# *Nature*'s reach: narrow work has broad impact

Alexander J. Gates, Qing Ke, Onur Varol and Albert-László Barabási

A scientific paper today is inspired by more disciplines than ever before, shows a new analysis marking the journal's 150th anniversary.

How knowledge informs and alters disciplines is itself an enlightening, and vibrant field[1]. This type of meta research into new findings, insights, conceptual frameworks and techniques is important, among other things, for policymakers who fund research in the hope of tackling society's most pressing challenges, which inevitably span disciplines.

Since its founding in 1869, *Nature* has offered a venue for publishing major advances from many fields. To mark its anniversary, we track here how papers cite and are cited across disciplines, using data on tens of millions of scientific articles indexed in Clarivate Analytics' Web of Science (WoS), a bibliometric database that encompasses many thousands of research journals starting from 1900. We pay particular attention to articles that appeared in *Nature*. In our view, this snapshot, for all its idiosyncrasies, reveals how scientific work is ever more becoming a mixture of disciplines.

Several caveats are important. The volatility of our metrics in the early twentieth century can be attributed, at least in part, to the fact that articles then typically had many fewer references and citations. Until the mid-1920s, *Nature* articles typically listed no references; today, they can have up to 50. Another caveat is that the number of disciplines recognized by WoS grew from 57 in 1900 to 251 in 1993, but this is only one factor contributing to the disciplinary trends we found.

Many scholars have developed methods and metrics to gauge how scientific publishing contributes to knowledge, and to assess influence. For more detailed explanations of our choices, along with essential qualifications, see Supplementary Information (SI).

Across the scientific literature overall, our analysis hints that articles are drawing from and influencing more disciplines than they did 100 years ago, although some disciplines have broader influence than others. As a journal, *Nature* publishes mostly specialized, or deeply disciplinary, papers; these tend to reference a narrower range of disciplines than does the average paper. Usually, however, *Nature* papers are cited by a broader range of disciplines than average.

## Colossal corpus

We extracted references for papers contained in the WoS publication database from 1900 to 2017, capturing close to 700 million citation relationships. We pinned subsequent analysis to the approximately 19 million articles that had at least one reference and one citation and that were published before 2010 (to give time for citations to accumulate). The resulting corpus integrated the discipline information for 38 million articles.

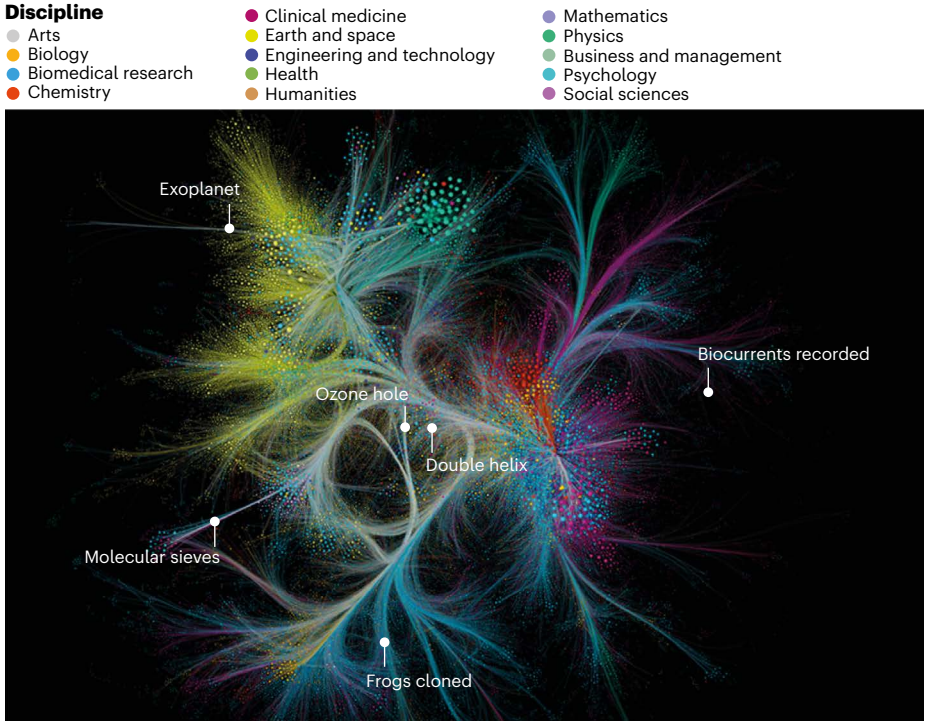
To identify disciplines, we relied on relatively broad categorizations from WoS. These are necessarily imperfect, but cumulatively reveal patterns of scholarship. Most journals are disciplinary, and so WoS assigns each article to one or more disciplines on the basis of the journal in which it is published. For instance, articles in the *Journal of Bacteriology* are categorized as microbiology.

We traced the conceptual journeys to each paper by identifying the inspiration for articles by their references: the works authors credited for their concepts, methods, techniques and insight. Similarly, we identified the impact of each publication by the citations it received in the corpus. Caution is required when using citation-based measures to assess the importance of individual papers or authors; still, the accessibility and quantity of such data provide one view — among many — of how scientific knowledge accumulates[1].

We explored how the 88,637 *Nature* articles in our data set mediate the metabolism of ideas using the broadest WoS disciplinary categories. A *Nature* article with references mainly from biomedical research will typically collect the largest proportion of its citations from other biomedical-research papers (see 'Knowledge flows'). About half of the papers that cite it will be spread across the other categories. By contrast, a paper with references mainly from engineering and technology is much more likely to be cited by papers in other fields (72%) than by other papers in the same field (28%). Engineering and technology papers also make up a very small proportion of the papers *Nature* opts to publish; those that are selected might be chosen for their broad appeal. At the other extreme, papers in Earth and space science are much more likely to be cited by papers in their own field (72%) than by other disciplines (28%).

## CO-CITATION NETWORK

Each *Nature* paper is a dot. Dots are linked if another paper cites both. Some articles (colourful clusters) are cited by many disciplines, others (monotone areas) are deeply embedded in their own disciplines. (See go.nature.com/n150int for an interactive version, including references to the six highlighted papers.)




SOURCE DATA: WEB OF SCIENCE. ANALYSIS BY A. J. GATES *ET AL.*

Another way to reveal intrinsic communities in and across disciplines is through co-citation analysis[2]. In this approach, each paper is represented by a node, shown as a dot. Two papers are linked if another paper cites both of them; the node size reflects the number of co-citations. Our visualization algorithm treats each link as a spring and arranges the nodes to make links as short as possible. This produces clusters of *Nature* papers that vary in their level of interdisciplinary connections (see go.nature.com/n150int).

The overall network structure echoes scientific perceptions of how publications relate to each other. Articles tend to bunch together according to age and topic, because authors usually reference recent articles related to their paper's subject[3]. Over its recent history, more than half of *Nature*'s papers have come from the life sciences. Consequently, clusters of biomedical-research papers appear throughout the network. Since 1930 (when it became reliable to use references to assign papers to disciplines), the proportion of physics papers has shrunk and Earth and space science has grown. Certain papers – such as the discovery of the first exoplanet orbiting a Sun-like star[4] – are deeply embedded in a cluster of papers in the same field. By contrast, the discovery of the ozone hole[5] is in a region where articles of many disciplines – chemistry, social sciences, Earth sciences – are found (see 'Co-citation network'). Our analysis shows that this paper's references are more diverse than 95% of *Nature* papers, and its citations are more diverse than 99% of *Nature* papers.

An analysis of the co-citation network from any more-specialized journal would probably look different. Still, distinct episodes from the history of science are apparent in the 3D view of *Nature*'s co-citation network (see go.nature.com/2patums). These include the study of radioactive elements in the 1930s, and how studies of superconducting materials flirted with diverse applications and then were intensely characterized deep within the physical sciences in the late 1980s and 1990s.

## Over time

The numbers of papers in every discipline grew exponentially over the past century[1]. Exact rates differ over time, although since about the 1960s, 48% of papers were in the life sciences (with 42% from 'hard' sciences and 10% from behavioural science).

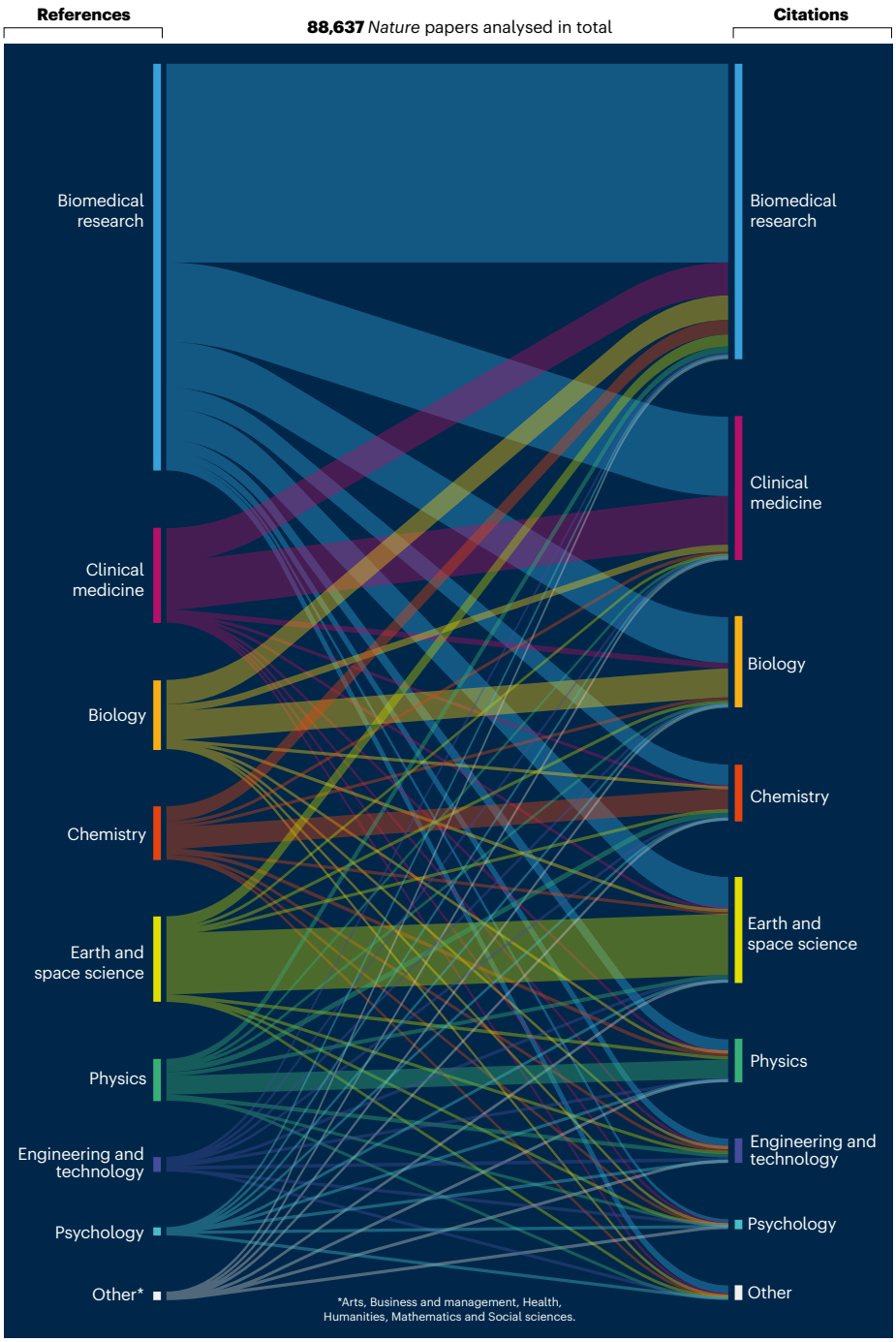
Scholars define and measure influences across disciplines in various ways. Multidisciplinarity usually refers to separate disciplines coming together yet remaining distinct: we define it for journals as the breadth of disciplines that are either inspiring or being impacted by the journal's articles. Interdisciplinarity refers to integration: we define it as the diversity in inspiration in an article's references, and the diversity in how an article's impact diffuses across disciplines. Although it is difficult to assess integration across an article's citations, this measure can capture how the knowledge communicated by the article had diverse impact[6]. This analysis indicates the extent of interactions across disciplines, but does not reveal the specific details of how those disciplines interact.

First, we explored the breadth of disciplines reflected in the references and citations across a journal, capturing the journal's multidisciplinarity (see 'Inspiration and impact'). We labelled each paper in a journal with the primary discipline assigned to its references (inspiration) or citations (impact), and measured multidisciplinarity on a scale of zero to one. Zero meant that all of an article's references or citations were in the same discipline; one meant that they were balanced evenly across all disciplines, using the normalized entropy measure (see SI). We found that this measure does not depend on the number of articles each journal published (see SI). It probably reflects other qualities of a journal, such as the pool of articles submitted and the editors' selection criteria.

For most journals, the breadth of impact and

## KNOWLEDGE FLOWS

*Nature* articles are mainly cited by their own disciplines, particularly in some fields, such as Earth and space science. (Each *Nature* paper was assigned to a discipline using its references, as was every paper in the Web of Science database that cited a *Nature* paper.)




SOURCE DATA: WEB OF SCIENCE. ANALYSIS BY A. J. GATES *ET AL.*

inspiration are highly correlated. This holds true for specialist journals such as *Cell* and *Physical Review Letters*. A typical journal today publishes articles inspired by and impacting about six disciplines.

The general-science journals *Nature* and *Science* both have a greater breadth of impact (citations) and inspiration (references) than 99.7% of other journals. The multidisciplinarity of *Nature* peaked in the 1960s and has remained relatively high since then, probably reflecting a combination of papers selected by *Nature* that are expected to have broad appeal, and the papers' greater visibility to the scientific community.

Second, we explored the interdisciplinarity of individual articles by measuring the diversity of disciplines in the references and citations[7–10]. Many measures have been proposed to assess interdisciplinarity, and can have inconsistent results (see, for example, refs 11,12). Scholars agree, however, that simply counting the number of disciplines that occur in references and citations is inadequate. For example, a paper that largely references biology and clinical science draws on less diversity than one inspired by biology and physics. We quantify this characteristic on a scale of zero to one using the Rao–Stirling diversity index, which captures the number of disciplines represented, how evenly they are distributed and their degree of difference[13].

Our analysis shows that the diversity of disciplines in articles' references and citations is increasing. Roughly speaking, a typical article is inspired by and impacts three times more disciplines this decade than it did 50 years ago.

Whereas a typical article published today references articles from the equivalent of 11 disciplines, a *Nature* publication references the equivalent of only 9 (SI, Fig. S5). This is in line with previous analyses suggesting that highly influential work tends to be grounded in deep expertise[14]. By contrast, the disciplinary diversity for the citations of articles in general-science journals has consistently been higher than for articles published elsewhere, suggesting that content in these journals reaches a broader swathe of the scientific community than it drew from. This observation makes sense, considering that these journals aim to reach a broader readership and to publish major advances.

Sometimes, the fields that inspire a paper differ markedly from those on which it has an impact. For example, 'The Digital Code of DNA', a 2003 *Nature* essay by systems biologists Leroy Hood and David Galas[15], takes most of its inspiration from molecular biology, yet is cited across computer science, clinical medicine and social science. We quantify cross-disciplinarity on a scale from zero to one. In this case, zero implies all disciplines that inspired an article and all those it impacts are identical; a score of one implies these lists differ completely (using the Jensen–Shannon divergence, a measure of the similarity between two probability distributions; see SI).

What we see is that in recent decades cross-disciplinarity has declined, with that of the general-science journals falling faster than the scientific literature overall. Perhaps this is because articles that bridge disciplines influence multiple fields, including those from which they arose. As works draw on a broader set of disciplines, there is less scope to influence a set of completely different disciplines.

Assessment of scientific work generally works best when contextualized within its specific discipline. For example, citation counts are more effective when comparing biomedical papers to other biomedical papers rather than to physics papers. But if interactions between disciplines are increasing, then a stringent, coherent assignment makes less sense. We speculate that considering how disciplines intermix within individual articles might allow better comparisons across disciplines or improve assessment of a paper's impact. What's more, strictly structured research departments and funding programmes make less sense if boundaries between disciplines are becoming less distinct. As network scientists, we relish the idea that science is becoming less siloed.

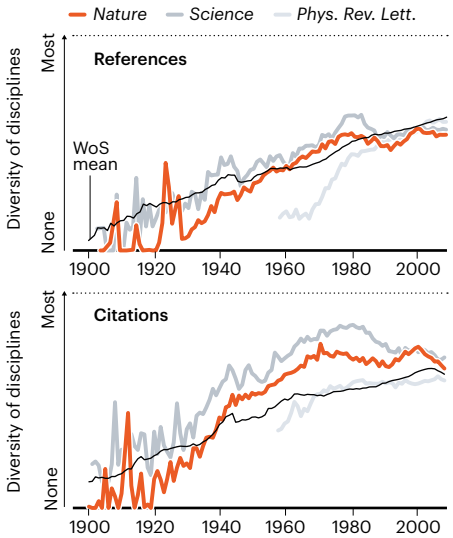
The increase we observe in interdisciplinary thinking is seen across disciplines (see SI) and shows no signs of slowing. With the population of researchers, scientific literature and knowledge ever growing, the scientific endeavour increasingly integrates across boundaries. Research institutions and funding bodies would do well to realize that interdisciplinarity is becoming the norm.

## INSPIRATION AND IMPACT

The diversity of disciplines in articles' citations (impact) and references (inspiration) is growing; the likelihood of articles crossing disciplines is not. Articles in *Nature* and *Science* are more broadly cited across disciplines.

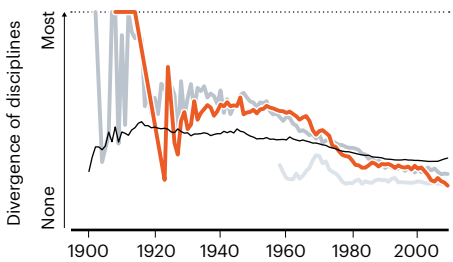
### Interdisciplinarity

How many, how diverse and how balanced disciplines are across an article's references and citations. This is growing across all of science.



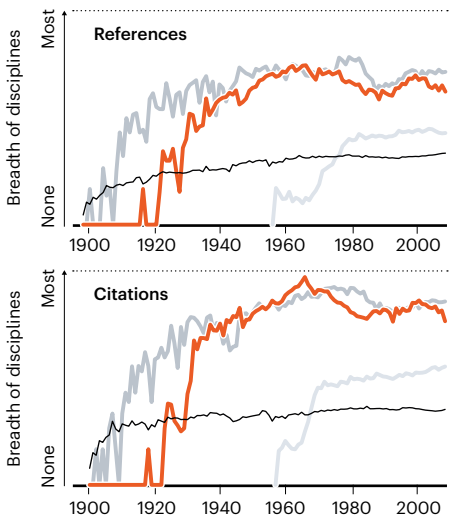

### Cross-disciplinarity

How much the disciplines in articles' references vary from those in their citations. The decline here is probably due to rising interdisciplinarity.




### Multidisciplinarity

How many disciplines are represented in a journal. Both *Nature* and *Science* consistently show broader impact than the average journal.




Volatile data pre-1930: papers had fewer citations and references, and indexing was less reliable.

SOURCE DATA: WEB OF SCIENCE. ANALYSIS BY A. J. GATES *ET AL.*

## The authors

**Alexander J. Gates**, **Qing Ke**, **Onur Varol** and **Albert-László Barabási** are researchers at the Network Science Institute, Northeastern University, Boston, Massachusetts, USA.
e-mail: a.barabasi@northeastern.edu

1. Fortunato, S. *et al. Science* **359**, eaao0185 (2018).
2. Small, H. *J. Am. Soc. Inf. Sci.* **24**, 265–269 (1973).
3. Mukherjee, S., Romero, D. M., Jones, B. & Uzzi, B. *Sci. Adv.* **3**, e1601315 (2017).
4. Mayor, M. & Queloz, D. *Nature* **378**, 355–359 (1995).
5. Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).
6. Leydesdorff, L., Wagner, C. S. & Bornmann, L. *Scientometrics* **114**, 567–592 (2018).
7. Choi, B. C. K. & Pak, A. W. P. *Clin. Invest. Med.* **29**, 351–364 (2006).
8. Porter, A. L. & Rafols, I. *Scientometrics* **81**, 719 (2009).
9. Wagner, C. S. *et al. J. Informetr.* **5**, 14–26 (2011).
10. Leydesdorff, L. & Rafols, I. *J. Informetr.* **5**, 87–100 (2011).
11. Wang, Q. & Schneider, J. W. Preprint at https://arxiv.org/abs/1810.00577 (2018).
12. Research Councils UK & Digital Science. *Interdisciplinary Research: Methodologies for Identification and Assessment* (RCUK/Digital Science, 2016).
13. Stirling, A. *J. R. Soc. Interface* **4**, 707–719 (2007).
14. Uzzi, B., Mukherjee, S., Stringer, M. & Jones, B. *Science* **342**, 468–472 (2013).
15. Hood, L. & Galas D. *Nature* **421**, 444–448 (2003).

Supplementary information accompanies this article: see go.nature.com/2wtoux3.