

Datové rozhraní Intergram – obecný popis (verze 1.14)

1. Úvod

Předmětem tohoto textu je popis formy datových souborů předávaných společností Intergram. Tento text popisuje základní pravidla a obecné zásady, na základě kterých jsou datové soubory vytvářeny. Popis konkrétních datových souborů určených příslušné organizaci poskytující data společnosti Intergram je uveden v textech navazujících.

2. Základní pojmy

| | |
|-----------------------------|---|
| <i>Dávka:</i> | skupina dat předaná v rámci jednoho realizovaného datového přenosu. |
| <i>Datový soubor:</i> | prvek dávky – dávka se sestává ze dvou nebo více datových souborů |
| <i>Protokolační soubor:</i> | zvláštní datový soubor, nese informace o ostatních datových souborech v dávce obsažených. |
| <i>Datová věta:</i> | je prvkem datového souboru, je určena pro zaznamenání jedné skupiny údajů (definovaného záznamu). |
| <i>Údaj věty:</i> | prvkem datové věty (atribut datové věty). |

3. Použité konvence

Všechny datové soubory jsou textového typu. Každý datový soubor může obsahovat specifikované typy datových vět zapisovaných v pořadí definovaném příslušným strukturogramem. Každá datová věta specifikovaného typu má pevný počet údajů oddělených definovaným oddělovačem. *Jména datových souborů:* protokolační soubor má jméno pevně určené – **protokol.t**, ostatní soubory mají jméno volitelné, avšak pevnou extenzi – **.t**. *Poznámka:* kódy znaků dle ASCII jsou uváděny v dekadické číselné soustavě.

3.1. Formát datové věty

<atr0>,<atr1>,<atr2>,...,<atrN>;

kde:

| | |
|-------------------|---|
| <atr0> | ... identifikuje typ věty |
| <atr1> ... <atrN> | ... údaje (prvky věty) |
| ' ; ' | ... oddělovač údajů věty (znak čárka, dle ASCII: 44) |
| ' ; ' | ... oddělovač datových vět (znak středník, dle ASCII: 59) |

3.2. Vymezení množiny povolených znaků v datových souborech

Množina povolených znaků v rámci datové věty je dána pravidly pro zápis údaje datové věty (viz. kap. 3.3.), definovaným oddělovačem údajů věty (znak čárka) a definovaným oddělovačem datových vět (znak středník). Pouze mezi definovaným oddělovačem datové věty a datovou větou bezprostředně následující se mohou vyskytovat znaky:

- sekvence pro odřádkování: CR (carriage return), LF (line feed) (dle ASCII ... CR: 13, LF: 10)
- znak LF (dle ASCII: 10)
- znak TAB (dle ASCII: 9)
- znak mezera (dle ASCII: 32)

Jiné znaky nejsou v přenášených souborech povoleny.

3.3. Forma zápisu údaje (prvku věty)

Forma zápisu údaje odpovídá jeho datovému typu. Údaj může být typu: ARRAY(n), INT(n), REAL(n,m), CHAR.

- a) typ ARRAY(n) – alfanumerický údaj
Údaj typu ARRAY(n) musí být ohraničen znakem " (uvozovky). Hodnota n udává jeho maximální délku.
Povolené znaky: dle ASCII ... 32,33,35–255 tj. vše od ASCII 32 (tj. znak mezera) výše mimo znaku uvozovky (ASCII 34)
- b) typ INT(n) – numerický údaj, přirozené číslo
Hodnota n udává maximální délku (počet cifer) údaje.
Povolené znaky: dle ASCII ... 48–57 tj. znaky: '0'...'9'
- c) typ REAL(n,m) – numerický údaj, reálné číslo
Hodnota n udává maximální délku údaje (včetně desetinné tečky a případného znaménka minus), hodnota m udává maximální povolený počet desetinných míst.
Povolené znaky: dle ASCII ... 45,46,48–57 tj. znaky: '-', '.', '0'...'9'
Znak minus ('-') je povolen jen na první pozici, znak tečka ('.') se smí vyskytovat pouze jedenkrát, za ním může následovat číslo max. délky m.
- d) typ CHAR – jednoznakový údaj
Povolené znaky: dle ASCII ... 33,35–43,45–58,60–127 tj. vše v rozsahu 33–127 ASCII mimo znaku uvozovky a definovaných oddělovačů (znaky čárka a středník)

Prázdný údaj libovolného typu je reprezentován řetězcem nulové délky (ihned následuje příslušný oddělovač).

Údaj významu rok (např. ROK_VYR INT(4) ... rok výroby): není-li uvedeno století (rok < 100), doplňuje se dle aktuálního data.

3.4. Forma popisu datové věty

Popis každé datové věty je členěn do dvou bodů:

- a) struktura věty – podrobný popis údajů datové věty z hlediska jejich významu
- b) formáty údajů věty – popis formy zápisu údajů datové věty

Formát datové věty je popisován tabulkou s následujícími sloupci:

- atribut ... jméno atributu (údaje)
- typ ... typ atributu (dle kap. 2.3.)
- hodnota ... pro konstantní údaje v rámci věty, nebo pro údaje, které mohou nabývat hodnot z definovaného výčtu

Povinně vyplněný údaj: údaj, jehož jméno atributu je označeno znakem "*", musí být vždy vyplněn.

Prvním atributem každé datové věty je atribut IndVeta (typ CHAR), který slouží k identifikaci typu věty (nabývá tedy pro každý typ datové věty konstantní hodnoty). Tento atribut je vždy uváděn jen v popisu formátů údajů věty (body b)).

3.5. Příklad zápisu datové věty

Definujme datovou větu V, jejíž údaje mají následující formát:

| | atribut | typ | hodnota |
|---|---------|-----------|---------|
| * | IndVeta | CHAR | 'V' |
| * | JMENO | ARRAY(12) | |
| | IND | CHAR | |
| | CISL | INT(6) | |
| | HONORAR | REAL(7,2) | |

Znak '*' před jménem atributu značí údaj (jak je uvedeno v kap. 2.4. b)), který musí být povinně vyplněn.

Dále mějme skupinu údajů:

| Jméno | Indikátor | Číslo | Doplatek |
|--------------|-----------|-------|----------|
| Petr Polák | A | 12345 | 100,- |
| Václav Kozel | B | - | 57,60 |
| Eva Novotná | - | - | - |
| Aleš Kopecký | C | 12 | 200,- |

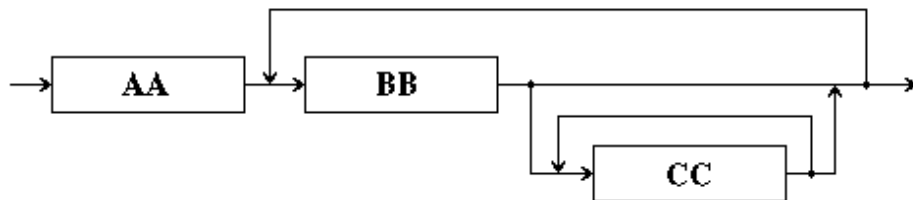
Uvedená skupina údajů se dle výše popsaných pravidel zapíše do datového souboru takto:

```
v,"Petr Polák",A,12345,100;  
v,"Václav Kozel",B,,57.60;  
v,"Eva Novotná",,,;  
v,"Aleš Kopecký",C,12,200;
```

3.6. Forma popisu datového souboru

Každý datový soubor je popsán pomocí strukturogramu. Strukturogram definuje pro každý typ datového souboru množinu povolených typů datových vět a povolenou následnost jednotlivých typů datových vět.

Příklad strukturogramu: mějme definovány tři druhy vět, větu typu AA, typu BB a typu CC. Povolené pořadí vět v datovém souboru pak definujeme strukturogramem:



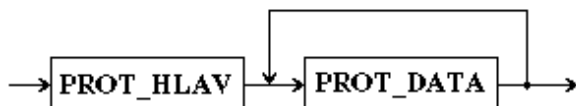
Z uvedeného strukturogramu pak vyplývají následující pravidla:

- věta typu AA se v datovém souboru vyskytuje pouze jednou, a to jako první v pořadí (hlavička souboru). Za větou typu AA musí následovat věta typu BB
- po větě typu BB může následovat věta typu CC, další věta typu BB nebo konec souboru
- po větě typu CC může následovat další věta typu CC, další věta typu BB, nebo konec souboru
- věta typu AA je větou 1.úrovně, věta typu BB – 2.úroveň, věta typu CC – 3.úroveň. Věty bezprostředně následující za větou úrovně o jeden stupeň vyšší jsou chápány jako další informace vztažené k této větě.

4. Protokol dávky

Je textový soubor fixního jména – **protokol.t** – popisující obsah dávky. Obsahuje dva typy datových vět.

4.1. Struktura protokolačního souboru (protokol.t)



4.2. Definice datových vět protokolu dávky

4.2.1. Věta: PROT_HLAV

Hlavička protokolačního souboru – základní údaje o dávce.

a) struktura věty

| | |
|------------|---|
| TypKodStr | Údaj identifikující použitou kódovou stránku pro textové údaje: M – MS Windows 3.1 EE Character Set (1250) |
| Dv, Mv, Rv | tj. den, měsíc, rok ... datum vytvoření dávky |
| Vysilatel | Jméno vysilatele |
| Ulice | |
| Místo | ... adresa |
| Telefon | |

b) formáty údajů věty

| | atribut | typ | hodnota |
|---|-----------|-----------|---------|
| * | IndVeta | CHAR | 'z' |
| * | TypKodStr | CHAR | 'M' |
| * | Dv | INT(2) | |
| * | Mv | INT(2) | |
| * | Rv | INT(4) | |
| * | Vysilatel | ARRAY(25) | |
| | Ulice | ARRAY(35) | |
| | Místo | ARRAY(35) | |
| | Telefon | ARRAY(12) | |

c) příklad

z,M,1,4,2023,"Český rozhlas","Vinohradská 12","Praha 2","274051 I.425";

tj. kódová stránka: CP 1250; datum vytvoření: 1.4.2023; od: Českého rozhlasu

4.2.2. Věta: PROT_DATA

Pro každý datový soubor, který je součástí dávky, existuje v protokolačním souboru právě jedna věta tohoto typu.

a) struktura věty

| | |
|---------|--|
| Soubor | Jméno příslušného datového souboru |
| Dat_typ | Typ datového souboru: A – typ TV C – typ rozhlas |

| | |
|----------|---|
| | Popis datových souborů je uveden v textech určených pro jednotlivé organizace poskytující společnosti Intergram data. |
| Typ | Nabývá hodnot: N – nově pořízená data O – oprava dříve zaslaných dat |
| | Je-li Typ – N, pak veškerá data v datovém souboru jsou považována za data nová; tzn. ukáže-li se, že stejná data již byla přijata dříve, je to považováno za chybu. |
| | Je-li Typ – O, pak veškerá data v datovém souboru se považují za opravná vzhledem k již dříve přijatým datům. Neexistují-li odpovídající dříve přijatá data, je to považováno za chybu. |
| Dv,Mv,Rv | tj. den, měsíc, rok ... datum vytvoření souboru (Datum_v) |
| Do,Mo,Ro | tj. den, měsíc, rok ... počáteční datum za soubor (Datum_od) |
| Dd,Md,Rd | tj. den, měsíc, rok ... konečné datum za soubor (Datum_do) |

Poznámka: údaje Datum_od – Datum_do vymezují období, za které jsou data v rámci příslušného souboru předávána.

b) formáty údajů věty

| | atribut | typ | hodnota |
|---|---------|-----------|---------|
| * | IndVeta | CHAR | 'y' |
| * | Soubor | ARRAY(14) | |
| * | Dat_typ | CHAR | 'A','C' |
| * | Typ | CHAR | 'N','O' |
| * | Dv | INT(2) | |
| * | Mv | INT(2) | |
| * | Rv | INT(4) | |
| * | Do | INT(2) | |
| * | Mo | INT(2) | |
| * | Ro | INT(4) | |
| * | Dd | INT(2) | |
| * | Md | INT(2) | |
| * | Rd | INT(4) | |

c) příklad

y," cro0323.t",E,N,16,4,2023,1,3,2023,31,3,2023;

tj. soubor: cro0323; typ: C – soubor typu rozhlas, N – nová data; datum vytvoření: 16.4.2023; data za období: od 1.3.2023 do 31.3.2023

5. Sestavení dávky

Jak je uvedeno výše, dávka se sestává ze souboru protokolačního a souborů datových. Pro zabezpečení jednoznačného způsobu zpracování souborů dávky je nutné definovat následující pravidla a omezení.

Více souborů stejného typu v rámci dávky: dávka může obsahovat více datových souborů stejného typu Údaje v těchto souborech na sebe musí z časového hlediska navazovat, tzn. je nutné zabezpečit časovou návaznost období (Datum_od – Datum_do) za která jsou jednotlivé soubory sestavovány.

Vícenásobný výskyt logicky stejných údajů: vícenásobný výskyt vět popisujících stejnou skupinu údajů v rámci datových souborů stejného datového typu se považuje za chybu.

6. Závěrečná poznámka

V nových verzích datových rozhraní je popisován pouze aktuální stav (do popisu nebudou zahrnovány popisy starších verzí). Programové vybavení bude respektovat starší verze datových rozhraní, pokud v textech popisu nebude uvedeno jinak.

INTERGRAM 24.5.2024