

Inovacije v arhitekturah podatkovnih prostorov

Nina Kliček, Martina Šestak, Muhamed Turkanović

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko,
Maribor, Slovenija
nina.klicek1@um.si, martina.sestak@um.si, muhamed.turkanovic@um.si

Podatkovni prostori predstavljajo odprto in povezano infrastrukturo za varno izmenjavo podatkov v skladu s skupnimi pravili, standardi in politikami. Evropska komisija je februarja 2020 predstavila Evropsko strategijo za podatke z jasnim ciljem ustvariti enoten trg za podatke, ki bi omogočil polno izkoriščanje vrednosti podatkov v korist evropske družbe in gospodarstva. Strategija določa smernice za oblikovanje skupnih evropskih podatkovnih prostorov na več ključnih področjih. Ključni izzivi so zagotavljanje interoperabilnosti na različnih ravneh ter osredotočenost na dostopnost in suverenost podatkov (angl. Data Sovereignty), kar lahko ovira implementacijo minimalno funkcionalnih podatkovnih prostorov. V prispevku bomo osvetlili pomen in kompleksnost arhitektur podatkovnih prostorov. Pregledali bomo trenutno stanje pristopov k vzpostavitvi podatkovnih prostorov, pri čemer se bomo osredotočili na identifikacijo in analizo izzivov v specifičnih domenah in glavnih deležnikov podatkovnih prostorov. Poudarek bo na IT arhitekturnih predstavitev, s podrobnim pregledom obstoječih pristopov in izzivov pri vzpostavitvi podatkovnih prostorov. Naša analiza bo vključevala ključne komponente, kot so integracija podatkov, upravljanje dostopa, zagotavljanje varnosti in skladnosti s predpisi, ter sodelovanje med različnimi deležniki. Tako bomo oblikovali celovito sliko o trenutnem stanju in prihodnjih usmeritvah na področju podatkovnih prostorov. Dodatno bomo predstavili inovativni pristop, ki omogoča identifikacijo podatkov in upravljanje dostopa do njih s pomočjo decentraliziranih tehnologij.

Ključne besede:

podatkovni prostor
referenčna arhitektura
interoperabilnost
IT arhitekture
varna izmenjava podatkov

1 Uvod

Podatki so ključna dobrina sodobne digitalne družbe. Prav tako pa lahko dandanes z uporabo interneta vsak posameznik ali organizacija med seboj deli podatke. Ko je v izmenjavo podatkov vpletenih več akterjev se hitro soočimo z različnimi izzivi, kot je zagotavljanje dostopnosti podatkov le pooblaščenim osebam in njihovo zaščito itn. Oblikovanje okvirov in reguliranih okolij za izmenjavo podatkov je zelo zapleteno, še posebej če je v izmenjavo podatkov vpletenih več strank, ki imajo različne interese in morajo slediti zapletenim predpisom o varovanju podatkov. Kot celovita rešitev za naslavljanje izzivov, ki se pojavljajo pri izmenjavi in povezovanju podatkov, so zasnovani podatkovni prostori.

Podatkovno gospodarstvo EU27 je leta 2019 doseglo vrednost skoraj 325 milijard evrov, kar predstavlja 2,6 % BDP. Po ocenah EU naj bi se ta vrednost do leta 2025 povečala na več kot 550 milijard EUR [1]. Temeljna načela podatkovnega gospodarstva temeljijo na vrednosti podatkov, interoperabilnosti, dostopnosti in suverenosti podatkov. Vsi ti elementi se združujejo v konceptu podatkovnih prostorov, ki predstavljajo odprto in zvezno infrastrukturo za varno izmenjavo podatkov v skladu s skupnimi pravili, standardi in politikami [2]. Podatkovni prostori imajo velik potencial za spodbujanje trajnostnega gospodarstva. Omogočajo namreč empirično reševanje problemov s celovitim pristopom, ki vključuje vse zainteresirane strani. To spodbuja inovacije, ki temeljijo na podatkih, in odkrivanje novih priložnosti [3].

Februarja 2020 je Evropska komisija (EK) predstavila Evropsko strategijo za podatke (angl. European Strategy for Data) z jasnim ciljem: ustvariti enoten trg za podatke, ki bi omogočil polno izkoriščanje vrednosti podatkov za koristi evropske družbe in gospodarstva, pri tem pa zagotavljati vso suverenost vseh vpletenih. Strategija določa smernice za oblikovanje skupnih evropskih podatkovnih prostorov na več ključnih področjih: večšine, zdravje, kmetijstvo, proizvodnja, energija, mobilnost, finance ter javna uprava [4].

Pričakuje se, da bodo podatkovni prostori ključna tehnološka platforma za izkoriščanje potenciala globalnega podatkovnega ekosistema. Da bi lahko izkoristili potencial podatkovnih prostorov, je potrebna široka razširjenost ideje in podprtost le te s tehnologijo. Trenutno poteka drugi val pobud za podatkovne prostore, ki se osredotoča na širše uvajanje podatkovnih prostorov, ki presegajo specifične primere uporabe. Spodbuja se splošna implementacija podatkovnih prostorov, kar bo prispevalo k širši uporabi in razvoju novih možnosti [5].

2 Podatkovni prostori

Podatkovni prostori delujejo kot združena platforma za deljenje in izmenjavo podatkov med različnimi entitetami, zagotavljajo potrebna orodja in varnostne ukrepe za varno izmenjavo in uporabo podatkov. Osrednji cilj in pomen podatkovnih prostorov pa je izboljšati dostopnost in uporabo podatkov s strani širokega kroga deležnikov, kot tudi spodbujanje inovacij in raziskav na področju digitalnega gospodarstva [6]. Podatkovni prostori tako omogočajo ustvarjanje novih vrednostnih verig, ki temeljijo na podatkih, ter omogočajo podjetjem in institucijam, da bolje izkoristijo svoje podatkovne vire.

Za lažjo predstavitev ključnih komponent podatkovnih prostorov, bi lahko elemente podatkovnega prostora razdelili v dve kategoriji: (1) **vloge** in (2) **temeljne komponente**.

2.1. Vloge

V podatkovnem prostoru lahko najdemo različne udeležence, kjer je vsak osredotočen na določeno področje delovanja. Mednarodna zveza podatkovnih prostorov (angl. IDSA) vloge, ki se pojavijo v podatkovnih prostorih razdeli v štiri kategorije:

- Kategorija 1: **Glavni/osrednji udeleženeec** (angl. *Core Participant*)
- Kategorija 2: **Posrednik** (angl. *Intermediary*)
- Kategorija 3: **Razvijalec programske opreme** (angl. *Software Developer*)

- Kategorija 4: **Organ upravljanja** (angl. *Governance Body*)

Osrednji udeleženci so vključeni in obvezni pri vsaki izmenjavi podatkov v mednarodnem podatkovnem prostoru. Vlogi, dodeljeni tej kategoriji, sta **dobavitelj** podatkov (angl. *Data Supplier*) in **odjemalec** podatkov (angl. *Data Customer*).

Dobavitelj podatkov je vloga, ki ponuja podatke v ekosistem podatkovnih prostorov. Odvisno od posameznega poslovnega in tehničnega modela delovanja poslovna vloga dobavitelj podatkov običajno prevzame osnovne vloge, kot so:

- **ustvarjalec podatkov** (Data Creator),
- **lastnik podatkov** (Data Owner) in/ali
- **ponudnik podatkov** (Data Provider).

Odjemalec podatkov prejme podatke od ponudnika podatkov. Z vidika modeliranja poslovnih procesov je odjemalec podatkov zrcalna entiteta ponudnika podatkov; dejavnosti, ki jih opravlja odjemalec podatkov, so zato podobne dejavnostim, ki jih opravlja ponudnik podatkov.

Posredniki, pogosto imenovani „platforme“, služijo kot zaupanja vredne entitete, ki imajo osrednjo vlogo pri izmenjavi podatkov med številnimi ponudniki in odjemalci. Ključne poslovne vloge posrednikov so:

- **posrednik podatkov** (angl. *Data Intermediary*),
- **posrednik storitev** (angl. *Service Intermediary*),
- **trgovina z aplikacijami** (angl. *App Store*),
- **posrednik besednjaka** (angl. *Vocabulary Intermediary*),
- **posredniška hiša** (angl. *Clearing House*) in
- **organ za ugotavljanje identitete** (angl. *Identity Authority*).

Te platforme pogosto združujejo več vlog, na primer delujejo kot posredniki podatkov in storitev.

Kategorija 3 vključuje IT podjetja, ki dobavljajo programsko opremo udeležencem podatkovnih prostorov. Glavni vlogi v tej kategoriji sta **razvijalec aplikacij in posredniških storitev** in **razvijalec povezovalnikov**. Ti vlogi ustvarjata vrednost z zagotavljanjem programske opreme, potrebne za izmenjavo podatkov v okviru podatkovnega prostora.

Organi upravljanja imajo nalogo, da določijo in uveljavijo smernice za standardizacijo izmenjave podatkov, vzpostavijo zaupanje in na koncu omogočijo trajnostno delovanje sistema.

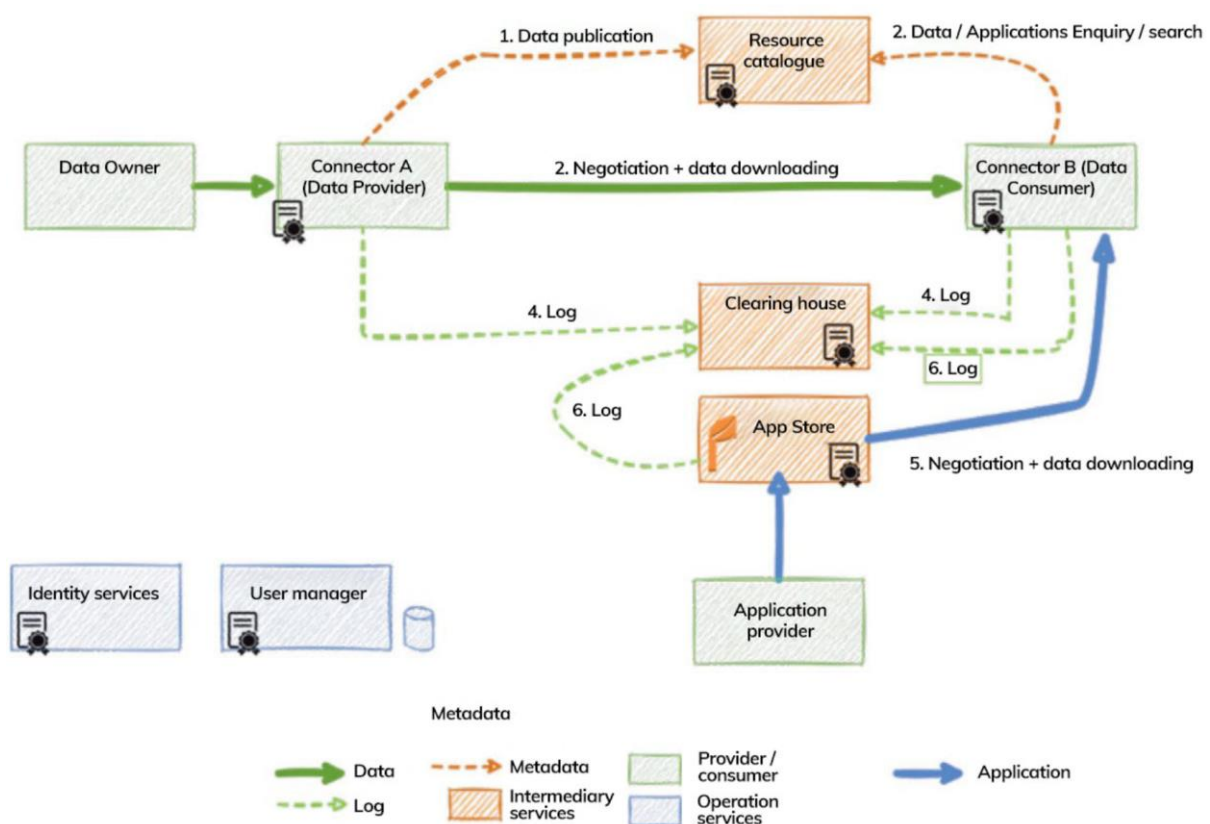
2.2. Temeljne komponente

Za varno in nadzorovano izvajanje dejavnosti (transakcij/operacij) v podatkovnem prostoru so potrebne naslednje **temeljne komponente** [7] :

- Komponente za dostop do podatkovnega prostora:
 - o **Povezovalnik (angl. *Connector*)** – Je eden od glavnih elementov podatkovnih prostorov, preko katerega udeleženci dostopajo do podatkovnega prostora in podatkov. Odgovoren je za ravnanje s podatki v skladu s politikami uporabe, ki jih opredeli lastnik podatkov, s čimer zagotavlja suverenost. Da bi preprečili zlonamerno manipulacijo, morajo biti povezovalniki podpisani s certifikatom, ki ga zagotovi upravljavec podatkovnega prostora.
- Komponente za posredovanje, ki omogočajo posredniške storitve te so:
 - o **Slovarji in ontologije (angl. *Vocabularies and Ontologies*)**, ki omogočajo sistematično organizacijo, kategorizacijo ali označevanje informacij, kar izboljšuje interoperabilnost.
 - o **Skladišča aplikacij (angl. *Application Stores*)**, ki vsebujejo seznam orodij, ki jih ponujajo ponudniki aplikacij, s čimer se zagotovi, da so uspešno prestala nadzor kakovosti.

- **Storitve posrednikov metapodatkov (angl. *Metadata Broker Services*)** poskrbijo za objavo kataloga ponudbe virov (podatkov in aplikacij) s čim več informacijami.
 - **Storitve orkestracije (angl. *Orchestration Services*)**, ki omogočajo avtomatizacijo različnih dejavnosti.
 - **Storitve za izmenjavo podatkov (angl. *Clearing House*)**, ki omogočajo nadzor nad izvedenimi dejavnostmi.
- Komponente za upravljanje identitete in varno izmenjavo podatkov:
 - Te komponente zagotavljajo identiteto udeležencev in varnost transakcij. Zato morajo udeleženci pogosto predložiti poverilnice (npr. s potrdili X.509).
 - Komponente za upravljanje podatkovnega prostora:
 - To so orodja, ki omogočajo normalno delovanje podatkovnega prostora, olajšujejo vsakodnevno delovanje, upravljanje udeležencev (registracija, izbris, preklic, začasna ukinitvev), spremljanje dejavnosti itd.

Vse predstavljene vloge in komponente sodelujejo druga z drugo (Slika 1). Najprej ponudnik podatkov (angl. *Data Provider*) registrira svojo ponudbo podatkov v katalogu, vključno z ustreznimi metapodatki, kot so politike uporabe. Potrošnik podatkov (angl. *Data Consumer*) v katalogu poišče zanj zanimive podatkovne zbirke in aplikacije. Ko jih najde, stopi v stik s ponudnikom in mu sporoči, katere vire želi pridobiti. V tem postopku lahko potekajo nadaljnja pogajanja o pogojih uporabe podatkov. Ko je dogovor dosežen, lahko potrošnik dostopa in/ali prenese podatke [7].



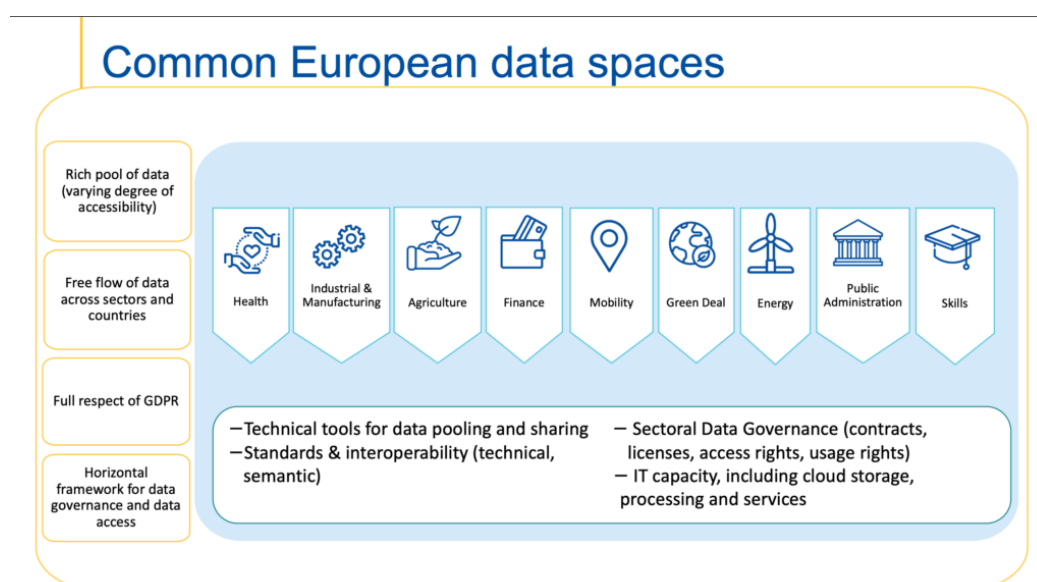
Slika 1: Scenarij izmenjave podatkov v podatkovnem prostoru. [7]

2.3. Evropska strategija za podatke in njen vpliv na oblikovanje podatkovnih prostorov

Evropska strategija za podatke, sprejeta leta 2020, je ambiciozen načrt Evropske komisije za ustvarjanje enotnega evropskega podatkovnega trga. Strategija potrjuje dejstvo, da podatki predstavljajo ključno sredstvo za inovacije, gospodarsko rast in izboljšanje javnih storitev.

Ta strategija stremi k varni in učinkoviti izmenjavi podatkov med različnimi sektorji in državami članicami. Ključni elementi strategije vključujejo ustvarjanje evropskih podatkovnih prostorov, zagotavljanje interoperabilnosti, spodbujanje podatkovne suverenosti ter podpora malim in srednje velikim podjetjem. Vizija Evropske unije (EU) je, da bo s tem povečala gospodarsko rast, izboljšala konkurenčnost evropskih podjetij v svetu in zagotovila boljše storitve za državljane, pri tem pa je pomembno omeniti, da je posebna pozornost posvečena tudi etičnim vidikom zbiranja podatkov in varovanja osebnih podatkov [8].

Razkritje podatkov zunaj organizacijskih meja odpira ogromno možnosti za skupne inovacije in soustvarjanje vrednosti z izmenjavo in ponovno uporabo podatkov. Kljub temu pa souporaba podatkov med partnerji pogosto ostaja omejena ali se izvaja skozi zaprte rešitve znotraj posameznih organizacij, kar preprečuje popolno izkoriščanje potenciala podatkov. Tveganja, povezana z deljenjem podatkov, vključujejo izgubo nadzora nad podatki, kršitve zasebnosti in možnosti zlorabe podatkov. Zaradi teh tveganj so bili uvedeni številni zakoni in regulative, ki urejajo varnost in deljenje podatkov. Med ključnimi zakonodajnimi okviri sta **Zakon o upravljanju podatkov** (angl. *Data Governance Act – DGA*) in Splošna uredba o varstvu podatkov (angl. *General Data Protection Regulation – GDPR*) [9]. DGA ima tudi kot referenco načrtovano ustvarjanje prvih skupnih evropskih podatkovnih prostorov, ki se vezana na ključne sektorje, kot so zdravstvo, kmetijstvo itn. (Slika 2). DGA določa smernice za razpoložljivost in deljivost podatkov, s čimer spodbuja ponovno uporabo podatkov med organizacijami ter med javnim in zasebnim sektorjem. Njegov cilj je olajšati izmenjavo podatkov z ureditvijo novih entitet, znanih kot posredniki podatkov, in spodbujanjem izmenjave podatkov iz altruističnih razlogov. DGA zajema osebne in neosebne podatke, pri čemer se GDPR uporablja, kadar gre za osebne podatke [10]. GDPR uvaja stroge zahteve za obdelavo osebnih podatkov, vključno s pridobivanjem privolitve posameznikov, pravico do dostopa do podatkov in pravico do pozabe. GDPR zagotavlja, da so osebni podatki obdelani na način, ki varuje zasebnost in pravice posameznikov [9].



Slika 2: Prvi načrtovani splošni evropski podatkovni prostori. [4]

Poleg tega je tu še Data Act, ki je začel veljati 11. januarja 2024. Je steber evropske strategije za podatke. Njegov glavni cilj je, da EU postane vodilna v podatkovnem gospodarstvu z izkoriščanjem potenciala vedno večje količine industrijskih podatkov, kar bo koristilo evropskemu gospodarstvu in družbi [11].

Skupaj ti zakoni tvorijo robusten regulativni okvir, ki omogoča varno in učinkovito izmenjavo podatkov ter štiti pravice posameznikov, hkrati pa spodbuja inovacije in ponovno uporabo podatkov na način, ki je skladen z evropskimi vrednotami in zakonodajo. Evropska strategija za podatke močno vpliva na oblikovanje in razvoj podatkovnih prostorov, saj spodbuja sprejetje standardov interoperabilnosti, varnosti in zasebnosti. Pobude, kot so Gaia-X in Mednarodni podatkovni prostori (IDS), so neposreden rezultat te strategije in delujejo kot model za vzpostavitev podatkovnih prostorov, ki podpirajo evropske cilje podatkovne suverenosti, gospodarske rasti in inovacij.

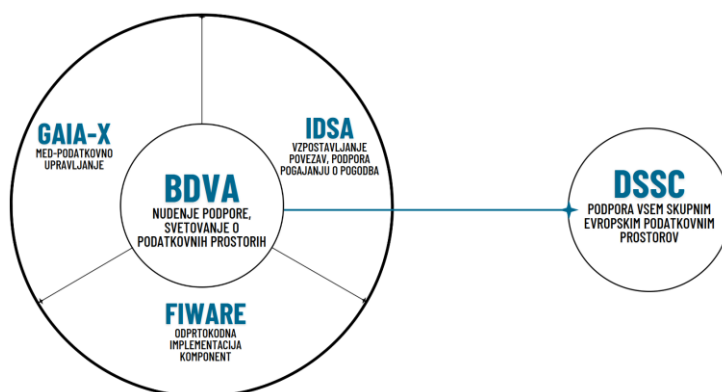
3 Trenutno stanje in pristopi k vzpostavitvi podatkovnih prostorov

Leta 2021 je bilo ustanovljeno poslovno združenje za podatkovne prostore (angl. DSBA – *Data Spaces Business Alliance*). DSBA sestavljajo GAIA-X, Asociacija vrednosti velepodatkov (angl. BDVA – *Big Data Value Association*), Fundacija FIWARE in Mednarodna asociacija podatkovnih prostorov (angl. IDSA – *International Data Spaces Association*). DSBA prispeva k podpornemu centru za podatkovne prostore (angl. DSSC – *Data Spaces Support Center*) pri usklajevanju in podpiranju vseh evropskih skupnih podatkovnih prostorov na različnih področjih [12].

Organizacije združenja DSBA so združile moči s ciljem, da bi zagotovile enoten tehnološki okvir za podatkovne prostore, pri čemer vse štiri nudijo različne ključne zmogljivosti za gradnjo skupnega ogrodja (Slika 3):

- BDVA je industrijsko usmerjena organizacija, ki si prizadeva za spodbujanje inovacijskega ekosistema, ki spodbuja digitalno preobrazbo evropskega gospodarstva in družbe, ki temelji na podatkih. Združenje BDVA, ki ima več kot 240 članov iz vse Evrope, vključno z velikimi, malimi in srednje velikimi podjetji, raziskovalnimi ustanovami in uporabniškimi organizacijami, se osredotoča na: tehnologije velikih podatkov, podatkovne platforme, industrijsko umetno inteligenco, ustvarjanje vrednosti na podlagi podatkov in razvoj spretnosti. BDVA sprejema nove člane, vključno z uporabniki, ponudniki in raziskovalci podatkov, ter podpira regionalno sodelovanje na evropski ravni, pri čemer poudarja soustvarjanje in eksperimentiranje. Čeprav BDVA ne pripravlja posebnih tehničnih standardov, zagotavlja svetovalne storitve, splošno svetovanje in podpirajo deljenje rešitev in rezultatov.
- Cilj fundacije FIWARE je vzpostaviti in spodbujati standarde odprtih platform za pametne rešitve v različnih sektorjih. S pristopom, ki temelji na odprti kodi in razvojno gnanem pristopu, fundacija prispeva k standardnim specifikacijam in se integrira z drugimi široko uporabljenimi tehnologijami. Fundacija FIWARE si skupaj z ostalimi člani prizadeva vplivati na razvoj specifikacij v ustreznih organizacijah in spodbuja njihovo hitro sprejetje na trgu s pomočjo odprto kodnih rešitev. Predvsem je vplivala na vmesnik NGSI-LD API za upravljanje podatkov digitalnega dvojčka.
- GAIA-X je globalno med-podatkovno upravljanje, ki temelji na evropskih vrednotah pomaga pri pripravi specifikacij, orodij in procedur za testiranje skladnosti izdelkov s temi specifikacijami. Gaia-X vzpostavlja skupni okvir upravljanja s posebnimi pravili za povezovanje podatkovnih in infrastrukturnih ekosistemov. Okvir Gaia-X vključuje funkcionalne specifikacije, tehnične specifikacije in kodo. Sodelovanje med podatkovnimi prostori je doseženo s skupnim upravljanjem in dodatnimi tehničnimi sredstvi, ki jih določa DSBA.

Mednarodno združenje za podatkovne prostore (IDSA) si prizadeva spodbujati zaupanja vredno okolje za izmenjavo podatkov z združenimi, globalno certificiranimi podatkovnimi prostori. S spodbujanjem izdelkov in sistemov, certificiranih s strani IDS. IDSA predvideva varno podatkovno gospodarstvo, v katerem lahko podjetja nemoteno izmenjujejo podatke v celotni vrednostni verigi. Ključni publikaciji sta Pravilnik IDSA o upravljanju in IDS-RAM V4, ki podrobno opisuje arhitekturo podatkovnega prostora. Protokol podatkovnega prostora omogoča interoperabilno izmenjavo podatkov in sporazume o uporabi, ki temeljijo na spletnih tehnologijah. Certificiranje, ki je ključno za doseganje podatkovne suverenosti, se upravlja prek IDS-Reference-Testbed. IDSA poudarja razvoj in potrjevanje povezovalnikov podatkovnega prostora za pogajanja o politikah in njihovo uveljavljanje v podatkovnih prostorih [13].



Slika 3: Organizacije združenja DSBA.

Za ustvarjalce podatkovnih prostorov je izziv spremljati vse potencialno uporabne tehnologije. Vsak podatkovni prostor ima različne tehnološke zahteve, ki so odvisne od poslovnega primera, izbire upravljanja in veljavnih predpisov.

Ustvarjalci podatkovnega prostora morajo biti pragmatični glede uporabe standardov in poskušati uporabiti obstoječe rešitve zunaj področja podatkovnega prostora za testiranje primerov uporabe, preden se zavežejo k razvoju [14].

Več akterjev na področju tehnologije podatkovnega prostora si prizadeva za oblikovanje standardov za podatkovni prostor in skupnih referenčnih tehnoloških okvirjev. Ko se začnemo seznanjati s podatkovnimi prostori, se kmalu seznanimo z naslednjimi akterji poleg že omenjenih, BDVA, IDSA, GAIA-X in FIWARE:

- Data Spaces Support Centre (DSSC) - projekt, financiran s strani EK v okviru programa Digitalna Evropa (DEP). DSSC raziskuje potrebe po pobudah podatkovnih prostorov, opredeljuje skupne zahteve in vzpostavlja najboljše prakse za pospešitev oblikovanja suverenih podatkovnih prostorov kot ključnega elementa digitalne preobrazbe na vseh področjih.
- iSHARE - je evropsko omrežje za mednarodno in suvereno izmenjavo poslovnih podatkov, upravljano s strani Fundacije iSHARE. iSHARE omogoča federirano upravljanje zaupanja podatkovnih prostorov. Ponuja komponente podatkovnih prostorov v skladu z načeli oblikovanja za podatkovne prostore iz projekta Open DEI, IDSA in Gaia-X.
- Eclipse Cross Federation Services Components (XFSC) - je odprtokodni projekt, ki razvija osnovne komponente, potrebne za vzpostavitev federiranih sistemov za izmenjavo podatkov. Pred prenosom na fundacijo Eclipse je bil projekt znan kot Storitve federacije Gaia-X (GXFS).
- Pontus-X, the Gaia-X Web3 ecosystem - si prizadeva zagotoviti decentraliziran in federativni pristop k upravljanju podatkov, kar omogoča, da se podatki, programska oprema, infrastruktura in federacijske storitve varno gradijo, zbirajo, delijo in monetizirajo.

- Eclipse Dataspace Components (EDC) - je odprtokodni projekt, ki si prizadeva implementirati standard Mednarodnih podatkovnih prostorov (IDS) in ustrezne protokole in zahteve, povezane z Gaia-X, ter s tem zagotoviti implementacijo in povratne informacije za te pobude

Komericalne ponudbe, ki se osredotočajo na podatkovne prostore, lahko zagotovijo dragoceno podporo pri obravnavi razvijajočih se različnih arhitektur in programske opreme. Trenutno se še majhno, vendar stalno naraščajoče število podjetij osredotoča predvsem na podatkovne prostore ali uvaja izdelke in storitve, specifične za podatkovne prostore, kot del večjega portfelja. Nekatera med njimi so: Advaneo, Dataspace Europe, deltaDAO, IONOS, nexyo, OKP4, sovity in TrustRelay [14].

Ko raziskujemo področje podatkovnih prostorov in implementacij le teh, moramo omeniti tudi radar podatkovnih prostorov (angl. *Data Space Radar*), ki je koristen vir za vse, ki so vključeni v razvoj, upravljanje in uporabo podatkovnih prostorov ter za tiste, ki želijo spremljati in razumeti dinamične spremembe v tem hitro razvijajočem se področju [15].

Radar podatkovnih prostorov, prikazan na sliki 4, je javno dostopno interaktivno orodje, namenjeno panoramskemu pregledu različnih pobud za podatkovne prostore po vsem svetu. Omogoča vpogled v sektorje, lokacije in razvojne faze teh pobud ter primere uporabe, ki jih omogočajo podatkovni prostori.



Slika 4: Interaktivno orodje Data Space Radar. [15]

V poročilu The Data Space Radar – Data Spaces Support Centre Edition, objavljeno marca 2024, najdemo predstavljajo izbrane člane skupnosti, ki prihajajo iz različnih sektorjev, kot so pametne skupnosti, proizvodnja, zdravje. Kar dokazuje raznolikost pokrajine podatkovnih prostorov in medpodročno poslanstvo skupnosti.

3.1. Identifikacija specifičnih izzivov na različnih področjih implementacije podatkovnih prostorov

Podatkovni prostori predstavljajo pomemben element v digitalizaciji različnih sektorjev, saj omogočajo učinkovito izmenjavo, upravljanje in uporabo podatkov. Vendar pa se vsako področje sooča z edinstvenimi izzivi, ki vplivajo

na dostopnost, varnost, interoperabilnost in skladnost podatkov. V nadaljevanju bomo predstavili nekatere izzive na področjih, ki jih je opredelila EK (Slika 2).

Zdravje (angl. Health)

Evropski zdravstveni podatkovni prostor (EHDS) je prvi skupni evropski podatkovni prostor, ki bo kmalu urejen z uredbo EU. Izzivi na tem področju vključujejo [17]:

- Pravna skladnost: Boljša uskladitev z zakonodajo EU in nacionalno zakonodajo za preprečevanje pravne negotovosti in zmede.
- Pravice in obveznosti: Jasna opredelitev obveznosti in pravic zdravstvenih delavcev glede podatkov, da ne bi bili odgovorni za podatke, ki jih neposredno ne nadzorujejo.
- Definicije: Jasne in koherentne definicije, kot so 'elektronski zdravstveni podatki', 'imetnik podatkov', 'sistem elektronskih zdravstvenih zapisov', za zagotovitev zakonodajne skladnosti z drugimi pravnimi akti EU.
- Varstvo poslovnih skrivnosti: Omogočiti imetnikom podatkov upravljanje s poslovnimi skrivnostmi in zagotoviti ustrezne ukrepe za njihovo zaščito.
- Prenosi podatkov: Omejitev obdelave zdravstvenih podatkov znotraj EU ter omogočiti prenose v tretje države le pod ustreznimi pogoji, skladnimi z GDPR.
- Mehanizmi za izključitev: Obvezen mehanizem za izključitev posameznika iz podatkovnega prostora, ki se uporablja enotno po vseh državah članicah.

Zdravstveni podatki bodo za sekundarno uporabo v EHDS večinoma deljeni v anonimizirani ali psevdonimizirani obliki. Vendar se pravne definicije in zahteve za anonimizacijo in psevdonimizacijo močno razlikujejo med državami, kar ovira delovanje farmacevtskih, digitalnih in zdravstvenih tehnoloških podjetij ter osnovni cilj EHDS - omogočiti nemoteno sekundarno uporabo podatkov čez meje. Na ravni EU so potrebna natančnejša navodila, kako izpolniti zahteve GDPR na način, ki uravnoteži tveganja in koristi ter odpravi pravno negotovost in strah pred tožbami za sekundarne uporabnike podatkov. Ključno je ohraniti pravice posameznikov glede varstva podatkov [18].

Kmetijstvo (angl. Agriculture)

V kmetijskem podatkovnem prostoru se pojavljajo specifični izzivi, kot so [19]:

- Povezljivost in infrastruktura: Številna podeželska območja nimajo hitrega interneta, kar zahteva zbiranje podatkov brez povezave ter kasnejšo sinhronizacijo.
- Digitalna izobrazba: Kmetje, večinoma strukturirani v malih in srednje velikih podjetjih, potrebujejo izobraževanje na področju digitalnih tehnologij.
- Podatkovna suverenost: Kmetje pogosto menijo, da pri uporabi digitalnih storitev izgubijo nadzor nad svojimi podatki. Rešitve morajo omogočati enostaven nadzor nad podatkovno suverenostjo.
- Interoperabilnost: Zagotoviti interoperabilnost vzdolž celotne vrednostne verige, da vse zainteresirane strani koristijo podatke za analize in podpora pri odločanju.

Proizvodnja (angl. Manufacturing)

Proizvodni sektor se sooča z izzivi, kot so [20]:

- Upravljanje podatkov: Naraščajoča količina podatkov, ki se zbirajo na vseh stopnjah (tovarna, izdelek, dobavna veriga), zahteva učinkovito upravljanje in analizo.
- Integracija strojnega učenja: Povezovanje lokalnih modelov strojnega učenja, usposobljenih na posameznih lokacijah, v enotno in natančnejšo podatkovno okolje.
- Deljenje podatkov: Različni uporabniki zbirajo podatke o različnih napakah in uspehih, kar vodi v različne učne nize. Učinkovito deljenje teh podatkov v varnem podatkovnem prostoru je izziv.

Energija (angl. Energy)

Energetski sektor se sooča z izzivi, kot so [21]:

- Sodelovanje med konkurenti: Velika podjetja v energetiki, pogosto konkurenti, morajo usklajeno sodelovati za nove storitve - usklajevanje prizadevanj za rast poslovanja, družbeni napredek in doseg ogljične nevtralnosti.
- Zbiranje heterogenih podatkov: Zahteva sodelovanje celotne vrednostne verige energetskega upravljanja, kar vključuje ponudnike energije, izvajalce, operaterje omrežij, agregatorje, potrošnike, upravljavce polnilnih postaj in varnostne organe.

Mobilnost (angl. Mobility)

Mobilnost se sooča z izzivi, kot so [22]:

- Konkurenčnost ponudnikov: Ponudniki različnih prevoznih možnosti (javni prevoz, taksiji, najem vozil, e-skuterji) ne zaupajo drug drugemu, kar preprečuje deljenje podatkov.
- Skladnost z GDPR: Uredba GDPR se uporablja kot izgovor za ne deljenje ključnih podatkov, kar ovira sodelovanje.
- Realno-časovno deljenje podatkov: Potreba po deljenju podatkov (lokacija, razpoložljivost, prometni pogoji) v skoraj realnem času zahteva zanesljivo infrastrukturo.

Finance (angl. Financial)

Finančni sektor se sooča z izzivi, kot so [23]:

- Pravna fragmentacija: Kljub prizadevanjem za harmonizacijo pravil obstaja fragmentacija na ravni držav članic pri implementaciji in nadzoru.
- Vrednost za stranke: Zagotavljanje dovolj velike vrednosti za stranke, da bi delile svoje podatke.
- Varovanje podatkov: Varovanje uporabe podatkov v kontekstu umetne inteligence in njihovega združevanja z drugimi viri podatkov.

Skupni izzivi vključujejo zagotavljanje pravne in varnostne skladnosti ter interoperabilnosti podatkov med različnimi sistemi in deležniki. Specifični izzivi v sektorjih, kot so zdravje, kmetijstvo, proizvodnja, energija, mobilnost in finance, so pogosto povezani z infrastrukturo, sodelovanjem in izobraževanjem. Podatkovni prostori za večšine in javno upravo trenutno ne izkazujejo posebej specifičnih izzivov, verjetno tudi zato, ker na tem področju še ni veliko raziskav.

4 Arhitektura

Podatkovni prostori določajo strukturo in način, kako se podatki zbirajo, shranjujejo, obdelujejo in delijo med različnimi deležniki. Eden od ključnih pristopov pri oblikovanju podatkovnih prostorov je poznavanje referenčne arhitekture, ki zagotavlja temeljne smernice za interoperabilnost, varnost in skladnost v podatkovnih prostorih. Te referenčne arhitekture služijo kot ogrodje za gradnjo podatkovnih prostorov, ki omogočajo varno izmenjavo podatkov med organizacijami, hkrati pa zagotavljajo visoko raven zaupanja in zaščite podatkov.

IDS-RAM predstavlja pet slojev referenčne arhitekture, katerih namen je zagotavljanje dobro strukturiranega in obširnega pristopa k podatkovnim prostorom (Slika 5). Sloji zagotavljajo jasno organizacijo opredelitve podatkovnih prostorov. Vsak sloj se osredotoča na določen vidik, pri tem pa ločuje izzive in spodbuja modularnost. To omogoča lažje razumevanje, razvoj in vzdrževanje podatkovnih prostorov. Čeprav vsak sloj naslavlja svoje področje, obstajajo pomembni vidiki, ki vplivajo na vse sloje. IDS RAM to rešuje z vključitvijo treh presečnih vidikov: varnost, certificiranje in upravljanje. To zagotavlja, da so ti pomembni vidiki upoštevanji na vseh slojih [24].



Slika 5: Sloji in perspektive referenčne arhitekture IDS RAM. [24]

Poslovni sloj:

- Opredeljuje in razvršča različne vloge, ki jih lahko prevzamejo udeleženci v mednarodnih podatkovnih prostorih in določa osnovne vzorce interakcije med temi vlogami.

Funkcionalni sloj:

- Opredeljuje funkcionalne in nefunkcionalne zahteve mednarodnih podatkovnih prostorov in funkcije, ki jih je treba zagotoviti.

Informacijski sloj:

- Določa informacijski model, skupni jezik, tj. besednjak mednarodnih podatkovnih prostorov, ki ni odvisen od domene kar omogoča združljivost in interoperabilnost, med udeleženci in ostalimi komponentami.

Procesni sloj:

- Določa interakcije med različnimi sestavnimi deli podatkovnih prostorov, ter vsebuje opise procesov in njihovih podprocesov. Ti procesi so povezani s ključnimi predlogi mednarodnega podatkovnega prostora in vključujejo večino vlog, predstavljenih v poslovni plasti.

Sistemski sloj:

- Vloge, ki so določene na poslovnem sloju, in procesi, opredeljeni v procesnem sloju, so preslikani na konkretno podatkovno in storitveno arhitekturo, to je tehnično jedro mednarodnih podatkovnih prostorov.

Perspektiva varnosti:

- Varnostna arhitektura IDS zagotavlja sredstva za identifikacijo naprav v IDS, zaščito komunikacije in transakcij izmenjave podatkov ter nadzor uporabe podatkov po njihovi izmenjavi.

Perspektiva certificiranja:

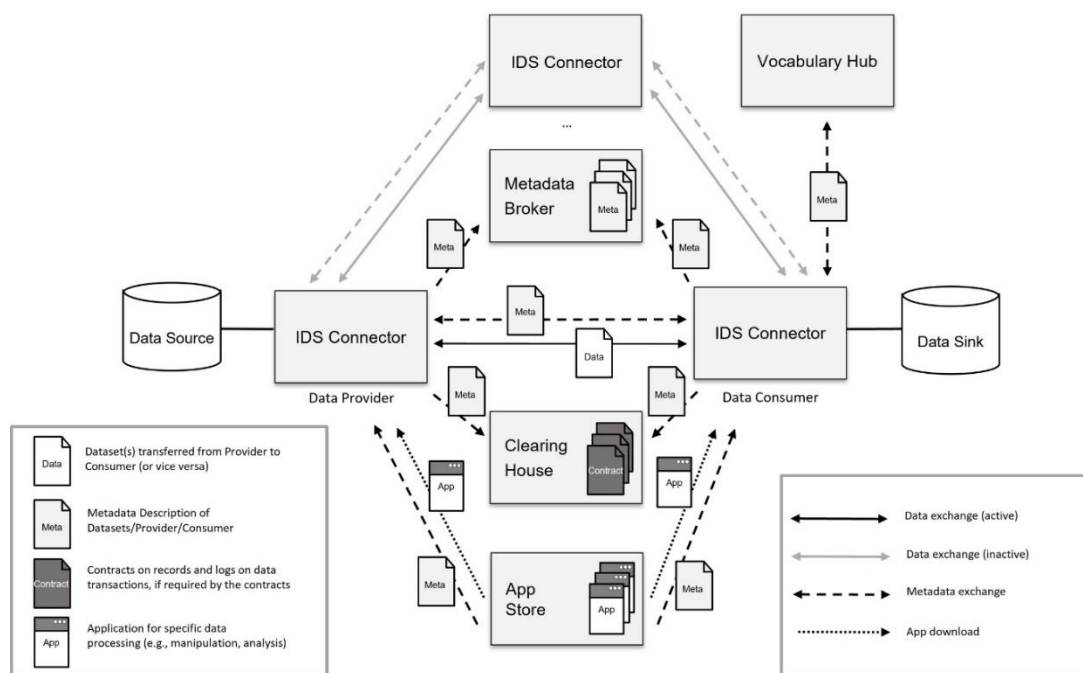
- Za omogočanje nadzora, mora vsak udeleženec upoštevati dogovorjena pravila, zato morajo komponente, preden jih je mogoče uporabiti v sistemu IDS, opraviti certificiranje.

Perspektiva upravljanja:

- Podpira mehanizme upravljanja na podlagi sodelovanja, tako da se dosežejo skupne storitve in ponudbe vrednosti ob hkratni zaščiti interesov vseh akterjev.

S temi ločenimi, a medsebojno povezanimi plastmi sistem IDS RAM zagotavlja celovit okvir za izgradnjo varnih, interoperabilnih in učinkovitih podatkovnih prostorov [24].

Procesi, opredeljeni v procesni plasti, so na sliki 6 povzeti kot interakcije med komponentami IDS. Pri tem pa moramo upoštevati, da ponudnik identitete na sliki ni prikazan zaradi ohranjanja berljivosti. Podatkovni prostori temeljijo na povezovanju različnih udeležencev, pri čemer imajo ti nameščene IDS povezovalnike ali druge osnovne komponente. IDS povezovalnik je odgovoren za sprožitev izmenjave podatkov med notranjimi podatkovnimi viri – podjetniškimi sistemi sodelujočih organizacij in mednarodnimi podatkovnimi prostori. Metapodatkovnemu posredniku zagotavlja metapodatke, kot je določeno v samoopisu IDS povezovalnika, vključno z opisom tehničnega vmesnika, mehanizmi avtentikacije in povezanimi politikami uporabe podatkov. Pogodbe o uporabi podatkov se lahko preko IDS povezovalnika prenesejo v klirinško hišo za zagotavljanje zaupanja. Besedišča je mogoče razlagati tako, da si v vozlišču slovarjev (angl. *Vocabulary Hub*) pridobite več podrobnosti. V IDS povezovalnik je mogoče prenesti dodatne IDS aplikacije za izvajanje različnih operacij na podatkih [25].



Slika 6: Interakcija tehničnih komponent. [25]

Glavna prednost referenčne arhitekture IDS in uporabe povezovalnika IDS je decentralizirano shranjevanje podatkov. To omogoča integracijo podatkov iz različnih virov podatkov in dostop do podatkov izključno prek drugih povezovalnikov IDS. Tako pa je zagotovljena tudi tehnična izvedba podatkovne suverenosti [26].

V podatkovnih prostorih podatkovne zbirke soobstajajo, vendar niso nujno popolnoma integrirane ali homogene v svojih shemah in semantiki. Namesto tega so podatki integrirani po potrebi. Jedro integracije je osredotočeno na entitete, ki predstavljajo entitete iz resničnega sveta. Poudarek je na tem, kako se te entitete in njihovi odnosi upravljajo znotraj podatkovnega prostora. Podatkovni prostori uporabljajo metodo sprotnega plačevanja (angl. *Pay-as-you-go*), kar pomeni, da je integracija postopna. Sistem se začne z osnovno integracijo in se nadgrajuje skozi čas. Ta pristop omogoča prilagodljivost in prilagajanje, ko se pojavijo novi podatki ali dodatne zahteve. Na

prvem mestu so torej podatki, model pa se prilagaja. To pomeni, da je začetni poudarek na zbiranju in integraciji razpoložljivih podatkov, nato pa se podatkovni model postopoma prečiščuje in prilagaja, da ustreza tem podatkom. V podatkovnem prostoru so heterogenimi podatki v porazdeljenem sistemu shranjevanja. To pomeni, da lahko obravnavajo različne vrste podatkov iz različnih virov, shranjenih na več lokacijah [27]. Za primerjavo integracije podatkovnega prostora je uporabljena podatkovna baza, kot je prikazano v tabeli spodaj.

Integracija podatkovnega prostora je prilagodljiva in inkrementalna (pay-as-you-go), saj omogoča obdelavo različnih vrst podatkov in porazdeljeno shranjevanje s poudarkom na entitetah in njihovih povezavah. Podatkovni model pa se po potrebi sproti prilagaja.

Tabela 1: Integracija podatkovnih prostorov in podatkovnih baz. [27]

Primerjava integracije	Podatkovni prostor	Podatkovna baza
Jedro integracije	Entitetni objekt	Poslovne potrebe
Metoda integracije	Pay-as-you-go	Pay-before-you-go
Način integracije	Najprej podatki, nato model	Najprej model
Integrirani podatki	Heterogeni podatki, porazdeljeno shranjevanje	Podatkovna shema, en sam podatkovni vir
Objekti integracije	Podatki in povezava podatkov	Podatki

4.1. Upravljanje dostopa

V mednarodnih podatkovnih prostorih (IDS), nadzor dostopa (angl. Access Control) uravnava zahteve za dostop do podatkovnih storitev. Lastniki podatkov določijo politike ABAC (angl. *Attribute-Based Access Control*) in zahtevane vrednosti atributov za dostop, kot na primer [28]:

- Identiteta povezovalnika (dostop imajo samo določeni povezovalniki (angl. Connectors)).
- Atributi povezovalnika (dostopajo lahko samo povezovalniki z določenimi atributi).
- Varnostni profili (dostopajo lahko samo povezovalniki, ki izpolnjujejo določene varnostne standarde).

Odločitev o nadzoru dostopa se sprejme v povezovalniku z uporabo tehnologij, kot sta XACML ali JAAS, odvisno od implementacije. Varnostna arhitektura IDS sicer ne predpisuje specifične tehnologije za izvajanje nadzora dostopa. Poleg nadzora dostopa pa varnostna arhitektura IDS spodbuja nadzor uporabe podatkov (angl. Usage Control), ki razširja nadzor dostopa z urejanjem tega, kaj se lahko stori s podatki po odobritvi dostopa. To vključuje določitev in uveljavljanje omejitev obdelave podatkov, pri čemer se ne obravnavajo zgolj določila o dostopu, temveč tudi kasnejše obveznosti. To je ključnega pomena za zaščito intelektualne lastnine, skladnost z zakonodajo in upravljanje digitalnih pravic. Nadzor uporabe podatkov v IDS deluje tako, da izmenjanim podatkom priloži informacije o politiki uporabe ter nenehno spremlja njihovo obdelavo in posredovanje. S tem se prepreči napačno ravnanje s podatki, na primer posredovanje osebnih podatkov nepooblaščenim končnim točkam. Služi kot orodje za preprečevanje kršitev varnosti in kot revizijski mehanizem za zagotavljanje skladne uporabe podatkov.

V okviru nadzora dostopa in uporabe sodelujejo različne vloge podatkovnega prostora, te so [28]:

- Posrednik metapodatkov: Upravlja povezovalnikov samo-opis, ki vsebuje tudi nadzor o uporabi.
- Povezovalnik: Izvaja nadzor uporabe. Povezovalci kot ponudniki podatkov morajo za podatke, zagotoviti tehnološko odvisne politike uporabe.

- Klirinška hiša: Spremlja uporabo podatkov in uveljavljanje omejitev uporabe prek sledenja izvora podatkov.
- Trgovina z aplikacijami: Zagotavljajo informacije o tem, ali aplikacije izvajajo tehnologijo za nadzor uporabe.

4.2. Zagotavljanje varnosti in skladnosti s predpisi

Zagotavljanje varnosti in skladnosti v podatkovnih prostorih je ključnega pomena za varovanje podatkov in izpolnjevanje zakonskih zahtev. Kot že omenjeno so lahko podatkovni prostori implementirani na različnih področjih, kot so veščine, zdravje, kmetijstvo, proizvodnja, energija, mobilnost, finance ter javna uprava. Vsako področje ima svoje specifične zahteve in zakonodajne okvirje, ki jih je treba upoštevati pri vzpostavljanju in upravljanju podatkovnih prostorov, da se zagotovi varnost, skladnost in učinkovito delovanje. Med tem ko so nekatere uredbe pomembne pri vseh področjih npr. GDPR za zagotavljanje varstva osebnih podatkov, pa so nekatere uredbe pomembne zgolj na določenih področjih.

Skladnost z zakonodajo v podatkovnih prostorih vključuje vrsto dejavnosti za zagotavljanje skladnosti z ustreznimi pravnimi okviri v celotnem življenjskem ciklu ekosistema za izmenjavo podatkov. To obsega vključitev regulativnih parametrov v zasnovo podatkovnega prostora, usmerjanje organov upravljanja in udeležencev glede implikacij vrednot in predpisov EU ter opredelitev vlog in odgovornosti. Skladnost zahteva tudi oblikovanje notranjih politik in stalno spremljanje upoštevanja predpisov.

5 Ključni izzivi pri implementaciji

Interoperabilnost podatkovnih prostorov je ključna za omogočanje učinkovitega in varnega izmenjevanja podatkov med različnimi organizacijami in sistemi. Interoperabilnost je mogoče doseči na različnih ravneh:

- **Pravna:** Zagotavljanje, da lahko sodelujejo organizacije, ki delujejo na osnovi različnih pravnih okvirov, politik in strategij.
- **Organizacijska:** Uskladitev procesov, komunikacijskih tokov in politik, ki različnim organizacijam omogočajo smiselno uporabo izmenjanih podatkov v njihovih procesih za doseganje skupno dogovorjenih in vzajemno koristnih ciljev.
- **Semantična:** Sposobnost različnih sistemov, da enotno razumejo podatke, ki se izmenjujejo.
- **Tehnična:** Sposobnost različnih sistemov, da komunicirajo in izmenjujejo podatke.

Okvirja, kot sta ISO/IEC 19941 in Evropski okvir interoperabilnosti (EIF), opredeljujejo te ravni ter zagotavljajo celovito povezovanje in usklajevanje podatkov [30].

Izmenjava podatkov med organizacijami pa zahteva interoperabilnost, kot tudi suverenost podatkov.

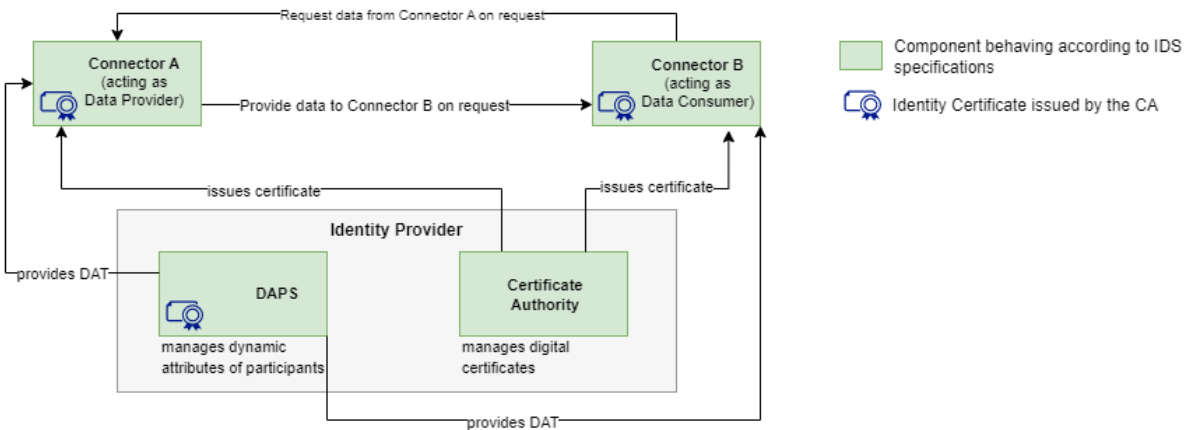
Suverenost podatkov, kot je opredeljena v podatkovnih prostorih, se nanaša na zmožnost posameznikov, organizacij in vlad, da imajo nadzor nad podatki, ki jih imajo, in uveljavljajo svoje pravice v zvezi s podatki, vključno z njihovim zbiranjem, shranjevanjem, izmenjavo in uporabo s strani drugih. Da bi zagotovili suverenost podatkov, morajo biti pravila upravljanja podatkov (politike uporabe podatkov, pogodbe o podatkih, pravice do uporabe itd.) jasno izražena in razumljiva za vse udeležence. IDSA teh pravil ne ustvarja, temveč zagotavlja okvir za opisovanje in uveljavljanje pravil, ki jih opredeli organ za upravljanje podatkovnega prostora, to so npr. upravljavci, vlade ali regulativni organi. Imetniki podatkov določijo pravila uporabe, podatkovni povezovalc IDSA pa z obdelavo teh pravil zagotavlja, da jih vsi udeleženci spoštujejo. [31]

Najosnovnejši sprejemljiv podatkovni prostor (angl. Minimum Viable Data Space - MVDS), lahko služi kot osnova za implementacijo v različnih primerih uporabe. MVDS je kombinacija komponent, ki omogočajo oblikovanje podatkovnega prostora z dovolj lastnostmi, da je uporaben za varno in neodvisno izmenjavo podatkov, kot določa

Mednarodno združenje za podatkovne prostore. MVDS mora vključevati vsaj dva povezovalnika, en mora predstavljati ponudnika podatkov drugi pa potrošnika podatkov (Slika 7). Opcijske komponente pa so vse tiste, ki omogočajo posredniške storitve, predstavljene v začetku [32].

MVDS lahko oblikujete tako, da (ponovno) uporabite in prilagodite odprtokodne komponente, navedene na Githubu IDS, ali pa začnete z razvojem nekaterih ali vseh komponent od začetka, tako da sledite specifikacijam komponent. V končni fazi obeh poti je priporočljiva uporaba testnega okolja IDS, da zagotovite združljivost in interoperabilnost komponent, ki jih boste uporabili v svojem MVDS [32].

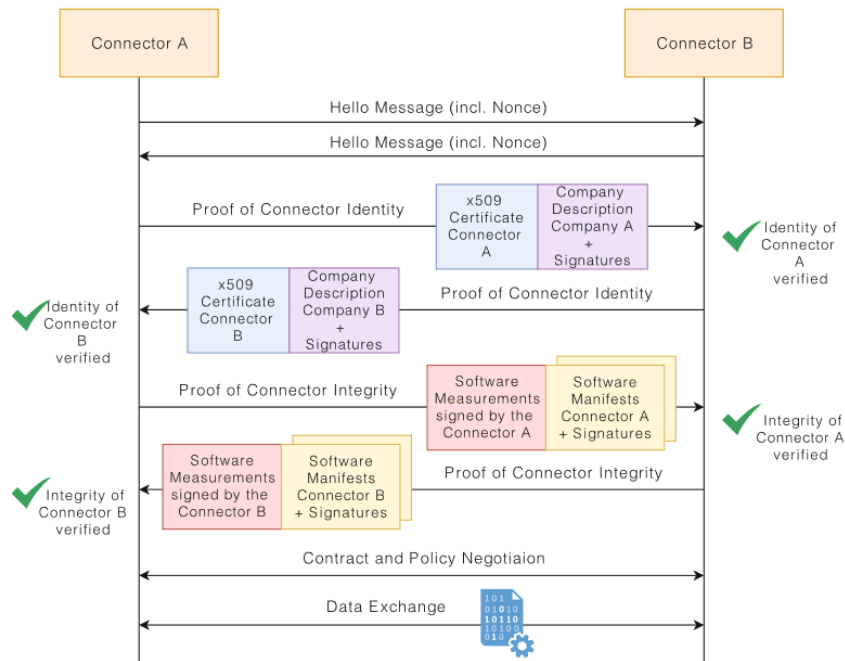
Vsi povezovalniki in njihovi upravljalci morajo zagotavljati zadostno raven varnosti (podatkov) in pravilno izvajati standarde IDS. Udeleženci podatkovnega prostora morajo zaupati, da drugi udeleženci, ki jih prej niso poznali, resnično izpolnjujejo te zahteve.



Slika 7: Najosnovnejši sprejemljiv podatkovni prostor.[32]

Slika 8 prikazuje poenostavljeno zaporedje vzpostavitve povezave in prve izmenjave podatkov. Koraki so naslednji [33]:

1. Pozdravno sporočilo: Pošlje se pozdravno sporočilo z naključno številko, da se zagotovi svežina in preprečijo napadi ponovitev.
2. Dokaz o identiteti povezovalca: Povezovalci se identificirajo s potrdili X.509 in opisi podjetij. Zasebni ključi vzpostavijo komunikacijski kanal, podpisi pa se preverijo pri organu za potrjevanje (angl. Organ za potrjevanje - CA). S tem se potrdita identiteta in raven certificiranja.
3. Dokaz integritete povezovalnika: Povezovalci preverijo celovitost s pošiljanjem podpisanih manifestov svoje baze zaupanja vrednega računalništva (Trusted Computing Base - TCB) in stanja sistema. Nato se preveri podpise, svežino in verigo artefaktov programske opreme ter preveri manifest pri organu za potrjevanje. S tem se potrdi zanesljivost in raven certificiranja TCB.
4. Pogajanja o pogodbah in politikah: Varna komunikacija partnerjem omogoča pogajanja o pogojih izmenjave podatkov in politikah nadzora uporabe.
5. Izmenjava podatkov: Podatki se lahko varno zamenjajo.

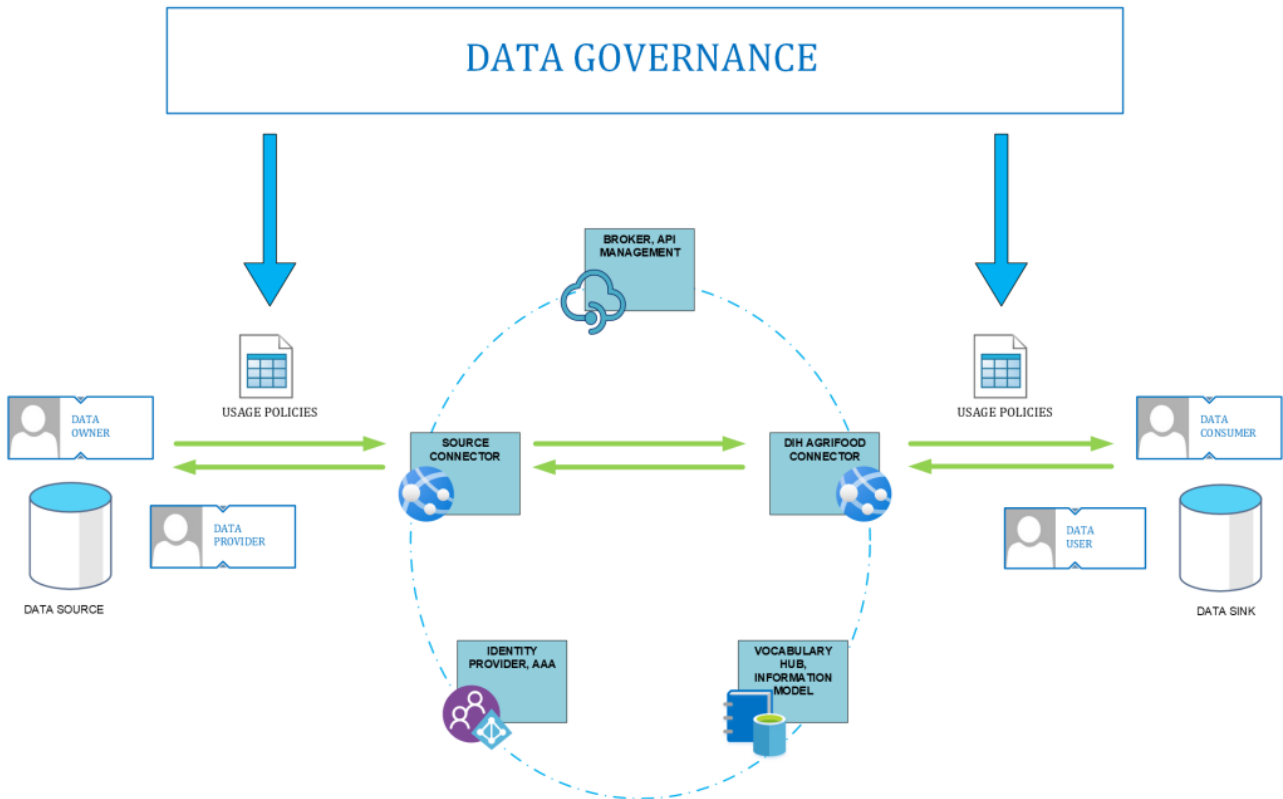


Slika 8: Komunikacija med povezovalniki. [33]

6 Primeri dobre prakse

Kljub temu, da je ideja podatkovnih prostorov na ravni EU šele že zaživela med ključnimi deležniki, ki se strinjajo s potrebo po novem pristopu upravljanja podatkov na bolj transparenten in varen način, raven zrelosti praktičnih rešitev za podatkovne prostore je šele v povojih. Občutno razliko se lahko opazi tudi med različnimi področji, ki je precej odvisna od ravni digitalizacije določenega področja. Kot primer je vzpostavitev podatkovnega prostora za agroživilski sektor zaradi nižje ravni digitalizacije bolj zahtevna kot za Industrijo 4.0, saj je potrebno vložiti precej več resursov v pripravo ustreznih podatkov in vključevanje ključnih deležnikov.

DIH Agrifood Data Space (v nadaljevanju DADS) predstavlja federativno platformo za deljenje podatkov, s čimer si prizadeva zagotoviti okoljsko, ekonomsko in družbeno trajnostnost. Za tehnično vzpostavitev in vzdrževanje infrastrukture DADS podatkovnega prostora skrbi Inovacijsko Tehnološki Grozd (angl. Innovation Technology Center, ITC) Murska Sobota, medtem ko za nadzor implementacije upravljaljskega modela skrbi digitalno inovacijsko stičišče za kmetijstvo in proizvodnjo hrane DIH AGRIFOOD sestavljeno iz več partnerjev.



Slika 9: Tehnična infrastruktura DADS podatkovnega prostora. [34]

Infrastruktura DADS pod. prostora omogoča katerem koli zainteresiranemu lastniku, ponudniku ali uporabniku podatkov v Sloveniji in širši regiji, da deli oz. uporablja podatke zbrane na podlagi aktivnosti deležnikov agroživilske verige (kmetje, predelovalci, dobavniki, HoReCa sektor in potrošniki). Na ta način DADS predstavlja temelj razvoja modela podatkovne ekonomije v slovenskem agroživilskem sektorju in pospešuje razvoj naprednih storitev na podlagi podatkov, ki prinašajo dodano vrednost za vse deležnike.

Vzpostavitev DADS pod. prostora vključuje vzpostavitev ključnih tehničnih in upravljaljskih gradnikov, ki sledi konceptom in metodologiji IDSA predstavljenimi v poglavju 2. Iz stališča tehničnih komponent, DADS pod. prostor vključuje naslednje gradnike:

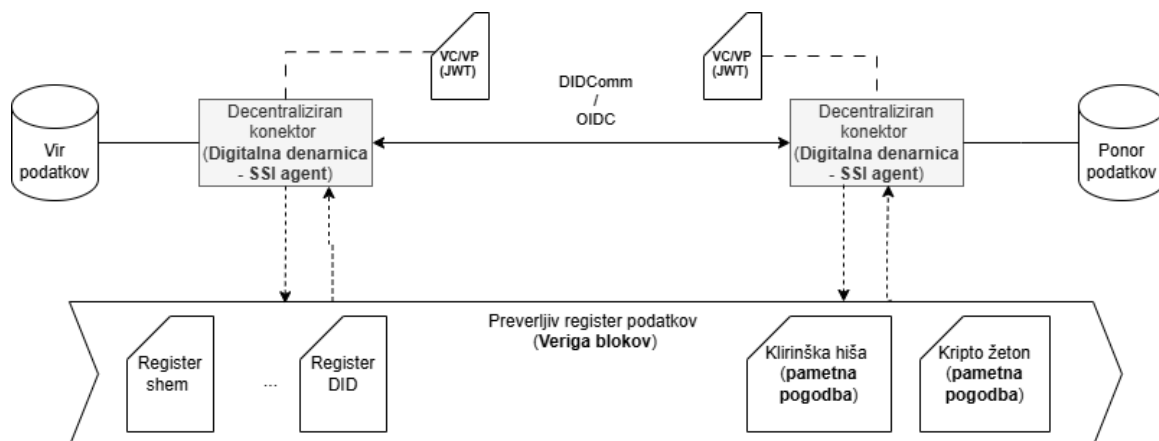
- DIH AGRIFOOD Identity Management – rešitev za upravljanje identitet po principu protokola OpenID Connect;
- Consent Management Platform – spletna platforma za lastnike in uporabnike podatkov, ki omogoča upravljanje dovoljenj in zahtev za deljenje podatkov;
- API Market Service – storitev za upravljanje programskih vmesnikov, ki omogoča pregled dostopnih podatkovnih množic in vmesnikov za njihovo deljenje;
- DIH AGRIFOOD IDS Connector – povezovalnik kompatibilen z IDSA protokolom, ki omogoča vključevanje in deljenje podatkov v podatkovnem prostoru;
- Posrednik (v razvoju) – zasebna vrsta IDS povezovalnika za registracijo in vzdrževanje samoopisov (angl. Self-Description) IDS povezovalnikov v podatkovnem prostoru;
- Posredniška hiša (v razvoju) – storitev za sledenje transakcij deljenja podatkov v pod. prostoru;
- Vozlišče slovarjev (v razvoju) – spletna platforma za deljenje in vzdrževanje semantičnih in podatkovnih modelov za podatkovne množice dostopne v podatkovnem prostoru.

Upravljaljski model DADS pod. prostora je skladen z relevantnimi smernicami in regulativami Evropske komisije, kot so Zakon o upravljanju podatkov, Zakon o podatkih (angl. Data Act) ter Evropska strategija za podatke (angl. EU Strategy for Data). Ker še vedno ne obstaja celovit upravljaljski model za podatkovne prostore (ali specifično

za agroživilski podatkovni prostor) je DADS upravljavski model razvit kot razširitev Pravilnika za FAIR podatkovno ekonomijo, oblikovanega s strani organizacije Sitra [35], hkrati pa upošteva predloge organizacij IDSA, OpenDEI in BDVA. Upravljavski model vsebuje definicijo vlog Ustanovnih članov (angl. Foundation Member), Upravnega odbora (angl. Steering Committee) ter ostalih članov, ki prispevajo k upravljanju DADS pod. prostora ter procedur za vključitev v pod. prostor, pravila za deljenje podatkov in jasne predloge za pogoje uporabe podatkov (angl. Dataset Terms of Use), ki zagotavljajo transparentnost in upoštevanje organizacijskega in pravnega okvirja pri deljenju podatkov.

7 Inovativni pristopi in tehnologije

Zaradi temeljne ideje podatkovnih prostorov, da se pospeši nadzirano deljenje podatkov, pri čemer pa lokacija podatkov ni nujno centralizirana, se kot potencialno zanimiv skupek tehnologij pojavljajo tehnologije decentralizacije. Web3 tehnologije, ki temeljijo na decentraliziranih principih, tehnologiji veriženja blokov itn., ponujajo napredne rešitve za vzpostavitev varnih, preglednih in zaupanja vrednih podatkovnih prostorov. To poglavje predstavlja, kako lahko različne Web3 komponente, kot so decentralizirane identitete (DID), pametne pogodbe, verige blokov, preverljive poverilnice in preverljiva dokazila, pripomorejo k doseganju teh ciljev (Slika). V nadaljevanju predstavimo tehnološke koncepte, ki se lahko uporabijo za doseganje MVDS, ki je skoraj v celoti decentraliziran.



Slika 10: Primer arhitekture podatkovnih prostor, temelječe na decentraliziranih tehnologijah.

Decentralizirane identitete (DID) predstavljajo temelj za vzpostavitev zaupanja vrednega okolja v podatkovnih prostorih. DID omogočajo posameznikom in organizacijam, da imajo popoln nadzor nad svojo digitalno identiteto brez potrebe po osrednji avtoriteti ali centralizirani infrastrukturi javnih ključev. S pomočjo DID lahko deležniki varno in enostavno vzpostavijo identiteto ter komunicirajo z drugimi deležniki znotraj podatkovnega prostora na osnovi protokolov komunikacije, kot so DIDComm. Ključne prednosti uporabe DID vključujejo:

- Samostojna identiteta: Deležniki lahko sami upravljajo svoje identitete, kar zmanjša tveganje zlorab in poveča zasebnost.
- Interoperabilnost: DID standardi omogočajo interoperabilnost med različnimi sistemi in platformami, kar olajša sodelovanje med različnimi deležniki.
- Zaupanje in varnost: DID uporabljajo kriptografske tehnike za zagotavljanje avtentičnosti in integritete identitet, kar povečuje zaupanje med deležniki.

Pametne pogodbe so samodejno izvajajoči se programi, ki tečejo na verigi blokov in omogočajo avtomatizacijo in varno izvajanje pogodb. V kontekstu podatkovnih prostorov lahko pametne pogodbe služijo kot klirinška hiša,

ki zagotavlja zanesljivo in transparentno izvajanje transakcij med deležniki. Uporaba pametnih pogodb prinaša naslednje prednosti:

- Avtomatizacija procesov: Pametne pogodbe avtomatizirajo izvajanje pogodb, kar zmanjšuje potrebo po posrednikih in povečuje učinkovitost.
- Transparentnost: Vse transakcije, izvedene s pametnimi pogodbami, so zapisane na verigi blokov, kar omogoča preglednost in sledljivost.
- Varnost in zanesljivost: Pametne pogodbe so odporne na manipulacije in zagotavljajo varno izvajanje pogodb, kar zmanjšuje tveganje napak in goljufig.
- Upravljanje dostopa: Lastniki podatkov lahko s pomočjo pogodb definirajo pogoje, pod katerimi dovolijo dostop do svojih podatkov.

Kripto žetoni, podprti s pametnimi pogodbami, igrajo pomembno vlogo v podatkovni ekonomiji znotraj podatkovnih prostorov. Ti žetoni omogočajo ustvarjanje ekonomskih spodbud za izmenjavo podatkov, olajšujejo transakcije in zagotavljajo pravično nagrajevanje za prispevek k ekosistemu. Ključne funkcije kripto žetonov v podatkovnih prostorih vključujejo:

- Spodbujanje izmenjave podatkov: Kripto žetoni lahko služijo kot nagrada za deležnike, ki delijo svoje podatke, s čimer spodbujajo prostovoljno izmenjavo podatkov in povečujejo obseg dostopnih informacij.
- Olajšanje transakcij: S pomočjo pametnih pogodb se lahko kripto žetoni uporabljajo za samodejno in varno izvajanje transakcij med deležniki, kar zmanjšuje potrebo po posrednikih in zmanjšuje transakcijske stroške.
- Nagrajevanje prispevkov: Deležniki, ki prispevajo k razvoju in vzdrževanju podatkovnega prostora, lahko prejmejo žetone kot nagrado za svoje delo, kar spodbuja aktivno sodelovanje in prispevanje k skupnosti.
- Zagotavljanje pravic in dostopa: Kripto žetoni se lahko uporabljajo za upravljanje dostopa do podatkov in storitev znotraj podatkovnega prostora. Deležniki lahko s pomočjo žetonov pridobijo dostop do specifičnih podatkov ali storitev, kar omogoča fleksibilno in pravično upravljanje pravic.
- Podpora decentraliziranim avtonomnim organizacijam (DAO): Kripto žetoni omogočajo vzpostavitev DAO, ki lahko upravljajo podatkovne prostore na decentraliziran način. Člani DAO lahko glasujejo o pomembnih odločitvah in upravljajo sredstva ekosistema s pomočjo žetonov, kar zagotavlja transparentnost in participacijo.

Preverljiv register podatkov (ang. Verifiable Data Registry - VDR) in **zaupanja vredni registri** (angl. trusted registries) sta ključni komponenti za upravljanje in hranjenje podatkovnih slovarjev v podatkovnih prostorih. VDR, ki predstavlja verigo blokov, ki ponuja različne zaupanja vredne register, omogoča beleženje in iskanje podatkovnih nizov ter zagotavlja zanesljivo in preverljivo shranjevanje podatkovnih shem in drugih ključnih seznamov. Te komponente prispevajo k:

- Organizaciji podatkov: Registri zagotavljajo centralizirano mesto za shranjevanje in upravljanje podatkovnih nizov in shem, kar olajša dostopnost in uporabo podatkov.
- Verifikacija podatkov: zaupanja vredni registry omogočajo preverjanje veljavnosti in celovitost, kar zagotavlja zanesljivost podatkov.
- Interoperabilnost: Standardizacija podatkovnih shem omogoča lažjo integracijo in interoperabilnost med različnimi sistemi in platformami.

Preverljive poverilnice (angl. Verifiable credentials - VC) in **preverljiva dokazila** (angl. Verifiable presentations - VP) predstavljajo napreden način za zagotavljanje nadzora dostopa do podatkov v podatkovnih prostorih. Preverljive poverilnice omogočajo deležnikom, da dokažejo svojo identiteto in pravice do dostopa do določenih podatkov, medtem ko preverljiva dokazila, ki se lahko tvorijo tudi v oblikah, kot je JWT, zagotavljajo dokazovanje

avtentičnosti in celovitosti podatkov ter nudijo ključ dostopa do podatkov. Ključne prednosti teh tehnologij vključujejo:

- Zaupanje in varnost: Preverljive poverilnice in dokazila temeljijo na kriptografskih tehnikah, ki zagotavljajo visoko stopnjo varnosti in zaupanja.
- Enostavnost uporabe: Deležniki lahko enostavno izdajo, preverijo in upravljajo poverilnice in dokazila, kar olajša nadzor dostopa do podatkov.
- Skladnost z regulativami: Te tehnologije omogočajo izpolnjevanje različnih regulativnih zahtev, kar zagotavlja skladnost in zanesljivost podatkovnih prostorov.

V ekosistemu podatkovnih prostorov igrajo **digitalne denarnice** in **SSI** (angl. Self-Sovereign Identity) **agenti** ključno vlogo pri upravljanju decentraliziranih identitet (DID) in preverljivih poverilnic (VC). Digitalne denarnice in SSI agenti služijo kot osrednja orodja za upravljanje identitet in poverilnic, ter lahko delujejo kot IDS (International Data Spaces) konektorji, ki omogočajo varno in učinkovito izmenjavo podatkov med različnimi deležniki.

Inovativni pristop k podatkovnim prostorom, ki temelji na Web3 tehnologijah, ponuja številne prednosti, vključno z večjo varnostjo, transparentnostjo, učinkovitostjo in decentralizacijo. Vendar pa je treba premagati tudi določene izzive, kot so regulativna skladnost, tehnična kompleksnost in uporabniška sprejemljivost. Kljub tem izzivom, potencial Web3 tehnologij za transformacijo podatkovnih prostorov predstavlja pomemben korak k prihodnosti digitalnega poslovanja, kjer bo sodelovanje in izmenjava podatkov med deležniki postala bolj varna, učinkovita in zanesljiva.

8 Zaključek

Podatkovni prostori predstavljajo odprto in povezano infrastrukturo za varno izmenjavo podatkov v skladu s skupnimi pravili, standardi in politikami. EK je februarja 2020 predstavila Evropsko strategijo za podatke z jasnim ciljem ustvariti enoten trg za podatke, ki bi omogočil polno izkoriščanje vrednosti podatkov v korist evropske družbe in gospodarstva. Strategija določa smernice za oblikovanje skupnih evropskih podatkovnih prostorov na več ključnih področjih. Kljub velikemu potencialu podatkovnih prostorov se soočamo z mnogimi izzivi. Ključni med njimi so zagotavljanje interoperabilnosti na različnih ravneh ter osredotočenost na dostopnost in suverenost podatkov, kar lahko ovira implementacijo minimalno funkcionalnih podatkovnih prostorov.

V prispevku smo osvetlili pomen in kompleksnost arhitektur podatkovnih prostorov ter pregledali trenutno stanje pristopov k vzpostavitvi podatkovnih prostorov. Osredotočili smo se na identifikacijo in analizo izzivov v specifičnih domenah in glavnih deležnikov podatkovnih prostorov, s poudarkom na IT arhitekturnih predstavitev. Predstavili smo tudi inovativne pristope, ki omogočajo identifikacijo podatkov in upravljanje dostopa do njih s pomočjo decentraliziranih tehnologij.

V naslednjih mesecih se bodo raziskave morale osredotočiti na razvoj standardov in protokolov, ki omogočajo interoperabilnost med različnimi podatkovnimi prostori in platformami. To vključuje semantično, organizacijsko, pravno in tehnično interoperabilnost. Prav tako bi se nadaljnje raziskave morale osredotočiti na razvoj metod za učinkovito upravljanje podatkovnih prostorov, vključno z avtomatizacijo procesov, nadzorom dostopa in upravljanjem identitet. Priporočamo izvedbo pilotnih projektov in študij primerov v različnih sektorjih, kot so zdravstvo, kmetijstvo, proizvodnja, energija, mobilnost in finance. To bo omogočilo vrednotenje in izboljšanje podatkovnih prostorov v realnih okoljih ter identifikacijo specifičnih izzivov in priložnosti.

Podatkovni prostori imajo potencial, da postanejo ključna platforma za izmenjavo in uporabo podatkov v prihodnosti. Z nadaljnjimi raziskavami, razvojem in sodelovanjem med deležniki lahko premagamo obstoječe izzive in izkoristimo vse priložnosti, ki jih podatkovni prostori ponujajo. S tem bomo prispevali k ustvarjanju digitalno preoblikovane družbe, ki temelji na podatkih, in spodbudili inovacije, gospodarsko rast in izboljšanje kakovosti življenja za vse.

Literatura

- [1] „Building a data economy — Brochure | Shaping Europe’s digital future“. Pridobljeno: 13. marec 2024. [Na spletu]. Dostopno na: <https://digital-strategy.ec.europa.eu/en/library/building-data-economy-brochure>
- [2] J. Conde, A. Pozo, A. Munoz-Arcentales, J. Choque, in Á. Alonso, „Fostering the integration of European Open Data into Data Spaces through High-Quality Metadata“, feb. 2024, [Na spletu]. Dostopno na: <http://arxiv.org/abs/2402.06693>
- [3] „Building data spaces for sustainability use cases - AWS Prescriptive Guidance“. Pridobljeno: 13. marec 2024. [Na spletu]. Dostopno na: <https://docs.aws.amazon.com/prescriptive-guidance/latest/strategy-building-data-spaces/introduction.html>
- [4] „Common European Data Spaces | Shaping Europe’s digital future“. Pridobljeno: 11. april 2024. [Na spletu]. Dostopno na: <https://digital-strategy.ec.europa.eu/en/policies/data-spaces>
- [5] E. Curry, „Real-time Linked Dataspaces Enabling Data Ecosystems for Intelligent Systems“.
- [6] Z. Boukhers, C. Lange, in O. Beyan, „Enhancing Data Space Semantic Interoperability through Machine Learning: a Visionary Perspective“, *ACM Web Conference 2023 - Companion of the World Wide Web Conference, WWW 2023*, str. 1462–1467, apr. 2023, doi: 10.1145/3543873.3587658.
- [7] „What are the main elements of a data space? | datos.gob.es“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://datos.gob.es/en/blog/what-are-main-elements-data-space>
- [8] „A European strategy for data | Shaping Europe’s digital future“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://digital-strategy.ec.europa.eu/en/policies/strategy-data>
- [9] „Data protection - European Commission“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: https://commission.europa.eu/law/law-topic/data-protection_en
- [10] „Data Governance Act explained | Shaping Europe’s digital future“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://digital-strategy.ec.europa.eu/en/policies/data-governance-act-explained>
- [11] „A European strategy for data | Shaping Europe’s digital future“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://digital-strategy.ec.europa.eu/en/policies/strategy-data>
- [12] J. Conde, A. Pozo, A. Munoz-Arcentales, J. Choque and, in A. Alonso, „Fostering the integration of European Open Data into Data Spaces through High-Quality Metadata“, Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://opendata.paris.fr>
- [13] „Technical Convergence Discussion Document“.
- [14] V. Takanen, P. J. Laszkowicz Advisor Futurice Teemu Toivonen, in A. Poikola, „THE TECHNOLOGY LANDSCAPE OF DATA SPACES Antti Poikola Leading Specialist Sitra The Technology Landscape of Data Spaces THE TECHNOLOGY LANDSCAPE OF DATA SPACES“, 2023, Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: www.sitra.fi
- [15] „Data Spaces Radar“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://www.dataspaces-radar.org/radar/>
- [16] „The Data Spaces Radar Data Spaces Support Centre Edition“.
- [17] „The EHDS: concerns and opportunities following ongoing trilogues“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://www.eucope.org/the-european-health-data-space-challenges-and-opportunities-with-respect-to-the-different-proposals-currently-subject-to-the-trilogue-negotiations/>
- [18] „Implementing the European Health Data Space Across Europe“.
- [19] M. Šestak in D. Copot, „Towards Trusted Data Sharing and Exchange in Agro-Food Supply Chains: Design Principles for Agricultural Data Spaces“, *Sustainability 2023, Vol. 15, Page 13746*, let. 15, št. 18, str. 13746, sep. 2023, doi: 10.3390/SU151813746.
- [20] E. Curry *idr.*, „Data Sharing Spaces: The BDVA Perspective“, v *Designing Data Spaces*, Cham: Springer International Publishing, 2022, str. 365–382. doi: 10.1007/978-3-030-93975-5_22.
- [21] M. Gouriet *idr.*, „The Energy Data Space: The Path to a European Approach for Energy“, v *Designing Data Spaces*, Cham: Springer International Publishing, 2022, str. 535–575. doi: 10.1007/978-3-030-93975-5_33.
- [22] C. Schlueter Langdon in K. Schweichhart, „Data Spaces: First Applications in Mobility and Industry“, v *Designing Data Spaces*, Cham: Springer International Publishing, 2022, str. 493–511. doi: 10.1007/978-3-030-93975-5_30.
- [23] „EU financial data space and cloud infrastructure“.

- [24] „1.2 Purpose and Structure of the Reference Architecture | IDS Knowledge Base“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-ram-4/introduction/1_1_goals_of_the_international_data_spaces/1_2_purpose_and_structure_of_the_document
- [25] „3.5 System Layer | IDS Knowledge Base“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-ram-4/layers-of-the-reference-architecture-model/3-layers-of-the-reference-architecture-model/3_5_0_system_layer
- [26] „Introduction | Dataspace Connector“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://international-data-spaces-association.github.io/DataspaceConnector/Introduction>
- [27] Y. Liu, D. Niu, S. Geng, J. Sun, in H. Zhang, „Application of Data Integration in Dataspace in Multi-value Chain Collaboration of Electric Power Manufacturing Industry“, *2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2022*, str. 292–298, 2022, doi: 10.1109/CSCWD54268.2022.9776190.
- [28] „4.1.6 Usage Control | IDS Knowledge Base“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-ram-4/perspectives-of-the-reference-architecture-model/4_perspectives/4_1_security_perspective/4_1_6_usage_control
- [29] „From Data Silos to Data Ecosystems: The role of Data Spaces Support Centre in addressing the legal challenges of the future single market for data - CiTiP blog“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://www.law.kuleuven.be/citip/blog/from-data-silos-to-data-ecosystems-the-role-of-data-spaces-support-centre-in-addressing-the-legal-challenges-of-the-future-single-market-for-data/>
- [30] „8 Interoperability - Blueprint v1.0 - Data Spaces Support Centre“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://dssc.eu/space/BVE/361660417/8+Interoperability>
- [31] „The significance of data spaces“.
- [32] „What is a Minimum Viable Data Space? | IDS Knowledge Base“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-reference-testbed/minimum-viable-data-space/mvds>
- [33] M. Huber, S. Wessel, G. Brost, in N. Menz, „Building Trust in Data Spaces“, v *Designing Data Spaces*, Cham: Springer International Publishing, 2022, str. 147–164. doi: 10.1007/978-3-030-93975-5_9.
- [34] „DIH AGRIFOOD Data Space – Technical Infrastructure“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://dataspace.dih-agrifood.com/technical-infrastructure>
- [35] „Rulebook for a fair data economy“. Pridobljeno: 31. julij 2024. [Na spletu]. Dostopno na: <https://www.sitra.fi/en/publications/rulebook-for-a-fair-data-economy/>