

Súčasn $\acute{e}$ metódy spracovania geologických informácií najmä na samočinných počítačoch

J. HRUŠKA — Geofond Praha

Exponenciálny rast geologických publikácií, správ, geologických pozorovaní, meraní a laboratórnych analýz nie je dnes možné zvládnuť komplexne tak, aby bolo možné rýchle a prehľadné informovanie o nových výsledkoch i pre syntetické štúdie. Jednou z ciest, ktorá uľahčuje zhromažďovanie, vyhľadávanie a vedeckú syntézu mnohých faktov, je použitie samočinných počítačov. Trend rozvoja týchto metód je zrejmý v posledných piatich rokoch v USA, Kanade, ZSSR a Francúzsku. Hlavná fáza prípravy, ukladanie, spracovanie informácií geologických odborov — ak sa používajú samočinné počítače — sa obyčajne rozdeľuje do troch kategórií:

1. Dokumentografické spracovanie informácií — zvlášť bibliografie registrov adresnej alebo profilovej informácie o dokumentoch, ako napr.: články dizertácie, mapy a pod. Ako príklad môžeme uviesť GEO-INDEX, periodikum Geofondu ODIS.
2. Spracovanie vecných (faktových) súborov informácií, a to práve tak základných alebo archívnych súborov, alebo len súborov informácií pre osobitné úlohy. Príkladom môžu byť údaje Data — banky, a to práve tak surovinového (ložiskového) charakteru, ako i pre komplexnú charakteristiku prostredia a pod.
3. Pre spracovanie kvantitatívnych dát — údajov — v prevažnej miere používajú sa štatistické metódy, matematické analýzy atď. Príkladom sú trendové analýzy, numerická klasifikácia, diskriminačné analýzy, testovanie hypotéz a pod.

V tomto článku zmieňujem sa predovšetkým o spracovaní faktových geologických informácií, ktoré rámcove patria do skupiny úloh spracovania dát na samočinných počítačoch.

OBSAH

záznamových súborov geologických informácií musí zodpovedať ich využiteľnosti pre najširší okruh geologických odborníkov a preto rozsah a charakter údajových položiek (napr. názov ložiska, stratigrafická epocha, petrografický charakter, predpokladané zásoby, technologická kategória a pod.) by sa mal čím skôr štan-
darizovať.

NOTÁCIA

musí zodpovedať nielen zvolenému alebo danému záznamovému prostrediu (dierkový štítok, dierková páska, magnetická páska a pod.) ale i vžitým vhodným formám vyjadrovania geologickej informácie. Kvantitatívne údaje majú sa zaznamenávať ako desatinné čísla s pevnou dĺžkou pred a za desatinnou čiarkou. Číselné kódovanie je veľmi ekonomické, ale pre základné súbory a pre Data —

banku je prípustné len vtedy, keď sa môžu použiť dobre určené a široko prijaté klasifikácie a číselníky. Mnemonické skratky a kódy (napr. GRNT = granit) mali by sa použiť práve tak pre skratku, ktorá je v obecnom používaní v geologických odboroch (symboly chemických prvkov, siete topografických máp), ako aj tam, kde základ terminológie je internacionálny (latina, gréčtina). Tu je výhodné používať mnemonické kódy v obecnom odborovom používaní (napr. genéza, species a subspecies v paleontológii alebo pri mineráloch). Otvorená reč mala by sa používať v jednoznačných terminologických schémach, pri miestnych menách a všade tam, kde jej význam je závislý na kontexte, v ktorom je použitá, alebo vtedy, ak informácia má slúžiť širokému okruhu užívateľov.

ŠTRUKTÚRA

strojovo spracovaných súborov informácií má byť jednoduchá, aby umožnila zostupovanie a zoradovanie informácií, pre potrebu rôznych užívateľov a ďalej pre analytické a následné štatistické podprogramy, ako i pre syntézu dát.

V súbore faktových dát o ložiskách nerastných surovín pri systéme programov SAFRAS, ktorý vyvinulo geologické oddelenie Univerzity Western Ontario London v Kanade, sú ilustrované hlavné problémy strojového spracovania informácií. Z nich je zrejme:

- a) že obsah záznamových súborov má najvyššiu početnosť pri týchto údajových položkách: názov ložiska — prvok alebo surovina — list mapovej siete — stupne a minúty zemepisnej dĺžky a šírky — hlavne minerály — obsah úžitkovej zložky — stav otvorenia a využitia ložiska — zásoby isté — zásoby odhadnuté — ťažba.
 - b) Notácia je dosť zmiešaná (kvantitatívne dáta — číselné a mnemonické kódy — otvorená reč) a to robí tento systém ekonomický i zrozumiteľný, relatívne ľahko kontrolovateľný.
 - c) Štruktúra je obmedzená na identifikáciu a zoskupenie informačných údajov (položiek). Počiatočné usporiadanie môže sa do určitej miery meniť novým usporiadaním, pridaním údajových položiek a tiež i možnou konverziou magnetickej pásky.
- Tento postup umožňuje predovšetkým vyšší programovací jazyk COBOL, ktorý sa v iných štátoch (Francúzsko, USA, V. Britania) ukázal ako veľmi vhodný pre obdobné systémy.

Niet pochyb o tom, že ani Československo sa nevyhne spracovaniu geologických informácií na samočinných počítačoch. Už dnes máme určité skúsenosti i úspechy s vydávaním registríkovej bibliografie GEO-INDEX, ktorá profituje informácie pre vrtnú techniku, prieskum, ťažbu ropy a plynu (OBIS-UGI Brno) a data-bank pre územné plánovania a pod. Nádejné projekty a experimenty pri spracovaní ložisiek vápencov a zosuvov ČSSR môžu posunúť našu geologickú službu a Geofond medzi významné pracoviská na tomto poli v Európe.

(Poznámka): Prednáška uskutočnená 3. marca 1970 v SGÚ, Bratislava.