



AWS-Whitepaper

# Bewährte Methoden für Amazon Connect Data Lake



# Bewährte Methoden für Amazon Connect Data Lake: AWS-Whitepaper

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Die Handelsmarken und Handelsaufmachung von Amazon dürfen nicht in einer Weise in Verbindung mit nicht von Amazon stammenden Produkten oder Services verwendet werden, durch die Kunden irregeführt werden könnten oder Amazon in schlechtem Licht dargestellt oder diskreditiert werden könnte. Alle anderen Marken, die nicht im Besitz von Amazon sind, gehören den jeweiligen Besitzern, die möglicherweise mit Amazon verbunden sind oder von Amazon gesponsert werden.

# Table of Contents

Zusammenfassung und Einführung .....	i
Überblick .....	1
Sind Sie Well-Architected? .....	1
Einführung .....	2
Amazon Connect .....	5
Prinzipien der Gestaltung von Data Lakes .....	7
Datentypen .....	8
„Kundenprofile“ .....	8
Kontaktaufzeichnungen .....	8
Gesprächsablaufprotokolle .....	9
Kontaktlinsen-Ausgabedateien .....	9
Streams von Agentenereignissen .....	10
Sprach- und Chat-Aufzeichnungen .....	10
Integration von Drittanbietern .....	10
Lebenszyklus eines Data Lakes .....	11
Speicher .....	12
Verschlucken .....	13
Katalogisierung .....	14
Sicherheit .....	15
Überwachen .....	16
Analysen .....	16
Machine Learning .....	17
Fazit und weitere Lektüre .....	19
Weitere Informationen .....	19
Dokumentenverlauf und Mitwirkende .....	20
Mitwirkende .....	20
Hinweise .....	21
.....	xxii

# Bewährte Methoden für Amazon Connect Data Lake

Datum der Veröffentlichung: 13. Mai 2021

## Überblick

Kundenservice ist ein entscheidendes Element für den Ruf einer Marke und für den Geschäftserfolg. Kontaktzentren sind von entscheidender Bedeutung für eine wechselseitige Interaktion zwischen Agenten und Kunden und für die Bereitstellung eines erstklassigen Kundendienstenerlebnisses. Umgekehrt kann eine schlechte Erfahrung zu einer Kundenabwanderung führen. Organizations investieren in Omnichannel-Kontaktzentren, um sich einen Wettbewerbsvorteil bei der Verbesserung des Kundenerlebnisses zu verschaffen.

Heute entwickeln Unternehmen Data-Lake-Strategien, um Informationen aus den vielfältigen und ständig wachsenden Daten zu nutzen. Das schnelle Wachstum des Datenvolumens stellt die Datenverwaltung und die Speicherkapazität vor Herausforderungen. Aus der Umfrage geht hervor, dass das organische Umsatzwachstum für Unternehmen, die einen Data Lake implementiert haben, um 9% gestiegen ist.

Um die fortschrittlichsten Analyse Vorteile nutzen zu können, benötigen Unternehmen eine robuste Plattform und eine kostengünstige Lösung, um ein erfolgreiches Kontaktzentrum zu betreiben. Amazon Web Services (AWS) bietet Kunden ein umfassendes Serviceangebot und eine skalierbare Plattform, um die hohe Verfügbarkeit, Sicherheit und Belastbarkeit eines Data Lakes in der Cloud zu gewährleisten.

In diesem Whitepaper werden die Best Practices für die Architektur eines Contact-Center-Data Lakes mit [Amazon Connect](#) beschrieben.

## Sind Sie Well-Architected?

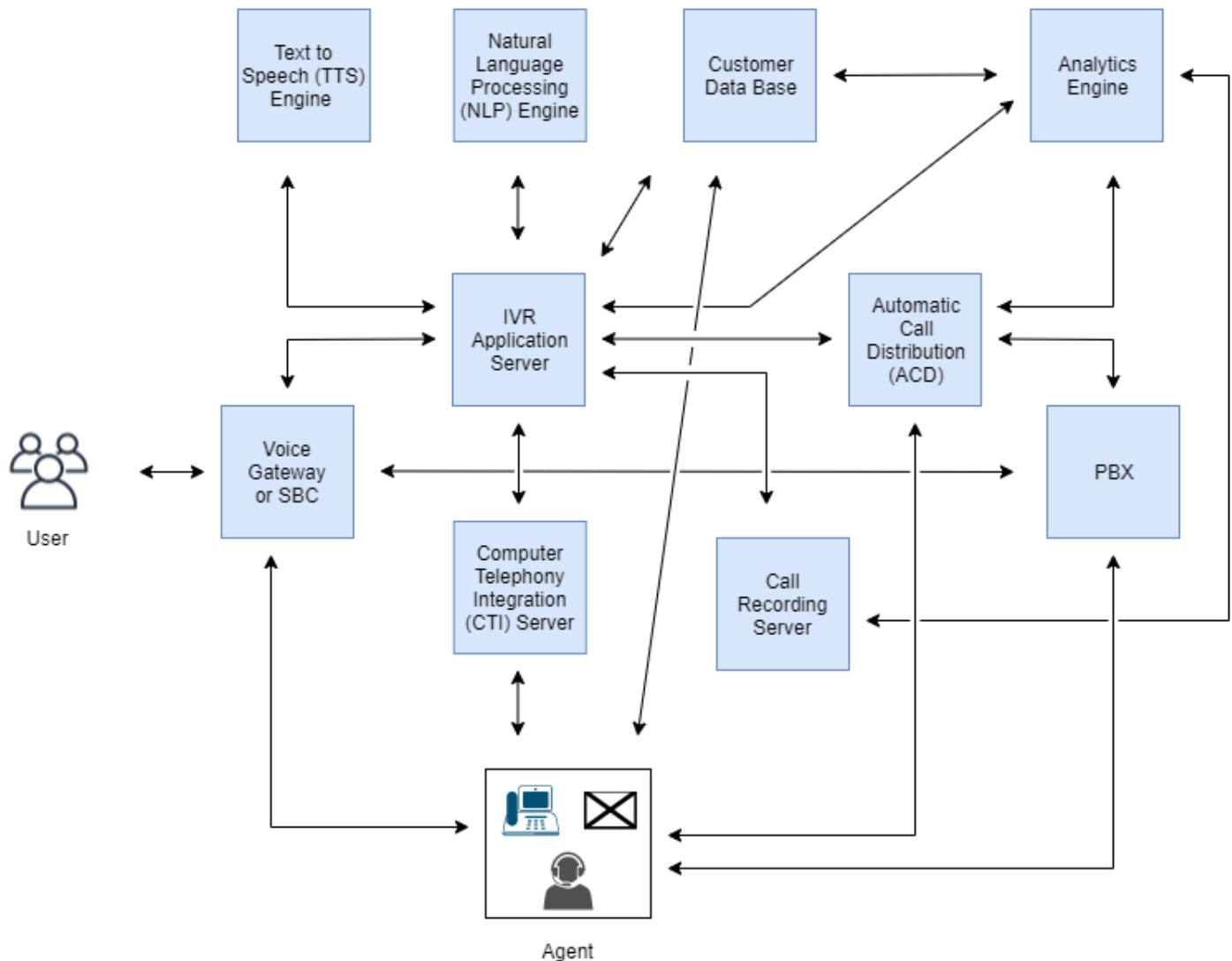
Das [Well-Architected Framework](#) hilft Ihnen dabei, die Vor- und Nachteile der Entscheidungen zu verstehen, die Sie beim Aufbau von Systemen treffen. AWS Mithilfe des Frameworks können Sie sich mit bewährten Architekturpraktiken für den Entwurf und Betrieb zuverlässiger, sicherer, effizienter und kostengünstiger Systeme in der Cloud vertraut machen.

In der [Machine Learning Lens](#) konzentrieren wir uns darauf, wie Sie Ihre Workloads für maschinelles Lernen in der AWS Cloud entwerfen, bereitstellen und gestalten können. Diese Linse ergänzt die im Well-Architected Framework beschriebenen Best Practices.

# Einführung

Herkömmliche Kontaktzentren vor Ort verwenden oft mehrere proprietäre Systeme, was zu unterschiedlichen Datenquellen mit Daten in verschiedenen Formaten führt. Herausforderungen bei der Standardisierung und Konsolidierung von Informationen verzögern die Entdeckung neuer Geschäftserkenntnisse oder möglicher betrieblicher Probleme.

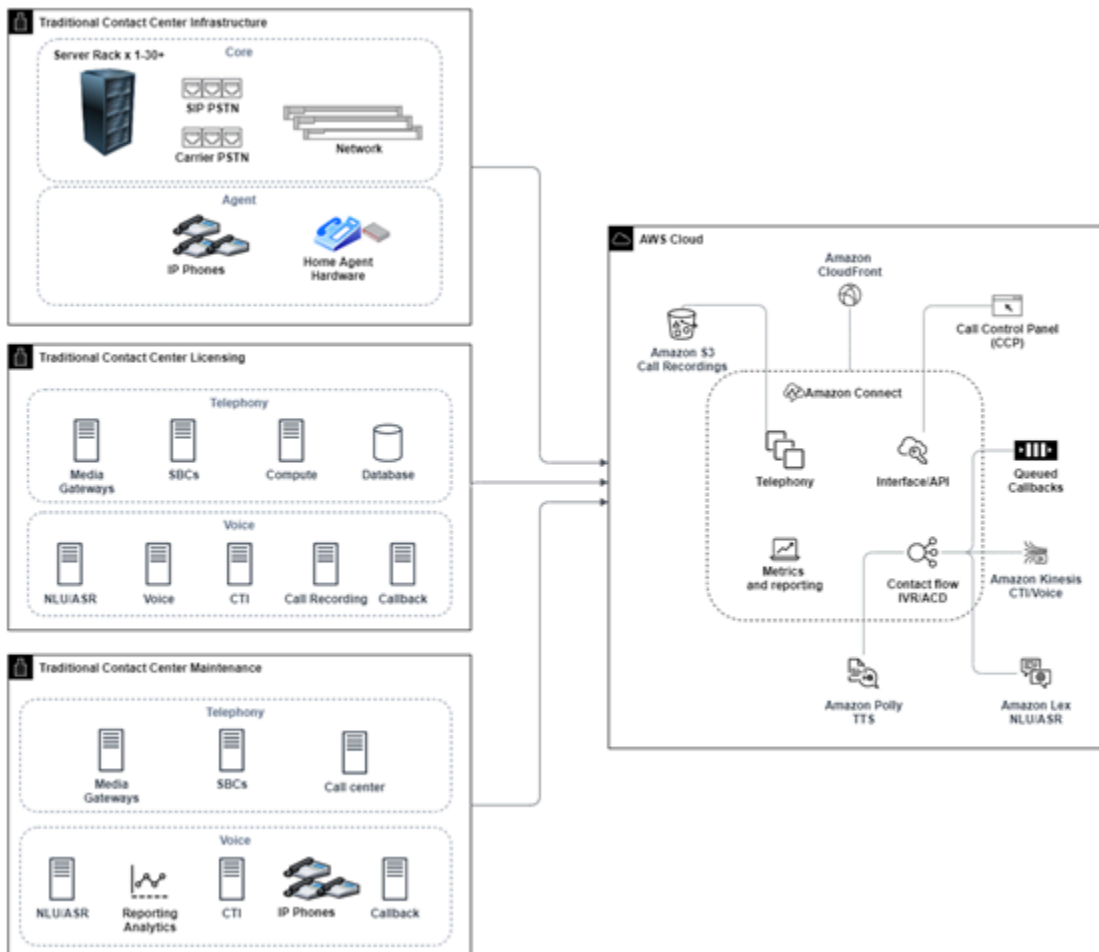
Die folgende Abbildung zeigt die Architektur eines herkömmlichen Contact Centers vor Ort.



## Lokale Contact-Center-Architektur

Daten spielen eine entscheidende Rolle für den Erfolg eines Contact Centers. Die Verwaltung von Datensilos ist ineffizient und kann zu einem schlechten Kundenerlebnis führen. Kunden erwarten ein reibungsloses Erlebnis, wenn sie sich mit einem Kontaktzentrum verbinden. Mit einer optimierten

Data-Lake-Lösung können Agenten ihren Kunden ein personalisiertes Erlebnis bieten. Die folgende Abbildung zeigt einen strategischen Ansatz zur Vereinfachung komplexer herkömmlicher Contact-Center-Daten, die sich über Infrastruktur-, Lizenz- und Wartungsumgebungen erstrecken, in [Amazon Connect](#).



Ein strategischer Ansatz zur Vereinfachung komplexer traditioneller Contact-Center-Daten in Amazon Connect

Ein Data Lake ist ein zentralisiertes, kuratiertes und sicheres Repository, das all Ihre strukturierten und unstrukturierten Daten in ihren nativen oder transformierten Formaten zur Analyse speichert und verwaltet. AWS bietet die Breite und Tiefe der Services für den Aufbau einer sicheren, skalierbaren, umfassenden und kostengünstigen [Data-Lake-Lösung](#). Sie können die AWS-Services verwenden, um Daten aus einer Vielzahl von Quellen aufzunehmen, zu speichern, zu suchen, zu verarbeiten und zu analysieren.

Dieses Whitepaper bietet architektonische Best Practices für Technologieverantwortliche wie Chief Technology Officers (CTOs), Architekten, Entwickler und Betriebsexperten beim Aufbau eines Contact-Center-Data Lakes mit Amazon Connect.

# Amazon Connect

[Amazon Connect](#) ist ein easy-to-use kostengünstiges Omnichannel-Cloud-Kontaktzentrum. Sie können innerhalb weniger Minuten mit einem vollständig verwalteten, cloudbasierten und auf künstlicher Intelligenz (KI) basierendem Kontaktzentrum beginnen. Mit dem pay-as-you-go Modell zahlen Sie nur, wenn der Service genutzt wird. Es gibt keine Infrastruktur zur Verwaltung oder Vorabkosten.

Forrester Research Consulting führte eine [Total Economic Impact \(TEI\) -Studie zu Amazon Connect durch und kam zu dem Schluss, dass sich Amazon Connect](#) drei Jahre lang finanziell darauf ausgewirkt hat, wie Amazon Connect Kunden dabei hilft, erhebliche Kosteneinsparungen zu erzielen, den Umsatz zu steigern und die Produktivität der Agenten zu verbessern. [Zu den wichtigsten Ergebnissen gehören:](#)

- Senkung der Kosten für Cloud-Technologie um 4,3 Millionen \$
- Einsparungen bei den Abonnementkosten um 31%
- Personaleinsparungen für Agenten durch ein reduziertes Anrufvolumen von 4,6 Millionen \$
- Steigerung des Betriebsergebnisses um 2,6 Millionen \$ bei verbessertem Kundenerlebnis
- Kapitalrendite (ROI) von 241%

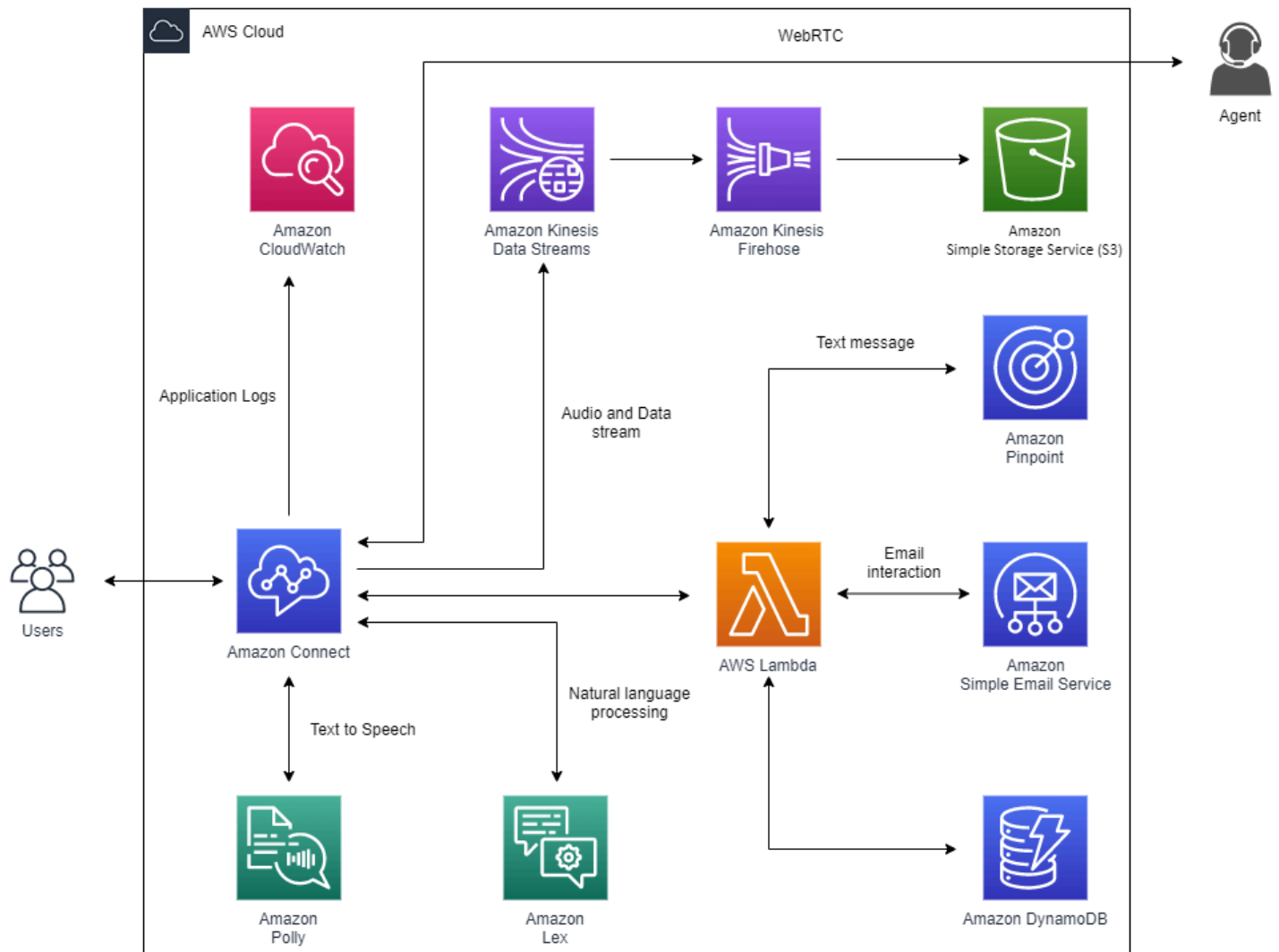
Amazon Connect bietet kompetenzbasiertes Routing, Aufgabenmanagement, leistungsstarke Echtzeit- und Verlaufsanalysen sowie intuitive Verwaltungstools. Mit Amazon Connect können Sie sich darauf konzentrieren, den Kundenservice zu verbessern und die Leistung Ihres Kontaktzentrums mühelos zu messen. Mit einem webbasierten Softphone können Agenten von jedem Standort aus schnell produktiv sein.

Mit integrierten Analysefunktionen wie [Contact Lens für Amazon Connect](#) kann das Aufsichtspersonal von Kontaktzentren die Stimmung in Bezug auf Kontaktinteraktion und betriebliche Effizienz ermitteln.

Amazon Connect ist eine offene Plattform. Mithilfe der umfangreichen veröffentlichten APIs Angebote von Amazon Connect können Sie programmgesteuert andere AWS-Services und Systeme von Drittanbietern integrieren, darunter CRM-Lösungen (Customer Relationship Management) und Betrugsbekämpfungslösungen.

Die folgende Abbildung zeigt eine Amazon Connect-Contact-Center-Architektur auf hoher Ebene. Amazon Connect bietet ein einheitliches und nahtloses Kundenerlebnis auf mehreren Kanälen.

Neben Sprach- und Webchat lässt sich Amazon Connect auch in [Amazon Pinpoint](#) und [Amazon Simple Email Service \(Amazon SES\)](#) integrieren, um die Funktionen des Kontaktzentrums für Textnachrichten und E-Mail-Zustellung zu erweitern. Amazon Connect ist in [Apple Business Chat](#) für Apple-Gerätebenutzer integriert.



Die Architektur des Amazon Connect Contact Centers

# Prinzipien des Data Lake-Designs

Der Aufbau eines Data Lakes kann Datensilos aufbrechen und Daten für die Wertextraktion demokratisieren. Ein zentraler Datenspeicher ermöglicht es Unternehmen, datengestützte Entscheidungen zu treffen und schnell Innovationen zu entwickeln.

Organizations benötigen eine kostengünstige und elastische Speicherkapazität, um unterschiedliche Datenquellen zu speichern, die exponentiell wachsen. Sie möchten riesige Datenmengen zentral verwalten und über verschiedene Geschäftsbereiche hinweg gemeinsam nutzen. Darüber hinaus möchten sie ihre Mitarbeiter und Stakeholder in die Lage versetzen, mit kürzeren Geschäftsinformationen zu gewinnen. time-to-value

Überlegungen beim Entwerfen eines Data Lakes:

- Wie sammelt, speichert und analysiert man Hochgeschwindigkeitsdaten verschiedener Datentypen, einschließlich strukturierter, unstrukturierter und halbstrukturierter Daten?
- Wie speichern und teilen Sie Petabyte an Daten auf Abruf weltweit und kostengünstig?
- Wie skalieren Sie IT-Ressourcen, um eine große Anzahl gleichzeitiger Abfragen Ihrer Daten zu unterstützen, und wie skalieren Sie sie automatisch, um Kosten zu sparen?
- Wie können Ihre Benutzer heute Abfragen in mehreren Datenrepositorien aufrufen, durchsuchen und ausführen?
- Wie leiten Sie future Erkenntnisse aus historischen Datenmustern und vergangenen Szenarien ab?

# Datentypen

Amazon Connect verwaltet eine Vielzahl von Contact-Center-Daten, darunter:

- Ressourcen und Konfigurationen wie Warteschlangen, Kontaktabläufe, Benutzer und Routing-Profile
- Kontaktmetadaten wie Verbindungszeit, Bearbeitungszeit, Quellnummer oder automatische Nummernidentifikation (ANI), Zielnummer oder Dialed Number Identification Service (DNIS) und benutzerdefinierte Kontaktattribute
- Agentenbezogene Leistungsdaten wie Anmeldezeit, Statusänderungen und behandelte Kontakte
- Audiostreams von Telefonanrufen, z. B. Anrufaufzeichnungen
- Chat-Transkripte
- Anlagen
- Integrationskonfiguration mit externen Anwendungen
- Wissensdokumente
- Sprachabdrücke zur Authentifizierung der Stimme des Kunden

Dieser Abschnitt gibt einen Überblick über die verschiedenen Datentypen, die in Amazon Connect verfügbar sind.

## „Kundenprofile“

[Amazon Connect Connect-Kundenprofile](#) ermöglichen es Agenten, einen effizienten und personalisierten Kundenservice zu bieten, indem Kundeninformationen aus verschiedenen Anwendungen in ein einheitliches Kundenprofil importiert werden. Sie können Kundendaten aus selbst entwickelten Anwendungen oder Drittanbieteranwendungen wie [Salesforce ServiceNow](#), [Zendesk](#) und [Marketo](#) mithilfe vorgefertigter Konnektoren in Ihren [Amazon Simple Storage Service \(Amazon S3\)](#) Data Lake aufnehmen.

## Kontakt Datensätze

Kontakt Datensätze erfassen Transaktionsmetriken wie Haltezeit, Wartezeit und Interaktionszeit mit Agenten im JSON-Format. Amazon Connect aggregiert Kontakt Datensatzdaten, um Metrikberichte zu

erstellen. Die Datenspeicherung für Kontaktdatensätze beträgt 24 Monate nach Kontaktabmahnung. Sie können Kontaktdatensätze zur längeren Aufbewahrung und erweiterter Analyse an [Amazon Kinesis](#) streamen. Das [Datenmodell für Kontaktdatensätze](#) beschreibt verschiedene Ereignistypen, die in Kontaktdatensätzen verfügbar sind.

## Gesprächsablaufprotokolle

[Amazon Connect Connect-Kontaktflussprotokolle](#) erfassen Ereignisse und Kennzahlen in Echtzeit darüber, wie Ihre Kunden mit Kontaktabläufen interagieren. [Amazon CloudWatch](#) erstellt eine Protokollgruppe für jede Amazon Connect Connect-Instance, wenn Sie die [Kontaktflussprotokollierung aktivieren](#), und fügen Sie einen [festgelegten Block für das Protokollierungsverhalten](#) für die Kontaktflüsse hinzu.

Kontaktflussprotokolle enthalten die Kontaktablauf-ID, die Kontakt-ID des Kunden und die Aktionen des Blocks. Mithilfe von Kontaktflussprotokollen können Sie die Interaktionen von Kunden mit verschiedenen Kontaktflussversionen vergleichen oder ihre Interaktionen in den einzelnen Kontaktabläufen verfolgen. Kontaktflussprotokolle helfen Ihnen beim Debuggen von Kontaktabläufen und beim Zurücksetzen auf frühere Versionen, falls Probleme auftreten sollten.

## Ausgabedateien für Kontaktlinsen

Mithilfe natürlicher Sprachverarbeitung (NLP) und speech-to-text Analytik bietet [Contact Lens for Amazon Connect](#) Einblicke zur Analyse der Kundenstimmung, zur Identifizierung von Gesprächstrends für Produktfeedback und zur Konformitätsprüfung für Standardgrüße und Abnahmen.

Mit der erweiterten Konversationssuche können Sie eine schnelle Volltextsuche nach relevanten Anrufen anhand von Stimmungswerten und Gesprächszeiten durchführen, um häufig vorkommende Äußerungen zu identifizieren, die zu einer positiven oder negativen Kundenstimmung führen. Contact Lens redigiert aus Datenschutzgründen automatisch sensible personenbezogene Daten (PII).

Sie können potenzielle negative Kundenerfahrungen abfangen, indem Sie Regeln für den Versand von Benachrichtigungen zu bestimmten Schlüsselwörtern oder Ausdrücken erstellen. Mitarbeiter können das Problem eskalieren und Anrufe weiterleiten und gleichzeitig die Transkripte in Echtzeit weiterleiten, um eine ordnungsgemäße Übergabe sicherzustellen.

Contact Lens speichert Metadaten für Anrufprotokolle, Stimmungsanalysen, Gesprächszeiten, Gesprächsgeschwindigkeit, Unterbrechungen und Kategorisierungskennzeichnungen in Amazon

S3. Sie können benutzerdefinierte Visualisierungs- oder ML-Modelle (Machine Learning) mithilfe von Daten aus Kontaktlisten und Kontaktdatensätzen, die in Amazon S3 gespeichert sind, erstellen.

## Streams von Agentenereignissen

[Amazon Connect Connect-Agenten-Event-Streams](#) erfassen und speichern Agentenaktivitäten in S3 über [Amazon Kinesis Data Streams](#). Sie können Dashboards für Agentenberichte fast in Echtzeit erstellen, z. B. für die Anmeldung von Agenten, die Abmeldung von Agenten, Verbindungen mit einem Kontakt und die Änderung des Agentenstatus.

Sie können Event-Streams für Agenten in Workforce Management (WFM) -Lösungen für die Personalverwaltung Ihrer Mitarbeiter integrieren oder Warnmeldungen zu bestimmten Agentenaktivitäten konfigurieren.

## Sprach- und Chat-Aufzeichnungen

Amazon Connect zeichnet eine Konversation nur auf, wenn ein Kunde eine Verbindung zu einem Agenten herstellt. Wenn der Kontakt unterbrochen wird, sind die Anrufaufzeichnungen in Ihrem S3-Bucket oder im Kontaktdatensatz des Kunden verfügbar.

Als Omnichannel-Kontaktzentrum ermöglicht [Amazon Connect Chat](#) Kunden, mit Agenten in Ihren Geschäftsanwendungen, im Internet oder auf Mobilgeräten zu chatten. Kunden können während des Chats Konversationen fortsetzen und das Gerät wechseln.

Amazon Connect redigiert, verschlüsselt und speichert Sprach- und Chat-Konversationen zwischen dem Agenten und dem Kontakt in Ihrem S3-Bucket für erweiterte Analysen.

## Integration von Drittanbietern

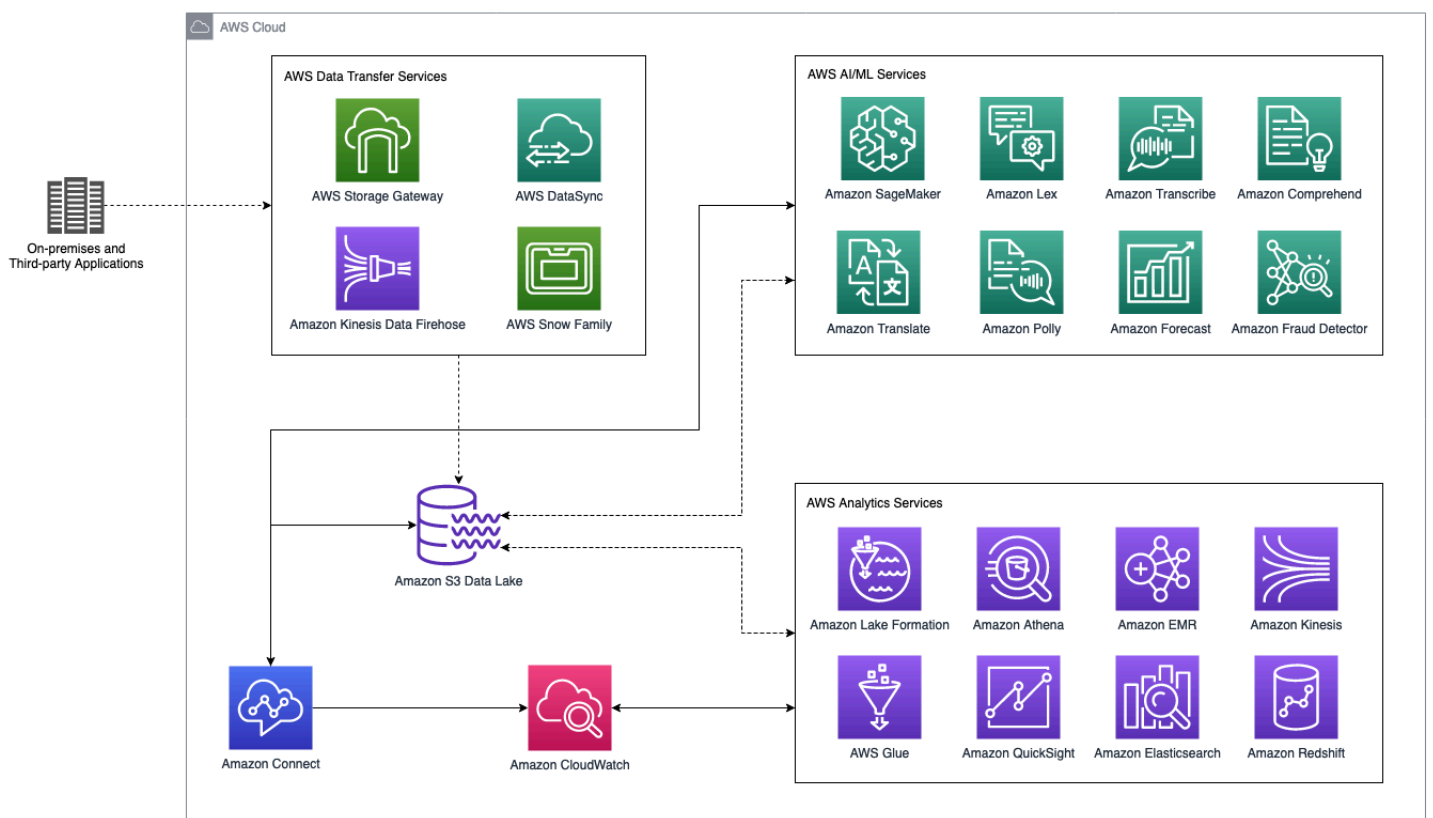
Wenn Sie [AWS-Partner](#) oder andere Drittanbieterlösungen mit Amazon Connect verwenden, können Sie Protokolle und externe Datenquellen in Amazon S3 konsolidieren.

# Lebenszyklus eines Data Lake

Der Aufbau eines Data Lakes umfasst in der Regel fünf Phasen:

- Speicher einrichten
- Daten verschieben
- Daten aufbereiten und katalogisieren
- Konfiguration von Sicherheitsrichtlinien
- Bereitstellung von Daten für den Konsum

Die folgende Abbildung zeigt ein allgemeines Architekturdiagramm eines Amazon Connect-Contact-Center-Data Lakes, der in AWS-Dienste für Analytik und künstliche Intelligenz und maschinelles Lernen (KI/ML) integriert ist. Der folgende Abschnitt behandelt die in dieser Abbildung gezeigten Szenarien und AWS-Services.



Amazon Connect Contact Center-Data Lake mit AWS-Analyse- und KI-/ML-Services

# Speicher

[Amazon S3](#) ist ein Objektspeicherservice, der branchenführende Skalierbarkeit, Datenverfügbarkeit, Sicherheit und Leistung bietet. S3 bietet eine Beständigkeit von 99,999999999% und eine Verfügbarkeit von 99,99% bei [starker](#) Konsistenz und unbegrenztem Datenspeicher weltweit. Sie können [Cross-Region Replication](#) (CRR) verwenden, um Daten zwischen S3-Buckets in mehreren Regionen zu kopieren, um gesetzliche Vorschriften einzuhalten und Anforderungen mit geringer Latenz zu erfüllen. S3 skaliert den Durchsatz automatisch für Leistung und betriebliche Effizienz.

S3-Buckets und -Objekte sind privat, wobei [S3 Block Public Access](#) standardmäßig für alle Regionen weltweit aktiviert ist. Sie können zentralisierte Zugriffskontrollen für S3-Ressourcen mithilfe von [Bucket-Richtlinien](#), [AWS Identity and Access Management](#) (IAM) -Richtlinien und [Zugriffskontrolllisten](#) (ACLs) einrichten. Mit [Access Analyzer for S3 können Sie alle Buckets mit öffentlichem Zugriff](#) auswerten und identifizieren. Mit Objektpräfixen und Tagging können Sie Zugriffskontrollen, Speichertiering und Replikationsregeln auf Objektebene detailliert verwalten.

[AWS CloudTrail protokolliert jeden API-Aufruf an die S3-Serverzugriffssprotokollierung](#). Das [S3-Inventar](#) prüft und meldet den Replikations- und Verschlüsselungsstatus Ihrer Daten.

[S3 Intelligent-Tiering](#) bietet automatische Kosteneinsparungen, indem Daten zwischen häufigen und seltenen Zugriffsstufen verschoben werden, wenn sich die Zugriffsmuster ändern, ohne dass sich dies auf die Leistung oder den Betriebsaufwand auswirkt. [S3 Glacier Deep Archive](#) spart bis zu 95% der Speicherkosten für Objekte, auf die selten zugegriffen wird und die langfristig aufbewahrt werden müssen.

[Das Speichern von Daten in Spaltenformaten wie Apache Parquet und Optimized Row Columnar \(ORC\) ermöglicht schnellere Abfragen und reduziert die Verarbeitungskosten mit Amazon Athena.](#) [Komprimierungsoptionen](#) wie [Snappy](#) with Parquet reduzieren den Kapazitätsbedarf und die Speicherkosten.

Mit [S3 Select](#) und [S3 Glacier Select](#) können Sie Objektmetadaten mithilfe eines SQL-Ausdrucks (Structured Query Language) abfragen, ohne die Objekte in einen anderen Datenspeicher verschieben zu müssen.

[S3 Batch Operations](#) automatisieren Massenoperationen an S3-Objekten, wie z. B. das Aktualisieren von Objektmetadaten und Eigenschaften, das Ausführen von Speicherverwaltungsaufgaben, das Ändern von Zugriffskontrollen und das Wiederherstellen archivierter Objekte aus [S3 Glacier](#).

[S3 Access Points](#) vereinfachen und bündeln den Zugriff auf gemeinsam genutzte Daten auf S3 durch verschiedene Teams und Anwendungen. Jeder Access Point ist mit einem eindeutigen DNS-Namen

für einen einzelnen Bucket verknüpft. Sie können [Richtlinien zur Servicesteuerung](#) (SCPs) erstellen, um Zugriffspunkte auf eine [Amazon Virtual Private Cloud](#) (Amazon VPC) einzuschränken und Daten in Ihren privaten Netzwerken zu isolieren.

[S3 Transfer Acceleration](#) ermöglicht die Dateiübertragung über große Entfernungen zwischen Ihrer Client-Umgebung und S3-Buckets.

Wenn Ihr Data Lake wächst, bietet [S3 Storage Lens](#) unternehmensweite Einblicke in die Objektspeichernutzung und die Aktivitätstrends mit umsetzbaren Empfehlungen zur Reduzierung von Kosten und Betriebskosten.

## Aufnahme

AWS bietet ein umfassendes Portfolio an Datenübertragungsservices, mit dem Sie Ihre vorhandenen Daten in einen zentralen Data Lake verschieben können. [Amazon Storage Gateway](#) und [AWS Direct Connect](#) können Hybrid-Cloud-Speicheranforderungen erfüllen. Erwägen Sie für die Online-Datenübertragung die Verwendung von [AWS DataSync](#) und [Amazon Kinesis](#). Verwenden Sie die [AWS Snow-Familie](#) für die Offline-Datenübertragung.

- AWS Storage Gateway erweitert Ihre lokalen Umgebungen auf AWS-Speicher, indem Bandbibliotheken durch Cloud-Speicher ersetzt werden, auf Cloud-Speicher gestützte Dateifreigaben bereitgestellt oder ein Cache mit niedriger Latenz für den Zugriff auf Ihre Daten in AWS aus lokalen Umgebungen erstellt wird.
- AWS Direct Connect stellt eine private Konnektivität zwischen Ihren lokalen Umgebungen und AWS her, um die Netzwerkkosten zu senken, den Durchsatz zu erhöhen und ein konsistentes Netzwerkerlebnis zu bieten.
- AWS DataSync kann Millionen von Dateien in S3, [Amazon Elastic File System](#) (Amazon EFS) oder [Amazon FSx for Windows File Server](#) übertragen und gleichzeitig die Netzwerkauslastung optimieren.
- Amazon Kinesis bietet eine sichere Möglichkeit, Streaming-Daten zu erfassen und in S3 zu laden. [Amazon Data Firehose](#) ist ein vollständig verwalteter Service für die direkte Übertragung von Echtzeit-Streaming-Daten an S3. Firehose passt sich automatisch dem Volumen und dem Durchsatz der Streaming-Daten an und erfordert keine laufende Verwaltung. Sie können Streaming-Daten mithilfe von Komprimierung, Verschlüsselung, Datenstapelung oder [AWS Lambda](#)-Funktionen in Firehose transformieren, bevor Sie Daten in S3 speichern. Die Firehose-Verschlüsselung unterstützt die serverseitige S3-Verschlüsselung mit [AWS Key Management Service](#) (AWS KMS). Alternativ können Sie die Daten mit Ihrem benutzerdefinierten Schlüssel

verschlüsseln. Firehose kann mehrere eingehende Datensätze zu einem einzigen S3-Objekt verketteten und bereitstellen, um die Kosten zu senken und den Durchsatz zu optimieren.

AWS Snow Family bietet einen Offline-Datenübertragungsmechanismus. [AWS Snowball Edge](#) bietet ein tragbares und robustes Edge-Computing-Gerät für die Datenerfassung, -verarbeitung und -migration. Für Datenübertragungen im Exabyte-Bereich können Sie [AWS Snowmobile](#) verwenden, um riesige Datenmengen in die Cloud zu verschieben.

[DistCp](#) bietet eine verteilte Kopierfunktion zum Verschieben von Daten im Hadoop-Ökosystem. [S3 DistCp](#) ist eine Erweiterung, die für die DistCp Übertragung von Daten zwischen Hadoop Distributed File System (HDFS) und S3 optimiert ist. [Dieser Blog](#) enthält Informationen zum Verschieben von Daten zwischen HDFS und S3 mithilfe von S3. DistCp

## Katalogisierung

Ein häufiges Problem bei einer Data-Lake-Architektur ist die mangelnde Kontrolle über den Inhalt der im Data Lake gespeicherten Rohdaten. Organizations benötigen Governance, semantische Konsistenz und Zugriffskontrollen, um die Fallstricke zu vermeiden, die mit der Schaffung eines Datensumpfs ohne Pflege einhergehen.

[AWS Lake Formation](#) kann die Datenaufnahme verwalten, [AWS Glue](#) indem Daten automatisch klassifiziert und Definitionen, Schemata und Metadaten in einem zentralen Datenkatalog gespeichert werden. Lake Formation verfügt über integrierte Funktionen für maschinelles Lernen zur Deduplizierung und zur Suche nach passenden Datensätzen, um die Datenqualität zu verbessern. Für schnellere Analysen konvertiert Lake Formation Daten in Apache Parquet und ORC, bevor sie in Ihrem S3 Data Lake gespeichert werden. Sie können Zugriffsrichtlinien definieren, einschließlich Zugriffskontrollen auf Tabellen- und Spaltenebene, oder Datenverschlüsselung im Ruhezustand erzwingen. Durch die konsequente Durchsetzung der Sicherheitsvorkehrungen können Ihre Benutzer mithilfe von Analyse- und Machine-Learning-Diensten ihrer Wahl auf einen kuratierten und zentralisierten Datensatz zugreifen und diesen analysieren.

[AWS Glue DataBrew](#), ein Tool zur visuellen Datenaufbereitung, ermöglicht es Dateneigentümern, Fachexperten oder Benutzern aller Fachrichtungen, am Datenaufbereitungsprozess teilzunehmen. Ohne Code schreiben zu müssen, können Ihre Teams aus über 250 vorgefertigten Transformationen wählen, um Datenvorbereitungsaufgaben zu automatisieren, darunter das Filtern von Datenanomalien, das Konvertieren von Daten in Standardformate und das Korrigieren ungültiger Werte. Die transformierten Daten sind bereit für fortgeschrittene Analysen und Machine-Learning-Projekte.

## Sicherheit

Amazon Connect trennt Daten nach AWS-Konto-ID und Amazon Connect Connect-Instance-ID, um den autorisierten Datenzugriff auf Amazon Connect Connect-Instance-Ebene sicherzustellen.

Amazon Connect verschlüsselt personenbezogene Daten (PII), Kontaktdaten und Kundenprofile im Speicher mithilfe eines zeitlich begrenzten Schlüssels, der für Ihre Amazon Connect Connect-Instanz spezifisch ist. Die serverseitige S3-Verschlüsselung schützt sowohl Sprach- als auch Chat-Aufzeichnungen im Ruhezustand mithilfe eines für jedes AWS-Konto eindeutigen KMS-Datenschlüssels. Sie behalten die vollständige Sicherheitskontrolle bei, um den Benutzerzugriff auf Anrufaufzeichnungen in Ihrem S3-Bucket zu konfigurieren, einschließlich der [Nachverfolgung, wer Anrufaufzeichnungen abhört oder löscht](#). Amazon Connect verschlüsselt die Stimmabdrücke der Kunden mit einem diensteeigenen KMS-Schlüssel, um die Kundenidentität zu schützen. Alle zwischen Amazon Connect und anderen AWS-Services oder externen Anwendungen ausgetauschten Daten [werden bei der Übertragung immer mit der branchenüblichen Transport Layer Security \(TLS\) - Verschlüsselung verschlüsselt](#).

Die Sicherung eines Data Lakes erfordert detaillierte Kontrollen, um den autorisierten Datenzugriff und die Nutzung sicherzustellen. S3-Ressourcen sind privat und standardmäßig nur für ihren Ressourcenbesitzer zugänglich. Der Ressourcenbesitzer kann eine Kombination aus ressourcen- oder identitätsbasierten IAM-Richtlinien erstellen, um Berechtigungen für S3-Buckets und -Objekte zu gewähren und zu verwalten. Ressourcenbasierte Richtlinien, wie z. B. Bucket-Richtlinien, sind an Ressourcen angehängt. ACLs Im Gegensatz dazu sind identitätsbasierte Richtlinien an die IAM-Benutzer, -Gruppen oder -Rollen in Ihrem AWS-Konto angehängt.

Wir empfehlen für die meisten Data Lake-Umgebungen [identitätsbasierte Richtlinien](#), um die Verwaltung des Ressourcenzugriffs und die Serviceberechtigungen für Ihre Data Lake-Benutzer zu vereinfachen. Sie können IAM-Benutzer, -Gruppen und -Rollen in AWS-Konten erstellen und sie identitätsbasierten Richtlinien zuordnen, die Zugriff auf S3-Ressourcen gewähren.

[Das AWS Lake Formation Berechtigungsmodell](#) funktioniert in Verbindung mit [IAM-Berechtigungen, um den Zugriff auf den Data Lake](#) zu steuern. Das Lake Formation Formation-Berechtigungsmodell verwendet einen GRANT- oder REVOKE-Mechanismus im Stil eines Datenbankmanagementsystems (DBMS). IAM-Berechtigungen enthalten identitätsbasierte Richtlinien. Beispielsweise muss ein Benutzer die Berechtigungsprüfungen sowohl durch IAM- als auch durch Lake Formation Formation-Berechtigungen bestehen, bevor er auf eine Data Lake-Ressource zugreifen kann.

AWS CloudTrail verfolgt Amazon Connect Connect-API-Aufrufe, einschließlich der IP-Adresse und Identität des Anfragenden sowie Datum und Uhrzeit der Anfrage im [CloudTrail Ereignisverlauf](#).

Die Erstellung eines AWS CloudTrail Trails ermöglicht die kontinuierliche Bereitstellung von AWS CloudTrail Protokollen an Ihren S3-Bucket.

[Amazon Athena Workgroups](#) können die Ausführung von Abfragen trennen und den Zugriff durch Benutzer, Teams oder Anwendungen mithilfe ressourcenbasierter Richtlinien kontrollieren. Sie können die Kostenkontrolle durchsetzen, indem Sie die Datennutzung in den Arbeitsgruppen [einschränken](#).

## Überwachen

Beobachtbarkeit ist unerlässlich, um die Verfügbarkeit, Zuverlässigkeit und Leistung eines Kontaktzentrums und eines Data Lakes sicherzustellen. [Amazon CloudWatch](#) bietet systemweiten Einblick in die Ressourcennutzung, die Anwendungsleistung und den Betriebszustand. Protokollieren Sie relevante Informationen aus den Amazon Connect Connect-Kontaktabläufen bei Amazon CloudWatch und erstellen Sie Benachrichtigungen in Echtzeit, wenn die Betriebsleistung unter vordefinierte Schwellenwerte fällt.

Amazon Connect sendet die Nutzungsdaten der Instance als CloudWatch Amazon-Metriken im Abstand von einer Minute. Die Datenspeicherung für CloudWatch Amazon-Metriken beträgt zwei Wochen. Definieren Sie frühzeitig die Anforderungen für die Aufbewahrung von Protokollen und Lebenszyklusrichtlinien, um die Einhaltung gesetzlicher Vorschriften sicherzustellen und Kosten für die langfristige Datenarchivierung zu sparen.

[Amazon CloudWatch Logs](#) bietet eine einfache Möglichkeit, Protokolldaten zu filtern und Verstöße zu identifizieren, um Vorfälle zu untersuchen und deren Lösung zu beschleunigen. Sie können die Kontaktabläufe anpassen, um Anrufer mit hohem Risiko oder potenziell betrügerische Aktivitäten zu erkennen. Sie können beispielsweise die Verbindung zu allen eingehenden Kontakten trennen, die auf Ihrer vordefinierten Ablehnungsliste stehen.

## Analysen

Ein Data Lake für Kontaktzentren, der auf einem Portfolio mit beschreibenden, vorausschauenden Analysen und Echtzeitanalysen basiert, hilft Ihnen dabei, aussagekräftige Erkenntnisse zu gewinnen und wichtige Geschäftsfragen zu beantworten.

Sobald Ihre Daten im S3-Datensee gelandet sind, können Sie alle speziell entwickelten Analysedienste wie Amazon Athena und [Amazon QuickSight](#) für eine Vielzahl von Anwendungsfällen

ohne arbeitsintensive ETL-Jobs (Extrahieren, Transformieren und Laden) verwenden. Alternativ können Sie Ihre bevorzugten Analyseplattformen in Ihren S3-Data Lake integrieren. [In diesem Blog](#) finden Sie eine exemplarische Vorgehensweise zur Analyse von Amazon Connect Connect-Daten mit Amazon Athena und Amazon. AWS Glue QuickSight

Für eine hoch skalierbare Data Warehousing-Lösung können Sie [Datenstreaming in Amazon Connect aktivieren](#), um Kontaktdatenätze über Amazon Kinesis in [Amazon Redshift](#) zu streamen.

## Machine Learning

Der Aufbau eines Data Lakes bringt ein neues Paradigma in die Contact-Center-Architektur und ermöglicht es Ihrem Unternehmen, mithilfe von Funktionen für maschinelles Lernen (ML) einen verbesserten und personalisierten Kundenservice zu bieten.

Die herkömmliche ML-Entwicklung ist ein komplexer und teurer Prozess. AWS bietet die Tiefe und Breite einer leistungsstarken, kostengünstigen, skalierbaren Infrastruktur und flexiblen [ML-Services für jedes ML-Projekt](#) oder jede Arbeitslast.

[Amazon SageMaker AI](#) ist ein vollständig verwalteter Service, der es Ihren Datenwissenschaftlern und Entwicklern ermöglicht, ML-Modelle für Contact-Center-Anwendungsfälle in großem Umfang zu erstellen, zu trainieren und bereitzustellen. Die Datenvorbereitung macht bis zu 80% der Zeit von Datenwissenschaftlern aus. [Amazon SageMaker AI Data Wrangler](#) vereinfacht und beschleunigt die Datenaufbereitung und Feature-Engineering aus verschiedenen Datenquellen mithilfe von über 300 integrierten Datentransformationen, ohne dass Code geschrieben werden muss. Sie können standardisierte Funktionen im [Amazon SageMaker AI Feature Store speichern](#), um sie wiederzuverwenden und mit dem Rest Ihrer Organisation zu teilen.

Um Kundenabwanderungen zu vermeiden, ist es wichtig, die Reibung während einer Kundenreise zu reduzieren. Um Ihr Kontaktzentrum intelligenter zu [gestalten, können Sie KI-gestützte Konversations-Chatbots](#) mithilfe der Funktionen für automatische Spracherkennung (ASR) und Natural Language Understanding (NLU) von [Amazon Lex](#) erstellen. Kunden können Self-Service-Aufgaben wie das Zurücksetzen von Passwörtern, die Überprüfung des Kontostands und die Terminplanung über Chatbots durchführen, ohne mit den menschlichen Mitarbeitern sprechen zu müssen. Um die häufig gestellten Fragen (FAQs) des Kontaktzentrums zu automatisieren, können Sie mit Amazon Lex [und Amazon Kendra einen Frage-und-Antwort-Chatbot \(Q&A\)](#) erstellen. Durch die Aktivierung der Textprotokollierung in Amazon CloudWatch Logs und das Speichern von Audioeingaben in S3 können Sie den Gesprächsfluss analysieren, das Gesprächsdesign verbessern und die Benutzerinteraktion erhöhen.

Um die allgemeine Servicequalität zu verbessern, ist es wichtig, die Dynamik zwischen Anrufer und Agent zu verstehen. [In diesem Blog erfahren Sie, wie Sie Sprachaufzeichnungen über Kinesis Video Stream zur Spracherkennung an Amazon Transcribe streamen und mit Amazon Comprehend Audio in Text umwandeln und Stimmungsanalysen für die Transkripte durchführen können.](#)

Für Unternehmen mit internationaler Präsenz können Sie mit Amazon [Polly](#) oder Amazon [Translate](#) für die Sprachübersetzung ein mehrsprachiges Spracherlebnis in Amazon [Connect](#) einrichten.

Herkömmliche Finanzplanungssoftware erstellt Prognosen auf der Grundlage historischer Zeitreihendaten, ohne inkonsistente Trends und relevante Variablen zu korrelieren. [Amazon Forecast](#) bietet eine um bis zu 50% höhere Genauigkeit mithilfe von maschinellem Lernen, um die zugrunde liegende Beziehung zwischen Zeitreihendaten und anderen Variablen wie Produktmerkmalen und Filialstandorten zu ermitteln. Da keine Erfahrung mit maschinellem Lernen erforderlich ist, können Sie ganz einfach eine Bedarfs- oder Inventarprognose für Agenten erstellen, indem Sie Amazon Forecast Zeitreihen und zugehörige Daten in Ihrem S3-Bucket zur Verfügung stellen. Sie können vertrauliche Inhalte mit Amazon Forecast verschlüsseln AWS KMS und den Zugriff auf Amazon Forecast mithilfe der IAM-Richtlinie kontrollieren. Amazon Forecast trainiert und hostet ein benutzerdefiniertes Modell für maschinelles Lernen in einer hochverfügbaren Umgebung. Sie können schnell hochgenaue Geschäftsprognosen erstellen, ohne eine Infrastruktur oder einen komplexen maschinellen Lernprozess verwalten zu müssen.

Amazon Connect stellt Anrufattribute von Telefonanbietern bereit, z. B. den geografischen Standort der Sprachausrüstung, um zu zeigen, woher der Anruf stammt, Telefongerätetypen wie Festnetz oder Mobilfunk, Anzahl der Netzwerksegmente, die der Anruf durchquert hat, und andere Informationen zur Anruferzeugung. Mit dem vollständig verwalteten [Amazon Fraud Detector](#) können Sie ein ML-Modell erstellen, um potenziell betrügerische Aktivitäten zu identifizieren, indem Sie Ihre Datensätze mit Amazon Connect Connect-Anrufattributen kombinieren. Sie können beispielsweise den Kontaktablauf so anpassen, dass Telefonanrufe mit potenziellen Betrugssignalen intelligent an einen spezialisierten Agenten weitergeleitet werden.

## Fazit und weitere Lektüre

Amazon Connect ist ein speziell entwickeltes Omnichannel-Cloud-Kontaktzentrum, das Ihren Kunden und Mitarbeitern ein nahtloses und reibungsloses Erlebnis bietet. Mit Amazon Connect können Sie Abläufe vereinfachen, die Effizienz Ihrer Agenten verbessern und die Kosten für Ihr Kontaktzentrum senken.

Amazon S3 ist ein skalierbarer, langlebiger und zuverlässiger Service zum Aufbau und zur Verwaltung eines sicheren Data Lake in großem Maßstab für Kontaktzentren. Sie können all Ihre Contact-Center-Daten unverändert im S3 Data Lake speichern, ohne die Daten neu strukturieren zu müssen, wodurch die Wertgewinnung mit kürzeren Daten beschleunigt wird. time-to-value Ihre Mitarbeiter und Stakeholder können verschiedene Analysen auf dem Data Lake des Kontaktzentrums durchführen, darunter Big-Data-Verarbeitung, Echtzeit-Dashboards und -Visualisierungen sowie maschinelles Lernen als Grundlage für datengestützte Geschäftsentscheidungen.

Ein effizienter und optimierter Contact-Center-Data Lake kann ein wichtiger Faktor für die Verbesserung des Kundenerlebnisses und die Steigerung der Marktakzeptanz sein. Mit einem umfassenden Portfolio an Analyseservices und einer skalierbaren Infrastruktur auf AWS können Sie die Leistung und Intelligenz Ihres Contact-Center-Data Lake nutzen, um das Geschäftswachstum zu beschleunigen.

## Weitere Informationen

Weitere Informationen finden Sie unter:

- [Data Lake-Speicher auf AWS](#)
- [Analytik auf AWS](#)

# Dokumentenverlauf und Mitwirkende

Abonnieren Sie den RSS-Feed, um über Aktualisierungen des Whitepapers benachrichtigt zu werden.

Änderung	Beschreibung	Datum
<a href="#">Erste Veröffentlichung</a>	Whitepaper zuerst veröffentlicht	13. Mai 2021

## Mitwirkende

Zu den Mitwirkenden an diesem Dokument gehören:

- Ankur Taunk, leitender spezialisierter Lösungsarchitekt: Amazon Connect, Amazon Web Services
- Cher Simon, Senior Partner Solutions Architect, Amazon Web Services

# Hinweise

Kunden sind dafür verantwortlich, Ihre eigene unabhängige Bewertung der Informationen in diesem Dokument vorzunehmen. Dieses Dokument: (a) dient nur zu Informationszwecken, (b) stellt aktuelle AWS-Produktangebote und -praktiken dar, die ohne vorherige Ankündigung geändert werden können, und (c) stellt keine Verpflichtungen oder Zusicherungen von AWS und seinen verbundenen Unternehmen, Lieferanten oder Lizenzgebern dar. AWS-Produkte oder -Services werden „wie sie sind“ ohne ausdrückliche oder stillschweigende Garantien, Zusicherungen oder Bedingungen jeglicher Art bereitgestellt. Die Verantwortung und Haftung von AWS gegenüber seinen Kunden werden durch AWS-Vereinbarungen geregelt. Dieses Dokument gehört, weder ganz noch teilweise, nicht zu den Vereinbarung von AWS mit seinen Kunden und ändert diese Vereinbarungen auch nicht.

© 2021 Amazon Web Services, Inc. oder seine Tochtergesellschaften. Alle Rechte vorbehalten.

Die vorliegende Übersetzung wurde maschinell erstellt. Im Falle eines Konflikts oder eines Widerspruchs zwischen dieser übersetzten Fassung und der englischen Fassung (einschließlich infolge von Verzögerungen bei der Übersetzung) ist die englische Fassung maßgeblich.