INFORMS Journal on Data Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Chronicles of a New Journal: Reflections by the Inaugural Editor-in-Chief

Galit Shmueli

To cite this article:
Galit Shmueli (2024) Chronicles of a New Journal: Reflections by the Inaugural Editor-in-Chief. INFORMS Journal on Data Science 3(1):1-5. https://doi.org/10.1287/ijds.2024.editorial.v3.n1

https://pubsonline.informs.org/journal/ijds

INFORMS JOURNAL ON DATA SCIENCE
Vol. 3, No. 1, April–June 2024, pp. 1–5
ISSN 2694-4022 (print), ISSN 2694-4030 (online)

Chronicles of a New Journal: Reflections by the Inaugural Editor-in-Chief

Galit Shmueli[a]

[a] Institute of Service Science, College of Technology Management, National Tsing Hua University, Hsinchu 30013, Taiwan
Contact: galit.shmueli@iss.nthu.edu.tw, https://orcid.org/0000-0002-0820-0301 (GS)

https://doi.org/10.1287/ijds.2024.editorial.v3.n1

Four years, four seasons
Crashing waves, slow clouds above
A raft, ship, or plane?

As my term as inaugural editor-in-chief (EIC) of *INFORMS Journal on Data Science* (*IJDS*) comes to a close, I reflect on our efforts, experiences, and achievements in these four years.[1] This journey through a newly evolving land was full of discovery: people, knowledge, landscape, process, and weather of all sorts. There were tabula rasa moments that called for innovation as well as well-trodden paths that provided structure and borders. How does one steer in such a world, and where did it take the journal?

Because there were no open seats at "EIC school" nor in "inaugural EIC" workshops, I had only my own experience, knowledge, and skills to set goals and then navigate through opportunities, expectations, requirements, and challenges. I onboarded in May 2020, as COVID was shocking the world and as we all struggled to manage fear, uncertainty, pain, and exhaustion in ourselves, families, friends, students, and communities. Starting a new journal in such a difficult time required extraordinary efforts from everyone, and for that, I am grateful to our inaugural team of senior editors (SEs),[2] INFORMS Vice President of Publications Cole Smith, and Managing Editor (ME) Chris Asher.

In this editorial, I will share some of the efforts and approaches we have taken at *IJDS*, challenges we have encountered, the ecosystem we designed, and successes we have achieved. I use "we" because it definitely takes a village to raise a journal.

The What and Who of *IJDS*

The first task was shaping the "personality" of the journal. This required careful contemplation and important decisions on a variety of sociotechnical issues. Given the goal of creating a top-tier journal for data science research, the first two critical steps were determining the types of papers of interest ("what") and relatedly, the journal's communities ("who").

"What"

From my own experience as a researcher and based on many conversations and discussions with scholars in different INFORMS data science societies and communities (including artificial intelligence (AI), data mining (DM), quality statistics and reliability (QSR), information systems, and simulation (ISS)) and beyond, a clear need emerged for an outlet for data science *methodological* work motivated by decision-making or real-world applications.[3] Thus, the first distinguishing feature of *IJDS* emerged, setting it apart from methodology-focused journals in statistics, operations research, machine learning, etc. and from domain-focused journals in management and engineering. Methodology-developing researchers (including myself) have continuously encountered the "in-limbo effect," where domain journals do not appreciate a methodologically focused paper (typically resulting in delegating the heart of the work to an appendix), whereas pure methodology journals care little about the importance of decision-making motivations and implications. Second, given the prevalence of machine learning and other computational approaches and techniques in contemporary data science and the fast pace of developments in these areas, methodological research of this type often relies on empirical validation rather than mathematical proofs. For example, the features of a new proposed algorithm might be analyzed, demonstrated, and validated using a variety of simulated data and real data sets, showcasing its practical decision-making/real-world usefulness. Lastly, data science researchers in business schools and industrial engineering departments share some similarities in terms of writing styles and coding practices; manuscripts are typically 20–30 pages long (as opposed to shorter computer science conference papers or much longer management-type manuscripts) and have a typical level of technical detail for describing the

method, data, method validation, and results. They typically use open-source software, such as LaTeX, R, and Python, wanting others to adopt their methods. Hence, we designed *IJDS* to accept submissions of this form and type to help authors write in their "native language" and in a way that is easily readable and usable by other data science researchers.

"Who"

IJDS was conceived thanks to efforts from multiple INFORMS societies and communities that value data science research. It was, therefore, clear who the initial *IJDS* audience was. Luckily, I was familiar with both the management and engineering research communities (my PhD work was in industrial statistics, and my later research moved to machine learning in healthcare, information systems, and marketing). When I approached and invited senior editors from these various communities, it seemed we all had the same goals and motivation in mind:[4] creating a new top-tier journal for our data science communities. With this excellent team of inaugural SEs, we started setting the process in motion. Looking through our monthly editorial board meeting minutes from July 2020 until the journal launch in November 2020, we discussed and set forth the *IJDS* "mission," submission and reviewing requirements, article types, and editorial board members' load and terms; ME Chris Asher set up the journal website and submission system to match these decisions. We invited submissions from a range of researchers, planned journal presentations on editorial panels at INFORMS workshops and conferences, and other journal promotions. Most importantly, we onboarded an incredible group of 34 associate editors (AEs) and a reproducibility editor representing our different communities, diverse in terms of research areas, location, gender, academia and industry, and editorial experience. This new editorial board reflected the "who" of the initial *IJDS*. Over the years, the editorial board evolved as we learned more about the topics of submissions and in an ongoing attempt to expand to new relevant communities.

Cultivating and Shaping Excellent Research

I urge my PhD students to always submit their best work to a top journal, with the warning that they are most likely to get rejected. This paradoxical recommendation highlights the critical role of top journals not only in publishing excellent research but also, in shaping and improving the quality of rejected manuscripts by providing authors high-quality reviews and editorial guidance.

At *IJDS*, all submissions (except for desk-rejected manuscripts) have typically received professional feedback from two or three reviewers, a reproducibility check, and importantly, editorial guidance by the associate and senior editors. Although this sounds quite standard, I invested efforts in three directions that in my experience were not standard in data science journals: creating clear and transparent guidelines for reviewers and AEs; sharing expectations and best practices with editorial board members and authors; and introducing the journal directly to the communities via collaborations with top conferences, talks at different societies and editorial panels, and author guidance in editorial roundtable sessions. I would like to elaborate on these three efforts.

Transparent Reviewer and AE Guidelines

As a reviewer and AE for several journals, I have been perplexed about the lack of guidance on what was expected from my report. What is important to the journal? What are the priorities? How many points should I raise? Importantly, as an author, I was never made aware of the criteria that reviewers would use to evaluate my work. I would search the journal's "mission" and its editorials for clues, but these by no means reflected what reviewers necessarily followed.

In 2016, Ron Kenett and I published the paper "Helping authors and reviewers ask the right questions: The [information quality (InfoQ)] framework for reviewing applied research" (Kenett and Shmueli 2016), where we created a suggested reviewer guideline checklist. This checklist is suitable for applied empirical research and is based on the concept of InfoQ. When I stepped into the *IJDS* EIC shoes, I used this idea to create publicly available guidelines delineating the questions that a reviewer should answer when reviewing a submission as well as questions the AE should consider for not only consolidating the reviewers' key points but also providing clear guidance for the authors. The final question for the AE is "What would it take to make this paper publishable?"[5] This critical question urges the AE to provide constructive guidance. Over time, we refined these guidelines to include separate questions for "review"-type articles.[6]

Transparent reviewer guidelines provide fairness and efficiency for authors, reviewers, and editors. Authors can direct their efforts more effectively in choosing the journal and in writing their manuscript accordingly. Reviewers and AEs better understand the journal's priorities and can more easily review a manuscript and customize their feedback. The shared language reduces friction, uncertainty, and misunderstandings. Importantly, it helps reduce the number of revisions (which we have strived to limit to a single major revision) and the time until the final decision.

Sharing Best Practices

The *IJDS* audience and board include researchers in management schools, industrial engineering departments,

and industry. Submissions by members of these groups revealed that the different communities have different styles of authoring, reviewing, and providing editorial guidance. Also, with our community diverse in terms of geography, age, culture, and editorial experience, several communication gaps emerged. This raised the need to strengthen the "socio" part of our sociotechnical journal to create a cohesive journal with excellent human practices. I believe that clear communication is key to improving alignment and avoiding unintended friction. I, therefore, initiated the following actions to improve uniformity on critical dimensions.

- A set of editorials describing what an *IJDS* paper is ("*INFORMS Journal on Data Science* (*IJDS*) editorial #1: What is an *IJDS* paper?" (Shmueli 2021) and "Congratulations, it's our inaugural issue!" (Shmueli 2022)).
- Appropriate paper structure: guideline for authors on components that should appear in their abstract (problem, decision making, methodology, data, and results) and their paper (see https://pubsonline.informs.org/page/ijds/submission-guidelines) in addition to the reviewer and AE guidelines mentioned.
- "Best editorial practices" for communication with authors that are shared with editorial board members during board meetings.
- In-person meetings with editorial board members: monthly board meetings with the senior editors, an annual meeting with all board members, and personal onboarding meetings with new AEs, reproducibility team members, and media team members.
- Slides and a recording with these points and tips for editorial board members to look up.

Interacting with the Research Communities

A journal is part of a web of research activities, and as such, it is very useful to maintain its connectivity with other related activities. To that end, we actively engaged with conferences, competitions, panels, and seminars.

In our first effort, we have been collaborating with program chairs at top research conferences and workshops[7] to channel "best paper" submissions into fast-track *IJDS* submissions. Such collaborations are win-win-win for the authors, reviewers, and scientific community as they expedite the publication of suitable research, utilizing reviewers' efforts more efficiently. Most of these collaborations have successfully continued on a regular basis, and we now have a streamlined process for transitioning best papers (and their reviewers) as fast-track *IJDS* submissions. A few such papers have been successfully published.

Second, several SEs and I have been delivering webinars and in-person talks introducing *IJDS* to different communities (some of the recordings are on YouTube). These talks are an effort to expose new audiences to the journal and to introduce the *IJDS* editors to the audience. The *IJDS* media team has also posted short interviews with several SEs on *IJDS* social media. As part of our outreach effort, SEs and I participated in editorial panels at different conferences, sharing *IJDS* information and answering questions about submission and reviewing. We used these venues to also highlight *IJDS* innovations and sometimes different-minded approaches. For example, we published our liberal approach to the use of ChatGPT and other large language models (LLMs) in the 2023 editorial "How can *IJDS* authors, reviewers, and editors use (and misuse) generative AI?" (Shmueli et al. 2023). Another related type of outreach activity is providing guidance to authors in "editorial roundtable" events, such as those organized by the INFORMS Data Science Workshop.

Finally, at the upcoming 2024 INFORMS Annual Conference, we will introduce a special "*IJDS* Spotlight Session"[8] showcasing a few highly cited *IJDS* papers.

Making Publications Impactful

In parallel with shaping the journal's identity and embedding it into its intended audiences, I invested efforts into the longer-term goal of making the published work impactful. The key effort, which has been a big success, is our focus on reproducibility. A second is dissemination. Both involve expanding the editorial board to include members with special knowledge and skills, integrating the use of contemporary and cutting-edge technologies, and embracing openness and transparency.

"Play It Again": Trailblazing Reproducibility

The reproducibility crisis that hit various fields raised the awareness of journal leadership of the need to evaluate the reproducibility of their journal's published work. The reproducibility of an *IJDS* paper means that readers should be able to easily implement its novel methodology. In my experience, it is often very difficult to reproduce a complex computational method that someone else published because either the authors do not provide their code or else, their code does not run for various reasons. Reproducibility of computational methodology is, therefore, critical for a paper to have impact in practice. Publicly available, easily runnable code encourages a broader dissemination of knowledge and further adoption and development of those methods. Note that reproducibility at *IJDS* focuses on the code for running the methodology, not on the data (which should be replaceable).

When I looked at reproducibility four years ago, I found that journals were not yet set up for reproducibility checks. Journal editors who did decide to introduce reproducibility checks into their pipeline typically found that manually checking each manuscript is extremely labor and time intensive, requiring a highly professional

and often very large team. After extensive research, along with our Reproducibility Editor Enric Junque de Fortuny, we identified a scalable and robust solution that allows us to perform fast reproducibility checks of each *IJDS* submission by an extremely small reproducibility team (Enric Junqué de Fortuny, Nicholas Danks, and (recently joined) Baidyanath Biswas). Most importantly and differently from most journals, the reproducibility check is done at the start of the reviewing process rather than upon acceptance, allowing reviewers to also review the code alongside the manuscript. *IJDS*'s cutting-edge reproducibility solution is based on a collaboration with CodeOcean—a reproducible and traceable computational science platform. The CodeOcean solution runs the authors' code in the cloud, producing outputs that can then be viewed by readers/reviewers. Readers can themselves use a browser to run or modify the code in the cloud without the need to install any software.

Introducing reproducibility based on cutting-edge technology combined with a new reviewing process posed various challenges to the "traditional" manuscript submission system and processes. These included the difficulty of integrating the two systems (requiring manual matching of manuscripts with their submitted code) and hindering anonymous code submission as well as integrating the reproducibility workflow into the journal submission system. As a new journal, we took a "start-up" approach of innovation. Clearly, the technology will continue evolving, and processes will hopefully be more integrated in the future.

The result is that every published *IJDS* research paper has a link to the CodeOcean "capsule," where the authors' code can be easily run directly in a browser (the capsule is also archived on the journal website).

If You Build It (but They Cannot See or Reach It), Will They Come?

Publishing top research is a required but not sufficient condition for impact. The fact that a paper is published in a top journal does not mean that it will be read. For that to happen, it should be easily accessible, and information about it should be broadcast over the right channels. Toward these goals, we made sure that all *IJDS* papers are free access. We also allow (and encourage) authors to post their initial submission on repositories, such as arXiv and SSRN, as is common practice in machine learning and other computational communities with fast-moving research.

Finally, learning from other journals, we created a small *IJDS* media outreach team that is composed of researchers from our communities (Sriram Somanchi, Thiyanga Talagala, Yeawon Yoo, and previously, Reema Gupta). This tech-savvy, motivated, and devoted team has been posting on a variety of social media platforms about new issues, recently accepted papers, and new editorial board members. As a new journal, we wanted to use this channel to also familiarize potential authors and readers with the *IJDS* community. The media team created and posted a series of interviews with editors and accepted authors discussing issues related to the submission process, the review process, what editors look for in a submission, and what it is like to be an *IJDS* author. I'm grateful to this terrific team for their wonderful contributions and initiatives!

In Closing

These four years of starting up *IJDS* and creating a sound basis for its future have confirmed some of my beliefs about sociotechnical systems and occasioned some lessons and surprises.

First, I believe that my collaborative approach to the EIC role has proved successful at *IJDS*. Our diverse editorial board spanning fields, cultures, geographies, and more has become a close-knit task force, and working together, we have been able to overcome many challenges and achieve numerous goals. I am extremely grateful for the wise and thoughtful feedback and suggestions from the SEs. Our nearly 50 monthly meetings crossed many time zones, including those in Asia, Europe, and North America, with western-most members (Nick Street, Maytal Saar-Tsechansky, and most extremely, Olivia Sheng) waking up very early. All the board members, including ME Chris Asher, the AEs, and the reproducibility and social media teams, have been exceptional in terms of combining professionalism with care, innovation, and adaptability. The volunteer efforts by so many tireless researchers toward building this new journal have been truly impressive.

I take huge pride in the top-notch editorial board and the high-quality reviews that *IJDS* editors and reviewers have provided on *IJDS* submissions. Some authors have even written especially to commend us for the professional and constructive reviewing process. Although no journal is perfect, I believe we have created a strong basis for a top journal.

In terms of a "unified *IJDS* personality," I realized that this is a learning process that takes time. We have invested efforts toward unifying important aspects of submissions, reviews, and editorial actions. Increasingly, AEs and reviewers are using the review guidelines and forms. SE and AE reports have become more guidance-providing and use more constructive language, although there is always room for improvement. Given the differences between the business and engineering research communities, I strived to integrate the best practices from each community into *IJDS*. Creating a shared standard and language across business and engineering data science researchers has been challenging, but I believe the mutual learning via editorial guidance has been a valuable and worthwhile path.

This editorial has not yet mentioned common statistics that editors love to report, such as turnaround time, numbers of submissions, etc. As a statistician, I was quite surprised to discover how ill defined these key metrics are; despite using an information technology (IT) submission system (ScholarOne), these "important statistics" are highly definition-sensitive. For example, measuring the "number of submissions AE X handled in the year 2023" can vary depending on what is defined as a submission and how dates are considered. Do we count only original submissions? Revisions? What about a reject-and-resubmit manuscript that returns as a new submission? Should the count include manuscripts the AE desk rejected or only those that went out for full review? Should we count manuscripts that were *submitted* in the year 2023 or that reached the AE in 2023? Should the AE's report be submitted in 2023 or the EIC's decision be issued in 2023? In short, a statistic that at first glance appears straightforward is in fact an extremely volatile measure that depends on many decisions. Although metrics are indeed useful, there is the danger of "the tyranny of metrics." Because of these shortcomings of automated metrics, I used multiple manual spreadsheets to ensure that I had a better picture. My approach was possible at *IJDS* because it is still a small journal, but it is not scalable once a journal has hundreds or thousands of yearly submissions.

Another surprise was journal management systems. I recently coauthored a paper entitled "Reimagining the journal editorial process: An AI-augmented versus an AI-driven future" (Shmueli and Ray 2024) and participated in a panel on "The Future Impact of AI on Academic Journals and the Editorial Process" at the 2023 International Conference on Information Systems. In our paper, we compared an editorial system that uses "simple" machine learning (mainly supervised learning) to assist decision making by editors, reviewers, and authors with an AI-driven system, where the system makes and shapes decisions in a more independent and automated way. While thinking about and writing the paper, I could not avoid seeing how distant the current journal management systems are from even the "simple machine learning" scenario. I also learned how much friction, inefficiency, and uncertainty an IT system can create if not continuously tested and updated with the users in the loop. It is prime time for a better journal management system.

I am happy to have shared with you some of the efforts, challenges, and successes that we have experienced at *IJDS*. I hope this encourages you to join the *IJDS* community. I once again thank all our readers, authors, reviewers, and editorial board members for your wonderful dedication and contributions. The end of my EIC term opens the door for new energy, innovations, initiatives, and opportunities that the incoming EIC will bring. I look forward to seeing the growth of *IJDS* in the years to come.

Finally, as you read through the rest of this issue, I hope you will find it interesting and useful as it presents another set of novel data science methodologies that support decision making.

Endnotes

[1] ChatGPT or other large language models (LLMs) were not used in the writing of this editorial.

[2] Senior editors who joined prior to the journal launch include Yu Ding, Nick Street, Rema Padman, David Martens, Shawndra Hill, Irad Ben-Gal, and Kwok Tsui. We collaborated intensively toward the *IJDS* launch in November 2020. Thank you each for your critical contributions at this critical stage of the journal.

[3] The term "decision making" is commonly understood by data scientists in business schools, whereas engineering researchers typically use the term "real world."

[4] Selection bias is, of course, plausible.

[5] Thanks to Maytal Saar-Tsechansky for this excellent phrasing.

[6] Thanks to Daniel Gartner for his help on structured reviews.

[7] We have collaborated with the INFORMS Quality Statistics & Reliability (QSR) Conference, the INFORMS Workshop on Data Mining and Decision Analytics (DMDA), the INFORMS Data Science Workshop, Workshop on Information Technologies and Systems (WITS), and the Institute of Industrial and Systems Engineers (IISE) Annual Conference, Quality Control & Reliability Engineering (QCRE) Division.

[8] The *IJDS* Spotlight session, organized by SE Nick Street, is invited by the Data Mining cluster and crosslisted with the QSR and AI clusters.

References

Kenett RS, Shmueli G (2016) Helping authors and reviewers ask the right questions: The InfoQ framework for reviewing applied research. *Statist. J. IAOS* 32(1):11–19.

Shmueli G (2021) *INFORMS Journal on Data Science* (*IJDS)* editorial #1: What is an *IJDS* paper? *INFORMS J. Data Sci.*, https://pubsonline.informs.org/pb-assets/What%20Is%20an%20IJDS%20Paper-1680176133.pdf.

Shmueli G (2022) Congratulations, it's our inaugural issue! *INFORMS J. Data Sci.* 1(1):1–3.

Shmueli G, Ray S (2024) Reimagining the journal editorial process: An AI-augmented vs. an AI-driven future. *J. Assoc. Inform. Systems* 25(1):10.

Shmueli G, Colosimo BM, Martens D, Padman R, Saar-Tsechansky M, Sheng ORL, Street WN, Tsui K-L (2023) How can *IJDS* authors, reviewers, and editors use (and misuse) generative AI? *INFORMS J. Data Sci.* 2(1):1–9.