

UDK 528.2:524:004.6:004.08

Izvorni znanstveni članak / Original scientific paper

Analiza i model poboljšanja prostornih metapodataka svemirskih istraživanja

Zvonimir NEVISTIĆ, Željko BAČIĆ – Zagreb¹

SAŽETAK. Podaci prikupljeni svemirskim misijama pohranjuju se i pružaju korisnicima na korištenje kroz arhive pojedinih svemirskih agencija te specijalizirane portale svemirskih misija. Korisnici često nailaze na brojne probleme prilikom pretraživanja i preuzimanja podataka od interesa, unatoč tome što je pristup podacima neograničen za sve skupine korisnika. Trenutačni načini pohrane ovog vrijednog skupa podataka usmjereni su na njihovo dugoročno arhiviranje te su u velikoj mjeri prilagođene za užu krug znanstvenika iz područja svemirskih istraživanja s neadekvatnim funkcionalnostima pristupa i pretraživanja koje ne zadovoljavaju potrebe šireg kruga korisnika. U trenutačnim načinima arhiviranja i pristupa podacima svemirskih istraživanja korisnici često nailaze na problem pronalaska podataka od interesa, raspršenost podataka te njihovo otežano tumačenje. Archive za opisivanje i pohranu podataka koriste zastarjele standarde (PDS 4) i neadekvatne metapodatke koji čine temelj uspješne IPP inicijative. Kako bi se ti podaci učinili efikasno dostupnim potrebno je razviti standarde i mehanizme koji će to omogućiti, drugim riječima razviti infrastrukturu prostornih podataka nebeskih tijela (IPPNT). U ovom radu opisani su i analizirani metapodaci trenutačnih arhiva podataka svemirskih istraživanja koji su uspoređeni s međunarodnim standardima zemaljskih inicijativa te na temelju čega su dane smjernice poboljšanja metapodataka u svrhu uspješne provedbe IPPNT inicijative.

Ključne riječi: infrastruktura prostornih podataka nebeskih tijela, metapodaci, podaci svemirskih istraživanja, Planetary Data System standard 4.

¹ dr. sc. Zvonimir Nevistić, Geodetski fakultet Sveučilišta u Zagrebu, Kačićeva 26, HR-10000 Zagreb, Hrvatska, e-mail: zvonimir.nevistic@geof.unizg.hr
prof. dr. sc. Željko Bačić, Geodetski fakultet Sveučilišta u Zagrebu, Kačićeva 26, HR-10000 Zagreb, Hrvatska, e-mail: zeljko.bacic@geof.unizg.hr

1. Uvod

Velika količina prostornih podataka koji su se pojavili i postali javno dostupni korisnicima na korištenje kroz različite web servise unazad 15-ak godina ne olakšava njihovu primjenu i pronalazak. S jedne strane, izazovno je pronaći i pristupiti bazama prostornih podataka koje se distribuiraju u raznim vladinim agencijama i na web-portalima (Li i dr. 2016). S druge strane, postoji velika redundantnost podataka pri čemu se troši novac i ljudski resursi kako bi se prikupili i održavali duplicirani podaci (Maguire i Longley 2005). To je pokrenulo razvoj koncepta infrastrukture prostornih podataka (IPP) koji rješava problem pronalaska i smanjenja redundantnosti prostornih podataka (Hu i Li 2017) te omogućuje bolje upravljanje podacima čime se može ostvariti ekonomska i ekološka korist.

IPP koncept nastoji povezati prostorne podatke u jedinstvenu mrežu, a da bi se mogao uspješno implementirati, neophodna je harmonizacija postojećih skupova podataka te njihova standardizacija i ujednačavanje kvalitete (Ključanin i dr. 2018). IPP osigurava osnovu za pretraživanje prostornih podataka, njihovu procjenu i primjenu na svim društvenim razinama. Nadalje, olakšava se integracija s drugim skupovima podataka čime se otvaraju nove inovativne poslovne mogućnosti, značajne uštede resursa, potiče održivi razvoj, bolje upravljanje okolišem i socijalna stabilnost (Cetl 2007, Rajabifard i dr. 2000). Coleman i McLaughlin (1998) definirali su osnovne komponente IPP-a: politike, tehnologije, standarde i ljudske resurse neophodne za efektivno prikupljanje, pristup, dostavu i iskorištavanje prostornih podataka. Ovakav pristup definiranja IPP-a orijentiran je na podatke, dok se u novije vrijeme, definicija IPP-a proširuje se komponentom ljudi, odnosno korisnika, čime je IPP orijentiran upravo na njihove potrebe pružanjem usluga i servisa. Temelj svakog IPP-a čine i metapodaci čija je svrha informirati korisnika o pogodnostima prostornih podataka za njihove potrebe, a sama distribucija prostornih podataka i pripadnih metapodataka definirana je različitim međunarodno priznatim standardima čime se omogućuje interoperabilnost između različitih organizacija.

Osim na Zemlji, prostorni se podaci danas sve više prikupljaju i svemirskim opažanjima. Danas svemirske letjelice posjećuju sve više planeta, satelita, komete i asteroida. Podaci prikupljeni na njima od posebnog su interesa jer njihova interpretacija omogućuje bolje razumijevanje Zemlje i njene dinamike te daju odgovore na aktualna i važna pitanja, poput utjecaja globalnog zatopljenja (Breuer i dr. 2009), omogućuju bolje razumijevanje Sunčevog sustava te se koriste za ublažavanje opasnosti života na Zemlji i doprinose razvoju znanosti općenito u pogledu novih tehnologija, inženjerstva i matematike (URL 1). Danas svemirski podaci sve više konvergiraju i s terestričkim geo-znanstvenim vizualizacijama i analizama poput GIS-a i web karata, a kako bi se ti podaci mogli što vjernije prikazati, sve više se prikupljaju i podaci s prostornom komponentom. The National Aeronautics and Space Administration (NASA) trenutno raspolaže s 2 petabajta prostornih podataka nebeskih tijela prikupljenih različitim misijama, a velike količine novih podataka arhiviraju se svake godine novim misijama. Podaci pohranjeni u raznim arhivama nisu podesni za neposrednu uporabu stoga uz podatke se pohranjuju i odgovarajući metapo-

daci kako bi se olakšala njihova uporaba i za korisnike izvan područja svemirskih istraživanja (URL 2). Povećanjem broja multidisciplinarnih svemirskih misija povećao se broj znanstvenika iz različitih domena i šire javnosti koji žele pristupiti tim podacima. Prikupljanje i distribucija podataka svemirskih istraživanja nailaze na brojne izazove. Jedan od izazova je standardizacija metoda pohrane podataka koja je ključna za točnu i preciznu analizu i znanstvena istraživanja. Danas je taj problem od velikog interesa s obzirom da je pristup podacima jednostavan i dostupan svima te još uvijek ne postoje adekvatni načini skladištenja podataka kao i njihove distribucije. Trenutačno postoji nekoliko arhiva u kojima je moguće pronaći podatke, a većina ih je koncipirana tako da je pristup prostornim podacima moguć na razini istraživačkih misija odnosno pojedinačnih instrumenata, a pretraživanje podataka unutar takvih arhiva (ili bilo koja druga vrsta semantičke pretrage) često je ograničeno. (Borden i Bishop 2019, Hare i dr. 2018). Prema tome, glavni je cilj razvoj standarda koji bi omogućili interoperabilnost i razmjenu podataka između različitih zajednica po primjeru na Zemaljske podatke te stvoriti arhive koji će udovoljiti korisnicima te koje će biti jednostavno pronaći, dijeliti i tumačiti.

S obzirom na spomenuti porast količine i raznolikosti podataka javlja se potreba za infrastrukturom prostornih podataka nebeskih tijela (IPPNT) kao proširenog koncepta tradicionalnog IPP-a, a koji bi trebao riješiti izazove prikupljanja, upravljanja, pronalaženja i uporabe podataka svemirskih misija. Razvoj ovakvog koncepta podržao bi istraživačke misije u svemirskoj zajednici te bi se maksimizirala vrijednost podataka prikupljenih na planetima i drugim nebeskim tijelima (Laura i dr. 2018). Kao jedan od preduvjeta za razvoj ovakvog koncepta nameću se adekvatni metapodaci koji olakšavaju pretragu i korištenje distribuiranih podataka za sve skupine korisnika. Podaci svemirskih istraživanja danas u najvećoj mjeri koriste PDS (Planetary Data System) standard za pohranu podataka i metapodataka. Iako ovaj standard u određenoj mjeri omogućuje interoperabilnost unutar svemirske zajednice, njegovo korištenje ne zadovoljava osnovna načela IPP-a i široku primjenu među različitim skupinama korisnika. U ovom radu dana je detaljan pregled i analiza PDS standarda za pohranu metapodataka svemirskih istraživanja kao i njegova usporedba za standardima zemaljskih IPP inicijativa na temelju čega su doneseni prijedlozi proširenja standarda kako bi zadovoljio potrebe razvoje IPPNT koncepta.

2. Distribucija podataka svemirskih istraživanja

Planetarna znanost predstavlja znanost o proučavanju planeta, mjeseca i ostalih svemirskih tijela u Sunčevom sustavu i svemiru (URL 3). Mnoge druge znanstvene discipline mogu se primijeniti u planetarnoj znanosti poput geologije, geografije i GIS-a, geodezije i daljinskih istraživanja, geofizike, mineralogije, vulkanologije, geomorfologije i drugih. Sve te discipline se primjenjuju kako bi se odredila točna procjena onoga što se događa na tijelima. Planetarni znanstvenici rade na poboljšanju našeg razumijevanja planeta i drugih

svemirskih tijela kroz proučavanje atmosfere, površine i njihove unutrašnjosti. Pokušava se razumjeti podrijetlo planeta i njihovih fizičkih procesa i sustava, kao i karakteristike asteroida i drugih objekata u svemiru (URL 4).

Važnu ulogu u planetarnoj znanosti imaju prostorni podaci dobiveni svemirskim istraživanjima. To su svi podaci s prostornom komponentom, dobiveni metodama daljinskog istraživanja, navigacijskim metodama, georeferenciranjem slika ili podaci (uzorci) prikupljeni roverima s pripadnom prostornom komponentom odnosno dovoljno prostornih informacija kako bi se mogli projicirati na pripadajuće tijelo. Ti podaci se prikupljaju u sirovom obliku nakon čega se obrađuju i stavljaju na web poslužitelje na raspolaganje korisnicima kroz arhive svemirskih agencija i specijalizirane web portale istraživačkih misija. Sami podaci koji su pohranjeni u raznim arhivama nisu omogućeni za neposrednu uporabu od strane nestručnih znanstvenika stoga se uz njih osiguravaju i odgovarajući metapodaci, zajedno sa slikovnim podacima kako bi se olakšalo njihovo korištenje i za širu javnost (URL 2).

Prikupljanje i distribucija podataka svemirskih istraživanja nailaze na brojne izazove. Jedan od njih je standardizacija kartografskih metoda i podataka koja je kritična za točnu i preciznu analizu i znanstvena istraživanja. Danas je ovaj problem od velikog interesa s obzirom da je pristup podacima jednostavan i dostupan svima. Iako je prikupljanje podataka (posebice onih s prostornom komponentom) naglo napredovalo, još uvijek ne postoje adekvatni načini skladištenja tih podataka kao i njihove distribucije korisnicima. Kao jedan od glavnih problema nameće se činjenica da je svemirska znanstvena zajednica iznimno decentralizirana te svaka organizacija ima svoje arhive i izvore podataka, ali i standarde i formate koje koriste. Glavni cilj koji se nameće je razvoj standarda koji bi omogućili interoperabilnost tih podataka između različitih zajednica te koji bi stvorili temelje uspostave IPPNT-a. Nadalje, veliki problem predstavlja i pronalazak te tumačenje dostupnih podataka (Hare i dr. 2018). Stoga, potrebno je stvoriti prostorne podatke koji će udovoljiti korisnicima (razvojem politika, standarda i dr.) te koje će biti jednostavno pronaći, dijeliti i tumačiti (metapodaci).

Najveća arhiva podataka svemirskih istraživanja, koja obuhvaća sve podatke prikupljene misijama koje su u nadležnosti NASA-e, ali i drugih svemirskih organizacija je Planetary Data System (PDS). On se sastoji od nekoliko čvorova (pristupnih točaka) od koji svaka ima zasebno sučelje što pronalazak ovih podataka ne čini jednostavnim. Svrha ovakve arhive je da podaci budu javno dostupni svima, znanstvenicima, inženjerima i drugim korisnicima. No, problem se javlja u činjenici da je određeni podatak od interesa korisnika teško pronaći te ne postoje odgovarajući metapodaci za njegovo tumačenje. Osim PDS-a postoje i druge pristupne točke poput ESA-ine Planetary Science Archive (PSA) u kojoj su arhivirani podaci Europskih svemirskih misija, ali i drugih misija poput NASA-inih s kojima ESA ima potpisan sporazum. Veliku količinu podataka, posebice kartografskih moguće je pronaći i u arhivi Astrogeology Science Center (ASC) pod vodstvom United State Geological Survey (USGS), ali i drugih svemirskih agencija poput Njemačkog DLR-a (Deutsches Zentrum für Luft- und Raumfahrt) ili Japanskog JAXA-a (Japan Aerospace Exploration Agency), Ruskog MIIGAiK-a (Moscow State University of Geodesy

and Cartography) i drugih. Postoji i veliki broj portala zasebnih misija gdje upravitelji misija stavljaju podatke korisnicima na korištenje, ali i inicijativa, poput MAPSIT inicijative za poboljšanje postojećih arhiva (Naß i dr. 2017).

Kako bi se standardizirao način arhiviranja podataka i metapodataka, NASA je za svoje arhive razvila PDS standard kojem je zadatak osigurati dosljedan opis podataka tako da korisnici znaju što očekivati prilikom rada s preuzetim PDS datotekama. Trenutno je važeći PDS 4 standard te se od 2011. godine svi podaci unutar NASA-inih arhiva arhiviraju upravo prema njemu (Milazzo 2018). NASA je ovaj standard razvila u suradnji s International Planetary Data Alliance (IPDA) čiji su svi članovi (npr. ESA, JAXA, DLR) prihvatili navedeni standard te ga koriste za pohranu i distribuciju vlastitih podataka čime se djelomično postiže interoperabilnost unutar svemirske zajednice.

No, unatoč tome, korisnici i dalje nailaze na brojne probleme prilikom pretraživanja i tumačenja podataka, posebice korisnici van domene svemirskih istraživanja. Podaci su raspršeni, obzirom na veliki broj pristupnih točaka gdje je moguće pristupiti podacima, podaci su ne dosljedni i ne razumljivi za znanstvenike izvan područja svemirskih istraživanja, a metapodaci nisu kreirani na način da bi olakšali njihovo razumijevanje i pretraživanje, pohranjeni formati podataka ne mogu se koristiti u standardnim alatima za obradu, pretraživanje podataka se temelji na činjenicama i potrebna su prethodna znanja o podacima i misijama koje su ih prikupile kako bi se podaci mogli pronaći, ne postoje mogućnosti pretraživanja cijele arhive već samo određenih dijelova prema zadanim kriterijima, opcije filtriranja i pretraživanja daju previše istih ili sličnih rezultata te ne postoje zahtijeva da podaci budu prostorno definirani (Nevistić i Bačić 2022)

S obzirom na porast i velika ulaganja u svemirske misije i tehnologije, navedeni problemi moraju se riješiti u što kraćem roku. Jedan od glavnih problema predstavljaju neadekvatni metapodaci što otežava tumačenje, pretraživanje i distribuciju samih skupova podataka. Kao jedno od rješenja za ove probleme nameće se uspostava infrastrukture prostornih podataka nebeskih tijela po uzoru na Zemaljski IPP.

3. Infrastruktura prostornih podataka nebeskih tijela

Definicija okvira IPPNT slična je onim zemaljskih inicijativa, ali je prilagođena svemirskoj zajednici i podacima. Definicija i koncept identificiraju prostorne podatke prikupljene svemirskim misijama, njihove korisnike i interoperabilnost kao pitanje od iznimne važnosti za svemirsku zajednicu, ali i širu javnost. Pristup ovom novom konceptu je usmjeren na korisnika i usluge za korisnike, a to će omogućiti učinkovito i dugoročno planiranje, razvoj, pristup i korištenje obuhvaćenih podataka. IPPNT će omogućiti online pristup širokom rasponu korisnika koji žele doći do prostornih podataka i usluga vezanih za svemirska istraživanja, omogućit će integraciju prostornih podataka s drugim vrstama prostornih informacija, suradnju multilateralnom razmjenom informacija i usklađenost različitih skupova podataka prikupljenih na nebeskim tijelima te olakšati definiranje i dijeljenje prostornih informacija vezanih za nebeska

tijela prikupljenih svemirskim misijama. Jedan od glavnih ciljeva IPP inicijativa općenito, a tako i ove je postizanje interoperabilnosti koja bi olakšala dijeljenje informacija i omogućila korisnicima da pronađu informacije, usluge i aplikacije kada je to potrebno, neovisno o fizičkoj lokaciji i bez potrebe za ulaganjem previše vremena i znanja.

Uz postizanje interoperabilnosti, glavni cilj IPPNT-a je stvoriti okruženje koje će na jednostavan način omogućiti učinkoviti pristup postojećim podacima i zainteresiranim korisnicima. Taj cilj ostvaruje se kroz niz aktivnosti kojima će se poboljšati trenutno stanje, a te aktivnosti obuhvaćaju prilagodbu postojećih skupova podataka i izradu kataloga metapodataka. Obzirom da trenutni standard metapodataka ne zadovoljava glavna načela i ciljeve IPPNT-a, isti je potrebno promijeniti ili proširiti. Implementacijom IPPNT očekuju se mnogo benefita poput jednostavnijeg pristupa podacima za sve grupe korisnika, agencija i vladinih organizacija, otvorit će se mogućnosti za mnoga nova istraživanja i projekte, razvit će se nove poslovne prilike, a korištenje podataka omogućit će bolje razumijevanje Sunčevog sustava i njegovih procesa.

4. Metapodaci svemirskih istraživanja

Prema mnogim literaturama, metapodaci se smatraju ključnim elementom za učinkovitu uporabu prostornih podataka u IPP-u. Nijedan skup prostornih podataka nije „savršen“, pa čak i ako postoji takav skup podataka, i dalje bi mogao biti beskoristan ili čak opasan za korištenje ako nije usklađen s potrebama korisnika. Sukladno tome, metapodaci omogućuju korisnicima ispravnu uporabu podataka i sprječavanje metodoloških i interpretativnih pogrešaka (Litwin i dr. 2011, Craglia i Campagna 2010, Philips i dr. 1998). Pristup metapodacima za korisnike potrebno je omogućiti kroz Web portal (katalog) i to potpuno besplatno. Zbog toga, metapodatke je potrebno pravilno strukturirati prema definiranom profilu kako bi se omogućilo njihovo indeksiranje, odnosno pretraživanje i preuzimanje kroz kataloge koji se distribuiraju putem interneta. Metapodaci kroz kataloge moraju biti dostupni korisnicima zajedno s jednostavnim alatima za pretraživanje koji im omogućuju da identificiraju podatke od interesa. Na taj način osigurava se široka primjena skupova podataka.

Prilikom pohrane podataka unutar arhiva svemirskih podataka, uz podatke, pohranjuju se i pripadni metapodaci koji korisnicima olakšavaju interpretaciju i daljnje iskorištavanje primarnih skupova podataka. NASA i PDS, kao vodeća i najveća arhiva, razvili su svoj vlastiti standard za pohranu podataka. Trenutačno aktivni standard naziva se PDS 4 te ga IPDA preporučuje svim svojim članicama prilikom pohrane podataka u arhivama. Iako neke agencije, poput USGS-a, razvijaju svoje vlastite profile metapodataka, većina misija i arhiva pohranjuje podatke prema PDS 4 preporukama i pravilima. NASA je razvojem PDS 4 standarda imala cilj poboljšati načine iskorištavanja, pretrage, pristupa i distribucije arhiviranih podataka.

PDS 4 za opisivanje pripadnih podataka koristi oznake (engl. Labels), a un-

utar tih oznaka nalaze se podaci o podacima koje oznaka opisuje. Metapodaci unutar oznaka mogu uključivati vrijeme i mjesto promatranja, instrument koji je obavio opažanje i uvjete opažanja, zatim informacije potrebne za čitanje digitalnog objekta, kao što su dimenzije i tip podataka, formati, sadržaj, opis, veličina i vrste podataka i dr. PDS4 oznake su napisane u XML-u koji je odabran jer je to jezik koji mogu čitati i ljudi i softver te zato što je široko korišteni međunarodni standard. Upravo uvođenje XML-a najveći je napredak PDS 4 standardna u odnosu na svoje prethodnike. PDS zahtijeva posebnu oznaku podatkovnog proizvoda za svaku pojedinačnu datoteku proizvoda. Ove različite oznake proizvoda mogu se izraditi na jedan od tri načina (slika 1): Priložene oznake (engl. Attached) gdje je oznaka priložena uz svaku datoteku podatkovnog proizvoda. Odvojene oznake (engl. Detached) gdje je oznaka odvojena od podataka i nalazi se u zasebnoj datoteci koja sadrži poveznicu na datoteku s podacima. Za svaku datoteku podataka postoji jedna odvojena datoteka oznake. Kombinirana odvojena (engl. Combined Detached) koja predstavlja jednu odvojenu datoteku s oznakom proizvoda, a koristi se za opisivanje sadržaja više od jedne datoteke podatkovnog proizvoda (URL 5). Za potrebe IPPNT-a odvojene i kombinirane odvojene oznake nisu preporučljive radi olakšavanja pretraživanja podataka korisnicima, stoga, potrebno je prilikom uspostave takve inicijative za svaki podatkovni proizvod pridružiti pripadajući skup oznaka (metapodataka), a jedna takva oznaka ne smije opisivati više od jednog proizvoda.



Slika 1. Vrste oznaka u PDS 4 standardu (URL 5).

PDS 4 oznaka sastoji se od pet osnovnih segmenata, a oznaka svakog pohranjenog proizvoda mora sadržavati sve elemente. Elementi pojedinog segmenta nisu jednaki kod svih proizvoda te se razlikuju ovisno o tome što prikazuju odnosno opisuju. Pojedini elementi obavezni su za sve podatke, dok je većina elemenata opcionalna i upisuje se ovisno o dostupnosti informacija. Obavezno je da oznake budu napisane u XML formatu, a slika 2. prikazuje opću strukturu s obaveznim segmentima svake oznake.

Svaka oznaka započinje deklaracijom XML datoteke, nakon čega slijedi područje identifikacije (engl. Identification area) koja se sastoji od jednog obaveznog elemenata dok su ostali elementi prema PDS 4 standardu opcionalni: LIVID i VID definicija, Informacije o autorima i citiranju (opcionalno), Povijest promjena proizvoda (opcionalno). LIVID služi za jedinstveno identificiranje proizvoda, a njegovo generiranje razvio je PDS. Pošto su metapodaci

pohranjeni upravo prema tom identifikatoru, često je u arhivama omogućeno i pretraživanje prema LIVID broju. Takav način pretraživanja nije prilagođen korisniku i zahtijeva puno uloženog vremena i truda da bi se podaci pretražili. LIVID brojevi su za korisnike koji nisu uključeni u misiju koja je prikupila i kreirala podataka često nerazumljivi stoga bi za potrebe IPPNT-a njegovo korištenje trebalo izbjegavati. Iako je prema PDS standardu informacija o povijesti promjena proizvoda opcionalna, neke arhive, poput PSA, ovaj element navode kao obavezi. Sljedeći segment oznake je područje promatranja (engl. Observation area). Unutar ovog segmenata nalaze se metapodaci koji najviše opisuju sam proizvod. Ovdje se mogu iščitati korisne informacije koje korisnicima olakšavaju korištenje podataka te se unutar ovog segmenta nalazi najviše obaveznih elemenata za pohranu, a to su: Vrijeme opažanja, Područje opažanja, Opis znanstvenog sadržaja proizvoda, Podaci o misiji i instrumentu, Specifični podaci misije, Disciplinsko područje. Vrijeme opažanja podrazumijeva početak i kraj opažanja misije, a moguće je opcionalno dodati i druge relevantne vremenske oznake u proizvodu. Područje opažanja opisuje mjesto na objektu za kojeg su pohranjeni podaci. Opis znanstvenog sadržaja proizvoda sastoji se od tri obavezna i jednog opcionalnog elementa. Obavezno je navesti svrhu opažanja koristeći unaprijed definirane izraze iz rječnika PDS-a, na isti način potrebno je navesti i razinu proizvoda koji se distribuira (neobrađeni podatak, djelomično obrađeni, kalibrirani, obrađeni i telemetrijski) te je potrebno unijeti tekstualni opis proizvoda. Opcionalno je moguće opisati i znanstveni aspekt samog proizvoda. Podaci o misiji i instrumentu opisuju sve relevantne podatke o svemirskoj misiji, njenim ciljevima i korištenim instrumentima koji su prikupili podatke, a specifični podaci misije najopsežniji je element koji se sastoji od niza podataka koji opisuju samu misiju i prikupljene podatke. Ovdje se navode različiti podaci o imenu misije, točnom vremenu prikupljanja podataka, imenu senzora na instrumentu koji je opažao podatak, podaci o samome senzoru (spektar opažanja, valna duljina opažanja, kalibracija, tip opažanja i dr.), tehnički podaci o proizvodu (npr. rezolucija slike, postavke geometrije, koordinatni sustavi i dr.). Unutar ovog segmenta, ako se radi o geometrijskim podacima, navode se i informacije o prostornoj lokaciji. Većinu informacija potrebno je opisati koristeći unaprijed definirane izraze iz PDS rječnika. Disciplinsko područje obuhvaća informaciju o tome za koja znanstvena istraživanja se proizvod može koristiti. Ovdje se odabire jedan ili više znanstvenih područja, definiranih PDS rječnikom, za koja opisani proizvod predstavlja relevantan izvor informacija. Znanstvena područja uglavnom se podudaraju s čvorovima PDS-a, no postoje i neka dodatna koja korisnici mogu navesti (npr. površina planeta, ionosfera, unutrašnja dinamika i dr.).

Sljedeći segment je lista referenci (engl. Reference list) koja daje informacije o drugim proizvodima, člancima i časopisima koji su neophodni za razumijevanje ovog proizvoda te predstavlja jedini dio strukture koji nije obavezan za sve proizvode. Sastoji se od elemenata: Poveznice na druge PDS 4 proizvode i vanjske publikacije vezane za podatke, Element područja (engl. File area observational) datoteke sadrži podatke, Ime datoteke, Statističke informacije (opcionalno), Format. Ovdje se opcionalno mogu navesti statistički opisi datoteke poput veličine, datuma kreiranja, a obavezno je navesti ime i format pohranjene datoteke.

Posljednji segment je dodatak područja datoteke (engl. File area observation supplemental) koji sadrži informacije o samoj strukturi datoteke. Unutar ovog elementa se nalaze informacije o duljina zapisa, opisi tablica i drugi strukturalni elementi koji nisu propisani već korisnik sam odabire koje će dodati.

Unutar arhiva svemirskih podataka nalaze se brojni proizvodi od kojih nisu svi prostorno definirani stoga u oznakama proizvoda podaci o prostornoj definiciji nisu obavezni već se preporučuju za sve proizvode gdje ih je moguće navesti. Kada se unose takvi tipovi opisa, potrebno je pratiti preporuke definirane zbirkom geometrije PDS-a (engl. Geometry collection) i unijeti ih unutar segmenta područja promatranja u XML oznaci. U trenutku pisanja ovog rada trenutno se razmatra minimalni skup geometrijskih podataka koji bi trebali biti uključeni u proizvode PDS-a. Pravila o izgledu i sadržaju geometrijskih metapodataka u PDS oznakama definirana su u rječniku PDS Geometrija. Geometrijski parametri koji se prikazuju unutar oznaka odnosno se na točke na površini tijela koji se promatra. Ovisno o tipu opažanja, točke mogu biti opisane kao kontrolne točke za opažanja površine, referentne točke za opažanje atmosfere ili in-situ točke za prikupljanje podataka na samoj površini tijela. Koordinate, odnosno pozicija točaka izražava se u planetocentričnom koordinatnom sustavu pomoću duljine i širine, a svi kutovi moraju biti izraženi u stupnjevima, a udaljenosti u kilometrima. Ukoliko se geometrija navodi, postoje minimalni atributi koji se moraju unijeti u oznake a to su: SPICE oznake geometrijskih vrijednosti, Referentno vrijeme opažanja točke u UTC (Univerzalno koordinirano vrijeme) vremenskoj skali, Proširene geometrijske vrijednosti (opcionalno). SPICE oznake predstavljaju obavezni element prikaza geometrije u metapodacima. Većini korisnika koji nisu uključeni u planiranje misija i svakodnevnu obradu svemirskih podataka ove oznake bit će ne razumljive. Iako SPICE za svemirske znanstvenike dobro opisuje geometriju, za potrebe IPPNT-a mora se uvesti drugačiji model opisivanja geometrije i prostorne definicije podataka koji će biti razumljiv za široki spektar korisnika. S obzirom na NASA-ine napore za definiranjem ovog sustava oznaka, SPICE bi trebao ostati u metapodacima za svemirske znanstvenike, dok bi se kao obavezna oznaka prostorne definicije i geometrije podataka trebala uvesti definicija koordinatnog sustava po primjeru na zemaljske IPP inicijative i njihove metapodatke. Uz SPICE oznaku, obavezan element je i referentno vrijeme opažanja točke koje mora biti naznačeno u UTC vremenskoj skali. Sve ostale vrijednosti geometrije trenutno su opcionalne, a one mogu obuhvaćati: početak i kraj vremena opažanja u UTC skali, centralna širina i duljina točke opažanja, najzapadnija, maksimalna i minimalna duljina i širina objekta opažanja, solarna duljina i širina, udaljenost objekta opažanja od instrumenta (senzora), podaci o poziciji senzora opažanja, visina objekta opažanja, pozicija Sunca u trenutku opažanja, kut opažanja, minimalni i maksimalni kut objekta prema senzoru koji ga je opažao (Nevistić 2022).

Kako bi se poboljšalo pretraživanje podataka prema lokaciji, što je i jedan od glavnih ciljeva koji se želi postići uspostavom IPPNT-a, potrebno je definirati minimalni skup podataka koji će sadržavati svaka oznaka (metapodatak), a na temelju kojih će biti moguće jednostavno prostorno definirati objekt. Kao najbolje rješenja nameće se obavezno uvođenje koordinata promatranih objekata (centralna točka objekta) prema IAU (International Astronomical Union) defin-

iciji koordinatnih sustava. Kada nije moguće primijeniti IAU definiciju, koordinate bi trebalo opisati geografskom širinom i duljinom u planetocentričnom koordinatnom sustavu. Uz koordinate središnje točke opažanja, potrebno je minimalno navesti i udaljenost objekta od senzora, te opisati rasprostiranje objekta opažanja (minimalnu i maksimalnu zapadnu i istočnu duljinu te sjevernu i južnu širinu objekta). Ovako definiranim skupom minimalnih podataka korisnicima se jasno daje informacija o prostornoj definiciji objekta koji se promatra. Uz obavezni skup prostornih podataka, kreatorima podataka potrebno je omogućiti i opcionalne opise koji mogu uključivati: poziciju senzora, udaljenost od senzora, kut opažanja i druge atributne podatke koji su i trenutno dostupni za unos u sklopu proširene geometrijske vrijednosti objekta.

4.1. Procjena iskoristivosti metapodataka

U današnje vrijeme u svijetu postoje mnogobrojni primjeri uspješno provedenih IPP inicijativa u kojima korisnici na jednostavan način pretražuju podatke od interesa. Kako bi podatke bilo jednostavno pronaći metapodaci moraju sadržavati određene elemente koji olakšavaju njihovu pretragu i na temelju kojih se alatima za pretraživanje i podaci od interesa za korisnika filtriraju iz mnoštva pohranjenih skupova podataka. U zemaljskim IPP inicijativama najčešće korišteni standardi za pohranu metapodataka su ISO 19115 i FGDC standard. U kontekstu IPPNT-a ovi standardi nisu direktno primjenjivi je se odnose na zemaljske podatke i ne zadovoljavaju potrebe svemirskih podataka. Primjerice, kod podataka svemirskih istraživanja korišteni su drugačiji koordinatni sustavi, neophodni su podaci o misiji i instrumentu ili na primjer informacije o disciplinskom području na koje se podaci odnose. No, slijedeći primjer uspješnih zemaljskih inicijativa, trenutno aktivni standardi za pohranu svemirskih podataka mogu se uspješno nadograditi kako bi se olakšalo njihovo pretraživanje, a time i uporaba kod širokog broja korisnika. Nadogradnja postojećeg PDS 4 standarda po primjeru na zemaljske standarde jednostavniji je proces od proširenja postojećih zemaljskih standarda na podatke svemirskih istraživanja. U tablici 1. nalazi se usporedba obaveznih i opcionalnih elemenata metapodataka (oznaka) u PDS 4 standardu s FDGC, INSPIRE i ISO standardom pohrane metapodataka. Prikazani podaci u tablici ukazat će koje elemente je potrebno nadograditi u PDS 4 standard za uspješnu provedbu IPPNT-a. Segmenti prikazani u tablici preuzeti su iz INSPIRE direktive, a pojedini elementi metapodataka po segmentu sumirani su zajedno za sva tri standarda. Primjerice, u INSPIRE direktiva kao zasebni segment koristi ključne riječi, dok je u FGDC standardu to element identifikacije. Svi segmenti i elementi iz FGDC, ISO i INSPIRE-a su obuhvaćeni tablicom, no u ISO i FGDC standardu pojedini element moguće pripada nekom drugom segmentu nego kao što je to slučaj kod INSPIRE-a koji koristi prilagođeni ISO standard za potrebe provedbe IPP inicijative u EU.

Tablica 1. *Usporedba elemenata metapodataka FGDC, INSPIRE, ISO i PDS 4 standarda (Nevistić 2022).*

Segmenti	Elementi metapodataka po pojedinom segmentu	INSPIRE	FGDC	ISO	PDS 4	
Identifikacija	Naslov	✓	✓	✓	✓	Korištenje stručnih termina (često naziv instrumenta koji je opažao)
	Sažetak	✓	✓	✓	✓	Opis sadržaja podataka
	Tip izvora	✓	∞	○	✓	
	Poveznica na izvor	∞	✓	○	∞	Dostupna ukoliko postoji
	Identifikator	✓	X	○	✓	LIVID oznaka
	Povezani izvor	○	○	○	∞	Izvor na stručne i znanstvene članke
	Jezik	∞	X	✓	X	
	Status	X	✓	X	X	
	Grafika/prikaz podataka	X	○	X	X	
	Osobe i institucije koje su doprinijele izradi podataka	X	○	X	○	
	Opis okoline u kojoj je podataka nastao (npr. Softver, sistem i dr.)	X	○	X	X	
Klasifikacija prostornih podataka i usluga	Kategorija teme	✓	✓	✓	X	Primjenjuje se znanstveno područje
	Vrsta usluge	✓	X	X	✓	Polje/grana za koje podatak primjenjiv
Ključne riječi	Vrijednost ključne riječi	✓	✓	○	X	
	Rječnik iz kojeg dolazi	○	X	○	X	
Geografski položaj	Geografski granični okvir (zemljopisna duljina i širina)	✓	X	∞	∞	SPICE oznaka
	Definicija referentnog koordinatnog sustava	X	∞	X	∞	Planetocentrični koordinatni sustav
	Definicija visinskog sustava	X	∞	X	X	
Vremenska referenca	Vremenski obuhvat	Δ	✓	Δ	✓	
	Datum objave	Δ	✓	Δ	✓	
	Datum promjena	Δ	✓	Δ	✓	
	Datum nastanka	Δ	✓	Δ	✓	

Kvaliteta i valjanost	Podrijetlo	✓	∞	○	✓	
	Prostorna rezolucija	∞	○	○	∞	Primjenjuje se za rasterske podatke
	Točnost atributa	X	∞	X	X	
	Logička konzistentnost	X	∞	X	X	
	Potpunost	X	∞	X	X	
	Položajna točnost	X	∞	X	X	
	Informacije o rasterskom/vektorskom skupu podataka	X	∞	X	✓	Formati podataka
Sukladnost	Specifikacija	✓	∞	○	✓	
	Stupanj sukladnosti	✓	X	○	X	
	Opis atributa	X	∞	X	X	
Ograničenja	Uvjeti o pristupu i uporabi	✓	✓	○	X	
	Ograničenja javnog pristupa	✓	✓	○	X	
Odgovorne organizacije	Odgovorna strana	✓	∞	Δ	∞	Moguće navesti u sklopu dodatnih informacija o datoteci
	Uloga odgovorne strane	✓	∞	Δ	∞	
Metapodaci o metapodacima	Kontakt	✓	○	○	∞	
	Datum nastanka	✓	✓	✓	∞	
	Jezik	✓	X	X	X	
	Vremenski period dostupnosti podataka	X	∞	X	X	
	Tehnički preduvjeti	X	∞	X	X	
	Proces preuzimanja	X	∞	X	X	
	Opis izvora	X	∞	X	✓	Opis instrumenta i misije
	Ime standarda i verzija	X	✓	X	✓	Verzija PDS standarda prema kojoj je podatak pohranjen
	Vremenska skala	X	○	X	X	
	Datum recenzije i datum buduće recenzije	X	○	X	X	

* Objašnjenje tablice: U tablici su oznakom „✓“ označeni elementi koji u pripadnom standardu spadaju pod obavezne, oznaka „O“ prikazuje opcionalne elemente, a oznakom „∞“ su prikazani obvezni/opcionalni elementi ovisno o nekom uvjetu (npr. navođenje jezika u INSPIRE direktivi obavezno je samo ukoliko postoje tekstualne informacije u sklopu metapodataka ili npr. poveznica na izvor je obavezna ukoliko je dostupan URL). Oznaka „X“ u tablici prikazuje da element metapodatka u pripadnom standardu ne postoji.

✓	Obavezan
O	Opcionalan
X	Ne postoji
∞	Obavezan ako je primjenjivo
Δ	Obavezan min. jedan od elementa u segmentu

Iz tablice je vidljivo da je INSPIRE direktiva prilikom definiranja elemenata metapodataka pratila ISO 19115 standard za geografske metapodatke. Gotovo svi elementi u ta dva standarda su jednaki, a razlika je u tome što INSPIRE zahtijeva više elemenata kao obaveznih, dok je u ISO standardu većina opcionalna ovisno o dostupnosti informacije. FGDC s druge strane je opširniji i zahtijeva više informacija (obaveznih i opcionalnih) prilikom pohrane. Obzirom da je PDS 4 standard razvijen od strane NASA-e, očekivano je da je prilikom razvoja praćen model FGDC-a koji je razvijen od strane Američkog nacionalnog geografskog odbora koji je odgovoran za distribuciju i promociju nacionalnih prostornih podataka.

PDS 4 u elementima metapodataka najviše se podudara s FGDC modelom, a uzimajući u obzir činjenicu da sva tri standarda prikazana u tablici 1. se uspješno koriste u IPP inicijativama i omogućuju korisnicima jednostavnu pretragu, pronalazak i korištenje metapodataka, jasno se može ukazati na nedostatak pojedinih elemenata unutar metapodataka PDS 4 standarda, a koje je potrebno uključiti za uspješnu provedbu IPPNT inicijative. Pojedine elemente koji postoje unutar PDS 4 standarda, potrebno je nadograditi, odnosno primijeniti pristup njihovog iskazivanja. Primjerice, u naslovu metapodataka PDS 4 standarda često su korišteni stručni izrazi i imena misija i instrumenta, a naslov ne opisuje dobro same podatke i ne ukazuje na to što korisnici mogu očekivati. Poveznice na izvor dostupne su samo u pojedinim slučajevima, a s obzirom da se ovdje radi o podacima koji su već arhivirani unutar neke od svemirskih arhiva, ovaj element trebao bi biti obavezan. LIVID oznaka je već ranije objašnjena te je preporuka da se ona pojednostavi ili da se pronađe novi model identifikacije podataka. Izvor na stručne i znanstvene članke neophodan je ukoliko postoji kako bi se korisnicima detaljnije ukazalo na mogućnosti korištenja skupova podataka. Ono što PDS 4 ne koristi su jezik, status, grafika/prikaz podatka i opis okoline u kojoj je podatak nastao kao što je to slučaj kod FGDC-a. Svi ovi elementi trebaju se, kao i kod drugih standarda, uključiti u elemente metapodataka ukoliko su dostupni, s izuzetkom jezika i opisa okoline koji bi trebali postati obavezni za svemirske podatke. Opis okoline u kojoj je podatak nastao vrlo je važan za kontekst IPPNT-a kako bi se korisnicima ukazalo o kakvom se podatku radi i kako je nastao s obzirom na brojne mogućnosti prikupljanja i obrade svemirskih podataka (in-situ, daljinska istraživanja, lab-

oratorijski podaci i dr.). Definiranjem IPPN-a definirat će se i teme podataka koje trebaju biti uključene u metapodatke, a koje će korisnicima omogućiti jednostavniju pretragu, jednako kao i ključne riječi koje su neophodan element u svim standardima, a trebali bi postati i u unutar PDS 4 zajedno s rječnikom iz kojeg se te riječi preuzimanju. Stoga, potrebno je razviti i pripadni rječnik s mnoštvom sinonima koji će korisnicima izvan svemirske domene omogućiti pretraživanje podataka. Kada govorimo o geografskom položaju, SPICE oznaka bi unutar IPPNT-a trebala postati obavezan element zajedno s definicijom geografskog okvira obuhvata podatka s pripadnim duljinama i širinama izraženim u koordinatnom sustavu prema preporukama IAU. Plentocentrične koordinate mogu se primjenjivati kada IAU definicija nije primjenjiva. Kada govorimo o kvaliteti podataka, potrebno je uključiti i položajnu točnost kada je to primjenjivo odnosno kada postoji dovoljno informacija kojima bi se ta točnost mogla iskazati. Uz dosadašnje elemente kvalitete i valjanosti dodavanjem položajne točnosti korisnicima bi se povećalo pouzdanje u korištenje skupova podataka. Ono što nedostaje u PDS 4 standardu su i elementi ograničenja, odnosno uvjeti o pristupu i uporabi samih podataka koje je potrebno uključiti kao obavezne elemente. Od preostalih informacija, uz već postojeće elemente, potrebno je uvesti datum nastanka metapodataka kao obavezan element što trenutno nije slučaj.

Kvaliteta metapodataka koje NASA distribuirala bili su predmet istraživanja Quaratija i dr. (2021) zajedno s metapodacima još četiri inicijative otvorenih podataka diljem svijeta. U navedenom istraživanju ispitivala se ukupna kvaliteta metapodataka pet institucija. Ukupna kvaliteta procijenjena je s obzirom na postojanje skupova metapodataka, njihovu otvorenost i usklađenost. NASA je u navedenom istraživanju postigla najlošiji rezultat s ukupnom ocjenom kvalitete od 0,35 pri čemu je maksimalna vrijednost iznosila 1,00. Za usporedbu, najbolji rezultat u ovom istraživanju postigao je nacionalni portal otvorenih podataka Republike Irske s ukupnom ocjenom kvalitete 0,77. NASA-ini metapodaci najbolje ocjene dobili su u pogledu postojanja metapodataka, s time da su prava korisnika, očuvanje, prostorna i vremenska komponenta u ovom segmentu ocjenjene ocjenom 0. Također, elementi postojanja kontakta, formati podataka i otvorene licence također su dobili najnižu ocjenu. Potreba za uvođenjem ili promjenom ovih elementa ocijenjenih s najnižom ocjenom jasno su iskazani i u prethodnoj tablici usporedbe s drugim međunarodnim standardima na kojima se temelje uspješne provedbe inicijativa otvorenih podataka i IPP-a.

Kako bi se detaljnije ukazala ocjena kvalitete metapodataka PDS 4 standarda u svrhu donošenja preporuka za njihovo poboljšanje, u sklopu ovog istraživanja provedena je procjena elemenata kvalitete metapodataka. Za procjenu je korišten MQA (Metadata Quality Assessment) model kojeg je razvio konzorcij data.europa.eu. Svrha ovog modela je pomoći kreatorima podataka i portalima da provjere kvalitetu svojih metapodataka i primaju prijedloge za njihovo poboljšanje. Elementi koje obuhvaća ova procjena odnose se na pristupačnost podataka navedenim u metapodacima putem URL-a za pristup i preuzimanje, strojna čitljivost podataka, korištenje licenci te mogućnosti pretraživanja. Ono što ovaj model uključuje, a nije primijenjeno prilikom procjene PDS 4 skupova metapodataka je usklađenost metapodataka s DCAT-AP (Application profile for data portals in Europe) i njihovim derivatima koji se odnose na podatkovne portale u Europi i opisivanje skupova podataka javnog sektora u Europi.

Njegova namjena je omogućiti pretraživanje skupova podataka na portalu s više podataka i učiniti podatke javnog sektora bolje dostupnima preko granica i sektora što se može postići razmjennom opisa skupova podataka između podatkovnih portala unutar Europe. Obzirom da se ovaj element procjene odnosi samo na podatke u Europi, isključen je iz istraživanja. MQA mjeri kvalitetu različitih pokazatelja, a svaki pokazatelj je objašnjen u tablici 2. U procjeni se uzima u obzir koji elementi metapodataka su obavezni, a koji opcionalni. Iskaz kvalitete metapodataka dan je ovisno o postignutim vrijednostima za pojedini element procjene, a nalazi se u tablicama 3 i 4. Obzirom da izvorni MQA iskaz kvalitete uključuje i usklađenost s DCAT-AP, ukupne maksimalne vrijednosti u ovoj procjeni umanjene su za maksimalni iznos koji donosi ovaj element, a ukupni iskaz kvalitete je izračunat na osnovu početnih postotaka umanjenih za navedeni iznos. Procjena se provodi u pet segmenata koji obuhvaćaju mogućnost pretraživanja, dostupnost, interoperabilnost, ponovno korištenje i kontekst podataka. Svaki segment sastoji se od nekoliko elemenata procjene pri čemu je za svaki element definirana maksimalna vrijednost procjene. Ukupna vrijednost procjene pojedinog segmenta dobiva se zbrojem dodijeljenih vrijednosti unutar segmenta, a ukupni iskaz kvalitete zbrojem dobivenih vrijednosti svih pet segmenata (URL 6).

Tablica 2. Procjena PDS 4 metapodataka primjenom MQA modela.

Segment	Element procjene	Opis	Maksimalna vrijednost	PDS 4 vrijednost
Mogućnost pretraživanja	Ključne riječi	Ključne riječi izravno podržavaju pretraživanje i time povećavaju mogućnost pronalaženja skupa podataka.	30	0
	Kategorije	Kategorije pomažu korisnicima da tematski istražuju skupove podataka.	30	15
	Pretraživanje na karti	Korištenje prostornih informacija omogućilo bi korisnicima da pronađu skup podataka pomoću pretraživanja na karti	20	5
	Pretraživanje na osnovi vremena	Korištenje vremenskih informacija omogućilo bi korisnicima bolje pretraživanje.	20	20
Dostupnost	Dostupnost URL-ova	URL nije nužno izravna veza na podatke, također se može odnositi na URL koji daje pristup skupu podataka ili gdje je dostupno više informacija o skupu podataka.	50	40
	Preuzimanje URL-ova		20	10
	Dostupnost preuzimanja URL-ova		30	15

Interoperabilnost	Formati	Ovo polje navodi format datoteke distribucije.	20	20
	Tip medija	Ovo polje navodi vrstu medija za distribuciju.	10	10
	Formati i tip izvora prema zadanom rječniku	Provjerava pripadaju li format i vrsta izvora rječniku.	10	5
	Ne-postojanje vlasnika	Provjerava je li format distribucije u nečijem vlasništvu.	20	10
	Strojno čitljivi formati	Provjerava je li format distribucije strojno čitljiv.	20	10
Ponovno korištenje	Licenca	Licenca je vrijedna informacija za ponovnu uporabu podataka.	20	0
	Vokabular licenci	Navođenje netočnih podataka o licenci. Na primjer, postoje mnoge CC licence kojima nedostaje verzija.	10	0
	Ograničenja pristupa	Ovo polje pokazuje je li pristup podacima javan ili ograničen.	10	0
	Vokabular ograničenja pristupa	Korištenje kontroliranog rječnika povećava mogućnost ponovne uporabe.	5	0
	Kontakt	Kontaktna točka sadrži informacije kome se obratiti u slučaju pitanja u vezi s podacima.	20	15
	Vlasništvo	Provjerava se je li svojstvo postavljeno ili ne.	10	5
Kontekst podataka	Prava	U nekim slučajevima, posebna licenca ne može se primijeniti na skup podataka. Polje 'Prava' može se koristiti za određivanje reference na resurs koji će informirati korisnika o pravima koja ima prilikom korištenja skupa podataka.	5	0
	Veličina datoteke	Veličina datoteke u bajtovima.	5	5
	Datum nastanka	Datum objave skupa podataka ili distribucije.	5	3
	Datum promjena	Datum posljednje promjene skupa podataka ili distribucije.	5	3

Pojedine vrijednosti elemenata procjene dodijeljene su na osnovu pretraživanja PDS arhive i pripadnih metapodataka te na temelju tablice 1. gdje su istraženi elementi metapodataka PDS 4 standarda. Prema preporukama MQA ukoliko neki element postoji ili ne postoji dodjeljuje se maksimalna odnosno minimalna vrijednost. Ukoliko pojedini element djelomično postoji ili je djelomično primjenjiv interpolira se vrijednost ovisno o njegovoj dostupnosti. U prvom segmentu koji se odnosi na mogućnosti pretraživanja za element ključnih riječi dodijeljena je vrijednost 0 obzirom da ih PDS 4 ne koristi, maksimalna vrijednost dodijeljena je za pretraživanje na osnovi vremena koje jedan od osnovnih kriterija pretraživanja svemirskih arhiva. Za kategorije je dijeljeno 15 od maksimalno 30 bodova obzirom da tematske kategorije ne postoje, ali postoje znanstvena područja unutar elemenata metapodataka koja donekle olakšavaju pretrag. Pretraživanje na karti moguće je samo u pojedinim alatima, a tamo gdje i dostupno ne daje željene rezultate te pretraživanje nije jednostavno i intuitivno stoga je vrijednost tog elementa 5 od maksimalnih 20 bodova. U metapodacima PDS 4 arhive dostupne su URL poveznice na izvorne podatke, skupove podataka te na znanstvenu i stručnu literaturu koja opisuje podatke i olakšava korištenje. Prvom elementu dostupnosti oduzeto je 10 bodova od maksimalne vrijednosti jer ovaj element nije obavezan već opcionalan u PDS 4 standardu. Preuzimanje URL-ova je ostvarilo polovičnu vrijednost dok je dostupnost preuzimanja ostvarilo 15 od maksimalno 30 bodova jer prilikom istraživanja u preuzetim metapodacima nisu bile točno navedene sve URL adrese, odnosno neke su vodile na nepoznate i nedostupne stranice. U segmentu interoperabilnosti provjerava se da li postoje podaci o formatima podataka što je u slučaju PDS 4 standarda obavezan element i dodijeljen mu je maksimalan broj bodova. Isto tako za tip medija za distribuciju dodijeljena je maksimalna vrijednost bodova. Za korištenje rječnika dodijeljeno je 5 od maksimalnih 10 bodova jer se u rječniku često koriste stručni termini i kratice ne razumljive svima. Iako se u većini metapodataka može pronaći informacija o vlasništvu (najčešće misija koja prikupila podataka), za dio podataka ova informacija ne postoji te nije obavezan element metapodataka te mu zato nije dodijeljena maksimalna vrijednost. Iako PDS 4 format preporučuje strojno čitljive formate i metapodatke u XML shemi i dalje u arhivama pronalazimo formate koji nisu prilagođeni za strojno čitanje radi čega je dijeljena ocjena 10 od maksimalnih 20. Segment ponovnog korištenja najlošije je ocijenjen s obzirom na ne postojanje obaveza stavljanja licenci i ograničenja pristupa u PDS 4 metapodacima. Za sve te elemente u procjeni dodijeljena je ocjena 0. Iako kontakt nije obavezan element, u većini slučajeva se navodi stoga je dodijeljena ocjena 15 od maksimalnih 20. Ista situacija je i s podacima o vlasništvu kojima je dodijeljena ocjena 5 od maksimalnih 10. U posljednjem segmentu, elementu prava dodijeljena je ocjena 0 s obzirom da ti elementi u metapodacima ne postoje, veličini datoteke dana je maksimalna vrijednost jer se u svim opisima metapodataka pronalazi informacija o veličini. Datum nastanka i promjena nisu obvezni elementi PDS 4 metapodataka iako se u pravilu navode stoga je

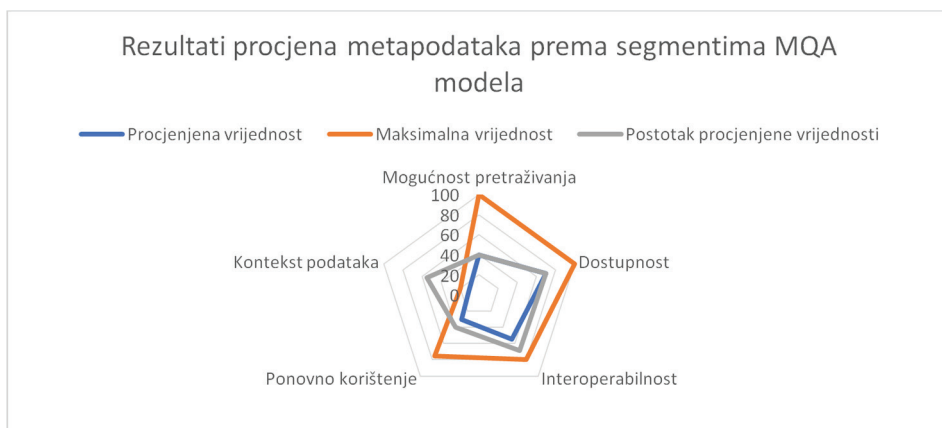
za ocjenu oba elementa dodijeljena vrijednost 3 od maksimalnih 5.

U Tablici 3. i na Grafu 1, koji prikazuju odnos vrijednosti dobivenih procjenom MQA metodom u odnosu na maksimalnu vrijednost te postotke procijenjene vrijednosti, vidljivi su rezultati procjene po pojedinim segmentima. Najbolje ocjene metapodaci postigli su u domeni dostupnosti i interoperabilnosti, slijedi kontekst podataka, a najlošija procjena je za segment ponovnog korištenja i mogućnosti pretraživanja.

Dobiveni rezultati pokazuju u kojem segmentu je potrebno nadopuniti trenutačne PDS 4 metapodatke, odnosno koje elemente dodati kao obavezne ili opcionalne, kako bi se poboljšala njihova efektivnost.

Tablica 3. Vrijednosti procjene PDS 4 metapodataka dobivenih MQA procjenom prema segmentima procjene.

	Procijenjena vrijednost	Maksimalna vrijednost	Postotak procijenjene vrijednosti (%)
Mogućnost pretraživanja	40	100	40
Dostupnost	65	100	70
Interoperabilnost	55	80	69
Ponovno korištenje	30	75	40
Kontekst podataka	11	20	55
SUMA Σ	206	375	



Graf 1. Vrijednosti procjene PDS 4 metapodataka dobivenih MQA procjenom prema segmentima procjene u odnosu na maksimalne vrijednosti i postotke procjene.

Prema MQA modelu procjene, zajednički zbroj svih pet procijenjenih segmenata daje ukupni iskaz kvalitete metapodataka koji se rangira prema ostvarenom iznosu u jedan od četiri ranga s definiranim rasponom bodova prikazanim u tablici 4. Ukupni zbroj vrijednosti za provedenu procjenu iznosi 201 što PDS 4 metapodatke stavlja na granicu ranga metapodataka koji se mogu okarakterizirati kao dobri odnosno dostatni. Za provedbu uspješne IPPNT inicijative metapodaci bi trebali ostvariti bolje rezultate procjene te biti svrstani, ako ne u najviši rang, tada barem na granici između prva dva ranga. Dobiveni rezultati procjene ukazuju na potrebu poboljšanja elemenata metapodataka za uspješnu provedbu koncepta IPPNT-a.

Tablica 4. Interpretacija ukupne kvalitete metapodataka prema MQA modelu.

Raspon bodova	Interpretacija kvalitete
0 – 112	Loši metapodaci
113 – 204	Dobri metapodaci
205 – 324	Dostatni metapodaci
325 – 375	Odlični metapodaci

4.2. Smjernice poboljšanja

Na osnovu provedenih istraživanja usporedbe PDS 4 standarda za metapodatke s učestalo korištenim standardima uspješnih IPP inicijativa te na temelju procjene metapodataka MQA modelom dobiveni su elementi metapodataka koji se moraju uključiti ili poboljšati u PDS 4 standard za uspješnu provedbu IPPNT-a, a to su:

- Naslov metapodataka mora biti jasnije opisan te se ne smiju koristiti kratice misija i instrumenata kako bi korisnici iz naslova mogli razaznati što podaci opisuju,
- Poveznice na izvor podataka moraju se uvesti kao obavezni elementi metapodataka, a dodatne poveznice na znanstvene i stručne literature moraju biti obavezne ukoliko postoje,
- Kao identifikator potrebno je uvesti jednostavniji model od LIVID broja koji se trenutačno koristi,
- Iako su gotovo svi metapodaci pisani engleskim jezikom, potrebno je za potrebe IPPNT-a uvesti element koji opisuje jezik metapodataka za tekstualne zapise,
- Po uzoru na FGDC standard, potrebno je uvesti element opisa okoline u kojem je podataka nastao kako bi se korisnicima dalo do znanja kako je podataka prikupljen i generiran što je od iznimne važnosti za svemirske podatke s obzirom na različite mogućnosti njegova prikupljanja i obrade,

- Grafički prikazi trebaju biti opcionalni, odnosno potrebno je uvesti mogućnosti grafičkog prikaza kada je to primjenjivo,
- Uspostavom IPPNT-a definirat će se teme podataka koja trebaju biti izvršene kao obavezni element u metapodatke,
- Metapodaci moraju uključivati ključne riječi te pripadni rječnik u kojem će iste biti definirane. Rječnik mora biti opširan kako bi bio razumljiv te kako bi olakšao pristup podacima velikom broju korisnika, posebice korisnicima izvan domene svemirskih istraživanja,
- Opis geografskog položaja u metapodacima mora postati obavezan element za sve podatke koji se distribuiraju u sklopu IPPNT-a. Uz SPICE oznaku koja će služiti svemirskim znanstvenicima, preporuča se uvođenje modela definiranja prostornog okvira unutar kojeg se nalaze podaci izraženi pomoću geografske širine i duljine po uzoru na INSPIRE direktivu, a preporuča se korištenje koordinatnih sustava prema smjernicama koje donosi IAU. Kada IAU smjernice nisu dostupne ili primjenjive, preporuča se korištenje planetocentričnih koordinata,
- Potrebno je uvesti obavezan element kvalitete podataka (npr. položajna točnost, rezolucija i sl.) kako bi se povećalo povjerenje korisnika u skupove podataka,
- Kao obavezan element potrebno je uvesti ograničenja i licence podataka po uzoru na druge standarde (uvjeti o pristupu i uporabi i ograničenja javnog pristupa).

Uvođenjem gore opisanih elemenata u PDS 4 metapodatke omogućilo bi se jednostavnije pretraživanje i korištenje podataka svemirskih istraživanja što trenutačno, posebice za korisnike izvan svemirske domene, predstavlja veliki izazov. Prilikom uspostave IPPNT inicijative, moraju se uključivati samo podaci koji zadovoljavaju sve dogovorene elemente, odnosno, trenutačni podaci i načini njihovog arhiviranja moraju se nadograditi i prilagoditi kako bi bili dostatni za uključivanje i korištenje u IPPNT-u. Jedino na taj način može se ostvariti puni potencijal inicijative i zadovoljiti glavni ciljevi koji se njenom implementacijom žele postići. Na slici 3 prikazana je usporedba trenutačne i nove predložene sheme XML dokumenta PDS4 oznake (metapodataka) proširena za gore navedene smjernice u svrhu uspostave IPPNT inicijative. Podcrtane vrijednosti su nove dodane vrijednosti, a oznake uz njih označavaju da li su iste obvezane (*) ili opcionalne (+). Oznake uz svaki postojeći element označavaju da je pojedini element bio do sada obavezan te takav i ostaje (*), ili je u prethodnoj verziji bio je opcionalan sada postaje obavezan (+), odnosno bio je opcionalan te takav i ostaje (#).

<Product Type>	<Geographical data> *
<Identification Area>	<IAU coordinates> +
<Alias List> <i>LIVID</i> <Alias List>*	<SPICE> *
<Citation Information> +	<Reference List> +
<Modification History> +	<External Reference> +
<Observation Area>	<Internal Reference> +
<Time Coordinates> *	<Source Product Internal> #
<Primary Result Summary> *	<Source Product External> #
<Investigation Area> *	<Data Quality> *
<Observing System> *	<File> *
<Target Identification> *	<File Area Observation>
<Mission Area> *	<File> #
<Discipline Area> +	<Data Structure> #
<Data environment> *	<File Area Observational Supplemental>
<Theme> *	<File> #
<Key Words> *	<Data Structure> #
<Language> *	<Licensing> *
<Graphical presentation> +	</Product Type>
<File> +	

Slika 2. Usporedba postojeće i nove XLM sheme metapodataka PDS-a.

5. Zaključak

Podaci svemirskih istraživanja predstavljaju vrlo vrijedan skup informacija koji se mogu primjenjivati u različite svrhe, a porastom broja svemirskih misija, raste i interes za podacima te se javlja potreba a uspostavom infrastrukture prostornih podataka nebeskih tijela (IPPNT). IPPNT se oslanja na adekvatne i dobro definirane metapodatke koji širokoj grupi korisnika omogućavaju jednostavno pretraživanje i tumačenje skupova podataka te olakšavaju daljnje analize i integraciju s drugim skupovima. Arhive danas većinom koriste specijalizirani PDS 4 standard za opis skupova podataka koji je u velikoj mjeri prilagođen znanstvenicima i domene svemirskih istraživanja. Ovim istraživanjem provedena je ocjena kvalitete metapodataka postojećih arhiva u pogledu mogućnosti njihovog pretraživanja, dostupnosti, interoperabilnosti, ponovnog korištenja i konteksta. Primjenom MQA metode, metapodaci su dobili ocjenu 206 od 375 što ih karakterizira kao dostatne s potrebnom za poboljšanjem. Usporedbom metapodataka s INSPIRE, ISO i FGDC standardima prema kojima se uspješno provode zemaljske IPP inicijative dobiven je uvid u kvalitetu metapodataka svemirskih istraživanja te su izdvojeni elementi metapodataka kojima je potrebno proširiti postojeći standard (PDS4) kako bi se povećala njihova dostupnost i omogućilo jednostavnije korištenje za potrebe IPPNT-a. Dakle, PDS3/4 standard potrebno je proširiti prema ISO i drugim kriterijima, odnosno postojeće standarde prostornih metapodataka potrebno je proširiti elementima koji će ih jednoznačno definirati u kontekstu svemirskih istraživanja (npr. načini prikupljanja, opis sadržaja, tematsko područje, IAU koordinatni sustav i dr.). Smjernice proširenja, koje su rezultat

ovog i drugih povezanih istraživanja, odnose se na poboljšanje rječnika metapodataka, korištenje intuitivnijih, jasnijih i razumljivijih opisnih riječi, dodavanje referenci na metapodatke u svrhu poboljšanja shvaćanja konteksta podataka i načina njihove uporabe te uvođenjem prostornog opisa podataka kao obaveznog elementa. Provedbom preporuka vezanih za metapodatke stvorit će se jedan od glavnih preduvjeta za implementaciju i ostvarivanje jednog od glavnih ciljeva IPPNT inicijative, a to je povećanje dostupnosti i iskoristivosti podataka svemirskih istraživanja.

Literatura

- Borden, R. M., Bishop, B. W. (2019): *Assessing Planetary Data Access and Use*, u zborniku: *4th Planetary Data Workshop*, Flagstaff, Arizona.
- Breuer, D., Grenfell, J. L., Grott, M., Hauber, E., Hussmann, H., Jänchen, J., Knapmeyer, M., Oberst, J., Sohl, F., Spohn, T., Stephan, K., Wagner, R., Wagner, F., Wählisch, M. (2009): *Astronomy and Astrophysics*, poglavlje u knjizi: *Astronomy, Astrophysics, and Cosmology Solar System*, Springer, Basel.
- Cetl, V. (2007): *Analiza poboljšanja infrastrukture prostornih podataka*, doktorska disertacija, Geodetski fakultet Sveučilišta u Zagrebu, Zagreb.
- Coleman, D. J., McLaughlin, J. (1998): *Defining global geospatial data infrastructure (GGDI): components, stakeholders and interfaces*, *Geomatica*, 52 (2), 129–144.
- Craglia, M., Campagna, M. (2010): *Advanced Regional SDIs in Europe: comparative cost-benefit evaluation and impact assessment perspectives*, *International Journal of Spatial Data Infrastructures Research*, 5, 145–167.
- Hare, T. M., Rossi, A. P., Frigeri, A., Marmo, C. (2018): *Interoperability in planetary research for geospatial data analysis*, *Planetary and Space Science*, 150, 36–42.
- Hu, Y., Li, W. (2017): *Spatial Data Infrastructures*, poglavlje u: *The Geographic Information Science & Technology Body of Knowledge*, Cornell University, Ithaca.
- Ključanin, S., Poslončec-Petrić, V., Bačić, Ž. (2018): *Osnove infrastrukture prostornih podataka*, Dobra knjiga, Sarajevo.
- Laura, J. R., Bland, M. T., Fergaso, R. L., Hare, T. M., Archinal, B. A. (2018): *Framework for the Development of Planetary Spatial Data Infrastructures: A Europa Case Study*, *Earth and Space Science*, 5, 486–502.
- Li, W., Wang, S., Bhatia, V. (2016): *PolarHub: A large-scale web crawling engine for OGC service discovery in cyberinfrastructure*, *Computers, Environment and Urban Systems*, 59, 195–207.

- Litwin, K.T., Georgiadou, Y., Bregt, A. (2011): *Geoinformation Metadata in INSPIRE and SDI: Understanding. Editing. Publishing*, Springer – Verlag, Berlin.
- Maguire, D., Longley, P. (2005): *The emergence of geoportals and their role in Spatial Data Infrastructures*, *Computers, Environment and Urban Systems*, 29, 3–14.
- Milazzo, M. (2018): *Introduction to PDS4*, Europa Clipper PSG Meeting, Pasadena.
- Naß, A., Di, K., Elgner, S., van Gasselt, S., Hare, T., Hargitai, H., Karachevtseva, I., Kersten, E., Manaud, N., Roatsch, T., Rossi, A. P., Skinner, Jr. J., Wählich, M. (2017): *Planetary Cartography And Mapping: Where We Are Today, And Where We Are Heading For?*, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-3/W1, 105–112.
- Nevistić, Z. (2022): *Poboljšanje dostupnosti i iskoristivosti prostornih podataka planetarnih istraživanja modeliranjem infrastrukture prostornih podataka nebeskih tijela*, doktorska disertacija, Geodetski fakultet Sveučilišta u Zagrebu, Zagreb.
- Nevistić, Z., Bačić, Ž. (2022): *Improving the Availability of Space Research Spatial Data*, *Interdisciplinary Description of Complex Systems*, 20 (2), 64–77.
- Philips, A., Williamson, I., Ezigbalike, C. (1998): *The Importance of Metadata Engines in Spatial Data Infrastructure*. AURISA 98 – The 26th Annual Conference of AURISA, Perth, Australia.
- Quarati, A., De Martino, M., Rosim, S. (2021): *Geospatial Open Data Usage and Metadata Quality*, *ISPRS Int. J. Geo-Inf.*, 10 (1), 30.
- Rajabifard, A., Williamson, I. P., Holland, P., Johnstone, G. (2000): *From Local to Global SDI Initiatives: a pyramid to building blocks*, u zborniku: 4th Global Spatial Data Infrastructure Conference, Cape Town, South Africa.

Mrežne adrese

- URL 1: United States (US) Congress, Commerce, Justice, Science, and Related Agencies Appropriations for 2011: Hearings before a Subcommittee of the Committee on Appropriations, House of Representatives, One Hundred Eleventh Congress, Second Session, https://www.nasa.gov/pdf/428154main_Planetary_Science.pdf, (1.9.2023.).
- URL 2: Mapping and Planetary Spatial Infrastructure (MAPSIT) Team, *Mapping and Planetary Spatial Data Infrastructure Roadmap 2019–2023*, <https://www.lpi.usra.edu/mapsit/roadmap/MAPSIT-Roadmap-2019-06-19.pdf>, (1.9.2023.).

- URL 3: Hunter, M. Supporting the Life-Cycle of planetary geospatial data, Esri User Conference 2019, <https://www.esri.com/content/dam/esrisites/en-us/about/events/media/UC-2019/user-presentations/UC18.pdf>, (28.8.2023.).
- URL 4: National Aeronautics and Space Administration (NASA), Planetary Science, Jet Propulsion Laboratory, [https://science.jpl.nasa.gov/Planetary Science/index.cfm](https://science.jpl.nasa.gov/Planetary%20Science/index.cfm), (1.9.2023.).
- URL 5: National Aeronautics and Space Administration (NASA), Planetary Data System Standard Reference, Jet Propulsion Laboratory, <https://pds.nasa.gov/datastandards/pds3/standards/>, (1.9.2023.).
- URL 6: Data.europa.eu, Metadata Quality Assurance, <https://data.europa.eu/mqa/methodology?locale=en>, (28.8.2023.).

Analysis and Improvement Metadata Model of Space Exploration Spatial Data

ABSTRACT. Data collected by space missions are stored and made available to users through archives of individual space agencies and specialized portals for space missions. Users often encounter numerous challenges when searching for and downloading data of interest, even though access to the data is unlimited for all user groups. Current methods of storing this valuable dataset are primarily focused on long-term archiving, and they are largely tailored to scientists in the field of space research, with inadequate access and search functionalities that do not meet the needs of a broader user base. To make this data efficiently accessible, it is necessary to develop a Spatial Data Infrastructure of Celestial Bodies (SDICB). In the current methods of archiving and accessing space data, users often face problems in finding data of interest, data dispersion, and the difficulty of interpreting them. The archives for describing and storing data use outdated standards (PDS 4) and inadequate metadata, which form the foundation of a successful IPP initiative. This paper describes and analyses the metadata of current space data archives, which are compared to international standards of terrestrial initiatives, and based on this, guidelines for improving metadata are provided for the successful implementation of the SDICB initiative.

Keywords: SDICB, metadata, space data, PDS 4.

Primljeno / Received: 2023-09-02

Prihvaćeno / Accepted: 2023-09-25