



(12) **Veröffentlichung**

der internationalen Anmeldung mit der
(87) Veröffentlichungs-Nr.: **WO 2021/090091**
in der deutschen Übersetzung (Art. III § 8 Abs. 2
IntPatÜbkG)
(21) Deutsches Aktenzeichen: **11 2020 004 652.4**
(86) PCT-Aktenzeichen: **PCT/IB2020/059698**
(86) PCT-Anmeldetag: **15.10.2020**
(87) PCT-Veröffentlichungstag: **14.05.2021**
(43) Veröffentlichungstag der PCT Anmeldung
in deutscher Übersetzung: **15.06.2022**

(51) Int Cl.: **G06F 16/172 (2019.01)**
G06V 20/13 (2022.01)

(30) Unionspriorität:
16/678,375 08.11.2019 US

(71) Anmelder:
**INTERNATIONAL BUSINESS MACHINES
CORPORATION, Armonk, NY, US**

(74) Vertreter:
**Richardt Patentanwälte PartG mbB, 65185
Wiesbaden, DE**

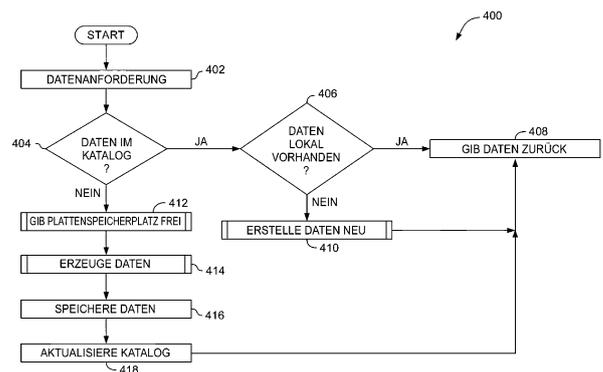
(72) Erfinder:
**Correia Villa Real, Lucas, São Paulo, BR; Santos,
Marcelo Nery dos, Rio De Janeiro, BR; Souza,
Renan Francisco Santos, Rio de Janeiro, BR**

Prüfungsantrag gemäß § 44 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **FORTLAUFENDE SPEICHERUNG VON DATEN IN EINEM SYSTEM MIT BEGRENZTER
SPEICHERKAPAZITÄT**

(57) Zusammenfassung: Techniken zum automatischen Wiederherstellen von aus einem Speicher entfernten Daten werden offenbart. Eine Anforderung für einen Zugriff auf Daten an einem ersten Speicherort wird empfangen. Es wird festgestellt, auf der Grundlage eines Katalogs und unter Verwendung eines Computerprozessors, dass die Daten aus dem ersten Speicherort entfernt wurden. Als Reaktion darauf werden die Daten an dem ersten Speicherort unter Verwendung des Computerprozessors automatisch wiederhergestellt. Das Wiederherstellen umfasst ein Angeben eines zweiten Speicherorts für die Daten auf der Grundlage des Katalogs, ein Abrufen der Daten aus dem zweiten Speicherort sowie ein Speichern der Daten an dem ersten Speicherort. Der Katalog wird aktualisiert, um das Wiederherstellen der Daten an dem ersten Speicherort widerzuspiegeln.



Beschreibung

HINTERGRUND

[0001] Die vorliegende Offenbarung betrifft die elektronische Datenspeicherung und insbesondere eine fortlaufende Speicherung von Daten in einem System mit begrenzter Speicherkapazität.

[0002] Es kommt immer häufiger vor, dass Forscher sehr große öffentliche Datensätze (wie beispielsweise Satellitenbilder) auswerten, um bei verschiedenen kritischen Aufgaben zu helfen. In einigen Fällen werden die öffentlichen Daten mit anderen Informationen kombiniert, um zusätzlichen Daten abzuleiten. Da diese Datensätze jedoch sehr groß sein können, kommt es häufig vor, dass der Speicherplatz für die Daten während der Auswertung knapp wird. Zum Beispiel können geologische oder Umweltforscher Satellitenbilder von einem ganzen Land auswerten. Versuche, aus den Satellitenbildern abgeleitete Informationen zu sichern, wie Daten über Pflanzengesundheit, den Status von verschiedenen Straßen und Gewässern usw., kann zu einem so großen Datensatz führen, dass dem Forscher der Speicherplatz ausgeht. In vorhandenen Systemen kann, wenn ein Problem wie dieses vorkommt, ein Fehler auftreten oder Benutzer müssen entscheiden, welche Daten aus dem Speichersystem gelöscht werden sollen, bevor neue Daten aufgenommen und verarbeitet werden können, was eine Koordinierung zwischen Benutzern erforderlich machen und dazu führen kann, dass unerwünschte Daten gelöscht werden. Dies ist zeitaufwendig und kann schwierig zu bewerkstelligen sein, da die Dateneigner möglicherweise nicht mehr da sind und den Benutzern, die Speicherplatz konsolidieren möchten, keine Informationen darüber zur Verfügung stehen, wie abgeleitete Informationen neu erstellt werden können.

KURZDARSTELLUNG

[0003] Ausführungsformen umfassen ein durch einen Computer ausgeführtes Verfahren. Das Verfahren umfasst ein Empfangen einer Anforderung für einen Zugriff auf Daten an einem ersten Speicherort. Das Verfahren umfasst des Weiteren ein Feststellen, auf der Grundlage eines Katalogs und unter Verwendung eines Computerprozessors, dass die Daten aus dem ersten Speicherort entfernt wurden, und als Reaktion darauf ein automatisches Wiederherstellen der Daten an dem ersten Speicherort unter Verwendung des Computerprozessors. Das Wiederherstellen weist ein Angeben eines zweiten Speicherorts für die Daten auf der Grundlage des Katalogs, ein Abrufen der Daten aus dem zweiten Speicherort sowie ein Speichern der Daten an dem ersten Speicherort auf. Das Verfahren umfasst des Weiteren ein Aktualisieren des Katalogs, um ein Wiederherstellen der Daten an dem ersten Speicherort widerzuspiegeln.

[0004] Ausführungsformen umfassen des Weiteren ein System. Das System weist einen Prozessor und einen Hauptspeicher auf, der ein Programm enthält, das, wenn es auf dem Prozessor ausgeführt wird, eine Operation durchführt. Die Operation weist ein Empfangen einer Anforderung für einen Zugriff auf Daten an einem ersten Speicherort auf. Die Operation weist des Weiteren ein Feststellen, auf der Grundlage eines Katalogs und unter Verwendung eines Computerprozessors, dass die Daten aus dem ersten Speicherort entfernt wurden, und als Reaktion darauf ein automatisches Wiederherstellen der Daten an dem ersten Speicherort unter Verwendung des Computerprozessors auf. Das Wiederherstellen weist ein Angeben eines zweiten Speicherorts für die Daten auf der Grundlage des Katalogs, ein Abrufen der Daten aus dem zweiten Speicherort sowie ein Speichern der Daten an dem ersten Speicherort auf. Die Operation weist des Weiteren ein Aktualisieren des Katalogs auf, um das Wiederherstellen der Daten an dem ersten Speicherort widerzuspiegeln.

[0005] Zu Ausführungsformen gehören des Weiteren ein Computerprogrammprodukt, um aus einem Speicher entfernte Daten automatisch wiederherzustellen, wobei das Computerprogrammprodukt ein durch einen Computer lesbares Speichermedium mit damit verkörpertem, durch einen Computer lesbaren Programmcode aufweist, wobei der durch einen Computer lesbare Programmcode durch einen oder mehrere Computerprozessoren ausführbar ist, um eine Operation durchzuführen. Die Operation weist ein Empfangen einer Anforderung für einen Zugriff auf Daten an einem ersten Speicherort auf. Die Operation weist des Weiteren ein Feststellen, auf der Grundlage eines Katalogs und unter Verwendung eines Computerprozessors, dass die Daten aus dem ersten Speicherort entfernt wurden, und als Reaktion darauf ein automatisches Wiederherstellen der Daten an dem ersten Speicherort unter Verwendung des Computerprozessors auf. Das Wiederherstellen weist ein Angeben eines zweiten Speicherorts für die Daten auf der Grundlage des Katalogs, ein Abrufen der Daten aus dem zweiten Speicherort sowie ein Speichern der Daten an dem ersten Speicherort auf. Die Operation weist des Weiteren ein Aktualisieren des Katalogs auf, um das Wiederherstellen der Daten an dem ersten Speicherort widerzuspiegeln.

Figurenliste

[0006] Ausführungsformen der Erfindung werden nun lediglich beispielhaft unter Bezugnahme auf die beiliegenden Zeichnungen beschrieben, bei denen:

Fig. 1 eine Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform veranschaulicht;

Fig. 2 Software-Services für eine fortlaufende Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform veranschaulicht;

Fig. 3 einen Server veranschaulicht, der Software-Services für eine fortlaufende Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform ausführt;

Fig. 4 ein Ablaufplan für eine fortlaufende Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform ist;

Fig. 5 ein Ablaufplan zum Neuerstellen von Daten als Teil einer fortlaufenden Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform ist;

Fig. 6 ein Ablaufplan zum Freigeben von Plattenspeicherplatz als Teil einer fortlaufenden Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform ist;

Fig. 7 ein Ablaufplan zum Neuerstellen von angeforderten Daten als Teil einer fortlaufenden Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform ist.

AUSFÜHRLICHE BESCHREIBUNG

[0007] Eine oder mehrere hierin offenbarte Ausführungsformen betreffen Software-Services, um es Benutzern zu ermöglichen, Daten, die aus der Ferne zugänglich sind, nahtlos an einer lokalen Speicherposition zu speichern, ohne sich über die Speicherkapazität Gedanken zu machen. In einer Ausführungsform werden ferne Datenquellen, aus denen Benutzer Daten abgerufen haben, in einem Katalogsystem aufgezeichnet. Statistiken und Informationen, die zu dem Datenabruf gehören (z.B. Zeitspanne für das Herunterladen, Dateigröße, in Anspruch genommene Ressourcen) werden ebenfalls in dem Katalog gespeichert. Die Software-Services verwenden diese Informationen, um, wenn sie mit begrenzter Speicherkapazität konfrontiert sind, automatisch festzustellen, welche Daten am leichtesten abgerufen und wiederhergestellt werden können. Die Software-Services können dann Speicherplatz freigeben, indem sie die leichter abgerufenen und wiederhergestellten Daten löschen. Wenn ein Benutzer diese Daten später anfordert, können sie nahtlos abgerufen und wiederhergestellt werden.

[0008] Des Weiteren können die Software-Services in einer oder mehreren Ausführungsformen Skripte überwachen, die verwendet werden, um heruntergeladene Daten zu verarbeiten und neue Daten abzuleiten. Informationen, die diese Skripte angeben, können zusammen mit den Skripten bereitzustellenden Eingabeparametern ebenfalls gesichert und bei der Feststellung, welche Daten gelöscht werden sollen, um Speicherplatz freizugeben, als ein Faktor verwendet werden. Als Teil der Wiederherstellung der Daten können die Services diese Skripte und beliebige Eingabeparameter verwenden, um die heruntergeladenen Daten zu verarbeiten und die abgeleiteten Daten wiederherzustellen.

[0009] **Fig. 1** veranschaulicht eine Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform. Eine Datenverarbeitungsumgebung 100 umfasst einen oder mehrere Benutzer 102. Diese Benutzer 102 greifen auf einen Datenabrufservice 122 zu, um Daten aus einer Datenquelle 110 abzurufen. In einer Ausführungsform ist die Datenquelle 110 öffentlich zugänglich. Zum Beispiel könnten die Benutzer 102 Forscher sein, die öffentlich verfügbare geografische Daten verwenden. Bei der Datenquelle 110 könnte es sich um eine öffentliche Datenrepository oder Website handeln. Als ein weiteres Beispiel kann die Datenquelle 110 eine Repository sein, die unter Verwendung des Internets oder eines anderen Weitverkehrsnetzes zugänglich ist. Ein Benutzer 102 kann eine Benutzereinheit (z.B. einen Laptop, einen Desktop, ein Smartphone, einen Tablet-Computer usw.) verwenden, um eine Anforderung für Daten aus der Datenquelle 110 zu übertragen. Diese Anforderung kann von einer beliebigen geeigneten Software auf der Benutzereinheit stammen, darunter einem Webbrowser, einer Forschungsanwendung usw.

[0010] Der Datenabrufservice 122 ruft Daten aus der Datenquelle 110 ab. Der Datenabrufservice 122 speichert die Daten dann in einem Speicher 150 (z.B. einer elektronischen Datenbank, einer Cloud-Speicherposi-

tion oder einem Dateisystem) und verwendet einen Katalog 140, um Einzelheiten über die Herkunft und den Abruf der Daten zu speichern. In einer Ausführungsform haben die in dem Katalog 140 gespeicherten Informationen die Funktion von Metadaten, die Einzelheiten über die Herkunft und den Abruf der Daten angeben. Bei dem Katalog 140 kann es sich zum Beispiel um eine Tabelle in einer elektronischen Datenbank (z.B. einer relationalen Datenbank, einer Graph-Datenbank oder einer beliebigen anderen geeigneten Datenbank) handeln. In einer Ausführungsform befindet sich der Speicher 150 auf einer durch die Benutzer 102 verwendeten Datenverarbeitungseinheit. Alternativ ist der Speicher 150 für die Benutzer 102 unter Verwendung eines lokalen Netzes (LAN, local area network), zum Beispiel eines Unternehmensnetzes, zugänglich. Als weitere Alternative ist der Speicher 150 für die Benutzer von einer zugeordneten fernen Speicherposition (z.B. einer fernen elektronischen Datenbank, einer öffentlichen Cloud oder einer privaten Cloud) oder einer anderen geeigneten Speicherposition aus zugänglich.

[0011] Die Datenabrufhistorie 130 kann in dem Katalog 140 gespeichert werden und Daten über die Herkunft (z.B. Quelle) von empfangenen Daten sowie Statistiken über das Abrufen der Daten umfassen. Zum Beispiel kann die Datenabrufhistorie 130 die Informationen in der nachstehenden Tabelle T1 umfassen:

Tabelle T1

Datenquelle(n):	<URL>, <URL:mirror>, ...
Datengröße:	<bytes>
Zeitspanne für das Herunterladen:	<seconds>
Zeitspanne für das Verarbeiten:	<seconds>
In Anspruch genommene CPU-Ressourcen:	<# of CPUs>, <use %>
In Anspruch genommener Hauptspeicher:	<% use>
Dateisystem-Pfad:	<path1>, <path2>, ...

[0012] In diesem Beispiel speichert die Zeile „Datenquelle“ die Adresse der öffentlich verfügbaren Quelle für die Daten. Diese Zeile kann zum Beispiel den Uniform Resource Locator (URL) und einen Spiegel-URL für die Datenquelle speichern. Die Zeile „Datengröße“ speichert die Größe der heruntergeladenen Daten (z.B. in Byte). Die Zeile „Zeitspanne für das Herunterladen“ speichert die zum Herunterladen der Daten aufgewendete Zeit (z.B. in Sekunden).

[0013] Die Zeile „Zeitspanne für das Verarbeiten“ speichert die Zeit für das Verarbeiten der Daten nach dem Herunterladen (z.B. in Sekunden). In einer Ausführungsform können die Benutzer 102 den Datenverarbeitungsservice 124 verwenden, um heruntergeladene Daten unter Verwendung eines Skripts zu verarbeiten (wobei z.B. Transformationen an den Daten durchgeführt oder die Daten gefiltert werden). Die ursprünglichen Daten und die verarbeiteten Daten oder beide können im Speicher 150 gespeichert werden.

[0014] Des Weiteren speichert der Datenverarbeitungsservice 124 in einer Ausführungsform Identifikationsinformationen (z.B. einen Dateipfad, einen Datenbankpfad, eine Netzadresse usw.) für beliebige Skripte, die zum Verarbeiten von heruntergeladenen Daten verwendet werden (wobei z.B. Transformationen an den Daten durchgeführt oder die Daten gefiltert werden). Zum Beispiel kann der Datenverarbeitungsservice 124 diese Skripte oder Informationen, die die Speicherposition der Skripte angeben, in dem Speicher 150, in dem Katalog 240 oder an einer beliebigen anderen geeigneten Speicherposition speichern. In einer Ausführungsform speichert der Datenverarbeitungsservice 124 auch den dem Skript bereitgestellten Eingabeparameter (z.B. in dem Katalog 140 oder an einer anderen geeigneten Speicherposition).

[0015] Wenn der Katalog in der Datenbank gesichert wird, kann die Zeile „Zeitspanne für das Verarbeiten“ die zur Durchführung dieser Verarbeitung benötigte Zeit speichern. Des Weiteren kann die Zeile „In Anspruch genommene CPU-Ressourcen“ Statistiken für die CPU-Auslastung während der Verarbeitung (z.B. die Anzahl der verwendeten CPUs und der Prozentsatz der in Anspruch genommenen verfügbaren Verarbeitungsleistung) und die Zeile „In Anspruch genommener Hauptspeicher“ kann Statistiken für die Hauptspeicherbelegung während der Verarbeitung (z.B. den Prozentsatz des in Anspruch genommenen verfügbaren Hauptspeichers) speichern. Die Zeile „Dateisystem-Pfad“ kann den Pfad für die heruntergeladenen Daten (und gegebenenfalls die verarbeiteten Daten) speichern. Dies sind lediglich Beispiele und beliebige geeignete Daten können in der Datenabrufhistorie 130 verwaltet werden.

[0016] In einer Ausführungsform können die Benutzer 102 auf den Datenabrufservice 122 und den Datenverarbeitungsservice 124 unter Verwendung einer geeigneten Netzverbindung zugreifen. Zum Beispiel, wie nachstehend mit Bezug auf **Fig. 3** ausführlicher erläutert wird, handelt es sich in einer Ausführungsform bei dem Datenabrufservice 122 und dem Datenverarbeitungsservice 124 um Software-Services, die auf einem oder mehreren Computerservern ausgeführt werden (die z.B. auf einem öffentlichen oder privaten Cloud-Computing-System ausgeführt werden). In dieser Ausführungsform können die Benutzer 102 auf den Datenabrufservice 122 und den Datenverarbeitungsservice 124 unter Verwendung eines LAN, eines Weitverkehrsnetzes (WAN, wide area network) oder eines anderen geeigneten Netzwerks zugreifen. Zum Beispiel können der Datenabrufservice 122 und der Datenverarbeitungsservice 124 über das Internet zugänglich sein und die Benutzer 102 können auf diese Services unter Verwendung einer geeigneten Webservice-Anwendungsprogrammierschnittstelle (API) zugreifen.

[0017] **Fig. 2** veranschaulicht Software-Services für eine fortlaufende Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform. Wie die in **Fig. 1** veranschaulichte Datenverarbeitungsumgebung 100 umfasst eine Datenverarbeitungsumgebung 200 einen oder mehrere Benutzer 102. Diese Benutzer 102 greifen auf einen Datenabrufservice 122 zu, um Daten aus einer Datenquelle 110 abzurufen. Zum Beispiel könnten die Benutzer 102 genau wie in **Fig. 1** Forscher sein, die öffentlich verfügbare geografische Daten verwenden. Bei der Datenquelle 110 könnte es sich um eine öffentliche Datenrepository oder Website handeln. Zum Beispiel kann die Datenquelle 110 eine Repository sein, die unter Verwendung des Internets oder eines anderen Weitverkehrsnetzes zugänglich ist.

[0018] Der Datenabrufservice 122 ruft Daten aus der Datenquelle 110 ab. Der Datenabrufservice 122 speichert die Daten dann im Speicher 150 (z.B. dem in **Fig. 1** veranschaulichten Speicher 150) und verwendet einen Katalog 240, um Einzelheiten über die Herkunft und den Abruf der Daten zu speichern. Zum Beispiel kann es sich bei dem Katalog 240 wie bei dem in **Fig. 1** veranschaulichten Katalog 140 um eine Tabelle in einer elektronischen Datenbank handeln.

[0019] Die Datenverarbeitungsumgebung 200 umfasst zwei zusätzliche Services: einen Dateiwiederherstellungsservice 226 und einen Dateientfernungsservice 228. In einer Ausführungsform ermöglicht der Dateientfernungsservice 228 das Entfernen von Daten aus dem Speicher 150, um eine fortlaufende Speicherung von Daten zu ermöglichen. Dies wird mit Bezug auf **Fig. 6** weiter unten ausführlicher erläutert.

[0020] In einer Ausführungsform ermöglicht der Dateiwiederherstellungsservice 226 eine Wiederherstellung von Dateien, die von dem Dateientfernungsservice 228 gelöscht wurden. Des Weiteren ermöglicht der Dateiwiederherstellungsservice 226 eine Verwendung von beliebigen Skripten durch den Datenverarbeitungsservice 124, um heruntergeladene Daten zu verarbeiten (z.B. von dem Datenverarbeitungsservice 124 oder einem anderen geeigneten Service angegebene Skripte). Dies macht das Entfernen von Dateien durch den Dateientfernungsservice für den Benutzer transparent. Das heißt, die Benutzer 102 können in der Datenverarbeitungsumgebung 200 ohne Kenntnis, welche Daten in dem Speicher 150 tatsächlich gespeichert sind und welche Daten entfernt wurden, um Speicherplatz zu sparen, arbeiten. Dies wird mit Bezug auf die **Fig. 4** bis **Fig. 7** weiter unten ausführlicher erläutert.

[0021] In einer Ausführungsform verwendet der Dateiwiederherstellungsservice 226 den Katalog 240, um eine Wiederherstellung von Dateien zu ermöglichen. Zum Beispiel kann der Katalog 240 eine Datenabrufhistorie 230 speichern. Die Datenabrufhistorie 230 kann wie die in **Fig. 1** veranschaulichte Datenabrufhistorie 130 Daten über die Herkunft (z.B. Quelle) von empfangenen Daten sowie Statistiken über das Abrufen der Daten umfassen. Die Datenabrufhistorie 230 kann des Weiteren Daten darüber umfassen, ob die zuvor abgerufenen Daten im Speicher 150 gespeichert sind. Diese Daten können als Wiederherstellungsparameter für den Wiederherstellungsservice verwendet werden. Zum Beispiel kann die Datenabrufhistorie 230 die Wiederherstellungsparameter in der nachstehenden Tabelle T2 umfassen:

Tabelle T2

Datenquelle(n):	<URL>, <URL:mirror>, ...
Datengröße:	<bytes>
Zeitspanne für das Herunterladen:	<seconds>
Zeitspanne für das Verarbeiten:	<seconds>
In Anspruch genommene CPU-Ressourcen:	<# of CPUs>, <use %>

In Anspruch genommener Hauptspeicher:	<% used>
Dateisystem-Pfad:	<path1>, <path2>, ...
Datei vorhanden?	<boolean>
Verarbeitungsskript(e):	<machine:path1:input_parameters>, <machine:path2:input_parameters>, ...

[0022] Wie gezeigt ist, umfasst die Tabelle T2 die gleichen Zeilen wie die vorstehend beschriebene Tabelle T1 und sie umfasst des Weiteren eine Zeile „Datei vorhanden?“. In einer Ausführungsform zeigt diese Zeile an, ob die zugehörigen Daten im Speicher 150 gespeichert sind (z.B., ob die Daten durch den Dateientfernungsservice 228 entfernt wurden).

[0023] Des Weiteren umfasst die Tabelle T2 eine Zeile „Verarbeitungsskript(e)“. In einer Ausführungsform wird diese Zeile verwendet, um die Speicherposition von Verarbeitungsskripten zu speichern, die von dem Datenverarbeitungsservice 124 zur Verarbeitung von Daten verwendet werden, nachdem diese abgerufen wurden. Zum Beispiel kann der Datenverarbeitungsservice 124 nach der Verarbeitung von abgerufenen Daten den Pfad zu den für die Verarbeitung verwendeten Skripten (einschließlich der Maschine, in der diese Skripte gespeichert sind) in der Datenabrufhistorie 230 speichern. Bei diesem Pfad kann es sich um einen Dateipfad, einen Datenbankpfad, eine Netzadresse oder einen URL oder einen beliebigen anderen geeigneten Pfad handeln, der die Speicherposition des Verarbeitungsskripts angibt. Des Weiteren speichert diese Zeile (oder eine zusätzliche Zeile) in einer Ausführungsform Eingabeparameter, die zum Betreiben des Skripts verwendet werden. Der Datenverarbeitungsservice 124 kann dem Skript diese Eingabeparameter bei der Verarbeitung der abgerufenen Daten bereitstellen.

[0024] Fig. 3 veranschaulicht einen Server 300, der Software-Services für eine fortlaufende Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform ausführt. Der Server 300 umfasst einen Prozessor 302, einen Hauptspeicher 310 und Netzkomponenten 320. Der Prozessor 302 ruft im Allgemeinen im Hauptspeicher 310 gespeicherte Programmieranweisungen ab und führt sie aus. Der Prozessor 302 wurde stellvertretend für eine einzelne zentrale Verarbeitungseinheit (CPU, central processing unit), mehrere CPUs, eine einzelne CPU mit mehreren Verarbeitungskernen, Grafikerarbeitungseinheiten (GPUs, graphics processing units,) mit mehreren Ausführungspfaden und dergleichen aufgenommen. Des Weiteren kann der Server 300 ein einzelner Server-Computer sein oder er kann einen Verbund von Server-Computern darstellen (die z.B. zusammen verwaltet werden).

[0025] Zu den Netzkomponenten 320 gehören die Komponenten, die der Server 300 zur Verbindung mit einem drahtlosen oder drahtgebundenen Übertragungsnetz benötigt (z.B. um mit den Benutzern 102 und der Datenquelle 110, die in den Fig. 1 und Fig. 2 veranschaulicht sind, Daten auszutauschen). Zum Beispiel können zu den Netzkomponenten 320 WiFi- oder Mobilfunkschnittstellenkomponenten und zugehörige Software oder Festnetzschnittstellenkomponenten (z.B. Ethernet-Komponenten, Glasfaserkomponenten usw.) gehören.

[0026] Obgleich der Hauptspeicher 310 als eine einzelne Entität gezeigt ist, kann der Hauptspeicher 310 eine oder mehrere Hauptspeichereinheiten mit Hauptspeicherblöcken umfassen, die physischen Adressen zugeordnet sind, wie beispielsweise Arbeitsspeicher (RAM, random access memory), Nur-Lese-Speicher (ROM, read only memory), Flashspeicher oder andere Arten von flüchtigem und/oder nicht flüchtigem Speicher. Der Hauptspeicher 310 umfasst im Allgemeinen Programmcode zur Durchführung von verschiedenen Funktionen in Zusammenhang mit der Verwendung des Servers 300. Der Programmcode wird im Allgemeinen als verschiedene Funktions- „Anwendungen“ oder „Module“ innerhalb des Hauptspeichers 310 beschrieben, obgleich alternative Ausführungen andere Funktionen und/oder Kombinationen von Funktionen haben können.

[0027] Der Hauptspeicher 310 umfasst einen Datenabrufservice 122. In einer Ausführungsform ist dies der in den Fig. 1 bis Fig. 2 veranschaulichte Datenabrufservice 122. Wie in diesen Figuren erläutert ist, ermöglicht der Datenabrufservice 122 einen Abruf und eine Speicherung von Daten (z.B. aus einer Datenquelle 110, wie in den Fig. 1 und Fig. 2 veranschaulicht ist). Der Hauptspeicher 310 umfasst des Weiteren einen Datenverarbeitungsservice 124. In einer Ausführungsform ist dies der in den Fig. 1 bis Fig. 2 veranschaulichte Datenverarbeitungsservice 124. Wie in diesen Figuren erläutert ist, verarbeitet der Datenverarbeitungsservice 124 aus einer Datenquelle heruntergeladene Daten (wobei z.B. Transformationen an den Daten durchgeführt oder die Daten gefiltert werden).

[0028] Der Hauptspeicher 310 umfasst des Weiteren einen Dateientfernungsservice 228. In einer Ausführungsform ist dies der in **Fig. 2** veranschaulichte Dateientfernungsservice 228. Wie in **Fig. 2** erläutert ist, ermöglicht der Dateientfernungsservice 228 eine Entfernung von zuvor heruntergeladenen Dateien, um Platz zu sparen, während er eine fortlaufende Speicherung von Daten ermöglicht.

[0029] Der Hauptspeicher 310 umfasst des Weiteren einen Dateiwiederherstellungsservice 226. In einer Ausführungsform ist dies der in **Fig. 2** veranschaulichte Dateiwiederherstellungsservice 226. Wie in **Fig. 2** erläutert ist, ermöglicht der Dateiwiederherstellungsservice 226 eine Wiederherstellung von Dateien, die heruntergeladen und dann entfernt wurden, um Speicherplatz zu sparen (z.B. unter Verwendung des Dateientfernungsservice 228).

[0030] Wie in **Fig. 3** veranschaulicht ist, werden die Services 122, 124, 226 und 228 im Server 300 ausgeführt. Dies ist lediglich ein Beispiel. Alternativ könnten die Services 122, 124, 226 und 228 jeweils auf einem separaten Server ausgeführt werden oder eine beliebige Kombination von Services kann auf einer beliebigen Anzahl von Servern ausgeführt werden. Des Weiteren kann ein bestimmter Service über mehrere Server verteilt sein (z.B. zur Arbeitslastverwaltung).

[0031] **Fig. 4** ist ein Ablaufplan 400 für eine fortlaufende Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform. Am Block 402 fordert ein Benutzer (z.B. einer der in den **Fig. 1** bis **Fig. 2** veranschaulichten Benutzer 102) Daten an (z.B. aus dem Speicher 150 oder einer Datenquelle 110, die in den **Fig. 1** bis **Fig. 2** veranschaulicht sind). In einer Ausführungsform überträgt der Benutzer eine Datenanforderung an einen Datenabrufservice (z.B. den in den **Fig. 1** bis **Fig. 3** veranschaulichten Datenabrufservice 122).

[0032] Am Block 404 prüft der Datenabrufservice, ob die angeforderten Daten über in einem Katalog (z.B. dem in **Fig. 2** veranschaulichten Katalog 240) aufgeführte Metadaten verfügen. In einem Katalog aufgeführte Daten zeigen zum Beispiel an, dass der Datenabrufservice die angeforderten Daten zuvor heruntergeladen hat (z.B. hat der Benutzer oder ein anderer Benutzer die Daten zuvor angefordert). Zum Beispiel kann der Datenabrufservice die in der Tabelle T2 veranschaulichte Zeile „Datei vorhanden?“ zur Feststellung verwenden, ob die Daten im Speicher 150 vorhanden sind. Wenn die Antwort am Block 404 „ja“ lautet, schaltet der Ablauf zum Block 406, wo der Datenabrufservice prüft, ob die angeforderten Daten lokal vorhanden sind. Verwendet man das Beispiel von Tabelle 2, kann der Katalog zum Beispiel eine Zeile „Datei vorhanden?“ in einer Datenabrufhistorie (z.B. der in **Fig. 2** veranschaulichten Datenabrufhistorie 230) umfassen.

[0033] Wenn dieser Block „wahr“ zurückgibt, sind die angeforderten Daten lokal vorhanden. Der Ablauf schreitet zum Block 408 fort und der Datenabrufservice gibt die angeforderten Daten zurück. Alternativ kann der Datenabrufservice den lokalen Dateipfad für die angeforderten Daten zurückgeben. In einer Ausführungsform kann dieser lokale Dateipfad als Teil der Datenabrufhistorie in dem Katalog gespeichert werden.

[0034] Zum Block 406 zurückkehrend, schaltet der Ablauf zum Block 410, wenn die angeforderten Daten nicht lokal vorhanden sind. Am Block 410 erstellt der Dateiwiederherstellungsservice (z.B. der in den **Fig. 2** bis **Fig. 3** veranschaulichte Dateiwiederherstellungsservice 226) die angeforderten Daten neu. In einer Ausführungsform lädt der Dateiwiederherstellungsservice die angeforderten Daten aus der zugehörigen Datenquelle herunter und führt eine beliebige zusätzliche Verarbeitung an den Daten aus. Des Weiteren kann der Dateiwiederherstellungsservice den Datenabrufservice und einen Datenverarbeitungsservice (z.B. den in den **Fig. 1** bis **Fig. 3** veranschaulichten Datenverarbeitungsservice 124) verwenden, um diese Vorgänge durchzuführen. Dies wird mit Bezug auf **Fig. 5** näher erläutert.

[0035] In einer Ausführungsform kann der Dateiwiederherstellungsservice eine Wiederherstellung von Daten überwachen. Zum Beispiel kann einem Systemadministrator ein Überwachungsservice bereitgestellt werden, der Daten angibt, die wiederhergestellt wurden. Dabei kann es sich um ein Echtzeit-Überwachungssystem, ein Protokoll oder ein beliebiges anderes geeignetes System handeln. Der Systemadministrator kann das Überwachungssystem zur Angabe von Daten verwenden, die wiederholt wiederhergestellt werden, um dem Systemadministrator die Lösung etwaiger Probleme zu ermöglichen (z.B. das Ändern von Dateientfernungsparametern, um Churn zu vermeiden). Das Überwachungssystem kann des Weiteren CPU-Auslastung, Speicherbelegung, Bandbreitennutzung usw. umfassen.

[0036] In einer Ausführungsform kann das Dateiwiederherstellungssystem einem Systemadministrator des Weiteren Benachrichtigungen bereitstellen, wenn Daten wiederhergestellt werden. Dies kann es dem Systemadministrator ermöglichen, festzustellen, dass der Speicher (z.B. der in den **Fig. 1** bis **Fig. 2** veranschau-

lichte Speicher 150) nicht über ausreichend Platz verfügt. Während zum Beispiel eine oder mehrere der hierin veranschaulichten Techniken eine fortlaufende Speicherung von Daten ohne Fehler ermöglichen kann/können, kann eine wiederholte Entfernung und Wiederherstellung von Daten Systemressourcen in Anspruch nehmen (z.B. Netzbandbreite, um die Daten abzurufen, und CPU-Ressourcen, um die Daten zu verarbeiten). Der Systemadministrator kann die Benachrichtigungen verwenden, um Daten anzugeben, die wiederholt entfernt und wiederhergestellt werden, und um festzustellen, ob zusätzlicher Platz hinzugefügt (oder zugeordnet) werden soll, um eine Entfernung und Wiederherstellung zu vermeiden.

[0037] Der Ablauf schaltet dann zum Block 408. Am Block 408 gibt der Dateiwiederherstellungsservice, wie vorstehend erläutert wurde, die angeforderten Daten oder eine Kennung für die angeforderten Daten (z.B. einen Dateipfad) zurück.

[0038] Zum Block 404 zurückkehrend, schaltet der Ablauf zum Block 412, wenn die Daten keine zugehörigen Metadaten in dem Katalog umfassen. Das bedeutet zum Beispiel, dass der Datenabrufservice die angeforderten Daten zuvor nicht heruntergeladen hat. Am Block 412 gibt ein Dateientfernungsservice (z.B. der in den **Fig. 2** bis **Fig. 3** veranschaulichte Dateientfernungsservice 228) Plattenspeicherplatz (bei Bedarf) frei. Dies wird mit Bezug auf **Fig. 6** weiter unten näher erläutert. In einer Ausführungsform kann dies Plattenspeicherplatzfehler vermeiden. Ohne Freigabe von Plattenspeicherplatz könnte das Abrufen und Verarbeiten der angeforderten Daten letztendlich zur Folge haben, dass mehr Plattenspeicherplatz in Anspruch genommen wird, als zur Verfügung steht, was zu einem Fehler führt. Der Dateientfernungsservice kann diese Fehler vermeiden, indem Plattenspeicherplatz am Block 412 freigegeben wird, bevor die Daten abgerufen und verarbeitet werden.

[0039] Am Block 414 erzeugt der Datenabrufservice die angeforderten Daten. In einer Ausführungsform lädt der Datenabrufservice die angeforderten Daten herunter und der Datenverarbeitungsservice führt eine beliebige zusätzliche Verarbeitung an den Daten aus. Dies wird mit Bezug auf **Fig. 7** weiter unten näher erläutert.

[0040] Am Block 416 speichert der Datenabrufservice oder ein anderer geeigneter Service (z.B. der Dateiwiederherstellungsservice) die Daten (z.B. in dem in den **Fig. 1** bis **Fig. 2** veranschaulichten Speicher 150). Wie vorstehend in Verbindung mit **Fig. 1** erläutert wurde, werden die Daten in einer Ausführungsform lokal auf der Client-Einheit gespeichert, die die Daten angefordert hat (z.B. auf dem lokalen Festplattenlaufwerk eines Benutzers). Alternativ oder zusätzlich werden die Daten in einem System gespeichert, das für den Benutzer zugänglich ist (z.B. über ein Übertragungsnetz). Zum Beispiel kann der Benutzer Teil einer Forschungsorganisation sein, die ein System zur Verwendung mit laufenden Projekten verwaltet (z.B. ein Netzdateisystem, eine Datenbank oder ein Cloud-Speichersystem). Der Speicherort kann von dem Benutzer entfernt, zugänglich für den Benutzer über ein Übertragungsnetz, sein.

[0041] Am Block 418 aktualisiert der Datenabrufservice den Katalog. Wie vorstehend in Bezug auf **Fig. 2** erläutert wurde, verwaltet der Datenabrufservice in einer Ausführungsform einen Katalog (z.B. den Katalog 240) mit einer Datenabrufhistorie (z.B. der Datenabrufhistorie 230). Diese Datenabrufhistorie umfasst verschiedene Daten in Bezug auf die abgerufenen Daten. Tabelle 2 oben ist ein Beispiel für geeignete Daten. Am Block 418 aktualisiert der Datenabrufservice die Datenabrufhistorie in dem Katalog (er aktualisiert z.B. die Download-Größe, die Download-Geschwindigkeit, Datenverarbeitungsstatistiken usw.). Der Ablauf schaltet dann zum Block 408 und gibt, wie vorstehend erläutert wurde, die Daten zurück.

[0042] **Fig. 5** ist ein Ablaufplan zum Neuerstellen von Daten als Teil einer fortlaufenden Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform. In einer Ausführungsform entspricht **Fig. 5** dem in **Fig. 4** veranschaulichten Block 410. Am Block 502 gibt ein Dateiwiederherstellungsservice (z.B. der in den **Fig. 2** bis **Fig. 3** veranschaulichte Dateiwiederherstellungsservice 226) die Datenquellen für die angeforderten Daten aus einem Katalog (z.B. dem in **Fig. 2** veranschaulichten Katalog 140) an. In einer Ausführungsform liest der Dateiwiederherstellungsservice den zu den angeforderten Daten (z.B. einer Datenbanktabelle) gehörenden Katalogeintrag und gibt die in dem Katalog gespeicherten Datenquelleninformationen an (z.B. die URL oder die Spiegel-URL). Zum Beispiel kann der Dateiwiederherstellungsservice die in der Tabelle T2 veranschaulichte Zeile „Datenquelle(n)“ verwenden.

[0043] Am Block 504 gibt der Dateiwiederherstellungsservice (oder der Dateientfernungsservice oder ein anderer geeigneter Service) den benötigten Plattenspeicherplatz frei. In einer Ausführungsform entspricht der Block 504 dem in **Fig. 4** veranschaulichten Block 412. Dies wird mit Bezug auf **Fig. 6** weiter unten näher erläutert.

[0044] Am Block 506 erzeugt der Dateiwiederherstellungsservice die Daten. In einer Ausführungsform entspricht der Block 506 dem in **Fig. 4** veranschaulichten Block 414. Zum Beispiel kann der Dateiwiederherstellungsservice die angeforderten Daten am Block 506 herunterladen und eine beliebige zusätzliche Verarbeitung an den Daten ausführen. Dies wird mit Bezug auf **Fig. 7** weiter unten näher erläutert.

[0045] **Fig. 6** ist ein Ablaufplan zum Freigeben von Plattenspeicherplatz als Teil einer fortlaufenden Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform. In einer Ausführungsform entspricht **Fig. 6** den Blöcken 412 (in **Fig. 4** veranschaulicht) und 504 (in **Fig. 5** veranschaulicht).

[0046] Am Block 602 stellt ein Dateiwiederherstellungsservice (z.B. der in den **Fig. 2** bis **Fig. 3** veranschaulichte Dateiwiederherstellungsservice 226) oder ein anderer geeigneter Service fest, ob Plattenspeicherplatz zur Verfügung steht. In einer Ausführungsform gibt der Service die Menge an freiem Speicherplatz, der in dem Speicher (z.B. dem in den **Fig. 1** bis **Fig. 2** veranschaulichten Speicher 150) zur Verfügung steht, und die Menge an Speicherplatz an, der für die abgerufenen Daten benötigt wird. Wenn die Daten zuvor abgerufen wurden und einen zugehörigen Katalogeintrag haben, kann der Katalog zur Feststellung der Menge an benötigtem Speicherplatz verwendet werden. Alternativ kann der Dateiwiederherstellungsservice die Menge an benötigtem Speicherplatz aus der Datenquelle (z.B. der in den **Fig. 1** bis **Fig. 2** veranschaulichten Datenquelle 110) feststellen.

[0047] Wenn genügend Plattenspeicherplatz zur Verfügung steht, endet der Ablauf. Wenn nicht genügend Plattenspeicherplatz zur Verfügung steht, schaltet der Ablauf zum Block 604. Am Block 604 gibt ein Dateientfernungsservice (z.B. der in den **Fig. 2** bis **Fig. 3** veranschaulichte Dateientfernungsservice 228) Daten zum Entfernen (z.B. eine oder mehrere Dateien) an. In einer Ausführungsform verwendet der Dateientfernungsservice den Katalog, um Daten zum Entfernen anzugeben. Der Dateientfernungsservice kann Daten angeben, die auf der Grundlage der in dem Katalog gespeicherten Wiederherstellungsparameter (z.B. Download-Größe, Download-Zeit, Verarbeitungszeit usw.) am einfachsten neu erstellt werden können.

[0048] Zum Beispiel kann der Dateientfernungsservice die in der Tabelle T2 veranschaulichte Zeile „Datengröße“ zur Angabe der Daten verwenden, die die kleinste Größe zum Herunterladen hatten. Der Dateientfernungsservice kann die in der Tabelle T2 veranschaulichte Zeile „Zeitspanne für das Herunterladen“ zur Angabe der Daten verwenden, die am schnellsten herunterzuladen waren. Der Dateientfernungsservice kann die in der Tabelle T2 veranschaulichte Zeile „Zeitspanne für das Verarbeiten“ zur Angabe der Daten verwenden, die am schnellsten zu verarbeiten waren. Der Dateientfernungsservice kann die Zeile „In Anspruch genommene CPU-Ressourcen“ in der Tabelle T2 zur Angabe der Daten verwenden, die die wenigsten CPU-Ressourcen zur Verarbeitung in Anspruch genommen haben. Der Dateientfernungsservice kann die in der Tabelle T2 veranschaulichte Zeile „In Anspruch genommener Hauptspeicher“ zur Angabe der Daten verwenden, die die wenigsten Hauptspeicherressourcen zur Verarbeitung in Anspruch genommen haben. Dies sind lediglich Beispiele. Der Dateientfernungsservice kann beliebige (oder alle) dieser Zeilen zusammen mit anderen Daten zur Angabe von zu entfernenden Daten verwenden.

[0049] Des Weiteren kann der Dateientfernungsservice Daten angeben, die selten verwendet werden, wahrscheinlich veraltet sind usw. Zum Beispiel kann der Katalog überwachen, wann zuletzt auf Daten zugegriffen wurde und wie oft auf die Daten zugegriffen wird (z.B. als eine Zeile in der Tabelle T2), und dies kann zur Angabe von Daten verwendet werden, die selten verwendet werden. Dies sind lediglich Beispiele und beliebige geeignete Parameter können verwendet werden.

[0050] In einer Ausführungsform umfasst der Dateientfernungsservice Standardparameter zur Angabe von zu entfernenden Daten. Ein Benutzer oder Administrator kann diese Parameter konfigurieren (z.B. unter Verwendung einer geeigneten Benutzerschnittstelle). Alternativ kann der Dateientfernungsservice Parameter umfassen, die von einem Benutzer nicht geändert werden können.

[0051] Als ein weiteres Beispiel kann der Dateientfernungsservice Kosten in Verbindung mit Cloud-Speicher bei der Angabe von zu entfernenden Daten berücksichtigen. Viele Cloudsysteme stellen Verarbeitungsleistung, Netzbandbreite, Speicher usw. in Rechnung. Der Dateientfernungsservice kann die in dem Katalog enthaltenen Daten verwenden, um die Kosten zu schätzen, die mit dem Wiederherstellen von entfernten Daten verbunden sind, und kann dies als einen Faktor in seiner Entscheidung berücksichtigen. Des Weiteren können die Cloud-Parameter (z.B. eine Markierung, die Cloud-Speicher, Verarbeitungskosten, Bandbreite-Kosten, Speicherkosten usw. angibt) selbst in dem Katalog oder an einer anderen geeigneten Speicherposition gespeichert und von dem Dateientfernungsservice zum Erzeugen der Kostenschätzung verwendet werden.

[0052] Des Weiteren kann der Dateientfernungs-service das Entfernen von Daten verhindern, die nicht wiederhergestellt werden können. Zum Beispiel haben die Daten möglicherweise keinen entsprechenden Eintrag in dem Katalog oder der Katalogeintrag ist möglicherweise unvollständig. Der Dateientfernungs-service kann das Entfernen dieser Daten verhindern. Als ein weiteres Beispiel können Daten in dem Katalog als aus einer Quelle beschaffte Daten markiert werden, die in Zukunft möglicherweise nicht mehr zur Verfügung steht. Zum Beispiel kann der Katalog ein zusätzliches Feld (z.B. eine Zeile oder Spalte) mit diesen Informationen umfassen. Der Dateientfernungs-service kann auch das Entfernen dieser Daten verhindern.

[0053] Der Dateientfernungs-service kann auch überprüfen, ob die Daten aus der Datenquelle immer noch zur Verfügung stehen, bevor er die Daten entfernt. Zum Beispiel kann der Dateientfernungs-service versuchen, mit dem Abrufen der angeforderten Daten (z.B. aus der in den **Fig. 1** bis **Fig. 2** veranschaulichten Datenquelle 110) zu beginnen, wobei er die in dem Katalog enthaltene Speicherposition verwendet. Wenn dies nicht erfolgreich ist (z.B. empfängt der Dateientfernungs-service einen Fehler), kann der Dateientfernungs-service das Entfernen der Daten verhindern.

[0054] Als weitere Alternative kann der Katalog ein Feld enthalten, das angibt, ob die Daten aus der Datenquelle 110 zur Verfügung stehen. Anstatt vor dem Löschen die Verfügbarkeit zu prüfen (oder zusätzlich dazu), kann der Dateientfernungs-service periodisch im Hintergrund arbeiten und dieses Feld aktualisieren, um Daten anzugeben, die nicht mehr zur Verfügung stehen. Zum Beispiel kann der Dateientfernungs-service arbeiten, wenn das System wenig genutzt wird oder zu vorgegebenen Zeiten, um die aktuelle Verfügbarkeit für in dem Katalog angegebene Daten zu prüfen.

[0055] Am Block 606 entfernt der Dateientfernungs-service die angegebenen Daten aus einem Speicher (z.B. dem in den **Fig. 1** bis **Fig. 2** veranschaulichten Speicher 150). Zum Beispiel kann der Dateientfernungs-service die in der Tabelle T2 veranschaulichte Zeile „Dateisystem-Pfad“ zur Angabe des Speicherorts für die angegebenen Daten verwenden. Der Dateientfernungs-service kann diese Daten dann entfernen.

[0056] Am Block 608 aktualisiert der Dateientfernungs-service den Katalog. Zum Beispiel kann der Katalog einen Eintrag umfassen, der angibt, ob die Daten in dem Speicher vorhanden sind (z.B. die in der obigen Tabelle 2 veranschaulichte Zeile „Datei vorhanden?“). Der Dateientfernungs-service kann diesen Wert von wahr (z.B. vorhanden) nach falsch (z.B. nicht vorhanden) umschalten. Der Ablauf kehrt dann zum Block 602 zurück.

[0057] **Fig. 7** ist ein Ablaufplan zum Erzeugen von angeforderten Daten als Teil einer fortlaufenden Speicherung von Daten in einem System mit begrenzter Speicherkapazität gemäß einer einzelnen Ausführungsform. In einer Ausführungsform entspricht **Fig. 7** den Blöcken 414 (in **Fig. 4** veranschaulicht) und 506 (in **Fig. 5** veranschaulicht). Am Block 702 ruft ein Datenabrufservice (z.B. der in den **Fig. 2** bis **Fig. 3** veranschaulichte Datenabrufservice 122) die angeforderten Daten aus der Datenquelle (z.B. der in den **Fig. 1** bis **Fig. 2** veranschaulichten Datenquelle 110) ab.

[0058] Am Block 704 gibt ein Datenverarbeitungsservice (z.B. der in den **Fig. 2** bis **Fig. 3** veranschaulichte Datenverarbeitungsservice 124) ein beliebiges Datenverarbeitungsskript an, das zu den angeforderten Daten gehört. Zum Beispiel hat ein Benutzer möglicherweise einen Filter, eine Transformation oder ein anderes Verarbeitungsskript angegeben, das an den Daten ausgeführt werden soll, nachdem sie abgerufen wurden, und er hat möglicherweise Eingabeparameter für das Skript bereitgestellt. In einer Ausführungsform wird ein Datenverarbeitungsskript als eine Reihe von Schritten gespeichert (z.B. ein von einem Menschen lesbares Programmiersprachenskript oder ein maschinenlesbares kompiliertes Skript). Das Skript kann in einer Datei, in einer Datenbanktabelle oder in einer beliebigen anderen geeigneten Weise gespeichert werden. Alternativ steht ein Skript an einem Netzspeicherort zur Verfügung und kann unter Verwendung des Datenverarbeitungsservices aufgerufen werden (z.B. unter Verwendung einer Netz-API). Beliebige Eingabeparameter können auch in dem Skript selbst, in einer Datenbanktabelle oder an einer beliebigen anderen geeigneten Speicherposition gespeichert werden. Zum Beispiel kann die in der Tabelle T2 veranschaulichte Zeile „Verarbeitungsskript(e)“ den Pfad (z.B. Dateipfad, Datenbankpfad oder Netzpfad) für die Skripte zusammen mit beliebigen Eingabeparametern angeben.

[0059] In einer Ausführungsform wurden die Daten zuvor erzeugt und werden jetzt neu erstellt (z.B. wie vorstehend mit Bezug auf **Fig. 5** erläutert wurde). Ein Dateiwiederherstellungsservice gibt das Skript unter Verwendung von Informationen in dem Katalog (z.B. einem Dateipfad oder einem Datenbankpfad, wie vorstehend mit Bezug auf **Fig. 2** erläutert wurde) an und ruft das Skript und beliebige mit dem Skript zu verwendende Eingabeparameter bei Bedarf ab. Alternativ kann der Dateiwiederherstellungsservice oder ein

anderer geeigneter Service (z.B. der Datenverarbeitungsservice) den Benutzer abfragen, um ein vorhandenes Skript anzugeben oder um die Schritte eines gewünschten Skripts einzugeben. In einer Ausführungsform kann der Dateiwiederherstellungsservice (oder ein anderer geeigneter Service) die Schritte des Skripts und den Speicherpfad in dem Katalog speichern, so dass das Skript verwendet werden kann, wenn die Daten später neu erstellt werden.

[0060] Am Block 706 führt der Datenverarbeitungsservice das Skript aus, um die abgerufenen Daten zu verarbeiten. Wenn das Skript zum Beispiel aus einem Speicher abgerufen wird, kann der Datenverarbeitungsservice das Skript mit den Eingabeparametern ausführen (z.B. ein Python-Skript oder ein beliebiges anderes geeignetes Skript ausführen). Als ein weiteres Beispiel kann der Datenverarbeitungsservice, wenn das Skript an einem Netzspeicherort zur Verfügung steht, den Netzspeicherort aufrufen, ihm die Eingabeparameter übergeben und die Datenausgabe empfangen (oder eine Kennung der Speicherposition der Datenausgabe empfangen). Der Datenverarbeitungsservice (oder ein anderer geeigneter Service) kann die verarbeiteten Daten zusammen mit den (oder anstelle der) abgerufenen Daten speichern.

[0061] Die Beschreibungen der verschiedenen Ausführungsformen der vorliegenden Erfindung erfolgten zum Zweck der Veranschaulichung, sollen jedoch nicht erschöpfend oder auf die offenbarten Ausführungsformen beschränkt sein. Viele Änderungen und Varianten sind für den Fachmann erkennbar, ohne vom Umfang und Wesen der beschriebenen Ausführungsformen abzuweichen. Die hierin verwendete Terminologie wurde gewählt, um die Grundgedanken der Ausführungsformen, die praktische Anwendung oder technische Verbesserung gegenüber auf dem Markt befindlichen Technologien am besten zu erklären bzw. um anderen Fachleuten das Verständnis der hierin offenbarten Ausführungsformen zu ermöglichen.

[0062] Im Vorstehenden wird Bezug auf in dieser Offenbarung aufgezeigte Ausführungsformen genommen. Der Umfang der vorliegenden Offenbarung ist jedoch nicht auf bestimmte beschriebene Ausführungsformen beschränkt. Vielmehr wird jede beliebige Kombination der vorstehend beschriebenen Merkmale und Elemente, ungeachtet dessen, ob sie auf verschiedene Ausführungsformen bezogen sind oder nicht, in Betracht gezogen, um in Betracht gezogene Ausführungsformen auszuführen und in die Praxis umzusetzen. Obgleich hierin offenbarte Ausführungsformen Vorteile gegenüber anderen möglichen Lösungen oder gegenüber dem Stand der Technik erzielen können, stellt die Frage, ob ein bestimmter Vorteil von einer gegebenen Ausführungsform erzielt wird oder nicht, überdies keine Einschränkung des Umfangs der vorliegenden Offenbarung dar. Die vorstehend erläuterten Aspekte, Merkmale, Ausführungsformen und Vorteile dienen somit lediglich der Veranschaulichung und werden nicht als Elemente oder Einschränkungen der beigefügten Ansprüche betrachtet, soweit dies in einem Anspruch oder in Ansprüchen nicht ausdrücklich anders angegeben ist. Ebenso ist eine Bezugnahme auf „die Erfindung“ nicht als Verallgemeinerung eines beliebigen hierin offenbarten Gegenstands auszulegen und nicht als ein Element oder eine Einschränkung der beigefügten Ansprüche zu betrachten, soweit dies in einem Anspruch oder in Ansprüchen nicht ausdrücklich anders angegeben ist.

[0063] Aspekte der vorliegenden Erfindung können die Form einer reinen Hardware-Ausführungsform, einer reinen Software-Ausführungsform (darunter Firmware, residente Software, Mikrocode usw.) oder einer Ausführungsform annehmen, die Software- und Hardware-Aspekte kombiniert, die hierin alle allgemein als eine „Schaltung“, ein „Modul“ oder ein „System“ bezeichnet werden können.

[0064] Bei der vorliegenden Erfindung kann es sich um ein System, ein Verfahren und/oder ein Computerprogrammprodukt handeln. Das Computerprogrammprodukt kann ein durch einen Computer lesbares Speichermedium (oder -medien) mit durch einen Computer lesbaren Programmanweisungen darauf umfassen, um einen Prozessor dazu zu veranlassen, Aspekte der vorliegenden Erfindung auszuführen.

[0065] Bei dem durch einen Computer lesbaren Speichermedium kann es sich um eine greifbare Einheit handeln, die Anweisungen zur Verwendung durch ein System zur Ausführung von Anweisungen behalten und speichern kann. Bei dem durch einen Computer lesbaren Speichermedium kann es sich zum Beispiel um eine elektronische Speichereinheit, eine magnetische Speichereinheit, eine optische Speichereinheit, eine elektromagnetische Speichereinheit, eine Halbleiterspeichereinheit oder jede geeignete Kombination daraus handeln, ohne auf diese beschränkt zu sein. Zu einer nicht erschöpfenden Liste spezifischerer Beispiele des durch einen Computer lesbaren Speichermediums gehören die folgenden: eine auswechselbare Computerdiskette, eine Festplatte, ein Direktzugriffsspeicher (RAM), ein Nur-Lese-Speicher (ROM), ein löschbarer programmierbarer Nur-Lese-Speicher (EPROM bzw. Flash-Speicher), ein statischer Direktzugriffsspeicher (SRAM), ein auswechselbarer Kompaktspeicherplatte-Nur-Lese-Speicher (CD-ROM), eine DVD (digital versatile disc), ein Speicher-Stick, eine Diskette, eine mechanisch kodierte Einheit wie zum Bei-

spiel Lochkarten oder erhabene Strukturen in einer Rille, auf denen Anweisungen gespeichert sind, und jede geeignete Kombination daraus. Ein durch einen Computer lesbares Speichermedium soll in der Verwendung hierin nicht als flüchtige Signale an sich aufgefasst werden, wie zum Beispiel Funkwellen oder andere sich frei ausbreitende elektromagnetische Wellen, elektromagnetische Wellen, die sich durch einen Wellenleiter oder ein anderes Übertragungsmedium ausbreiten (z.B. Lichtwellenleiterkabel durchlaufende Lichtimpulse) oder durch einen Draht übermittelte elektrische Signale.

[0066] Hierin beschriebene, durch einen Computer lesbare Programmanweisungen können von einem durch einen Computer lesbaren Speichermedium auf jeweilige Datenverarbeitungs-/Verarbeitungseinheiten oder über ein Netzwerk wie zum Beispiel das Internet, ein lokales Netzwerk, ein Weitverkehrsnetz und/oder ein drahtloses Netzwerk auf einen externen Computer oder eine externe Speichereinheit heruntergeladen werden. Das Netzwerk kann Kupferübertragungskabel, Lichtwellenübertragungsleiter, drahtlose Übertragung, Leitwegrechner, Firewalls, Vermittlungseinheiten, Gateway-Computer und/oder Edge-Server aufweisen. Eine Netzwerkkarte oder Netzwerkschnittstelle in jeder Datenverarbeitungs-/Verarbeitungseinheit empfängt durch einen Computer lesbare Programmanweisungen aus dem Netzwerk und leitet die durch einen Computer lesbaren Programmanweisungen zur Speicherung in einem durch einen Computer lesbaren Speichermedium innerhalb der entsprechenden Datenverarbeitungs-/Verarbeitungseinheit weiter.

[0067] Bei durch einen Computer lesbaren Programmanweisungen zum Ausführen von Arbeitsschritten der vorliegenden Erfindung kann es sich um Assembler-Anweisungen, ISA-Anweisungen (Instruction-Set-Architecture), Maschinenanweisungen, maschinenabhängige Anweisungen, Mikrocoded, Firmware-Anweisungen, zustandsetzende Daten oder entweder Quellcode oder Objektcode handeln, die in einer beliebigen Kombination aus einer oder mehreren Programmiersprachen geschrieben werden, darunter objektorientierte Programmiersprachen wie Smalltalk, C++ o.ä. sowie herkömmliche prozedurale Programmiersprachen wie die Programmiersprache „C“ oder ähnliche Programmiersprachen. Die durch einen Computer lesbaren Programmanweisungen können vollständig auf dem Computer des Benutzers, teilweise auf dem Computer des Benutzers, als eigenständiges Softwarepaket, teilweise auf dem Computer des Benutzers und teilweise auf einem fernen Computer oder vollständig auf dem fernen Computer oder Server ausgeführt werden. In letzterem Fall kann der entfernt angeordnete Computer mit dem Computer des Benutzers durch eine beliebige Art Netzwerk verbunden sein, darunter ein lokales Netzwerk (LAN) oder ein Weitverkehrsnetz (WAN), oder die Verbindung kann mit einem externen Computer hergestellt werden (zum Beispiel über das Internet unter Verwendung eines Internet-Diensteanbieters). In einigen Ausführungsformen können elektronische Schaltungen, darunter zum Beispiel programmierbare Logikschaltungen, vor Ort programmierbare Gatter-Anordnungen (FPGA, field-programmable gate arrays) oder programmierbare Logikanordnungen (PLA, programmable logic arrays) die durch einen Computer lesbaren Programmanweisungen ausführen, indem sie Zustandsinformationen der durch einen Computer lesbaren Programmanweisungen nutzen, um die elektronischen Schaltungen zu personalisieren, um Aspekte der vorliegenden Erfindung durchzuführen.

[0068] Aspekte der vorliegenden Erfindung sind hierin unter Bezugnahme auf Ablaufpläne und/oder Blockschaltbilder bzw. Schaubilder von Verfahren, Vorrichtungen (Systemen) und Computerprogrammprodukten gemäß Ausführungsformen der Erfindung beschrieben. Es wird darauf hingewiesen, dass jeder Block der Ablaufpläne und/oder der Blockschaltbilder bzw. Schaubilder sowie Kombinationen von Blöcken in den Ablaufplänen und/oder den Blockschaltbildern bzw. Schaubildern mittels durch einen Computer lesbare Programmanweisungen ausgeführt werden können.

[0069] Diese durch einen Computer lesbaren Programmanweisungen können einem Prozessor eines Universalcomputers, eines Spezialcomputers oder einer anderen programmierbaren Datenverarbeitungsvorrichtung bereitgestellt werden, um eine Maschine zu erzeugen, so dass die über den Prozessor des Computers bzw. der anderen programmierbaren Datenverarbeitungsvorrichtung ausgeführten Anweisungen ein Mittel zur Umsetzung der in dem Block bzw. den Blöcken der Ablaufpläne und/oder der Blockschaltbilder bzw. Schaubilder festgelegten Funktionen/Schritte erzeugen. Diese durch einen Computer lesbaren Programmanweisungen können auch auf einem durch einen Computer lesbaren Speichermedium gespeichert sein, das einen Computer, eine programmierbare Datenverarbeitungsvorrichtung und/oder andere Einheiten so steuern kann, dass sie auf eine bestimmte Art funktionieren, so dass das durch einen Computer lesbare Speichermedium, auf dem Anweisungen gespeichert sind, ein Herstellungsprodukt aufweist, darunter Anweisungen, welche Aspekte der/des in dem Block bzw. den Blöcken des Ablaufplans und/oder der Blockschaltbilder bzw. Schaubilder angegebenen Funktion/Schritts umsetzen.

[0070] Die durch einen Computer lesbaren Programmanweisungen können auch auf einen Computer, eine andere programmierbare Datenverarbeitungsvorrichtung oder eine andere Einheit geladen werden, um das

Ausführen einer Reihe von Prozessschritten auf dem Computer bzw. der anderen programmierbaren Vorrichtung oder anderen Einheit zu verursachen, um einen auf einem Computer ausgeführten Prozess zu erzeugen, so dass die auf dem Computer, einer anderen programmierbaren Vorrichtung oder einer anderen Einheit ausgeführten Anweisungen die in dem Block bzw. den Blöcken der Ablaufpläne und/oder der Blockschaltbilder bzw. Schaubilder festgelegten Funktionen/Schritte umsetzen.

[0071] Die Ablaufpläne und die Blockschaltbilder bzw. Schaubilder in den Figuren veranschaulichen die Architektur, die Funktionalität und den Betrieb möglicher Ausführungen von Systemen, Verfahren und Computerprogrammprodukten gemäß verschiedenen Ausführungsformen der vorliegenden Erfindung. In diesem Zusammenhang kann jeder Block in dem Ablaufplan oder in den Blockschaltbildern bzw. Schaubildern ein Modul, ein Segment oder einen Teil von Anweisungen darstellen, das bzw. der eine oder mehrere ausführbare Anweisungen zur Ausführung der bestimmten logischen Funktion(en) aufweist. In einigen alternativen Ausführungen können die in dem Block angegebenen Funktionen in einer anderen Reihenfolge als in den Figuren gezeigt stattfinden. Zwei nacheinander gezeigte Blöcke können zum Beispiel in Wirklichkeit weitgehend gleichzeitig ausgeführt werden, oder die Blöcke können manchmal je nach entsprechender Funktionalität in umgekehrter Reihenfolge ausgeführt werden. Es ist ferner anzumerken, dass jeder Block der Blockschaltbilder bzw. Schaubilder und/oder der Ablaufpläne sowie Kombinationen aus Blöcken in den Blockschaltbildern bzw. Schaubildern und/oder den Ablaufplänen durch spezielle auf Hardware beruhende Systeme umgesetzt werden können, welche die festgelegten Funktionen oder Schritte durchführen, oder Kombinationen aus Spezial-Hardware und Computeranweisungen ausführen.

[0072] Ausführungsformen der Erfindung können Endbenutzern durch eine Cloud-Computing-Infrastruktur bereitgestellt werden. Cloud-Computing bezieht sich im Allgemeinen auf die als ein Dienst über ein Netz erfolgende Bereitstellung von skalierbaren Datenverarbeitungsressourcen. Formaler ausgedrückt, Cloud-Computing kann als eine Möglichkeit der Datenverarbeitung definiert werden, die eine Abstraktion zwischen der Datenverarbeitungsressource und ihrer zugrunde liegenden technischen Architektur (z.B. Server, Speicher, Netzwerke) bereitstellt, was einen komfortablen, bedarfsweisen Netzwerkzugriff auf einen gemeinsam genutzten Bestand an konfigurierbaren Datenverarbeitungsressourcen ermöglicht, die schnell bereitgestellt und mit minimalem Verwaltungsaufwand oder minimaler Interaktion eines Diensteanbieters freigegeben werden können. Folglich ermöglicht Cloud-Computing einem Benutzer den Zugriff auf virtuelle Datenverarbeitungsressourcen (z.B. Speicher, Daten, Anwendungen und sogar ganze virtualisierte Datenverarbeitungssysteme) in „der Cloud“, ohne Rücksicht auf die zugrunde liegenden physischen Systeme (oder Standorte dieser Systeme), die zur Bereitstellung der Datenverarbeitungsressourcen verwendet werden.

[0073] Üblicherweise werden Cloud-Computing-Ressourcen einem Benutzer nutzungsorientiert bereitgestellt, wobei Benutzern nur die tatsächlich verwendeten Datenverarbeitungsressourcen berechnet werden (z.B. durch einen Benutzer verbrauchter Speicherplatz oder eine Anzahl von virtualisierten, durch den Benutzer instanziierten Systemen). Ein Benutzer kann jederzeit und von überall über das Internet auf beliebige der Ressourcen zugreifen, die sich in der Cloud befinden. Im Kontext der vorliegenden Erfindung kann ein Benutzer auf Anwendungen (z.B. eine oder mehrere der in **Fig. 3** veranschaulichten Services 122, 124, 226 und 228) oder zugehörige, in der Cloud verfügbare Daten zugreifen. Zum Beispiel könnte der Datenabrufservice 122, der Datenverarbeitungsservice 124, der Dateiwiederherstellungsservice 226 und der Dateientfernungsservice 228 auf einem Datenverarbeitungssystem in der Cloud ausgeführt werden und eine fortlaufende Speicherung von Daten in einem System mit begrenzter Speicherkapazität ermöglichen. Dadurch wird einem Benutzer ermöglicht, von einem beliebigen Datenverarbeitungssystem aus, das an ein mit der Cloud verbundenes Netzwerk (z.B. das Internet) angeschlossen ist, auf diese Informationen zugreifen.

[0074] Während das Vorstehende auf Ausführungsformen der vorliegenden Erfindung gerichtet ist, können andere und weitere Ausführungsformen der Erfindung entwickelt werden, ohne vom grundlegenden Umfang der Erfindung abzuweichen, und der Umfang der Erfindung wird von den Ansprüchen, die folgen, bestimmt.

Patentansprüche

1. Durch einen Computer ausgeführtes Verfahren, das umfasst:
Empfangen einer Anforderung für einen Zugriff auf Daten an einem ersten Speicherort;
Feststellen, auf der Grundlage eines Katalogs und unter Verwendung eines Computerprozessors, dass die Daten aus dem ersten Speicherort entfernt wurden, und als Reaktion darauf automatisches Wiederherstellen der Daten an dem ersten Speicherort unter Verwendung des Computerprozessors, wobei das Wiederherstellen aufweist:
Angaben eines zweiten Speicherorts für die Daten auf der Grundlage des Katalogs;

Abrufen der Daten aus dem zweiten Speicherort; und
Speichern der Daten an dem ersten Speicherort; und
Aktualisieren des Katalogs, um das Wiederherstellen der Daten an dem ersten Speicherort widerzuspiegeln.

2. Verfahren nach Anspruch 1, wobei der erste Speicherort nicht öffentlich zugänglich ist, wobei der zweite Speicherort öffentlich zugänglich ist und wobei der Katalog aufweist: (i) eine Tabelle in einer elektronischen Datenbank und/oder (ii) zu den angeforderten Daten gehörende Metadaten.

3. Verfahren nach Anspruch 1 oder 2, wobei das Wiederherstellen der Daten des Weiteren aufweist:
Angabe einer ersten Menge an Speicherplatz, die zur Speicherung der Daten an dem ersten Speicherort erforderlich ist, und einer zweiten Menge von verfügbarem Speicherplatz an dem ersten Speicherort;
Entscheiden, auf der Grundlage der ersten Menge und der zweiten Menge, Speicherplatz an dem ersten Speicherort freizugeben;
Angabe, auf der Grundlage des Katalogs, einer aus dem ersten Speicherort zu entfernenden Datei;
Entfernen der Datei aus dem ersten Speicherort; und
Aktualisieren des Katalogs, um das Entfernen der Datei aus dem ersten Speicherort widerzuspiegeln.

4. Verfahren nach Anspruch 3, wobei das Angeben, auf der Grundlage des Katalogs, der aus dem ersten Speicherort zu entfernenden Datei des Weiteren aufweist:
Feststellen, dass der Katalog einen öffentlich zugänglichen dritten Speicherort für die Datei aufweist.

5. Verfahren nach Anspruch 4, wobei das Angeben, auf der Grundlage des Katalogs, der aus dem ersten Speicherort zu entfernenden Datei des Weiteren aufweist:
Auswählen der zu entfernenden Datei aus einer Mehrzahl von in dem Katalog angegebenen Dateien auf der Grundlage von einem oder mehreren Wiederherstellungsparametern, die zu der Datei gehören und in dem Katalog gespeichert sind.

6. Verfahren nach Anspruch 5, wobei die Wiederherstellungsparameter eines oder mehr aufweisen von:
(i) ein letztes Zugriffsdatum für die Datei, (ii) eine zu der Datei gehörende Größe, (iii) eine Zeitspanne für das Herunterladen der Datei, (iv) eine Zeitspanne für das Verarbeiten der Datei, (v) eine Menge an Verarbeitungsressourcen, die zum Verarbeiten der Datei in Anspruch genommen werden, oder (vi) eine Menge an Hauptspeicherressourcen, die zum Verarbeiten der Datei in Anspruch genommen werden.

7. Verfahren nach einem der vorhergehenden Ansprüche, wobei ein automatisches Wiederherstellen der Daten an dem ersten Speicherort des Weiteren aufweist:
Angabe eines zur Verarbeitung der Daten verwendeten Skripts;
Verarbeiten der abgerufenen Daten unter Verwendung des Skripts; und
Speichern der verarbeiteten Daten an dem ersten Speicherort.

8. Verfahren nach Anspruch 7, wobei ein Verarbeiten der abgerufenen Daten unter Verwendung des Skripts des Weiteren aufweist:
Angabe von mindestens einem zu der Verarbeitung gehörenden Eingabeparameter;
und
Bereitstellen des mindestens einen Eingabeparameters für das Skript.

9. Verfahren nach Anspruch 7 oder 8, wobei das Skript aufweist: (i) eine Datei, die einen oder mehrere Verarbeitungsschritte aufweist, und/oder (ii) einen an einem Netzspeicherort zugänglichen Software-Service.

10. Verfahren nach Anspruch 7, 8 oder 9, wobei ein Angeben des Skripts auf in dem Katalog gespeicherten Informationen beruht.

11. System, das aufweist:
einen Prozessor; und
einen Hauptspeicher, der ein Programm enthält, das, wenn es auf dem Prozessor ausgeführt wird, eine Operation durchführt, wobei die Operation aufweist:
Empfangen einer Anforderung für einen Zugriff auf Daten an einem ersten Speicherort;
Feststellen, auf der Grundlage eines Katalogs und unter Verwendung des Prozessors, dass die Daten aus dem ersten Speicherort entfernt wurden, und als Reaktion darauf automatisches Wiederherstellen der

Daten an dem ersten Speicherort unter Verwendung des Prozessors, wobei das Wiederherstellen aufweist: Angeben eines zweiten Speicherorts für die Daten auf der Grundlage des Katalogs; Abrufen der Daten aus dem zweiten Speicherort; und Speichern der Daten an dem ersten Speicherort; und Aktualisieren des Katalogs, um das Wiederherstellen der Daten an dem ersten Speicherort widerzuspiegeln.

12. System nach Anspruch 11, wobei der erste Speicherort nicht öffentlich zugänglich ist, wobei der zweite Speicherort öffentlich zugänglich ist und wobei der Katalog aufweist: (i) eine Tabelle in einer elektronischen Datenbank und/oder (ii) zu den angeforderten Daten gehörende Metadaten.

13. System nach Anspruch 11 oder 12, wobei das Wiederherstellen der Daten des Weiteren aufweist: Angeben einer ersten Menge an Speicherplatz, die zur Speicherung der Daten an dem ersten Speicherort erforderlich ist, und einer zweiten Menge von verfügbarem Speicherplatz an dem ersten Speicherort; Entscheiden, auf der Grundlage der ersten Menge und der zweiten Menge, Speicherplatz an dem ersten Speicherort freizugeben; Angeben, auf der Grundlage des Katalogs, einer aus dem ersten Speicherort zu entfernenden Datei; Entfernen der Datei aus dem ersten Speicherort; und Aktualisieren des Katalogs, um das Entfernen der Datei aus dem ersten Speicherort widerzuspiegeln.

14. System nach Anspruch 13, wobei das Angeben, auf der Grundlage des Katalogs, der aus dem ersten Speicherort zu entfernenden Datei des Weiteren aufweist: Feststellen, dass der Katalog einen öffentlich zugänglichen dritten Speicherort für die Datei aufweist; und Auswählen der zu entfernenden Datei aus einer Mehrzahl von in dem Katalog angegebenen Dateien auf der Grundlage von einem oder mehreren Wiederherstellungsparametern, die zu der Datei gehören und in dem Katalog gespeichert sind.

15. System nach Anspruch 14, wobei die Wiederherstellungsparameter eines oder mehr aufweisen von: (i) ein letztes Zugriffsdatum für die Datei, (ii) eine zu der Datei gehörende Größe, (iii) eine Zeitspanne für das Herunterladen der Datei, (iv) eine Zeitspanne für das Verarbeiten der Datei, (v) eine Menge an Verarbeitungsressourcen, die zum Verarbeiten der Datei in Anspruch genommen werden, oder (vi) eine Menge an Hauptspeicherressourcen, die zum Verarbeiten der Datei in Anspruch genommen werden.

16. System nach einem der Ansprüche 11 bis 15, wobei ein automatisches Wiederherstellen der Daten an dem ersten Speicherort des Weiteren aufweist: Angeben eines zur Verarbeitung der Daten verwendeten Skripts; Verarbeiten der abgerufenen Daten unter Verwendung des Skripts; und Speichern der verarbeiteten Daten an dem ersten Speicherort.

17. System nach Anspruch 16, wobei ein Verarbeiten der abgerufenen Daten unter Verwendung des Skripts des Weiteren aufweist: Angeben von mindestens einem zu der Verarbeitung gehörenden Eingabeparameter; und Bereitstellen des mindestens einen Eingabeparameters für das Skript.

18. System nach Anspruch 16 oder 17, wobei das Skript aufweist: (i) eine Datei, die einen oder mehrere Verarbeitungsschritte aufweist, und/oder (ii) einen an einem Netzspeicherort zugänglichen Software-Service.

19. System nach Anspruch 16, 17 oder 18, wobei das Angeben des Skripts auf in dem Katalog gespeicherten Informationen beruht.

20. Computerprogrammprodukt zum automatischen Wiederherstellen von aus einem Speicher entfernten Daten, wobei das Computerprogrammprodukt aufweist: ein durch einen Computer lesbares Speichermedium mit damit verkörpertem, durch einen Computer lesbaren Programmcode, wobei der durch einen Computer lesbare Programmcode durch einen oder mehrere Computerprozessoren ausführbar ist, um eine Operation durchzuführen, die aufweist: Empfangen einer Anforderung für einen Zugriff auf Daten an einem ersten Speicherort; Feststellen, auf der Grundlage eines Katalogs, dass die Daten aus dem ersten Speicherort entfernt wurden, und als Reaktion darauf automatisches Wiederherstellen der Daten an dem ersten Speicherort, wobei das Wiederherstellen aufweist:

Angeben eines zweiten Speicherorts für die Daten auf der Grundlage des Katalogs;
Abrufen der Daten aus dem zweiten Speicherort; und
Speichern der Daten an dem ersten Speicherort; und
Aktualisieren des Katalogs, um das Wiederherstellen der Daten an dem ersten Speicherort widerzuspiegeln.

21. Computerprogrammprodukt nach Anspruch 20, wobei der erste Speicherort nicht öffentlich zugänglich ist, wobei der zweite Speicherort öffentlich zugänglich ist und wobei der Katalog aufweist: (i) eine Tabelle in einer elektronischen Datenbank und/oder (ii) zu den angeforderten Daten gehörende Metadaten.

22. Computerprogrammprodukt nach Anspruch 20 oder 21, wobei das Wiederherstellen der Daten des Weiteren aufweist:

Angeben einer ersten Menge an Speicherplatz, die zur Speicherung der Daten an dem ersten Speicherort erforderlich ist, und einer zweiten Menge von verfügbarem Speicherplatz an dem ersten Speicherort;
Entscheiden, auf der Grundlage der ersten Menge und der zweiten Menge, Speicherplatz an dem ersten Speicherort freizugeben;

Angeben, auf der Grundlage des Katalogs, einer aus dem ersten Speicherort zu entfernenden Datei;
Entfernen der Datei aus dem ersten Speicherort; und
Aktualisieren des Katalogs, um das Entfernen der Datei aus dem ersten Speicherort widerzuspiegeln.

23. Computerprogrammprodukt nach Anspruch 22, wobei das Angeben, auf der Grundlage des Katalogs, der aus dem ersten Speicherort zu entfernenden Datei des Weiteren aufweist:

Feststellen, dass der Katalog einen öffentlich zugänglichen dritten Speicherort für die Datei aufweist; und
Auswählen der zu entfernenden Datei aus einer Mehrzahl von in dem Katalog angegebenen Dateien auf der Grundlage von einem oder mehreren Wiederherstellungsparametern, die zu der Datei gehören und in dem Katalog gespeichert sind.

24. Computerprogrammprodukt nach Anspruch 20, wobei ein automatisches Wiederherstellen der Daten an dem ersten Speicherort des Weiteren aufweist:

Angeben eines zur Verarbeitung der Daten verwendeten Skripts;
Verarbeiten der abgerufenen Daten unter Verwendung des Skripts; und
Speichern der verarbeiteten Daten an dem ersten Speicherort.

25. Computerprogramm, das Programmcode-Mittel aufweist, die so ausgelegt sind, dass sie das Verfahren nach einem der Ansprüche 1 bis 10 durchführen, wenn das Programm auf einem Computer ausgeführt wird.

Es folgen 7 Seiten Zeichnungen

Anhängende Zeichnungen

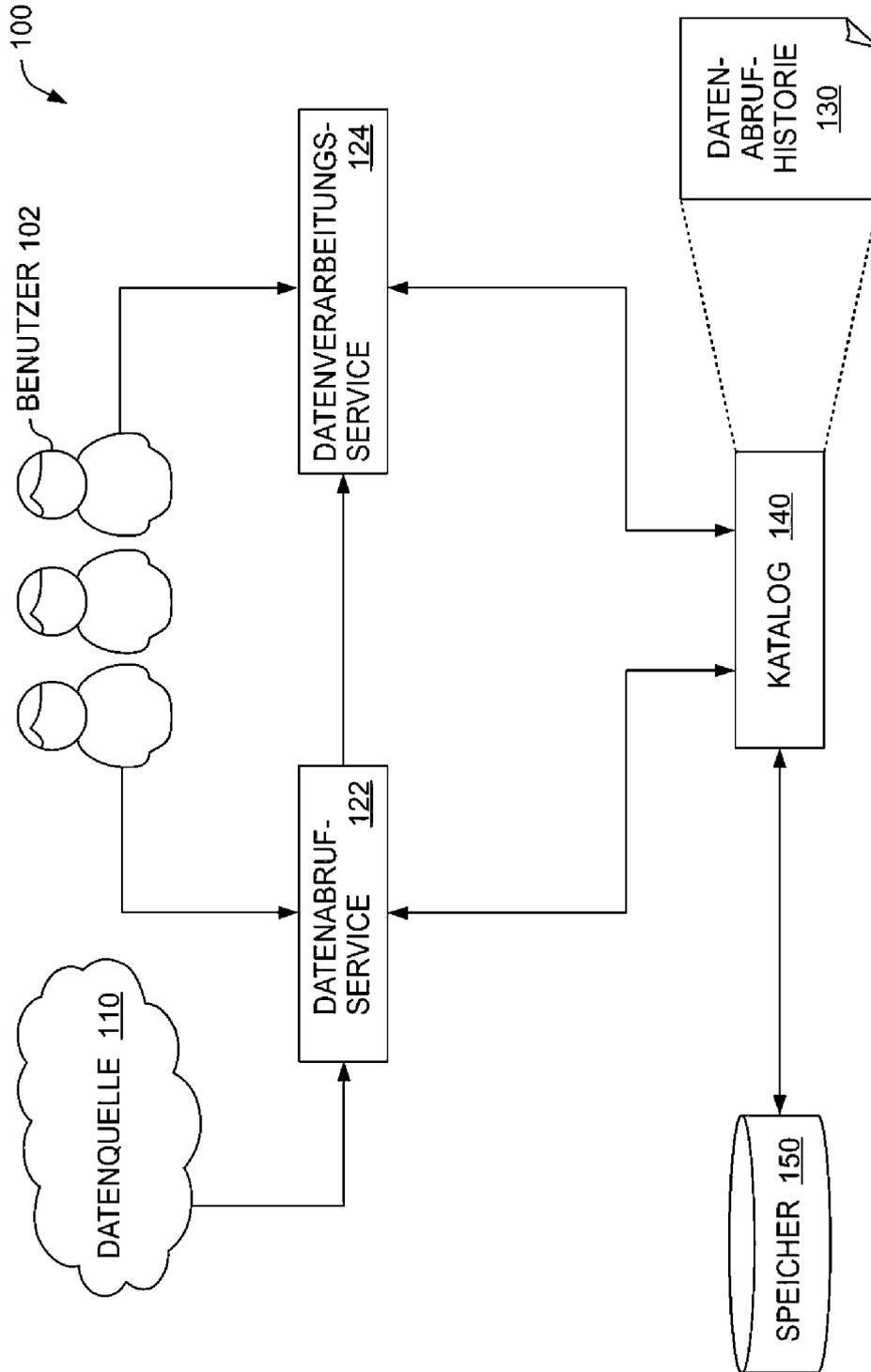


FIG. 1

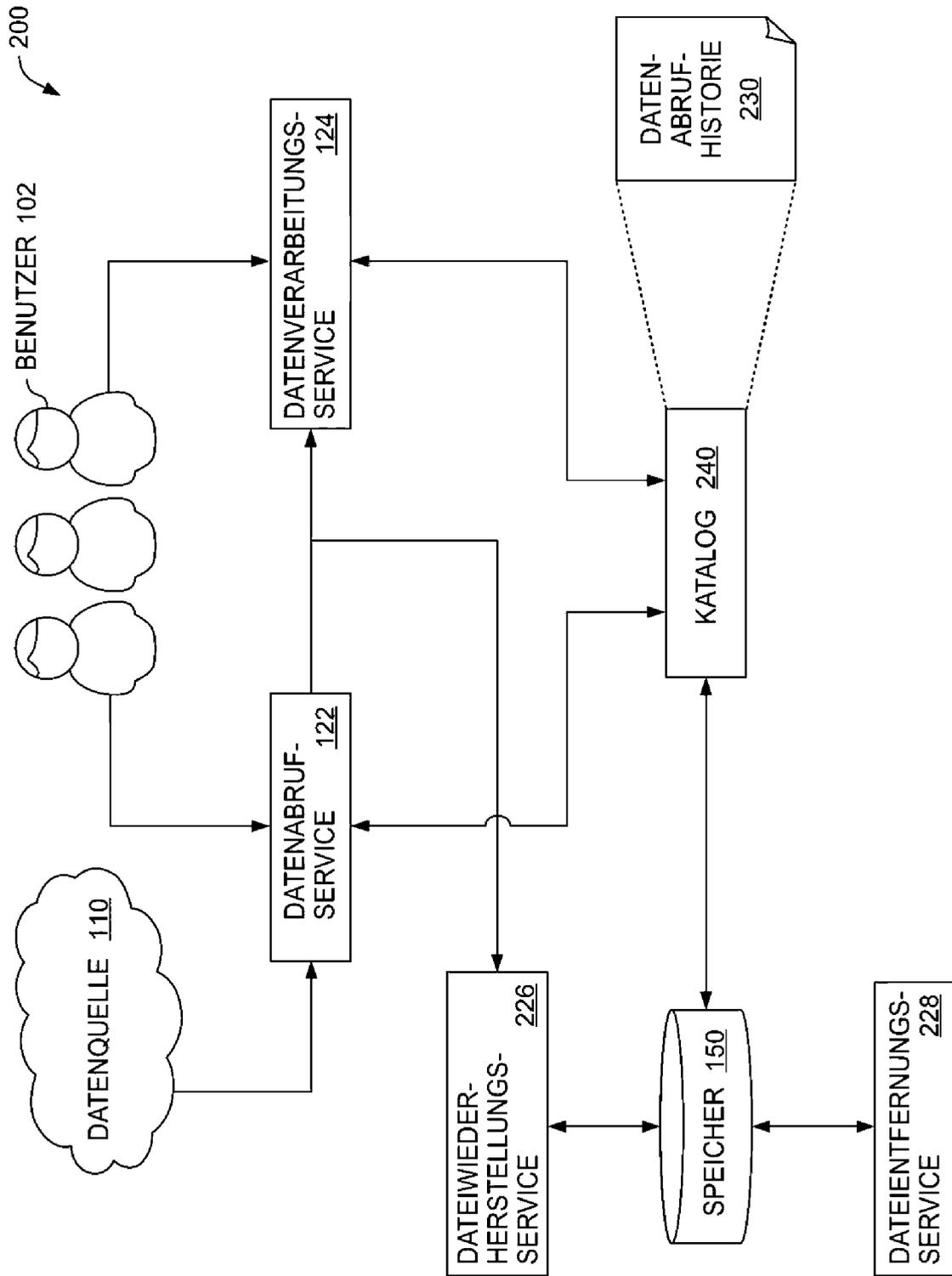


FIG. 2

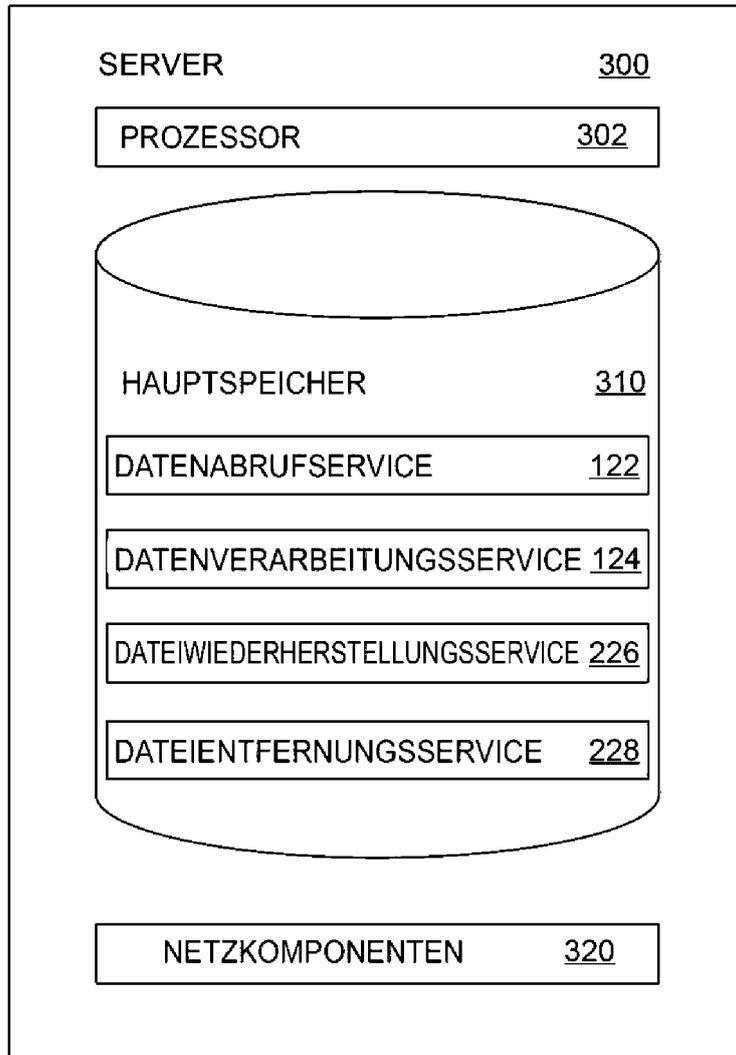


FIG. 3

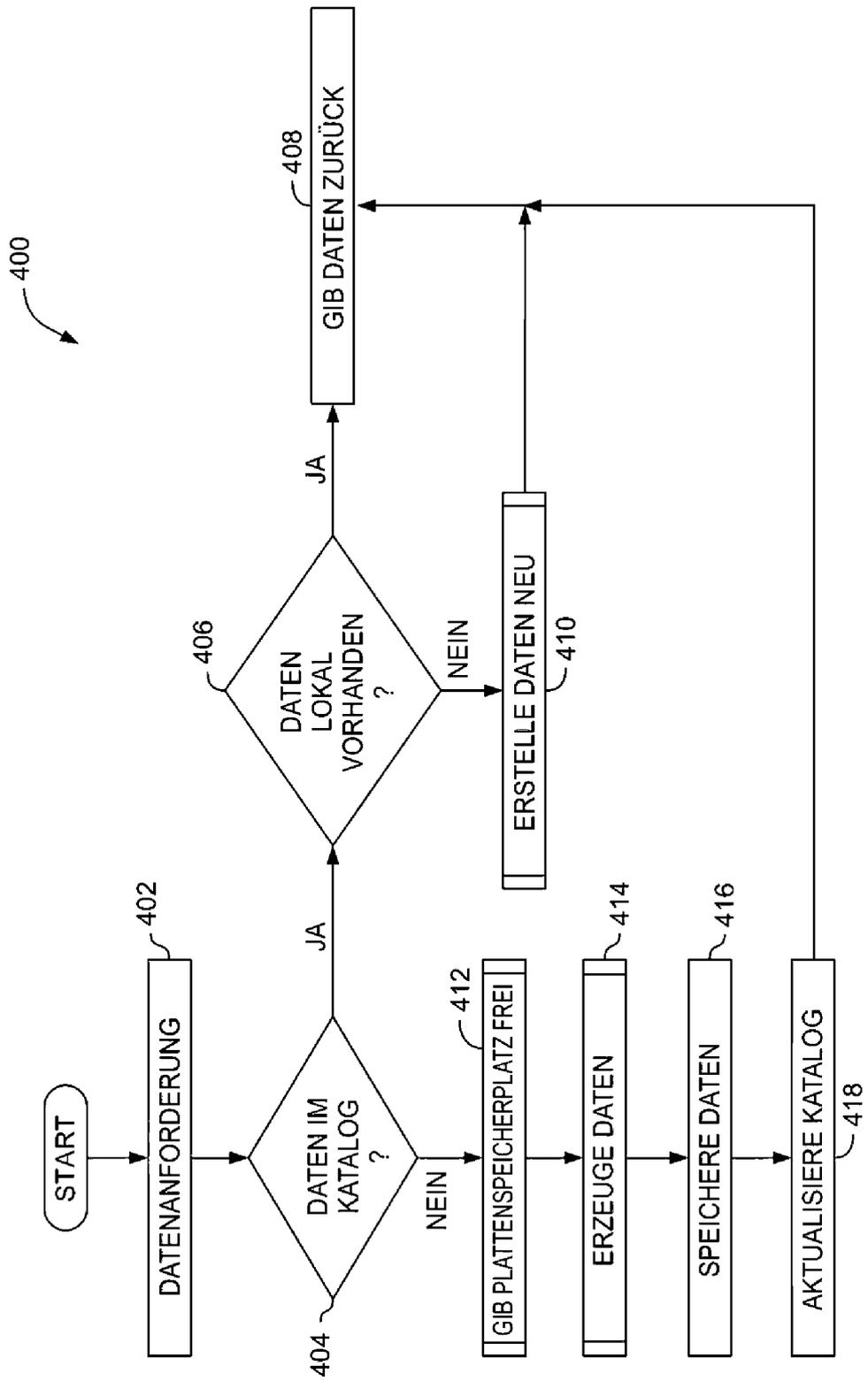


FIG. 4

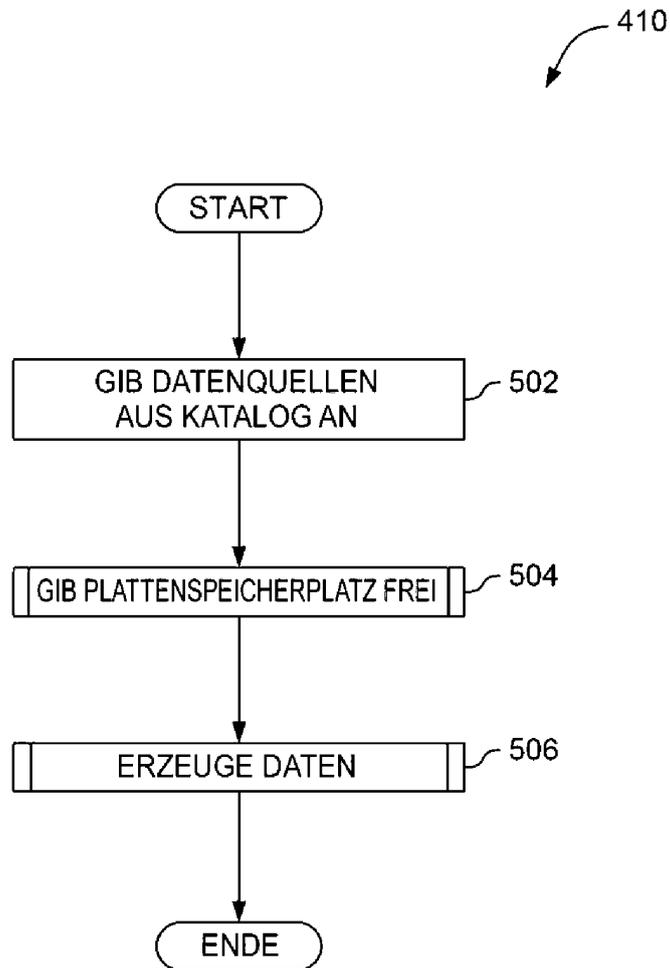


FIG. 5

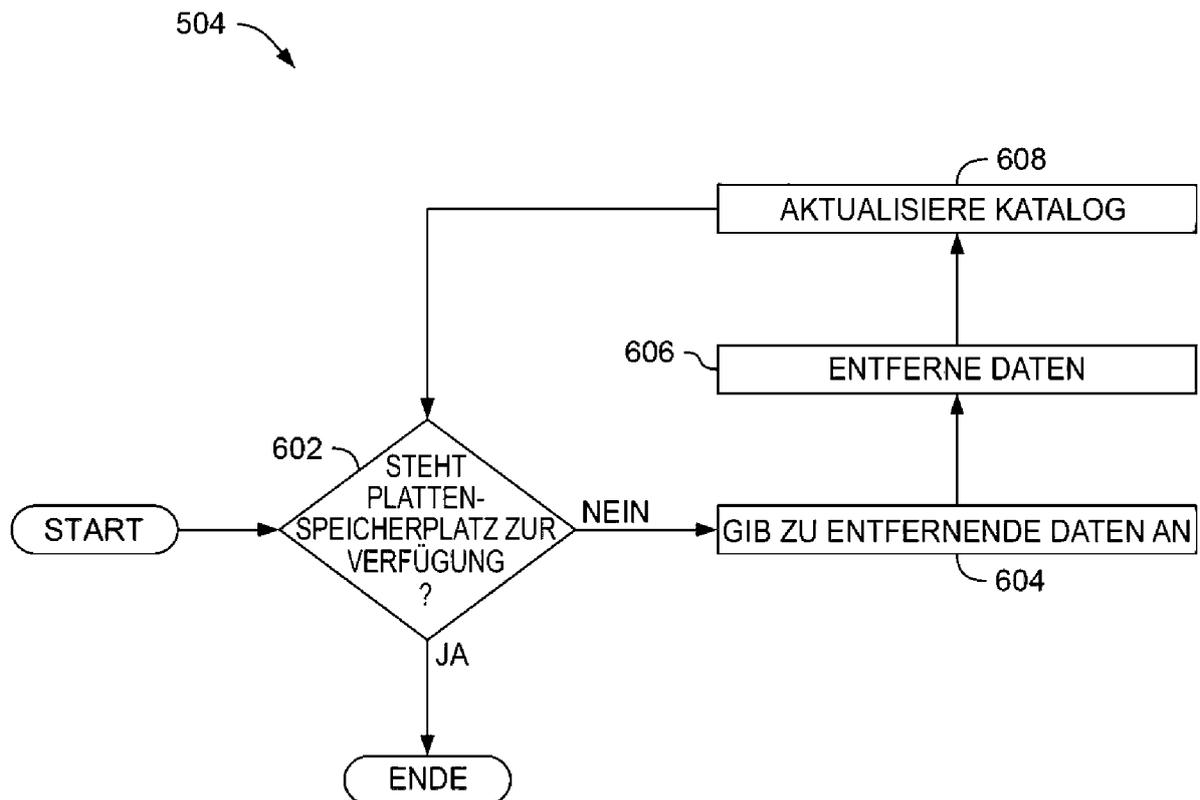


FIG. 6

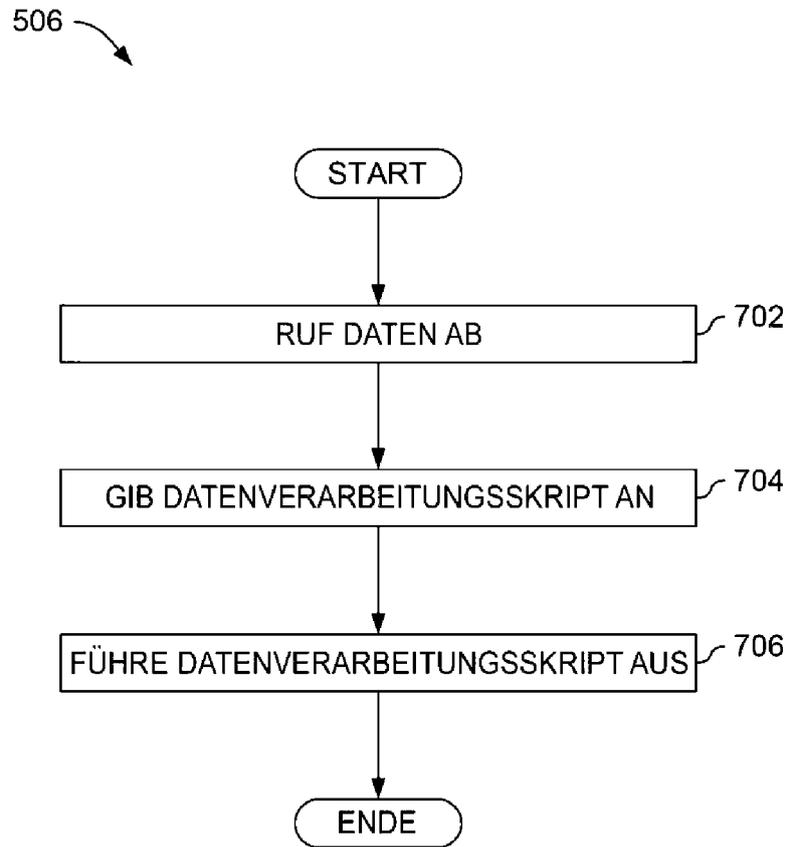


FIG. 7