

Helsana



Vom DWH-Testing zum Data Quality Management

Alex Häfliger, Bridge Consulting GmbH

Felix Sauter, Helsana

Inhaltsverzeichnis/Agenda

- 1. Aufbau des DWH Testings**
- 2. Data Quality Management**
- 3. BiG EVAL, was die Software sonst noch bietet**

Vorstellungsrunde

Bridge Consulting
Alex Häfliger
www.bridge-consulting.ch



Langjähriger externer Dienstleister und Berater der Helsana Versicherungen
Aufbau Testing DWH

BiG EVAL
Thomas Bolt
www.bigeval.com/de/



Software Lieferant BiG EVAL

IT-Logix
IT-LOGIX
BUSINESS INTELLIGENCE



Externer Dienstleister im Projekt IG WhereScape

Helsana
Helsana
Engagiert für das Leben.



Die Helsana AG mit Sitz in Dübendorf ist ein auf die Krankenversicherung spezialisierter Schweizer Versicherungskonzern

Aufbau des DWH-Testings

Der Auftrag

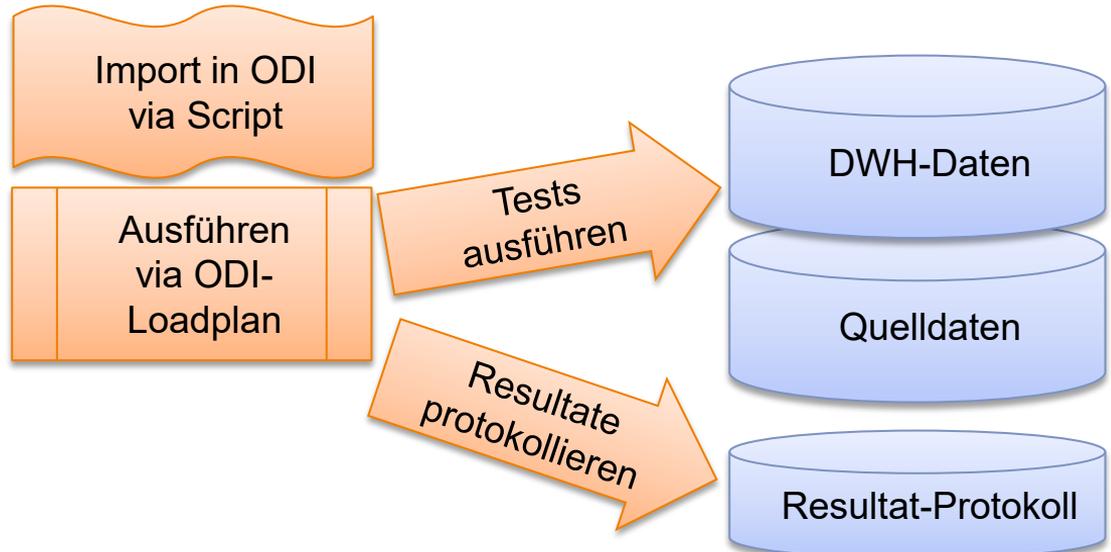
DWH-Projektleiter nach den ersten Sprints:

«Wir haben jetzt die ersten Ladestrecken implementiert. Alex, könntest du bitte prüfen, ob die Daten korrekt ankommen»

Zunächst stellte ich mir die folgenden Fragen:

- Wie kann ich die Ladestrecken effizient überprüfen?
- Was kann ich ohne tiefes fachliches Wissen überhaupt Prüfen?
- Wie kann ich die Prüfungen immer wieder wiederholen, da ja laufend noch Anpassungen an den Fakten und Dimensionen gemacht werden?

Initialer Aufbau der Tests



Testmethoden

Vergleiche
Objektvergleiche
Prüfsummen

Konsistenz-Prüfungen
Dubletten-Prüfungen
Zeitachsen-Prüfungen
Referenzen
Datumswerte
Stringlängen

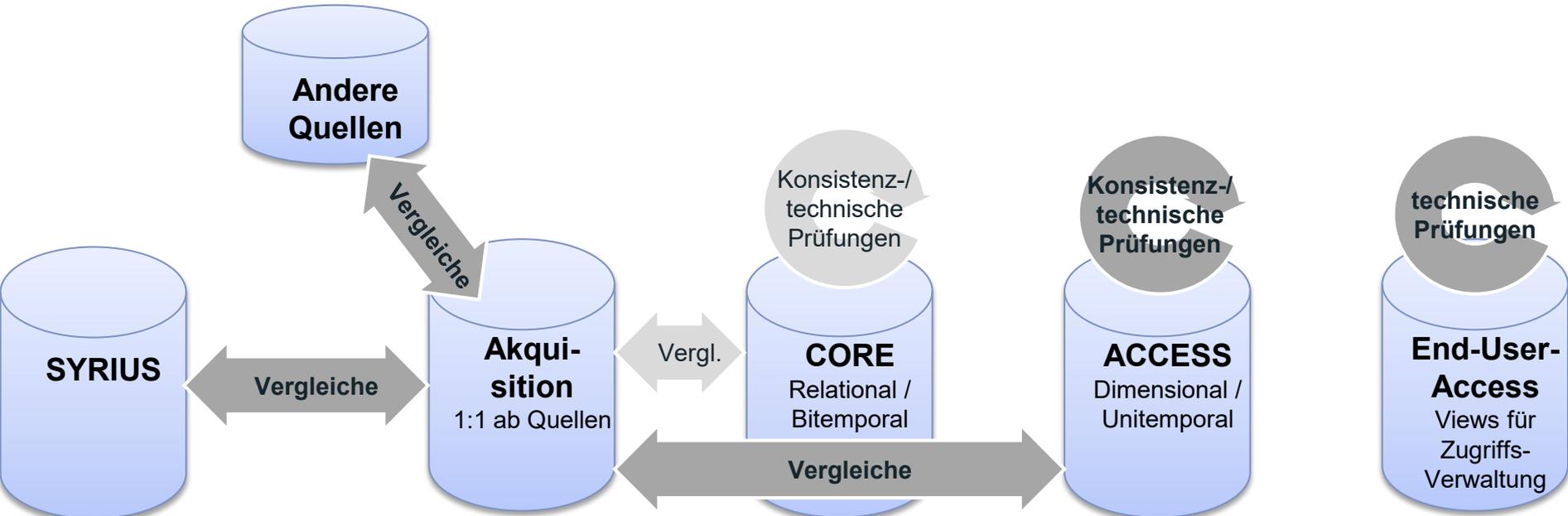


Technische Prüfungen
Leere Tabellen
Views ausführbar
Views vorhanden



Historien-Vergleiche
Historie-Werte
Historie-Zeilen
Verteilung

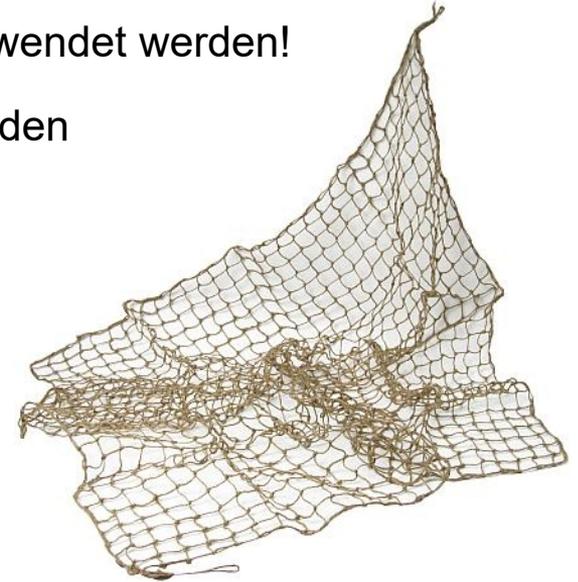
DWH-Layer



Testnetz aufbauen

Die aufgebauten Methoden nützen nichts, wenn sie nicht angewendet werden!

- Tests konsequent auf alle neu entwickelten Tabellen anwenden
- Testabdeckung immer wieder überprüfen
- Tests regelmässig ausführen
 - Testfenster in Beladungsplanung berücksichtigen
 - Testausführung muss selbstverständlich werden

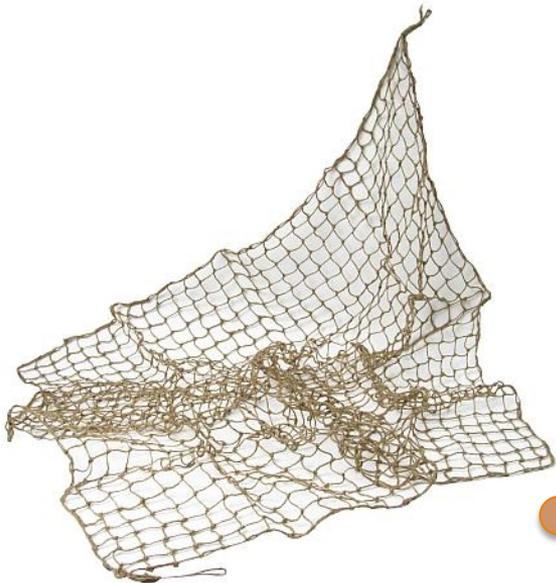


Die Kombination von allen verfügbaren Testmethoden über alle vorhandenen Tabellen ergibt ein Testnetz, welches kontinuierlich dichter wird.

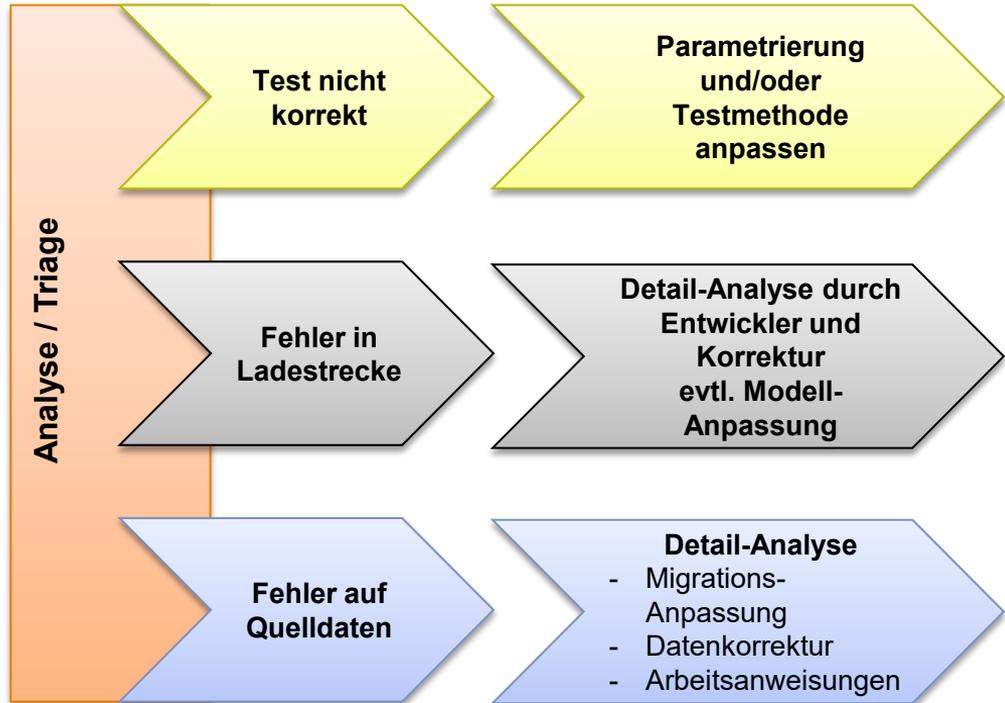
Resultate analysieren und Verteilen

Im Testnetz

bleibt auch mal was hängen!



Befund
aus Test-
ausführung



Verwendung der Tests

Im Laufe des Projektes wurden die Tests, bzw. die Testresultate immer weiter verwendet

1. Entwicklung der DWH-Ladestrecken
2. Indirekt in Migration und Parametrierung des Quellsystems SYRIUS
3. Integrations-Tests des gesamten Systemverbundes
4. Release-Freigabe
5. Release-/Produktions-Verifikation
6. Operativer Betrieb

Die Tests werden nach Projektabschluss weitergeführt. Somit muss die Wartung an interne Mitarbeiter übergeben werden können.

Wartbares Testsystem gesucht

Nachteile bisherige Lösung

- Komplizierter Aufbau mit Scripts in unterschiedlichen Systemen
- System wurde zu schwerfällig
 - Generierung von geänderten Tests dauert gegen eine Stunde
 - Deployment in ODI dauert mehrere Stunden
- Verknüpft mit Releaseprozess
 - Deployment von Änderungen nur im Rahmen eines Releases möglich
 - Nur komplettes Testpaket austauschbar
- Zentrale Datenhaltung der Test-Administration in Excel

Das System ist auf Dauer nicht wartbar.

Wartbares Testsystem gesucht

BiG EVAL überzeugt durch folgende Punkte

- Einfache Systemarchitektur
- Intuitive und ansprechende Benutzeroberfläche
- Erweiterbar dank Integration von eigenen C#-Methoden und SQL-Statements

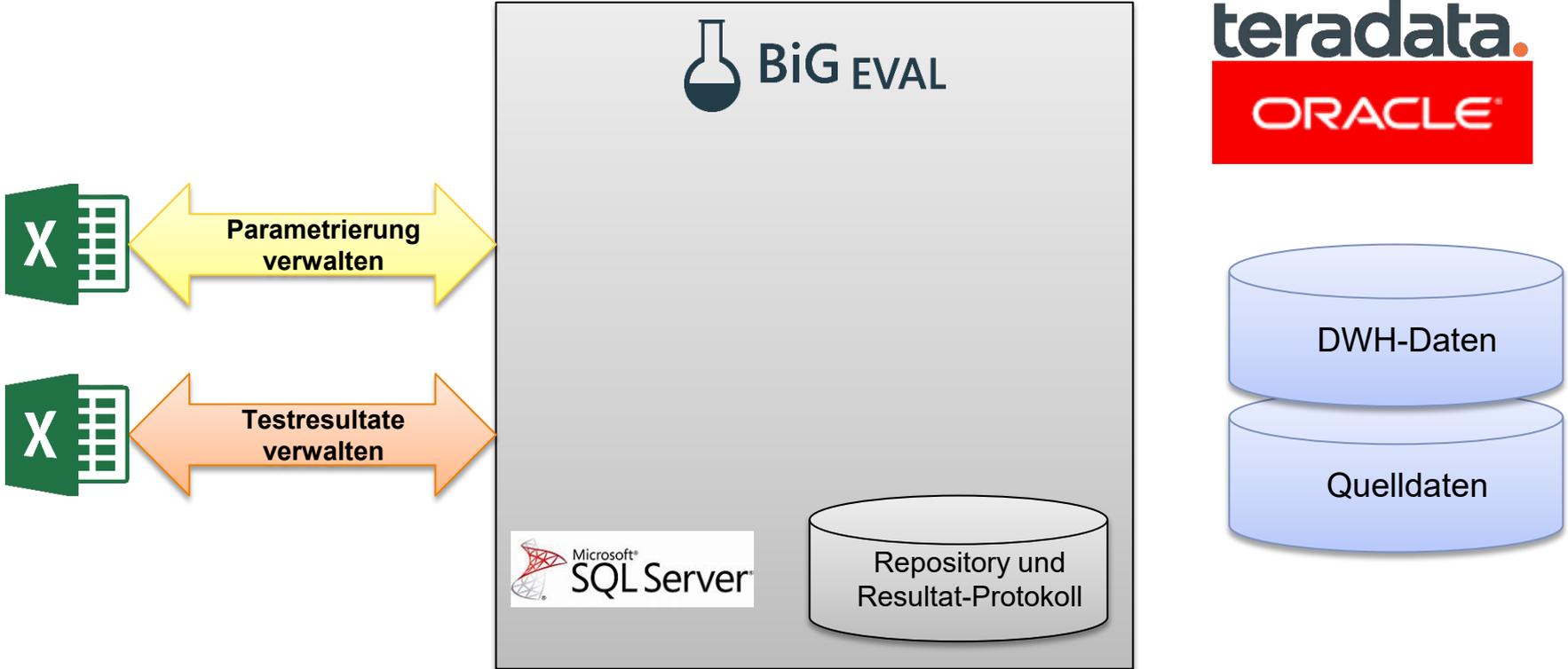


Der Proof of Concept zeigt folgende Resultate

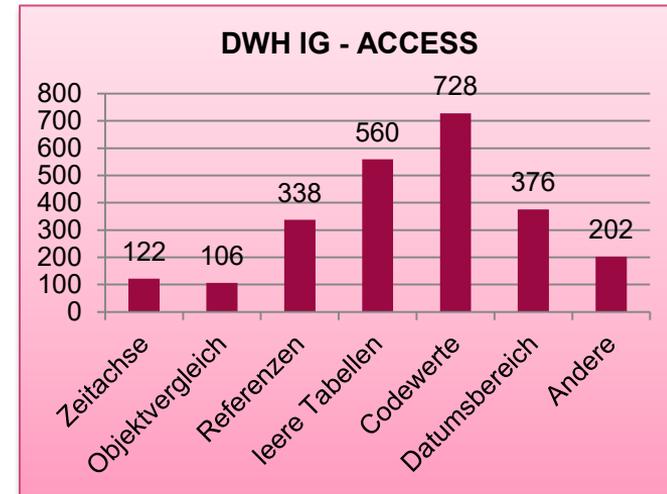
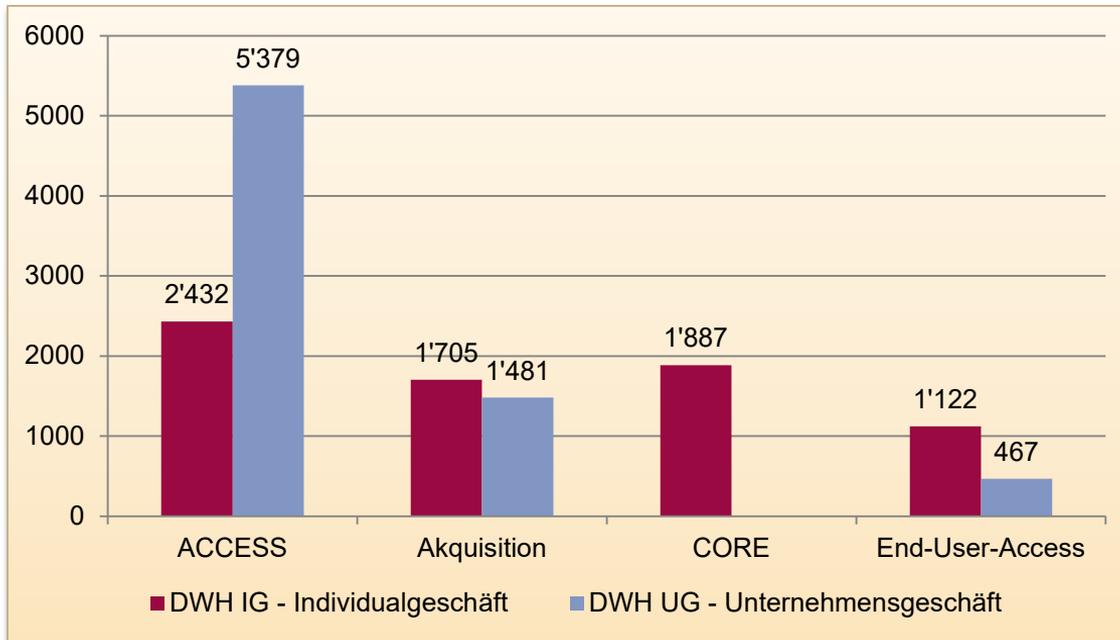
- Die bestehenden Testmethoden konnten im PoC mit eigenen Erweiterungen umgesetzt werden (Die notwendigen Erweiterungen wurden anschliessend direkt in der Software integriert)
- Administration, Ausführung und Dokumentation der Tests können in einem System zentralisiert werden
- Automatisierung der Testläufe durch ControlIM konnte umgesetzt werden.
- Konstruktive und unkomplizierte Zusammenarbeit mit Hersteller Bolt Technology Consulting

Keine weitere Evaluation notwendig. BiG EVAL wird bei Helsana eingeführt!

Aufbau der Tests mit BiG EVAL



Statistik der ausgeführten Tests

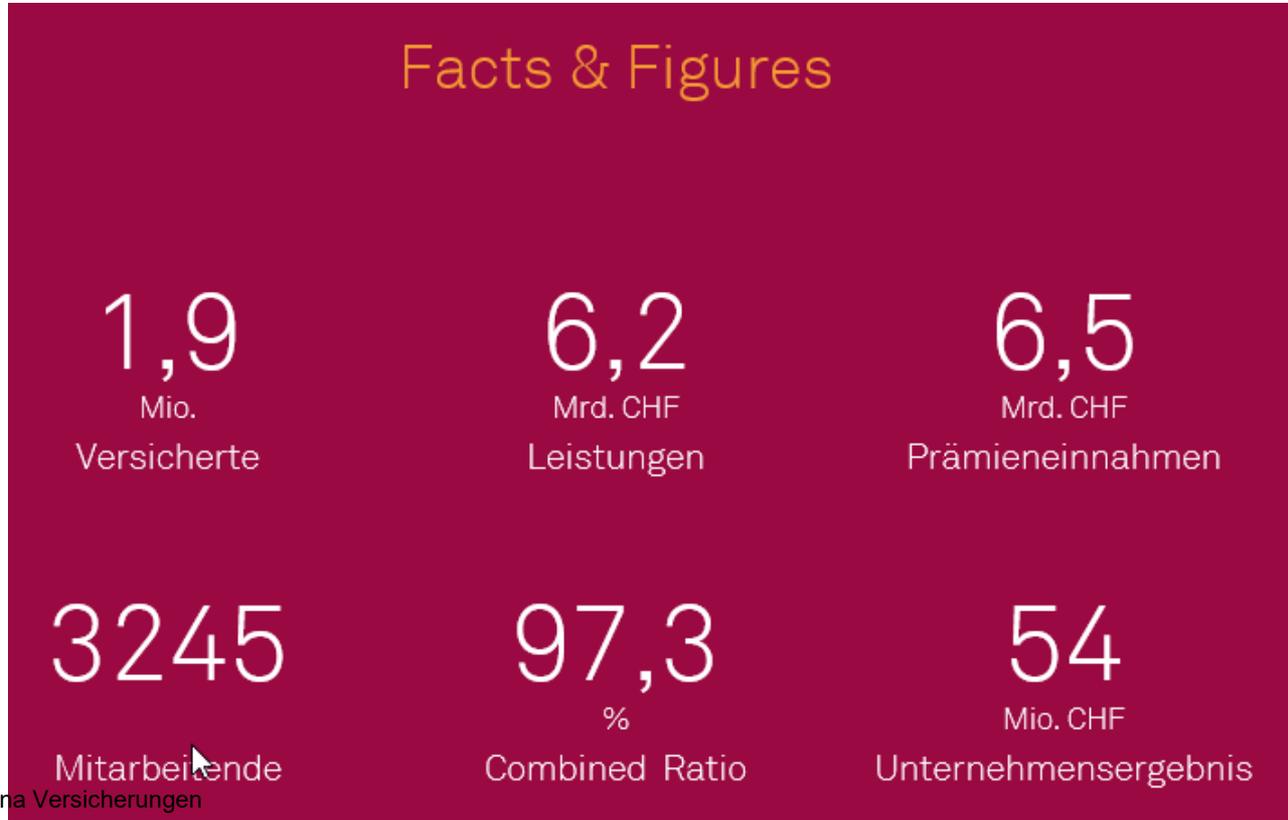


Data Quality Management

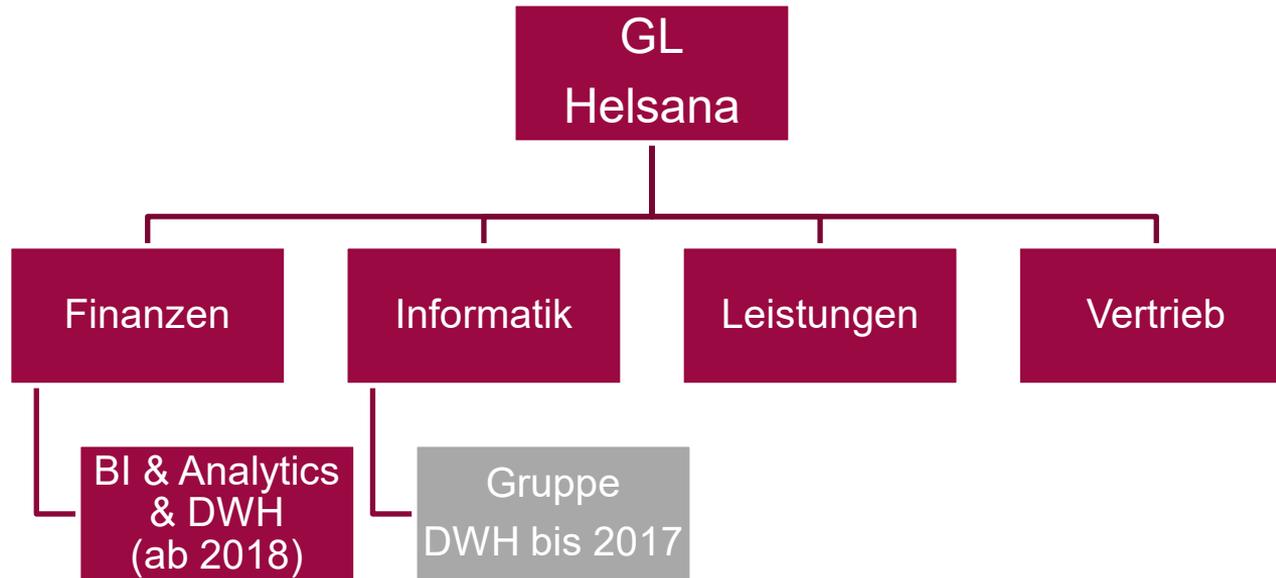
Helsana Versicherungen

Nachfolgend im Markt vergleichbare Fakten

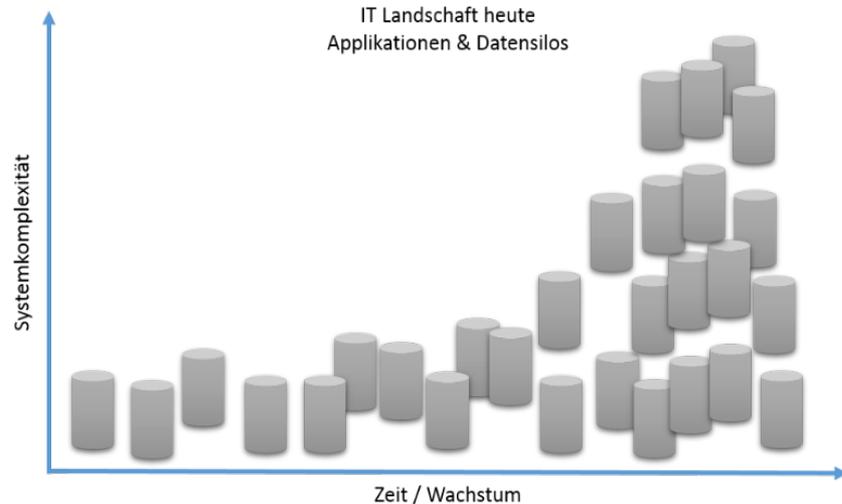
Facts & Figures



Organisation Helsana / BI & Analytics und DWH



Qualität und ihr Verständnis



Quelle: Sarsfield, The Data Governance Imperative, 2009

- Entwicklung der zunehmenden Systemkomplexität
- Fehlende Stosskraft für das umfassende Datenqualitätsmanagement
- Fehlende Sensibilisierung auf allen Stufen
- Fehlende Unterstützung aus dem Management
- Unnötige Kunden-Beschwerden
- Keine qualitativen und finanziellen Aussagen zu Datenqualität

Qualität und ihr Verständnis

Das Ziel "Excellent Customer Experience" wird nur in der Kombination verschiedenster, eindeutiger Verbesserungen und Anpassungen in Prozessen und Daten erreicht werden

- Helsana hat von 2015-2017 das Projekt Zielbild für das Individualgeschäft umgesetzt
 - Neue Systeme als Quellen
 - Neues Datawarehouse
 - CRM in der Cloud und verschiedenste, weitere Plattformen für den internen und externen Datenaustausch wurden mit dem Ziel aufgebaut, den Kunden in einer 360° Sicht zu begleiten.
- Datenqualität (DQ) war in dieser Phase keine effektive Planung, Detaillierung, Umsetzung und Überprüfung wert
 - Fehlende durchgängige Wahrnehmung betreffend des Einflusses schlechter DQ auf den Wertschöpfungsprozess, der entsprechenden Auswertungen und der direkten und indirekten Kosten für DQ Verbesserungen belasten heute die Zielsysteme für die 360° Kundensicht und die Business Intelligence Datenlandschaft.

Kritische Erfolgsfaktoren (KEF)

AKV = Aufgaben, Kompetenzen, Verantwortung

KEF	Beschreibung
K1	<p>Einheitliche Datenbasis</p> <p>Eine systematisch strukturierte Datenbasis soll übergreifend die Aussagen zur Datenqualität und Interpretation derselben sicherstellen. Die Systematik sichert ebenfalls den als real eingeschätzten Erweiterungsbedarf der Datenbasis</p>
K2	<p>Rollen, Personen und AKV</p> <p>Die Klärung von Rollen, AKV und der Besetzung mit relevanten Mitarbeitenden ist zentral für die Akzeptanz von Verständnis Datenqualität und dessen ökonomischer Auswirkungen.</p> <p>Der bereits bestehende aber kaum gelebte Prozess DQ erfährt einen Relaunch</p>
K3	<p>Aufbereitung gem. Stakeholder</p> <p>Der Unternehmensbezug differenter Stakeholder erfordert verschiedene Aggregationsstufen aufbereiteter Kennzahlen. Falsche Aufbereitungen resp. Aggregationen führen zu Fehlentscheidungen und haben ggf. einen Akzeptanzverlust zur Folge</p>
K4	<p>Einbezug der Stakeholder</p> <p>Um die zukünftig geplante, ausgeweitete Abdeckung von Datenqualitätskennzahlen entlang dem Wertschöpfungsprozess sicher zu stellen, ist der Einbezug der Stakeholder, langfristig und von BI getrieben, sicherzustellen</p>

Stakeholder

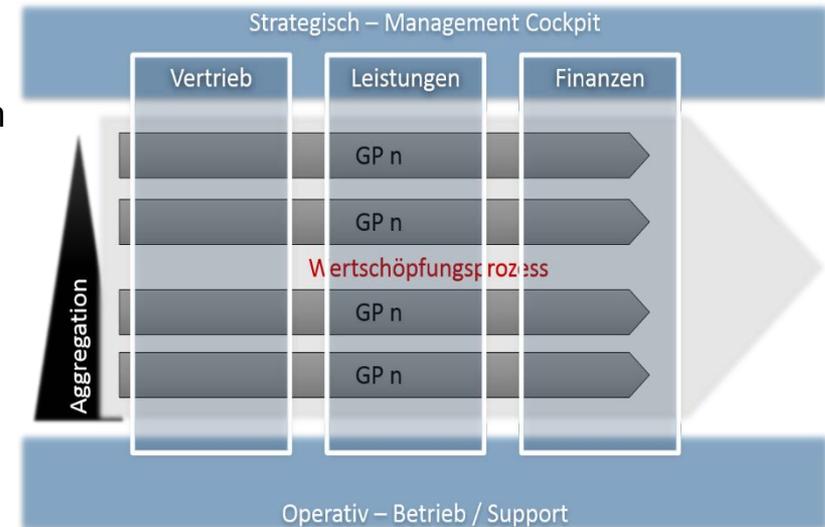
Unternehmensbezug	Stakeholder
Strategisch	Leiter Finanzen (Bezug GL) Leiter Business Intelligence & Analytics (BI&A)
Koordinativ	Datenschutz Risikomanagement Interne Revision (IKS) Bereichsleiter (Finanzen, Vertrieb, Leistung, Kundendienst, Gesundheitsökonomie) Beschwerdemanagement
Operativ	Verantwortlicher DQ im Bereich BI Fachverantwortliche DQ Kundenservice Generalagenturen (GA) Backoffices (Regionenbezug)

4 Quality Gates für DQ

QG	Titel	Beschreibung
1	Prozesse anpassen, stabilisieren und verantworten	Die bestehenden Prozesse werden so vernetzt, dass eine konzernübergreifende, effiziente und zielorientierte Zusammenarbeit möglich wird
2	Organisation vervollständigen	Die vorhandenen Strukturen werden eingebunden. Die Ressourcen mit DQ Auftrag werden somit optimal vernetzt
3	KB übergreifendes Vorgehen	Konzernübergreifende Probleme zu Datenqualität werden gemeinsam gelöst. Dies stellt die übergeordnete Lösung sicher
4	Tool unterstützte Prozesse	Tool Stammdatenprüfung -> IML Tool Technische Daten Prüfung -> BiGEVAL

Kennzahlensystem

- Ein dreistufiges Kennzahlensystem inkl. Quality Gates, das stufengerecht Informationen an verschiedene Empfänger liefert. Zusammen mit den zukünftigen Nutzern werden die Kennzahlen definiert und konsolidiert. Die Einzigartigkeit der Kennzahlen und die logische Kontrolle erfolgt in einem zweiten Schritt
- Die Aufbereitung der Kennzahlen erfolgt auf der aktuellen Modelllandschaft
- Die Integration der Kennzahlen ist in der Prozesslandkarte (Produkt TopEase) vorgesehen
- Vorgesehen als Vorstufe zu Dataownership



Helsana BSC DQ (geplant)

Finanzen

- Reduktion direkte und indirekte Kosten DQ
- ROI der eingesetzten Mittel
- Zusatzgeschäfte

Kunden

- Kundenzufriedenheit
- Rückgang Kundenbeschwerden
- Wiederempfehlungsrate (Net Promoter Score)

Prozesse

- Prozesskosten
- Prozesseffektivität
- Organisationssicherung

Potential/ MA

- Zeit für DQ Handlungen
- Einhaltung Arbeitsvorgaben
- Mitarbeitenden Zufriedenheit

Eine sinnvolle Darstellung und Bewirtschaftung der KEF und der KPI ist bezugnehmend auf den Aggregations- resp. Darstellungslevel die Business Scorecard (BSC).

Vorteile der BSC sind die eindeutigen Ursache-Wirkungs Informationen und die nachvollziehbar abgeleiteten Massnahmen.

Reifegradmodell Datenqualität

Ein Reifegradmodell gibt darüber Auskunft, wie ein Unternehmen Datenqualität umfassend sicherstellt und es die Basis für KVP und dessen abgeleitete Massnahmen. Die Durchführung von RG Bewertungen führt zu Etablierung und Entwicklung von DQM im Unternehmen

Optimierend		<ul style="list-style-type: none"> Datenqualitätsmanagement (DQM) wird optimiert 	Hoch
Steuernd		<ul style="list-style-type: none"> Vollständige Dokumentation der Informationsstruktur DQ Governance etabliert, Problemlösung im Prozess 	Nutzen
Messend		<ul style="list-style-type: none"> Information wird gemanaged und als Potential gepflegt DQ Metriken und Kennzahlensystem aufgebaut 	Nutzen
Reaktiv		<ul style="list-style-type: none"> Informationsprozesse dokumentiert, DQ Prozess vorhanden Bewusstsein betr. DQ initiiert 	Nutzen
Chaotisch		<ul style="list-style-type: none"> Prozesse vorhanden, keine Dokumentation Kein oder begrenztes Bewusstsein betr. DQ, Ignoranz 	Tief

-  = Ansatz
-  = Definiert
-  = Eingeführt

DQM@Helsana

	Operativ		Taktisch		Strategisch
Organisation	CC DQ (StDa)	CC DQ (StDa)	CC DQ	HCB (DQ)	HCB
Prozesse	DQ StDa	DQ Fach/Techn	DQ (PLK)	DQ (PLK)	Kernprozesse
Toolunterstützung	IML (ab 2012)	IML/BiGEVAL	IML/BiGEVAL	IML/BiGEVAL	
Kennzahlensystem					
Visualisierung / Dashboard					
Kritische Erfolgsfaktoren (KEF)					
Verbreitung im Unternehmen					
Bewusstsein	Ignoranz	Initiiert	Integriert	Etabliert	Entwickelt
Organisation	Funktional	Weisung	Matrix	Prozess	Service
Reifegradmodell	Chaotisch	Reaktiv	Messend	Steuernd	Optimierend
	2015/17	2018/19		2020 folgende	



BiG EVAL

High Data Quality.
Mastered.

BiG EVAL

Testautomatisierungs- und DQM-Suite

Was die Software sonst noch bietet

Business Regeln

not empty?

Valid
Range?
Max 50%
Discount

Name	Birthdate	Gender	Email	Discount
Lauren Walker	1971-10-06	F	lauren@walker.com	0%
Julio Ruiz		M	ruiz@contoso.com	90%
Amanda Rivera	1984-08-26	W	ar@google.com	5%
Marcus Harris	1986-22-45	M	Marcus.harris@	10%

Valid Format?

M or F only?

Valid Email?

Edit Rule

Name: Gender not empty

Description: This rule checks whether the gender-fields of all employees are filled-in.

Active:

Checks probe: 1016.1 - Employees

Max. Rule Breaches: 0 Rows

Expression: *Data("Gender").IsNotNull

RULE	STATUS	SUCCESSRATE	PROCESSED	SUCCEEDED	FAILED
Birth Year > 1970 This rule checks whether all employees have a birth year of 1970 and above. Probe: 1016.1 - Employees Fault Tolerance: 0 rows	✗	87%	290	252	38 Show
VacationDays depending on age (using Script-Syntax) Checks whether the employee has the right amount of vacation-days depending on his/her age. Probe: 1016.1 - Employees Fault Tolerance: 0 rows	✗	0%	290	0	290 Show
Gender M or F (using Script-Syntax) This rule uses the script-syntax to check whether the gender-field has allowed values only. Probe: 1016.1 - Employees Fault Tolerance: 0 rows	✓	100%	290	290	0
Gender not empty This rule checks whether the gender-fields of all employees are filled-in. Probe: 1016.1 - Employees Fault Tolerance: 0 rows	✓	100%	290	290	0
Gender M or F (using Expression-Syntax) This rule uses the expression-syntax to check whether the gender-field has allowed values only. Probe: 1016.1 - Employees Fault Tolerance: 0 rows	✓	100%	290	290	0

Testset Addons

- Breite Testabdeckung mit wenigen Testpatterns.
- Schnelle Testentwicklung mit guter Wartbarkeit.
- Schnelle Ergebnisse mit wenig Aufwand zur Testentwicklung.
- Direkte Anwendung bewährter Tests aus live Projekten mit Parametrierung.
- Anpassungen und Weiterentwicklung der Testpattern nach Bedarf.



Vergleiche

Objektvergleiche

Prüfsummen

Konsistenz-Prüfungen

Dubletten-Prüfungen

Zeitachsen-Prüfungen

Referenzen

Datumswerte

Stringlängen



Quell-DB



Vergleiche



Ziel-DB



Historien-
Vergleiche

Technische Prüfungen

Leere Tabellen

Views ausführbar



technische
Prüfungen

Historien-Vergleiche

Historie-Werte

Historie-Zeilen

Verteilung

Unterstützung von vor- und nachgelagerten Prozessen

Anforderungs- Management

- Dokumentieren von Fachanforderungen für Tests
- Verknüpfungen mit ausführbaren Tests
- Sammeln und visualisieren von Testresultaten je Fachanforderung
- Verknüpfung mit Prozess-Tools ...

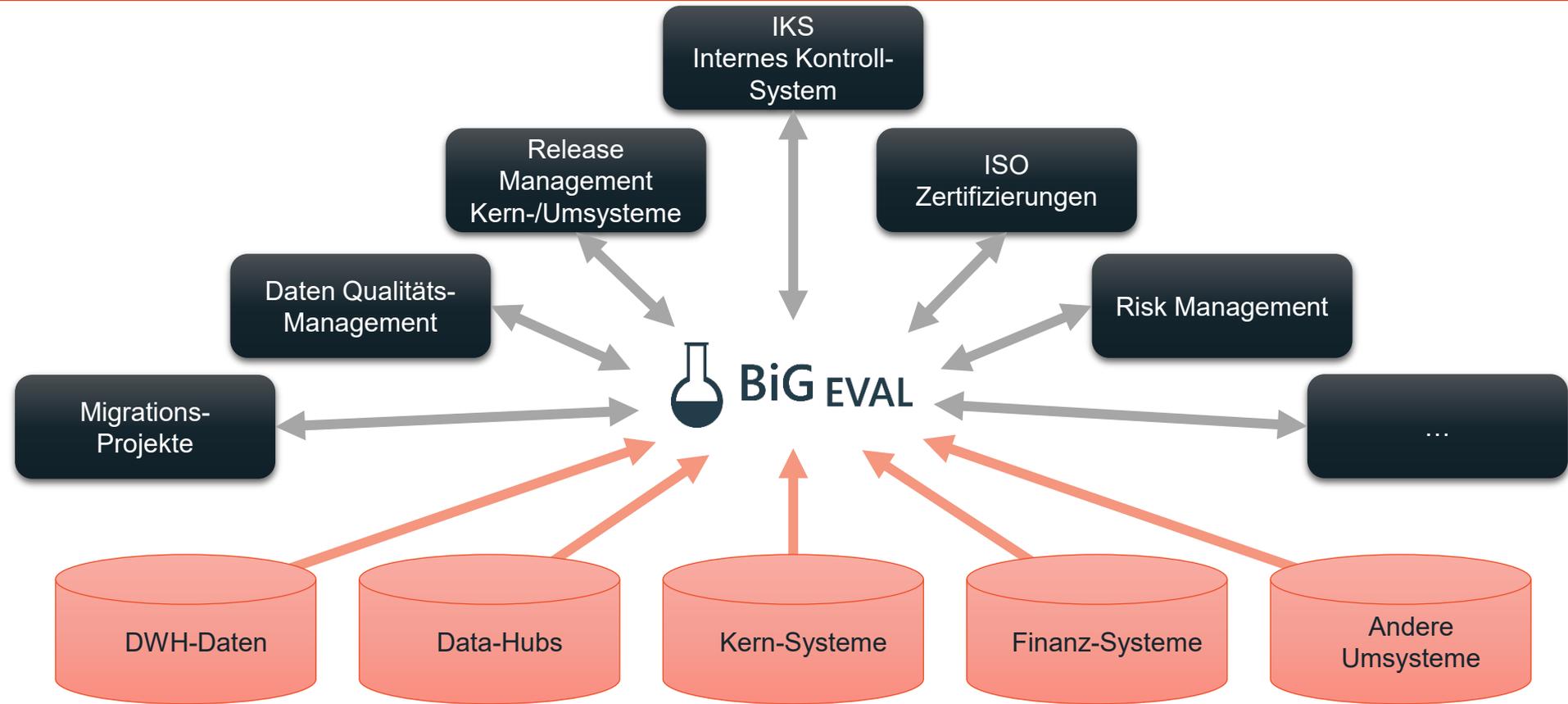
Automatisiertes Testing

- Kern-Funktionalität für automatisiertes Testing
 - erstellen
 - verwalten
 - ausführen
 - protokollieren
 - orchestrieren

Resultat- Management

- Identifizieren und verwalten von bekannten Feststellungen
- Festlegen von Limiten zur bekannten Feststellung
- Dokumentation von Massnahmen
- Schnittstelle zu Ticket-Software (z.B. Jira)

Einsatzmöglichkeiten von BiG EVAL



Vielen Dank!
Fragen?
