

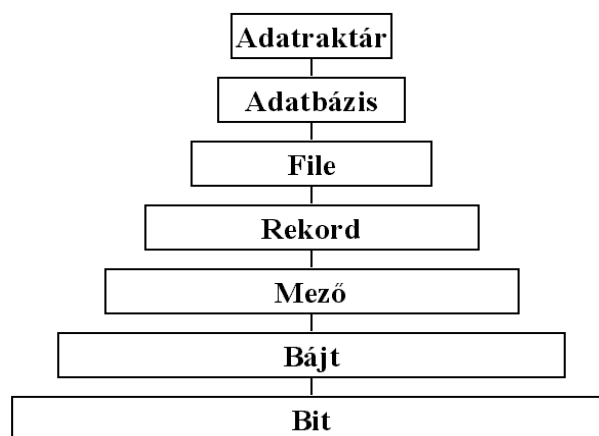
Adatszervezés

Alapvetően kétféle adatszervezési mód létezik:

- a **fájlkezelő rendszerre** és az
- **adatbázis-kezelő rendszerre** alapozott.

Vizsgáljuk meg a két rendszer különbségét! Mielőtt azonban ezt az összehasonlítást megtennénk, tegyünk egy kis kitérőt! Ismételjük át az adatok hierarchiájának kialakításáról tanultakat!

Az adathierarchia ábrája



1. ábra Adathierarchia

A hierarchia legalsó szintjén a bit áll, a legkisebb egysége az adatnak, amely mint tanultuk két értéket vehet fel (0 vagy 1). A manuális rendszerben ez a szint még hiányzik, ez csak a számítógépes rendszereknél fontos.

- A bitek alkotják a bájtot és ezek alkotják a következő hierarchikus szintet. Egy **bájt** a manuális rendszerben egy karaktert reprezentál.
- A **mező** összetartozó karakterek egy csoportját jelenti, amely lehet egy szó, több összetartozó szó illetve szám, (például a hallgatók „Iratkozási lap”-ján a hallgató neve, lakcíme).
- A logikailag egymáshoz tartozó mezők csoportját, tehát pl. egy hallgatói „Iratkozási lap” összes adatát **rekordnak** nevezzük.
- Minden hallgató „Iratkozási lap”-on lévő adatát együtt (összetartozó rekordok halmaza) **adattáblának vagy fájl**nak nevezzük.
- Ha feltétezzük, hogy a többi adattábla (fájl) tartalmazza a hallgatók egyéb adathalmazát (például vizsgaeredmények, befizetések), akkor a hallgatói nyilvántartás minden adattáblája együttesen jelenti az **adatbázist**.

A hierarchia legfelső foka sokáig az adatbázis volt, néhány éve azonban megjelentek az adatraktárok, más néven adattárházak (data warehouse), melyek helyet követeltek maguknak. Az adattárházak valójában az adatbázisokból különböző eljárásokkal (elsősorban aggregációval) előállított többdimenziós adathalmazok, melyek szerkezete a lekérdezési, felhasználói cél szerint kialakított. (Az adatszerkezeteknél részletesebben írunk az adattárházakról.)

Fájlkezelő rendszer

A fájlkezelő rendszerben a felhasználók az általuk használt programok fájljaiban tárolják az adatokat. Az adathierarchia legmagasabb szintje a fájl.

Egy gazdasági rendszerben több tíz, száz fájl keletkezhet, amelyek csak a saját programjaikkal kezelhetők.

Az adatok egy része több fájlban is megtalálható egyszerre (ez elsősorban a **törzsadat** jellegű adatokra jellemző), másrészt több felhasználónak is szüksége lehet ugyanarra az adatra.

A hagyományos fájlkezelő rendszerek hátrányai:

- **Programfüggőség**, ami azt jelenti, hogy minden fájl csak adott programhoz tartozhat, csak azzal kezelhető.
- **Redundancia**, mely egy adat több helyen való tárolását jelenti. A redundancia többféle problémát jelent, melyek közül a legfontosabb: nagyobb lesz a tárolóhely-igény.
- **Inkonzisztencia**, melyet a különböző fájlokban tárolt azonos adatok eltérő időben történő módosítása, aktualizálása okoz. Más-más értékeket vehet fel egy adat különböző időpontokban, attól függően, hogy mely fájl tartalmazza.

Adatbázis-kezelő rendszer

A hagyományos fájlkezelő rendszerrel említett hátrányok kiküszöbölésére dolgozták ki az adatbázis-kezelő rendszer (DataBase Management System) koncepcióját.

Az adatok hierarchiájának legmagasabb szintje az adatbázis, mely jelenti a logikailag összetartozó fájlok (adattáblák) csoportját.

Az adatbázis független azoktól a programoktól, melyek a benne tárolt adatokat használják.

A teljes függetlenség abból adódik, hogy a programok nem közvetlenül, hanem az adatbázis-kezelő rendszeren keresztül jutnak hozzá az adatokhoz.

Az adatbázis-kezelő rendszereket kétféle módon csoportosíthatjuk.

Az egyik csoportosítás szerint megkülönböztetünk:

- **tény-adatbázis kezelőket és**
- **dokumentum típusú** (szöveges információ-visszakereső) rendszereket.¹

A gyakorlatban adatbázisnak általában a tény-adatbázisokat szokták tekinteni.

¹ A szöveges információ-visszakereső rendszerek az adatbázisok speciális kategóriáját alkotják. Ezzel a témával e tananyagban nincs lehetőségünk foglalkozni.

A (tény) adatbázis-kezelő rendszereknek biztosítani kell:

- különféle felhasználói igények hatékony kielégítését
- adatfüggetlenséget
- adatok közötti komplex kapcsolatok ábrázolását
- redundancia mentességet illetve annak ellenőrzését
- egyszerű használatot
- az adatok védelmét
- az adatok **integritását**, hogy a hozzáférésre jogosultak se ronthassák el lehetőleg az adatbázist
- a helyreállíthatóságot, hogy bármilyen hiba esetén az eredeti állapotot vissza lehessen állítani
- több felhasználós adatbázisoknál az egyidejű hozzáférés lehetőségét, illetve annak ellenőrzését
- osztott adatbázisoknál lehetővé kell tenni az adatok szétosztását és megtalálását a különböző helyekről, az egyidejű hozzáférés és adatforgalom optimalizációját.

További ismereteket szerezhet az adatbázisok témakörét feldolgozó tananyagokból, melyeket szintén a Centroszet weboldalán találhat meg.