Annals of Science and Inquiry
Vol. 2 No. 11 (2026)

Lowercasing a Journal DOI Can Hide Its Preprint Link

Elizabeth Parker · Lauren Arnold · Noah W. Miller

Published: November 26, 2026 | Article 1208

Abstract

Crossref's journal and preprint records provide different directions for tracing a publication relationship. We tested two journal-DOI lookup routes on 369 exact preprint–journal pairs drawn from three Crossref namespaces. The journal's backlink field recovered 297 pairs, while a search using lowercase publication DOIs as relationship objects recovered 325. The union recovered 350. All 44 misses from the object search involved publication identifiers deposited with uppercase characters. Retrying those searches with the original spelling recovered all 44 pairs. Searching by an identifier and comparing returned identifiers therefore required different treatment of letter case in this public interface.

Lowercasing a Journal DOI Can Hide Its Preprint Link

Abstract

Crossref's journal and preprint records provide different directions for tracing a publication relationship. We tested two journal-DOI lookup routes on 369 exact preprint–journal pairs drawn from three Crossref namespaces. The journal's backlink field recovered 297 pairs, while a search using lowercase publication DOIs as relationship objects recovered 325. The union recovered 350. All 44 misses from the object search involved publication identifiers deposited with uppercase characters. Retrying those searches with the original spelling recovered all 44 pairs. Searching by an identifier and comparing returned identifiers therefore required different treatment of letter case in this public interface.

1. Work identity and the direction of discovery

When a reader starts with a published article DOI, an associated preprint is most useful if it can be found directly from that starting point. A preprint server may already record the journal DOI on its own record. The reader does not know that server's preprint DOI when beginning the search. The question is whether the registry publishes the link in a way the reader can traverse backward. A deposited relationship is an ordered statement with a subject and an object. A bibliographic graph may derive an inverse link for the object, or it may require the reader to search for records that name that object. Both paths can be tested against the same known ordered pairs.

This distinction matters because a DOI'd preprint and its eventual journal article are separate digital objects. Paskin (2010) describes how the DOI system identifies objects and maintains resolution information. A DOI's persistence makes the identifier suitable for the starting key. It supplies little information about where in a metadata service a version relationship will appear. Hendricks et al. (2020) describe the distributed deposit and retrieval mechanisms through which Crossref makes bibliographic metadata available. The Crossref relation vocabulary distinguishes a preprint-to-publication statement from its inverse (Crossref, n.d.). A publication-facing user can ask for the deposited inverse on the article record or use the relationship-object search filter, which returns records whose metadata name the article (Crossref, 2025). The results can, in principle, differ even when a connection is known at the preprint side.

The measurable research questions are how often each publication-seeded route retrieves the exact preprint–article DOI pair and what additional information the corresponding registry fields preserve. A second outcome uses the documented version suffixes in two of the repositories to ask whether an apparent failure at exact DOI level can be explained by the article returning a different identifier in the same recognizable preprint family. Last, we summarize how each returned article-side relation identifies the record from which the assertion came. This distinguishes discovery of a connection from observing claims from two independently deposited sides.

A simple article outside the main frame helps explain why conditioning matters. The Crossref journal record for Yi et al. (2024), DOI 10.1016/j.esmoop.2024.102239, resolves as a journal article and supplies no relationship field in the selected public record. Its DOI can identify the published item. That public response leaves the preprint linkage question open. We do not use it as a negative control, because the study has no separately established preprint pair for it. The primary denominator instead starts from preprints with a registered forward link to a named article DOI. This places a concrete, publicly recorded pair behind every eligible discovery test.

2. Which relationship fields can answer the question?

Preprint sources and publication stages serve different functions. Bourne et al. (2017) describe how researchers can post a manuscript while it is still moving through journal submission and peer review. The National Information Standards Organization (2008) distinguishes different stages in the life of a journal article. Klump et al. (2021) make a related point about identifiers and version granularity in research datasets. An identifier assigned to an individual release enables more precise matching than an identifier used to describe the whole evolving object. In the preprint settings used here, EarthArXiv (n.d.) describes assigning persistent identifiers to its repository submissions. The Center for Open Science (n.d.) documents version-specific preprint DOIs across hosted platforms including PsyArXiv and SocArXiv. DataCite also distinguishes versioning practice from the way preprints or manuscripts are related to a separate publication DOI (DataCite, n.d.-a, n.d.-b). We use posted Crossref records and their stated relation targets, rather than inferring publication stage from DOI syntax.

Several empirical studies use linked preprints to answer substantive questions about publication. Lin et al. (2020) investigated the relationship between computer science preprints and subsequently published articles. Fraser et al. (2020) examined preprint availability alongside measures of later attention. Klein et al. (2016) compared text in preprint and published versions, and Akbaritabar et al. (2022) studied how reference lists changed across matched preprint–publication pairs. The choice of linkage procedure can affect the set that such comparisons can observe. A record found by querying one side may be invisible to a workflow arriving from the other side. The current analysis stops at public administrative metadata and tests two exact-identifier lookup procedures. It makes no comparison of publication contents or outcomes and does not estimate the proportion of preprints becoming journal articles.

Two design considerations follow. First, a journal record can return a useful backlink even when the assertion was established on the preprint record. The fields report where the relationship was asserted. Reciprocity in the graph is separate from independence of evidentiary sources. Second, the possible result is a property of the registry response at collection time. Rittman (2026) reported a case in which reciprocal relationships were removed from the Crossref REST API and were later being restored. That account motivates a snapshot without attributing differences found here to that particular incident. The source provides one acquisition interval, and the study lacks historical field observations with which to track the underlying change. Public Europe PMC documentation also describes its own preprint and publication search paths (Europe PMC, n.d.). This illustrates that differently constructed discovery routes exist. The study deliberately holds the registry fixed to isolate the effect of two public interfaces and two evidence structures within one system.

3. Registering the comparison population

3.1. A positive DOI-pair frame

We enumerated the Crossref REST results for posted content under three DOI prefixes during January through March 2024. The prefixes identify the EarthArXiv, PsyArXiv and SocArXiv strata used in the source queries. Each query restricted the Crossref publication-date field and required the returned record to be indexed under the relation type is-preprint-of. This release quarter is bounded by dates of the source records. It does not describe the time a preprint relation was deposited, when a journal peer-reviewed a manuscript, or when a later platform discovered an existing deposit. We retained all records returned by the queries and verified their totals and DOI uniqueness. The sampling frame is deliberately positive for a current registry declaration. It supports conditional retrieval assessments within those administrative observations.

We extracted the relation records on each posted-content DOI and kept the ones with the expected predicate and a syntactically usable DOI object. A string found in a generic relation field did not qualify. The DOI keys were matched by case-folded text after removing surrounding whitespace and recognized DOI resolver prefixes. The primary normalization preserved suffix characters, so a versioned identifier remains distinct from its base DOI. Cioffi et al. (2022) investigate errors in citation DOIs. That work illustrates why validity checks should be explicit even when the field advertises a DOI. A prefix or case correction is a narrow operational transformation. It does not recover a malformed identifier or declare a match between two different DOI strings.

From the resulting preprint relations, we separately retrieved the target DOIs through a grouped exact DOI lookup. The selected Crossref fields gave the target's registered type and relation entries. The 369 eligible pairs attached to 360 unique articles. A pair was eligible for the primary comparison only if the exact target could be fetched and its registered type was journal article. This protects the search denominator from combining absent target records with retrieved article records that have an absent backlink. It also keeps a preprint-to-other-object relation from being judged against an outcome meant for published journal items. If several distinct preprint DOIs pointed to one journal DOI, each exact pair remained an observation. Publication-level summaries would conceal that a pipeline might recover one version but miss another.

3.2. Two routes starting at the article

The first route inspected each retrieved journal article's has-preprint entries. A match required the exact originating preprint DOI. We also recorded the backlink's asserted-by value. In a journal-side relation, an object-side origin indicates that the connected record supplied the assertion. It therefore matters when interpreting a seemingly reciprocal pair. A source might label its graph bidirectional while documenting only one originating deposit. A second DOI on the same journal article contributes only if its exact key matches the pair being tested. Found linked siblings and matching version families are tracked separately from exact matches.

The second route queried the relationship-object index starting with the known journal DOI. It requested posted content within the preprint's prefix and specified the is-preprint-of type. It did not restrict the dates of returned preprints. A reader starting from a publication typically does not know when a particular source record was posted. For each search group, we inspected the returned records for the exact typed pair and accepted it only if its publication DOI was among the requested object keys. This post-check matters because API query eligibility and the explicit relationship in a returned record are different statements. It also distinguishes additional returned pairs outside the original quarter from matches to the forward frame. The output records the union of discoverable pairs across both routes and the incremental gain of searching the indexed field.

The sensitivity measure collapsed the terminal underscore, v and positive integer suffix only for OSF-format DOIs in the two named prefixes with documented versioning. It left EarthArXiv identifiers and other DOI text untouched. It then rematched the observed article-side relations by this simplified string plus the exact journal DOI. A family match can show why some workflows recover a usable preprint connection while exact matching misses the specified version. The rule is vocabulary-specific, and it neither tests content equivalence nor replaces the primary DOI-level result. No names, article titles, abstracts or text bodies were needed for the linked comparison.

The object route originally supplied normalized lowercase DOI keys, applying the same routine identifier representation as the pair match. We then tested the query input as a separate question. For all eligible pairs whose deposited relation object differed in spelling from the normalized DOI, we reran the object

query with the unchanged deposited string, grouped by source prefix. We kept the same explicit pair validation after receiving each response. This is a diagnostic with access to the original relation field. It isolates whether the result is sensitive to how the search parameter was supplied. A reader with only the article DOI may lack that exact spelling, so the diagnostic does not define the primary publication-only retrieval rate.

3.3. Interpreting the proportions

All outcomes are descriptive counts and shares of eligible ordered DOI pairs. These pairs can share source or target records. They are enumerated in one collection so repeated targets do not need to be modeled as independent draws. A proportion from this design measures the fraction recovered among exact paired records supported by a forward relation in the selected quarter. It measures neither precision against independent ground truth nor prevalence across preprints, publishers, years or indexes. Methods research by Park and Tosaka (2010) identifies the importance of testing metadata semantics within repository workflows. Batini et al. (2009) discuss how data quality measures depend on how the observation and criterion are defined. The exact DOI pair provides both a stable test target and a transparent explanation of what either route must recover.

4. Results

4.1. Publication lookup and pair recovery

The three queries returned 405 posted-content records with 405 distinct exact preprint–target DOI pairs and 396 distinct target DOIs. Thirty-six pairs did not meet the journal-side inclusion check. In 13 pairs the published DOI was not returned by the grouped exact Crossref lookup. In 23 it was returned with another registered type. Because the eligibility test concerns the registry at collection, the absent records are reported as unresolved in the measured route. It makes no determination about the DOI's registration status elsewhere. The remaining 369 pairs formed the comparison sample in Table 1.

A direct exact backlink recovered 297 of the 369 eligible pairs, or 80.5%. The relationship-object route recovered 325, or 88.1%. The two results overlapped on 272 pairs. Twenty-five pairs appeared only through the article record, and 53 appeared only through the indexed object route. Nineteen were not recovered by either compared route, despite their presence in the forward frame. The combined procedure recovered 350, or 94.9%. Its gain over the direct field was 53 cases, and its gain over the object query was 25. Retrieval counts by themselves do not establish that one method is nested in the other. Indeed, the smaller of the two returned sets supplied pairs absent from the larger one.

Table 1. *Observed exact preprint–journal DOI pair recovery from a publication DOI.*

Preprint prefix	Forward pairs	Eligible journal pairs	Exact journal backlink	Indexed object search	Found by both	Found by either	OSF-family journal backlink
EarthArXiv	68	63	0	44	0	44	0
PsyArXiv	218	200	191	181	172	200	194
SocArXiv	119	106	106	100	100	106	106
All three	405	369	297	325	272	350	300

Note. The direction shown by a direct article backlink is checked against the exact forward DOI pair. The indexed object search has no posting-date restriction, and the table counts only its exact matches in the eligible frame. The version-family rule leaves EarthArXiv DOIs unchanged.

EarthArXiv accounted for 44 of the 53 pairs recovered only by the object search and all 19 pairs absent from both compared routes. In the matched EarthArXiv journal records, none of the 63 exact source DOIs appeared as a direct backlink, while 44 appeared in forward records recovered by using the publication as an object. In PsyArXiv the direct route recovered 191 of 200 eligible pairs, compared with 181 through the object route. The union recovered all 200 because those routes missed different pairs. SocArXiv exhibited complete exact recovery through the displayed backlink field and 100 of 106 pairs through the object search. The pattern is visible at the same time in one registry. Variations in deposits, indexing and the composition of published targets remain candidate explanations, with no separate process test here.

4.2. Provenance and version granularity

The 297 exact article backlinks reported the preprint-side object as the assertion source for 295 pairs. The remaining two reported the journal-side subject. On the preprint side, 367 relations reported the preprint itself as the assertion source, while two reported the article as the origin. Those two pairs are precisely the two with a journal-side subject assertion. The paired provenance fields are therefore consistent with one reported origin per relationship, even where the pair is visible in both directions. A database export that retains only the presence of a relationship drops this evidence about whether reciprocity came from a derived inverse or a separately reported assertion.

The OSF-family sensitivity recovered three additional article-side pairs, all in PsyArXiv. Overall recovery through the direct article field rose from 297 to 300 out of 369. The added allowance changes a small number of exact misses even though version-specific identifiers are relevant in the largest stratum. Independently, the indexed object lookups discovered 14 valid typed pairs for the queried publications outside the original forward set. Thirteen shared an explicit OSF preprint family with a pair already in that set. These additional pairs are evidence that a returned result inventory can extend beyond its counted denominator. They remain separate from the exact matches in Table 1. The measurement question determines whether several version DOIs should be preserved as separately identified releases or deduplicated when constructing a list intended for discovery.

4.3. The query's literal text determined 44 reverse-search misses

Every one of the 325 pairs with a deposited publication-object string already matching the normalized form was recovered by the normalized object search. The other 44 pairs had deposited strings with uppercase characters and none was recovered by that lowercase search. Resubmitting the original object strings recovered all 44, including 19 eligible EarthArXiv pairs, 19 PsyArXiv pairs and six SocArXiv pairs. Neither lowercasing the returned DOI values for pair comparison nor allowing an OSF version suffix can explain this contrast. The changed input is the literal object key in the API search request, while the criterion for validating the returned pair stays fixed.

Adding the exact-spelling retries to the original search recovers the 369 eligible pairs, including all 19 originally absent from both article-seeded routes. This is a data-grounded diagnostic rather than a universal discovery method. Its successful spellings come from previously captured forward relations. Journal DOI responses did not provide those spelling variants for the 44 retries. The observation narrows the interpretation of the EarthArXiv result. The 19 cases absent from both primary pathways are still part of a recoverable registry relationship, provided the caller uses the recorded object spelling. The audit does not identify the internal search implementation that produces the measured sensitivity.

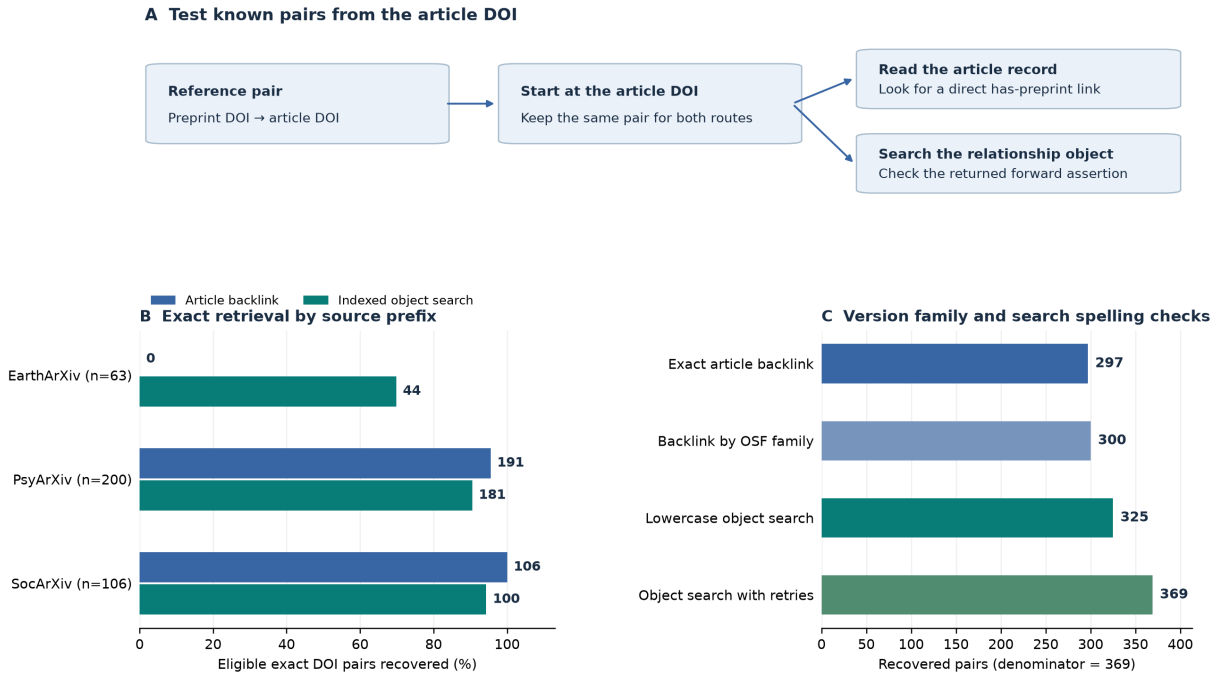


Figure 1. *Panel A shows the reference pair and two article-seeded lookup routes. Panel B compares exact DOI-pair recovery within each prefix stratum using the denominators in Table 1. Panel C compares the small version-family sensitivity with the gains from retaining the deposited spelling in the search parameter. All indexed matches are validated as exact normalized DOI pairs.*

5. What directional metadata can support

A reader starting with a journal DOI sees the subset of registry information exposed by the chosen route. This difference has consequences for computing a linked corpus. When the article response has an exact backlink, that response works as a direct join to the designated preprint. If it has another versioned DOI from the same family, it may help a reader locate a related version. An analysis at exact release level still needs the appropriate version key. If there is no relevant article-side backlink, a typed object search remains a separate way to test for the connection. The grade of evidence increases when the candidate returned by that search has the declared predicate and exact article object on the same record.

This approach fits with previous evidence about bibliographic interfaces. Visser et al. (2021) show that scholarly data sources can vary at the document and citation levels. Delgado-Quirós and Ortega (2025) compare the accessibility of research entity information across databases. Culbert et al. (2025) distinguish indexing a document from covering its cited references, a similar separation of a node from available edges. Massari et al. (2024) show how persistent bibliographic entity representation serves a graph built around scholarly relationships. Within one registry, an interface that exposes the target of a public relation search can display a larger effective edge inventory for the same starting identifier. The current study measures that question for one relation type and one explicitly bounded population.

Results in the source strata also need interpretation at record level. Deposits can originate in different systems, and de Jonge and Kramer (2026) document patterns of Crossref metadata completeness associated with production workflows. Our strata contain different preprint records and different journal targets. A prefix comparison describes those combinations in the specified quarter. It cannot identify an effect of a repository, publisher or later registry processing step. The literal-string retry demonstrates how one visible part of that heterogeneity can arise from a retrieval parameter. A broader comparison could separate the spelling distribution across source deposits from the state of journal-side backlink

availability. The current data offer no deposit timeline or platform assignment with which to adjudicate those explanations.

The distinction between an exact backlink and a route capable of recovering the pair is relevant to reproducibility. Klein et al. (2014) document instability in finding the context of web references over time. Even well-maintained identifiers can be accompanied by mutable registry relationships. Wilkinson et al. (2016) emphasize discoverable, reusable supporting metadata, and Peng (2011) describes the role of data and executable analyses in reproducing a computation. Gebru et al. (2021) apply similar reasoning to documenting the composition, intended use and collection of datasets. Freezing the route-specific records and their retrieval timestamps is enough for a third party to replicate the comparisons in this study. A future live query can then treat changes as new observations instead of requiring the historical count to persist.

The corpus is also limited by constructed eligibility. Forward claims may contain mistakes, several sources may point to the same publication, and some published versions will be absent from the initial Crossref predicates entirely. Pair recovery is neither a confirmation that both texts are equivalent nor a guarantee that a reader can access them. Because the two routes share Crossref as an underlying service, agreement offers evidence about complementary routes to the same public metadata. It supplies no independent attestation of publication history. For a bibliography dataset that needs exact preprint provenance, the result of a lookup can be stored as a pair of identifiers with the typed relation and route. That representation keeps the exact registered observation separable from assumptions about editions or platform behavior.

Data and code availability

The reproduction record identifies the public metadata routes, frozen projections and empirical analysis. The study retains only identifier, date, registered type and relationship fields used in the administrative comparison. Contributions, funding and competing-interest declarations await author confirmation.

References

Akbaritabar, Aliakbar; Stephen, Dimity; Squazzoni, Flaminio (2022). A study of referencing changes in preprint-publication pairs across multiple fields. *Journal of Informetrics, 16(2), 101258*. DOI.

Batini, Carlo; Cappiello, Cinzia; Francalanci, Chiara; Maurino, Andrea (2009). Methodologies for data quality assessment and improvement. *ACM Computing Surveys, 41(3), 1–52*. DOI.

Bourne, Philip E.; Polka, Jessica K.; Vale, Ronald D.; Kiley, Robert (2017). Ten simple rules to consider regarding preprint submission. *PLOS Computational Biology, 13(5), e1005473*. DOI.

Center for Open Science (n.d.). Welcome to OSF Preprints. *OSF Support*. Source.

Cioffi, Alessia; Coppini, Sara; Massari, Arcangelo; Moretti, Arianna; Peroni, Silvio; Santini, Cristian; Shahidzadeh Asadi, Nooshin (2022). Identifying and correcting invalid citations due to DOI errors in Crossref data. *Scientometrics, 127(6), 3593–3612*. DOI.

Crossref (2025). REST API filters. *Crossref documentation*. Source.

Crossref (n.d.). Relationships. *Crossref documentation*. Source.

Culbert, Jack H.; Hobert, Anne; Jahn, Najko; Haupka, Nick; Schmidt, Marion; Donner, Paul; Mayr, Philipp (2025). Reference coverage analysis of OpenAlex compared to Web of Science and Scopus. *Scientometrics, 130(4), 2475–2492*. DOI.

DataCite (n.d.-a). Versioning. *DataCite Support*. Source.

DataCite (n.d.-b). Preprints, post-prints and author manuscripts. *DataCite Support*. Source.

Yi, L.; Ning, Z.; Xu, L.; Shen, Y.; Zhu, X.; Yu, W.; Xie, J.; Meng, Z. (2024). The combination treatment of oncolytic adenovirus H101 with nivolumab for refractory advanced hepatocellular carcinoma: an open-label, single-arm, pilot study. *ESMO Open, 9(2), 102239*. DOI.

de Jonge, Hans; Kramer, Bianca (2026). Manuscript submission systems and metadata completeness in Crossref: Patterns and associations. *PLOS One, 21(3), e0345417*. DOI.

Delgado-Quirós, Lorena; Ortega, José Luis (2025). Research entity information and coverage in eight free access scholarly databases. *Online Information Review, 49(6), 1135–1153*. DOI.

EarthArXiv (n.d.). About EarthArXiv. *EarthArXiv*. Source.

Europe PMC (n.d.). Europe PMC preprint FAQs. *Europe PMC plus*. Source.

Fraser, Nicholas; Momeni, Fakhri; Mayr, Philipp; Peters, Isabella (2020). The relationship between bioRxiv preprints, citations and altmetrics. *Quantitative Science Studies, 1(2), 618–638*. DOI.

Gebru, Timnit; Morgenstern, Jamie; Vecchione, Briana; Vaughan, Jennifer Wortman; Wallach, Hanna; III, Hal Daumé; Crawford, Kate (2021). Datasheets for datasets. *Communications of the ACM, 64(12), 86–92*. DOI.

Hendricks, Ginny; Tkaczyk, Dominika; Lin, Jennifer; Feeney, Patricia (2020). Crossref: The sustainable source of community-owned scholarly metadata. *Quantitative Science Studies, 1(1), 414–427*. DOI.

Klein, Martin; Van de Sompel, Herbert; Sanderson, Robert; Shankar, Harihar; Balakireva, Lyudmila; Zhou, Ke; Tobin, Richard (2014). Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot. *PLoS ONE, 9(12), e115253*. DOI.

Klein, Martin; Broadwell, Peter; Farb, Sharon E.; Grappone, Todd (2016). Comparing Published Scientific Journal Articles to Their Pre-print Versions. *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries, 153–162*. DOI.

Klump, Jens; Wyborn, Lesley; Wu, Mingfang; Martin, Julia; Downs, Robert R.; Asmi, Ari (2021). Versioning Data Is About More than Revisions: A Conceptual Framework and Proposed Principles. *Data Science Journal, 20, 12*. DOI.

Lin, Jialiang; Yu, Yao; Zhou, Yu; Zhou, Zhiyang; Shi, Xiaodong (2020). How many preprints have actually been printed and why: a case study of computer science preprints on arXiv. *Scientometrics, 124(1), 555–574*. DOI.

Massari, Arcangelo; Mariani, Fabio; Heibi, Ivan; Peroni, Silvio; Shotton, David (2024). OpenCitations Meta. *Quantitative Science Studies, 5(1), 50–75*. DOI.

National Information Standards Organization (2008). Journal Article Versions (JAV): Recommendations. *NISO RP-8-2008*. Source.

Park, Jung-Ran; Tosaka, Yuji (2010). Metadata Quality Control in Digital Repositories and Collections: Criteria, Semantics, and Mechanisms. *Cataloging & Classification Quarterly, 48(8), 696–715*. DOI.

Paskin, Norman (2010). Digital Object Identifier (DOI®) System. *Encyclopedia of Library and Information Sciences, 3rd ed., 1586–1592*. DOI.

Peng, Roger D. (2011). Reproducible Research in Computational Science. *Science, 334(6060), 1226–1227*. DOI.

Rittman, Martyn (2026). Missing reciprocal relationships are being added back in. *Crossref Community Forum*. Source.

Visser, Martijn; van Eck, Nees Jan; Waltman, Ludo (2021). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quantitative Science Studies, 2(1), 20–41*. DOI.

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; Gonzalez-Beltran, Alejandra; Gray, Alasdair J.G.; Groth, Paul; Goble, Carole; Grethe, Jeffrey S.; Heringa, Jaap; 't Hoen, Peter A.C; Hooft, Rob; Kuhn, Tobias; Kok, Ruben; Kok, Joost; Lusher, Scott J.; Martone, Maryann E.; Mons, Albert; Packer, Abel L.; Persson, Bengt; Rocca-Serra, Philippe; Roos, Marco; van Schaik, Rene; Sansone, Susanna-Assunta; Schultes, Erik; Sengstag, Thierry; Slater, Ted; Strawn, George; Swertz, Morris A.; Thompson, Mark; van der Lei, Johan; van Mulligen, Erik; Velterop, Jan; Waagmeester, Andra; Wittenburg, Peter; Wolstencroft, Katherine; Zhao, Jun; Mons, Barend (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3(1), 160018*. DOI.