

A KUTATÁSI ADATOK MEGOSZTÁSA

Koltay Tibor

PhD, Dr. habil.,

Szent István Egyetem Alkalmazott Bölcsészeti és Pedagógiai Kar, Jászberény

Koltay.Tibor@abpk.szie.hu

A digitális kutatási adatok megosztása lehetővé teszi ezen adatok újrafelhasználását és eredetük igazolását. A megosztás célja az is lehet, hogy a közpénzekből finanszírozott kutatások eredményeit hozzáférhetővé tegyünk a nyilvánosság számára. Végso soron célba vehetjük a hatékonyabb kutatást is.

Az adatmegosztás szónak számos jelentése van. Meghatározhatjuk úgy, mint az adatok közreadását abból a célból, hogy mások is használhassák őket. Ennek a formája lehet a kérésre történő magánjellegű csere, és egészen addig az igényig terjedhet, hogy az adatok nyilvános adatgyűjteményekbe kerüljenek. Cél lehet az is, hogy vizsgálatok adatait össze tudjuk egymással hasonlítani, ami viszont nem mindig kivitelezhető vagy kívánatos. Megosztásnak minősül az is, ha egy nyilvános weboldalon helyezük el az adatokat, vagy mellékletként egy folyóirat rendelkezésére bocsátjuk őket. A hasznosság, a megbízhatóság mértéke és az adatok értékes volta azonban sokféle lehet. Bizonyos adatok jellemzője a gazdag strukturáltság és gondozottság, míg mások lehetnek minimális dokumentációval ellátott nyers fájlok. Hasonlóképpen, a célközönség egy szűk szakterület kutatóitól a nagyközönségig terjedhet (Borgman, 2012).

Aligha kétséges, hogy a tudományos kutatásban mindig is kulcsszerepük volt az adatoknak. Sőt, nyilvánvaló, hogy a legtöbb

kutató munkájában alapvető jelentősége van az adatokra alapozott gondolkodásnak, ami összekapcsolódik annak képességével, hogy az adatokat absztrakt fogalmakra fordítsuk le (Davies et al., 2011). Azt viszont, hogy mit használunk fel adatként, nagymértékben az adott szakmai érdekközösség határozza meg, viszont egy-egy kutató akár több ilyen közösség tagja is lehet, továbbá ezeknek a közösségeknek eltérőek lehetnek az adat fogalmáról alkotott elképzelései és a velük kapcsolatos gyakorlata. Ezeknek az érdekközösségeknek a határai nem egyértelműek és nem is tartósak (Borgman, 2012).

A közelmúltig a kutatást finanszírozó szervezetek általában nem ösztönözték arra a kutatókat, hogy gondozzák és megőrizzék a tevékenységük során keletkezett adatokat, ezért a kutatók nem is sokat foglalkoztak ezekkel a tevékenységekkel. A kiadók is csak egy töredéküknek szántak figyelmet azzal, hogy a publikációkban, egy-egy érvelést alátámasztandó, táblázatok és ábrák formájában közöljék őket. Ennélfogva a szövegek, táblázatok, laboratóriumi feljegyzések, kérdőívek, kép- és hangfelvételek, fehérje- és génszekvenciák, modellek stb. formáját öltő adatok zöme nem kerül napvilágra, ott maradnak a kutatók és kutatóhelyek számítógépein (Murphy, 2014).

A kutatási adatok megosztása az utóbbi néhány évben került a figyelem középpontjába

ba, pedig Christine Borgman már 2007-ben úgy látta, hogy az adatok, és azok megosztása a publikált cikkekhez hasonló tudástöket jelentenek (Borgman, 2007). Ugyanabban az évben, idehaza Z. Karvalics László (2007) éppen a *Magyar Tudomány* folyóirat hasábjain hívta fel a figyelmet az adatintenzív tudományra.

Ezen a téren mindenesetre nagy változásnak lehetünk szemtanúi. Miközben a kutatók körében kulturális természetű hagyománya van annak, hogy inkább önmagukra és megbízható kollégáikra támaszkodjanak, mint valamilyen központi szolgáltatásra (Pryor, 2012), a kutatásokat finanszírozó ügynökségek kezdik megkövetelni az adatok közreadását, bár ennek mértéke és a végrehajtás kötelező jellege változó. A követelmények egy része nem is olyan új. Az Amerikai Egyesült Államokban a *National Institutes of Health* például 2003 óta követeli meg, hogy az 500 ezer dollár feletti támogatásokhoz adatkezelési terv készüljön. A *National Science Foundation* 2001 óta írja elő az adatok megosztását az általa finanszírozott kutatások esetében, azonban nem volt következetes abban, hogy ezt meg is követelje. 2010-ben viszont bejelentették, hogy a jövőben minden pályázatnak tartalmaznia kell egy kétoldalas adatkezelési tervet. Egyes folyóiratok is elvárják, hogy a szerzők helyezték el náluk a cikkekhez kötődő adatokat és más kutatási dokumentációt. Az erre vonatkozó irányelvek szigorodtak is az idők folyamán (Borgman, 2012). Nagy-Britanniában hasonló a helyzet.

A *Horizon2020* programban az Európai Unió is nagy jelentőséget tulajdonít az adatok megosztásának (Murphy, 2014). Az uniós kutatási és innovációs pályázatoknak – amennyiben az releváns – adatkezeléssel kapcsolatos részt is kell tartalmaznia, amelyben kitér-

nek arra, hogy milyen típusú adatokat használnak fel, vagy hoznak létre a projekt során; milyen szabványokat használnak ezek kezeléséhez; miként osztják meg ezeket az adatokat, vagy ha nem lehet őket nyilvánosságra hozni, annak mi az oka; és hogyan gondoskodnak az adatok kezeléséről és megőrzéséről (EC Guidelines, 2013). Emellett az EU az adatok megosztásának lehetőségeit vizsgáló kutatási projekteket is finanszíroz. A *RECODE* (Policy Recommendations for Open Access to Research Data in Europe, URL1) projekt például az érintettek, tehát a kutatást finanszírozó és a kutató szervezetek, intézmények, az adatokat kezelő szakemberek (adat-kurátorok, adat-könyvtárosok stb.), valamint a kiadók szemszögéből és ezeknek az érdekelteknek a bevonásával vizsgálta az adatokhoz való nyílt hozzáférés számos kérdését.

Tovább is sorolhatnánk a példákat, de csak azt említsük még meg, hogy a nagy finanszírozók mellett olyan tudományos társaságok is súlyt helyeznek az adatok megosztására, mint a *British Ecological Society*. Annak érdekében, hogy tagjaik kellően felkészülhessenek a változásokra, már 2012-ben bejelentették, hogy 2014-től kötelezővé teszik az adatok archiválását (Norman, 2014).

A kutatási adatok megosztásának tágabb kontextusát a *nyitott tudományról* alkotott elképzelések adják meg. A *Royal Society* a nyitott tudományt úgy határozta meg, mint a nyílt adatok, a tudományos publikációkhoz való nyílt hozzáférés, valamint a tudományos tartalmak hatékony kommunikációjának kombinációját. Ebben a keretben a nyílt adatoknak elérhetőnek, használhatóknak, értékelhetőnek és értelmezhetőnek kell lenniük (Boulton et al., 2012).

Az adatokban gazdag világgal szemben óriási a várákozás, legyen szó új gyógyszerek

felfedezéséről, a világ időjárásának, történelmünknek vagy kultúránknak jobb megismeréséről. Az adatok mennyiségének megnövekedése az olyan „nagy tudományok” területén, mint a csillagászat és a fizika, *negyedik paradigma* néven új modellek létrejöttét, továbbá olyan új tudományágak megszületését hozta magával, mint az asztroinformatika, a bioinformatika vagy a digitális bölcsészet.

Az adatok megosztása azonban viszonylag kevés tudományterületre korlátozódik, és a gyakorlat még ezeken a területeken sem következetes. Az Amerikai Egyesült Államokban, a *National Science Foundation (NSF) Science and Technology Center* adatokkal kapcsolatos gyakorlatának vizsgálata azt mutatta, hogy az ezeket az adatokat létrehozó *teamek*en túl kevés adat kerül be valamilyen tágabb körforgásba, sőt kevesen is kéri ezeket az adatokat, aminek számos oka van. A kutatóknak például többnyire nincs meg az ezirányú szakértelmük, nincsenek eszközeik hozzá, valamint ösztönzést sem éreznek arra, hogy kutatási adataikat másokkal megosszák. Bizonyos adatok (a már részben említett) etikai, valamint ismeretelméleti okokból sem oszthatók meg. Néhány esetben pedig nem világos, hogy melyek az adott kutatási projekthez tartozó adatok. Az adatmegosztási készség hiánya abból is ered, hogy a kutatás területén meglevő versenyhelyzetben sok kutató attól fél, hogy adatait nem megfelelően használják fel mások (Murphy, 2014).

Annak ellenére tehát, hogy kutatásokat finanszírozó szervezetek nyomást gyakorolnak a kutatókra az adatok megosztása érdekében, továbbá kutatások eredményei azt mutatják, hogy az adatok megosztása növelheti a hivatkozások mértékét, valószínű, hogy kevés adat megosztására kerül sor (Borgman, 2012). A kutatókat viszont ösztönözni kellene adataik

megosztására, ami hatékonyabb lenne, ha ezért valamilyen ellenszolgáltatást kapnának, például úgy, hogy az adatok publikálása és a rájuk történő hivatkozás a tudományos teljesítmény elismerésének része lenne. Egyelőre azonban ritkán és véletlenszerűen találkozunk vele, és ezen a területen egyelőre nincsenek kikristályosodott szabványok. Várható azonban, hogy a folyóiratcikkeket URL-jüktől függetlenül azonosító DOI (Digital Object Identifier, URL₂) metaadat-szabványra épülő szabványok lesznek a leghasznosabbak. Közben azért létezik közvetett ösztönzés is. Ennek jó példája a *Holland Adatdíj* (Dutch Data Prize, URL₃), amellyel az adataikat megosztó kutatókat jutalmazták.

A „nagy adatok”

Az adatok megosztása kapcsán szót kell ejtenünk a *big data* jelenségről. A mai értelemben vett nagy adatok esetében nem annyira a méret, az adatok mennyisége a lényeg, hanem az, hogy a számítástechnika hatékonyságának maximalizálása folytán lehetővé vált ezen adatállományok összegyűjtése, elemzése, összekapcsolása és összevetése.

A digitális eszközök elterjedtsége és népszerűsége folytán a nagy adatok egy részét a közösségi média használói állítják elő. Ezzel új utak nyíltak a társadalmi és a kulturális folyamatok tanulmányozásában. Azonban, más társadalmi-technikai jelenségekhez hasonlóan, a nagy adatok jelensége egyszerre vált ki az utópiára és a disztópiára (antiutópiára) épülő retorikát. Egyrészt ott van a mítosz, hogy ezek az adatállományok a tudás és az intelligencia eddiginél fejlettebb formáját kínálják olyan meglátásokat lehetővé téve, amelyek eddig nem voltak elképzelhetőek. Mindezt az igazság, az objektivitás és a pontosság aurája veszi körül. Másrészt viszont sokan

úgy tekintenek a nagy adatokra, mint ami lehetővé teszi a magánélet titkosságának megsértését, a szabadságjogok csorbítását, az állam és a cégek által gyakorolt ellenőrzés megnövekedését. Kétségtelen, hogy a nagy adatok felhasználása új utat nyit a humán tudományok számára ahhoz, hogy a kvantitatív és objektív módszertannal dolgozó tudomány státuszát követeljék maguknak. Tény az is, hogy a nagy adatok számos társadalmi jelenséget tesznek számszerűsíthetővé, ami azonban továbbra sem zárja ki a szubjektivitást, ráadásul pusztán számszerűsítéssel nem feltétlenül kerülünk közelebb az objektív igazsághoz. Ráadásul az internetről vett nagy adatállományok gyakran megbízhatatlanok. Ha egy adatállomány mérete milliós nagyságrendű, az nem garancia arra, hogy reprezentatív is. Ahhoz, hogy statisztikailag érvényes következtetéseket vonjunk le egy-egy adatállományból, tudnunk kell, hogy honnan származik, és melyek a gyengeségei. Tudatában kell lennünk azoknak a tényezőknek, amelyek az értelmezést torzítják. A nagy adatok ezenkívül arra hajlamosítanak, hogy összefüggéseket lássunk ott is, ahol nincsenek. A több adat tehát nem mindig jobb adat. A nagy adat pedig nem azonos az adatok teljességével.

Gondolnunk kell arra is, hogy a nagy adatokkal kapcsolatos etikai kérdésekről keveset tudunk. Az a tény, hogy valami elérhető, nem feltétlenül teszi a felhasználását etikussá. Az adatközlők névtelensége például hamar elillanhat, ami különösen összetett kérdéseket vet fel a közösségi oldalokról vett adatok esetében. Felmerül az a kérdés is, hogy kell-e ezek felhasználásához engedélyt kérnünk az érintettektől, továbbá mit jelent az egyén számára, ha tudtán kívül, vagy az eredetitől eltérő kontextusba helyezve elemzik a tevékenységét (Boyd–Crawford, 2012).

Az adatmegosztás további feltételei

A hatékony megosztás megköveteli, hogy az adatok konzisztens mutatók segítségével megtalálhatóak legyenek. Ehhez meg kell őriznünk, és hosszú távon hozzáférhetővé kell tennünk őket. Minőségüknek lehetővé kell tennie, hogy használhatóak legyenek (Kowalczyk – Shankar, 2011).

Az adatok megosztásához tudnunk kell a következőket, bár a kérdésekre nem mindig tudunk igen-nem válaszokat adni: *Létezik-e az adatkészlet?* • *Hol található?* • *Sérült vagy elavult-e?* • *Van-e engedélyünk a használatára?* • *Elégé szabványos-e ahhoz, hogy elfogadható mértékű erőfeszítéssel használni tudjuk?* • *Elég világos-e, hogy mit reprezentál?* • *Eredete és hibaarányai ismertek és elfogadhatók-e, tehát megbízhatónak tekinthető-e?* • *Célszerű-e felhasználnunk céljaink elérésére?*

Ezek a kérdések láncolatot alkotnak, tehát ha létezik az adatkészlet, lehet, hogy nem férhetünk hozzá. Ha hozzáférünk, előfordulhat, hogy nem használható. Ha használható, nincs használatára engedélyünk, és így tovább. Bármelyik probléma magakadályozhatja a felhasználást. Amikor pedig elhárítjuk az egyik akadályt, egy másik még mindig lehetlenné teheti a felhasználást (Buckland, 2011).

Az adatok nyilvánosságra kerülésének önmagában nincs értelme, ha nincs, aki használja őket. Miközben bizonytalanság van abban a tekintetben, hogy miként vállaljanak részt az adatok kezeléséből az egyetemek vagy a kutatás finanszírozásában részt vevő szervezetek, a könyvtárak elvileg jó helyzetben vannak ahhoz, hogy kulcsszereplővé váljanak ebben, mivel tapasztalatuk van a válogatásban, a különböző gyűjtemények működtetésében, a megőrzésben, a gondozásban és a hozzáférés biztosításában (Stuart, 2011).

Nem téveszthetjük szem elől, hogy bizonyos adatok megosztható formában állnak rendelkezésre, míg mások nem. Bizonyos adatok értékét felismeri az adott közösség, míg másokét nem. Egyes kutatók minden adatukat bármikor hajlandók megosztani, míg mások soha semmilyen adatukat nem gondolnák megosztani. A leggyakoribb azonban az, hogy bizonyos adataikat néha hajlandóak megosztani (Borgman, 2012).

Az adatok minősége sokdimenziós jellemző. Része a bizalom, amelynek mértéke számos szubjektív tényezőtől függ. Ezek közé tartozik, hogy az adatokat autentikusnak ítéljük-e meg, elfogadhatónak találjuk-e felhasználásukat vagy alkalmazásukat. Hasonló tényező azoknak a jó híre, akik az adatokat létrehozzák. A minőség megítélését befolyásolják az adatok értékelőinek elfogultságai és előítéletei. Annak ellenére, hogy igen összetett kérdés, a hitelességen alapuló bizalom problémáját is megemlíthetjük, különösen azért, mert felülírhat más szempontokat. A hitelesség ebben a kontextusban annak körülbelüli mértéke, hogy az adatok mennyire a „jó tudományt” tükrözik, és olyan kérdéseket állít a középpontba, mint az adatgyűjtés eszközeinek megbízhatósága, az elméleti alapok megfelelő volta, az adatok teljessége, pontossága és érvényessége, valamint az adatok ontológiai következetessége. Annak érdekében, hogy a hitelességet meg tudjuk ítélni, az adatoknak érthetőnek kell lenniük.

Az érthetőség értékeléséhez nélkülözhetetlen, hogy az adatokat leíró dokumentáció, az azonosításukra szolgáló, egységes szerkezetben készült metaadatok vagy az adatok eredetére vonatkozó információk formájában elegendő kontextus álljon rendelkezésre, valamint az, hogy az adatok használhatók legyenek.

A használhatóság megköveteli, hogy az adatok megtalálhatók és hozzáférhetőek legyenek, még hozzá megfelelő fájlformátumokban. Az adatok minőségét megítélő egyéneknek megfelelő eszközökkel kell rendelkezniük az eléréshez, továbbá biztosítani kell az adatok megfelelő mértékű integritását.

Az integritás az adatok minőségével szembeni elvárás, tehát az a bizonyosság, hogy az adatok teljesek és hiánytalanok, konzisztensek és helyesek intellektuálisan és technikai szempontból is. Az integritást a létrehozás és a használat bármely fázisában veszélyeztetetik emberi hibák. Mivel az adatok javítása mindig költséges, a legjobb gyakorlat az, ha kezdettől fogva helyes adatokkal dolgozunk.

Bizonyos tudományterületeken nem elégséges a nyers adatokat megosztanunk, a használhatósághoz szükség lehet az adatok létrehozásához használt, elemzéséhez és megosztásához használható eszközök és módszerek leírására is (Kowalczyk – Shankar, 2011).

Az adatfogalom értelmezése azért is eltérő, mert bizonyos megfigyelések eredményeit adatoknak tekintheti az, aki gyűjtötte őket, amit azonban a befogadók nem így fognak fel. Fordított esetben, a kutató birtokában lehet anyagoknak anélkül, hogy tudatosulna benne, hogy mint adatok milyen értéket képviselnek (Borgman, 2012).

Annak érdekében, hogy az adatok megtalálhatók, elérhetőek és használhatók legyenek, ki kell alakítanunk a tárolás és a visszakeresés megfelelő technikai és szervezeti infrastruktúráit. Bár a nagy adatállományok archiválásának megvan a technológiája, ez nem egyszerű feladat, mert adatainkat könnyen hozzáférhetővé és jól leírhatóvá kell tenni (Kowalczyk – Shankar, 2011).

A gondozásnak a digitális források jövőbeni használhatóságát olyan módon kell

biztosítania, hogy a kutatók tudhassák, mi-
da és honnan származik az adott objektum,
miért fontos és hogyan kell használni. Az
ehhez szükséges kontextust dokumentációk
vagy metaadatok, valamint ezek keveréke
adja meg.

A gondozás során szükség lehet arra, hogy
döntsünk adatok törléséről, ami bizalmas
adatok esetében biztonságos megsemmisíté-
süket jelentheti. Az adatok gondozásának
költségei meg is követelik, hogy időről időre
felülvizsgáljuk, mely adatokat őrzünk meg,
és ezt kell tennünk akkor is, amikor az ada-
tokat új formára alakítjuk, vagy az elavulástól
védendő, új hardver- és szoftverkörnyezetbe
visszük át (Pryor, 2012).

A kutatási adatok talán legtágabb szakmai
kontextusát a (néha adatinformációs írástudás-
nak nevezett) adat-írástudás jelenti, amelyet
legegyszerűbben úgy határozhatunk meg,
mint az adatok megértésének, használatának,
kezelésének képességét (Qin–D’Ignazio, 2010).

Összegzés

A nyíltságra és átláthatóságra olyan trendként
tekintenek a politikusok, a finanszírozó szer-
vezetek, valamint a kutatók, amely elhossa a
jobb és gyorsabban fejlődő tudományt és
innovációt. Sokak meggyőződése, hogy a
közpénzekből finanszírozott kutatások során
keletkezett adatok újrafelhasználása gazdasá-
gi előnyökkel is járhat. A (már említett) RE-
CODE projekt eredményei azonban azt
mutatták, hogy erre egyelőre nincsen elege-
dő bizonyíték. A megkérdezett kutatók véle-
ménye pozitív volt ugyan, viszont a megva-
lósíthatóságot sokan megkérdőjelelik, többek
között a feladatok összetettsége okán.

Kulcsfontossága van tehát az adatok értel-
mezésének, ami szakértelmet igényel, és ami
nélkül az adatok nem használhatók fel továb-
bi kutatásokhoz. Végül, de nem utolsósorban,
minden tudományág más és más megköze-
lítést igényel (Sveinsdottir, 2014).

Kulcsszavak: *kutatási adatok, nagy adatok, adatmegosztás, nyílt adatok*

IRODALOM

- Borgman, Christine L. (2007): *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. MIT Press, Cambridge, MA
- Borgman, Christine L. (2012): The Conundrum of Sharing Research Data. *Journal of the American Society for Information Science and Technology*. 63, 6, 1059–1078. DOI: 10.1002/asi.22634 • <http://onlinelibrary.wiley.com/doi/10.1002/asi.22634/pdf>
- Boulton, R. et al. (2012): *Science as an Open Enterprise*. The Royal Society, London • https://royalsocietypublishing.org/~/media/royal_society_content/policy/projects/sape/2012-06-20-saoe.pdf
- Boyd, Danah – Crawford, Kate (2012): Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon, *Information, Communication & Society*. 15, 5, 662–679. DOI: 10.1080/1369118X.2012.678878 • <http://www.tandfonline.com/doi/full/10.1080/1369118X.2012.678878#tabModule>
- Buckland, Michael (2011): Data Management as Bibliography. *Bulletin of the American Society for Information Science and Technology*. 37, 6, 34–37. • https://www.asis.org/Bulletin/Aug-11/AugSept11_Buckland.pdf
- Davies, Anna – Fidler, D. – Gorbis, M. (2011): *Future Work Skills, 2020*. Institute for the Future, Palo Alto, CA • <http://www.iff.org/our-work/global-landscape/work/future-work-skills-2020/>
- EC Guidelines (2013): *Guidelines on Data Management in Horizon 2020*. European Commission • http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- Kowalczyk, Stacy – Shankar, Kalpana (2011): Data Sharing in the Sciences. *Annual Review of Information Science and Technology*. 45, 1, 247–294. DOI: 10.1002/aris.2011.1440450113 • https://kalpanashankar.files.wordpress.com/2010/06/arist_data_sharing.pdf

- Murphy, Fiona (2014): Data and Scholarly Publishing: The Transforming Landscape. *Learned Publishing*, 27, 5, 3–7. DOI: 10.1087/20140502 • <http://www.ingentaconnect.com/content/alpsp/lp/2014/00000027/00000005/art00002>
- Norman, Hazel (2014): Mandating Data Archiving: Experiences from the Frontline. *Learned Publishing*, 27, 5, 35–38. DOI: 10.1087/20140507 • <http://www.ingentaconnect.com/content/alpsp/lp/2014/00000027/00000005/art00007>
- Pryor, Graham (2012): Why Manage Research Data? In: Pryor, Graham (ed.): *Managing Research Data*. Facet, London, 1–16.
- Qin, Jian – D’Ignazio, John (2010): Lessons Learned from a Two-year Experience in Science Data Literacy Education. In: *Proceedings of the 3rd Annual IATUL Conference*, June, 20–24, 2010 • <http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1009&context=iatul2010>
- Stuart, David (2011): *Facilitating Access to the Web of Data*. Facet, London
- Sveinsdottir, Thordis (2014): The Meaning of Data: On the Development of Open Access to Research Data. *Methodological Challenges*, February 17, 2014, • http://methodologicalchallenges.group.shef.ac.uk/?page_id=57
- Z. Karvalics László (2007): A Cyber-infrastruktúra mint aktuális kihívás és mint tudományozsziológiai probléma. *Magyar Tudomány*, 167, 4, 475–489. • <http://www.matud.iif.hu/07apr/11.html>
- URL1: RECODE – Policy Recommendations for Open Access to Research Data in Europe • <http://recodeproject.eu/>
- URL2: DOI – Digital Object Identifier • <http://www.doi.org/>
- URL3: Holland Adatdij (Dutch Data Prize) • <http://www.researchdata.nl/en/activities/data-prize/>

