TUGboat's 100 issues—Basic statistics and random gleanings

David Walden and Karl Berry

Abstract

TUG was founded to provide an organization for people who are interested in typography and font design, particularly those interested in Don Knuth's TeX typesetting system. TUG's print journal TUGboat, now at its $100^{\rm th}$ issue over 32 years (certainly a noteworthy run), has been and remains an important component in carrying out TUG's mission.

1 Remembering Knuth's 3:16 book

Casting about for an appropriate article for the 100th issue of TUGboat, we remembered Donald Knuth's book entitled 3:16 Bible Texts Illuminated. In that book Knuth describes studying the Bible by looking at chapter 3, verse 16 of each of 59 books of the Old and New Testaments of the Bible (he left out books that were so short they don't have a verse 3:16if the book was long enough but stopped short of verse 16 in chapter 3, he kept counting from the last verse of chapter three into chapter four until he got to the sixteenth verse). For each such arbitrarily (randomly) selected verse, Knuth's book has four pages: (1) a sketch of the book as a whole; (2) a calligraphic transcription of the verse (each from a different renowned calligrapher), (3–4) Knuth's restatement of the verse in contemporary English and a description of his research and analysis of the verse and placing it in context.

In the book's foreword and afterword, Knuth discusses how his random sampling approach (a mathematician and computer scientist's approach) might, and in fact did, produce a revealing picture of the Bible more generally. This suggested to us that a random sampling approach might also be an interesting way to get an overall picture of the first 100 issues of TUGboat. Also, there would be a certain symbolism in using a method promulgated by Knuth.

2 Our random sampling approach

We first considered choosing the 100th page (symbolic of this 100th issue) of each yearly volume. However, that had the problem that a couple of volumes didn't have as many as 100 pages. Furthermore, the order of papers in issues is not random, being organized in categories, such as "Macros" and "Fonts", in an order more or less consistent from issue to issue. Always using page 100 reduces the chances of selecting a

wal	********	total	physical		
1	1980	pages 23	issues 1	page 21	notes
2	1981	$\frac{23}{267}$	3	219	
3	1982	88	2	31	
4	1983	132	$\frac{2}{2}$	118	
5	1984	168	$\frac{2}{2}$	79	
6	1985	174	2	140	
7	1986	198	3	125	
8	1987		3	93	
9	1988	342	3	256	
10	1989	765	4	102	
11	1990	693	4	494	
12	1991	588	3	167	four logical issues
13	1992	544	4	396	1041 1081001 100400
14	1993	445	4	213	
15	1994	508	4	359	
16	1995	443	4	110	
17	1996	411	4	263	
18	1997	321	4	245	
19	1998	440	4	427	
20	1999	404	4	286	
21	2000	440	4	427	
22	2001	376	3	60	
23	2002	359	3	41	four logical issues
24	2003	624	3	285	
25	2004	232	2	106	excludes TUG'04
		202		2.0	conference preprints
26	2005	302	3	26	
27	2006	268	3	268	
28	2007	384	3	208	
29	2008	488	3	38	
30	2009	183	3	40	
31	2010	340	3	249	2011
32	2011	128	1	37	one 2011 issue
Total pages:			1143	80	
Total issues:				00	
Average pages/issue:				.7	
Average pages/year: 365					

Figure 1: *TUGboat* statistics and information (average page figures are based on extrapolation of first 2011 issue size to the full year).

page in a category that typically appears before page 100 of the first issue of a volume.

Thus we decided to select a random page from each volume, summing the page totals on the rare occasions when an issue 2–4 started over at page 1, and for the remaining volumes using the largest page number in the table of contents (TOC) of the last issue of the year, as found in the online TUG-boat TOCs (http://tug.org/TUGboat/Contents). These preparations for finding a random page in each issue immediately gave us some statistics about TUGboat. This is all summarized in Figure 1.

We also patched our program that generates the

 $^{^{1}}$ A-R Editions, Inc., Middleton, WI, 1991.

all-TUGboat tables of contents and lists (see our companion article in this issue) to count the number of titles and authors across all 100 issues. The result was 3,127 titles and 2,446 authors which equals, for what it's worth, an average of 31 titles and 24 authors per issue. These numbers are generous, since they include each independent item on the TOC pages, such as the front cover, complete issue PDF, etc.

Volumes of *TUGboat* include the TUG annual conference proceedings for 1989–2010, excluding 2004 which was published as a Springer-Verlag book, three PracTEX conference proceedings, four EuroTEX conference proceedings, one combined EuroBachoTEX conference proceedings, and one NorthEast U.S. conference proceedings.

The described random selection method has the bias of leaving out unnumbered pages at the ends of some issues which at various times included TUG job opening notices, the TUG mailing list, TUG membership and ordering information, TUGboat submission information, TEX and METAFONT errata, the membership list, an order form for the book Joy of TEX, other advertisements, an AMS-TEX panel discussion, profiles of TEX installations, Computers and Typesetting errata, changes and supplements, and other miscellany. However, sticking to numbered pages was easier than having to logically renumber all pages in a volume to include unnumbered pages.

Also, the largest page number listed in the online TOCs might leave out a follow-on page to the last TOC entry since article start pages and not page intervals are given in the online TOCs. The cover pages (c1–c4) were also ignored. Finally, as of issue 100 we only have the first issue of volume 32, and we did the random page selection from the 124 pages of that single issue (once we were pretty sure of the number of pages in the issue including this article).

For the computations, we used an online random number generator (http://quantitativeskills.com/sisa/calculations/random.htm), rather than programming, for instance, the linear congruence pseudo-random number generator described by Knuth in The Art of Computer Programming, Volume 2: Seminumerical Algorithms, Chapter 3, "Random Numbers"—just because it was the easier path.

The reader might well ask, "Would it not have been more symbolic to choose a random page from each of the 100 issues rather than from each of the 32 volumes, and this could also have provided a usefully larger sample?" We answer, "Perhaps, but we were writing a journal article, not a book; it would also have taken considerably more effort—though we acknowledge that deliberately choosing the less-effort approach departs from the Knuthian way."

3 The selections and notes

In this article for a single journal issue we cannot follow Knuth's model of four pages for each sample. Instead we made a copy of each article containing a randomly selected page, and then we thoughtfully (no longer randomly) made appropriate comments on the randomly selected page, the paper more generally, the author, and/or what it suggested about TUG and TUGboat.

Even someone who has never seen another issue of TUGboat may get a reasonable idea of its coverage from the following set of samples. On the other hand, long-term readers of TUGboat may be reminded of favorite articles, or other TUGboat articles not included in the random samples. We have also included notes on and references to current developments relating to the selected papers.

The randomly selected page numbers shown in column 5 of Figure 1 result in the papers below being chosen for each volume. Links to this set of selected articles may be found on the TOC page for this issue, http://tug.org/TUGboat/tb32-1.

 Questions & Answers; Letters; Miscellaneous, 1:1, p. 21.

On pages 2–3 of the first issue of *TUGboat* (the only issue in the first volume), editor Robert Welland explains, "The TEX Users Group (TUG) met at Stanford University in February of 1980 . . . and among other things decided that the group would publish a newsletter to assist the distribution of TEXpertise." These pages are worth reading (http://tug.org/TUGboat/tb01-1/tb01edit.pdf) and include some justification for the name *TUGboat*. The first issue appeared in October 1980, and covered a variety of early TEX and *TUGboat* topics, including reports from the February users group meeting.

Our randomly selected page from that issue contains requests that committees organized at the users group meeting identify themselves and that people send reports on TEX sites, introduces two letters from satisfied users, and mentions a TEX Errata list and the TUG mailing list. The "newsletter" was up and was finding its way.

- **2.** "TEX for the HP 3000", by Lance Carnes, 2:3, pp. 25–26.
- **3.** "TEX-news from Pisa", by L. Aiello and S. Pavan, 3:1, pp. 31-32.

The randomly selected pages for both volumes 2 and 3 were in the Site Reports area (http://www.tug.org/TUGboat/Contents/listkeyword.html#CatTAGSiteReports), an important TUGboat feature for a decade or more as the TEX community spread The selected article in volume 2 covered 13

TeX sites in 15 pages, and the volume 3 report covered 5 sites in 5 pages.

Lance Carnes has remained a notable figure in the TEX community ever since, first as the "small TEX" department editor for TUGboat and later (and still) as the vendor of PCTEX, a well-known commercial distribution of TEX.

4. "Summary of AMS-TEX", by Michael Spivak, 4:3, pp. 103–126.

The first issue of *TUGboat* included an article on AMS-T_EX, one of the first large macro formats created for T_EX. by its creator, Michael Spivak. Here, he gives an update consistent with T_EX82.

5. "First principles of typographic design for document production", by Richard Southall, **5**:2, pp. 79–90.

Richard Southall was a type designer and typographer, invited to be part of the digital typography group at Stanford in the early 1980s. Of this paper he says, "Leslie Lamport and I taught a two-day course on 'First principles of typographic design for document production' as a preliminary to the TUG meeting at Stanford University, August 13–14, 1984. What follows is an expansion, somewhat revised and restructured, of my lecture notes." We can speculate that this content influenced the way that Lamport thought about the design of LATEX.

6. "Assembling a moderately-priced, high-performance clone of the IBM PC for running T_EX", by M. Pfeffer and A. Hoenig, **6**:3, pp. 14–145.

Running TEX on smaller and smaller computers was an ongoing topic in *TUGboat* until personal computers became logically big. This article is the first instance of a column entitled "Typesetting on Personal Computers", and it starts at the beginning by providing an instruction manual for buying and assembling the parts of an appropriate personal computer.

7. Title page, 7:3, p. 125.

TUGboat title pages are generally given regular page numbers in sequence, and this one was randomly selected for our sample (Figure 2). We can see that Barbara Beeton was already three years into her almost thirty year (and still counting) tenure as TUGboat editor. From TUG's address, we also have a hint of the major role AMS played in the early years of TEX (and continues to play, in that Barbara and other TEX developers are AMS employees).

8. Advertisements, 8:1, pp. 87–96.

This issue of *TUGboat* included 10 pages of TEX-related advertisements. For many years TEX was a leading-edge development, and many people hoped to make money from TEX.

... books look like books because, though it sounds a bit simple to say so, that's what they are. And that's what they're supposed to look like.

Richard Hendel
"A book designer's odyssey"
Scholarly Publishing,
July 1986, p. 350

TUGBOAT

THE TEX USERS GROUP NEWSLETTER
EDITOR BARBARA BEETON

VOLUME 7, NUMBER 3 • OCTOBER, 1986 PROVIDENCE • RHODE ISLAND • U.S.A.

Figure 2: Title page of volume 7, issue 3.

9. "TEX output devices", by Don Hosek, **9**:3, pp. 251–260.

In the days before PDF, a big issue was making TEX (and DVI) work with a wide variety of output devices; thus, for a while TUGboat had regular reports on output devices for TEX (http://tug.org/TUGboat/Contents/listkeyword.html#CatTAGOutputDevices). This particular report consisted of 10 pages of charts and tables about various output devices.

10. "Contents of archive server as of 16 January 1989", by Michael DeCorte, **10**:1, pp. 97–102.

With the advent of the Internet and FTP, naturally it made sense to have online archives of content, in this case relating to LATEX. DeCorte wrote five articles from 1988–1990 on the content and organization of this archive, hosted at Clarkson University, along with providing contact information to acquire TEX material on diskettes, etc. The TEX archives spread, through collaboration and independent development, becoming CTAN today (http://www.ctan.org).

- 11. "Comments on the future of TEX and META-FONT", by Nelson Beebe, 11:4, pp. 490–494.
- 11. "Editorial Comment", by Barbara Beeton, 11:4, pp. 494–496.

There are two articles on the randomly selected page (page 494) of volume 11—one article ending and one starting. In the context of Knuth's thenrecent announcement that he had finished his work with TEX and METAFONT, Nelson Beebe discussed TUG's ongoing role, TEX's place in the world, the need for continuing TEX and TEX-related developments, and so on.

As if to emphasize the continuing viability of TEX and related activities, Barbara Beeton touched on a significant number of TUG and TEX meetings, TUGboat activities, etc. TEX, TUG, and TUGboat were continuing with or without Knuth.

12. "Some T_EX manuals", by Angela Barden, **12**:1, pp. 166–170.

Author Barden critiques five early well-known books on TeX and LaTeX (she likes Leslie Lamport's LaTeX manual) and mentioned a couple of others. She discusses her struggle to learn TeX and summarizes her philosophy of what would make a good tutorial book on TeX. In other words, her article is indicative of a long-standing problem with TeX, that has never fully been solved, particularly given the open-endedness of the TeX world (so different from a highly specified commercial product).

13. The Donald E. Knuth Scholarship: 1992 Scholar and 1993 announcement, **13**:3, pp. 395–396.

Named for Don Knuth, the scholarship was aimed at recognizing and promoting the use of TeX by "support" personnel (as opposed to professors and programmers). Up to \$2000 was provided toward attending the TuG conference. In keeping with other changes in the TeX world, the scholarship has not been awarded since 1997 (http://tug.org/committees.html).

14. "A format compilation framework for European languages", by Laurent Siebenmann, **14**:3, pp. 212–221.

The author developed a package to enable TEX formats (Knuth's word for a set of macros such as those that define plain TEX) to include hyphenation patterns for many languages, along with other multilingual support. Over the years TUGboat has published papers on many interesting ideas or experiments that never saw wide-spread use and where the perceived problem was later more or less solved by a more general capability. (In the TEX distributions of today, all suitable hyphenation patterns are included by default.)

15. "TeX innovations at the Louis-Jean printing house", by Maurice Laugier and Yannis Haralambous, **15**:4, pp. 438–443.

Over the years since the creation of TeX, various publishers, typesetters, and printers have made use of TeX, particularly as a component in automating their processes. Some of these entities have created their own tools to move between TeX and other document processing systems, and this article describes one such case.

16. "A practical introduction to SGML", by Michel Goossens and Janne Saarela, **16**:2, pp. 103–145.

The authors of this 43-page article explained, "This article discusses the basic ideas of SGML and looks at a few interesting tools. It should provide the reader with a better understanding of the latest developments in the field of electronic documents in general, and of SGML/HTML in particular." The article was one of seven articles that made up TUG-boat 16:2 (June 1995), grappling with how TEX would fit into the rather new world of SGML, HTML, hyperlinks, and so on.

17. "TeX in Russia: ab ovo, or About the TeXnical evolution in Russia", by Irina A. Makhovaya, 17:3, pp. 259–264.

As TEX spread and became multi-lingual, the TEX world became regionalized to a considerable extent. This paper sums up the situation in Russia ca. 1995.

18. "Typographers' Inn", by Peter Flynn, **18**:4, pp. 242–245.

This was the first of now a dozen or so columns (still continuing) with this title, about typography as much TEX, that Peter Flynn has written for *TUG-boat*. In this first installment, Flynn addressed some concerns that *TUGboat* should be exclusively focused on TEX.

19. "Hey—It Works!", by Jeremy Gibbons, **19**:4, pp. 426–427.

This column contained hints and tricks for doing things with TEX with items from a variety of people. This column had been in the separate publication TEX and TUG News publication (http://tug.org/pubs.html), and continued for four more years in TUGboat. The complete collection is available online at http://tug.org/TUGboat/hiw.

Our random selections for volumes 18 and 19 are to columns that appeared over a period of time. A few other columns that have appeared more of less regularly for a period of time are: the already mentioned "Site Reports" and "Output Devices"; Victor Eijkhout's "Bag of [macro] Tricks"; Peter Wilson's "Glisterings"; "The Treasure Chest" with

several editors over the years; and Aditya Mahajan's "ConTFXt Basics for Users".

20. "*MathKit*: Alternatives to Computer Modern mathematics", by Alan Hoenig, **20**:3, pp. 282–289.

Much activity in the TEX world has been about using other fonts than Knuth's Computer Modern set. Math fonts have been of particular interest to the community. Author Hoenig developed a tool to allow people to create their own math fonts to match existing fonts, e.g., Baskerville. Like a number of other ideas suggested in *TUGboat*, this interesting approach didn't survive the years, as more fonts included math from the outset. (At http://mirror.ctan.org/info/Free_Math_Font_Survey is a nearly-comprehensive survey of free math fonts currently available for TeX.)

21. "Micro-typographic extensions to the TEX typesetting system", by Hàn Thế Thành, 21:4, pp. 317–434.

Hàn Thế Thành's success in making T_EX use PDF as a native output format, in parallel with DVI, was perhaps the most important step in sustaining T_EX in the world since the creation of I^AT_EX. It was fitting that TUGboat publish his full Ph.D. thesis on this work as an issue of TUGboat. Now, a decade later, his micro-typographic extensions are finally available and commonly used in all the major T_EX distributions.

22. "The status quo of the $\mathcal{N}_{\mathcal{T}}\mathcal{S}$ project", by Hans Hagen, **22**:1–2, pp. 58–66.

TUGboat has often included items on various distributions of TEX, TEX engines, and so forth. NTS was a reimplementation of TEX in Java using an object-oriented approach that had significant support from the TEX user groups. This thoughtful article by Hans Hagen reviews the project and draws a number of interesting conclusions. We speculate that this analysis by Hans helped him sort out his own thinking about opening up and extending TEX, with resulting major developments in LuaTEX, MetaPost, and ConTEXt.

23. "FarsiT_EX and the Iranian T_EX community", by Behdad Esfahbod and Roozbeh Pournader, **23**:1, pp. 41–45.

Among the more difficult directions TEX has been pushed has been the desire for type setting Persian and Arabic text.

24. Abstracts from *MAPS* 28, Fall 2002, **24**:2, pp. 283–285.

As the T_EX world expanded, many other T_EX user groups came into being, especially in Europe, some with their own journals, typically with all or most articles in the language of the particular coun-

try. Thus *TUGboat* was (and is) no longer the only publication with a focus on TEX and related things, and it was appropriate for *TUGboat* to include abstracts of articles from these other journals, such as these from *MAPS*, the journal of the Dutch group NTG (Nederlandstalige TEX Gebruikersgroep). A list of all such journals, past and extant, is at http://tug.org/pubs.html.

25. "The ℵ (Aleph) project", by Giuseppe Bilotta, **25**:1, pp. 105–107.

 ε -TeX was an extension of the TeX engine, primarily implemented by Peter Breitenlohner, that added right-to-left typesetting among a variety of other new features. Omega was an effort by John Plaice and Yannis Haralambous to move TeX more into a multi-cultural, multi-lingual world, including native support for Unicode. Aleph was an effort by Giuseppe Bilotta to produce a more stable version of Omega. The lasting impact of Omega and Aleph now resides in LuaTeX, while the ε -TeX extensions were incorporated into pdfTeX and are widely used and available.

26. "Using the RPM package manager for (IA)TEX packages", by Tristan Miller, **26**:1, pp. 17–28.

Managing the plethora of (IA)TEX packages and organizing distributions has long been an issue. Author Miller proposed an approach using the RPM tools standard in the GNU/Linux world.

27. TeX consulting and production services, **27**:2, pp. 285.

A regular feature in *TUGboat* is advertisements for companies and people wishing to provide TEX-related consulting and production services for a fee. Some ads have run for many years. One can find the original ads from each issue in the whole-issue PDFs on the *TUGboat* web site (http://tug.org/TUGboat). (The links from each issue's table-of-contents page go to the current list of advertisers, http://tug.org/consultants.html.)

28. "Installing ConTEXt expert fonts: Minion Pro", by Idris Samawi Hamid, 28:2, pp. 200–209.

Installing fonts and accessing them within TEX is another common theme through TUGboat's run. These days, there are at least four domains of such font installation articles: for plain TEX, standard LATEX, XATEX and LuaTEX with their OpenType support, and the ConTEXt system. Author Hamid is also deeply involved with the current Oriental TEX project for Arabic typesetting on which he works closely with Hans Hagen, creator of ConTEXt.

29. "Do we need a font system in TEX?", by Hans Hagen, **29**:1, pp. 28–33.

Hans Hagen and Taco Hoekwater, along with a cadre of other workers, are probably pushing the opening up of TEX (and in effect contemporary rewriting) harder than anyone else. Hans is fond of intriguing and provocative titles for his papers, which are typically thoughtful and full of new ideas and wise ways of thinking about things. As Hans says in the introductory paragraph of the present article, "... working on LuaTEX and ConTEXt MkIV... gives us much time and opportunity to explore new frontiers and reconsider existing ConTEXt features."

30. "Managing bibliographies with LATEX", by Lapo Mori, **30**:1, pp. 36–48.

Managing bibliographies well is one of TEX's strong suits, thanks to the BibTeX program developed by Oren Patashnik as part of the original Stanford TeX project. Several definitive (for the time) articles have appeared in *TUGboat* about managing bibliographies. Author Mori provided a reasonably comprehensive overview of the available methods.

31. TUG 2010 abstracts, **31**:2, pp. 248–249.

Not infrequently, presenters at TUG annual conferences decline to provide a written paper to be included in the *TUGboat* proceedings for the conference. In most of these instances, an abstract is included in its place. Fortunately, in recent years some TUG conferences have been videotaped (thanks to the efforts of Kaveh Bazargan at River Valley Technologies), and thus videos of some unpublished presentations are available on the Internet (http://river-valley.tv).

32. "IATEX at Distributed Proofreaders and the electronic preservation of mathematical literature at Project Gutenberg", by Andrew Hwang, **32**:1, pp. 32–38.

Our thirty-second and final randomly selected paper is a particularly appropriate "choice". It brings TUGboat's reporting about TFX and the TFX world back around to Knuth's original purpose for creating TeX: the creation of well-typeset books, especially math books (a capability that was being lost as early computer typesetting systems replaced pre-computer typesetting processes). In this article, Andrew Hwang describes a project to create, using LATFX, electronic versions of classic mathematics books published in the 19th and early 20th centuries. Knuth used T_FX to preserve the quality of pre-computer typesetting in newly-written math books. This part of the Distributed Proofreaders project is using TEX's high quality of typesetting in the electronic preservation of pre-computer mathematics books themselves.

4 Reflections

These random samples from 32 volumes of *TUGboat* suggest the breadth of the journal's coverage of the TEX world, while still leaving many specific categories unmentioned. http://tug.org/TUGboat/Contents points to a comprehensive list of topic areas used over the years by *TUGboat* to categorize its articles, and full author and title lists, as well as the issue-by-issue tables of contents. Scanning these gives an even greater sense of *TUGboat*'s diversity.

Of course TUGboat has served as a newsletter for the TeX Users Group (and the broader TeX world) about projects, events, and people. It has simultaneously provided tutorial material for various levels of T_FX practitioners, a forum for new ideas to be suggested and experiments to be described, and a place where major new developments in the T_FX world can be permanently documented. TUGboat articles have been peer-reviewed, but it has never been a journal of pure academic scholarship; it has served the typical role of a scientific or engineering journal in allowing participants in the field to learn about and build on (or create alternatives to) the work of others, thus extending to practice and education as well as the occasional theoretical article. Furthermore, it has played a role beyond TEX, regularly dealing with non-TeX issues of typography, design, document preparation and display.

Don Knuth has stated that computer science departments had to exist because there was a group of people who thought in a certain way which didn't fit within the confines of either traditional math or engineering departments. Perhaps because it evolved out of a creation by Knuth, *TUGboat* serves a similarly unique role for a collection of people who have interests in or needs for typography, font design, and a powerful typesetting system, and who in many cases want to go beyond the capabilities allowed to users of commercial typesetting systems. Users of T_EX and related technologies are a somewhat self-selecting community, and *TUGboat* exists and is edited to serve broadly the interests of that community.

Acknowledgments

Thanks to Barbara Beeton, who edited and proofread this article, as always. Also, over its 32 years literally thousands of people have contributed to TUGboat as authors, columnists, editors, members of the production team, creators of the web site, and so on. The TUGboat vessel is a significant community of its own within the worldwide TFX community.