

Elektronsko poslovanje in standardi

Branko Šafarič
GS1 Slovenija
branko.safaric@gs1si.org

Povzetek

Medsebojna izmenjava podatkov o asortimanih je ključni element preskrbovalnih verig, zato je zelo pomembno, kako so podatki o izdelkih – predmetu trgovine – vpisani v podatkovne baze. Praksa je pokazala, da imajo podjetja, ki so razvijala svoje informacijske sisteme, pogosto težave z usklajevanjem podatkov prav zaradi neustrezne strukture podatkovne baze. Članek se navezuje na sistem standardov GS1 (po starem EAN) za označevanje artiklov in sistem sinhronizacije podatkov prek GDSN (Global Data Synchronization Network), ki je prav tako eden od temeljnih standardov sistema GS1. Članek pokaže na nekaj problemov v zvezi strukture zapisa podatkovne baze.

Ključne besede: elektronsko poslovanje, elektronska izmenjava podatkov, podatkovne strukture baz podatkov, sinhronizacija podatkov

Abstract

E-COMMERCE AND STANDARDS

The exchange of assortment data is the key element of supply chains, therefore it is of vital importance how the data about products – the object of commerce – are registered in the database. The practice has shown that the enterprises developing their own information systems often experience difficulties in adjusting the data owing to inadequate structure of database. The paper is referring to a system of standards GS1 (formerly EAN) for labeling the articles and to a system of data synchronization with GDNS (Global Data Synchronization Network), which is also one of the basic standards within the system GS1. The paper points out a few problems regarding the structure of database records.

Keywords: e-Commerce, Electronic Data Interchange, Master Data Database Structure, Data Synchronization

1 Struktura baze izdelkov

1.1 Uvod

Standardi so v našem življenju prisotni tako zelo, da postajajo transparentni. To, da vsi uporabljamo enako električno napetost, opazimo šele, ko pridemo na področje, kjer imajo drugačne vtikače ali drugačano napetost. Hočem reči, da prave standarde opazimo, šele ko jih ni.

Podobno velja v preskrbovalnih verigah. Ko stvari tečejo gladko, se prisotnosti standarda tako rekoč ne zavedamo.

Pri optimizaciji poslovnih procesov vedno bolj prihaja do združevanja različnih procesov, pri čemer moramo na neki način povezovati različne procese ali računalniške rešitve. Tipični primer je povezovanje finančnega poslovanja z materialnim tokom.

Da bo analogija bolj konkretna, si lahko predstavljamo probleme, ki jih imamo pri združevanju dveh različnih informacijskih sistemov. Kako težko je vzpostaviti interakcijo med različnimi informacijskimi sistemi, po možnosti različnega izvora, vedo samo tisti, ki so to kdaj počeli. Pa ni toliko problem tehnologija (npr. različni programski jeziki ali različne podatkovne baze, se pravi različne platforme). Tehnologija je manjša težava. Najhujši problem je vsebinski –

različni šifranti, različna uporaba identifikacijskih oznak ipd. Pa seveda tudi različna struktura zapisa v sami podatkovni bazi.

Preprost primer tega je recimo telefonska številka. Sestavljena je iz klicne številke države, omrežne številke in naročnikove telefonske številke. Pa seveda ne smemo pozabiti na možnost zapisa več števil in njihovega tipa (mobilna, službena, zasebna ...).

Vsak projektant informacijskega sistema ima svoje predstave o tem, kakšen zapis telefonske številke je najbolj primeren. Nekateri bodo rezervirali dolgo besedilno polje, v katerega bodo vpisali dolg niz znakov, v katerem bodo vse številke skupaj z morebitnimi opisi in komentarji.

Drugi bodo naredili posebno tabelo, v kateri bo v enem zapisu mogoče vpisati več števil. Ali bodo pri tem upoštevali tudi številko države kot ločeno polje ali ne, je stvar odločitve, ki je enkrat taka, drugič drugačna.

Moj namen ni, da bi zagovarjal strogi puritanski sistem pri oblikovanju podatkovne baze, v kateri mora imeti vsaka entiteta svojo tabelo, v kateri ne sme prihajati do redundanc ali kaj podobnega. Oblikovanje strukture podatkovne baze je kot avtorsko delo, vsak se mora držati pravil procesa in ga optimirati tako, da

je strukturno pravičen in maksimalno odziven. Rad bi samo na kratko prikazal problematiko definicije podatkovne strukture za artikle.

Prav vsako podjetje, ki nastopa v procesu preskrbovalne verige, ima v svojem sistemu vpisane svoje artikle – izdelke ali storitve na svoj način.

Praksa vpogleda v načine zapisa teh podatkov s strani GS1 je pokazala, da je tu veliko napak, ki povzročajo probleme pri povezovanju različnih sistemov, tako znotraj istega podjetja, še bolj pa pri povezovanju z drugimi podjetji – procesi (kot je recimo globalni katalog artiklov).

Ta prispevek je namenjen vpogledu v problematiko zapisovanja podatkov o artiklih s stališča tipa artikla, uporabe artikla v maloprodaji, transportu in elektronskem poslovanju.

Ta problematika je mnogo hujša, kot jo lahko opišemo v tem prispevku, zato bo tu samo orientacijski pogled.

1.2 Je zapis o artiklu preprost?

Ko si v prvem trenutku predstavljamo, kako zapisati podatke o artiklu, se nam to zdi tako preprosto, da je o tem kar škoda razmišljati ali izgubljati besede. Saj je to samo spisek, v katerem so neki podatki, ki opisujejo artikel, ne? Saj je to samo šifra artikla, njegov opis, cena, pa mogoče še kaj, ne?

A resničnost je povsem drugačna in praksa neprijetno hujša. Problematika strukture podatkovne baze je mnogo kompleksnejša kot izgleda na prvi pogled. V kombinaciji z zahtevami o skladnosti s standardi GS1 pa je to še mnogo huje.

Že samo retorično vprašanje »kaj pa sploh še lahko spada k podatkom o artiklu« je poseben problem. Standardi GS1 (GDS – Global Data Synchronization) in globalni katalogi artiklov, kot so recimo GXS, ISync, Sinfos ipd., se ukvarjajo s tem in v praksi se to izkazuje kot izjemno hud problem.

Poleg najosnovnejših podatkov o artiklu, ki vsakemu od nas takoj padejo v glavo, je množica drugih, ki jih potrebujemo v logistiki, naročanju, posredovanju, pakiranju ... V standardu je približno 150 vseh atributov = podatkov o artiklu. Seveda niso vsi nujni, nekaterih ne potrebujemo, nekateri se med seboj celo izključujejo, nekateri se pogojujejo, nekateri imajo vsebinske omejitve ... Kateri točno so podatki, ki jih potrebujejo vsi, tudi ni povsem jasno. Ta najnujnejši presek podatkov je približno 30 atributov.

Pri reševanju problematike in obdelavi podatkov o artiklih se srečamo s tremi področji:

- področje fizičnega zapisa na artikel oziroma transportno enoto – označevanje artiklov,
- področje zapisa v podatkovno bazo,
- področje elektronskega poslovanja.

1.3 Katere podatke potrebujemo

Kogar koli vprašamo, katere attribute (podatke) si predstavlja, da spadajo k enemu artiklu, nikoli ne dobimo povsem istih odgovorov. Struktura podatkovnega zapisa o artiklu je v svojih temeljih sporna in odvisna od artikla do artikla. Zato ni čudno, da podjetja, ki sama razvijajo svoje informacijske sisteme, pogosto pridejo do medsebojno povsem nezdržljivih podatkovnih struktur.

Izbira nujnih atributov, ki v minimalni definiciji opišejo artikel, je bolj ali manj jasna, a minimalni nabor artiklov je tudi prav to – minimalen. Zato je za praktično rabo nezadosten.

1.4 Glavna identifikacija artikla

Vsak artikel mora imeti svojo unikatno identifikacijo, sicer zapis v podatkovni bazi ni kaj prida uporaben. Na prvi pogled se zdi idealno, da za unikatno identifikacijo – ključ zapisa v podatkovni bazi – uporabimo GTIN, tj. identifikacijsko številko GS1. Identifikacijske številke GS1 so globalne in univerzalne, zato so v svoji funkciji unikatnega ključa tudi idealne pri prenosu podatkov med različnimi informacijskimi sistemi in podatkovnimi bazami – lahko rečemo pri stiku z zunanjim svetom.

A pri internih obdelavah in reševanju problematike referenčne integritete svojih podatkovnih baz je optimalna unikatna identifikacija INTERNA identifikacijska številka. Se pravi neka poljubna identifikacija, ki jo določi vsako podjetje zase.

Zakaj taka heretična trditev, ko pa vsi v GS1 trdimo, da imamo najboljši sistem za označevanje na svetu? Žal, razlaga ni povsem preprosta.

Velja tudi, da ta definicija ni v vseh primerih absolutno veljavna. Prav lahko se namreč zgodi, da je GTIN za nas optimalna identifikacija, ker v naših podatkovnih bazah pač ni nobene hierarhije med artikli. A pri večjih zbirkah artiklov, v katerih uporabljamo tudi družine izdelkov, v katerih imamo vpisane tudi izdelke s spremenljivo vsebino, v katerih nam procesi narekujejo potrebo po nekakšni nadrejeni identi-

fikaciji, iz katere izvirajo tudi drugi artikli, je interna številka najboljši možni identifikator.

Značilnost problema identifikacije je pri artiklih s spremenljivo vsebino (siri, meso ipd. – vse kar tehta-mo). Le kako naj identifikacijo artikla s spremenljivo vsebino označimo kot unikatni ključ? To ni mogoče. To pomeni, da v podatkovni bazi potrebujemo neko generično identifikacijo, ki je hierarhično nadrejena vsem izpeljankam tega artikla. Izpeljanke so lahko drugačna pakiranja, drugačna embalaža, drugačne oznake ipd.

! Pomembno (celo zelo pomembno): GTIN mora biti v vseh podatkovni bazah zapisan kot 14-mestno alfanumerično polje, s predznačenimi ničlami in prirejeno kontrolno številko. Se pravi GTIN-8 za recimo žvečilni gumi 42101925 v podatkovni bazi zapišemo kot: 00000042101925.

Interna identifikacija je samo pomoč, samo nekakšno sistemsko reševanje praktičnega problema/dejstva, da so v praksi artikli, ki imajo več vzporednih identifikacijskih števil. Abstraktno lahko sicer rečemo, da za interno označevanje lahko tudi uporabimo GTIN, vendar je potem treba ločiti namen notranjih in zunanjih identifikacij.

Tabela 1: Uporaba identifikacije po načinu uporabe zapisa

	Označevanje artikla (črtna koda)	Označevanje transportne enote (logistika)	Uporaba identifikacije v eCOM
Interna identifikacija	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> *
Variabilna vsebina	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
GTIN	<input checked="" type="checkbox"/> **	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

* Interno identifikacijo v elektronskih sporočilih uporabljamo kot vzporedno – podrejeno glavni identifikacijski oznaki, ki je praviloma GTIN.

** Artikle s spremenljivo vsebino označujemo z GTIN samo takrat, ko so pakirani v vedno isti količini (npr. sir vedno po 200 g). V tem primeru tak artikel obravnavamo kot vsak drug (se pravi, kot da ne bi bil spremenljive vsebine).

1.4.1 Problematika vpisa spremenljive identifikacije v podatkovno strukturo

Vedno mora veljati pravilo, da lahko na podlagi vsake identifikacijske številke, ki se pojavlja v kakršni koli obliki zapisa (ali samo številka, ali črtna koda, ali oznaka EPC-RFID ...), lahko identificiramo artikel. Ne smemo dovoliti duplikatov niti neidentificiranih artiklov. Običajno takšno vrsto omejitve postavlja že definicija podatkovne strukture v bazi (referenčna integriteta).

Pri obdelavi spremenljive količine se moramo zavedati razlik pri identifikaciji glede na mesto uporabe – na terminalih POS, v logistiki ali v elektronskih sporočilih.

Za spremenljivo vsebino je značilno, da del identifikacije sestoji iz fiksne in spremenljive komponente. Fiksno komponento, ki določa konkretni artikel (npr. sir ementalec določenega proizvajalca), praviloma dodeli organizacija GSI. Drugi, spremenljivi del pa je odvisen od količine konkretnega artikla. To seveda pomeni, da takšne identifikacije ne moremo v celoti zapisati v podatkovno bazo, temveč moramo poiskati parcialni ključ – fiksno komponento identifikacije. Terminali POS oziroma njihova programska oprema mora znati ločevati identifikacije spremenljive vsebine od običajnih in temu primerno spremeniti način iskanja podatkov o artiklu (uporabi drugačen ključ – običajno fiksni del, ki mu sledijo same ničle).

Kadar pa uporabljamo artikle s spremenljivo vsebino v logistiki ali v elektronskih sporočilih, moramo uporabljati generično identifikacijo.

1.5 Nadrejenost in podrejenost

Hierarhija artikla oziroma družina artikla je prisotna tako rekoč pri vsakem prodajnem artiklu. Tipični primer je različno pakiranje fizično istega artikla. Ista marmelada je lahko pakirana v različnih embalažah in vsaka od njih ima praviloma svoj – različen GTIN. Seveda to velja tudi za skupinska pakiranja – zaboj z več (vedno istim) številom (vedno enakih) artiklov ima svoj GTIN. To pomeni, da so ti artikli med seboj vsebinsko povezani, saj gre vedno za isto vrsto marmelade. Pri definiciji artikla v podatkovni bazi moramo imeti tudi eksplicitno možnost, da se navežemo na predhodni in naslednji artikel v hierarhiji pakiranja.

Vsaka hierarhija ima nekje na začetku svoj konec oziroma začetek. (Winston Churchill je imel baje na svoji pisalni mizi tablico z napisom: »Tu se prelaganje odgovornosti konča.«) Jasno je, da imamo zato tudi nekje začetek verige soodvisnih artiklov oziroma njihovih zapisov. Prvi artikel v hierarhiji je običajno tisti, ki je hkrati tudi najmanjša logična enota – po navadi je to maloprodajni artikel.

V zapisu o artiklu moramo imeti oznako nadrejenosti in podrejenosti, kar je v standardu GSI tudi predvideno (slika 1).

Povezavo navzgor in navzdol seveda označimo kar z identifikacijsko oznako ustreznega zapisa. To je lahko GTIN ali pa interna oznaka artikla. V zapisu elek-



Slika 1

tronskega kataloga artiklov hierarhijo praviloma označujemo z GTIN-i nadrejenih oziroma podrejenih artiklov.

1.6 Povezava med interno in globalno identifikacijo

Podatkovno bazo o artiklih uporabljamo pretežno za to, da kar najhitreje najdemo podatke o artiklu. Tipični primer je blagajna (terminal POS) v trgovini. Ob vsakem odčitavanju artikla je treba ugotoviti, za kateri artikel gre, in izbrati ceno, da se postopek na blagajni lahko nadaljuje. Na sliki 2 je razvidno, da (v tem predlogu rešitve) vedno iščemo (ali po interni identifikaciji ali po GTIN) po tabeli »Artikel_Sekundarni_ID«, na

podlagi tega dobimo drugi identifikator (če smo iskali po interni številki, dobimo GTIN ali obratno). Z obema ključnima podatkomah lahko res enolično dostopamo do vseh nastopanj – do cele družine artiklov s skupno interno identifikacijo.

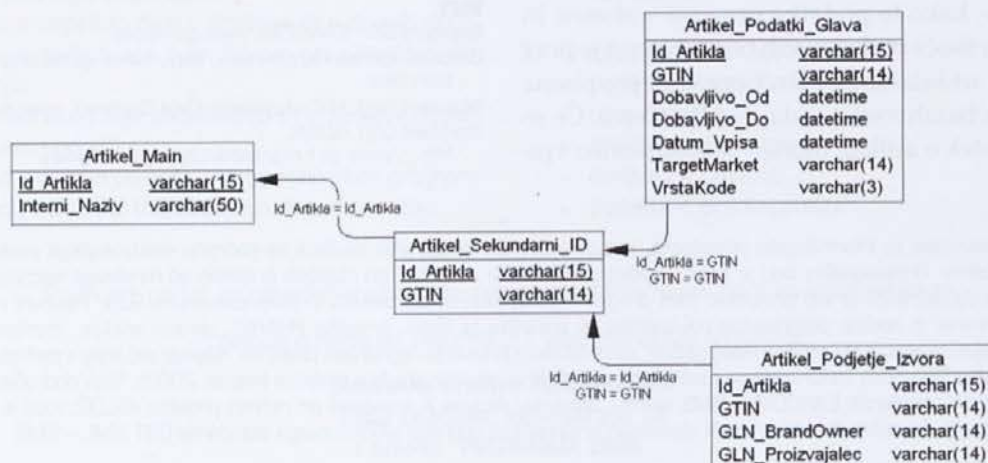
Mnogo laže kot opišemo, lahko to ponazorimo s primerom.

Recimo, da imamo sir ementalec, ki ima interno identifikacijo SIR_123. Ta sir se lahko pojavlja v individualnem pakiranju s spremenljivo vsebino, kjer ima GTIN 2611111xxxxC (saj vemo: 26 pomeni teža, 11111 je identifikacijska številka tega artikla, ki jo dodeli GS1 Slovenija, xxxx je prostor, kamor se vpiše teža na tri decimalke in C je seveda kontrolna številka). Ta sir tudi pošiljamo/prejemamo v zabojih po približno 3, 5 in 20 kg. Vsaka od teh vrst pakiranja ima svoj GTIN.

Tabela 2: Vrste oznak za interni artikel SIR_123

Vrsta pakiranja	Oznaka	Zapis v bazi
Variabilna vsebina	2611111xxxxC	Id_Arikla: SIR_123 GTIN: 02611111000001
Generični GTIN	3831111100009	Id_Arikla: SIR_123 GTIN: 03831111100009
Zaboj 3kg	3831111100030	Id_Arikla: SIR_123 GTIN: 03831111100030
Zaboj 5kg	3831111100054	Id_Arikla: SIR_123 GTIN: 03831111100054

Jasno je, da gre v vseh primerih za ISTI artikel – sir ementalec. Očitno je, da ima vsaka vrsta pakiranja svoj GTIN, kar pomeni, da brez nadrejene, skupne



Slika 2: Groba skica podatkovnih atributov artikla

šifre, ki ji rečemo primarna identifikacija, ne moremo združiti različnih nastopanj tega artikla, kar nujno potrebujemo zaradi analiz naročanja in prodaje oziroma planiranja proizvodnje/prodaje.

Drugi problem je to, da spremenljive oznake ne moremo uporabljati za transportne enote in v elektronskem poslovanju, to pa posledično pomeni, da potrebujemo neko generično številko. Generična številka tu pomeni neki GTIN, ki se nikoli ne pojavlja v obliki nalepke, ampak vedno samo kot pomožna identifikacija v elektronskem poslovanju.

1.6.1 Označevanje artiklov s spremenljivo vsebino v elektronskih sporočilih

Pri elektronski obdelavi artiklov s spremenljivo vsebino moramo uporabljati generično identifikacijsko številko – se pravi neki GTIN, ki se ga sicer nikoli ne uporablja v obliki črtne kode. Natančneje to govori formulacija iz standarda EANCOM, poglavje 4.1.1.

1.6.2 Kvaliteta podatkov

Mimogrede se splača omeniti tudi kakovost podatkov. Kakovost podatkov pomeni ne samo pravilnost temveč tudi ažurnost. V sistemih z več tisoči, celo deset tisoči artiklov je tovrstno zahtevo precej težko izpolniti, saj ažurnost zahteva stalno, in res stalno delo – nadzor nad podatki, to pa pomeni tudi angažiranje ustreznega števila ljudi in s tem povezane stroške. Idealni sistem je, kadar za kakovost podatkov o artiklu skrbijo »lastniki artikla«. Lastnik artikla je lahko predvsem proizvajalec konkretnega artikla ali pa morda kakšni zastopniki oziroma posredniki.

Idealno je torej, da vse spremembe, dopolnila, popravke ... artikla vpiše lastnik. Ostane samo »malenkosten problem«, kako te podatke prenesti v stotine in morda tisoče in tisoče podatkovnih baz, v katerih je prav tako vpisan ta artikel. Na primer, Coca Cola je vpisana v podatkovnih bazah vseh prodajalcev po svetu. Če se spremeni podatek o artiklu, morajo to spremembo vpi-

sati prav vsi udeleženci. Če podatkov ne posodabljammo redno, se nam v podatkovni bazi nabere balast nekakovostnih in zastarelih podatkov, ki jih je zelo težko prečistiti. Izkušnje večjih podjetij kažejo, da je kakovost podatkov lažje vzdrževati, kot pa periodično očistiti.

2 Sklep

O načinu zapisa artiklov bi lahko napisali knjigo, ne samo članek. Večine problemov se tem članku sploh nismo dotaknili (npr. klasifikacije, načinov prenosa podatkov, sinhronizacije, povezave na elektronski katalog artiklov ...). Namen članka je pretežno ozaveščanje o problemu nadzora nad podatkovno bazo artiklov, ki mora biti kakovostna, če hoče podjetje razmišljati o vstopu v svet sinhronizacije, kar prej ali slej in brez dvoma čaka večino pomembnih podjetij.

Kadar v svojih podatkovnih bazah vodimo družine artiklov, se moramo zavedati, da potrebujemo nekakšen vpis primarnega – generičnega – artikla, na katerega so podrejeno vezani vsi drugi vpisi. Takšen princip omogoča tvorjenje družin artiklov, preprosto prenašanje že vpisanih podatkov, predvsem pa zagotovljeno unikatnost pri različnih kombinacijah identifikacijskih oznak posameznih artiklov. Med generičnim artiklom in ostalimi nastopnji tega artikla moramo vzpostaviti identifikacijsko povezavo, za katero je najbolje izbrati neko interno šifro. Ta interna šifra ne nastopa nikjer drugje, nikoli se ne prenaša v druge podatkovne baze. Seveda lahko namesto te interne šifre izberemo tudi neki prosti GTIN, a pri tem se moramo zavedati, da takega GTIN-a ne moremo uporabiti nikjer drugje (ne na nalepkah, še manj pa pri naročanju, elektronskem poslovanju ali logistiki).

Viri

Standard GS1 EANCOM, www.gs1.si.org.

Globalni uporabniški priročnik, <http://www.gs1.si.org/doc/GUM/start.html>.

Standard GS1 ADC (Automatic Data Capture), www.gs1.si.org.

Standard GS1 GDSN,

<http://www.gs1.org/productssolutions/gdsn/>.

Branko Šafarič, svetovalec za informacijsko tehnologijo, ima petindvajset let delovnih izkušenj na področju elektronskega poslovanja in razvoja informacijskih sistemov. Profesionalno delo je začel v podjetju IBT, kjer je sodeloval pri različnih projektih od mrežnega načrtovanja do uvajanja pisarniške avtomatizacije. Vodil je več projektov, med drugim mednarodno borzo ponudb in povpraševanj na GZS, Poslovni register na GZS, avtomatsko generiranje in nadzor dokumentov pri sodniku za prekrške (v okviru projekta PHARE), sistem analize, meritev in optimizacije avtobusnega prometa, razvil je sisteme za video nadzor, avtomatsko odčitavanje registrskih tablic itn. Napisal več knjig s področja računalništva. Je tudi soavtor knjige Skrivnosti elektronskega poslovanja: priročnik za mala in srednja podjetja (marec 2002). Vodi področje za računalniško izmenjavo podatkov po standardih EANCOM in BMS na GS1 Slovenija. Aktivno je sodeloval pri razvoju projekta eSLOG; vodil je delovno skupino za vsebinske standarde. Trenutno dela na pripravi slovenskih priporočil za uporabo elektronskega standarda GS1 XML – BMS.