CTAN goes multi-lingual: Additional language support for the Web portal

Gerd Neugebauer

Abstract

\TeX{} is used for many languages: support in the \TeX{} engines and macro packages is abundant, but the CTAN portal has been available in English only. Now we are conducting an experiment to provide the Web presentation in German as well.

1 Introduction

Modern Web frameworks are equipped with means for internationalizing the presentation. We are currently attempting to make use of these means for the CTAN portal, https://www.ctan.org.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{The cover page in German}
\end{figure}

Whenever you visit the portal, the language is automatically selected for you. This magic happens in a negotiation between the browser and the server in the background. In the browser you can configure your preferred languages (cf. figures 2–4).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Language configuration in Firefox}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Language configuration in Chrome}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Language configuration in IE}
\end{figure}

When you request a page, the browser passes these language preferences to the server, where the list is compared with the list of supported languages. The best fit — or the default fallback — is then chosen. Currently, the CTAN portal supports English and German. Thus it is usually sufficient for you to navigate to the CTAN portal to see your preferred language from these alternatives.

You can also explicitly choose the language on the settings page by clicking on the appropriate flag (figure 5). This selection is valid only for the current session. When you return later, this setting is forgotten.

2 Features

Internationalization includes the localized presentation of several types of content:

- The page frame which includes header and footer as well as the static parts of the dynamic pages
- The page content for the “static” pages
- The dynamic content from the \TeX{} archive
The dynamic content from the Catalogue

The dynamic content from the search index

For the current internationalization experiment, only the frame and some of the static pages are offered in German, as a starting point. It should primarily help to make it more comfortable for visitors in their first steps.

Some pages have intentionally not been translated. For instance, the upload page (https://www.ctan.org/upload) is provided in English only since the communication language with the CTAN team is English. Thus, an uploader should not be encouraged to try another language since this might not be understood on our side.

3 Limitations

The major limitations are derived from missing data in the back-end. The files in the \TeX{} archive are provided by the numerous authors. The language is whatever the author provides — for instance in the package documentation.

The \TeX{} catalogue is mainly in English. It is partially prepared to carry texts in different languages. For instance the descriptions are to a small degree already present in languages other than English. The CTAN team is presently not able to provide translations, due to limited resources and skills.

If no appropriate localized text is available the English version is used. This can be seen in figure 1. Here the text for the topic teaser is in English while the other parts of the page are in German.

Another limitation applies to the search. The search is currently language-agnostic. It should in the future be based on the language settings, in that entries with the proper language should be ranked higher than non-matching entries. Maybe entries for languages not understood by the user should be suppressed completely. But this already belongs into the next section on “visions”.

4 Visions

In the future we could consider to develop the internationalization of the CTAN portal in several directions. First, of course there are more languages to be supported. Right now the language-specific texts are contained in 31 files. These files are either mapping files which map keys to language-specific texts, or complete pages which are present in separate incarnations.

To support a new language is initially a matter of providing these 31 files and adding a configuration option for the new language. However, this isn’t enough. The portal is not static. The content changes over time. Thus we would need a commitment that we have volunteers for the new supported language to guarantee continuity.

For the support of even more languages the existing administration interface could be extended. Then maintenance of the different languages could be performed via a Web interface by different persons.

The other side has already been mentioned. The \TeX{} catalogue has to be augmented with new language-specific texts. The same applies here. The support has to be guaranteed for future changes. Thus volunteers for a long-term engagement would be required.

We must also reconsider the processes in the CTAN team. They are currently not designed for parallel processing of a change for a single package. These processes work well for the small team which is currently active. The more languages that are supported, the more people have to work on one change. Unfortunately this seems to be beyond the capabilities of the CTAN team right now.

5 Epilogue

I hope that the experiment with the German language for the CTAN portal succeeds. Feedback in any direction is welcome. For discussions, please consider using the mailing list ctanweb@dante.de. You can subscribe via https://lists.dante.de/mailman/listinfo/ctanweb.

Keep on \TeX{}ing — in many languages.

Gerd Neugebauer
Im Lerchelsböh 5
64521 Groß-Gerau (Germany)
gene@gerd-neugebauer.de
www.gerd-neugebauer.de

CTAN goes multi-lingual: Additional language support for the Web portal