

A közigazgatásban rendelkezésre álló adatvagyon kezelésének, felhasználásának tervezéséhez és szervezéséhez segítséget nyújthatnak azok a módszertani kezdeményezések is, amelyek tudományos kutatási eredményként jelennek meg. Mindezek mellett külön figyelmet érdemelnek az Európai Uniónak a témával összefüggésben keletkezett politikai dokumentumai nyomán születő gyakorlati segédletek.

Kulcsszavak: tudományos információk, kutatási adat, adatvagyon-gazdálkodás, nyílt hozzáférés, adatmenedzsmentterv

Bevezetés

Az Európai Unió (a továbbiakban: EU) *Horizont 2020* kutatási és innovációs keretprogramja nagy hangsúlyt helyez a tudományos információk – tudományos közlemények és kutatási adatok – minél szélesebb körű és minél gyorsabb megismerhetőségének biztosítására.

A tudományos információk megismerhetővé tételének szükségességével az elmúlt években több EU-dokumentum foglalkozott. Jelen tanulmány az EU két iránymutatását emeli ki. A *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020* (a továbbiakban: *Nyílt hozzáférés útmutató*) a közzététel elméleti kérdéseit tekinti át a keretprogramban finanszírozott projektek kedvezményezettjei számára értelmezhető módon, a *Guidelines on Data Management in Horizon 2020* (a továbbiakban: *Adatmenedzsment-útmutató*) a keretprogram résztvevői számára kíván segítséget nyújtani az adatminőségre, adatmegosztásra és adatbiztonságra vonatkozó kötelezettségek teljesítéséhez. [1] [2]

A fenti iránymutatások nyomán három francia egyetemi szakember 2015 januárjában tett közzé egy adatmenedzsment-tervre vonatkozó ajánlást – *Réaliser un plan de gestion de données: guide de rédaction* –, amely már konkrét mintát ad az érintettek számára az adatok sorsának tervezéséhez és dokumentálásához. [3]

A dokumentumok áttekintése – az adatvagyon-gazdálkodás, adatmenedzsment egy részterületének értékelésével – a közigazgatási adatvagyon-gazdálkodás számára is nyújthat elméleti, módszertani segítséget.

Az adat és az információ szerepe a gazdasági növekedésben

Az EU 2010-ben elfogadott stratégiája (*Európa 2020 – Az intelligens, fenntartható és inkluzív növekedés stratégiája*) tudáson és innováción alapuló, az erőforrásokat hatékonyan kezelő és használó gazdaság megteremtését, növekedésének elősegítését tűzte ki célul. [4] A stratégiát követő kezdeményezések, programok, felmérések, egyéb dokumentumok meghatározott részében az egyik alapvető erőforrásként definiált adat- és információ-tömeg hasznosításával kapcsolatos elképzelések, elvárások fogalmazódnak meg, szoros összefüggésben „a digitális adatok, a számítástechnika és az automatizálás új ipari forradalmának” kihívásaira adandó válaszokkal. [5]

Az Európai Bizottság adatközpontú gazdaságról szóló 2014-es közleménye – részben a fent is említett előzményekre hivatkozva – ismételten megerősíti, hogy „az adatközpontú gazdaságbeli világszintű versenyképesség megteremtése érdekében [többek között az Európai Unió] nyilvános adatforrásainak és kutatási adatinfrastruktúráinak széles körben történő megosztása, alkalmazása és fejlesztése” szükséges. [5] Az adattömeg gazdasági jelentőségét hangsúlyozandó a dokumentum rögzíti: „az óriási méretű adathalmazokhoz kapcsolódó technológia és szolgáltatások értéke 2015-ben várhatóan eléri majd a 16,9 milliárd USD-t, 40%-os összesített éves növekedést követően, amely hozzávetőleg hétésszerese az információs és kommunikációs technológiák teljes piaci növekedésének.” [5]

Az adatok hasznosításának egyik kiemelt területe a *közadatok* további felhasználása. A 2003-ban elfogadott PSI irányelv „a közsféra dokumentumainak felhasználására vonatkozó nemzeti szabályok és gyakorlatok minimum harmonizációját” kívánta elősegíteni. [6] Az irányelvet 2013-ban módosították, mivel az adatmennyiség növekedése, az újabb típusú adatok megjelenése, az adatok kiaknázására irányuló technológiák ugrásszerű fejlődése miatt a tagállamok számára újabb – szigorúbb, ösztönzőbb – magatartás-előírások rögzítése vált szükségessé. [7]

További ígéretes terület a *tudományos információk* megismerésének és kiaknázásának biztosítása.

A tudományos információk megismerhetősége

Az adatmennyiség robbanásszerű növekedése (big data, magyarul – szebben – „adatrengeteg”, „adatözn”) a tudományos információk esetében is egyre nagyobb lehetőségeket, ugyanakkor számos új kihívást jelent.

A lehetőségek és az ezek kiaknázáshoz szükséges teendők rövid összefoglalását adja – részben a tárgyban korábban született EU-dokumentumok összegzésével – a *Nyílt hozzáférés útmutató*, amely szerint „a modern kutatás a széles körű tudományos párbeszéd

alapul és a korábbi munkák tökéletesítése révén fejlődik” [1] A tudományos közleményekhez és a kutatási adatokhoz való nyílt hozzáférés lehetővé teszi

- a korábbi kutatási eredmények felhasználását (az eredmények minőségi javulását),
- a jobb együttműködést és az erőforrások párhuzamos felhasználásának elkerülését (a hatékonyság növelését),
- az innováció felgyorsítását (a gyorsabb piacra jutást és így a gyorsabb növekedést),
- a polgárok és a társadalom bevonását (a tudományos folyamatok átláthatóságának növelését). [1]

Bár az útmutató nem említi, a tudományos információk nyilvánosságának elősegítése melletti érvként jelenhet meg (nem indokolatlanul) a tudományos csalások lehetőségének visszaszorítása is. [8]

A nyílt hozzáférés biztosítása alapvető elvárás a közfinanszírozású kutatások esetében, egyrészt azért, hogy a társadalom által egyszer már finanszírozott kutatás eredményeiért az érintett közösség tagjainak ne legyen szükséges még egyszer fizetniük, másrészt azért, mert az európai kutatók, polgárok és az innovációban élen járó gazdasági társaságok, vállalatok további költségek nélkül így tudják hasznosítani az eredményeket.

A nyílt hozzáférés biztosítása a tudományos közleményekhez történhet közvetlen folyóiratbeli közzéléssel (gold open access, 'arany út') vagy online repozitóriumban történő elhelyezéssel (green open access, „zöld út”). Az EU által 2012-ben közzétett adatok szerint a tudományos szakirodalom 20%-a nyílt hozzáférésű: 8% folyóiratok („arany út”), 12% online repozitóriumok („zöld út”) révén; a cél az, hogy 2016-ra a nyílt hozzáférés keretében elérhető közfinanszírozású tudományos cikkek aránya 60%-ra emelkedjen. [8] [9]

A kutatási adatok esetében – azok online repozitóriumban történő elhelyezését követően, indokolt esetben 6–12 hónapos késleltetés (embargóidőszak) után – szabad hozzáférést és felhasználást, illetve különleges indokok alapján korlátozott hozzáférést és/vagy felhasználást szükséges biztosítani. A nyílt hozzáférés biztosításától akkor lehet eltekinteni, ha az eredmények szabadalmaztatása, üzleti célú hasznosítása várható, a személyes adatok védelme megköveteli, nemzetbiztonsági előírások lehetővé vagy szükségessé teszik, stb.

A tudományos információk felhasználásával és megosztásával kapcsolatban felmerülő nehézségek – ahogyan manapság minden tudással, tudásmegosztással összefüggő élethelyzetben – lehetnek szakmai, jogi, információtechnológiai vagy pénzügyi vonatkozásúak. Ezek számbavételével számos EU-dokumentum, cikk, tanulmány foglalkozik. Így például az említett bizottsági közlemények közül a 2012-es az adathozzáféréssel és -felhasználással kapcsolatos felelősségi körök szervezetlenségét és tisztázatlanságát, a hosszú távú hozzáférést biztosító finanszírozási modellek, továbbá az országok és a tudományterületek közötti interoperabilitás hiányát, a 2014-es a jelenlegi jogi környezet összetett jellegét, a nagy adatkészletekhez és a mögöttes infrastruktúrához való hozzáférés hiányát, az új technológiák (az Amerikai Egyesült Államokhoz képest) lassú meghonosítását emeli

ki. [5] [9] Holl András két tanulmánya a kutatói szellemi befogadó kapacitás korlátai és a rendelkezésre álló adattömeg egyre növekvő mérete közötti ellentmondáson és a kutatók publikálástól való vonakodásán túl a publikálás, az adatelemzés és az adattömegben történő keresés, valamint a hosszú távú megőrzés információtechnológiai problémáit tárgyalja. [10] [11]

Az egyes problémák részletes elemzése meghaladná jelen tanulmány kereteit; mindössze annak az általánosan levonható következtetésnek a rögzítése szükséges, hogy a kutatási adatok, információk hatékony felhasználása érdekében alapvető elvárássá válik az éppen végzett tudományos tevékenységhez szükséges és elégséges adatok, információk kiválasztásának (megtalálásának) és feldolgozásának támogatása.

Az ehhez a támogatáshoz szükséges eszközök ugyancsak jellemezhetők a már említett szempontok alapján (azaz lehetnek szakmai, jogi, információtechnológiai vagy pénzügyi eszközök). A szakmai, jogi, információtechnológiai eszközök jellemzői alapján egyértelművé válik, hogy egységesített – szabványos – hozzáférhetővé tételre van szükség, tartalmi és formai szempontból egyaránt. Egyértelmű továbbá az is, hogy az egységességnek, a szabványosságnak az adatok létrehozásának/gyűjtésének folyamatában és az eredmények közzététele során is érvényesülnie kell.

Módszertani elvárás: adatmenedzsment

A *Nyílt hozzáférés útmutatóban* nagy hangsúlyt kap a *Horizont 2020* egyes területein megvalósuló *Open Research Data Pilot*, amelynek célja a keretprogram projektjeiben összegyűjtött vagy létrehozott kutatási adatokhoz való szabad hozzáférésnek és az adatok újrahasznosításának elősegítése.

Az *Open Research Data Pilot* hatókörébe tartozó projektek résztvevői – a korábban kifejtettekkel összhangban – kötelesek

- a kutatási adatokat egy tematikus, intézményi vagy központi online repozitóriumban eltárolni;
- harmadik személyek részére lehetővé tenni a kutatási adatok elérését, az adatbányászatot, az adatok hasznosítását, reprodukálását és terjesztését, a már idézett kivételekre tekintettel (szellemi tulajdonjogok védelme, jogi és biztonsági megfontolások), továbbá mindezek érdekében az adatok eléréséhez szükséges technikai információkat szolgáltatni.

Új – módszertani – elvárásként fogalmazódik meg az az előírás, amely szerint a projektek résztvevői kötelesek adatmenedzsment tervet készíteni, és azt folyamatosan aktualizálni.

Az *adatmenedzsment terv* (a továbbiakban: AMT) alkalmazását a *Horizon 2020 Work Programme for 2014–2015* dokumentum tette kötelezővé az *Open Research Data Pilot*

ban résztvevők számára. Az *Adatmenedzsment-útmutató* szerint az AMT rögzíti, hogy a projektben gyűjtött vagy létrehozott adatokat hogyan fogják kezelni a kutatási projekt során, majd annak lezárását követően, azaz

- milyen adatokat fognak gyűjteni, illetve létrehozni,
- milyen szabványokat fogak követni,
- hogyan lesz biztosított ezen adatok hasznosítása, megosztása, és/vagy hozzáférhetővé tétele a verifikálás vagy az újrahasznosítás érdekében (illetve amennyiben erre nem kerül sor, ennek indokai), végül
- hogyan lesz biztosított ezen adatok gondozása¹ és megőrzése.

Az AMT első verzióját a lehető leghamarabb, de legkésőbb a projekt hatodik hónapja végéig el kell készíteni. Legalább a projekt félidejében és végén ajánlott az AMT aktualizálása, hogy megtörténhessen a finomhangolás azon adatok és felhasználási módok tekintetében, amelyek a projekt indításakor még nem voltak láthatók. Ugyancsak módosítható az AMT, ha jelentős változás következik be a projektben, akár új adatkészletek bevonása, akár a projektben résztvevők adatmenedzsment-politikájának változása, akár külső tényezők miatt.

Az *Adatmenedzsment-útmutató* két függeléke további gyakorlati segítséget ad. Az *AMT-sablon* az AMT elvárt fő fejezeteit tartalmazza, amelyek az alábbiak:

- a létrejövő adatkészlet megnevezése és azonosítója,
- az adatkészlet leírása (a létrehozandó és összegyűjtendő adatok, ez utóbbiak forrása, az adatok jellege, mennyisége, lehetséges hasznosítói, tudományos publikálása, hasonló adatok rendelkezésre állására vonatkozó információk, felhasználási lehetőségek),
- szabványok és metaadatok (a tudományterületen alkalmazható szabványok leírása, amennyiben ilyenek nincsenek, áttekintő leírás arról, hogyan és milyen metaadatokat fognak létrehozni),
- adatmegosztás (az adatok elérhetővé tételének leírása; hozzáférési eljárások, esetleges embargóidőszak, a közzététel technikai megvalósítása, az esetleges újrahasznosításhoz szükséges szoftver és egyéb eszközök; annak meghatározása, hogy az adatok bárki vagy csak egyes csoportok számára lesznek elérhetőek; az adatok tervezett repozitóriuma, az igénybe venni tervezett, már létező repozitórium típusa – intézményi, tudományterületi; amennyiben az adatcsoport nem megosztható, ennek okai – pl. etikai fenntartások, személyes adatok védelme, szellemi tulajdon védelme, kereskedelmi, nemzetbiztonsági szempontok stb.),
- archiválás és megőrzés (az adatok hosszú távú megőrzésének módja, a megőrzés időtartama, az adatok várható mennyisége, a járulékos költségek és ezek tervezett biztosításának módja).

1 Az angol *curation* kifejezés – újabb divatszó – az online elérhető adatok szűrését, feldolgozását, „szervezését”, használatra ajánlását jelenti.

Az *AMT kiegészítő útmutató* öt pontban rögzíti az adatok elvárt tulajdonságait, továbbá az egyes elvárásokhoz kapcsolódó ellenőrző kérdéseket tartalmaz, amelyek megválaszolása a projektben keletkező összes adatkészlet esetében szükséges. A dokumentum szerint a kutatási adatok legyenek könnyen

- fellelhetők,
- hozzáférhetők,
- értékelhetők és érthetők,
- használhatók az eredeti összegyűjtési célon kívül is,
- együttműködők speciális minőségi szabványokkal.

Módszertani segédlet: adatmenedzsment terv

Az EU említett iránymutatásaiban megjelenő elvárások teljesítésének elősegítése érdekében francia egyetemi szakemberek egy AMT-szerkesztési segédletet készítettek. [3] A szerzők jelezték, hogy a dokumentum célja az (előbbieken részletezett) EU-elvárásoknak megfelelő, francia egyetemi kutatási környezetben létrehozandó AMT összeállításának támogatása, így a dokumentum néhány nemzeti és szakterületi sajátosságot is tartalmaz. Mindazonáltal a segédlet – némi általánosítással – más területeken, így a közigazgatásban is hasznosítható.

A segédlet bevezető része kitér arra is, hogy az AMT-k angolszász országokban született gyakorlatát nem könnyű a szerzők francia környezetére alkalmazni, hiszen bár az AMT kidolgozásához szükséges kompetenciák a francia egyetemeken is megtalálhatóak, azok különböző szereplők, szakterületek között oszlanak meg, és az érintettek „mostanáig nem kényszerültek arra, hogy együtt dolgozzanak”. [3] Erre tekintettel a segédlet azt is tartalmazza, hogy az egyes fejezetek kidolgozásában az AMT kidolgozásáért, aktualizálásáért és végrehajtásáért felelősséggel tartozó adatmenedzsment-felelőst és a projektben részt vevő kutatókat mely szakterületeknek (informatikusok, tudományos és műszaki tájékoztatási szakemberek, levéltárosok) kell támogatniuk.

A segédlet az AMT tartalmát – az AMT-re vonatkozó bevezetőt követően – öt fejezetben javasolja rögzíteni, az egyes fejezetek kidolgozásához a kifejtendő kérdések megnevezésével, azok tartalmi magyarázatával, illetve egyes esetekben konkrét (szövegszerű) példákkal nyújt segítséget.

A segédletnek az AMT tartalmára és az AMT-fejezetek kidolgozóira vonatkozó részeit egyszerűsített formában jelen tanulmány *Melléklete* tartalmazza.

Összegzés: tanulságok

Jelen tanulmánynak nem célja a tudományos kutatás során folytatott adatkezelés és adatfelhasználás változásainak fentieket meghaladó értékelése, inkább az azzal összefüggő gondolkodásmód változásából fakadó módszertani tanulságok hangsúlyozása – annak reményében, hogy így egy, a közigazgatásban is meghonosodóban levő tevékenység tervezéséhez és szervezéséhez hasznosítható tapasztalatok összegzésére kerülhessen sor.

Ha azt szükséges meghatározni, hogy egy adatvagyon-gazdálkodásban, adatmenedzsmentben még nem járatos közigazgatási szerv ezt a tevékenységét hogyan tudja megkezdni, a fenti dokumentumok alapján az alábbiak megfontolása javasolt.

- 1) Az adatvagyon-gazdálkodás, az adatmenedzsment „összművészet”. Komplex megközelítést igényel és különböző szakterületek együttműködését. Az adatvagyon-gazdálkodás, adatmenedzsment (még?) nem önálló „szakma”, hanem sokszor nagyon eltérő mentalitású személyek, szervezetek közötti, meghatározott szempontok alapján, meghatározott szabályok mellett végzett, erős koordináció, közös vagy összefüggő eredmények létrehozása érdekében. Érintett benne a közigazgatási szerv alaptevékenységét végző szakmai oldal, bármilyen közigazgatási alaptevékenység vitathatatlan információtechnológiai meghatározottsága okán a közigazgatási szerv tevékenységét támogató informatikai szakterület, a közigazgatási szerv által végzett adatkezelés tényére, illetve az ennek szabályozottságára vonatkozó elvárásra tekintettel a jogi szakterület, a közigazgatási szerv működési módjából következően az ügyvitelszervezési szakterület (beleértve az iratkezelési, levéltári megőrzéssel kapcsolatos tevékenységeket is), valamint az adat- és információvédelemért felelős szakterület(ek). Tartalmából, összeállításának módjából is következően fontos az elfogadottsága és a támogatottsága; a szervezetben szokásos felsővezetői jóváhagyási/engedélyezési folyamatot követően kezdhető meg a végrehajtása.
- 2) Ennek megfelelően – a hiedelmekkel ellentétben – nem a támogató informatikai eszköz a siker kritériuma, hanem a tervszerű és következetes gondolkodás és együttműködés: „a sokat emlegetett hozzáadott érték [...] csak a megfelelő szakértelem és a befektetett munka révén keletkezik.” [12]
- 3) Szükséges a fogalmi következetesség a tevékenység tárgyát – a menedzselt adadfajtákat, adattípusokat, adatkészleteket – érintően. Tudatában kell lenni azonban annak, hogy az adatokra irányuló tevékenység leírására használt fogalmak² nem mindig egyértelműek, a szakirodalomban nyelvenként, országonként változóak, a

² Az adatvagyon-gazdálkodás és az adatmenedzsment fogalma sincs elhatárolva egymástól; a két fogalom tartalma nagyrészt fedi egymást, a közigazgatási tevékenység leírására – részben a nemzeti adatvagyon körébe tartozó állami nyilvántartások fokozottabb védelméről szóló 2010. évi CLVII. törvény (Adatvagyon tv.) miatt – az előbbi kifejezés az elterjedtebb.

jogszabályokban nem feltétlenül következetesek. Folyamatosan változó, felgyorsult világunkban az adat- és információkezelő tevékenység körében felmerülő újabb és újabb kifejezések lehetnek lényeges módosulást jelző, meghonosodó (szak)kifejezések, amelyek akár jogi normákban is létjogosultságot nyernek, de lehetnek kérészerű divatszavak is. Éppen ezért ezekkel a fogalmakkal óvatosan kell bánni. A normatív rendelkezésekben megjelenő, kötelezően alkalmazandó fogalmak használatát össze kell tudni egyeztetni (szükség esetén korrekciókat javasolva) a „való világ” jelenségeivel, felfogásával és ezt tükröző szóhasználatával, és legalább az adott közigazgatási szerv tevékenysége során következetesen alkalmazni.

- 4) Központi fogalmi elem az adatokért való felelősség. Ez részben az adatminőség, részben az adatokkal végzett tevékenység jogszerűségének és szakszerűségének garanciája. Egy közigazgatási szerv esetében mindez alapfeltétel az állampolgárok bizalmának megőrzéséhez.
- 5) A tevékenység szabályozottságát, működési kereteit külföldi és hazai normák és szokások – jogszabályok, közjogi szervezetszabályozó eszközök, szakmai és technológiai szabályok, szabványok, jó gyakorlatok stb. – biztosít(hat)ják. Ezek kölcsönhatása, illetve a környezet változásából és a tevékenység fejlődéséből következő újabb és újabb szabályozási kényszer lehetővé teszi az egyre pontosabb elvárás- és feltételrendszer rögzítését.
- 6) Az adatvagyon-gazdálkodás, az adatmenedzsment az adatok életciklusát végigkövető tevékenység, amely önmagában is ciklikus – a PDCA ciklus³ elemei szükségszerűen megjelennek művelése során –, véglegesen nem lezárható feladat.
- 7) Előbbi tulajdonságaiból következően egységesíthető és formalizálható. Ez szükséges is, mivel jelenleg még nincsenek teljesen kiforrott gyakorlatok, inkább különböző, figyelmet érdemlő irányzatok, kezdeményezések.
- 8) Tartalmát tekintve a józan arányosság elve követendő. Legyen átfogó (mindent egyben látó), ugyanakkor kellően részletes (mindent leíró). Ne legyen túl elméleti, mert a mindennapokban nem lesz hasznosítható. Ne legyen túl részletes, mert művelői el fognak veszni benne, az áttekintéséhez pedig az egyéb érintetteknek nem lesz elég idejük és energiájuk.
- 9) Erőforrásigénye, költségessége a fentiek fényében változhat, azonban tervezhető és módosítható.
- 10) A legfontosabb: lehetőség szerint minél előbb meg kell kezdeni és (valószínűleg folyamatos korrekciókkal) csinálni kell!

³ A PDCA ciklus egy négylépéses, spirálszerűen ismétlődő, és ezáltal egyre magasabb minőségi szintet biztosító általános irányítási-fejlesztési modell. A megnevezés az angol Plan (Tervezés) – Do (Végrehajtás) – Check (Ellenőrzés) – Act (Beavatkozás) szavak kezdőbetűiből alkotott betűszó. A modellt W. Edwards Deming amerikai mérnök-statisztikus népszerűsítette az 1950-es években. Előzménye a szintén amerikai mérnök-statisztikus W. Andrew Shewhart által kidolgozott háromlépcsős, statisztikailag ellenőrzött gyártási folyamat (specifikáció, gyártás, értékelés).

Melléklet: AMT-szerkesztési útmutató

Forrás: Cartier, Aurore – Moysan, Magalie – Reymonet, Nathalie: Réaliser un plan de gestion de données: guide de rédaction (V1, 09/01/2015).

0. Bevezető

- AMT verziószáma, kelte
- AMT 1. verziójának kelte

1. Projektinformációk

A fejezet célja azon projekt adminisztratív adatainak a rögzítése, amelyhez az AMT kapcsolódik

Felelősök és közreműködők: Projektmérnökök.

- Projektfelhívás azonosítója
- Téma
- Finanszírozási szerződés azonosítója
- Kutatási program megnevezése
- Projekt megnevezésének rövidítése (betűszó)
- Projekt címe
- Projekt céljai

A projekt jellegének, céljainak és lefolyásának leírása. Lehetővé teszi a projekt során létrejövő, illetve összegyűjtendő adatok típusainak és összefüggéseinek a megértését.

- Projekt kulcsszavai
- Koordinátor / kedvezményezett szervezet
- Projekt tudományos felelőse (projektfelelős)
- Projektfelelős szervezete, szervezeti egysége

2. Felelősség az adatokért

A fejezet célja azon személy(ek) azonosítása, aki(k) az AMT létrehozásáért és aktualizálásáért felel(nek) a projekt során, valamint az adatok tulajdonjogának a rögzítése.

Felelősök és közreműködők: Adatmenedzsment-felelős,

Kutatók,

Projektmérnökök.

- Az adatmenedzsment-felelős neve a projektben

Az adatmenedzsment-felelős a projektvezetésben végig jelen van, a létrehozástól a lezárásig. Felelős az AMT-ért, összeállításáért, követéséért, esetleges módosításáért a kutatási projekt során. Nem feltétlenül azonos a projektfelelőssel.

Amennyiben a projektben több terület vagy intézmény vesz részt, a felelőségeket pontosan meg kell határozni.

- ♦ Az adatok tulajdonjoga

Kihez tartoznak a létrejövő/felhasznált adatok? Milyen jog, jogszabályok szabályozzák az adatbázisokat és tartalmukat? Hogyan oszlik meg az adatok szellemi tulajdonjoga? stb.

Az adatok tulajdonjogát a projektben közreműködők között létrejött megállapodás rögzíti. A projekt és az AMT tartalma feletti megegyezés alapvető jelentőségű, hiszen részben ez határozza meg az adatok későbbi megosztásának és közzétételének feltételeit.

Pl.: Saját eredmények: az eredmények azon résztvevő tulajdonába kerülnek, aki létrehozta (vagy alvállalkozóval a többi résztvevő finanszírozása nélkül létrehozta) azokat a tanulmányokat, munkákat, melyek az eredményekhez vezettek. A saját eredmények tulajdonosa fogja – kizárólagosan – meghatározni az eredmények védelmének és felhasználásnak módját és idejét. Közös eredmények: ha az eredmények egynél több résztvevő tanulmányainak/munkájának következményei (vagy olyan alvállalkozóé, akit közösen finanszíroztak), minden résztvevőnek azonos arányban kell részesednie az eredmények tulajdonjogából.

3. Az AMT megvalósításához szükséges erőforrások

A fejezet célja az AMT megvalósításához szükséges kompetenciákra és ennek költségeire vonatkozó becslés rögzítése – adatmenedzsment, adatgondnokság (válogatás, tisztítás, normalizálás, data enrichment), hosszú távú megőrzés.

Felelősök és közreműködők: Projektmérnökök.

- ♦ Infrastruktúra
- ♦ Személyzet
- ♦ Képzés, oktatás
- ♦ Költségek

4. Adatkészletek

A projekttől függően keletkezhet egy vagy több adatkészlet, amely(ek) lehet(nek) technikailag homogének vagy logikailag koherens(ek), ugyanakkor technikailag heterogén(ek).

- ♦ Adatkészletek száma

Az alábbi 4.1. és 4.4. közötti szakaszok adatkészletenként ismétlődnek

4.1. Az adatok leírása – ... sz. adatkészlet

A fejezet célja a projekt keretei között létrehozandó és/vagy összegyűjtendő adatkészletek leírása.

Felelősök és közreműködők: Adatmenedzsment-felelős,

Kutatók.

- ♦ Adatkészlet azonosítója és megnevezése

- ♦ Adatok jellege
Pl.: mérési adat, mintavétel, logikai kód, szöveg, fénykép stb.
- ♦ Már rendelkezésre álló adatok felhasználása
Ez a rész rögzíti, ha a projekt már létező, esetleg más kutatócsoportok által létrehozott adatokra, adatkészletekre támaszkodik.
- ♦ Adatok létrehozásának módja
Ez a rész pontosítja az adatok létrehozásának vagy kidolgozásának folyamatát: mérés, elemzés, megfigyelés, kompiláció, szimuláció stb. Jelezni kell, hogy a projektben létrehozott adatokról, átvett adatokról vagy mindkettőről van-e szó.
- ♦ Adatformátumok, szabványok
Rögzíteni kell az adatok formátumát. Amennyiben az adatok konvertálásra kerülnek, az eredeti formátum egyszerű említése elegendő, azonban a konverziós formátum részletes leírása kötelező. Előnyben kell részesíteni a nyílt vagy széles körben elterjedt formátumokat a megosztás elősegítése érdekében.

4.2. A projekt folyamata

A fejezet célja a projekt alatti tárolás, védelem és hozzáférés módozatainak rögzítése.

Felelősök és közreműködők: Informatikusok.

Tárolás és rögzítés:

- ♦ Adathordozó
Papír vagy egyéb.
- ♦ Tervezett mennyiség
Tárolókapacitás méretében; a projekt során felülvizsgálatra kerülhet.
- ♦ Adattárolás
A tároló infrastruktúra és a tárolás fizikai helyének leírása. A projekt során az adattároláshoz igénybe vesznek-e külső szolgáltatót? Szükség van-e a tároláshoz speciális szoftverre?

Adatbiztonság, adatvédelem:

- ♦ Veszélyek és fenyegetések az adatok vonatkozásában
*Jelezni kell az adatok bizalmasságát, rendelkezésre állását és integritását érintő veszélyeket, fenyegetéseket a projekt teljes aktív szakaszában.
Pl.: véletlen vagy jogosulatlan adattörlés, nem engedélyezett hozzáférés, kockázatos környezet stb.*
- ♦ Az adatok bizalmasságának garanciái
*Pontosan meg kell határozni azokat az intézkedéseket, amelyeket az adatok bizalmasságának a megőrzése érdekében hoztak, figyelemmel a feltárt kockázatokra.
Pl.: külön engedélyezési eljárás minden jövőbeli felhasználó esetében.*
- ♦ Az adatok integritásának és hitelességének garanciái
Pontosan meg kell határozni azokat az intézkedéseket, amelyeket az adatok integ-

ritásának és hitelességének a megőrzése érdekében hoztak, figyelemmel a feltárt kockázatokra.

Pl.: kutatási protokollok, tárolási nyilvántartások, naplózás, autentikáció az eszközök használatakor.

Hozzáférés az adatokhoz:

- ♦ Az adatok olvasása
Szükséges-e speciális szoftver vagy eszköz az adatok olvasásához, ha igen, mi?
- ♦ Az adatok rendelkezésre állásának garanciái
A felhatalmazott felhasználók számára biztosított az adatokhoz és a funkciókhoz történő hozzáférés, amikor számukra szükséges.
- ♦ Az adatokhoz történő hozzáférés kezelése
Milyen módon fognak hozzáférni az adatokhoz a projektben résztvevők a kutatás alatt?
Pl.: biztonságos nem hálózati környezetben, desktop virtualizáció használatával.
- ♦ Adatcsere, adatmegosztás
Az adatokat cserélik-e, megosztják-e harmadik személyekkel?

4.3. Metaadatok – az adatok dokumentálása és szervezése

A fejezet célja annak meghatározása, hogy a projekt alatt létrehozott és gyűjtött adatokat hogyan fogják leírni, szervezni.

Felelősök és közreműködők: Tudományos és műszaki tájékoztatási szakemberek, Levéltárosok.

- ♦ A metaadatokra vonatkozó formátumok, szabványok
Amennyiben a tudományágnak vannak saját szabványai, elsősorban ezeket kell használni az adatok különböző rendszerek és felhasználók közötti átjárhatóságának biztosítása érdekében.
Pl.: A metaadatok XML-ben lesznek rögzítve, a Data Documentation Initiative (DDI) formátum használatával. A projektben gyűjtött klinikai adatok a CDISC szabvány szerint lesznek dokumentálva.
- ♦ A metaadatok meghatározásának módja és felelősei
Rögzíteni kell, hogy kik, milyen módon és esetleg milyen eszközökkel látják el metaadatokkal az adatokat.
- ♦ Osztályozási fa
Mielőbb gondoskodni kell az adatok osztályozási fájának elkészítéséről, hogy megkönnyítsék az adatokhoz való hozzáférést és a tárolást a projekt során, majd az archiválási időszak végéig.
- ♦ Az adatkészletek elnevezésének szabályai (névkonvenció)
- ♦ Kapcsolódó dokumentáció
Ezt a részt kell feltölteni minden egyéb információval, amelyek hosszú távon biz-

tosítják az adatok érthetőségét, értelmezését: kódok, rövidítések, olvasáshoz szükséges szoftververziók stb.

4.4. A projekt lezárása – Terjesztés

Megosztás, terjesztés, újrafelhasználás

A fejezet célja azon eljárások és esetleges etikai, jogi, technikai garanciák, intézkedések meghatározása, amelyekkel az adatok hozzáférhetővé tételét fogják biztosítani.

Felelősök és közreműködők: Tudományos és műszaki tájékoztatási szakemberek.

- ♦ A megosztás általános elve
Rögzíteni kell az adatok többsége esetében alkalmazandó, általános megosztási elvet, a vonatkozó (jog)szabállyal együtt.
- ♦ A licence típusa
Rögzíteni kell a megosztás és az újrafelhasználás feltételeit, így a licence és az esetleges pénzügyi ellentételezés szabályait.
- ♦ Az újrahazsnosítás lehetőségei
A célzott felhasználói kör (nagyközönség, tudományos közösség, privát szektor stb.) és az alkalmazás, az adatokra épülő fejlesztés lehetőségeinek meghatározása.
- ♦ Az adatokra vonatkozó publikációk léte
A projekt végén határozható meg: igen vagy nem.
- ♦ Repozitórium és hozzáférés
Létezik-e egy tudományos online repozitórium, ahol az adatok tárolhatók? Amennyiben nincs, a Horizont 2020-ban a Zenodóban lehet tárolni az adatokat.

Az érzékeny adatok védelme:

Különböző etikai, jogi, pénzügyi vagy éppen technikai megfontolásokból egyes adatok különleges védelmet igényelnek és így kivételt képeznek a közzététel elvei alól. A fejezet célja, hogy azonosítsa és meghatározza a projekt során létrehozott vagy összegyűjtött érzékeny adatok védelmi kritériumait.

- ♦ Az érzékeny adatkészletek azonosítása
Itt kell meghatározni az érzékeny, közzététel alól kiveendő adatkészleteket (pl. védett tudományos adatok, személyes adatok, stratégiai adatok stb.).
- ♦ Az általános közzététel alóli kivétel igazolása
Itt szükséges részletezni a kivételképzés szükségességét (pl. biztonsági érdek, magánélet védelme, szellemi tulajdon védelme stb.).
- ♦ Védelmi intézkedések
Pl. személyes adatok anonimizálása.
- ♦ Embargóidőszak
Átmeneti védelem, amely meghatározott időszakra lehetővé teszi a közzététel és/

vagy az újrahasznosítás alóli kivételt. Alkalmazása az adatok tulajdonosának döntésétől függ.

5. Az adatok kiválasztása és archiválása

A fejezet célja a projekt során létrehozott vagy gyűjtött, hozzáférhetővé tett vagy nem tett adatok összességére vonatkozik. A fejezet kidolgozása során **megkerülhetetlen** levéltáros közreműködése..

Felelősök és közreműködők: Levéltárosok.

- ♦ Az adatok sorsa a projekt lezárásakor
Meg kell határozni az adatoknak a projekt aktív fázisa utáni sorsát (hozzáférés, eredmények értékelése).
- ♦ Az adatok szelekciója
Nem minden adatot szükséges hosszú távon megőrizni, mindazonáltal a levéltáros szakember véleményét célszerű kikérni az adatok nagy tömegű megsemmisítése előtt. Jelezni kell, hogy mely adatkészleteket célszerű megőrizni közép- vagy hosszú távon, tudományos, jogi okból vagy azért, mert a közvagyon része. Kétség esetén levéltáros szakértőhöz kell fordulni.
- ♦ Az adatok végleges mennyisége/terjedelme
Jelezni kell a végleges becslést mennyiséget.
- ♦ Javasolt megőrzési időtartam
Az időtartam erősen változó, néhány hónaptól a határidő nélküli megőrzésig terjedhet. Figyelembe kell venni a jogszabályi előírásokat és a kiegészítő szabályozást is.
- ♦ Archiválási platform
Milyen platformon lesznek az adatok megőrizve hosszú távon? Az elektronikus tárolóeszközöket el kell fogadnia a felügyeleti szervnek is. Egyes egyetemek saját archiválási módszertannal rendelkeznek.

Irodalomjegyzék

- [1] *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020. Version 1.0.* 11 December 2013. http://www.gsrt.gr/EOX/files/h2020-hi-oa-pilot-guide_en.pdf (Letöltés: 2016. 03. 12.) (Letöltés: 2015. 08. 17.)
- [2] *Guidelines on Data Management in Horizon 2020. Version 1.0.* 11 December 2013. http://www.gsrt.gr/EOX/files/h2020-hi-oa-data-mgt_en.pdf (Letöltés: 2015. 08. 17.)
- [3] Cartier, Aurore – Moysan, Magalie – Reymonet, Nathalie: *Réaliser un plan de gestion de données: guide de rédaction* (V1, 09/01/2015). http://www.univ-paris-diderot.fr/DocumentsFCK/recherche/Realiser_un_DMP_V1.pdf (Letöltve: 2016. 03. 12.) (Letöltés: 2015. 08. 17.)
- [4] *Európa 2020 – Az intelligens, fenntartható és inkluzív növekedés stratégiája.* COM (2010) 2020 végleges. Európai Bizottság, Brüsszel, 2010. 3. 3. <http://eur-lex.europa.eu/legal-content/HU/TXT/PDF/?uri=CELEX:52010DC2020&from=HU> (Letöltés: 2016. 03. 12.) <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:Hu:PDF> (Letöltés: 2015. 08. 17.)
- [5] *Úton a prosperáló adatközpontú gazdaság felé.* COM(2014) 442 final, Európai Bizott-

- ság, Brüsszel, 2014. 7. 2. <http://ec.europa.eu/transparency/regdoc/rep/1/2014/HU/1-2014-442-HU-F1-1.PDF> (Letöltés: 2016. 03. 12.) (Letöltés: 2015. 08. 17.)
- [6] *Az Európai Parlament és a Tanács 2003/98/EK Irányelve (2003. november 17.) a közzféra információinak további felhasználásáról.* <http://eur-lex.europa.eu/legal-content/HU/TXT/PDF/?uri=CELEX:32003L0098&qid=1459919550356&from=HU> (Letöltés: 2015. 08. 17.)
- [7] *Az Európai Parlament és a Tanács 2013/37/EU Irányelve (2013. június 26.) a közzféra információinak további felhasználásáról szóló 2003/98/EK irányelv módosításáról.* <http://eur-lex.europa.eu/legal-content/HU/TXT/PDF/?uri=C ELEX:32013L0037&from=HU> (Letöltés: 2016. 03. 12.) (Letöltés: 2015. 08. 17.)
- [8] *Memo/12/565.* European Commission, Brussels, 17 July 2012. http://europa.eu/rapid/press-release_MEMO-12-565_en.htm (Letöltés: 2016. 03. 12.) http://europa.eu/rapid/press-release_IP-12-790_hu.htm (Letöltés: 2015. 08. 17.)
- [9] *Úton a tudományos információkhoz való jobb hozzáférés felé. A közfinanszírozású kutató-sok előnyeinek megsokszorozása.* COM(2012) 401 final. Európai Bizottság, Brüsszel, 2012. 7. 17. <http://ec.europa.eu/transparency/regdoc/rep/1/2012/HU/1-2012-401-HU-F1-1.Pdf> (Letöltés: 2016. 03. 12.) (Letöltés: 2015. 08. 17.)
- [10] Holl András: Információáradat és hullámlövágás. *Magyar Tudomány*, 2013/4, 473–478. <http://www.matud.iif.hu/2013/04/13.htm> (Letöltés: 2016. 03. 12.) (Letöltés: 2015. 08. 17.)
- [11] Holl András: Adatok, lehetőségek, feladatok a kutatási adatok hozzáférhetőségéről. *Magyar Tudomány*, 2013/10, 1208–1212. <http://www.matud.iif.hu/2013/10/09.htm> (Letöltés: 2016. 03. 12.) (Letöltés: 2015. 08. 17.)
- [12] Pécsi Ferenc: *Házi tartalomgyarak.* <https://www.mediapiac.com/blogzona/mediablog/Hazi-tartalomgyarak/7521/> (Letöltve: 2015. 08. 17.)

Data asset management, data management – Simplified methodology of a subdivision

NAGYNÉ TAKÁCS VERONIKA

The asset management, planning and coordination of available data of public administration is supported by the methodological initiatives which are results of scientific research. In this regard, the policy documents and practical manuals of the European Union also have special importance.

Keywords: scientific information, research data, data asset management, open access, data management plan