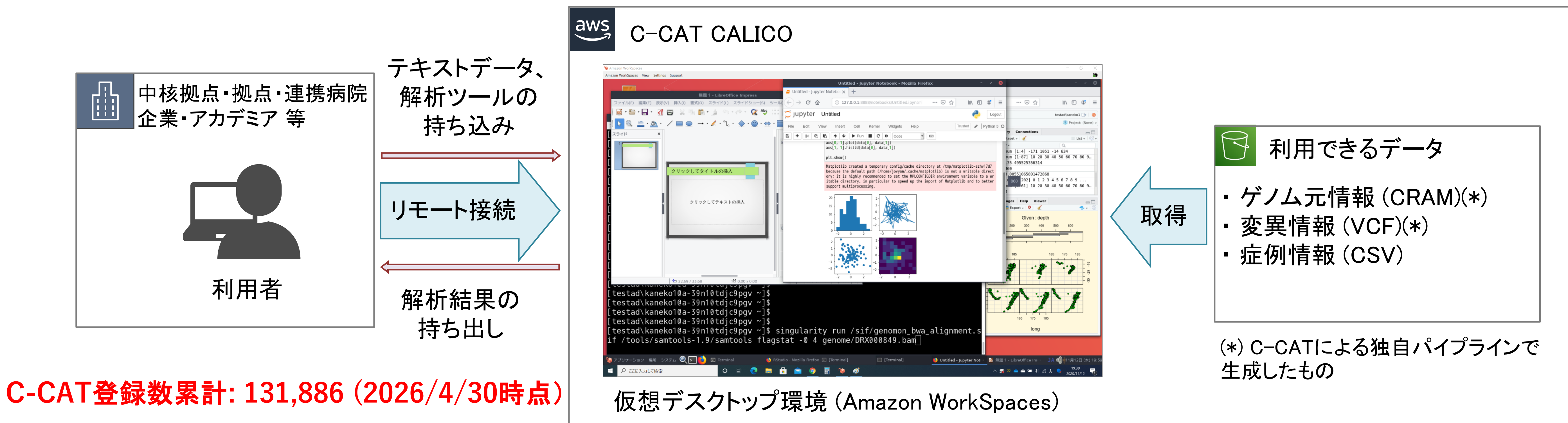


利活用クラウド C-CAT CALICO

千葉 健一(1)、岡田 愛(1)、玉井 郁夫(2)、河野 隆志(3)、温川 恭至(4)、白石 友一(1)

(1) 国立がん研究センター がんゲノム情報管理センター ゲノム解析室、(2) 国立がん研究センター がんゲノム情報管理センター ネットワーク・システム管理室、
(3) 国立がん研究センター がんゲノム情報管理センター センター長、(4) 国立がん研究センター がんゲノム情報管理センター 情報利活用戦略室

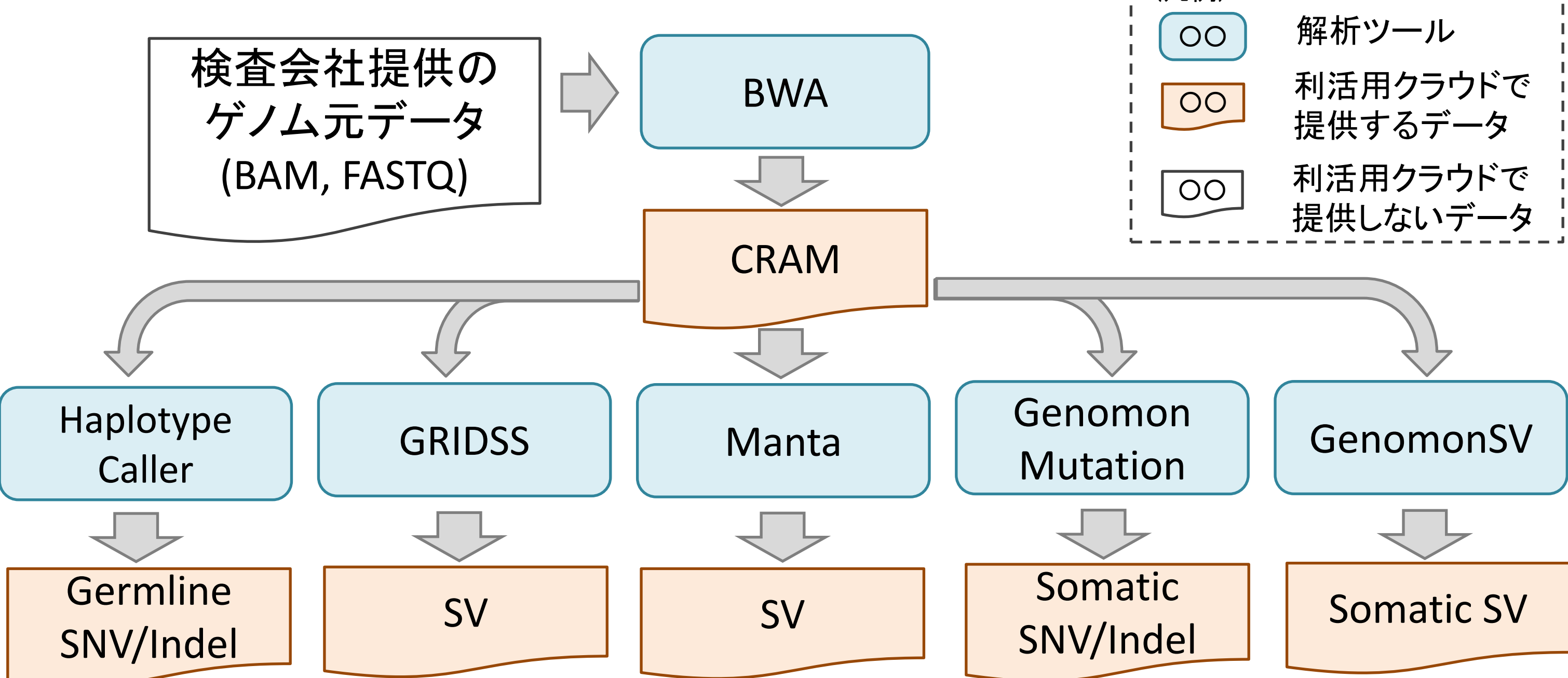
C-CAT CALICO (CALculation & Investigation CIOud) とは



遺伝子パネル検査の**ゲノム元情報 (FASTQ、CRAM)**に基づいて**多種多様な解析するための仮想デスクトップ環境**を、企業、アカデミア向けに提供するサービスです。仮想デスクトップ環境内にゲノム元情報や変異情報をダウンロードし、**C-CATが準備した解析ツール**を使って、解析を行うことができます。

C-CAT CALICO で利用できるデータ

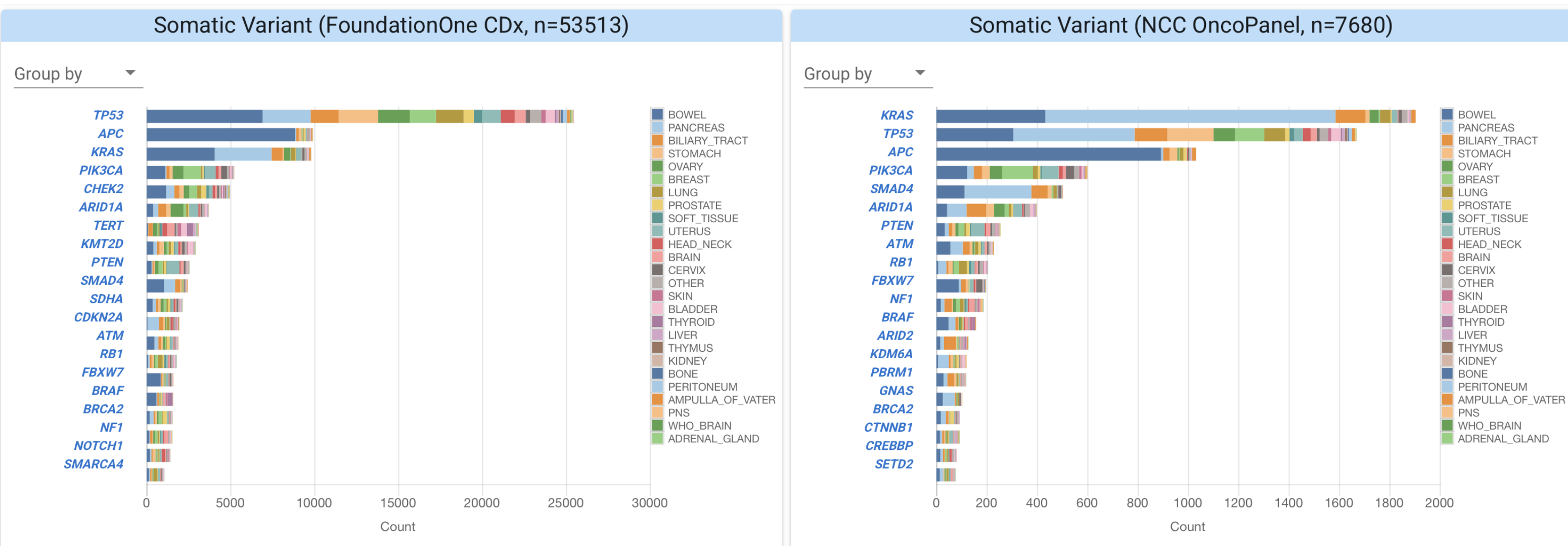
GCATWorkflowCloud パイプラインの解析例



検査会社から提供されたゲノム元データを、統一化された解析パイプラインで変異検出・偽陽性除去・アノテーションの付与をしています。

GCATWorkflowCloud: <https://github.com/ncc-ccat-gap/GCATWorkflowCloud>

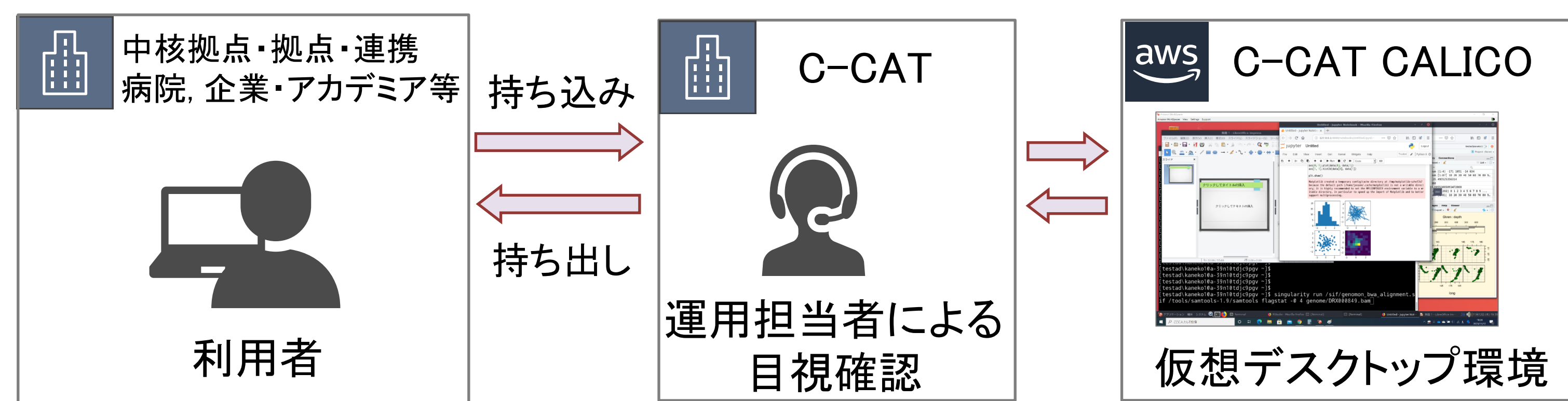
利活用で利用できるデータを使用した集計



利用者は、変異情報(VCF)、症例情報(CSV)、解析ツールなどを使用し、すぐに様々な集計を開始することができます。

FoundationOne CDx: 53,513、NCC OncoPanel: 7,680のゲノム元情報 (CRAM)、変異情報(VCF) が使用できます (2026/7/21時点)。

C-CAT CALICO でのデータの持ち込み・持ち出し



C-CAT CALICOの規定を満たす**データを仮想デスクトップ環境への持ち込み、持ち出しが可能**です。C-CAT CALICOの運用担当者の目視による確認を行っています。

- 持ち込み可能ファイルはテキストファイル、Apptainerイメージファイルです。
- 持ち出し可能ファイルはテキストファイル、画像ファイルです。

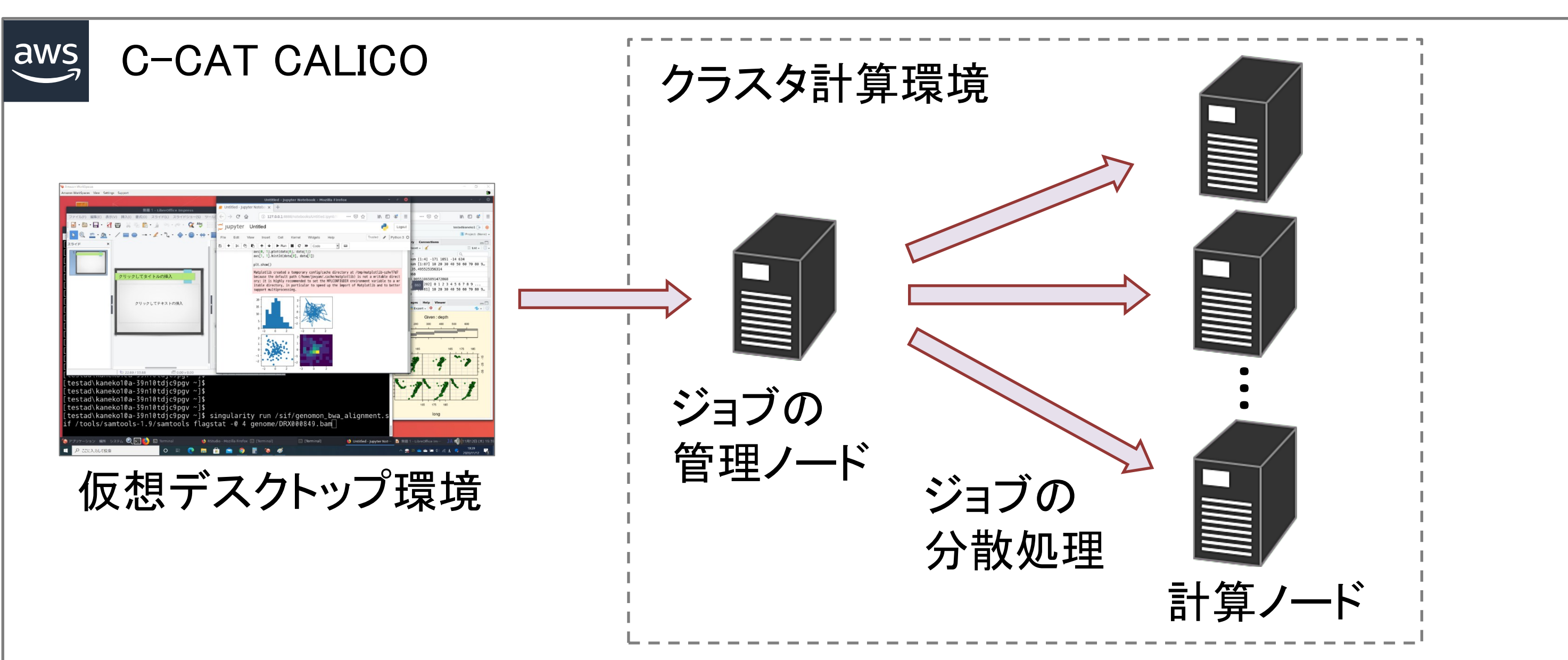
C-CAT CALICO で利用できる解析ツール

解析ツール一覧

用途	ツール名	用途	ツール名	用途	ツール名
マッピング	BWA	変異データ操作	SAMtools	プログラミング言語	Perl - BioPerl
	Bowtie2		HTSlib		Ruby - BioRuby
	BBMap		Picard		Python - pandas - pysam - pybedtools - scikit-learn - BioPython
	STAR	VCF操作	BCFtools		R - tidyverse
	HISAT2		VCFtools		
QCツール	FastQC		cyvcf2		
	RSeQC	FASTQ操作	SeqKit		
	MultiQC		Seqtk		
変異検出	GATK		fastp		
統計解析	RStudio				

仮想デスクトップ環境には**解析ツールがあらかじめインストール**されています。**利用者が自作したコンテナイメージを持ち込んで解析に利用**することもできます。

C-CAT CALICO での大量解析



大規模な計算に対応できるように、仮想デスクトップから接続できる**クラスタ計算環境の整備**を進めています。

C-CAT CALICO の利用方法

現在、33のアカデミアグループがC-CAT CALICOを利用しています。がんゲノム医療病院の場合、あるいはその他のアカデミアに関連する公的研究費を獲得している場合、**C-CAT CALICO利用料は無償**です(クラウド環境利用の実費分は有償です)。C-CAT CALICO の利用には「C-CATデータ利活用(新規・変更・延長)申請書」に、倫理審査委員会へ提出の「研究計画書」および「倫理審査結果の写し」、もしくは「調査概要書」を添付して、C-CAT情報利活用戦略室に提出する必要があります。詳しくは以下のWebページをご確認ください。

C-CAT 利活用のページ > 提供する情報と各種利活用システムについて



本コンテンツ利用について (<https://www.ncc.go.jp/jp/copyright.html>)

