What do today's newcomers want?

Jim Hefferon

Abstract

Social media gives us a chance to hear directly from today's newcomers about what they are working on and what hurdles they have.

1   Introduction

Helping users is a goal of the TeX Users Group. So insight into what today's beginners need is potentially useful. Here we shall argue that social media gives us a chance to listen in on beginners, and that Reddit is a good place to do that. Then we shall present some statistics from that site about what these newcomers say.

Social media has many aspects that are like the posting boards that people in the TeX and LaTeX community have used for years. One thing that is new is that interacting in this way has become mainstream so we can expect that many people will be comfortable speaking up there. This includes people who are newcomers to our community. Some of the things they discuss are surprising.

2   Where are they?

Reddit is a news aggregation and social web site, at `http://www.reddit.com`. Members submit content to the site such as links, images, or text posts. Posts are organized by subject into user-created subreddits, which cover a variety of topics. Site members vote these up or down and submissions with more votes are displayed at the top of their subreddit. In addition, over time new posts replace older ones.

There are many subreddits, more than a hundred thousand. They are named with the prefix `r/`. One is `r/latex`, at `www.reddit.com/r/latex`, for discussions about LaTeX and TeX in general. (There is also `r/tex` but it gets very little traffic.)

The `r/latex` page looks similar to other boards that TeX and LaTeX users have seen. A typical day has a list of posts, which are usually questions, for a visitor to read and for site members to vote on. They can also comment on the post, perhaps by answering the question.

2.1   Demographics

For us, the key point about Reddit is that it is the sixth most visited website in US and twenty-first in the world, with 542 million monthly visitors (as of March 2019). The site is predominantly in English: 54% of users are from the United States, then 8% from the UK, and 6% from Canada. The `r/latex` subreddit has 19,000 members. There are a small number of posts each day and the atmosphere is relaxed and polite.

Reddit attracts young people, as this comparison with the general US shows.

	18–29	30–49	50–64	65+
US	22%	34%	25%	19%
Reddit	64%	29%	6%	1%

Users average fifteen minutes per day on the site, usually lurking.

So the first argument for the presence of newcomers here is simply the clustering of a good number of people of the right age.

2.2   Architecture

The second argument is about the alternatives. For many *TUGboat* readers the first board that comes to mind is Stack Exchange, `tex.stackexchange.com`. Here too, the page contains a list of posts consisting of questions on which site members can vote, answer, or comment.

But the culture is very different. The *About* page says, "The goal ... is to be an exhaustive and curated library of detailed answers to every question related to TeX." It strives to be all business, "This site is all about getting answers. ... There's no chit-chat" and, "Questions that need improvement may be closed until someone fixes them, or just closed." The success of the site shows that this social engineering is very effective, indeed.

However for newcomers this can be discouraging. Being told that your query is closed can be off-putting, even though you are told this politely. Beginners may feel that they don't know enough to be able to state a precise question or to search for one in the past that is related to theirs. And, a person who lurks will see lots of stuff that is far beyond them.

So, an 18–29 year old new LaTeX-er who occasionally scans the contents of `r/latex` may find more of interest, and perhaps a more amenable atmosphere, than at Stack Exchange.

A word about two additional familiar English-language boards. The Usenet group `comp.text.tex` has been around for ages. But it doesn't attract newcomers because it has become mostly a CTAN announcement list. Another long time board, and a great resource, is `texhax`. But it is not as well-known as the others to newcomers and it is low traffic so it would not reward lurking.

Thus, a second reason that `r/latex` has a disproportionate number of beginners is that the history or engineering of other places may nudge those beginners over.

3 Results

I have been a regular on `r/latex` for more than a decade. Some of the things I have found there are surprising. To quantify them I collected some data.

3.1 Data

I grabbed one thousand posts, covering `r/latex` from 2018-Nov-10 through 2019-July-08, and characterized each post in a few ways. There is a good deal of judgment involved in these characterizations but despite this noise, the numbers tell an interesting story.

3.2 Findings

First, many people make clear that they are beginners, often simply by saying it. I counted 265 authors as beginners, 29 as experienced, and 703 posts were not clear. (The numbers do not add to 1000 because of some spam.) That is, many posts begin like, "I'm a total noob ...," supporting the earlier analysis that this site attracts this group.

I suspected that many of today's beginners do not install TeX on their computer but rather start at an online site such as *Overleaf* or *CoCalc*, so I also characterized the posts by computing platform. As TeX and LaTeX are in many ways platform-independent, the great majority of posts, 817, did not name the platform. Of the remaining, 17 were using GNU/Linux, 22 were Macintosh, 63 were Windows, and the largest number, 78, were online.

The biggest challenge was to characterize the post's subject. In some cases there was more than one subject and I judged what was the main one. Here are the numbers; I'll expand on the keys below.

Key	Number
wrapper	92
biblio	81
graphics creation	78
resume	31
thesis	16
article	14
pandoc	13
book	13
presentation	12
classwork	12
unknown	169
other	466

The "wrapper" refers to editors or other creation environments, from `vim` to TeXworks to *Overleaf.* Thus, close to ten percent of all posts are from folks, often newcomers to LaTeX, who say they are struggling with an inability to do something that is, in some sense, not TeX or LaTeX.

Often beginners have trouble distinguishing the wrapper from TeX so they may ask, "how to get LaTeX to find and replace?" They are often not sure whether they are using `pdflatex`, or `xelatex`, etc., because that is hidden in a submenu. This is an example where the sophisticated systems available to beginners today can at least to some extent prevent them from understanding what they need to do to move forward.

Number two on the most-asked list is bibliographies. This is an area where we can perhaps do a better job helping people. For instance, on `r/latex` I often urge people to get started with LaTeX by reading the *Not So Short Introduction* [3] but I note that this document has less than two pages on this topic. Also perhaps of use would be a web page like the *LaTeX Font Catalogue* [2], but instead containing a selection of bibliography styles.

The next most asked topic is another one that many experienced users also have trouble with, creating graphics. This usually takes the form of, "In Ti*k*Z, how do I ... ?"

After that, the next item is a surprise, at least for me: today many people, as an early LaTeX encounter, write a resume. One factor may be that because resumes have rather complicated formatting, they could lead to a disproportionate number of posts.

Following those are subjects that many readers may have expected. This includes questions about thesis and article styles, styles for books, and questions about presentations using Beamer. It also includes using TeX and LaTeX for class notes or homework. (Pandoc is a program to convert files among markup languages.)

The "other" category is large but also scattered. Subjects that appeared include fonts, figures, and tables. Also there are links to blog posts about LaTeX topics.

3.3 Observations

I will close with a few comments comparing newcomers on `r/latex` with those in the past.

First, delightfully, missing from the board discussion are many things that in the past gave newcomers trouble. There are not many questions about installing an entire system, rarely are people stuck on the "Hello World" problem of getting that first document out the other end, and no one ever asks about tuning font parameters for a printer.

Surprising to me is that many posters introduce themselves as undergrads. No longer is the first encounter with TeX and LaTeX, and our community, restricted to professionals and graduate students.

Perhaps related to the prior point is that often posters acknowledge a sense that TeX does the best documents. So there is widespread awareness among this young group of the power of TeX and LaTeX and friends.

These beginners often ask for a "template," by which they mean a file into which they can drop their content. For these, people are often pointed to *Overleaf*.

Finally, related to that, newcomers have typically not looked on CTAN, or even heard of CTAN. And if they have gone there, they can be stymied by a paradox of choice.

3.4   Just ask them

As a follow-up to the survey, I posted: *If you are a beginner then what would help you in TeX and LaTeX? . . . What do you find to be the biggest hurdle?* [1] There were about a dozen responses, which make interesting reading (see the link in the citation). These are largely in line with the description above so that respondents described problems with packages, including finding a suitable one or understanding interactions and conflicts among packages. And, they expressed struggling with Ti*k*Z.

4   Summary

There are reasons to suppose that social media can help us understand what beginners today are working on and struggling with. Analysis shows that some real stumbling blocks from the past are not present today and tells us a little about what does give newcomers trouble. It also shows that today's beginners are younger than has been traditional.

References

[1] JimH10. Are you a newcomer to LaTeX? `https://www.reddit.com/r/LaTeX/ comments/ccc93c/are_you_a_newcomer_ to_latex/`, 2019. [Online; accessed 2019-August-13].

[2] Palle Jørgensen. The LaTeX Font Catalogue. `https://tug.org/FontCatalogue/`, 2019.

[3] Tobias Oetiker. The Not So Short Introduction to LaTeX 2$_\varepsilon$. `https://ctan.org/pkg/lshort`, 2019.

⋄ Jim Hefferon
   Saint Michael's College
   `jhefferon (at) smcvt dot edu`