

**A Vision for
ACM Transactions on Database Systems (TODS)**

Z. Meral Özsoyoğlu

July 27, 2007

1. Introduction

ACM TODS is the premier journal in databases, and I am honored to be considered a candidate for the next ACM TODS Editor-in-Chief (EiC). This document briefly summarizes what I see as the current status, challenges and opportunities for ACM TODS, and my vision for maintaining its status as one of the best scientific journals across all disciplines in terms of prestige and impact, and moving it forward in the 21st century: the age of Internet and the changing times.

2. Relevant Experience

I have been an ACM member since 1973, starting as a student member. I have been a reader of ACM TODS and SIGMOD papers since 1976. Over a span of 32 years, my relationship with both SIGMOD and TODS continued to evolve: as an author for both, as an associate editor for TODS, and as a program committee member, organizing committee member of SIGMOD conferences, and vice chair for SIGMOD. I have been an associate editor for TODS for six years, joining the editorial board when Rick Snodgrass became the EiC in July 2001. I was an associate editor for both IEEE TKDE and IEEE Data Engineering Bulletin. Over the years, I served in program committees of SIGMOD, PODS (as well as VLDB and ICDE), and chaired program committees for PODS, ICDE and SSDBM. I was also a vice-program chair for ICDE several times in the past. Finally, I served as SIGMOD vice chair between 1997 and 2001.

3. Current Status

ACM TODS is indisputably the most prestigious and premier database journal. In terms of citations and impact, it leads all database journals. As stated in the ACM TODS charter, “TODS helped to define the field of database research”. TODS is founded by David Hsiao, the first Editor-in-Chief (EiC), with “the aim of being the focal point for integrated dissemination of database research and development results” in a wide range of areas, and striving to “publish high quality papers on all aspects of computer database systems” [H78].

TODS is further improved during the second EiC, Won Kim, who maintained the leadership of TODS as the premier database journal, and also decreased the turnaround time for papers for increased relevance.

The third and the most recent EiC, Rick Snodgrass, has also made significant improvements to TODS during his term as EiC¹ for TODS to maintain its status as the premier journal, as well as to improve the effectiveness and efficiency of handling papers. All of these had a positive impact on authors and readers without sacrificing the selectivity and high quality of papers that are published in TODS. These improvements are all documented in the form of editorials and SIGMOD Record articles that are available online [Snod]. I will mention two of these here: the editor’s manual and adapting ACM Manuscript Central and tuning it for ACM TODS. These changes in turn had direct implications on reducing the turnaround time, making TODS more relevant and timelier as a journal, and increasing its impact on database research. Rick Snodgrass also championed the double blind reviewing for TODS papers, which is approved as a trial for two years by the ACM Publication Board.

¹ I am more familiar with the improvements on ACM TODS made during Rick Snodgrass’s term as the EiC since I had the chance of observing them first hand and collaborating with him as a member of the TODS editorial board.

4. Challenges and Opportunities

The main challenge for the next EIC of TODS is to maintain the status of TODS as the premier database journal, *and* to expand its reader base and make it even more relevant into the 21st century. The continued evolution of TODS and its forward movement needs to continue and even speed up. As a member of the database community for more than three decades, I am deeply interested in the well-being of TODS. In this section I will outline a list of issues, some of which had been perceived as challenges previously [Snod], and some are new opportunities to retain TODS as the most forward-looking journal in ACM. And, in the next section, I will outline how I plan to approach these challenges if I am named as the next EIC of TODS.

Below I first list a number of challenges selected from those elaborated by Rick Snodgrass in his quarterly columns in SIGMOD record.

- *Declining individual subscription base for TODS.* This decline is mostly due to the increases in institutional subscriptions to the ACM Digital Library, and, thus, the decrease in individual subscriptions to TODS is to be expected. Consequently, enlarging the reader base is a more important metric than increasing the individual subscription base to measure how viable TODS is.
- *Making ACM the preferred publisher.* As a response to the ACM publication board's policy "Rights and Responsibilities in ACM Publishing", TODS under the guidance of Rick Snodgrass has taken a number of specific steps, including
 - *Timely review and clear feedback.* The latest average manuscript processing time (submission to editorial decision) is 3.0 months.
 - *Improved turnaround performance and reduced end-to-end time.* Maximum turnaround time is now 6 months. End-to-end time now is also the lowest in over two decades, which is currently around 16 months.
 - *Policies/Guidelines on reviewing, choosing referees and referee loads.* There are several guidelines, policies and statistics that are made available online with the ACM Manuscript Central. There is a new "Editors Only" section on the TODS web site, which is a great resource that helps editors in every aspect of processing papers, and has a positive impact on improved turnaround time performance, and more effective and more efficient processing of submissions.
 - *Automated help for editors in processing papers.* Current ACM Manuscript central is very useful to editors in keeping track and the processing of submissions, by sending automated reminders and acknowledgements.

My own personal opinion is that during Rick Snodgrass' term as the EIC, TODS have made significant accomplishments towards facing this challenge.

- *Timeliness and originality of the work published in TODS.* As with any archival journal, this is an ongoing challenge. Clearly, reducing end-to-end turnaround time is a great step towards making TODS papers timelier. And, as Rick Snodgrass also noted [Snod], in terms of the timeliness of its results, it is difficult for TODS to compete with the top-class and international conferences such as SIGMOD/PODS, VLDB, ICDE, EDBT and ICDT. Moreover, many of the TODS submissions are now based on work published in such conferences, reducing the originality and impact factor of the submissions. While TODS has the originality requirement that "each new submission must contain at least 30% new material over the version published in an earlier conference", this may still give the perception that, while publishing excellent articles, main research results reported in TODS are delayed and have lesser-impact; and this may have a negative effect on the TODS reader base. I believe that, on this issue, we can adapt approaches used other fields to reduce this "delay",

and move TODS to a unique position by making it an indisputably must-read journal in data management and databases as well as being the premier archival journal. I elaborate more on this in the next section.

Next, I list what I perceive to be additional opportunities for TODS.

- *Relevancy in core database research.* As with any maturing field, the research reported in database conferences has become quite diverse and rich, and continues to get richer. In any year, in the top database conferences, there are many must-read new and high-impact papers for researchers and practitioners in various areas of databases. It will be great to publish a selected set of best papers from these conferences in TODS without much delay. TODS publishes annually between 20-30 papers² (with an average of about 5-7 publications per issue), and is limited by the page restrictions of hard copy publications. A new track for TODS which is fully electronic, i.e., online only, can be an opportunity for publishing full versions of such high impact papers without much delay, in addition to the regular, hard-copy TODS publications, similar to other journals in other fields.
- *Relevancy to other fields with major data management and database issues.* Presently, there are other fields that are faced with major data management and database problems, and need new and innovative data management/database solutions to their problems. Here I will give an example from biology, a discipline I have become familiar. There are significant biological data management issues--so significant in fact that (a) the National Science Foundation Biology division has the BDI (Biological Databases) Initiative, and the DB&I program (Databases and Informatics), and (b) the National Institute of Health³ now refuses to fund projects if they have data management/data exchange research components and such components are not clearly spelled out using precise database terminology. As a premier journal in databases and data management, what should TODS' role be in fields with data management problems, such as biology, and how can this role be promoted and achieved?
- *Reproducible experimental results or systems work in publications.* At the present time, the current practice in the database community is that the authors of submitted experimental or systems papers are not expected to provide an independent verification of the validity of their results. I refer to this as "reproducibility", which may have multiple parts, such as submitting code, data, and constraints used to conduct experimental or systems work. Currently, there are no specific "reproducibility" requirements of experimental work or software systems, though TODS welcomes ancillary unrefereed information, including code, from authors. In comparison, such is not the case for theoretical parts of submitted papers, and the authors of such papers are required to fully submit proofs of their theorems, lemmas, etc., for review. This is a broad issue that needs to be taken up by the SIGMOD leadership and the TODS editorial board; but, I believe that time has come to develop "reproducibility standards" for experimental/systems publications. In my career as a database researcher over three decades, I have personally encountered cases of (a) "software systems" papers, where, when I needed to compare the performances of my own relevant system, it became clear that the "system" in the published paper really did not exist anymore, or was a set of non-usable and disconnected pieces of code/scripts, or (b) "experimental work code" that was no longer available.

² TODS has been originally planned to have 5-6 papers per issue [H 76]. While the number of papers per issue has dropped to less than 20 for some years in the past, it has been increasing during the last 7 years, and reached to a maximum of 36 in 2006. This increase is mainly due to the increased submissions as a consequence of more visibility, and shorter papers by replacing long papers with shorter ones with online supplements.

³ Actually there are several database researchers whose research is funded by NIH and who routinely serve as members of NIH research review panels.

- *Outreach to other database conferences.* Database community, and TODS as a premier journal, will gain from synergistic collaborations with other database conferences. Previous EiC Rick Snodgrass has already started working closely with the executive committees of ICDT, EDBT and PODS to publish selected best papers from these conferences in TODS with accelerated reviews. Extending this outreach will enhance the visibility of TODS, increase the buy-in from the community, and facilitate collaboration to enhance the visibility of our field across other disciplines, as well as setting the standards for scientific publishing for our research community.

5. My Vision

If named the next Editor in Chief for ACM TODS, I will work closely with the TODS editorial board, SIGMOD executive committee, and ACM publications board to respond to the challenges confronting scientific publishing and ACM TODS. Meeting with the TODS editorial board members face to face annually, possibly during SIGMOD conferences, can be very effective for reviewing the current year, assessing the current status, identifying challenges and opportunities, and strategic planning to evolve the journal to meet the needs of our community better. This will be in addition to communication by email and online discussions throughout the year with the editorial board on regular issues, and announcements. Below, I discuss my position on the challenges and opportunities listed in the previous section.

5.1 Relevancy in Core Database Research: Conference Special Track for TODS

At the present time, within the database community, with some exceptions, the most innovative and the most groundbreaking work first gets published in database conferences, usually, the top premier conferences such as SIGMOD, PODS, VLDB, ICDE, EDBT, and ICDT. And, usually, some of these papers find their ways to TODS in much more complete and stronger forms. To help with this transition, Rick Snodgrass as the EiC instituted a process by which a small number of selected best papers from conferences are invited for submission to TODS—after significant extensions and revisions. Such papers go through a thorough review and revision process, and get published in regular issues of TODS.

An alternative approach can be attracting the best database conference papers to TODS as a new “conference special issue” track for TODS, in extended form (with proofs and more complete experimental and systems work) and *without losing anytime* so that they retain their originality. This of course needs to be discussed and agreed upon by the TODS editorial board, the SIGMOD and PODS leadership, and the ACM publication board before it can be adapted. One possible approach⁴ may be as follows: (1) Premier conferences (such as SIGMOD and PODS), with the involvement of members of the TODS editorial board, select a number of papers directly to TODS, subject to the revisions recommended by the reviewers. The conference reviews themselves become more rigorous and detailed for the selected TODS conference special issue track papers. (2) Authors of such selected TODS papers are requested to “complete” their papers with proofs, etc. (if applicable), as well as with additional experimental work (if any), and that the revisions be completed in 2-3 months, matching the conference time. Each such revised TODS paper has no or

⁴ This is similar to the approach used in computational biology, another fast-moving scientific field. The two premier conferences in computational biology, namely, ISMB (rated number 1), and RECOMB (rated number 2), each is associated with a journal: ISMB with the *Bioinformatics* journal (rated as among the “highest-impact” journals), and RECOMB with the *Journal of Computational Biology* (another high-impact journal). When ISMB conference papers get accepted, a selected number of conference papers are also accepted to a “supplemental ISMB issue” track of the *Bioinformatics* journal, subject to a very small number of changes as requested by the reviewers. Such papers from ISMB are published in the *Bioinformatics* journal with a two-month delay after the ISMB conference gets held.

little size limitations (but, satisfies the formatting requirements of TODS). (3) The reviews of revised papers get handled by a TODS associate editor, who makes an effort to expedite the process. (4) The paper “usually” gets completed by conference time, and becomes available in full form to only the conference attendees (to encourage conference attendance), and (5) two months later, the paper and the conference special issue becomes available to all TODS readers.

There can be refinements to the above-summarized proposal for “conference special track papers” of TODS. For example, one practice I have observed in biology journals is that different types of papers have pre-specified section heading formats such as “methodology”, “experiments”, etc.

Such a policy, if adapted, has several advantages: It is likely to increase the readership for TODS and the awareness of TODS as the place to read full versions of best conference papers; it is more desirable to read the extended versions of best papers of SIGMOD/PODS at the same time as (or a very short time thereafter) the conference itself; accessing the online supplemental TODS issue can be an additional reason to visit the TODS site and browse the contents of other TODS issues including the current one.

5.2 Timeliness and Originality of the Work Published in TODS

During the last 6 years, there have been significant improvements in speeding up the processing of traditional TODS submissions with the leadership of Rick Snodgrass. Further end-to-end publication time improvements to traditional TODS submissions, while difficult to achieve, should continually be pursued. We also need to look for new and refreshing approaches to increase timeliness and originality. As an example of this latter direction, I believe that the approach listed above in section 5.1, if implemented carefully, effectively, and thoughtfully, is a candidate for immediately increasing the timeliness and originality of the work published in TODS.

5.3 Relevancy in Core Database Research: Main Paper Track for TODS

Currently, TODS papers have a page limitations of 45 pages (the statistics up to 2006 can be found in [B+05]). These statistics on the length of TODS articles, while showing a great improvement over the recent past⁵, limits the total number papers that can be published in hardcopy publications per volume. Even though, the current trend is very encouraging in this respect (12 articles in 1996 to 31 in 2006 [B+05]) there is still room for more modest increases in the number of TODS papers published in a year without sacrificing the quality of publications that TODS is known for. This will also help make TODS more vibrant, and enlarge its readership base.

5.4 Relevancy to Other Fields

I believe that TODS can and should attract some of the highly innovative works of database researchers in other fields. Such an approach will also be highly effective to open up new research directions and new research areas for other young database researchers. I will use again the “bio” related fields as an example, such as genetics, biophysiology, physiology, various sub-fields of biochemistry ranging from nutritional biochemistry to metabolic biochemistry, various up-and-coming “omics” fields (e.g., proteomics, genomics, and metabolomics), and, ultimately, the emerging field of systems biology. These fields are experiencing data management/database renaissances, and this is only expected to intensify in the 21st century. Moreover, there are already a select few database researchers who are active in biology-related database/data-management research, publishing their database research in other biology-only outlets. Encouraging such research to appear in TODS, in the form of supplemental online-only special issues to foster research in new areas can be an effective first step towards being “friendly”, and, thus, relevant, to other fields with increasingly large database problems.

⁵ In 2000, max number of pages in a TODS article was 80, the average was 44, and the minimum was 34. In 2005, these statistics are: max. 60, average 40, and min. 20.

In summary, I think that there is a role to be played in the field of biology by database community in general and TODS in particular. TODS can and should attract excellent and groundbreaking data management/database research results into its issues. Such a new direction cannot of course be taken by a single individual, and should be a consensus decision involving the editorial board of TODS as well as the leadership of SIGMOD. My personal opinion is that TODS should be viewed as being among the premier publication outlets for data management and database-related research in other fields such as biology.

5.5 Reproducibility of Experimental/Systems Papers

I believe that the TODS editorial board and SIGMOD leadership should lead an effort to develop reproducibility standards involving experimental/systems paper submissions. Several issues immediately come to mind: protecting the ownership rights of authors of systems publications; what needs to be submitted and in what formats; how long should the reproducibility data of accepted papers need to be maintained; how should the reproducibility data of submitted papers be guarded and/or be made available, etc. Once these issues are ironed out, it is my opinion that TODS should take a leadership role in implementing such a standard.

5.6 Increased visibility and transparency.

With a series of articles in the form of ACM TODS editorials and quarterly columns in SIGMOD record by the EIC, the readership is kept up-to-date about what is happening with TODS. I think that this is a great practice helping TODS, and should continue.

5.7 Controversial issues in Database Community on Scientific Publishing

Currently, due to increasingly large numbers of submissions to major database conferences (SIGMOD, PODS, VLDB, ICDE and others) and despite the fact that we as database community are spending much time⁶ for paper reviewing, there is room for improvement and new creative solutions for publishing papers in conferences. There is usually a very large group of papers that end up in the weak accept/weak reject category, and it is very costly in terms of time to sort out these papers so that a small number of papers can be accepted and included reasonably for presentation in the conference. There are strong views on these issues, and a wide range of solutions are being discussed. While some researchers support keeping the conferences as-is despite the existing problems, others support having an online journal instead of conferences for paper submissions, and change the conferences as places for a few invited presentations (selected from the online journal).

There is also the controversy between pros and cons of double blind reviewing (see the latest SIGMOD record articles). After long discussions and debates among the editorial board of TODS, double blind reviewing has now been recommended, and the ACM Publication Board has approved the adaption of double blind reviewing for TODS on a trial basis. However, even during the trial basis, I envision changes to double blind reviewing, especially for papers that are originally published in major conferences, and papers that are selected from conferences and invited for publication in TODS after additional rounds of reviewing. For such papers, double blind reviewing does not really seem to be applicable.

Database community as a whole needs to establish guidelines, and standards for scientific publishing using a variety of publication venues, from SIGMOD record, IEEE DE bulletin to conference and workshop proceedings to archival journals, including ACM TODS, VLDB journal, and IEEE TKDE. TODS is a very important component in this closely interleaving spectrum of publication outlets, and, I believe, TODS EIC and the TODS editorial board should collaborate closely with executive committees of SIGMOD, PODS, and other conferences, VLDB Endowment, EIC's of

⁶ The PC committees are becoming larger and larger, and yet the load for each PC member stays almost unchanged.

other major journals, and the ACM publication board, in setting standards and providing leadership for resolving issues with unifying support and buy-in from the community across all levels.

6. Conclusion

TODS is in great state both in terms of prestige and impact. The database community and TODS have been lucky to have hard-working and innovative TODS EiCs and TODS editorial boards, who elevated TODS to its current state. Especially, with the innovations and hardwork of the previous EiC Rick Snodgrass, TODS now has the much needed guidelines and infrastructure for processing papers, and has great online resources to help authors, readers and editors.

In summary, it is an honor to be considered for the position of Editor-in-Chief of ACM TODS. If named, I am ready to face the challenges and opportunities, and assume the responsibility to maintain and move forward the status of TODS as the premier journal in the field. To do so, I will work with the database community peacefully and energetically towards making TODS more vibrant, dynamic and relevant-- with new initiatives and without sacrificing its technical quality.

7. References

- [H 76] D. Hsiao, "Aim and Scope", ACM TODS, Vol. 1, No.1, 1976.
- [H 78] D. Hsiao, "TODS—The first Three Years (1976-1978)", ACM TODS, Vol. 3, 1978.
- [Snod] R. Snodgrass, Editorials, <http://www.acm.org/tods> .
- [B+ 05] P. Bernstein, et al., "An Apples to Apple Comparison of Two Database Journals", SIGMOD Record, Vol. 34, No.4, December 2005.