

Quick scan data kwaliteit

Andre Bal

Toepassingsgebieden

Datakwaliteit is vaak een bepalende factor om structureel successen te kunnen boeken, denk hierbij aan:

- Inrichtingen van ERP omgevingen zoals b.v. ITSM, HRM, SAP
- Data migraties /consolidaties bij applicatie vervangingen (b.v. migratie van HPSM naar Service-Now, of opzetten van IAM omgeving)
- Complexe data model veranderingen en gerelateerde migraties in applicatie landschappen.
- Voorbereidingen op interne of externe audits (DNB, ECB FDA audits)
- Verificatie correcte werking interfaces vanuit ERP systemen naar andere applicaties

Use cases

- Mijn leidende ICT leverancier onderhoud mijn CMDB, ik wil meetbaar zien hoe goed dat gebeurt
- Wij moeten migreren van HP Service Manager naar Service-Now, daarbij komt dat deze verschillende datamodellen hebben. Ik wil zekerheid omtrent de kwaliteit van de data migratie
- Mijn leveranciers sturen mij detail facturatie data, hoe kan ik automatisch bepalen of deze wel klopt
- Onze in en uit dienst processen zijn slecht geregeld, ik wil de impact daarvan zichtbaar maken
- Wij willen een Single identity management system opzetten, hoe kan ik zichtbaar maken of dit consistent is, en aansluit op onze HR processen
- Wij verwachten een audit en ik wil vooraf aantoonbaar hebben waar onze potentiële problemen zitten

Doelen van een quickscan

- Bedrijfsbrede verificatie van belangrijkste data kwaliteit zoals:
 - Compleetheid over compleetheid
 - Correctheid / Consistentie
 - Voldoen aan specifieke Business rules en kwaliteit eisen
- Het op feiten gebaseerd, aantonen van de basale datakwaliteit om potentieel voor verbeter acties te identificeren.
- Potentieel langdurig en geautomatiseerd uitvoeren
 - Om verantwoordelijken in bronssystemen concrete en meetbare correcties te laten uitvoeren
 - Het structureel volgen en dashboards hoe deze kwaliteit zich ontwikkeld in de tijd

Wat behelst het?

- Aanleveren van relevante bron data (structured data)
- Verbanden bepalen tussen diverse bron gegevens
- Aanvullen van ontbrekende bron gegevens
- Vaststellen Business Rules
 - als in bron A... dan moet in bron B..
 - Velden moeten voldoen aan specifieke criteria
- Dashboard rapportage van feitelijke datakwaliteit
 - Wat is aantoonbaar goed
 - Wat is aantoonbaar fout
 - Welke concrete fouten moeten hersteld worden en waar.

Hoe werkt het technisch?

- Bron data wordt aangeleverd (met contact person per bron)
- Data wordt in een DB geladen
- Verbanden worden gezocht en gelegd
- Queries om de data kwaliteit meetbaar te maken worden gemaakt
- Dashboards om te rapporteren worden geautomatiseerd gemaakt
- Het geheel is automatisch herhaalbaar bij periodieke aanlevering van nieuwe bron data
- Op basis van de dashboards ontstaan inzicht in problemen en kunnen deze structureel in de bron systemen worden opgelost

Van Quickscan naar structurele scan

Quickscan toont aan in hoeverre een organisatie “In Control” is

Quickscan kan ook langduriger ingezet worden om tijdens verbetertrajecten of migraties op feiten via dashboarding de daadwerkelijk voorgang van verbetering te rapporteren (fact based reporting).

Dit kan op week / maandbasis en vindt geautomatiseerd plaats totdat de kwaliteit op aanvaardbaar niveau is.

Hoeveel werk is het?

- Is afhankelijk van het aantal bronnen (data sources) dat betrokken is
- Een basale quickscan kan veelal met een doorlooptijd van 4-6 weken en een capaciteit van 3-4 weken worden gerealiseerd.
- Resultaat zal bepalen of verdere uitbreiding van controles in de scan noodzakelijk is.
- Resultaat zal bepalen of inzet voor een langere periode noodzakelijk is
- Vaak kiest men voor een periode van 1 jaar waarbij 2-4 wekelijks de brondata wordt aangeleverd en dashboards worden gemaakt om de kwaliteit te volgen.
- Vaak kiest men ervoor om gedurende langduriger inzet steeds meer kwaliteit controles toe te voegen.

Hamvraag

- Vaak denkt men de datakwaliteit op orde te hebben
- Onze vraag:
 - Waarmee en hoe toon je dit concreet aan?
- Onze quickscan geeft concreet inzicht in de kwaliteit

Voorbeeld dashboard CMDB data

Signal message	Object Type	Object status	Hits	IB IBM managed	% hits on Install base	% Correct on Insta
ASQ00010 Asset with empty Computerroom	Asset*	In use	611	4218	14.486%	85.514%
ASQ00020 Asset with empty coordinates	Asset*	In use	115	4218	2.726%	97.274%
ASQ00030 Asset with empty serial number	Asset*	In use	3	4218	0.071%	99.929%
ASQ00040 Asset with empty asset tag	Asset*	In use	3	4218	0.071%	99.929%
CIQ00005 CIs without any relations	Asset*	In use	24	4218	0.569%	99.431%
CIQ00005 CIs without any relations	Middleware*	In use	95	17276	0.550%	99.450%
CIQ00005 CIs without any relations	Network printer	In use	13	4536	0.287%	99.713%
CIQ00005 CIs without any relations	Software CI	In use	28	628	4.459%	95.541%
CIQ00005 CIs without any relations	Storage CI	In use	43	493	8.722%	91.278%
CIQ00005 CIs without any relations	Workstation*	In use	19	37619	0.051%	99.949%
CIQ00010 CI without relation to business application	HMC	In use	6	69	8.696%	91.304%
CIQ00010 CI without relation to business application	Middleware*	In use	1000	17276	5.788%	94.212%
CIQ00010 CI without relation to business application	Server	In use	54	6741	0.801%	99.199%
CIQ00020 Server without relation to asset	Server	In use	66	6741	0.979%	99.021%
CIQ00040 Server without OS	Server	In use	100	6741	1.483%	98.517%
CIQ00045 Middleware without OS	Middleware*	In use	52	17276	0.301%	99.699%
CIQ00050 CI with duplicate serial number	Asset*	In use	43	4218	1.019%	98.981%
CIQ00050 CI with duplicate serial number	Network printer	In use	205	4536	4.519%	95.481%
CIQ00050 CI with duplicate serial number	Peripheral device	In use	82	2261	3.627%	96.373%
CIQ00050 CI with duplicate serial number	Workstation*	In use	1	37619	0.003%	99.997%

Voorbeeld CMDB kwaliteit details

CMDB fouten i.r.t. Data center management systeem

TBL CMDB Signals					
Object name	Object model	Object status	Relation sysid	Signal message	
4000300882-FDO1923EOMB	Cisco Catalyst 3650			W2W00300A AssetTagSerial not found in CMDB	
4000300883-FDO1923EOMF	Cisco Catalyst 3650			W2W00300A AssetTagSerial not found in CMDB	
4000601256-JAB050584B4	Cisco Catalyst 3650			W2W00300A AssetTagSerial not found in CMDB	
4000601671-FCZ16302057	Cisco Catalyst 3750G-12S-S WS-C3750G-12S-S			W2W00300A AssetTagSerial not found in CMDB	
4000601503-CAT0815N0D1	Cisco Catalyst 3750G-12S-S WS-C3750G-12S-S			W2W00300A AssetTagSerial not found in CMDB	
4000601251-FDO1132X0UW	Cisco Catalyst 3750G-12S-S WS-C3750G-12S-S			W2W00300A AssetTagSerial not found in CMDB	
4000601252-FOC1129Y0PK	Cisco Catalyst 3750G-12S-S WS-C3750G-12S-S			W2W00300A AssetTagSerial not found in CMDB	
4000601253-FOC1207Y10H	Cisco Catalyst 3750G-12S-S WS-C3750G-12S-S			W2W00300A AssetTagSerial not found in CMDB	
4000601254-FOC1207Y10V	Cisco Catalyst 3750G-12S-S WS-C3750G-12S-S			W2W00300A AssetTagSerial not found in CMDB	
4000601567-CAT0951Z0VH	Cisco Catalyst 3750G-12S-S WS-C3750G-12S-S			W2W00300A AssetTagSerial not found in CMDB	
4000602405-FCZ162720MZ	Cisco Catalyst 3750G-12S-S WS-C3750G-12S-S			W2W00300A AssetTagSerial not found in CMDB	
4000601064-FCZ162720N3	Cisco Catalyst 3750G-12S-S WS-C3750G-12S-S			W2W00300A AssetTagSerial not found in CMDB	
4000601514-CAT1023Z11T	Cisco Catalyst 3750G-24PS			W2W00300A AssetTagSerial not found in CMDB	
4000401667-CAT1023Z12W	Cisco Catalyst 3750G-24PS			W2W00300A AssetTagSerial not found in CMDB	
4000601515-FDO1234Y3JT	Cisco Catalyst 3750G-24PS			W2W00300A AssetTagSerial not found in CMDB	
4000601516-FDO1234Y3JT	Cisco Catalyst 3750G-24PS			W2W00300A AssetTagSerial not found in CMDB	