THE GENERATION OF KNOWLEDGE · TRUST, QUALITY AND THE FUTURE

Transparency, Provenance, and the Traceability of Knowledge

What transparency and provenance mean for knowledge, how the ability to trace a claim back to its sources protects against error and distortion, what happens when citations drift from their sources, the standards that support traceability, and how research organisations can make their findings traceable.

ModQu research article · 1,906 words · 20 references (APA 7th edition)

A claim is only as trustworthy as the path that leads back to its source. When a newspaper reports a statistic, a policy paper cites a study or a client presentation quotes a survey result, the reader is asked to accept something they cannot see being produced. If the path back to the source is clear, the reader, or someone acting on their behalf, can check it. If the path is broken, the claim must be taken on trust. **Transparency** is the openness of the process by which knowledge is produced. **Provenance** is the record of where a piece of knowledge came from and how it was produced. **Traceability** is the practical ability to follow that record back to the origin. This article examines why these qualities matter and how they can be built into the production of knowledge.

Provenance: where knowledge comes from

The term provenance comes from the art and archive worlds, where it refers to the history of ownership of a work or a record. Establishing provenance is how experts distinguish authentic works from forgeries and how archivists show that records have not been tampered with. The historian Marc Bloch, in a classic account of historical method, emphasised that the first task of the historian is to examine sources critically: to ask who produced a document, when, for what purpose and how it reached the present, before trusting what it says (Bloch, 1953).

In computing, provenance has become a field of study in its own right. Buneman et al. (2001) distinguished between **why-provenance**, which explains which source data justify a given result, and **where-provenance**, which identifies the locations in the source data from which a result was copied. Simmhan et al. (2005) reviewed provenance systems in scientific computing and argued that recording the derivation of data products, including the processes and parameters used, is essential for judging their quality, reproducing them and assigning credit.

The World Wide Web Consortium developed a standard model for recording provenance, known as PROV. It describes provenance in terms of three kinds of things: **entities**, such as datasets, documents and results; **activities**, the processes that use and generate entities; and **agents**, the people, organisations or software responsible for activities (Moreau & Missier, 2013). Relationships between them record, for example, that a table was derived from a dataset, that the dataset was generated by a weighting procedure, and that the procedure was carried out by a particular analyst. Such records can be read by both people and machines.

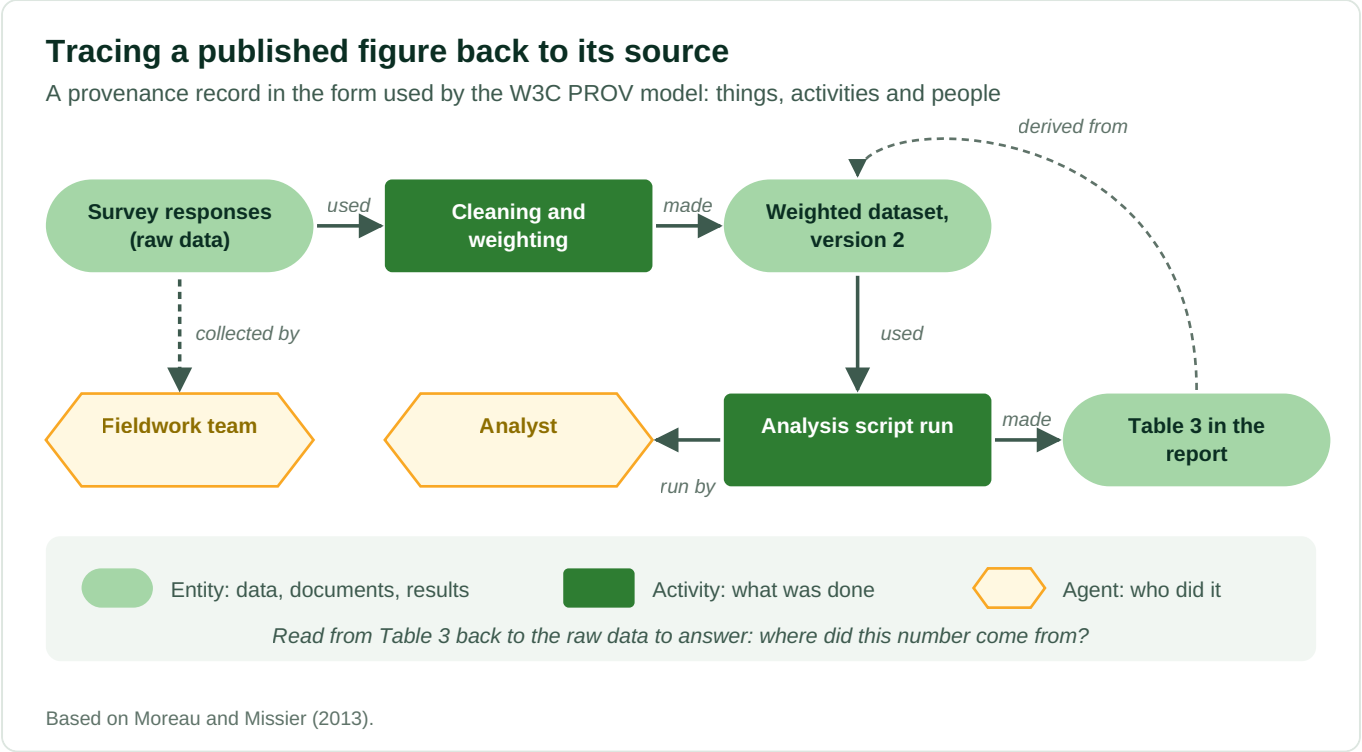


Figure 1. Tracing a published figure back to its source. Based on Moreau and Missier (2013); diagram by ModQu.

When traceability breaks

Citations are the main means of tracing claims in scholarship, but they do not always lead where they appear to. Greenberg (2009) analysed the network of citations surrounding a claim in biomedical research about a protein associated with a muscle disease. He found that the belief in the claim had grown largely through citations that amplified supportive studies, ignored contradictory ones and, in some cases, cited papers as support for statements they did not contain. He called this **citation distortion** and showed how it could create unfounded authority for a claim.

Mogull (2017) reviewed studies of the accuracy of quotations in medical research, that is, whether cited sources actually support the statements they are cited for. Across many studies, a substantial proportion of citations misrepresented their sources, some seriously.

A striking example concerns opioid painkillers. In 1980, Porter and Jick published a one-paragraph letter in the *New England Journal of Medicine* reporting that, among hospitalised patients who had received narcotics in their records, few had developed addiction (Porter & Jick, 1980). The letter described a specific group of patients treated in hospital, with no details of how addiction was assessed. Leung et al. (2017), a team of Canadian researchers, found that the letter had been cited hundreds of times, and that most citing articles used it as evidence that addiction was rare in patients treated with opioids, generally without describing what the letter actually reported. They concluded that this pattern of citation contributed to the opioid crisis in North America by helping to persuade doctors that opioids carried little risk of addiction. A reader who traced the claim back to its source would have found a brief letter about hospital patients, not evidence about long-term treatment in the community.

Provenance in intelligence and investigation

Intelligence and investigative work have long treated provenance as central. Analysts are trained to assess two things separately: the reliability of a source, based on its record and its access to the information, and the credibility of the particular information, based on whether it is consistent with other evidence. A widely used scheme, sometimes called the Admiralty system, grades sources with letters and information with numbers, so that a report can carry a compact record of how far it should be trusted. Irwin and Mandel (2019) reviewed such methods and found that they are often applied inconsistently, that the two dimensions are easily confused, and that grades can be lost or changed as reports are summarised and passed on. Their review shows that recording provenance is not enough. The record must be clear, consistently applied and kept with the information as it travels.

Provenance of digital media

Digital media raise new problems of provenance. Photographs, audio and video can now be altered or entirely generated by software in ways that are hard to detect. Chesney and Citron (2019) warned that such synthetic media, often called deepfakes, threaten not only individuals who may be falsely depicted but public trust in evidence generally, since genuine recordings can also be dismissed as fakes. One response has been to attach provenance records to media at the point of creation. The Coalition for Content Provenance and Authenticity (C2PA, 2022), formed by technology and media companies, published a technical standard for attaching signed records to images and other files that describe how they were created and edited. Such records do not prove that content is true, but they allow users to see where it came from and what has been done to it.

Transparency in research

Transparency means making the methods, data and reasoning behind findings open to inspection. The Royal Society (2012), in a report on science as an open enterprise, argued that mere disclosure is not enough. It called for **intelligent openness**: information should be accessible, intelligible, assessable and usable. Data dumped without explanation may be technically open but practically opaque. O'Neill (2002) made a similar point: transparency that floods people with information they cannot interpret does not build trust. What is needed is communication that people can follow and check.

Social scientists have pressed for greater transparency in their own fields. Miguel et al. (2014) argued that social science research should adopt three practices: disclosure of all measures, conditions and analyses; registration of studies and analysis plans in advance; and open sharing of data and materials. Christensen and Miguel (2018) reviewed evidence of problems in economics, including publication bias and the selective reporting of results, and argued that transparency practices are the most promising way of improving the credibility of research.

Nosek et al. (2015) proposed the Transparency and Openness Promotion (TOP) guidelines, which set out standards that journals can adopt at different levels of stringency. They cover the citation of data and code, the sharing of data, materials and analysis code, the preregistration of studies and analysis plans, and the publication of replications. Kidwell

et al. (2016) studied the effect of a simple incentive: awarding articles a badge for open data. After one psychology journal introduced the badges, the proportion of its articles that made data available increased sharply, while comparable journals without badges showed little change.

Transparency also extends to values. Elliott and Resnik (2014) argued that research relevant to policy inevitably involves value judgements, for example about how much evidence is enough to act on, and that scientists should make these judgements transparent so that users can see where they might reasonably have chosen differently.

Transparency in surveys and statistics

Survey research has a long tradition of methodological disclosure. The American Association for Public Opinion Research (2021) requires its members to disclose key information about any survey they publish, including the sponsor, the organisation that conducted it, the exact question wording, the population, the sampling method, the dates of collection, the sample size, the precision of estimates and how the data were weighted. These requirements allow readers to judge whether a survey's findings can support the conclusions drawn from them.

For statistics, traceability means that each published figure can be linked to the data, definitions and methods that produced it. When a number is reused in a report, a presentation or a news article, its link to the original source should travel with it, so that anyone who doubts it can check. The FAIR principles express the same aim for data in general: data should be findable, accessible, interoperable and reusable, which requires rich metadata and persistent identifiers that link data to their origins (Wilkinson et al., 2016).

Consider how traceability works for a single survey figure. Suppose a report states that 42% of small businesses in a province expect to hire more staff next year. A traceable figure allows a reader to find the survey that produced it, the date of fieldwork, the exact question asked and the answer options, the number of businesses that answered, how the sample was drawn, how results were weighted to represent all small businesses, and the margin of error. If the figure is later quoted in a news article as "most small businesses plan to hire", a reader who follows the trail back can see that the original figure was well below half, and that the question concerned expectations, not plans. Without the trail, the distortion would be invisible.

Building traceability into practice

Several practices make knowledge traceable:

- **Cite the original source.** When using a statistic or finding, cite the source that produced it, not a secondary source that repeated it, and check that the source says what it is cited for.
- **Record provenance as work is done.** Keep a record of where each dataset came from, which processes were applied to it and by whom, using version control and scripts rather than manual steps.
- **Link outputs to inputs.** Make it possible to trace each table, chart and figure in a report back to the data and code that produced it.
- **Disclose methods with findings.** Publish the essential information about how data were collected and analysed alongside the results.
- **Use persistent identifiers.** Give datasets and reports stable identifiers so that references to them remain valid.

- **Disclose interests and judgements.** State who funded the work and what value judgements shaped it.

Conclusion

Knowledge is trustworthy when its path back to the source can be followed. Provenance records where knowledge came from; transparency opens the process of producing it to inspection; traceability makes it possible to check. The history of citation distortion, including the opioid letter that was cited far beyond what it showed, shows how claims can drift from their sources when no one traces them back. Standards such as PROV, reporting and disclosure requirements, transparency guidelines and the FAIR principles provide practical means of keeping the path open. For organisations that produce research, traceability is not only a technical feature. It is a commitment to let others check, and therefore a foundation of the trust on which their work depends.

References

American Association for Public Opinion Research. (2021). *AAPOR code of professional ethics and practices*. AAPOR.

Bloch, M. (1953). *The historian's craft* (P. Putnam, Trans.). Knopf.

Buneman, P., Khanna, S., & Tan, W.-C. (2001). Why and where: A characterization of data provenance. In J. Van den Bussche & V. Vianu (Eds.), *Database theory: ICDT 2001* (Lecture Notes in Computer Science, Vol. 1973, pp. 316–330). Springer. https://doi.org/10.1007/3-540-44503-X_20

C2PA. (2022). *C2PA technical specification*. Coalition for Content Provenance and Authenticity.

Chesney, R., & Citron, D. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(6), 1753–1820.

Christensen, G., & Miguel, E. (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature, 56*(3), 920–980. https://doi.org/10.1257/jel.20171350

Elliott, K. C., & Resnik, D. B. (2014). Science, policy, and the transparency of values. *Environmental Health Perspectives, 122*(7), 647–650. https://doi.org/10.1289/ehp.1408107

Greenberg, S. A. (2009). How citation distortions create unfounded authority: Analysis of a citation network. *BMJ, 339*, b2680. https://doi.org/10.1136/bmj.b2680

Irwin, D., & Mandel, D. R. (2019). Improving information evaluation for intelligence production. *Intelligence and National Security, 34*(4), 503–525. https://doi.org/10.1080/02684527.2019.1569343

Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., Kennett, C., Slowik, A., Sonnleitner, C., Hess-Holden, C., Errington, T. M., Fiedler, S., & Nosek, B. A. (2016). Badges to acknowledge open practices: A simple, low-cost, effective method for increasing transparency. *PLoS Biology, 14*(5), e1002456. https://doi.org/10.1371/journal.pbio.1002456

Leung, P. T. M., Macdonald, E. M., Stanbrook, M. B., Dhalla, I. A., & Juurlink, D. N. (2017). A 1980 letter on the risk of opioid addiction. *New England Journal of Medicine, 376*(22), 2194–2195. https://doi.org/10.1056/NEJMc1700150

Miguel, E., Camerer, C., Casey, K., Cohen, J., Esterling, K. M., Gerber, A., Glennerster, R., Green, D. P., Humphreys, M., Imbens, G., Laitin, D., Madon, T., Nelson, L., Nosek, B. A., Petersen, M., Sedlmayr, R., Simmons, J. P., Simonsohn, U., & Van der Laan, M. (2014). Promoting transparency in social science research. *Science*, *343*(6166), 30–31. https://doi.org/10.1126/science.1245317

Mogull, S. A. (2017). Accuracy of cited "facts" in medical research articles: A review of study methodology and recalculation of quotation error rate. *PLoS ONE, 12*(9), e0184727. https://doi.org/10.1371/journal.pone.0184727

Moreau, L., & Missier, P. (Eds.). (2013). *PROV-DM: The PROV data model* (W3C Recommendation). World Wide Web Consortium.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

O'Neill, O. (2002). *A question of trust: The BBC Reith Lectures 2002*. Cambridge University Press.

Porter, J., & Jick, H. (1980). Addiction rare in patients treated with narcotics. *New England Journal of Medicine*, *302*(2), 123. https://doi.org/10.1056/NEJM198001103020221

Royal Society. (2012). *Science as an open enterprise*. The Royal Society.

Simmhan, Y. L., Plale, B., & Gannon, D. (2005). A survey of data provenance in e-science. *ACM SIGMOD Record*, *34*(3), 31–36. https://doi.org/10.1145/1084805.1084812

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*, Article 160018. https://doi.org/10.1038/sdata.2016.18