Voice of Evidence

Voice of Evidence

A Look Back

Forrest Shull, Tore Dybå, Helen Sharp, and Rafael Prikladnicki

An underlying assumption of the articles that have appeared in Voice of Evidence (VoE) is that software researchers and practitioners can benefit from accumulated evidence and experience about specific software technologies. This evidence is usually expressed as quantitative and qualitative data from multiple sources. VoE’s challenge has always been, can we use this body of information to extract meaningful and generalizable lessons for professionals?

VoE has now appeared in *IEEE Software* for a decade and has been overseen by four editors. All of us editors have shared the assumption (bias?) that good research should create knowledge that will be of use to the professionals who create software. The challenge in many cases has been to translate precise, focused research articles into more robust takeaways that connect with the concerns of practitioners, mostly software developers or managers. This must be done without overgeneralizing. That is, we need to understand where the research relies too much on laboratory experiments and toy problems, detached from a more practical context, or too much on a specific context. Another part of this challenge lies in understanding where gaps in the research are—issues of concern to practitioners that just haven’t been examined in the literature.

So, in this article celebrating the 200th issue of *IEEE Software*, we examine whether we can take our own medicine and use data to determine where VoE has successfully navigated the constraints we just mentioned. Relying on available data can be hard; it’s open to multiple interpretations and doesn’t always tell us what we’d like to be true. Hopefully, though, it tells us what really is. So let’s see what it tell us about where VoE has been the most successful at reaching an audience.

# How Do You Measure Impact?

As in many VoE articles, the thing we want to measure—in this case, impact—is hard to get at directly. To assess impact, the best we can do is to rely on metrics that *IEEE Software* and other scientific publications often use: the number of citations and downloads. The number of citations is a reasonable measure of impact in the scientific community: how often has something we’ve published been cited as evidence in another scientific article?

However, citations give no sense of how the articles might have influenced practice and not even how many practitioners might have read the articles. To understand this, we must rely on a more problematic measure: the number of downloads from the IEEE Computer Society Digital Library (CSDL), on the theory that this is a reasonable proxy for the number of readers of an article.

But this metric doesn’t measure real impact—that is, that these articles have changed the practice. In addition, it doesn’t count downloads from other sources (for example, IEEE Xplore) or the number of practitioners who read the articles as part of their subscription to the magazine. It also doesn’t exclude researchers, who after all also might want electronic copies of specific articles. Nevertheless, problematic as this metric is, it remains our best proxy for practitioner interest.

Of course, both citations and downloads are biased in favor of older articles, which simply have had more time to accumulate use.

Tables 1 and 2 show the top 10 VoE articles ranked by the number of citations and downloads. We collected the citation data from Google Scholar and the download data from the CSDL. The citation data is cumulative up to April 2017; the download data is cumulative up to the end of 2016.

|  |
| --- |
| Table 1. The Voice of Evidence articles with the most citations (as a proxy for researcher interest). |
| Rank | Article title | Year | No. of citations |
| 1 | “What Do We Know about Agile Software Development?” | 2009 | 160 |
| 2 | “Are Two Heads Better than One? On the Effectiveness of Pair Programming” | 2007 | 127 |
| 3 | “To Game or Not to Game?” | 2009 | 76 |
| 4 | “Understanding the Customer: What Do We Know about Requirements Elicitation?” | 2008 | 55 |
| 5 | “Improving Evidence about Software Technologies: A Look at Model-Based Testing” | 2008 | 54 |
| 6 | “What Do We Know about Test-Driven Development?” | 2010 | 44 |
| 7 | “Software Metrics: Progress after 25 Years?” | 2008 | 44 |
| 8 | “A Whisper of Evidence in Global Software Engineering” | 2011 | 43 |
| 9 | “What Do We Know about Knowledge Management? Practical Implications for Software Engineering” | 2009 | 50 |
| 10 | “Patterns in Effective Distributed Software Development” | 2010 | 35 |

|  |
| --- |
| Table 2. The Voice of Evidence articles with the most downloads (as a proxy for practitioner interest). |
| Rank | Article title | Year | No. of downloads |
| 1 | “What Do We Know about Agile Software Development?” | 2009 | 4,188 |
| 2 | “Understanding the Customer: What Do We Know about Requirements Elicitation?” | 2008 | 1,371 |
| 3 | “Software Project Management: Learning from Our Mistakes” | 2015 | 1,258 |
| 4 | “What Do We Know about Test-Driven Development?” | 2010 | 1,217 |
| 5 | “Five Facts on the Adoption of Open Source Software” | 2011 | 1,094 |
| 6 | “Creating Software Process Capability/Maturity Models” | 2010 | 1,029 |
| 7 | “Software Product Management” | 2014 | 1,003 |
| 8 | “Design Patterns: Magic or Myth?” | 2013 | 961 |
| 9 | “Does Involving Users in Software Development Really Influence System Success?” | 2013 | 948 |
| 10 | “A Whisper of Evidence in Global Software Engineering” | 2011 | 933 |

# The Most Popular Topics

Reviewing the two tables, we notice several things. First, four articles are on both lists. The two lists include articles on many different topics, which might reflect that these two audiences (practitioners and researchers) are interested in different things. Topics of interest to both communities reflect important recent trends in practice such as agile development (including test-driven development) and globally distributed software engineering.

On the research side (the number of citations), we see topics, such as educational gaming, that seem to have been of significant interest to researchers but haven’t (yet) attracted high attention from practitioners. Topics such as metrics and knowledge management tend to be of interest to researchers but generate many high-level principles that aren’t always easy to apply directly. But some topics, such as pair programming, have made a difference in practice.

The articles in the practitioner-oriented (the number of downloads) list are related to management (of projects and products), which isn’t usually a direct focus of software engineering research. These articles reframed research results into a format that could be more actionable for software managers. Another topic, the creation of maturity models, has again historically not been of significant interest to researchers but often proves useful in practice as a way to understand whether practices are being applied.

In short, at least on the basis of our own biases and not backed up by rigorous analysis, it’s plausible that the two lists indeed reflect the interests of different communities.

The articles in the citations table tend to be older than the ones in the downloads table. There are several possible reasons for this, which we can’t distinguish among. Perhaps the articles published during this time frame were of more interest to researchers for one reason or another. Or, the results might just reflect a special dynamic that gave an advantage to older articles. For example, maybe articles accumulate citations slowly and steadily, as researchers see articles cited in other papers and find them interesting to cite in future work.

On the basis of our quick-and-dirty analysis, we can see some trends that help us understand what might be useful to cover for different communities going forward. **//1. Rafael: Regarding the previous sentence, to what trends are you referring?// We are refereeing to Tables 1 and 2 and the popular topics.**

We wrote this article over the weekend in which the March for Science occurred in cities around the world. For those of us who participated, this event was inspiring—a chance to reflect on the sense of excitement and social responsibility of many people who work in the sciences and who see the importance of rigorous observation and analysis as a way to improve the human experience. A key rallying cry was “evidence matters.” We, as past and current VoE editors, look forward to many more years of the hard but rewarding work of translating scientific evidence and real-world experience into takeaways that stand to improve software practice.

This look back would be remiss if we didn’t convey our thanks to Hakan Erdogmus, who was editor in chief from 2007 to 2010. Hakan first pushed the idea for this department and supported it during its early years. It’s clear that VoE wouldn’t have happened without him.

Finally, we invite you to share your experiences on how a VoE article changed your practice. Please send your reminiscences to Rafael. We’d love to write about them.

**//2. Rafael: I added the following short bios; are they accurate?// Yes**

Forrest Shull is the assistant director for empirical research at Carnegie Mellon University’s Software Engineering Institute. He’s editor in chief emeritus of IEEE Software. Contact him at fshull@computer.org.

Tore Dybå is a Chief Scientist at SINTEF Digital. He’s on IEEE Transactions on Software Engineering’s Review Board. Contact him at tore.dyba@sintef.no**.**

Helen Sharp is a professor of software engineering at The Open University, Milton Keynes. She’s on IEEE Software’s advisory board. Contact her at helen.sharp@open.ac.uk.

Rafael Prikladnicki is an associate professor in the Computer Science School and director of the Science and Technology Park (Tecnopuc) at the Pontifical Catholic University of Rio Grande do Sul, where he also leads the MuNDDoS research group. He’s on IEEE Software’s editorial board and chairs the magazine’s advisory board. Contact him at rafaelp@pucrs.br.

Section Title: Voice of Evidence

Article Title: Voice of Evidence: A Look Back

Abstract: To celebrate the 200th issue of *IEEE Software*, Voice of Evidence uses data to examine whether it has helped bridge the gap between research and practice by extracting actionable lessons from the body of research.

Keywords: Voice of Evidence, software development, software engineering

Content Type: orig-research