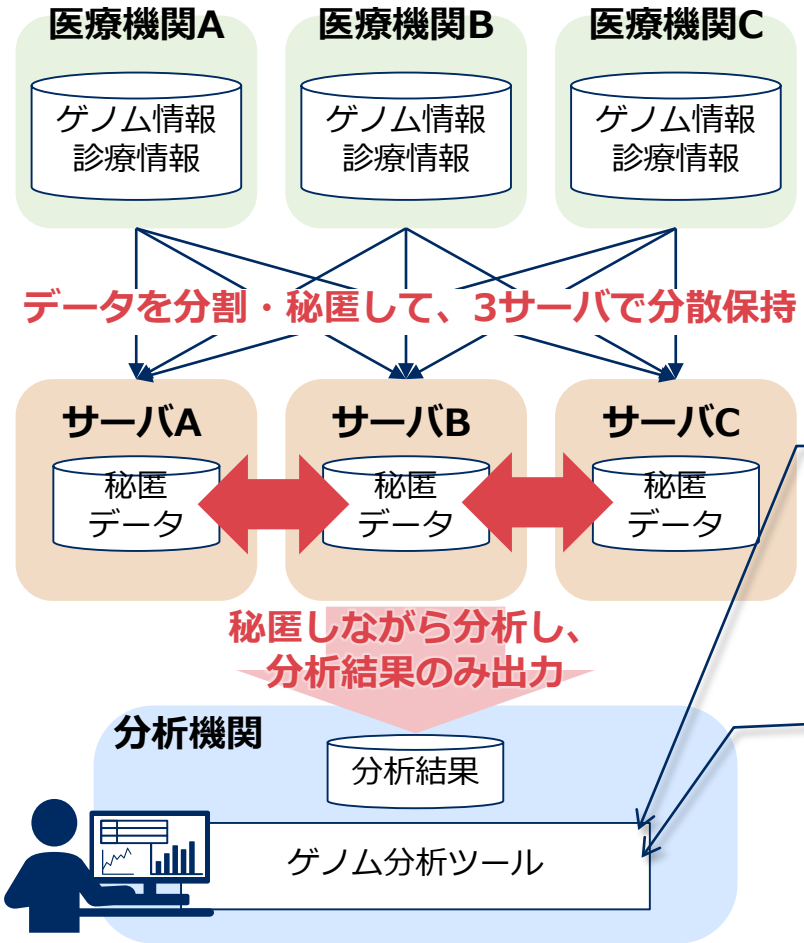
- 秘密計算によって、ゲノム情報等を安全に収集・分析できることを示す
  ⇒ 医学研究の促進に貢献
- 秘密計算が実用レベルであることを検証する

検証項目

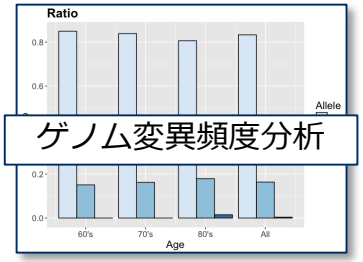
- ①高速性：実用レベルの処理速度であるかの検証
- ②開発容易性： ゲノム解析に秘密計算を容易に適用可能であるかの検証

実証結果の概要

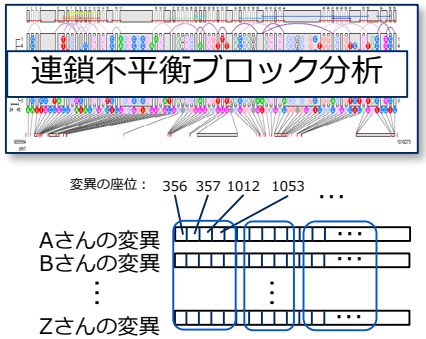
大阪大学とNECと共同でゲノム分析の実証実験を行い、**①高速性**（典型的な集計を約1秒で処理）と、**②開発容易性**（独自アルゴリズムに容易に対応）を確認



■典型的ゲノム分析
（ゲノム変異分析）
ゲノム変異が起こる割合を
年代や性別毎に集計グラフ化
⇒ 実証①：
処理の高速性を実証



■非典型的ゲノム分析
（連鎖不平衡ブロック生成）
世代間の遺伝のし易さを
変異場所のブロックで示す
独自の分析手法
⇒ 実証②：
容易にNECの秘密計算を
適用できることを実証



実証結果の詳細①：高速性の実証

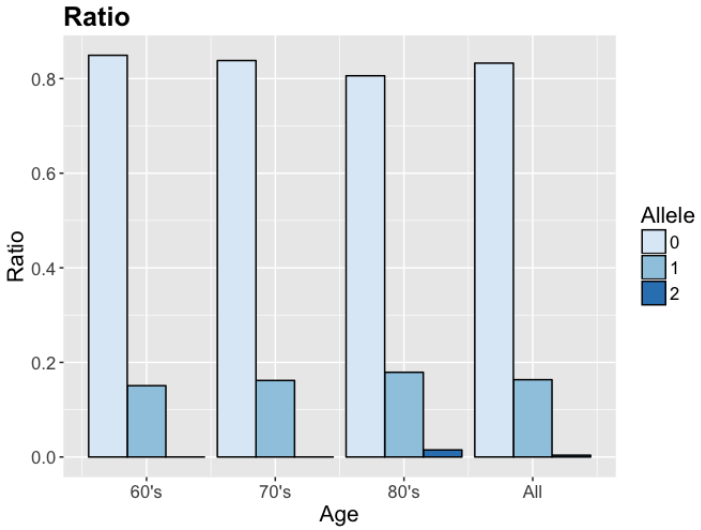
典型的なゲノム解析を行う、分析・可視化ツールに秘密計算を適用し、
約8000人のゲノムデータの集計処理を約1秒で処理可能なことを確認

処理内容

- 年代(4パターン)と変異(3パターン)の組み合わせ(計12パターン)の集計

集計処理結果のグラフ表示例

遺伝子型(Alleleタイプ)の年代別割合



秘密計算サーバのスペック概要

- CPU: Xeon E5-2650v4
- Memory: 128GBytes
- Network: 1Gbps Ethernet
- Disk: SAS HDD 2TBytes

処理速度

- 12並列：約0.8秒
- 4並列：約1.2秒
- 並列化無し：約2.1秒

実証結果の詳細②：開発容易性の実証

大阪大学の独自のゲノム解析の手法に秘密計算を適用し、約400行の処理記述にて、容易に適用可能なことを確認

処理内容：

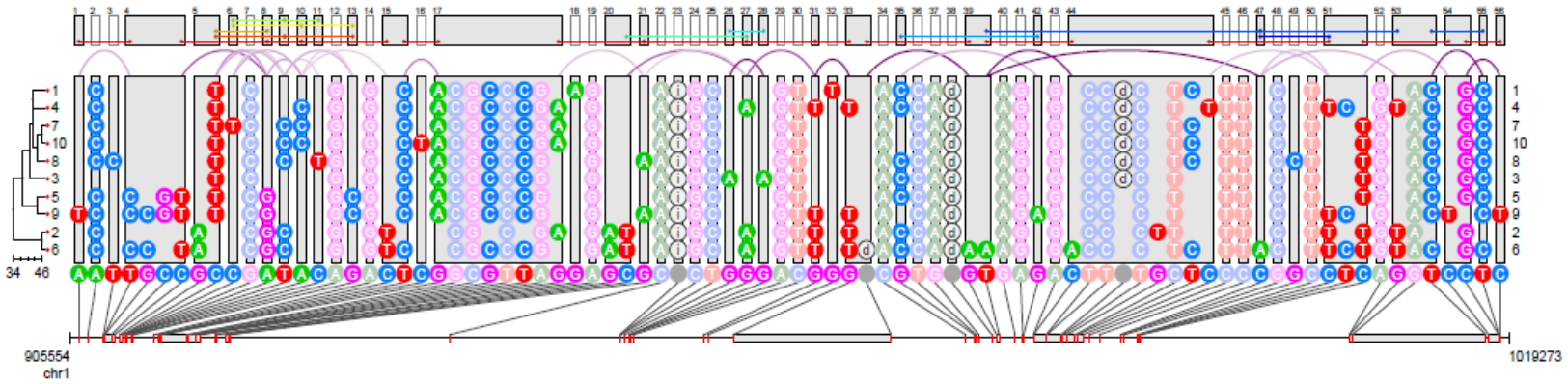
- 相関する変異パタンが連続する染色体領域をブロック化する処理
  （連鎖不平衡ブロックの生成処理）

処理プログラム行：

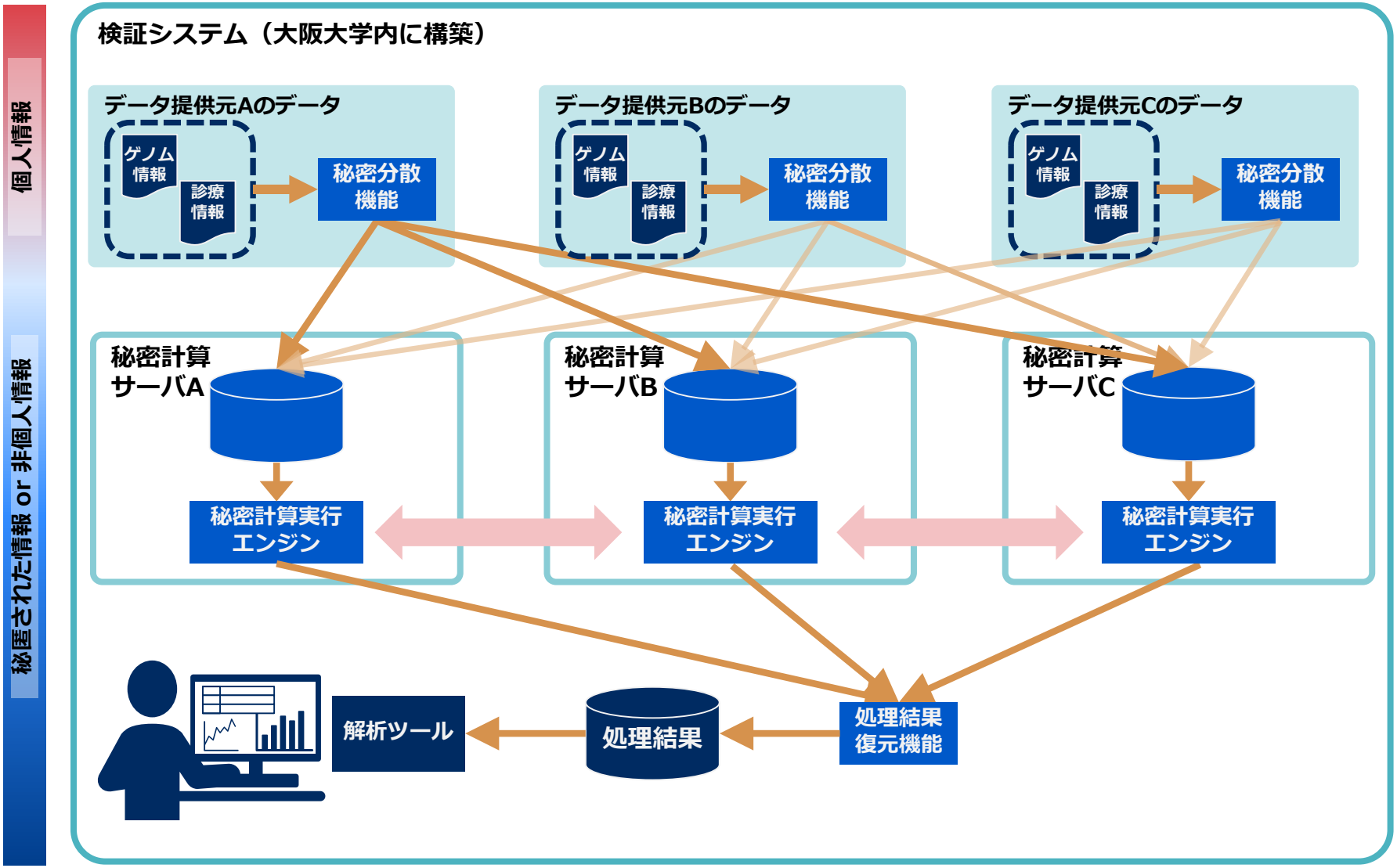
- 相関する変異パタン抽出処理：約130行
- ブロック抽出/結合処理：約250行

⇒ 合計約400行であり、数日程度で実装可能なレベル

大阪大学の独自のゲノム解析の処理結果の表示例



【ご参考】検証システム構成



付録：NECの秘密計算

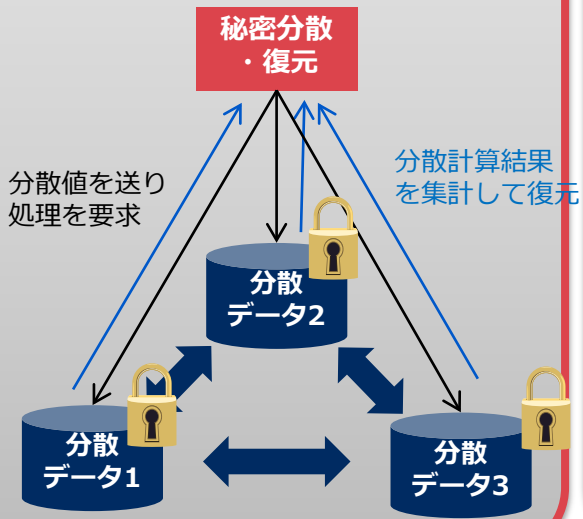
秘密計算技術の様々な方式

秘密計算には様々な方式が存在するが、秘密分散を利用した秘密計算は高速化を実現し、実用化への期待が大きい

秘密計算

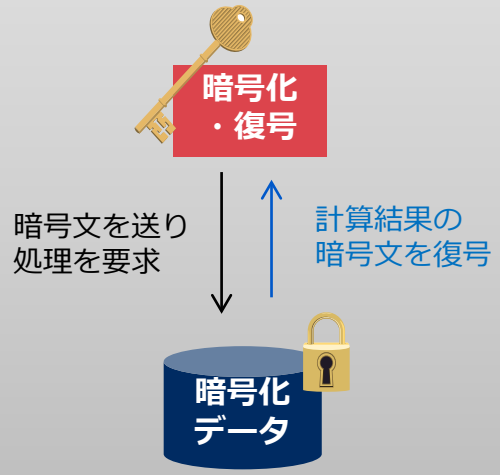
秘密分散を利用した方式

データを秘密分散したまま処理



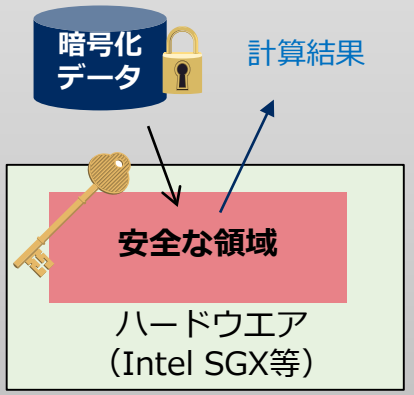
準同型暗号を利用した方式

データを暗号化したまま処理



ハードウエアを利用した方式

ハードウエア上の安全な領域で処理



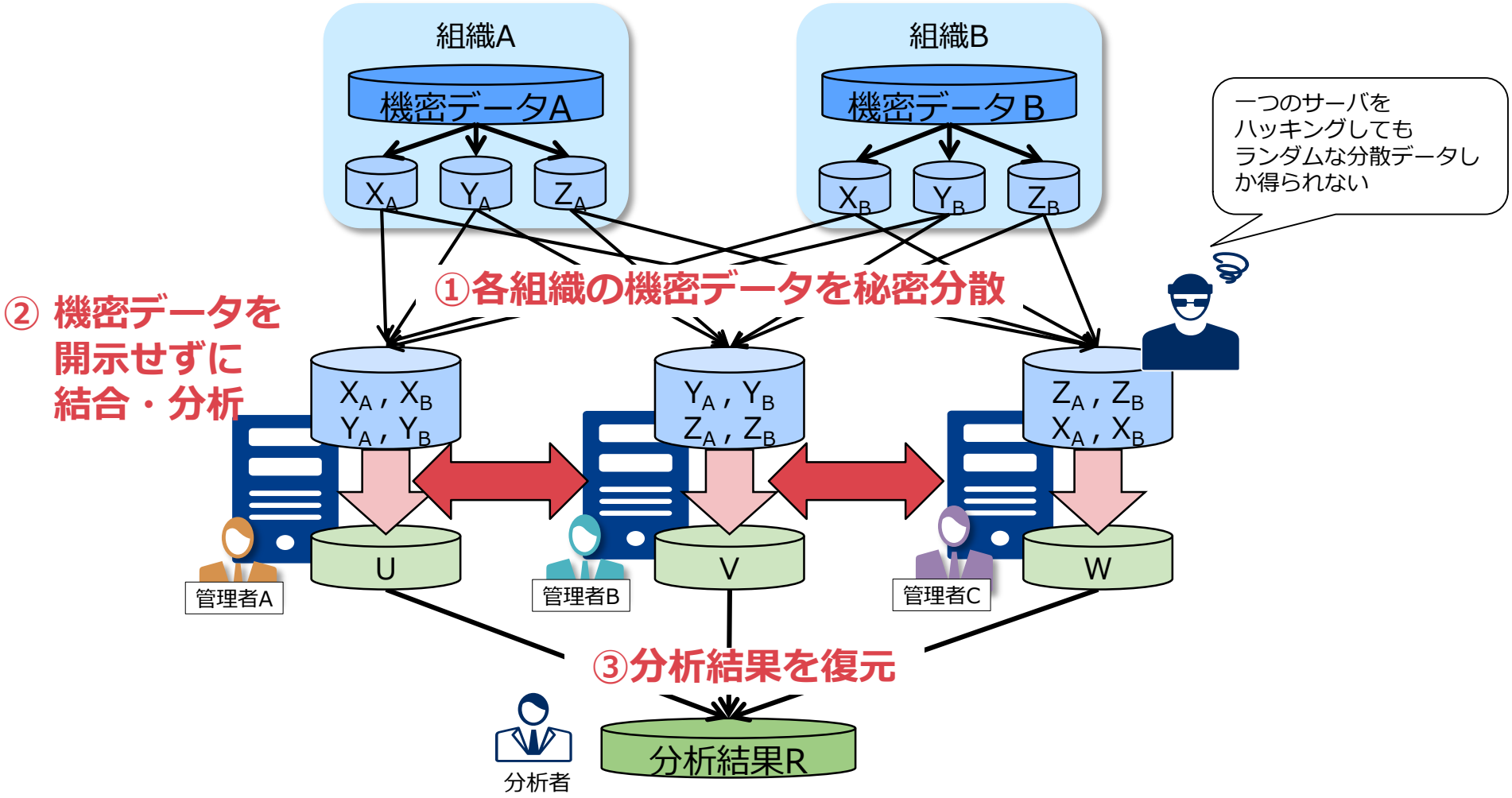
安全な領域に保持している鍵で、データを復号して処理する

※1 Trusted Execution Environment等

・・・ その他方式も存在

秘密分散を利用した秘密計算とは

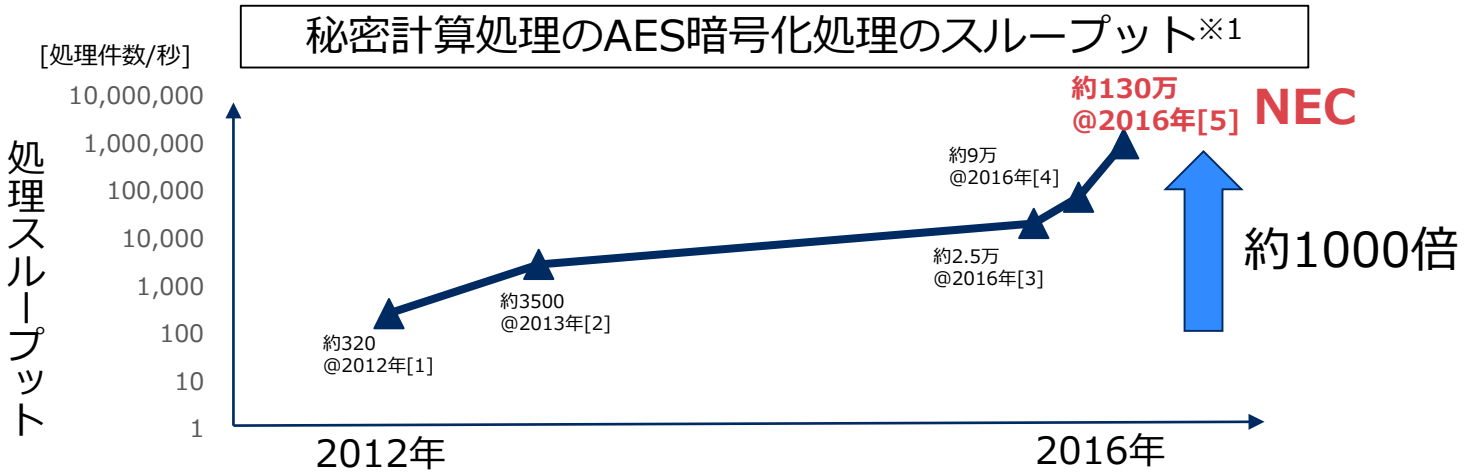
複数サーバに秘匿したデータを分散保持し、**秘匿したまま処理が可能**
たとえ分散した一部データが漏洩しても元データは復元できないため安全



NECの秘密計算技術の特長：①高速性

NECは**高速な秘密計算**を実現し、一部処理で実用的な性能を達成

秘密計算は近年桁違いに高速化を実現（2012年以降で約1000倍）



例：顔特徴量の照合処理やDNA編集距離計算で、現実的な性能を達成※2

- 顔特徴量の照合処理： 1000次元の特徴量の照合が約45,000件/秒※3
- DNA編集距離の計算処理： 長さ100のDNA配列同士の距離計算：約145/秒

国際的に高い評価を得て、難関国際学会に採択。
CCS2016(Best Paper)**、Eurocrypt2017、S&P2017、CCS2018**

[1] J. Launchbury, I.S. Diatchki, T. DuBuisson and A. Adams-Moran. "Effcient lookup-table protocol in secure multiparty computation". ACM ICFP2012.
[2] S. Laur, R. Talviste and J. Willemson. "From Oblivious AES to Effcient and Secure Database Join in the Multiparty Setting", ACNS2013.
[3] R. Talviste. "Applying Secure Multi-Party Computation in Practice", Ph.D dissertation, Univ. of Tartu, 2016.
[4] J. Randmets. Personal comm. AES performance on the new Sharemind cluster. May, 2016.
[5] Toshinori Araki, Jun Furukawa, Yehuda Lindell, Ariel Nof, Kazuma Ohara,"High-Throughput Semi-Honest Secure Three-Party Computation with an Honest Majority", ACM CCS2016.

※1 semi-honest安全な3パーティの秘密計算で比較。[5]論文のTable.1を参考にグラフ化
※2 詳細は「土田 他, “不正検知可能なマルチパーティー計算による生体情報と遺伝子情報の保護”, SCIS2018」を参照
※3 VISAのピーク時トランザクション数：秒間47,000件

NECの秘密計算技術の特長：②開発容易性

NECが開発した秘密計算の「開発支援ツール」によって、一般のエンジニアでも容易に秘密計算のアプリケーション構築が可能

Before:

秘密計算の回路（論理回路、算術回路）を設計し、記述

秘密計算の専門家

```
…
ldsi                s100, 0
adds                s102, s101, s100
movs  s102, s100
…
ge_startmult 2, sg5, sg7
…
ge_stopmult 1, sg506
gadds   sg505, sg2, sg503
gaddsi  sg496, sg505, 1
…
```

数万行

After:

Pythonに似た
プログラミング言語で
処理を記載

一般的なSE

```
def mean(value, num):
sum = sint(0)
for i in range(num):
   sum = sum + value[i]
     …
```

数10～100行

開発支援ツール
(“コンパイラ”)

```
…
ldsi                s100, 0
adds                s102, s101, s100
movs  s102, s100
…
ge_startmult 2, sg5, sg7
…
ge_stopmult 1, sg506
gadds   sg505, sg2, sg503
gaddsi  sg496, sg505, 1
…
```

数万行 ※最適化処理によって、処理の行数の削減も可能


Orchestrating a brighter world
NEC