DSX Engine: Decentraliziran povezovalnik podatkovnih prostorov

Martin Ferenec, Teo Lah, Muhamed Turkanović

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko,
Maribor, Slovenija
martin.ferenec@student.um.si, teo.lah@student.um.si, muhamed.turkanovic@um.si

Članek obravnava izzive suverene in varne med-organizacijske izmenjave podatkov, ki jih rešujejo podatkovni prostori. Temelji na evropski strategiji za podatke in standardu International Data Spaces Association, ki predvideva uporabo povezovalnikov kot ključnih tehničnih gradnikov. DSX Engine nadgrajuje referenčno implementacijo Eclipse Dataspace Components z mehanizmi za decentralizirano zaupanje: decentralizirane identitete, pametne pogodbe in podporo trgovanja podatkov s pomočjo žetonov ERC-20. S tem odpravlja osrednje točke odpovedi, avtomatizira sklenitev podatkovnih pogodb ter omogoča sledljiv, kriptografsko preverljiv ekosistem za deljenje in trgovanje podatkov. Funkcionalnost DSX Engine je potrjena v pilotnem podatkovnem prostoru DIH AGRIFOOD Data Space, ki povezuje kmetijske, okolijske in logistične podatke različnih deležnikov, hkrati pa ohranja njihovo suverenost. Članek prispeva metodološke in tehnične smernice za nadaljnji razvoj podatkovnih prostorov v industriji in javnem sektorju.

Ključne besede:

podatkovni prostori,
povezovalniki podatkovnih prostorov,
Eclipse Dataspace Components,
DSX Engine,
decentralizacija,
Web 3.0,
tehnologije veriženja blokov,
pametne pogodbe,
podatkovna suverenost,
varna, zaupanja vredna izmenjava podatkov.

Univerzitetna založba Univerze v Mariboru

ISBN 978-961-299-037-4
DOI https://doi.org/10.18690/um.feri.7.2025.11

1 Uvod

V zadnjem desetletju smo bili priča premiku od paradigme masovnih podatkov k paradigmi zaupanja vrednih podatkov. Medtem ko so se prejšnje rešitve osredotočale predvsem na količino in hitrost obdelave podatkov, današnje pobude vse bolj poudarjajo kakovost, integriteto, sledljivost in predvsem zaupanje v podatke in njihove vire. Ta sprememba je tesno povezana z razvojem decentraliziranih tehnologij, ki odpravljajo odvisnost od enega ponudnika infrastrukture ter uvajajo preverljive, kriptografsko zaščitene mehanizme za upravljanje dostopa in uporabe podatkov. V sodobnem podatkovno vodenem gospodarstvu postajajo podatkovni prostori (angl. data spaces) ključni za varno in zanesljivo deljenje podatkov med organizacijami. Evropska unija je v okviru svoje strategije za podatke prepoznala potrebo po oblikovanju *skupnih evropskih podatkovnih prostorov* v različnih sektorjih, ki bodo omogočali nadzorovano izmenjavo podatkov ob zagotavljanju podatkovne suverenosti deležnikov [1]. Klasični pristopi k deljenju podatkov, kot so centralizirana skladišča ali ad-hoc izmenjava pogosto ne zagotavljajo zadostnega nadzora nad tem, kdo in pod kakšnimi pogoji dostopa do podatkov. Prav tako je velikokrat nerazrešeno finančno vprašanje oz. finančna kompenzacije. Podatkovni prostori rešujejo te izzive z uvedbo federativnega modela – podatki ostajajo pri viru, deljenje pa poteka preko standardiziranih vmesnikov in pogodb, kjer imajo lastniki podatkov ves čas nadzor nad uporabo svojih podatkov [2]. V takem okolju osrednjo tehnološko vlogo igrajo povezovalniki podatkovnih prostorov, ki predstavljajo prehode (ang. gateway) med različnimi deležniki ter omogočajo povezovanje različnih informacijskih sistemov v podatkovni prostor ob uveljavljanju pravil uporabe in varnosti [3]. V nadaljevanju članka najprej predstavimo koncept podatkovnih prostorov ter okvir, ki ga je razvilo Mednarodno združenje za podatkovne prostore (IDSA), nato pa se osredotočimo na **Eclipse Dataspace Components (EDC)** kot referenčno implementacijo povezovalnika podatkovnih prostorov. Osrednji del prispevka je posvečen naši rešitvi **DSX Engine**, ki EDC nadgrajuje z vrsto decentralizacijskih mehanizmov in novimi funkcionalnostmi. Članek zaključujemo s predstavitvijo **pilotne implementacije v podatkovnem prostoru DIH AGRIFOOD Data Space (DADS)** in razpravo o izzivih, pridobljenih izkušnjah ter možnih prihodnjih nadgradnjah.

2 Podatkovni prostori

2.1. Potreba po podatkovnih prostorih in pobude Evropske komisije

Obseg podatkov, ki jih ustvarjajo podjetja in druge organizacije, eksponentno narašča, s tem pa tudi potencial za ustvarjanje dodane vrednosti z njihovo delitvijo in analiziranjem. V praksi izmenjavo podatkov onemogočajo številne ovire: pravna in varnostna tveganja, pomanjkanje zaupanja, tehnološka nezdružljivost ter strah pred izgubo nadzora nad podatki. Podatkovni prostori so se pojavili kot koncept, ki naj bi naslovil te izzive in omogočil zaupanja vredno med-organizacijsko izmenjavo podatkov ob ohranjanju suverenosti vseh vpletenih deležnikov [3]. Ključna ideja je, da udeleženci sami nadzorujejo svoje podatke in pogoje deljenja: vsak udeleženec odloča, s kom bo podatke delil in pod katerimi vnaprej določenimi pogoji uporabe, identiteta in pravice vsakega udeleženca pa so nadzorovane [3]. Tako podatkovni prostori zagotavljajo, da se podatki lahko uporabijo za nove storitve in inovacije, ne da bi lastniki izgubili nadzor nad njihovo uporabo.

Evropska komisija je prepoznala strateški pomen takšnega pristopa in v evropski strategiji za podatke začrtala vzpostavitev skupnega evropskega podatkovnega prostora, ki bo okrepil konkurenčnost in podatkovno suverenost Evrope [4]. V praksi to pomeni vzpostavitev več tematskih podatkovnih prostorov (npr. industrijski, zdravstveni, kmetijski, finančni, energetski itd.), kjer bodo deležniki posameznih sektorjev lahko učinkoviteje izmenjevali vsebinsko in domensko povezane podatke pod skupnimi pravili. Za podporo temu cilju je EU sprejela tudi regulatorne ukrepe, kot sta *Akt o upravljanju podatkov* (angl. Data Governance Act - DGA) in *Akt o podatkih* (angl. Data Act), ki vzpostavljata pravila za podatkovne posrednike, deljenje podatkov med podjetji in zagotavljanje zaupanja [4]. Te pobude poudarjajo vlogo nevtralnih posrednikov in standardov pri izmenjavi podatkov. V okviru

industrije sta nastala tudi pomembna projekta, International Data Spaces Association (IDSA) in Gaia-X, ki s tehničnimi standardi in referenčnimi arhitekturami podpirata vizijo decentraliziranih podatkovnih ekosistemov na evropski ravni.

2.2. Standardizacija podatkovnih prostorov – IDSA: sloji, komponente, procesi

Za uresničitev podatkovnih prostorov je ključno skupno razumevanje arhitekture in komponent. Mednarodno združenje za podatkovne prostore (IDSA) je že od leta 2015 vodilno pri definiranju takšne arhitekture. Rezultat tega je *Referenčni arhitekturni model* (angl. Reference Architecture Model – IDSA RAM), ki podaja več nivojski opis gradnikov podatkovnega prostora. Uvaja pet slojev arhitekture [5]:

Poslovni sloj

Opredeljuje poslovne vloge udeležencev v podatkovnem prostoru ter njihove medsebojne interakcije in poslovne modele. Te vloge so ponudnik podatkov, potrošnik podatkov, posrednik, skrbnik podatkovnega prostora itd. Ta sloj zagotavlja okvir, znotraj katerega lahko nastajajo nove digitalne storitve in poslovni modeli na osnovi podatkov.

Funkcionalni sloj

Določa potrebne funkcionalnosti in zahteve sistema za zagotovitev zaupanja, varnosti in suverenosti podatkov. Sem sodijo mehanizmi za upravljanje identitet in dostopa, certificiranje udeležencev, posredovanje metapodatkov (ang. broker), uveljavljanje politik uporabe podatkov ter zanesljivo komunikacijo (Slika 1) [6]. Opredeljene so tudi podatkovne aplikacije (Data Apps) za izvedbo obdelav nad podatki in koncept klirinške podatkov (ang. Clearing House), ki omogoča monetizacijo podatkov z vključenimi mehanizmi nadzora in obračunavanja. Naslavlja šest vidikov sistema: zaupanje, varnost in podatkovna suverenost, ekosistem podatkov, standardizirana interoperabilnost, podatkovne aplikacije ter podatkovne tržnice.

Slika 1: IDSA RAM funkcijski sloj. [5]

Informacijski sloj

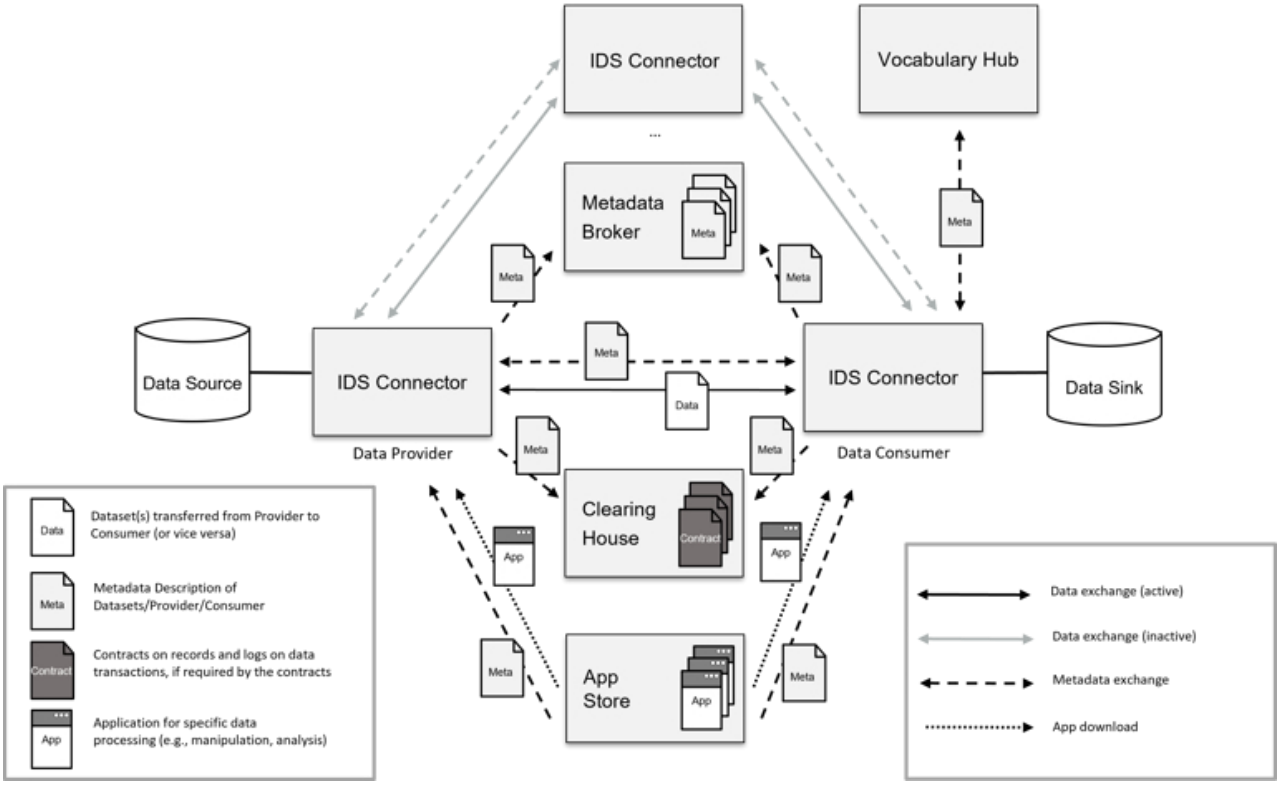
Ureja semantični model podatkov in standardizirane podatkovne formate. Da lahko različni deležniki učinkovito izmenjujejo podatke, morajo uporabljati skupne pojme in ontologije – informacijski sloj tako definira, kako so podatki opisani in predstavljeni, vključno z metapodatki (npr. opis ponudbe podatkov, pogojev uporabe itd.).

Procesni sloj

Podaja dinamični vidik delovanja podatkovnega prostora – opisuje osnovne procese in poteke dela, kot so priključitev novega udeleženca (angl. onboarding), objava ponudbe podatkov, pogajanje in sklenitev podatkovne pogodbe, izmenjava podatkov ter izključitev udeleženca [6]. Ta sloj opisuje zaporedje interakcij med komponentami, kot so povezovalnik, posrednik in ponudnik identitet, v posameznem scenariju znotraj podatkovnega prostora.

Sistemski sloj

Opisuje tehnično infrastrukturo in arhitekturo komponent (Slika 2). Določa, katere tehnične komponente sestavljajo podatkovni prostor (npr. **povezovalniki**, posredniki, ponudniki identitet, beleženje transakcij, hranjenje digitalnih potrdil) in kako te komponente med seboj komunicirajo, vključno s protokoli in vmesniki (API-ji). Ta sloj ponuja načrt za implementacijo podatkovnega prostora v obliki programske opreme in storitev [6].



Slika 2: IDSA RAM sistemski sloj – arhitektura. [5]

Čez vse sloje referenčnega modela se kot prečne teme obravnavajo še varnost, zasebnost, zaupanje in upravljanje (angl. governance) [6]. Poseben poudarek v International Data Spaces (IDS) je na certificiranju – ključne komponente, predvsem povezovalniki in udeleženci morajo pridobiti certifikate zaupanja, kar zagotavlja, da vsi v podatkovnem prostoru spoštujejo skupna pravila in varnostne zahteve [6]. IDS je tako vzpostavil celovit okvir, ki naslavlja tako organizacijske kot tehnične vidike podatkovnih prostorov.

Pomembno je, da vsi udeleženci spoštujejo enotne protokole in semantiko – le tako lahko dosežemo interoperabilnost med različnimi implementacijami podatkovnih prostorov. IDSA aktivno vzdržuje te specifikacije (trenutna različica referenčnega modela je 4.0) in jih usklajuje z drugimi pobudami (npr. Gaia-X), da se prepreči razdrobljenost standardov.

IDSA knjiga pravil

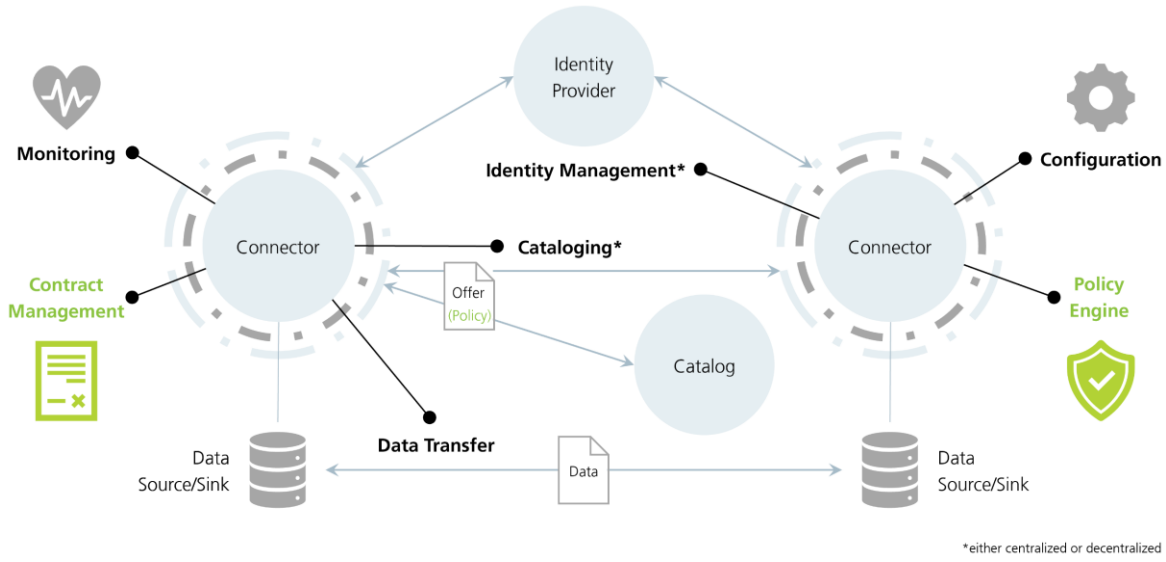
Predstavlja osrednji normativni dokument za standardizacijo v podatkovnih prostorih, saj določa enotna pravila za vključevanje udeležencev, certificiranje povezovalnikov in uveljavljanje politik uporabe. V njem so natančno opredeljene poslovne vloge (ponudnik, potrošnik, posrednik, upravljalec), minimalne varnostne zahteve, postopki revizije ter merila za izdajanje zaupanja vrednih poverilnic (angl. Verifiable Credentials - VC). Pravilnik (ang. rulebook) vzpostavlja tudi mehanizem reševanja sporov in sankcij v primeru neskladnosti, kar omogoča dosledno uveljavljanje suverenosti podatkov in medsebojne interoperabilnosti [7]. S tem služi kot obvezen referenčni okvir, ki ga morajo spoštovati vsi člani podatkovnega prostora IDS vključno z njihovimi tehničnimi implementacijami, kot so EDC ali DSX Engine.

3 Eclipse Dataspace Components (EDC)

Z namenom pospešitve razvoja podatkovnih prostorov je skupina industrijskih partnerjev (npr. v okviru pobude Catena-X v avtomobilski industriji) sprožila odprtokodni projekt pod okriljem Eclipse Foundation, imenovan Eclipse Dataspace Components (EDC). EDC je ena izmed referenčnih implementacij osnovnih komponent povezovalnika, združljivega s standardi IDS in prilagojenega evropskim pobudam za podatkovne prostore. Gre za ogrodje (angl. framework), ki organizacijam omogoča enostavnejšo vzpostavitev lastnega povezovalnika in vključitev v podatkovni prostor, kjer lahko varno delijo podatke z ostalimi udeleženci. Vsaka organizacija namesti svoj prilagojen EDC povezovalnik, ki deluje kot končna točka za deljenje podatkov in kot čuvaj, ki zagotavlja upoštevanje korporativnih politik in predpisov o suverenosti podatkov [1].

Prvi cilj projekta EDC je bil implementirati koncept IDS povezovalnika v praksi ter hkrati le tega utemeljiti na konceptih decentralizacije (Slika 3). Zgrajen je tako, da lahko komunicira z ostalimi osnovnimi storitvami v podatkovnem prostoru, kot so ponudniki identitet za preverjanje identitet udeležencev, posredniki (ang. Brokerji) za iskanje podatkovnih ponudb ali klirinška hiša (ang. Clearing House) za beleženje dogodkov, skladno z IDS referenčnim modelom [1]. S tem EDC lahko prevzame vlogo osnovne infrastrukture podatkovnega prostora. Projekt se je ustanovil leta 2021 in hitro pritegnil pozornost različnih domen – od industrije 4.0, pametne mobilnosti, do finančnega in kmetijskega sektorja – skratka povsod, kjer se pojavi potreba po zaupanja vredni izmenjavi podatkov med partnerji.

3.1. Arhitektura in komponente EDC



Slika 3: Arhitektura EDC. [8]

Arhitektura EDC sledi načelom modularnosti, razširljivosti in robustnosti. Sestavljena je iz več komponent oziroma modulov, od katerih vsak opravlja določeno funkcijo v procesu deljenja podatkov [1]. Ključni elementi arhitekture EDC so [8]: (1) krmilna ravnina (ang. control plane), (2) podatkovna ravnina (ang. data plane), (3) upravljalnik identitet (ang. identity hub), (4) izdajatelj poverilnic oziroma člen zaupanja, (5) mehanizem za upravljanje politik (ang. policy engine) in drugi moduli za upravljanje ter revizijo.

Krmilna ravnina je ključni del povezovalnika, ki skrbi za upravljanje in pogajanje glede dostopa do podatkov. Tu potekajo procesi, kot so registracija podatkovnih ponudb v katalog, iskanje podatkov, pogajanje med ponudnikom in potrošnikom ter sklenitev pogodbe o podatkih. Krmilna ravnina vsebuje ustrezne module za izvedbo poizvedb, usklajevanje izmenjave, preverjanje in uveljavljanje politik uporabe, spremljanje in revizijo transakcij. Pomembno je poudariti, da EDC uporablja asinhron način komunikacije med komponentami in ne uvaja nobene centralne baze podatkov, kar izboljšuje razpršenost in odpravlja eno točko odpovedi [1]. Preko nadzornika organizacije prav tako upravljajo in definirajo svoje kataloge podatkov, politike uporabe (angl. policy), definicije pogodb (angl. contract defnition) in dostopajo do lokalno predpomnjenega kataloga podatkovnega prostora, ki ga povezovalnik gradi s pomočjo pajka (ang. crawler) komponente, ki se nahaja v EDC Federated Catalog rešitvi.

Pajek je komponenta EDC, ki periodično razpošilja poizvedbe po katalogih vsem udeležencem podatkovnega prostora iz lokalnega seznama udeležencev. Nadzornik posameznega udeleženca ob zahtevi vrne prilagojen podatkovni katalog, ki vsebuje le ponudbe podatkov s katerimi politikami dostopa (angl. access policy) je povezovalnik, ki poizveduje po katalogu, skladen. Pajek nato shrani kataloge udeležencev v svoj lokalni podatkovni katalog, ki predstavlja katalog podatkovnega prostora. Glede na število udeležencev podatkovnega prostora, krmilna ravnina samodejno prilagodi, koliko pajkov je razposlanih sočasno, za namene pohitritve posodobitev lokalnega kataloga podatkov [8] [9].

Podatkovna ravnina je ločen del, ki skrbi za dejanski prenos podatkov od ponudnika do potrošnika. Ko je v krmilni ravnini sklenjena pogodba in dodeljena avtorizacija, podatkovna ravnina vzpostavi neposredno komunikacijo (npr. prenos datotek, pretok podatkov preko API) med povezovalnikoma. EDC privzeto zagotavlja revizijsko sled za prenos podatkov – vse transakcije se beležijo v dnevnik (angl. audit log), ki ga lahko preverita obe strani, vsaka pri svojem povezovalniku. Za povečanje varnosti EDC hrani občutljive informacije (npr. ključe, žetone za dostop) v varovanem trezorju (Hashicorp Vault) in uporablja začasne poverilnice, da minimizira možnosti zlorabe [1]. Vsak nadzornik lahko ima registriranih več podatkovnih posrednikov. S tem zagotovimo horizontalno skalabilnost, kar omogoča večji pretok podatkov.

Upravljalnik identitet je ločena komponenta EDC, odgovorna za shranjevanje, upravljanje in izmenjavo zaupanja vrednih poverilnic, upravljanje decentraliziranih identifikatorjev, certifikatov in ostalih z varnostjo ter identiteto povezanih zadev med udeleženci podatkovnega prostora. Izvaja naslednje funkcije [8]:

- Shranjuje in upravlja poverilnice, ki jih izdajajo zaupanja vredni izdajatelji poverilnic;
- Na podlagi poverilnic, generira preverljive predstavitve (angl. Verifiable Presentations - VP), ki omogočajo predstavitve poverilnic drugim udeležencem, pri čemer preverja in upravlja dostop na osnovi pravil. S tem pripomore pri ohranjanju popolnega nadzora nad vsebino, obsegom in kontekstom razkritih informacij;
- Upravlja DID organizacije;
- Upravlja s kriptografskimi pari ključev, pri čemer poskrbi tudi za njihovo rotacijo;

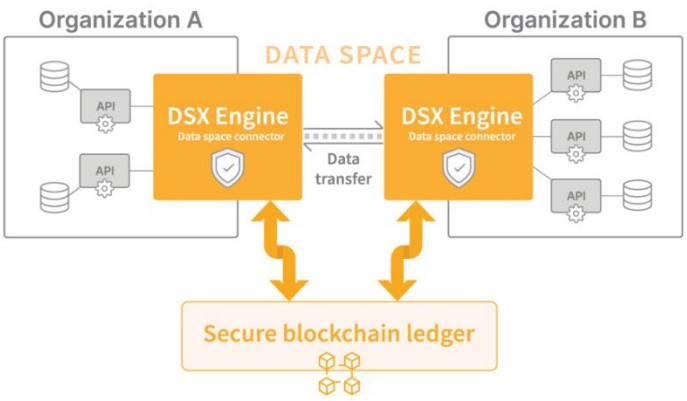
Deluje kot decentralizirana denarnica identitet, ki omogoča varno in nadzorovano izmenjavo verodostojnih trditev (angl. claim), s čimer podpira dinamično overjanje in avtorizacijo udeležencev v skladu z Eclipse Decentralized Claims Protocol (DCP) specifikacijo [8].

Izdajatelj poverilnic v arhitekturi EDC predstavlja člen zaupanja, katerega naloga je ustvarjanje, podpisovanje in posredovanje zaupanja vrednih poverilnic udeležencem podatkovnega prostora. Izdaja poverilnic je definirana v specifikaciji Decentralized Claims Protocol (DCP) kot asinhron proces, v katerem upravljalnik identitet povezovalnika pošlje zahtevo za poverilnico na vnaprej znan končni naslov izdajatelja, ki mu zaupa. Ta zahteva vključuje samopodpisan identitetni žeton (angl. identity token), ki vsebuje tudi dostopni žeton (angl. access token),

s katerim lahko izdajatelj po postopku odobritve dostavi podpisano poverilnico nazaj v upravljalnik identitet pošiljatelja [8] [10].
Postopek odobritve lahko vključuje preverjanje dodatnih verodostojnih trditev pošiljatelja. Na primer, izdajatelj lahko pred izdajo zahteva dokazilo, da prejemnik že razpolaga z drugo zanesljivo poverilnico kot je dokaz o članstvu v podatkovnem prostoru. Za identifikacijo udeležencev DCP uporablja decentralizirane identifikatorje (DID). EDC podpira DID z metodo did:web, ki omogoča, da organizacije določijo svojo identiteto prek HTTPS infrastrukture. DID je uporabljen pri podpisovanju poverilnic in verifikaciji ter overjanju izvora zahtevkov v podatkovnem prostoru [8] [10].
EDC ima vgrajen **mehanizem za izvrševanje politik uporabe podatkov** (ang. policy engine), ki omogoča, da ponudnik ob objavi podatkovne ponudbe določi pogoje (npr. kdo lahko dostopa, časovna omejitev dostopa, namen uporabe itd.). Te politike se prenesejo do povezovalnika potrošnika ob sklenitvi pogodbe. Povezovalnika na obeh straneh nato avtomatizirano uveljavljata ta pravila – npr. blokirata dostop, če pogoji niso izpolnjeni, ali omejita število dostopov. S tem je zagotovljeno spoštovanje načel rabe v celotni življenjski dobi podatkov, tudi po prenosu [1].
Kot del nadzornika, EDC ponuja **module za spremljanje poteka izmenjave in revizijsko sled**. Vsako dejanje – od zahtevka za podatke, pogodbe, do prenosa – je mogoče zabeležiti, kar omogoča poznejše preverjanje skladnosti, obračunavanje porabe podatkov in reševanje sporov. Centralizirane podatkovne shrambe v EDC ni. Namesto tega vsaka organizacija hrani svoje podatke na svoji infrastrukturi, EDC pa poskrbi, da se ti podatki na zanesljiv in varen način pretakajo do pooblaščenih potrošnikov.
Čeprav EDC omogoča modularno razširitev in interoperabilnost skladno z IDS specifikacijami, v svoji osnovni obliki **ni v celoti decentraliziran**. Tipične implementacije se pri registraciji udeležencev in beleženju dogodkov še vedno opirajo na osrednje komponente ali zaupanja vredne posrednike. To pomeni, da v primeru izpada ali kompromitacije teh centralnih točk lahko pride do motenj ali celo onemogočanja delovanja podatkovnega prostora.
Ena od prednosti EDC je njegova razširljivost – zasnovan je kot ogrodje, ki ga je mogoče prilagajati specifičnim potrebam. Ima podporo za vtičnike (plug-in točke preko SPI, angl. Service Provider Interface) za integracijo različnih sistemov za avtentikacijo in avtorizacijo, podporo različnim protokolom prenosa podatkov, povezovanje z različnimi katalogi in storitvami. Skupnost razvijalcev EDC je do danes dodala že več komponent, npr. modul za decentralizirane identitete (angl. decentralized identity) in integracijo z denarnicami za potrebe Gaia-X [11], ter načrtuje nadaljnje izboljšave v smeri skladnosti z Gaia-X specifikacijami (npr. infrastruktura za preverjanje udeležencev in storitev). EDC tako predstavlja temeljni gradnik, ki sicer omogoča minimalno delujoč podatkovni prostor, vendar na osnovi številnih statičnih spremenljivk in modulov, vendar nam kot ogrodje omogoča, da gradimo prilagojene rešitve povezovalnikov – ena takšnih nadgradenj osnovne arhitekture EDC je naš DSX Engine.

4 DSX Engine

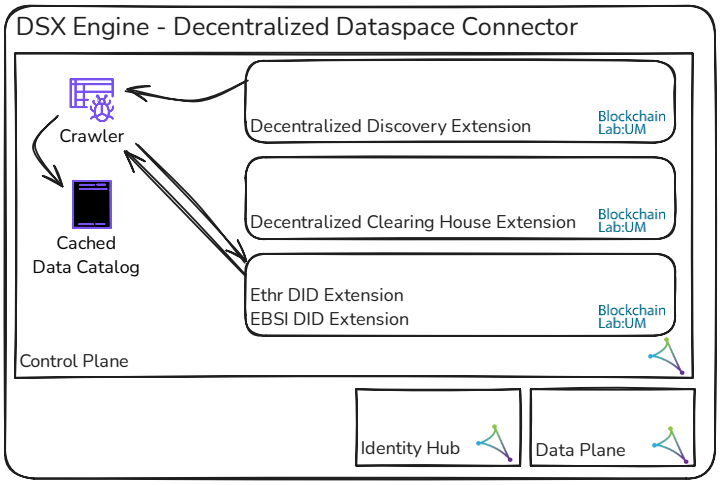
DSX Engine je nastal z namenom razširitve in nadgradnje obstoječih povezovalnikov podatkovnih prostorov, za potrebe še bolj decentraliziranega, zaupanja polnega in funkcionalno bogatega podatkovnega prostora. Osnovna arhitektura podatkovnih prostorov, kot jo opredeljuje IDSA in implementira EDC, predstavlja odlično izhodišče, vendar so se v praksi pokazale potrebe po dodatnih funkcionalnostih, zaradi katerih je bila razvita rešitev DSX Engine. Slika 4 prikazuje konceptualno osnovo povezovalnika DSX Engine.



Slika 4: Konceptualni prikaz podatkovnega prostora z DSX Engine povezovalniki.

4.1. Implementirane razširitve

Z namenom doseganja polno delujočega, dinamičnega in samostojnega povezovalnika, ki bo izpolnjeval dodatno zastavljene kriterije decentralizacije, smo obstoječi EDC nadgradili s tremi osrednjimi razširitvami in drugimi podpornimi funkcionalnostmi, ki jih v nadaljevanju podrobneje opišemo. Slika 5 prikazuje visokonivojsko arhitekturo DSX Engine, kjer so izpostavljeni osrednje implementirane razširitve.



Slika 5: Arhitektura DSX Engine.

Razširitev decentralizirane poizvedbe

Z namenom podpreti popolnoma decentraliziran način povezovanja med povezovalniki, smo razvili t. i. razširitev decentralizirane poizvedbe (ang. Decentralized Discovery Extension - DDE), ki omogoča dinamično registracijo in odkrivanje udeležencev (tj. drugih povezovalnikov) v podatkovnem prostoru, ki temelji na infrastrukturi verige blokov. DDE je implementiran kot EDC razširitev, ki vzpostavi neposreden most med EDC povezovalnikom in pametno pogodbo v okolju Ethereum. Pametna pogodba, ki smo jo implementirali in objavili na verigo blokov, deluje kot javni in preverljiv register udeležencev.

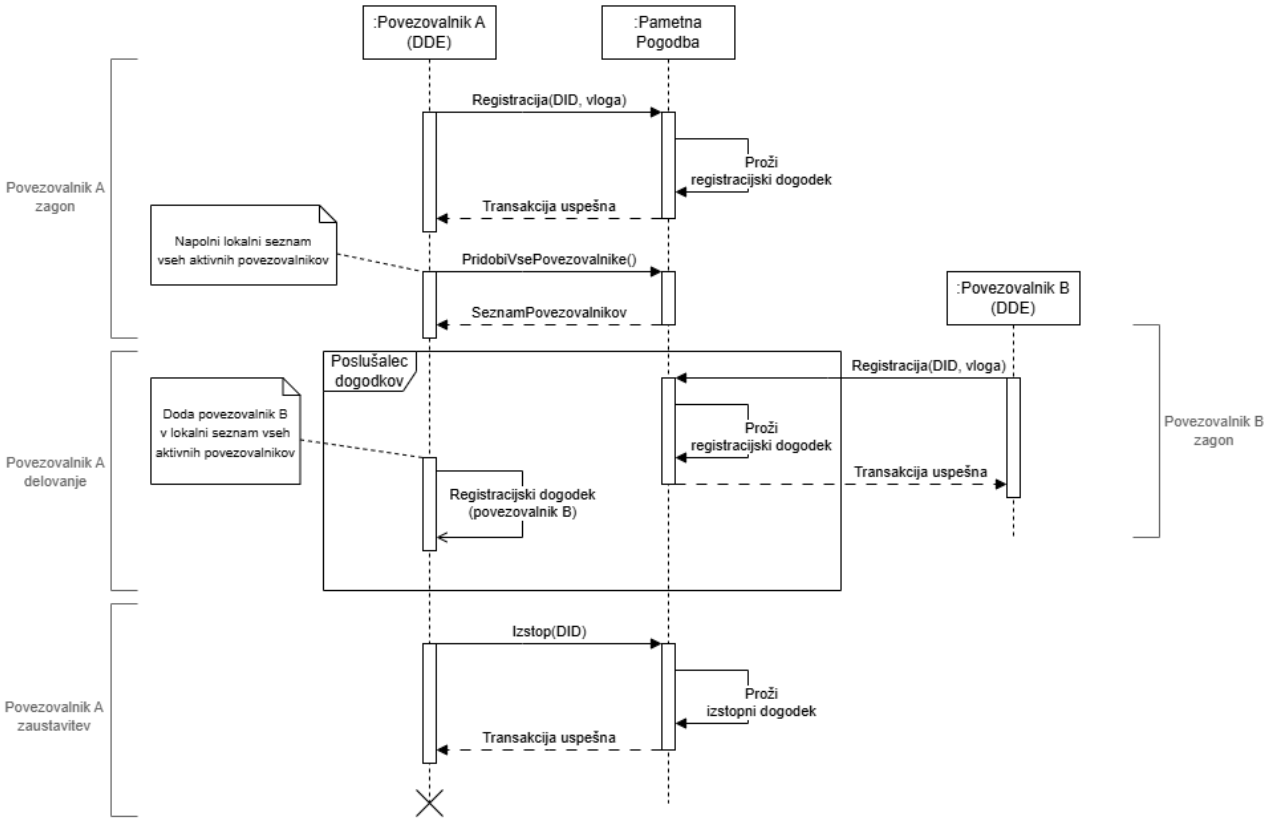
DDE v povezovalniku dinamično gradi lokalni seznam vseh aktivnih ponudnikov v podatkovnem prostoru. Ponudniki so povezovalniki (udeleženci), ki imajo vlogo nastavljeno na *ponudnik*. Organizacija pa ima možnost ročno dodati na seznam tudi udeleženca, ki ni ponudnik. EDC komponenta pajek izvede iskanje podatkovnih katalogov pri odkritih udeležencih (povezovalnikih) in jih shrani v lokalni predpomnjen seznam ter s tem omogoči poizvedovanje in procesiranje vsebine katalogov [8].

Povezovalnik se ob vzpostavitvi samodejno registrira v pametni pogodbi, pri čemer objavi svoj decentraliziran identifikator (DID) in vlogo ali je ponudnik ali potrošnik. Vloga potrošnik sporoči ostalim udeležencem, da ta povezovalnik ne nudi podatkov, kar pomeni, da ga ostali udeleženci ne bodo samodejno poizvedovali po katalogu. S tem se zmanjša število nepotrebnih klicev med udeleženci podatkovnega prostora. **Pametna pogodba ob registraciji povezovalnika sproži dogodek, ki ga ostali povezovalniki zaznajo**.

Po zagonu DDE pridobi vse udeležence in nato aktivira poslušalca dogodkov, ki posluša dogodke v pametni pogodbi (vstop ali izstop udeleženca). Tako samodejno periodično posodablja seznam ponudnikov. Vsak povezovalnik tako upravlja svojo lokalno, a usklajeno kopijo seznama vseh aktivnih ponudnikov, brez potrebe po centraliziranem imeniku. Proces je predstavljen na sliki 6.

DSX Engine vključuje izvorno podporo za izvajanje pametnih pogodb v **Ethereum** omrežjih. To pomeni, da lahko deluje tako v **javnih** omrežjih (npr. Ethereum mainnet, testneti) kot tudi v **zasebnih ali konzorcijskih** EVM-kompatibilnih omrežjih. Ta fleksibilnost omogoča prilagoditev različnim regulativnim in varnostnim zahtevam, obenem pa odpira možnost integracije z obstoječimi rešitvami in orodji verig blokov.

Vse decentralizirane razširitve DSX Engine lahko uporabljajo katerokoli Ethereum-kompatibilno infrastrukturo kot temelj za izvajanje pametnih pogodb, kar omogoča prenosljivost in ponovljivo uvajanje v različnih okoljih.



Slika 6: Diagram zaporedja razširitve DDE.

DDE uporablja decentralizirane identifikatorje (DID), kjer je identiteta povezovalnika zgrajena v skladu s specifikacijo W3C DID [12]. DSX Engine trenutno podpira DID:ethr, v razvoju pa je DID:ebsi. DID:ethr je osrednji identifikator, saj se za odkrivanje uporablja omrežje in veriga blokov Ethereum. DID je možno razrešiti (angl. resolve) in pridobiti DID dokument z atributi, ki med drugim vključujejo *verificationMethod*, ki vsebuje javni kriptografski ključ povezovalnika in *service* polje, kjer je definiran javni naslov (URL). Ta naslov definira, kje je izpostavljen povezovalnikov aplikacijsko programski vmesnik (API) do protokola podatkovnega prostora (Dataspace Protocol - DSP) [13]. Vsak povezovalnik ima lasten par kriptografskih ključev (javni in zasebni ključ), s katerimi podpisuje transakcije na verigi blokov ter omogoči varno in preverljivo komunikacijo med samimi

povezovalniki. Vsaka organizacija skrbi za lastno identiteto s čimer se zagotavlja avtentikacija in avtorizacija povezovalnikov brez centralne avtoritete.
Razširitev zagotavlja:

- Popolno decentralizacijo, brez potrebe po centralnem katalogu udeležencev, avtoriteti ali posredniku;
- Avtomatizirano integracijo, saj se udeleženci v podatkovni prostor vključijo zgolj z registracijo DID-a in vloge, ostali povezovalniki ga nato samostojno odkrijejo preko tehnologije veriženja blokov;
- Samostojnost udeležencev, ki sami skrbijo za svojo prisotnost v omrežju;
- Samoorganizacijo podatkovnih prostorov na osnovi pametnih pogodb;
- Interoperabilnost med povezovalniki;
- Usklajevanje udeležencev brez medsebojnega zaupanja (angl. trustless coordination).

Razširitev decentralizirane klirinške hiše

Namen razširitve decentralizirane klirinške hiše (ang. Decentralized Clearing House – DCH) je vzpostavitev zanesljivega, transparentnega, neizpodbitnega in decentraliziranega mehanizma za beleženje dogodkov, ki nastajajo znotraj podatkovnega prostora, saj je to ključna funkcionalnost za vse nadaljnje korake in funkcionalnosti, povezane s podatkovno ekonomijo, ki jo podpiramo s tokenizacijo. Povezovalniki beležijo vse svoje dogodke obeh ravnin (nadzorniška in podatkovna) na pametno pogodbo, ki smo jo implementirali in objavili na Ethereum verigi blokov in ima vlogo decentralizirane, transparentne, nespremenljive in zaupanja vredne shrambe dogodkov. DCH razširitev temelji na specifikaciji IDS Clearing House [14]. Vsak povezovalnik ima svoj DCH, ki spremlja relevantne interne EDC dogodke in jih pošilja na verigo blokov. Ti podatki se lahko na primer uporabijo za obračunavanje porabe podatkov ali sledljivost transakcij med dvema povezovalnikoma.
Primeri internih EDC dogodkov so:

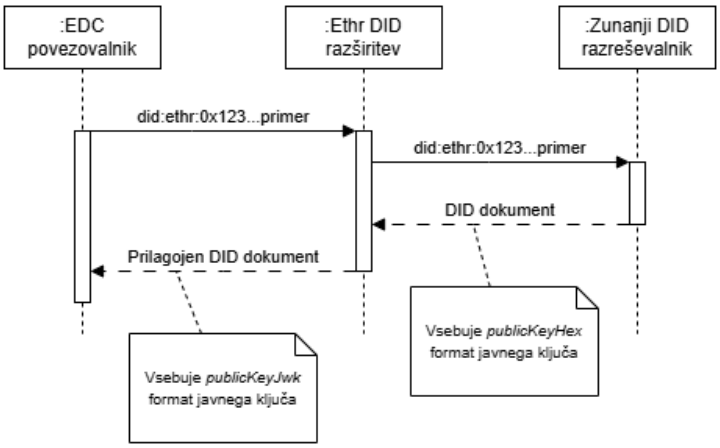
- Ponudbe podatkov (angl. asset): ustvarjanje, posodabljanje, brisanje
- Definicije politik (angl. policy definition): ustvarjanje, posodabljanje, brisanje
- Definicije pogodb (angl. contract definition): ustvarjanje, posodabljanje, brisanje
- Pogajanje pogodb (angl. contract negotiation): sprememba stanja (zahtevano, ponujeno, sprejeto, dogovorjeno, preverjeno, prekinjeno, zaključeno) [13]
- Procesi prenosa podatkov (angl. transfer process): sprememba stanja (zahtevano, začeto, suspendirano, zaključeno, prekinjeno) [13]

Ob zaznavi dogodka razširitev izlušči relevantne metapodatke (npr. ID ponudbe podatkov, ID pogajanja itd.) in jih trajno objavi na pametno pogodbo.

Razširitev Ethr DID

Zaradi potrebe po uporabi decentraliziranih identifikatorjev (ang. decentralized identifiers - DID) temelječih na infrastrukturi Ethereum, smo razvili Ethr DID razširitev, ki omogoča podporo za DID metodo »did:ethr«. Osnovni EDC podpira le metodo did:web [8].
Ob zagonu EDC povezovalnika se razširitev interno registrira kot razreševalnik za *did:ethr* in s tem povezovalniku omogoči obravnavo vseh DID identifikatorjev z metodo »did:ethr«.
Ko udeleženec *A* želi začeti komunikacijo z drugim že odkritim udeležencem *B* znotraj podatkovnega prostora, EDC povezovalnik pridobi DID in v primeru, da uporablja metodo »did:ethr« uporabi Ethr DID razširitev za pridobitev spletnega naslova (URL), kjer se nahaja DSP API povezovalnika udeleženca *B* in njegov verifikacijski javni ključ. Razširitev uporabi poljubni ustrezni zunanji DID razreševalnik (angl. DID resolver), ki vrne DID dokument, ki med drugim mora vsebovati vsaj eno verifikacijsko metodo z atributom »publicKeyHex« in URL do DSP API povezovalnika (Slika 7).
EDC trenutno podpira le »publicKeyJwk« ali »publicKeyMultibase« kot tipe verifikacijskih metod kot jih definira W3C DID [12]. Ker »did:ethr« običajno uporablja »publicKeyHex« format, Ethr DID razširitev izvede

transformacijo iz tega formata v publicKeyJwk format. »publicKeyJwk« je javni kriptografski ključ zapisan v JSON Web Key (JWK) formatu, medtem ko je »publicKeyHex« zapisan v šestnajstiškem formatu [12].
Podobno kot Ethr DID razširitev pripravljamo tudi EBSI DID razširitev, ki bo omogočala podporo »did:ebsi« metode decentraliziranih identifikatorjev, in sicer javni DID:ebsi, kjer bi se DID dokument shranjeval v javnem registru DID na infrastrukturi EBSI (European Blockchain Service Infrastructure). To smatramo ključnega pomena, saj je EBSI primarno namenjena vpeljavi decentraliziranih tehnologij v javni sektor Evropske unije in širše, kar posledično omogoča tudi lažjo integracijo in vpeljavo povezovalnika podatkovnega prostora, kot je DSX Engine v javni sektor.



Slika 7: Diagram zaporedja delovanja razširitve DID:ethr.

Razširitev registracije

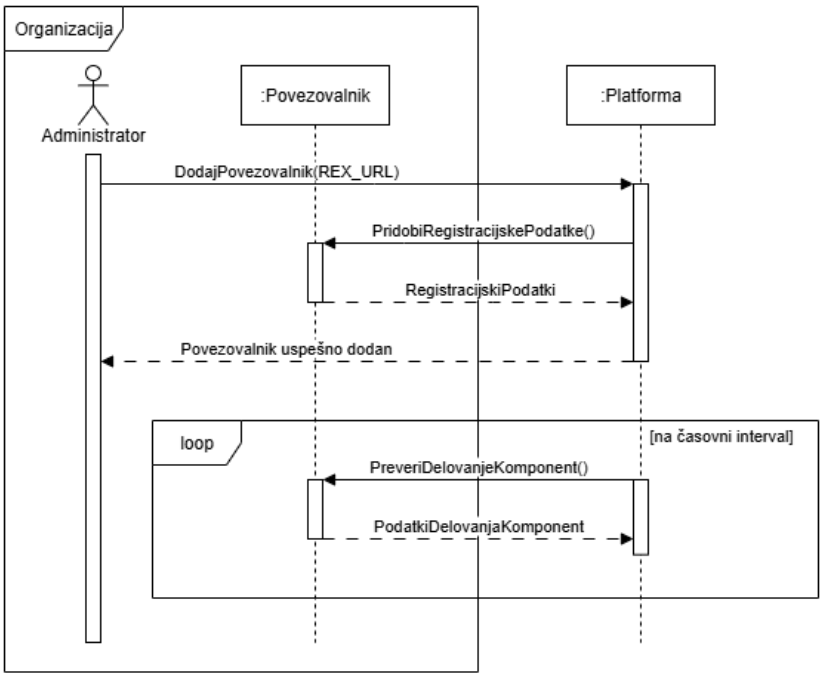
Namen razširitev registracije (ang. Registration Extension - REX) je omogočiti izpostavitev osnovnih informacij o povezovalniku, z namenom preprostejše povezave z nadzornimi ploščami, oziroma platformami podatkovnega prostora. Platforme nudijo uporabniški grafični vmesnik preko katerega organizacija lahko upravlja svoj povezovalnik, pregleduje podatkovni katalog, sklepa pogodbe in ostale povezane zadeve.
Osnovna namestitev EDC povezovalnika ne ponuja mehanizma za razkrivanje podatkov o povezovalniku, kot so identiteta udeleženca, naslovi API-jev ter stanje delovanja ali napak internih komponent (Izsek kode 1). To predstavlja oviro za hitro, enostavno in potencialno avtomatizirano vzpostavitev novega povezovalnika ter sprotno spremljanje delovanja.
REX izpostavi naslednje informacije o povezovalniku:

- Osnovne podatke (DID, vlogo);
- URL naslove, kjer se nahajajo API-ji določenih komponent, kot sta API za upravljanje povezovalnika in API za upravljanje podatkovnega kataloga;
- Stanje delovanja komponent (primeri stanj: zagnano, pripravljeno, delujoče);
- Podatke in stanje vseh podatkovnih posrednikov (Data Plane).

REX pridobiva povezovalnikove interne informacije o delovanju in jih izpostavi preko REST strežnika. Glavni končni točki sta */getRegistrationData*, ki vrne osnovne informacije o povezovalniku in */getStatusInfo*, ki prikaže trenutno stanje sistema. Končne točke so zavarovane z avtentikacijo na podlagi Bearer žetonov oziroma ključev (angl. API key). Proces delovanja REX je prikazan na sliki 8.

```
{
  "participantId": "did:ethr:0x3C44CdDdB6a900fa2b585dd299e03d12FA4293BC",
  "isProvider": true,
  "managementUrl": "http://10.5.0.10:39193/api/management",
  "catalogUrl": "http://10.5.0.10:38181/api/catalog"
}
```

Izsek kode 1: Primer odgovora Registration Extension komponente za osnovne metapodatke povezovalnika



Slika 8: Diagram zaporedja uporabe komponente REX

Politike po meri

Za podporo decentraliziranih podatkovnih prostorov, ki temeljijo na zaupanju med udeleženci, smo razvili razširitev in podporo politikam po meri, ki omogočajo vrednotenje politik na osnovi poverilnic v skladu z Decentralized Claims Protocol (DCP), ki ga implementira EDC [8]. Razširitev omogoča registracijo in uporabo politik po meri v podatkovnem prostoru.

Eden izmed ključnih ciljev je zagotoviti mehanizem, ki omogoča preverjanje članstva povezovalnikov v podatkovnem prostoru na decentraliziran način – ena izmed politik po meri. Ta politika je podprta na način, da ko povezovalnik želi ovrednotiti člansko poverilnico drugega povezovalnika izmed vseh ustrezno podpisanih poverilnic, ki mu jih je drugi povezovalnik poslal, filtrira/pridobi le poverilnico, ki je tipa članska poverilnica. Tukaj govorimo o formatu preverljive poverilnice (ang. Verifiable Credential - VC), ki je skladna s standardom VC po W3C. Iz te poverilnice prebere podatke, kot so status udeleženca, čas začetka in konca veljavnosti članstva. Te podatke preveri in kot rezultat vrne ali je poverilnica veljavna ali neveljavna.

Podpora ostalih politik po meri je implementirana po enakem principu.

4.2. Delovanje v praksi

Na visoki ravni DSX Engine deluje podobno kot standardni IDS povezovalnik: vsaka organizacija, ki želi sodelovati v podatkovnem prostoru, poganja instanco DSX Engine, preko katere ponuja svoje podatke ali dostopa do podatkov drugih.

Spodaj opisani potek ponazarja avtomatiziran, decentraliziran in zaupanja vreden način izmenjave podatkov v podatkovnemu prostoru, ki ga omogoča DSX Engine. Ključna razlika v primerjavi z drugimi povezovalniki in s klasičnim pristopom je v tem, da brez ročnih ali centraliziranih korakov (preverjanje certifikatov, podpis pogodbe,

beleženje transakcije, izvedba plačila) prevzame tehnologija – bodisi znotraj samega povezovalnika, bodisi v obliki pametnih pogodb na infrastrukturi tehnologij veriženja blokov. S tem se znižuje možnost napak, zmanjšuje potreba po posrednikih in dokumentih zunaj sistema, predvsem pa se povečuje zaupanje, saj vsak korak za seboj pusti kriptografsko preverljiv dokaz. DSX Engine je s svojim delovanjem primer vzpostavitve decentraliziranega podatkovnega povezovalnika, ki združuje najboljše prakse standardov IDS/EDC ter nove pristope, ki jih omogoča razvoj tehnologije.

Pridružitev podatkovnemu prostoru

Da se organizacije lahko pridružijo podatkovnemu prostoru, morajo pri sebi vzpostaviti povezovalnik DSX Engine in ga registrirati kot udeleženca v podatkovni prostor, pri tem pa potrebujejo veljavne digitalno podpisane poverilnice (VC) udeležbe, ki jih izdajajo skrbniki podatkovnega prostora.

Pred prvo pridružitvijo v podatkovni prostor morajo organizacije od skrbnikov podatkovnega prostora pridobiti svojo zaupanja vredno poverilnico formata VC, ki izkazuje članstvo organizacije v podatkovnem prostoru. Skrbniki podatkovnega prostora organizacije preverijo in jim odobrijo ali zavrnejo članstvo. Skrbniki podpisane poverilnice o članstvu, ki so zaupanja vredne vrnejo organizacijam, ki si jih shranijo v svoje upravljalnike identitet. Ena organizacija ima lahko eno veljavno poverilnico o članstvu. Potreba po takšnih pridružitvenih poverilnicah izhaja iz ideje podatkovnih prostorov in IDSA smernic, kjer ima vsak manjši ali večji podatkovni prostor svoja pravila igra in delovanja, ki jih je potrebno na organizacijski in tehničen način sprejeti, preden se omogoči dotičnemu povezovalniku povezovati z vsemi ostalimi.

Ker ponudbe podatkov v podatkovnem prostoru lahko zahtevajo skladnost z različnimi politikami, si organizacije lahko ustvarjajo mnoge poverilnice, v katere zapisujejo, s čim so skladne. Primeri skladnosti so: da organizacije podatke hranijo le znotraj Evropske unije, imajo določen certifikat (npr. ISO/IEC/IEEE 90003:2018), se strinjajo z nečim in ostale skladnosti, ki jih definirajo poslovne potrebe organizacije. Poverilnice morajo biti skladne z oblikovanjem, ki ga definira podatkovni prostor. Organizacije ustvarijo svoje poverilnice, ki so lastnoročno podpisane in niso zaupanja vredne, dokler jih ne podpiše član zaupanja v podatkovnem prostoru s poljubnimi trditvami (angl. claim), ki so tehnično podprte znotraj podatkovnega prostora (npr. organizacija je znotraj EU). Po tem, ko jih podpiše član zaupanja, si organizacije shranijo poverilnice v svoje upravljalnike identitet. Z ustrezno podpisanimi veljavnimi zaupanja vrednimi poverilnicami povezovalniki organizacij lahko izkazujejo skladnost z različnimi politikami ostalim udeležencem.

Po uspešni namestitvi in registraciji v podatkovni prostor DSX Engine povezovalniki samodejno z uporabo razširitve DDE periodično odkrivajo povezovalnike v podatkovnem prostoru. S tem si gradijo lokalne sezname vseh aktivnih ponudnikov podatkovnega prostora, katerega uporablja EDC komponenta pajek.

Objava podatkov v podatkovni katalog

Organizacije, ki želijo objaviti podatke v podatkovni prostor z namenom delitve z ostalimi udeleženci, morajo objaviti eno ali več ponudb (angl. asset) v svoj podatkovni katalog. Ponudbe so metapodatki dejanskih podatkov, ki jih želijo objaviti. Metapodatke v podatkovnem katalogu si lahko predstavljamo kot prikazan izdelek v katalogu neke trgovine. Opisuje kaj ta izdelek je, kje ga lahko kupimo/pridobimo in s čim moramo biti skladni (npr. nakup alkohola možen le polnoletnim osebam). Ko stranka izbere izdelek, ki ga želi, gre v ustrezno trgovino, izvede transakcijo in šele nato pridobi dejanski izdelek. Po podobnem principu deluje tudi podatkovni katalog, ki hrani le ponudbe podatkov, ki med drugim podajo tudi informacije potrebne za pridobitev dejanskih podatkov.

Objave posameznih ponudb v podatkovni katalog so sestavljene iz treh delov:

Politik (politike dostopa in pogodbene politike);

Ponudb (angl. Assets);

Povezave politik s ponudbami – definicije pogodb (angl. Contract Definitions).

Prvi del, potreben za objavo ponudb, so ustvarjene politike. Vsaka politika lahko definira tri različne kategorije zahtev [15]:

- Dovoljenja (angl. permissions) – kaj se lahko dela s podatki (npr. dovoljena je uporaba za strojno učenje);
- Prepovedi (angl. prohibitions) – kaj se ne sme delati s podatki (npr. ni dovoljena komercialna uporaba);
- Obveznosti (angl. obligations) – dolžnosti ravnanja s podatki (npr. skladnost z GDPR).

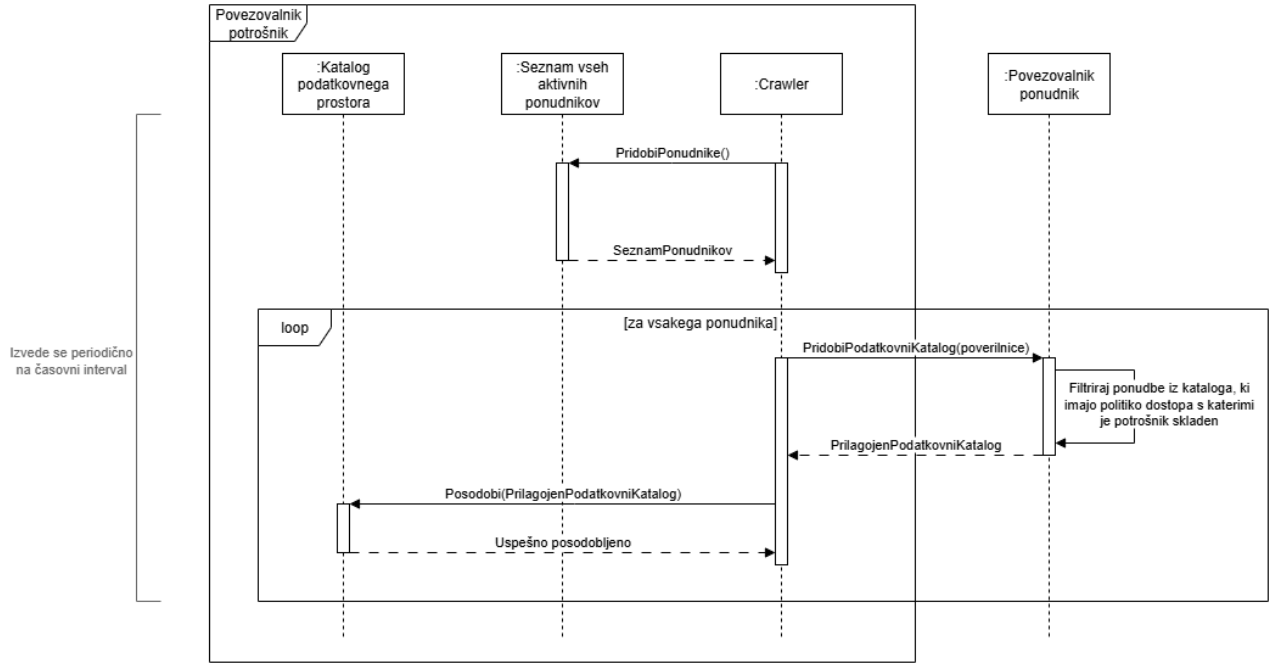
V politikah se lahko definirajo zahteve, kot so časovne omejitve (npr. dostop od 1.1.2025 naprej), lokacijske omejitve (npr. potrošnik je v Evropski uniji), omejitve shranjevanja in obdelave podatkov (npr. podatki ne smejo zapustiti Evropske Unije), cenik (npr. vsaka poizvedba stane 1 EUR ali DSX žeton), pripadnost (npr. potrošnik pripada določeni organizaciji), certifikati (npr. potrošnik mora imeti določene certifikate), namen uporabe (npr. dovoljena uporaba le za raziskovalne namene).
V teoriji so zahteve v posameznih politikah lahko karkoli, potrebno je le, da podatkovni prostor delovanje takšnih zahtev implementira. Organizacije lahko definirajo dve vrsti politik: politiko dostopa (angl. access policy) in pogodbeno politiko (angl. contract policy).
V politikah dostopa se definirajo zahteve, s katerimi morajo udeleženci podatkovnega prostora biti skladni, da jim povezovalniki ponudnikov prikažejo oz. pošljejo določene ponudbe.
Pogodbene politike definirajo zahteve s katerimi se potrošniki morajo strinjati oz. biti skladni, da bodo lahko pridobili dostop do dejanskih podatkov, ki jih želijo od ponudnikov. Politike so zapisane v jeziku ODRL (Open Digital Rights Language), ki je W3C priporočilo za izražanje pravic uporabe in omejitev za digitalne vsebine [15]. V EDC se ODRL uporablja za definiranje tako politik dostopa kot pogodbenih politik v JSON-LD formatu [8]. Organizacije lahko imajo več politik dostopa in pogodbenih politik, ki se kasneje vežejo na posamezne ponudbe.
Drugi del objave ponudb so ponudbe (angl. asset), ki jih ustvarijo organizacije za dejanske podatke, ki jih želijo deliti v podatkovnem prostoru. Ponudbam je potrebno definirati metapodatke kot so osnovna URL pot (angl. base URL) in dodatne poljubne atribute (npr. opis). Osnovna URL pot vodi do vira dejanskih podatkov znotraj organizacije, to je na primer končna točka REST strežnika znotraj organizacije. Ta pot se uporablja samo v podatkovni ravnini (ang. data plane) in ni nikoli razkrita ostalim udeležencem v podatkovnem prostoru. Poljubni atributi imajo dve kategorije, lahko so zasebni, kar pomeni, da jih vidi samo lastnik povezovalnika, ali pa javni, kar pomeni, da so vidni v podatkovnih katalogih udeležencev.
Tretji del, potreben za objavo ponudb, so definicije pogodb (angl. contract definitions), kjer se definirajo vzorci pogodb, ki se navezujejo na politiko dostopa, pogodbeno politiko in ponudbo. Več definicij pogodb z morebitnimi različnimi ponudbami se lahko navezuje na isto politiko, kar omogoča, da imajo organizacije nekaj politik, ki jih uporabljajo pri ponudbah. Definicija pogodbe v EDC je objekt *Criterion* v JSON-LD formatu [16]. Šele ob uspešnih definicijah pogodb so ponudbe, oziroma definicije pogodb, ki se navezujejo na ponudbe objavljene v podatkovni katalog povezovalnikov, kjer jih lahko odkrijejo ostali udeleženci podatkovnega prostora. Vsak povezovalnik ima en javno dostopni podatkovni katalog.

Pridobivanje podatkov

Po uspešni pridružitvi povezovalnikov v podatkovni prostor, lahko organizacije začnejo pridobivati dejanske podatke, ki jih ponujajo ostali udeleženci. Proces pridobitve podatkov je sestavljen iz naslednjih korakov [17] in prikazan na sliki 9:
Samodejno pridobivanje podatkovnih katalogov;
Pogajanje pogodbe;
Vzpostavitev procesa prenosa;
Prenos podatkov.

Prvi korak je pridobitev podatkovnih katalogov od odkritih udeležencev. EDC komponenta pajek na povezovalnikih potrošnikov, ki na podlagi lokalnega seznama vseh aktivnih ponudnikov podatkovnega prostora od vsakega ponudnika pridobi podatkovni katalog. Ko pajek od ponudnikov zahteva podatkovne kataloge,

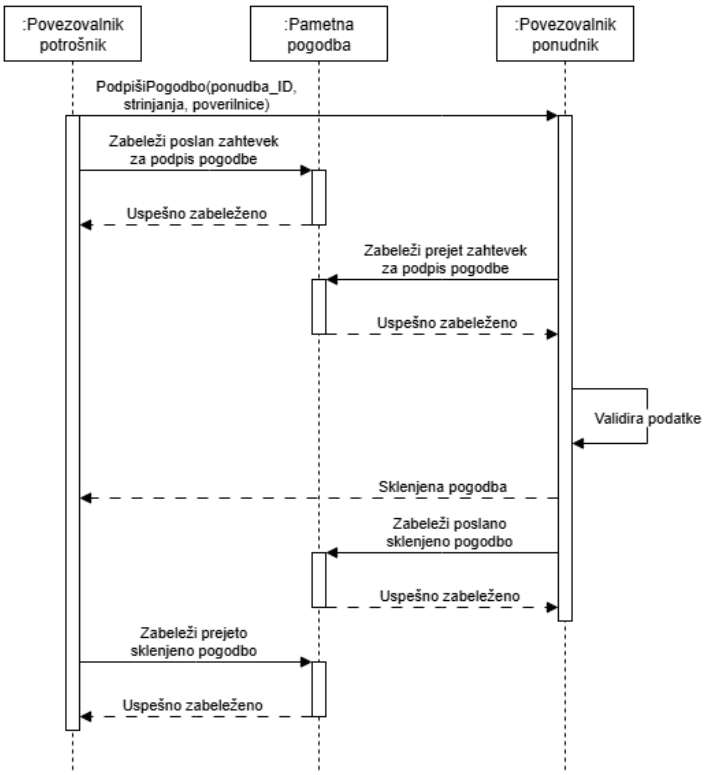
ponudniki vrnejo prilagojene podatkovne kataloge le s ponudbami, s katerimi politikami dostopa so potrošniki skladni. Ostalih ponudb posamezni potrošniki ne dobijo. Tako vsak DSX Engine povezovalnik periodično na nastavljene časovne intervale gradi lokalni predpomnjen katalog podatkovnega prostora. Glede na število aktivnih ponudnikov se dinamično paralelno sproži več instanc komponente pajek, za učinkovitejše in hitrejše posodobitve lokalnega kataloga podatkov.



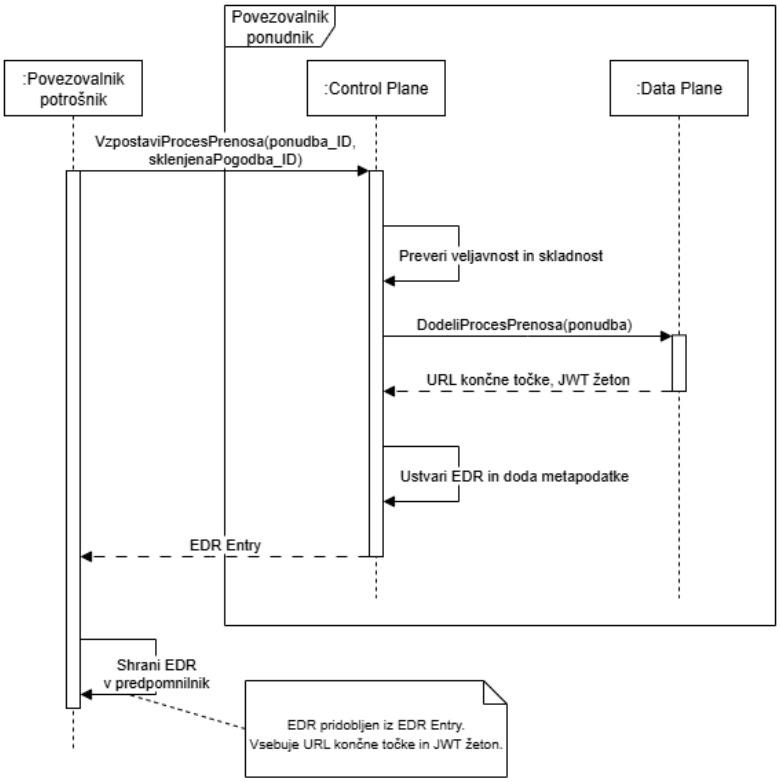
Slika 9: Posplošen diagram zaporedij samodejnega pridobivanja podatkovnih katalogov.

Drugi korak je pogajanje pogodb, med ponudniki in potrošniki. Posamezna pogodba se navezuje na natanko dva povezovalnika (eden v vlogi ponudnika, drugi v vlogi potrošnika). Ob izbiri ponudbe, potrošnik pošlje ponudniku zahtevek za podpis pogodbe. V tem zahtevku definira na katero ponudbo podatkov se navezuje in s katerimi politikami se strinja oz. je v skladu - to so dovoljenja, prepovedi, obveznosti. S tem potrošnik ponudniku posreduje tudi potrebne poverilnice. Ponudnik primerja pogodbeno politiko v definiciji pogodbe, ki se navezuje na ponudbo s politiko in poverilnicami iz potrošnikovega zahtevka. Ponudnik pričakuje, da se potrošnik strinja z vsemi zahtevami, ki so v pogodbeni politiki in prav tako preveri veljavnost in ustreznost podpisa pogodbe s strani izdajatelja poverilnic ter vsebino potrošnikovih poverilnic. V kolikor so poverilnice veljavne in vsebinsko ustrezne ter se je potrošnik strinjal z vsemi zahtevami v pogodbeni politiki, je pogodba med ponudnikom in potrošnikom uspešno sklenjena. Ta postopek se imenuje pogajanje pogodbe (angl. Contract negotiation). Pogajanja kot tudi sklenjene pogodbe (angl. Contract Agreement) razširitev DCH v povezovalnikih tako ponudnikov kot potrošnikov sproti beleži na verigo blokov. Vsak povezovalnik beleži svoje dogodke na verigo blokov in s tem zagotovi transparentnost in neizpodbitnost transakcij. Slika 10 prikazuje zaporedje sklepanja pogodbe.

Tretji korak pridobivanja podatkov je vzpostavitev procesa prenosa med določenim ponudnikom in potrošnikom. Po sklenjeni pogodbi, potrošnik ponudniku pošlje zahtevek za vzpostavitev procesa prenosa, kjer navede ponudbo in sklenjeno pogodbo na katero se navezuje. Ko ponudnik prejme zahtevek in ga uspešno preveri, dodeli proces prenosa enemu izmed razpoložljivih podatkovnih ravnin (ang. Data Plane) znotraj organizacije, preko katerega bo možno od zunaj dostopati do dejanskih podatkov. Ob uspešno vzpostavljenem procesu prenosa ponudnik ustvari Endpoint Data Reference (EDR), ki vsebuje informacije kje se nahajajo podatki (vključuje URL naslov do končne točke na podatkovnem posredniku ponudnika in JWT avtorizacijski žeton). EDR se vključno z dodatnimi metapodatki zapakira v zapis EDR Entry in kot odgovor pošlje potrošniku. Pridobljen zapis si potrošnik shrani lokalno. V kolikor JWT žetonu poteče veljavnost, potrošnik lahko ponovno ponudniku pošlje zahtevek in pridobi nov JWT žeton. Proces je predstavljen na sliki 11.



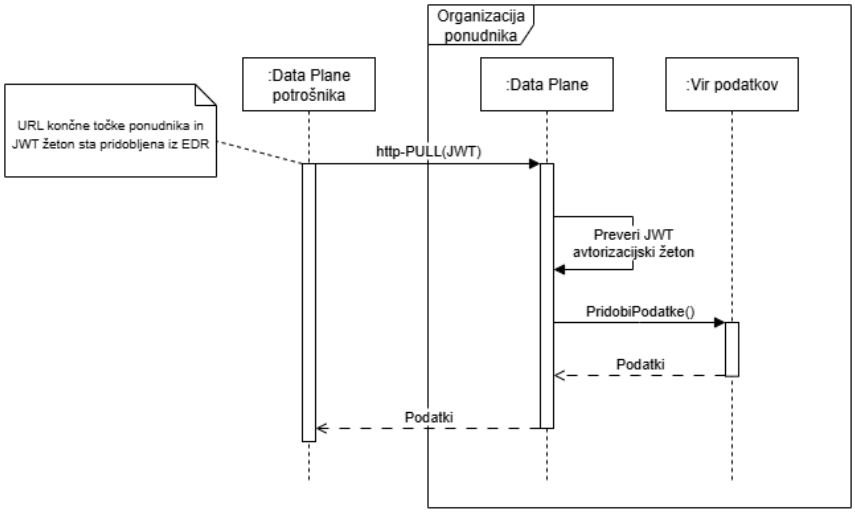
Slika 10: Posplošen diagram zaporedij pogajanja pogodbe.



Slika 11: Posplošen diagram zaporedij vzpostavitve procesa prenosa.

Četrti, zadnji korak pridobivanja podatkov v podatkovnem prostoru je prenos podatkov od ponudnika do določenega potrošnika. Ko potrošnik iz EDR razbere URL naslov končne točke podatkovnega posrednika in

avtorizacijski žeton JWT za dostop, lahko na prejet URL naslov pošlje zahtevek GET, kjer v glavi HTTP zahtevka doda atribut *Authorization*, katerega vrednost nastavi na prejet žeton JWT. Podatkovna ravnina povezovalnika ponudnika prejme zahtevek, preveri avtorizacijski žeton in po uspešni avtorizaciji pridobi dejanske podatke iz podatkovnega vira znotraj organizacije (npr. iz REST strežnika ali podatkovne baze, kjer so dejanski podatki shranjeni) ter jih posreduje potrošniku. Ta tip pridobivanja podatkov se imenuje Consumer-PULL in v zgoraj opisanem primeru poteka preko protokola HTTPS. Podprt je tudi Provider-PUSH, kjer ob vzpostavitvi procesa pridobitve podatkov podamo URL do našega ponora podatkov, na katerega bo ponudnik samodejno naložil podatke na način, kot je na primer pretakanje v realnem času (angl. streaming) ali pa periodično odlaganje podatkov. Slika 12 predstavlja proces prenosa podatkov.



Slika 12: Diagram zaporedij prenosa podatkov.

4.3. Tokenizacija

Koncept

Ena izmed osnovnih vodil podatkovnih prostorov je omogočanje t. i. **podatkovne ekonomije**. Znotraj podatkovnega prostora lahko ponudniki podatke ne le ponudijo temveč prodajajo potrošnikom. DSX Engine omogoča monetizacijo podatkov znotraj podatkovnega prostora z uporabo pametne pogodbe za DSX žetone in pametne pogodbe za izvajanje plačil na tehnologiji veriženja blokov.

Podrobnosti kako deluje tokenizacija znotraj podatkovnega prostora so odvisne od implementacije in odločitev podatkovnega prostora.

DSX žeton (ERC-20)

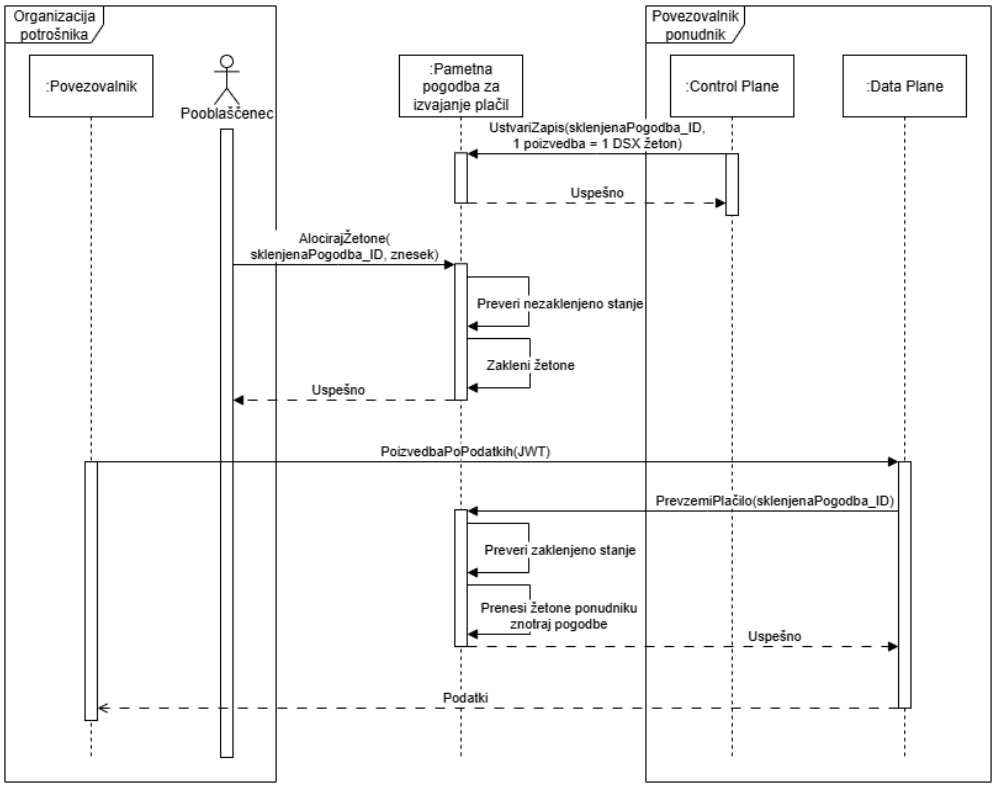
DSX žeton (angl. DSX token) temelji na pametni pogodbi, ki temelji na standardu zamenljivih žetonov OpenZeppelin ERC-20 na verigi blokov Ethereum [18]. DSX žetoni delujejo kot digitalna valuta za namene prodaje podatkov med ponudniki in potrošniki znotraj podatkovnega prostora. S tem zagotovimo transparentnost in ne-zanikanje transakcij.

Skrbniki podatkovnega prostora opredelijo, kako bodo DSX žetone delili oz., kako jih bodo predajali/prodajali udeležencem podatkovnega prostora. V spodaj opisanem primeru udeleženci podatkovnega prostora lahko DSX žetone zamenjajo z žetoni ETH in obratno, po tečaju, ki ga določi podatkovni prostor.

Ob objavi podatkov v podatkovni katalog ponudniki v pogodbeni politiki (angl. Contract Policy) pod obveznosti (angl. obligation) določijo med drugimi zahtevami tudi način obračunavanja podatkov. Spodaj je opisan primer obračunavanja na poizvedbo, na primer 1 DSX žeton na poizvedbo.

Potrošniki, ki želijo dostopati do plačljivih podatkov v podatkovnem prostoru, morajo na pametno pogodbo za izvajanje plačil naložiti žetone DSX (npr. 3 DSX žetone).

Pri pridobivanju plačljivih podatkov povezovalniki potrošnikov in ponudnikov sklenejo pogodbe (angl. Contract Agreement), nakar povezovalniki ponudnikov samodejno v pametno pogodbi za izvajanje plačil kličejo funkcijo, ki ustvari zapis na podlagi katerega potekajo plačila.
Vsak zapis se navezuje na natanko eno sklenjeno pogodbo, kjer je zabeležen tudi način obračunavanja (npr. 1 poizvedba = 1 žeton DSX).
Potrošniki morajo pred poizvedbo za podatke v pametni pogodbi klicati funkcijo s katero dovolijo, da se poljuben znesek izmed vseh žetonov, ki so jih naložili na pametno pogodbo, alocira za plačila določene sklenjene pogodbe za podatke. Potrošniki lahko kadarkoli ta sredstva tudi odklenejo.
Ko povezovalniki potrošnikov pošljejo zahtevke po podatkih, podatkovna ravnina (ang. Data Plane) ponudnikov v pametni pogodbi kličejo funkcijo za prevzem plačila. Funkcija v pametni pogodbi preveri, ali je v zaklenjenem znesku, ki se navezuje na določeno sklenjeno pogodbo, dovolj žetonov. V kolikor je žetonov dovolj, se ustrezna količina žetonov, glede na način obračunavanja, prenese ustreznemu ponudniku znotraj pogodbe. Podatkovni posredniki ponudnikov nato posredujejo podatke potrošnikom. Proces je predstavljen na sliki 13.



Slika 13: Diagram zaporedij trgovanja podatkov.

Udeleženci podatkovnega prostora lahko nezaklenjen znesek, ki ga imajo na pametni pogodbi zamenjajo v ETH žetone po tečaju, ki ga je določil podatkovni prostor (npr. 1 DSX žeton = 0,0005 ETH žetona).
V prihodnosti bo možno nastaviti tudi druge načine obračunavanja podatkov. Na primer 1 DSX žeton za 1 GB prenešenih podatkov, 10 DSX žetonov za en mesec poizvedovanja in ostale načine, ki jih potrebujejo ponudniki za njihovo poslovanje.

4.4. Primerjava

Za boljše razumevanje razlike med osnovno implementacijo EDC in našo nadgradnjo DSX Engine, v nadaljevanju podajamo primerjalni pregled ključnih funkcionalnosti in pristopov. Primerjava jasno pokaže, kako DSX Engine gradi na temeljnih zmožnostih EDC in jih nadgrajuje z dodatnimi mehanizmi decentralizacije, odpornosti in podpore novim poslovnim modelom.

Tabela 1: Primerjava med osnovno implementacijo EDC in DSX Engine.

Lastnost / Funkcionalnost	EDC (osnovna implementacija)	DSX Engine (nadgradnja EDC)
Registracija udeležencev	Centralizirana ali preko posrednika	Decentralizirana registracija preko pametnih pogodb (DDE)
Odkrivanje udeležencev	Federativno, odvisno od vnaprej znanih vozlišč	Dinamično, bazirano na pametnih pogodbah, avtomatsko odkrivanje
Upravljanje identitet	Podpora did:web	Podpora DID:ethr, načrtovana DID:ebsi; decentralizirana denarnica identitet
Beleženje dogodkov / revizijska sled	Lokalno pri vsakem udeležencu	Decentralizirana klirinška hiša (DCH) na verigi blokov
Podpora pametnim pogodbam	Ni vgrajene neposredne podpore	Neposredna podpora pametnim pogodbam na platformi Ethereum
Tokenizacija in plačila	Ni privzete podpore	DSX ERC-20 žeton in pametne pogodbe za obračunavanje
Politični mehanizmi	Standardne politike EDC	Politike po meri na osnovi preverljivih poverilnic (VC)

5 Pilotiranje

Kot pilotni primer implementacije konceptov podatkovnih prostorov in uporabe povezovalnika DSX Engine izpostavljamo DIH AGRIFOOD Data Space (DADS) – podatkovni prostor za agroživilski sektor. DADS je pobuda Digitalnega inovacijskega stičišča za kmetijstvo in prehrano (DIH Agrifood) v Sloveniji, ki vzpostavlja federativno platformo za deljenje podatkov v agroživilstvu. Po definiciji gre za "federalno organizirano in upravljano platformo za deljenje podatkov iz agriživilskega sektorja, kjer se podatki delijo na osnovi vnaprej definiranih pogojev uporabe" [19]. Ključni cilj DADS je povezati različne deležnike v agroživilskem ekosistemu – od kmetijskih gospodarstev, zadrug, svetovalnih služb, živilskopredelovalnih podjetij do javnih ustanov – in jim omogočiti nadzorovano izmenjavo podatkov v skupno korist (npr. za optimizacijo pridelave, sledljivost hrane, okoljsko spremljanje itd.).

DADS je zgrajen skladno z evropskimi smernicami za podatkovne prostore: ima upravljavski model (governance), utemeljen na etičnih načelih in evropskih regulativah (npr. Data Act, Evropska strategija za podatke), uporablja standarde, ki jih definira IDSA, ter sledi načelom, opredeljenim v okviru projekta OPEN DEI (Open Digital Ecosystems, ki je pripravil oblikovalska načela za podatkovne prostore) [20]. Tehnično gledano DADS vključuje vse osnovne gradnike podatkovnega prostora: identitetni sistem, katalog ponudb, povezovalnike (DSX Engine) pri udeležencih, modul za upravljanje soglasij in politik uporabe ter mehanizme za beleženje in sledljivost. DIH Agrifood kot iniciator projekta nastopa tudi kot koordinator podatkovnega prostora – skrbi za vzdrževanje pravil (t. i. knjiga pravil podatkovnega prostora) in nudi podporo udeležencem pri priključitvi. Pomembno je izpostaviti, da DADS ni centralizirana baza podatkov: podatki ostajajo pri posameznih ponudnikih podatkov (npr. pri kmetih, podjetjih ali institucijah), med tem ko DADS poskrbi za infrastrukturo, da ti podatki najdejo pot do porabnikov podatkov pod znanimi in nadzorovanimi pogoji [21].

V okviru DADS je uporabljen povezovalnik DSX Engine. Platforma DADS se ponaša s tem, da ne gre zgolj za deljenje statičnih naborov podatkov, temveč omogoča povezovanje do živih podatkovnih virov (npr. realno časovni podatki preko API-jev). To je bilo demonstrirano, ko je DADS uspešno povezal podatke Mestne občine Murska Sobota (MOMS). MOMS je preko svojega povezovalnika, ki teče na njeni infrastrukturi, registriraa REST API storitve z okolijskimi in GIS podatki v katalog DADS. Prav taka pa lahko potrošniki podatkov le te v realnem času uporabljajo preko DADS ob priključitvi v podatkovni prostor in uporabi povezovalnika DSX Engine [21].
V pilotnem primeru DADS uporabljamo konzorcijsko omrežje EduCTX, ki temelji na Hyperledger Besu. Omrežje je že vzpostavljeno v okviru DIH Agrifood in služi kot zaupanja vredna platforma za izvajanje pametnih pogodb, beleženje dogodkov in upravljanje identitet. Prednost te izbire je, da so vsi tehnični in organizacijski procesi za upravljanje konzorcijskega omrežja verige blokov že vpeljani, kar je skrajšalo čas implementacije in zmanjšalo stroške pilotiranja.
DIH Agrifood prav tako zagotavlja centralni portal, kjer najprej vsaka organizacija registrira oziroma poveže svoj DSX Engine povezovalnik. Nato ga preko tega portala upravlja. Povezovalnik vsake organizacije skrbi za pridobitev kataloga celotnega podatkovnega prostora, omogoča objavo podatkov organizacije v podatkovni prostor, določitev pravil in pogojev dostopa do podatkov, ter sklepanje pogodb za pridobitev podatkov z drugimi udeleženci podatkovnega prostora DADS, ki uporabljajo DSX Engine povezovalnik.
Primer DADS ponazarja vrednost podatkovnih prostorov v praksi. Kmetijski podatki, ki so bili prej razpršeni in težko dostopni, so zdaj na voljo vsem udeležencem, ki izpolnjujejo pogoje uporabe, hkrati pa kmetje kot lastniki podatkov ohranijo suverenost – sami odločajo, katere podatke delijo in pod kakšnimi pogoji. Implementacija DSX Engine povezovalnikov v DADS je pokazala, da je tehnologija dovolj zrela za realne primere uporabe, hkrati pa je razkrila tudi nekatera praktična vprašanja. Med njimi so pomembna interoperabilnost (povezovanje z drugimi podatkovnimi prostori – npr. v prihodnosti povezava DADS z evropskim podatkovnim prostorom za kmetijstvo), uporaba standardne semantike (v DADS so se morali dogovoriti o skupnih podatkovnih shemah za npr. podatke o pridelkih, kar kaže na potrebo po mednarodnih standardih) in pravno zagotavljanje skladnosti (pogodbe, sklenjene preko sistema, je treba uskladiti z obstoječim pravnim okvirom varstva podatkov in obligacij).
DADS sodeluje tudi v mednarodnih projektih (Horizon Europe projekti, kot so DIVINE, ZeroW, ipd. [21]), kjer deluje kot primer dobre prakse pri postavljanju podatkovnih prostorov. Izkušnje, pridobljene z DADS, se vračajo v nadaljnji razvoj DSX Engine – pilotiranje je namreč razkrilo, katere funkcionalnosti so najbolj koristne in katere je treba še izboljšati. Tako je DADS pomemben korak na poti od koncepta do polnopravne operativne platforme za podatkovno ekonomijo v kmetijstvu.

6 Zaključek

V tem članku smo predstavili DSX Engine kot primer decentraliziranega povezovalnika podatkovnih prostorov, ki nadgrajuje referenčno odprtokodno ogrodje Eclipse Dataspace Components z mehanizmi za decentralizirano zaupanje in avtomatizirano sklepanje pogodb. Pokazali smo, kako uvedba pametnih pogodb, decentraliziranih identitet in tokenizacije omogoča odpravo osrednjih točk odpovedi ter povečuje stopnjo zaupanja in varnosti v med-organizacijski izmenjavi podatkov. Predstavljene razširitve, kot so DDE, DCH in podpora za decentralizirane identifikatorje omogočajo popolnoma decentralizirano registracijo in odkrivanje udeležencev, neizpodbitno revizijsko sled ter interoperabilnost med heterogenimi sistemi in organizacijami.
Praktično vrednost in zrelost DSX Engine smo prikazali na primeru implementacije v pilotnem podatkovnem prostoru DIH AGRIFOOD Data Space (DADS), kjer rešitev omogoča kmetijskim in okoljskim deležnikom varno, nadzorovano in sledljivo izmenjavo podatkov ob ohranjanju njihove suverenosti.
V prihodnje nameravamo nadaljevati z razvojem DSX Engine v smeri večje skladnosti z evropskimi specifikacijami, kot so Gaia-X in EBSI DID, razširiti podporo semantičnim standardom za interoperabilnost med sektorji, omogočiti več načinov prenosa podatkov in trgovanja s podatki ter izboljšati uporabniško izkušnjo upravljanja in nadzora nad podatkovnimi povezovalniki. Prav tako bomo na podlagi izkušenj iz pilotnih

implementacij naslavljali odprta vprašanja pravne skladnosti in standardizacije pogodb ter prispevali k vzpostavitvi zaupanja vredne podatkovne ekonomije v evropskem in mednarodnem prostoru.

Literatura

[1] M. Spiekermann, „Eclipse Dataspace Connector: Trusted Data Sharing With Sovereignty," 20. oktober 2021. [Elektronski]. Dostopno na: https://newsroom.eclipse.org/eclipse-newsletter/2021/october/eclipse-dataspace-connector-trusted-data-sharing-sovereignty. [Poskus dostopa 26. maj 2025].

[2] T. Dam, L. D. Klausner, S. Neumaier in T. Priebe, „A Survey of Dataspace Connector Implementations," 9. januar 2024. [Elektronski]. Dostopno na: https://arxiv.org/pdf/2309.11282. [Poskus dostopa 26. maj 2025].

[3] N. Kliček, M. Šestak in M. Turkanović, „Inovacije v arhitekturah podatkovnih prostorov," september 2024. [Elektronski]. Dostopno na: https://plus.cobiss.net/cobiss/si/sl/bib/ktfmb/207213827. [Poskus dostopa 26. maj 2025].

[4] European Commision, „A European strategy for data," 9. april 2025. [Elektronski]. Dostopno na: https://digital-strategy.ec.europa.eu/en/policies/strategy-data. [Poskus dostopa 26. maj 2025].

[5] International Data Spaces Association, „IDS-RAM 4," [Elektronski]. Dostopno na: https://docs.internationaldataspaces.org/ids-knowledgebase/ids-ram-4. [Poskus dostopa 26. maj 2025].

[6] I. Chulani, „Understanding the IDS Reference Architecture Model," 2. oktober 2024. [Elektronski]. Dostopno na: https://internationaldataspaces.org/understanding-the-idsa-reference-architecture-model. [Poskus dostopa 26. maj 2025].

[7] International Data Spaces Association, „IDSA Rulebook," [Elektronski]. Dostopno na: https://docs.internationaldataspaces.org/ids-knowledgebase/idsa-rulebook. [Poskus dostopa 26. maj 2025].

[8] Eclipse Foundation, „EDC Documentation," [Elektronski]. Dostopno na: https://eclipse-edc.github.io/documentation/. [Poskus dostopa 26. maj 2025].

[9] J. Marino, „The Eclipse Dataspace Connector Architecture and Principles," 31. januar 2022. [Elektronski]. Dostopno na: https://raw.githubusercontent.com/eclipse-edc/Collateral/main/Events/Conferences/2022-01%20EDC%20Conference/2022-01-31%20EDC%20-%20Architecture%20and%20Concepts.pdf. [Poskus dostopa 30. junij 2025].

[10] A. Bertagnolli, P. Koen, M. Kollenstart, P. Latzelsperger, J. Marino, J. Pampus, R. Rajani, E. Risa, M. Ruland, B. Scholtes, N. Schulte, M. Spiekermann, S. Steinbuss, A. Weiß in H. Yildiz, „Eclipse Decentralized Claims Protocol," Eclipse Foundation, 19. junij 2025. [Elektronski]. Dostopno na: https://eclipse-dataspace-dcp.github.io/decentralized-claims-protocol/v1.0-RC4. [Poskus dostopa 3. julij 2025].

[11] Gaia-X, „Building a Dataspace: Technical Overview," april 2023. [Elektronski]. Dostopno na: https://www.gaia-x.at/wp-content/uploads/2023/04/WhitepaperGaiaX.pdf. [Poskus dostopa 26. maj 2025].

[12] M. Sporny, D. Longley, M. Sabadello, D. Reed, O. Steele in C. Allen, „Decentralized Identifiers (DIDs) v1.0," World Wide Web Consortium (W3C), 19. julij 2022. [Elektronski]. Dostopno na: https://www.w3.org/TR/did-1.0. [Poskus dostopa 20. junij 2025].

[13] P. Koen, M. Kollenstart, J. Marino, J. Pampus, A. Turkmayali, S. Steinbuss in A. Weiß, „Dataspace Protocol 2025-1-RC2," Eclipse, 19. junij 2025. [Elektronski]. Dostopno na: https://eclipse-dataspace-protocol-base.github.io/DataspaceProtocol/2025-1-RC2/. [Poskus dostopa 24. junij 2025].

[14] D. i. H. Bastiaansen, G. Bramm, J. Ceballos, M. Gall in M. Kollenstart, „Specification: IDS Clearing House," International Data Spaces Association, 1. april 2020. [Elektronski]. Dostopno na: https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-White-Paper-Specification-IDS-Clearing-House-.pdf. [Poskus dostopa 23. junij 2025].

[15] W3C, „ODRL Information Model 2.2," World Wide Web Consortium (W3C), 15. februar 2018. [Elektronski]. Dostopno na: https://www.w3.org/TR/odrl-model/. [Poskus dostopa 26. junij 2025].

[16] Eclipse Foundation, „EDC Developer's Handbook," 29. marec 2024. [Elektronski]. Dostopno na: https://github.com/eclipse-edc/docs/blob/main/developer/handbook.md. [Poskus dostopa 26. junij 2025].

[17] Eclipse Foundation, „EDC transfer samples,“ 25. junij 2025. [Elektronski]. Dostopno na: https://github.com/eclipse-edc/Samples/tree/main/transfer. [Poskus dostopa 27. junij 2025].

[18] OpenZeppelin, „ERC-20,“ OpenZeppelin, [Elektronski]. Dostopno na: https://docs.openzeppelin.com/contracts/5.x/erc20. [Poskus dostopa 4. julij 2025].

[19] DIH AGRIFOOD, „DIH AGRIFOOD Data Space (DADS) PowerPoint,“ 18. januar 2024. [Elektronski]. Dostopno na: https://www.gov.si/assets/ministrstva/MDP/Dokumenti/Podatkovni-prostori/DIH-Agrifood.pdf. [Poskus dostopa 26. maj 2025].

[20] DIH AGRIFOOD, „DIH AGRIFOOD Data Space (DADS),“ 2025. [Elektronski]. Dostopno na: https://dih-agrifood.com/dih-agrifood-data-space. [Poskus dostopa 26. maj 2025].

[21] M. Bunderla, „DADS Milestone - MOMS Collaboration,“ december 2024. [Elektronski]. Dostopno na: https://www.linkedin.com/posts/miran-bunderla-a11a47b5_important-announcement-from-the-dih-agrifood-activity-7262825514682269696-bd8L. [Poskus dostopa 26. maj 2025].