

ÖSSZEFOGLALÓK

Adatbőség

VEROSZTA ZSUZSANNA

ADMINISZTRATÍV ADATOK TÁRSADALOMKUTATÁSI KEZELÉSE

A tanulmány célja számba venni az adminisztratív adatok társadalomkutatási felhasználásának elméleti és gyakorlati feltételeit és következményeit. Ennek során összefoglaljuk azokat a tényezőket, amelyek az adminisztratív adatok iránti megnövekedett figyelmet táplálják, majd megvizsgáljuk ezen adattípus természetét és viszonyát a kutatási céllal gyűjtött adatokhoz. Ezt követően az adminisztratív adatokon nyugvó társadalomtudományi kutatás szükséges validálási kritériumait és eljárásait tekintjük át, illusztratív példákat is beemelve. Az elemzés során emellett az adatbázisok összekapcsolásában rejlő lehetőségek tipizáló áttekintését végezzük el.

KULCSSZAVAK: ADMINISZTRATÍV ADAT, ADATÖSSZEKAPCSOLÁS, ÉRVÉNYSÉG, ADATBÁZIS, ELEMI ADAT

NAGY PÉTER TIBOR

A RÉGI ÉS ÚJ TÖRTÉNETI ADATOK TÁRSADALMI TERMÉSZETE

A tanulmány három - eltérő méretű - altémát fejt ki: a régi típusú, de új adatok keletkezését; az új típusú, új adatok keletkezését és a régi adatok újrahasznosulását. Minden egyes altémán belül először azzal foglalkozik, hogy az „adatok új világa” (mely valószínűleg a számítástechnikával, s különösen az internettel vette kezdetét) hogyan alakítja át - funkcionalistán szólva - a tudományos témákat és módszereket, majd arra, hogy - konfliktusos paradigmában fogalmazva - a tudományos közösségek milyen érdekvizonyok alapján kezelik, illetve ismerik el egyáltalán releváns adatnak az új információs helyzet „termékeit”. A vizsgálat során tehát az adatfeltárás, adatrendezés és adatelemzés jelenségvilágát egy komplex tudományszociológiai összefüggésrendszer részeként törekszünk értelmezni.

KULCSSZAVAK: ADAT, ADATBŐSÉG, ADATROBBANÁS, TUDOMÁNYSZOCIOLÓGIA, KONFLIKTUSOS PARADIGMA, FUNKCIONALISTA PARADIGMA, KUTATÁSMÓDSZERTAN, DIGITALIZÁLÁS

GÁRDOS ÉVA**ADATOK ÉS KEZELÉSÜK A HIVATALOS STATISZTIKÁBAN**

Jó minőségű statisztikai mutatók előállításához nagyon fontos a megfelelő adatforrás kiválasztása. A tanulmány bemutatja a hivatalos statisztikai munkában használható adatforrás-típusokat és alkalmazási területeiket, kitérve a másodlagos adatokra, különös tekintettel az adminisztratív adatok egyre növekvő jelentőségére. Részletes képet kaphat az olvasó a különböző fajta adatok felhasználási feltételeiről, ezek között is nagy nyomatékot kap a hazai és az EU-s jogszabályi háttér. Emellett röviden szó esik a statisztikai termékek és adat-előállítási folyamat minőségi kritériumairól, valamint az adatvédelem szabályairól, és a felfedés elleni védelem során figyelembe veendő szempontokról. A tanulmány átfogó képet ad a hivatalos statisztikai rendszer működéséről és bepillantást enged a KSH tevékenységét meghatározó legfontosabb irányelvekbe is.

KULCSSZAVAK: STATISZTIKAI ADAT, MÁSODLAGOS ADAT, ADMINISZTRATÍV ADAT, STATISZTIKAI MINŐSÉG, STATISZTIKAI FELFEDÉS ELLENI VÉDELEM

SZÉKELY IVÁN**KÖZADATOK ÉS NYILVÁNOS ADATBÁZISOK: A HOZZÁFÉRÉS KÉRDÉSEI**

A közadatokhoz való szabad hozzáférés a demokratikus jogállamok alapértéke és alapvető joga, mégis állandó érdekütközések színtere és gyakorlati problémák felvetője. A tanulmány bemutatja a nemzetközi és magyar jog fogalomrendszerét és megközelítésének korlátait a közadatok hozzáférhetősége terén, ismerteti legfontosabb jogintézményeit, majd külön fejezetet szentel az ingyenesség kérdésének, esettanulmányok elemzésével. A szerző bemutat néhány releváns hazai és nemzetközi közadat-hozzáférést segítő technológiát és alkalmazást, majd felvet néhány ritkábban elemzett szempontot a közadatok, nyilvános adatbázisok felhasználási lehetőségei terén. A tanulmány zárásképpen röviden kitér az adatbőség oktatási vonatkozásaira, mind a diákok, mind a pedagógusok szempontjából.

KULCSSZAVAK: INFORMÁCIÓSZABADSÁG, KÖZÉRDEKŰ ADAT, ADATBÁZIS, INGYENESSÉG, KÖZADAT-HOZZÁFÉRÉST SEGÍTŐ TECHNOLÓGIÁK

HALÁSZ GÁBOR**ADATOKKAL TÖRTÉNŐ KORMÁNYZÁS: DILEMMÁK ÉS PERSPEKTÍVÁK**

E tanulmány az adatokkal vagy indikátorokkal történő kormányzás dilemmáit és perspektíváit elemzi, különös tekintettel a szakpolitikai koordináció Európai Uniót jellemző gyakorlatára, valamint a tényeken alapuló szakpolitikai döntéshozatal európai normáira. Bemutatja az adatok és indikátorok kormányzási eszközként történő használatával járó előnyöket és az ennek nyomán jelentkező kihívásokat. A tények és adatok közötti különbséget hangsúlyozva bemutatja a tényekre épülő politikai nyomásgyakorlás és meggyőzés gyakorlatának előnyeit, hangsúlyozva ennek korlátait is. A tanulmány utal azokra az új lehetőségekre, amelyek a nagy mennyiségű mikro-szintű, longitudinális adatot gyűjtő adatrendszerek létrejötte nyomán keletkezhetnek.

KULCSSZAVAK: OKTATÁSPOLITIKA, KORMÁNYZÁS, ADATOKKAL TÖRTÉNŐ KORMÁNYZÁS, TÉNYEKEN ALAPULÓ SZAKPOLITIKA, EURÓPAI UNIÓ, SZAKPOLITIKAI KOORDINÁCIÓ, NAGY ADATRENDSZEREK

SZÉLL KRISZTIÁN**AZ ADATHOZZÁFÉRÉS ÉS FELHASZNÁLÁS NEMZETKÖZI GYAKORLATAI**

A tanulmány egy angolszász (Anglia) és egy skandináv (Svédország) kultúrkörből származó országban tárja fel a közoktatási célú adatgyűjtések alapján előálló adatbázisokhoz való hozzáférés és az adatfelhasználás gyakorlatait. A tanulmány arra keresi a választ, hogy a vizsgált nemzeti adatrendszerek (1) milyen célból gyűjtenek adatokat és fejlesztenek indikátorokat, (2) a célok elérése érdekében milyen szintű és tartalmú adatgyűjtések valósulnak meg, valamint (3) milyen hozzáférést biztosítanak az egyes szereplők számára a háttéradatbázisokhoz. A tapasztalatok alapján a fejlett adatrendszerekben az adatgyűjtések és indikátorfejlesztések céljai között a költségvetés-tervezési és tanügyigazgatási szempontok mellett hangsúlyosan megjelennek a fejlesztési, kutatási és értékelési szempontok is. Az egységes alapelvekre épülő és több évtizedre visszatekintő adatkultúrával rendelkező nemzeti adatrendszerek elsősorban egyéni szintű adatokat gyűjtenek. Adott tehát a lehetőség – és a szándék – az anonimizált formában megvalósuló adatintegritációra, hozzáátve, hogy a fejlett adatrendszerek különös figyelmet fordítanak az adatvédelem és az adathozzáférés közötti egyensúly megteremtésére. Összességében elmondható, hogy az adatgyűjtési és hozzáférési alapelvek lefektetése és társadalmi elfogadottsága erősíti az adatok célhoz kötött és széles körű felhasználását.

KULCSSZAVAK: ADATHOZZÁFÉRÉS, ADATGYŰJTÉS, INDIKÁTOROK, FEJLESZTÉS, ÉRTÉKELÉS

SALOMVÁRI GYÖRGY**KÖZNEVELÉS AZ ADATOK BŰVÖLETÉBEN – AZ ADATTERMELÉS FORRADALMA**

Az adattermelés forradalma a pedagógia területét sem hagyja érintetlenül. A hatására újonnan megjelenő komplex adatvagyon-kezelő rendszerek a felszint karcoló statisztikai megközelítés helyett alapvetően egyéni szintű adatokra épülnek, ezáltal kifejezetten alkalmasak közvetlen pedagógiai célú alkalmazásra: olyan részletezettséggel ismerhetjük meg a tanítás – tanulás folyamatát, eredményességét és hatékonyságát jellemző adatok időbeli változásait, melyre korábban soha nem volt lehetőség.

A lehetőség egyben felelősség is, az adatvagyon létrehozásán túlmenően a felhasználás körülményeinek és korlátainak a meghatározása a legfontosabb feladatunk. Sikert akkor érünk el, ha a pedagógia gyermek és tanulóközpontú komplex megközelítést alkalmaz, melynek integráns eleme a pedagógusok adatvagyon felhasználásra épülő tevékenysége - szakértők és oktatáskutatók által „kívülről” támogatva.

KULCSSZAVAK: ADATTERMELÉS, ADATVAGYON, ADATKAPCSOLATI HÁLÓ, ELEMÍ ADATOK, EGYÉNI SZINTŰ ADATOK, ADATVAGYON HASZNOSÍTÁSRA ÉPÍTŐ PEDAGÓGIAI RENDSZEREK

BOGÁROMI ESZTER - MÁTH ANDRÁS**SOKFORRÁSÚ ADATBÁZIS ÉPÍTÉS - BUKTATÓK, NEHÉZSÉGEK, MEGOLDÁSI KÍSÉRLETEK**

A tanulmány a külső forrásból származó adatbázisok felhasználásának folyamatát mutatja be egy példán keresztül.

Az első szakaszban megismerhetjük az ingyenesen és fizetősen hozzáférhető nyilvános, adminisztratív és egyénsoros adatbázisok körét, azok adatfelvételének hullámain, alap-

egységeit, változóit és további tulajdonságait. Továbbá ebben a szakaszban mutatjuk be azt is, milyen megoldások szolgálhatnak arra, ha a számunkra szükséges adatbázisok nem nincsenek előállítva a piacon. A második szakaszban a különböző adatbázisok összefűzésének buktatóit és nehézségeit mutatjuk be, megoldási javaslatokkal együtt. A példában település szinten történt az adatbázisok összekapcsolása. A harmadik szakaszban a kutatás alkalmával tervezett elemzésekben használt módszertanok tükrében vizsgáljuk a rendelkezésre álló változókat és eseteket. A negyedik fejezetben összefoglaljuk a fent bemutatott adatsorok összekapcsolásának kutatási felhasználásának lehetőségeit, eredményeit.

KULCSSZAVAK: TELEPÜLÉS, ADATHASZNOSÍTÁS, GYAKORLATI PÉLDA, KSH, TGI, ADATBÁZISOK ÖSSZEKAPCSOLÁSA

Data abundance

ZSUZSANNA VEROSZTA

THE GROWING USE OF ADMINISTRATIVE DATA IN THE SOCIAL SCIENCES

The study aims to list and evaluate the theoretical and practical conditions and also the consequences of using administrative data in the social sciences. In this context, the first step is to summarize the factors that have fuelled the increased attention being given to administrative data gathering; secondly, we examine the nature of this type of data and compare it with data collected directly for research purposes. Next, we review the validity criteria pertaining to administrative data-based research, using illustrative examples. In addition, during the analysis we provide an overview of and also categorize the possibilities inherent in database linkage.

KEYWORDS: ADMINISTRATIVE DATA, RECORD LINKAGE, VALIDITY, DATABASE, MICRO DATA

PÉTER TIBOR NAGY

THE SOCIAL NATURE OF OLD AND NEW DATA

This study seeks to provide an explanation for three sub-problems. 1. The social determination of creation/the arising of “old-type” new historical data (for example determinative factors for excavations, etc.); 2. the social circumstances of creation/the arising of “new-type” new historical data (for example, historical medicine) 3. re-interpretation/re-structuring of old data (for example, the digitalisation of sources). There are two topics within this area of sub-problems: 1. how the new “data universe” (the internet) has changed the topics and methods utilized in research; 2. how academic communities have perceived and integrated the new data types and how the face of a community has become altered within this process. The first sub-problem is interpreted from a functionalist perspective, the second from a conflict-theoretical one.

KEYWORDS: DATA, ABUNDANCE OF DATA, SCIENTIFIC CLASSIFICATION, SOCIOLOGY OF SCIENCE, SOCIAL CONFLICT THEORY, FUNCTIONALIST THEORY, RESEARCH METHODOLOGY, DIGITALIZATION

ÉVA GÁRDOS

DATA AND DATA MANAGEMENT IN OFFICIAL STATISTICS

From the perspective of producing statistical indicators of good quality, it is important to select the most appropriate data sources. This study presents data sources of different types which can be used in official statistical work as well as the areas of their application, going into detail regarding secondary data, with special attention being given to administrative data and its growing significance. The reader may get a detailed picture of conditions needing to be taken into account when using data of different types; while major emphasis is given to both Hungarian legislation and that of the European Union. Moreover, the qualitative criteria of the statistical product along with the data production

process are also briefly mentioned, as well as data protection rules and also viewpoints that one should consider in the course of disclosure control. The study provides a broad overview of the functioning of the official statistical system and gives insight into the most important policies and guidelines defining the activities of the Hungarian Central Statistical Office.

KEYWORDS: STATISTICAL DATA, SECONDARY DATA, ADMINISTRATIVE DATA, STATISTICAL QUALITY, STATISTICAL DISCLOSURE CONTROL

IVÁN SZÉKELY

PUBLIC DATA AND PUBLIC DATABASES: QUESTIONS OF ACCESS

Free access to public data is a basic value and fundamental right of a constitutional democracy, though it invariably generates both conflicts of interest and practical problems. This study presents the notions used in international and Hungarian law together with the limitations in their approach and, by analysing case studies, it devotes a separate section to discussing questions of 'access free of charge'. The author presents several relevant technologies and applications supporting access to public data and databases, and he raises some less analysed issues relating to public data and database usage. Finally, the study briefly touches upon the educational potential of data abundance both from the students' and the educators' point of view.

KEYWORDS: FREEDOM OF INFORMATION, DATA OF PUBLIC INTEREST, DATABASES, ACCESS FREE OF CHARGE, TECHNOLOGIES SUPPORTING ACCESS TO PUBLIC INFORMATION

GÁBOR HALÁSZ

GOVERNING BY DATA – DILEMMAS AND PERSPECTIVES

This paper analyses the dilemmas and perspectives of governing by data and indicators in the education sector with a special focus being put on the policy coordination practices of the European Union and on the emerging European norm of evidence-based policy-making. It presents some of the most relevant issues related to the use of data and indicators as instruments of governance, exhibiting the advantages and the challenges underpinning this relatively new practice. Stressing the difference between facts and data, it presents the benefits of using facts as tools of political pressure and persuasion, revealing also the limitations of this practice. The paper makes reference to the new opportunities created by the emergence of big databases containing large quantities of micro level, longitudinal data.

KEYWORDS: EDUCATION POLICY, GOVERNANCE, GOVERNING BY DATA, EVIDENCE-BASED POLICY, EUROPEAN UNION, POLICY COORDINATION, BIG DATA

KRISZTIÁN SZÉLL

DATA ACCESSIBILITY IN AN INTERNATIONAL CONTEXT

The study explores the possibilities of data accessibility and the practices of using data in an Anglo-Saxon (England) and in a Scandinavian country (Sweden). The paper tries to answer the following questions: (1) What are the purposes of data collection and indicator development? (2) What level and content of data are collected in order to attain goals? (3) What accessibility is provided regarding databases for stakeholders and publicity? Experience shows that besides budget planning and education administration aspects, research, development and assessment factors are also emphasized among data

collection and indicator development aims. National data systems that have a unified set of principles and a history of data collection looking back several decades mainly gather individual-level data. So the opportunity – and the intent – to integrate anonymized data is given, though developed data systems do put great emphasis on the balance between data protection and data accessibility. All in all, creation and social acceptance of the principles of data collection and accessibility strengthen the aim-assigned and extensive use of data.

KEYWORDS: DATA ACCESS, DATA COLLECTION, INDICATORS, DEVELOPMENT, EVALUATION

GYÖRGY SALOMVÁRI

THE DATA REVOLUTION IN PUBLIC EDUCATION

The data production revolution has not left untouched the area of education. This has caused newly-emerging complex data asset management systems that are fundamentally based on individual-level data instead of the surface abrasive statistical approach, and this is especially suitable when it comes to usage with direct pedagogical purposes: we can now obtain a level of detail regarding teaching-learning process effectiveness and also see the efficacy of typical data changes occurring over time, something which has never previously been possible.

The chance to do this additionally points to a responsibility along with the creation of data asset conditions and limitations for which the ability to determine outcomes is our most important task. Success is achieved when education applies a complex child- and learner-centred approach, which is an integral part of activity-based educators' use of data assets – and this is supported by education experts and researchers from “outside”.

KEYWORDS: DATA PRODUCTION, DATA ASSETS, DATA NETWORK OF RELATIONSHIPS, PRIMARY DATA, INDIVIDUAL-LEVEL DATA, UTILIZATION OF DATA ASSETS BUILT ON EDUCATIONAL SYSTEMS

ESZTER BOGÁROMI – ANDRÁS MÁTH

THE POSSIBILITIES OF DATA RECOVERY, LIMITATIONS, PROFITS – PRESENTING AN EXAMPLE

The study demonstrates the process of using databases in a specific example, and our focus is on databases that are available from external sources. In its first part, we reveal the list of free and paid-for databases (used in research) as well as the range of public, administrative and one-line databases; and the section presents different periods of data collection, the basic units, variables and other characteristics. Also shown here is how one may proceed if a data set is not at one's disposal. In our example, we make use of a self-created database that includes all settlements in Hungary – which initially required us to merge several databases, which was not an easy task as many pitfalls and difficulties were encountered. The study also presents some of the possible 'solutions' regarding the way of proceeding and also the most useful results.

KEYWORDS: SETTLEMENT, DATA RECOVERY, PRACTICAL EXAMPLES, HUNGARIAN CENTRAL STATISTICAL OFFICE, TGI, DATABASE MERGING