User Support for the Digital Humanities

Heidemarie Sambale   **Hanna Hedeland**   **Tommi Antero Pirinen**

Hamburg Centre for Language Corpora (HZSK)
Universität Hamburg, Germany
`{firstname.lastname}@uni-hamburg.de`

Abstract

In this article, we describe a user support solution for the digital humanities. As a case study, we show the development of the CLARIN-D Helpdesk from 2013 into the current support solution that has been extended for several other CLARIN-related software and projects and the DARIAH-ERIC. Furthermore, we describe a way towards a common support platform for CLARIAH-DE, which is currently in the final phase. We hope to further expand the help desk in the following years in order to act as a hub for user support and a central knowledge resource for the digital humanities not only in the German, but also in the European area and perhaps at some point worldwide.

1 Introduction

For both the ongoing digitalisation of humanities research in general and the CLARIN infrastructure in particular, the non-technical aspects of adequate training and user support are crucial for acceptance and involvement from the research communities. Many humanities researchers come from a rather non-technical background, and the use of digital tools and resources has not yet entered the curriculum to an appropriate extent. Researchers thus face various problems when confronted with tools and platforms for digital humanities research, many of which might not be predictable to the developers. Improving the usability of such kinds of tools and providing comprehensive documentation is undoubtedly very important, but in the end there is no replacement for a reliable help desk to assist users when they for various reasons struggle with digital resources, tools and services or need qualified advice in methodological questions.

Since digital humanities is multidisciplinary, the user support becomes a very central and important resource for exchanging information and experience, and for gathering expertise from various contexts. Apart from providing users with reliable support, the direct interaction with the users also provides valuable input about the users and their behaviour and on existing problems with the tools and platforms for infrastructure providers and developers. In our paper, we describe the development of a comprehensive resource based on a sustainable platform with re-usable workflows, for which we have also developed various strategies for scalability.

The rest of the article is structured as follows: In section 2 we describe the background of our approach to user support and our help desk. In section 3 we present the current developments. In section 4 we introduce the scalability strategies and in section 5 we describe the technical implementation of various

challenges regarding management and scalability of the help desk. Finally, in section 6, we describe our plans in the future, based on the current development.

2 The CLARIN-D Helpdesk

The CLARIN-D Helpdesk (Lehmberg, 2014; Lehmberg, 2015) was first launched in 2013 to provide the necessary user support for the emerging CLARIN-D infrastructure. After a thorough review of current ticketing systems, the open source OTRS platform[1] was considered the best solution in order to meet the needs of the infrastructure's users and developers. In OTRS, the support requests are managed by using support *tickets*. A ticket represents and documents the entire conversation between agent and customer chronologically. The support is provided by *agent* users that are active within one or more support areas. The tickets are organised into *queues*, which represent the various support areas. Figure 1 shows the arrangement of the queues: every centre (e.g. EKUT, HZSK) has its own queue with several subqueues for every section they offer support.
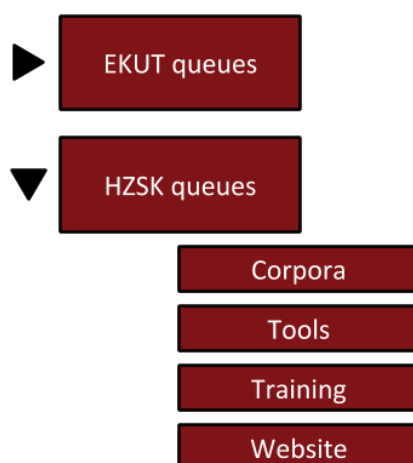

Figure 1: Structure of queues representing support areas in the CLARIN-D Helpdesk.

Two agent types can be assigned to each ticket, responsible and owner. The owner handles the customer requests, whereas the responsible monitors the overall progress. While the owner of the ticket can change in the course of different working stages, the responsible remains the same agent. The information about which agent is responsible and which one is the owner of the ticket is visible to other agents. Apart from the communication with the customer, it is possible to communicate internally in a structured and documented manner within the ticket conversation in order to find an appropriate answer to an inquiry collaboratively.

This technical infrastructure and the related concepts and workflows comprise the relevant functionality required to reliably answer and document incoming queries. The use of a ticketing system solves most of the issues related to providing user support via email, e.g.

- relevant information (such as previous answers, templates etc.) is not accessible for other agents due to the use of private mail accounts,
- duplicate answers are sent when a common mail account is used by several agents and coordination fails,
- comprehensive documentation of every working step is not created,
- workflows cannot be standardised and simplified, e. g. by the use of templates and automatic answers
- long response time owing to holiday or illness.

---

[1] https://otrs.com/

In any case, without some kind of ticketing system, status and responsibilities for inquiries need to be managed individually. By using the OTRS, the support progress is clear to other agents, so that another agent can easily take over if the current owner of the ticket is absent.

As shown in Figure 2, tickets arrive to the help desk from multiple sources. Depending on the origins, the tickets may need further sorting in the queues and assigning to the agents. The ticketing system records metadata about the issues, such as first response times and closing times that are used to gather statistics relevant for the goal of providing efficient user support. The textual content of the help desk, i.e. questions and answers, can also be searched as a semi-structured knowledge base, or used as the basis for edited FAQ articles, which are also distributed from the help desk.
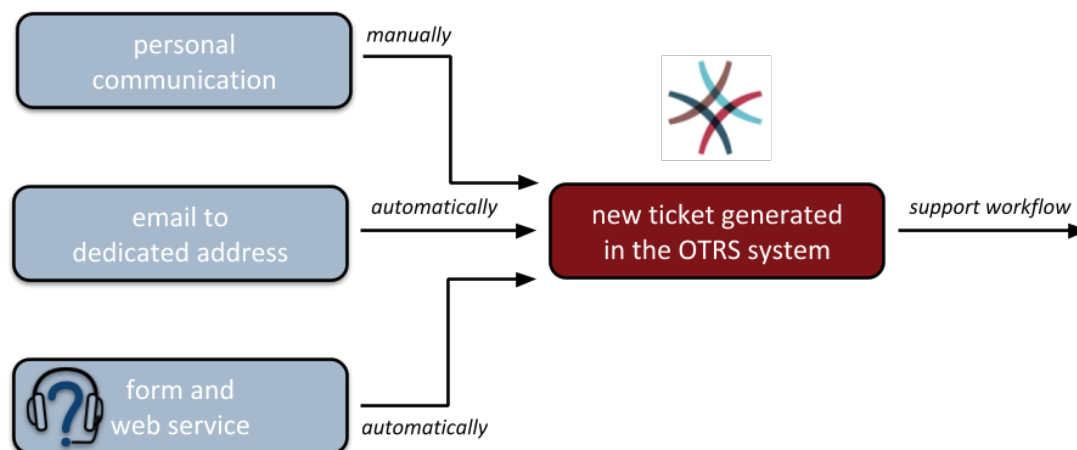

Figure 2: Tickets can be generated in various ways, allowing for pre-sorting and delegation through queue-specific addresses or web service parameters.

Since the CLARIN-D infrastructure is distributed and centre-based, the queue structure of the CLARIN-D Helpdesk models the centres and the services they provide, and the help desk is used to distribute the support requests to the relevant experts. Most of the tickets are automatically sorted and delegated using parameters from email addresses, web forms or keywords. Not automatically sorted tickets are manually assigned to the right queue by the first line support agents, as shown in Figure 3. The first line support is carried out by experienced student assistants, who receive specific training. The training involves monitoring and managing tickets, answering common and general questions, and delegating incoming inquiries that have not already been assigned to a queue automatically. Administrators and first line support have a complete overview of all queues and agents whereas most expert agents of the second line support will only be able to see tickets and queues relevant to them. The experts comprising the second line support are researchers and developers of the participating centres and projects. Before an inquiry is successfully closed, several experts might contribute with their respective expertise on a complex matter, requiring the ticket to be moved across queues and reassigned several times. According to our quality guidelines, at least 95 % of the tickets should be answered within two days. This target was not only met, the percentage of tickets answered within two days has been increasing steadily in recent years.
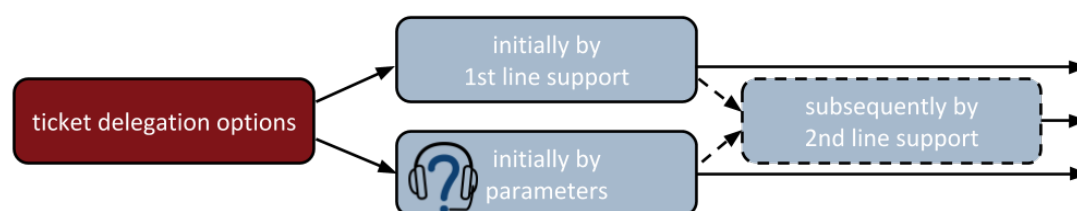

Figure 3: Tickets are sorted by queues and assigned to the right persons automatically or manually.

2.1 Inquiries at the HZSK

There is a huge range of the inquiries content arriving the CLARIN-D Helpdesk and also by looking only on inquiries for the Hamburg Centre for Language Corpora. The type of inquiries at the HZSK can be categorised into four main areas:

- **Corpora** – mainly requests to access one of the corpora in our repository[2], but also inquiries about the handling and terms of use of the corpora
- **Services** – this section contains many different types of questions, mainly from researchers in different projects, about data management, data format, best practice, hosting corpora, corpus creation and corpus curation
- **Tools** – inquiries regarding the EXMARaLDA[3] software, mainly questions about handling, issues and requests for future developments
- **Training** – registrations or questions about trainings, organised and carried out by the HZSK

It is not always easy to draw the line between these four sections. Especially the area service has intersections with every other area. Nevertheless, the categorisation helps to get an overview of the distribution of the inquiries. As we can see in Figure 4, more than 80 % of the 236 incoming inquiries at the HZSK in 2019 belong to the sections tools and corpora. Answers and further communication with the customer are not included in that number, so it is also important to note that this does not reflect the invested time. While inquiries in the sections Corpora, Tools and Training can often be solved within a few messages, the questions in section Service require intensive support over a period of time.
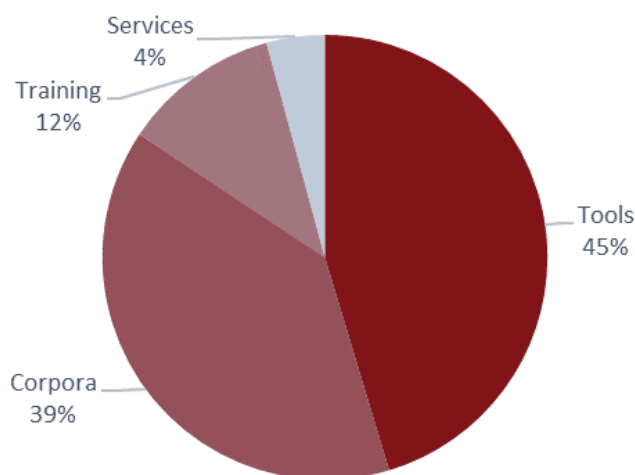

Figure 4: Distribution of inquiries at the HZSK in 2019.

2.2 Distributed Support within CLARIN-D

For complex inquiries, the contribution of several centres with different areas of expertise is useful. An example of an area often requiring coordination between multiple experts and centres is support for projects interested in creating digital spoken language resources using transcription software. The creation and analysis of spoken language resources usually requires a great deal of technical and methodological support, both due to the mostly non-technical background of the users, and due to the highly complex nature of the task at hand. Spoken corpora comprise various interrelated data types and file formats and complex metadata valid across the corpus' components. For the creation of transcripts several software systems exist that are specialised for specific scenarios. Apart from expertise regarding

---

[2]https://corpora.uni-hamburg.de/hzsk/en/repository-search
[3]https://exmaralda.org

certain tools, centres also provide expertise according to the raw data, e.g. depending on the language or modality of the resource.

Based on the areas of expertise depicted in Figure 5, a simple technical question regarding the transcription software EXMARaLDA might be initially answered by the first line support at the HZSK centre. If the reply of the customer contains a more complex question regarding the software's scope of use, the request would then be forwarded to the second line support, who might find that this user should rather try the tools provided by another centre (e.g. IDS or BAS) and forward the ticket accordingly.
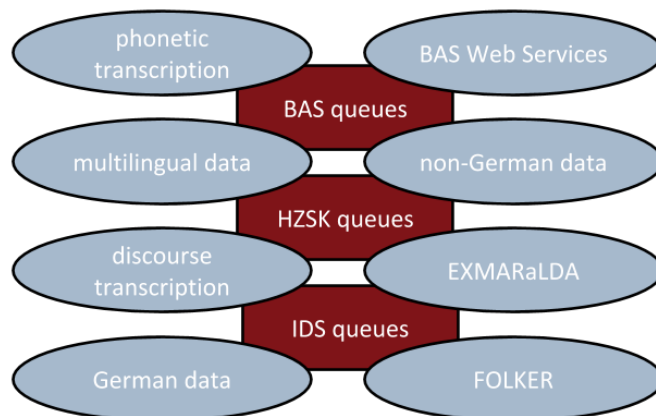

Figure 5: Professional support within the complex area of spoken language resources requires coordination and cooperation between several centres providing complementary expertise.

3 Support beyond CLARIN-D

Beyond CLARIN-D, the CLARIN-D Helpdesk is used to provide support on the European CLARIN level; for the CLARIN Virtual Language Observatory (VLO)[4] user feedback and outreach - including both generic feedback about user experience or reports of erroneous metadata - and for the CLARIN Federated Content Search Aggregator (FCS)[5]. For the VLO, the user feedback is divided into tickets regarding the VLO application and tickets regarding the metadata and the resources it describes. The tickets regarding the VLO application are handled by the VLO developers and the metadata related tickets can be handled by first line support or by the SCCTC Metadata Curation Taskforce as a part of their quality assurance work.

Apart from tools and services directly integrated into CLARIN, some support workflows for related tools from other contexts have also been successfully integrated into the CLARIN-D Helpdesk. This allows for interaction between developers and users of these tools across expertise and support areas and thus for a wider outreach. This could be a first step for these partners to become a part of the emerging Knowledge Sharing Infrastructure[6] even though they are not certified CLARIN centres (yet).

4 CLARIAH-DE - joining forces

In the course of merging CLARIN-D and DARIAH-DE into CLARIAH-DE, the support was also merged into one comprehensive help desk based on the CLARIN-D Helpdesk in November 2019. The integration of the DARIAH-DE support differed from the other extensions in the help desk, hence not only a new module of queues and agents needed to be added, but a fusion has to be made. Apart from uniform external presentation and communication as well as user information, the challenge is to bring the two different internal organisation structures together.

As previously mentioned, CLARIN-D is centre-based. Therefore, in OTRS every CLARIN-D Centre has its own queue, with respective agents for each queue. Some centres have a wide range of tasks or

---

[4]https://vlo.clarin.eu/
[5]https://www.clarin.eu/content/content-search
[6]https://www.clarin.eu/content/knowledge-centres

multiple software/applications, therefore it is necessary for them to organise the centre queue into several subqueues, as shown in Figure 1. In CLARIN-D, the agents are mainly managed in groups, whereas every centre has its own group. This means that despite the subdivision within the individual centre queues, every agent is part of one centre group and can work with every ticket in that group (see Figure 6). This organisation is possible because there are only a few agents in each centre, belonging to the respective groups.
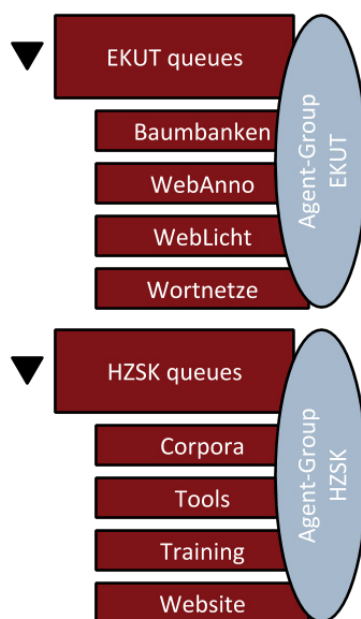

Figure 6: Queues and agent responsibility in CLARIN-D

In contrast, DARIAH-DE is much more centralised which leads to agents with varying overarching responsibilities for different queues. Figure 7 shows a section of the DARIAH-DE queues and responsibilities schematically.
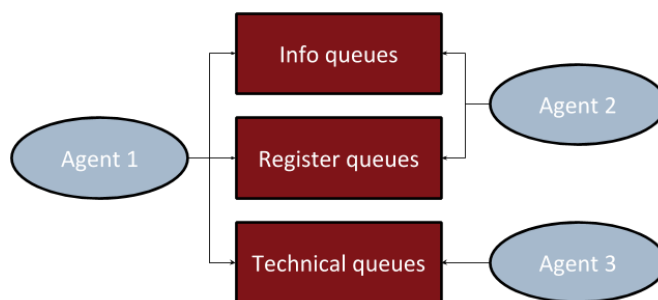

Figure 7: Queues and agent responsibility in DARIAH-DE

If we compare the help desk of CLARIN-D (Figure 6) and DARIAH-DE (Figure 7) there is a structural difference regarding responsibilities of agents within queues. In recent years, both help desk structures had developed their own workflows based on the respective system and both structures have proven their value in user support. So the goal was to keep up both structures. To achieve this, the best way to handle the increased number of agents with its partly overlapping responsibilities was to extend the role system in OTRS. Thereby it was possible to integrate the two different structures within one OTRS system in an elegant way.

The setting options for generating statistics have also been expanded. It is now possible for certain agents without admin status to generate statistics on the tickets in their queues. The use of general statistic templates such as 'tickets created last month/quarter/half/year' makes it possible to use them regularly.

In addition, special statistic templates can also be created to meet individual needs. Thanks to the restructuring, it is now very easy to expand the group of people who want to generate individual statistics. We can now offer this feature to all previous and future users of the CLARIAH-DE Helpdesk.

The main part of the migration has been completed successfully, but the CLARIAH-DE Helpdesk will be further optimized and adapted to the needs of the users.

5 Customizing ticket handling for increased flexibility

To be able to integrate support workflows for different tools and services from different contexts, we aim to make the help desk platform and all related components highly flexible. The support system is provided to the end users typically as a part of a website or web application. For seamless integration of the support, there are various technologies that can be used for a certain service, tool or other support area, usually the implementation will be carried out by the respective website managers. The most basic model of integration can be achieved with a support email address that forwards to the OTRS system. The CLARIN-D Helpdesk also has a web API based on SOAP that can be used to create and assign support requests. We provide a reference implementation of an HTML form for this API in our public GitHub repository that can be customised according to the requirements for the various support areas.[7] The integration of DARIAH-ERIC has been built on top of this API for WordPress and is likewise available at their GitHub repository.[8] Further information about the implementation of DARIAH-ERIC are available in Raciti et al.(Raciti et al., 2019). Each of the ticket creation workflows in the schematic presentation in Figure 2 come with benefits and costs. Furthermore, some of the technical improvements have been made to enhance the automatic ticket assignment in order to minimise the overhead of the first line support in dealing with spam and to categorise more tickets automatically, e.g. by analysing and state more precise keywords in the filter options. By this, the first line support can use their time more efficiently in actual user support tasks. To retain this in future, the integration of further support areas should be kept as simple as possible.

6 Outlook

We have described a scalable help desk system with sustainable workflows introduced for CLARIN-D but already used far beyond its original designation. In the future, we hope to extend the help desk and the areas of support catered for even further. On the European level, we have already integrated support for several CLARIN services and the DARIAH-ERIC Helpdesk into our system, hence we are currently operating a national and international CLARIAH-DE Helpdesk. While in the past the administration of the agents and their authorisations could largely be managed via groups, the merge and the help desk expansion required the development of an extensive role system, which allows for additional flexibility when integrating further organisational units. Apart from the challenge ahead in optimizing the merged CLARIN-D and DARIAH-DE user support, we are looking forward to integrating complementary support areas and workflows and to further enhancing the usage of the help desk as a central knowledge resource for the digital humanities.

References

Timm Lehmberg. 2014. The CLARIN-D Help Desk. In *Papers, Posters and Demos CAC2014*. CLARIN ERIC: Utrecht, The Netherlands.

Timm Lehmberg. 2015. Wissenstransfer und Wissensressourcen: Support und Helpdesk in den Digital Humanities. In *FORGE*, pages 25 – 27.

Marco Raciti, Yoann Moranville, Raisa Barthauer, Stefan Buddenbohm, and Dorian Seillier. 2019. D5.4 - Implementation of a centralized helpdesk and marketplace mockup. Research report, DARIAH, March.

---

[7]https://github.com/hzsk/clarind-helpdesk
[8]https://github.com/DARIAH-ERIC/contact-helpdesk