



David Pennock Receives Two Test of Time Awards

[December 2020] Two of DIMACS Director David Pennock's papers were honored with Test of Time awards this year. Both papers, published in the early to mid-2000's, helped shape today's online search experience. The first made progress in web search algorithms and the second in monetizing search through advertisement auctions.

Pennock's paper with [Kushal Dave](#) (Scroll.com) and [Steve Lawrence](#), [Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews](#), received the [2020 Seoul Test of Time Award](#) at The Web Conference in April. The paper



was first presented at the Twelfth International World Wide Web Conference in 2003—a time that was still early in the development of online commerce. Pennock and his collaborators were among the first to recognize the trove of information contained in online product reviews and the potential value of distilling those reviews into simple classifications accessible to online consumers. To do this, they combined a variety of approaches in information retrieval and sentiment analysis to obtain what they described as “fairly good results for the review classification task”. Despite this modest claim, their groundbreaking paper now has more than 2,800 citations and has become foundational research in the area of opinion mining on the web.



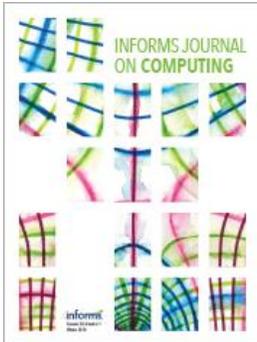
In presenting the award, Dame Wendy Hall, Conference Committee Chair, said: “With the huge progress of ecommerce in recent years, the value of product reviews and more specifically the sentiments expressed by customers in their reviews is now obvious to all. This paper was the first to demonstrate the viability of such an approach and is a very worthy winner of the Seoul Test of Time award 2020.”

In accepting the award, the authors reflected on why this work has stood the test of time:

“There are three things that feel interesting about this paper in hindsight. First, the marriage of user interface and machine learning. By looking at the attribute-level instead of the whole document level, what you're aiming for with classification becomes more clear. And when you miss, it's easier to forgive machine learning for being imperfect since it's still delivering a whole new level of insight to the user. Second, we learned as much about what didn't work as what does. We tried a lot, a lot...a lot of variations, and we weren't afraid to try simple or counterintuitive things based on insight from the raw data—patterns starting with the word ‘the’, turning off stemming, discarding complex algorithms in favor of simple [information retrieval] techniques. Steve often repeated ‘perfect is the enemy of good’, and that philosophy is at the heart of today's applied machine learning—build it and see what it works, even if you don't have a theory yet. Third, never underestimate the power of giving your paper a funny title.”

The Seoul Test of Time Award has been awarded annually since 2015 to the authors of a paper presented at a previous Web Conference. The first Award was presented to Google founders Sergey Brin and Larry Page for their world-changing paper, [The Anatomy of a Large-Scale Hypertextual Web Search Engine](#), at the Seventh World Wide Web Conference in 1998. Other winners include the authors of the famous “[bow-tie](#)” [web structure paper](#) and a [classic recommender systems paper](#).

Also in April 2020, Pennock’s paper with [Juan Feng](#) (City University of Hong Kong) and [Hemant Bhargava](#) (University of California, Davis), [Implementing Sponsored Search in Web Search Engines: Computational Evaluation of Alternative Mechanisms](#), received the INFORMS Journal on Computing Test of Time Award. The [award was announced](#) in the Spring 2020 issue of the journal.



Today, Google earns over \$100 billion per year in revenue—more than 80% from advertising—but at the time, researchers like Pennock and his collaborators puzzled over how free search engines could make money. The answer came from Overture, a startup company that was ultimately acquired by Yahoo!. In their [comments upon receiving the award](#), the authors reflected on this connection saying, “*Overture pioneered text-based ads that were priced, sold, and implemented much differently than the “banners” in newspapers, magazines, and early websites, and Google perfected the practice. Instead of negotiating bulk sales, Google and Overture allocated ad slots one at a time using large-scale, continuous, real-time auctions.*

The setting allowed instantaneous measurement of the relevance and performance of ads, enabling search engines to collect money only when users actually clicked—about 40 cents per click on average in 2003. Our paper focused on a few key issues needed to make auction-based, pay-per-click advertising work well.” The authors elaborated on these issues in their comments for the [award announcement](#). A more detailed reflection appears in [Hemant Bhargava’s blog post](#).

The test of time award is given annually to a paper published in the INFORMS Journal on Computing (IJOR) within a rolling five-year window. The paper of Pennock and his colleagues was selected from among all papers published in IJOR between 2003-2007 and was only the second paper to receive the award. The award citation (contained in the announcement) describes contributions of the paper this way: “*This paper from Feng, Bhargava, and Pennock is among the earliest publications to study mechanisms for the allocation of ads by search engine providers. The paper reports extensive computational simulations, allowing the authors to compare expected revenue obtained from a variety of ad placement strategies. The study provides insights into the impact of attention decay and other key factors. This highly cited paper continues to be an important reference for the community of researchers studying this rapidly growing sector of the advertising market. The work is a very nice example of the reach of operations research techniques into the new economy.*”

References:

Kushal Dave, Steve Lawrence, and David M. Pennock, Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews, *Proceedings of the Twelfth International World Wide Web Conference*, May 2003, Pages 519–528.

Juan Feng, Hemant K. Bhargava, and David M. Pennock, Implementing Sponsored Search in Web Search Engines: Computational Evaluation of Alternative Mechanisms, *INFORMS Journal on Computing*, Vol. 19, No. 1, Feb 2007, Pages 137–148.

Related Links:

- 2020 Seoul Test of Time: <https://www.iw3c2.org/ToT/PressRelease-ToT-Award-20200421.pdf>
- IJOR award announcement: <https://pubsonline.informs.org/doi/full/10.1287/ijoc.2020.0968>
- Hemant Bhargava’s blog post: <https://gsm.ucdavis.edu/blog/look-back-how-web-search-ads-emerged-transformed-business>