

## Tematikus összeállítás

# SZÁMÍTÓGÉPES TÁRSADALOMTUDOMÁNY

## COMPUTATIONAL SOCIAL SCIENCE

VENDÉGSZERKESZTŐK: RUDAS TAMÁS, PÉLI GÁBOR

### BEVEZETŐ

### INTRODUCTION

Rudas Tamás<sup>1</sup>, Péli Gábor<sup>2</sup>

<sup>1</sup> az MTA doktora, az MTA Társadalomtudományi Kutatóközpont főigazgatója, rudas.tamas@tk.mta.hu

<sup>2</sup> az MTA külső tagja, az MTA Társadalomtudományi Kutatóközpont főigazgató-helyettese, peli.gabor@tk.mta.hu

### ÖSSZEFOGLALÁS

Bevezetőnkben elsőnek az MTA keretében a közelmúltban létrejött *Számítógépes Társadalomtudomány* témacsoport céljait ismertetjük, majd röviden áttekintjük, hogy a tematikus blokk írásai miként járulnak hozzá a célok megvalósulásához.

### ABSTRACT

First, we introduce the goals of the *Computational Social Science* thematic group that have been recently established within the organizational frame of the Hungarian Academy of Sciences. Then we briefly survey how the papers of this thematic bloc contribute to the fulfillment of these goals.

**Kulcsszavak:** számítógépes társadalomtudomány, transzdiszciplináris kutatások, elméletépítés

**Keywords:** computational social science, transdisciplinary research, theory building

Az MTA Társadalomtudományi Kutatóközpontja kezdeményezésére 2017 őszén megalakult az MTA *Számítógépes Társadalomtudomány (Computational Social Science)* témacsoportja. Hivatalos bemutatkozásunkra 2017. november 14-én, a Magyar Tudomány Ünnepe keretében került sor. Tematikus összeállításunkban az akkor elhangzott előadások kibővített és szerkesztett változatai szerepelnek.

Témacsoportunk erőfeszítéseinek középpontjában a nagy teljesítményű számítástechnikai módszerek társadalomtudományi kutatásokban való felhasználásának elősegítése áll. A Magyar Tudományos Akadémia fő küldetésének megfelelően különös hangsúlyt helyezünk a feltáró jellegű alapkutatásokra. Döntő fontosságúnak tartjuk azonban, hogy ezek kiegészüljenek az alkalmazásra irányuló kutatásokkal, valamint a társadalmi és kutatói felhasználóknak szánt algoritmikus eszköztár fejlesztésével. Segítségünkre lehet az a szerencsés körülmény, hogy a magyar társadalom-, természet- és adattudományi szakemberek körében már mintegy két évtizedes hagyománya van a számítógépes támogatású társadalomkutatás egyes kérdéskörében, leginkább a társadalmi hálózatkutatás terén való együttműködésnek. Tapasztalataink szerint ez a laza szálakból fonódó, de folytonos transzdiszciplináris kapcsolat számottevő versenyelőnyt jelent a magyar tudományosságnak a nemzetközi mezőnyben. Előnyünk azonban illanékony. Hogy tartósan megőrizhessük, a kialakult tudományközi hálózat laza kapcsolódásait módszeres együttműködési rendszerré szükséges alakítani, mely egyszersmind jól azonosítható entitásként, elismert szereplőként jelenik meg a nemzetközi tudományosságban is. Ezt elősegítendő kiemelt figyelmet kívánunk fordítani a módszertani vonzatú kutatásokra, melyek átjárást biztosítanak az egyes tudományágak történetileg rögzült diszciplináris megközelítései között. Ezzel témacsoportunk a magyar kutatók egy olyan új generációjának a képzéséhez is hozzá kíván járulni, melynek tagjai képesek az egyes, ma még egymástól nagy módszertani távolságra eső szaktudományos ismeretanyagok hatékony ötvözésére. Külön öröm számunkra, hogy tematikus blokkunk sokszerzős cikkeinek hozzájárulói között nagy számban találunk tudományos pályafutásuk első harmadában-negyedében járó fiatal kutatókat. E változatos diszciplináris háttérű kutatók szakmai szocializációja szakterületeik meghatározó kutatóinak intézményi műhelyeiben folyik. Ezek az intézmények egyúttal a Számítógépes Társadalomtudomány témacsoport alapító tagjai is.

A tematikus összeállításunkat alkotó cikkek sorához Vámos Tibornak, a számítógépes társadalomtudomány hazai nesztorának írása (*Akadémiánk és az információs világ*) adja meg a hangütést. A társadalmi algoritmizáció mindenütt jelen lévő folyamata során az emberi tényező, az értelmezési kereteket kiépítő intellektus szerepe döntő. E szerep betöltése azonban egy transzdiszciplináris együttműködést lehetővé tevő közös szakmai nyelv kialakítását feltételezi. Tematikus blokkunk egy ilyen, a természet-, társadalom- és adattudományokat összekötő,

a számítógépes társadalomtudomány együttes művelését segítő közös tudásalap kialakításához kíván hozzájárulni.

A következő négy írás mindegyike valamely, a számítógépes társadalomtudomány témakörébe tartozó területtel és az ehhez kapcsolódó kutatói megközelítésmódokkal ismerteti meg. Novák Attila, Siklósi Borbála és Prószéky Gábor cikke (*Segíthetnek-e a szóbeágyazási modellek a társadalomtudósoknak?*) az MTA–PPKE Magyar Nyelvtechnológiai Kutatócsoportjának szójelentés-reprezentációs kutatásait ismerteti. Neurális hálózatokon alapuló, egyben grafikus megjelenítéssel támogatott modelljeik meglepő „kreativitással” ragadják meg a vizsgált kifejezések tágabb stílári és csoportnyelvi vonatkozásait. Ács Judit, Borbély Gábor, Makrai Márton, Nemeskey Dávid, Recski Gábor és Kornai András írása (*Hibrid nyelvtechnológiák*) a hagyományos, szabályalapú nyelvészeti megközelítéseknek a nagyméretű adatállományok gépi tanulós módszerekkel történő átszűrésén alapuló módszerekkel való összekapcsolása mellett érvel, az MTA SZTAKI Nyelvtechnológiai Kutatócsoportja által követett hibrid megközelítésen alapuló kutatásokat ismerteti. Galántai Júlia, Pápay Boróka, Kubik Bálint György, Szabó Martina Katalin és Takács Károly első ránézésre meglepőnek tűnő című írása (*A pletyka a társas rend szolgálatában. Az informális kommunikáció struktúrájának mélyebb megértéséért a Computational Social Science eszközeivel*) a pletykának a kooperáció és a társadalmi normák karbantartását szolgáló szerepéről szól. A pletykakutatás a számítógépes támogatású társadalmi hálózatok kutatás egy gyorsan fejlődő irányzatának, az úgynevezett negatív kapcsolatok vizsgálatának egyik érdekes leágazása. Kertész János, Roberta Sinatra és Vedres Balázs írása (*A társadalom hálózati jelenségeinek adatvezérelt vizsgálata*) három, a CEU Hálózattudományi Központjához kapcsolódó szimulációs módszertanú vizsgálatot ismerteti. Az első a társadalmi hatások kaszkádszerű terjedését szemlélte, egyebek között a hazánkban egykor népszerű iWiW-szolgáltatás összeomlásának példáján. A másik vizsgálat a *teamek* eltérő kreativitási fokára keres magyarázatot a hálózati szerkezeti gyűrődések (structural folds) alapján, míg a harmadik a tudományos hatás életkorfüggési mintázatát tárja fel a közlemények citációs hálózatára támaszkodva.

Az ezeket követő három írás az adatinfrastruktúra alakulásának a tudomány művelésében betöltött szerepét járja körül. Lévai Péter és Telcs András hozzájárulása (*Új módszerek régi kérdések megválaszolására az Akadémiai Felhőben. Hálózatok és oksági kapcsolatok felderítése a társadalomtudományokban*) bemutatja, hogy az akadémiai felhő szolgáltatásai miként támogathatják az adatigényes társadalomtudományi kutatásokat. Noha a számítógépes felhőt az MTA Wigner Fizikai Kutatóközpontjában, illetve az MTA SZTAKI-ban hozták létre, annak szolgáltatásai valamennyi tudományterület számára nyitva állnak. A Wigner Kutatóközpontban egy újonnan létrehozott osztály kifejezetten azt hivatott elősegíteni, hogy e lehetőséggel hazánk számítástechnikai részletekben kevésbé járatos

kutatói is élni tudjanak. Kovács László, az MTA SZTAKI Elosztott Rendszerek osztályának vezetője az emberi tényező tudományterületenként változó szerepét veszi szemügyre (*Adatkezelés. A [kutatási] adatok kezelése a magyar tudományos- és memóriaintézményekben*). Az egyes diszciplínák adatkezelési szokásrendszerei egyedileg, sokszor egy jóval korábbi technikai szinthez igazodva alakultak ki. Ezeknek a meggyökeresedett rutinoknak a *state-of-the-art* adatkezelési lehetőségekkel való felváltása döntő fontosságú, bár korántsem egyszerű feladat. Simon Eszter és Váradi Tamás munkája (*Magyar nyelvtechnológiai infrastruktúra a társadalomtudományok szolgálatában*) az MTA Nyelvtudományi Intézetében kifejlesztett magyar nyelvű szövegeket feldolgozó eszközláncot – nyelvi elemző eszközök egyetlen koherens technológiai láncba rendezett együttesét – ismerteti társadalomtudományi alkalmazásokon keresztül. A különféle digitális formában előálló nyelvi korpuszok már ma is számos szociológiai, politikatudományi, lélektani kontextusú empirikus kutatás számára szolgálnak lehetséges bemenetként. Így az adatfolyamok elemzését támogató felhasználóbarát eszközlánc az adattudományokban kevésbé járatos kutatók széles köre számára teheti elérhetővé a fejlett nyelvtechnológiai eszközök kutatásaikban való alkalmazását. Végül Muraközy Balázs tanulmánya (*Gépi tanulás, predikció és okság a közgazdaságtudományban*) a metaszintet feszegeti. Hogyan változik a hagyományos oksági, elméletalapú és a gépi tanulásra építő statisztikai-predikciós fókuszú kutatások viszonya a társadalomtudományok, jelesül a közgazdaságtan területén? A szerző a két megközelítés komplementaritását és a két megközelítést hatékonyan ötvöző empirikus módszerek kikísérletezésének szükségességét emeli ki.

Ezzel visszaértünk tematikus blokkunk központi gondolatához. A számítógépes társadalomtudomány talán legnagyobb kihívása, hogy az értelmezési kapacitás lépést tud-e tartani a számítógépes módszerek *output*jainak rohamos bővülésével. Ha igen, úgy ez nagy teherbírású transzdiszciplináris értelmezési keretek kiépítését feltételezi. Olyan módszerek megalkotását, melyekkel a hagyományos oksági-analitikus kutatás „megmagyarázott szabályszerűségekké” alakítja azokat az empirikus mintázatokat, melyeket a gépi-statisztikai masinéria az adattömegben felfedez. A nagy teljesítményű számítógépes módszerek hozadéka így nem a társadalomra vonatkozó elméletalkotás kiüresítése, hanem hatékony támogatása lehet.