



NodeGraph – (Qlik Sense®/ QlikView® Add-on)

Die Datenqualität steigern, während sich der manuelle Aufwand reduziert.

November 2018

Autor:

Max Gebauer, Consultant | Product Owner NodeGraph

www.evaco.de

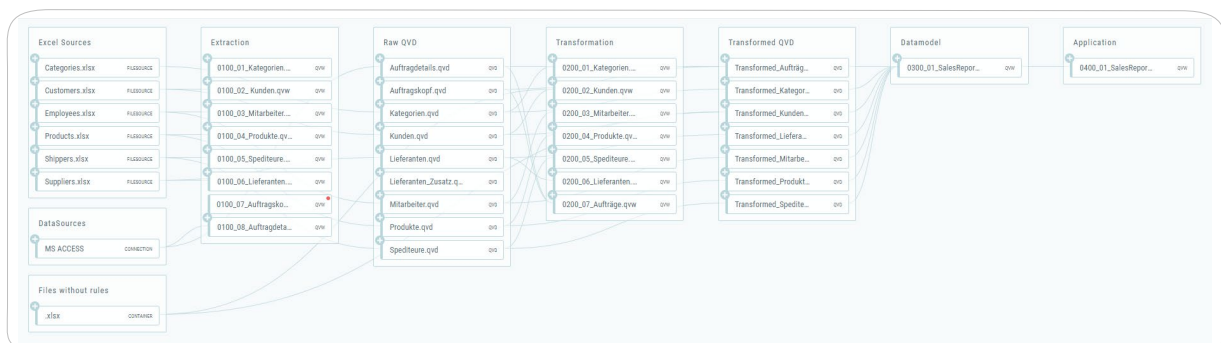
A large, light gray gear graphic is positioned in the bottom right corner of the page, partially overlapping the footer area.

Qlik Sense® / QlikView® Projekte – Daily Issues

Eine Herausforderung, der Qlik® User in der täglichen Arbeit gegenüberstehen, ist die hohe Investition von unverhältnismäßig viel Zeit in die Validierung der Daten. Die Steigerung der Datenqualität bedarf einen extrem hohen manuellen Aufwand, damit diese vertrauenswürdig sind: aktuelle Dokumentationen, Nachvollziehen der Datenherkunft, Beseitigen bzw. Vermeiden der Datenredundanz und Verstehen der Datenzusammenhänge. Darüber hinaus sind die Zugriffsberechtigungen oft undurchsichtig, dies ist mit der neuen DSGVO nicht mehr vereinbar.

Was ist NodeGraph?

Bei dem Qlik® Add-on NodeGraph handelt es sich um ein Data Lineage-Tool mit zusätzlichen Funktionen, welche in erster Linie dazu dienen die Entwicklungsumgebung zu visualisieren. Darüber wird Transparenz erzeugt, die dabei hilft Herkunft, Speicherung, Verarbeitung und Verwendung der Daten einfacher und somit schneller zu verstehen.



Qlik® Umgebung: Dependency Explorer

Neben der grafischen Übersicht hilft NodeGraph folgende Fragestellungen zu beantworten:

- Welche Veränderungen haben bestimmte Felder im Laufe des ETL-Prozesses durchlaufen?
- Fließen Daten in die Applikationen ein, die eventuell nicht dokumentiert sind?
- Gibt es redundante Daten innerhalb der Entwicklungsumgebung?
- Gibt es Dateien, Applikationen oder Transformationsschritte, die nicht mehr benötigt werden?
- Funktioniert die Sicherheit über die implementierte Zugriffsberechtigung auf die Qlik Sense® bzw. QlikView® Applikationen?

1. #Umsatz

Fakten
Process: c:\nodegraph\northwind\0300_data_model\0300_01_salesreport.qvw
🔍(1)
Fakten:
Load
[*] #Umsatz
FROM [..\0000_Resources\0020_QVD_Transform\Transformed_Aufträge.qvd] (qvd)
Created by: 0200_07_Aufträge.qvw

Transformed_Aufträge.qvd
Process: c:\nodegraph\northwind\0200_transformation\0200_07_aufträge.qvw
🔍(1)
Join (Aufträge)
Load
UnitPrice * Quantity as [#Umsatz]
FROM [..\0000_Resources\0010_QVD_Raw\Auftragdetails.qvd] (qvd)
Created by: 0100_08_Rename.qvw

Auftragdetails.qvd
Process: c:\nodegraph\northwind\0100_extraction\0100_08_auftragdetails.qvw
🔍(1)
Auftragdetails:
Load
[*] UnitPrice
[*] Quantity
SELECT OrderID,
ProductID,
UnitPrice,
Quantity,
Discount
FROM "Order Details"
Connection: MS ACCESS->MS ACCESS->Order Details

Modul: Field Explorer – ETL-Prozess



Für wen ist NodeGraph interessant?

Das Vertrauen in die Korrektheit der Daten spielt in der BI-Welt eine entscheidende Rolle. Deshalb bietet NodeGraph nicht nur dem Entwickler einen besseren Überblick über die Qlik® Landschaft, sondern bildet ebenfalls eine Brücke zwischen visualisierten Kennzahlen und der Datengrundlage. Durch das Auslesen der Applikationsinhalte ist es auch Power Usern oder Business Analysten möglich, die Formeln und Logiken der jeweiligen Objekte zu kontrollieren, zu verstehen und ebenfalls bis zur Datenquelle zurückzuverfolgen.



Application Content und Formel

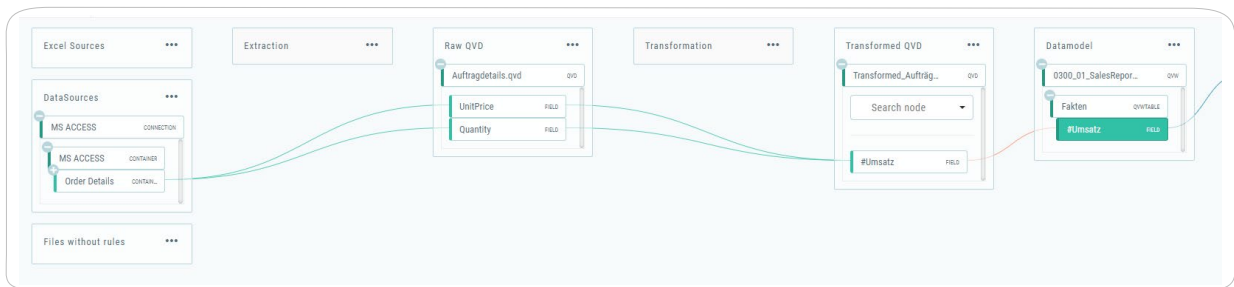
Mit NodeGraph sind die User stets auskunftsfähig bezüglich der Qlik® Landschaft, sparen Kosten und haben mehr Zeit, sich auf relevante Analysen zu fokussieren.

Die Qlik® Umgebung entspricht bereits der Best Practice ETL-Architektur.

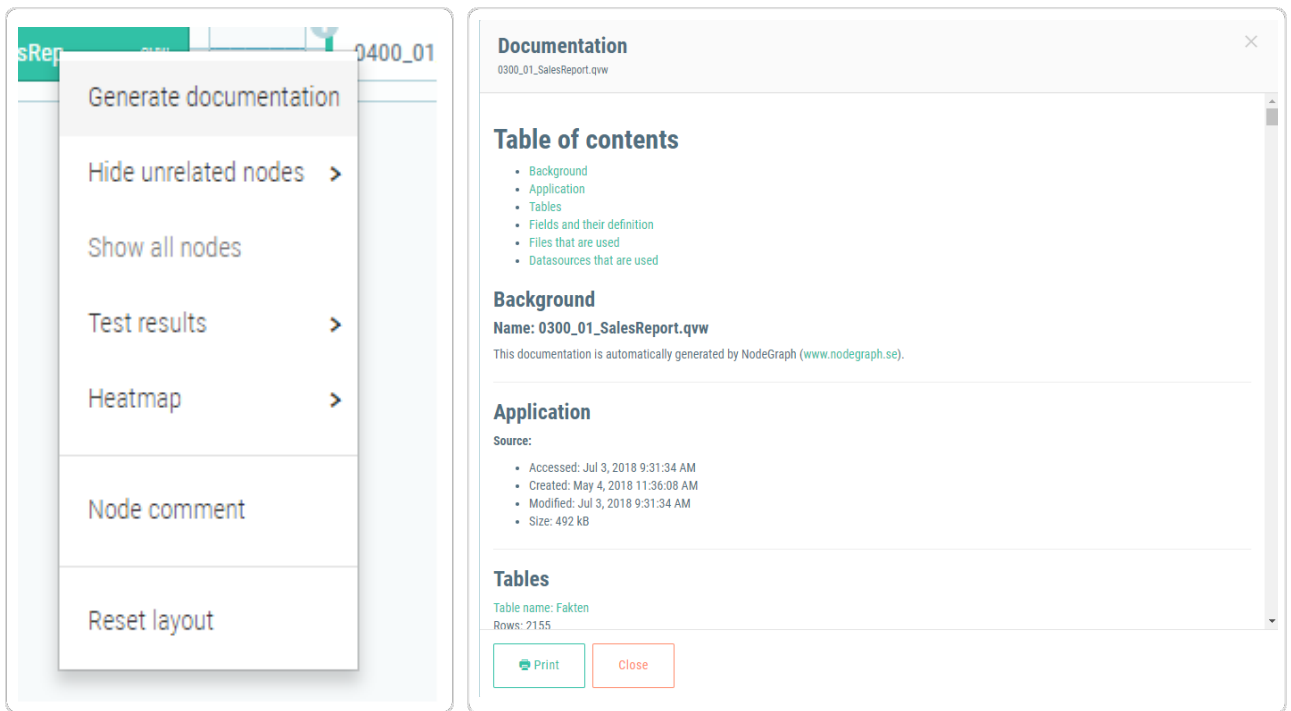
Lohnt sich NodeGraph trotzdem?

Data Lineage & Dokumentation

Auch die strukturierteste Entwicklungsumgebung bleibt, ab einem gewissen Grad der Komplexität, von Transparenzproblemen nicht verschont. Gerade in diesem Fall ist die regelmäßige Aktualisierung einer technischen Dokumentation notwendig. Die automatisierte Dokumentation von NodeGraph ermöglicht, stets den Überblick über die Qlik® Landschaft zu behalten, sodass Sorgen über die Aktualität keine Rolle spielen.

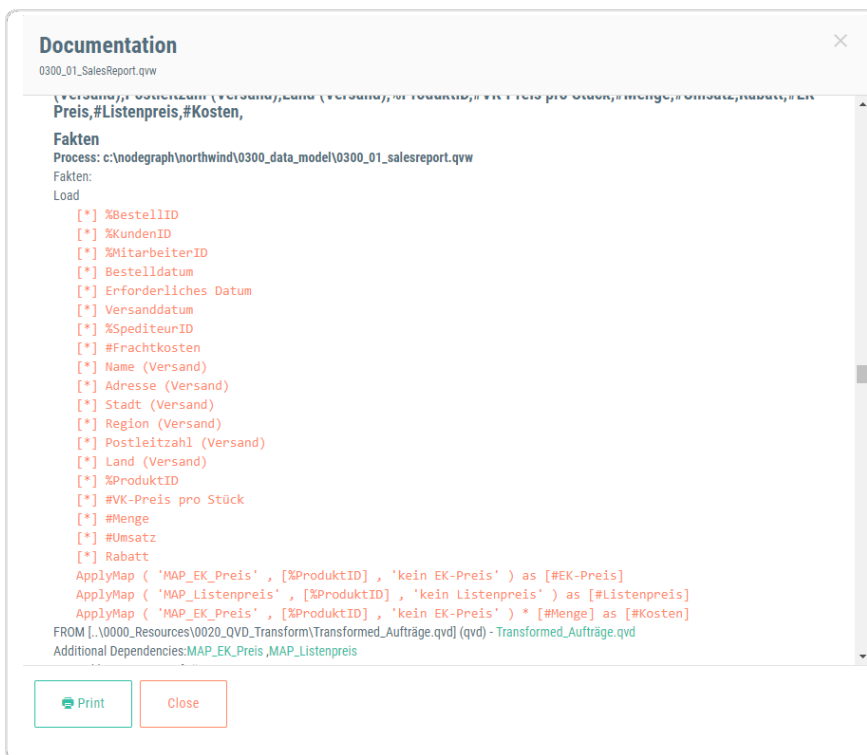


Datenherkunft und Datenverarbeitung



Dokumentation mit einem Mausklick

Strukturierte und detaillierte Dokumentation

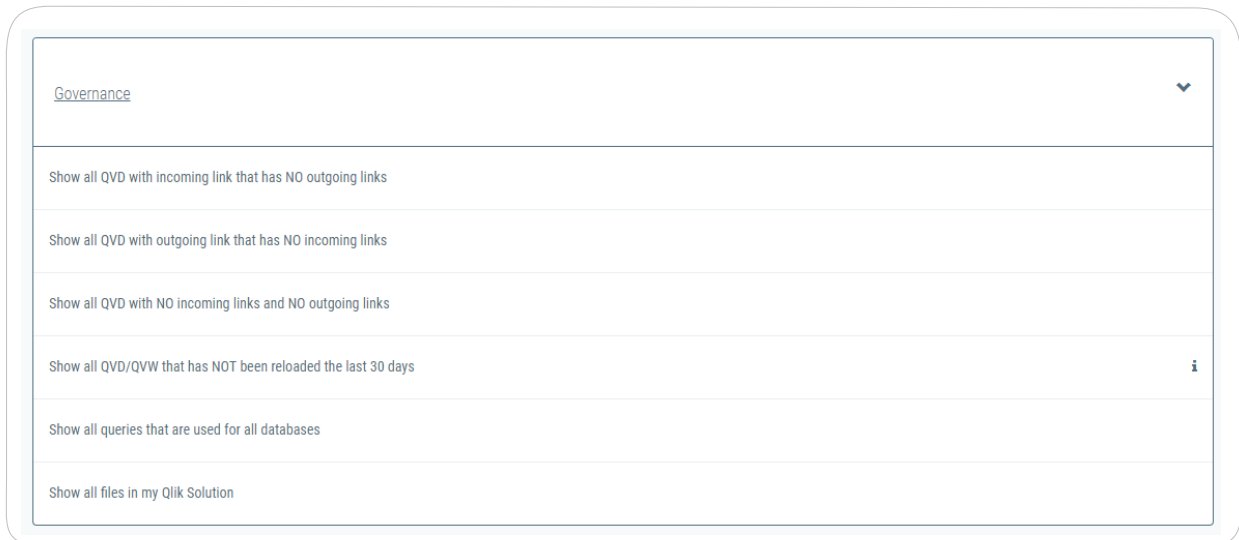


Dokumentation der einzelnen Transformationen



Governance

Neue Anforderungen bedeuten Anpassungen und Erweiterungen. Anpassungen können zu potenziellen Fehlerquellen und Erweiterungen zu Datenredundanz führen. Das Governance-Modul behält die Aktualität und die Verwendungszwecke der Daten im Auge. Diverse Ad-hoc-Berichte zeigen beispielsweise, ob Daten lediglich extrahiert, aber nicht verarbeitet werden oder ob verarbeitete Daten weiterhin aktuell sind.



Veränderungsprozesse über das Governance-Modul beobachten

DSGVO, auch General Data Protection Regulation (GDPR)

Die neue Datenschutzgrundverordnung, welche seit dem 25. Mai 2018 in Kraft getreten ist, erhöht den Fokus in der BI-Welt auf personenbezogene Daten deutlich. Das GDPR-Modul unterstützt bei der Überprüfung der Zugriffsberechtigung auf die Applikation. Wenige Klicks reichen, um zu erkennen, wer Zugriff auf die Applikation hat und wer bereits darauf zugegriffen hat. Die automatisierte Ausgabe des GDPR-Berichts liefert stets einen aktuellen Überblick über das Berechtigungskonzept und damit die Kontrolle über sensible sowie personenbezogene Daten. So spart man auch an dieser Stelle erheblich viel Zeit, durch wegfallende, zeitintensive Analysen.

Transparente Dokumentation der Verwendung:

- ✓ aktuelle Zugriffe auf die Daten
- ✓ Zugriffe in der Vergangenheit
- ✓ Transparenz der Datenquellen
- ✓ Speicherort der Daten
- ✓ Zweck der Datenspeicherung
- ✓ Verarbeitung der Daten
- ✓ Veränderungen der Felder
- ✓ lückenlose Dokumentation

3 Tracked Fields

Field: Kontaktname (Kunde)

Location: c:\nodegraph\northwind\0400_application\0400_01_salesreport.qvw->Dim_Kunden

Description:

Field: Name (Kunde)

Location: c:\nodegraph\northwind\0400_application\0400_01_salesreport.qvw->Dim_Kunden

Description:

Field: %KundenID

Location: c:\nodegraph\northwind\0400_application\0400_01_salesreport.qvw->Dim_Kunden

Description:

Field: %KategorienID

Location: c:\nodegraph\northwind\0200_transformation\0200_01_kategorien.qvw->Kategorien

Description:

3.1 Kontaktname (Kunde)

Location: c:\nodegraph\northwind\0400_application\0400_01_salesreport.qvw->Dim_Kunden

Description:

Kontaktname (Kunde) is used in 1 applications

3.1.1 0400_01_SalesReport.qvw

Location: C:\NodeGraph\Northwind\0400_Application\0400_01_SalesReport.qvw

3.1.1.1 Users with access

NT-AUTORITÄT\SYSTEM
VORDEFINIERT\Administratoren
VORDEFINIERT\Benutzer
EVACO\mgr

3.1.1.2 Users that has accessed

No usage found.

DSGVO-Auflagen über vorbereitete Berichte erfüllen



Datenvalidierung

Als weitere Instanz eingesetzt, dient das Test-Modul dazu, erste Plausibilitätsprüfungen der Daten automatisiert durchzuführen. Tests erfolgen über eine benutzerfreundliche Oberfläche, die zeigt, ob Daten vollständig, plausibel, aktuell und somit vertrauenswürdig sind.

^ Auftragdetails.qvd - QVD	Not run	⚙
Item exists - ?	Not run	⚙
Item updated - ?	Not run	⚙
Fields exist - ?	Not run	⚙
QVD record count - ?	Not run	⚙
Field aggregation - ?	Not run	⚙
∨ 0100_07_Auftragskopf.qvw - QVW	Not run	⚙

Zeit und Kosten durch einfaches Testing reduzieren



Wie aufwändig ist es, NodeGraph für meine Qlik® Umgebung zu konfigurieren?

Installation und Konfiguration von NodeGraph sind so simpel wie die Bedienung selbst. Durch das Angeben von Überverzeichnissen, Streams, Dateitypen oder expliziten Dateinamen bzw. Applikationen, wird NodeGraph entsprechend der vorhandenen ETL-Landschaft konfiguriert. Zusätzlich besteht die Freiheit, die Qlik® Umgebung innerhalb von NodeGraph neu zu erfinden und nach eigenen Vorstellungen zu konfigurieren bzw. zu visualisieren. Bereits während der Konfiguration lassen sich Optimierungsmöglichkeiten erkennen, um Prozesse sauber, strukturiert und transparent zu halten.

PATH ●

C:\NodeGraph\Northwind\0100_Extraction

[Browse path](#)

Use rule name as top level

Scan subfolders

Scan subfolders as levels

Include content

INCLUDE FILES ●

*.qvw

EXCLUDE FILES & FOLDERS

Optional

Installation und Konfiguration von NodeGraph



Fazit: NodeGraph verringert den manuellen Aufwand und steigert die Datenqualität. Der User handelt dabei stets datenschutzkonform.

Mit dem Qlik® Add-on NodeGraph lassen sich alle Abhängigkeiten visualisieren: Der User hat einen Überblick über Herkunft, Speicherung, Verarbeitung und Verwendung der Daten. So lassen sich vorhandene Daten besser verstehen. Redundante Daten deckt NodeGraph auf, sodass sich diese beseitigen bzw. vermeiden lassen. Darüber hinaus hat der User entsprechend der DSGVO stets die Kontrolle über Zugriffsberechtigungen. Installation und Konfiguration des Add-ons sind denkbar einfach.

EVACO berät Sie gerne individuell. Hier erreichen Sie uns:

EVACO GmbH

Philosophenweg 31-33

47051 Duisburg

Tel.: +49 203 709002-0

info@evaco.de

www.evaco.de

Duisburg • Hamburg • Leipzig • München

Folgen Sie uns! Aktuelle News auch hier:



Online austauschen und kennenlernen: Mit unserer [#datatalk congress und meetup](#) Gruppe bieten wir eine Online-Plattform für den regelmäßigen fachlichen Austausch und Networking.

