

Illumina Connected Analytics

Zur Transformation Ihrer
Bioinformatikverfahren

- Optimierte Integration in Sequenziersysteme
- Skalierbarer Durchsatz für die bedarfsgerechte Probenverarbeitung
- Intuitive Benutzeroberfläche zum Auswählen, Erstellen und Anpassen von Workflows

Einleitung

Fortschritte bei den NGS-Technologien (Next-Generation Sequencing, Sequenzierung der nächsten Generation) haben die Maßstäbe in den Life Sciences und der klinischen Forschung dramatisch verändert. Vor dem Hintergrund der wachsenden Geschwindigkeit der Sequenzierung und der sinkenden Kosten werden die Kapazitäten zur Generierung von Daten die Kapazitäten zu deren Überführung in biologische und klinische Erkenntnisse bei Weitem übertreffen. Für die Bewältigung der Herausforderungen hinsichtlich der Entwicklung und Implementierung entsprechender Pipelines, der Skalierung von Informatik-Workflows und der Gewährleistung eines sicheren Datenmanagements bedarf es einer flexiblen und umfassenden Plattform. Mit Illumina Connected Analytics können Anwender flexible Analysepipelines erstellen, überarbeiten und bereitstellen, während zugleich Datenschutz, Sicherheit und Compliance gewährleistet bleiben.

Illumina Connected Analytics ist eine sichere, cloudbasierte Bioinformatik-Datenplattform zur Operationalisierung von Informatikverfahren und zur Förderung wissenschaftlicher Erkenntnisse (Abbildung 1, Tabelle 1). Mit Illumina Connected Analytics können Kunden Workflows erstellen, die folgende Vorteile bieten:

- Vernetzt, mit optimierter Integration in Systeme und Software von Illumina
- Skalierbar, mit der Möglichkeit, enorm umfangreiche Datensätze zu verwalten, zu analysieren und abzufragen
- Flexibel, lassen sich mit DRAGEN™-Anwendungen und entsprechenden benutzerdefinierten Analysepipelines erstellen und anpassen

Optimierter Workflow

Illumina Connected Analytics ist eine zentrale Komponente für Labore, die NGS-Studien mit Illumina-Sequenziersystemen durchführen. Dank der durch Cloud-Computing ermöglichten flexiblen Ressourcennutzung eignet sich Illumina Connected Analytics für den Einsatz in jeder Größenordnung auf derselben Architektur – von gelegentlichen Screenings bis hin zu Anwendungen mit Zehntausenden Zellen in komplexen Einzelzellprojekten und der populationsweiten Genomsequenzierung. Anwender können eigene Geräte problemlos in Illumina Connected Analytics integrieren.

Daten lassen sich in Illumina Connected Analytics abhängig vom gewählten Workflow automatisch mithilfe gebrauchsfertiger DRAGEN-Sekundäranalyse-Pipelines oder mit anwendungsspezifischen Pipelines analysieren. Die breite Palette an Analyseoptionen reicht von Qualitätskontrolle und Datenaggregation bis hin zu fortschrittlichen Data-Science-Tools zur schnellen, skalierbaren Datenverarbeitung. Illumina Connected Analytics bietet eine

erweiterbare Plattform mit einem umfangreichen Satz von RESTful-APIs (Application Program Interfaces, Anwendungsprogrammchnittstellen) und einem CLI-Tool (Command-Line Interface, Befehlszeilenschnittstelle). Diese APIs maximieren die Workfloweffizienz, da Daten über ihren gesamten Lebenszyklus übertragen, abgerufen und verwendet werden.

Tabelle 1: Illumina Connected Analytics im Überblick

	Merkmals	Vorteil
Sicherheit und Datenschutz	Compliance	Entwickelt mithilfe des Qualitätsmanagementsystems (QMS) von Illumina und konform mit lokalen, regionalen sowie global gültigen regulatorischen Standards wie DSGVO ² , HIPAA ³ , ISO 27001/27701 ^{4,5} und ISO 13485 ⁶
	Sicherheitsmaßnahmen	Aufrechterhaltung einer strikten Datentrennung, Verschlüsselung bei der Übertragung (TLS 1.2) und Speicherung (AES 256)
	Auditpfad	Tracking von Aktivitäts- und Ereignisprotokollen zur Zugriffsüberwachung
	Multifaktor-Authentifizierung (MFA)	Nutzung einrichtungs-spezifischer Anmeldedaten zur Verwaltung und Steuerung des Systemzugriffs
Ressourcen-zuweisung	Bedarfsabhängige Skalierung	Skalierung der Cloudspeicher- und der entsprechenden Computing-Ressourcen abhängig vom aktuellen Bedarf
	Bedarfsabhängige Computing-Ressourcen	Einsparungen, da nur Kosten für die Computing-Ressourcen in der Pipeline-Engine anfallen
Management	Projekt- und Anwendermanagement	Granularer Datenschutz durch Steuerung von Anwenderzugriff und -aktivität
	Gemeinsame Nutzung von Daten	Überbrückung von Datensilos für die globale Zusammenarbeit und Unterstützung der Datenbereitstellung
	Archivierung von Daten	Kostensenkung für langfristiges Datenmanagement durch Archivierung auf spezifischen Speicherebenen
	Nutzung des eigenen AWS-Cloud-Buckets	Zugriff auf in einem privat verwalteten AWS-Cloudkonto gespeicherte Daten
Benutzerfreundlichkeit und Zugänglichkeit	Integration von Sequenziersystemen	Nahtlose Datenübertragung von Illumina-Sequenziersystemen zu Illumina Connected Analytics-Speicherorten und -Analysen
	Tools und Pipelines	Nutzung sofort einsatzbereiter DRAGEN-Pipelines oder Import vorhandener Tools
	APIs und CLI	Programmgesteuerte Interaktion mit Plattform, Tools und Daten über CLI oder RESTful-API
	Intuitive grafische Benutzeroberfläche	Zugriff auf Daten und Tools über eine grafische Weboberfläche; Anzeige von Daten mit R- und Python-Paketen
Fortschrittliche Tools	Unterstützung von Docker, Nextflow und CWL	Pipeline-Programmierung in einer gemeinsamen Workflowsprache und einfacher Start von Analysen in der Cloud
	JupyterLab-Integration	Ausführung benutzerdefinierter Skripte, Erstellung und Training von KI-/Machine-Learning-Modellen und Interaktion mit Daten innerhalb der Plattform über Notebook-Arbeitsbereiche für die Zusammenarbeit
	Datenaggregation und -abfrage	Organisation und Abfrage strukturierter Multiomik-Datensätze zur Unterstützung der Tertiäranalyse

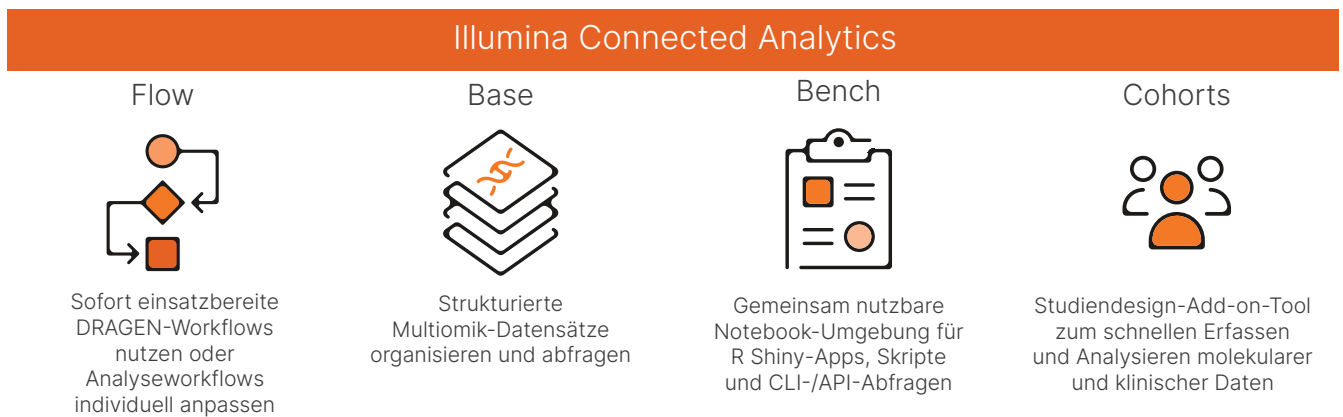


Abbildung 1: Illumina Connected Analytics-Modulkomponenten: Verarbeitung, Aggregation und Analyse von Daten in einer sicheren Datenmanagementumgebung mit kontrolliertem Zugriff.

BaseSpace™ Sequence Hub

BaseSpace Sequence Hub ist in allen Illumina Connected Analytics-Abonnements enthalten und fungiert als direkte Erweiterung Ihrer Illumina-Geräte. Dank der Integration von BaseSpace Sequence Hub können Sie Läufe konfigurieren, die Laufqualität überwachen sowie Sequenzierungsdaten direkt in die Cloud streamen und so die Laboreffizienz steigern. Die Daten werden verschlüsselt vom Gerät in BaseSpace Sequence Hub übertragen und lassen sich so in einer sicheren Umgebung einfach verwalten und in ausgewählten Apps analysieren.

Überführung von Reads in Daten

Illumina Connected Analytics bietet mehrere Optionen für die Sekundäranalyse von Daten und optimiert den Workflow von den Reads bis zu den Ergebnissen. Illumina Connected Analytics bietet die Flexibilität zur Nutzung vorgefertigter Pipelines oder zur Erstellung und Konfiguration anwendungsspezifischer Pipelines. Damit eignet sich Illumina Connected Analytics für praktisch jede Informatikanwendung.

Gebrauchsfertige Optionen

Illumina Connected Analytics bietet leistungsstarke, sofort einsatzbereite Tools und Pipelines für die Datenverarbeitung, einschließlich Zugriff auf die DRAGEN-Sekundäranalyse¹, die eine genaue, effiziente und umfassende Sekundäranalyse von Sequenzierungsdaten ermöglicht (Abbildung 2).

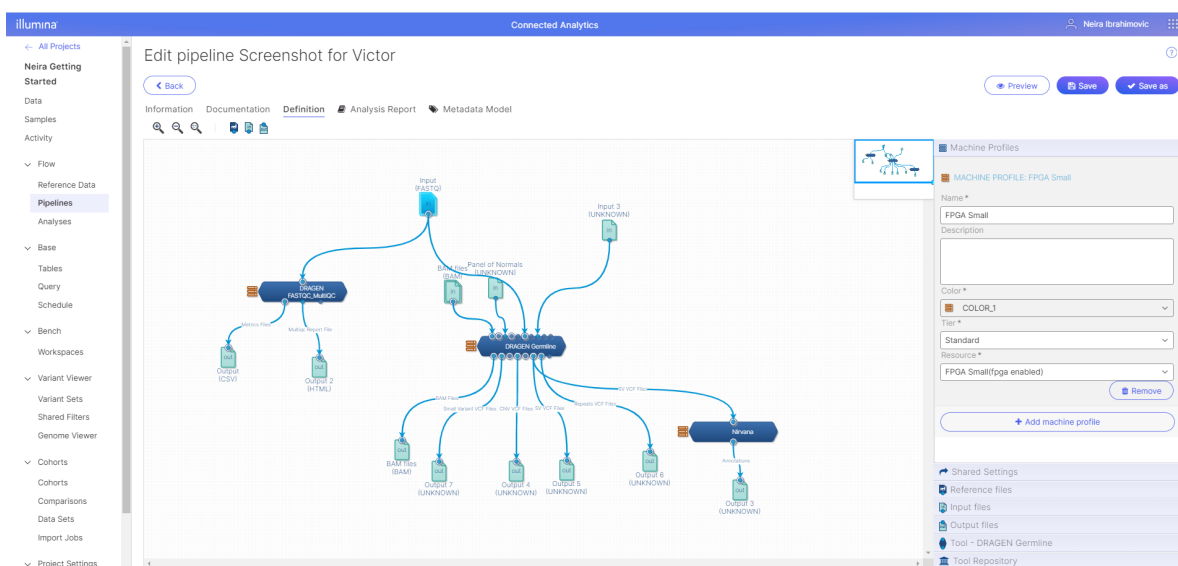


Abbildung 2: Illumina Connected Analytics Flow: Flexible Erstellung, Bereitstellung und Versionierung von Analysepipelines.

Anpassung von Pipelines

Bioinformatiker können in einem Docker Hub-Image-Repository⁷ vorhandene Analysetools importieren oder mit Nextflow, Common Workflow Language (CWL) und dem grafischen Pipeline-Editor neue Pipelines erstellen und bearbeiten. Labormitarbeiter und andere Wissenschaftler können Pipelines ganz einfach über die intuitive Benutzeroberfläche starten.

Datenmanagement und -kontrolle

Da immer größere Datenmengen generiert werden, wächst auch der Bedarf an Infrastruktur für die gemeinsame Nutzung, Wiederverwendung und Integration von Daten innerhalb der Wissenschaftsgemeinde, damit sich einzelne Datensätze optimal auswerten lassen. Illumina Connected Analytics umfasst hierzu mehrere Funktionen, die die Befolgung von Best Practices für das Datenmanagement ermöglichen.

Zugriffskontrolle

Dank der granularen Zugriffskontrolle können Administratoren den Zugriff über Berechtigungen und in der Einrichtung bereits genutzte Anmeldeinformationen steuern. Ereignisse und Änderungen werden in einem Auditprotokoll festgehalten. Dabei werden sämtliche Anwender erfasst, die auf die Plattform zugreifen, sowie deren Aktionen während der Nutzung der Plattform, wodurch sich Compliance und Rechenschaftspflicht gewährleisten lassen.

Offenes Format

Illumina Connected Analytics ist als Plattform für alle Daten konzipiert. Die Plattform ermöglicht die Analyse zahlreicher Datentypen, darunter Molekular-, Klinik- und Phänotypdaten sowie unstrukturierte Daten wie Bilder.

Zusammenarbeit

Illumina Connected Analytics ermöglicht eine standortunabhängige Zusammenarbeit unter Einhaltung sämtlicher Compliance-Anforderungen. Daten und Tools lassen sich unmittelbar bereitstellen und gemeinsam mit anderen Anwendern nutzen. Die Integrität und der Schutz der Daten bleiben hierbei gewährleistet. Darüber hinaus können in einer externen Cloudquelle gehostete Daten- und Analysetools zur Analyse und gemeinsamen Nutzung in Illumina Connected Analytics importiert werden.

Datenaggregation und -abfrage

Illumina Connected Analytics schafft anhand der Automatisierung komplexer Aggregations- und Integrationsschritte ein funktionales Wissensmanagementsystem, das Daten zu Millionen von Proben umfasst. Hierbei werden praktisch alle verfügbaren Datentypen (Genotyp-, Phänotyp-, Meta- und Annotationsdaten) sowie weitere zugehörige Informationen erfasst. Anwender können eigene Datenmodelle sowie Abfragen erstellen und Zusammenhänge zwischen den Datensätzen wie gewünscht untersuchen. Die auf Illumina Connected Analytics aggregierten Daten bilden eine breite Basis für die Suche nach neuartigen Biomarkern, die Stratifizierung von Patientenpopulationen, die Überwachung der Assayleistung im Laufe der Zeit und vieles mehr.

Verwendung von Cohorts zur Untersuchung klinisch-genomischer Daten in großem Maßstab

Illumina Connected Analytics Cohorts ist ein Studiendesign-Tool zum schnellen Erstellen und Untersuchen von Kohorten ([Abbildung 3](#)). Illumina Connected Analytics Cohorts beschleunigt die Überführung von genomischen Entdeckungen in aussagekräftige Studien und ermöglicht beim Studiendesign eine bessere Entscheidungsfindung, indem Multiomik- und Phänotypdaten an einem einzigen Ort zusammengeführt werden.

Kohorten lassen sich mithilfe enthaltener harmonisierter öffentlicher Datensätze wie The Cancer Genome Atlas (TCGA), Broad Rare Genomes Project, 1000 Genomes Project und Gene Expression Omnibus ergänzen. Dank eines integrierten Kohortenbrowsers können Sie Hypothesen in Minuten statt Wochen erstellen. Außerdem stehen Ihnen umfangreich kuratierte öffentliche Datensätze zur Verfügung. Nutzen Sie einfache visuelle Tools, mit denen sich Kohorten auswählen und visualisieren lassen, ohne dass eine komplexe Abfrage erforderlich ist.

Sichere Notebook-Umgebung zur Auswertung

Die Entwicklung und Anpassung von Algorithmen sind wesentliche Komponenten von Illumina Connected Analytics zur Unterstützung einer tiefgreifenden Datenauswertung. Ein interaktives Programmiermodul, das gängige Jupyter-Notebooks, Python und R nutzt, ermöglicht Datenwissenschaftlern die Analyse aggregierter Daten in einer unkomplizierten und sicheren Umgebung ([Abbildung 4](#)).

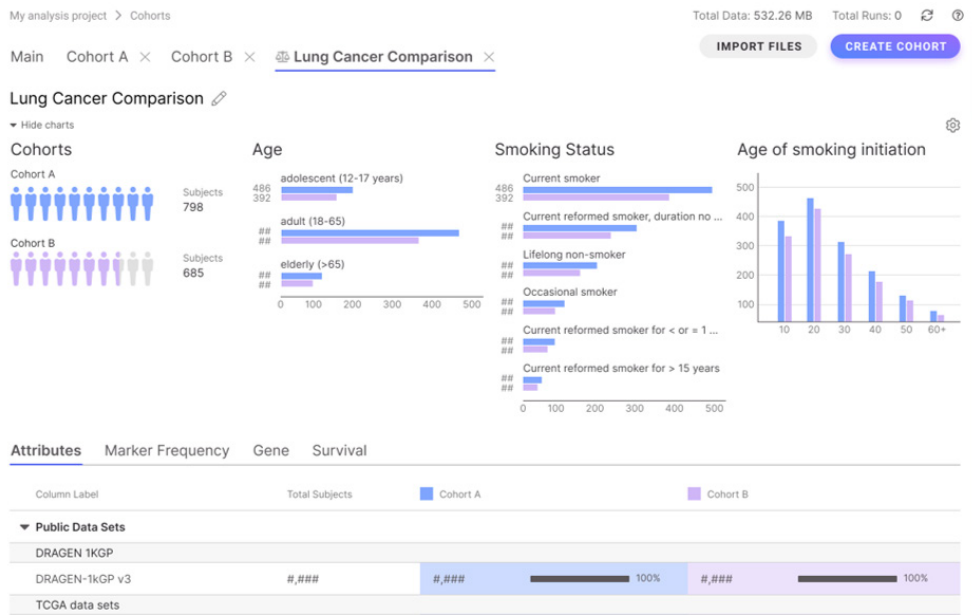


Abbildung 3: Illumina Connected Analytics Cohorts: Schnelle Erfassung und Untersuchung von molekularen und klinischen Daten mithilfe dieses Add-on-Moduls in Illumina Connected Analytics.

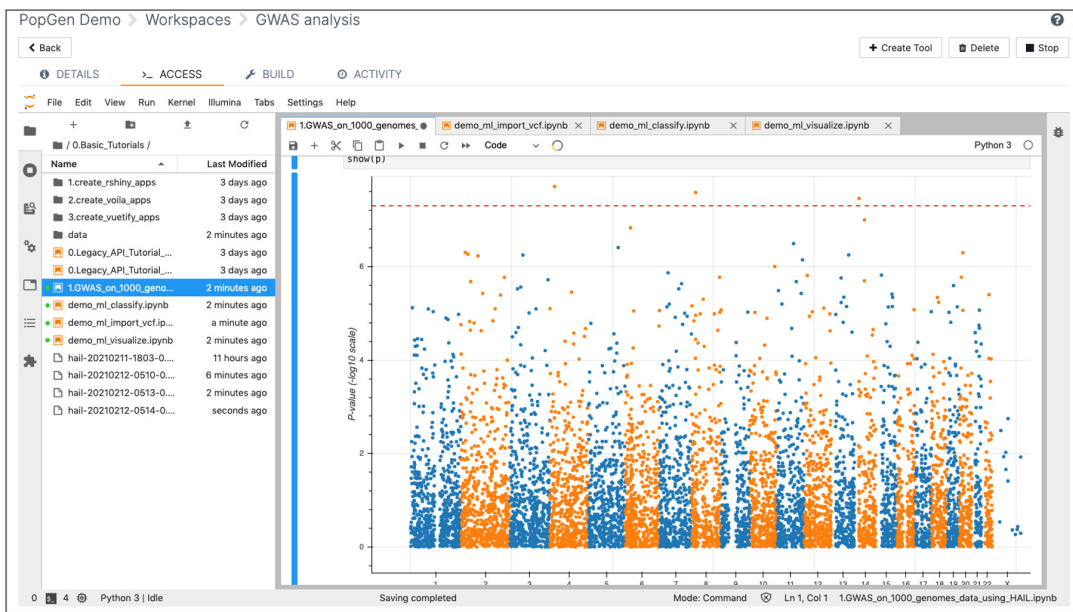


Abbildung 4: Illumina Connected Analytics Bench: In Illumina Connected Analytics Bench sind flexible und interaktive Computing-Tools wie Jupyter-Notebooks, Python und R integriert.

Anwender profitieren während der Phase der Entwicklung von Methoden und Algorithmen von einer Sandbox-Umgebung, in der sich Pipelines entwickeln und bearbeiten lassen. Dort haben Anwender Zugriff auf Standardbibliotheken und können ganz einfach eigene benutzerdefinierte Bibliotheken wie TensorFlow⁸ oder scikit-learn⁹ integrieren, um benutzerdefinierte und komplexe Skripte zu erstellen, mit denen sich Daten innerhalb von Illumina Connected Analytics gemeinsam verarbeiten lassen. Sobald Anwender zur Produktionsphase übergehen möchten, ermöglicht Illumina Connected Analytics die Umwandlung von Notebooks in Tools. Diese Tools sind anschließend im Illumina Connected Analytics-Tools-Repository zur Integration in Produktionspipelines verfügbar.

Sicherheit und Compliance im Mittelpunkt

Sicherheit ist bei der Arbeit mit Genomikdaten in Zusammenhang mit der Forschung, klinischen Therapeutika und der Humandiagnostik von größter Bedeutung. Illumina Connected Analytics gewährleistet mithilfe unterschiedlicher digitaler und administrativer Maßnahmen die Einhaltung selbst der strengsten Datensicherheitsvorschriften:

- Von Sequenzierungsgeräten hochgeladene Daten werden mit dem Standard AES-256¹⁰ verschlüsselt und durch Transfer Layer Security (TLS) geschützt.
- Die Daten in Illumina Connected Analytics werden bei Amazon Web Services (AWS) gehostet. Hierbei sorgen AWS Well-Architected-Best-Practices für die Einhaltung zahlreicher branchenüblicher Sicherheitsstandards.¹¹
- Illumina Connected Analytics wird in mehr als 10 globalen AWS-Regionen gehostet, um die Einhaltung lokaler Vorschriften zum Umgang mit genomischen Daten in bestimmten geografischen Regionen zu ermöglichen.
- Der Authentifizierungsdienst basiert auf Security Assertion Markup Language (SAML) 2.0 zur Verwaltung der Benutzer und Kennwörter in der Einrichtung (optional).
- Auditberichte ermöglichen die Rückverfolgung der Herkunft von Daten.

Illumina Connected Analytics eignet sich auch für Kunden, die in regulierten Bereichen tätig sind und strenge Anforderungen einhalten müssen. Illumina Connected Analytics wurde in Übereinstimmung mit dem auf das Illumina-Qualitätsmanagementsystem (QMS) abgestimmten Illumina-Softwareentwicklungsprozess entwickelt.

Darüber hinaus folgen die Prozesse innerhalb des Illumina-QMS branchenspezifischen Best Practices und relevanten Normen, darunter:

- ISO 27001:2013 (Informationssicherheits-Managementssystem)⁴ und ISO 27701:2019 (Sicherheitstechniken)⁵ der Internationalen Organisation für Normung
- Best Practices gemäß ISO 13485 für ein Qualitätsmanagementsystem (QMS) für Unternehmen⁶
- Datenschutz-Grundverordnung (DSGVO)²
- Health Insurance Portability and Accountability Act (HIPAA)³
- Garantierte Data Residency zur Erfüllung lokaler Bestimmungen und Compliance-Anforderungen

Bestellinformationen

Produkt	Katalog-Nr.
ICA Enterprise Annual Subscription	20038994
ICA Professional Annual Subscription	20044876
ICA Training and Onboarding	20049422
ICA Cohorts Annual Subscription	20065842
ICA Cohorts Annual Subscription	20065842
Illumina Analytics - 1 iCredit	20042038
ICA Enterprise Svc and Compliance Add-on	20066830

Abkürzungen: ICA, Illumina Connected Analytics

Weitere Informationen

[Illumina Connected Analytics](#)

Quellen

1. Illumina. Illumina DRAGEN secondary analysis. www.illumina.com/products/by-type/informatics-products/dragen-secondary-analysis.html. Aufgerufen am 1. März 2024.
2. General Data Protection Regulation (GDPR) Compliance Guidelines. GDPR-Website. gdpr.eu. Aufgerufen am 1. März 2024.
3. US Department of Health and Human Services. Health Information Privacy. HHS-Website. hhs.gov/hipaa/index.html. Aufgerufen am 1. März 2024.
4. International Organization for Standardization. ISO-ISO/IEC 27001—Information security management. ISO-Website. iso.org/isoiec-27001-information-security.html. Aufgerufen am 1. März 2024.
5. International Organization for Standardization. ISO/IEC 27701:2019—Security techniques. iso.org/standard/71670.html. Aufgerufen am 1. März 2024.
6. International Organization for Standardization. ISO 13485—Medical devices. ISO-Website. iso.org/iso-13485-medical-devices.html. Aufgerufen am 26. März 2024.
7. Docker. Docker Hub Image Container Library. <https://hub.docker.com>. Aufgerufen am 26. März 2024.
8. TensorFlow. TensorFlow-Website. tensorflow.org. Aufgerufen am 1. März 2024.
9. scikit-learn: machine learning in Python. scikit-learn-Website. scikit-learn.org/stable/. Aufgerufen am 1. März 2024.
10. National Institute of Standards and Technology. Advanced Encryption Standard (AES). nist.gov/publications/advanced-encryption-standard-aes-0. Aufgerufen am 26. März 2024.
11. Cloud Security—Amazon Web Services (AWS). Amazon-Website. aws.amazon.com/security. Aufgerufen am 1. März 2024.
12. Illumina. iCredits for Data Storage & Analysis. illumina.com/products/by-type/informatics-products/icredits.html. Aufgerufen am 1. März 2024.



1 800 8094566 (USA, gebührenfrei) | +1 858 2024566 (Tel. außerhalb der USA)
techsupport@illumina.com | www.illumina.com

© 2024 Illumina, Inc. Alle Rechte vorbehalten. Alle Marken sind Eigentum von Illumina, Inc. bzw. der jeweiligen Inhaber. Spezifische Informationen zu Marken finden Sie unter www.illumina.com/company/legal.html.
M-GL-00684 DEU v3.0.