Vlaamse Toezichtcommissie
voor het
elektronische bestuurlijke gegevensverkeer

4 november 2015

Principes voor het opstarten, gebruiken, onderhouden en stopzetten van een datawarehouse met persoonsgegevens.

Het begrip datawarehouse (afgekort DWH): een DWH is een techniek van gegevensverzameling die in een dusdanige vorm is gebracht dat terugkerende en ad-hoc vragen in relatief korte tijd beantwoord kunnen worden zonder dat de operationele bronsystemen zelf daardoor overmatig belast worden. De betreffende gegevens zijn afkomstig van en worden op geautomatiseerde wijze onttrokken aan de operationele bronsystemen. Gegevens kunnen in het DWH niet worden ingevoerd of aangepast door gebruikers zelf.

Een DWH wordt typisch gebruikt voor:

a. Management: beleidsevaluatie en beleidsvoorbereiding
b. Ondersteunen wetenschappelijk onderzoek
c. Informatieverlenende functie die elke overheidsdienst te vervullen heeft

Deze nota betreft niet fraudeopsporing of repressie: het DWH mag niet leiden tot individuele beslissingen.

Operationele bronsystemen (databanken) zijn niet het meest geschikt voor terugkerende en ad-hoc ondervragingen

- de gegevens uit operationele bronsystemen worden gekopieerd op een apart computersysteem via aangepaste software en in een gecontroleerde omgeving.
- de databanken, waaruit de gegevens gehaald worden die men wil koppelen, zijn operationele bronsystemen waarbij de historiek niet (voldoende) wordt bijgehouden.
- de operationele bronsystemen worden niet ontwikkeld voor het uitvoeren van statistische verwerkingen met grote hoeveelheden informatie. De organisatie van het operationeel bronsysteem is gericht op de efficiëntie van de uit te voeren operationele taken.
- door de “historisatie” van gegevens (opslag DWH op ijkmomenten) kan opnieuw teruggegrepen worden naar ervaringen uit het verleden.

De gegevens die in het DWH worden opgenomen:

- komen zoveel mogelijk uit authentieke bronnen.
- zijn enkel deze gegevens die nodig voor de gestelde doeleinden van het DWH: niet meer en niet meer gedetailleerd dan nodig (zoveel mogelijk in klassen).

De entiteit moet altijd alle verplichtingen nakomen inzake gegevensbescherming, ook bij opname van gegevens van de entiteit in het DWH

De entiteit blijft altijd ultiem aansprakelijk voor wat in het DWH wordt opgenomen en wat er met het DWH gebeurt, ook als dit DWH bij een externe partij staat.

De entiteit duidt minstens één DWH beheerder aan

- de entiteit zal minstens één interne medewerker of één externe partij formeel aanduiden als DWH beheerder en dit tijdig communiceren aan alle relevante betrokken partijen.
- de DWHbeheerder zal een deontologische code en een confidentialiteitsverklaring onderschrijven.

- de DWHbeheerder zorgt voor de ontwikkeling en analyse van de DWHbasistoepassing(en) en coördineert de kwaliteit en de planning, het opladen, het bewaren, het onderhouden en het archiveren of vernietigen van de gegevens in een DWH.
- de DWHbeheerder verzorgt ook een regelmatige rapportering naar alle relevante betrokken partijen.

De minimale veiligheidsnormen worden gerespecteerd, ook voor een DWH:

Het DWH wordt minstens even veilig behandeld als de operationele bronsystemen

De gegevens in het DWH worden beveiligd op basis van een risico analyse opgesteld door het DWH beheerder en formeel gevalideerd door het management van de entiteit.

Codering (encryptie/versleuteling) van gegevens in het DWH gebeurt

- volgens het principe van functiescheiding: een andere entiteit dan die van het management, onderzoekers of de operationele entiteiten
- voor de opname van de gegevens in een DWH of erna vooraleer de gegevens aan de DWH gebruikers worden bezorgd.
- in principe door een "trusted third party" (TTP).
- belangrijke keuzemogelijkheid: anonimisatie van de gegevens in het DWH

Toegang tot het DWH wordt gecontroleerd, beperkt en beveiligd

- de DWH systemen werken met authenticatie (toegang) en autorisatie (rollen / machtigingen)
- de DWH zijn bereikbaar voor een zo klein aantal personen als werkbaar (inhoudelijk en technisch[i])
- de entiteit zal regelmatig (minstens jaarlijks) een formele controle op de toegang tot het DWH laten uitvoeren en rapporteren aan het management van de entiteit.

Het vier-ogen principe wordt gehanteerd voor persoonsgegevens in het DWH

- de DWH gebruikers mogen geen toegang hebben tot de niet-gecodeerde gegevens: er moet een scheiding ingebouwd zijn: een "trusted third party" (TTP) of een persoonii met voldoende kennis van de materie oefent een controle uit op de proportionaliteitiii van de gevraagde gegevens en waakt over de degelijkheid van de codering.
- om anonimiteit te garanderen moet een "small cell size" bepaald zijn voor DWH analyses:
    - "hoe groot moet het resultaat minimaal zijn van een DWH vraagstelling om het resultaat door te geven aan de DWH gebruiker(s)?"
    - de standaard is minstens 5 personen in het resultaat.

Transparantie rond het DWH geschiedt via continue monitoring[iv] en tracering[v] op alle relevante activiteiten (inhoudelijk en technisch) in het DWH

Er bestaat een formele, gevalideerde procedure voor DWH incidenten (inclusief escalatie, communicatie en rapportering van "lessons learned") die regelmatig getest en effectief uitgevoerd wordt.

Bij het stopzetten van het DWH zal de DWH beheerder dit direct melden aan alle relevante betrokken partijen en zal de correcte en veilige archivering en vernietiging van het DWH geschieden.

---

[i] De beheerder van de infrastructuur mag slechts een beperkte en gecontroleerde toegang hebben tot de gegevens.
[ii] Dit is geen taak van de informatieveiligheidsconsulent, maar van een (onafhankelijke) operationele medewerker.
[iii] Bijvoorbeeld de populatie tot of van een steekproef beperken.
[iv] Methode waarbij alle activiteiten worden opgevolgd in de DWH : wie doet wat wanneer op welke gegevens en welke systemen.
[v] Methode waarbij actief gezocht wordt naar het verloop van activiteiten door diverse informatiebronnen te combineren (via vragen / queries).