

Podatkovne semantične smernice kot temelj upravljanja podatkov in semantične interoperabilnosti v javni upravi

Martina Furlan, Miha Jesenko, Karmen Kern Pipan

Ministrstvo za notranje zadeve in javno upravo Republike Slovenije, Direktorat za informatiko, Ljubljana, Slovenija
martina.furlan@gov.si, miha.jesenko@gov.si, karmen.kern-pipan@gov.si

Prispevek prikazuje temeljne poudarke dokumenta »Podatkovne semantične smernice«, ki jih je marca 2026 objavilo Ministrstvo za digitalno preobrazbo (zdaj Ministrstvo za notranje zadeve in javno upravo). Smernice so nastale ob podpori strokovnjakov s Fakultete za računalništvo in informatiko Univerze v Ljubljani ter v sodelovanju z različnimi organi javne uprave. Smernice na enem mestu združujejo vsebine upravljanja podatkov in semantične interoperabilnosti. Za kakovostno odločanje in oblikovanje politik na podlagi podatkov potrebujemo kakovostne in interoperabilne podatke, dobro upravljanje in s tem možnost izrabe novih tehnologij. Smernice predstavljajo temelj za nadaljnji razvoj in boljše, tudi čezmejno izmenjavo podatkov v skladu z evropskimi in mednarodnimi priporočili ter so standard za razvijalce aplikacij in druge uporabnike v organih javne uprave.

Ključne besede:

podatki

podatkovne semantične smernice

upravljanje podatkov

semantična interoperabilnost

nove tehnologije

1 Uvod

Količina podatkov v digitalnem svetu iz leta v leto strmo narašča, zlasti v zadnjem desetletju. Medtem ko je leta 2010 znašala približno 2 zetabajta¹, je do leta 2015 narasla na 15,5 zetabajta. Leta 2025 je dosegla že 181 zetabajtov, za leto 2026 pa se ocenjuje, da bo znašala kar 221 zetabajtov. Vsak dan se po svetu ustvari približno 402,74 milijona terabajtov podatkov, kar ustreza približno 0,4 zetabajta dnevno. Ocenjuje se, da je bilo kar 90 % vseh podatkov na svetu ustvarjenih v zadnjih dveh letih. Pri tem imajo pomembno vlogo video vsebine, ki predstavljajo približno 82 % celotnega globalnega podatkovnega prometa, vključno s socialnimi omrežji in drugimi spletnimi platformami [1]. Podatki torej nastajajo in nas spremljajo na vseh področjih, nove tehnologije pa odpirajo skoraj neomejene možnosti njihove uporabe ter podatke postavljajo med ključne vire sodobne družbe in gospodarstva.

Pojav novih tehnologij in njihovo vse večje uvajanje v javni upravi sta sovpadla z eksponentno rastjo količine in uporabe podatkov kot posledico digitalizacije, s čimer so se razširile možnosti analize podatkov [1, str. 3]. Okolje, v katerem se izmenjujejo podatki med organi državne uprave, je kompleksno in vključuje številne ovire in izzive za zagotavljanje storitev. Ovire lahko vključujejo različno interpretacijo podatkov, pomanjkanje skupnih in široko uporabnih podatkovnih modelov, odsotnost skupnih šifrantov, pomanjkanje interoperabilnostnih orodij in formatov itd. Semantična interoperabilnost zagotavlja, da se bosta oblika in pomen podatkov ohranila, tako da bodo podatki razumljeni na enak način pri vseh uporabnikih (npr. pri elektronski izmenjavi podatkov). V Strategiji digitalnih javnih storitev 2030 je opredeljen strateški cilj »Sodobna informacijska tehnologija ter upravljanje zaupanja vrednih podatkov«, ki podpira vpeljavo podatkovnih standardov in smernic [2]. Tudi cilj Digitalne strategije Evropske unije (EU) je zagotoviti, da bo digitalna preobrazba koristila državljanom in podjetjem ter da bo pripomogla k doseganju cilja podnebno nevtralne Evrope do leta 2050. Evropska komisija (EK) je v tem okviru opredelila obdobje 2020–2030 kot »digitalno desetletje« Evrope. Obenem želi EK okrepiti digitalno suverenost EU in prispevati k oblikovanju standardov, pri čemer je poseben poudarek namenjen podatkom, tehnologiji in infrastrukturi [3]. Podatkovne semantične smernice na krovni ravni vzpostavljajo okvir za semantično interoperabilnost in predstavljajo pomemben korak na poti do cilja boljše izmenjave podatkov. Dokument je nastal v sodelovanju s strokovnjakoma dr. Slavkom Žitnikom in dr. Dejanom Lavbičem s Fakultete za računalništvo in informatiko Univerze v Ljubljani. Med ciljnimi uporabniki teh smernic so organi javnega sektorja ter drugi deležniki, ki so kakor koli povezani z javnim sektorjem, kot denimo razvijalci, raziskovalci in drugi uporabniki podatkov oziroma informacijskih rešitev. Smernice so bile pripravljene v okviru projekta »Razvoj novih dinamičnih e-storitev (DES)«, ki sta ga sofinancirali Republika Slovenija in Evropska unija iz Načrta za okrevanje in odpornost [4]. Dokument je bil predstavljen na desetih delavnicah več kot 100 udeležencem iz več kot 30 državnih organov, vključeni so bili tudi prejeti predlogi za izboljšave [5]. Verjamemo, da se bo vsebina v sodelovanju z deležniki sčasoma ustrezno nadgrajevala in izboljševala. V nadaljevanju bomo predstavili nekaj poudarkov iz pomembnejših vsebin dokumenta.

2 Upravljanje podatkov

Upravljanje podatkov (angl. *data management*) obsega proces zbiranja, shranjevanja, zaščite in zagotavljanja kakovosti podatkov v celotnem življenjskem ciklu. MIT (Massachusetts Institute of Technology) ugotavlja, da je upravljanje podatkov ključnega pomena za ustvarjanje okolja, v katerem se podatki lahko uporabljajo v celotni organizaciji. Učinkovito upravljanje podatkov zmanjšuje težave, ki jih povzročajo napačni podatki, kot so netočne napovedi in celo preprosto dejstvo, da podatki niso dostopni, v idealnem primeru še preden se te težave pojavijo [6]. Gartner navaja, da mora upravljanje podatkov organizaciji prinašati koristi ob pravem času. Sedanje prakse upravljanja podatkov so pogosto pretoge in ne upoštevajo poslovnega okolja. Ocenjuje se, da kar 60 % organizacij ne bo

¹ Merske enote velikosti bajtov: <https://www.geeksforgeeks.org/computer-science-fundamentals/understanding-file-sizes-bytes-kb-mb-gb-tb-pb-eb-zb-yb/>, obiskano 27. 7. 2026

uspelo uresničiti pričakovane vrednosti svojih možnosti uporabe umetne inteligence do leta 2027 zaradi neskladnih sistemov upravljanja podatkov. Nadalje Gartner izpostavlja, da je treba vključiti najboljše prakse v načrt za upravljanje podatkov in vzpostaviti učinkovito strukturo upravljanja, oblikovati in uvesti učinkovite politike in standarde, oceniti in izboljšati učinkovitost upravljanja podatkov ter vzpostaviti proces ponavljanja in učenja [7]. V Sloveniji je pri upravljanju podatkov treba upoštevati tudi zakonodajne akte (npr. Zakon o varstvu osebnih podatkov (ZVOP-2) [8], Zakon o dostopu do informacij javnega značaja (ZDIJZ) [9], Zakon o izvajanju Uredbe (EU) o evropskem upravljanju podatkov (ZIUUEP)) [10], Akt o interoperabilni Evropi [11], ki določajo pravila za zbiranje, uporabo, izrabo in izmenjavo podatkov, ter strateške dokumente (npr. Strategija digitalnih javnih storitev 2030, Strategija Digitalna Slovenija 2030 in evropsko digitalno desetletje), usmerjene v krepitev interoperabilnosti ter učinkovitejšo izrabo podatkov.

2.1. Načela FAIR

Pomemben temelj upravljanja podatkov predstavljajo načela FAIR (FAIR – Findable, Accessible, Interoperable, Reusable).

F – Findable (najdljivi) – podatki in metapodatki morajo biti enostavno najdljivi tako za ljudi kot za računalniške sisteme. Vsak podatkovni zapis mora imeti trajni in enolični identifikator (npr. DOI), biti ustrezno opisan z metapodatki ter indeksiran v iskalnih sistemih.

A – Accessible (dostopni) – podatki morajo biti dostopni prek standardiziranih komunikacijskih protokolov. Tudi kadar podatki niso več na voljo, morajo metapodatki ostati dostopni.

I – Interoperable (povezljivi oziroma interoperabilni) – podatki morajo uporabljati standardizirane formate, besednjake in ontologije, da jih je mogoče povezovati ter kombinirati z drugimi podatki in sistemi.

R – Reusable (ponovno uporabni) – podatki morajo biti opisani tako dobro, da jih lahko drugi raziskovalci ali organizacije ponovno uporabijo. Jasno morajo biti navedeni pogoji uporabe, izvor podatkov in standardi kakovosti [12]. Načela FAIR:

- omogočajo dolgotrajno in učinkovito rabo podatkov v različnih sektorjih in sistemov,
- zmanjšujejo podvajanja in povečujejo zaupanje v podatke,
- omogočajo povezovanje in ponovno uporabo podatkov za raziskovalne, analitične in upravne namene,
- spodbujajo transparentnost, odgovornost in vrednotenje učinkov javnih politik [4, str. 13].

Nadalje prikazujemo nekaj smernic in praktičnih primerov iz dokumenta »Podatkovne semantične smernice«:

Smernica 1 – Sledenje zakonodaji in smernicam: organizacije morajo redno spremljati predpise ter strateške usmeritve pri upravljanju podatkov.

Primer 1 – Zakon o varstvu dokumentarnega in arhivskega gradiva ter arhivih (Uradni list RS, št. 30/06, 24/14 – odl. US in 51/14) definira podobna načela, kot jih naslavlja načela FAIR, na primer načelo trajnosti (4. člen), načelo celovitosti (5. člen) in načelo dostopnosti (6. člen). Pri tem je treba upoštevati tudi Uredbo o varstvu dokumentarnega in arhivskega gradiva (UVDAG, Uradni list RS, št. 42/2017) in Pravilnik o enotnih tehnoloških zahtevah za zajem in hrambo gradiva v digitalni obliki (Uradni list RS, št. 118/20). Upoštevanje načel FAIR ne bi pomenilo le boljše enotnosti podatkov in večjih možnosti njihove izmenjave, temveč bi tudi pri arhiviranju zmanjšalo obseg dodatnega dela, potrebnega za pretvorbo podatkov v ustrezno obliko.

Smernica 2 – Upoštevanje načel FAIR: ob ustvarjanju novih ali posodabljanju obstoječih podatkovnih virov je treba zagotoviti najdljivost, dostopnost, interoperabilnost in ponovno uporabljivost podatkov.

Primer 2 – Pozitivni učinki upoštevanja načel FAIR: pri modeliranju podatkov za pomoč uporabnikom z različnimi oblikami oviranosti lahko neprimerno definiran podatkovni standard vodi do izključenosti ali diskriminacije [4, str. 13].

2.2. Etika

Etika postavlja moralna načela in usmeritve za odgovorno ravnanje z vse bolj raznovrstnimi in obsežnimi podatki (zbiranje, shranjevanje, analiza, izmenjava in uporaba podatkov, pri čemer uporaba vključuje tudi strojno interpretacijo ali predstavitev podatkov). Etika ravnanja s podatki zajema moralne obveznosti v zvezi z zbiranjem, varovanjem in uporabo osebnih podatkov ter vpliv, ki ga to ima na posameznike [13]. Gre za ohranjanje zaupanja uporabnikov, strank, zaposlenih in deležnikov pri ravnanju s podatki, kar je ključni vidik sodobnega upravljanja, glede na vse večjo odvisnost od podatkov v procesih odločanja [14]. Pri delu s podatki je ključno upoštevati etična načela, med katerimi so zaščita zasebnosti, transparentnost in preprečevanje diskriminacije.

V nadaljevanju prikazujemo nekaj praktičnih primerov in smernic iz Podatkovnih semantičnih smernic o etičnem ravnanju pri upravljanju podatkov:

Smernica 3 – Pravičnost in preprečevanje diskriminacije: pri modeliranju podatkov je treba prepoznati in odpraviti pristranskost, zagotavljati ažurnost in celovitost podatkov.

Primer 3 - Preprečevanje nepoštenih izidov modela: slabo zasnovan model lahko pri avtomatiziranem odločanju na podlagi podatkov o družini ali zaposlenosti povzroči nepoštene izide.

Smernica 6 – Etično ravnanje pri ponovni uporabi podatkov: Pri ponovni uporabi podatkov iz različnih virov je treba spoštovati prvotni namen uporabe, predvsem pravne podlage in kontekst, v katerem so bili podatki zbrani.

Primer 6 – Uporaba podatkov za drug namen: v primeru osebnih in drugih pravno varovanih podatkov je zelo pomembno, da se podatki zbirajo skladno s pravno podlago, pri kateri je jasno opredeljen namen zbiranja. Zbranih podatkov ni dovoljeno uporabiti za druge namene brez ustrezne pravne podlage [4, str. 14–16].

2.3. Življenjski cikel podatkov

Da bi znanost o podatkih postavili v ustrezen kontekst, je treba predstaviti faze življenjskega cikla podatkov, od njihovega nastanka do njihove interpretacije. Te faze surove podatke pretvarjajo v vrednost za končnega uporabnika. Poleg tega je treba na vsaki stopnji življenjskega cikla podatkov upoštevati varstvo podatkov in etiko ravnanja s podatki [15]. Pri tem je zelo pomembno tudi upravljanje podatkov, pri katerem se uporabljajo uveljavljeni modeli za upravljanje podatkov (kot denimo DAMA-DMBOK (angl. Data Management Body of Knowledge), CMMI DMM (angl. Data Management Maturity) ter standard ISO 8000), ki omogočajo strukturirano upravljanje skozi celoten življenjski cikel podatkov – od zajema podatkov, njihove obdelave, shranjevanja, upravljanja, analiz, interpretacije do njihovega arhiviranja ali izbrisa. Tabela 1 prikazuje stopnje življenjskega cikla podatkov skupaj s kratkim opisom.

Za vpeljavo življenjskega cikla podatkov v organizacijo je lahko v pomoč tudi Smernica 22:

Smernica 22 – Določanje življenjskega cikla podatkov: opredeliti je treba faze življenjskega cikla podatkov in pravila ravnanja s podatki v vsaki fazi. Posamezne stopnje življenjskega cikla podatkov naj predstavljajo »testni okvir« s katerim lahko preverjamo ustreznost ravnanja s podatki, določamo namene zbiranja, hranjenja in arhiviranja podatkov ali izbrisa podatkov, s čimer tudi sledimo trajnostnemu upravljanju podatkov [4, str. 35–36].

Tabela 1: Stopnje življenjskega cikla podatkov

STOPNJA	KLJUČEN PROCES	NOSILEC
Generiranje	Nastanek podatkov v transakcijskih bazah, senzorjih itd.	Uporabniki naprav, skrbniki informacijskih sistemov.
Zbiranje in obdelava	Čiščenje, transformacija, zagotavljanje varnosti.	Skrbniki podatkov (angl. data stewards), IT služba, varnostni strokovnjaki.
Shranjevanje	Varno hranjenje, zagotavljanje dostopnosti do podatkov. Predvsem upoštevanje zakonodaje glede dostopnosti in varnosti podatkov.	Skrbnik podatkov, IT služba.
Upravljanje	Beleženje dostopov do podatkov, zagotavljanje revizijskih sledi in sledenje spremembam, upravljanje s podatkovnimi politikami.	Skrbniki podatkov, IT služba, imetniki podatkovnih zbirk.
Analiza in vizualizacija	Izvajanje statistik, algoritmov, gradnja modelov, priprava grafov, poročil, nadzornih plošč.	Analitiki, podatkovni znanstveniki, inženirji poslovne inteligence.
Interpretacija	Strokovna presoja, strateško odločanje.	Vsebinski strokovnjaki, odločevalci, imetniki podatkov.
Arhiviranje ali izbris	Zagotavljanje trajnosti ali sprostitev podatkovnih virov.	Imetniki podatkov, vsebinski strokovnjaki, odločevalci, IT služba

Vir: [4, str. 32].

2.4. Kakovost in varnost podatkov

Organizacije javnega sektorja potrebujejo ustrezne podatke za zagotavljanje kakovostnih storitev, sprejem pravih odločitev in oblikovanje učinkovitih politik. Podatki so lahko različnih vrst, med drugim proračunske številke, podatki o zaposlenih, odgovori na ankete in podatki, zbrani v okviru digitalnih storitev. Vsak (zaposleni) ima svojo vlogo pri zagotavljanju kakovosti podatkov [16].

Kakovost podatkov se meri z dimenzijami, kot so točnost, celovitost, doslednost, zanesljivost in ažurnost. Točnost pomeni natančnost glede na dejansko stanje, celovitost pomeni prisotnost vseh potrebnih podatkov, doslednost pomeni usklajenost med različnimi viri, zanesljivost pomeni stabilnost, zaupanje in preverljivost, ažurnost pa njihovo časovno ustreznost. Te lastnosti skupaj določajo, ali so podatki primerni za uporabo oziroma v kolikšni meri. Slaba kakovost podatkov lahko vodi do napačnih odločitev, zamujenih priložnosti in povečanih stroškov. Pomembno vlogo ima tudi izobraževanje zaposlenih in vzpostavitev organizacijske kulture, ki ceni kakovost podatkov kot skupno odgovornost. Ohranjanje in zagotavljanje kakovosti podatkov v organizaciji je stalen in neprekinjen proces.

Za izboljšanje kakovosti podatkov v organizaciji sta lahko v pomoč tudi Smernici 25 in 27:

Smernica 25 – Opredelitev in spremljanje dimenzij kakovosti podatkov: za vsak podatkovni vir je treba opredeliti ključne dimenzije kakovosti (točnost, celovitost, doslednost, zanesljivost, časovna ustreznost) in jih redno spremljati z ustreznimi metrikami.

Smernica 27 – Jasna določitev odgovornosti za kakovost podatkov: za vsak podatkovni vir je treba določiti osebe ali vloge (npr. skrbnik kakovosti podatkov), ki so zadolžene za spremljanje, poročanje in izboljševanje kakovosti.

Varnost podatkov pomeni proces ohranjanja zaupnosti, celovitosti in razpoložljivosti podatkov organizacije na način, ki je skladen s postavljenimi strategijo obvladovanja tveganj. Preden pride do incidenta, morajo organizacije imeti vzpostavljeno varnostno arhitekturo in načrt odzivanja. Ko se incident zgodi, morajo biti sposobne zaznati dogodek in se nanj ustrezno odzvati. Po incidentu mora biti organizacija sposobna učinkovito in uspešno obnoviti delovanje [17]. Varnost podatkov je ključna in nepogrešljiva sestavina učinkovitega upravljanja podatkov. Zaščita pred nepooblaščenim dostopom, razkritjem, spremembami ali uničenjem je bistvena za ohranjanje zaupanja, celovitosti in zanesljivosti podatkovnih sistemov. Za učinkovito zaščito je najprimernejši večplasten pristop. Ta vključuje stroge nadzore dostopa, omrežna in druga varnostna pravila, uporabo naprednih tehnologij, kot so požarni zidovi, protivirusna programska oprema, sistemi za zaznavanje vdorov ter sisteme za upravljanje

varnostnih informacij in dogodkov. Prav tako so pomembni redni varnostni pregledi, uporaba šifriranja in tehnik zakrivanja podatkov ter priprava načrtov za odzivanje na varnostne incidente [4, str. 37-38].

Praktični primer zagotavljanja varnosti podatkov iz Podatkovnih semantičnih smernic:

Primer 24 - Zaščita pred nepooblaščenim dostopom: javne evidence in storitve (npr. GURS ali ZZZS) morajo zagotavljati, da so osebni podatki državljanov, kot so zdravstveni podatki ali zemljiškoknjižni zapisi, zaščiteni pred nepooblaščenim dostopom. V ta namen uporabljajo večplastni sistem preverjanja pristnosti, šifriranje podatkov in redne varnostne preglede, skladno z zakonodajo, kot je npr. Zakon o varstvu osebnih podatkov (ZVOP-2).

3 Semantično opisovanje podatkov

Semantična interoperabilnost zagotavlja, da se natančna oblika in pomen izmenjanih podatkov ter informacij ohranita in razumeta med izmenjavami med stranmi, z drugimi besedami: »kar se pošlje, se tudi razume«. V okviru EIF² semantična interoperabilnost zajema tako semantične kot tudi sintaktične vidike [18]. Semantična interoperabilnost pomeni, da sistemi podatke ne le izmenjujejo, temveč jih tudi enako razumejo, kar je ključno za učinkovito delovanje javnih storitev. Semantično opisovanje podatkov temelji na semantičnih modelih, kot so ontologije, besednjaki in konceptualne sheme, ki omogočajo strukturiran opis informacij. Semantično opisani podatki omogočajo tudi vsebinsko povezovanje, ponovno uporabo in avtomatizirano obdelavo podatkov, kar zmanjšuje potrebo po ročnem usklajevanju pomenov in povečuje učinkovitost ter zanesljivost [4, str. 58-59].

3.1. Semantično modeliranje

Semantično modeliranje je zelo pomembno pri organizaciji informacij na spletu, saj pomaga pri medsebojni komunikaciji različnih akterjev in njihovih pogledov. Pri tem nam pomagajo semantični modeli (npr. konceptualni modeli, leksikalni modeli, sheme, besednjaki, ontologije), ki omogočajo strukturiran zajem kompleksnih konceptov. Semantični modeli so uporabni tako za opisovanje kot za razlago in napovedovanje. Semantično opisovanje temelji predvsem na standardu RDF, ki z uvedbo trojiškega zapisa (subjekt, predikat, objekt) predstavlja ogrodje za opis povezanih podatkov (angl. Resource Description Framework), RDFS, ki predstavlja shemo RDF (angl. Resource Description Framework Schema) ter OWL, ki znatno širi besednjak RDF in predstavlja spletni jezik za opisovanje ontologij (angl. Web Ontology Language). RDF je podatkovni standard oziroma jezik za formalno opisovanje virov in njihovih metapodatkov ter skupaj z RDFS in OWL predstavlja osnovne jezike semantičnega spleta za semantično modeliranje [4, str. 58-59].

Sposobnost zagotavljanja strukturirane semantike (strukturiran pomen) kot konteksta za različne agente in klepetalne robote je široko priznana kot ključni dejavnik v tekmi na področju generativne umetne inteligence, saj zmanjšuje »halucinacije« in izboljšuje natančnost. Konceptualno modeliranje podatkov je odlična metoda za ugotavljanje, kakšne naj bi bile entitete in odnosi.

Čeprav je mogoče z orodji umetne inteligence iz nestrukturiranih podatkov in različnih baz znanja izluščiti semantične koncepte, je veliko organizacijskega in poslovnega znanja implicitno znanje, ki ga imajo poslovni strokovnjaki v glavi. Konceptualno modeliranje je preizkušena metoda za pridobivanje tega implicitnega znanja. Tako lahko povežemo semantiko in umetno inteligenco na primer na način, da uporabimo priznane podatkovne standarde in obstoječe baze znanja na običajnih podatkovnih strukturah v neki organizaciji. Nato sledi konceptualno modeliranje, ki pokrije splošno implicitno znanje (tudi stroke) in posebno implicitno znanje organizacije. Standarde, baze znanja in konceptualne modele nato uporabimo v okviru AI agentov, ki kot semantični pomočniki obdelajo velike količine besedil in poiščejo podobnosti, ki jih dodajo konceptualni in drugi strukturirani semantiki [19].

² Evropski interoperabilnostni okvir (European Interoperability Framework)

Medtem ko standarde za podatke razvijajo v okviru konsenzualnih postopkov, obstajajo metodologije, ki pospešujejo uvajanje in usklajevanje teh uveljavljenih standardov v številnih in različnih zbirkah podatkov. Uporaba velikih jezikovnih modelov (angl. Large Language Models) bi lahko pospešila proces obdelave podatkov (za en red velikosti), saj avtomatizira naporne naloge usklajevanja, ki bi sicer zahtevale ročno delo.

Implementacija podatkovnih standardov na različne podatkovne nize se običajno opira na ročno urejanje in na tehnike leksičnega ujemanja ter na pravilih temelječega preslikavanje za standardizacijo podatkovnih elementov iz že obstoječih, strukturiranih podatkovnih naborov. Čeprav so te metode učinkovite za določena področja s standardiziranimi vhodnimi podatki, so po naravi omejene zaradi potrebe po obsežnem ročnem posredovanju in običajno pokrivajo le majhen podsklop področij. V nasprotju s tem bi pristop, podprt z umetno inteligenco, znatno zmanjšal ročno delo in se razširil tudi na nestandardizirane podatkovne nize, kar raziskovalcem in drugim omogoča, da se učinkoviteje osredotočijo na znanstveno raziskovanje namesto na obdelavo podatkov [20].

Tudi v organih javnega sektorja obstaja problem samostojnega upravljanja podatkov in podatkovnih produktov, opredelitve metrik in rešitev z umetno inteligenco. To se odraža v razpršeni pokrajini definicij, standardov in metodologij in se posledično kaže v bolj ali manj kompleksnih prilagoditvah podatkov in informacijskih sistemov za medsebojno povezovanje oziroma komuniciranje. Zaupanje v deljene podatke je nizko zaradi različnih definicij, opisov in predpostavk v podatkih. Prav zato so pomembni skupni standardi, definicije in upravljanje podatkov v širšem smislu, da se oblikuje enoten podatkovni ekosistem in se zagotovi ustrezna kakovost podatkov za odločanje, kar je s pojavom naprednih tehnologij postalo še bolj pomembno.

3.2. Metapodatki in opis podatkov

Metapodatki omogočajo razumevanje, iskanje in ponovno uporabo podatkov ter predstavljajo ključen element semantične interoperabilnosti. Spremljati morajo vsak nabor podatkov. Evropska komisija navaja, da metapodatki opredeljujejo in opisujejo značilnosti drugih podatkov ter se uporabljajo za izboljšanje poslovnega in tehničnega razumevanja podatkov ter s podatki povezanih procesov [21]. Metapodatki so dodatne informacije, ki pojasnjujejo pomen in strukturo podatkov ter podajajo podrobnosti o njihovi kakovosti, izvoru, licencah, dostopnosti in posodabljanju. Učinkovitost uporabe metapodatkov na več podatkovnih virih hkrati povečujejo metapodatki, ki niso odprtih zalog vrednosti, temveč imajo omejene zaloge vrednosti oziroma imajo omejene zaloge vrednosti nad razreda. Če so podatki urejeni hierarhično po razredih oziroma pojmi z medsebojnimi relacijami, je na primer enostavno narediti analizo ali drugo obdelavo podatkovnega vira, ki ugotavlja obstoj osebnih podatkov v različnih podatkovnih virih na podlagi podatkovnega modeliranja s temeljnim besednjakom za fizične osebe, ali npr. vseh pravnih oseb, ki imajo pooblaščen osebno za podatke (prek relacije vloge v organizaciji), na podlagi besednjaka organizacije W3C³. Če so na podlagi standardov modelirane tudi notranje strukture podatkovnih virov, lahko izvajamo še podrobnejše vsebinske analize in obdelave podatkovnih virov, na primer izbor vseh podatkov, ki obravnavajo promet ali prometne nesreče. Na podlagi pravil (npr. SHACL⁴), se v semantično obogatenih podatkovnih ekosistemih analize in obdelave izvajajo tudi samodejno, oblikovno in vsebinsko pravilno. Pravila lahko obravnavajo strukturne omejitve in zagotavljajo, da podatki ustrezajo določeni shemi ali modelu, kot so zahtevane lastnosti, dovoljene vrednosti lastnosti ali določene hierarhije razredov. Lahko obravnavajo omejitve vrednosti, ki jih lahko prevzamejo lastnosti, vključno s podatkovnimi tipi, razponi vrednosti in ujemanjem vzorcev. Nadalje gre lahko za omejitve kardinalnosti, torej opredeljevanje najmanjšega in največjega števila pojavov lastnosti, ali za logične omejitve, na primer uporaba logičnih pogojev za lastnosti in vrednosti, kot so enakost ali neenakost ter njihove kombinacije s pomočjo logičnih operatorjev [22].

Spodnji smernici opisujeta primer načina objave metapodatkov in nabora metapodatkov podatkov:

Smernica 47 – Metapodatki v strojno in človeku berljivi obliki: metapodatki naj se objavijo v dveh oblikah: kot del spletne strani ali v ločeni besedilni datoteki za potrebe človeških uporabnikov in v strojno berljivi obliki, kot sta RDF zapis

³ W3C - angl. World Wide Web Consortium, ki združuje posamezne delovne skupine, ki razvijajo standarde in smernice.

⁴ SHACL - angl. Shapes Constraint Language – jezik za preverjanje semantičnega opisovanja

ali JSON-LD, ki predstavlja zapis objektov JavaScript za povezane podatke (angl. JavaScript Object Notation for Linked Data).

Smernica 48 – Opisni metapodatki za nabor podatkov in njegove distribucije: za vsak nabor podatkov naj se zagotovijo metapodatki, kot so naslov, opis, ključne besede, datum objave, kontaktni podatki, izdajatelj, prostorska pokritost, časovno obdobje, kategorija [4, str. 63–64].

3.3. Identifikatorji

Identifikatorji so temeljni gradniki informacijskih sistemov, na spletu pa v tej vlogi prevladujejo naslovi URI, ki so globalno enolični identifikatorji (angl. Uniform Resource Identifier) in se uporabljajo za identifikacijo virov. Uporaba globalnih identifikatorjev (URI) omogoča enolično identifikacijo virov in njihovo povezovanje v širše podatkovne ekosisteme. Pomembno je, da so naslovi URI uporabniku prijazni in lahko berljivi.

Uporabo metapodatkov opisujeta smernici 50 in 51.

Smernica 50 – Nespremenljivi URI identifikatorji: pri identifikaciji nabora podatkov je treba uporabiti nespremenljive URI naslove.

Smernica 51 – Ponovna uporaba nespremenljivih URI identifikatorjev: pri objavi podatkov, kjer se sklicujemo na že opredeljene koncepte, dosledno uporabljamo obstoječe, nespremenljive URI identifikatorje. Takšna uporaba krepi učinke mrežnega povezovanja, kar je osrednja ideja povezanih podatkov [4, str. 65–66].

3.4. Podatkovni slovarji

Za semantično interoperabilnost je ključna uporaba podatkovnih slovarjev, ki zmanjšujejo nejasnosti in omogočajo povezovanje podatkov. Podatkovni slovarji opredeljujejo koncepte in povezave med njimi ter jih poznamo v različnih oblikah, kot so ontologije, nadzorovani besednjaki, slovarji sopomenk in druge

Smernica 53 priporoča uporabo konceptov iz obstoječih besedišč.

Smernica 53 – Ponovna uporaba standardiziranih slovarjev: pri objavi podatkov in metapodatkov naj se uporabljajo koncepti iz obstoječih in po možnosti standardiziranih besedišč. Priporoča se uporaba standardiziranih slovarjev, kot so denimo: besednjak podatkovnih katalogov – DCAT (angl. Data Catalog Vocabulary), Dublin Core, shema za enostaven sistem organizacije znanja – SKOS (Simple Knowledge Organization System) in CNB (Centralni nadzorovani besednjak (javne uprave)). S tem se poveča skladnost z obstoječimi praksami v skupnosti, zmanjša podvajanje in spodbuja ponovna uporaba [4, str. 77].

3.5. Licenca, izvor in kakovost podatkov

Licenca, izvor in kakovost podatkov so ključni elementi metapodatkov, saj omogočajo uporabnikom razumevanje pravic uporabe, porekla in zanesljivosti podatkov. Objavljeni podatki morajo biti razumljivi ter pravno in vsebinsko zanesljivi. To pomeni, da morajo uporabniki jasno vedeti, pod kakšnimi pogoji lahko podatke uporabljajo, od kod podatki izvirajo in kako kakovostni so.

Smernice 68, 69 in 70 opisujejo ključne vidike objave podatkov, ki vplivajo na njihovo pravno veljavnost, sledljivost in uporabnost.

Smernica 68 – Licenca podatkov: Pri objavi podatkov je treba zagotoviti jasno informacijo o licenci, ki določa pogoje za uporabo, deljenje in ponovno uporabo podatkov. Licenca naj bo jasno določena za celoten nabor podatkov ali posamezne distribucije, če se pravice razlikujejo.

Smernica 69 – Izvor podatkov: Pri objavi podatkov je treba jasno navesti izvor podatkov, vključno z vsemi spremembami in transformacijami, ki so bile izvedene.

Smernica 70 – Kakovost podatkov: Pri objavi podatkov se navedejo informacije o njihovi kakovosti, kar opišemo z različnimi dimenzijami (npr. točnost, ažurnost, dostopnost, celovitost) [4, str. 77].

4 Zaključek

Podatki nastajajo ves čas, podnevi in ponoči, tako v povezavi z ljudmi kot s stvarmi. Na drugi strani ves čas prihajajo nove napredne tehnologije, ki odpirajo nova področja in neslutene možnosti obdelave, izrabe in izkoriščanja v različne namene. Možnosti obdelave podatkov se odpirajo povsod: tako na ravni posameznika kot na ravni organizacij (zasebnega in javnega sektorja) in širše družbe. V javni upravi, kjer skrbimo za velike (kritične) informacijske sisteme ter hranimo in obdelujemo velike količine pravno varovanih podatkov, je skrb za zasebnost, varnost in kakovost posebnega pomena.

Z novimi tehnologijami so se tudi v javni upravi možnosti izrabe podatkov za odločanje in oblikovanje politik izjemno povečale. Da bi jih lahko v celoti izkoristili, je izjemnega pomena, da so ti urejeni, skladni in upravljeni. S tem namenom so nastale Podatkovne semantične smernice, ki predstavljajo celovit okvir za upravljanje in semantično opisovanje podatkov v javni upravi.

Uvajanje semantične interoperabilnosti pa ni enkratni dogodek, temveč stalen proces, ki zahteva postopno prilagajanje obstoječih podatkovnih virov in sistemov. Zato bo ključno, da se njeno uvajanje izvaja sistematično in dolgoročno, ob razvoju kompetenc ter tesnem sodelovanju z vsemi deležniki.

Literatura

- [1] <https://explodingtopics.com/blog/data-generated-per-day>, »Amount of Data Created Daily«, obiskano 21. 7. 2026.
- [2] Strategija digitalnih javnih storitev, Ministrstvo za javno upravo, Ljubljana, 2023.
[Strategija_digitalnih_javnih_storitev_2030-1.pdf](#) obiskano 21. 7. 2026.
- [3] <https://digital-strategy.ec.europa.eu/en/policies/europes-digital-decade>, »Europe's Digital Decade«, obiskano 21. 7. 2026.
- [4] »Podatkovne semantične smernice«, Ministrstvo za digitalno preobrazbo, Ljubljana, 2026,
<https://nio.gov.si/products/podatkovne%2Bsemanticne%2Bsmernice>, obiskano 20. 7. 2026.
- [5] FURLAN Martina, JESENKO Miha, KERN PIPAN Karmen, ŽITNIK Slavko, LAVBIČ Dejan »Podatkovne semantične smernice«, Dnevi slovenske informatike 2026: Zbornik 33. konference (in prezentacija), Portorož, Slovensko društvo informatika, str. 66.
- [6] DAVENPORT Thomas H., REDMAN Thomas C. »How AI Is Improving Data Management«, MIT,
<https://sloanreview.mit.edu/article/how-ai-is-improving-data-management>, obiskano 15. 7. 2026.
- [7] <https://www.gartner.com/en/data-analytics/topics/data-governance>, Gartner, obiskano 13. 7. 2026.
- [8] Zakon o varstvu osebnih podatkov (ZVOP-2), Uradni list RS, št. 163/22, 40/25 – ZInFV-1 in 10/26 – ZP-1L.
- [9] Zakon o dostopu do informacij javnega značaja (ZDIJZ), Uradni list RS, št. 51/06 – uradno prečiščeno besedilo, 117/06 – ZDavP-2, 23/14, 50/14, 19/15 – odl. US, 102/15, 7/18, 141/22 in 40/25 – ZInFV-1.
- [10] Zakon o izvajanju Uredbe (EU) o evropskem upravljanju podatkov (ZIUUEUP), Uradni list RS, št. 102/24.
- [11] Uredba (EU) 2024/903 Evropskega parlamenta in Sveta z dne 13. marca 2024 o določitvi ukrepov za visoko raven interoperabilnosti javnega sektorja v Uniji (Akt o interoperabilni Evropi) (UL L št. 2024/903 z dne 22. 3. 2024).
- [12] WILKINSON Mark D., DUMONTIER Michel, AALBERSBERG Ijsbrand Jan in drugi »The FAIR Guiding Principles for Scientific Data Management and Stewardship«, Scientific Data, letnik 3, številka 1, 2016.
<https://doi.org/10.1038/sdata.2016.1>, obiskano 21. 7. 2026
- [13] <https://online.hbs.edu/blog/post/data-ethics>, Harvard Business School, »5 Principles of Data Ethics for Business«, obiskano 15. 7. 2026
- [14] <https://standards.education.gov.uk/standard/data-ethics>, »Data Ethics«, obiskano 21. 7. 2026.
- [15] WING Jeannette M., »The Data Life Cycle«, <https://hdr.mitpress.mit.edu/pub/577rq08d/release/4>, obiskano 16. 7. 2026
- [16] <https://www.gov.uk/government/news/what-is-data-quality>, »What is Data Quality«, obiskano 16. 7. 2026.

- [17] <https://www.nccoe.nist.gov/data-security>, NIST, obiskano 20. 7. 2026
- [18] <https://interoperable-europe.ec.europa.eu/collection/iopeu-monitoring/glossary/term/semantic-interoperability>, Interoperable EuropePortal, obiskano 20. 7. 2026.
- [19] KORPELA Juha, Modeling Semantics: »How Data Models and Ontologies Connect to Build Your Semantic Foundations«, <https://moderndata101.substack.com/p/semantic-foundations-with-data-models-or-ontology>, obiskano 24. 7. 2026.
- [20] LONG Rodney Alan, BALLARD Shannon, SHAH Syed. et al. »A New AI Assisted Approach Aligns Data Standards and Accelerates Interoperability in Biomedical Research«. *npj Digital Medicine*. (2026). <https://doi.org/10.1038/s41746-026-02795-z>, obiskano 24. 7. 2026.
- [21] <https://interoperable-europe.ec.europa.eu/collection/digital-ready-policymaking/glossary/term/metadata>, Interoperable Europe Portal, obiskano 21. 7. 2026
- [22] DOUBLEDAY Kevin, Fluree, PBC. »What Is SHACL? (With Examples)« <https://medium.com/fluree/what-is-shacl-with-examples-2697f659d465>, obiskano 24. 7. 2026.