

De datakant van het 'IT-Deltaplan voor de overheid'

Jan Klaver (27 februari 2020)

Het centraliseren van de data met een constant wijzigende informatiebehoefte is een idee dat wat mij betreft deels achterhaald is.

Persoonlijk zie ik meer in een opzet van informatiearchitectuur per datadomein, dat lijkt meer realistisch qua haalbaarheid.

Daarbij zouden de datadomeinen (ordegrootte 10-20) wel centraal gekozen en gedefinieerd dienen te zijn.

Voor een datadomein dient een accountable rol (Data Owner) toegewezen te worden. Verder zullen er de nodige standaarden voor de ontwikkeling van datadomeinen afgesproken moeten worden (Denk aan naamgeving, afkortingen, data kwaliteitsstandaarden en data design patronen).

Wat in het IT-Deltaplan niet naar voren komt, is hoe je van huidige situatie naar toekomstige situatie "gecentraliseerde data" denkt te komen. Dat kon nog wel eens een hele uitdaging blijken te zijn.

Een stap van AS-IS naar TO-BE per datadomein zou een stuk haalbaarder kunnen blijken te zijn. Blijft staan hoe je software migratie naar het nieuwe landschap uit zou moeten voeren.

Over raakvlakken tussen datadomeinen dienen bij ontwikkeling afgestemd te blijven. Daar ligt een zeer belangrijke coördinerende taak voor iedere Data Owner. Hierbij geldt wel hoe meer domeinen hoe meer raakvlakken afgestemd dienen te blijven, dus ergens tussen 10 en 20 moet een optimum gevonden worden.

Ik doe hier geen uitspraken over hoe/waar de data opgeslagen dient te worden, maar ik vind persoonlijk wel dat jouw statement over een Nederlandse Cloud mogelijk onnodig restrictief is, waar een Europese Cloud al meer mogelijkheden biedt en de kans op een afwijkend wettelijk kader beperkt is.

Informatiebehoeften die over domeinen heen gaat en waarbij de data niet in dezelfde dataplatformen terug te vinden is, kan tegenwoordig ook eventueel met data virtualisatietechnieken worden opgelost.

Verder vind ik dat er rekening mee gehouden moet worden dat data pas van toegevoegde waarde is als het eenvoudig kan worden geconsumeerd toegesneden op gebruiksdoelen.

Daar zit ook een mogelijke strijdigheid met datacentralisatie. De gebruiksdoelen zijn veelal aan wijziging onderhevig, terwijl de onderliggende data niet noodzakelijk van structuur hoeft te veranderen, hooguit verrijkt mits het basisdesign goed is opgezet, daar komen de eerder genoemde data design patronen om de hoek kijken.

Ook voor de genoemde gebruiksdoelen biedt data virtualisatietechniek slimme en snelle design-mogelijkheden om data gebruiksgereed en relatief snel aan te passen voor zulke nieuwe requirements zonder de onderliggende data structuren iedere keer overhoop te moeten halen.