



b.telligent

München • Düsseldorf • Frankfurt • Hamburg • Stuttgart • Zürich

# DWH- Automatisierung

Zeit und Geld sparen  
sowie Risiken reduzieren

## Mit Data-Warehouse-Automatisierung Zeit und Geld sparen sowie Risiken reduzieren

**Data-Warehouse-Automatisierung ist eine Kombination von neuen Data-Warehousing-Tools und Methodiken, welche die Lösungen zur Effizienz- und Effektivitätssteigerung in Data-Warehouse-Prozessen bieten. Es handelt sich allerdings um mehr als die reine Automatisierung des Designs und der Erstellung des Data Warehouses. Vielmehr umfasst sie den gesamten Lebenszyklus von Planung, Analyse und Design über Entwicklung bis hin zu Betrieb, Wartung, Änderungsmanagement und Dokumentation.**

Durch die Anwendung der Data-Warehouse-Automatisierung ändert sich die Denk- und Vorgehensweise in Data-Warehouse-Projekten grundlegend. Die weit verbreitete Wasserfall-Methodik der umfangreichen Analyse mit anschließendem vollständigem Design und Modellierung wird der Vergangenheit angehören. Es ist nicht mehr Ziel, in einer einzigen großen und langen Entwicklungsphase das komplette Data Warehouse direkt vollständig richtig zu erstellen, sondern einzelne Funktionalitäten und Bereiche in sehr vielen, sehr kurzen, sich sehr oft wiederholenden Entwicklungsphasen bereitzustellen. Das Data Warehouse entsteht somit iterativ. Dieser Ansatz spiegelt moderne agile Softwareentwicklungsprozesse wider und ermöglicht es somit, Anforderungen im Projektverlauf ändern zu können.

Durch die Automatisierung lässt sich im Data-Warehouse-Lebenszyklus viel Zeit sparen, denn einen Code auf Knopfdruck zu generieren ist in jedem Fall schneller, als der beste Entwickler ihn programmieren kann. Die gesparte Zeit kann in den Test investiert werden. Da ein generierter Code außerdem deutlich weniger fehleranfällig ist, wird hier weitere Zeit eingespart und zusätzlich das Fehlerrisiko erheblich gesenkt. Beides gemeinsam führt zu kürzeren Projektlaufzeiten und somit zu wesentlich geringeren Kosten.

### Wie sind die steigenden Anforderungen an modernes und effizientes Data Warehousing in Zeiten von Big Data erfüllbar?

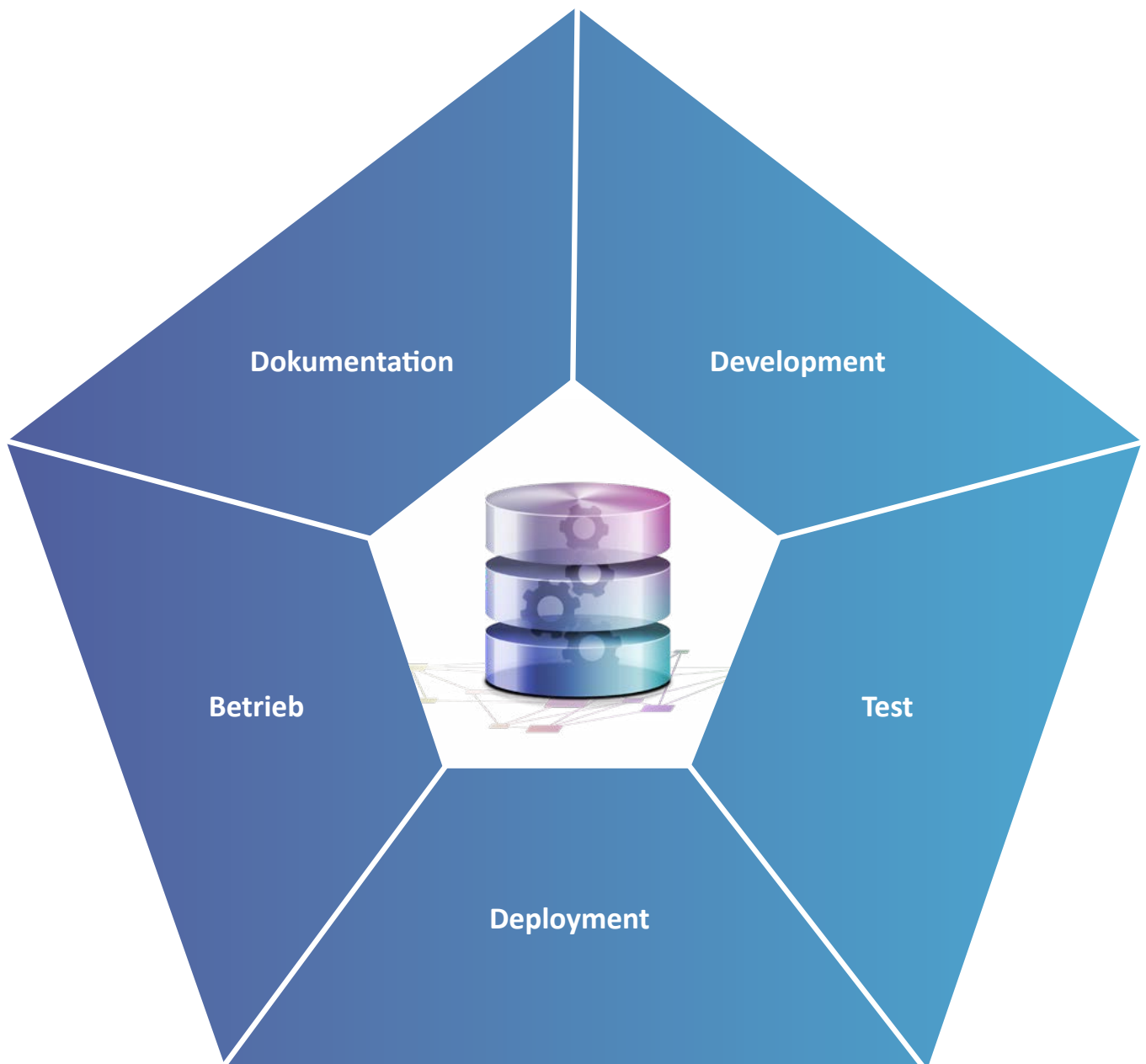
Auch in Zeiten von Big Data, Data Lakes und Datenmanagementplattformen ist das Data Warehouse heute weiter eine unverzichtbare Datenquelle für konsolidierte, übergreifende, qualitätsgesicherte Daten und Voraussetzung für tägliche Analysen, Auswertungen und die operative sowie strategi-

sche Steuerung von Prozessen bis hin zum gesamten Unternehmen. Aufgrund der heutigen sich immer schneller drehenden Welt und der kontinuierlich wachsenden Anforderungen auf regulatorischer Ebene steigen die Anforderungen an das Data Warehouse permanent an. Gleichzeitig wächst durch die fortschreitende digitale Transformation die Menge an Datenquellen und Daten überproportional und dementsprechend steigen auch das Volumen und die Komplexität der Daten, die in das Data Warehouse integriert werden müssen.

Aufgrund dessen befindet sich das moderne Data Warehouse in einer bisher nie da gewesenen Situation und ist den neuen Anforderungen und Änderungen rigoros ausgesetzt. Data-Warehouse-Teams stehen somit mehr denn je vor der Herausforderung, diesem Veränderungsdruck mit den bisher bewährten Methoden, Werkzeugen, knappen Ressourcen sowie den hohen Qualitätsanforderungen standzuhalten. Dies führt oft dazu, dass zeitaufwändige, aber essentielle Themen wie die fachliche Interpretation der Daten, fachliches Testing, Dokumentation oder auch die Qualitätssicherung der Prozesse vernachlässigt werden. Abstriche in den oben genannten Bereichen führen jedoch unweigerlich zu erhöhtem Aufwand bei zukünftigen Entwicklungen, komplexeren Deployments sowie mehr Aufwand im Betrieb und tragen damit zu einer weiteren Verlangsamung der Weiterentwicklung bei.

Ein wirksamer Ansatz, dieses Dilemma nachhaltig zu lösen, besteht in dem kombinierten Einsatz von agiler Entwicklungsmethodik gepaart mit Data-Warehouse-Automatisierungsmethoden und -Technologien, welche den vollständigen Data-Warehouse-Lebenszyklus von der Konzeption, Entwicklung über Test und Deployment bis hin zu Dokumentation und Betrieb optimieren. Wie das genau funktioniert und welchen Mehrwert und welche Vorteile die Data-Warehouse-Automatisierung mit sich bringt, soll im Folgenden detailliert dargestellt werden.

## Die 5 Facetten der Data-Warehouse-Automatisierung



Data-Warehouse-Automatisierung basiert – wie jede Art von Automatisierung – auf der konsequenten Nutzung von Metadaten. In den Metadaten, die in einem zentralen Repository gespeichert werden, werden alle Informationen über die Strukturen des Data Warehouses und der Datenquellen gespeichert sowie Ladestrecken und Transformationen definiert. Aus diesen Metadaten können dann sowohl SQL-Sta-

tements als auch Prozeduren automatisch generiert werden. Große Teile an bisher manueller Programmierung werden so durch den aus den Metadaten generierten Code abgelöst. Damit sind Änderungen nicht nur wesentlich schneller implementiert, sondern auch deutlich robuster. Der generierte Code folgt immer den gleichen einheitlichen Standards, ist dokumentiert und damit leicht lesbar.

Der Fokus der Entwicklung kann sich durch die Ausnutzung der Automatisierungsmethodik auf die Umsetzung und das Testen der fachlichen Logik konzentrieren. Auf diese Weise sind ausreichend Ressourcen für das äußerst wichtige Durchführen von fachlichen Tests verfügbar. Hierzu werden die definierten Testfälle mit Ausgangssituation und Soll-Ergebnis in einer Testdatenbank hinterlegt, woraufhin diese dann zu jeder Zeit automatisch durchlaufen werden können. In der Praxis hat sich das Prinzip des Test-Driven Development bewährt, bei dem aus einer Anforderung zunächst konkrete Testfälle entwickelt werden und erst danach die eigentliche Umsetzung erfolgt. Veränderungen und Weiterentwicklungen können auf diese Weise mit sehr geringem Testaufwand in kurzer Zeit getestet und validiert werden.

Die Metadaten dienen darüber hinaus aber auch zur Erstellung der optimalen Ladereihenfolge („Scheduling“). Da die Metadaten die Abhängigkeiten der Objekte und Ladestrecken enthalten, kann ein Vorschlag für eine optimale Ladereihenfolge automatisiert – inklusive Logging und Monitoring – erstellt werden.

Im Anschluss an das Scheduling erfolgt ein ebenso automatisiertes Deployment, das wesentlich einfacher als in einem nach klassischem Ansatz entwickelten Data Warehouse ist. Grund hierfür ist, dass in einem klassischen Data Warehouse alle geänderten Objekte in einer bestimmten Reihenfolge und unter Beachtung des aktuellen Versionsstandes der Zielumgebung eingespielt werden müssen. Außerdem muss genau darauf geachtet werden, dass der Versionsstand von Funktionen und Prozeduren mit dem Versionsstand der DB-Struktur kompatibel ist. In einem automatisierten Data Warehouse müssen dagegen lediglich die Metadaten in der gewünschten Version in die Zielumgebung kopiert werden. Alle weiteren benötigten Statements sowie sonstige Änderungen werden anschließend direkt in der Zielumgebung generiert. Durch dieses sehr einfache und extrem schnelle Deployment werden agile Vorgehensweisen mit kurzen Sprints überhaupt erst möglich.

Die Dokumentation wird allzu oft in Data-Warehouse-Projekten aufgrund des Änderungs- und Zeitdruckes vernachlässigt. Aus Erfahrung ist der Großteil der im Betrieb befindlichen

Data-Warehouse-Systeme bereits nach dem ersten Release nicht oder nur zum Teil dokumentiert. Wird die Dokumentation nur einmal vernachlässigt, ist es sehr aufwändig bzw. teilweise auch unmöglich, die durchgeführten Änderungen am System im Nachhinein zu dokumentieren.

Ein automatisiertes Data Warehouse dokumentiert sich hingegen (fast) von selbst. Aus den Metadaten lässt sich jederzeit eine aktuelle Systemdokumentation erzeugen, und zwar für jede einzelne Umgebung (DEV, TEST, PROD). Ebenso schnell können Ad-hoc-Impact- und -Datenflussanalysen auf einzelnen Objekten ausgeführt werden und Datenflüsse von der Extraktion aus dem Quellsystem bis zur analytischen Kennzahl sichtbar gemacht werden – ein unschätzbare Mehrwert nicht nur bei großen Data-Warehouse-Systemen.

Die jederzeit aktuelle und vollständige Dokumentation stellt sicher, dass die Auswirkungen von Änderungen bereits im Vorfeld viel besser eingeschätzt werden können und es nicht zu bösen Überraschungen während der Umsetzung kommt. Der Aufwand kann also besser abgeschätzt werden, was ein wesentlicher Erfolgsfaktor für die erfolgreiche Umsetzung von Anforderungen innerhalb agiler Sprints ist.

Auch im betrieblichen Bereich bieten sich deutliche Vorteile durch Data-Warehouse-Automatisierung. Neben der oben beschriebenen Möglichkeit zur automatisierten Erstellung des „Schedulings“ umfasst dies auch das Logging und Monitoring. Hier werden Beginn und Ende eines Ladeschrittes, die Anzahl verarbeiteter Datensätze sowie natürlich auch der Prozessstatus getrackt. Der dazu notwendige Logging-Code wird vom Code-Generator automatisch in die generierten Prozeduren eingefügt.

Nicht weniger wichtig ist die Überwachung und Plausibilisierung der Data-Warehouse-Inhalte. Entsprechende Statements und Funktionsaufrufe zur Ermittlung und Prüfung von Tabellengrößen, der Anzahl unterschiedlicher Werte, von Extremwerten, Stringlängen etc. lassen sich ebenso einfach mittels Code-Generator an den richtigen Stellen einbauen. Dies kann bis zur automatisierten Erstellung eines Tickets im Ticketingsystem direkt durch Ereignisse in Ladeprozessen gesteuert werden.

## Mehrwert durch Data-Warehouse-Automatisierung auf einen Blick

- Größerer Fokus auf Fachlichkeit
- Gleichbleibende und einheitliche Code-Qualität
- Stets aktuelle Dokumentation
- Kurze Time-To-Market
- Höhere Standards und Best Practices

### Schnellere Umsetzung

### Bessere Qualität

- Schnellere Analysen durch End-to-End-Impact-Analyse
- Schnellere Feedback der Fachbereiche zur Eignung des Lösungsansatzes
- Automatisierte Ladesteuerung

- Einfache Änderbarkeit
- Automatische Versionierung
- Einfaches Deployment durch Kopie von Metadaten

### Erhöhte Flexibilität

### Zeit- und Kostenersparnis

- Erhöhung des Return on Investment
- Verwendung der bestehenden Ressourcen für neue Innovationen und Projekte
- Kürzere Coding-Zeiten durch Code-Generierung

## Mit b.telligent erfolgreich die Vorteile von Data-Warehouse-Automatisierung erzielen

b.telligent setzt bereits seit Jahren sein eigenes Data-Warehouse-Framework erfolgreich in Kundenprojekten ein und verfügt dadurch über weitreichende Erfahrungen in den genannten Themen. Darüber hinaus pflegt b.telligent eine enge Partnerschaft mit WhereScape, dem Pionier im Bereich Data-Warehouse-Automatisierung. Das aus Neuseeland stammende Unternehmen stellt mit seinen Produkten WhereScape RED und WhereScape 3D professionelle und ausgereifte Werkzeuge bereit, die die Data-Warehouse-Automatisierung hervorragend unterstützen.

Gerne unterstützen wir Sie und beraten Sie dabei, wie Sie mit Hilfe von Data-Warehouse-Automatisierung Ihre Entwicklungsgeschwindigkeit vervielfachen und dadurch erhebliche Kosten sparen können. WhereScape-Kundenreferenzen zeigen, dass sich durch den Einsatz von Data-Warehouse-Automatisierung die Entwicklungsgeschwindigkeit vervielfacht, die Projektrisiken erheblich reduziert werden und der „Return on Investment“ damit schon innerhalb weniger Monate erreicht wird.





### Über den Autor

Klaus Blaschek ist Gründer und Geschäftsführer von b.telligent und berät seit über 14 Jahren Kunden der verschiedensten Branchen in den Themen Business Intelligence, Data Warehouse, Customer Intelligence, CRM und E-Commerce.

### Über b.telligent

Die Unternehmensberatung b.telligent hat sich auf die Einführung und Weiterentwicklung von Business Intelligence, Customer Relationship Management, DWH und Big Data in Unternehmen in Massenmärkten spezialisiert.

Der Fokus liegt dabei auf der kontinuierlichen Optimierung

von Geschäftsprozessen, Kunden- und Lieferantenbeziehungen durch den Erkenntnisgewinn aus der Verdichtung, Analyse und Prognose von systemübergreifenden Geschäftsdaten. So lassen sich Margen erhöhen, Kosten senken und Risiken besser kontrollieren. [www.btelligent.com](http://www.btelligent.com)

#### b.telligent

Walter-Gropius-Straße 17 • 80807 München  
info@btelligent.com • Tel: +49 89 122 281 110

#### Klaus Blaschek

Managing Director  
klaus.blaschek@btelligent.com

#### Follow us

