



Podatkovni prostori

Pogled iz Slovenije

Jure Jeraj

Big Data Specialist
vodja ekipe podatkovnih inženirjev
Result d.o.o.
IKT horizontalna mreža

Martin Puhar

mag. informacijskih ved
vodja IKT HM horizontale GIS-T
Igea d.o.o.

Flavio Fuat

samostojni svetovalec
IKT horizontalna mreža
Gospodarska zbornica Slovenije

Ljubljana, 17.3.2022

Evropska strategija za podatke

Evropsko strategijo za podatke (Evropska komisija, 2020) je Evropska komisija objavila v začetku 2020. Cilj strategije je, da EU postane vodilna družba, ki bo temeljila na podatkih. Ustvarjanje enotnega trga za podatke bo omogočilo prost pretok podatkov znotraj EU, kjer bo še posebej poudarek pretoku in uporabi podatkov med sektorji v korist podjetij, prebivalcev, raziskovalcev in javnih uprav. Implementacija strategije skozi zakonodajne in tehnične ukrepe bo opolnomočila ljudi, podjetja in organizacije za sprejemanje boljših odločitev na podlagi vpogledov iz neosebnih podatkov, ki bi morali biti na voljo vsem pod določenimi pogoji.

Ključni zakonodajni ukrepi s tega področja so Akt o upravljanju podatkov, Akt o podatkih in Akt o umetni inteligenci. Ta zakonodajni okvir bo postavil jasna pravila za upravljanje s podatki, deljenje podatkov ter reguliral uporabo umetne inteligence. Jasna pravila bodo podprla zaupanje v podatke in s tem bo podprta rast podatkovne ekonomije.

Podpora podatkovni ekonomiji je eden od ključnih ciljev Evropske strategije za podatke, kar je razvidno tudi iz sredstev, ki so temu namenjena v finančni perspektivi 2021-2027.

Podatkovni ekosistem

Na osnovi načel postavljenih v strategiji, se je že začel oblikovati Evropski podatkovni ekosistem. Krovni elementi ekosistema so podatkovne platforme (Data Platforms), podatkovni prostori (Data Spaces) in podatkovne tržnice (Data Marketplaces).

Podatkovne platforme sestavljajo osnovno infrastrukturo za deljenje in procesiranje podatkov. So okolje, ki daje možnost izrabe ekonomske vrednosti podatkov. So tehnična rešitev, temeljujoč na tehnologijah za obvladovanje masovnih podatkov (ang. Big Data Stack), večinoma pa se rešitve izvajajo z odprtokodnimi rešitvami. Kar se izvaja na nivoju Evrope, so referenčne arhitekture, ki na visokonivojskem nivoju shematsko prikazujejo sestavne elemente in procese platforme, sama izvedba pa je prepuščena naročniku/izvajalcu/ponudniku.

Podatkovne prostore si lahko predstavljamo kot dejanske hrambe podatkov, ki so strukturno in vsebinsko (semantično) vodeni in urejeni. Del podatkovnih prostorov so tudi vgrajena pravila in pogoji za deljenje in procesiranje podatkov, tako osebnih kot neosebnih. Ena od lastnosti podatkovnih prostorov je težnja, da podatki raje ostanejo v izvorni podatkovni platformi, obstajajo pa katalogi, ki platforme povezujejo (federated data spaces). Podatki se potem prenesejo oz. obdelajo na zahtevo.

Ti katalogi so osnova za delovanje **podatkovnih tržnic**, ki so okolje za semantično razumevanje podatkov. S tem vzpostavimo dejanski mehanizem za deljenje podatkov. To je ključni korak pri vzpostavitvi podatkovne ekonomije. Podatkov je vse več, so različno segmentirani in je navigiranje skozi te podatke bo vse bolj zapleteno. Podatkovna tržnica na eni strani omejuje in regulira dostope in uporabo, na drugi strani pa se nam omogoča prost, avtomatiziran, pretok podatkov upoštevajoč zakonske okvire in pogoje uporabe, ki jih je predpisal lastnik podatkov.

Priložnost podatkovnih prostorov

Pravni okviri, finančni instrumenti in predvsem prevzem teh konceptov s strani **gospodarstva, raziskovalnih organizacij, držav članic in civilne družbe** so ključni za vzpostavitev podatkovne ekonomije, gospodarsko rast in napredek civilne družbe. Zaradi tega so opredeljene priložnosti za ključne deležnike v tem procesu: gospodarstvo, prebivalstvo, znanost, vlade in javne ustanove. Namen te opredelitve je poudariti, da za vsakega deležnika deljenje in uporaba podatkov prinaša določene priložnosti oz. dodano vrednost (BDVA, 2019). Nekaj takih priložnosti:

Priložnosti za gospodarstvo:

- Podatkovne tržnice odprtih podatkov, ki bodo podjetjem bolj izenačila in poenostavila izmenjavo podatkov.
- Večja razpoložljivost obsežnih in heterogenih podatkovnih ekosistemov za potrebe umetne inteligence (umetna inteligenca namreč zahteva ogromno množico podatkov za učenje, ki jih podjetja sama mogoče celo ne zmorejo zagotoviti).
- Inovativni poslovni modeli na podlagi podatkov.
- Priložnost za dostop do varnih osebnih podatkov (z ustrezno zavarovanimi podatkovnimi prostori in jasno opredeljenimi pravili, bodo na voljo tudi osebni podatki, ki niso del GDPR-ja; ti bodo lahko postali tudi predmet vseh prejšnjih treh točk).

Priložnosti za prebivalstvo:

- Popolna kontrola na osebnimi podatki (podatkovni prostori bodo morali v celoti spoštovati GDPR, ščitil jih bo tudi akt o podatkih, akt o umetni inteligenci in podobno).
- Z ustrezno izmenjavo podatkov se bo prebivalstvu omogočila boljša kakovost življenja (ki jih bodo lahko zagotavljajo gospodarstvo ali civilna sfera na podlagi boljših možnosti inovativnih modelov).
- Dostop do prilagojenih in medsektorskih storitev B2C (s podatkovno digitalizacijo se bo lahko ponudba mnogo bolj približala potrošniku, mnogo več bo personalizacije).
- Povečanje možnosti monetizacije osebnih podatkov (spodbujalo se bo deljenje lastnih podatkov z uvedbo poslovnih modelov C2B s kombinacijo pravičnega lastnega nadzora na podatki ter prejemanjem ustreznih nadomestil za pripravljenost deljenje podatkov).
- Nove poklicne možnosti.

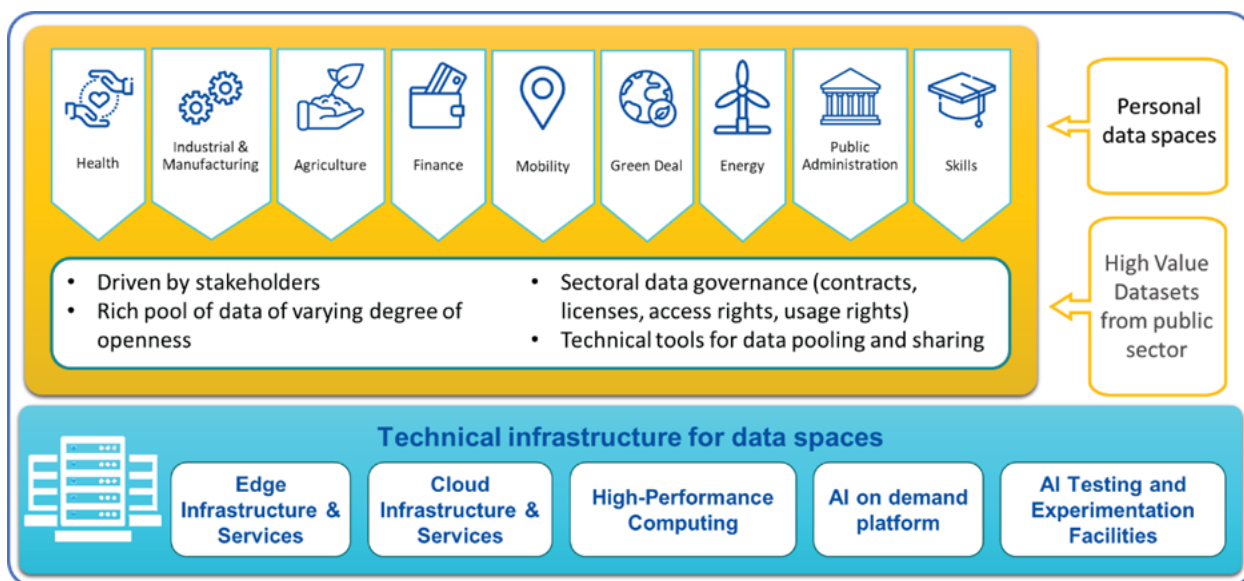
Priložnosti za znanost:

- Povečanje družbeno-ekonomskega vpliva raziskovalnih podatkov prek vertikalnih domen in geografskih meja.
- Splošni napredek in inovacij na podlagi boljše razpoložljivosti podatkov.
- Priložnosti za razvoj podlag za nove poslovne modele na podlagi boljšega izkoriščanja podatkov.

Priložnosti za javno upravo

- Izboljšan javni servis na meddržavnem nivoju (s prostim pretokom je pomembno tudi čim bolj poenostaviti javni servis na tem pretoku).
- Izboljšane javne storitve z uporabo umetne inteligence.
- Evropska statistika v realnem času.

Skupni evropski podatkovni prostori



Slika 1: Slika: Ustvarjanje podatkovne ekonomije (Evropska komisija, 2021)

Evropska komisija definira devet podatkovnih prostorov, ki so ključni za razvoj družbe in gospodarstva: zdravstvo, industrijska proizvodnja, kmetijstvo, finance, mobilnosti, zeleni dogovor, energija, javna uprava in kompetence.

Podatke je potrebno združiti oz. zagotoviti interoperabilnost in dostopnost. V podatkovne prostore se vključujejo osebni, neosebni in industrijski podatki, ustrezno zaščiteni za zagotovitev varovanja osebnih podatkov, varovanja poslovne skrivnosti in varovanja intelektualne lastnine.

Na drugi strani pa je pomembna infrastruktura, ki zagotovi zbiranje, obdelovanje in deljenje podatkov.

Finančni instrumenti programa Obzorje Evropa in Digitalna Evropa - DIGITAL usmerjajo razvoj in zagotavljajo pogoje za raziskovalno dejavnost, inovacije in implementacijo končnih rešitev na teh področjih v celotni EU.

INSPIRE - zametki podatkovnega prostora zelenega dogovora¹

INSPIRE je evropska direktiva (polni naslov: *Direktiva 2007/2/ES Evropskega parlamenta in Sveta z dne 14. marca 2007 o vzpostavitvi infrastrukture za prostorske informacije v Evropski skupnosti - INSPIRE*), ki je bila sprejeta leta 2007 in se nanaša na prostorske podatke. Namen direktive je zagotovitev prostorskih podatkov za podporo izvajanja okoljskih politik na ravni Evrope. V ta namen naj bi bili s strani držav članic EU in štirih držav EFTA (Norveška, Švica, Lihtenštajn in Islandija) harmonizirani prostorski podatki in zagotovljene standardizirane omrežne storitve za dostop do prostorskih podatkov za celotno področje Evrope. INSPIRE direktiva opredeljuje vsebine podatkov v okviru 34 podatkovnih tematik, ki so predmet aneksov št. I., II. in III. Tehnološka izhodišča se naslanjajo na OGC²/ISO standarde, ki so uveljavljeni na področju prostorskih podatkov (nastali v obdobju 2000-2006). Podatkovne specifikacije, ki so osnova za harmonizacijo podatkov in izvedbo storitev, so pripravljene v UML modelirnem jeziku, kot standardni izmenjevalni format je uporabljena XML/GML notacija. Ključne so predvsem storitve iskanja podatkov oz. metapodatki (OGC CSW standard), storitve vpogleda (OGC WMS standard) in storitve prenosa (OGC FWS standard).

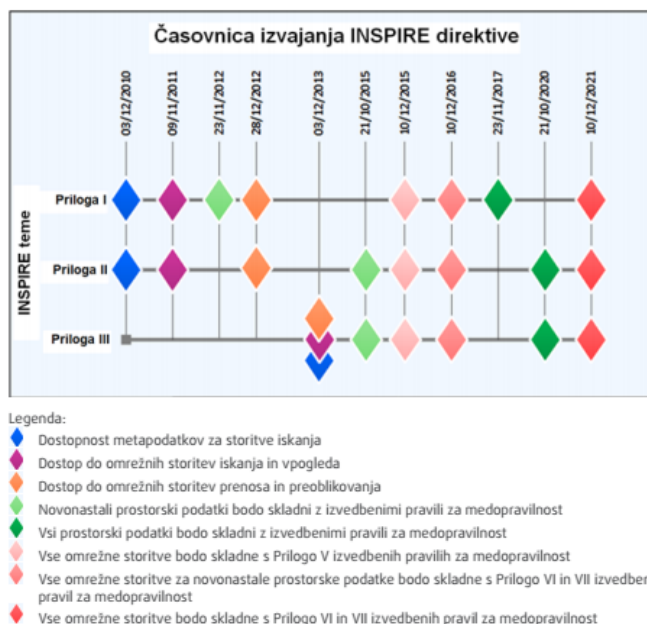
Annex I 1. Coordinate reference systems 2. Geographical grid systems 3. Geographical names 4. Administrative units 5. Addresses 6. Cadastral parcels 7. Transport networks 8. Hydrography 9. Protected sites	Annex III 1. Statistical units 2. Buildings 3. Soil 4. Land use 5. Human health and safety 6. Utility and governmental services 7. Environmental monitoring facilities 8. Production and industrial facilities 9. Agricultural and aquaculture facilities 10. Population distribution – demography	11. Area management/ restriction/regulation zones & reporting units 12. Natural risk zones 13. Atmospheric conditions 14. Meteorological geographical features 15. Oceanographic geographical features 16. Sea regions 17. Bio-geographical regions 18. Habitats and biotopes 19. Species distribution 20. Energy Resources 21. Mineral resources
Annex II 1. Elevation 2. Land cover 3. Ortho-imagery 4. Geology		

Slika 2: INSPIRE tematike po sklopih I, II, III

Časovnica uvedbe je bila v času sprejema direktive videti optimistična in je bila odmaknjena globoko v desetletje 2010-2020, zadnji rok je potekel 10.12.2021 in se nanaša na uvedbo vseh predvidenih omrežnih storitev.

¹ Green Deal Data Space

² <https://www.ogc.org/>



Slika 3: Časovnica uvedbe INSPIRE

Vse zbirke, ki sodijo v okvir 34 INSPIRE tematik in pripadajoče storitve morajo biti vključene v Evropski INSPIRE Geoportal, ki je osrednja točka za dostop do podatkov (dostopen je na spletnem naslovu: <https://inspire-geoportal.ec.europa.eu/>). Sistem temelji na posredovanju metapodatkov s strani posameznih držav preko t.i. sistema harvesting (zrcaljenje).

Izkazalo se je, da je bila časovnica zastavljena precej optimistično, saj države zaradi tehnološke in vsebinske zahtevnosti niso dosegale rokov. Večji napredek pri vzpostavljanju storitev je bil storjen šele v zadnjih nekaj letih, kar je dobro razvidno tudi iz števila zbirk in storitev s področja prioriternih zbirk podatkov, to so zbirke, ki so podpora za okoljsko poročanje (npr. natura 2000, zavarovana območja, ekološko pomembna območja, ...).

V letu 2018 je bilo na INSPIRE geoportalu registriranih 1031 prioriternih zbirk, za 453 zbirk so bile dostopne storitve za prenos podatkov. V začetku leta 2022 je registriranih že 2389 zbirk, za 1343 zbirk so dostopne storitve za prenos. Slovenija je imela leta 2018 registrirane 4 zbirke in 1 storitev prenosa, v letu 2022 pa 9 zbirk, za vseh 9 zbirk so na voljo tudi storitve prenosa podatkov.

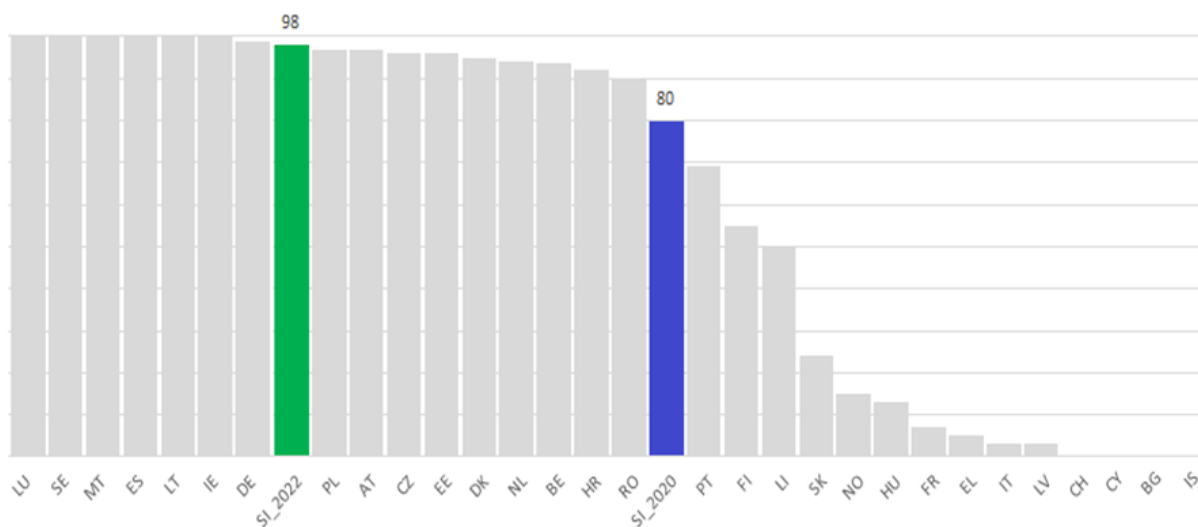
Priority Data Sets - EU & EFTA Country overview



Slika 4: Prioritetne zbirke na INSPIRE geoportalu v začetku leta 2022

Trenutno (stanje 14.03.2022) je na INSPIRE geoportalu registrianih 90912 zbirk, 9473 storitev prenosa in 55724 storitev vpogleda.

V tem obdobju je pri zagotavljanju storitev napredovala tudi Slovenija. Od 98 zbirk, ki so registrirane na INSPIRE geoportalu, jih ima 48 na voljo storitve prenosa, 59 pa storitve vpogleda, kar je nekoliko nad povprečjem EU. Povsem v vrhu pa je Slovenija pri zagotavljanju skladnosti storitev, saj zagotavlja 98% skladnih omrežnih storitev.

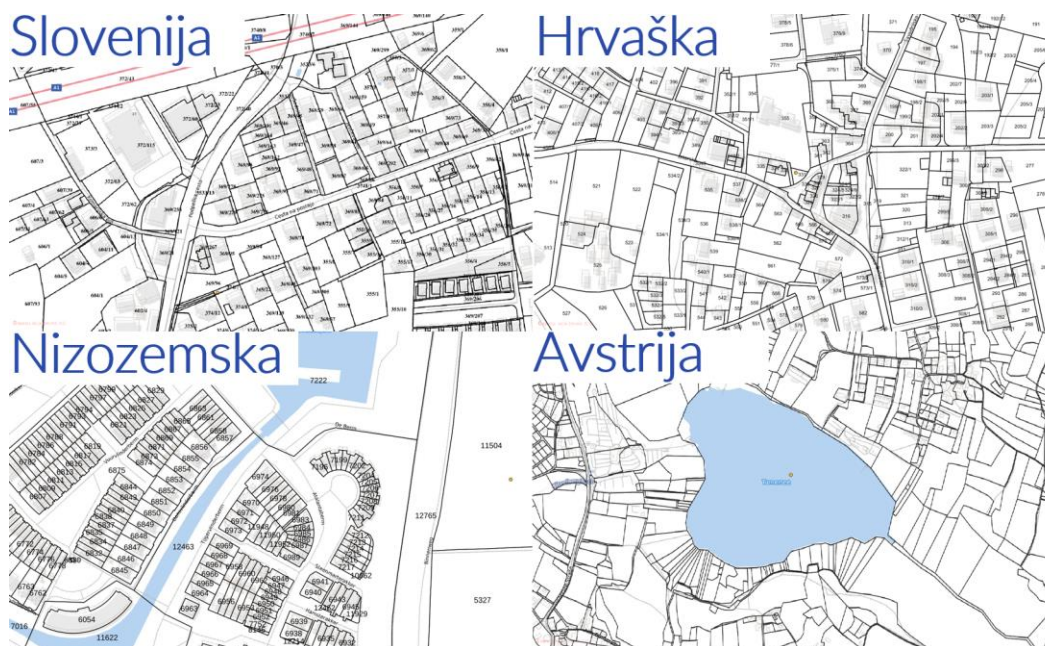


Slika 5: Stopnja skladnosti INSPIRE storitev po posameznih državah

Direktiva INSPIRE je v tem času doživela tudi nekaj kritik in vprašanj glede njenih praktičnih koristi. Večji del vprašanj je vezan na že omenjeno vsebinsko in tehnično zahtevnost pri implementaciji ter na precejšnjo okornost glede posodabljanja specifikacij, kar je posledica dokaj stroge ureditve v okviru EU direktive. Prav tako je bilo večkrat mogoče slišati predloge, da naj

INSPIRE infrastruktura zagotovi pan-evropske podatkovne produkte, npr. skupni register naslovov, karto administrativnih enot, skupni katastrski načrt in osnovno karto, pri čemer snovalci infrastrukture ves čas zagovarjajo, da je namen INSPIRE zagotoviti infrastrukturo, ki omogoča izdelavo tovrstnih produktov, a ne tudi njihovo izdelavo.

Vzpostavljane infrastrukture po posameznih državah in tudi na ravni EU je na drugi strani prinesla kar nekaj pozitivnih učinkov. Močno se je npr. dvignil nivo zagotavljanja metapodatkov o prostorskih zbirkah podatkov, uvedli so se različni skupni registri, očitna je standardizacija omrežnih storitev. Opazen je napredek glede povečanja dostopnosti prostorskih podatkov, kar je v precejšnji meri posledica dejstva, da INSPIRE močno promovira koncept odprtih javnih podatkov. INSPIRE je tudi odlična podlaga za harmonizacijo prostorskih podatkov, npr. na projektih in aplikacijah čezmejnega značaja. Primer dostopnosti podatkov različnih držav je prikazan na podatkih zemljiškega katastra, kjer večina držav že zagotavlja standardizirane storitve vpogleda, pa tudi prenosa podatkov, kar je bilo še nekaj let nazaj popolnoma nedosegljivo.

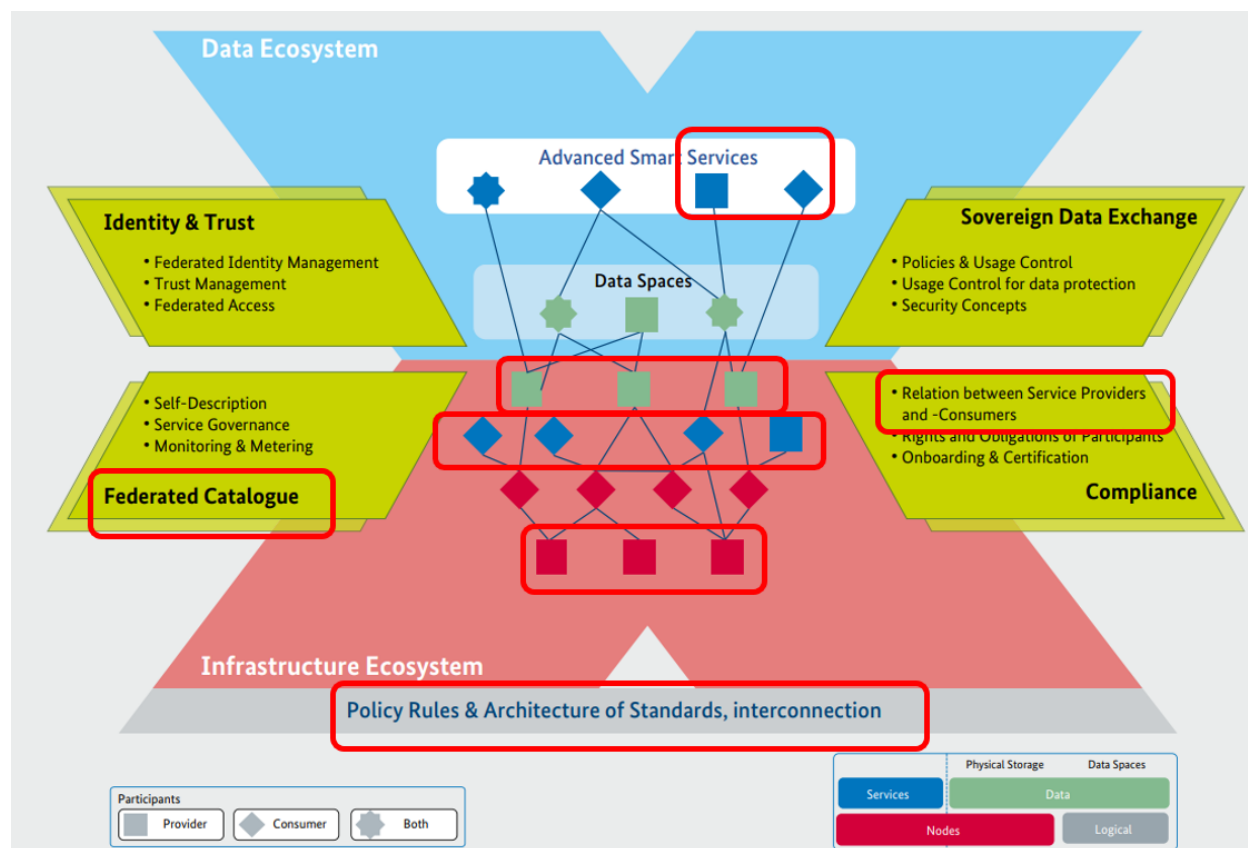


Slika 6: Prikaz storitev za vpogled v podatke zemljiškega katastra po različnih državah

Glede na to, da je v okviru INSPIRE direktive vzpostavljeno že velik del podatkovne in storitvene infrastrukture, so vprašanja glede bodočnosti INSPIRE direktive aktualna tudi v luči novo nastajajočih okvirov, kot je Evropski zeleni dogovor in koncept uvajanja skupnih podatkovnih prostorov.

Evropska komisija se je s tem vprašanjem ukvarjala že na INSPIRE konferenci leta 2020 (Evropska komisija, 2020), v letu 2021 pa je Skupno raziskovalno središče Evropske komisije objavilo poročilo na temo vključitve INSPIRE podatkov v skupni evropski prostor zelenega

dogovora (Evropska komisija, 2021). Na INSPIRE konferenci v letu 2021³ je bilo podanih kar nekaj izhodišč glede nadaljnega razvoja INSPIRE, med drugim je izhodišče, da naj se INSPIRE izogiba tehničnemu “nad-inženiringu”, opredeljeni so dodatni oz. alternativni načini kodiranja prostorskih podatkov (OGC Features API, Geodatabase). Na tej konferenci je bilo precej časa posvečeno tudi iskanju sinergije med INSPIRE in evropskim projektom GAIA-x⁴, v okviru katerega poteka praktična implementacija konceptov skupnih podatkovnih prostorov.



Slika 7: Gaia-X referenčna arhitektura in koncepti INSPIRE (rdeči okvirji)

Ob hitri primerjavi referenčne tehnične arhitekture Gaia-X in INSPIRE ugotovimo, da je precejšen del gradnikov, ki so sestavni del referenčne arhitekture za skupne podatkovne prostore, poznan tudi iz INSPIRE arhitekture: politike, pravila, standardi so ključni del INSPIRE direktive; federacija katalogov poteka v INSPIRE preko povezovanja metapodatkovnih sistemov; povezava med uporabniki in ponudniki prostorskih podatkov je osnovna ideja INSPIRE; pomembna je skladnost storitev; INSPIRE obravnava prostorske podatke, vozlišča in seveda tudi storitve. Gradniki, ki jih trenutno zagotavlja INSPIRE (še) niso na tehnološki ravni, kot so pričakovani v arhitekturi podatkovnih prostorov. Kaže, da bo tudi ta težava oziroma ovira postopoma odpravljena, saj v tej smeri že potekajo aktivnosti recimo v sodelovanju z OASC (Open & Agile Smart Cities), kjer se gradijo minimalni interoperabilni mehanizmi (MIMs⁵). MIMs so orodje za doseg interoperabilnosti in so že vključeni v iniciative povezane z graditvijo podatkovnih prostorov. V letu 2021 se je število

³ <https://inspire.ec.europa.eu/conference2021>

⁴ <https://www.gaia-x.eu/>

⁵ <https://mims.oascities.org/>

teh mehanizmov razširilo na 10, pri čemer je za prostorske podatke in storitve pomemben MIM7 (Places), kjer je INSPIRE opredeljen kot primer dobre prakse na ravni Evrope, predvsem zaradi standardizacije in harmonizacije prostorskih podatkov na ravni EU.

Omenjeni MIM7 se osredotoča tudi na integracijo prostorskih omrežnih storitev. Glavno vlogo še naprej glavno igra OGC, kjer so v pripravi specifikacije novih tipov standardnih storitev, ki temeljijo na REST API tehnologiji, npr. OGC API Features specifikacija kot nadgradnja OGC WFS standarda in OGC SensorThingsAPI specifikacija, ki je namenjena povezovanju IoT naprav.

Ekosistem evropskih in slovenskih organizacij

Evropska komisija zagotavlja pogoje za implementacijo podatkovnih prostorov, aktivno izvajanje, postavljanje tehnoloških standardov in razvoj tehničnih rešitev pa prepušča deležnikom podatkovnega ekosistema. S tega vidika je ustanovljena DSBA (Data Spaces Business Alliance), ki združuje štiri najbolj pomembna evropska združenja na področju podatkov: BDVA (Big Data Value Association), FIWARE, Gaia-X in IDSA (International Data Spaces Association). BDVA že objavlja skupna stališča in referenčne arhitekture za podatkovne prostore. Začeli so se prvi pilotski projekti implementacije osnovnih komponent in primerov uporabe. BDVA se osredotoča na vrednost podatkov in podatkovno ekonomijo, FIWARE s implementacijo tehničnih rešitev in standardov, Gaia-X z upravljanjem podatkov in protokoli za posredovanje med ponudnikom in uporabnikom podatkov, IDSA pa postavlja referenčne arhitekture in pravila za združevanje podatkov. Infrastruktura INSPIRE bo v tem kontekstu zagotovila predvsem harmonizirane prostorske podatke in storitve, ki bodo dobra začetna osnova za vzpostavitev skupne evropske podatkovne infrastrukture.

IKT HM je član FIWARE, BDVA in Gaia-X. Vzpostavljen je tudi Gaia-X Hub Slovenia (GXH-SI)⁶. GXH-SI je glas uporabniških ekosistemov na nacionalni ravni. Glavni cilji so razviti ekosisteme, povezati nacionalne pobude in zagotoviti osrednjo kontaktno točko zainteresiranim stranem v Sloveniji. GXH-SI se povezuje z ostalimi stičišči evropskih držav, z namenom, da bi zagotovili mednarodno usklajenost glede svojih dejavnosti, opredelitve zahtev in prepoznavanja regulativnih ovir. IKT HM je tudi soustanovitelj Stičišča odprtih podatkov Slovenije – OPSI HUB, ki se prizadeva za celostno sodelovanje k spodbujanju in promociji ter razvoju odpiranja podatkov ter ponovne uporabe odprtih podatkov z namenom, da se razvijejo podjetniške, družbene in socialne vrednosti, ki jih nudijo odprti podatki.

IKT HM je bila aktivna pri standardizaciji podatkovnih modelov za pametna mesta in skupnosti⁷, kjer so na voljo priporočila, referenčna arhitektura in standardi. Trenutno so v pripravi smernice za uporabo GIS gradnikov za doseg interoperabilnosti na področju pametnih mest. IKT HM je bila tudi aktivna pri vzpostavitvi povezave lokalnih meta podatkovnih sistemov z državnim (INSPIRE) meta podatkovnim sistemom in preko slednjega z OPSI.

⁶ <https://gaia-x.gzs.si/>

⁷ <https://smartsociety.gzs.si/standardi/standardizacija>

IKT HM se bo tudi v prihodnje aktivno vključevala v razvoj osnovnih tehnoloških gradnikov, standardizacijo in gradnjo skupnih evropskih infrastruktur, saj želi Slovenija postati referenčna država na področju platform in rešitev za pametna okolja.

Reference in literatura

- BDVA. (4 2019). BDVA Position Paper: TOWARDS A EUROPEAN DATA SHARING SPACE. Pridobljeno iz https://www.bdva.eu/sites/default/files/BDVA%20DataSharingSpace%20PositionPaper_April2019_V1.pdf
- Evropska komisija. (3. 6 2020). Building the Green Data Space. Pridobljeno iz https://inspire.ec.europa.eu/sites/default/files/inspire2020_greendataspace_green_deal_data_space.pdf
- Evropska komisija. (19. 2 2020). Evropska strategija za podatke. Pridobljeno iz <https://eur-lex.europa.eu/legal-content/SL/TXT/?qid=1593073685620&uri=CELEX:52020DC0066>
- Evropska komisija. (1. 1 2021). Building a data economy — Brochure. Pridobljeno iz <https://digital-strategy.ec.europa.eu/en/library/building-data-economy-brochure>
- Evropska komisija. (21. 10 2021). INSPIRE, a public sector contribution to the European green deal data space. Pridobljeno iz <https://op.europa.eu/sl/publication-detail/-/publication/510cd3aa-32e4-11ec-bd8e-01aa75ed71a1>
- FIWARE. (6 2021). Fiware for Data Spaces - Position Paper. Pridobljeno iz <https://www.fiware.org/marketing-material/fiware-for-data-spaces/>
- Gaia-X AISBL. (21. 12 2021). Gaia-X - Architecture Document. Pridobljeno iz https://gaia-x.eu/sites/default/files/2022-01/Gaia-X_Architecture_Document_2112.pdf
- IDSA. (4 2019). International Data Spaces Association - Reference Architecture Model. Pridobljeno iz <https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf>
- IDSA. (4 2021). Design Principles for Data Spaces - Position Paper. Pridobljeno iz <https://design-principles-for-data-spaces.org/>
- IDSA. (9 2021). International Data Spaces Association - Enabling Data Economy. Pridobljeno iz https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-brochure-International-Data-Spaces-Enabling-Data-Economy.pdf