



Expertblog

De waarde van actief data kwaliteit beleid

Geschreven door
Rinus Broekmans

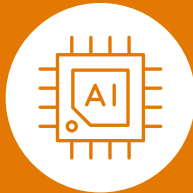
17-02-2023





Introductie

Data wordt door de meeste organisaties gezien als een belangrijke grondstof. Het wordt vaak zelfs de meest waardevolle asset genoemd. Toch blijkt uit onderzoek dat 50% van de organisaties geen actief beleid voert op het gebied van datakwaliteit. Met alle gevolgen van dien: veel data issues, hoge kosten, proces verstoringen en ontevreden klanten. In deze blog bespreken we wat de waarde is van een actief databeleid en geven we tips uit de praktijk van organisaties die hier al mee werken.





Wat is data quality management (DQM)?

Data quality management (DQM) is in feite de operationele uitvoering van de visie en het beleid op het gebied van data.

Wat zijn de data definities?

Wie heeft welke rol in het verwerven en verwerken van data?

Hoe wordt de inhoudelijke kwaliteit van data geborgd?

Wat wordt er gedaan als de kwaliteit van data niet voldoet aan de verwachtingen?

Organisaties hebben vaak in het beleid opgenomen dat ze meer data-driven willen worden. Ze willen hiervoor graag tools als data analytics en AI inzetten. Dit heeft alleen zin als de basis in order is: de data moet beschikbaar, betrouwbaar, bruikbaar en bij-de-tijd zijn. Actief DQM is hiervoor onmisbaar.





Reactief data kwaliteit beheer

Bijna elke organisatie onderkent dat goede data een basisvoorziening is. Toch weten wij dat bij het overgrote deel van de organisaties problemen zijn met de beschikbaarheid van data. Tegelijkertijd heeft bijna de helft ook regelmatig het tegenovergestelde probleem, dat data redundant is en daardoor dus niet altijd goed bruikbaar is.

In de praktijk wordt dan ook veel tijd en geld geïnvesteerd aan het oplossen van data quality (DQ) issues. Zo zijn data analisten vaak meer dan de helft van hun tijd kwijt aan het opschonen en structureren van de data waarmee ze moeten werken. In de operatie kunnen DQ issues tot verborgen kosten leiden, niet alleen direct voor dubbel werk in databeheer of het opsporen en verbeteren van foute data. De indirecte kosten voor stilstanden in de primaire processen, verlies van klanten en imagoschade kunnen nog veel hoger zijn. De werkelijke kosten voor DQ issues zijn dus moeilijk in te schatten.

Metten is weten

Het is dus kostbaar om data reactief te corrigeren. Een tip uit de praktijk is daarom om een verbetercyclus voor data correcties in te richten. Registreer DQ issues op een manier dat past bij jouw organisatie. Leg de (directe) kosten voor analyse en herstel vast en noteer de oplossing. Zo ontstaat inzicht in de kosten en kun je terugvallen op vastgelegde kennis en ervaring.



Actief data kwaliteit beheer

Meeste van de bedrijven die wij bezoeken tijdens onze consults hebben een actief beleid op het gebied van data kwaliteit. Deze organisaties maken hun medewerkers duidelijk wat het belang is van goede data, de verantwoordelijkheid hiervoor wordt eenduidig belegd en medewerkers worden hierop ook getraind en aangesproken. Zo voorkom je dat de mooie functie- en procesbeschrijvingen een papieren tijger worden. De meeste organisaties besteden al veel aandacht aan de invoer van data. Gegevens worden zo dicht mogelijk bij de bron ingevoerd en gecontroleerd. Consistentie van data wordt gewaarborgd door geautomatiseerde validatie en verwerking van de gegevens. Dit zijn de belangrijkste maatregelen die meestal wel genomen worden om de datakwaliteit te waarborgen.

Data bewuste organisaties controleren daarnaast procesmatig de kwaliteit van de opgeslagen gegevens, bijvoorbeeld met periodieke rapportages of een (near) real-time DQ dashboard. Zij herkennen issues sneller, kunnen deze vaak zonder tijdsdruk analyseren en oplossen en kunnen extra kosten door verstoringen in het primaire bedrijfsproces vermijden. De beste resultaten worden hierbij geboekt als data uit meerdere bronnen in een aparte data analytics omgeving (datalake) gecombineerd kunnen worden. Zo kunnen inconsistenties tussen systemen gevonden worden en worden operationele systemen niet belast. Data bewuste organisaties realiseren zich ook dat eenmaal ingevoerde data meerdere gebruiksdoelen kan dienen. Zij voeren actief beleid om data goed te kunnen delen, bijvoorbeeld door het vastleggen en delen van heldere en bruikbare datadefinities. Voor organisaties die als beleid hebben om meer data-driven te worden, is het noodzaak om data te delen en zij zullen actie moeten ondernemen om dit te faciliteren. Tot slot is er in de meeste gevallen nog veel te winnen door er voor te zorgen dat de data ook bij-de-tijd blijft. Archiveren en verwijderen van data is nog wel eens een beetje een ondergeschoven kindje. Het is niet alleen fijn voor de dagelijkse werkzaamheden en voor data analytics dat de dataset up-to-date is. In het kader van wet- en regelgeving is het ook belangrijk dat dit goed geregeld is. Denk bijvoorbeeld aan de AVG en "het recht om vergeten te worden".





De waarde van actief DQM

Data wordt door organisaties vrij algemeen gezien als het nieuwe goud. Het is duidelijk dat voor het waarborgen van een hoge data kwaliteit er significante kosten gemaakt worden. De verborgen kosten voor matige data kwaliteit kunnen nog veel hoger zijn. Voor organisaties met een actief data kwaliteitsbeheer is kostenbeheersing dan ook niet de belangrijkste drijfveer. Zij investeren vooral in de waarde van goed beschikbare, betrouwbare en actuele data, zodat de gegevens intern én extern met een gerust hart gedeeld kunnen worden. Dit betaalt zich uiteindelijk uit in minder verstoringen in operationele bedrijfsprocessen en een betere ervaring voor de klant.



Wilt u uw data kwaliteit verhogen?

Neem dan contact op met de experts van Squadra voor een nadere kennismaking.