

# 秘密計算による複数機関が保有するゲノム情報の安全統合解析の実証

竹之内隆夫<sup>1)†</sup>, 中谷明弘<sup>2)</sup>, 菊地正隆<sup>2)</sup>, 小林香織<sup>2)</sup>, 平松直人<sup>3)</sup>  
大原一真<sup>1)</sup>, 土田光<sup>1)</sup>, 荒木俊則<sup>1)</sup>, ○岡村利彦<sup>1)‡</sup>

1) NEC 2) 大阪大学大学院 医学系研究科 3) NECソリューションイノベータ  
†) 現在, 株式会社デジタルガレージ所属 ‡) e-mail: t-okamura@da.jp.nec.com

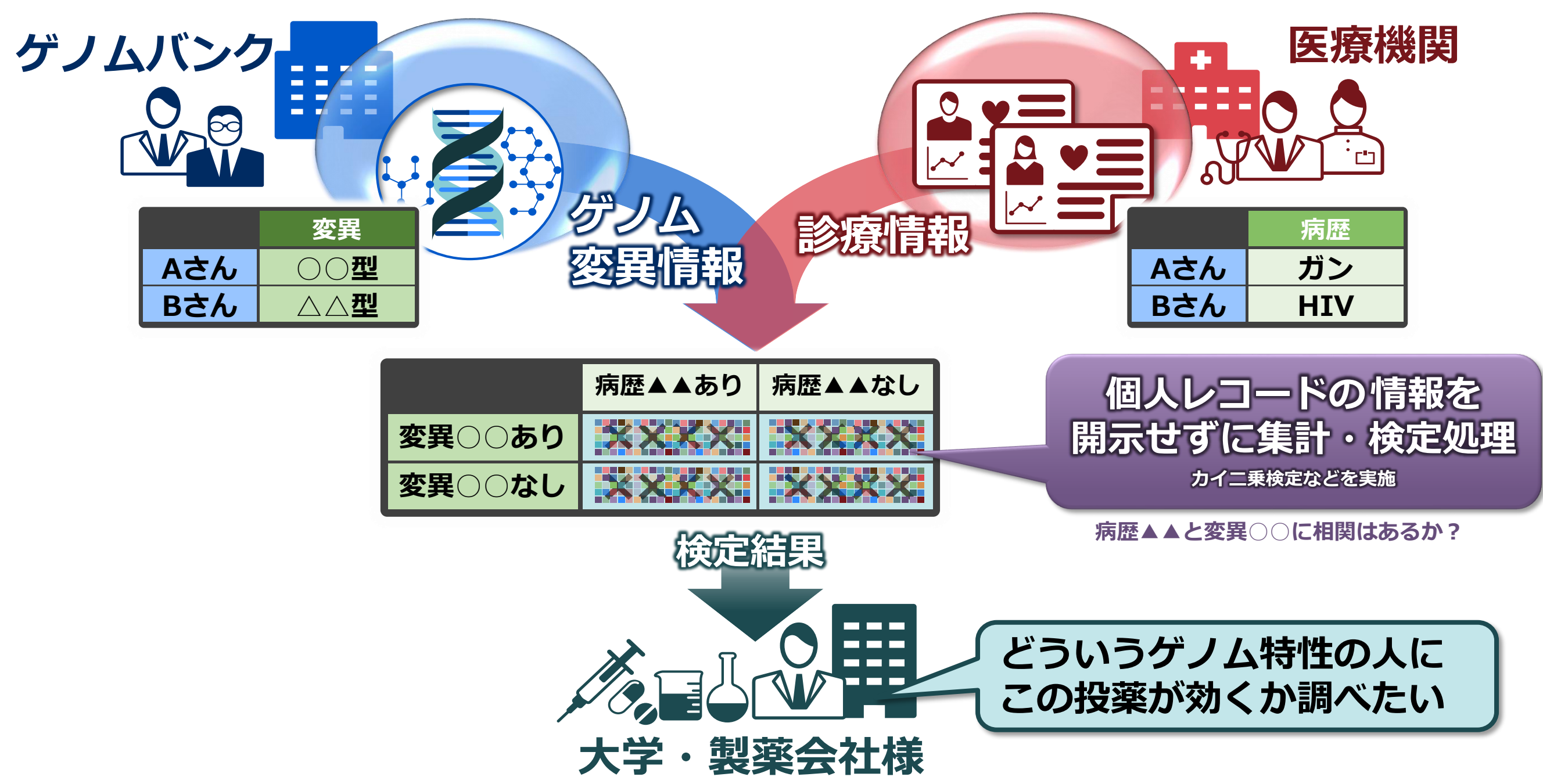
## 実証の概要

秘密計算技術による「複数機関が保有するゲノム情報の安全な統合解析」が、  
① 実行速度の高速性、② 開発容易性の両面で実用レベルにあり、  
秘密計算技術が医学研究の促進に貢献可能なことを実証

## 背景

複数機関が保有するゲノム情報や診療情報を統合して解析することで医学研究の促進が期待できるが、  
これらは機微情報であるため機関外への開示は困難

安全で実用的なデータ統合解析の手法として  
秘密計算への期待が高まっている

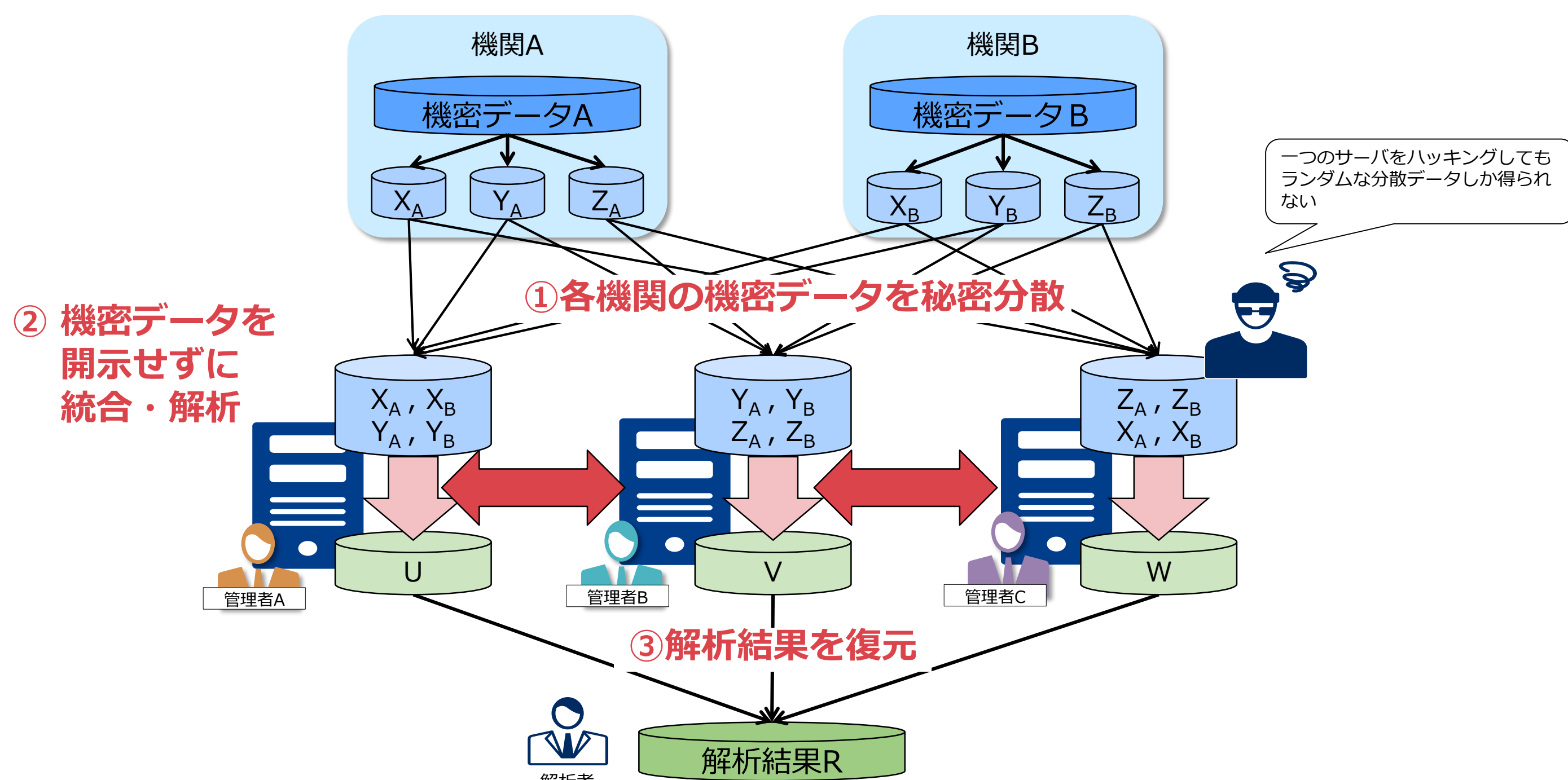


## 秘密計算技術とは

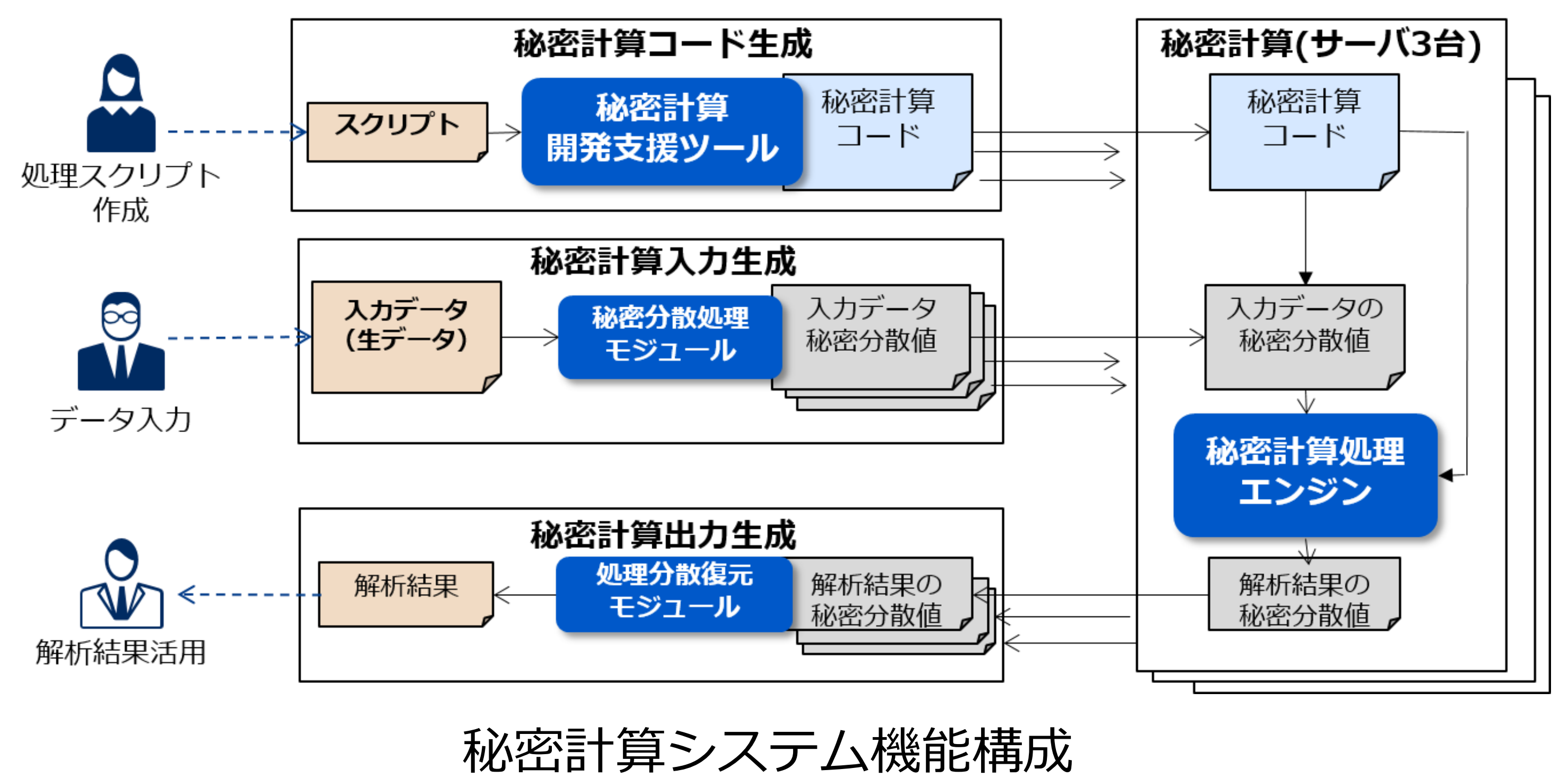
データを秘匿したまま統計解析等の処理ができる技術であり、異なる機関のデータをお互いに一切、開示せずに統合した解析を実現できる。

実証に用いた秘密計算[1]の特徴:

- ① 高速性: 秘密計算の課題である性能を改善し、世界トップレベルを達成
- ② 開発容易性: “開発支援ツール”によって、Pythonライクなスクリプトから秘密計算用コードを生成することが可能



## 実証実験の詳細



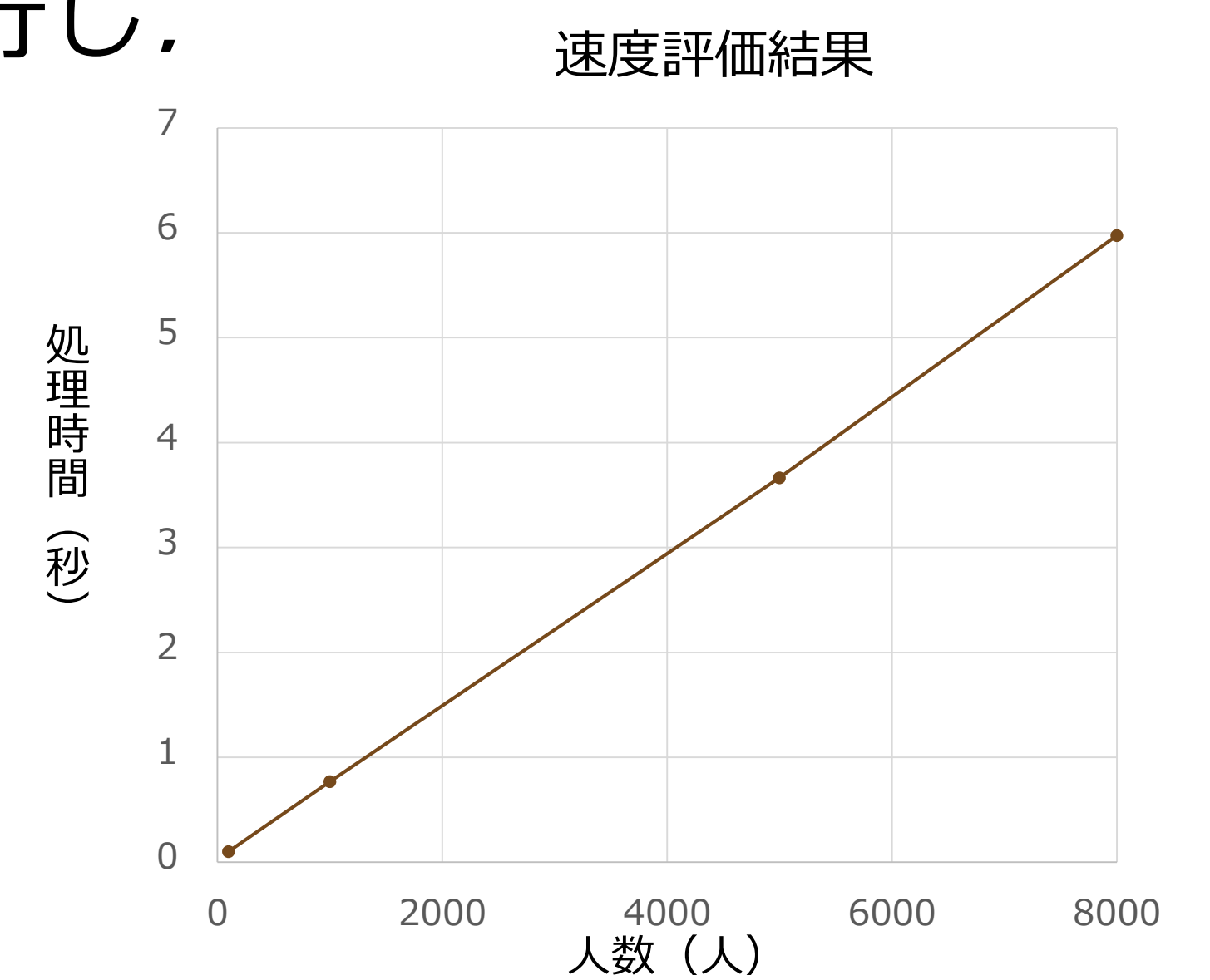
### ① 高速性

#### ◆ 実証内容

分析・可視化ツールに秘密計算を適用し、性別、年代(4パターン)、Alleleタイプ(3パターン)および疾患あり/なしを秘匿したままの状態でもクロス集計処理を実行し、  
処理時間を計測

#### ◆ 結果

8,000人分のデータのクロス集計を6秒以下で実現



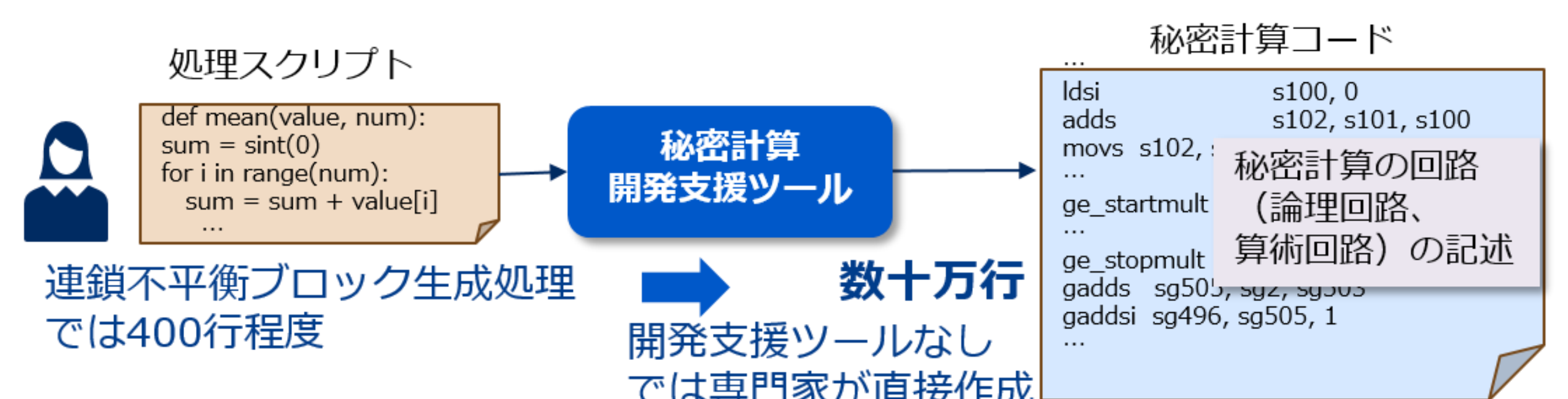
### ② 開発容易性

#### ◆ 実証内容

相関する変異パターンが連続する染色体領域を、カイ二乗検定を用いてブロック化する、  
大阪大学独自のゲノム解析手法である“連鎖不平衡ブロック生成処理”に対して、変異情報を秘匿したまま処理可能な秘密計算コードを生成する

#### ◆ 結果

専門家が1か月程度かかる秘密計算用コード生成を、  
開発支援ツールを用いることで一般のエンジニアが数日間で完了



[1] Araki, T., Barak, A., Furukawa, J., Keller, M., Lindell, Y., Ohara, K. and Tsuchida, H. Generalizing the SPDZ Compiler For Other Protocols. ACM on Computer and Communications Security (ACM CCS) 2018: 880-895

## まとめと今後の展開

秘密計算を用いたゲノム解析の実証を行い、高速性と開発容易性の面で十分実用可能であることを示した。  
今後は様々な医療解析や、ヘルスケア等のデータを用いた解析などへの適用を検討している。