

Informační rešeršní systém micro CDS/ISIS

Informační rešeršní systém CDS/ISIS (Computerized Documentation Service/Integrated Set of Information Systems) je programový systém, vytvořený ve spolupráci několika specializovaných agentur OSN a v současné době spravovaný UNESCO. CDS/ISIS je využíván různými agenturami OSN (UNESCO, ILO, MAEA, UNIDO), univerzitami, výzkumnými a vývojovými pracovišti na celém světě. V ČSSR je v současné době okolo 100 instalací.

UNESCO pověřila distribucí systému CDS/ISIS v ČSSR ústav pro automatizaci řízení v průmyslu INORGA Praha, který systém rozšiřuje ve verzích pro počítače IBM PC XT/AT a sálové počítače pracující s operačním systémem OS.

Funkce systému

CDS/ISIS je obecný systém navržený pro počítačové zpracování strukturovaných nenumernických databází.

Umožňuje zejména:

- definovat databáze obsahující požadované datové položky
- zaznamenávat nové věty do příslušné databáze
- modifikovat, opravovat nebo rušit existující věty
- automaticky vytvářet a udržovat soubory pro rychlý přístup do každé databáze
- vyhledávat věty pomocí uměle vytvořeného vyhledávacího jazyka
- libovolně třídit věty
- zobrazit věty nebo jejich části podle požadavků
- tisknout částečné nebo úplné katalogy nebo indexy z jakékoliv databáze
- provádět výměnu dat mezi instalacemi.

CDS/ISIS se skládá ze sady programů provádějící různé funkce systému. Uživatelské programy (tj. programy pracující s již vytvořenými databázemi) jsou:

- ISIS - provádí všechny služby, týkající se údržby databází a vyhledávání
- ISISPRT - poskytuje tiskové výstupy
- ISISINV - umožňuje údržbu invertovaných souborů a různé pomocné funkce

Systémové programy (tj. programy, určené pro definování nových databází a provádění systémových služeb):

- ISISDEF - umožňuje definování nových databází a modifikaci již vytvořených definic
- ISISUTL - poskytuje pomocné funkce systému
- ISISXCH - poskytuje prostředky pro výměnu dat mezi instalacemi a pomocné programy pro základní soubor.

Jednotlivé programy se volají jejich jménem z operačního systému. Výběr požadovaných operací se provádí volbou z menu. Po výběru z menu systém operaci vykoná nebo se dotazuje na další potřebné informace, nutné k vykonání operace. V některých případech je jako odpověď na výběr jiné menu, ze kterého se provádí další výběr.

Všechny programy jsou plně interaktivní. Systém rovněž umožňuje volbu jazyka, ve kterém budou systémová menu a zprávy zobrazovány (za předpokladu, že zvolená jazyková verze je v instalaci zahrnuta).

Vstup dat je prováděn přes vstupní formuláře (obrazovky). Systém používá dva typy formulářů, tzv. "systémové formuláře" (pro získávání parametrů pro vykonání požadované funkce, např. tisk) a formuláře pro sběr dat pro příslušnou databázi.

Struktura databáze v CDS/ISIS

CDS/ISIS je navržen pro práci s větami proměnné délky.

Pole v jednotlivých větách jsou nepovinná, mohou obsahovat podpole nebo mohou být opakovatelná.

Parametry databáze (platí pro CDS/ISIS micro verze release 1.0):

max. počet vět	32 000
max. počet polí ve větě	99
max. délka pole	1 650 znaků.

Každá databáze se jeví navenek jako jeden soubor informací, ale fyzicky se skládá z několika souborů:

Základní soubor - obsahuje všechny věty dané databáze

Soubor křížových referencí - indexy obsahující umístění každé věty v základním souboru

Invertovaný soubor - obsahuje termíny, které se dají použít jako přístupové body pro vyhledávání a odkazy na věty základního souboru, ve kterých je termín obsažen. Pro umožnění rychlého vyhledávání se invertovaný soubor skládá z několika fyzických souborů.

Tyto soubory spolu s dalšími, které se automaticky zakládají při definování databáze, patří mezi tzv. "povinné soubory", které musí pro danou databázi vždy existovat. Systém dále vytváří pomocné soubory, které slouží k provádění určitých systémových funkcí. Uživatel má rovněž možnost podle potřeby vytvářet tzv. "uživatelské soubory" (např. výstupní formáty a pod.).

Vyhledávání

Věty základního souboru je možno vyhledávat buď přímo pomocí jejich čísel (prostřednictvím souboru křížových referencí), a nebo zadáním přístupových termínů, které se vyhledávají v invertovaném souboru. V invertovaném souboru se určí

seznam čísel vět ze základního souboru, odpovídající každému termínu. Tyto seznamy jsou pak zpracovány programem s použitím booleovských operátorů, zadaných při formulaci dotazu. Nakonec je vytvořen seznam čísel vět, které splňují dané podmínky. V případě požadavku systém vyhledá každou větu uvedenou v "hit seznamu", formátuje ji a zobrazí na obrazovce, nebo ji uchová pro pozdější tisk.

Přístupové termíny je možno zadávat přímo z klávesnice, nebo je možno vybírat ze slovníku klíčových slov, který systém na požádání zobrazí. Jednotlivé termíny je možno spojovat do dotazového výrazu. Každý dotazový výraz je možno libovolně upravovat (rozšiřovat, zužovat a pod.) tak, aby na jeho výstupu byly vybrány relevantní záznamy.

Třídění

Před tiskem požadovaných vět je možno zadat požadavek na třídění výstupu. Pro třídění lze zadat až 4 třídící klíče o maximální délce 256 znaků.

Tiskový formátovací jazyk a dotazovací jazyk CDS/ISIS

Tiskový formátovací jazyk umožňuje přesně definovat formátovací požadavky pro věty databáze. Pomocí tohoto jazyka je možné vybírat jednu nebo více datových položek v libovolném pořadí a volitelně vkládat konstantní text. Formátovací jazyk je v systému využíván nejen ke specifikaci vět pro zobrazení nebo tisk, ale je také široce použit v celém systému, kdykoliv je nutné specifikovat data, se kterými se má pracovat.

Dotazovací jazyk je založen na booleovské algebře, která využívá metody vyjádření logické operace mezi třídami. Klíčové slovo nebo termín pro popis obsahu nebo jiné charakteristiky záznamu může být považován za třídu všech záznamů

obsahujících daný přístupový termín. Kromě základních booleanových operátorů systém využívá operátory speciální, jako např. operátory úrovně a prohledání sousedství.

Instalace CDS/ISIS (micro verze)

CDS/ISIS (micro verze) je určen pro instalaci na počítačích IBM PC XT/AT a jim kompatibilních, pracujících pod operačním systémem MS (PC) DOS. Pro využití všech možností CDS/ISIS je vhodná konfigurace s pevným diskem, barevným monitorem (EGA) a tiskárnou s možností grafického výstupu. Systém je však možno provozovat s určitými omezeními (především se do velikosti databáze) i bez pevného disku pouze s využitím dvou disketových jednotek. Otázka barevného monitoru a rychle tiskárny s možností grafického výstupu je více méně záležitostí komfortu uživatele.

Využití CDS/ISIS

Jednou z hlavních výhod systému CDS/ISIS je to, že jeden soubor programů je schopný zpracovávat neomezený počet databází, skládajících se z naprosto odlišných datových elementů (resp. omezení je dáno jen kapacitou pevného disku) a uživatel není nucen navrhovat a sestavovat programy pokaždé, kdy je vyžadováno použití vyhledávací a třídící techniky.

Z hlediska praktického využití je aplikace systému CDS/ISIS vhodná zejména tam, kde převládají textové informace, např. pro faktografické informační systémy, informační systémy a softwarových produktech, provádění automatizovaných knihovnických služeb VTEI, poradensko-informační služby, sledování vynálezeckého a zlepšovateleckého hnutí, patentová a výstřížková služba a pod.