

Helena Wallenfelsová, Jitka Štěpánová

katedra informatiky, PEF, tel: (02) 3382277
Vysoká škola zemědělská, 165 21 Praha 6 - Suchbátka

Anotace:

Článek informuje o procesu standardizace IS ve světě i u nás. Vymezuje metodiku tvorby datových prvků pro IS v zemědělství a navrhuje základní atributy těchto datových prvků.

Summary:

The article gives information about process of information system standardization. It describes the principles and rules for identification of data elements and indicates basic data elements attributes.

Klíčová slova:

standardizace informací, EDIFACT, normy EN, normy ISO, datový prvek, JÚZO, SIS, atribut datového prvku

Key words:

standardization of data, EDIFACT, standard ISO, standard EU, data element, Uniform Data Base, Government Information System, attribute of data element

Trvale roste podíl informací, které si uživatelé vyměňují elektronickou cestou. Informace, které jsou předmětem této výměny, mohou být efektivně využity jen za podmínky, že jednotlivé datové prvky jsou na straně odesílatele i příjemce shodně vymezeny.

Snaha vytvářet standardizované informace pro potřeby elektronické výměny dat byla na mezinárodní úrovni zahrnuta do projektu EDIFACT [1].

Model systému EDIFACT je dnes otevřený vůči národním normám pro sestavování a přenos zpráv. Jde tedy o to, aby národní normy byly vytvořeny, ať již přijetím mezinárodních standardů nebo definováním vlastních. Závaznost přijetí mezinárodních norem se týká pouze tzv. norem Evropských.

Evropské normy (EN) jsou normy, které zásadně platí jako národní normy ve všech účastnických státech CEN.

Na rozdíl od EN existují ještě normy mezinárodní (tzv. ISO normy), které vypracovává Mezinárodní organizace pro normalizaci ISO.

Základním rozdílem mezi normami ISO a CEN je, že ISO vydává normy zásadně jako doporučení, zatímco normy CEN jsou závazné pro členské státy EU.

Zavádění projektu EDIFACT v České republice má na starosti Ministerstvo hospodářství a Česká obchodní a průmyslová komora.

Úkolem pracovníků, kteří se zabývají aplikacemi informačních technologií v zemědělství, je i tuto oblast učinit způsobilou využívat elektronické komunikace na mezinárodní úrovni.

Po přípravných koncepčních krocích bylo v letošním roce přikročeno k etapě vytváření datových slovníků, které jsou nezbytným základem pro následné návrhy standardizovaných informačních systémů v zemědělství.

Datové standardy pro vyváření obsahu informačních systémů v zemědělství je třeba tvořit v návaznosti na normy EDI a s ohledem na horizontální a vertikální integraci dat.

Celou problematiku lze rozdělit do čtyř hlavních oblastí:

názvosloví, tvorbu číselníků, tvorbu datových prvků, tvorbu sborníků datových prvků.

Metodika tvorby datových prvků není prozatím stabilizovaná v České republice ani ve světě. Při tvorbě datových prvků se lze řídit jen některými vzory existujících sborníků: ČSN EN 27373 Sborník datových prvků pro obchod, katalogy JÚZO (viz. [1]), Standard Státního informačního systému (SIS) k územní identifikaci, Katalog datového standardu SIS. Na mezinárodní úrovni je metodika rozpracována v návrhu ISO 11179. Také je již navržen šestimístný formát kódu pro sborníky prvků (ICSI - International Coding Scheme Identifier), a to v návrhu ISO 7826. Problémem však zůstává, jak postupovat v současné době, kdy obě normy jsou teprve v návrhu a lze v nich očekávat změny.

V návrhu ISO 11179-3 je popsáno 23 atributů, jimiž musí či může být datový prvek charakterizován (M označuje povinné atributy, C podmíněné a O volitelné), viz. Tabulka 1:

Kategorie atr.	Název	Závaznost
Identifikační	Název datového prvku	M
	Identifikátor	C
	Verze	C
	Synonyma	O
	Kontext	O
Definiční	Definice	M
	Přípustné instance	M
Relační	Kvalifikátor čeho	C
	Kvalifikován čím	C
	Klasifikační schéma	O
	Klíč(e)	O
	Relační prvky	O
Reprezentační	Kategorie reprezentace	M
	Forma reprezentace	M
	Datový typ hodnot datových prvků	M
	Min. rozsah hodnot datových prvků	M
	Max. rozsah hodnot datových prvků	M
	Plán reprezentace	C
	Přípustné hodnoty datových prvků	C
Administrativní	Zodpovědná organizace	O
	Registrační status	M
	Předkládající organizace	O
	Komentář	O

Kromě těchto základních atributů definuje uvedená norma ještě řadu doplňujících atributů, které mohou být užity podle uvážení tvůrců datových slovníků.

Pro tvorbu našich národních datových slovníků byly prozatím vybrány jen některé atributy, doplněn byl v souladu s národní zvyklostí atribut Akronym (v seznamu doplňujících atributů normy ISO 11179-3 uveden jako Zkrácené jméno), Informace o číselníku, Právní předpisy a Výskyt, viz. Tabulka 2:

Název atributu	Závaznost
Název datového prvku	M
Identifikátor	C
Verze	C
Akronym	C
Definice a/nebo popis	M
Formát	M
Informace o číselníku a/nebo číselník	C
Poznámka	C
Právní předpisy	C
Výskyt	C
Synonyma	C

Název datového prvku

Metodika tvorby názvů datových prvků je v návrhu ISO 11179-5. Vzhledem k jazykovým rozdílům mezi češtinou a angličtinou nelze tento návrh u nás aplikovat bez úprav.

Identifikátor datového prvku

Podle návrhu ISO 11179-3 není povinným atributem, ale je povinný v mezinárodních sbornících datových prvků a tam, kde název datového prvku není jednoznačný. Identifikátor, na rozdíl od akronymu, má být vyjádřen kódem nazávislým na jazykových hlediscích. Obecné zásady pro tvorbu a přidělování těchto kódů jsou v návrhu ISO 11179-5. V České republice přicházejí v úvahu dvě možnosti. Tvorba identifikátorů podle zásad ČSN EN 27372, a nebo podle zásad JÚZO.

Verze datového prvku

Předpokládá se, že bychom u nás mohli využít označení verze ve formátech standardu UN/EDIFACT.

Akronym datového prvku

Akronym by měl být pracovním zkráceným názvem nebo zkratkou názvu datového prvku.

Definice a/nebo popis datového prvku

Definice je nejdůležitějším atributem a při její tvorbě je třeba vycházet z analýzy informačních potřeb (viz závěr článku).

Metodika tvorby definice datového prvku je v návrhu ISO 11179-4.

Formát datového prvku

Formát datových prvků je popsán v návrhu ISO 11179-3.

Informace o číselníku nebo číselník

V případě, že je v popisu datového prvku použit identifikátor, je třeba uvést informaci, z jakého číselníku tento identifikátor je. Lze přitom užít odkaz na normu ČSN, jiný veřejně dostupný číselník), při malém rozsahu číselníku jej lze uvést přímo u datového prvku.

Poznámka

Slouží k doplnění informací, které nelze zařadit k ostatním atributům.

Právní předpisy

Zde se citují předpisy, které datový prvek zavádějí, definují, používají nebo určují způsob jeho používání.

Výskyt datového prvku

Tento atribut spolu s atributem Zdroj údaje ze zákona používá SIS.

Synonyma

Odlišné názvy téhož datového prvku, které je zvykem užívat v různých oborech nebo souvislostech.

Jako podklad ke stanovení datových prvků pro tvorbu datových sborníků oblasti zemědělství bude třeba provést analýzu

- informačních potřeb dané oblasti
- obecně závazných právních předpisů
- mezinárodně platných norem
- již zpracovaných katalogů prvků
- informačního systému a systému řízení příslušné organizační úrovně.

Na závěr poznámku k definici datového prvku. Datový prvek je definován jako nedělitelná jednotka dat, pro kterou identifikace, význam, reprezentace a přípustné hodnoty jsou specifikovány prostřednictvím množiny atributů. Je však nutné si uvědomit, že charakter "nedělitelný" se vztahuje vždy pouze k určité organizační a sémantické úrovni informačního procesu, tedy to, co je na určité úrovni chápáno jako jednoduchý element, může být pro jinou úroveň složitým objektem vyžadujícím popis celou řadou jednoduchých datových prvků. A toto je třeba také mít na zřeteli při vytváření datových slovníků.

Literatura:

[1] Wallenfelsová H., Aspekty elektronické výměny dat, čl., Sborník konference VŠZ, Praha 1992

[2] Draft International Standard ISO/IEC DIS 11179

[3] Roleček J., Zabezpečení projektu elektronické výměny dat, úkol č. RU/0702/940