

KORMÁNYZATI ADATPOLITIKA A MESTERSÉGES INTELLIGENCIA KORÁBAN

Áttekintés a mesterséges intelligenciában rejlő lehetőségek kiaknázásához szükséges kormányzati adatpolitikai feladatokról

SIKOLYA ZSOLT¹



Abstract

A tanulmány arra tesz javaslatot, hogy a készülő magyar mesterséges intelligencia stratégiának milyen kormányzati adatpolitikai feladatokat kell kitűznie ahhoz, hogy a mesterséges intelligencián alapuló megoldások számára szükséges hatalmas mennyiségű adat rendelkezésre álljon, és ezáltal megteremtődjön az egyik feltétele annak, hogy az ilyen megoldások gazdasági, kormányzati és társadalmi alkalmazása az ország versenyképességének növeléséhez szükséges mértékben terjedjen Magyarországon.

Governmental Data Policy in the Time of Artificial Intelligence – Overview of Tasks of Governmental Data Policy in Order to Exploit the Potential of Artificial Intelligence

The study proposes the tasks of governmental data policy which the upcoming Hungarian AI Strategy should address in order to provide the enormous amount of data needed for AI based solutions, thereby creating one of the preconditions for such solutions to boost their application in the Hungarian economy, government and society to the extent necessary to increase the country's competitiveness.

1. Bevezető

A mesterséges intelligencia megszületését a múlt század '50-es éveire datálják, de igazi áttörést a szükséges és a rendelkezésre álló számítási kapacitás közti jelentős különbség miatt évtizedekig nem tudott elérni ez a technológia. Ezen a területen a mostani évtized hozott jelentős előrelépést, amikor az igényelt olcsó számítási teljesítmény rendelkezésre állása következtében tömegessé válhatott a mesterséges intelligencia alkalmazása. A prognózisok a következő évekre-évtizedekre még nagyobb felfutást jeleznek: azt jelzik, hogy ez a technológia a mindennapi életünknek is nélkülözhetetlen részévé

válik. Joggal beszélhetünk tehát arról, hogy a mesterséges intelligencia korába léptünk.

A mesterséges intelligenciára épülő megoldásoknak hatalmas mennyiségben van szükségük adatokra – legyenek azok a gépi tanuláshoz használt tanítóadatok vagy bementi, illetve tesztadatok. Számos országnak van már nemzeti mesterséges intelligencia stratégiája. Ezek közül például a német stratégia a mesterséges intelligencia alkalmazásával elérhető eredmények központi feltételének és meghatározó tényezőjének tartja az adatok rendelkezésre állását és minőségét. A dán stratégia szerint az adatok minőségén és mennyiségén múlik, hogy mire juthatunk a mesterséges intelligenciával. Ezek alapján érthető, hogy a legtöbb ilyen stratégia külön fejezetben foglalkozik a szükséges adatok rendelkezésre állásával, illetve ebben a kormány szerepével.

A mesterséges intelligencia alkalmazásának tömegessé válása egybeesik az adatgazdaság és az adatvezérelt közigazgatás kiépítésének igényével. Adatokra nyilván nem csak a mesterséges intelligencián alapuló megoldásoknak van szükségük. Ma már elengedhetetlen, hogy a gazdaságban és a kormány-

¹ Villamosmérnök, alkalmazott matematikus. 1967-től az INFELOR-nál számítástechnikai területen dolgozott, 1975-től 1993-ig a VILATI-nál, majd 1996-ig a Geoview Systems Kft.-nél vezetett szoftverfejlesztési projekteket. 1996-tól 2002-ig a Miniszterelnöki Hivatalnál kormányzati informatikával, 2006-ig pedig az Informatikai és Hírközlési Minisztériumnál e-közigazgatással foglalkozott. 2007-től 2013-ig az eGov Tanácsadó Kft. vezető tanácsadója volt. Jelenleg a Magyar Zoltán E-közigazgatástudományi Egyesület elnökségi tagja és a Nemzeti Hírközlési és Informatikai Tanács szakértője.

zásban is adatok alapján szülessenek meg a döntések – függetlenül attól, hogy milyen eszközöket alkalmaznak a döntések előkészítéséhez. Ugyanígy a gazdaság és a kormányzás napi működése sem képzelhető el adatok nélkül. Az Európai Unió már évek óta hangsúlyozza az adatgazdaság kiépítésének fontosságát, és 2017-ben közzé is tette az erre vonatkozó stratégiáját. Az elmúlt évtizedben számos uniós és más ország adott közre adatstratégiát, és az ezekben deklarált adatvezérelt gazdaságépítési és kormányzási elvek alapján hozzák meg sorra intézkedéseiket. Ezek mentén az intézkedések mentén az igencsak „adatéhes” mesterséges intelligencia tömeges alkalmazása elől is elhárulnak az akadályok.

Magyarország még csak a lassú eszmélés állapotában van. A korábbi jogalkotási eredmények és intézkedési tervek eddig alig hoztak értékelhető eredményt. Reményre jogosít, hogy a Kormány az utóbbi időben felismerte a mesterséges intelligencia elterjedésének és egy korszerű adatpolitika megvalósításának az ország versenyképességét jelentősen befolyásoló hatását. Készül az ország mesterséges intelligencia stratégiája, és ebből akcióterv formájában már áttörést jelentő adatpolitikai intézkedésekre vonatkozó tervek is napvilágra kerültek.

Az alábbiakban az imént vázolt kérdésekkel foglalkozunk részletesebben. Megállapításaink alapját részben az a részletes tanulmány képezi, amelyet a Mesterséges Intelligencia Koalíció Adatipar, adatvagyon munkacsoportja készített ebben az évben a Magyar Zoltán E-közigazgatástudományi Egyesület jelentős közreműködésével².

Bár nem kívánunk a tanulmányban használt sok fogalom definiálásában elveszni, mielőtt a részletekben elmélyednénk, egy rövid meghatározást adunk a tanulmány központi fogalmára, az adatpolitikára. Ezt a fogalmat nemzeti szinten értelmezzük – bár értelmezhető szervezeti szinten is –, és azoknak a döntéseknek és intézkedéseknek az összességét értjük alatta, amelyek célja az országban megteremteni az adatok előállításának, kezelésének és felhasználásának koherens feltételeit. Az adatpolitikában a mindenkorinak kormányának meghatározó feladata van, de megvan a maga szerepe a gazdasági, tudományos és civil szférának is. Mindezen szereplők tervszerű műszaki, jogi, tudományos, oktatási és finanszírozási együttműködéséből alakul ki a nemzeti adat-ökoszisztéma.

2. Az Európai Unió kezdeményezései

Az Európai Bizottság 2018-ban tette közzé mesterséges intelligencia stratégiáját³. Ebben leszögezi, hogy az adatokhoz való hozzáférés kulcsfontosságú alkotóeleme a mesterséges intelligencia versenyképes alkalmazásának. A Bizottság elő kívánja segíteni a szükséges adatok elérhetőségét, amihez megfelelő jogszabályi alapot teremt a korábbi PSI-irányelv 2019. évi átdolgozásával létrejött nyíltadat-irányelv⁴, valamint az angol

rövidítésük nyomán FFD-rendeletként⁵ és a GDPR-ként⁶ emlegetett két rendelet. A Bizottság további erőfeszítéseket is kíván tenni annak érdekében, hogy a közadatok és közfinanszírozású kutatási eredmények hozzáférhetővé váljanak. A magánfelek adataira a Bizottságnak és az államoknak nincs ugyan közvetlen ráhatásuk, de ösztönözniük kell azok rendelkezésre állását is. A Bizottság egy adatmegosztást támogató központot is létre fog hozni. Ennek igénybevételel a mesterséges intelligenciára épülő megoldások fejlesztői és használói egyetlen ponton keresztül férhetnek hozzá jogi és technikai szempontból is akadálymentesen a szükséges adattárakhoz.

Amint azt már jeleztük, az adatvezérelt tervezés és működés a mesterséges intelligencia igényein túl is fontos versenyképességi kérdés a gazdaság és a kormányzás számára. Az Európai Bizottság már 1998-ban kiadott egy Zöld Könyvet⁷ a közszféraadatok (magyar terminológiával: közadatok) elérhetőségének fontosságáról, 2011-ben pedig egy nyíltadat-stratégiában⁸ foglalta össze az ezzel kapcsolatos céljait és a tagállamok számára megfogalmazott javaslatát. A 2014-ben közzétett adatközpontú gazdasági stratégia⁹ a nagy adattömegekben (big data) rejlő lehetőségek hangsúlyozásával hívja fel a tagállamok figyelmét az adatközpontú gazdaság megvalósítására, és vállal maga is konkrét európai bizottsági feladatokat. Az adatvezérelt gazdaság meghatározó része az adatgazdaság, amelyet a digitális technológiák révén keletkező adatok létrehozása, gyűjtése, tárolása, feldolgozása, terjesztése, elemzése, kidolgozása, szolgáltatása és felhasználása alkot. A 2017-ben közzétett adatgazdaság-kiépítési stratégia¹⁰ felhívja a figyelmet arra, hogy a már 2015-ben 272 milliárd euró összértéket képviselő uniós adatgazdaság évi 5,6%-os növekedést produkált, és 2020-ra akár 7,4 millió embert is foglalkoztathat. Ez a stratégia elsősorban az adatgazdaság fejlődését hátráltató problémákat veszi számba, és tervet dolgoz ki ezen problémák leküzdésére.

Végül 2018-ban jelent meg az az adatpolitikai dokumentumcsomag, amelynek része a korábban említett mesterséges intelligencia stratégia, központi dokumentuma pedig egy olyan közlemény¹¹, amelyik egy közös európai adattér létrehozását tűzi ki célul. A csomag külön dokumentumokban foglalkozik többek között a tudományos adatokhoz való hozzáféréssel¹², a különböző típusú szervezetek közti adatmeg-

⁵ Az Európai Parlament és a Tanács (EU) 2018/1807 rendelete (2018. november 14.) a nem személyes adatok Európai Unióban való szabad áramlásának keretéről.

⁶ Az Európai Parlament és a Tanács (EU) 2016/679 rendelete (2016. április 27.) a természetes személyeknek a személyes adatok kezelése tekintetében történő védelméről és az ilyen adatok szabad áramlásáról, valamint a 95/46/EK irányelv hatályon kívül helyezéséről.

⁷ Public Sector Information: a Key Resource for Europe. Green Paper on open Public Sector Information in the Information Society COM(1998) 585.

⁸ Nyílt adatok – az innováció, a növekedés és az átlátható kormányzás mozgatórugói. A Bizottság COM(2011) 882 közleménye.

⁹ Úton a prosperáló adatközpontú gazdaság felé. A Bizottság COM(2014) 442 közleménye.

¹⁰ Az európai adatgazdaság kiépítése. A Bizottság COM(2017) 9 közleménye.

¹¹ A közös európai adattér kialakítása felé. A Bizottság COM(2018) 232 közleménye.

¹² A Bizottság C(2018) 2375 ajánlása a tudományos információkhoz való hozzáféréssel és azok megőrzéséről.

² Adatpolitikai stratégiai javaslat az MI-alapú innováció beindítására Magyarországon.

³ Mesterséges intelligencia Európa számára. A Bizottság COM(2018) 237 közleménye.

⁴ Az Európai Parlament és a Tanács (EU) 2019/1024 irányelve (2019. június 20.) a nyílt hozzáférésű adatokról és a közszféra információinak további felhasználásáról (átdolgozás).

osztás lehetőségeivel¹³ és az egészségügyi adatok megosztásának jelentőségével¹⁴.

Az Európai Bizottság tevékenysége azonban nem merül ki stratégiák kidolgozásában, hanem számos konkrét adatpolitikai intézkedéssel is segíti az európai adatgazdaság fejlődését. Ilyen intézkedés volt például a Nyílt hozzáférésű adatok EU-portáljának¹⁵ a létrehozása 2012-ben, amelyen ma már az uniós intézmények mintegy 13 000 magáról a portálról vagy az intézményi rendszerekből letölthető adatkészletének metaadatai találhatóak meg. Az adatok értelmezését, elemzését, hasznosítását interaktív keresési lehetőségek, alkalmazásprogramozási interfészen (API) keresztül elérhető adatszolgáltatások, szabványos, az adatok összekapcsolását is támogató adatmodell, valamint különféle alkalmazások segítik. Ugyancsak jelentős intézkedés volt 2015-ben az Európai Adatportál¹⁶ elindítása, amely 35 európai ország különféle adatportáljairól gyűjti össze és teszi elérhetővé, kereshetővé a közzétett – mára már közel 900 ezer – adatkészletének metaadatait.

Az Európai Unió adatpolitikai intézkedései jogszabályok megalkotásában csúcsozottak ki. Ezek közül is a nyíltadat-irányelvet¹⁷ kell kiemelni, amelynek 2003-ban elfogadott eredeti¹⁸, majd 2013-ban módosított¹⁹ változatát a „*public sector information*” (közzétett adatok) angol rövidítése után még PSI-irányelvnek nevezték, de miután a legújabb átdolgozás már kimondottan ösztönzi, hogy a közadatok „*beépített és alapértelmezett módon*” nyílt, tehát jogi és műszaki szempontból is akadálymentesen elérhető adatok legyenek, az irányelv rövid megnevezését is nyíltadat-irányelvre változtatták. Az irányelv megalkotásának célja a közadatok gazdasági célú újrahasznosításának szabályozása. Alapelvei a digitális innovációt és különösen a mesterséges intelligencia terjedését segítő nyílt adatpolitika szorgalmazása mellett a diszkrimináció és a kizárólagosság elkerülése, az átláthatóság, valamint a közadatok lehetőség szerint ingyenes vagy minimális költségek melletti újrahasznosításának a lehetővé tétele. A legújabb átdolgozás a közzétett adatok szervezeten túl már kiterjed az ún. közvállalkozásokra is, és kötelezővé teszi a külön meghatározandó ún. nagy értékű adatkészletek API-kon keresztüli ingyenes elérhetővé tételét. Az irányelv korábbi változatait már áttültették a tagállamok saját jogrendjükbe, a legutóbbi átdolgozás esetében az áttültetési határidő 2021. július 17.

Az egységes európai téradat-infrastruktúra megteremtését célzó ún. INSPIRE²⁰-irányelv²¹ 2007. évi elfogadása idején ugyan még nem volt hangsúlyos szempont a mesterséges intelligencia igényeinek figyelembevétele, de mivel egyre több mesterséges intelligenciát alkalmazó megoldás használ téradatokat, ezért fontos ezt a jogszabályt is megemlíteni, amit ugyancsak minden tagállamnak át kellett ültetnie saját jogrendjébe. A tagállamoknak 2020-ig fokozatosan 34 témakörben kell geoportálon keresztül, egységes specifikációk szerint kereshetővé és elérhetővé tenniük téradatokat, ill. azok metaadatait.

A nem személyes adatok szabad áramlásáról szóló ún. FFD²²-rendelet²³ igyekszik nevesítetten a mesterséges intelligencia igényeit is szolgálni azzal, hogy tiltja a nem személyes adatok uniós országok közti szabad áramlását akadályozó tagállami korlátozásokat, és meg kívánja könnyíteni az adatok hordozhatóságát.

Külön kell foglalkoznunk az EU általános adatvédelmi rendeletével, amelyre közismert angol rövidítéssel GDPR-ként szoktak hivatkozni. Ez a rendelet nem tekinthető adatpolitikai célú jogszabálynak, mivel az uniós állampolgárok egyik alapjogát, a személyes adataikhoz fűződő jogaik védelmét szolgálja. Ugyanakkor számos tekintetben hatással van a személyes adatok mesterséges intelligencia célú használhatóságára. A mesterséges intelligencia alkalmazása jelentős részben igényel személyes adatokat vagy személyes jellegűtől megfosztott, ún. anonimizált adatokat. Ez vonatkozhat akár a gépi tanulási módszerek esetében szükséges adatokra, vagy akár az alkalmazások működésekor használt bemeneti adatokra is. Ezért a Mesterséges Intelligencia Koalíció Adatipar, adatvagyon munkacsoportja külön tanulmányban²⁴ elemezte a GDPR hatását a mesterséges intelligenciára épülő megoldásokra. Terjedelmi okokból itt most csak a tanulmány néhány eredményére térhetünk ki.

A tanulmány megállapítja, hogy a GDPR 6. cikk (1) bekezdésében megengedett adatkezelési jogalapok többsége elvi vagy gyakorlati okokból nem vagy csak nehézkesen alkalmazható személyes adatok mesterséges intelligencia számára történő felhasználására. A problémákat többek között az alábbiak okozzák:

- a jogos érdek alkalmazásának nehézségei (elsősorban a gépi tanulás során);
- a hozzájárulás visszavonhatósága és beszerzésének nehézsége;
- az automatizált döntéshozatalra, a profilalkotásra és az érintettek tájékoztatására vonatkozó szabályozás, ill. kapcsolódó iránymutatás erős korlátozásai;

¹³ Iránymutatás a magánszektorbeli adatoknak az európai adatgazdaságban történő megosztásáról SWD(2018) 125

¹⁴ A Bizottság COM(2018) 233 közleménye az egészségügy és az ellátás digitális átalakításának a digitális egységes piacon való lehetővé tételéről, a polgárok szerepének erősítéséről és egy egészségesebb társadalom megteremtéséről.

¹⁵ <https://data.europa.eu/euodp/hu/home>

¹⁶ <https://www.europeandataportal.eu/hu/>

¹⁷ Az Európai Parlament és a Tanács (EU) 2019/1024 irányelve (2019. június 20.) a nyílt hozzáférésű adatokról és a közzétett információinak további felhasználásáról (átdolgozás).

¹⁸ Az Európai Parlament és a Tanács 2003/98/EK irányelve (2003. november 17.) a közzétett információinak további felhasználásáról.

¹⁹ Az Európai Parlament és a Tanács 2013/37/EU irányelve a közzétett információinak további felhasználásáról szóló 2003/98/EK irányelv módosításáról.

²⁰ Infrastructure for Spatial Information in the European Community.

²¹ Az Európai Parlament és a Tanács 2007/2/EK irányelve (2007. március 14.) az Európai Közösségen belüli térinformációs infrastruktúra (INSPIRE) kialakításáról.

²² Free Flow of Data.

²³ Az Európai Parlament és a Tanács (EU) 2018/1807 rendelete (2018. november 14.) a nem személyes adatok Európai Unióban való szabad áramlásának keretéről.

²⁴ GDPR konform adatfelhasználás, újrahasznosítás, anonimizálás, és más adatbiztonságot szolgáló technikák, hozzájárulás alkalmazhatósága és a releváns technikák vizsgálata (nyilvánosan nem közzétett tanulmány).

- a statisztikai célú adatkezeléssel kapcsolatos szabályozás határozatlanságai (például hogy milyen adatkezelők esetében értelmezhető a statisztikai célú adatkezelés, és hogy ha a hivatalos statisztikán kívül is értelmezhető, akkor annak feltételeit mi szabályozza);
- a jelentős közérdek fogalmának határozatlansága.

Hasonlóan problémás az adatgyűjtés céljától eltérő célból történő adatkezelésre vonatkozó szabályok alkalmazása. Az adatkezelés feltételei még szigorúbbak a különleges adatok esetében, ami például az egyébként igen fontos egészségügyi alkalmazások esetében jelent gondot.

A tanulmány külön foglalkozik az anonimizálás kérdéssel, ugyanis azok az eredetileg személyes adatok, amelyeket anonimizálnak, vagyis amelyek érintettjei többé nem azonosíthatók, már nem tartoznak a GDPR hatálya alá. Azon túl, hogy a tanulmány részletesen elemzi a különböző anonimizálási eljárások jellemzőit, sőt felhívja a figyelmet arra is, hogy egyes esetekben fennáll annak veszélye, hogy éppen mesterséges intelligencián alapuló algoritmusokkal felfedhető az anonimizáltak tekintett adatok kapcsolata az eredeti érintettekkel, felveti azt az érdekes kérdést, hogy vajon szükség van-e az adatkezelés eredeti célját meghatározó jogalaptól különböző jogalapra magának az anonimizálásnak az elvégzésére. A GDPR külön nem szabályozza ezt a kérdést, és a tanulmány szerzői a GDPR-ből vett analógiák és egyes rendelkezéseiből levont következtetések alapján arra jutnak, hogy maga az anonimizálás nem minősíthető olyan adatkezelésnek, amely csak külön jogalappal végezhető el. A jogszerűséget annak az adatkezelésnek a jogszerűsége dönti el, amelyhez az anonimizálást eszközként használják.

A tanulmány végső soron azt javasolja, hogy a mesterséges intelligencia alkalmazását indokolatlanul nehezítő problémákat, értelmezési bizonytalanságokat a GDPR módosításával, hazai jogalkotással vagy a szükséges állásfoglalások kiadásával orvosolják.

A GDPR kapcsán még megjegyezzük, hogy a szabályozás a személyes adatok esetére tiltja azok szabad áramlásának akadályozását az EU-n belül csupán arra hivatkozva, hogy azok személyes adatok, és ezzel segíti a mesterséges intelligenciát használó alkalmazások terjedését.

3. Egyes uniós tagállamok gyakorlata

Az alábbiakban az uniós tagállamok adatpolitikai gyakorlatát foglaljuk össze. A terjedelmi korlát mellett azért nem terjesztjük ki figyelmünket más országokra is, mert az ímént ismertetett, Magyarországra nézve is kötelező uniós stratégiákhoz és jogszabályokhoz természetesen a tagállamok igazodnak leginkább.

Amint említettük, az uniós országok már elkészült nemzeti mesterséges intelligencia stratégiái mind rendelkeznek adatpolitikai értékeléssel és intézkedési tervekkel, mivel az adatoknak megkülönböztetett szerepet tulajdonítanak a mesterséges intelligencia kutatásában, fejlesztésében és alkalmazásában, ezért először ezeket a terveket összegezzük.

Az összegzéshez Belgium²⁵, a Cseh Köztársaság²⁶, Dánia²⁷, az Egyesült Királyság²⁸, Finnország²⁹, Franciaország³⁰, Hollandia³¹, Málta³², Németország³³, Portugália³⁴ és Svédország³⁵ stratégiáit használtuk fel.

A stratégiák közös jellemzője, hogy nagy hangsúlyt helyeznek a közszféra jogszabály által nem védett adatainak – különösen a téradatoknak – akadálymentes, nyílt, infrastruktúra-szerű elérhetőségére. Vannak országok (például Dánia, Finnország, Svédország), amelyek saját megállapításuk szerint élen járnak ezen a területen, de azok is törekednek a folyamatos előrelépésre, fejlődésre. Ugyanakkor fontosnak tartják, hogy a tudományos és a privát szféra adataiból is minél több legyen elérhető – akár nyílt, akár kereskedelmi adatként. A közpénzekből finanszírozott kutatások által előállított adatoknak lehetőség szerint ingyenesen kell mindenki számára rendelkezésre állniuk. Központi megoldásokkal kívánják támogatni a különböző szereplők közti kooperációt, adatcserét. Mindezt egy nemzeti adatmegosztási platform – esetenként szektorális adatplatformok – keretében tervezik megvalósítani, de nagyon fontosnak tartják az EU kezdeményezéseiben való részvételt is (közös európai adattér, adatmegosztást támogató központ stb.).

A stratégiák szerint kiemelt figyelmet kell fordítani az adatok interoperabilitására, mivel a mesterséges intelligencián alapuló megoldásoknak gyakran van szükségük több forrásból származó adatokra. Adatszabványokat, egységes adatformátumokat kell meghatározni, amihez uniós szintű együttműködésre van szükség. Az országok a közszférában javítani kívánják az adatok minőségét, strukturáltságát is. Fontosnak tartják a nagy adattömegek (big data) technológiájának mesterséges intelligenciabeli használatát – különösen a szenzoradatok esetében. Kiemelt jelentőséget tulajdonítanak az anonimizált személyes adatok használatának, mivel annak révén nagyon sok tanítóadathoz lehetne jutni. Ezért kutatni kívánják a megfelelő anonimizálási technikákat. Több országnál is felmerül az adatokkal kapcsolatos szabályozási, etikai, bizalmi, adatvédelmi kérdések fontossága.

Az egyes országok adottságaiknak megfelelően különböző ágazatokat emelnek ki a mesterséges intelligencia alkalmazásának – és ezért az adatok megosztásának – elsődleges területeiként. Ezen ágazatok közül leggyakoribb az egészségügy, a mobilizáció és a környezet, de érdemes megemlíteni Dánia azon törekvését, hogy mesterséges intelligenciát alkalmazó megoldásokkal támogassák a dán nyelvnek az angollal legalább egyenrangú nyelvtechnológiai használatát.

A tagállamok jelentős része rendelkezik nyíltadat-stratégiával, amelyre a mesterséges intelligencia stratégiát is készít

²⁵ AI 4 Belgium

²⁶ National Artificial Intelligence Strategy of the Czech Republic

²⁷ National Strategy for Artificial Intelligence

²⁸ Artificial Intelligence Sector Deal

²⁹ Finland's Age of Artificial Intelligence

³⁰ For a Meaningful Artificial Intelligence. Towards a French and European Strategy

³¹ Strategisch Actieplan voor Artificiële Intelligentie

³² Malta: The Ultimate AI Launchpad

³³ Artificial Intelligence Strategy

³⁴ AI Portugal 2030

³⁵ National approach to artificial intelligence

országok eleve tudtak építeni. Röviden összefoglaljuk ezen stratégiák közös jellemzőit is, mert Magyarországnak jelenleg ilyen stratégiája sincs, és ezért adatpolitikájának meghatározásakor célszerű ezeket is figyelembe venni. Az alábbi összefoglalást elsősorban az Egyesült Királyság³⁶ (azon belül is külön Észak-Írország³⁷ és Skócia³⁸), Írország³⁹, Lengyelország⁴⁰, Szlovákia⁴¹ adatstratégiái, valamint az uniós tagállamoknak a PSI-irányelv átültetése során létrejött nemzeti jogszabályokból kiolvasható adatpolitikai törekvései alapján készítettük. Megjegyezzük, hogy számos országban nemcsak nemzeti adatstratégia készült, hanem szövetségi államoknak, tartományoknak, egyes régióknak és városoknak is van saját adatstratégiájuk, sőt bizonyos ágazatok is kidolgozták saját stratégiájukat. Az európai nagyvárosok közül Berlin⁴², Edinburgh⁴³, Helsinki⁴⁴ és London⁴⁵ ilyen anyagait is figyelembe vettük.

A nemzeti adatstratégiák visszatérő, közös célja az adatvezérelt gazdaság és közzféra létrejötte, vagyis hogy a gazdaságban és a közzférában a döntések adatokra alapozva szülessenek meg. A közzférában ehhez társul még a polgárok bizalmának növelése is mint fontos cél. Ehhez szükségesnek tartják, hogy a közzféra adatai – különös hangsúllyal a nyers, tehát eredeti, nem feldolgozott adatokra – alapértelmezés szerint nyíltak, vagyis akadálymentesen elérhetők legyenek. Ugyanakkor reálisan számolnak azzal, hogy ennek a célnak a teljes megvalósításához idő kell. A folyamat átláthatósága érdekében több ország is rendszeresen publikálja a közadataik nyílt tételére kidolgozott éves terveit. Egyes országok speciális kormányzati nyíltadat-licenceket dolgoztak ki, mások a Creative Commonst használják. Fontosnak tartják a nyílt adatok gépi feldolgozhatóságát, összekapcsolhatóságát, minőségét, aktualitását, interoperabilitását, elérésükhöz az API-k, szabványos protokollok használatát, az elérhető adatkészletekről metaadatok közzétételét, továbbá mind az adatok, mind a metaadatok tekintetében nyílt szabványok használatát.

Valamennyi adatstratégia egyik alappillére egy nyíltadat-portál működtetése. Ezek a portálok kereshető metaadatokat tesznek közzé az elérhető adatkészletekről, sok esetben maguk az adatok is elérhetők, letölthetők a portálokon keresztül. Vannak portálok, amelyek különféle alkalmazások használatát is lehetővé teszik az adatok közvetlen hasznosítására, esetleg a publikált adatokat hasznosító mobilalkalmazások is letölthetők, vagy ajánlanak letöltésre applikációkat, ill. használatra weboldalakat. Sok esetben ismeretterjesztő anyagok, tananyagok, adatpolitikai jelentőségű hírek is elérhetők az adatportálon, és fórumok is állnak rendelkezésre különböző témák megvitatására, észrevételek, javaslatok té-

telére – a legtöbb adatstratégia ugyanis nagy hangsúlyt helyez a fejlesztők, az adatgazdák alkalmazottainak és a közzféra vezetőinek adatpolitikai tájékoztatására, képzésére, véleménycseréjére.

A stratégiák változatos területek adatainak elérhetőségét tekintik különösen sürgetőnek (közlekedési adatok, környezeti adatok stb.), de abban egységesek, hogy a téradatok akadálymentes rendelkezésre állása rendkívül fontos.

Az adatstratégiákhoz különböző részletezettségű intézkedési terveket tettek közzé. A készítőik tisztában vannak azzal, hogy a kitűzött célokat, feladatokat csak megfelelő szervezettel lehet megvalósítani. Ezért mindegyik ország deklarálja és általában jogszabályba is foglalja az adatpolitikai szervezetrendszerét, felelősségi rendjét. Az alkalmazott megoldások természetesen különbözők.

Még nem tekinthető általánosnak, de több adatstratégia is törekszik a közös adatvagyon-hasznosításba bevonni – akár kereskedelmi jelleggel is – a privát és a tudományos szféra, valamint a közösségi adatgyűjtések (crowdsourcing) adatait, hogy sokféle platform adatait használó integrált alkalmazások jöhessenek létre. Ezen adatok közül egyre nagyobb hangsúlyt kapnak a különféle eszközök szenzoradatai (l. dolgok internete).

4. A hazai adatpolitika helyzete

Előjáróban le kell szögeznünk, hogy a magyar kormányzatnak jelenleg nincs elfogadott adatpolitikája és mesterséges intelligencia stratégiája sem, amely adatpolitikai intézkedéseket helyezett volna kilátásba. A vonatkozó uniós irányelvek átültetése megtörtént. Voltak korábban adatpolitikai kezdeményezések, sőt jogszabályokban, kormányhatározatokban előírt intézkedések is; ezek egy része már elavult, többségük pedig megvalósíthatatlan maradt.

Az utóbbi évek kezdeményezéseiből ki kell emelnünk a 2016-ban a Nemzeti Hírközlési és Informatikai Tanács által készített, általában csak Fehér könyvként hivatkozott stratégiai tervdokumentumot⁴⁶, amely – megalapozó tanulmányával⁴⁷ együtt – elsőként elemezte átfogóan a nemzetközi és magyar adatpolitikai helyzetet, és fogalmazott meg részletes javaslatokat egy egészségesen működő nemzeti adat-ökoszisztéma megteremtésére. A Fehér könyv azóta is hivatkozási alap, a Digitális Jólét Program is átvette több javaslatát, de kormány által elfogadott stratégia vagy intézkedési terv nem lett belőle, és a gyakorlatban sem valósultak meg javaslatai. Ez részben arra vezethető vissza, hogy a magyar kormányzatban jelenleg nincs gazdája az adatpolitikának.

Az európai törekvések nyomán a Fehér könyv is elsősorban a közadatok sokkal jobb hasznosulását célozta meg. A közzférán belüli hasznosulást, amely az adatvezérelt közigazgatás alapja, elsődleges, míg a közzférán kívüli, ún. újrahasonosítást, amely az adatgazdaság kiépülésének egyik meghatározó hajtómotorja, másodlagos hasznosításnak nevezte a stratégiai dokumentum.

³⁶ Open Data White Paper. Unleashing the Potential

³⁷ Open data strategy for Northern Ireland 2015–2018

³⁸ Open Data Strategy

³⁹ Open Data Strategy 2017–2022

⁴⁰ Open Government Data Review of Poland

⁴¹ Stratégia a akčný plán sprístupnenia apoužívania otvorených údajov verejnej správy

⁴² Berlin Open Data Strategy

⁴³ Empowering Edinburgh. Strategy for Open Data

⁴⁴ Helsinki Region Infoshare service opens up city data

⁴⁵ Data for London. A city Data Strategy

⁴⁶ Fehér könyv a nemzeti adatpolitikáról

⁴⁷ Megalapozó tanulmány a nemzeti adatpolitikáról szóló Fehér könyvhöz

Az elsődleges hasznosítás egyik legfontosabb célja, hogy az állam, illetve az államigazgatási szervek döntéseiket adatok alapján tudják előzetesen mérlegelni és utólagosan vizsgálni. Ennek feltételeit egy 2007-ben meghozott törvény⁴⁸ és annak végrehajtási rendelete⁴⁹ szabályozza. A törvény elrendeli, hogy a közzféra adatkezelői a kezelésükben lévő közérdekű adatot ingyenesen továbbítsák az azt közfeladatának ellátása érdekében igénylő államigazgatási szervnek. Személyes adatról, adótitokról vagy egyedi statisztikai adatról is igényelhető másolat, de csak személyes vagy bizalmas jellegűtől visszafordíthatatlan módon megfosztva – tehát anonimizálva. Lehetőséget ad arra is a törvény, hogy több adatkezelő személyes vagy egyedi statisztikai adatot is tartalmazó adatbázisából összekapcsolva kapjon meg adatokat az igénylő szerv. Ilyenkor a személyes vagy bizalmas jellegűtől megfosztott adatokat az ún. „*adatbázis létrehozásáért felelős szerv*” – amely jelenleg a Nemzeti Infokommunikációs Szolgáltató Zrt. (NISZ) – végzi el az összekapcsolást. Az adat alapú döntéshozatalhoz szükséges szabályozás tehát megvan, de minden jel arra utal, hogy a folyamat nem, vagy csak nagyon ritka esetekben működik. A szabályozás szerint például a NISZ-nek számos információról – többek között mintavételi és anonimizálási eljárásokról – kellene tájékoztatást adnia a honlapján az adatkezelők és az adatkérők számára, de ilyen információk nem találhatók a NISZ honlapján. Ami a honlapon szerepel, az 25 olyan anonimizált adatbázis megnevezése, amely a törvény által lehetővé tett összekapcsolással jött létre, illetve amelyet a Költségvetési Tanács adott át közadatként, de felülvizsgálat miatt ezeket az adatbázisokat jelenleg nem áll módjukban kiadni.

A másodlagos hasznosítást a PSI-irányelvet és annak 2013. évi módosítását átültető törvény⁵⁰ szabályozza. A törvény előremutató módon – túllépve a PSI-irányelv rendelkezéseiben – néhány kivétellel minden közadatra vonatkozóan előírja, hogy újrahasonosítási igény esetén az igénylő rendelkezésére kell bocsátani az irányelvben meghatározottak szerint. Felmérések eredményei mutatják, hogy a törvény előírásai nem teljesülnek, és a közzféra szerveinek többsége nincs is felkészülve a rendelkezések teljesítésére. Jellemző, hogy a törvény különféle részletszabályok megállapítására hatalmazza fel a minisztereket, de ilyen részletszabályok mind a mai napig nem születtek.

Már egy 2015-ben elfogadott kormányhatározat⁵¹ elrendelte a nemzeti közadatportál – nemzetközi szóhasználattal nyíltadat-portál – létrehozását, de az elkészültéről hivatalos bejelentés mind a mai napig nem történt. Az utóbbi hónapokban elhangzott szakmai fórumokon, hogy elkészült a portál próbaváltozata, és az interneten meg is található egy 9 adatkészletet publikáló modell⁵², de ezt hivatalosan még

nem emelték nemzeti portál rangjára. A nemzeti adatportál hiánya is oka lehet annak, hogy a közel 900 ezer adatkészlet metaadatait publikáló Európai Adatportálon 35 országból Magyarország az utolsó 51 adatkészlettel.

Sajnos több más nemzetközi adatpolitikai értékelési rangsorban is a sereghajtók között szerepel az országunk. Az Európai Bizottság által évente kiadott Digitális Gazdasági és Társadalom Index (közismert angol rövidítéssel: DESI) Jelentés 2019. évi kimutatásában a Digitális közszolgáltatások dimenzió⁵³ nyílt adatok eleménél Magyarország egyedüli tagállamként úgy jelenik meg, mint amelyik nem közölt magáról értékelhető információkat erről a területről. (A 2018-ban közzétett jelentésben⁵⁴ a 28 tagországból 26-dikak voltunk.) Ugyanez szerepel Magyarország nyíltadat-politikájáról az Európai Bizottság idei e-kormányzati jelentésében is.⁵⁵

Ami mégis reményt keltő, az az, hogy 2018-ban megalakult a mára már több mint 200 szervezeti tagot számláló Mesterséges Intelligencia Koalíció, amelyik – felismerve, hogy az adatok meghatározó szerepet töltenek be a mesterséges intelligenciában – létrehozta az Adatipar, adatvagyon munkacsoportját. A munkacsoport a nemzeti mesterséges intelligencia stratégia adatpolitikai vonatkozású részeinek kidolgozásához igyekszik minél több hasznos anyagot szolgáltatni. Két ilyen tanulmányra már hivatkoztunk. Ki kell emelni még a B2B adat alapú együttműködés című anyagot, valamint a Technológia és biztonság munkacsoport Mesterséges intelligencia kutatási felhő című tanulmányát, amely többek között egy felhőben működő adatpiactér alapjait fektette le. (Ezek a dokumentumok jelenleg még nyilvánosan nem elérhetők.)

A koalíció 2019. október 15-én bejelentett egy mesterséges intelligencia akciótervet, amely a készülő stratégia része is lesz. Ebben szerepel egy GDPR kompatibilis adatpiaci platform létrehozása a keresletet és kínálatot összekötő adatpiacterek építésére. Az akcióterv kilátásba helyezte egy Nemzeti Adatvagyon Ügynökség létrehozását, valamint azt, hogy 2020 tavaszáig kereshetővé váljanak a nem személyhez kötött vagy teljesen anonimizált közadatok. Egy adatvagyon-törvény előkészítése is folyamatban van.

Ugyancsak bizakodásra ad okot, hogy a Kormány 2019. február végén közzétett versenyképességi programjában⁵⁶ a mesterséges intelligencia stratégia kialakításának részeként szerepel a Fehér könyv adatpolitikájának kiterjesztése. Ez arra utal, hogy a Kormány újabb versenyképességi tényezőként tekint az adatpolitikára.

5. Javaslatok a mesterséges intelligencia használatát támogató hazai adatpolitikára

Az alábbi javaslatok nagy része a már hivatkozott Adatpolitikai stratégiai javaslat az MI-alapú innováció beindítására

⁴⁸ 2007. évi CI. törvény a döntéselőkészítéshez szükséges adatok hozzáférhetőségének biztosításáról

⁴⁹ 335/2007. (XII. 13.) Korm. rendelet a döntéselőkészítéshez szükséges adatok hozzáférhetőségének biztosításáról szóló 2007. évi CI. törvény végrehajtásáról

⁵⁰ 2012. évi LXIII. törvény a közadatok újrahasonosításáról

⁵¹ 1310/2015. (V. 21.) Korm. határozat a közadatok széles körű újrahasonosításához szükséges intézkedésekről.

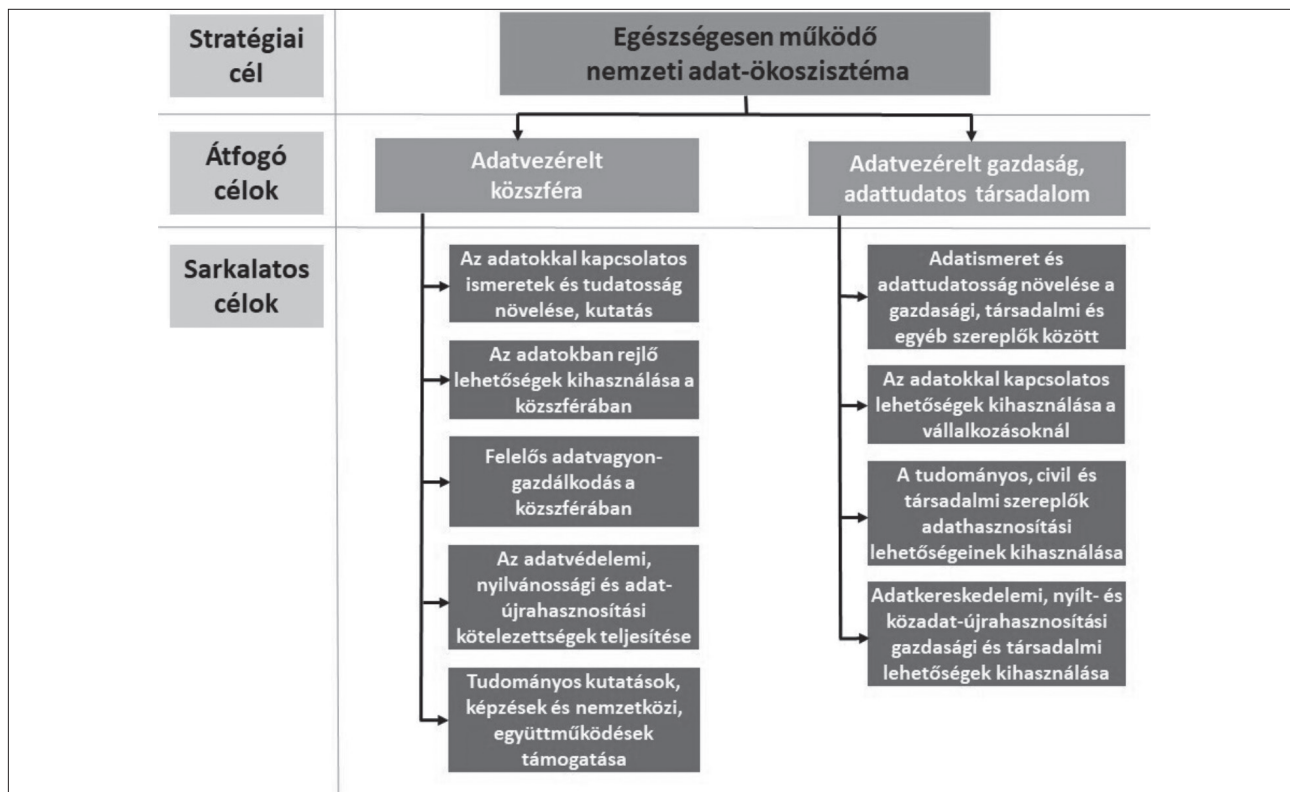
⁵² <http://vm.ik.bme.hu:4339/>

⁵³ Digital Economy and Society Index Report 2019 – Digital Public Services.

⁵⁴ Digital Economy and Society Index Report 2018 – Digital Public Services.

⁵⁵ eGovernment Benchmark 2019 – Factsheets.

⁵⁶ Program a versenyképesebb Magyarorszáért, 2019.



Az adatpolitika célrendszerének felső három szintje

Magyarországon című dokumentum javaslatának összefoglalása. A dokumentum a Fehér könyvben megfogalmazott célokból indult ki, és azokat a mesterséges intelligencia és más, új kezdeményezések igényeihez igazítva dolgozott ki egy olyan célrendszert, amelyben a központi stratégiai cél továbbra is egy egészségesen működő nemzeti adat-ökoszisztéma létrehozása. Ez a központi cél két átfogó célra bomlik: az adatvezérelt közsféra, valamint az adatvezérelt gazdaság, adattudatos társadalom kialakulására. Mivel mindkét átfogó célterületen egyre növekvő szerepet kap a mesterséges intelligencia, ezért a teljes célrendszer megvalósítása a mesterséges intelligenciára épülő megoldások elterjedését is segíti. Az átfogó célok további sarkalatos célokra, azok pedig specifikus célokra bomlanak. A célrendszer felső három szintjét a fenti ábra szemlélteti.

A célok megvalósításának alapfeltétele, hogy a Kormány fogadjon el egy, a javasolt célrendszert is magában foglaló adatpolitikai stratégiát és intézkedési tervet. Mivel az adatpolitikai célok megvalósítására nemcsak a mesterséges intelligencia elterjesztéséhez van szükség, ezért célszerű lenne, ha az adatstratégia önálló lehetne, de a jelenlegi törekvésekre tekintettel megoldás lehet az is, ha a mesterséges intelligencia stratégia részstratégiájaként fogadják el.

A következő feladat a stratégia megvalósításához szükséges szabályozási környezet megteremtése. A szabályozási feladatok közül itt csak néhányat sorolunk fel. A megszabott határidőre természetesen át kell ültetni a hazai jogrendbe a nyíltadat-irányelv átdolgozott és új rendelkezéseit. Az FFD-rendelet is szükségessé teheti jogszabályok módosítását. Szabályozást igényel az adatpolitikai szervezetrendszer létrehozása, a kötelezettségek, feladat- és hatáskörök kijelölése. Lehetnek olyan

kötelezettségek is, amelyek megszegésének esetére szankciók alkalmazásának lehetőségét is el kell rendelni. Egy bizonyos időtávon belüli feladatok végrehajtásának számonkérhetőségéhez szükség van feladatkitűző szabályozásra. A kilátásba helyezett új adatvagyon-törvény feltehetőleg tisztázni fog olyan jogdogmatikai kérdéseket is, mint hogy egyáltalán milyen jogintézmények alkalmazhatók az adatokra, azon belül is a közadatokra. Várható, hogy az új és a módosított jogszabályok olyan fogalmi kérdéseket is tisztába fognak tenni, amelyek körül jelenleg fogalmi zavar van. A közsféra adatai közül a hatósági nyilvántartások vezetésére is célszerű lenne közös normákat kidolgozni. Korábban az anonimizálással kapcsolatosan is említettünk tisztázatlanságokat, amelyek hátráltatják mesterséges intelligenciára alapuló megoldások fejlesztését, ill. alkalmazását. Célszerű lenne ezeket jogszabályi úton – vagy legalább állásfoglalások formájában – elosztatni. Az adatokkal kapcsolatos etikai, felelősségi, bizalmi kérdések egy része is szabályozásra szorul.

Az adatstratégia végrehajtásának sikerességét alapvetően határozza meg a létrehozandó szervezetrendszer. A szervezeti hierarchia tetején annak a miniszternek kell állnia, akinek az ilyen irányú felelősségi és hatáskörét jogszabály írja elő. Mivel a kormányzati adatpolitika gyakorlatilag valamennyi tárca érinti, célszerű egy tárcaközi bizottság vagy munkacsoport létrehozása is, amelyben a többi tárca is képviseltetni tudja magát. De ez a szakterület kihat más hatalmi ágakra, a gazdasági, tudományos és civil élet szinte egészére, ezért egy szélesebb körű szakmai tanács működtetése is kívánatos lenne. Adott esetben ez lehet akár a Mesterséges Intelligencia Koalíció Adatipar, adatvagyon munkacsoportja is. A konkrét kormányzati adatpolitikai feladatok végrehajtását egy külön

szervezetre kellene bízni. Erre alkalmas lehet a mesterséges intelligencia akciótervben bejelentett Nemzeti Adatvagyon Ügynökség. A szervezetnek hatósági, szakmai, koordinációs, tudatosítási, ismeretterjesztési és az uniós feladatok miatt nemzetközi kapcsolattartási feladatokat is el kell látnia. A közadatokkal kapcsolatos feladatok minden közfeladatot ellátó szervet érintenek, ezért ezeknél a szerveknél – legalább is a nagyobbaknál – adatpolitikai felelősöket kell kijelölni.

Említettük már, hogy más országokban az adatpolitika egyik meghatározó pillére a nemzeti adatportál. Ennek létrehozásáról, működtetéséről és az érintetteknek a portállal kapcsolatos kötelezettségeiről jogszabályban kell rendelkezni – mint ahogy arról is, hogy a portálnak együtt kell tudni működni az Európai Adatportállal. Ez a portál adhat helyet az egész közszféra teljes adatkataszterének is, amelyre ugyancsak nagy szükség lenne. Rendezni kell a portál viszonyát a jelenlegi közadatkereső rendszerrel⁵⁷, amelyről az Infotv.⁵⁸ rendelkezik, valamint a mesterséges intelligencia akcióterv keretében bejelentett adatpiaci platformmal. A portál megvalósításánál át kell venni az igen nagy számban rendelkezésre

álló nemzetközi tapasztalatokat, és minden jó ötletet fel kell használni.

Az Adatpolitikai stratégiai javaslat az MI-alapú innováció beindítására Magyarországon című dokumentum még további 20 alapvetően fontos intézkedésre tesz javaslatot, amelyek között szerepelnek pl. kutatási, képzési, ösztönzési, pénzügyi, szervezési, kommunikációs, nemzetközi, szabványosítási és egyéb szakmai feladatok is. Az intézkedésekhez a felelősségre, finanszírozásra és az indikátorokra vonatkozó javaslatok is társulnak.

Az adatigények változása és a technika, azon belül is a mesterséges intelligencia gyors fejlődése miatt nagyon fontos az adatpolitikai stratégia megvalósításának folyamatos monitorozása – ehhez nyújtanak segítséget a javasolt indikátorok –, és a tapasztalatok alapján szükség lehet a stratégia időnkénti módosítására is. A stratégia megvalósításának kiszámíthatóságát és nyomon követését is segítené, ha a közszférában közzéteendő adatkészletekre vonatkozóan éves tervek, a stratégia megvalósításának előrehaladásáról pedig évenkénti értékelések készülnének, amelyek nyilvánosan is megjelenének.

⁵⁷ <http://kozadat.hu/kereso/>

⁵⁸ 2011. évi CXII. törvény az információs önrendelkezési jogról és az információszabadságról