Esemény Alapú Adatszolgáltatási Platform Pilot Projekt

Az EMAP rendszer architektúra variánsainak felderítése és kiértékelése

2023. november

EMAP
1051 Budapest, József nádor tér 2-4.
Telefon: (1) 428-5501 • E-mail: emap@pm.gov.hu

TARTALOMJEGYZÉK

Vezetői összefoglaló

A foglalkoztatással összefüggő adatszolgáltatási követelményeknek való maradéktalan megfelelés jelentős erőforrásokat igényel a magyar vállalkozások részéről. Az Esemény Alapú Adatszolgáltatási Platform (a továbbiakban EMAP) Pilot projekt elsődleges célja ezen *adminisztrációs terhek csökkentése*.

Ennek kapcsán a foglalkoztatók adatszolgáltatási folyamatainak megkönnyítése érdekében *egyszeri, egycsatornás* adatszolgáltatási folyamat kerül kialakításra, mely a jelenlegi bérszámfejtési szoftverek megoldásaihoz és munkafolyamataihoz igazodna. Ez lehetővé tenné, hogy az adatszolgáltatás a vállalkozások üzleti folyamataihoz alkalmazkodjon, a jelenlegi hatósági szakrendszerek eljárásai helyett.

A rendszer fő funkciói

Az EMAP rendszer felelőssége lesz a szolgáltatott adatok elosztása a megfelelő szakrendszerek felé, az adatokból a szakrendszerek által befogadható formátumú és tartalmú nyomtatványok előállítása, valamint – a nyomtatványok tartalmi változásainak tekintetében – a *jogszabályok követése*. Ez utóbbival csökken a foglalkoztatókra és a bérszámfejtő szoftverek fejlesztőire nehezedő szabálykövetési feladattömeg is, mivel az általuk beküldendő alapesemények köre nagyon ritkán változik. A platformtól elvárt feladat továbbá, hogy már a folyamat elején, az adatok bevitelekor *formai és tartalmi ellenőrzést* végezzen, azonnali visszajelzést adva az adatszolgáltatók felé. Ettől az adatszolgáltatások minőségi javulása várható, amivel egyrészt nőne az *adatkonzisztencia*, másrészt az adminisztrációs terhek csökkenése így nem csak a foglalkoztatók oldalán jelentkezne, hanem a szakhatóságok esetében is, ugyanis elkerülhető a hibásan befogadott adatokból keletkező javíttatási és hiánypótlási eljárások megindítása.

A rendszer a szolgáltatott adatok alanyainak (mint például a természetes személyeknek) is képes lesz a rájuk vonatkozó adatokat visszamutatni. Ennek megfelelően a külső *lekérdezések* támogatása és hatékony megvalósítása is fontos funkció. Végezetül, a projekt egy fokozatos integráció keretében tervezi bevezetni az esemény alapú adatszolgáltatást.

Ennek eszköze az úgynevezett *nyomtatvány transzformáció* funkció, amely a rögzített adatok alapján a szakrendszerek által jelenleg is használt nyomtatványokat előállítja, és azokat adja át az érintett szervezeteknek a közvetlen integráció helyett. A nyomtatvány alapú integráció ezen kívül átmenetileg (a Pilot időtartama alatt) egy *önellenőrzés* célú párhuzamos működést tesz lehetővé a jelenleg is működő adatszolgáltatással.

Szisztematikusan, részletes szempontrendszer alapján megvizsgálásra került, hogy a fenti funkcionalitások milyen technológiai megoldások mellett valósíthatók meg.

A rendszer által kezelt adatok

Az EMAP rendszer egyik fontos (a technológiai javaslatot leginkább befolyásoló) aspektusa a rendszer által kezelt és mások felé megosztott, illetve előállított adatok köre. A rendszer által kezelt adatokat a következő főbb típusokba oszthatjuk fel forrásuk, tartalmuk és felhasználásuk szempontjából:

- *Entitás adatok*: a rendszerben megjelenő entitásoknak (például az adatszolgáltatások alanyainak) azonosító adatai.
- *Esemény adatok*: a definiált esemény típusoknak megfelelő, az adatszolgáltatás során keletkezett adatok.
- *Állapot adatok*: az események eredményeképp előálló, jogi relevanciával bíró állapotok és göngyölt adatok.

Az entitás és állapot adatok esetén vizsgálandó a DIMOP 1.3.7-ben létrehozásra kerülő "*Komplex Jogviszony-nyilvántartás*" központi elemével való átfedés, annak integrációja, illetve közvetlen felhasználása. A szóban forgó fejlesztés egy, a közigazgatási szolgáltatások automatizálását kiszolgáló "entitásállapot-mátrix", ami a hatósági szolgáltatások üzleti logikájában vizsgálandó kritériumok kiszolgálását támogatja. A szolgálgatás ugyanakkor egy adatszolgáltatás tartalmának tartalmi ellenőrzéséhez is felhasználható lehet.

A rendszer által kezelt adatok jellege mentén két fő technológiai irányt, és ennek megfelelő rendszer architektúra variánst határoztunk meg: a megszokott, *centralizált* adattárolási módszeren túl javaslatot teszünk egy innovatív, *blokklánc* alapú elosztott adattárolásra is.

Kiemelt fontossággal vizsgáltuk az egyes megoldásokat az adatkezelési jogszabályoknak (pl. *GDPR*) való megfelelésük szempontjából. Az érzékeny, személyes entitás adatoknak egy központi fél által megvalósított, centralizált tárolása egy olyan szereplőt vezetne be a rendszerbe, akinek az adatok teljes egésze felett adatkezelési felhatalmazással kellene rendelkeznie. Kerülve az ilyen mértékű központosítást, az egyik fő tervezési szempont az volt, hogy minden érzékeny adatot kizárólag azok a szervezetek tudjanak kezelni, akiknek jelenleg is joguk van rá.

A szervezetek közötti komplex adatmegosztás és -kezelés funkciók (valamint a funkciókon túlmutató követelmények, mint például a rendelkezésreállás) megvalósítására hatékony megoldást kínálnak a *zárt és jogosultságkezelt blokklánc* platformok.

Az innovatív, blokklánc alapú megoldás

A privát blokklánc platformok az elosztott főkönyvi technológiák egyik típusa, amelyet speciális felhasználási esetekre terveztek. Egy privát blokkláncban a hálózathoz való hozzáférés a résztvevők egy kiválasztott csoportjára korlátozódik, így olyan szervezetek vagy konzorciumok számára alkalmas, ahol a bizalom és a résztvevők azonosítási mechanizmusai már adottak. Emellett lehetővé teszik a hálózatirányítás és az adatok titkosságának finomabb ellenőrzését, így különösen alkalmasak *kollaboratív vállalati alkalmazásokhoz*.

A blokklánc platformok decentralizált felépítésüknek köszönhetően eleve magas rendelkezésre állású, infrastrukturális szempontból hibatűrő megoldások. Ezen kívül kiemelt hangsúlyt fektetnek a szereplők közötti hatékony és biztonságos adatcserére, amely adatok szakterülettől függően számtalan dolgot reprezentálhatnak. A beépített kriptográfiai mechanizmusoknak köszönhetően a blokklánc platformok elsődleges funkcióként biztosítják az adatok integritását, az adattal való műveletek folyamatos naplózását, a napló elemek utólagos módosíthatatlanságát, illetve a módosítások letagadhatatlanságát. Mindezen funkciók továbbá lehetőséget biztosítanak a rendszer folyamatos auditálhatóságára.

Az érzékeny adatok blokkláncon történő tárolásához kettő fő kihívásra kell megoldást adni: 1) csak azok a szereplők legyenek képesek *feldolgozni* adott típusú személyes adatokat, akiknek joguk van rá, illetve 2) a személyes adatoknak *törölhetőnek* kell lenniük.

Mindkét pontra megoldásként szolgál a tervezés során kidolgozott *deidentifikációs* módszer. A deidentifikációs lépés során a személyes adatokat megfosztjuk azonosíthatóságuktól úgy, hogy a személyes azonosítókat *egyedi véletlen számokra (EVSZ)* cseréljük minden adatpont esetén. Ez a módszer nem csupán pszeudonimizálja az adatpontokat, hiszen minden adatpont külön EVSZ-t kap, akkor is, ha ugyanahhoz az adatalanyhoz tartoznak. Ezáltal az adatok *korrelációjának* lehetőségét is *minimalizálja* a módszer.

A deidentifikált adat nyugodtan megosztható (gyors és konzisztens módon) a blokkláncban résztvevő szervezetek, mint adatfeldolgozók között. Az EVSZ-k eredeti azonosítókhoz rendelését az úgynevezett *kapcsolótábla* tárolja, minden szervezetnél saját hatáskörben, hagyományos tárolási módszerekkel. Ez teszi lehetővé az érzékeny adatok blokkláncból való „*törlését*”. Ha egy kapcsolótábla bejegyzést törlünk, akkor a hozzá tartozó, blokkláncon tárolt adat végleg *anonimmá* válik, ezáltal megszűnik személyes adatnak lenni, effektíve megvalósítva annak törlését („felejtését”). A kapcsolótábla bejegyzések *szervezetek közötti szétosztása* fogja meghatározni azt, hogy melyik szervezet melyik személyes adat (re-identifikáláson keresztüli) feldolgozására képes.

A blokklánc alapú architektúra tovább egyszerűsíthető fejlesztési és üzemeltetési szempontból is (*technológiai homogenizáción* keresztül), ha blokklánc technológiaként a

Hyperledger Fabric platform kerül kiválasztásra. A Fabric egy nyílt forráskódú blokkláncplatform, amelyet vállalati alkalmazásokhoz terveztek. Ennek megfelelően privát és jogosultságkezelt, azaz a blokkláncot olvasni és módosítani tudó résztvevők köre szigorú szabályok mentén definiált és felügyelt. A platform *moduláris, skálázható* és az *adatvédelemre* összpontosít. A Fabric programozható *okosszerződéseivel* és beépített adatvédelmi módszereivel olyan iparágak számára is alkalmas, mint a pénzügy, egészségügy és az ellátási láncok. Továbbá nagy hangsúlyt fektet az engedélyezett hozzáférésre, az identitás-kezelésre és a robusztus, decentralizált irányításra a résztvevő szervezetek között. A teljesen blokklánc alapú, de továbbra is titkosság megőrző megoldás támaszkodik a Fabric *privát adat kollekció* képességére. A privát adat kollekció a Fabric olyan funkciója, amely lehetővé teszi az érzékeny adatok *szelektív megosztását* egy blokklánc-hálózaton belül, miközben az adatokat bizalmasan és a nyilvános főkönyvön kívül tartja. Ezek a kollekciók fogják betöltetni a szervezet szintű érzékeny kapcsolótáblák szerepét, kihasználva a mögöttes hagyományos adatbázis technológiát.

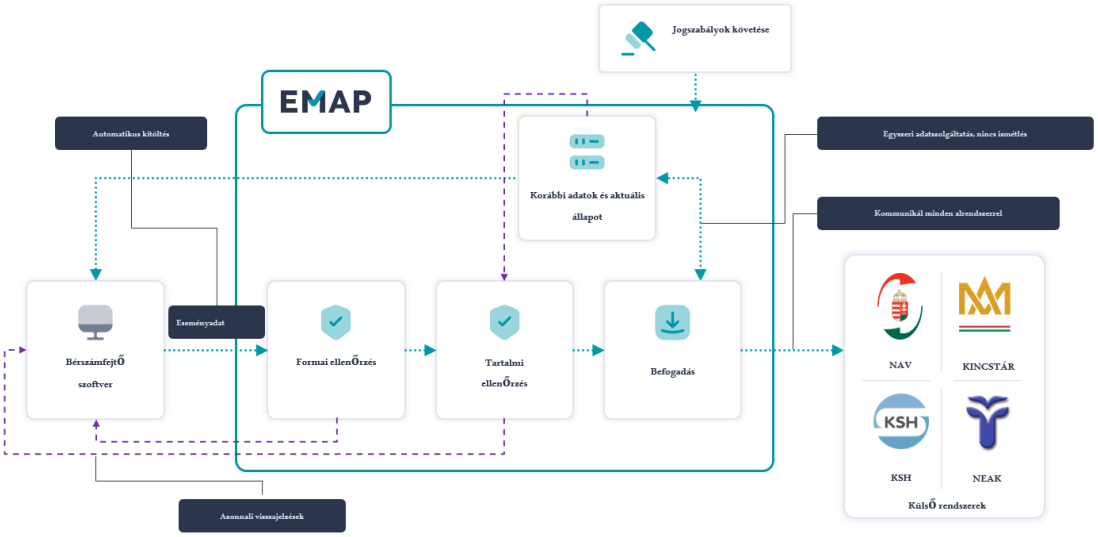
A rendszer funkciókon túlmutató szempontok mentén is (mint például a *hibatűrés* és *ellenállóképesség*) összehasonlítottuk az architektúra variánsokat. Habár mindkét megoldás képes ellátni a szükséges funkciókat, ugyanakkor vannak bizonyos szempontok, amelyek esetén egyértelműen megmutatkozik a blokklánc megoldás előnye. Tisztán funkcionális előnyként említhető, hogy a blokklánc megoldás beépítve támogatja és biztosítja az *adat integritását*, *biztonságát*, *konzisztenciáját* szervezetek között, *minőségét*, és gyors elosztását. Ugyanakkor a funkción túlmutató, perspektivikus szempont, hogy a megoldás lehetővé teszi az adatkezelés *transzparens* működését, valamint biztosítja a közös felelősséggel, jól definiált konszenzusfolyamatok mentén történő rendszer fejlesztést és üzemeltetést.

A szempontrendszer részletes kiértékelése után, valamint a fenti kiemeléseket figyelembe véve, a blokklánc alapú megoldást tartjuk a hosszabb távon perspektivikusabb, fenntarthatóbb és innovatívabb variánsnak, mivel ez gyorsabb és konzisztensebb adatelosztást tesz lehetővé és technológiailag biztosítja annak a problémának az elkerülését, hogy legyen olyan szervezet, aki minden adatot egyben lát és kezel.

1. Az EMAP rendszer magas szintű funkciói

***Vezetői összefoglaló**. A fejezet egy rövid áttekintést ad az EMAP rendszer magas szintű funkcióiról. Ezek a funkciók az architektúra variánsok kiértékelése során központi szerepet kapnak, hiszen ezek támogatása kötelező eleme bármilyen szóba kerülő megoldásnak.*

A rendszer főbb funkciói közé tartozik a szolgáltatatott adatok formai és tartalmi ellenőrzése, a felhasználók lekérdezéseinek kiszolgálása, valamint a szolgáltatott adatokból a különböző szakrendszerek által elvárt nyomtatványok előállítása.



1. ábra: Az EMAP rendszer főbb funkciói

Az 1. ábra bemutatja az EMAP rendszer funkció-központú felépítését. A szolgáltatott adatok életciklusa alapján az alábbi fő támogató rendszer funkciókat azonosíthatjuk:

- Adat validáció és automatikus kitöltés
- Lekérdezések kiszolgálása
- Nyomtatvány transzformáció

A szolgáltatott adatok azonnali formai és tartalmi validációja kiemelt előfeltétele az adatminőség biztosításának. A validáció során a rendszer a már rögzített adatok alapján bírálja el a további adatok helyességét, megelőzve ezáltal, hogy valamilyen szempontból inkonzisztens adat kerüljön a rendszerbe, elkerülve a későbbi hibakezelési folyamatokat.

A rendszer a szolgáltatott adatok alanyainak (mint például a természetes személyeknek) is képes a rájuk vonatkozó adatokkal szolgálni. Ennek megfelelően a külső lekérdezések támogatása és hatékony megvalósítása is fontos funkció.

Végezetül, a projekt egy folyamatos integráció keretében szeretné bevezetni az esemény alapú adatszolgáltatást. Ennek eszköze az úgynevezett nyomtatvány transzformáció funkció, amely a rögzített adatok alapján a szakrendszerek által jelenleg is használt nyomtatványokat előállítja, és azokat adja át az érintett szervezeteknek a közvetlen integráció helyett. A nyomtatvány alapú integráció ezen kívül a Pilot során egy önellenőrzés célú párhuzamos működést tesz lehetővé a jelenleg is működő adatszolgáltatással.

2. A rendszer által kezelt adatok

Vezetői összefoglaló. *A fejezet ismerteti a rendszer funkcióinak megvalósításához szükséges szakterületi adatok körét, valamint különböző adatmodellezési megközelítéseket azok digitális reprezentációjához.*

Az elsődleges adattípusok közé soroljuk az entitások (pl. természetes személyek), állapotuk, valamint a vonatkozó EMAP események halmazait. Ezen szakterületi adatok különböző modellezési és tárolási módjai alkotják majd az egyes rendszer architektúra variánsok közötti fő különbséget.

2.1. A kezelt adatok köre

Az EMAP rendszer egyik fontos (a technológiai javaslatot leginkább befolyásoló) aspektusa a rendszer által kezelt és mások felé megosztott, illetve előállított adatok köre. A rendszer által kezelt adatokat a következő főbb típusokba – és ennek megfelelően logikai tárakba – oszthatjuk fel forrásuk, tartalmuk és felhasználásuk szempontjából:

- **Elsődleges tárak**
  - **Entitás tár**: a rendszerben megjelenő entitások halmaza, akik lehetnek például az adatszolgáltatás alanyai (természetes személyek vagy cégek), illetve maguk az adatszolgáltatók (azaz a foglalkoztatók).
  - **Entitás állapot tár**: az azonosíthatóságon kívül az entitások különböző állapotokkal is rendelkeznek, amelyek időben változhatnak.
  - **Esemény tár**: a rendszer által definiált esemény típusoknak megfelelő, az adatszolgáltatás során keletkezett adatok halmaza.
- **Másodlagos/segéd tárak**
  - **Formanyomtatvány tár**: a rendszer által automatikusan előállított formanyomtatványok tára.
  - **Jogosultság tár**: a rendszerben részt vevő szereplők jogosultságainak tára, amelyek adatkezelési irányelvek formájában meghatározzák, hogy ki milyen adat (vagy akár irányelv) kezelésére (például módosítására) képes.
  - **Metaadat tár**: a többi tár adatainak integritását, letagadhatatlanságát és verziókezelését támogató adatok.

A fejezet további részei részletesen bemutatják az egyes adat típusok rendszerben elfoglalt helyét, kitérve a validációs feladatokra és az adatok érzékenységére.

2.1.1. Entitás tár

Az entitás tár tartalmazza a külső szakrendszerekből periodikusan becsatornázott személyi és céges azonosító adatokat, azaz egy regiszterként szolgál. A EMAP események alanyai, amennyiben vannak, ebből az adathalmazból kerülnek ki.

Az EMAP rendszer a következő alapvető validációs feladatokhoz használja a tárat:

- Természetes személyek azonosítása, azaz létezésük ellenőrzése a megadott azonosító halmaz alapján (például, TAJ szám).
- Jogi entitások (pl. foglalkoztatók) azonosítása, azaz létezésük ellenőrzése a megadott azonosító halmaz alapján (például, cégjegyzékszám).

Megjegyzendő, hogy a tár bizonyos adatai periodikusan, kötegelt formában fognak frissülni abban az esetben, ha a vonatkozó külső szakrendszer nincs felkészítve valós idejű adatszolgáltatásra. Ez a képesség minden külső szakrendszer esetén validálandó.

Ugyanakkor valószínűsíthető, hogy ritkán kerül be a szakrendszerekbe olyan entitás (és olyan időzítéssel), amire az EMAP rendszernek azonnal (tehát a következő frissítési időpont előtt) szüksége lenne események validálásához. Továbbá az eset valószínűsége hangolható (tipikusan elhanyagolhatóan alacsonyra csökkenthető) a frissítési periódusidő módosításával, figyelembe véve a külső szakrendszer képességeit.

Ha ez az eset mégis előfordul (tehát egy olyan entitásra hivatkozik egy adatszolgáltató, amely nem létezik az EMAP rendszeren belül), akkor a hiánypótlás – a külső szakrendszer képességeitől függően – azonnal megtehető egy oldalirányú lekérdezéssel. Így nem kell az esemény bejelentéssel megvárni a következő becsatornázás/frissítés időpontját, tehát a külső szakrendszer integrációja transzparens az adatszolgáltatók szempontjából.

2.1.2. Entitás állapot tár

Az entitás állapot tár tartalmazza a külső szakrendszerekből periodikusan becsatornázott személyes és céges állapotjelző adatokat, amelyek túlmutatnak az entitás létezésének igazolásán. A EMAP események bizonyos attribútumainak validációja ezen állapotjelzőktől függ.

Az EMAP rendszer számos, esemény típus függő validációs feladatokhoz használja az entitás állapot tárat:

- Családi kedvezményre való jogosultság;
- Rokkantság ténye;
- Táppénz időtartama;
- Stb.

A különböző állapot leírásokhoz kapcsolódó adatstruktúrák felépítésében a következő szabályosságok valószínűsíthetők az esetek többségében (követve az entitás-reláció modellezési megközelítéseket):

- Tartalmaznak az állapot alanyát azonosító, elsődleges attribútumokat (például, TAJ szám).
- Tartalmaznak a konkrét állapot típushoz köthető állapotjelölő, másodlagos attribútumokat (például, biztosított).

Megjegyzendő, hogy a tár bizonyos adatai periodikusan, kötegelt formában fognak frissülni abban az esetben, ha a vonatkozó külső szakrendszer nincs felkészítve valós idejű adatszolgáltatásra. Ez a képesség minden külső szakrendszer esetén validálandó. Ugyanakkor valószínűsíthető, hogy ritkán kerül be a szakrendszerekbe olyan adat (és olyan időzítéssel), amire az EMAP rendszernek azonnal (tehát a következő frissítési időpont előtt) szüksége lenne események validálásához.

Ha mégis bekövetkezne a fenti (előre nem detektálható, tehát oldalirányú bekérdezéssel nem kezelhető) eshetőség, akkor egy esemény elavult adat alapján kerülhet validálásra és elfogadásra (az elutasítás adatintegritás szempontjából nem kritikus). Ugyanakkor az így rögzített események az utólagos becsatornázás során automatikusan detektálhatóak az esemény típusok közötti validációs függőségek alapján és korrigálásra/felülvizsgálatra kijelölhetők. Ezen események választás szerint ügyintézőhöz irányíthatók, amennyiben a korrekció algoritmikusan nem leírható.

Például előfordulhat, hogy az óránkénti becsatornázás előtt fél órával beengedésre kerül egy olyan esemény, amelynek a validációjához szükséges adatok közben (az előző fél órában) megváltoztak a külső szakrendszerben, de még nem látszanak az EMAP rendszerben. A következő becsatornázáskor a változott adatok és az esemény típusok validációs kritériumai alapján eldönthető, hogy mely rögzített események lehetnek érintettek a megváltozott adatok alapján. Ezek az események vagy kijelölhetőek manuális felülvizsgálatra egy ügyintéző által, vagy akár automatikusan is elvégezhető az esemény ismételt validációja a már friss adatok alapján. A két megoldás természetesen ötvözhető is önellenőrzés céljából.

2.1.3. Esemény tár

Az esemény tár az adatszolgáltatás során létrejött események adatait tartalmazza, tehát az adatszolgáltatók/foglalkoztatók rendszeréből becsatornázott, vagy általuk közvetlenül előállított adatokat tárolja. A különböző eseménytípusokhoz kapcsolódó adatstruktúrák felépítésében a következő szabályosságok figyelhetők meg az esetek többségében:

- Tartalmaznak az esemény alanyait (ha vannak) azonosító, elsődleges attribútumokat (például, TAJ szám).
- Tartalmaznak a konkrét eseménytípushoz köthető állapotjelölő, másodlagos attribútumokat (például, kifizetés jogcíme).
- Tartalmaznak más eseményeket azonosító adatokat (azaz referenciákat) tartalmi függőségek mentén (például, vonatkozó levont járulékok esemény azonosítója).

Megjegyzendő, hogy az esemény tár és az entitás állapot tár adatainak felépítése a hagyományos adatmodellezési konvencióknak (3.1 fejezet) köszönhetően hasonló. Ennek távlati kiaknázását a 3.3 fejezet tárgyalja.

A másodlagos attribútumok validálása során a rendszer (eseménytípustól függően) ellenőrizhet bejegyzéseket az entitás állapot tárban (például: jogosultság táppénz kifizetésére). Fontos felmérni a biztosítható adatkonzisztencia mértékét minden olyan esetben, amikor a validációhoz szükséges adatok nem egy logikai tárban helyezkednek el az esemény adatokkal. Hasonló konzisztencia problémák merülhetnek fel (hasonlóan kis valószínűséggel), mint a periodikus becsatornázás miatt.

2.1.4. Formanyomtatvány tár

A rendszer az eseményadatok alapján időnként (explicit kérésre, periodikusan, vagy bizonyos feltételek teljesülése esetén automatikusan) olyan további, származtatott adatokat állít elő, amelyek felhasználhatók az egyes szervezetek által jelenleg is használt formanyomtatványok generálására. Ezt a folyamatot nyomtatványtranszformációnak nevezzük.

Fontos kiemelni, hogy ez a tár az EMAP rendszer sajátja, tehát nincs szó közvetlen integrációról az érintett szervezetek szakrendszereivel. Ezáltal lehetővé válik, hogy az érintett szervezetek saját hatáskörben eldöntsék a generált formanyomtatványok felhasználási módját. Vagy közvetlenül felhasználják a saját szakrendszereikben az eddigi adatszolgáltatási munkafolyamatok új bemeneteként, vagy önellenőrző módon összevetik a hagyományos módon szolgáltatott nyomtatványokat a generált nyomtatványokkal egy külön folyamat részeként. Ez az elkülönített feldolgozás rugalmas integrációs lehetőségeket biztosít az EMAP rendszerrel a szervezeteknek.

Az eseményadatok szigorú validációja miatt a nyomtatványtranszformációs lépés ismételt bemeneti adat validációt csak indokolt esetben igényel (például a rendszer üzemeltetésének korai fázisában, amikor hasonló többszörös ellenőrzések képesek látens hibák felderítésére). Ugyanakkor az esetlegesen későn frissülő entitás (állapot) tár változásai kihathatnak az események, és így a generált formanyomtatványok konzisztenciájára. Ilyen esetben a teljes adattranszformációs lánc felülvizsgálata szükséges az érintett események mentén.

Ez ugyanaz a jelenség, ami a külső entitás állapot adatok becsatornázásakor is előfordulhat. Az eset utólagos kezelésére ugyan az a mechanizmus használható: az új adatok által érintett nyomtatványok meghatározása és kijelölése manuális vagy automatikus felülvizsgálatra.

2.1.5. Jogosultság tár

A rendszer által kezelt adatokhoz való hozzáférés jogosultsága központi kérdés, amely jogi úton, illetve esetenkénti felhatalmazások által szabályozott. A rendszer rugalmassága (hatékony frissíthetősége) érdekében a vonatkozó irányelveket érdemes nem pusztán szoftverként kodifikált formában alkalmazni, hanem azokat adatként tárolni és interpretálni. Ez a megközelítés lehetővé teszi a jogosultságok későbbi, dinamikus frissítését a rendszerszoftver módosítása (és a rendszer leállítása) nélkül.

2.1.6. Metaadat tár

A metaadat tár nem tárol közvetlenül szakterületspecifikus adatot, hanem a többi tár adatkészletéhez nyúlt bizonyos támogató funkciókat:

- **Digitális lenyomatot** az adatintegritás támogatásához.
- **Digitális aláírást** az adattartalmak letagadhatatlanságához.
- **Időbélyegeket** az adatok adott időben történő létezésének igazolásához.
- **Verziókezelést** az adattartalmak megváltozásának követéséhez.

Az alábbi kiegészítő funkciók szakterület függetlenek, így tetszőleges (akár több) szakterületi tár kiegészítésére is használhatóak.

2.1.6.1. Digitális lenyomat

Az adatok digitális ujjlenyomata kulcsfontosságú technika az információbiztonság területén, amely megbízható eszközt kínál a digitális tartalom egyedi azonosítására és nyomon követésére különböző platformokon és hálózatokon keresztül. Ez a módszer egy fájl vagy adatkészlet kompakt, kriptográfiai ujjlenyomatának létrehozását jelenti, amely a tartalom digitális „összefoglalásaként” szolgál.

Ezek az ujjlenyomatok megváltoztathatatlanok, és megbízható módot biztosítanak az adatok integritásának és hitelességének ellenőrzésére, a jogosulatlan módosítások felderítésére és a digitális eszközök eredetének nyomon követésére. A digitális ujjlenyomatok tipikus megoldásai közé tartoznak az olyan kriptográfiai hash-algoritmusok, mint a SHA-256 és BLAKE2. Ezeket a megoldásokat széles körben alkalmazzák a tartalomkezelésben, a szerzői jogvédelemben és a kiberbiztonságban, biztosítva az adatok integritását és elszámoltathatóságát az egyre inkább összekapcsolt digitális világban.

2.1.6.2. Digitális aláírás

A digitális aláírás a modern kriptográfia sarokköve, amely biztonságos és hamisíthatatlan módszert kínál a digitális dokumentumok, tranzakciók és kommunikáció hitelesítésére. Ezek a kriptográfiai aláírások virtuális jóváhagyási pecsétként működnek, megerősítve egy üzenet vagy fájl eredetét és sértetlenségét a digitális világban. Azáltal, hogy egy privát kriptográfiai kulcs segítségével minden egyes adathoz egyedi aláírást hoz létre, az egyének és szervezetek ellenőrizhetik mind a feladó személyazonosságát, mind a tartalom megváltoztathatatlanságát.

A digitális aláírások tipikus megoldásai közé tartoznak az aszimmetrikus titkosítási algoritmusok, mint például az RSA és az elliptikus görbe kriptográfia, valamint az olyan digitális aláírási szabványok, mint a PKCS #1 és a PGP. Ezek a megoldások kulcsfontosságú szerepet játszanak az adatok hitelességének, bizalmas jellegének és letagadhatatlanságának biztosításában különböző területeken, többek között az online bankolásban és a digitális szerződéskezelésben.

2.1.6.3. Időbélyeg

Az adatok időbélyegzése a digitális világ alapvető gyakorlata, amelynek során egy ellenőrizhető időbélyeget helyeznek el egy információval kapcsolatban, amely biztosítja, hogy pontosan rögzítsék, mikor hozták létre, módosították vagy továbbították azt. Ez az időbeli metaadat döntő fontosságú az események időrendi sorrendjének megállapításához, az adatok törvényszéki vizsgálatának segítéséhez, valamint a jogi és szabályozási megfelelés megbízható alapjának biztosításához.

Az adatok időbélyegzésére szolgáló tipikus megoldások közé tartoznak megbízható harmadik felek időbélyegző szolgáltatásai, amelyek kriptográfiai technikákat használnak az aktuális idő digitális lenyomatának létrehozására, és ezzel bizonyítják, hogy egy adott dokumentum vagy adat egy adott időpontban létezett. A blokklánc-technológia szintén hatékony eszközként jelent meg a decentralizált és hamisításálló időbélyegzéshez, lehetővé téve a felhasználók számára az adatokkal kapcsolatos tevékenységek biztonságos rögzítését és időzítésének ellenőrzését. Legyen szó jogi szerződésekről, pénzügyi tranzakciókról vagy adatellenőrzésről, az időbélyegzés kulcsfontosságú szerepet játszik az adatok integritásának és elszámoltathatóságának biztosításában a digitális korban.

2.1.6.4. Verzió

Az adatok verziózása az adatkezelés és a (digitális) kollaborációk alapvető gyakorlata, amely biztosítja a digitális információban az idő múlásával bekövetkező változások szer-

vezett és szisztematikus nyomon követését. Ez magában foglalja egy adatkészlet, dokumentum vagy szoftver iterációinak/verzióinak létrehozását és karbantartását, és minden egyes verzió egy pillanatfelvételt őriz az adatok állapotáról a fejlődés egy adott pontján.

Az adatok verziókezelésének tipikus megoldásai számos eszközt és módszertant foglalnak magukban. A dokumentumkezelő rendszerek és felhőalapú tárolási platformok gyakran tartalmaznak verziókezelési funkciókat, amelyek segítségével a felhasználók nyomon követhetik a dokumentumok szerkesztéseit és revízióit. Az adatok verziózása nemcsak az együttműködést és a hibák helyreállítását segíti elő, hanem az adatok integritását és ellenőrizhetőségét is javítja a különböző területeken, a szoftverfejlesztéstől kezdve a tartalomkészítésen át a tartalom megosztásáig.

3. Adatmodellezési módszerek

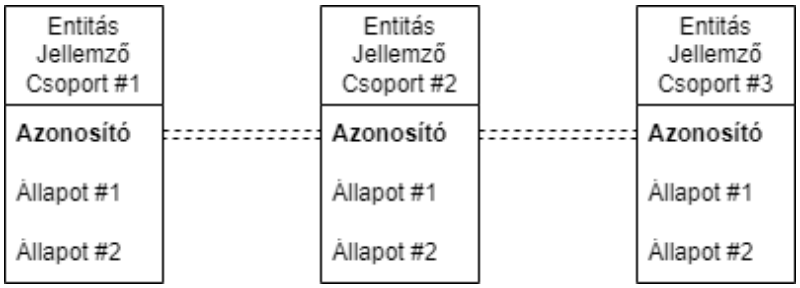
Vezetői összefoglaló. *Az alábbi alfejezetek a rendszer architektúra variánsai szempontjából kritikus, két alapvetően különböző adatmodellezési és tárolási módszert mutatnak be: a hagyományos entitás-reláció alapút centralizált rendszerekhez, és egy deidentifikációt használó megközelítést blokklánc alapú rendszerekhez.*

A deidentifikációt használó megoldás lehetővé teszi az adatok szélesebb körű megosztását és jogszabályi szempontból egyszerűbb kezelését. Végezetül egy távlati, perspektivikus adatszervezési megközelítés is bemutatásra kerül, amely segítheti a szakrendszer közötti adatkonzisztencia biztosítását.

3.1. A hagyományos, centralizált megközelítés

Egy általános, állapottal rendelkező entitást leíró adatstruktúra felépítésében a következő szabályosságok valószínűsíthetők, igazodva a hagyományos entitás-reláció modellek gyakorlatához (2. ábra):

- Tartalmaznak elsődleges, az állapotjelzők alanyait azonosító attribútumokat.
- Tartalmaznak másodlagos, konkrét állapotjelölő attribútumokat.



2. ábra: Entitás jellemzők korrelációja szakrendszereken keresztül közös azonosító által

3.1.1. Elsődleges attribútumok

Elsődleges adatnak tekinthető minden olyan azonosítót vagy relációt kódoló attribútum, amely a másodlagos attribútumokat a rendszer entitásaihoz köti. Ilyen adatok például a foglalkoztatott vagy foglalkoztató azonosítók (amelyek egy entitást azonosítanak), vagy egy jogviszony azonosító, amely közvetve azonosítja az előbbi entitásokat. Az, hogy önmagukban (másodlagos attribútumok nélkül) az azonosítók (vagy csoportjuk) érzékeny adatnak minősülnek-e, minden felhasználási esetben megvizsgálandó.

3.1.2. Másodlagos attribútumok

Az adatstruktúrák másodlagos attribútumai alatt azokat az adatokat értjük, amelyek az azonosítás helyett az entitás saját állapotát írják le és nem relációként/hivatkozásként viselkednek más entitások felé. Ilyen lehet például egy természetes személy társadalombiztosítottságának státusza, vagy az információ, hogy nyugdíjas-e.

Az állapotjelző attribútumokat tipikusan külön csoportokra (és logikai tárakra) osztva kezelik. A csoportosítás történhet szakterületi megfontolások (például külön kezelendő egy természetes személy egészségügyi és banki adata), vagy bármilyen hasznosnak bizonyuló irányelvek mentén (például az adat feldolgozási helye által meghatározva). Ez esetben minden csoport rendelkezik az entitás azonosításához szükséges attribútumokkal.

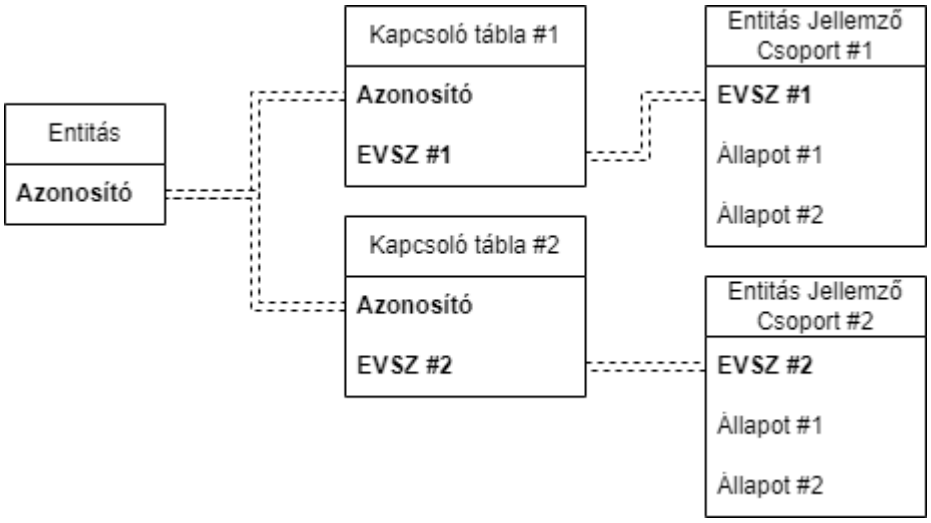
Megjegyzendő, hogy ezek az azonosítók nem szükségszerűen egyformák, az egyes tárakban használhatunk például szakterület specifikus, ugyan arra az entitásra vonatkozó, de minden csoportban eltérő azonosítókat. Megegyező azonosítók esetén a különböző adatok fúziója triviális (mint a 2. ábra), míg különböző azonosítók esetén további információ szükséges azok korrelálásához. Felhasználási eset függő, hogy melyik a kívánt megközelítés.

A másodlagos attribútumok érzékeny adatnak minősülnek, hiszen az entitás létének tényén (tehát az azonosítókon) túlmutató adatokat csatolnak az entitáshoz. A 2. ábra szerinti közös szakrendszeri azonosítók magában hordozzák annak a veszélyét, hogy minél több jellemző áll rendelkezésre egy entitásról, annál könnyebben beazonosítható az entitás még akkor is, ha a technológiai/szakterületi azonosítója közvetlenül nem fedi fel az entitás kilétét (azaz pszeudonimizált). Ez az úgynevezett „singling out”, vagy kiemelés probléma, amelynek megoldását a következő fejezet részletezi).

3.2. Az érzékeny adatok deidentifikációja blokklánc alapú tároláshoz

Az érzékeny adatok (főleg statisztikai célokra történő) megosztása során bevett módszer azok pszeudonimizálása, azaz a személyhez/entitáshoz köthető azonosítók lecserélése egy azoktól független, de (például kriptográfiai módszerekkel) egyértelműen előállítható azonosítóra. Ugyanakkor ez a megközelítés nem véd a korreláció alapú beazonosíthatóság („singling out”) problémája ellen, amely során több pszeudonimizált adatpont segítségével mégis beazonosítható az eredeti entitás annak közvetlen azonosítóinak ismerete nélkül is.[1]

[1] Például, ha tudható, hogy egy nyugdíjas, de ma is aktív színésznőről van szó, aki nemzet színésze juttatásban részesül és 1945-ben született, akkor ezekből kitalálható, hogy Molnár Piroskáról van szó.



3. ábra: Az érzékeny adatok kapcsolótáblákon keresztüli deidentifikálása

Erre az esetre nyújt megoldást a 3. ábra bemutatott teljeskörű deidentifikálás, amely a pszeudonimizálás lépést minden egyes adatcsoportra úgy hajtja végre, hogy azok mindig különböző „azonosítót” kapnak (például egyedi véletlen számok, EVSZ, formájában), akkor is, ha amúgy ugyan arra az entitásra vonatkoznak. Tehát az entitások azonosítóját és állapot attribútumait azután egy kapcsolótábla fogja egymáshoz rendelni.

Innentől kezdve a korreláció nem lehetséges pusztán az adatcsoportok alapján, hiszen még azt sem eldönthető, hogy két adatcsoport ugyanarra az entitásra vonatkozik-e. Szükség esetén az adatcsoport kisebb darabokra is partícionálható, egészen akár addig a szintig, hogy minden attribútumhoz tartozik egy saját kapcsolótábla. Mivel az adatok innentől kezdve nem köthetőek például személyhez, nem minősülnek már érzékeny adatnak, így tárolásukra is megengedőbb feltételek vonatkoznak.

A megoldás további előnye, hogy az entitáshoz köthető adatok innentől kezdve szelektíven anonimizálhatóak a megfelelő kapcsolótáblák törlésével. Ez a lépés továbbra is meghagyja az anonimizált adatokat statisztikai célú feldolgozásra, tehát csak az entitással való kapcsolata, azaz beazonosíthatósága törlődik.

Ezen a ponton az érzékeny adatok blokklánc alapú tárolásának két nagy problémáját sikerült kiküszöbölni:

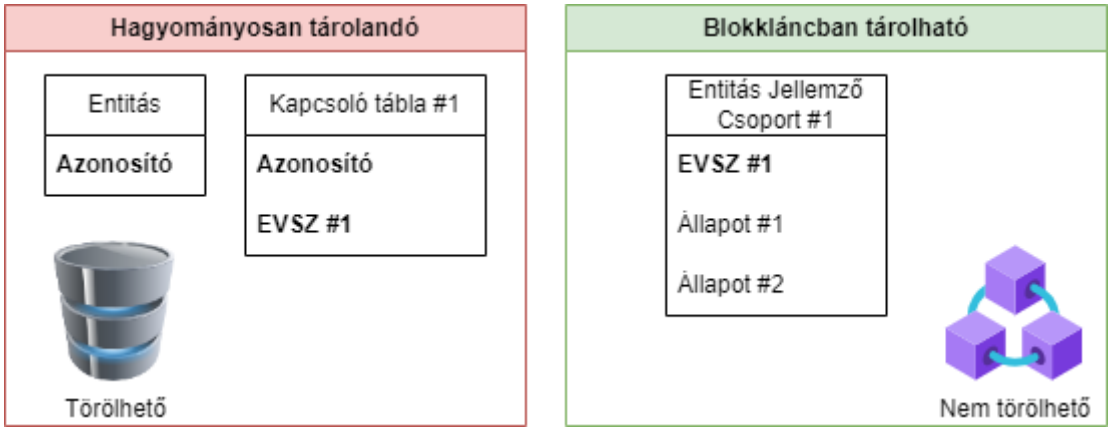
- Az állapot adatok teljesen deidentifikáltak, így azok nyugodtan tárolhatók egy több résztvevő között megosztott tárban (aminek az elérése továbbra is jogosultságkezelt, tehát nem mindenki számára elérhető) anélkül, hogy minden szervezetnek adatkezelési felhatalmazást kellene adni.

- Az állapot adat „törölhetővé” (vagyis anonimizálhatóvá, azaz a GDPR hatálya alá nem tartozóvá) válik a módosíthatatlan blokkláncból úgy, hogy a láncon kívül tárolt, az adatot entitáshoz kötő kapcsolótábla rekordot megsemmisítjük, ezáltal a blokkláncon tárolt állapot adatok visszamenőleges beazonosíthatósága is megszűnik, eleget téve egy törlés művelet adatalany oldali értelmezésének.

Fontos megjegyezni, hogy a kapcsolótáblák a deidentifikáló módszer (és a re-identifikáló lépés) központi elemei, így biztonságos tárolásuk és védelmük kiemelt fontosságú! Az adatvédelmi incidensek érzékeny adatok felfedését és/vagy adatvesztést eredményezhetnek.

A 4. ábra bemutatja a deidentifikált adathalmazok esetében tárolandó adatrészeket és a vonatkozó tárolási lehetőségeket és elvárásokat. A megfelelően deidentifikált adatok esetén – az egyszerű adatmegosztás és adatkonzisztencia biztosítása miatt – azok blokklánc alapú tárolása megengedett anélkül, hogy minden résztvevő szervezetnek adatkezelési jogosultságot kellene adnia. Az azonosíthatóságot lehetővé tevő érzékeny adatok esetén azonban egy maradéktalan felejtésre képes technológia használata a megkövetelt.

Megvizsgálandó a 4.1.2.2 alfejezetben felvetett kérdés, miszerint pusztán az entitások létét igazoló, „telefonkönyv jellegű” adatok (például TAJ szám) valóban érzékeny adatnak minősülnek-e önmagukban. A választól függően a 4. ábra Entitásai akár a blokklánc oldalra is kerülhetnének. Ugyanakkor, ez a különbség nem befolyásolja az alfejezet további tartalmát.



4. ábra: Az adatok technológia független tárolási kényszerei és lehetőségei

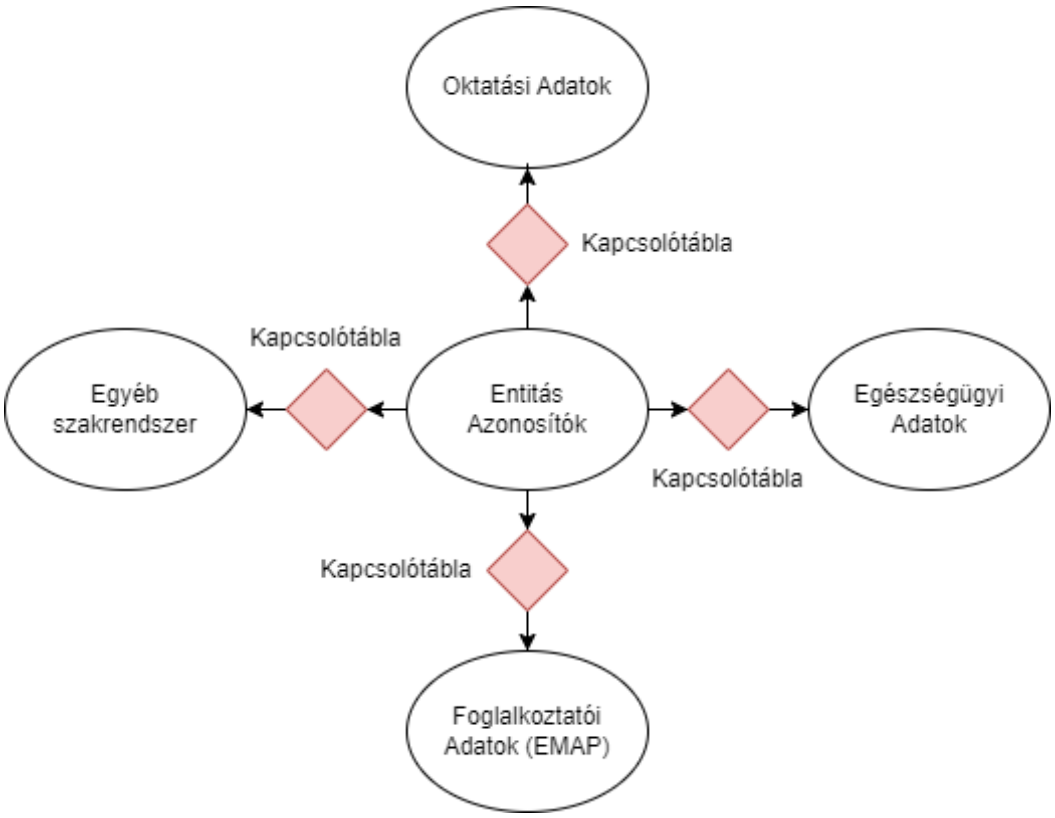
3.3. Egy hiteles entitás tár ökoszisztéma irányába

Ha a különböző szakrendszerek adatainak (az EMAP projekttől független) egészét tekintjük, akkor az 5. ábra látható csillag szervezési struktúra hosszú távon perspektivikus megoldást nyújthat a dedikált szakrendszerek adatainak egyszerűbb, lazább, és konzisztenciát

biztosító integrációjára. A csillag struktúra közepén a létező entitások (azonosítókon keresztüli) felsorolásáért felelős regiszter található.

Ezek az entitások határozzák meg szinte az összes szakterületi rendszer adatainak alanyait (például a természetes személyeket). Jelenleg ezeket az entitásokat (és a vonatkozó alap adatokat) a különböző szakrendszerek saját hatáskörben, többszörözve tárolják, felvetve a rendszerek közötti adatkonzisztencia kérdését. Egy, az adathitelesség szempontjából (tehát technológiailag nem feltétlenül) központosított entitás regiszter lehetőséget biztosítana a szakrendszerekben található redundancia csökkentésére és ezáltal a konzisztencia növelésére.

Az adatok érzékenysége szempontjából az ábrán pirossal jelölt kapcsolótáblák (amelyek a szakrendszeri adatok közötti korrelációt teremtik meg) kritikusak. Ezek megfelelő kezelése (például tárolása) lényeges kérdés, hiszen akár az entitás állapotok statisztikai alapú korrelációját is lehetővé teszik. Ennek elkerülésére a 3.2 alfejezetben bemutatott deidentifikációs módszer használható, amely akár a blokklánc alapú tárolással is kompatibilis.



5. ábra: Központosított, hiteles entitás tár a szakrendszerek támogatásához

4. Az architektúra variánsok feltérképezése

Vezetői összefoglaló. *A fejezet összegezi a teljes rendszerre kiterjedő megvalósítási lehetőségeket az egyes komponensek megvalósítási módja alapján. A variánsok azonosításában az entitás állapot és esemény adatok tárolási megoldásai voltak a meghatározó tényezők.*

Attól függően, hogy a két adattár centralizált vagy blokklánc alapon kerül megvalósításra, négy potenciális architektúra kerülhet szóba. Azonban fejlesztési és üzemeltetési szempontok mentén ez kettőre redukálódik: egy tisztán centralizált és egy tisztán blokklánc alapú megoldásra.

A fejezet végezetül bemutatja a Hyperledger Fabric blokklánc technológiát, amellyel a fejlesztési és üzemeltetési komplexitás csökkenthető a rendszer homogén technológián való megvalósítása által.

4.1. A főbb rendszer modulok és megvalósítási lehetőségeik

A rendszer architektúra kialakításának egyik fő szempontja a teljes központosítás elkerülése, azaz egy olyan szereplő kijelölése, aki minden adat birtokában van. Ugyanezen megfontolás mentén a skála másik oldala, azaz a „mindenki mindent lát” megoldás is kerülendő. Hangsúlyozandó továbbá, hogy a rendszer üzemeltetésében/használatában potenciálisan több, egymástól különálló szervezet vesz majd részt, így a közös felelősségvállalás lehetőségét is meg kell vizsgálni.

Az alábbi alfejezetek a rendszer által kezelt adatok (2. fejezet) típusai mentén áttekintik a rendszer főbb komponenseit és azok megvalósítási lehetőségeit centralizáltság szempontjából. Az egyes komponens szintű megvalósítások az adatkezelési kérdések szempontjából is összehasonlításra kerülnek, ahol releváns.

4.1.1. Egycsatornás végpont

Az EMAP rendszer egyik célja, hogy az adatok bejelentését (amellett, hogy elemi eseményekre bontja őket), illetve a lekérdezések és konzisztenciavizsgálatok eredményéről való visszajelzéseket egyetlen (de kétirányú!) csatornán keresztül tegye lehetővé a foglalkoztatók számára. Ennek megfelelően a bejelentők felé (amik akár szoftverek is lehetnek) egyetlen szolgáltatásként (azaz web tartományként) kell látszania a rendszernek, függetlenül a mögöttes megvalósítástól.

A tisztán végponti funkciók közül talán az adatszolgáltatók azonosítását (azaz bejelentkeztetését) célszerű megemlíteni, mint legfontosabbat. Az azonosítási mechanizmus konkrét módja eltérhet az egyes végfelhasználói megvalósítások esetén. A végpont egy

másik hasonlóan fontos funkciója a beérkező kérések irányítása a rendszer megfelelő komponensei felé.

Mivel az EMAP végponti szolgáltatások csak minimális saját állapottal rendelkeznek, így a centralizált (de a magas rendelkezésre állás miatt több példányban üzemelő) és a szervezetek között elosztott telepítés leginkább a közös felelősségvállalás szempontjából különbözik. A centralizált megoldás egyetlen szervezet üzemeltetése alatt futtatja az adott modult több példányban. Az elosztott verzió esetében minden szervezet futtat egy vagy több példányt az adott modulból, felelve a saját példányok megfelelő üzemeltetéséért. A megoldások adatfüggetlen mivolta miatt ez a modul a későbbiekben nem játszik szerepet a teljes architektúra variánsok kiértékelése során.

A választott megvalósítástól függetlenül egy központi kezelésű tartomány név (weboldal cím) fogja azonosítani a teljes szolgáltatást, így a rendszeren belüli tervezői döntések teljesen transzparensek a szolgáltatást használóknak (és mivel a felhasználók kizárólag ezen a végponton keresztül érik el a rendszert, így ez az állítás a többi komponensre is igaz).

4.1.2. Entitás adatok

Az entitás azonosító és állapot adatok több eseménytípus validációja szempontjából is központi szerepet kapnak. Ugyanakkor a két tár érzékeny személyes adatokat tartalmaz, amelyek hozzáférését, illetve fizikai elhelyezkedését/tárolását szigorúan felügyelni kell.

4.1.2.1. Centralizált megoldás

Az adatszervezés szempontjából legegyszerűbb megoldás egy centralizált megvalósítás, amely során egyetlen kijelölt szervezet felelős az érzékeny entitás adatok menedzseléséért. A szervezet egy központosított (például szerepalapú) hozzáférés szabályozáson keresztül bocsátja a szükséges adatokat a további érdekelt felek rendelkezésére.

Ebben az esetben az adatok törlése/felejtése is triviálisan megtehető a centralizált adatbázisból való törléssel. Megjegyzendő, hogy ha a szóban forgó adatok hiteles forrása esetleg egy másik szakrendszer, akkor pusztán az EMAP rendszerből való törlés nem elégséges (de továbbra is szükséges). Ilyenkor megvizsgálandó, hogy nem célszerűbb-e közvetlenül a külső szakrendszer integrálása annak (részleges) duplikálása helyett.

A megoldás nyilvánvaló hátránya az, hogy egyetlen szervezetnél összpontosul a felelősség az entitás adatok kezeléséért. Ugyanakkor, érdemes felmérni a párhuzamosan futó digitalizációs projekteket (mint például a DIMOP), ugyanis a jövőbeli szinergia érdekében érdemes olyan architektúra variánst választani, amely rövid távon ugyan nem feltétlenül optimális, ugyanakkor megkönnyíti a későbbi integrációt az új digitalizált szakrendszerekkel.

4.1.2.2. Blokklánc alapú megoldás

Ha megosztott felelősségű adatkezelést szeretnénk a szervezetek között (elkerülve a központosított megoldásokat), akkor a megoldásba célszerű a blokklánc technológia integrálása. A 3.2 alfejezetben bemutatott módszer alapján deidentifikált (és korreláció rezisztens) entitás állapotok egy az egyben kerülnek blokkláncra, így azon keresztül megosztásra a rendszerben résztvevő szervezetekkel.

Az entitás azonosítók és a kapcsolódó re-idenfitikációs információk viszont a láncon kívül kerülnek tárolásra, hiszen érzékeny adatnak minősülnek. Ennek megfelelően nem oszthatóak meg szabadon bárkivel, illetve a maradéktalan törölhetőségüket is garantálni kell. Emiatt hagyományos adatbázis technológiák alkalmazása szükséges a tárolásukhoz, minden szervezet esetén saját hatáskörben. Így csak azoknál a szervezeteknél létezik fizikailag is az adat, akiknek joga van a kapcsolótáblákon keresztül re-identifikálni azt.

Megvizsgálandó, hogy valóban szükséges-e eleget tenni a felejtés jogának minden adat esetén, vagy bizonyos (például a létezést hitelesítő) adatok önmagukban nem esnek ezen igény alá.

4.1.3. EMAP esemény adatok

Az esemény adatok érzékeny elsődleges és másodlagos attribútumai között hasonló relációk állnak fenn, mint az entitás azonosítók és állapotok között. Az eseményben található azonosítók ráadásul többnyire az entitás azonosítók közül kerülnek ki (ezáltal az esemény alanyaira hivatkozva). Tehát az előző alfejezet érveléseit követve itt is a következő kettő variáns képzelhető el:

- Teljesen centralizált megoldás egy kijelölt szervezet felügyelete alatt.
- Blokklánc alapú megoldás, ahol az érzékeny kapcsolótábla adatokat szervezetenként egy hagyományos adatbázis kezeli, a deidentifikált adatok pedig blokklánc segítségével kerülnek megosztásra a résztvevők között.

4.1.4. Formanyomtatvány tár

A formanyomtatvány tár azokat az adatokat tárolja, amelyeket az EMAP rendszer a bejelentett események alapján készít el. Ezen származtatott (például időszakosan aggregált) adatok tárolására a legkézenfekvőbb módszer, ha minden szervezet saját hatáskörben tárolja és kezeli azokat az EMAP rendszer keretében.

Ismét megjegyzendő, hogy ez a tár még nem a szervezet szakrendszerével való közvetlen integrációt jelenti, hanem az EMAP rendszer fennhatósága alatt lévő belső tár. A nyom-

tatvány tár és a szervezeti rendszerek közötti laza csatolásnak köszönhetően minden szervezet a saját igényei szerint, fokozatosan térhet át a generált nyomtatványok feldolgozására.

Mivel a nyomtatvány alapú adatszolgáltatás az adaptáció megkönnyítése miatt kerül átmenetileg megtartásra, hosszútávú perspektíva valószínűsíthetőleg nincs egy elosztott megoldás kidolgozásában:

- A felmerülő eseti nyomtatvány megosztás igények pont-pont jellegű, vagy központosított integrációval is kezelhetőek a szervezetek között.
- A nyomtatványok bármikor újragenerálhatóak a rendszerben tárolt adatok alapján (így a tár leginkább csak egy gyorsítótár funkcióját látja el).
- A nyomtatványt valószínűsíthetőleg csak addig kell tárolni, ameddig a létező szakrendszer vissza nem igazolja annak befogadását feldolgozásra.

A fenti megfontolások miatt ez a modul a későbbiekben nem játszik szerepet a teljes architektúra variánsok kiértékelése során. Megjegyzendő, hogy ha valamelyik másik tár esetén is blokklánc alapú megoldás kerül implementálásra, akkor alacsonyabb költséggel elvégezhető a nyomtatványok tárolásának és feldolgozásának blokklánc alapú megoldása is.

4.1.5. Jogosultság tár

Mivel a jogosultság tár közvetlenül nem érintett a hozzáférésszabályozott adatok tartalmának tárolásában, így az adat érzékenység probléma ebben az esetben nem központi kérdés. Továbbá, egy ilyen jellegű tár akár a szakterületi adatok tartalmától függetlenül is megvalósítható, azaz rendszerek között újrahasználható, vagy egy már létező megoldás integrálható. Ennek megfelelően ez a modul a későbbiekben nem játszik szerepet a teljes architektúra variánsok kiértékelése során. Javasolt a rendszer egyéb komponenseihez használt technológiai/szervezési megoldásokhoz igazítani a tár kialakítását.

Megjegyzendő, hogy létező blokklánc komponens esetén egy újabb blokklánc komponens megvalósításának és működtetésének inkrementális költségét elhanyagolhatónak tekintjük a rendszer teljes komplexitásához képest. Ez esetben érdemes megvizsgálni egy decentralizált (tehát transzparens, ellenőrizhető) jogosultságkezelő komponens lehetőségét, amely akár különböző szakrendszereket is támogathat egyszerre (ahogy az 5. ábra szervezése is szemlélteti).

4.1.6. Metaadat tár

Mivel a metaadat tár szakterület független, és az alap adat nélkül nem interpretálható adatokat tartalmaz, így az adat érzékenység probléma ebben az esetben nem központi

kérdés. Továbbá, egy ilyen jellegű metaadat tár az EMAP projekttől függetlenül is használható, azaz tetszőleges szakterület adatának leírása felcsatolható rá.

Habár nincs akadálya, hogy egy ilyen rendszer egy központi *megbízott* fél irányítása alatt kerüljön megvalósításra, a blokklánc megoldás több szempontból is kézenfekvőbb, hatékonyabb, és perspektivikusabb megoldás:

- A metaadatok alapján az eredeti adat nem azonosítható, így az adatmegosztás problémája egy (blokkláncokon) triviális adatreplikációvá egyszerűsödik.
- A metaadatok új féllel történő megosztásához elég hozzáadni a felet a blokklánc hálózathoz, az adat szinkronizációja onnantól kezdve automatikus.
- Tetszőleges szektor szervezete becsatolható a blokklánc hálózatba, segítve ezáltal a szektorok közötti, láncon kívüli adatcserék ellenőrizhetőségét a láncon tárolt metaadatok segítségével. Például, egy biztosító és egy egészségügyi szervezet közötti, közvetlen adatcsere után a biztosító képes ellenőrizni a kapott adatok különböző tulajdonságait, amelyeket a 2.1.6 fejezet mutatott be. Ezt a metaadat alapú működési modellt támogatja az EBSI (European Blockchain Services Infrastructure) rendszer is.
- A blokklánc alapú metaadat kezelés és megosztás EU szinten perspektivikus irány, amelynek úttörő megvalósítása az EBSI projekt.

A fenti megfontolások miatt az architektúra variánsok kiértékelése során a modult blokklánc alapúnak vesszük, és csak ott térünk ki rá külön, ahogy egyéb blokklánc alapú komponens nem szerepel a rendszerben. Létező blokklánc komponens esetén ugyanis egy metaadat tár – mint újabb blokklánc komponens – megvalósításának inkrementális költségét elhanyagolhatónak tekintjük a rendszer teljes komplexitásához képest.

4.2. Rendszer architektúra variánsok

Az 1. táblázat összefoglalja a lehetséges modul megvalósítási kombinációkat, amelyek a teljes rendszer architektúra variációit eredményezik. Az egyes tárak vagy centralizált módon (C), vagy blokklánc alapon (B), a szükséges deidentifikációs módszerrel (3.2 fejezet) kerülhetnek megvalósításra.

		Esemény tár	
		Centralizált	*Blokklánc*
Entitás tár	*Centralizált*	CC (+ B meta)	CB
	Blokklánc	BC	BB

1. táblázat: A rendszer architektúra variánsok szisztematikus áttekintése

A metaadat tárat a 4.1.6 alfejezetben leírtaknak megfelelően érdemes mindenképpen blokklánc alapon megvalósítani, többek között a jövőbeli digitalizációs irányokkal való szinergia miatt (6. fejezet). Ennek megfelelően a metaadat tár technológiai megvalósítása nem játszik közre a variánsok közötti különbségtételben (csak a tár léte).

Megjegyzendő, hogy ha több tár is blokklánc alapon kerül megvalósításra, akkor a láncok akár ugyanazon technológia és infrastruktúra felett is menedzselhetők. Mivel a metaadat tár blokklánc alapú, így a további blokklánc megvalósítások implementációs és üzemeltetési ráfordításai inkrementális jellegűek, azaz jelentősen alacsonyabbak.

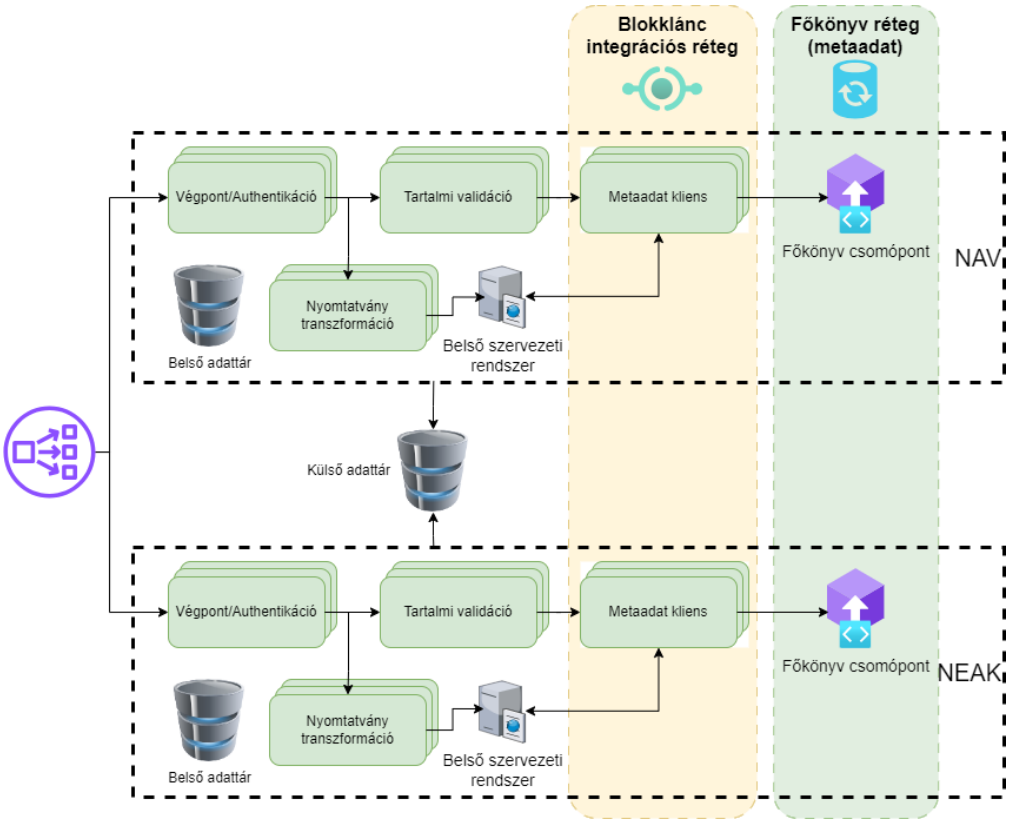
A 1. táblázatban felsorolt variánsok számát csökkenthetjük az alábbi megfontolások és eliminációk segítségével:

- **BC**: Ha az egyik tár kapcsán már megtervezésre és megvalósításra kerül egy elosztott, blokklánc alapú érzékeny adat tárolás megoldás, akkor ezt érdemes alkalmazni (a funkcionalitás szintjén újrahasználni) a másik tárra is egy centralizált megoldás helyett a költséghatékonyság miatt.
- **CB**: A fenti pont alól kivételt képezhet az az eset, amikor egy létező (centralizált) digitalizációs rendszerrel való integráció stratégiai kérdés.

A reprezentatív kombinációkat a következő tervezői döntés által meghatározott tengely mentén tudjuk felsorolni: az adatok mekkora része kerüljön blokkláncra az (különböző funkciókkal támogatott) adatmegosztás érdekében. A fennmaradó kombinációkat ennek megfelelően a centralizált és decentralizált pólusok mentén oszthatjuk két csoportba.

4.2.1. Centralizált rendszer architektúra

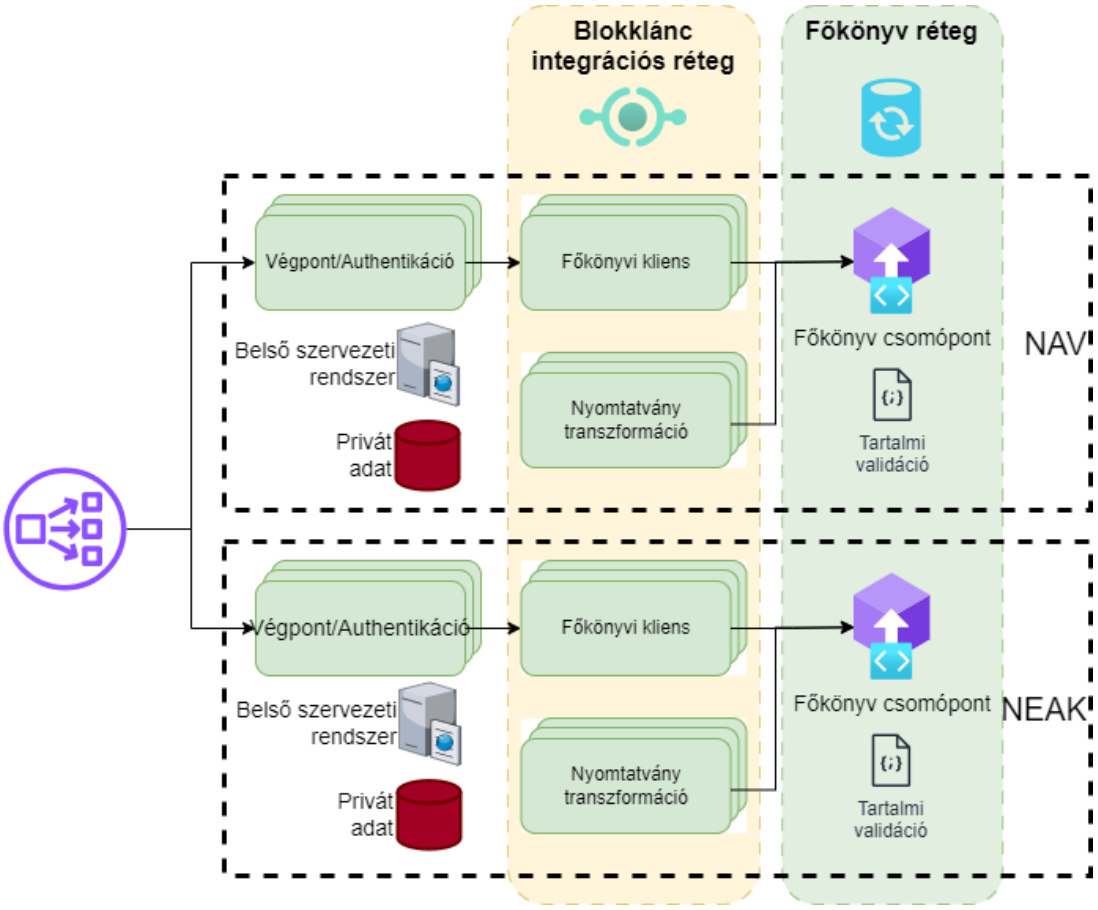
Ebben az esetben a rendszer központi funkciói centralizált szolgáltatásokra támaszkodva kerülnek kialakításra, hagyományos fejlesztési és üzemeltetési módszereket használva. A metaadatokat érdemes blokkláncra helyezni (4.1.6 fejezet), lehetővé téve a szakterületek közötti centralizált adatcsere esetén az elosztott bizalmat és a retrospektív adat ellenőrzést az érintett felek számára. A sematikus architektúra tervet az 6. ábra szemlélteti.



6. ábra: Egy centralizált architektúra blokklánc alapú metaadat támogatással

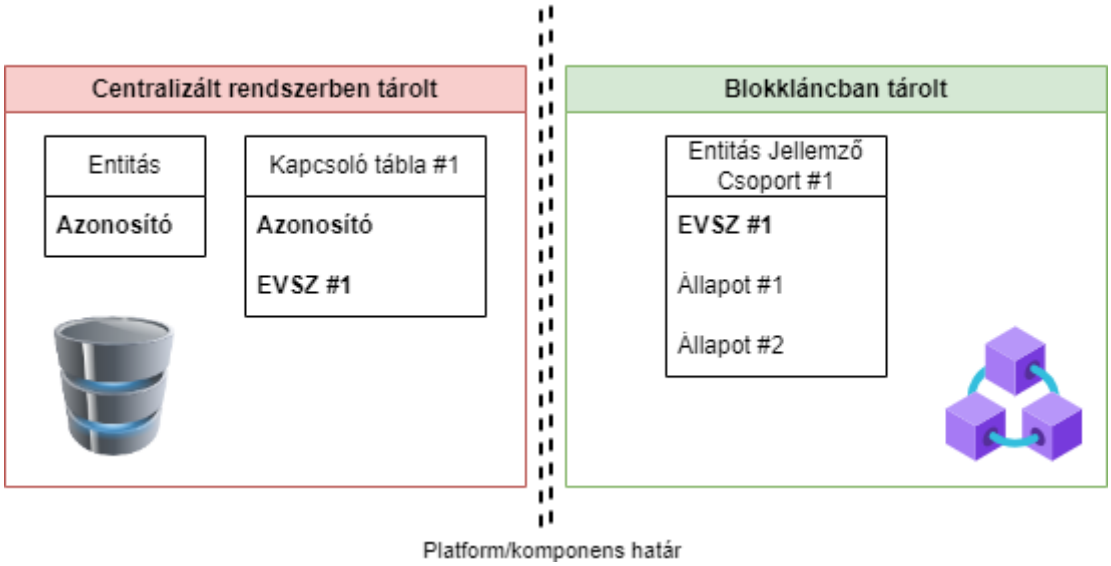
4.2.2. Decentralizált rendszer architektúra

A decentralizált esetben az entitásoktól leválasztott, deidentifikált entitás állapotok blokkláncon kerülnek tárolásra (7. ábra). A re-identifikáláshoz szükséges kapcsolótábla adatokat továbbra is hagyományos adatbázisban kell tárolni (aminek egy technológiailag homogén megoldását mutatja be a 4.3 fejezet).



7. ábra: Egy blokklánc alapú architektúra centralizált kapcsolótábla tárolással

Az 8. ábra egy, a fenti adattárolási követelményeknek eleget tevő megoldást mutat be. Az érzékeny adatok tárolását a blokklánc platformtól független technológia, egy elkülönülő centralizált adatbázis végzi. A hagyományos adatbázis lehetővé teszi a felejtéshez való jog implementációját, eleget téve ezáltal a blokkláncok által (közvetlenül) nem támogatott kritikus adatkezelési szabályozásnak.

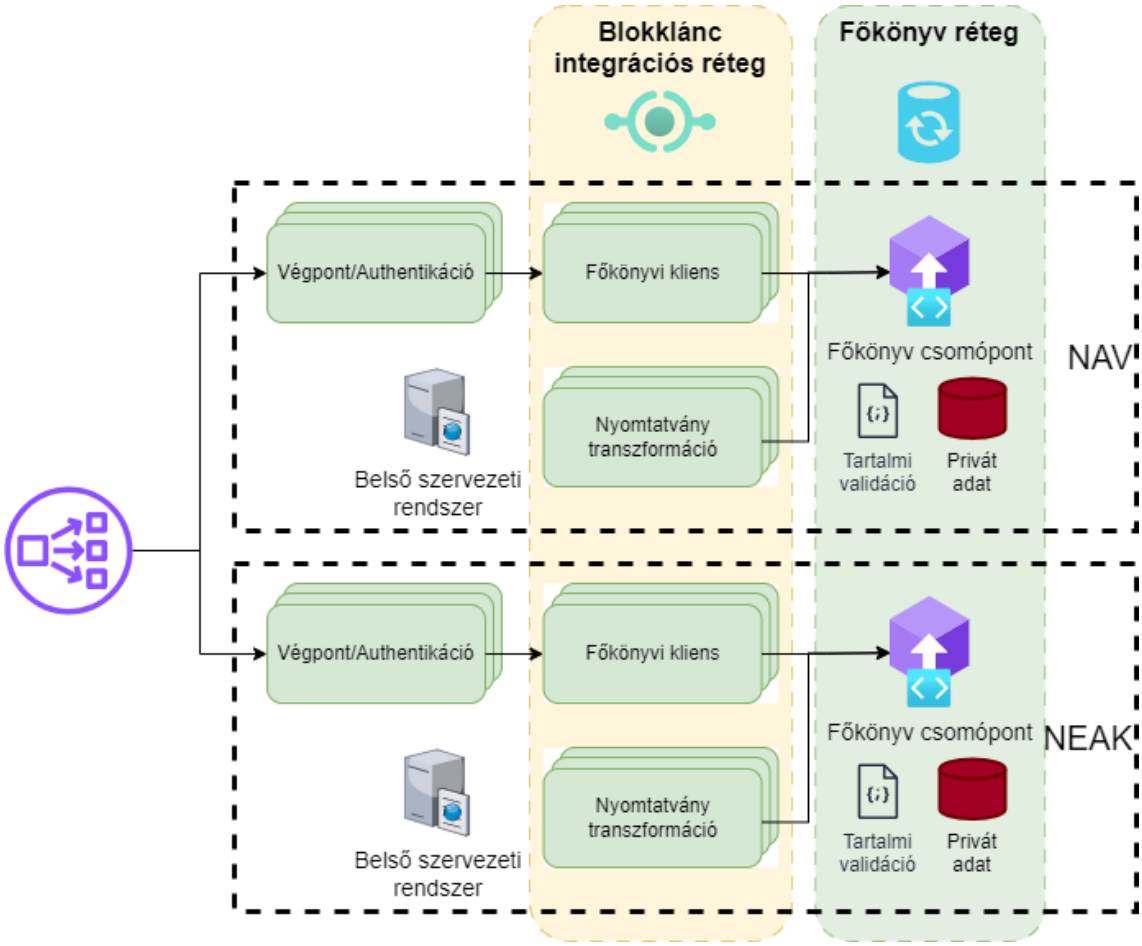


8. ábra: Érzékeny adatok tárolása egy hibrid centralizált-decentralizált architektúrával

4.3. Homogén blokklánc megoldás Hyperledger Fabric alapon

A bemutatott blokklánc alapú architektúra tovább egyszerűsíthető fejlesztési és üzemeltetési szempontból is, ha blokklánc technológiaként a Hyperledger Fabric platform kerül kiválasztásra. A Fabric egy nyílt forráskódú blokkláncplatform, amelyet vállalati alkalmazásokhoz terveztek. Ennek megfelelően privát és jogosultságkezelt, azaz a blokkláncot olvasni és módosítani tudó résztvevők köre szigorú szabályok mentén definiált és felügyelt (ellentétben a világméretű, kriptopénzt kezelő láncokkal).

A platform moduláris, skálázható és az adatvédelemre összpontosít. A Fabric programozható okosszerződéseivel és beépített adatvédelmi módszereivel olyan iparágak számára is alkalmas, mint a pénzügy, az egészségügy és az ellátási lánc. Továbbá nagy hangsúlyt fektet az engedélyezett hozzáférésre, az identitás-kezelésre és a robusztus, decentralizált irányításra a résztvevő szervezetek között.



9. ábra: Egy homogén, Hyperledger Fabric alapú architektúra

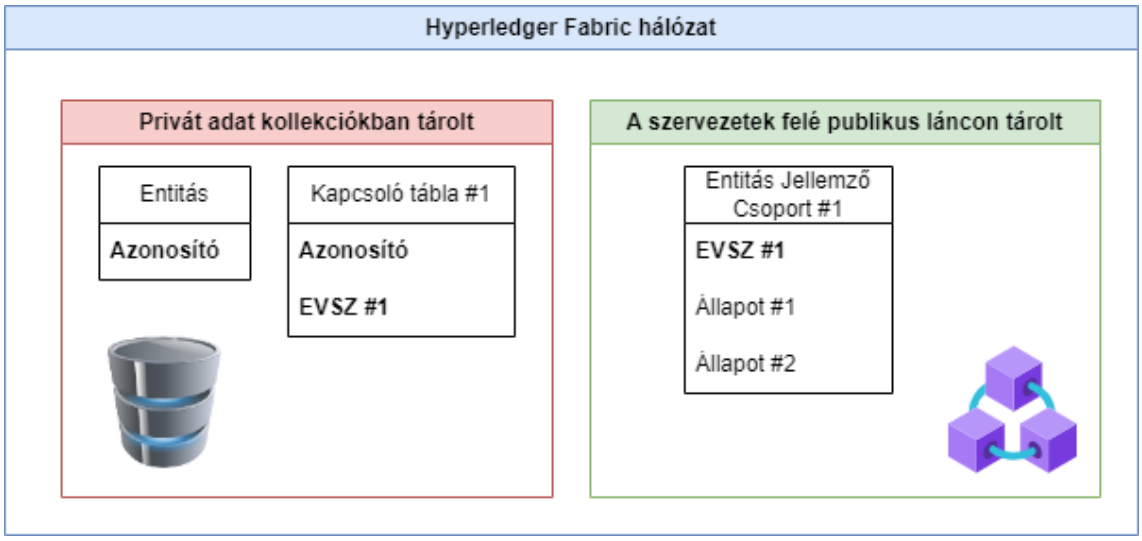
A teljesen blokklánc alapú, de továbbra is titkosság megőrző megoldás támaszkodik a Fabric privát adat kollekció képességére (9. ábra). A privát adat kollekció a Fabric olyan funkciója, amely lehetővé teszi az érzékeny adatok szelektív megosztását egy blokklánc-hálózaton belül, miközben az adatokat bizalmasan és a nyilvános főkönyvön kívül tartja. Amikor egy tranzakciót benyújtanak a blokkláncra, az tartalmazhat nyilvános és privát adatokat is. A privát adatokat, amelyeknek bizalmasnak kell maradniuk, az érintett szervezetek elkülönítik a nyilvános adatoktól, azok nem szerepelnek a blokklánc főkönyvében tárolt tranzakció adatai között.

Minden olyan szereplő, amely egy privát adat kollekció része, saját privát adattárat tart fenn, amely a szervezetek számára publikus főkönyvtől elkülönül, így biztosítva a titkosságot. A hozzáférés szabályozási irányelvek az okosszerződésben vannak elkódolva annak meghatározására, hogy ki olvashatja vagy írhatja a privát adatokat. A tranzakciók publikus eredményében csak a privát adatok kriptográfiai lenyomatai (hash) szerepelnek.

Továbbá a privát adatok fizikai elosztását (kiküldését) a jogosult szervezetek között a platform automatikusan elvégzi, konfigurálható adat rendelkezésreállás feltételek mentén. Mielőtt az adatokat a privát adatgyűjteményekben tárolnánk, az érzékeny információk védelme érdekében azok szükség esetén tovább rejthetők (például titkosítási módszereket használva).

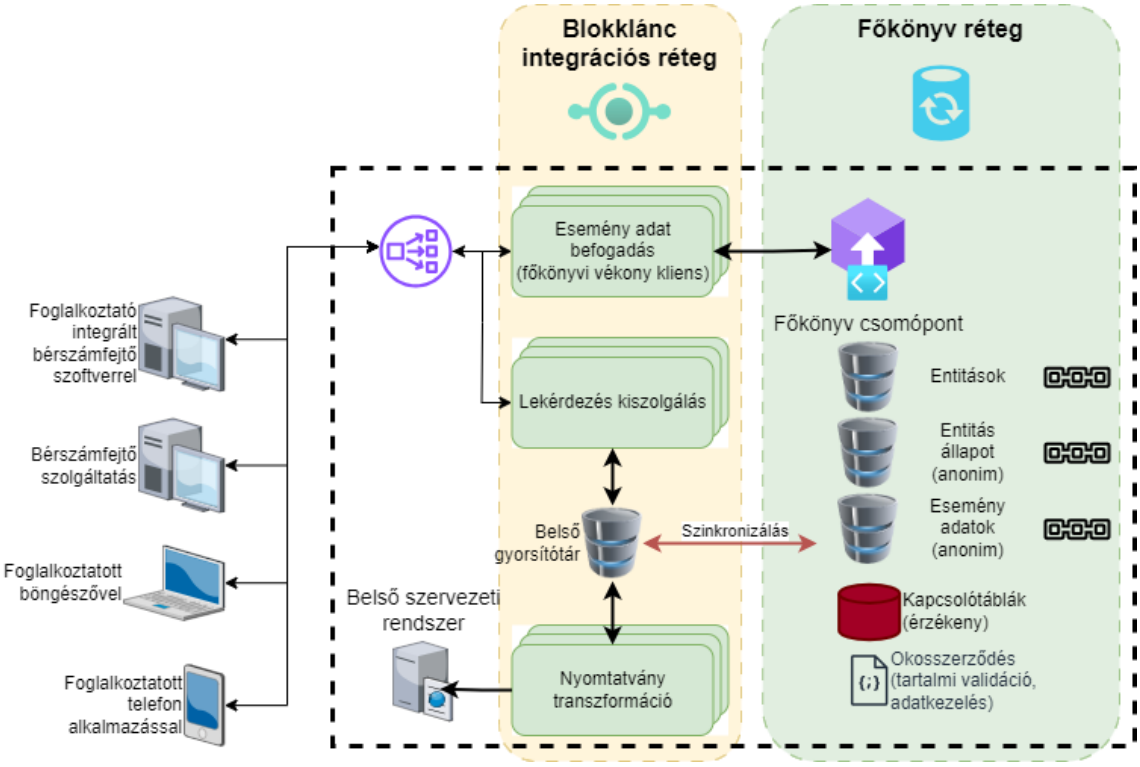
A privát adatkollekciók további fontos tulajdonsága az, hogy a privát adatok módosításai is a Fabric konszenzus protokoll hatásköre alá esnek, azaz megegyezést igényel az adathoz hozzáférő felek részéről. Továbbá, a privát adatoknak csak a digitális lenyomata tárolódik blokklánc struktúraként, maga a nyers adat hagyományos adatbázis módszerekkel kerül tárolásra. Ez azt jelenti, hogy az adat törlése/felejtése és mozgatása is megtehető úgy, hogy a publikus láncon ne maradjon dekódolható nyoma, ugyanakkor maga a művelet megtörténte (és utólag akár a pontos tartalma is) bizonyítható legyen.

A 10. ábra felfedi a privát adat kollekciók működésének egy fontos implementációs részletét. A Hyperledger Fabric platform a privát adatokat egy arra a célra elkülönített, a szervezeti csomópontok szintjén centralizált, hagyományos adatbázis technológiával tárolja. Továbbá az okosszerződés programozási interfészen keresztül lehetőséget biztosít a kollekcióban tárolt adatok teljeskörű törlésére (aktuális, és historikus verziókkal együtt).



10. ábra: Érzékeny adatok megfelelő tárolása tisztán a Hyperledger Fabric platformmal

Az 8. ábra és 10. ábra összevetve megállapíthatjuk, hogy a két megoldás csak a magas szintű leírás szintjén jelent lényegi különbséget, technológiai szinten megfeleltethető egymásnak. Továbbá, a tisztán blokklánc alapú megoldás (amelynek részletesebb technológiai kibontását a 11. ábra mutatja) egy homogén technológiai megoldást jelent, amely konszolidálja a fejlesztési és üzemeltetési feladatokat.



11. ábra: A Hyperledger Fabric alapú architektúra részletesebb technológiai kifejtése

4.4. Hatékony jogosultságkezelési irányelvek kialakítása

A blokklánc technológiák talán leglényegesebb tulajdonsága – amely megkülönbözteti őket a hagyományos, centralizált rendszerektől – az, hogy a rendszer irányítási feladatai nem egy központilag kijelölt szereplő kezében összpontosulnak, hanem a hálózat résztvevői akár egyenrangú felenként, többségi alapon is hozhatnak döntéseket a rendszert érintő változásokról (például új résztvevő felvétele a hálózatba).

Fabric esetén a rendszerirányítási kérdések több aspektusa is részletesen testreszabható. Egymástól függetlenül szabályozható például, hogy kiknek a jóváhagyása szükséges például egy-egy tranzakció validációjához és befogadásához, okosszerződés telepítéséhez, vagy résztvevő szervezetek hozzáadásához, illetve eltávolításához. Ezen jóváhagyás szabályok későbbi módosítása is az aktuális konszenzus beállítások által szabályozott. Az egyes funkciókhoz tartozó szabályok kifejezőereje nagy, akár komplex jóváhagyási feltételek is megfogalmazhatóak. Például, megkövetelhetjük, hogy egy módosítás eszközöléséhez elegendő az ***A*** szervezet egyedüli, *vagy* a ***B*** és ***C*** szervezet együttes jóváhagyása.

A fenti példa alapján látható, hogy nem szükségszerű a jóváhagyási irányelvekben való részvételt egyenletesen elosztani a résztvevők között. Ezeket a beállításokat tipikusan a szakterületi igények és szabályozások mentén kell (akár jogi szinten is) definiálni, majd a definícióknak megfelelő Fabric beállításokat és okosszerződés logikát implementálni. Tehát az *irányítási keretrendszer* (governance framework) definíciója fogja meghatározni a technológiai konfigurációt (hacsak nem kell figyelembe venni esetleges technológiai korlátokat).

A projekt kapcsán a hatékony, technológiai szintű jogosultságkezelés (tehát nem fizikai üzemeltetés!) és az adatvédelmi szabályoknak való megfelelés érdekében definiálható egy dedikált szereplő, amely kizárólag blokkláncidentitással és ahhoz kapcsolt jogosultságokkal rendelkezik a rendszerben. Egy ilyen szereplőnek például joga (azaz technológiai szintű hozzáférése) lehet szereplők hozzáadásához a rendszerhez, vagy az okosszerződések frissítéséhez. Fontos kiemelni, hogy ez a szereplő nem kell, hogy saját csomóponttal (ezáltal érzékeny adat replikákkal) rendelkezzen a rendszerben. Irányítási jogosultságait más szervezetek csomópontjainak címzett kérések mentén is gyakorolni tudja (hiszen minden csomóponton ugyanazok a jogosultsági szabályok érvényesek).

Egy ilyen konfigurációban a résztvevő szervezetek a jogosultságkezelést delegálják egy erre a funkcióra kijelölt, adatfeldolgozási jogosultsággal rendelkező szereplőre. Ugyanakkor a végrehajtandó funkciókról nem az adatfeldolgozó fél dönt, ő csak a végrehajtó szerepét tölti be. A döntések a végrehajtandó műveletekről különböző forrásokból érkezhetnek, mint például jogszabályi módosítás, vagy a szervezetek közös (blokkláncon kívüli) megegyezése.

5. Az architektúra variánsok kiértékelése

***Vezetői összefoglaló**. A fejezet egy funkcionális és funkciókon túlmutató szempontrendszer mentén összehasonlítja a centralizált és blokklánc alapú megoldásokat. A kiértékelés alapján az a szakmai javaslatunk, hogy a blokklánc alapú megoldás fenntarthatóbb és hosszabb távon perspektivikusabb.*

A 2. táblázat magas szinten összehasonlítja a centralizált és blokklánc alapú megoldásokat egy adott szempontrendszer mentén. Megjegyzendő, hogy a centralizált megoldás is tartalmaz egy blokklánc alapú metaadat tárat az adatok utólagos, reaktív integritásellenőrzésének támogatásához. Ezáltal minden, a blokklánc technológia nehézségeiből eredő megjegyzés a centralizált verzióra is vonatkozik (szempontonként más-más mértékben).

A lenti szempontrendszert figyelembe véve a blokklánc alapú megoldást tartjuk a hosszabb távon perspektivikusabb, fenntarthatóbb és innovatívabb variánsnak.

	Architektúra	
Szempont	**Centralizált (+ blokklánc alapú metaadat tár)**	**Hyperledger Fabric blokklánc alapú**
Nyomtatvány transzformáció helye	Egyetlen helyen implementálva és üzemeltetve → *Felmerül a központosított adatösszekapcsolás jogi akadálya*	Nyomtatványonként szervezetekre bontva → *Megoldja a központosított adatösszekapcsolás jogi akadályát.*
Lekérdezések kiszolgálása	A központi adatbázisból → *Egyszerűbb, de felmerül a központosított adatösszekapcsolás jogi akadálya*	Re-identifikáció után a vonatkozó szervezetnél gyorsítótárból, vagy teljesen kliens oldalon → *A fókuszált lekérdezések gyorsak és hitelesek. Az aggregát lekérdezések bonyolultabbak, de megoldhatók. Megoldja a központosított adatösszekapcsolás jogi akadályát.*
Fejlesztéshez szakembertudás	Bőséges, de a metadattár igényli a blokklánc szakértelmet.	Szűk, de hagyományos kompetenciákra épít

Üzemeltetéshez szakembertudás	Bőséges, de a metadattár igényli a blokklánc szakértelmet.	Szűkebb, de elégséges, kevés a platform specifikus igény
Skálázhatóság	Hagyományos módszerekkel, rugalmas	Platform specifikus tudás, limitáltabb
Adat minőség biztosítása	Reaktív, utólagos.	Proaktív, azonnali.
Adatok tartalmi ellenőrzése	Egyszerű, de központosított. → *Felmerül a központosított adatösszekapcsolás jogi akadálya*	Több adat forrás (lánc), konszenzusos hitelesítés.
Integrálhatóság	Eseti alapon biztosítandó	A rendszer természetes kiterjesztésével, egyszerű mechanizmusként – a platform alapvető funkciójának részeként.
Adatok eljuttatása a felekhez	Bonyolultabb	Egyszerű, replikáció
Adatkezelés	Egyszerű	Bonyolultabb (deidentifikált), de rugalmasabb és hosszabb távon könnyen kiterjeszthető.
Adatvédelmi jogok biztosítása	Egyszerű	Deidentifikálással megoldott, technikailag hagyományos adatbázisban tárolva
Adatok archiválhatósága	Egyszerű	Több rétegű, megoldott és jól dokumentálható.
Kliensoldali válaszidők hossza	Másodperc nagyságrend (kvázi valós idejű)	Másodperc nagyságrend (kvázi valós idejű)
Hibatűrés, rendelkezésre állás, ellenállóképesség (robosztusság)	Tervezés során biztosítandó	Eredendően magas, így a rendszer védett és ellenálló a külső behatásoknak.

Adatsérülés felderítése	Reaktív, manuális, detektív jellegű, látens hibák maradhatnak.	Proaktív, szinte azonnali, a technológia által biztosított a teljeskörű felderítés az adatciklus teles fázisában.
Adatok visszaállíthatósága	Tervezés során biztosítandó	A technológia maga biztosítja, ha a node-ok nem egy szervezetnél vannak elhelyezve.
Adatkezelési jogosultságok követése	Egyszerű	Tervezés alapján, jól dokumentálva és konszenzus által hitelesítve.
Verziófrissítések komplexitása	Alacsony	Magasabb, de jól dokumentált és konszenzus által hitelesített.
Teljes komponens komplexitás	Alacsony, de a metadattár igényli a blokklánc szakértelmet.	Közepes, továbbra is homogén a rendszer. De a további bővíthetőség és üzemeltethetőség költségei alacsonyak.
Rendszer és/vagy elemének elérhetetlenségi kockázata	Magas (tervezendő hibatűrés és centralizált elemek)	Minimális (beépített redundancia és decentralizált működés)

2. táblázat: A centralizált és blokklánc alapú megoldások összehasonlítása

6. Kitekintés: a digitális adatkezelés új irányai

Vezetői összefoglaló. *Az általános személyes adatok decentralizált azonosítókkal (DID) és ellenőrizhető tanúsítványokkal való felvértezése számtalan előnyt jelent a digitális személyazonosság és az adatkezelés területén. Először is, ez a megközelítés fokozott biztonságot és adatvédelmet kínál, mivel a DID-k és az ellenőrizhető tanúsítványok kriptográfiai módszerekre támaszkodnak az adatok biztonsága és hitelesítése érdekében, csökkentve az adatvédelmi incidensek és a személyazonosság-lopás kockázatát. Másodszor, az egyéneknek nagyobb ellenőrzést és tulajdonjogot biztosít személyes adataik felett, lehetővé téve számukra, hogy szelektíven osszák meg az adatokat bizonyos felekkel, miközben megmarad a hozzáférés bármikor történő visszavonásának lehetősége.*

Emellett a DID-k és az ellenőrizhető tanúsítványok elősegítik az interoperabilitást azáltal, hogy lehetővé teszik a különböző szolgáltatások és platformok közötti zökkenőmentes adatcserét, csökkentik a redundanciát és egyszerűsítik a személyazonosság-ellenőrzési folyamatokat. Továbbá ez a megközelítés racionalizálhatja az adminisztratív eljárásokat, például a pénzügyi szolgáltatásokhoz való csatlakozás vagy a foglalkoztatás ellenőrzése során, mivel digitális, hamisíthatatlan nyilvántartást biztosít az egyének képzettségéről és tulajdonságairól. A blokklánc alapú EMAP architektúra egy minőségbiztosított és hiteles adatforrást biztosít a fenti digitalizációs folyamatokhoz.

6.1. Az önrendelkezésű identitások paradigmája

Az önrendelkezésű identitások (Self-sovereign Identity, SSI) forradalmi koncepció a digitális személyazonosság-kezelés területén, amely az egyének számára nagyobb ellenőrzést és rendelkezést biztosít személyes adataik felett. A hagyományos személyazonossági rendszerekkel ellentétben, ahol különböző intézmények központosított módon tárolják és hitelesítik a felhasználói adatokat, az SSI lehetővé teszi az egyének számára, hogy saját digitális személyazonosságukat biztonságos, decentralizált és felhasználó-központú módon kezeljék.

Az SSI alapvetően blokklánc technológiát és kriptográfiai elveket használva egy olyan rendszert hoz létre, amelyben a felhasználók szükség szerint tárolhatják, megoszthatják és ellenőrizhetik személyazonossági adataikat, anélkül, hogy közvetítőkre támaszkodnának. Ez a paradigmaváltás az egyéneket teszi felelőssé digitális személyiségükért, javítva az adatvédelmet, a biztonságot és az interoperabilitást a különböző online szolgáltatások és alkalmazások között, miközben csökkenti a központosított adattárakkal és a személyazonosság-lopással kapcsolatos kockázatokat. Az SSI képes átalakítani a digitális világgal való kapcsolatunkat, mivel a személyazonosság ellenőrzésének és hitelesítésének inkluzívabb, átláthatóbb és felhasználóbarátabb megközelítését kínálja.

Az SSI a decentralizált identitásokra (Decentralized Identifier, DID), az ellenőrizhető tanúsítványokra (Verifiable Credential, VC) és az ellenőrizhető prezentációkra (Verifiable Presentation, VP) épül. A DID-k a felhasználók által létrehozott és kezelt, egyedi azonosítók. A VC-k megbízható szervezetek által kiadott digitális tanúsítványok, a VP-k pedig lehetővé teszik az egyének számára, hogy szükség szerint szelektíven megosszák a VC-ket személyazonosságuk vagy tulajdonságaik bizonyítására.

6.1.1. Elosztott identitások

A decentralizált azonosítók (DID-k) jelentős előrelépést jelentenek a digitális személyazonosság-kezelés területén, mivel technikai keretet kínálnak egyedi, önálló azonosítók létrehozásához és kezeléséhez, amelyek nem függenek a központi hatóságoktól. Ezek az azonosítók egy globálisan egyedi karakterláncon alapulnak, amely jellemzően egy URL, és egy DID-dokumentumhoz kapcsolódik. A DID-dokumentum olyan adatstruktúra, amely kulcsfontosságú információkat tartalmaz, beleértve a kriptográfiai kulcsokat, a szolgáltatás végpontjait és a kapcsolódó DID-hez kapcsolódó egyéb metaadatokat. Ezek a kriptográfiai kulcsok kulcsfontosságú szerepet játszanak a kommunikáció és az interakciók biztosításában, az ellenőrizhető állítások lehetővé tételében, valamint az egyének vagy szervezetek digitális identitásuk feletti ellenőrzésének biztosításában. A DID-ek és a hozzájuk tartozó DID-dokumentumok képezik a decentralizált és a magánélet védelmét biztosító személyazonossági ökoszisztéma alapját, elősegítve a különböző digitális alkalmazások és szolgáltatások interoperabilitását és bizalmát.

6.1.2. Ellenőrizhető tanúsítványok

Az ellenőrizhető tanúsítványok forradalmi paradigmaváltást jelentenek a digitális személyazonosság és a személyes adatok kezelésének és megosztásának módjában. Ezek a tanúsítványok egy személyre, szervezetre vagy egyéb entitásokra vonatkozó információk vagy állítások digitális reprezentációját jelentik, és hitelességük és eredetük kriptográfiai bizonyítékkal van ellátva. Az ellenőrizhető tanúsítványok biztonságos, az adatvédelmet megőrző módját kínálják az információcserének, mivel lehetővé teszik, hogy a birtokosuk csak a releváns adatokat ossza meg az érdekelt felekkel, felesleges részletek felfedése nélkül.

Ez a megközelítés nemcsak az adatvédelmet és a biztonságot növeli, hanem egyszerűsíti a személyazonosság-ellenőrzési folyamatokat is, csökkentve a hagyományos papír alapú dokumentáció szükségességét, és hatékonyabb, bizalomalapú interakciókat tesz lehetővé különböző területeken, többek között a pénzügyekben, az oktatásban, az egészségügyben és más területeken. Az ellenőrizhető tanúsítványok egyre szélesebb körű elterjedésével

az ígéretek szerint az egyének kezébe kerülhet az adataik feletti ellenőrzés, és az interneten a digitális személyazonosság és a bizalom felhasználó-központúbb és decentralizáltabb megközelítését segítik elő.

6.1.3. Ellenőrizhető prezentációk

Az ellenőrizhető prezentációk kritikus elemet jelentenek az önrendelkezésű identitások és a digitális bizalom kialakulóban lévő tájképében. Ezek a prezentációk az ellenőrizhető tanúsítványok koncepciójára épülnek, lehetővé téve az egyének számára, hogy tanúsítványokat és állításokat egyetlen, ellenőrizhető csomagként csomagolják és osszák meg egymással. Az ellenőrizhető prezentációk kriptográfiai technikák segítségével lehetővé teszik, hogy a birtokosok meggyőző bizonyítékot szolgáltassanak személyazonosságukról vagy tulajdonságaikról a megbízó felek számára anélkül, hogy felfednék az egyes hitelesítő okmányok mögöttes részleteit.

Ez nemcsak a személyazonosság-ellenőrzési folyamatokat egyszerűsíti, hanem az érzékeny adatok védelmét és az ellenőrzést is fokozza, lehetővé téve az egyének számára, hogy az interakció kontextusa és bizalmi szintje alapján szelektíven adjanak ki információkat. Az ellenőrizhető prezentációkkal a digitális személyazonossági ökoszisztéma egy sokoldalú eszközzel bővül a biztonságos, felhasználó-központú adatmegosztáshoz, ami elősegíti a bizalmat és a hatékonyságot az online tranzakciók széles skáláján, a szolgáltatások elérésétől a képesítések és jogosultságok ellenőrzéséig.

6.2. Az ellenőrizhető EMAP adatok irányába

A javasolt, blokklánc alapú EMAP architektúra biztosítja az adatok minőségét, hitelességét és azonnali konzisztenciáját több résztvevőn keresztül is. Ezen kívül a használt blokklánc platform beépített módon lehetővé teszi a rendszer könnyű kiterjesztését további szervezeteken keresztül is. Ugyanakkor, ha a kívánt szakrendszerek közötti integráció valamilyen szempontból nem előnyös (például nagyon ritka az adatmozgás ahhoz, hogy megérje a közvetlen integráció lefejlesztése), akkor az EMAP által tárolt adatok közvetve is eljuttathatók a cél rendszerbe, az adat alanyán (például a természetes személyen) keresztül.

Az SSI paradigma pontosan ezt a fajta interakciót támogatja. Ahelyett, hogy a szakrendszerek közvetlen, pont-pont összeköttetése jönne létre (amely az érintett szakrendszerek számával rosszul skálázódik), az adat alanya fogja biztosítani az összeköttetést a két rendszer között:

1. Az EMAP rendszer egy előre definiált formátumban kibocsát a felhasználó számára egy ellenőrizhető tanúsítványt a szükséges adatok alapján. A tanúsítvány hitelességét és integritását a kibocsátó a digitális aláírásával biztosítja.

2. A felhasználó eltárolja a tanúsítványt a saját digitális tárcájában, amelynek innentől kezdve nem kell összeköttetésben állnia a kibocsátó rendszerrel (hasonlóan egy letöltött, onnantól kezdve offline elérhető fájlhoz).
3. A célrendszer megkéri a felhasználót, hogy egy ellenőrizhető prezentáció formájában mutassa be az aktuális ügyintézéshez szükséges hiteles adatait. A felhasználó a digitális tárcája segítségével előállítja az ellenőrizhető tanúsítványból a prezentációt, és feltölti a célrendszerbe, annak saját (például egyszerű fájl feltöltés) megoldását használva.
4. A célrendszer egy hiteles regiszteren keresztül ellenőrzi, hogy az adatokat valóban a szükséges szervezet (például a NAV) írta alá digitálisan. Ha az ellenőrzés sikeres, akkor az ügyintézési folyamat folytatódhat tovább.

Ez a fajta információáramlás hasonló ahhoz, ahogy például a személyazonosító igazolványt használják természetes személyek. Az okmányt az állam bocsátja ki az előírt feltételek teljesülése esetén. A hitelességét az okmányon elhelyezett hologramok (és hasonló megoldások) biztosítják. A személy innentől kedve bárki felé igazolhatja az okmányon szereplő adatok érvényességét egyszerűen az okmány felmutatásával. Az ellenőrző fél a hitelesség ellenőrzése után elfogadja az adatok tartalmát, hiszen megbízható félként tartja számon az államot.

A folyamat digitális verziójában egy központi szerepet játszik a 4. pontban előkerült hiteles regiszter, amihez az ellenőrző felek fordulnak. Ha egy központi szereplő biztosítja ezt a regisztert, akkor felmerülhet az ellenőrzés eredményének hitelessége és integritása. Ugyanakkor, ha egy ilyen regiszterhez több szereplő is egyszerű, közvetlen hozzáférést kap blokklánc technológia segítségével, akkor az ellenőrzések szempontjából teljes lesz a transzparencia. Akár olyannyira is, hogy egy ilyen blokkláncot széles körben hozzáférhetővé teszünk, mint például az EU tagállomok átívelő EBSI platform. A blokklánc technológia tehát egy jól skálázódó, transzparens megoldást nyújt a hiteles és minőségbiztosított adatmegosztásra.