



Bausteine einer modernen Analyseplattform

Von Dan Kogan und Jen Underwood

Was ist „moderne Analytik“ und warum brauchen wir sie?

Wir leben in einer aufregenden Zeit beschleunigter Innovation, eines verstärkten globalen Wettbewerbs und einer beispiellosen Chance für bahnbrechende Änderungen und neue Erfindungen. Das exponentielle Wachstum der digitalen Technologie, die allgegenwärtige Internetanbindung und die Verbreitung intelligenter Geräte, die alle eine Flut von Daten erzeugen, die einen Wettbewerbsvorteil für jeden bieten, der sie schnell und genau verstehen kann, ermöglichen eine vierte industrielle Revolution.

Das Potenzial moderner Technik ist nicht zu unterschätzen. Die konkrete Grenze zwischen physischer und virtueller Welt wird aufgehoben, die Hindernisse für den Markteintritt werden abgebaut und völlig neue Geschäftsmodelle werden geschaffen. Jeder Industriezweig ist reif für Umwälzungen, selbst jene, die vor zehn Jahren Umwälzungen ausgelöst haben. Betrachten Sie die Auswirkung des Crowdfunding auf die Kreditsituation von Kleinbetrieben, des Online-Einzelhandels auf den erweiterten Einzelhandelsmarkt oder der Medienabonnements auf das traditionelle Kabelfernsehen. Überall in Ihrem Umfeld beschleunigt sich dieser Wandel: Automatisierte Prozesse, Messenger-Bots und künstliche Intelligenz sind nur eine Handvoll neuer Technologien, die den Weg in eine spannende, wenn auch unsichere Zukunft ebnen.



Wenn Sie mehr über die veränderte Bedeutung des beliebtesten BI-Sammelbegriffs erfahren möchten, lesen Sie unser Whitepaper „**Analytics neu definiert**“, in dem die Schlüsselbegriffe rund um die Analytik von heute beleuchtet werden.

Doch all diese Trends haben eine gemeinsame Basis: riesige Datenmengen.

Wenn Daten ein Rohstoff sind, dann ist Analytik der Prozess, um sie zu nützlichen Informationen zu veredeln und Ihrem Unternehmen letztendlich einen Wettbewerbsvorteil zu verschaffen. Daten sind heute lebenswichtiger als je zuvor. Wenn Unternehmen mit dem Wandel der Branche Schritt halten wollen, müssen sie sich auf immer mehr präzise und zeitnahe Daten verlassen, um schnelle und intelligente Entscheidungen treffen zu können. Doch die Analysetechnologie hat sich traditionell nur langsam und mühsam an das Wachstum und die sich ändernden Datenformen angepasst. Es scheint, als gäbe es jede Woche einen neuen Datenbanktyp, Daten, die aus einer neuen Gerätesuite generiert werden, und alles ist mit einer völlig neuen Technologie aufgebaut. Die Analytik hat einfach nicht Schritt gehalten.



Um mehr darüber zu erfahren, wie wichtig die Zusammenarbeit zwischen IT und Geschäftsbetrieb ist, um einen modernen Ansatz für die Unternehmensanalyse zu ermöglichen, lesen Sie unser Whitepaper: **Wie eine Kultur für Selfservice-Analysen aufgebaut wird.**

Um eine datengesteuerte Kultur für das digitale Zeitalter von heute zu schaffen, die bereit ist, die Herausforderungen der heutigen Geschäftswelt mit Genauigkeit und Schnelligkeit zu meistern, müssen Unternehmen nicht nur in neue Technologien investieren, sondern auch in neue Wege der Bereitstellung dieser Informationen, einschließlich der Personen, die für die Leitung des Vorhabens erforderlich sind. Dieser parallele kulturelle Wandel ist eine grundlegende Veränderung der Beziehung zwischen IT und Geschäftsbetrieb. Sie sind zwei Partner, die daran arbeiten, Daten zu sammeln und zu nutzen, aber auch zu verfeinern und bei Bedarf die richtigen Informationen bereitzustellen. Wenn IT



Wenn IT-Organisationen für die Umstellung auf Selfservice-Analysen verantwortlich sind, sind Governance und Sicherheit in hohem Maße gewährleistet. Wenn IT-Abteilungen darüber hinaus Geschäftsanwendern Möglichkeiten für datengesteuerte und agile Entscheidungen an die Hand geben, werden sie gleichzeitig zu vertrauensvollen Partnern des Geschäftsbetriebs.

- COLIN REES, CIO, DOMINO'S PIZZA GROUP LIMITED

und Geschäftsbetrieb zusammenarbeiten, können Unternehmen den Wunschtraum einer Selfservice-Analysen-Kultur in die Realität umsetzen.

Dieses Whitepaper untersucht die Bausteine, die eine moderne Analyseplattform ausmachen, mit der Geschäftsbetrieb und IT gemeinsam Daten, Werte und Entscheidungen für das gesamte Unternehmen bereitstellen können. Dazu gehören die neuen, glänzenden Tools von heute ebenso wie die traditionellen Werkzeuge, die seit Jahrzehnten die Grundlage für Business Intelligence bilden. Wir zeigen Ihnen, wie jeder Baustein in den Gesamtprozess der Umwandlung von Daten in Erkenntnisse passt, ob es sich dabei um die Tools handelt, die Sie für die Erfassung und Auswertung von Daten benötigen oder um neue Technologien, mit denen Sie interaktive Erkenntnisse austauschen können. Wir werden auch darüber diskutieren, wie Tableau sowohl als Grundlage für eine moderne Analyseplattform als auch als Katalysator für die Schaffung einer neuen datengetriebenen Analysekultur dienen kann.

Inhaltsverzeichnis

Was ist „moderne Analytik“ und warum brauchen wir sie?	2
1. Die drei aktuellen Datenherausforderungen	4
Daten gibt es überall	4
Jeder braucht Daten	5
Daten sind ständig in Bewegung	5
2. Die Bausteine einer modernen Analyseplattform	6
IT-aktiviert	7
Verfassen und verbrauchen	9
3. Alles miteinander verbinden	12
4. Anhang	13
Einbeziehung von Streams	13
Integration der Hub-Orchestrierung	13
Weitere Informationen zu unstrukturierten Daten, NoSQL und Datenseen	13
Daten als Dienstleistung	15
Logisches Data Warehouse	15
Maschinelles Lernen	17
Natürliche Sprache	17
Empfohlene Datenuntersuchung	17
Suche	17
Warnungen	18
Storytelling	18
Über Tableau	19

1. Die drei aktuellen Datenherausforderungen

Daten sind überall und werden sekundenschnell im Überfluss generiert. Zum Beispiel ist eine Zahnbürste jetzt ein intelligentes Gerät, das erfasst, wann und wie lange Sie putzen, und das den Status der Komponenten protokolliert. Sie kann all diese Daten auch an Ihren Zahnarzt übermitteln. Eine einfache Aufgabe besteht jetzt aus tausenden von Datenpunkten. Multiplizieren Sie dieses eine Beispiel mit Millionen von Geräten, und Sie haben jetzt eine einzige, kleine Branche, die Berge von Daten erzeugt, die es vor einigen Jahren noch nicht gab. Fügen Sie Ereignisprotokolle, APIs, Soziale Medien, Website-Tracking und eine Vielzahl anderer Internettechnologien hinzu, und Ihre Daten explodieren jedes Mal, wenn Sie mit den Augen blinzeln.

Dieses moderne Systemumfeld weist drei geschäftliche Herausforderungen auf:

1. Daten gibt es überall
2. Jeder braucht Daten
3. Daten sind ständig in Bewegung

Daten gibt es überall



Abbildung 1 Tableau stellt von überall Verbindungen zu Daten her

Traditionell haben die meisten Unternehmen ihre Daten am Standort aufbewahrt. Sie arbeiteten hart daran, alle Daten zu kontrollieren, die sowohl erstellt als auch in einem scheinbar klar definierten Data Warehouse gespeichert wurden. Wenn es Daten gab, die Sie nicht erfasst haben, waren sie unbedeutend und wahrscheinlich nicht der Mühe wert.



Weitere Informationen über die aktuellen Trends bei der Cloud-Migration finden Sie in [Der Bericht zu Cloud-Daten](#).

Durch diese Mentalität kann ein Unternehmen in der heutigen Zeit ausgelöscht werden, in der Websites, Mobilgeräte und Cloudanwendungen Daten außerhalb eines Unternehmens generieren. Denken Sie an Google Analytics, Splunk, ServiceNow und Salesforce, um nur einige zu nennen. Dieser Trend wird sich weiter beschleunigen, da immer mehr nützliche Daten in der Cloud von Drittanbietern generiert werden. Um diesen Hang zur Cloud zu ergänzen, verschieben einige Unternehmen ihre eigene lokale Infrastruktur auch in die Cloud.

Jeder braucht Daten

Innerhalb der letzten zwei Jahrzehnte haben wir eine Marktverschiebung in Richtung Digital Business und Cloud erlebt. Dies ist ein Schlüssel zur modernen Analytikrevolution. Zum anderen geht es darum, eine datengesteuerte Kultur mit Selfservice-Business Intelligence zu ermöglichen. Eine Kultur der Analytik durchdringt die innovativsten Unternehmen von heute, in denen der Geschäftsanwender mit einer Frage die Antwort selbst herausfinden kann. Damit werden Unternehmen zu Meistern der Verfeinerung von Daten in Informationen, in digitales Gold, mit rasanter Geschwindigkeit. Damit der Wandel zu einer Analysekultur gelingt, muss ein Unternehmen Mitarbeiter und Daten zusammenführen, die zweifellos zu den wichtigsten Ressourcen eines Unternehmens gehören. Jeder muss auf die richtigen Daten zugreifen können und motiviert sein, diese zu untersuchen und zusammenzuarbeiten.

Bei einem modernen Analyseansatz arbeiten IT und Geschäftsbetrieb Hand in Hand. Die IT bietet eine zentralisierte Umgebung, in der Geschäftsanwender vertrauenswürdige Daten und Inhalte finden. Sie ermutigt alle, diese sicher zu nutzen, Fragen zu stellen, zu experimentieren und Entscheidungen schnell zu treffen. Dabei handelt es sich um eine Bottom-up-Methodik, die sich aus Fachexperten zusammensetzt, die Metadaten, Geschäftsregeln und Berichtsmodelle für flüssige Agilität und kontinuierliche, beschleunigte Verbesserung generieren.



Als wir mit Tableau angefangen haben, haben wir nur an Dashboards und Berichte gedacht. Wir hätten nie gedacht, dass Tableau die DNS unserer Organisation so grundsätzlich verändern würde. Es geht nicht nur um eine Lösung oder eine Technologie, sondern darum, wie sich die Haltung gegenüber Daten gewandelt hat.

- ASHISH BRAGANZA, DIRECTOR OF GLOBAL BUSINESS INTELLIGENCE

Daten sind ständig in Bewegung

Wir alle wissen es: Die einzig wahre „Konstante“ ist die „Veränderung“. Eine moderne Analyseplattform setzt auf Flexibilität, die Möglichkeit, Daten plattformübergreifend zu verschieben, die Infrastruktur bei Bedarf anzupassen, neue Datentypen zu nutzen und neue Anwendungsfälle zu ermöglichen. Zusätzlich scheint es, dass jeden Tag eine neue Technologie zur Datenanalyse für die Welt freigesetzt wird – Technologien wie maschinelles Lernen, Sprachassistenten und natürlichsprachliche Abfragen. Einige davon scheinen zumindest vorerst eher fiktiv als praktisch zu sein, aber neue Methoden und Techniken müssen reifen und ihren Nutzen für Ihre Kunden und Ihr Unternehmen unter Beweis stellen.

In einer Welt sich schnell entwickelnder Daten ist Flexibilität von höchster Bedeutung, sowohl für Ihren wachsenden Infrastrukturbedarf als auch für neue Technologien. Flexibilität ist entscheidend für die Schaffung und Erhaltung eines klaren Wettbewerbsvorteils. Wenn Sie eine zukunftsorientierte Analytics-Architektur in Betracht ziehen, sollten Sie eine Herstellerbindung durch proprietäre Architekturen vermeiden, da dies Ihre Agilität in der Zukunft erheblich beeinträchtigen kann.

2. Die Bausteine einer modernen Analyseplattform

Die drei Herausforderungen, vor denen Unternehmen heute stehen, sind nicht so unüberwindbar, wie sie scheinen. Wenn Daten in der heutigen sich entwickelnden Geschäftswelt der rote Faden sind, dann ist eine moderne Analyseplattform der Schlüssel zur Erschließung ihres Potenzials. Aber eine moderne Analyseplattform ist nicht eine Infrastruktureinheit: Sie umfasst mehrere, unabhängige Bausteine. Einige davon sind traditionelle Elemente der Business Intelligence, die gerade erst eingeführt wurden (z. B. Data Warehouses). Andere sind völlig neue Konzepte, die die Art und Weise revolutioniert haben, wie Unternehmen die Datenanalyse in erster Linie angehen (z. B. Visual Analytics). Beide Arten von Bausteinen bilden zusammen eine Analyseplattform, die jedem Unternehmen helfen kann, sich den Herausforderungen heutiger Unternehmen zu stellen.

Eine moderne Analyseplattform kann auf zwei verschiedene Hälften reduziert werden:

- **IT-aktiviert**, umfasst das Sammeln, Verwalten und Aufbereiten von Daten.
- **Verfassen und verbrauchen**, was die Analyse von Daten und die Vermittlung von Erkenntnissen an die richtigen Stakeholder beinhaltet.

Traditionell wurden beide Hälften zu einem Prozess zusammengeführt, der sich auf den IT-Bereich beschränkt, und zwar nur auf diesen. Wir betrachten jetzt die erste Hälfte – das Erstellen und Verarbeiten von Datenquellen – als IT-fähig an. Die zweite Hälfte – Analyse und Bereitstellung – wird nach wie vor von der IT-Abteilung unterstützt, jedoch von den Geschäftsanwendern selbst durchgeführt.

Diese Kombination aus beidem ist eine echte Partnerschaft zwischen Geschäftsbetrieb und IT und die moderne Art und Weise, jede Organisation zu führen, die schnell datengesteuerte Entscheidungen treffen möchte. Es wird manchmal auch als bimodale BI bezeichnet, die das Beste der herkömmlichen BI und des operativen Berichtswesens beibehalten und gleichzeitig den Selfserviceansatz moderner Analytik übernehmen.

In dieser Beziehung zwischen IT und Business entwirft die IT die Datenarchitektur und ermöglicht eine angemessene Datensicherheit und Zugriffskontrolle. Geschäftsexperten erstellen dann die analytischen Assets, die sie benötigen, wann immer sie sie brauchen. Daraus resultiert eine IT, die es jedem ermöglicht, seine eigenen kritischen Fragen effizient zu beantworten, und Geschäftsanwender, die sofort eine Frage beantworten können, wodurch eine agile Organisation entsteht, die bereit ist, die Herausforderungen der modernen Unternehmenslandschaft anzupacken.

Wir werden uns auf hoher Ebene mit den verschiedenen Bausteinen befassen, die beide Seiten einer modernen Analyseplattform umfassen, einige der Top-Trends innerhalb jeder Plattform und wichtige Konzepte, die wir im Auge behalten sollten. Für einen tiefen Einblick in spezifische Komponenten – einschließlich marktführender Anbieter, die diese bereitstellen – und um zu sehen, ob sie auf Sie zutreffen, lesen Sie die entsprechenden Abschnitte im [Anhang](#) zu den einzelnen Bausteinen.

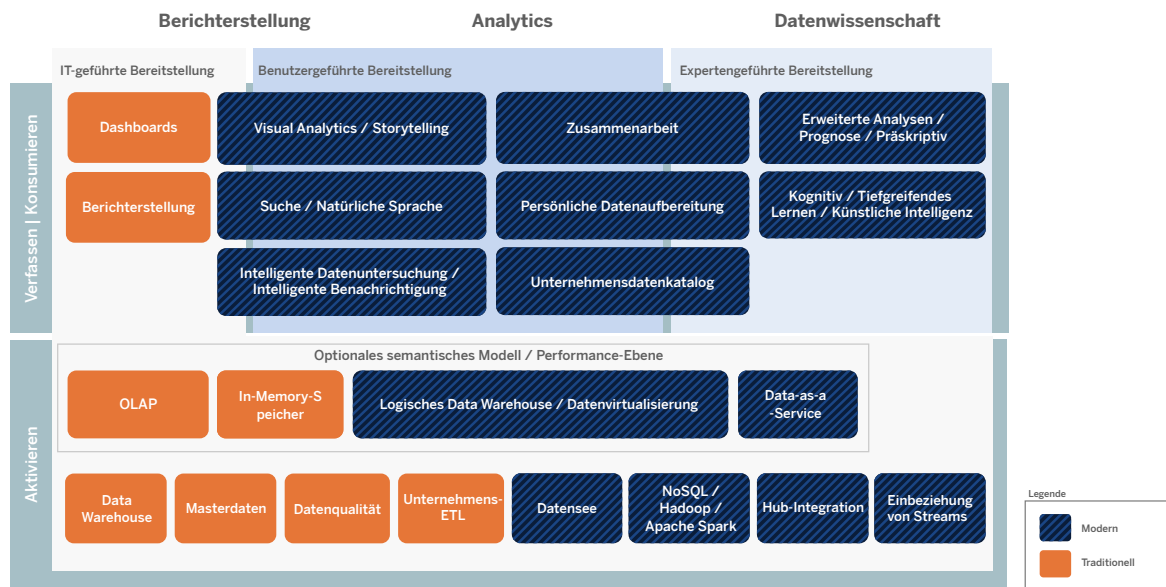


Abbildung 2 Grundbausteine

IT-aktiviert

Anders als herkömmliche IT-gesteuerte Business Intelligence konzentrieren sich die effektivsten IT-Organisationen heute auf die Bereitstellung analytischer Datenquellen, indem sie Daten für Benutzer und Experten orchestrieren, organisieren und vereinheitlichen, um sie zu erstellen und zu konsumieren. Diese Rolle liegt auf der Hand, sollte jedoch nicht überbewertet werden. Das Sammeln von Daten, das Verwalten ihrer Quellen und die Verarbeitung dieser Daten, um sie für andere nutzbar zu machen, war schon immer von entscheidender Bedeutung für Business Intelligence und bleibt das Herzstück einer modernen Analyseplattform. Welche Erkenntnisse können entdeckt werden, wenn es keinen Rohstoff für die Veredelung gibt?

Das Alleinstellungsmerkmal einer modernen Analyseplattform ist die Partnerschaft zwischen Geschäftsbetrieb und IT. Wenn Geschäftsanwender über Tools verfügen, die es ihnen ermöglichen, Daten selbständig zu analysieren, können sie Fragen im Handumdrehen beantworten und wissen, dass sie den eigentlichen Daten vertrauen können. Dies führt zu genauen, agilen Berichten und Dashboards. Und die IT-Abteilung kann ohne Dashboard und Änderungsanforderungen endlich die Daten selbst priorisieren: Schutz der Datenkontrolle und -sicherheit, Gewährleistung der Datengenauigkeit und Einrichtung der effizientesten Pipelines für Sammlung, Verarbeitung und Speicherung der Daten.

Die Priorisierung von Daten könnte zu keinem besseren Zeitpunkt erfolgen. Ihr Unternehmen, unabhängig von seiner Größe, sammelt bereits Daten und analysiert höchstwahrscheinlich einen kleinen Teil davon – der Rest sind **dunkle Daten**. Es gibt eine Milliarde Orte, an denen Daten gesammelt werden können, und weitere Tools werden auf den Markt kommen, die Ihnen helfen, so viele wie möglich davon zu sammeln. Heute finden Sie eine Reihe von Technologien, mit denen verschiedene Merkmale wie hohes Volumen, Datenortung und eine Vielzahl von Datenquellentypen bearbeitet werden. Jedes Unternehmen ist wirklich einzigartig, und Sie sollten sich die Zeit nehmen, Prioritäten zu setzen, welche Komponenten für Sie heute und in Zukunft am besten geeignet sind.

Hier sind einige Punkte zu beachten.

Weitere Einzelheiten zu bestimmten Technologien, wie z. B. Streaming-Daten und Data-as-a-Service, einschließlich spezifischer Anbietersauswahl, finden Sie im **Anhang**.

Datenbanken und Data Warehouses

Datenbanken und Data Warehouses waren jahrzehntlang die Grundpfeiler der Business Intelligence. Einige von ihnen haben nach wie vor ihren Platz in der modernen Analysearchitektur, während andere ihre Relevanz zu verlieren beginnen.

Einige der ältesten Datenbanken sind als OLAP (Online Analytical Processing) bekannt. Sie begannen als Reaktion auf Datenbanktechnologie, die langsam war, und verwendeten Aggregationen und Caching, um die Antwortzeiten für vorhersehbare Abfragen zu beschleunigen. Da die Fragen eines Unternehmens jedoch immer komplexer und schwieriger zu prognostizieren sind, kann OLAP nicht mehr mithalten und erfordert oft völlig neue Aggregationen. Zudem verliert sie an Relevanz angesichts verbesserter Datenbanktechnologie.

Heutige Datenbanken nutzen die Vorteile von Computing-Technologien, wie In-Memory- und massive parallele Verarbeitung (MPP). So können Datenbanken extrem schnelle Leistung bei linearer Skalierbarkeit liefern und gleichzeitig werden Datenspeicherung, Hardware-Speichernutzung und manchmal sogar integrierte Rechen- und Datenerfassungsfunktionen optimiert.

Darüber hinaus hat die Einführung der Cloud den Datenbanktechnologien neues Leben eingehaucht, das bei Vor-Ort-Versionen einfach nicht mithalten kann. Dies schließt die Fähigkeit ein, ohne Beschaffung von Hardware zu beginnen, elastisch zu skalieren, wenn sich geschäftliche Anforderungen ändern, und ohne ein Team zur Verwaltung der Infrastruktur aufbauen zu müssen.

Datenbanken und Data Warehouses werden in der modernen Analysearchitektur immer einen Platz haben. Sie spielen weiterhin eine entscheidende Rolle bei der Bereitstellung von kontrollierten, akkuraten und konformen Dimensionsdaten im gesamten Unternehmen für die Selfservice-Berichterstellung. Selbst Unternehmen, die andere Technologien einsetzen (z. B. Hadoop, Datenseen), behalten in der Regel relationale Datenbanken als Teil ihrer Datenquellenmischung.

NoSQL, unstrukturierte Daten und Datenseen

Datenbanken und Data Warehouses sind besonders leistungsfähig bei der Unterstützung von Analysen, wenn die Daten aus vorhersagbaren Quellen und Formaten stammen. Natürlich sind nicht alle Daten vorhersehbar. In der modernen Analytics-Architektur werden NoSQL-Datenbanken zum Bestandteil des Arsenal eines jeden Unternehmens, da sie Daten aus beliebigen Quellen schnell laden können, einschließlich Datenquellen, die nicht über genau definierte Schemata oder Formate verfügen. NoSQL-Datenbanken – gelegentlich auch Non-SQL- oder Not-Only-SQL-Datenbanken genannt – bieten im Vergleich zu herkömmlichen relationalen Datenbanken alternative Arten der Datenspeicherung, einschließlich Spalten-, Dokument-, Schlüsselwert und Diagrammspeichertypen.

Im Zusammenhang mit unstrukturierten Daten steht das Konzept der Big Data und der Datenseen. Daten werden überall erzeugt, manchmal an zufälligen Orten, und es kann frustrierend sein, alles zu sammeln und in ein brauchbares Format zu bringen. Es wurden Technologien entwickelt, damit analytische Tools eine Verbindung zu den Rohdaten herstellen können, anstatt die Daten erst in ein bestimmtes Format bringen zu müssen.

Eines davon ist der Datensee, ein Speicher-Repository, das riesige Datenmengen in ihrem nativen Format, strukturiert oder nicht, speichern kann. Die Daten können dann mit optimierten Verarbeitungsmechanismen wie APIs oder SQL-ähnlichen Sprachen analysiert werden, um sie umzuwandeln, ohne sie zuvor in ein bestimmtes Format aufbereiten zu müssen.

All diese Tools werden häufig im Zusammenhang mit dem Internet der Dinge, der Datenwissenschaft, mit Streaming-Daten und anderen unstrukturierten Analysen verwendet, wenn die Datengenerierung unvorhersehbar ist, sowohl hinsichtlich des Umfangs als auch des Ortes.

Eine Liste der Datentechnologien im Zusammenhang mit NoSQL, Hadoop und Datenseen finden Sie im [Anhang](#).

Unstrukturierte Dateien

Die Glanzzeiten von Excel- und CSV-Dateien sind noch lange nicht vorbei. Ganz gleich, ob Sie ein kleines oder ein großes Unternehmen sind, diese unstrukturierten Dateien, die scheinbar aus dem Nichts erscheinen, werden weiterhin existieren, wahrscheinlich für immer. Sie sind sogar an mehr Orten als je zuvor anzutreffen. Sie existierten bisher auf einem physischen Computer. Jetzt sind sie in Cloudspeichersystemen wie Google Drive oder Dropbox zu finden. Externe Anbieter produzieren unstrukturierte Dateien im Rahmen von Data-as-a-Service. Sie sind nützlich als Legenden für verschiedene Datenfelder, als zusätzliche Kundenrecherche oder als zusätzliche Angaben, die einen vorhandenen Datenbestand ergänzen, weil sie schnell erstellt werden können.

Wenden Sie außerdem zur richtigen Zeit die richtigen Sicherheitsmaßnahmen auf unstrukturierte Dateien an. Fördern Sie deren Verwendung bei Bedarf, insbesondere in Einzelfällen. Wenn bestimmte Dateien an Popularität gewinnen, wenden Sie die richtigen Sicherheitsprotokolle an, um sicherzustellen, dass sie sicher und nur für die richtigen Personen zugänglich sind.

Verfassen und verbrauchen

Das Markenzeichen moderner Business Intelligence ist die Einführung der Geschäftsanwender in die Business Intelligence-Plattform. Vorbei sind die Zeiten, in denen ein Entscheidungsträger einen Bericht von der IT anfordern und mehrere Tage warten musste, um einen veralteten Bericht zu bekommen, der die eigentliche Frage nicht ganz beantwortet. Heute stehen dem fragenden Entscheidungsträger auch die Tools zur Verfügung, mit denen er die Antwort selbst herausfindet. Da die IT im gesamten Unternehmen für vertrauenswürdige Daten gesorgt hat, können Geschäftsanwender bei Bedarf intelligente, datengesteuerte Entscheidungen treffen, ohne programmieren zu müssen.

Der entscheidende Baustein dabei ist das eigentliche Analysetool, das bereits am ersten Tag einsatzbereit ist. Schon bevor Sie herausfinden, welche Bausteine in diesem Abschnitt für Ihr Unternehmen am relevantesten sind, müssen Sie eine Verbindung zu Ihren Daten herstellen, wo immer sie sind. Sei es nur, um Ihre Datenquellen zu validieren und schnell zu untersuchen, während Sie sich darauf vorbereiten, Ihre gesamte Architektur aufzubauen.

Auch wenn die Datenanalyse aus verschiedenen Komponenten besteht, sind wir überzeugt, dass die visuelle Analyse der Kern ist. Ein Mittel, mit dem jeder in Ihrem Unternehmen – mit oder ohne Programmiererfahrung – eine Direktverbindung zu einer Datenquelle herstellen und Erkenntnisse daraus gewinnen kann. Das Bereitstellen dieser Art von Tools für Lehrer, Ärzte und Vertriebsmitarbeiter verwandelt Ihr Unternehmen von einer änderungsbedürftigen in eine gut geölte Maschine.

Neu bei der Business Intelligence liegt der zusätzliche Fokus darauf, wie Sie Erkenntnisse für andere freigeben. Die Benutzer sind nicht mehr nur auf Dashboards und Berichte beschränkt, sondern können komplette interaktive Anwendungen oder lange Artikel gestalten, Daten, Texte und Bilder einfügen oder sogar für den Mobileinsatz optimierte Anzeigemöglichkeiten schaffen.

Wenn Unternehmen und ihre kleineren Abteilungen wachsen, werden diese Tools zu intelligenten Produktivitätswerkzeugen für schnellen Informationsaustausch, um Datenquellen zu untersuchen, mit Dashboards auf dem neuesten Stand zu bleiben und die wichtigsten Kennzahlen zu verfolgen.

In diesem Abschnitt betrachten wir einige Komponenten der besten modernen Analysetools, die die Datenbereitstellung erleichtern. Weitere Informationen, etwa über Storytelling und Alarme, finden Sie im [Anhang](#).

Visuelle Analysen

Das menschliche visuelle System ist eines der mächtigsten Werkzeuge der Welt. Und heute ist es endlich ein integraler Bestandteil der Datenanalyse. Basierend auf Mustererkennungen, die das Gehirn bereits täglich verwendet, können visuelle Analysen ebenfalls Muster in Daten aufdecken, wie z. B. nach oben und unten geneigte Trends, unregelmäßige Aktivitätsspitzen oder bestimmte Ausreißerdatensätze.

Herkömmliche Spreadsheets erforderten eine Datenanalyse in Zeilen und Spalten, die Auswahl einer Teilmenge für die Freigabe und die Erstellung eines Diagramms. Ob durch umständliche Assistenten oder textbasierte Befehle, diese Diagramme beantworteten manchmal die Frage und ließen manchmal neue aufkommen, waren aber immer die letzte Sackgasse. Im Gegensatz dazu bieten visuelle Analysen ein elegantes, intuitives analytisches Erlebnis mit einfachen Mausektionen, bei denen Visualisierungen zu einem Teil des Prozesses werden, mehr als nur ein Diagramm als Endergebnis. Der Weg zur Erkenntnis ist ebenso wertvoll wie die Antwort.

Visuelle Analysen sind nicht nur ein hübsches Visualisierungswerkzeug. Sie sind eine Sprache, mit deren Hilfe Sie Daten kombinieren, Anomalien erkennen und Daten mit Berechnungen, Gruppierungen, WENN-Bedingungen und vielem mehr erweitern können, ohne Unterstützung durch Programmierer zu benötigen.

Traditionelle BI und Berichterstellung

Traditionelle BI, Dashboards und Berichterstellung sind immer noch vorhanden, auch wenn sie heute anders erstellt werden. Viele statische Berichte, wie z. B. Dashboards für Führungskräfte oder Finanzaudits, erforderten technische Entwicklungskompetenz für analytische Abfragen, die im Vorfeld abgerufen wurden und häufig die Änderung der zugrunde liegenden Datenmodelle erforderlich machten. All dies kann Tage, Wochen oder Monate dauern.

In einer modernen Analyseplattform beginnen viele dieser Dashboards und Berichte als Ad-hoc-Fragen, die aufgrund der Art der von ihnen beantworteten Fragen durch IT- und Datenverwalter erhärtet und verifiziert werden und letztlich herkömmliche statische Reports ablösen. Dieser aktualisierte Prozess nutzt das Fachwissen der Geschäftsanwender, während sie die Daten durchforsten, um die richtigen Antworten zu finden, wenn sich Fragen entwickeln, ändern und zu völlig neuen Fragen führen. Die Flexibilität der modernen Analytik ersetzt herkömmliche Tools, auch wenn die traditionellen Berichtsanforderungen bestehen bleiben.

Persönliche Tools zur Datenaufbereitung

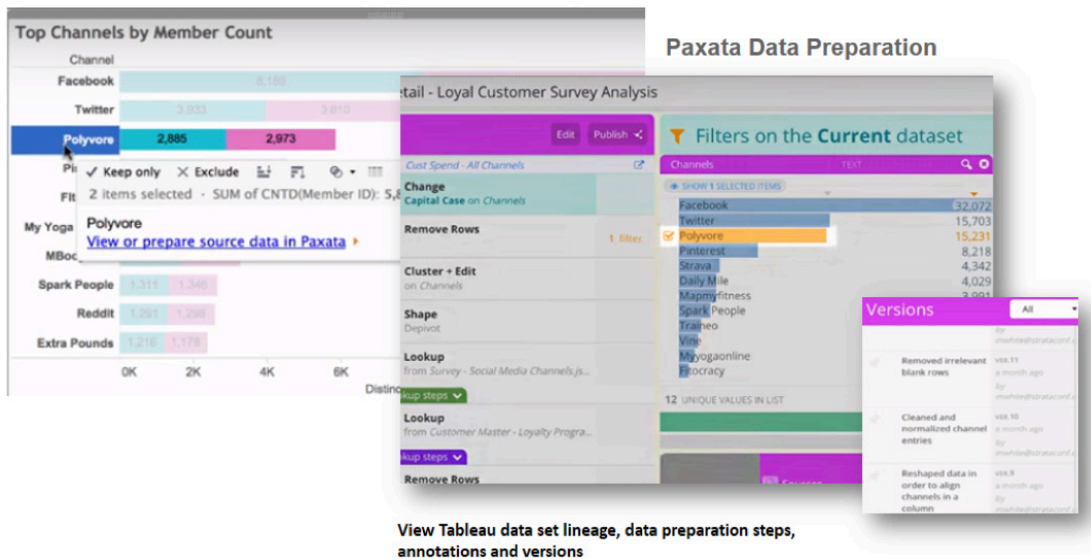
Nicht zu verwechseln mit ETL sind die Tools zur Datenvorbereitung einfache Anwendungen, die Nicht-IT-Anwendern helfen, Daten leistungsfähig und präzise zu bearbeiten. Sie basieren auf dem Prinzip der Benutzerfreundlichkeit, Schnelligkeit und Agilität als Tools für die visuelle Analyse, die es alltäglichen Geschäftsanwendern ermöglichen, Datenbestände zu kombinieren, Verknüpfungen zu automatisieren, Felder umzubenennen und weitere Verbesserungen an den Daten vorzunehmen, mit dem Ziel, sie für die Analyse aufzubereiten.

Diese sollten etwas weiter hinten auf Ihrer Prioritätenliste stehen, da es oft schwer ist zu wissen, wie Sie die Daten ändern müssen, bevor Sie sie verwenden können. Aber über die gesamte Lebensdauer eines Unternehmens wird ein Großteil der Zeit zur Beantwortung einer Frage darauf verwendet, die Daten in die richtige Form zu bringen. Persönliche Tools zur Datenaufbereitung sind eine leistungsfähige Möglichkeit, diesen Zeitaufwand zu reduzieren, ohne die Notwendigkeit einer von der IT entwickelten, vordefinierten semantischen Schicht mit einzubeziehen oder diese zu negieren.



Gemäß „**What's Your Data Strategy?**“ in der Harvard Business Review „verbringen die Analysten 80 % ihrer Zeit mit der Untersuchung und Aufbereitung der Daten.“

Tableau



View Tableau data set lineage, data preparation steps, annotations and versions

Abbildung 3 Integration der visuellen Analyse mit den Persönlichen Tools zur Datenaufbereitung von Praxata

Erweiterte Analysen

Erweiterte Analysen haben sich zu einem kritischen Bestandteil moderner Analysearchitekturen entwickelt. Unternehmen greifen dabei auf Statistiken, Prognosealgorithmen und maschinelles Lernen zurück, um den Wert sehr großer Datenbestände zu maximieren.

In der Vergangenheit waren erweiterte Analysen nur geschulten Datenwissenschaftlern mit Programmen wie R, Python, SPSS oder SAS zugänglich. Verbesserungen in der visuellen Analytik erweitern heute als integrierte Funktionen die Reichweite der erweiterten Analysen für alle Benutzer von Selfservice-Analysen. Funktionen wie Box-Plots, Baumkarten und grundlegende prädiktive Modellierungstechniken sind nun alltäglich und per Mausklick zugänglich.

Es gibt immer noch viele Anwendungsfälle für dedizierte statistische Analysetools. Unternehmen, die sich ständig weiterentwickelnde Algorithmen entwickeln möchten, um ihren Kunden eine Vorschaufunktion zu bieten oder Funktionen schreiben möchten, um Betrug mit der Kreditkarte eines Kunden zu entdecken, sollten vielleicht dedizierte Tools in Betracht ziehen, um den Kernbereich ihrer Tools zur visuellen Analyse zu ergänzen. Der Einsatz dieser Tools erfordert Schulung und toolspezifische Programmiererfahrung, die sich über Monate erstrecken kann.

Freigabe und Zusammenarbeit

In modernen BI-Plattformen ist das Freigeben, Zusammenarbeiten und Weitergeben von Erkenntnissen eine Kernkompetenz. Von der Bewertung des Kontexts bis hin zur Priorisierung der besten nächsten Maßnahme – die Wirkung von Erkenntnissen wird durch Zusammenarbeit maximiert. Moderne Analyseplattformen bieten Diskussionsforen, Anmerkungen, Kommentare, Favoriten, Likes und andere soziale Konzepte, die an führende Produktivitäts- und Portalanwendungen angepasst sind. Die Möglichkeit, die Erkenntnisse einer Analyseplattform direkt im Tool freizugeben, vereinfacht den Analysefluss erheblich und fördert die weitere Untersuchung und Diskussion wertvoller Erkenntnisse. Durch Einbettung kann die Zusammenarbeit auch auf externe Anwendungen und Portale ausgeweitet werden.

Eingebettete Analysen

Eines der leistungsstärksten und häufig übersehenen Geschäftskonzepte ist der Ablauf. Statt Geschäftsanwender aus ihren Standard-Betriebsprozessen herauszulösen, um Antworten auf Daten zu suchen, fügen Sie diese Erkenntnisse nahtlos in die etablierten Abläufe ein und integrieren sie in bestehende Prozesse.

Bei moderner Analytik finden Sie Daten und Dashboards direkt in Unternehmensportalen, anderen Anwendungen oder integriert in Produktivitätswerkzeuge. Die besten Analyseplattformen unterstützen all diese Szenarien mit ausgereiften APIs, Software Developer Kits und flexiblen Bereitstellungsmechanismen. So können Sie mühelos von einem Tool zum nächsten springen und sogar alle Tools in einem einzigen Portal kombinieren.

Der Ablauf bezieht sich auch auf den physischen Standort. Mitarbeiter sind heute oft unterwegs, ohne Zugang zu internen Ressourcen hinter Firewalls. Moderne Analytik unterstützt den Zugriff auf Daten von überall und auf jedem Gerät. So kann ein Verkäufer eine fundierte, datengesteuerte Entscheidung von seinem mobilen Gerät aus treffen, ohne einen Laptop öffnen zu müssen. Es bedeutet auch, dass ein Bauleiter vor Ort sein und auf kritische Informationen zugreifen kann, ohne per VPN auf das Firmennetzwerk zugreifen zu müssen. Mobilgeräte und Cloud-Computing haben die Art und Weise, wie ein Unternehmen seine Geschäftsabläufe steuern kann, nachhaltig verändert. Moderne Analyseplattformen müssen Unternehmen in die Lage versetzen, diese Vorteile zu nutzen.

3. Alles miteinander verbinden

Diese Bausteine sind die Grundlage für eine moderne Analyseplattform, die Unternehmen befähigt, jede Herausforderung anzunehmen. Zusammen mit einer echten Partnerschaft zwischen Geschäftsbetrieb und IT bietet diese Plattform jedem in Ihrem Unternehmen das Vertrauen, dass er für die erforderlichen Entscheidungen über die notwendigen Tools verfügt und dass seine Entscheidung auf vertrauenswürdigen Daten beruhen.

Eine ganze moderne Analyseplattform zu gestalten, kann an sich schon eine Herausforderung sein. Die gute Nachricht ist, dass nicht erst das gesamte Systemumfeld komplett aufgebaut werden muss, bevor begonnen werden kann. Tatsächlich sind Unternehmen heutzutage am erfolgreichsten, wenn sie es nicht tun. Stattdessen beginnen sie den Wandel mit kleinen Schritten, aus denen sich letztlich ergibt, wo als Nächstes investiert werden muss. Unternehmen können mit Pilotversuchen beginnen, bevor sie Technologielösungen in größerem Umfang bereitstellen. Sie müssen nicht jede einzelne Komponente integrieren, um eine neue Strategie auf den Weg zu bringen. Sie können beispielsweise mit Ihrem Tool für visuelle Analysen Lücken in Ihrer Datenpipeline entdecken, bevor Sie Ihr gesamtes Lager fertigstellen. Dies hilft Ihnen, sofortigen Nutzen aus Analysen zu ziehen, Lücken und Fehler in Ihren Daten zu finden und letztendlich ein präziseres und besser funktionierendes Data Warehouse aufzubauen.

Der Schlüssel liegt darin, mit Tools zu arbeiten, die diese Art von inkrementellen Änderungen ermöglichen. Eine moderne Analyseplattform besteht letztlich aus einer Reihe von Bausteinen, die sich einzeln zusammensetzen lassen. So wird Ihr Unternehmen zugänglicher, agiler und befähigt, aus einer Vielzahl von Datenquellen Erkenntnisse zu gewinnen. Dies ist genau das, was führende Weltklasseanalysten tun: Sie nutzen visuelle Analysen in Kombination mit anderen Spitzenlösungen für Big Data-Analysen, Internet of Things und Datenwissenschaften auf strategische Weise.

So hat zum Beispiel **Netflix eine umfangreiche Big Data-Plattform** und einen Datensee aufgebaut, um die enormen Datenmengen zu unterstützen, die der Betrieb generiert. Tableau ist die wesentliche Komponente, die es ihnen ermöglicht, die unterschiedlichen Tools wie S3, EMR und Spark zu einer zusammenhängenden Analyseplattform zu kombinieren, die das Geschäft stützt.

Unabhängig davon, wo Sie sich auf dem Weg des digitalen Wandels befinden, ist es entscheidend, sofort mit den Daten zu beginnen, die Sie heute haben. Unternehmen müssen angesichts der nächsten großen Marktumwälzung nun noch schneller handeln. Nehmen Sie eine Komponente moderner Analytik und befähigen Sie das Unternehmen, datengesteuerte Entscheidungen zu treffen und der Umwälzer zu sein.

4. Anhang

Es gibt viele verschiedene Analysetechnologien und Lösungsoptionen, jede mit eigenen enggefassten Zielen und Vorteilen. In diesem Anhang schlüsseln wir Optionen auf, die in der Übersicht nicht aufgeführt sind.

Einbeziehung von Streams

Stream-Daten werden kontinuierlich von angeschlossenen Geräten und Apps generiert, die sich überall befinden, wie z. B. Soziale Netzwerke, Intelligente Zähler, Heimautomatisierung, Videospiele und IoT-Sensoren. Häufig werden diese Daten über Pipelines halbstrukturierter Daten gesammelt. Echtzeit-Analysen und Prognosealgorithmen können zwar auf Streams angewendet werden, jedoch werden die Streamdaten für die Analysen üblicherweise in Rohformaten mithilfe von [Lambda-Architektur](#) und in einen Datensee wie Hadoop geleitet und dort gespeichert.

Die Lambda-Architektur ist eine Datenverarbeitungsarchitektur, die entwickelt wurde, um riesige Datenmengen zu handhaben und dabei sowohl Batch- als auch Stream-Verarbeitungsmethoden zu nutzen. Das Design gleicht Latenz, Durchsatz und Fehlertoleranz aus.

Es gibt heute eine Vielzahl von Optionen für das Streaming von Daten, darunter Amazon Kinesis, Storm, Flume, Kafka und Informatica Vibe Data Stream.

Integration der Hub-Orchestrierung

Hub-and-Spoke-Integrationsmuster sind ein leicht verständliches und weitverbreitetes Datenarchitekturdesign. Hubs entkoppeln Datenquellen, die sich an beliebigen Orten befinden, und ermöglichen eine flexiblere Integration, da die Anzahl der zu verwaltenden Punkt-zu-Punkt-Schnittstellen reduziert wird. Integrationsdrehscheiben mit Publish/Subscribe-Funktionen fördern die Wiederverwendung von Daten und bieten eine zentrale Kontrolle für Optimierungs-, Datenstandard- und Governance-Zwecke. Die zentrale Verwaltung sorgt für eine bessere Transparenz über alle orchestrierten Datenbewegungs Pipelines, die sich über alle Datenquellen erstrecken.

Datenintegrationsknoten der neuen Generation erweitern traditionelle Funktionen auf Benutzer von Selfservice-Analysen. Jeder kann mit nur geringer Unterstützung der IT-Abteilung Datenfeeds in modernen Integrations-Hubs veröffentlichen oder sie abonnieren. Datenkonsumenten können zertifizierte Daten nutzen, Einblick in Herkunft und Integrationsprozesse nehmen. Weitere Vorteile der modernen Datenintegrations-Hubs sind nahtlose Datenqualitätsfunktionen, beschleunigte Einbeziehung von Datenquellen und die Just-in-time-Bereitstellung kleiner oder massiver Datenvolumen.

Informatica und Cisco sind Marktführer im Bereich der Hub-Technologie für Datenintegration. Durch die tiefe Integration von Tableau in Informatica können Sie Hunderte verschiedener Datenquellen in Tableau-Datenextrakte integrieren, die auf dem [Tableau-Datenserver](#) gespeichert und auf dem neuesten Stand gehalten werden – für jedermann im Unternehmen.

Weitere Informationen zu unstrukturierten Daten, NoSQL und Datenseen

Datenseen unterstützen moderne Anforderungen für die Big Data-Analyse durch schnellere und flexiblere Datenaufnahme und -speicherung, damit jeder Rohdaten auf vielfältige Weise schnell analysieren kann. Datenseen ersetzen keine Data Warehouses.

Bei modernen Konzepten der Dateneingliederung und -bereitstellung ist das Ziel für Rohdaten jeder Größe und Form oft ein Datensee. Ein Datensee ist ein Speicher-Repository, das eine riesige Datenmenge in nativem Format enthält: strukturiert, halb- und unstrukturiert. Datenseen bieten auch optimierte Verarbeitungsmechanismen über APIs oder SQL-ähnliche Sprachen für die Umwandlung von Rohdaten mit „Schema on Read“-Funktion.

Obwohl Hadoop seit dem ersten Hadoop Distributed File System (HDFS) aufgrund seiner Robustheit und niedrigen Kosten für Datenseen verwendet wird, ist es nicht die einzige Datensee-Implementierungsoption. Als Datenseen können auch Objektspeicher wie Simple Storage Service (S3) von Amazon Web Services und NoSQL-Datenbanken mit flexiblen Schemata verwendet werden. Tableau **unterstützt jetzt den Athena**-Datendienst von Amazon zur Verbindung mit Amazon S3 und verfügt über verschiedene Tools, die eine direkte Verbindung zu NoSQL-Datenbanken ermöglichen.

In modernen Analysearchitekturen werden NoSQL-Datenbanken zur Normalität, weil die Daten von überall schneller geladen werden und aufgrund der schemalosen Datenbankkonzepte. NoSQL-, Nicht-SQL- oder Nicht-nur-SQL-Datenbanken bieten alternative Arten der Datenspeicherung. Übliche NoSQL-Speichertypen sind Spalten, Dokumente, Schlüsselwerte und Diagramme.

Beispiele für NoSQL-Datenbanken, die häufig mit Tableau verwendet werden, sind u. a. MongoDB, Datastax und MarkLogic.

Hadoop wird oft als Big Data-Plattform genutzt, ist aber keine Datenbank. Hadoop ist ein Open-Source Software-Framework für die Speicherung von Daten und das Ausführen von Anwendungen auf Clustern von Standardhardware. Es bietet massiven Speicherplatz für jede Art von Daten, enorme Rechenleistung und die Fähigkeit, extreme Mengen an gleichzeitigen Aufgaben oder Aufträgen zu bewältigen.

In einer modernen Analysearchitektur bietet Hadoop kostengünstige Speicher- und Datenarchivierungslösungen für die Entladung alter historischer Daten aus dem Data Warehouse in Online-Kühlhäuser. Es wird auch bei Analysen in den Bereichen IoT, Datenwissenschaft und für unstrukturierte Analysen eingesetzt.

Innerhalb des Hadoop-Frameworks umfassen verwandte Technologien zum Laden, Organisieren und Abfragen von Daten unter anderem Folgendes:

- **Apache Spark** – Open-Source Cluster Computing Framework mit hochperformanter In-Memory-Analyse und einer wachsenden Anzahl dazugehöriger Projekte.
- **Apache Impala** – die analytische Open-Source-MPP-Datenbank für Apache Hadoop. Dies ist die Datenverbindung, die am häufigsten in erfolgreichen Hadoop-Projekten mit Tableau verwendet wird.
- **Apache Presto** – eine verteilte Open-Source-SQL-Abfrage-Engine für die Ausführung von interaktiven Abfragen in Datenbeständen jeglicher Größe. **Tableau unterstützt Presto seit der Version 10**
- **MapReduce** – ein Software-Framework für Parallelverarbeitung, das Inputs aufnimmt, in kleinere Probleme partitioniert und an Worker-Knoten verteilt.
- **Hive** – eine SQL-ähnliche Abfragesprache für Data Warehousing. Hive 2.0 enthält auch LLAP (Live Long and Process), das die Performance der Hive-Abfrage deutlich verbessert.
- **Hadoop Distributed File System (HDFS)** – das skalierbare System, das Daten ohne vorherige Organisation auf mehreren Rechnern speichert
- **YARN** – (Yet Another Resource Negotiator) übernimmt das Ressourcenmanagement für die Prozesse auf Hadoop
- **Ambari** – ein Web-Interface zur Verwaltung von Hadoop-Diensten und -Komponenten
- **Cassandra** – ein verteiltes Datenbanksystem
- **Flume** – Software zum Streamen von Daten in HDFS
- **HBase** – eine nicht-relationale, verteilte Datenbank, die auf Hadoop läuft.
- **HCatalog** – eine Tabellen- und Speicherverwaltungsebene

- **Oozie** – ein Planer für Hadoop-Aufträge
- **Pig** – eine Plattform für die Handhabung von in HDFS gespeicherten Daten
- **Solr** – ein skalierbares Suchwerkzeug
- **Sqoop** – verschiebt Daten zwischen Hadoop und relationalen Datenbanken
- **Zookeeper** – eine Anwendung, die verteilte Verarbeitung koordiniert

In den letzten beiden Jahren hat sich Apache Spark von einem Bestandteil des Hadoop-Systemumfelds zu einer eigenständigen Big Data-Analyseplattform erster Wahl für eine Reihe von Unternehmen entwickelt. Spark bietet eine drastisch erhöhte Datenverarbeitungsgeschwindigkeit im Vergleich zu Hadoop. Spark selbst hat viele verwandte Projekte wie die Apache Spark Runtime, Spark SQL, Spark Streaming, MLlib, ML und GraphX. Es ist jetzt das größte Open-Source-Projekt für Big Data mit mehr als 1000 Mitarbeitern aus über 250 Organisationen.

Tableau ist einer der Marktführer im Bereich Big Data speziell für Analyseanbindung und visuelle Datenanalyse. Klassenbeste Analyseprogramme für Big Data verwenden Tableau mit Cloudera, Spark SQL, Amazonas EMR, Hortonworks, Microsoft HDInsight/Data Lake und MapR. Viele andere Big Data-Technologien können über diese offen unterstützten Technologien oder deren Treiber an Tableau angebunden werden.

Daten als Dienstleistung

In einer digitalen Welt, in der Daten Gold sind, sind Daten auch ein Produkt, das jeder konsumieren kann. Kunden-, Finanz-, Finanz-, Markt-, Wetter-, geografische und demografische Daten werden bereits heute als Dienstleistung für den Einkauf in Datenmärkten und Handelsplattformen angeboten.

Data-as-a-Service wendet ein flexibles SOA-Muster (Service Oriented Architecture) für die Datenbereitstellung über die Cloud an. Dieser Ansatz bietet extreme Agilität, da die SOA-Architektur sehr einfach ist. Heute bieten ISVs, CRMs und ERPs standardmäßige REST-APIs für Data-as-a-Service zum Integrieren oder für externe Berichterstellung an.

Mit dem **Webdaten-Connector SDK von Tableau** können Verbindungen zu Daten aufgebaut werden, die sich außerhalb der bestehenden Konnektoren befinden. Benutzer von Selfservice-Analysen können Verbindungen zu fast allen über HTTP zugänglichen Daten herstellen, auch über interne Webservices, JSON-Daten und REST-APIs.

Logisches Data Warehouse

Führende Analytik-Organisationen liefern flexible, logische und einheitliche Dimensionsansichten von Daten über Datenvirtualisierungstechnologien von Anbietern wie Cisco und Denodo. Für Anwender von Analysefunktionen sieht ein logisches Data Warehouse aus und verhält sich wie ein relationales Data Warehouse. Tableau-Benutzer können über standardmäßige ODBC-Treiber eine Verbindung zu diesen Tools herstellen.

Eine der Kernkompetenzen der Datenvirtualisierung ist die Optimierung dezentral verteilter, heterogener Abfragen in einer Vielzahl von verschiedenen Datenquellen und REST-APIs. Das logische Data Warehouse dient auch als semantische Schicht zur Pufferung von Reporting-Anwendungen gegenüber Datenquellenänderungen. Logische Data Warehouses werden häufig mit einem Unternehmensdatenkatalog verwendet.

Stammdatenverwaltung

Analytik ist nur so gut wie die Qualität der verwendeten Daten und Menschen können nur dann die richtigen Entscheidungen treffen, wenn ihre Daten korrekt sind. Da immer mehr Fachexperten Daten in einem Bottom-up-Ansatz erstellen, lässt sich ein wieder wachsendes Interesse an traditioneller Datenqualitäts- und Stammdatenverwaltung erkennen, die wiederum von entscheidender Bedeutung sind, um sicherzustellen, dass die Berichtsdatenquellen aktuell, bereinigt, konsistent und genau sind.

Zu den beliebtesten Angeboten für die Stammdatenverwaltung gehören unter anderem Informatica, IBM und Stibo. Einige der am häufigsten mit Tableau verwendeten Datenqualitätslösungen sind Trillium, Informatica Data Quality, Talend Data Quality 6.0 und Tamr Eisenhower.

Unternehmensdatenkatalog

Eine weitere neue Technologie ist der Unternehmensdatenkatalog. Mithilfe der Datenkataloge von Unternehmen können Anwender selbstständiger Berichte die richtigen Daten für die Entscheidungsfindung aus genehmigten Datenquellen finden. Unternehmensdatenkataloge gibt es innerhalb von Visual Analytics-Lösungen und sie sind auch als eigenständige Lösungen für die nahtlose Integration mit Tableau erhältlich.

Unternehmensdatenkataloge werden mit Metadaten aus Tabellen, Ansichten und gespeicherten Prozeduren gefüllt, indem einbezogene Datenquellen gescannt werden. Mit der automatisierten Erkennung neuer Datenquellen, intelligenter Datenklassifizierung und datenquellenübergreifendem Entity-Mapping dient ein Datenkatalog im Wesentlichen als unternehmensweites Glossar mit Datenquellen und gemeinsamen Datendefinitionen. Fachexperten erweitern den Kontext von Katalogdatenquellen durch Anmerkungen, Versionen und Dokumentationen.

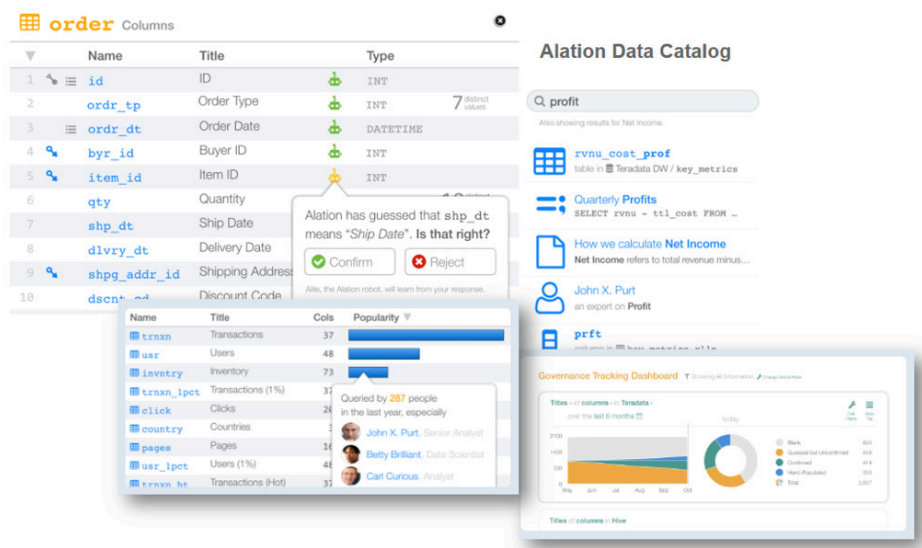


Abbildung 4 Ein Unternehmensdatenkatalog von Alation

Datenkataloglösungen fördern die Datenpflege und effiziente Wiederverwendung vorhandener Daten. Darüber hinaus bieten sie eine sehr wünschenswerte Datenabstammung und eine zusätzliche Ebene für Datenverwaltung, Sicherheit, Protokollierung und Prüfung.

Zu den Anbietern mit umfangreichen Datenkatalogen, die sich sehr gut in Tableau integrieren lassen, gehören Alation, Colibra, Attivio, Informatica und Waterline.

Maschinelles Lernen

Kognitives Lernen, tiefes Lernen und künstliche Intelligenz ziehen Rückschlüsse aus vorhandenen Daten und Mustern, ziehen Schlussfolgerungen auf der Grundlage vorhandener Wissensdatenbanken und fügen diese dann in fortlaufenden, kontinuierlichen, selbstlernenden Kreisläufen in eine Wissensbasis ein.

Der Verbrauch dieser Art der Analysen bedeutet in der Regel, dass die Ausgabe über eine eingebettete API in einem Bericht oder eine integrierte Anwendung überprüft wird. Tableau wird heute verwendet, um die Ausgabe von CognitiveCode, Digital Reasoning und anderen Anbietern zu visualisieren.

In Tableau 10.3 haben wir [die von uns empfohlenen Tabellen und smarten Verknüpfungen vorgestellt](#), um Ihnen Zeit beim Herstellen der Verbindung und bei der Vorbereitung Ihrer Daten zu sparen. Mit Hilfe von maschinellem Lernen verbessern sich die Empfehlungen im Laufe der Zeit, da Datenbanken häufiger genutzt werden.

Natürliche Sprache

Natürliche Sprache und Sprachabfrage beginnen, das Erzählen von Datenstorys weiterzuentwickeln, sodass die Datenuntersuchung auf flexible Weise weiterentwickelt wird. Natürliche Sprache sorgt dafür, dass Analysen auf jeder Plattform leichter zugänglich werden. Es können automatisierte kontextbezogene Beschreibungen der wichtigsten Ergebnisse abgerufen, Prognosen angefordert oder Mengen von Textdokumenten analysiert werden.

Heute kann die visuelle Analyse von Tableau mit führenden Lösungen für die Generierung natürlicher Sprache (NLG) kombiniert werden, etwa mit Yseop, Narrative Science und Automated Insight. Da diese Technologien den Kontext von Tableau-Visualisierungen weitgehend interpretieren, werden die Integrationen am häufigsten im Natural Language Tool selbst oder als Erweiterung durch JavaScript durchgeführt. Darüber hinaus wird die [Akquisition von ClearGraph](#) die intelligente Datenuntersuchung und -analyse direkt in Tableau ermöglichen und die Interaktion mit Daten durch natürliche Sprache erleichtern.

Empfohlene Datenuntersuchung

Intelligente Datenuntersuchung nutzt Algorithmen des maschinellen Lernens, um viel tiefere analytische („was wird geschehen“) und präskriptive („auf welche Weise optimieren“) Fähigkeiten mit zu diesem Zweck formatierten Daten bereitzustellen. Mit der Weiterentwicklung der visuellen Analyse werden neue automatisierte Erkenntnisse und Empfehlungen hinzugefügt. Diese Fähigkeiten wurden in der Roadmap-Keynote auf der [Tableau Conference 2016 gezeigt](#). Durch die Integration von R, Python, APIs und analytischen Datenbankfunktionen können die Ergebnisse der erweiterten Analysefunktionen in Tableau visualisiert und erforscht werden.

Suche

Mit modernen Analysearchitekturen können Nutzer und Experten die Daten mit Google-ähnlicher Leichtigkeit durchsuchen. Anstatt Daten zu modellieren, ordnet die analytische Indizierungstechnologie für Suchmaschinen automatisch verschiedene Datenquellen basierend auf Feldnamen, Datentypen und maschineller Lernintelligenz zu. Im Laufe der Zeit werden dynamische Suchvorschläge basierend auf historischen Abfragen und gemeldeter Nutzung generiert. Seit dem Aufkommen der Sprachtechnologien wie Siri und Alexa tauchen Sprachabfragefähigkeiten auf, die mit analytischer Suche kombiniert werden. Ein großartiges Beispiel dafür war der [Gewinner des jährlichen Hackathon](#) auf der Tableau Conference 2016 von Automated Insight, der Alexa Voice Control zu Tableau brachte.

Warnungen

Eine moderne Analysearchitektur umfasst konfigurierbare, intelligente, datengesteuerte Alarmmeldungen, die kontinuierlich ein wertvolles Signal im digitalen Datenmeer überwachen. Es ist für einen Menschen unmöglich, jeden wichtigen Wert rund um die Uhr zu kontrollieren. Hier sind Automatisierung und Alarmierung sehr hilfreich im modernen Analyse-Arsenal.

Einige Tools bieten Momentaufnahmen in regelmäßigen Abständen, andere führen Protokolle, um zu sehen, ob Zahlen bestimmte Schwellenwerte überschritten haben. Es gibt Gründe für beide. Einige Dashboards sind informativ und Sie möchten sie einfach täglich überprüfen. Andere wiederum sind die Basis für kritische Maßnahmen, aber das tägliche Überprüfen eines Dashboards ohne praxisrelevante Erkenntnisse ist keine effiziente Nutzung der Zeit.

Mit Tableau können Sie mit [datengesteuerten Warnmeldungen für Tableau Server](#) auf dem Laufenden bleiben. Wählen Sie einfach einen Schwellenwert, damit Sie selbst oder Ihr gesamtes Team eine E-Mail-Benachrichtigung erhalten.

Storytelling

Manchmal ist die Erkenntnis – bzw. das „Was“ – nicht ausreichend. Die Benutzer möchten das „Warum“ hinter den Daten verstehen. Warum ist der Umsatz gestiegen? Was hat den Ausreißer beim Website-Traffic verursacht? Warum ist es so schwierig, den medizinischen Nachschub zu lagern?

Die Unternehmen versuchen seit langem, dieses Problem zu lösen, indem sie Analytik mit anderen Kommunikationsformen kombinieren: Text, Bilder und sogar Videos. Analysten erstellen Präsentationen mit Powerpoint, schreiben lange Berichte als PDFs oder noch umständlicher, drucken Seiten und Seiten von Dokumenten aus und fügen sie in einem Ordner zusammen.

Moderne Analyse-Tools von heute nutzen das Beste aus diesen Storytelling-Konzepten und integrieren sie als erstklassige Features. Damit können Sie interaktive Dashboards erstellen, spezifische Schnappschüsse von Daten senden, die automatisch im Hintergrund aktualisiert werden, nachdem neue Daten hinzugefügt wurden, oder sogar Berichte erstellen, die interaktive Diagramme mit Text und Bildern kombinieren. Storytelling ermöglicht es den Benutzern, die Datenanalyse zu erklären, anstatt nur eine Zahl anzugeben.

Tableau legt großen Wert auf Auswahlmöglichkeiten und offene Standards. Wir investieren viel in Forschung und Entwicklung, um die Analysen zu beschleunigen und zu vereinfachen. Das bedeutet auch, dass wir mit unserem Systemumfeld von Partnern innovieren. So wird sichergestellt, dass die Welt der Analysen sich weiterentwickelt und neue Technologien ins Spiel kommen. Analytik-Führungskräfte werden immer in der Lage sein, Tableau in ihre Wahl aktueller und zukünftiger Datentechnologien zu integrieren.

Über Tableau

Tableau unterstützt Ihre Mitarbeiter dabei, ihre Daten zu sehen und zu verstehen – unabhängig davon, wie umfangreich sie sind, aus welchem Kanal sie stammen oder in welcher Datenbank sie gespeichert sind. Das nahtlose Zusammenspiel von PC und iPad ermöglicht dabei das schnelle Verknüpfen, Zusammenführen und Visualisieren von Dashboards. Erstellen und veröffentlichen Sie Marketingdashboards mit automatisierten Datenupdates und teilen Sie Ihre Erkenntnisse in Echtzeit mit Kollegen, Teams, dem Management, Partnern oder Kunden – ganz ohne Programmierkenntnisse. [Gleich heute kostenlos testen!](#)

Weitere Materialien

[Wie eine Kultur für Selfservice-Analysen aufgebaut wird](#)

[Analytics neu definiert](#)

[Analytikansatz](#)

[Der Bericht zu Cloud-Daten](#)

[Fortgeschrittene Analytik mit Tableau \(Advanced Analytics with Tableau\)](#)

