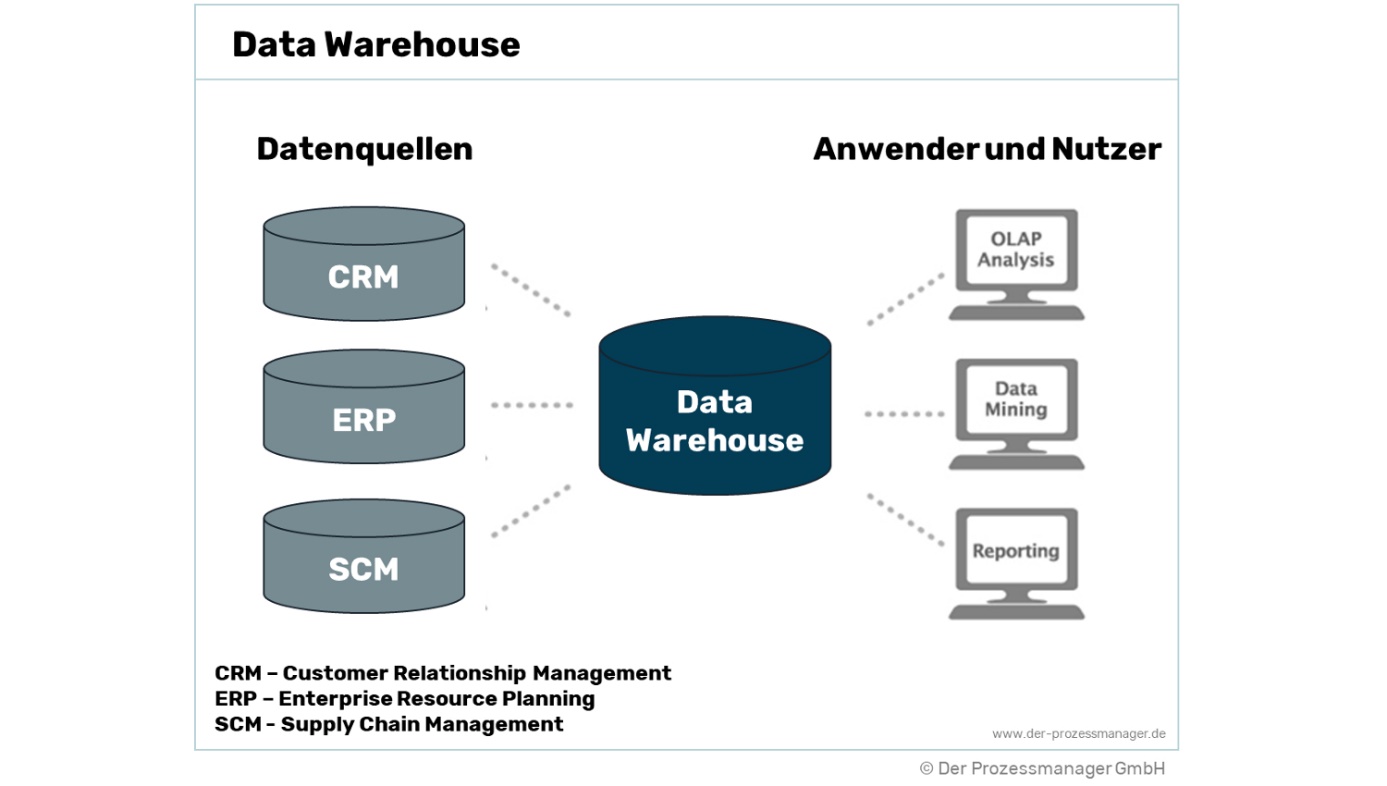
**Was ist ein Data Warehouse (DWH, Unterschied zu einem ERP-System, transaktionsorientiertem System)**

Ein Data Warehouse ist ein "unternehmensweites Konzept, das als logisch zentraler Speicher eine einheitliche und konsistente Datenbasis für die vielfältigen Anwendungen zur Managementunterstützung bietet und losgelöst von den operativen Datenbanken betrieben wird". Es sammelt als zentrale Datenbank in regelmäßigen Intervallen Einträge aus den operativen Systemen und aus externen Quellen, konsolidiert sie, filtert Unwichtiges heraus, ordnet und verdichtet sie, versieht sie mit beschreibenden Zusatzinformationen, so genannten Metadaten, und zeigt mit Hilfe von Analysewerkzeugen Relevantes an.

Data Warehousing ist ein evolutionärer Prozess: Ein Data Warehouse kann man nicht kaufen. Man muss es bauen. Es gibt deshalb kein einfaches Standarddesign für ein Data Warehouse, sondern Größe und Aussehen eines Data Warehouse variieren gemäß der Unternehmensstruktur und den Anforderungen der Benutzer. Konzeption und Erstellung können somit zu einem langwierigen und teuren Prozess auswachsen.

Bei der Implementierung einer Data-Warehouse-Lösung ist zu beachten, dass die für die Unternehmensführung notwendigen Informationen aus den verschiedenen betrieblichen Bereichen (z.B. Produktion, Vertrieb) oder organisationsexternen Quellen wie z.B. Nachrichten- oder Online-Diensten stammen. Entsprechend dem Data-Warehouse-Grundgedanken sind diese Informationen zunächst zusammenzufügen und umzuformen, bevor sie transformiert und schließlich im "Datenlager" gesammelt und aufbereitet werden.

Vorteil des Datenlagers ist, dass eine globale Sicht auf Daten aus unterschiedlichen Datenbeständen entsteht. Gleichzeitig vereinfacht sich der Zugriff auf die Daten für Anwender, da sie in einer zentralen Datenbank konsistent und strukturiert bereitgestellt sind.



Das Data Warehousing ist in vier Teilprozesse aufteilbar:

* **Datenbeschaffung**: Beschaffung und Extraktion der Daten aus verschiedenen Datenbeständen
* **Datenhaltung**: Speicherung der Daten im Datenlager inklusive Langzeitarchivierung
* **Datenversorgung**: Versorgung der nachgelagerten Systeme mit den benötigten Daten, Bereitstellung von Data Marts
* **Datenauswertung**: Analysen und Auswertungen der Datenbestände

|  |  |
| --- | --- |
| ERP | Data Warehouse |
| Entwickelt um Daten zu speichern | Entwickelt um Daten zu analysieren |
| Speichert detailierte Daten | Speichert zusammengefasste Daten |
| Benutzt Online Transkationsverarbeitung OLTP | Benutz Online Analyseverarbeitung OLAP |
| führt grundlegende Geschäftsvorgänge und Transaktionen durch | Ermöglicht dem Benutzer die Analyse von Geschäftsdaten |
| Daten sind in Echtzeit verfügbar | Daten müssen bei Bedarf aktualisiert werden |
| Anwendungsorientierte Datenerfassung | Subjektorientierte Datenerfassung |
| Limitiert auf eine Anwendung | Zieht Daten aus einer Reihe anderer Anwendungen |

**Organisatorische Rahmenbedingungen zur erfolgreichen Einführung und Betrieb DWH**

Ein Data Warehouse Analyst analysiert und verwaltet alle relevanten Daten des jeweiligen Unternehmens. Hier besteht die wichtige Aufgabe darin die Daten so zu bereinigen, aufzuarbeiten und einzupflegen, dass jeder Mitarbeiter des Unternehmens Zugriff darauf hat und dass zu möglichst jeder Zeit.

Bei einem Data Warehouse Analyst handelt es sich um ein Bindeglied zwischen Management und IT. Somit erfordert seine Position auch eine gute Zusammenarbeit mit dem Team.

Unter anderem arbeitet er dabei mit Kundendaten, die wichtige Aufschlüsse über die aktuelle Marktlage aber auch über die zukünftige Marktentwicklung geben.

Zum weiteren Aufgabengebiet gehört, dass die gesammelten Daten über eine Business-Intelligence Software nach ihrem Nutzen und ihrer Wichtigkeit eingeordnet werden. Das hat wiederum zur Folge, dass Fachwissen über verschiedene Datenbanksysteme wie SAS oder Oracle vorhanden sein muss.

Zum weiteren Aufgabengebiet gehört das Bewerten und Erarbeiten von Strategien und Konzepten zur Informationsnutzung sowie die Unterstützung des Unternehmens bei der Qualitätssicherung. Das macht wiederum deutlich, dass ein Data Warehouse Analyst ein großes und vielfältiges Fachwissen in vielen Bereichen aufweisen muss.

Bei der Einführung von DWH-Systemen können etliche Probleme auftauchen und Erfolge beeinträchtigen. Viele lassen sich aber durch eine professionelle Aufklärung und Planungsphase verhindern. Folgende Punkte sollten beachtet werden:

**Realistische Planung**

* das Aufstellen eines realistischen Zeit- und Budgetplanes
* die genaue Bestimmung aller benötigter Ressourcen für die Einführung
* die Berücksichtigung von wartungsbedingten Kosten und Zeitaufwand

Werden diese Punkte nicht beachtet, kann es zu einem beträchtlichen finanziellen Schaden kommen.

**Korrekte Datengrundlage**

Die richtige Auswahl der Datenquellen, um ein DWH zu versorgen, ist entscheidend zur Bildung einer vernünftigen Entscheidungsgrundlage. Hierbei gilt es, nur qualitativ hochwertige Daten zu verwenden und eine stetige Zuführung aktueller Daten zu gewährleisten. Das Verknüpfen falscher Datenquellen oder die Benutzung veralteter Daten kann zu fatalen Fehlentscheidungen in den Unterstützungsprozessen führen.

**Richtige Interpretation**

Selbst bei den besten DWH-Systemen kann es zu Fehlinterpretationen der Daten kommen. In Folge dessen kann es passieren, dass Kunden mit für sie irrelevanten Informationen überflutet werden. In diesem Fall ist es unumgänglich, den Fehler so schnell wie möglich zu erkennen und die Daten manuell anzupassen, um weiteren Schaden zu vermeiden.

**Fundierte Einführung der Benutzer**

Die richtige Schulung der Mitarbeiter ist bei der Einführung neuer Systeme eine Grundvorrausetzung, um diese effizient nutzen zu können. Fehlende Kenntnisse im Umgang mit der Software lassen schnell Frust aufkommen und führen zu einer verminderten Produktivität und falschen Entscheidungen.

**Technische Rahmenbedingungen zur erfolgreichen Einführung und Betrieb DWH**

**Test und die Dokumentation zur Planung von Data-Warehouse-Test bestimmen**

Da sich Data-Warehouse-Tests von den meisten anderen Softwaretests unterscheiden, empfiehlt es sich, den Test- und Validierungsprozess in mehrere klar definierte, übergeordnete Bereiche für Data-WarehouseProjekte zu unterteilen. Dies ermöglicht eine gezielte Planung für jeden einzelnen Bereich, z. B. für die Integration und Datenvalidierung.

**Den exakten Fokus von ETL-Tests bestimmen**

Gut geplante Extraktions-, Transformations- und Lasttests sollten eine hohe Priorität haben und entsprechend fokussiert sein. Der Abgleich von Warehouse-Daten mit den Daten des Quellsystems (Daten, die in das Data Warehouse eingespeist werden) ist von entscheidender Bedeutung, damit die Anwender sicher sind, dass alle gemeldeten Informationen den Geschäftsanforderungen und Datenzuordnungsregeln entsprechen.

**Auswahl qualifizierter Tester für die Qualitätssicherung im Data Warehouse**

Gerade in der finalen Lieferphase besteht oft starker Druck, Kosten zu senken. Ein häufiger Fehler besteht darin, die Testverantwortung an Personal mit begrenzter Erfahrung in der Geschäfts- und Datenprüfung zu delegieren.

**Identifizieren und Korrigieren von Qualitätsmängeln der Quelldaten**

Das erstmalige Laden eines Data Warehouse ist oft sehr zeitaufwändig. Es ist wichtig, sicherzustellen, dass Sie mit der Qualität der Daten, die in das Warehouse geladen werden, zufrieden sind, bevor Sie mehrere Tage, in einigen Fällen sogar Wochen, für ETL-Prozesse aufwenden.

**Sichern der Unterstützung der Projektbeteiligten**

Mit der Dokumentenflut ist es ähnlich wie mit der Polizei. Jeder meidet sie, bis sie benötigt wird – wenn er in allen Phasen des Daten- und Berichtstests eine Freigabe erreichen muss. Ein Stakeholder ist eine Person, deren Interessen direkt oder indirekt durch das Data-Warehouse-Projekt tangiert werden. Aus Projektsicht können die Stakeholder nach verschiedenen Verantwortlichkeiten eingeteilt werden:

* **Rechenschaftspflichtig:** Einige Stakeholder sind für den Erfolg des gesamten Projekts oder einer bestimmten Phase des Projekts verantwortlich. In der Regel sind dies die Abteilungs- oder Bereichsleiter.
* **Verantwortlich:** Andere Interessengruppen sind für die zu erbringenden Leistungen verantwortlich. Im Allgemeinen wurden sie vom Management damit beauftragt, ein endgültiges Projektergebnis vorzulegen.
* **Beratend:** Einige Stakeholder werden als Berater eingesetzt, beispielsweise Fachexperten (Subject Matter Expert, SME).
* **Informationsberechtigt:** Viele Stakeholder müssen auf dem Laufenden gehalten werden. In der Regel umfasst dies auch das Management, das wissen muss, wofür das Investitionsbudget verwendet wurde, oder auch externe Beteiligte.