Proposal: Host a Discourse site for DITA users (v0.2)

Contents
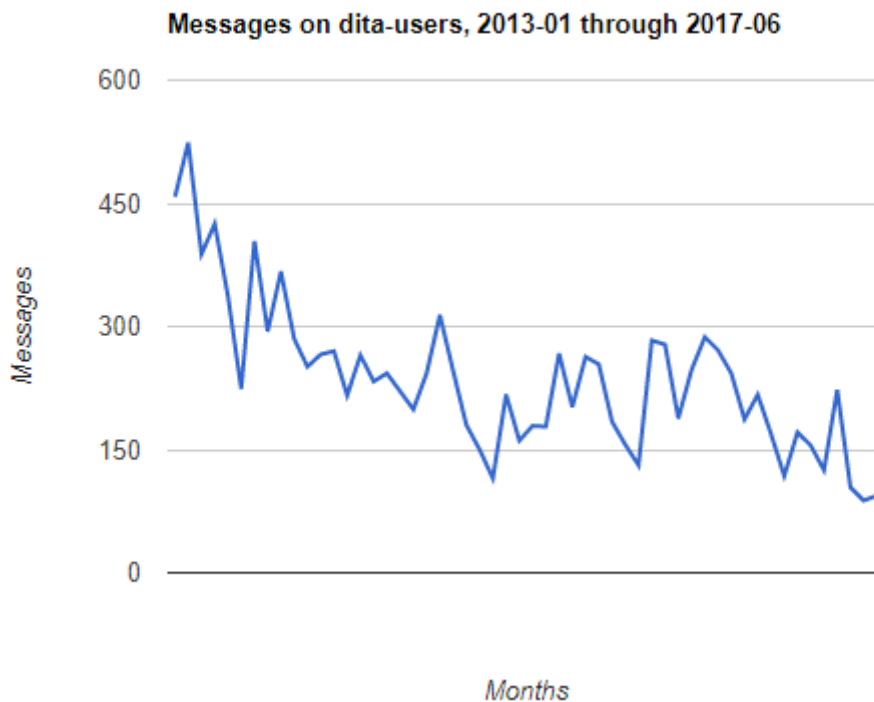
Summary

The Yahoo! Groups web interface is frustrating, and participation in the DITA community forum hosted there is losing steam. A replacement that facilitates communication among the DITA community is needed. Discourse is an open-source internet forum and mailing list platform backed by a healthy, for-profit company, and a healthy community. Hosting a Discourse site requires active administrative and financial support, both of which are forthcoming. The risks of hosting seem manageable, and the cost is reasonable. Using Google Groups is briefly discussed as an alternative.

Background

The dita-users group on Yahoo! Groups has been the community forum for DITA users since 2004. The Yahoo! Groups service is free and reliable but limited or just plain frustrating. A 2013 effort to move from Yahoo! Groups to a forum on Stack Exchange did not succeed, and more recently, other options were considered but rejected as too expensive. For unknown reasons, participation in the forum has dropped since 2013, as shown in the following chart.

Messages on dita-users, 2013-01 through 2017-06

Disposition: / Status:
New paragraph:

Finally, should Lightweight DITA attract throngs of new users, it'd be nice to have a modern forum for them to join.

Goals

The overall goal of replacing Yahoo! Groups is to improve communication among the users, implementers, and developers of DITA technologies. The following list highlights additional goals that support the long-term mission of the DITA Adoption TC; the list is numbered solely to facilitate discussion:

1. The software license of the replacement should be open-source.

    Open-source projects do not rely on one vendor for changes.
2. The replacement should support data export.

    A sweeter solution could come along.
3. Additional goals TBD

Discourse software

This proposal advocates installing and hosting a dedicated Discourse discussion site. Discourse is an open-source internet forum and mailing list platform maintained by Civilized Discourse Construction Kit, Inc. Significantly, the company's strategy is the "freemium" business model: The open-source version is actively developed and maintained, with extra features reserved for paying customers.

History

Civilized Discourse Construction Kit, Inc., (CDCK) is a privately-held company co-founded by Jeff Atwood and Sam Saffron. Jeff writes a prominent blog about programming, and he founded Stack Exchange in 2008. In 2012, Jeff began prototyping an internet discussion forum platform, which was released in February 2013. He decided to do so because he found the available options feature-poor, and he realized that Stack Exchange was for Q&A only, not discussions. Currently, the company boasts a healthy list of prominent customers.

Features

Discourse is an internet discussion forum platform with email support. That is, the web interface is the primary means of creating and consuming information, rather than a layer on top of email. The features are nicely documented on the Discourse web site. Highlights include (order not significant):

1. Spam protection and community moderation
2. Trusted members can summarize or condense topics
3. Single sign-on and "social" logins (Google, GitHub, LinkedIn, and so on)
4. Comprehensive API and plug-in architecture
5. Free iOS and Android apps
6. Private spaces (for vendors?)
7. Reply to posts via email
8. Translations for 25 languages

The following list provides a few examples of open-source users; the appearance of the last example includes some customization:

*   Let's Encrypt
*   GitLab
*   Fair Phone

The Discourse community forum includes a healthy marketplace of paid installation, migration, development, and administration services. The hosting area features generous help from experienced users, including CDCK employees, for do-it-yourself users.

Hosting

Disposition: / Status:
Edits in this paragraph:

CDCK makes its money by hosting Discourse sites. The DITA Adoption TC very likely qualifies for a 50% discount, but the basic plan plus TLS/SSL costs $120/month. Either way, the annual cost is prohibitive without firm commitments of vendor support. Hosting one's own Discourse site is the alternative. Importantly, a self-hosted site can be moved to CDCK hosting very quickly, with most of the time spent waiting for the new IP address to propagate through the internet.

Installing Discourse requires the following resources:

1. A discrete Linux server with a minimum of 1GB of RAM.

   Ubuntu LTS is a common choice, but any 64-bit Linux distribution with kernel version 3.10+ is fine.
2. An SMTP (email) server.

   Only the foolhardy administer their own SMTP servers. Discourse recommends several that play nice with their platform.
3. **Disposition: / Status:**
   Added uploads to this item

   Online storage for backups and uploads.

   Discourse includes an integration with Amazon S3 for backups. (A less expensive option is Backblaze B2.) Low-cost VPS hosts typically provide relatively little storage space, so image and file uploads are typically configured to move to S3 automatically.

Also, hosting requires Linux system administration skills. Discourse provides a well-tested installation script and procedure, so the initial setup is not difficult. However, DIY users are responsible for securing and maintaining the operating system, and for updating the Discourse software.

Disposition: / Status:
New paragraphs

Typically, Discourse sites are subdomains of another site (because Discourse requires its own server host). Configuring a subdomain is a common, straightforward administration task.

It is possible, though uncommon and somewhat indirect, to pay in advance for the required resources:

1. Most hosting providers accept payment in advance for their VPS instances.
2. For an Amazon S3 account, one can pay more than is billed and then have subsequent usage charges subtracted from the credit balance.
3. For email providers, the answer is unknown.

Paying in advance for a self-hosting strategy should be weighed against the potential for vendor support of CDCK hosting. The payments may not be recoverable.

Moderation

Disposition: / Status:
New topic

The moderators of dita-users—Don Day, Kristen James Eberlein, Michael Priestley, and Keith Schengili-Roberts—will be invited to step into the same roles on the Discourse forum. Discourse tracks user participation and can grant

"trusted" status to users, which includes some moderation privileges. This may or may not be desirable, but does provide an organic way to develop leadership and community standards.

When the DITA Adoption TC reaches consensus about this proposal, it should be shared with the dita-users moderators, to solicit their feedback. Then, this proposal can be posted on the dita-users list.

Risks

Issue	Risk	Mitigation
The annual hosting costs can't be covered indefinitely.	Low	One vendor has volunteered to contribute already, and individual users may also be inspired to do so, too. Vendors may be interested in having their own private areas on the site, for a fee.
The history in Yahoo! Groups will be lost.	Certain	The history can be mined for best practices. The history is already at risk (Yahoo! could discontinue the service).
No qualified operating system or Discourse software administrator is willing to succeed the first volunteer.	Medium	Document the required procedures, so that a replacement can step in with confidence. Hire a Discourse community member (from their marketplace) to provide the services.
The forum will be less responsive for users outside the U.S.	High	Use a content delivery network (CDN). Unfortunately, they're pricey. Using an east coast hosting facility helps European users somewhat.

Budget

If we could sign up some vendors for private areas, we would most likely be able to afford one of CDCK's hosting packages. Otherwise...

Resource	Cost	Comments
Virtual private server (VPS)	$10/month	$10/month is a common price for a VPS on a hosting service like Digital Ocean.
SMTP server (15,000 emails per month)	$2.50/month	The costs vary quite a bit by provider. The $2.50 cost is for Mailgun; their pricing is based on volume.
Storage for backups	$2.50/month (a guess)	This cost is *profoundly* difficult to estimate. It all depends on the size of each backup, which will grow over time.
Administration services	$0 to $100 per task	The author is confident but just one person.

The known costs are $120 to approximately $180, and the unknown costs are unknown.

Alternative: Google Groups

With deep gratitude for the person or persons who wrote the article about Google Groups and shallow interest in the equivalent original research, the author presents the following introductory quote from Wikipedia:

Google Groups...provides discussion groups for people sharing common interests [and]...also provides a gateway to Usenet newsgroups via a shared user interface....Google Groups offers at least two kinds of discussion group; in

both cases users can participate in threaded conversations, either through a web interface or by e-mail. The first kind are forums specific to Google Groups which are inaccessible by NNTP [Network News Transfer Protocol] and act more like mailing lists. The second kind are Usenet groups, for which Google Groups acts as gateway and unofficial archive...

The first kind of discussion group that Google Groups provides is similar to what Yahoo! Groups provides, and if we're lucky, there's no interest in the second kind, because Usenet postings are unmoderated and a serious usability problem for casual users. (For example, read this post in comp.lang.python.) Significantly, both kinds of discussion groups are designed around message exchange protocols (NNTP resembles SMTP). The web interface is a layer on top of collections of messages.

In the Google Groups web interface, messages are presented chronologically, and message threading is supported. The web interface provides a nice set of filtering criteria for displaying messages of interest, and searching in a group is excellent. There is no direct support for files or photos; presumably, one's own Google Photos or Google Drive are used and the contents linked, if one has a Google account. There is no support for calendars, or for customized landing pages. One simply arrives at the list of messages.

Google Groups provides several management options, including message moderation, two levels of categories, group-wide message controls, and selective banning. Overall, the controls are more sophisticated than those offered by Yahoo! Groups. However, neither service provides an API for exporting content.

The Google Groups service is free and the risk of losing it is very low. Google does shut down services when they no longer suit their purposes; the circumstances that would precipitate closing Groups are difficult to imagine.