



DOI errors and possible solutions for Web of Science

Junwen Zhu¹ · Guangyuan Hu² · Weishu Liu³

Received: 12 November 2018 / Published online: 19 December 2018
© Akadémiai Kiadó, Budapest, Hungary 2018

Abstract

As unique and permanent alphanumeric strings to identify objects, digital object identifier (DOI) has been increasingly used to identify academic publications. Previous studies have reported the incorrect assignment of a single DOI name to multiple papers in the Scopus database, yet it remains unknown if this also holds in other datasets. In this paper we found incorrect DOI names are also problematic in the Web of Science but with different errors of duplicate DOI names. Tentative solutions are proposed in the end.

Keywords Digital object identifier · Database errors · Solution · Bibliometric database

Introduction

DOI, namely digital object identifier, is a permanent alphanumeric string to uniquely identify objects. The DOI name is a case-insensitive string and is made up of a prefix beginning with “10.” and a suffix separated by a forward slash. The DOI is widely used to identify academic publications (Boudry and Chartron 2017; Gorraiz et al. 2016). The DOI name, like an ID card number, can be used to identify a specific publication within a huge database.

Franceschini et al. (2015) revealed that quite a few single DOI names were incorrectly assigned to multiple papers indexed in the Scopus database. Yet it remains unknown whether this also holds for other leading standardized publication datasets (Tang and Shapira 2011). In this paper, we developed a special search query and to experimentally probe the DOI name error problem in the Web of Science.

✉ Weishu Liu
wslu08@163.com

Junwen Zhu
jwzhu@ed.ecnu.edu.cn

Guangyuan Hu
hu.guangyuan@shufe.edu.cn

¹ Faculty of Education, East China Normal University, Shanghai 200062, China

² School of Public Economics and Administration, Shanghai University of Finance and Economics, Shanghai 200433, China

³ School of Information Management and Engineering, Zhejiang University of Finance and Economics, Hangzhou 310018, Zhejiang, China

Method

Let us start with admitting that there is no silver line which can systematically identify and retrieve all records indexed in Web of Science with wrong DOI names. Yet intuitively, also based on our past research experience that the number “0” is most often be treated as the letter “O”. Thus we use a special string (within the DOI name, the letter “O” appears between two 0–9 numbers) as an illustrating case to explore if such errors also happened to DOI, an increasingly heavily relied tool to identify research articles.

$$DO = \left(\begin{array}{l} 10. * 000* OR 10. * 001* OR 10. * 002* OR 10. * 003* OR 10. * 004* OR 10. * 005* OR \\ 10. * 006* OR 10. * 007* OR 10. * 008* OR 10. * 009* OR 10. * 100* OR 10. * 101* OR 10. * 102* \\ OR 10. * 103* OR 10. * 104* OR 10. * 105* OR 10. * 106* OR 10. * 107* OR 10. * 108* OR \\ 10. * 109* OR 10. * 200* OR 10. * 201* OR 10. * 202* OR 10. * 203* OR 10. * 204* OR 10. * 205* \\ OR 10. * 206* OR 10. * 207* OR 10. * 208* OR 10. * 209* OR 10. * 300* OR 10. * 301* OR \\ 10. * 302* OR 10. * 303* OR 10. * 304* OR 10. * 305* OR 10. * 306* OR 10. * 307* OR 10. * 308* \\ OR 10. * 309* OR 10. * 400* OR 10. * 401* OR 10. * 402* OR 10. * 403* OR 10. * 404* OR \\ 10. * 405* OR 10. * 406* OR 10. * 407* OR 10. * 408* OR 10. * 409* OR 10. * 500* OR 10. * 501* \\ OR 10. * 502* OR 10. * 503* OR 10. * 504* OR 10. * 505* OR 10. * 506* OR 10. * 507* OR \\ 10. * 508* OR 10. * 509* OR 10. * 600* OR 10. * 601* OR 10. * 602* OR 10. * 603* OR 10. * 604* \\ OR 10. * 605* OR 10. * 606* OR 10. * 607* OR 10. * 608* OR 10. * 609* OR 10. * 700* OR \\ 10. * 701* OR 10. * 702* OR 10. * 703* OR 10. * 704* OR 10. * 705* OR 10. * 706* OR 10. * 707* \\ OR 10. * 708* OR 10. * 709* OR 10. * 800* OR 10. * 801* OR 10. * 802* OR 10. * 803* OR \\ 10. * 804* OR 10. * 805* OR 10. * 806* OR 10. * 807* OR 10. * 808* OR 10. * 809* OR 10. * 900* \\ OR 10. * 901* OR 10. * 902* OR 10. * 903* OR 10. * 904* OR 10. * 905* OR 10. * 906* OR \\ 10. * 907* OR 10. * 908* OR 10. * 909* \end{array} \right)$$

Indexes = SCI - EXPANDED, SSCI, A&HCI Timespan = 1900 – 2017

We conducted the searching on November 5th, 2018 from the library of Xi’an Jiao Tong University. 319 records were hit by the search strategy. Only 310 records published in English were selected for further analysis. All the bibliographic information of these records is downloaded and further imported to Excel for further process. We used the DOI system (<http://dx.doi.org/>) as the golden rule to resolve the DOI names.

Analyses

Wrong DOI names

Among these 310 records, we find 119 DOI names downloaded from the Web of Science (the third column of Table 1) cannot be found in the DOI System. These unresolved DOI names cover records published from 2001 to 2017. For comparison, we also manually collected each record’s DOI name provided on the publisher’s webpage of each record (the fifth column of Table 1). We assumed that the DOI names provided by the publishers on the webpage is correct. Luckily, all the DOI names provided by the publishers can be found in the DOI System. The DOI names provided by the publishers

Table 1 A sample of incorrect DOI names in Web of Science

| Rank | Year | DOI from Web of Science | | DOI from the publisher | | Both two DOI names work in Web of Science |
|------|------|---|------------|---|------------|---|
| | | DOI name | DOI system | DOI name | DOI system | |
| 1 | 2001 | https://doi.org/10.1106/606r-6mlh-6mqa-hpr2 | Unfound | https://doi.org/10.1106/606r-6mlh-6mqa-hpr2 | Found | Yes |
| 2 | 2002 | https://doi.org/10.1191/0748233702th167o2 | Unfound | https://doi.org/10.1191/0748233702th167o2 | Found | Yes |
| 3 | 2002 | https://doi.org/10.1021/ma02o4271 | Unfound | https://doi.org/10.1021/ma02o4271 | Found | Yes |
| 4 | 2002 | https://doi.org/10.1177/088506602237106 | Unfound | https://doi.org/10.1177/088506602237106 | Found | Yes |
| 5 | 2002 | https://doi.org/10.1012/s0024610703004587 | Unfound | https://doi.org/10.1012/s0024610703004587 | Found | Yes |
| 6 | 2003 | https://doi.org/10.1021/jp0035232 | Unfound | https://doi.org/10.1021/jp0035232 | Found | Yes |
| 7 | 2003 | https://doi.org/10.1080/00926230390195506 | Unfound | https://doi.org/10.1080/00926230390195506 | Found | Yes |
| 8 | 2003 | https://doi.org/10.1093/jae/dkg2o3 | Unfound | https://doi.org/10.1093/jae/dkg2o3 | Found | Yes |
| 9 | 2003 | https://doi.org/10.1097/01.blo.0000074409o99625.ca | Unfound | https://doi.org/10.1097/01.blo.0000074409o99625.ca | Found | Yes |
| 10 | 2003 | https://doi.org/10.1021/ma02o3290 | Unfound | https://doi.org/10.1021/ma02o3290 | Found | No |
| 11 | 2003 | https://doi.org/10.1080/09286580490888807 | Unfound | https://doi.org/10.1080/09286580490888807 | Found | Yes |
| 12 | 2004 | https://doi.org/10.1017/s030500090406518 | Unfound | https://doi.org/10.1017/s030500090406518 | Found | Yes |
| 13 | 2004 | https://doi.org/10.1309/vbxbmk5r78o96v090 | Unfound | https://doi.org/10.1309/vbxbmk5r78o96v090 | Found | No |
| 14 | 2005 | https://doi.org/10.1021/joo5o9556 | Unfound | https://doi.org/10.1021/joo5o9556 | Found | Yes |
| 15 | 2005 | https://doi.org/10.1279/174367605x52o68 | Unfound | https://doi.org/10.1279/174367605x52o68 | Found | Yes |
| 16 | 2005 | https://doi.org/10.1309/np7fctc5o98nty3w | Unfound | https://doi.org/10.1309/np7fctc5o98nty3w | Found | Yes |
| 17 | 2005 | https://doi.org/10.1309/apo4o9452gnlr8w | Unfound | https://doi.org/10.1309/apo4o9452gnlr8w | Found | No |
| 18 | 2005 | https://doi.org/10.1017/s0022226704003o68 | Unfound | https://doi.org/10.1017/s0022226704003o68 | Found | Yes |
| 19 | 2005 | https://doi.org/10.1017/s044686x042429o8 | Unfound | https://doi.org/10.1017/s044686x042429o8 | Found | Yes |
| 20 | 2005 | https://doi.org/10.1179/1743284o5x20996 | Unfound | https://doi.org/10.1179/1743284o5x20996 | Found | Yes |
| 21 | 2005 | https://doi.org/10.1309/2exqv1gnhebx7o5 | Unfound | https://doi.org/10.1309/2exqv1gnhebx7o5 | Found | Yes |
| 22 | 2006 | https://doi.org/10.1191/1362168806ir2o3o2a | Unfound | https://doi.org/10.1191/1362168806ir2o3o2a | Found | No |
| 23 | 2006 | https://doi.org/10.1191/1362168806ir2o3o2a | Unfound | https://doi.org/10.1191/1362168806ir2o3o2a | Found | Yes |

Table 1 (continued)

| Rank | Year | DOI from Web of Science | | DOI from the publisher | | Both two DOI names work in Web of Science |
|------|------|---|------------|---|------------|---|
| | | DOI name | DOI system | DOI name | DOI system | |
| 24 | 2006 | https://doi.org/10.1021/la06o368v | Unfound | https://doi.org/10.1021/la060368v | Found | Yes |
| 25 | 2006 | https://doi.org/10.1021/txo5o358e | Unfound | https://doi.org/10.1021/txo50358e | Found | Yes |
| 26 | 2006 | https://doi.org/10.1070/mc2006v016n01abeh00o2228 | Unfound | https://doi.org/10.1070/mc2006v016n01abeh002228 | Found | Yes |
| 27 | 2007 | https://doi.org/10.1021/jo7o2o76j | Unfound | https://doi.org/10.1021/jo702076j | Found | Yes |
| 28 | 2007 | https://doi.org/10.1021/ol7o2280q | Unfound | https://doi.org/10.1021/ol702280q | Found | Yes |
| 29 | 2007 | https://doi.org/10.2754/avb2o0776040553 | Unfound | https://doi.org/10.2754/avb200776040553 | Found | Yes |
| 30 | 2007 | https://doi.org/10.1007/s11606-007-0307-0 | Unfound | https://doi.org/10.1007/s11606-007-0307-0 | Found | Yes |
| 31 | 2007 | https://doi.org/10.1007/s11606-007-0293-2 | Unfound | https://doi.org/10.1007/s11606-007-0293-2 | Found | Yes |
| 32 | 2007 | https://doi.org/10.1007/s11606-007-0303-4 | Unfound | https://doi.org/10.1007/s11606-007-0303-4 | Found | Yes |
| 33 | 2007 | https://doi.org/10.1007/s11606-007-0316-z | Unfound | https://doi.org/10.1007/s11606-007-0316-z | Found | Yes |
| 34 | 2007 | https://doi.org/10.1021/jf07o453o | Unfound | https://doi.org/10.1021/jf070453o | Found | Yes |
| 35 | 2007 | https://doi.org/10.1021/jf07o8600 | Unfound | https://doi.org/10.1021/jf0708600 | Found | Yes |
| 36 | 2007 | https://doi.org/10.1021/cg07o6o2v | Unfound | https://doi.org/10.1021/cg070602v | Found | Yes |
| 37 | 2007 | https://doi.org/10.1310/tst4o6-45 | Unfound | https://doi.org/10.1310/tst1406-45 | Found | No |
| 38 | 2007 | https://doi.org/10.1021/cmo7o628 | Unfound | https://doi.org/10.1021/cm070628+ | Found | Yes |
| 39 | 2007 | https://doi.org/10.1021/cmo7o976b | Unfound | https://doi.org/10.1021/cm070976b | Found | Yes |
| 40 | 2007 | https://doi.org/10.1017/s0022112007007604 | Unfound | https://doi.org/10.1017/s0022112007007604 | Found | No |
| 41 | 2007 | https://doi.org/10.1093/rheumatology/kem2o4 | Unfound | https://doi.org/10.1093/rheumatology/kem204 | Found | Yes |
| 42 | 2007 | https://doi.org/10.1021/jo07o439a | Unfound | https://doi.org/10.1021/jo070439a | Found | Yes |
| 43 | 2007 | https://doi.org/10.1080/073283o0701787198 | Unfound | https://doi.org/10.1080/07328300701787198 | Found | Yes |
| 44 | 2007 | https://doi.org/10.1021/ac07o362q | Unfound | https://doi.org/10.1021/ac070362q | Found | Yes |
| 45 | 2007 | https://doi.org/10.1039/b7o8744f | Unfound | https://doi.org/10.1039/b708744f | Found | Yes |
| 46 | 2007 | https://doi.org/10.1309/ghtewlvocd8o25jl | Unfound | https://doi.org/10.1309/ghtewlvocd8025jl | Found | Yes |

Table 1 (continued)

| Rank | Year | DOI from Web of Science | | DOI from the publisher | | Both two DOI names work in Web of Science | |
|------|------|---|------------|---|------------|---|------------|
| | | DOI name | DOI system | DOI name | DOI system | DOI system | DOI system |
| 47 | 2007 | https://doi.org/10.1111/j.1399-3062.2006.00162.x | Unfound | https://doi.org/10.1111/j.1399-3062.2006.00162.x | Found | Found | Yes |
| 48 | 2007 | https://doi.org/10.1179/136404607x268085 | Unfound | https://doi.org/10.1179/136404607x268085 | Found | Found | Yes |
| 49 | 2007 | https://doi.org/10.1021/es070908q | Unfound | https://doi.org/10.1021/es070908q | Found | Found | No |
| 50 | 2008 | https://doi.org/10.1021/jp803219x | Unfound | https://doi.org/10.1021/jp803219x | Found | Found | Yes |
| 51 | 2008 | https://doi.org/10.1002/c.c.20406 | Unfound | https://doi.org/10.1002/c.c.20406 | Found | Found | No |
| 52 | 2008 | https://doi.org/10.1021/jf800406v | Unfound | https://doi.org/10.1021/jf800406v | Found | Found | Yes |
| 53 | 2008 | https://doi.org/10.1021/ma7o2222a | Unfound | https://doi.org/10.1021/ma7o2222a | Found | Found | Yes |
| 54 | 2008 | https://doi.org/10.1017/s136898008002206 | Unfound | https://doi.org/10.1017/s136898008002206 | Found | Found | No |
| 55 | 2008 | https://doi.org/10.1021/cm7o3530n | Unfound | https://doi.org/10.1021/cm7o3530n | Found | Found | Yes |
| 56 | 2008 | https://doi.org/10.1021/bi7o24904 | Unfound | https://doi.org/10.1021/bi7o24904 | Found | Found | Yes |
| 57 | 2008 | https://doi.org/10.1021/bi7o2223c | Unfound | https://doi.org/10.1021/bi7o2223c | Found | Found | Yes |
| 58 | 2008 | https://doi.org/10.1021/bi7o2238b | Unfound | https://doi.org/10.1021/bi7o2238b | Found | Found | Yes |
| 59 | 2008 | https://doi.org/10.1021/bi7o2537s | Unfound | https://doi.org/10.1021/bi7o2537s | Found | Found | Yes |
| 60 | 2008 | https://doi.org/10.1021/bi7o25485 | Unfound | https://doi.org/10.1021/bi7o25485 | Found | Found | Yes |
| 61 | 2008 | https://doi.org/10.1021/bi7o25084 | Unfound | https://doi.org/10.1021/bi7o25084 | Found | Found | Yes |
| 62 | 2008 | https://doi.org/10.1021/bi7o2222x | Unfound | https://doi.org/10.1021/bi7o2222x | Found | Found | Yes |
| 63 | 2008 | https://doi.org/10.1021/bi7o2492n | Unfound | https://doi.org/10.1021/bi7o2492n | Found | Found | Yes |
| 64 | 2008 | https://doi.org/10.1021/es7o27394 | Unfound | https://doi.org/10.1021/es7o27394 | Found | Found | Yes |
| 65 | 2008 | https://doi.org/10.7021/es7o2861u | Unfound | https://doi.org/10.7021/es7o2861u | Found | Found | Yes |
| 66 | 2008 | https://doi.org/10.1007/s10198-008-0099-0 | Unfound | https://doi.org/10.1007/s10198-008-0099-0 | Found | Found | Yes |
| 67 | 2008 | https://doi.org/10.1017/s0025315408000453 | Unfound | https://doi.org/10.1017/s0025315408000453 | Found | Found | No |
| 68 | 2008 | https://doi.org/10.1002/sml.200701206 | Unfound | https://doi.org/10.1002/sml.200701206 | Found | Found | Yes |

Table 1 (continued)

| Rank | Year | DOI from Web of Science | | DOI from the publisher | | Both two DOI names work in Web of Science | |
|------|------|---|------------|---|------------|---|------------|
| | | DOI name | DOI system | DOI name | DOI system | DOI system | DOI system |
| 69 | 2008 | https://doi.org/10.1017/s0047404508080512 | Unfound | https://doi.org/10.1017/s0047404508080512 | Found | No | |
| 70 | 2008 | https://doi.org/10.1021/cm7o2283u | Unfound | https://doi.org/10.1021/cm7o2283u | Found | Yes | |
| 71 | 2008 | https://doi.org/10.1021/jo7o2570q | Unfound | https://doi.org/10.1021/jo7o2570q | Found | Yes | |
| 72 | 2008 | https://doi.org/10.1021/ac7o2408u | Unfound | https://doi.org/10.1021/ac7o2408u | Found | Yes | |
| 73 | 2008 | https://doi.org/10.1213/ane.0b013e318163f602 | Unfound | https://doi.org/10.1213/ane.0b013e318163f602 | Found | No | |
| 74 | 2008 | https://doi.org/10.1179/174602208x269409 | Unfound | https://doi.org/10.1179/174602208x269409 | Found | Yes | |
| 75 | 2008 | https://doi.org/10.1017/s0030500410700076x | Unfound | https://doi.org/10.1017/s0030500410700076x | Found | No | |
| 76 | 2008 | https://doi.org/10.1021/bi7o23817 | Unfound | https://doi.org/10.1021/bi7o23817 | Found | Yes | |
| 77 | 2008 | https://doi.org/10.1016/j.yexcr.2o07.11.012 | Unfound | https://doi.org/10.1016/j.yexcr.2o07.11.012 | Found | Yes | |
| 78 | 2008 | https://doi.org/10.1016/j.yexcr.2o07.11.012 | Unfound | https://doi.org/10.1016/j.yexcr.2o07.11.012 | Found | No | |
| 79 | 2008 | https://doi.org/10.1021/jo7o2o43g | Unfound | https://doi.org/10.1021/jo7o2o43g | Found | Yes | |
| 80 | 2008 | https://doi.org/10.1021/jo7o2o62n | Unfound | https://doi.org/10.1021/jo7o2o62n | Found | Yes | |
| 81 | 2008 | https://doi.org/10.1163/156855508x292437 | Unfound | https://doi.org/10.1163/156855508x292437 | Found | Yes | |
| 82 | 2008 | https://doi.org/10.1134/s002315840801014x | Unfound | https://doi.org/10.1134/s002315840801014x | Found | Yes | |
| 83 | 2008 | https://doi.org/10.1007/s12o28-007-9021-3 | Unfound | https://doi.org/10.1007/s12o28-007-9021-3 | Found | Yes | |
| 84 | 2008 | https://doi.org/10.1309/1mvo1xrb9o3jzv | Unfound | https://doi.org/10.1309/1mvo1xrb9o3jzv | Found | Yes | |
| 85 | 2008 | https://doi.org/10.1185/0300799o8x261078 | Unfound | https://doi.org/10.1185/0300799o8x261078 | Found | Yes | |
| 86 | 2008 | https://doi.org/10.1185/0300799o8x261050 | Unfound | https://doi.org/10.1185/0300799o8x261050 | Found | Yes | |
| 87 | 2008 | https://doi.org/10.1185/0300799o8x261014 | Unfound | https://doi.org/10.1185/0300799o8x261014 | Found | Yes | |
| 88 | 2009 | https://doi.org/10.1071/zo08065 | Unfound | https://doi.org/10.1071/zo08065 | Found | No | |
| 89 | 2010 | https://doi.org/10.1017/s0022112009993922 | Unfound | https://doi.org/10.1017/s0022112009993922 | Found | No | |
| 90 | 2010 | https://doi.org/10.1021/c9o2o44g | Unfound | https://doi.org/10.1021/c9o2o44g | Found | No | |

Table 1 (continued)

| Rank | Year | DOI from Web of Science | | DOI from the publisher | | Both two DOI names work in Web of Science |
|------|------|---|------------|---|------------|---|
| | | DOI name | DOI system | DOI name | DOI system | |
| 91 | 2010 | https://doi.org/10.1309/ajpc2o7yswkctgs | Unfound | https://doi.org/10.1309/ajpc2q7yswkctgs | Found | No |
| 92 | 2011 | https://doi.org/10.1525/jm.2o11.28.1.81 | Unfound | https://doi.org/10.1525/jm.2011.28.1.81 | Found | No |
| 93 | 2012 | https://doi.org/10.1039/c2ob26019k | Unfound | https://doi.org/10.1039/c2ob26019k | Found | No |
| 94 | 2012 | https://doi.org/10.1016/j.jfoodmicro.2o12.08.002 | Unfound | https://doi.org/10.1016/j.jfoodmicro.2012.08.002 | Found | No |
| 95 | 2012 | https://doi.org/10.1021/jp3o6051w | Unfound | https://doi.org/10.1021/jp3o6051w | Found | No |
| 96 | 2012 | https://doi.org/10.1603/en11o72 | Unfound | https://doi.org/10.1603/en11o72 | Found | No |
| 97 | 2013 | https://doi.org/10.1021/es4o3767y | Unfound | https://doi.org/10.1021/es4o3767y | Found | No |
| 98 | 2013 | https://doi.org/10.1021/ac4o214o2 | Unfound | https://doi.org/10.1021/ac4o214o2 | Found | No |
| 99 | 2013 | https://doi.org/10.1021/ac4o2745w | Unfound | https://doi.org/10.1021/ac4o2745w | Found | No |
| 100 | 2013 | https://doi.org/10.1016/j.physletb.2o13.08.021 | Unfound | https://doi.org/10.1016/j.physletb.2013.08.021 | Found | No |
| 101 | 2014 | https://doi.org/10.1021/ac5o2627w | Unfound | https://doi.org/10.1021/ac5o2627w | Found | No |
| 102 | 2014 | https://doi.org/10.1016/j.ejogrb.2o14.10.002 | Unfound | https://doi.org/10.1016/j.ejogrb.2014.10.002 | Found | No |
| 103 | 2014 | https://doi.org/10.1158/1538-7755.disp13-0o3 | Unfound | https://doi.org/10.1158/1538-7755.disp13-c03 | Found | No |
| 104 | 2014 | https://doi.org/10.1158/1538-7755.disp13-0o2 | Unfound | https://doi.org/10.1158/1538-7755.disp13-c02 | Found | No |
| 105 | 2014 | https://doi.org/10.5343/bms.2o14.1o17 | Unfound | https://doi.org/10.5343/bms.2014.1o17 | Found | No |
| 106 | 2014 | https://doi.org/10.3347/kjp.2o14.52.4.449 | Unfound | https://doi.org/10.3347/kjp.2014.52.4.449 | Found | No |
| 107 | 2014 | https://doi.org/10.1179/14787318147.00000o000056 | Unfound | https://doi.org/10.1179/1478731814z.000000000056 | Found | No |
| 108 | 2014 | https://doi.org/10.5343/bms.2o13.1o77 | Unfound | https://doi.org/10.5343/bms.2013.1o77 | Found | No |
| 109 | 2014 | https://doi.org/10.1525/hsns.2o14.44.3.277 | Unfound | https://doi.org/10.1525/hsns.2014.44.3.277 | Found | No |
| 110 | 2014 | https://doi.org/10.1080/14029251.2o14.936763 | Unfound | https://doi.org/10.1080/14029251.2014.936763 | Found | No |
| 111 | 2015 | https://doi.org/10.1309/ajcpa01o4dwbkqbm | Unfound | https://doi.org/10.1309/ajcpa0o4dwbkqbm | Found | No |
| 112 | 2015 | https://doi.org/10.1016/j.prevetmed.2o15.o1.007 | Unfound | https://doi.org/10.1016/j.prevetmed.2015.o1.007 | Found | No |

Table 1 (continued)

| Rank | Year | DOI from Web of Science | | DOI from the publisher | | Both two DOI names work in Web of Science |
|------|------|---|------------|---|------------|---|
| | | DOI name | DOI system | DOI name | DOI system | |
| 113 | 2015 | https://doi.org/10.1016/j.matlet.2014.10.134 | Unfound | https://doi.org/10.1016/j.matlet.2014.10.134 | Found | No |
| 114 | 2015 | https://doi.org/10.1039/c5ob000959f | Unfound | https://doi.org/10.1039/c5ob000959f | Found | No |
| 115 | 2016 | https://doi.org/10.1016/j.jedc.2016.10.005 | Unfound | https://doi.org/10.1016/j.jedc.2016.10.005 | Found | No |
| 116 | 2016 | https://doi.org/10.1016/j.jhazmat.2015.11.017 | Unfound | https://doi.org/10.1016/j.jhazmat.2015.11.017 | Found | No |
| 117 | 2017 | https://doi.org/10.1016/j.jsbmb.2016.09.000 | Unfound | https://doi.org/10.1016/j.jsbmb.2016.09.000 | Found | No |
| 118 | 2017 | https://doi.org/10.1039/c7ob00178a | Unfound | https://doi.org/10.1039/c7ob00178a | Found | No |
| 119 | 2017 | https://doi.org/10.1525/ncl.2017.71.4.516 | Unfound | https://doi.org/10.1525/ncl.2017.71.4.516 | Found | No |

Accessed on November 5th, 2018 from the library of Xi'an Jiao Tong University

can be linked to the papers via the DOI System. That is to say, the 119 DOI names downloaded from the Web of Science are incorrect. By comparing the DOI names in the third and fifth columns of Table 1, we find that not only the letter “O” is easily confused with the number “0”, but also similar characters such as “b” and “6”, “O” and “Q” are easily to be confused.

One paper with two different DOI names

Interestingly, for 73 of all the 119 records, both the DOI names downloaded from the Web of Science and those collected from the publishers can be used to search the specific record in Web of Science (last column of Table 1). For example, the query DO = (“<https://doi.org/10.1106/606r-6mlh-6mq-a-hpr2>” and “<https://doi.org/10.1106/6o6r-6mlh-6mq-a-hpr2>”) can be used to search the first record in the Web of Science. Put it in other words, for some records, the Web of Science has stored both the wrong DOI names and right DOI names. However, we are provided with the wrong DOI names when we download the bibliographic data.

We also find one article with two correct DOI names. For example, both the DOI names <https://doi.org/10.1109/tgrs.20o4.826811> and <https://doi.org/10.1109/tgrs.2004.826811> can be resolved to this article via the DOI System but only the second DOI name works in the Web of Science. By a closer checking, we find the first DOI name appears on the publisher’s webpage of this article, however, the second DOI name appears on the PDF file of this article.

Discussion

As noted above, there is no simple way to identify and thus to assess the extent of DOI errors in the Web of Science dataset. This paper, yet, with a special search query reminds us the existence of DOI name errors in the Web of Science, though we do not know how representative this common error is when compared to other possible mistypes. For some records, Web of Science may contain two different DOI names for one record.

Based on our analysis, we argue that since the wrong DOI names cannot be resolved via the DOI System, Web of Science can cooperate with the DOI System to identify potentially wrong DOI names and then recollect the correct DOI names from the publishers. For those articles with two different DOI names problem, Web of Science should identify the correct DOI names and keep the correct DOI names only. We hope this problem can be fixed in the near future.

Both the Web of Science and Scopus are widely used in bibliometric analysis (Liao et al. 2019; Yu et al. 2018). However, these two databases are not free of errors/limitations (Franceschini et al. 2016; Hu et al. 2012; Liu 2017; Liu et al. 2018). For example, the funding is a great concern for academia (Liu et al. 2019), however, the funding acknowledgement information collected by Web of Science is not free of error/limitation (Tang et al. 2017). Users should be aware of these errors/limitations and at the same time the database providers should try to solve or improve the data quality problem.

Acknowledgements This research is partially supported by Zhejiang Provincial Natural Science Foundation of China (Grant numbers LQ18G010005 and LQ18G030010) and the National Natural Science Foundation of China (Grant numbers 71801189, 91746202 and 71433006).

References

- Boudry, C., & Chartron, G. (2017). Availability of digital object identifiers in publications archived by PubMed. *Scientometrics*, *110*(3), 1453–1469. <https://doi.org/10.1007/s11192-016-2225-6>.
- Franceschini, F., Maisano, D., & Mastrogiacomo, L. (2015). Errors in DOI indexing by bibliometric databases. *Scientometrics*, *102*(3), 2181–2186. <https://doi.org/10.1007/s11192-014-1503-4>.
- Franceschini, F., Maisano, D., & Mastrogiacomo, L. (2016). The museum of errors/horrors in Scopus. *Journal of Informetrics*, *10*(1), 174–182. <https://doi.org/10.1016/j.joi.2015.11.006>.
- Gorraiz, J., Melero-Fuentes, D., Gumpenberger, C., & Valderrama-Zurián, J. C. (2016). Availability of digital object identifiers (DOIs) in Web of Science and Scopus. *Journal of Informetrics*, *10*(1), 98–109. <https://doi.org/10.1016/j.joi.2015.11.008>.
- Hu, G., Carley, S., & Tang, L. (2012). Visualizing nanotechnology research in Canada: Evidence from publication activities, 1990–2009. *Journal of Technology Transfer*, *35*(4), 550–562. <https://doi.org/10.1007/s10961-011-9238-3>.
- Liao, H., Tang, M., Li, Z., & Benjamin, L. (2019). Bibliometric analysis for highly cited papers in operations research and management science from 2008 to 2017 based on essential science indicators. *Omega*. <https://doi.org/10.1016/j.omega.2018.11.005>.
- Liu, W. (2017). The changing role of non-English papers in scholarly communication: Evidence from Web of Science's three journal citation indexes. *Learned Publishing*, *30*(2), 115–123. <https://doi.org/10.1002/leap.1089>.
- Liu, F., Chen, Y. W., Yang, J. B., Xu, D. L., & Liu, W. (2019). Solving multiple-criteria R&D project selection problems with a data-driven evidential reasoning rule. *International Journal of Project Management*, *37*(1), 87–97. <https://doi.org/10.1016/j.ijproman.2018.10.006>.
- Liu, W., Hu, G., & Tang, L. (2018). Missing author address information in Web of Science—An explorative study. *Journal of Informetrics*, *12*(3), 985–997. <https://doi.org/10.1016/j.joi.2018.07.008>.
- Tang, L., & Shapira, P. (2011). China-US scientific collaboration in nanotechnology: patterns and dynamics. *Scientometrics*, *88*, 1–16. <https://doi.org/10.1007/s11192-011-0376-z>.
- Tang, L., Hu, G., & Liu, W. (2017). Funding acknowledgment analysis: Queries and caveats. *Journal of the Association for Information Science and Technology*, *68*(3), 790–794. <https://doi.org/10.1002/asi.23713>.
- Yu, D., Xu, Z., & Wang, W. (2018). Bibliometric analysis of fuzzy theory research in China: A 30-year perspective. *Knowledge-Based Systems*, *141*, 188–199. <https://doi.org/10.1016/j.knosys.2017.11.018>.