

Dátové sklady Metadata Management

Juraj Caňo

1.10.2010

Pokročile databazové technológie, FIIT STU



Obsah

Čo sú to metadáta

Metadata Management

Príklady, Riešenia



- Čo sú to metadata?
- Definícia metadát:
 - *Metadáta sú dáta opisujúce iné dáta. Môžu reprezentovať jednoduché pomenovanie informácií o zdroji, ale aj zložitejšie štruktúrované záznamy.*
 - *Metadáta sú štruktúrované informácie, ktoré nám dovoľujú nájsť informácie o dátach, spravovať ich, kontrolovať a porozumieť im.*

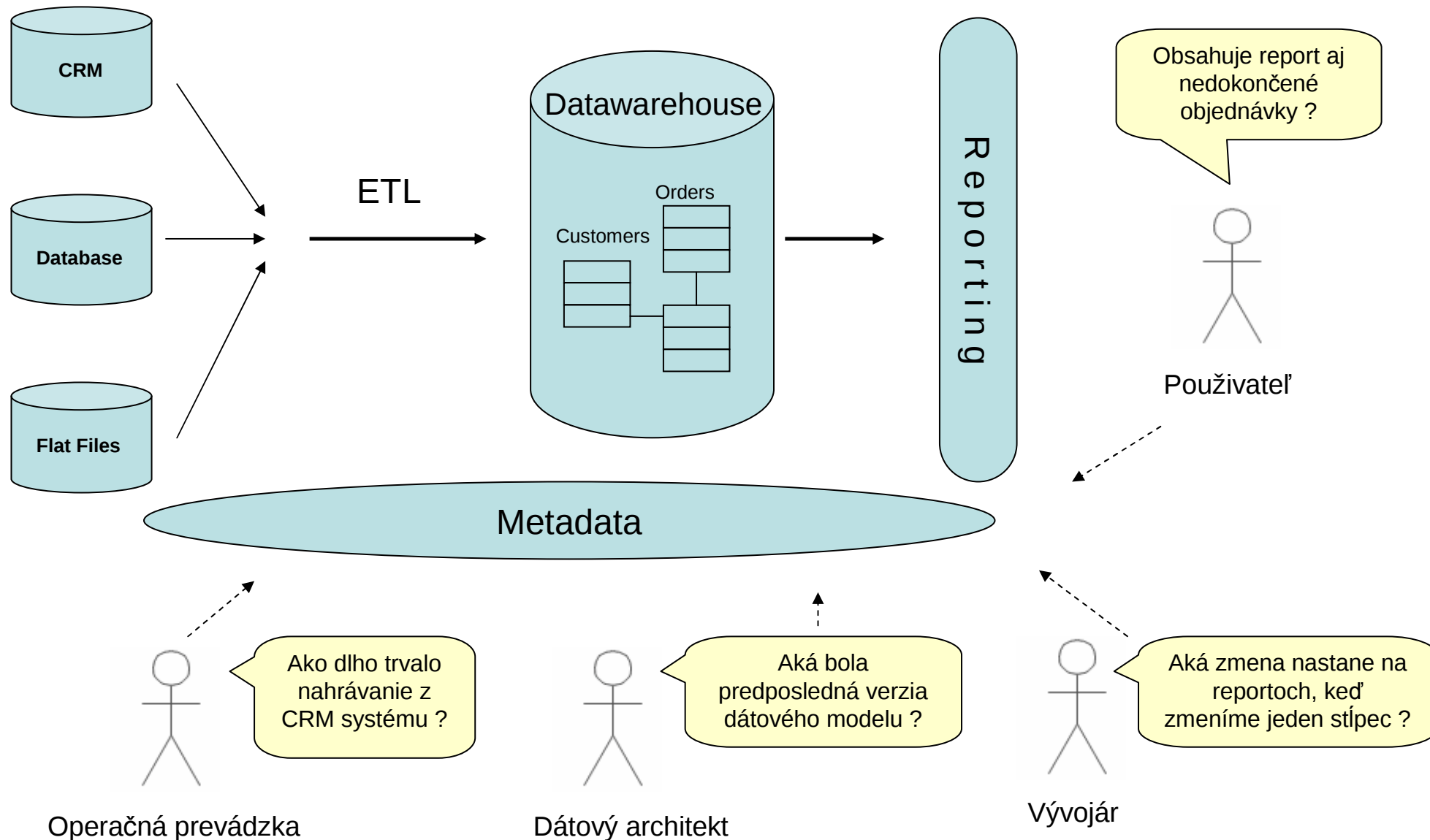
Citát, David Marco:

„Metadáta predstavujú fyzické dáta a know-how z každej firemnej oblasti, vrátane informáciach o fyzických dátach, technických a business procesov, pravidiel a obmedzení nad dátami a štruktúre dát použitej v danej organizácii”

História Metadát

- Metadáta vznikli ako dôsledok komplexnosti dátových skladov a nemožnosti udržať všetky informácie o dátových tokoch iba v hlavách analytikov.
- Ako prvé vznikli metadáta o jednotlivých transformáciách dátových tokov a neskôr sa k nim pridali technické a operačné metadáta.

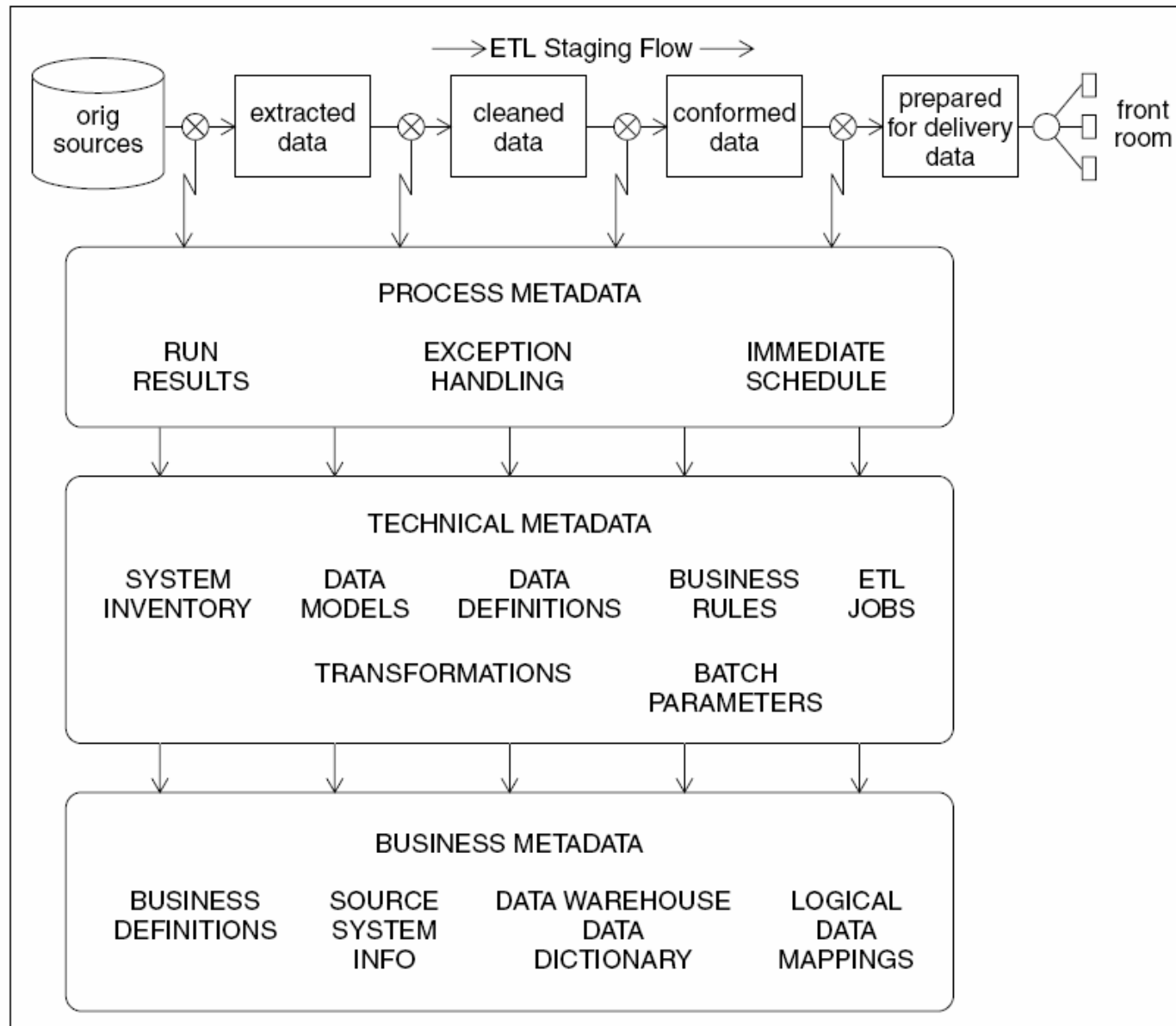
Metadáta



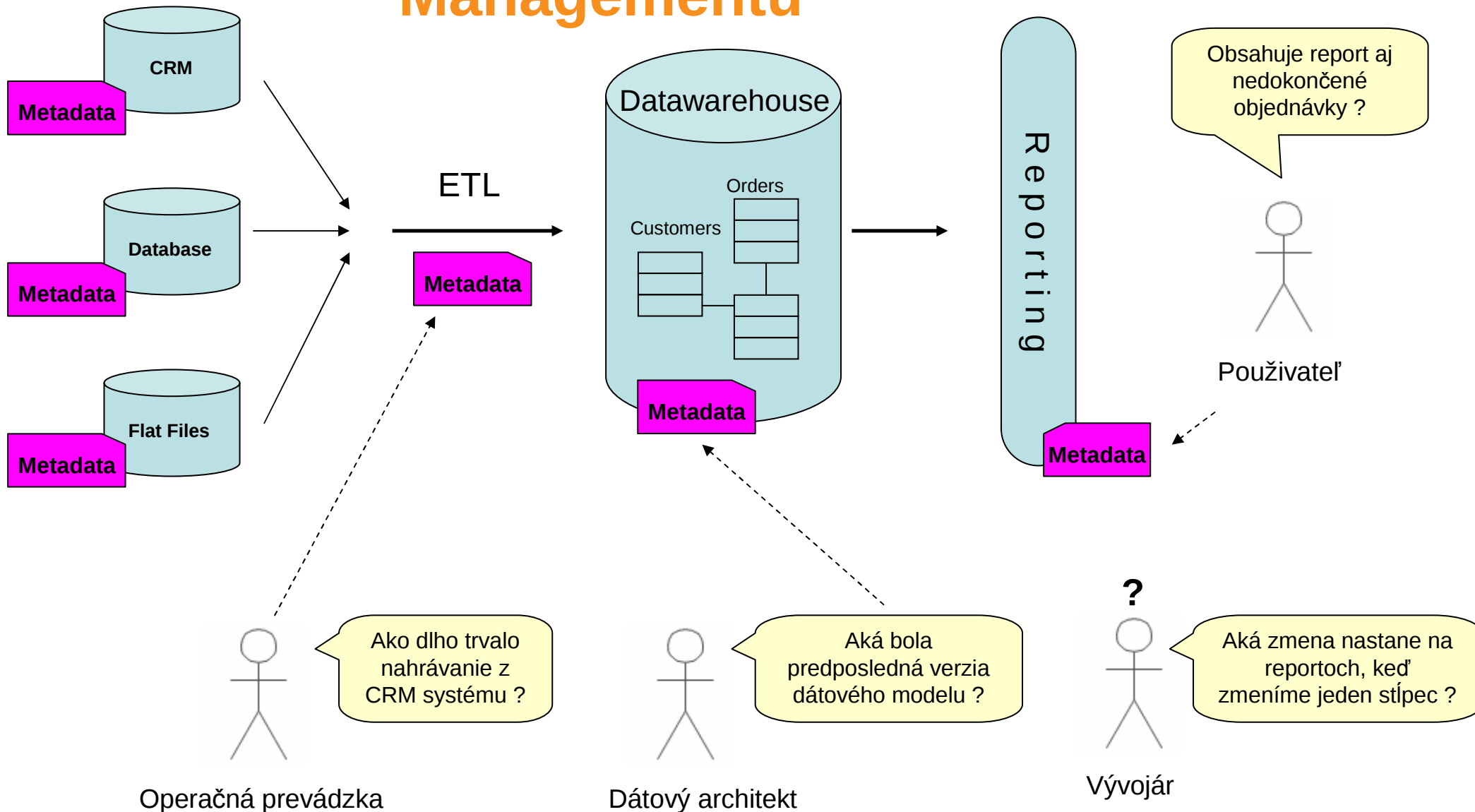
Rozdelenie metadát

- Technické
 - Tabuľka
 - dátum vytvorenia, autor, názvy stĺpcov, názov databázy
 - Stĺpec
 - dátový typ stĺpca, maximálna dĺžka stĺpca, tabuľka ku ktorej patrí
 - Report
 - Typ reportu, dátum vytvorenia, zdroje daného reportu
- Business
 - Business názvy, business pravidlá
 - Názov tabuľky, popis tabuľky, popis ako sa majú dané informácie pripraviť
- Process Execution
 - Dĺžka spracovania dát, úspešnosť spracovania dát

Metadáta



Riešenie bez Metadata Managementu



Obsah

Čo sú to metadáta

Meta-data Management

Príklady, Riešenia



Metadata Managent

- Automatické extrahovanie metadát
- Integrácia dostupných metadát
- Prepojenie súvisiacich metadát
- Centralizovaná správa metadát

Automatické extrahovanie metadát

- Podpora pre väčšinu používaných systémov
 - Databázy, modelovacie nástroje, ETL, reportovacie nástroje
- Možnosť nastavenia automatického schedulingu
 - Periodická aktualizácia metadát
- Možnosť definovať si vlastný model metadát
 - Možnosť definície vlastných pravidiel a štruktúry pre metadáta

Integrácia metadát

- Integrácia metadát zo všetkých systémov
- Jednotný katalóg metadát
 - Vyhľadávanie, triedenie, exportovanie
- Konzistencia metadát
 - Všetci používatelia používajú rovnakú databázu metadát
- Vytvorenie reportingovej vrstvy nad metadátami
 - Možnosť generovania reportov nad metadátami

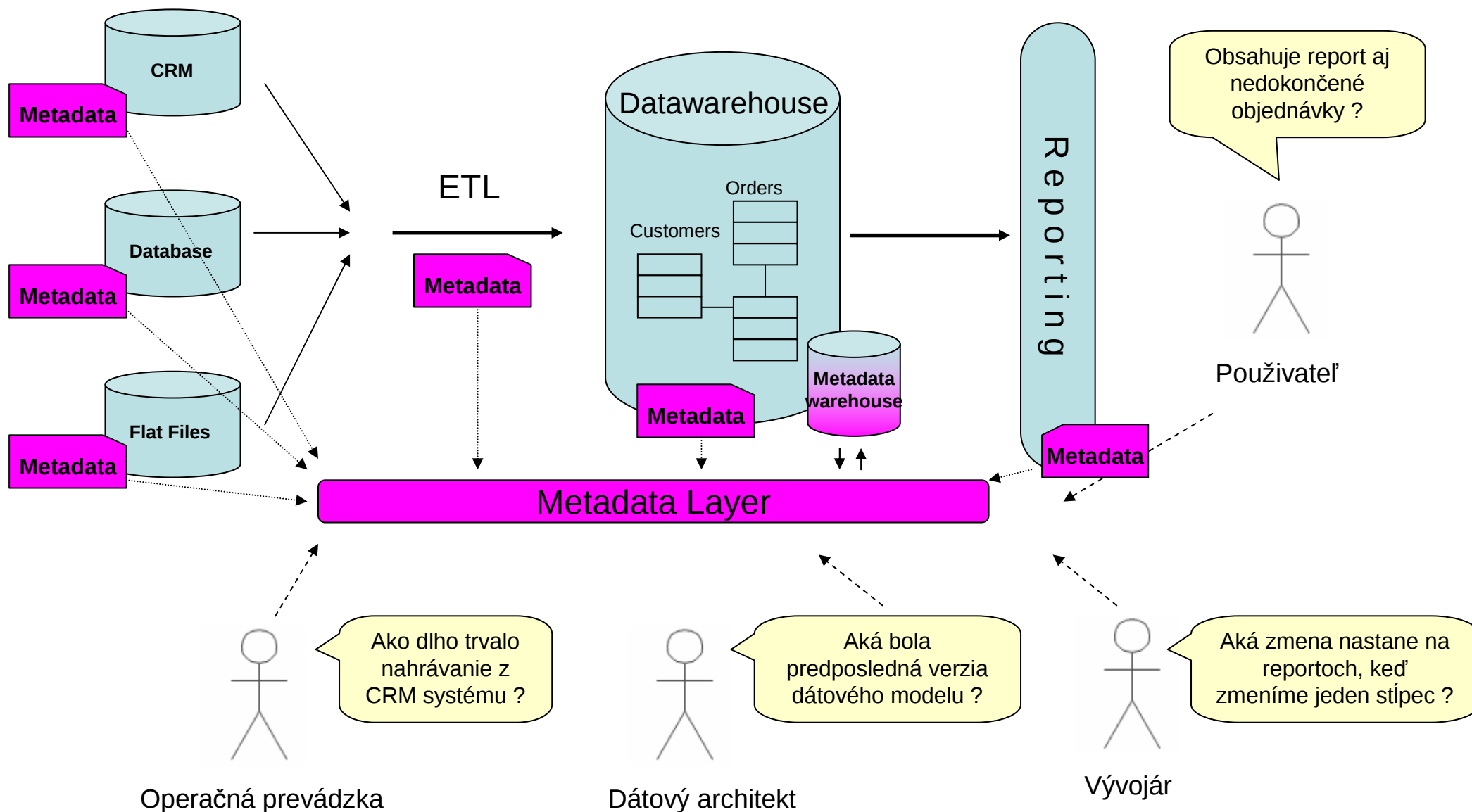
Prepojenie súvisiacich metadát

- Prepojeniu súvisiach metadát sa hovorí aj **Data Lineage**
- Zobrazuje transformáciu dát od zdroja cez ETL, dátový sklad až po reporty
- Kvôli komplexnosti dátových skladoch je to jedna z najdôležitejších funkcií

Centralizovaná správa metadát

- Poskytnutie jednotného portálu pre správu metadát
- Používatelia majú prístup ku všetkým metadátam na jednom mieste
- V prípade zmeny v metadátach sa katalóg aktualizuje pre všetkých

Riešenie s Metadata Managementom



Problémy v Metadata Managemente

- Neustále zmeny v dátovom sklade
 - Neaktuálne metadáta
 - Nemožno sa na 100% spoľahnúť, že metadáta sú správne
- Nie všetky metadáta sa dajú extrahovať automaticky

Obsah

Čo sú to metadáta

Meta-data management

Príklady, Riešenia



Riešenia pre Metadata Management

- Asg Software Solutions

- Rochade



- Informatica

- Metadata Manager



- Data Advantage Group

- MetaCenter



Diskusia

