

KSH/ 3138-3/2019

Emlékeztető

**az Országos Statisztikai Tanács (OST) és a Nemzeti Statisztikai Koordinációs Testület (NSKT)
2019. december 9-ei együttes üléséről**

Napirendi pontok	Időkeret
1. A Mesterséges Intelligencia Központban elért eredményekről, különös tekintettel az Adatpiac megoldásokra Előterjesztő: Szertics Gergely (Mesterséges Intelligencia Központ)	15.00 - 15.25
2. Mesterséges Intelligencia Központ statisztikai célú használatának lehetőségeiről Előterjesztő: Mag Kornélia (KSH)	15.25 - 15.50
3. Az Oktatási Hivatal adatházisai kutatási célokra:	15.50 - 16.10
3.1. A Diplomás Pályakövetési Rendszer fejlesztési eredményei. A kutatási lehetőségek feltárása a társadalmi hatások kutathatósága Előterjesztő: Goldfárthné Veres Edit főosztályvezető (Oktatási Hivatal)	
3.2. Az Oktatási Hivatal kompetencia-mérés adatházisából bővebb háttér adatok átvételi igénye, kutatási célokra Előterjesztő: Balácsi Ildikó főosztályvezető (Oktatási Hivatal)	16.10 – 16.30

Az ülésen részt vettek

az OST tagjai, valamint az NSKT tagjai, illetve néhány esetben a képviselőtüket ellátó személyek (jelenléti ív szerint).

a KSH képviselőiben: Freid Mónika a KSH elnökhelyettese, Németh Eszter a KSH elnökhelyettese, Mag Kornélia a KSH elnökhelyettese, továbbá a KSH érintett főosztályvezetői vagy helyettesei, az OST titkára, az OST és NSKT Titkárságának munkatársai.

Dr. Papp Attila az OST Elnöke köszöntötte az OST tagjait és a megjelenteket.

Freid Mónika a KSH elnökhelyettese az OST újonnan kinevezett tagjainak átadta a megbízóleveleket:
Dr. Dargay Eszter (a Belügyminisztérium képviselőiben) **újonnan kinevezett OST tag** távollétében megbízott helyettesének, Dr. Szabó Eszternek,
Győri Péter Pál (a Honvédelmi Minisztérium képviselőiben) **újonnan kinevezett OST tagnak.**

Majd az NSKT újonnan kinevezett tagja **Szabó-Gulyás Krisztina** (a Pénzügyminisztérium képviselőiben) részére is átadta a megbízólevelet.

1. napirendi pont: A Mesterséges Intelligencia Központban elért eredményekről, különös tekintettel az Adatpiac megoldásokra

Szertics Gergely a Mesterséges Intelligencia Koalíció (MIK) szakmai vezetője röviden bemutatta a Koalíciót, amely 2018 októberében alakult meg és jelenleg 225 tagszervezettel és több mint 800 delegált szakértővel rendelkezik. Célja, hogy létrehozza Magyarország Mesterséges Intelligencia Stratégiáját, amelynek elfogadása 2020 január elején várható. Az előadásban a Stratégia néhány szempontjának ismertetésére kerül sor, bemutatásra kerül, hogy melyek azok a területek, amelyek az adatgazdaság szempontjából különösen érdekesek lehetnek.

A Stratégia általános struktúrája széles körű alapozó pillérei, a társadalom és gazdaság egészét célzó alapintézkedések: a Mesterséges Intelligencia (a továbbiakban: MI) értéklánc, amely magában foglalja az adatgazdaság beindítását, a kutatás- fejlesztés – innovációt valamint az alkalmazások ösztönzését. Mindezek kiegészülnek a teljességet célzó pillérekkel, mint a szabályozás és etikai keretek, az infrastruktúra fejlesztés valamint a kompetencia fejlesztés. Ezt követően ismertetésre kerültek azok a szektorok, amelyek esetében a Stratégia kiemelt figyelmet fordít a mesterséges intelligencia alkalmazására vagy annak fejlesztésére. Vannak olyan szektorok, amelyeknél a MI technológia fejlesztése a kitűzött cél (pl.: gépi látás, hálózatkutatás), még más szektorok esetében a szektorális hatékonyságfejlesztés kap hangsúlyt a Stratégiában pl.: gyártás, energetika. Az agrárszektor, az egészségügy és a közlekedés-logisztika területén mind a két szempontból történő fejlesztés szükséges.

Az adatgazdaság beindítása stratégiai pillér bemutatása során Szertics Gergely kiemelte a közadatportált és az adatpiacot, amelyek kifejlesztése már megkezdődött. A közadatportál az államilag létrehozott adatok gyűjtőhelye lesz, ahol megkereshető lesz, hogy mely adatot, hol, melyik szervtől lehet kikérni. Az adatpiac a magánpiaci szereplők adatvagyonának megosztására fókuszál. A közadatportál létrehozása jogszabályi előírások alapján kötelező, míg a magánadatok hozzáférhetővé tétele az adatpiac platformon önkéntes. Cél az adatok másodlagos felhasználásának ösztönzése. Első körben olyan módon, hogy ismertté váljon, melyik adatgazdához lehet fordulni az adatkérésekkel, majd infrastruktúra építése, hogy az adatgazdaság szereplői (adatkeresleti, adatkínálati és adat - transzformációs szereplők) egymásra találhassanak.

Kialakításra kerül egy ajánlási keretrendszer is, amely segítséget nyújt abban, hogy milyen típusú adatokat, milyen célokra, milyen transzformációt követően lehet felhasználni, főleg, ha személyes adatok is érintettek. Cél, hogy a nem személyes vagy teljes mértékben anonimizált adatokat lehessen a piacra kivinni. Készül egy adatvagyon kerettörvény is, amely kialakítja az adatpiaccal, adatok vagyonelemként történő kezelésével kapcsolatos fogalomrendszert, ennek magánjogi alapjait, valamint az adat vagyonosításának pénzügyi jogi kereteit.

Az infrastruktúra fejlesztés kapcsán átfogó cél a széleskörű stratégiai pillérek alap infrastruktúrájának biztosítása, szem előtt tartva azt, hogy a jövőben az elérhetővé tett adatok is infrastruktúrájának minősülnek majd. Fontos a fizikai infrastruktúra rugalmas használatának megteremtése.

Végül az Adattárca transzformatív projekt került bővebb bemutatásra az előadás során. Egy-egy állampolgárról jelenleg nagyon sok tranzakciós dimenzióban keletkezik adat (pl.: egészségügy, oktatás, banki adatok). Jelenleg, a gyakorlatban a személyes adatokról való önrendelkezésre nincs sok lehetőség. Cél, hogy az egyes állampolgárok is részt tudjanak venni tudatosan az adatgazdaságban, amelynek jelenleg teremtik meg a kereteit (másodlagos adatfelhasználás). Az adatok másodlagos felhasználásba való bekapcsolódása az adatgazdaság alapjait teremti meg, mind a céges szektorban, mind a közsférában, mind a magánemberek körében, amelyeket különböző időtávokon belül lehet megvalósítani.

Dr. Papp Attila kérdésére Szertics Gergely elmondta, hogy a Belügyminisztérium szándéka szerint a közadatportál átadására már az év végén sor kerül. Jelenleg tesztelés alatt áll, jövő év elejére előreláthatóan működni fog. Ehhez kapcsolódóan készül egy közadatkataszter, amely a különböző közadatok gyűjteményét állítja össze, első körben a jogszabályok alapján. Az adatok kutatási célra való hozzáférése, ennek részletei a feleken múlnak, a MIK a kommunikációs csatorna megteremtését, a felek egymásra találását szolgálja. A közadatok igénylése az eddigi gyakorlatnak megfelelően zajlik tovább.

Páll Zsombor (Agrárminisztérium) az iránt érdeklődött, hogy van-e olyan adatbázis, ahol feldolgozott formában megtalálhatók, letölthetők az adatok. A kérdésére Szertics Gergely kifejtette, hogy jelenleg az megoldott, hogy az adatot kereső megtalálja az adatgazdát, akitől közvetlenül elkéri majd a szükséges adatokat, nincs olyan platform, ahol a szükséges adatok feltöltésre kerülnek és azt a döntéselőkészítő állami szervek közvetlenül, megkeresés nélkül elérik. Egy nagy fejlesztés folyamatban van az állami adatok állami szervek egymás közti kikérése tekintetében, amelynek célja egy olyan csatorna építése, amelyen keresztül közvetlenül lehet átküldeni egymás részére adatokat. Ennek megvalósulásához azonban előtte Stratégia majd egy akcióterv elfogadására van szükség.

2. napirendi pont: A Mesterséges Intelligencia Központ statisztikai célú használatának lehetőségeiről

Mag Kornélia a KSH elnökhelyettese előadásának célja, a hivatalos statisztika szempontjából bemutatni az előző előadásban elhangzottakat. A Központi Statisztikai Hivatal alapító tagja a Mesterséges Intelligencia Koalíciónak, lehetőségként tekint a Stratégiára és elemeire. Prezentációjában röviden bemutatta a digitális statisztikát. A digitalizáció jelent egyrészt egy új jelenséget, amelyet a statisztikának mérnie kell a gazdaság és társadalom részeként, másrészt a digitalizáció új adatforrásokat hozott létre, együtt jár vele az adat vagyonosítása, értékesítése. A másodlagos adatforrások köre kibővült az adminisztratív adatforrásokon kívül a magántulajdonban lévő adatforrásokat is magában foglalva. Emellett a másodlagos adatforrások jellege is bővült (aggregált adat, mikroadat, big data). Harmadrészt új technológiák és módszertanok jelentek meg (mesterséges intelligencia különböző ágai, gépi tanulás), amelyek alkalmazása új távlatokat nyit meg a hivatalos statisztika szervei számára. Fentiek tükrében kell vizsgálni azt, hogy a mesterséges intelligencia mit jelent a hivatalos statisztika számára.

A hivatalos statisztika viszonylag korán reagált a digitalizációra, így az új jelenségek mérésére (pl.: OECD - digitális transzformáció mérése; uniós szinten DESI – gazdasági és társadalmi digitalizációs index; társadalmi és gazdasági szervezetek digitalizációjának hazai mérése, azok bővítése – PIAAC, Agrárcenzus).

Mag Kornélia ezt követően a Big Data jelenségével kapcsolatban, a hivatalos statisztikában történeteket ismertette röviden világ-és európai szinten. Kiemelendő uniós szinten a 2013-ban kelt Scheveningen Memorandum, amely uniós szintű akciótervet hozott létre a témakörben, azóta számos képzésbe épült be, versenyek, kutatások, általános szakmai anyagok tárgyát képezi a Big Data.

A mesterséges intelligencia és a gépi tanulás a hivatalos statisztika adatelőállítási és adatközlési gyakorlatában hozott jelentős változásokat, illetve a Big Data adatforrások használatának alapját képezik. A Szertics úr előadására visszautalva kiemelte a széleskörű alapozó pillérek elkészültének fontosságát, főként a szabályozási környezetet a hivatalos statisztika szempontjából. A stratégia kiemeli az állami szolgáltatások hatékonyságának növelését, amely szintén jó lehetőség a hivatalos statisztika számára. Ezt követően ismertetésre került a hivatalos statisztika szempontjából legfontosabb szempontok az adatgazdaság beindítása stratégiai pillérhez kapcsolódóan, kiemelve a magánadatok piaci alapon való kereskedelmének megteremtését, a rendelkezésre álló magánadatok megismerését.

Mag Kornélia előadása végén összefoglalta a Mesterséges Intelligencia Stratégia által a hivatalos statisztika számára hozott változásokat: a szabályozott környezet elősegíti majd a másodlagos felhasználást, a közadatkataszter pedig az adminisztratív adatok statisztikai célú felhasználását, átfogóbb képet ad a rendelkezésre álló magánadatokról valamint keretet biztosít azok statisztikai felhasználásához. Ezáltal a hivatalos statisztikai szolgálatok eredményesebbé válhatnak. Ismertette az „okos statisztika” fogalmát is.

3.1. napirendi pont: A Diplomás Pályakövetési Rendszer fejlesztési eredményei. A kutatási lehetőségek feltárása a társadalmi hatások kutathatósága

Goldfárthné Veres Edit az Oktatási Hivatal főosztályvezetője a kutatási adatbázisokról és eredménytermékekről adott tájékoztatást. Az Oktatási Hivatal két elnökhelyetteségén folyik kutatási adatok gyűjtése illetve szolgáltatása köznevelés és felsőoktatás témakörében. A Felsőoktatási Információs Rendszer önálló, komplex adatrendszer, amely OSAP statisztikát is ad. A rendszer felsőoktatási intézményekre vonatkozó adatelemeket, hallgatókra vonatkozó adatelemeket tartalmaz. A Felsőoktatási Elemzési Főosztály nem végez hatósági feladatokat, kutatási tevékenységet lát el, a kollégák többsége szociológus és gazdasági informatikus végzettségű. A főosztály által kezelt adatbázisok többsége a diplomás pályakövetéshez kapcsolódik. A Felsőoktatási Elemzési Főosztály kétféle jellegű adatbázissal dolgozik, a legtöbb adatforrás adatintegrációból származik, ugyanakkor van néhány kérdőíves, anonim survey kutatásuk is. Ismertetésre kerültek egyes kiemelt kutatási témák.

A hallgatói életút-kutatás kezdete a felvételhez kötődik, a felsőoktatási intézménybe felvett hallgatóról a köznevelési adatbázisból máris rengeteg információ rendelkezésre áll, de vannak információk a hallgatók szociális háttéréről, OKM eredménypontszámairól, amelyek felhasználhatóak a kutatásaikhoz. Az elképzelések szerint a felsőoktatásba való belépéskor bemeneti-kompetenciamérés valósul meg, a tanulmányi életút alatt tematikus hallgatói kutatások zajlanak, a végzés után 1, 5 és 7 évvel pedig kérdőíves vizsgálatokkal illetve adatintegráción belül teljes alapsokasággal munkaerő-piaci pályakövetést végeznek. Jelenleg csak azokról tudnak információt nyerni és szolgáltatni, akik a felsőoktatás bármely ágán beléptek a felsőoktatási rendszerbe.

A Diplomás Pályakövetési Rendszer (DPR) 2009-ben, kérdőív formában született meg, a jelenlegi DPR tartalmilag teljesen megújult az évek során. Ismertetette a DPR megújításának folyamatát és az ahhoz vezető okokat is (elszigetelten fejlődő kutatási modulok).

Röviden bemutatásra került a DPR Adminisztratív Adatbázisok Egyesítése kutatási modul valamint a Hallgatói DPR felmérés, amellyel 2020-tól szeretnének képet biztosítani arról, hogy milyen lesz a jövő munkavállalója. Ismertette a jelen munkavállalójának – álláskeresési, állásvállalási attitűd - feltérképezését végző kérdőíves kutatást is (Frissdiplomás Kutatás).

Az Eurostudent VII nemzetközi kutatás online kérdőív formában, 7500 válaszadóval zajlott le, 2019. tavaszán. A felmérés leginkább a hallgatók szociális helyzetére fókuszált.

Goldfárthné Veres Edit beszámolt a főosztály fejlesztéséről (www.diplomantul.hu) amellyel nyilvánossá tették kutatási adatbázisának az alapadatokként definiált körét. A beérkező adatkérések 90-95%-a ugyanarra az adatkörre vonatkozik, ezért hozták létre a nyilvános keresőt, amelynek célközönsége a hallgatók, szülők, felsőoktatási intézmények, sajtó, amelyek érdeklődési igénye kielégíthető ebből a forrásból.

Molnár György (MTA) kérdésére Goldfárthné Veres Edit elmondta, hogy a DPR kutatási célú igénybevételére kidolgozott protokoll létezik, adatkérési lapon keresztül kell benyújtani az adatigényt,

felajánlva a személyes konzultáció lehetőségét is. A négy óránál több időt igénybe nem vevő, teljesíthető adatkérések igényesek.

3.2. napirendi pont: Az Oktatási Hivatal kompetencia-mérés adatbázisából bővebb háttér adatok átvételi igénye, kutatási célokra

Balácsi Ildikó az Oktatási Hivatal főosztályvezetője röviden bemutatta, hogy milyen köznevelés – értékelési programok léteznek ma Magyarországon. A nemzetközi mérések közül kiemelte a PISA mérést. Majd az országos programok, azok részprogramjai kerültek ismertetésre, melyek között szerepel az előadás témája, az Országos Kompetenciamérés (OKM).

A kompetenciamérés hatékony eszközt biztosít az intézményi önértékeléshez: visszajelzést ad arról, hogy az iskola milyen eredménnyel közvetíti a társadalom által elvárt tudást, biztosítja, hogy képes legyen összehasonlítani magát más hasonló iskolákkal, a tanárok pedig visszajelzést kaphatnak arról, hogy a tanulók az általuk átadott tudást mennyire tudják alkalmazni. Cél továbbá a mérési – értékelési kultúra terjesztése, a mérés módszereinek megismertetése a pedagógusokkal, valamint az iskolafenntartók informálása saját intézményeik eredményeiről. Az elsődleges terméke a mérésnek az iskolai és telephelyi jelentés. Az eredmények nyilvánosan is elérhetőek. A tanárok bővebb tartalomban, feladatonként is vizsgálhatják az eredményeket. A jelentések értelmezését segítő, az alkalmazott mérésmetodikai módszereket bemutató leírások is elérhetőek a visszajelzés mellett. Az országos szintű eredményekről elemzések készülnek.

Ezt követően a kompetenciamérés alapdokumentumainak ismertetésére került sor (jogszabályok, mérések szakmai dokumentumai: tartalmi keret - bemutatja mely az a két terület, amelyet a vizsgálat mér, illetve az ehhez szükséges feladatok jellegét, típusát; a mérés minden szakaszára kiterjedő technikai leírás; tesztfüzetek, javítókulcsok, háttérkérdőívek – a mérés után nyilvánosságra kerülnek; a Feladatok és jellemzőik kötetekben a feladatok ismertetése, ezt követően pedig a megoldókulcs és a megoldottsági mutatók kerülnek bemutatásra. A FIT jelentések (fenntartói – intézményi telephelyi jelentések) a mérés fő termékei. Ezeket egészítik ki a mérés adminisztrációs dokumentumai.

Ezt követően az OKM háttérkérdőiveiről esett szó, amelyek a kutatói fájlok alapját képezik. A tanulók nemcsak tesztfüzetet töltenek ki, háttérkérdőívek is kapcsolódnak a vizsgálatokhoz, amelyeket a tanulók valamint a telephelyi – és intézményvezetők is kitöltenek. Céljuk a tanulói teljesítménnyel összefüggő tényezők vizsgálata. A tanulói kérdőív alapvetően a tanulói háttér szociológiai – gazdasági – kulturális jellemzőit vizsgálja, kitöltésük önkéntes, a kitöltési arány 70-80%-os. A telephelyi kérdőív az iskola általános jellemzőire kérdez rá (fizikai állapot, felszereltség, erőforrások) illetve az egyes képzési formák jellemzőit vizsgálja. Az intézményi kérdőív az intézményigazgató jellemzőire, az alkalmazottak számára (pedagógusok fluktuációja) kérdez rá, az iskola költségvetésére vonatkozó kérdések 2014 –től a Klebelsberg Intézményfenntartó Központ létrejöttével kikerültek ebből a kérdőívből.

A tanulói kérdőívből készül a családi háttérindex, amely a szülők iskolai végzettségét, az otthon található könyvek számát és egyéb kulturális - vagyoni tényezőket foglal össze egyetlen mutatószámába.

Végül a kutatói adatbázisra tért ki előadásában. A kutatói adatbázis négy szinttel rendelkezik: fenntartó, intézmény, telephely és tanuló szinttel, mindegyiknek vannak azonosító adatai, amelyek segítségével össze lehet kötni az adatbázisokat. Kiemelte, hogy fenntartói szintű adatfájlok külön nem léteznek, ezek az intézményi szintű adatfájlokban kerülnek megjelenítésre. Ezt követően röviden összefoglalta a tanulói szintű, a telephelyi szintű valamint az intézményi szintű adatfájlok főbb részeit. A GDPR előírásainak megfelelően, mivel a tanulói szintű adatfájlok nem anonim, hanem álnevesített adatok, nem adhatóak ki jelenleg adatvédelmi okból. (Az intézményi és telephelyi szintű adatok közadatok) Az anonimizálás nagyon nehéz a tanulói szintű adatfájlok tekintetében, megoldást jelenthet a tanulók azonosítását lehetővé tévő adatok törlése valamint a folytonos változók kategorizálása, így csökkentve a felfedési kockázatot. Emellett törvénymódosítási javaslat irányul a kutatói adatfájlok kutatási célból történő hozzáférhetővé tételére.

Horn Dániel (MTA) kérdésére Balácsi Ildikó elmondta, hogy az Oktatási Hivatal vizsgálta annak lehetőségét, hogy kutatószobai körülmények között tegyék elérhetővé kutatási célra adataikat, de jelenleg nincs megfelelő kapacitásuk (emberi erőforrás) és nem született elvi döntés ennek megvalósítására, ezért választották a korábban ismertetett megoldást a tanulói szintű fájlok elérhetővé tételére.

Dr. Papp Attila kérdésére Balácsi Ildikó beszámolt arról, hogy a Köznevelési törvény 80.§-a 2014. évtől úgy rendelkezik, hogy az Oktatási Hivatal, belső használatra, az adatbázisában lévő adatokat mérési azonosítóval szcenélyazonosításra alkalmatlan módon összekötheti és elemezheti az oktatási rendszer eredményességének javítása céljából.

Zárás:

Dr. Papp Attila tájékoztatást adott arról, hogy az emlékeztetővel együtt az előadás anyagokat elektronikus úton az OST Titkárság és az NSKT Titkárság munkatársai megküldik az OST és NSKT tagjai és állandó meghívottjai részére. Megköszönte a tagok és a jelenlévők részvételét, munkáját és az ülést bezárta.

Budapest, 2019. december 19.



Dr. Papp Attila
az OST elnöke



Mezősiné Rózsár Erika
az OST titkára



Dr. Nagy Eszter
az NSKT titkára

Az emlékeztetőt kapják:

*Dr. Vukovich Gabriella a KSH elnöke
Freid Mónika a KSH elnökhelyettese
Mag Kornélia a KSH elnökhelyettese
Németh Eszter a KSH elnökhelyettese
Dr. Kincses Áron a KSH elnökhelyettese
az OST és NSKT tagjai és állandó meghívottjai*