Data-Driven Governance in Crises: Topic Modelling for the Identification of Refugee Needs

Sprenkamp, Kilian ; Zavolokina, Liudmila ; Angst, Mario ; Dolata, Mateusz

DOI: https://doi.org/10.1145/3598469.3598470

Originally published at:
Sprenkamp, Kilian; Zavolokina, Liudmila; Angst, Mario; Dolata, Mateusz (2023). Data-Driven Governance in Crises: Topic Modelling for the Identification of Refugee Needs. In: 24th Annual International Conference on Digital Government Research, Gdansk, Poland, 11 July 2023 - 14 July 2023. ACM Digital library, 1-11.
DOI: https://doi.org/10.1145/3598469.3598470

Data-Driven Governance in Crises: Topic Modelling for the Identification of Refugee Needs

Kilian Sprenkamp
University of Zurich
kilian.sprenkamp@uzh.ch

Liudmila Zavolokina
University of Zurich
zavolokina@ifi.uzh.ch

Mario Angst
University of Zurich
mario.angst@dsi.uzh.ch

Mateusz Dolata
University of Zurich
dolata@ifi.uzh.ch

ABSTRACT

The war in Ukraine and the following refugee crisis have recently again highlighted the need for effective refugee management across European countries. Refugee management contemporarily mostly relies on top-down management approaches by governments. These often lead to suboptimal policies for refugees and highlight a need to better identify and integrate refugee needs into management. Here, we show that modern applications of Natural Language Processing (NLP) allow for the effective analysis of large text corpora linked to refugee needs, making it possible to complement top-down approaches with bottom-up knowledge centered around the current needs of the refugee population. By following a Design Science Research Methodology, we utilize 58 semi-structured stakeholder interviews within Switzerland to develop design requirements for NLP applications for refugee management. Based on the design requirements, we developed *R2G – "Refugees to Government"*, an application based on state-of-the-art topic modeling to identify refugee needs bottom-up through Telegram data. We evaluate *R2G* with a dedicated workshop held with stakeholders from the public sector and civil society. Thus, we contribute to the ongoing discourse on how to design refugee management applications and showcase how topic modeling can be utilized for data-driven governance during refugee crises.

CCS CONCEPTS

• **Computing methodologies**; • **Information extraction;**; • **Human-centered computing**; • **Systems and tools for interaction design.**;

KEYWORDS

Natural Language Processing, Topic Modeling, Refugee Management, Design Science Research

ACM Reference Format:

Kilian Sprenkamp, Liudmila Zavolokina, Mario Angst, and Mateusz Dolata. 2023. Data-Driven Governance in Crises: Topic Modelling for the Identification of Refugee Needs. In *24th Annual International Conference on Digital Government Research - Together in the unstable world: Digital government and solidarity (DGO 2023), July 11–14, 2023, Gdańsk, Poland.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3598469.3598470

DGO 2023, July 11–14, 2023, Gdańsk, Poland

ACM ISBN 979-8-4007-0837-4/23/07.
https://doi.org/10.1145/3598469.3598470

1 INTRODUCTION

Currently, 103 million people are classified by the United Nations High Commissioner for Refugees (UNHCR) as refugees, a number that has been growing since the second world war [29]. The ongoing war in Ukraine displaced roughly 8 million people [38]. Within Switzerland, the country we focus on in this study, approximately 70,000 refugees have been registered [21]. Refugee management thus has become a constant task in most countries. However, decision-makers often struggle to adequately address the incoming stream of refugees, which burdens refugees and citizens of the host country. As a consequence, refugee management has been discussed by the media, e.g., Swiss refugee management is described as repressive towards refugees [23]. Further criticism includes the Swiss government's missing capacities of asylum centers for the incoming stream of refugees [43].

Refugee management can be categorized into top-down and bottom-up approaches [12]. Typically, in top-down approaches, decisions are based on the leading political body [2], whereas bottom-up approaches actively incorporate citizens or even refugees [7]. After the second world war, refugee management has primarily been based on top-down approaches, leading to suboptimal planned policies [12]. Consequently, the UNHCR's refugee protection discourse has been accused of functioning as a global police of population with an authoritarian nature [33].

Complementing top-down approaches is possible through bottom-up approaches starting with the citizen or the refugee and putting them at the center of potential policies [2, 7, 36]. Betts et al. [7] argue that the economic lives of refugees are treated distinctively by institutions and show that more integration and involvement in the host countries can help to better the lives of refugees. Therefore, refugee management benefits from methods that effectively integrate refugee needs into management. However, this raises the issue of how bottom-up refugee management can be done by incorporating a large and diverse number of refugees. Thus, we ask the following research question:

RQ: How can we support workers in public sector organizations to enable bottom-up refugee needs identification?

Techniques proposed in crisis informatics [25], like information and communication technology (ICT) and the integration of open-source data, enable the incorporation of a large and diverse number

of refugees. Applications within crisis informatics are often powered by machine learning, especially natural language processing (NLP), allowing for the analysis of large amounts of unstructured text data [41]. Since the popularization of transformer-based [40] models, the field of NLP has gained increasing attention due to models like BERT [11] being able to analyze text with improved accuracy compared to legacy approaches. This promises the ability to gain insights from large amounts of unstructured text, allowing for data-driven decision-making within eGovernment [3] to tackle challenges like refugee crises. Topic modeling, an unsupervised machine learning technique for identifying clusters in a given dataset, has evolved similarly, yielding new models like BERTopic [14], which use the transformer architecture as well [40]. We claim for refugee management, state-of-the-art topic modeling can identify needs within large text corpora representing the refugee population.

To answer the given RQ, we utilized a Design Science Research Methodology (DSRM) [26]. We start with identifying the needs of policymakers and refugees based on semi-structured interviews with stakeholders within Switzerland involved in refugee management to formalize the problem identification and solution objective. Based on the solution objectives, we designed a dashboard application for identifying refugee needs within Switzerland via public Telegram messages aided by state-of-the-art topic modeling. We named the application *"Refugees to Government" – R2G*, which we evaluated through a dedicated workshop.

The paper is structured as follows. The subsequent chapter presents related work of the study, focusing on the current application of NLP for refugee management. Chapter 3 explains our methodology. We define current problems in refugee management in chapter 4. Chapter 5 provides possible solutions and design requirements for the identified problems. In chapter 6, we present our designed application, *R2G*, the automatic refugee needs identification application. The evaluation of the application is done in chapter 7. Chapter 8 discusses the implications of *R2G*. We conclude the study with an outlook in Chapter 9.

2 RELATED WORK

Top-down refugee management approaches are characterized by direct decision-making through the governing political body regarding refugees [2]. To complement top-down approaches, bottom-up approaches that consider citizens' or even the refugees' perspectives are utilized to enhance the quality and efficiency of political decisions [36]. As an enabler of bottom-up refugee management, ICTs, like mobile phone applications or information technology infrastructure (e.g. databases), have gained momentum to aid displaced groups [28]. Andrade and Doolin [5] state that the core capabilities of ICTs are to enable refugees to participate in an information society, communicate effectively, understand a new society, be socially connected, and express cultural identity. Realizing these capabilities allows the refugee to exercise agency, increase well-being and regain control during disruption [5].

Design guidance for using the above-named ICTs, to address the needs of refugees are well-researched [4, 31, 34]. Still, to our best knowledge, none of the existing DSRMs focused on finding a solution to support the public sector organizations in refugee management by using NLP for a bottom-up management approach. This aligns with the current research agenda by AbuJarour et al. [1] on how the public sector can leverage emerging techniques like machine learning to aid refugee management.

Current approaches in crisis informatics, promise to close this gap but lack practical design guidance. While crisis informatics originates from simple crowd-based applications like wikis [25], the field increasingly relies on NLP applications [41]. Modern examples of crisis informatics include the observation and analysis of online sentiment during disease outbreaks like COVID-19 [42] or the classification of tweets during natural and man-made disasters to aid rescue workers [44]. The public sector can thus monitor and respond to emergencies by detecting patterns within messages sent on various social media platforms [19]. NLP techniques have progressed technically on various fronts, increasing their potential for refugee management applications. A key improvement has been the popularization of transformer-based model architectures [40], increasing the accuracy of various modeling techniques [18]. One specific transformer-based model is BERT [11], which is pre-trained on a large corpus of text and can be fine-tuned to be used on various NLP tasks, like topic modeling, text classification, named-entity recognition, or question answering. Thus, BERT [11] can be utilized as a basis for various powerful tools within the public sector analyzing large numbers of unstructured text and allowing for data-driven governance [3, 10].

Unsupervised learning can be easily utilized by the public sector as it does not require labeling the obtained data. Next to the initial low setup cost of unsupervised approaches, like topic modeling, derive classes directly from the text. Compared to supervised techniques, topic modeling thus allows analyzing unexplored and rapidly changing data without prejudice from predetermined classes. However, unsupervised approaches are also difficult to evaluate on a technical level as a comparison to the true class of an instance cannot be done as in supervised learning [6][1]. One method to evaluate unsupervised algorithms like topic modeling is to measure coherence scores and topic diversity within the found topics [14]. However, the metrics can only be approximated and are based on subjective judgment [14].

Within topic modeling, the largely popular Latent Dirichlet Allocation (LDA) [8] has become a legacy approach compared to modern transformer-based techniques like BERTopic [14], building up on BERT [11]. Advantages of BERTopic [14] compared to LDA include the utilization of contextual embeddings allowing the model to capture context within the input text [14]. BERTopic achieves this by leveraging sentence transformers, which are pre-trained neural network models designed to generate dense vector representations (embeddings) of sentences or paragraphs, capturing their semantic meaning [14]. This results in enhanced topic modeling performance, producing more coherent and contextually relevant topics[14]. Further, BERTopic [14] utilizes a modular structure, meaning if new advances in language models, e.g., sentence transformers, are created, they can be easily utilized [14]. Moreover, LDA cannot incorporate multi-lingual text data without prior modifications to the algorithm [9], making it difficult to use for refugee management as the data set often contains multiple languages (e.g., the language of the host country and the language of the refugee).

[1] Consider Berry et. al [6] to learn more about supervised and unsupervised learning

Within refugee management, there are limited attempts to utilize online participation and topic modeling, as done in other applications of crisis informatics [42, 44]. Current research primarily relies on legacy approaches like LDA and does not consider transformer-based architecture. Lind et al. [20] explore media framing of migration, meaning how the media portrays migration. The authors compare the media of 7 European Union countries, creating a multi-lingual text corpus containing Spanish, English, German, Swedish, Polish, Hungarian and Romanian texts. However, while the authors name possible extensions to enable a multi-lingual topic model, they create a monolingual topic model for each country utilizing LDA. Hence, the topics found in the media data are likely biased toward a single country, and relationships between countries are harder to analyze. Similarly, while focusing on sentiment analysis, Kamyab et al. [17] analyzed Twitter posts regarding the Syrian refugee crisis in Turkish and English. The results, while comparing text from both languages, only allow for a limited joint view of the refugee crisis. We aim to build on these prior studies, by exploring new BERT-based [11] approaches, e.g., BERTopic [14], enabling the modeling of multi-lingual text corpora, which were not available at the time these studies were conducted.

Further, the data utilized for topic modeling does not necessarily represent the needs of refugees but often focuses on the perception of refugees within the host country. Heidenreich et al. [15] utilize topic modeling to investigate the media framing of refugees within five European countries. The analysis reveals dynamics and shifts in the refugee debate, showing the importance of the geo-location of a country regarding media framing of refugees. Similarly, Pöyhtäri et al. [27] analyzed the Finnish hybrid media, consisting of traditional and social media, via topic modeling. The authors thus identified overemphasis on topics like crime on social media, quoting traditional media sources. Hence, the analysis suggests that the hybrid media setting leads to the polarization of locals within the host country, which political parties can utilize. A top-down approach to refugee management applying topic modeling to historical policy documents was implemented by Grant et al. [13], revealing trends over time within refugee management. The studies [12, 14, 26] provide valuable insights into the perception of refugees and historical policies utilizing topic modeling. However, we see an opportunity to further build upon this research and explore how bottom-up data-driven governance [3, 10] of refugees can be done.

Summarizing, we identified that (1) currently, there exists no design guidance for applying NLP, especially topic modeling for bottom-up refugee management, (2) applications utilizing topic modeling often focus on legacy approaches, and (3) the text corpora analyzed do not allow for bottom-up refugee management. This leads to negative effects for governments and, consequently, refugees in the real world. First, governments cannot follow design guidance for implementing NLP applications. This can result in non-data-driven policies, potentially suboptimal for refugees. Second, the potential use of legacy approaches by governments can lead to applications being unable to deal with needed requirements, e.g., the ability to automatically understand a multi-lingual text. Third, solely focusing on top-down approaches can lead to poorly planned policies, which can be enhanced by data-driven bottom-up knowledge.

3 METHOD

This study is a part of a larger project that deals with the Ukrainian refugee crisis in Switzerland and the role of digital platforms therein. In our project, to create a solution for a real-world problem [35], we employed the DSRM [16, 24, 26]. DSRM is a widely accepted approach in information systems and digital government research, and it aims to create IT artifacts such as design products, methods, or design processes. Our project resulted in creating a proof-of-concept prototype [24] of an application intended to be used by the Swiss public sector employees involved in the management or operations during the Ukrainian refugee crisis. Proof-of-concept prototypes are valuable for understanding and early insight into the problem space and the feasibility of a potential solution and its functionality. Using the DSRM framework, we followed a six-step process outlined by Peffers et al. [26], including (1) problem identification, (2) solution objectives, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication.

To (1) identify the emergent problems of different stakeholders in the Ukrainian refugee crisis and (2) derive the objectives for our solution, we conducted 58 semi-structured interviews [22]. We interviewed 20 Ukrainian refugees (abbreviated as *R1-R20*). Further, we interviewed representatives of 15 public sector organizations (*P1-P15*), 15 non-profit and civil society organizations (*NP1-NP15*), and eight interviews with individuals who volunteered, e.g., by translating, providing housing to refugees, or meeting arriving refugees at train stations (*V1-V8*). These stakeholders were chosen as each interacts with the public sector and has specific demands towards public management. We decided to conduct semi-structured interviews since they allow for collecting and analyzing stories of peoples' experiences and perspectives, naturally told in their own words. For this, we prepared open-ended questions about the recent experiences of an interviewee regarding what went well and badly for them and what digital technologies they used. In the interviews with organizations, we also asked about their feedback on using digital platforms and applications such as data analytics in crisis situations. The interviews were recorded, transcribed, and analyzed using qualitative coding techniques [32].

Based on the qualitative analysis of interview data, we identify key problems in the public sector which arose during the refugee crisis. Based on the problem identification, we define solution objectives for utilizing NLP applications within the Swiss public sector. For the (3) design and development phase, we iteratively created our prototype, an interactive dashboard application called *R2G*, that users could navigate independently. *R2G* visualizes actual data from 33 open Telegram groups related to the Ukrainian refugee crisis in Switzerland. These groups were used for communication and coordination with and between refugees, volunteers, and sometimes organizations. *R2G* then clusters the Telegram data via the BERTopic algorithm [14], creating distinct topics that directly relate to refugees' needs. As per the DSRM methodology, it is essential to evaluate the artifacts created to determine if the research objectives have been met. Hence, (4) demonstration and (5) evaluation were carried out as the next steps.

The evaluation aimed to assess the prototypes' performance in a realistic setting within a particular organizational context. As per the suggestions of Sonnenberg et al. [37], this evaluation followed a

"prototyping pattern," which includes naturalistic and ex-post evaluation. In this pattern, the researchers specify the artifact's design, develop a prototype, select real users and a relevant organizational context, monitor the use of the prototype, and evaluate if the set goals of the prototype were met during its usage. We conducted intermediate evaluations to ensure the usability and usefulness of the design within the research team, partner organizations, and several interviewed organizations. The evaluation results we report were collected in the final workshop. The workshop's goal was to gather feedback on the general usefulness and potential specific aspects based on the design requirements and concerns of *R2G* for refugee management. Consequently, in the final workshop, we presented *R2G* to four participants: two were government executives, and two were representatives of civil society organizations in Switzerland. Participants had a chance to try out the provided tool. Afterward, we discussed the challenges and benefits of the solution in a one-hour-long discussion. The discussion was captured using post-its and whiteboards. Finally, the (6) communication of this research happens with the article at hand. While we believe that the study at hand already provides valuable insight into the potential of NLP and topic modeling for refugee management, this work will be continued to reach more mature design phases of proof-of-use and proof-of-value [24].

4 PROBLEM IDENTIFICATION

We utilized the interviews to identify refugees' problems and how the public sector can build applications using NLP to act quickly regarding varying refugee needs.

(1) Understanding Refugee Concerns: To provide refugees with high-quality services, the public sector needs to understand the problems that refugees face well. However, from our refugee interviews, we can identify numerous services needed by refugees, which the Swiss public sector was not aware of or did underestimate and consequently mishandled. Examples of needed services are childcare (*R15*), temporary and long-term accommodation (*R4, R2, R16, R19*), travel (*R8*), medical system (*R11*) or work permits (*R2, R16, R17*). Our public sector interviews verified this claim of misunderstanding concerns made by the refugee community, *P13* stated that: *"Bringing the right people together and making sure that they don't only think in their own structures and organize information just how they understand it and need it [. . .]."* Thus, it can be indicated that the public sector acts based on their insights, not necessarily on the refugee perspective. This leads to services that are not optimally designed for the refugee, as stated by *V5: "You know, the processes at the beginning were so, so inaccurate, and we don't know exactly what happened [. . .]."*

(2) Inadequate exchange across federal level: The federal political system of Switzerland delegates decision-making power in many policy domains to cantons and municipalities, following subsidiarity principles. During the Ukrainian refugee crisis, this resulted in an unclear distribution of responsibilities between different federal levels. We especially observed frustration with the federal level from the municipalities and the cantons, as they did not perceive adequate support in refugee management, e.g., *P13: "Many things are organized on the municipal level. And only certain things are organized on the cantonal or even federal level. So, you know, it's a complex thing. Especially when it comes to who pays which fees and stuff."* Federalism requires solutions to work on different federal levels independently, which creates the danger of multiple different solutions being designed concurrently, e.g., *P8: "a solution should be initialized nationally, because if every canton creates something, then it needs interfaces again afterward. A smart solution would be to repel that from the federal government and suggest a solution all cantons can use."* Our data suggest that inadequate exchange across the federal level led the federal government not to provide municipalities and cantons with a coordinated but independently usable set of tools to deal with refugee management.

(3) Language Barrier: Multiple refugees stated that they could not access services offered by the public sector due to language barriers, e.g., *R16: "I do not have Swiss acquaintances at the moment. I think that might be because of the language barrier"* Interviews with both the public sector and volunteers confirmed the problem of the language barrier, stating that understanding refugees and communicating their services in a foreign language is challenging, e.g., *P4: "[. . .] because the language barrier is very big and then everything must be translated."* While if both parties speak a common language like English, communication would be easier, e.g., *V5: "she [the refugee] just could speak English, quite well, so that one could communicate well."* Hence, we deduce that the existing language barrier is a problem in refugee management.

(4) Data exchange: The public sector struggles with legal and technical issues in relation to data exchange. On a technical level, there is no API for data exchange between different public sector organizations within Switzerland, compare *P3: "Currently, not even one office of the Canton (. . .) can share personal information with another office of the Canton internally."* Next to technical challenges, legal requirements for data exchange can hinder efficient working, *P8: "Due to data security [. . .] connections with family circumstances, where they live, what they can do, etc., and this data is not available to us."* Our interviews suggested that the inability of data exchange is likely a main factor hindering effective government services.

(5) Scalability: Processes and tools used for refugee management often did not scale well in crisis situations, e.g., *P13: "In regular time administrative processes work well, but they are not scalable. And once there are that many people the system fails"*. In addition, it was highlighted by *NP7* that technical resources are often missing and need to be bought from third parties: *"we are forced to look for third-party providers, companies in Switzerland or overseas to store this data reliably because there is no public service for community projects like ours."* This highlights the necessity to create digital tools, which can be easily scaled if demand increases abruptly.

5 SOLUTION OBJECTIVES AND DESIGN REQUIREMENTS

We derive solution objectives and propose design requirements from the five problems identified in the prior section (Table 1). To solve problem *(1) Understanding Refugee Concerns* we define the solution objective (*i) Design services within refugee management on insights from the refugee community* and propose the design requirement *(A) Choose a data source reflecting the refugee population*, we thus can ensure that refugee services are created in a bottom-up manner. Further, to solve problem *(2) Inadequate exchange across*

Table 1: Problems, Solution Objectives and Design Requirements

Problems	Solution Objective	Design Requirements
(1) Understanding Refugee Concerns	*(i) Design services within refugee management on insights from the refugee community*	*(A) Choose a data source reflecting the refugee population*
(2) Inadequate exchange across federal level institutions	*(ii) Use on various government levels independently is possible*	*(B) Map data points to geo-location*
(3) Language Barrier	*(iii) Understand multiple languages*	*(C) Utilize multi-lingual sentence transformer*
(4) Data exchange	*(iv) Share no data across governmental organizations*	*(D) Utilize open-Source Data*
(5) Scalability	*(v) Resile against varying traffic*	*(E) Scalable architecture design*

Table 2: Design Requirements and Design Features of R2G

Design Requirements	**Design Features of R2G**
(A) Choose a data source reflecting the refugee population	*(a) Utilize Telegram channels by the Ukrainian refugee population in Switzerland*
(B) Map data points to geo-location	*(b) Approximate the geo-location of messages within Telegram*
(C) Utilize multi-lingual sentence transformer	*(c) Select multilingual sentence transformer within BERTopic during model training*
(D) Utilize open-source Data	*(d) Choose open Telegram channels for data collection*
(E) Scalable architecture design	*(e) Utilize Streamlit for hosting*

federal level institutions, we specify the solution objective *(ii) Use on various government levels independently is possible*. To achieve this, we define the design requirement *(B) Map data points to geo-location*. Hence, we can ensure that the provided services work on various levels independently in the Swiss government (e.g., municipality, canton, federal) and each organization's application is based on data points relevant to their constituency. While this requirement is most important for the Swiss government as it gives municipalities and cantons large amounts of freedom, we estimate it also applies to other federalist government structures, e.g., Germany. Moreover, to solve problem *(3) Language Barrier*, we define the solution objective *(iii) understand multiple languages* for the application. We meet this solution objective through design requirement *(C) Utilize multi-lingual sentence transformer*. Thus, the designed application can work with multi-lingual input text. Additionally, for Problem *(4) Data exchange* and to not interfere with the national data privacy law, we define solution objective *(iv) Share no data across governmental organizations*. While we cannot alter Swiss legislation or the data exchange within the Swiss public sector to achieve the solution objective, we propose a workaround *(4) Utilize open-source data* as a design requirement. Finally, to resolve Problem *(5) Scalability* the application must be *(v) Resile against varying traffic*, we recommend the design requirement *(E) Scalable architecture design*.

6 ARTEFACT DESIGN AND DEVELOPMENT

Based on the defined solution objectives, we designed and developed a dashboard application called *R2G* for identifying refugee needs in a bottom-up manner. Table 2 summarizes how different Design Features of *R2G* fulfill each design requirement. Policymakers and public sector agencies can utilize *R2G* to make data-driven decisions by quantifying the needs of refugees. Moreover, refugees can hope for better service delivery enabled by the public sector using *R2G*. The Python code used to develop *R2G* can be viewed in the following anonymized GitHub repo[2]. Within the GitHub repo, we also provide an architecture diagram for a better overview of how the *R2G* works.

To meet design requirement *(A) Choose a data source reflecting the refugee population,* we use data from Telegram as it is the main communication channel of refugees from Ukraine. Using the Telethon Python package, we scrapped data from 33 open Telegram groups, meeting design requirement *(D) Utilize Open-Source Data* active in various cantons across Switzerland. A list of the used Telegram groups can be found in the GitHub repo. In total, 476,352 messages dating from 25.02.2022 to 24.11.2022 were scrapped. From each message, we scraped the following attributes: the group name the message was posted in, the message text, and the date the message was written on. While Telethon allows us to scrape user information like the Telegram user ID and potentially name and phone number, we refrained from obtaining this data to shield the refugee's privacy.

Next, we utilized the BERTopic [14] algorithm on the data set obtained. However, we exclude messages shorter than 100 characters long, as these tend to have little value for forming clusters. This left us with 170,915 messages for further analysis. Moreover, we remove stop words using the NLTK Python package, e.g., articles, prepositions, pronouns, and conjunctions, so the topic modeling algorithm does not process low-information content. For Ukrainian, as it is a low-resource language and not included in the NLTK package, we utilized an open-source list of stop words [39], which we evaluated by a native Ukrainian speaker. We set the algorithm

[2]https://github.com/sprenkamp/DGO_2023

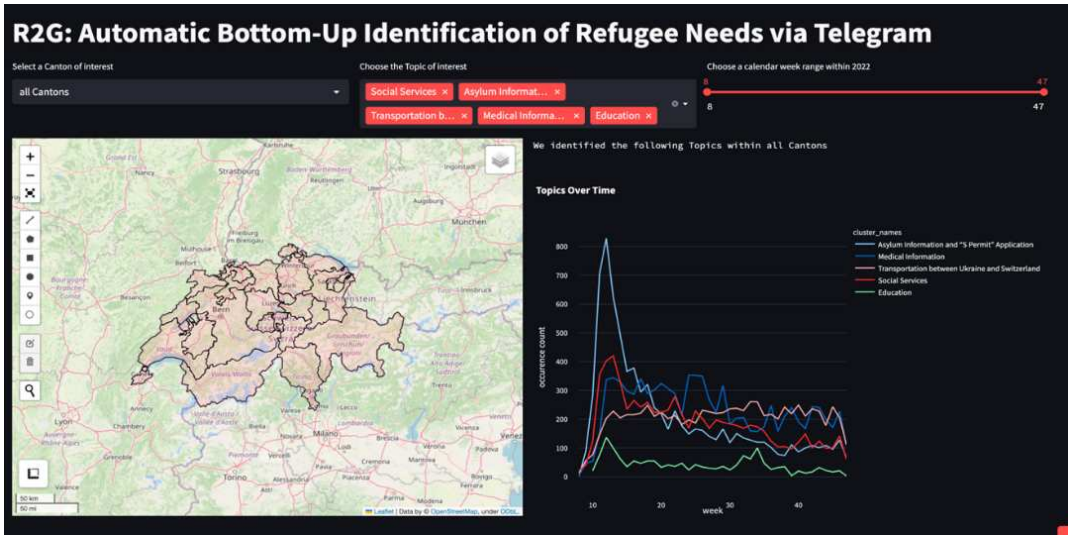

Figure 1: Example Usage of R2G

to *'multilingual'* thus, different languages are encoded utilizing the same sentence transformer model [30], enabling the model to cluster text from different languages, e.g., Ukrainian, Russian, German or English. We thus meet design requirement *(C) Utilize multi-lingual sentence transformer*. Apart from the steps above, we utilized the base settings of the BERTopic [14] algorithm and did not provide special hyperparameter tuning, as optimization within unsupervised learning tasks is not straightforward.

We chose to generate the top 25 topics based on the balance between cluster quality and human comprehensibility, as our goal is to develop a tool for decision-makers. This approach was recommended by the creator of BERTopic [14]. We evaluated varying numbers of clusters (e.g., 10, 25, 50, and 100) and the *'auto'* setting in BERTopic [14], but found that the latter produced too many clusters, making it difficult for human understanding. We then let the obtained model predict the class of each of the 476,352 messages. As BERTopic [14] automatically forms a class of non-assignable textual inputs, most messages belong to this class, as they are noise not relating directly to refugee needs. The model thus assigns 65,154 to the 25 topics. Afterward, we evaluate each topic in a qualitative manner by selecting 100 random samples. Based on the given samples, we name each cluster and merge similar clusters, resulting in a total of 17 topics directly relating to refugee needs. In the appendix, we present Table **??**, showing the cluster names, description of the refugee problems, and the distribution of messages belonging to a given topic. We can see that the top 5 topics referring to refugee needs are: Medical Information, Asylum Information and 'S Permit' [3]Application, Transportation between Ukraine and Switzerland, Social Services, and Integration Problems.

Last, from the results of the given BERTopic model, we utilize the Streamlit Python package to create an interactive dashboard[4], which can be used to identify the needs of refugees in a bottom-up manner and thus allows for data-driven governance [3, 10]. Next to the interactive dashboard, which the reader can utilize, we further provide an image showing an example of usage of *R2G* in Figure 1. While Streamlit is a rather simple tool for Python-based web development, and *R2G* is currently using only using a rather slow server, it allows the hosting of servers and databases on different services like Amazon Web Services. Hence, we can easily increase server capacities or storage needed, meeting design requirement *(E) Scalable architecture design*. Further, within *R2G*, we utilize the name of the open Telegram groups to indicate where a message was written. We also create a lookup table to detect city names within a given message and assign them to a canton. The latter technique of the lookup table overwrites the prior made approximation through the group name. Hence, we can approximate the geo-location of a Telegram message to the cantonal level and meet design requirement *(B) Map data points to geo-location.* In theory, users of the tool could also view single messages and thus qualitatively analyze each topic. However, we decided not to enable this in the public-facing version of *R2G*, to shield the refugees' privacy.

7 EVALUATION

The participants of the evaluation workshops underline the general usefulness of *R2G* as it can be set up efficiently. Thus, refugee needs can be analyzed in near real-time, with each topic of the BERTopic algorithm relating to a particular need of refugees, allowing the public sector to adapt to a crisis. The public sector can automatically monitor refugee needs and is not dependent on other means of identifying needs, e.g., interviews, which may require more effort and resources and will be potentially biased.

A similar workflow can easily be replicated for different crisis events utilizing data from various open-source resources, e.g., Twitter, Facebook, or traditional media. Each data stream would potentially have a different value for various public sector stakeholders. Hence, the workshop participants regarded the tool as adaptable to other crises, which can be set up to aid information retrieval and decision-making for public sector agencies.

[3]S Permit refers to the permit given to Ukrainians in need of protection.
[4]https://dgosubmission2023-r2g.streamlit.app/

Moreover, the workshop participants saw value in following chosen design features within *R2G*. Participants appreciated that *R2G* utilizes geo-tagging of messages to a given location, e.g., Swiss cantons, allowed through design feature *(b) Approximate the geo-location of messages within Telegram.* Thus, refugee needs can be quantified across various regions within Switzerland. Similarly, we can filter based on the date allowing the decision-makers to view certain periods within a given geo-location. Moreover, the possibility of algorithms to understand multiple languages at once was not known to participants. Hence, they valued that *R2G* can work with multi-lingual text inputs, enabled through design feature *(c) Select multilingual sentence transformer within BERTopic during model training.*

The workshop participants also discussed *R2G*'s major drawbacks. While *R2G* only utilizes open-source data, participants feared that *R2G* does not act within the correct ethical boundaries. It is unclear if the refugees appreciate using their messages to identify their needs. We see a conflict between better service quality, which can be provided through *R2G*, and the refugees' privacy. However, the workshop participants agreed that the users of the instant messenger need to improve their digital literacy. Thus, public channels should only be utilized for content that can be consumed and analyzed by everyone.

Concerning ethical issues, the workshop participants also were concerned about the potential abuse of *R2G*. For example, one could easily spot a negative sentiment of single Ukrainian refugees against the Swiss government by looking into the messages within a given topic. *R2G* could thus be used to monitor refugees' activities rather than just identify their needs.

Further, the workshop participants feared the impact of wrong information provided by the refugee crowd could harm the public sector. However, in the current stage, incorrect information is not a major problem as the aggregated information is only used to form manually labeled topics. These are not altered by wrong information provided through a single message.

8 DISCUSSION

This study aimed to address the research gap in how state-of-the-art topic modeling can be utilized to identify refugee needs. We did so by following a DSRM [26], starting with the identification of problems of refugees and the public sector. Following the design requirements, we implemented *R2G* for the automated identification of refugee needs, which we evaluated with a workshop. Hence, we can answer our *RQ: How can we support workers in public sector organizations to enable bottom-up refugee needs identification?*

First, the effect of ICTs, like mobile phone applications or information technology infrastructure, on refugee management is well-researched [5, 28], and design principles for using these ICTs [4, 31, 34] to aid refugee needs exist. However, it remained unclear how machine learning can be utilized by the public sector, as stated by the research agenda of AbuJarour et al. [1]. We contribute to this discourse by deriving design requirements (Table 1) for developing NLP applications based on 58 stakeholder interviews. Further, by creating *R2G,* we show how these requirements can be converted into design features (Table 2). Therefore, we offer an example of how data-driven governance [3, 10] can be designed and developed in Switzerland, from which governments and refugees can profit. The obtained knowledge can be used to create different NLP applications within crisis informatics [25], with potentially only a minor adaption.

Second, existing topic modeling applications within refugee management do not allow for bottom-up refugee management. Instead, their focus is on understanding the media framing of refugees [15, 27], allowing us to get insights into the inclusion of refugees within their host country. While analyzing public policies [13] in a top-down manner yields insights into the evolution of refugee policies, the authors cannot guarantee that the information obtained through topic modeling reflects the refugee's needs. Contrarily, *R2G* adds a new perspective to refugee management, putting the refugee at the center of the analysis, allowing us to identify the needs of refugees, e.g., Medical Information, Asylum Information and 'S Permit' Application, or Transportation between Ukraine and Switzerland. We thus create a bottom-up refugee management application, which can be used to complement top-down approaches. Further, through temporal and geospatial analysis, we can monitor and effectively prioritize the needs of refugees. Using *R2G,* the public sector can, for example, see that in calendar week 12 of 2022, the demand for "Asylum Information and 'S Permit' Application" peaked roughly four weeks after 24.02.2022. Similarly, while the need for "Asylum Information and 'S Permit' Application" steeply declined after calendar week 12, the need for "Medical Information" and "Transportation between Ukraine and Switzerland" remained constant through the observed period (Figure 1). Thus, governments can build and prioritize their services based on data-driven evidence [3]. However, while topic modeling can identify and prioritize the needs of refugees, as shown by *R2G,* we cannot measure the sentiment within each topic. Previous research, such as Kamyab et al. [17] have shown the value of further NLP applications in crises like sentiment analysis.

Third, our tool utilizes state-of-the-art topic modeling based on the transformer architecture [40], while prior applications [13, 15, 20, 27] are built with the LDA technique [8]. Utilizing BERTopic [14] allows us to leverage contextual embeddings to capture the notion of the input text better. Moreover, due to the modular structure within BERTopic [14], parts of the model, e.g., the sentence transformer [30], can be updated. We expect that the quality of our tool has the chance to increase, given further developments within NLP. Additionally, our system based on BERTopic [14] allows for multi-lingual topic modeling through the use of multi-lingual sentence transformers [30]. *R2G* can thus capture relationships between textual inputs in multiple languages extending on prior work, creating a single model for each language in a multi-lingual text corpora [17, 20], which can miss capturing relationships between text in multiple languages. Thus, we can cope with the unique challenge of the Ukrainian refugee crisis, where refugees communicate in Ukrainian and Russian as well as German and English. We can easily adapt our solution to different refugee crises, like the Syrian refugee crisis, as the used sentence transformer [30] supports Arabic. Supposing the language is not supported by the current sentence transformer [30], we could still train a sentence transformer independently and use the modular structure of BERTopic [14] to support a refugee crisis with refugees speaking a low-resource language. For other crisis events, we see

that such an application may be valuable in cases where there is a need to interpret multi-lingual text corpora. For example, an application dealing with the effect of natural disasters in the United States will mostly handle texts written in English and maybe a couple in other languages like Spanish, making the effect of the multi-lingual analysis marginal. However, regarding other crises, the multi-lingual analysis could prove useful. For each crisis, if the language is supported by the sentence transformer [30], we can use an application like *R2G*, powered by BERTopic [14], out of the box.

Still, *R2G* has several weaknesses we identified through the external workshop. First, while we utilize open-source data to solve the problem of data sharing between administrations within the Swiss public sector, it remains questionable if it is ethically correct to utilize Telegram messages to gain insights on refugee needs, as the workshop participants criticized it. Ethical concerns are a further issue as the workshop participants see the possibility of misconduct of *R2G*. Thus, governments can abuse the system to identify "misbehaving" refugees rather than identifying refugee needs. While we only retrieved necessary information for *R2G*, a malicious government could utilize personal information about a "misbehaving" refugee and use it against the refugee. Additionally, our system exclusively depends on Telegram data as the primary information source, raising concerns among workshop participants about potential inaccuracies in the data influencing *R2G*. In the current iteration, we use messages solely to identify refugee needs. For instance, incorrect information about a topic merely emphasizes its importance, but the government would not actively use this misinformation for communication, only for monitoring. However, we did not foresee the potential impact of bots or malicious agents (e.g. far-right political groups), that could skew the analysis and negatively affect the perception of refugees. Integrating additional traditional and social media sources could provide a more comprehensive understanding of the Ukrainian refugee crisis.

9 CONCLUSION

Due to geopolitical tension, the number of refugees has been growing since the second world war [29], resulting in an immense task for the public sector. To combat this development, we utilized a DSRM [26] to create *R2G*, an application based on state-of-the-art topic modeling to identify refugee needs bottom-up, allowing governments to monitor and prioritize needs and services for refugees effectively.

The contribution of this study is twofold. First, through the DSRM, we identify problems, solution objectives, and design requirements (Table 1) based on 58 interviews. We hope these provide a basis for designing NLP-based applications for refugee management and that the design knowledge gathered can be successfully applied to other crises. Second, we designed the first system to automatically identify refugee needs in a bottom-up approach. We hope *R2G* will spark further research into NLP systems that function in a bottom-up manner. These contributions affect multiple stakeholders. First, the public sector benefits from an application they can employ. Second, refugees can hope for a timelier response to their problems. Third, researchers in digital government gain insights into designing NLP systems for bottom-up refugee management.

However, our study shows several limitations. First, interviews and workshops were held focusing solely on Switzerland. Still, we assume that the DSRM approach would yield similar results within other European countries we cannot exclude that in different regions of the world, the design requirements given cannot be applied. Second, for the workshop, only four participants were used. Further, studies should focus on a larger group of evaluation participants to discuss the benefits and challenges of *R2G*. Moreover, a technical evaluation of the BERTopic model was not done due to the subjective nature of evaluating unsupervised learning algorithms [14]. A further technical evaluation could include measure like topic similarity and topic coherence. These measures could be further used to compare BERTopic to traditional approaches like LDA [8]. Third, our tool only utilizes data from telegram, creating a bias towards the messenger service.

Summarizing, we see substantial potential for using topic modeling to enable bottom-up refugee management. We suggest that future research should explicitly investigate the ethics of utilizing open-source data. Additionally, since our evaluation involved only four participants, future studies could aim for a more extensive assessment to refine design requirements and determine the specific categories of decision-makers (e.g., political leaders, public administration) who would primarily benefit from using *R2G*. Furthermore, our study focused mainly on topic modeling. Other NLP algorithms, like sentiment analysis [17], will give further insights for the public sector. Last, our study solely utilized a single data source -Telegram- within a single country. We believe that extending this to further traditional and social media sources and expanding data collection to more countries, potentially all of Europe, will generate a comprehensive view of refugee management for the public sector. We see these recommendations as opportunities for future extensions to our existing application, *R2G*.

ACKNOWLEDGMENTS

We thank the University of Zurich and the Digital Society Initiative for (partially) financing this study under the Digitalization Initiative of the Zurich Higher Education Institutions (DIZH) postdoc fellowship of Liudmila Zavolokina. We further thank the DIZH for funding the rapid action 'Government as/is a platform - Orchestrating those involved in the refugee crisis through digital platforms in the Canton of Zurich'. Additionally, we like to thank our project partners 'Campax' and 'Staatskanzlei Kanton Zürich'. Last, we especially like to thank the team members, including: Prof. Gerhard Schwabe, Dr. Inna Vashkite, Dzmitry Katsiuba, Ivan Volkov, Michalina Preisner, and Zoia Katashinskaia.

REFERENCES

[1] AbuJarour, S., Wiesche, M., Díaz Andrade, A., Fedorowicz, J., Krasnova, H., Olbrich, S., Tan, C.-W., Urquhart, C. and Venkatesh, V. 2019. ICT-enabled Refugee Integration: A Research Agenda. *Communications of the Association for Information Systems*. (2019), 874–891.

[2] Afouxenidis, A., Petrou, M., Kandylis, G., Tramountanis, A. and Giannaki, D. 2017. Dealing with a humanitarian crisis: Refugees on the Eastern EU border of the island of Lesvos. *Journal of Applied Security Research*. 12, 1 (2017), 7–39.

[3] Alexopoulos, C., Lachana, Z., Androutsopoulou, A., Diamantopoulou, V., Charalabidis, Y. and Loutsaris, M.A. 2019. How machine learning is changing e-government. *Proceedings of the 12th international conference on theory and practice of electronic governance* (2019), 354–363.

[4] Amard, A., Hoess, A., Roth, T., Fridgen, G. and Rieger, A. 2022. Guiding Refugees Through European Bureaucracy: Designing a Trustworthy Mobile App for Document Management. *International Conference on Design Science Research in Information Systems and Technology* (2022), 171–182.

[5] Andrade, A.D. and Doolin, B. 2016. Information and Communication Technology and the Social Inclusion of Refugees. *MIS Quarterly*. 40, 2 (2016), 405–416.

[6] Berry, M.W., Mohamed, A. and Yap, B.W. 2019. *Supervised and unsupervised learning for data science*. Springer.

[7] Betts, A., Omata, N. and Bloom, L. 2017. Thrive or survive? Explaining variation in economic outcomes for refugees. *Journal on Migration and Human Security*. 5, 4 (2017), 716–743.

[8] Blei, D.M., Ng, A.Y. and Jordan, M.I. 2003. Latent dirichlet allocation. *Journal of machine Learning research*. 3, Jan (2003), 993–1022.

[9] De Smet, W., Tang, J. and Moens, M.-F. 2011. Knowledge transfer across multilingual corpora via latent topics. *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (2011), 549–560.

[10] Dencik, L., Redden, J., Hintz, A. and Warne, H. 2019. The'golden view': data-driven governance in the scoring society. *Internet Policy Review*. 8, 2 (2019), 1–24.

[11] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. (2018).

[12] Easton-Calabria, E.E. 2015. From bottom-up to top-down: The 'pre-history'of refugee livelihoods assistance from 1919 to 1979. *Journal of Refugee Studies*. 28, 3 (2015), 412–436.

[13] Grant, P., Sebastian, R., Allassonnière-Tang, M. and Cosemans, S. 2021. Topic modelling on archive documents from the 1970s: global policies on refugees. *Digital Scholarship in the Humanities*. 36, 4 (2021), 886–904.

[14] Grootendorst, M. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*. (2022).

[15] Heidenreich, T., Lind, F., Eberl, J.-M. and Boomgaarden, H.G. 2019. Media framing dynamics of the 'European refugee crisis': A comparative topic modelling approach. *Journal of Refugee Studies*. 32, Special_Issue_1 (2019), i172–i182.

[16] Hevner, A.R., March, S.T., Park, J. and Ram, S. 2004. Design science in information systems research. *Management Information Systems Quarterly*. (2004). DOI:https://doi.org/10.2307/25148625.

[17] Kamyab, M., Tao, R., Mohammadi, M.H. and Rasool, A. 2018. Sentiment Analysis on Twitter: A text Mining Approach to the Afghanistan Status Reviews. *AIVR 2018*. (2018). DOI:https://doi.org/10.1145/3293663.3293687.

[18] Karita, S., Chen, N., Hayashi, T., Hori, T., Inaguma, H., Jiang, Z., Someki, M., Soplin, N.E.Y., Yamamoto, R., Wang, X., and others 2019. A comparative study on transformer vs rnn in speech applications. *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)* (2019), 449–456.

[19] Kavanaugh, A., Fox, E.A., Sheetz, S., Yang, S., Li, L.T., Whalen, T., Shoemaker, D., Natsev, P. and Xie, L. 2011. Social media use by government: From the routine to the critical. *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times* (2011), 121–130.

[20] Lind, F., Eberl, J.-M., Galyga, S., Heidenreich, T., Boomgaarden, H.G., Jiménez, B.H. and Berganza, R. 2019. A bridge over the language gap: Topic modelling for text analyses across languages for country comparative research. *University of Vienna: Working Paper of the REMINDER-Project*. (2019).

[21] More than 70,000 Ukrainian refugees flee to Switzerland: https://www.swissinfo.ch/eng/politics/more-than-70-000-ukrainian-refugees-flee-to-switzerland/48036534. Accessed: 2023-01-19.

[22]]Myers, M.D. and Newman, M. 2007. The qualitative interview in IS research: Examining the craft. *Information and Organization*. 17, 1 (Jan. 2007), 2–26. DOI:https://doi.org/10.1016/j.infoandorg.2006.11.001.

[23] Neff, B. 2022. Ruud Koopmans: "Das heutige Asylsystem ist marode." *Neue Zürcher Zeitung*.

[24] Nunamaker, J.F., Briggs, R.O., Derrick, D.C. and Schwabe, G. 2015. The Last Research Mile: Achieving Both Rigor and Relevance in Information Systems Research. *Journal of Management Information Systems*. 32, 3 (Jul. 2015), 10–47. DOI:https://doi.org/10.1080/07421222.2015.1094961.

[25] Palen, L. and Anderson, K.M. 2016. Crisis informatics—New data for extraordinary times. *Science*. 353, 6296 (2016), 224–225.

[26] Peffers, K., Tuunanen, T., Rothenberger, M.A. and Chatterjee, S. 2007. A design science research methodology for information systems research. *Journal of management information systems*. 24, 3 (2007), 45–77.

[27] Pöyhtäri, R., Nelimarkka, M., Nikunen, K., Ojala, M., Pantti, M. and Pääkkönen, J. 2021. Refugee debate and networked framing in the hybrid media environment. *International Communication Gazette*. 83, 1 (2021), 81–102.

[28] Qureshi, S. 2015. Are we making a better world with information and communication technology for development (ICT4D) research? Findings from the field and theory building. *Information Technology for Development*. Taylor & Francis.

[29] Refugees under UNHCR's mandate: 2022. https://www.unhcr.org/refugee-statistics/download/?url$=$2bxU2f. Accessed: 2023-01-07.

[30] Reimers, N. and Gurevych, I. 2020. Making monolingual sentence embeddings multilingual using knowledge distillation. *arXiv preprint arXiv:2004.09813*. (2020).

[31] Rosenbaum, J., Zepic, R., Bracklo, C., Schreieck, M., Wiesche, M. and Krcmar, H. 2018. Integreat: An information ecosystem for refugees. (2018).

[32] Saldaña, J. 2015. *The coding manual for qualitative researchers*. Sage.

[33] Scheel, S. and Ratfisch, P. 2017. Refugee protection meets migration management: UNHCR as a global police of populations. *International Organisations and the Politics of Migration*. Routledge. 60–77.

[34] Schreieck, M., Zitzelsberger, J., Siepe, S., Wiesche, M. and Krcmar, H. Supporting Refugees in Every Day Life – Intercultural Design Evaluation of an Application for Local Information. 12.

[35] Sein, M.K., Henfridsson, O., Purao, S., Rossi, M. and Lindgren, R. 2011. Action design research. *MIS quarterly*. (2011), 37–56.

[36] Siviş, S. 2021. Integrating Bottom-up into Top-down: The Role of Local Actors in Labour Market Integration of Syrian Refugees in Turkey. *International Migration*. 59, 4 (2021), 190–206.

[37] Sonnenberg, C. and vom Brocke, J. 2012. Evaluations in the science of the artificial–reconsidering the build-evaluate pattern in design science research. *International Conference on Design Science Research in Information Systems* (2012), 381–397.

[38] Ukraine emergency: https://www.unhcr.org/ukraine-emergency.html. Accessed: 2023-01-13.

[39] Ukrainian-Stopwords: 2022. https://github.com/skupriienko/Ukrainian-Stopwords. Accessed: 2023-01-14.

[40] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser,\ Lukasz and Polosukhin, I. 2017. Attention is all you need. *Advances in neural information processing systems*. 30, (2017).

[41] Verma, S., Vieweg, S., Corvey, W., Palen, L., Martin, J., Palmer, M., Schram, A. and Anderson, K. 2011. Natural language processing to the rescue? extracting" situational awareness" tweets during mass emergency. *Proceedings of the International AAAI Conference on Web and Social Media* (2011), 385–392.

[42] Wang, T., Lu, K., Chow, K.P. and Zhu, Q. 2020. COVID-19 sensing: negative sentiment analysis on social media in China via BERT model. *Ieee Access*. 8, (2020), 138162–138169.

[43] Warnung vor Flüchtlingskrise – Darum sind die Asylzentren schon jetzt voll: https://www.tagesanzeiger.ch/darum-sind-die-asylzentren-schon-jetzt-voll-902272860197. Accessed: 2022-12-03.

[44] Zahera, H.M., Elgendy, I.A., Jalota, R. and Sherif, M.A. 2019. Fine-tuned BERT Model for Multi-Label Tweets Classification. *TREC* (2019), 1–7.

A APPENDIX

Table 3: Topic Description and Distribution

Topic Name	Summary of Needs and Problems identified within the Topic	Number of Messages assigned
Medical Information	*How to use doctoral services within Switzerland?* *How to find specialist within the medical field?* *Information on which medical services are insured in Switzerland*	*9266*
Asylum Information and 'S Permit' -Application	*Where to find the nearest refugee camp?* *How to apply for the 'S Permit' within Switzerland for a temporary stay?* *How to prepare for the interview, granting the 'S Permit'?* *Is Switzerland still intaking refugees for Asylum?* *If an 'S Permit' is obtained, what are the rights of a refugee, e.g., can I travel abroad and return to Switzerland?*	*8262*
Transportation between Ukraine and Switzerland	*How to travel as a person between Switzerland and Ukraine?* *How to transport goods between Switzerland and Ukraine?* *How can a person return to Ukraine?* *Formation of private -sometimes commercial- initiatives organizing travel between Switzerland and Ukraine.*	*8123*
Social Services	*How and where to apply for social services?* *How much does Swiss Federation pay single refugees as well as families?* *Need of expert knowledge of 'special' rules regarding Swiss social services, e.g., "I have a Ukrainian Pension, how does this effect my social services?"*	*7393*
Integration Problems	*How to integrate into Swiss society?* *Problems related to communication with locals* *Problems related to communication with other refugees offline and online, especially Russian speaking Ukrainians*	*5291*
Banking	*How to transfer funds between Ukraine and Switzerland and vice versa* *How to exchange Ukrainian Hryvnia and Swiss Francs?* *What is the exchange rate between Ukrainian Hryvnia and Swiss Francs?* *How to open a bank account?*	*4519*
Veterinarian	*How to find a veterinarian?* *What are regulations regarding pets?*	*3709*
Consulate Services	*How to reach the Ukrainian consulate within Switzerland?* *How to use services of the consulate, e.g., apply for a new passport?*	*3298*
Accommodation	*How to find a permanent accommodation within Switzerland?* *Are there any requirements to obtain permanent accommodation?*	*2603*
Train and Bus Transport in Europe	*How to plan public transportation journeys within Europe, e.g., how can I travel from Zürich to Madrid?*	*2335*
Dentistry Information	*How to find a dentist in Switzerland?* *Are dentist services insured in Switzerland?*	*1877*
Quality Refugee Camp	*What is quality of the refugee camp?*	*1600*

Topic Name	**Summary of Needs and Problems identified within the Topic**	**Number of Messages assigned**
Education	*How to send children to a Swiss School?* *How can children attend further courses outside of school?*	*1585*
Personal Care	*How to obtain services for personal care e.g., hairdresser?* *How to participate in educational or entertainment (e.g. sports) course?*	*1464*
Employment Opportunities	*How to find a job in Switzerland?* *What are the requirements to work in Switzerland?*	*1455*
Train and Bus Transport in Switzerland	*How does the Swiss public transportation system work?* *Is the Swiss public transportation system free for refugees?*	*1387*
Vaccination	*What are the vaccination requirements for COVID-19 within Switzerland and other countries in the European Union?* *How to get vaccinated against a infectious or malignant disease, especially COVID-19?*	*987*