

Ing. Vlastimil ČEVELA

VS Ingstav, n.p., Brno

TVORBA JMEN V ANALYTICKÉ A PROGRAMOVÉ DOKUMENTACI

1. Úvod

Zpracování a realizace integrovaných projektů v oblasti hromadných dat vyžaduje zajištění jednoznačných definic a informačních návazností, platných pro automatizované i ručně prováděné činnosti. Při plnění tohoto záměru vystupuje do popředí problematika tvorby jmen, a to nejen dokumentačních a realizačních celků jako projektů, programů a souborů dat, ale i např. označování jednotlivých ručně prováděných činností, a v neposlední řadě též detailních prvků, tj. jednotlivých datových položek, programových modulů, algoritmů apod.

Předkládaný příspěvek je pokusem o konstruktivní diskuzi otázek souvisejících s touto tematikou, a o prezentaci některých vlastních představ a námětů. Materiál je rozdělen do několika částí. Nejprve je shrnut výčet objektů, které by měly být předmětem pojmenování. Po krátkých úvahách o vztahu mezi tvorbou jmen a postupem integrace ve vývoji i při realizaci automatizovaného zpracování pak následuje obsáhlejší stať o různých přístupech k řešení. Dále je předložen návrh identifikace pro metasystém ASŘP, a závěrem jsou shrnuty základní these.

Celá práce se opírá o řadu poznatků, získaných v průběhu přípravy a při vlastní realizaci automatizovaného zpra-

cování hromadných dat ve VS Ingstav n.p. Brno, a proto je zaměřena na úroveň podniku jako systému.

2. Přehled pojmenovaných objektů

Při nejširším pohledu můžeme objekty, které jsou předmětem našeho zájmu, rozdělit na procesy a informační soubory. Je to ovšem míněno skutečně v nejširším slova smyslu, tj. např. činnost odborného útvaru při zajišťování všech jeho funkcí je proces, /3/ zrovna tak jako dokumentace programu či algoritmu je vlastně popisem procesu zpracování /ať již na počítači nebo ručně/ apod. Informačním souborem je při tomto přístupu pak nejen určitý typ věty na počítačovém médiu, ale i různé formalizované doklady, stejně tak jako běžným textem psané řídicí normy a předpisy, anebo soubor dokumentace projektu či agendy apod.

S informačními soubory i procesy jsou pak v úzkém vztahu jejich posílelé /vykonávané procesy, posíle informace/, avšak tyto vazby mohou, ale v řadě případů spíše nemají mít určující význam při pojmenování, poněvadž jsou často zcela nezávislé.

Posledním důležitým faktorem je čas. V té souvislosti je nutné nejen zachytit jednotlivé fáze a pořadí procesů apod. a periodicitou jejich provádění, ale především pojmenovat generace informačních souborů, aby bylo možno popisovat systém v jeho jednotlivých stavech.

Je vhodné zde připomenout, že se zabýváme procesy zpracování informací, tj. pracujeme v rámci informačního systému, který je pouze zobrazením výrobního a dalších procesů, reálně probíhajících v podniku.

Při částečném zjednodušení můžeme říci, že procesy jsou realizovány prostřednictvím jednotlivých automatizovaných nebo

ručně prováděných činností /funkcí/. Činnosti jsou pak agregovány do dílčích zpracování, a ta pak do příslušných subsystemů, resp. dále do celého systému. Jako vykonavatelé činností pak fungují buď jednotlivé útvary podniku /příp. organizace z okolí podniku/, anebo programy /podprogramy, moduly/ počítače. S vazbou na čas jsou činnosti prováděny dle stanoveného pořadí po jednotlivých fázích a v potřebné periodicitě.

Informační soubory jsou realizovány prostřednictvím informačních vět /záznamů/, které mohou mít svoji vnitřní skladbu uspořádanou dle různých typů vět. V rámci určitého typu věty je pak nejmenším informačním prvkem datová položka. Za typy vět je možno považovat i neformalizovaná písemná nebo ústní sdělení. Jako nosiče souborů se vyskytují nejen strojevě zpracovatelné děrné a magnetické pásky, disky, diskety, děrné štítky, sestavy, mikrofiše, ale i papírové formuláře dokladů apod. Čas je v případě informačních souborů uplatněn v označení jejich generace anebo plnosti. Přitom je nutno poznamenat, že i dokumentace procesů /např. programů/ je vlastně informačním souborem, a proto i zde musí být uvažováno s generací, resp. verzí.

Vzhledem k rozsáhlosti celého informačního fondu a složitosti jeho zpracování je nutno uvažovat ještě s dalšími pojmy. Z hlediska fyzického uložení souborů jsou oblasti, které zahrnují několik souborů, a usnadňují provádění potřebných manipulací. K popisu i praktickému využívání informačního fondu z různých hledisek pak slouží tzv. logická sestavení, která se uplatňují především jako SET v databance, ale i jako různé rejstříky a další soubory průřezového charakteru /7/.

3. Tvorba jmen a procesy integrace

V souvislosti s vývojem automatizovaného zpracování hromadných dat lze vysledovat několik dlouhodobě probíhajících integračních procesů, které pro naše potřeby můžeme nazvat

vnitřní a vnější datová integrace a integrace procesů zpracování. Pokusme se nyní uvažovat nad jejich vlivem na tvorbu jmen.

Nejvýraznější je zatím patrně proces vnitřní datové integrace, který se projevuje v přechodu od zpracování izolovaných agend k jejich propojování a k postupné tvorbě stále více integrovaných systémů. To sebou samozřejmě přináší nutnost široké výměny datových souborů mezi jednotlivými agendami či úseky zpracování, resp. programy. Zatímco dříve si mohl každý programátor individuálně nadeklarovat svoje věty, nejen pokud jde o jména datových položek, ale často i strukturou, dnes se stále ve větší míře uplatňuje v souvislosti s vyššími programovacími jazyky též tendence k využívání centrálních definic typů vět, přetahovaných ze zdrojových knihoven. Je zřejmé, že má-li být takový rozsáhlejší systém alespoň minimálně unosným způsobem zdokumentován a řízen z hlediska informačních toků, musí být jednotná nejen jména informačních souborů, ale i datových položek, aby bylo možno vysledovat jejich zpracování.

Proces vnější datové integrace charakterizuje důsledky rozvoje aplikace vypočetní techniky, tj. postupné začleňování jejich automatizovaně prováděných funkcí do běžné činnosti podniku či organizace, která počítač využívá. V souvislosti s výše zmíněnou vnitřní integrací to ovšem přináší podstatné zvýšení potřeby jednotně řízených informačních vazeb jak mezi automatizovanými a ručně prováděnými činnostmi, tak řadu nových koordinačních vztahů mezi dílčími realizatory v odborných útvarech, i v rámci spolupráce "odb. útvar-analytik-programátor-provoz". Přitom není třeba zdůrazňovat, že používání různých jmen např. pro datové položky, definiční nebo vazební číselníky, či algoritmy zpracování, je nejen neekonomické, ale vnáší zákonitě do celého systému řadu chyb a zbytečných nedorozumění.

Perspektivně nejvýznamnější je patrně integrace procesů zpracování. S vývojem technických a programovacích prostředků dochází ke specializaci, tj. k vytváření různých integrovaných úseků zpracování, např. v oblasti pořizování dat jedním proudem, jednotné výstupní programy apod., s vrcholem v tvorbě databankových systémů. Tento trend spolu s rozšiřující se dělbou práce v oblasti programování vede až k unifikaci při provádění určitých funkcí, a ve svých důsledcích, jak již ukazují některé výsledky, pravděpodobně směřuje až k automatizaci programování a projektování vůbec. /4/, /7/, /8/ V takových podmínkách jsou pak jména všech prvků v systému zcela jednoznačně normativní záležitostí.

Z uvedeného vyplývá, že již nyní stojíme před reálným tlakem nejen na využívání počítače pro zpracování dílčích částí dokumentace, ale na tvorbu informačních subsystémů, které jediné budou schopny popsat celé obrovské množství prvků, které figurují v automatizovaném systému zpracování dat, a jejich potřebné vazby. Jména všech těchto prvků se tím sama stanou předmětem zpracování dat, a proto k nim musíme i z tohoto hlediska přistupovat. Jinými slovy to, co se snažíme uplatňovat v rámci námi vytvářených systémů, musíme aplikovat i na svoji vlastní práci. /4/, /8/

4. Diskuze různých přístupů k tvorbě jmen

Pomineme-li odlišnosti, vyplývající z různosti objektů, které jsou pojmenovávány, můžeme vysledovat v zásadě několik základních hodnotících hledisek pro přístupy k tvorbě jmen, která jsou předmětem našeho zájmu.

Především je to, dá se říci klasický rozpor ~~mezi~~ technika-pořadí, resp. mezi významovými jmény, která nějakým způsobem navozují asociaci buď přímo na daný jev, nebo na určitou kategorizaci - a jmény, vytvořenými prostým způsobem pořadového číslování.

Pokud se pokusíme hodnotit oba přístupy objektivně, nutně musíme každému z nich přiznat řadu výhod, ale i nevýhody. A bohužel jsou to právě významově tvořená jména, která přinášejí nejvíce problémů. Už sám fakt, že jsou spojena s významem, který jimi není zpravidla exaktně přesně definován, ale uplatňován pouze pomocí asociace, tj. přes psychické představy, nese sebou silné omezení jejich obecné použitelnosti vzhledem k tomu, že navození takových představ je velice individuální.

To co je zřejmé a jednoznačné pro jednoho člověka, ani zdaleka nemusí znamenat stejnou souvislost pro druhého. Shodnou asociací kromě toho mohou navodit i podobná jména /např. ZV, ZA, ZAV = ZAVOD/, takže jejich používání ve vztahu k striktně jednoznačnému počítači nutně vede k chybám. V tomto směru je asi nejvíce problematické tvoření akronymů, tj. zkratkových slov, složených z počátečních slabik několika slov /např. POST = počáteční stav/, které při rozsáhlejším použití mohou vytvořit pro neobeznámeného spolupracovníka dokonalý kryptogram./1/

Mnemotechnická jména jsou asi nejpřijatelnější v plné, obecně používané podobě /SKLAD, ZAVOD, CTENI-I1 atd/, avšak vzhledem k jejich různé délce je obtížná jak jejich formalizace, tak přehledné uspořádání do počítačového zpracování. Z uvedeného tedy vyplývá, že významová jména by měla být používána velice uvážlivě.

Dalším hodnotícím hlediskem je vztah identifikace - klasifikace. Je až s podivem, jak málo zatím v obecném podvědomí zakotvila nesporně správná téze, že všechny objekty zkoumání je nutno nejprve jednoznačně identifikovat, abychom je pak mohli různým způsobem klasifikovat. Bohužel příliš často se setkáváme se složitými systémy pro označování dokumentace, které sice řeší všechny možné vztahy příslušností k jednotlivým subsystémům, řešitelům, seborům apod, ale při

popisu problematiky systémem řešené zaberou tolik místa, že někdy odsunou do pozadí i skutečné jádro věci.

Různá hierarchická označování /mluvící číselníky/ mají sice řadu předností, avšak v případě vyčerpání příslušné dílčí řady jsou okamžitě znehodnocena, a s velikými rezervami jsou zase příliš neefektivní. Z tohoto pohledu se asi jeví jako skutečně nejvhodnější, využít je pouze pro každou základní kategorii jevů, a dále pokračovat s pořadovým číslováním /5/.

Velice důležitým momentem při rozhodování o tvorbě jmen je obor jejich platnosti. Pokud se pohybujeme v oblasti lokální /např. datové položky deklaraace pracovního souboru nebo názvy odstavců v programu/, není otázka volby jmen tak závažná, jako při pojmenovávání prvků s působností globální /datových položek, které jsou jednoznačné v celém systému, programů, číselníků, algoritmů, základních souborů apod./. V té souvislosti je ovšem nutno připomenout, že s vazbou na průběh výše citovaných integračních procesů je stále více prvků, které mají globální platnost /7/.

Souhrnem tedy lze říci:

- A - Tvorba jmen by se měla řídit především cílem, kterému budou jména sloužit. Proto by např. ve jménech vyšších celků /vět, programů, projektů, dokladů apod./ mělo kromě nějaké základní kategorizace převládat pořadové číslování. Většina pracovišť ať již ve vývojových týmech nebo v odborných útvarech totiž nepracuje najednou s příliš mnoha objekty tohoto typu, takže práce s čísly bez zřejmého významu, by u nich neměla činit problémy. Pokud se na některých pracovištích / správa dat, provoz VS/ soustřeďuje práce s více objekty, je zde zpravidla možnost dokonalejší dokumentace, které umožňuje rychlou orientaci. Naproti tomu v oblastech, kde se musí řada lidí rychle a často orientovat ve větších množstvích jmen /datové

položky, názvy odatavců v programu/, bude asi užitečné vhodné použití významových jmen, která umožní vytváření asociací na podporu paměti.

- B - Současný stav vývoje integračních procesů již vyžaduje, aby se při úvahách o tvorbě jmen důsledně uvažovalo s jejich počítačovým zpracováním, ať již v rámci informačních metasystémů, při využití různých generátorů, nebo v DATA DIRECTORY databankového systému.
- C - K tvorbě jmen je nutno přistupovat z pohledu celého systému, tj. nelze např. brát v potaz pouze hledisko programátora, ale je třeba vzít v úvahu též jejich stávající i budoucí využití jak v jednotlivých stádiích analýzy, tak v realizačních příručkách pro odborné útvary i provoz VS. S tím též samozřejmě souvisí používání jednotných jmen ve všech řešených subsystémech. Pokud se nepodaří takový přístup zajistit, nutně dojde vlivem duplicitních označení a s tím související nejednotnosti výkladu k řadě nedorozumění a problémům nejen při vývoji, ale i v průběhu realizace.

5. Návrh identifikace pro metasystém ASRP

Dále předložený návrh rámcově popisuje představu, jak by mohla být řešena identifikace v konkrétním informačním metasystému ASRP. Vychází přitom ze zkušeností, získaných s jednotným číslováním prvků v systému práce VS /5/, které je velice úspěšně prakticky beze změn již 10 let využíváno ve VS Ingstav při práci s počítačem TESLA 200, z různých pokusů o jednotné označování datových položek a z řady prací v oblasti počítačové tvorby dokumentace z poslední doby /3/, /4/, /7/, /8/, /9/.

Hlavní zásady:

- Informační metasystém bude jednotným způsobem popisovat celý systém, tj. jak VS, tak procesy a inf. soubory v OÚ

- S výjimkou lokálních jmen odstavců ve zdrojových programech /předp. jazyk COBOL/ a lokálně přímo v programu využívaných řídicích a pomocných datových položek, budou všechna jména důsledně normalizována
- Počínaje analýzou, přes programování, až po přípravu realizace v provozu VS i v odborných útvarech, bude veškerá formalizovatelná dokumentace tvořena programovým systémem pro manipulaci s daty /9/ s využitím počítače TESLA 200, příp. mikrografického inf. systému, a důsledně bude používáno normalizovaných jmen
- Hierarchické, resp. klasifikační číslování bude využito pouze pro označení útvarů v rámci podniku tak, aby vyjadřovalo organizační strukturu, a pro označení jednotlivých kategorií dokumentačních prvků ASŘP /tj. rozlišení, zda jde o program, projekt či soubor dat atd/
- Jména prvků v rámci jedné kategorie budou tvořena známkou kategorie a prostým pořadovým číslem /celkem 6 míst/
- Jména datových položek budou tvořena normovanými zkratkami jednotného formátu /patrně 2 písmena/ s tím, že u nejvíce frekventovaných položek bude snaha o využití asociace s jejich významem /např. ZA = závod, ST = stavba, CM = číslo materiálu apod./
- Při více výskytech shodné datové položky v jednom typu věty bude ke zkratce připojeno pořadové číslo
- Datové položky budou definovány buď jako globální /2 písmena/, tj. s definicí jednotně platnou v rámci celého systému, nebo jako lokální /1 písmeno/, definované pouze v rámci určitého typu věty nebo inf. souboru, takže při odvolávce na jejich jméno bude nutná kvalifikace. Dle dosavadních ověřování je oprávněný předpoklad, že v uvedené kombinaci by mělo být možno skutečně všechny potřebné datové položky pojmenovat. /Lokální jména označují širší pojmy, a další rozvinutí je zmíněným pořadovým číslováním/
- Jména inf. souborů budou zásadně vyjadřovat pouze jejich skladbu, označení generace bude písmenem jako doplněk.

Hierarchická skladba procesů:

- 01. systém /dokumentační jednotka/
- 02. subsystém /dokumentační jednotka/
- 03. dílčí zpracování /realizační jednotka/
programový chod na počítači, skupina úloh
ruční zpracování v podnikovém útvaru
- 04. činnost /provádění určité funkce/
automatická realizace v programu, podprogramu
či modulu, úloha
ruční realizace v podnikovém útvaru
- 05. operace
instrukce pro počítačovou operaci
instrukce pro ruční operaci

Hierarchická skladba informačních souborů:

a/ fyzicky

- 01. informační fond
- 02. oblasti
- 03. soubory
- 04. typy vět
- 05. datové položky
- 06. jednotlivé znaky

b/ logicky

- 01. informační fond
- 02. logické sestavení
- 03. typy vět
- 04. datové položky

Hlavní kategorie prvků:

- A - agendy /subsystémy, projekty z různých dekompozičních hledisek/, rozdělené na programové chody /doplňek = max 2 písmena/ a na dílčí ruční zpracování /doplňek = max 2 číslice/

F - činnosti /funkce na různé hierarchické úrovni,
podobně jako agendy, ruční i automatizované,
algoritmy /

W - typy vět /pojmenovaná skupina datových položek,
identifikovaná bez ohledu na nosič informací/

Datové položky /dle centrální definice, dělené na globální
a lokální/

Dále pak pomocí znaku kategorie a pořadového číslování
vytvářejena jména ^{pro} programy, podprogramy, soubory dle médií,
vlastní fyzické nosiče informací, útvary, doklady atd.

6. Závěry

Tvorba jmen má dopady v rámci celého systému, a nelze
ji proto řešit izolovaně.

Mnemotechniku je vhodné využívat řízeným způsobem a to
pouze tehdy, kde jsou její asociace obecněji platné, a její
výhodnost skutečně odůvodněná.

Základem při tvorbě jmen většiny kategorií prvků v do-
kumentaci ASŘP by mělo být označení kategorie a pořadové
číslování, bez uvažování dalších souvislostí.

Veškerá jména je nutné tvořit s ohledem na to, že se
mohou sama stát předmětem počítačového zpracování.

Základním pravidlem pro tvorbu jmen by měl být postup
dle co nejjednodušší filozofie, protože jen tak je možno
zajistit jednotnou realizaci v rámci celého systému.

Předložené náměty, ani uvedené závěry si v žádném přípa-
dě nečiní nárok ani na originalitu, ani na to, že jsou jediné
správné. Cílem příspěvku bylo upozornit na některé souvislos-
ti v dané problematice a pokusit se formulovat prakticky pou-
žitelná a ucelená řešení.

7. Literatura

- /1/ R.M.Armstrong: Modular Programming in Cobol, 1973
- /2/ J. Hradil, J. Klečka, T. Marek, K. Metzl: Formální úprava zdrojových programů, Standard č. 3, ZVT OKD Ostrava 1975
- /3/ R. Firla, F. Herout, B. Šimeček: Metodika projektu komplexní organizační a systémové analýzy, MLVH ČSR, 1977
- /4/ Sborník ze semináře "Výpočtové středisko '78", Košice
- /5/ V. Čevela: Identifikace jednotlivých prvků v systému práce VS, MAA 10/78, str. 399
- /6/ B. Lacko: Standardizované řešení v programech ASŘP, Softin 31, TOS Kuřim, 1978
- /7/ Sborník ze semináře "Používání jazyka COBOL", DT ČS-VTS Pardubice, 1979
- /8/ Sborník ze semináře "Výpočetní středisko '79", Karlovy Vary
- /9/ V. Čevela, F. Knotek: Programový systém pro manipulaci s daty, VS Ingstav Brno, 1980