

Informationssystem 4

Del 2: Beslutstödssystem, inklusive data warehousing och data mining

Erik Perjons

DSV, Stockholms Universitet



IT-stöd för beslutsfattande/analys

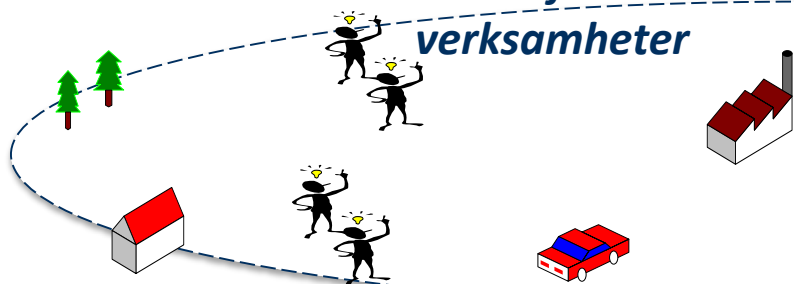


Beslut i verksamheter



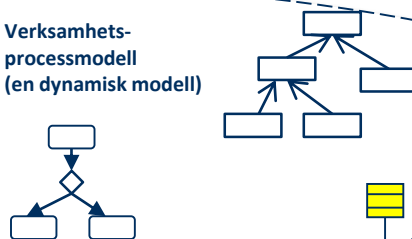
Stockholms
universitet

*Beslut fattas i
verksamheter*



Målmodell (en statisk modell)

Verksamhets-
processmodell
(en dynamisk modell)



Konceptuell
modell över
verksamhetens
termer/begrepp
(en statisk modell)

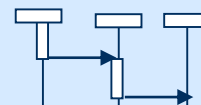


Systemfunktioner:

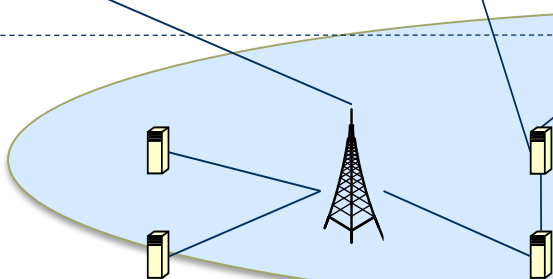
- Söka kund
- Lägga beställning



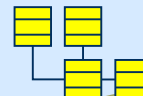
Interaktion
mellan mjukvaruobjekt



*IT/IS-stöd för
beslut/analys*



Modell över
klasser av
mjukvaruobjekt



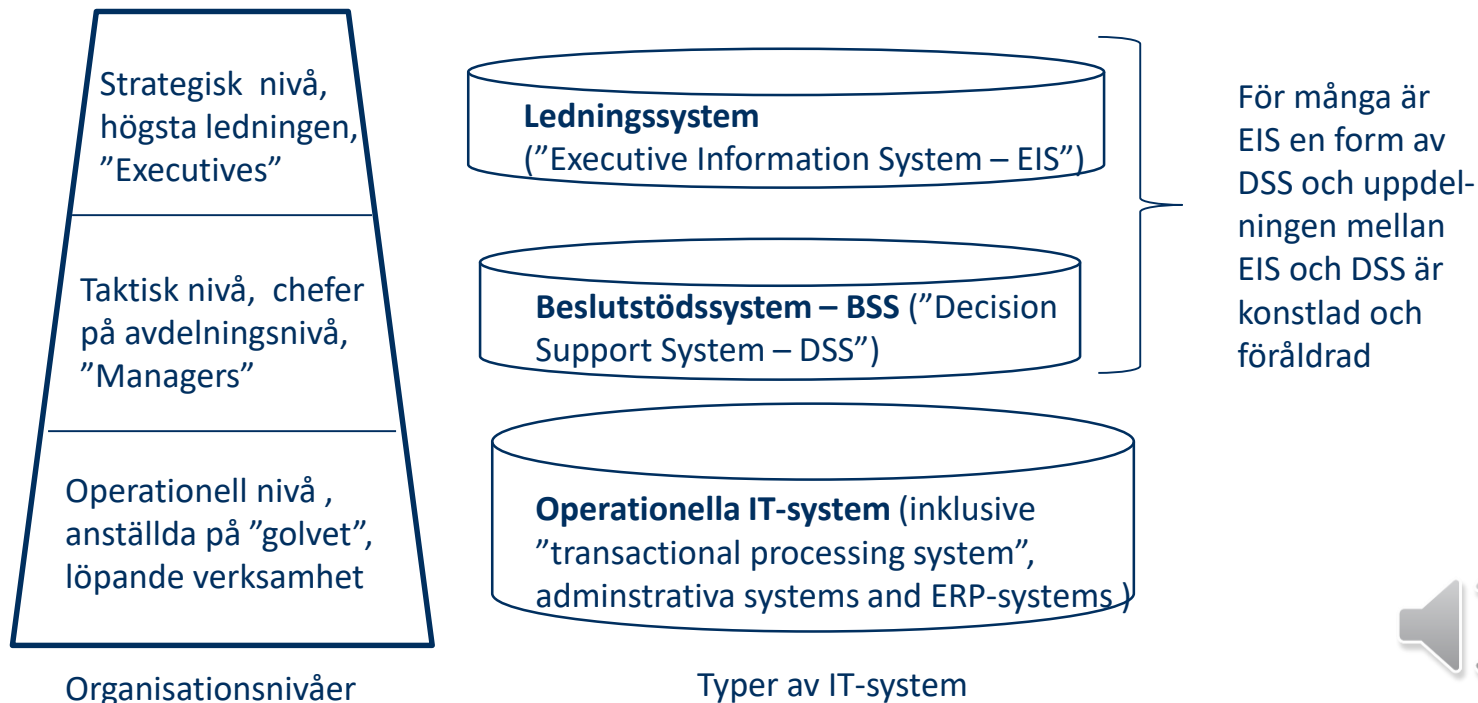
Verkligheten

Grafiska modeller/diagram



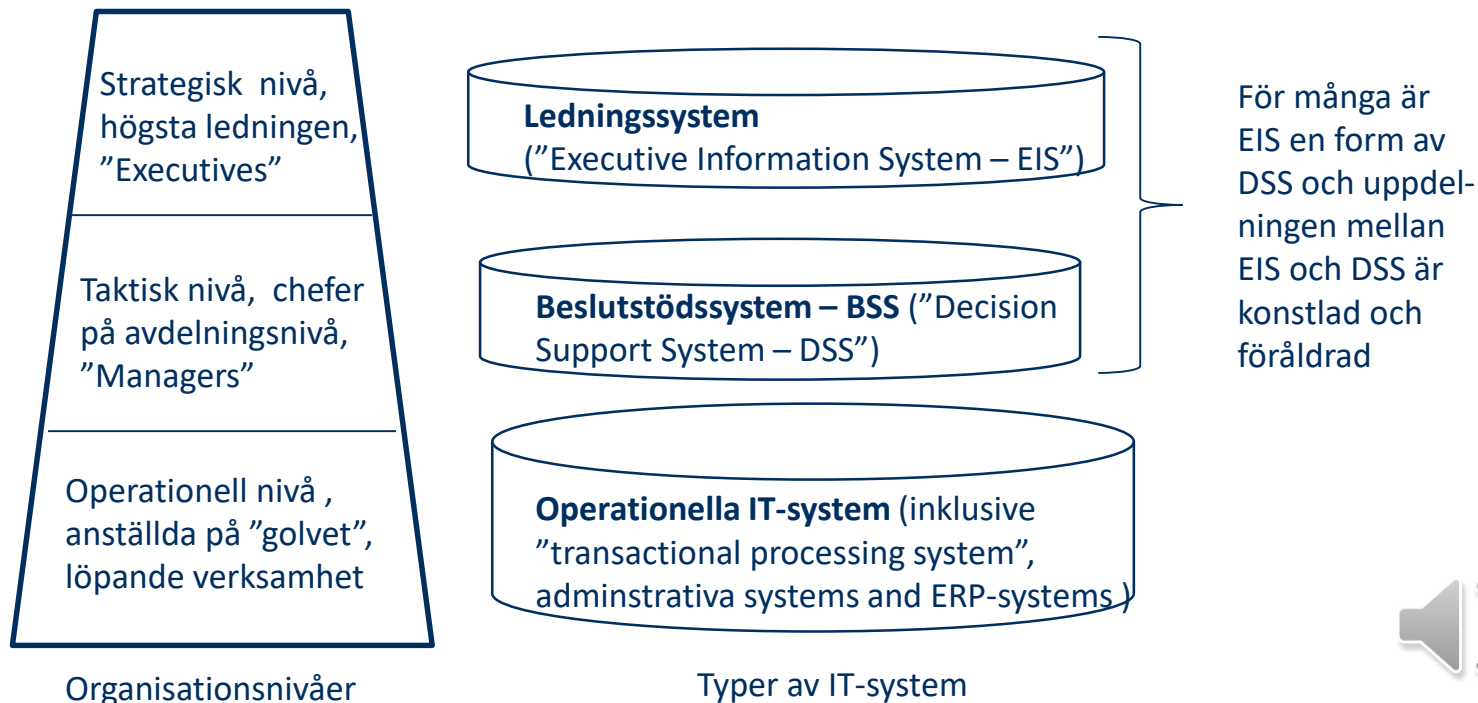
IT-system för beslut på olika nivåer

- Notera att beslut på en organisationsnivå kan stödjas av alla typer av IT – men ofta är en typ av IT-system central på en nivå



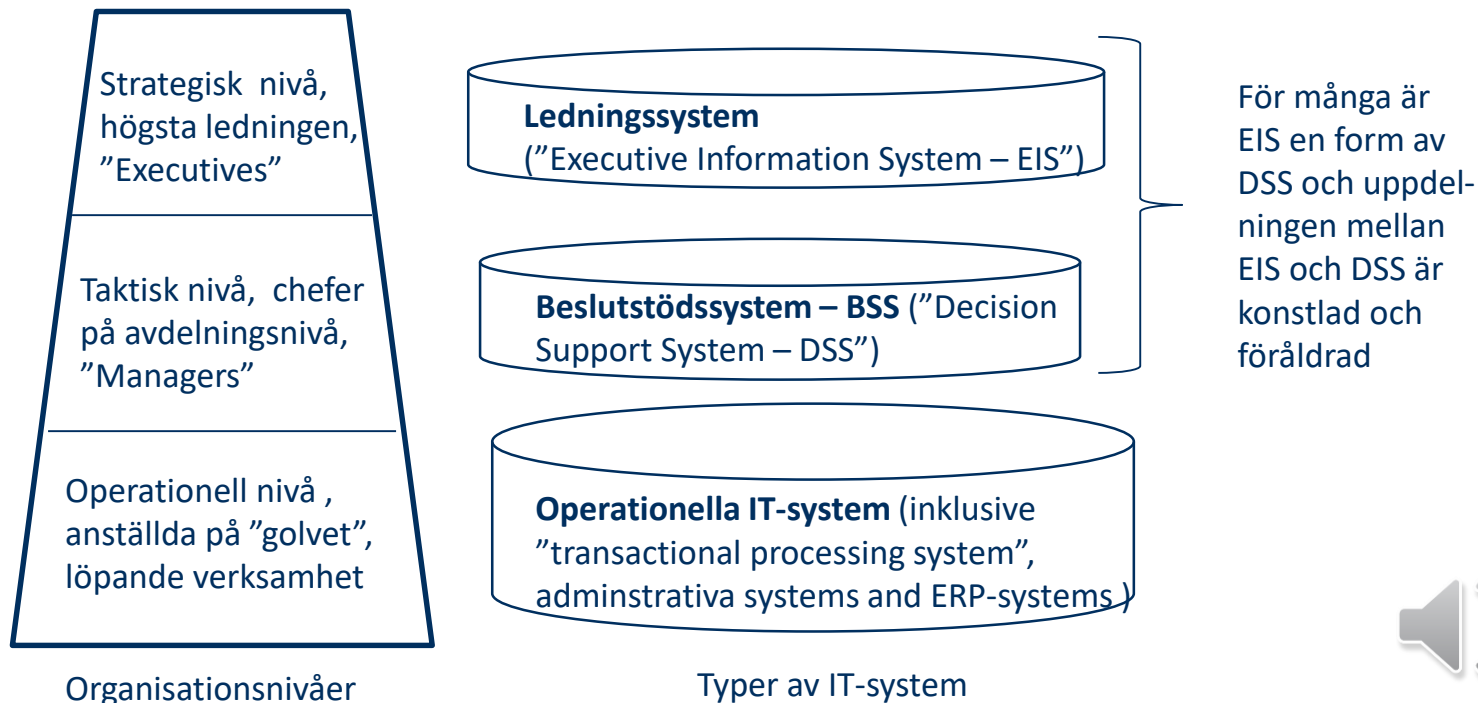
IT-system för beslut på olika nivåer

- Notera att beslut på en organisationsnivå kan stödjas av alla typer av IT – men ofta är en typ av IT-system central på en nivå

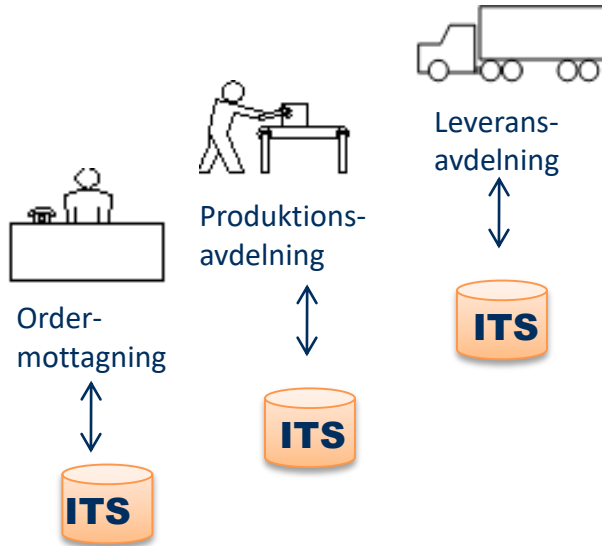


IT-system för beslut på olika nivåer

- Notera att beslut på en organisationsnivå kan stödjas av alla typer av IT – men ofta är en typ av IT-system central på en nivå



Typer av IT som stöd för beslut

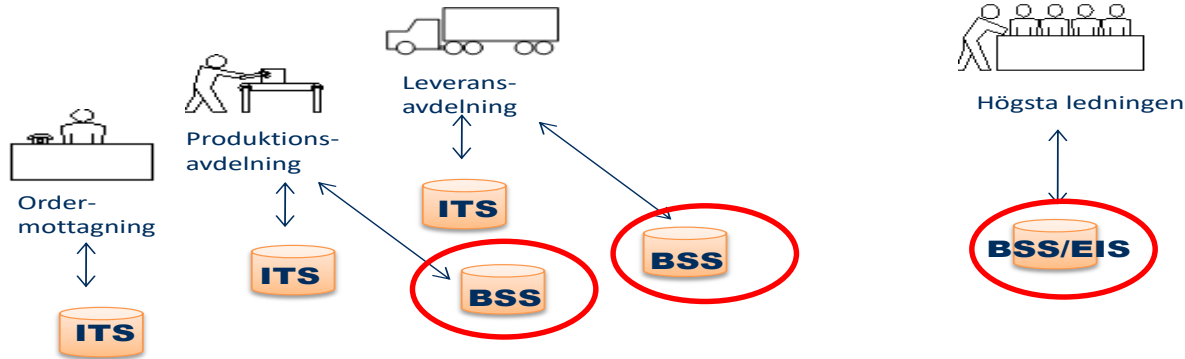


Operationella IT-system (ITS) – stödjer bland annat

dagliga, ofta rutinmässiga, beslut. En del enkla beslut kan **automatiseras**, det vill säga ITS kan mer eller mindre automatiskt ta beslut baserat på viss information.



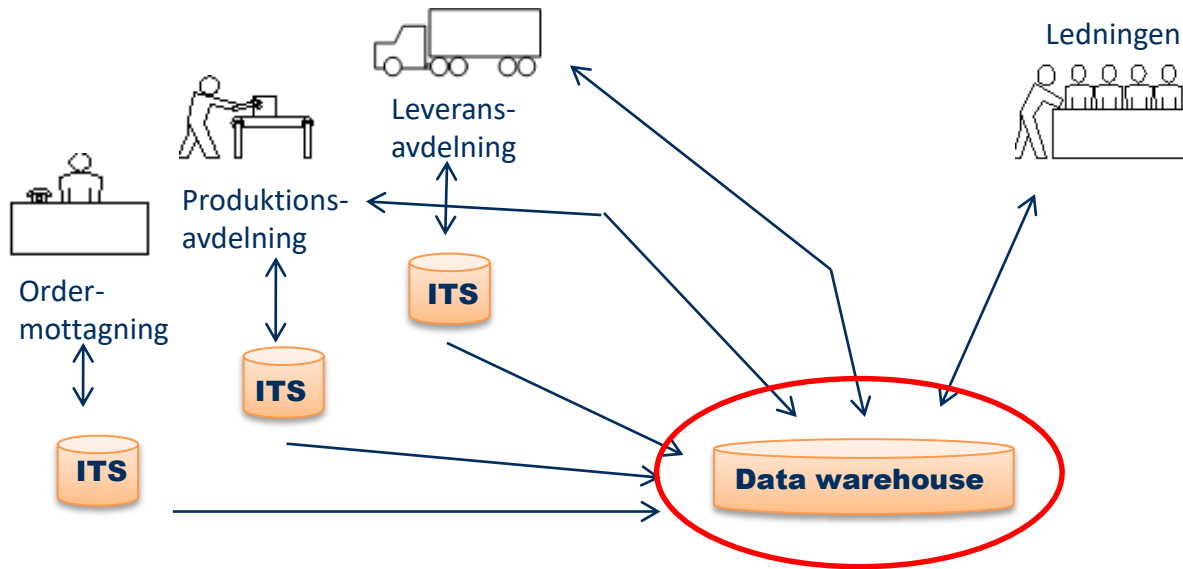
BSS: Typer av IT som stöd för beslut



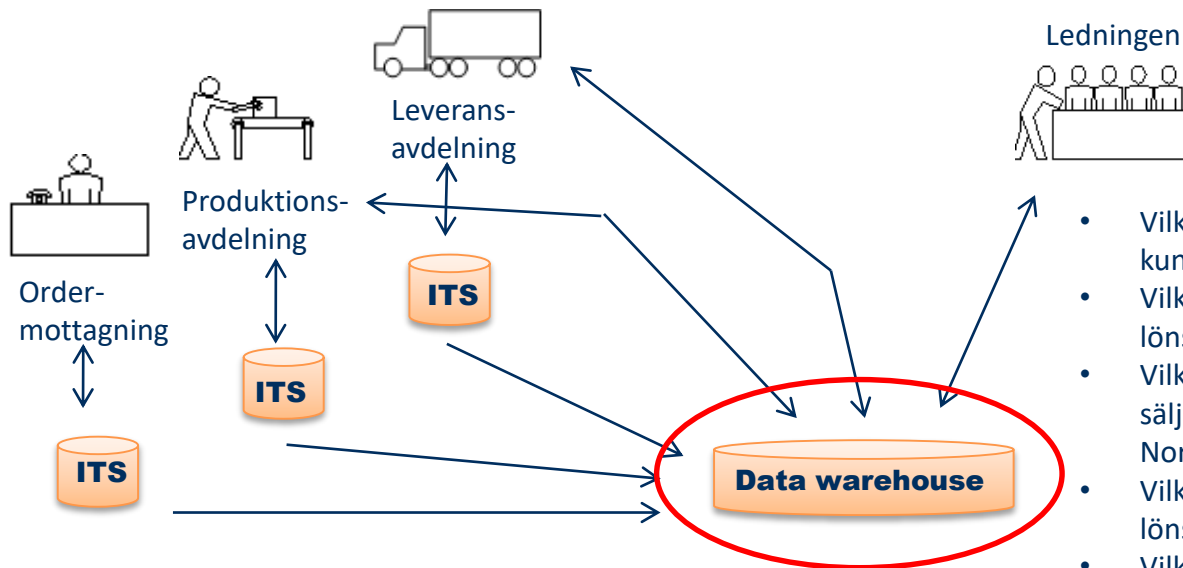
Beslutstödssystem (BSS) – är system utvecklade för
att mer specifikt stödja beslutsfattandet



Data warehouse-arkitektur 1



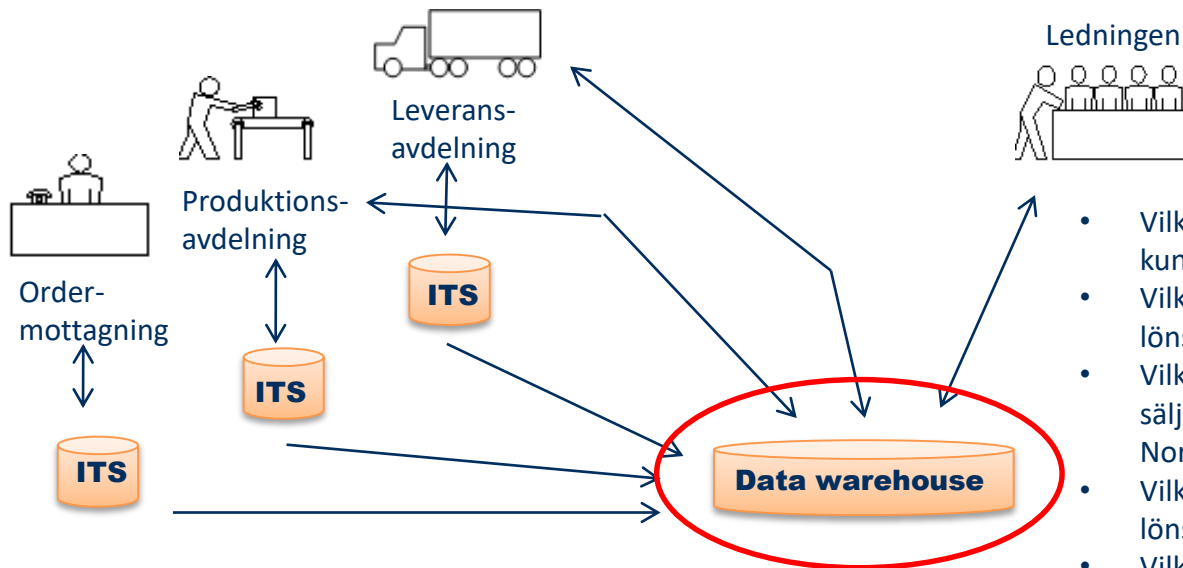
Data warehouse-arkitektur 1



- Vilka är våra mest lönsamma kunder?
- Vilka produkter köper de lönsamma kunderna mest av?
- Vilka produkter och tjänster säljs det mest och minst av i Norge?
- Vilka kontor i Sverige är mest lönsamma?
- Vilka produkter säljs inte alls under vintern i Stockholm?



Data warehouse-arkitektur 1



För att få svar på dessa frågor måste masterdata och i synnerhet transaktionsdata finnas nedladdade i DW

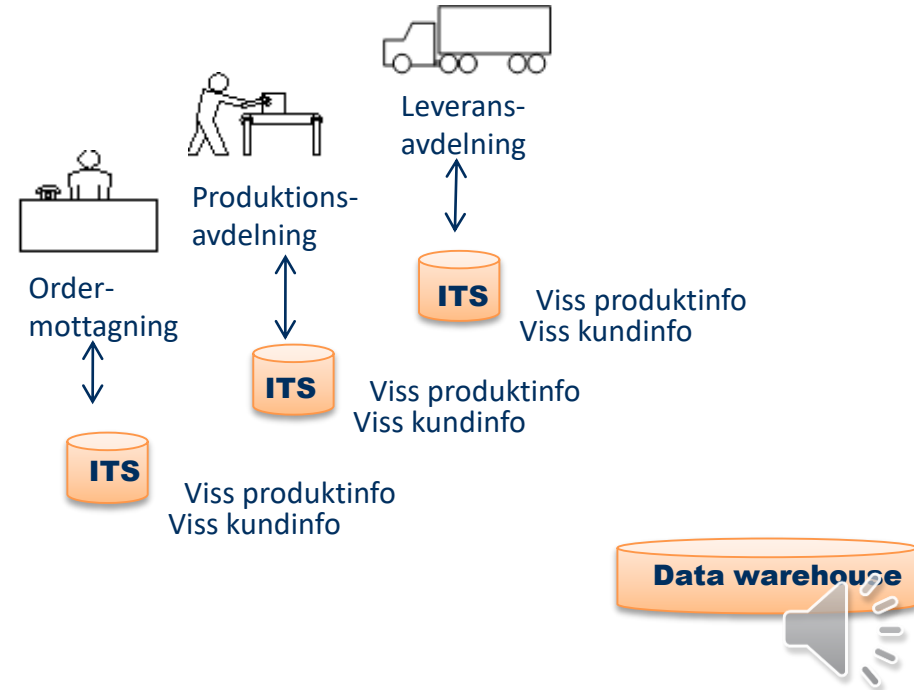


- Vilka är våra mest lönsamma kunder?
- Vilka produkter köper de lönsamma kunderna mest av?
- Vilka produkter och tjänster säljs det mest och minst av i Norge?
- Vilka kontor i Sverige är mest lönsamma?
- Vilka produkter säljs inte alls under vintern i Stockholm?



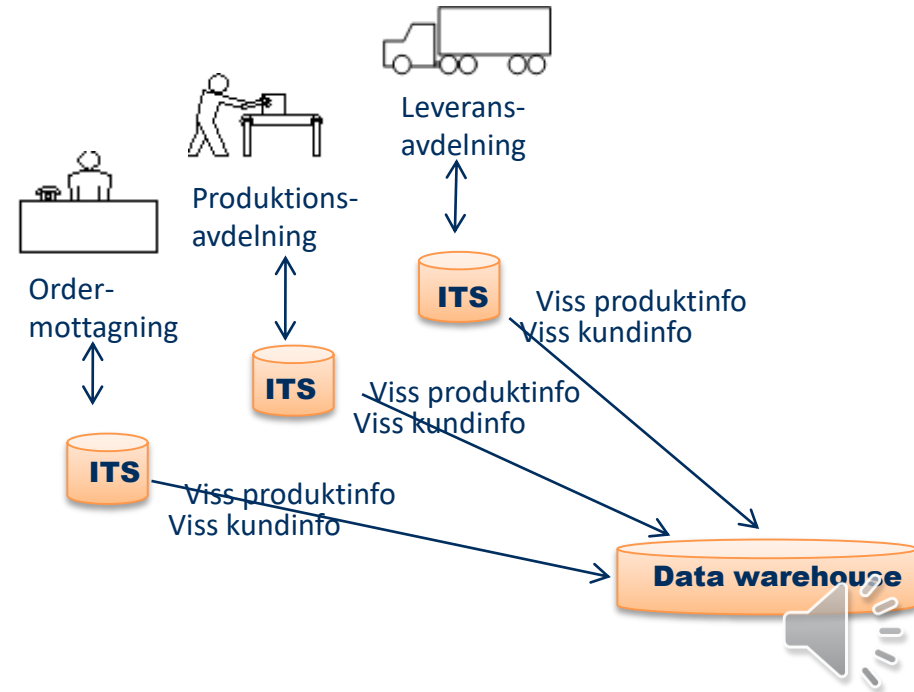
Varför DW?

- Data om produkter, kunder, leverantörer - och relaterad transaktionsdata - kan finnas utspridd i flera system, vilket inte möjliggör en helhetsbild av verksamheten, vilket i sin tur försvårar beslutsfattande
- De operationella systemen är inte uppbyggda för att snabbt besvara frågor som ska stödja beslutsfattande



Varför DW?

- Data om produkter, kunder, leverantörer - och relaterad transaktionsdata - kan finnas utspridd i flera system, vilket inte möjliggör en helhetsbild av verksamheten, vilket i sin tur försvårar beslutsfattande
- De operationella systemen är inte uppbyggda för att snabbt besvara frågor som ska stödja beslutsfattande
- **Bättre då att ladda ned denna information i DW som är uppbyggt för att svara på sådana frågor**



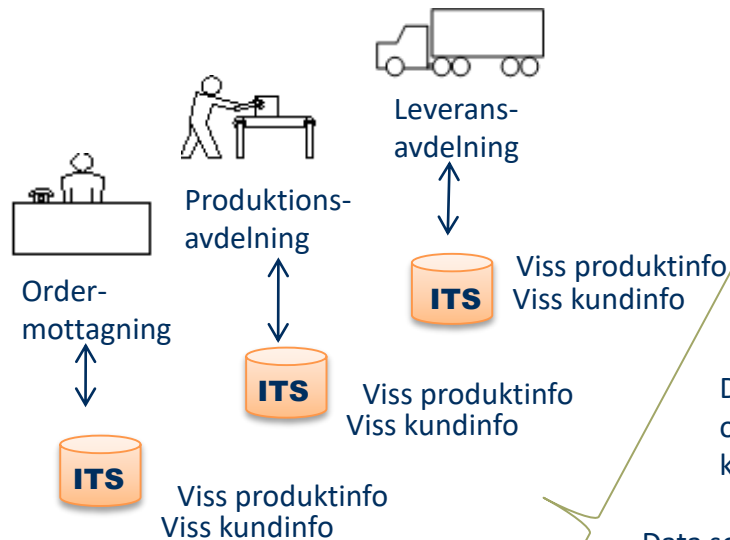
Data Warehouse

- **Data Warehouse (DW) ("datalager")** – består av en **samling av data** insamlad från flera **operationella system (ITS)** (och externa källor). Detta DW är avsett att **stödja analys och beslutsfattande**
- **Vilken data som ska extraheras ur de operationella systemen** måste bestämmas i förväg genom att ta fram krav på DW
- **Hur och när data ska laddas ned i DW** måste också bestämmas i förväg, och sedan programmeras detta så att detta görs automatiskt
- Data från de operationella systemen extraheras och laddas sedan ned i DW **vanligtvis en gång per dygn och då på natten**



Problem som DW måste adressera

- Men det är inte bara att ladda ned data i DW från de operationella systemen:
Innan nedladdning av data i data warehouset kan ske måste en rad problem ha hanterats



Problem som
måste
adresseras:

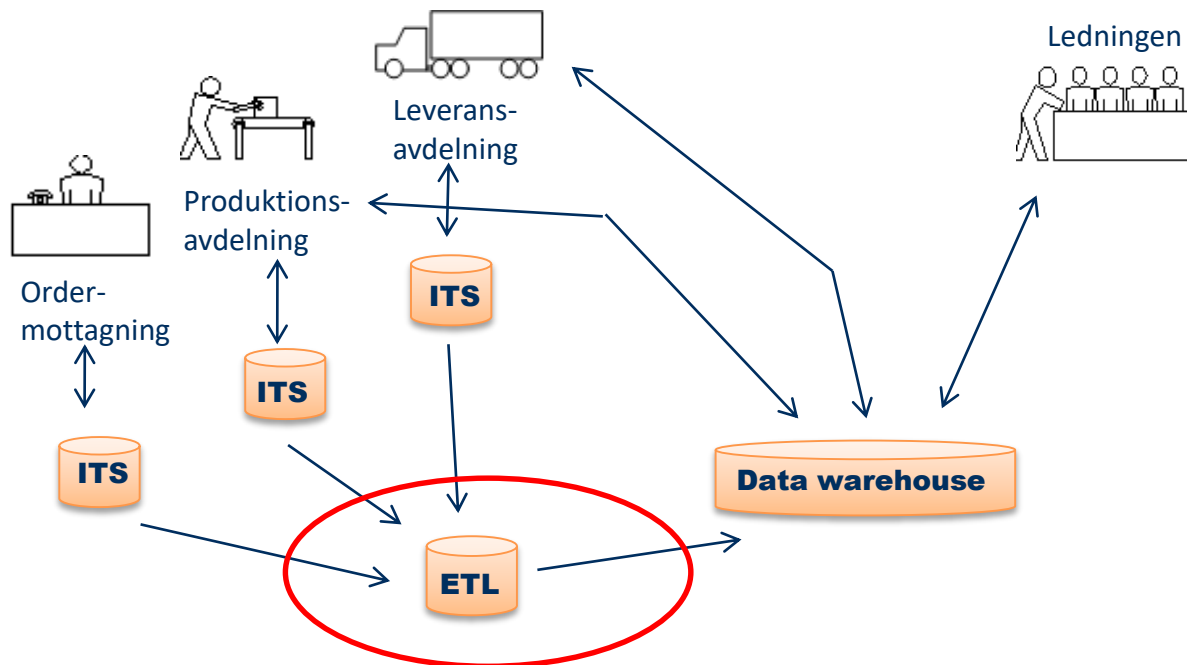
Termer som kund och produkt kan betyda olika saker i de olika operationella system

Data kan i de operationella systemen kan innehålla felaktigheter

Data som borde finnas kan saknas i de operationella systemen



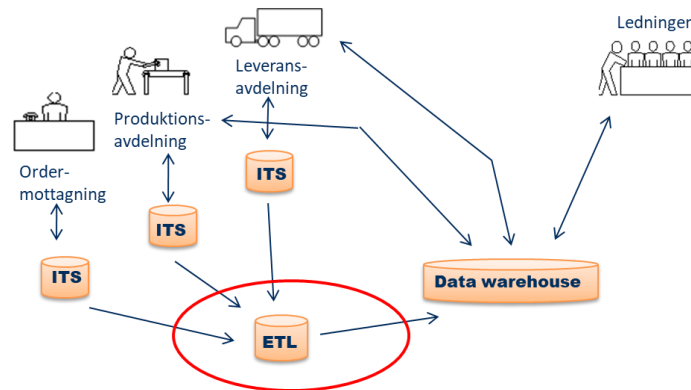
Data warehouse-arkitektur 2



ETL-verktyg



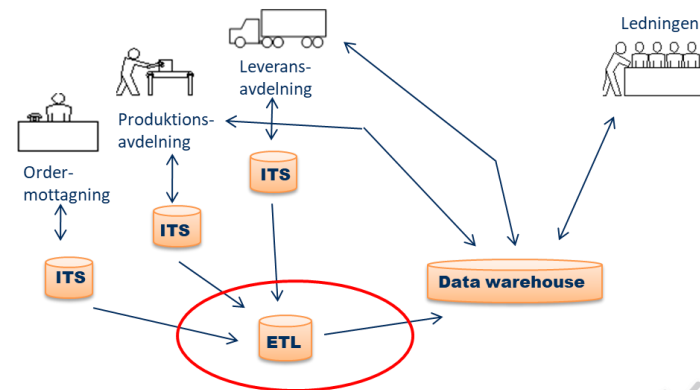
- Ett ETL-verktyg är ett system som stödjer **"extract"**, extrahering av data från de operationella systemen, men också från externa källor om de innehåller viktig info, **"transform"**, transformering av data, och **"load"**, laddning av data in i DW



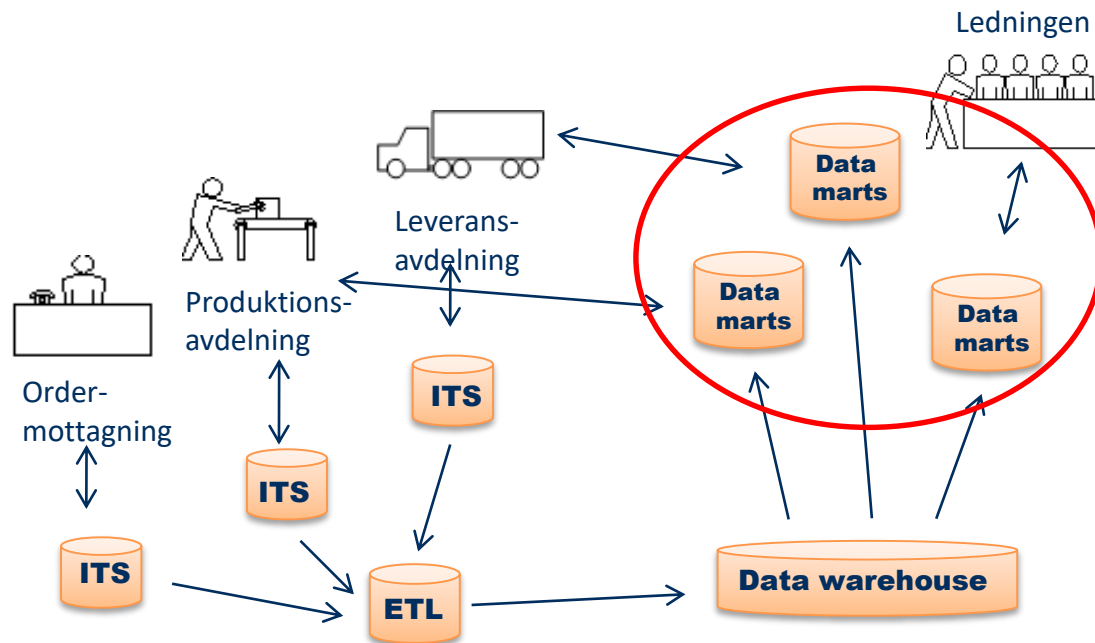
ETL-verktyg



- **"Transform"**, transformering av data, kan innebära flera saker:
 - Transformera data till en gemensam datastruktur
 - Transformera data så att felaktigheter korrigeras ("tvätta data")
 - Transformera data så centrala begrepp som kan ha olika betydelser i de olika operationella system får samma betydelse i DW



Data warehouse-arkitektur 3



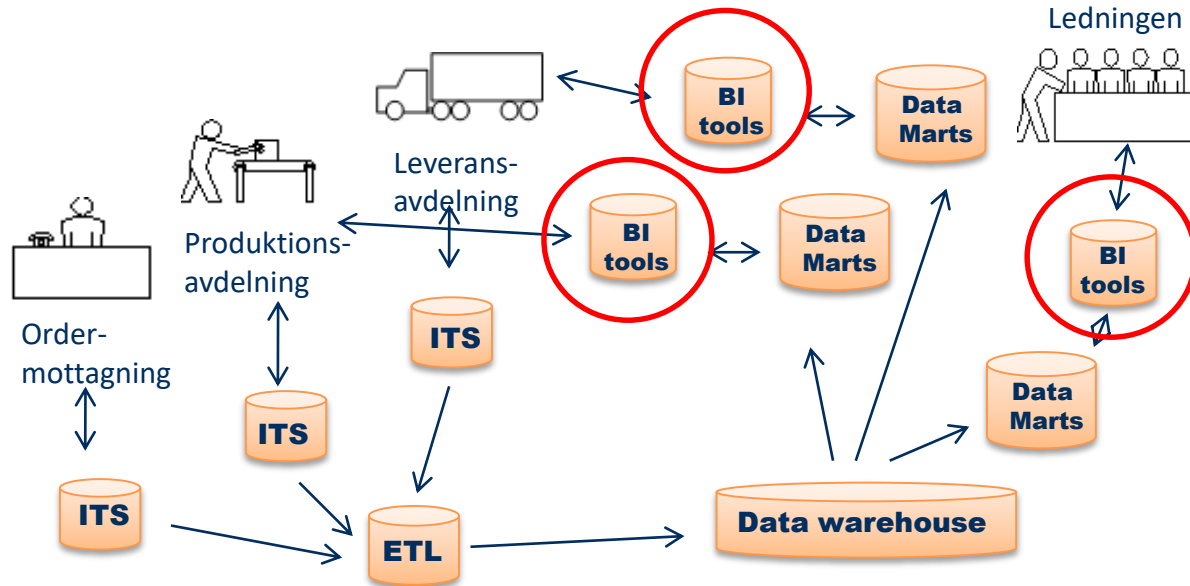
Data marts



- Ett data marts - är ett mindre DW som är anpassat för att stödja en viss avdelning eller process för enkel och snabb access.



Data warehouse-arkitektur 5



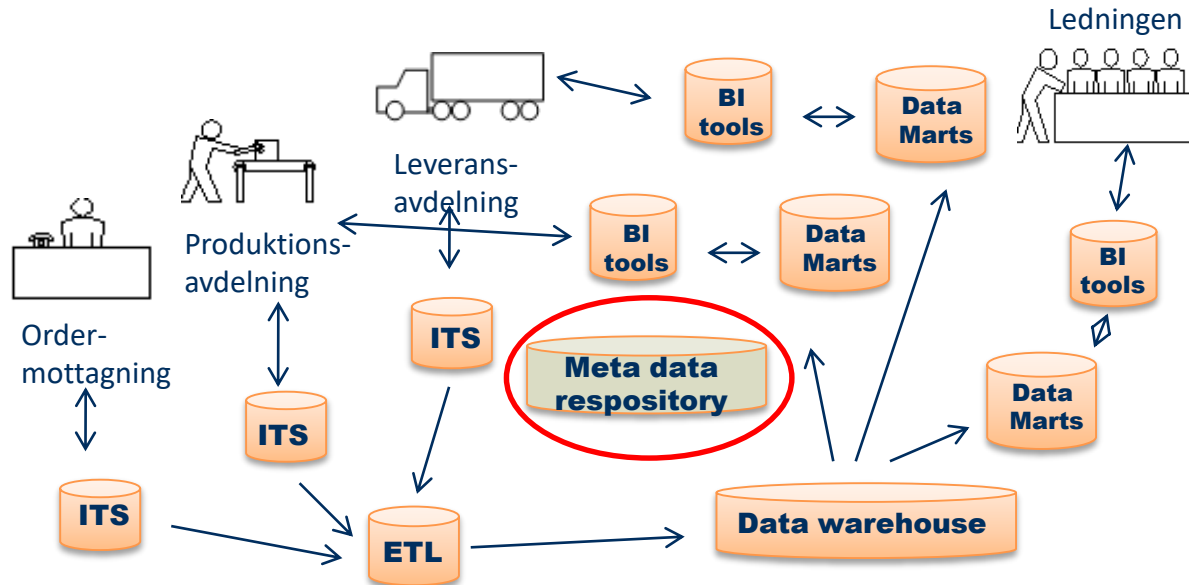
BI-verktyg



- **Business Intelligence-verktyg (BI-verktyg)** – är de system som används för att accessa informationen i DW och data marts
- Olika former av BI-verktyg:
 - **Rapport/Rapportdesign-verktyg**
 - IT-system som används av utvecklare eller användare för att skapa användaranpassade rapporter
 - **OLAP-verktyg**
 - IT-system som möjliggör för användare att utföra multidimensionell analys
 - **Data mining-verktyg**
 - IT-system som möjliggör för statistiker och skickliga verksamhetsanalytiker att upptäcka mönster i stora datamängder genom att använda olika algoritmer och modeller



Data warehouse-arkitektur 6



Meta data repository



- Innehåller information om datastrukturen i data warehouse och i data marts
- Innehåller information om vilken data som hämtas från vilka källsystem (ITS, operationella system)
- Innehåller information om transformeringar och tvättningar i ETL-verktyget
- Innehåller information om BI-verktyg som används
- Innehåller information om hur användarna använder systemen (BI-verktygen, data marts och DW) i DW-arkitekturen



BSS: Fördelar med data warehousing 1(2)

- **Infokvaliteten kan öka**
 - Korrekthet – felaktig info har ofta korrigerats i ETL-systemet
 - Fullständighet – info från olika operationella systemen är samlade i ett system (DW) samt att man kan hämta data från externa källor
 - Konsistens – de regler som används av ETL-systemet ser till att info är konsistent i DW
- **Snabba svar på frågor**
 - DW är anpassat för att ge snabba svar på frågor från användaren



BSS: Fördelar med data warehousing 2(2)

- **Stödjer multidimensionell analys (OLAP – Online Analytical Processing)**
 - Info kan enkelt analyseras ur olika dimensioner genom att användaren väljer vilka dimensioner som han/hon vill studera
- **Möjliggör historisk analys**
 - Alla info sparas i DW över tiden vilket möjliggör analys av historiska trender
- **Avlastar operativa system från frågor av analytisk typ**
 - DW gör att det inte finns samma behov av att fråga de operationella systemen och störa dessa i sitt dagliga arbete eftersom information laddas ned från de operationella systemen under natten

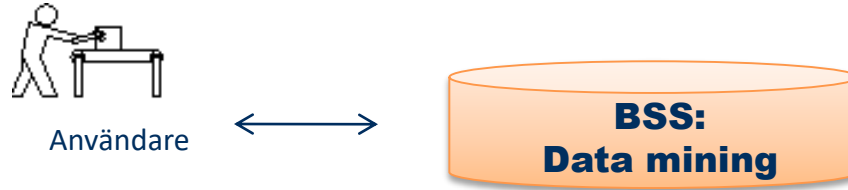


BSS: Nackdelar med data warehousing

- DW-projekt är stora, dyra och riskfyllda projekt
- Kostnad för drift och förvaltning av DW-system kan överstiga nyttan
- Risk att väsentlig data för beslutsfattande inte finns tillgänglig - varken källsystem eller i form av extern data
- Risk att de operationella systemen har data av dålig kvalitet så att datakvalitet i DW också blir låg
- ETL-process tar flera dagar att processa istället för några timmar



BSS: Data mining



Data mining – möjliggör för statistiker och skickliga verksamhetsanalytiker att upptäcka mönster i stora mängder data genom att använda olika algoritmer och modeller, utan att ha någon direkt hypotes om samband mellan data – endast en idé om vilken data som ska korreleras

Till exempel kan en data mining-algoritm tillämpas på kvitton hos en varuhuskedja för att se vilka varor som köps tillsammans

