

Data Warehouses en Business Intelligence

code/duur/doelgroep/leerdoelen

code
DWBI

duur
2 dagen

doelgroep

iedereen die bij de bouw, het gebruik of het beheer van een Data Warehouse en bij het toepassen van Business Intelligence betrokken is of gaat worden, zoals projectleiders, ontwikkelaars, beheerders, programmeurs en gebruikers. Daarnaast is de workshop ook bedoeld voor informatieanalisten, database ontwerpers en gegevensbeheerders.

leerdoelen

De deelnemer is na de workshop in staat aan te geven:

- wat een Data Warehouse is en hoe een organisatie profijt kan trekken uit het hebben van een Data Warehouse door middel van Business Intelligence
- wat Business Intelligence is en hoe Business Intelligence zich verhoudt met querying, OLAP, data mining, EIS, statistical processing
- welke punten een rol spelen bij het opzetten, gebruiken en beheren van een Data Warehouse
- welke problemen overwonnen moeten worden bij het opzetten en gebruiken van een Data Warehouse
- welke plaats een Data Warehouse inneemt tussen andere in een organisatie voorkomende informatiesystemen
- welke technologie ingezet wordt bij het opzetten en gebruiken van een Data Warehouse en bij het toepassen van Business Intelligence

onderwerpen

In de workshop komen de volgende onderwerpen aan de orde:

- wat is een Data Warehouse
- wat is Business Intelligence
- hoe vul je de Data Warehouse (download bronsystemen en upload in de data warehouse)
- soorten bronnen
- waar zitten moeilijkheden
- technologieaspecten
- aansluiting op de database problematiek van de organisatie
- multidimensionele databases
- ontwerpaspecten (niet alleen van de Data Warehouse zelf, maar ook van bijvoorbeeld multidimensionele databases)
- op wat voor manieren bewaar je de gegevens (oa. aggregatieniveau's, technische problemen)

- externe gegevens (gegevens die ingekocht worden van derden)
- aansluiting met bronsystemen (o.a. problemen met onvergelykbaarheid van gegevens over de jaren heen)
- hoe gebruiken we de inhoud naar toepassingen toe (opdelen/ data marts/ waarom maar een deel van de Data Warehouse in een data mart en waarom zitten data marts op aparte machines)
- Business Intelligence: analyse op de gegevens in de Data Warehouse (bijvoorbeeld querying, data mining, on line analytical processing / OLAP, statistical processing)
- soorten eindgebruikers
- waarvoor worden resultaten gebruikt
- risico's van verkeerd gebruik resultaten
- beveiligen inhoud van de Data Warehouse
- beheren van het data model van de Data Warehouse en het beheren van de Data Warehouse zelf
- spelers rondom de Data Warehouse en Business Intelligence (mensen die de Data Warehouse maken, beheren of gebruiken)
- eindgebruiker tools (vb. allerlei grafische mogelijkheden, maar kun je daar ook wat mee)

algemeen

tweedehands gegevens

Organisaties vergaren en gebruiken grote hoeveelheden gegevens. Dit gebeurt met allerlei (operationele) informatiesystemen door elkaar. De gegevens worden, wanneer ze niet meer de actuele situatie van de organisatie beschrijven, meestal weggegooid. Dat is zonde. Een organisatie kan veel nut hebben van deze gegevens. Zo kunnen gegevens die volledig afgehandelde bestellingen beschrijven inzicht verschaffen in het koopgedrag en het betaalgedrag van de klanten. De gegevens laten ook zien hoe lang klanten al klant zijn en of ze steeds meer gaan kopen of juist niet. Voor marketing is zulke informatie van onschatbare waarde. De gegevens kunnen ook inzicht verschaffen in de te verwachten verkopen in een bepaalde periode. Op basis daarvan kan productiecapaciteit worden ingepland. Op basis daarvan kan ook besloten worden of de productiecapaciteit uitgebreid moet worden. Voor een goed inzicht moeten de gegevens van (het liefst) een groot aantal jaren (bijvoorbeeld 10 jaren) worden bewaard. Dit geeft de mogelijkheid om het verloop van allerlei indicatoren (zoals KPI's: key performance indicators) te volgen over een langere termijn. We kunnen ze zelfs achteraf bepalen.

enorme hoeveelheden gegevens

Gegevens worden echter niet voor niets weggegooid. Zouden we de gegevens niet verwijderen wanneer ze niet meer gebruikt worden in de informatiesystemen, dan

Data Warehouses en Business Intelligence

loopt de ruimte die deze informatiesystemen hebben langzaam maar zeker vol. Daarnaast kunnen deze informatiesystemen de genoemde vragen niet beantwoorden, indicatoren niet bepalen en analyses niet uitvoeren. Een ander probleem is dat voor dit alles vaak de gegevens uit meer dan één informatiesysteem nodig zijn.

Kortom, er is behoefte aan wat we Business Intelligence noemen. Met Business Intelligence onderzoeken we:

- wat er gebeurd is in het verleden (door querying)
- waarom het gebeurd is (door OLAP: on line analytical processing)
- of we interessante verbanden in de gegevens kunnen vinden en of deze verbanden trends bevatten waardoor we mogelijk de toekomst voorspelbaar maken (data mining)
- of we voor management interessante gegevens/parameters kunnen vinden (EIS: executive information systems)

Data Warehouse

Business Intelligence kunnen we meestal niet toepassen op basis van wat in de databases van de operationele systemen is opgeslagen. De gegevens zijn niet geïntegreerd en gaan niet ver genoeg terug in het verleden. De oplossing voor deze problemen is het opzetten van een Data Warehouse. Hierin bewaren we de gegevens van vele jaren (tot wel zo'n 10 jaar) en maken we deze gegevens geschikt voor gebruik. De Data Warehouse voeden we met de gegevens van vaak vele verschillende informatiesystemen (bronnen genoemd), soms zelfs van buiten de organisatie zelf (externe bronnen).

problemen

Het samenbrengen van gegevens in de Data Warehouse levert allerlei problemen op. Zo vormen de in de Data Warehouse vergaarde gegevens een berg waar met

gewone databasetechnologie niet goed mee om te gaan is. Daarnaast zijn gegevens niet altijd goed vergelijkbaar. In de gegevens uit het ene informatiesysteem worden werknemers ingedeeld in drie soorten, terwijl een ander informatiesysteem met tien soorten werknemers werkt. Voorraden worden in het ene informatiesysteem op basis van inkoop prijs bijgehouden, terwijl een ander informatiesysteem de voorraden bijhoudt op basis van de verkoopprijs. Ook het bijhouden van gegevens uit vele jaren levert problemen op. Hoe moeten bijvoorbeeld gegevens van voor en na een fusie met elkaar worden vergeleken?

Een geheel ander probleem is dat het meestal niet doenlijk is de hele historie van wijzigingen van gegevens over een bepaald iets vast te leggen. Vaak wordt er voor gekozen met geregelde tussenpozen de gegevens over de op dat moment geldende situatie vast te leggen. Dit gaat goed voor het vastleggen van gegevens over werknemers, maar dit kan bij het vastleggen van verkopen leiden tot het weglaten van verkopen of dubbel vastleggen van verkopen. Wat precies kan en moet is iets wat bij het ontwerp van de Data Warehouse heel nauwkeurig uitgezocht wordt.

de workshop

In de workshop besteden we ruime aandacht aan wat een Data Warehouse is, hoe we een Data Warehouse moeten opzetten en hoe we Business Intelligence toepassen. De hiervoor genoemde problemen komen breed aan bod. Het ontwerp, de technologieaspecten, het gebruik door middel van data marts, multidimensionele databases, het downloaden van bronsystemen, het uploaden naar de Data Warehouse en het toepassen van Business Intelligence behandelen we aan de hand van een uitgebreide casus. De deelnemer maakt hierdoor kennis met de Data Warehouse en Business Intelligence aan de hand van een praktijksituatie.