

TrendMD: Using AI to enhance discovery and achieve publisher goals

Bert Carelli*

Director of Partnerships, TrendMD, Inc., 2068 Braemar Road, Oakland, CA 94602, USA

Abstract. TrendMD’s recommendation engine uses Artificial Intelligence (AI) to connect ideas, subjects, and people. We help researchers discover new content related to their interests, within the context of their research workflow. Publishers use our service to grow their website traffic, build readership, find new users, and increase citations. With libraries reducing budgets for subscriptions amid greater availability of open access content, publishers are under greater pressure than ever to grow their audience and ensure that their content is finding its way to users who will value it the most. This paper includes research and case studies that demonstrate how TrendMD is helping publishers achieve their goals through enhanced discovery.

Keywords: Discovery, bibliometrics, article usage, academic journals, marketing, reach, impact, trendmd, knowledge dissemination, research, related article

1. About TrendMD

TrendMD is an article discovery platform that generates related articles on journals, blogs, and other sites that academics, doctors, and researchers use in their day-to-day work. Scholarly publishers, authors, and funders use TrendMD to increase site traffic, reach their target audiences, and drive article impact. TrendMD uses AI technology to generate recommendations through a combination of semantic enrichment, collaborative filtering (e.g. ‘users that read X, also clicked on Y’), and personalization (what users have read in the past).

TrendMD was founded in 2014 by a group of professionals with backgrounds in academic research, scholarly publishing, and digital technology. Early funding and nurturing came from Y Combinator - the startup accelerator that incubated Reddit, Dropbox, and Airbnb. The company has twenty-one employees, with management in Toronto, Canada and California, U.S.A. The company’s services are used by over three hundred publishers, including scholarly societies and scholarly commercial publishers. More than one hundred and fifty million readers per month view TrendMD recommendations on nearly five thousand sites.

*E-mail: bert.carelli@trendmd.com.

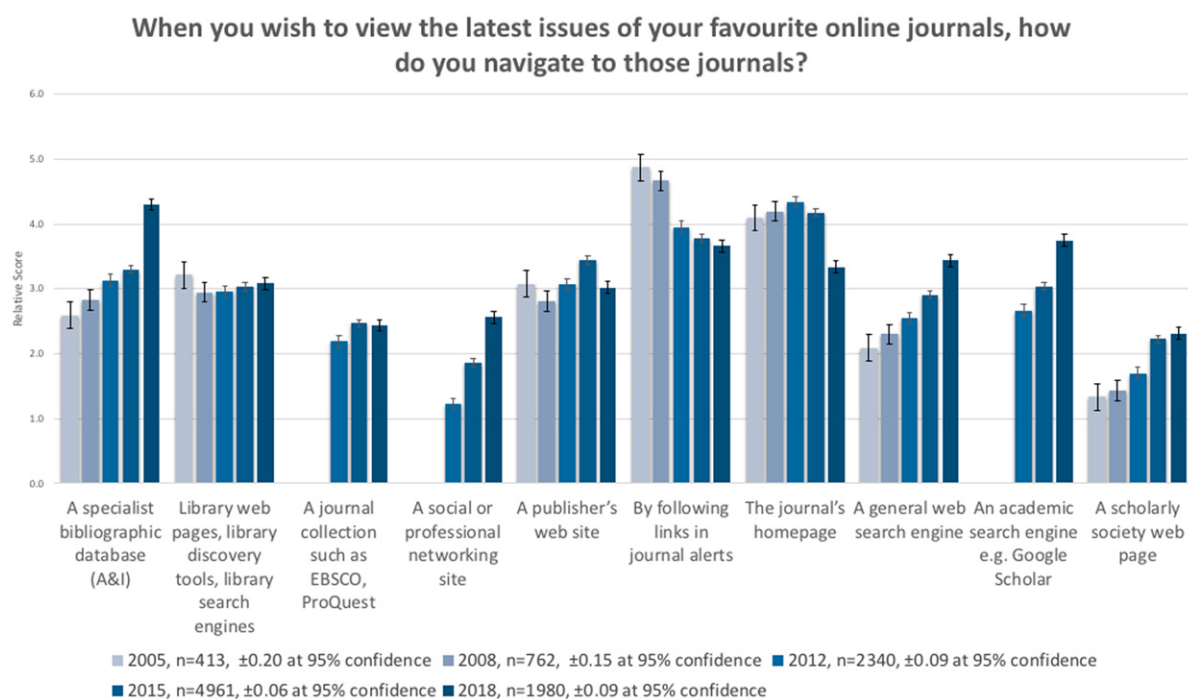


Fig 1. From the report *How Readers Discover Content in Scholarly Publications* [2], used by permission.

2. Background

TrendMD is focused on a familiar problem in scholarly publishing: with over 2.5 million scholarly articles published each year - more than eight thousand each day - the competition for discovery is only getting tougher. The problem poses a challenge to readers, as well as publishers. TrendMD's business is focused on helping match readers to the correct articles, with personalized marketing solutions tailored for scholarly publishing organizations.

Traditionally, in order to ensure that their research would be noticed, authors would have to rely on getting their work published in prominent journals, while publishers leveraged a journal's reputation, its Impact Factor, and their relationships with the library community to promote the utilization of their content. With the ever-increasing dominance of electronic distribution, abstracting and indexing (A&I) services, and search engines, users are more likely to discover individual articles independent of the journal or issue. The report, *How Readers Discover Content in Scholarly Publications* [1], which has been produced every three years since 2005 by Simon Inger and Tracy Gardner, demonstrates that discovery has evolved to coincide with the increasing dominance of electronic content distribution. A&I databases and search engines have become the first place users look for articles on a specific subject - rising in popularity with each successive survey - while other methods such as Table of Contents alerts, publisher website searches, and library search engines, have either declined or stayed consistent over the last fifteen years.

The same survey also recorded a change in the way that researchers use search results. Typically, the use of a search results page doesn't end with the user finding one specific article; users are also likely to browse the other articles that surface in a search. Researchers are showing an increasing enthusiasm

for serendipity, as the authors concluded: “Browse and search are becoming more blurred and more similar.” This is consistent with the feedback we have received from individual researchers. During the 2018 webinar held by the Professional and Scholarly Publishing Division of the Association of American Publishers (AAP/PSP) “The Changing Discovery Landscape - Part 1,” David McCandlish, an Assistant Professor at the Simons Center for Quantitative Biology’s Cold Spring Harbor Laboratory, discussed how he discovers new content that is relevant and useful to his work. He stated that he is less reliant on the library, and much more dependent on interactive tools available on the Web, including personalized recommendations on article pages [3].

For publishers, this has meant that promoting scholarly content is increasingly similar to the way that commercial publishers and other retailers promote their products. When purchasing nearly any product online, users today are accustomed to finding recommendations for additional products related to their purchase or browsing history. In the words of Professor McCandlish, “Netflix and Amazon have that magic factor of enabling discovery of content you had no idea existed, but is valuable and interesting.” Similar services for the distribution of content in the consumer web, including Outbrain (see: www.outbrain.com) and Taboola (see: www.taboola.com) generate the “From the Web” and “You may Like” recommendations seen alongside content on many popular websites such as CNN or The *New York Times*. Therefore, it is not surprising that when users were asked what features of a publisher’s website they use most often, the related articles functionality showed consistently increasing popularity over successive surveys. This is in stark contrast to other website features, such as publisher-produced news, site search, saved search, and alerting, which have declined.

3. How it Works

TrendMD uses an AI technique called Collaborative Filtering [4] to generate recommendations, which is similar to the way that services such as Amazon suggest additional products to users when they are browsing or purchasing. By looking at the behavior of similar users, TrendMD can anticipate which articles readers will find interesting and create better recommendations than those created by text analysis alone. For example, in Amazon’s case, a user who has purchased a teapot will not be offered additional teapots. Instead, they will be recommended products that previous buyers of teapots have bought in the past; this could include items such as recipe books, teacups, or cakes. In TrendMD’s case, click behavior is used to rank the potential relevance of related articles by using the data on what other users have clicked on in the past to predict what will most interest the current reader. For example, a medical practitioner reading an article on smoking hazards in the *British Medical Journal (BMJ)* may see recommended articles on cancer treatment, while a public health policy specialist may see articles on secondhand smoke, and a biochemical researcher may see articles on tumor formation.

The following illustration shows an example of how TrendMD’s recommendations algorithm works. Initially, TrendMD uses the standard PubMed similar articles algorithm [6], which matches keywords and context to identify similarities between articles. The system then tracks what types of articles readers click on (e.g. articles in medicine, physics, economics, computer science, etc.). After that, the system makes predictions about recommendations the user hasn’t yet clicked on. These predictions are built upon the existing data of other users who share similar data with the active user. In this simple example, the system has made a prediction about which users would have a similar interest in a medical article.

In a 2014 study [8] the *Journal of Medical Internet Research (JMIR)* (see: <https://www.jmir.org>) compared TrendMD’s use of recommendations generated by using collaborative filtering and user behavior

What features do you find useful in publishers' journal web sites? 2005 to 2018 comparison.

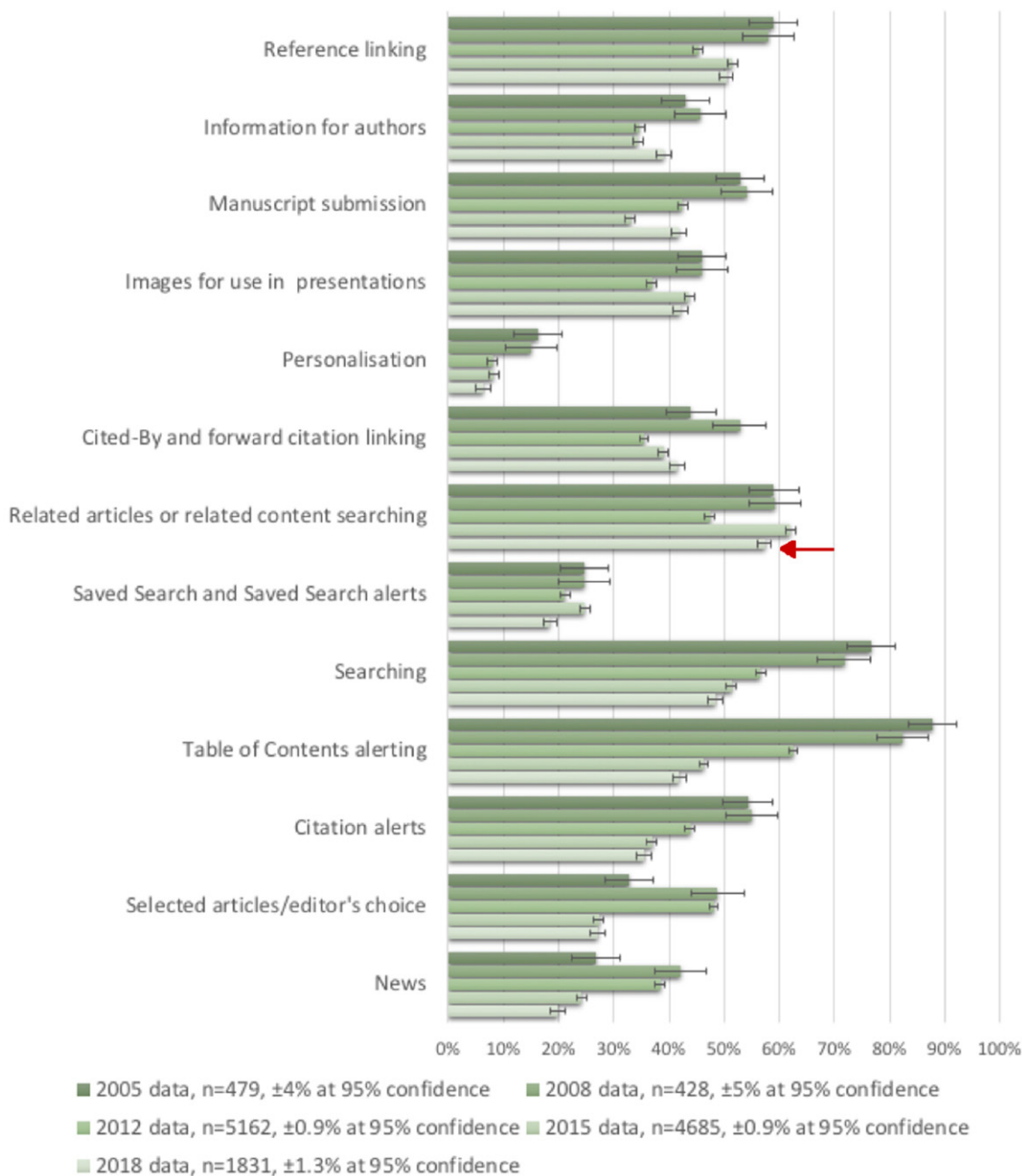


Fig 2. Gardner and Inger, reprinted with permission (See [1]).

with recommendations generated by semantic similarity alone. In a six-week A/B test, researchers measured aggregate click-through rate (CTR) for all article recommendations displayed by the widget. Equal

AAAS Become a Member Log In ScienceMag.org Search

Science Contents News Careers Journals

SHARE **REPORT**

Evaluation of an antibody to $\alpha_4\beta_7$ in the control of SIVmac239-*nef-stop* infection

M. Di Mascio^{1,*}, J. D. Lifson², S. Srinivasula³, I. Kim⁴, P. DeGrange⁵, B. F. Keele², A. J. Belli⁶, K. A. Reimann⁶, Y. Wang⁸, M. ...
 * See all authors and affiliations

Science 06 Sep 2019;
 Vol. 365, Issue 6457, pp. 1025-1029
 DOI: 10.1126/science.aav6695

Article Figures & Data Info & Metrics eLetters PDF

An antibody is not the antidote

An HIV therapeutic that would give long-term remission without sustained antiretroviral therapy (ART) is a long-term goal. Byrreddy *et al.* [*Science* **354**, 197 (2016)] reported that treating simian immunodeficiency virus (SIV)-positive macaques with an antibody against integrin $\alpha_4\beta_7$ during and after ART results in sustained virologic control after stopping all treatment. Three studies in this issue question the reproducibility of that result. Di Mascio *et al.* sequenced the virus used in the 2016 study and found that it was a variant with a stop codon in the *nef* gene rather than a wild-type virus. Abbink *et al.* used the same antibody for $\alpha_4\beta_7$ as before but tested control of a more commonly used pathogenic virus. Iwamoto *et al.* used the same *nef-stop* virus as in the earlier paper but combined the antibody against the integrin with an antibody against the SIV envelope glycoprotein, which also blocks viral binding of the integrin. None of these three new studies found that treating with the antibody had any effect on virologic control after stopping ART treatment.

Science, this issue p. **1025**, p. **1029**, p. **1033**

Abstract

Treatment of SIV-infected rhesus macaques with short-term antiretroviral therapy (ART) and partially overlapping infusions of antibody to integrin $\alpha_4\beta_7$ was reported to induce durable posttreatment viral suppression. In an attempt to replicate those observations, we treated macaques infected with the same virus and with the same ART and monoclonal antibody (mAb) regimens (anti- $\alpha_4\beta_7$ versus control mAb). Sequencing demonstrated that the virus used was actually SIVmac239-*nef-stop*, not wild-type SIVmac239. A positive correlation was found at 2 weeks after infection between the frequency of repair of attenuated *Nef-STOP* virus to pathogenic *Nef-OPEN* and plasma SIV RNA levels. Levels of plasma viremia before the first antibody infusion and preinfection levels of $\alpha_4\beta_7^{\text{hi}}$ CD4⁺ T cells, but not treatment with antibody to $\alpha_4\beta_7$, correlated with levels of viral replication upon discontinuation of all treatments. Follow-up plasma viremia, peripheral blood CD4⁺ T cell counts, and lymph node and rectal tissue viral load were not significantly different between anti- $\alpha_4\beta_7$ and control mAb groups.

<http://www.sciencemag.org/about/science-licenses-journal-article-reuse>

This is an article distributed under the terms of the [Science Journals Default License](#).
[View Full Text](#)

Recommended articles from TrendMD

<p>Blocking $\alpha_4\beta_7$ integrin binding to SIV does not improve virologic control Nami Iwamoto <i>et al.</i>, <i>Science</i>, 2019</p> <p>Lack of therapeutic efficacy of an antibody to $\alpha_4\beta_7$ in SIVmac251-infected rhesus macaques Peter Abbink <i>et al.</i>, <i>Science</i>, 2019</p> <p>An open-label phase 1 clinical trial of the anti-$\alpha_4\beta_7$ monoclonal antibody vedolizumab in HIV-infected individuals Michael C. Sneller <i>et al.</i>, <i>Sci Transl Med</i>, 2019</p> <p>Sustained virologic control in SIV+ macaques after antiretroviral and $\alpha_4\beta_7$ antibody therapy Siddappa N. Byrreddy <i>et al.</i>, <i>Science</i>, 2016</p> <p>Virion incorporation of integrin $\alpha_4\beta_7$ facilitates HIV-1 infection and intestinal homing Christina Guzzo <i>et al.</i>, <i>Sci Immunol</i>, 2017</p>	<p>Early treatment of SIV+ macaques with an $\alpha_4\beta_7$ mAb alters virus distribution and preserves CD4+ T cells in later stages of infection ↗ P J Santangelo <i>et al.</i>, <i>Mucosal Immunology</i>, 2017</p> <p>A 30-year journey of trial and error towards a tolerogenic AIDS vaccine ↗ Jean-Marie Andrieu <i>et al.</i>, <i>Archives of Virology</i>, 2018</p> <p>MAdCAM costimulation through Integrin-$\alpha_4\beta_7$ promotes HIV replication ↗ Fatima Nawaz <i>et al.</i>, <i>Mucosal Immunology</i>, 2018</p> <p>Combination anti-PD-1 and antiretroviral therapy provides therapeutic benefit against SIV ↗ Geetha H. Mylvaganam <i>et al.</i>, <i>JCI Insight</i>, 2018</p> <p>Manipulating the Interferon Signaling Pathway: Implications for HIV Infection ↗ Krystelle Nganou-Makamdop <i>et al.</i>, <i>Virologica Sinica</i>, 2019</p>
--	---

Fig 3. Example of the TrendMD recommendations widget on *Science Magazine* [5], reprinted with permission from the journal.

Match

Fig 4. Adapted from Wikipedia contributors. (2019, June 13). “Collaborative filtering,” In *Wikipedia, The Free Encyclopedia* [7].

numbers of recommendation sets were generated through each of these methodologies and presented to the readers. The user interface was identical in all instances and all previous click data was deleted at the start of the experiment. In the graph below, the CTR for each methodology is depicted. As expected, the CTR was similar in the first week, before TrendMD had gathered the click data necessary to further personalize the semantic similar article filtering, which is used by both approaches. This is known as the “cold start” condition in studies of this type. The results show that by the second week, after data had begun to accumulate, the recommendations generated using collaborative filtering began to perform better, leading to a nearly three-fold higher CTR by the sixth week.

4. The TrendMD network

An important feature of TrendMD is that recommendations that are shown in the widget are both from a publisher’s own content as well as from an equal number of recommendations from external sites. A publisher’s first reaction might be “why would I want to present my readers with links to another publisher’s site - even to sites that I might see as competitive?” The reason is due to another innovation that TrendMD has introduced to the industry - the TrendMD credit system that gives publishers traffic back in return for those who click on external links.

Figure 5 shows a simplified look behind the scenes at how the TrendMD widget drives traffic and page views. The left column (or the top half of recommendations, if it’s a single-column display) directs the reader to additional articles from a publisher’s own content. These links can be from the same journal, or for publishers with multiple journals, a chance to cross-promote within their family of journals. The right column (or the lower half of recommendations in a single-column display) displays links to external sites and the links that appear here are facilitated by a credit exchange. Publishers earn credits when users click on those external links, while they spend credits when a reader clicks on one of their articles appearing on external sites. Links to a publisher’s content only appear on external sites in our network when there is a positive credit balance.

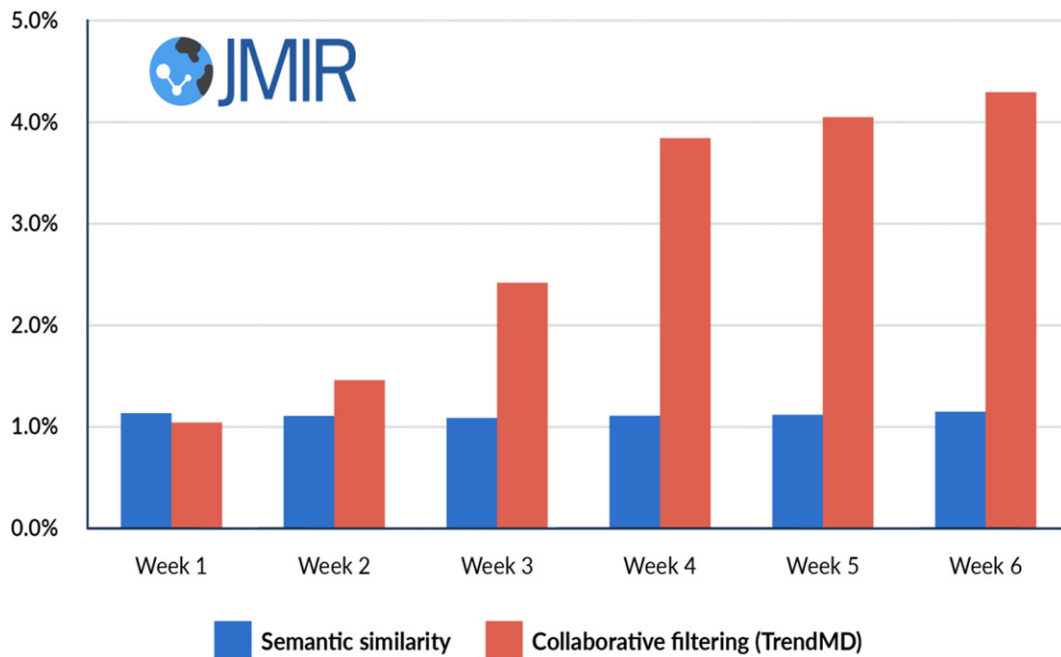


Fig 5. Illustration of results described in Kudlow et al., 2014.

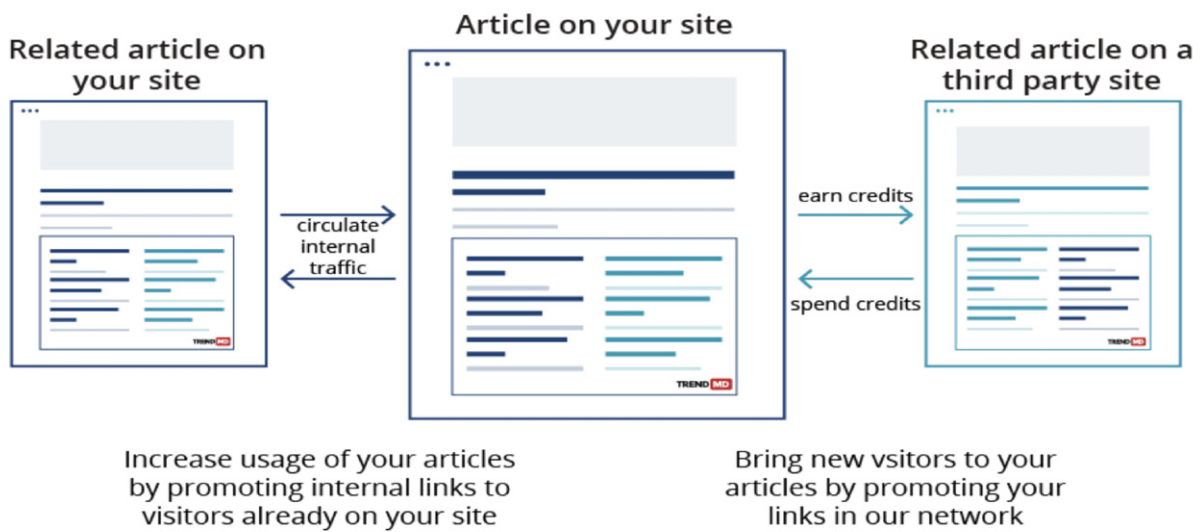


Fig 6. Illustration of the TrendMD credit system.

The publisher earns a half traffic credit for each click on an external link that takes place on their site. The publisher spends one credit for each reader TrendMD sends from an external site. This ability to earn credits means that publishers can get new readers just by installing the widget, at no cost. Most publishers start out using TrendMD under the free plan; in some cases publishers will choose to purchase additional credits to get more incoming readers and unlock additional features, but this is optional. As publishers

TrendMD unique widget impressions by country

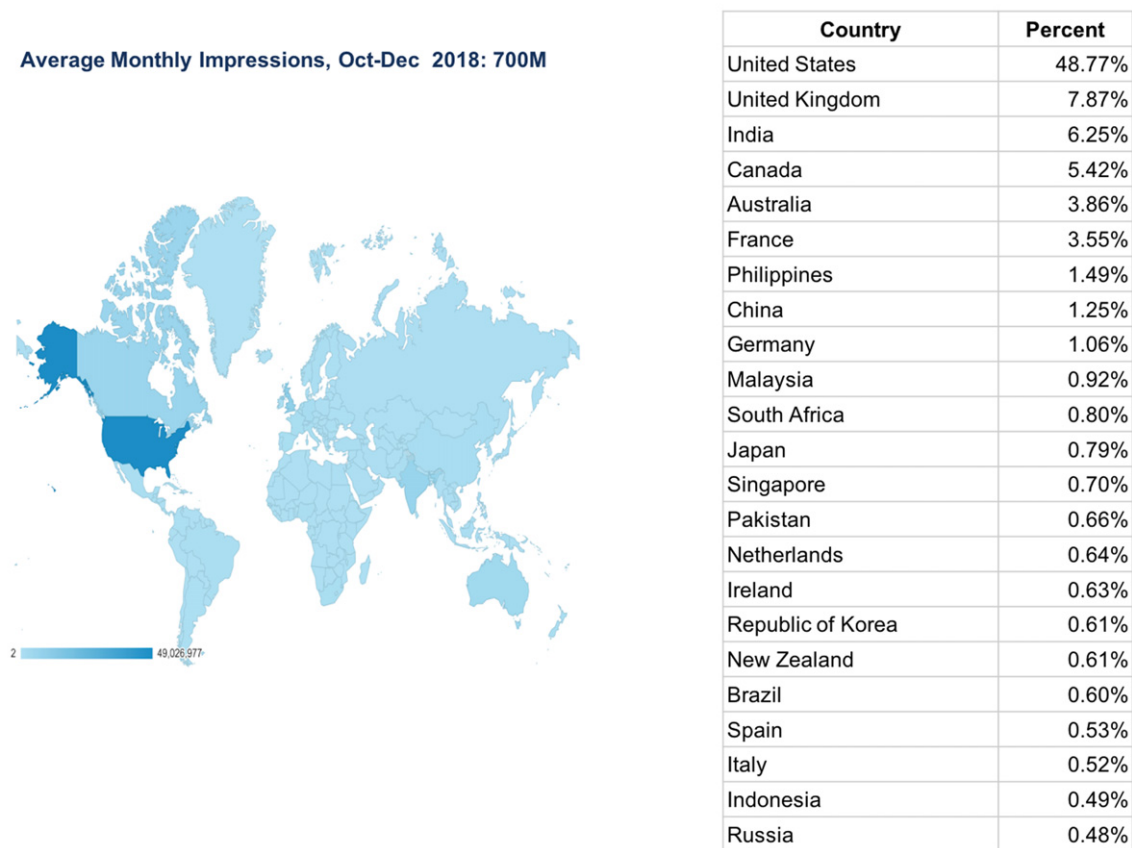


Fig 7. TrendMD global reach.

buy credits, all credits within the system thereby have a monetary value; the use of credits as fiat currency works very similarly to the way that blockchain systems create and exchange cryptocurrency.

5. Benefits to publishers

Increasing Discovery and Reach: A key reason for TrendMD's rapid growth in the scholarly market is our ability to help publishers maximize their growth by expanding their reach. TrendMD's reach, as illustrated in Fig. 6 from recent data, enables scholarly publishers to attract new readers and authors from all across the world.

Traffic Shaping: In addition to adding more traffic credits, the paid plans, known as Professional and Enterprise plans, unlock additional features that enable publishers to use their traffic credits strategically to "shape" their traffic, and thereby accomplish specific business goals. Some examples of traffic shaping include sponsoring the last two years' of content to focus on raising a journal's Impact Factor, raising awareness of a newly-launched journal, or aiming to increase author submissions. The following section highlights a few examples showing just how this can be done, and how effective these strategies have been.

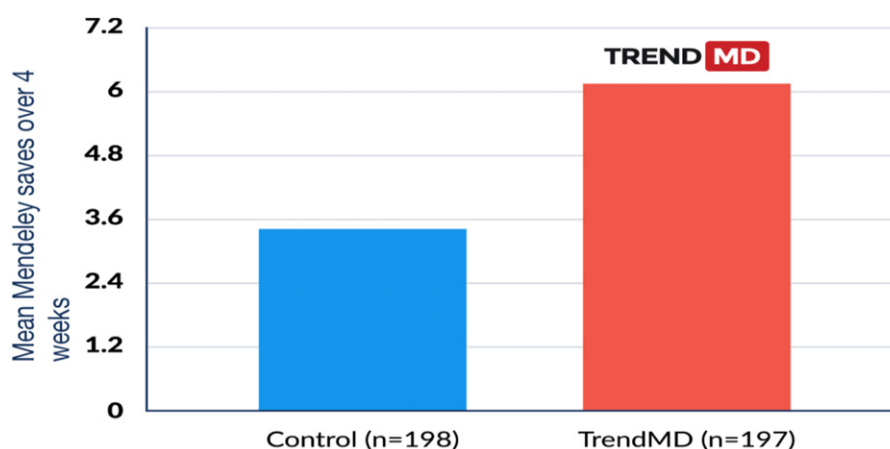


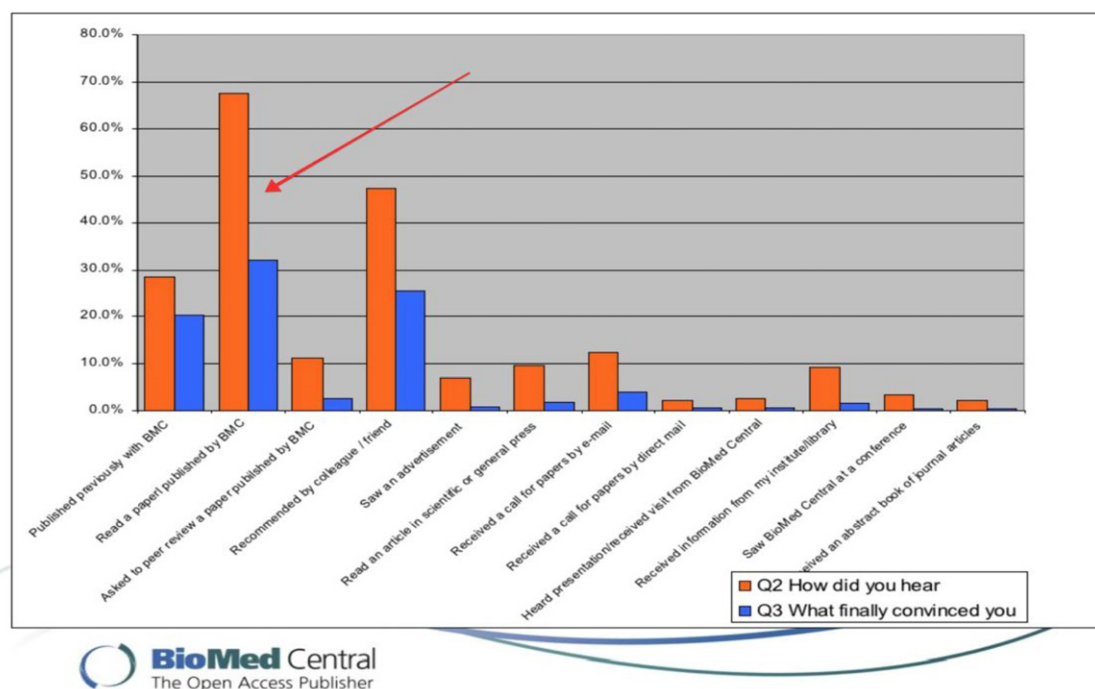
Fig 8. Mean Mendeley saves over 4 weeks: TrendMD versus control.

Increasing Impact: A 2017 study [9] in the journal *Scientometrics* showed that articles promoted through TrendMD were 77% more frequently saved to Mendeley by readers. This is important because unlike other measures of article usage, such as total page views or Altmetric scores, Mendeley saves have been shown to be a more accurate predictor of future citations (Priem et al. 2012 [10]; Lin and Fenner 2013 [11]; Zahedi et al., 2014 [12], 2015 [13]; Ebrahimi et al. 2016 [14]; Maflahi and Thelwall 2016 [15]; Thelwall and Wilson 2016 [16]; Li and Thelwall 2012 [17].)

Some interesting secondary effects also were revealed in the same study:

- Articles promoted through TrendMD received a 95% increase in total page views relative to the control group over the four-week trial.
- The TrendMD-promoted articles also had higher organic page views. One possible explanation for this is that discovery of articles via TrendMD leads to individuals visiting articles recommended via TrendMD more frequently. This could include readers coming back to articles independently (e.g. saving them as bookmarks on an internet browser and visiting them later), sharing articles with their colleagues over email, or spreading awareness via word of mouth.
- TrendMD visitors were more engaged when compared to the control group and additional sharing of articles took place.
- Site visitors who clicked on TrendMD links visited a greater number of pages per session than visitors referred through other channels.
- The TrendMD-recommended articles had lower bounce rates.

Increasing author submissions: The graph below shows the results of a survey done by *BioMed Central* (BMC) (see: <https://www.biomedcentral.com/>) of new authors to determine what led them to submit papers. Overwhelmingly, authors responded that the key influence was discovery of other papers published in their field by BMC.



To test using TrendMD to increase submissions, JMIR Publications (see: <https://www.jmirpublications.com/>), a leading eHealth publisher of twenty-four Open Access journals - promoted five hundred articles to targeted researchers in the TrendMD network, while at the same time running similar tests using Google AdWords and Facebook. TrendMD outperformed Google AdWords by more than two times and Facebook by more than five times in generating new article submissions.

Targeting Users: While the above traffic shaping strategies are all about the publisher choosing which subset of their content they want to promote - e.g. a journal, a range of volumes, a collection - another strategy is choosing what *readers* the publisher wants to target. Using IP recognition, TrendMD is able to detect a reader's region of the world, institutional affiliation, and in some cases even their profession, and use this data to filter which recommendations the user will see. Healthcare professionals in the U.S.A. are identified by a cookie in their browsers that includes NPI numbers, which are the unique identifiers of profession - doctor, nurse, etc. - and specialty.

Elsevier's *Practice Update* is a free publication for clinicians, but requires email signup to receive full text. The publisher launched a campaign targeted to clinicians, with the goal of increasing signups. Practice Update content was directed to approximately seventy-five hundred Healthcare professionals, with five hundred signing up to receive emails in the first month, a conversion rate of over 5%.

In another example of targeted promotion in the medical profession, Merck compared several channels, including TrendMD, to direct a targeted list of physicians to a microsite with information on a new drug. TrendMD's conversion rate of 3.7% was the highest of all channels used.

6. Conclusion

TrendMD is a prime example of the effective use of AI technology in scholarly publishing, helping researchers to discover new articles based on their interests, and helping publishers and authors promote their content. The TrendMD technology, combined with its growing network, have made TrendMD a highly cost-effective way for publishers to reach their most valued audiences.

About the Author

Bert Carelli, Director of Partnerships for TrendMD, helps scholarly publishers find new readers through increased discovery of contextually relevant content across a rapidly growing network of nearly five thousand scholarly sites. He is a veteran of the online information industry, having worked with many scholarly publishers previously as Senior Publication Manager at HighWire Press, and as head of business development for Access Innovations and for DeepDyve. Earlier in his career, he led content licensing teams for Dow Jones/Factiva and Dialog - at the time two of the three largest global professional information services. Throughout his career, Mr. Carelli has focused on building value for content providers by leveraging new technologies for finding, managing, and productizing information. He received his BA degree from Stanford University and earned an MBA degree from St. Mary's College of California. Email Address: bert.carelli@trendmd.com.

References

- [1] T. Gardner and S. Inger, *How Readers Discover Content in Scholarly Publications*. Renew Publishing Consultants, 2018, available at: <https://renewconsultants.com/wp-content/uploads/2018/08/How-Readers-Discover-Content-2018-Published-180903.pdf>.
- [2] T. Gardner and S. Inger, *op.cit.*
- [3] A. Stone, Do journal article recommendation features change reader behavior?, *The Scholarly Kitchen* (2018), Available at: <https://scholarlykitchen.sspnet.org/2018/04/09/guest-post-journal-article-recommendation-features-change-reader-behavior/>, accessed September 7, 2019.
- [4] Collaborative Filtering, *Wikipedia* https://en.wikipedia.org/wiki/Collaborative_filtering, accessed September 7, 2019.
- [5] See: <https://science.sciencemag.org/content/365/6457/1025.abstract>, accessed September 10, 2019.
- [6] See: https://www.nlm.nih.gov/bsd/disted/pubmedtutorial/020_190.html, accessed September 10, 2019.
- [7] See: https://en.wikipedia.org/w/index.php?title=Collaborative_filtering&oldid=901694091.
- [8] P. Kudrow, A. Rutledge and G. Eysenbach, TrendMD: Reaching, more targeted audiences by distributing scholarly content online, *Editor's Bulletin* **10**(1) (2014), 11–15. doi:10.1080/17521742.2015.1031965, available at: <https://www.tandfonline.com/doi/full/10.1080/17521742.2015.1031965>.
- [9] P. Kudlow, M. Cockerill and D. Toccalino, Online distribution channel increases article usage on mendeley: A randomized trial, *Scientometrics* **112**(3) (2017), 1537–1556. <https://doi.org/10.1007/s11192-017-2438-3>.
- [10] J. Priem, H.A. Piwowar and B.M. Hemminger, Altmetrics in the wild: Using social media to explore scholarly impact. 2012, arXiv preprint at: <https://arxiv.org/abs/1203.4745>, accessed September 7, 2019.
- [11] J. Lin and M. Fenner, Altmetrics in evolution: defining and redefining the ontology of article-level metrics, *Information Standards Quarterly* **25**(2): summer 2013, available at: https://www.niso.org/sites/default/files/stories/2017-08/IP_Lin_Fenner_PLOS_altmetrics_isqv25no2.pdf, accessed September 7, 2019.
- [12] Z. Zahedi, R. Costas and P. Wouters, How well developed are altmetrics? a cross-disciplinary analysis of the presence of “alternative metrics” in scientific publications, *Scientometrics* **101**(2) (2014), 1491–1513. doi:10.1007/s11192-014-1264-0.
- [13] Z. Zahedi, R. Costas and P. Wouters, Do Mendeley Readership Counts Help to Filter Highly Cited WoS Publications Better than Average Citation Impact of Journals (JCS)? July 8, 2015, available at: <https://arxiv.org/abs/1507.02093>, accessed September 7, 2019.

- [14] S. Ebrahimi, J. Mehrad, F. Setareh and M. Hosseinchari, Path analysis of the relationship between visibility and citation: the mediating roles of save, discussion, and recommendation metrics, *Scientometrics* **109**(3) (2016), 1497–1510. doi:[10.1007/s11192-016-2130-z](https://doi.org/10.1007/s11192-016-2130-z).
- [15] N. Maflahi and M. Thelwall, When are readership counts as useful as citation counts? scopus versus mendeley for LIS journals, *Journal of the Association for Information Science and Technology* **67**: (2015), 191–199. doi:[10.1002/asi.23369](https://doi.org/10.1002/asi.23369), preprint available at: <https://core.ac.uk/download/pdf/77012075.pdf>, accessed September 7, 2019.
- [16] M. Thelwall and P. Wilson, Mendeley readership altmetrics for medical articles: an analysis of 45 fields, *Journal of the Association for Information Science and Technology* **67**: 1962–1972. [10.1002/asi.23501](https://doi.org/10.1002/asi.23501), May 2015, preprint available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.706.4019&rep=rep1&type=pdf>, accessed September 7, 2019.
- [17] X. Li and M. Thelwall, F1000, mendeley and traditional bibliometric indicators, in: *Proceedings of the 17th International Conference on Science and Technology Indicators*, Vol. 2, pp. 451–551. See Google Scholar at: <https://scholar.google.com/scholar?q=Li%2C%20X.%2C%20%26%20Thelwall%2C%20M.%20%282012%29.%20F1000%2C%20Mendeley%20and%20traditional%20bibliometric%20indicators.%20In%20Proceedings%20of%20the%2017th%20international%20conference%20on%2522>, accessed September 7, 2019.