**Databáze** (DB= báze dat = datový sklad) = organizovaný a uložený soubor dat.

Rozsáhlejší systémy pro zpracování dat se nazývají informační systémy.

**Informační systém** (IS) = systém pro sběr, uchovávání, vyhledávání a zpracovávání dat za účelem poskytování informací.

Databáze a informační systémy:

dříve: v papírové podobě např. kartoréka        x    dnes: v počítačové podobě např. počítačová databáze

**Historie zpracování dat:**

Historie zpracování dat je dlouhá a zahrnuje mnoho významných událostí a pokroků v oblasti technologií a informačních systémů. Zde je stručný přehled historických událostí v oblasti zpracování dat:

1. **Předhistorie:**
   * Lidé po tisíce let uchovávali a předávali informace ústně a prostřednictvím zápisů na kamenech, hliněných tabulkách a papyru.
2. **16. a 17. století:**
   * Předchůdci moderních počítačů jako Pascalina a analytický stroj od Charlese Babbageho byly navrženy pro zpracování matematických výpočtů.
3. **19. století:**
   * Charles Babbage vyvíjel koncept programovatelného stroje – analytického stroje, který byl předchůdcem moderních počítačů.
4. **20. století:**
   * 40. léta: První elektronický počítač ENIAC (rok 1943) byl sestaven na University of Pennsylvania. Byl schopen provádět složité výpočty a analýzy.
   * 50. léta: Vznikají první programovací jazyky, jako je Fortran a COBOL, které umožnily psát instrukce pro počítače.
   * rok 1959 COBOL první jazyk pro počítačové zpracování dat
   * 60. léta: Koncept databázových systémů se začal rozvíjet. První generace komerčních počítačů byla nasazena v podnicích a institucích.
   * rok 1965 - první systémy řízení báze dat (SŘBD, RDBMS)
   * 70. léta: Vznik relačního modelu dat a relačních databázových systémů. Vývoj SQL (Structured Query Language) umožnil snadnější manipulaci s daty.
   * rok 1974 - jazyk SEQUEL (předchůdce SQL)
   * 80. léta : Růst počítačových sítí a vznik internetu přispěly k větší dostupnosti a sdílení dat. Další SQL databáze (DB2, Informix, SyHase, Progres, ...)
   * rok 1980 - SQL databáze Oracle
   * rok 1986 první standardizace SQL86 (organizace ANSI), další rozšířené standardy SQL89, SQL-2 (92), SQL-3(99), SQL2003, SQL2006
   * 90. léta: Růst počtu osobních počítačů a pokroky v technologiích úložiště dat přinesly revoluci v uchovávání a zpracování dat.
   * 2000+: Nástup big data – zpracování a analýza obrovských datových souborů, rozvoj cloudových technologií a umělé inteligence.
5. **21. století:**
   * Růst významu datové analýzy, strojového učení a umělé inteligence. Data se stávají strategickým zdrojem pro podnikání a rozhodování.

Historie zpracování dat jasně ukazuje, jak rychle se technologie vyvíjejí a jaký vliv mají na společnost, ekonomiku a vědu. Způsob, jakým zpracováváme a využíváme data, se neustále mění a přizpůsobuje novým výzvám a možnostem.

Hromadné zpracování dat - přístup k datům: Souborový nebo Databázový

**Souborový přístup:**data jsou uložena do jednoho nebo více datových souborů uložených na vhodném paměťovém médiu (například magnetické pásky). Součástí souboru je i jejich popis. Aplikace je v tomto případě úzce vázána na strukturalizaci dat.

Nevýhody souborového přístupu:

- obtížná dosažitelnost dat = každý nový požadavek uživatele vyžaduje vytvořit speciální program, což vede k časovým prodlevám a omezení uživatele

- nezabezpečení proti nedůslednosti uživatele

- uživatelský přístup = problémy se současným přístupem více uživatelů. Navíc není umožněno sdílení dat mezi více různými uživateli (například mezi jednotlivými odděleními ve společnosti)

- programy a data jsou vzájemně závislá: pokud je nutné změnit organizaci dat, je třeba tyto změny promítnout do všech programů, které s daty pracují

- redundance dat = opakování dat ve vztahu k problémovému využívání týchž dat pro různé aplikace, dochází k duplicitě uložených dat a tím k růstu celkového objemu dat

 - izolovanost dat - neexistuje struktura mezi jednotlivými soubory

- nekonzistence dat - problém s časovými změnami dat. Například je-li některý údaj redundantní, pak musí mít ve všech souborech stejnou hodnotu (= konzistence dat). Při změně údajů se mohou data stát nekonzistentní, protože se údaj nezmění všude, kde je zapsán.

- nezajištění integrity dat - omezená pravdivost dat a jejich celistvost, všechna data v datových souborech musí odpovídat stavu reálného světa a součástí aplikačních programů musí být kontrola vstupních dat

**Databázový přístup**: jsou odstraněny nevýhody souborově orientovaného přístupu. Vzniká komplikovanější struktura, která je řízena Systémem řízení báze dat (SŘBD). V tuto chvíli mluvíme o Databázových systémech (DBS), které vznikají spojením Systému řízení báze dat (SŘBD) a vlastní Databáze (BD).

**Databázový systém:  🔴 DBS je součástí IS.**

**🔴 DBS = SŘBD + DB**

Výhody databázového přístupu:

- sdílení dat = vytvoření možnosti paralelního přístupu k datům s případným omezením dle autorizovaného přístupu

- ochrana dat před zneužitím

- nezávislost dat na aplikaci - datové soubory jsou striktně odděleny od aplikační části a spravují  se společně

- přístup k datům je možný pouze prostřednictvím databázových programů

- data je možné vyhodnocovat různými způsoby včetně grafických přehledů

- možnost ukládání velkých objemů dat v závislosti na kapacitě úložného prostoru

- využití jazyka SQL - jazyk SQL tvoří standard, který umožňuje použití datových zdrojů z různých databázových systémů

- zamezení redundance dat - správně navržená databáze eliminuje redundanci, ale s největší pravděpodobností ji nikdy neodstraní úplně (některé záznamy je vhodné duplikovat pro zvýšení výkonu DB systému nebo kvůli zpětné kontrole)

- zajištění konzistence dat - právě zamezením redundance = při omezení počtu stejných položek je menší riziko, že se při aktualizaci na některou z duplikovaných položek zapomene

- integrita dat - SŘBD může díky správně nastaveným pravidlům a integritním omezením vynucovat na uživateli zadávání takových dat, která dané podmínky splní a zároveň zabezpečí jejich celistvost