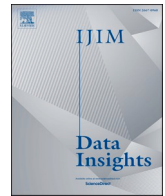



Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# International Journal of Information Management Data Insights

journal homepage: [www.elsevier.com/locate/jjime](http://www.elsevier.com/locate/jjime)

## Beyond tickets and texts: Addressing challenges in AI-driven customer service

Szabina Fodor<sup>a,\*</sup> , Csaba Csáki<sup>b</sup>, Andrea Kó<sup>b</sup><sup>a</sup> Department of Computer Science, Corvinus University of Budapest, Fővám tér 13-15, Budapest, 1093, Hungary<sup>b</sup> Department of Information Systems, Corvinus University of Budapest, Fővám tér 13-15, Budapest, 1093, Hungary

## ARTICLE INFO

## Keywords:

Natural language processing  
Text analytics  
Artificial intelligence  
Customer service  
Morphologically rich languages

## ABSTRACT

Adapting AI-driven solutions in customer services (CS) is becoming increasingly important to counter the combined impact of increased workload and management expectations. Addressing copious text volumes, text analytics (TA) assumes an important supporting role within CSs. However, the absence of TA and NLP application guidelines slows progress. Utilization of AI advancements in linguistically intricate languages and less digitally developed regions presents challenges.

Using a systematic in-depth review of the literature and expert feedback, this study presents the TA/NLP landscape in the context of CS. The paper discusses challenges, open issues, and active research areas in TA/NLP, mainly in customer service, and validates its findings with a country example and industry experts. In addition to the theoretical and practical value of the AI-driven customer service landscape, a unique methodological approach, which includes AI-assisted content mining with BERTopic, systematic literature review, domain expert validation, and the example contributes to methodology development.

### 1. Introduction

As a result of the networked digital economy of the 21<sup>st</sup> century, the role of customer services has widened, and their importance has increased in both reaching customer satisfaction and delivering higher-quality service (Cassandra et al., 2019). The changing role and the appearance of new technologies and communication channels (web forms, online feedback, social media, chat, video chat, etc.) brought on new responsibilities, increasing pressure on this organizational function (Saber et al., 2017). The new channels, especially the general availability of various forms of social media, have increased the number of incoming messages, requests, complaints and questions. In addition, depending on the business profile, the number of inquiries to be handled can be exceptionally high for days or weeks at a time during specific "campaign" periods. Moreover, these incoming messages often need to be dealt with at short notice precisely because of the periodic nature of the issues.

One possible way out is to digitalize customer services (Correani et al., 2020). While most digitalization projects have their strategic aim to improve the quality of customer experience, they often focus on sales, marketing, and logistics aspects, while back-end support of customer

care often lags behind (Hryhorak et al., 2020; Kim & Lee, 2024). The latest artificial intelligence (AI) developments offer many new opportunities for customer service (Chung et al., 2020; Davenport et al., 2020; Inavolu, 2024; Lazo & Ebarido, 2023; Xu et al., 2020), including CRM customization, robo-advisors, recommendation systems, chatbots, self-service options, faster response time and issue resolution, operational efficiency enhancement and fraud detection. According to Huang and Rust (Huang & Rust, 2018), there are two dominant characteristics of AI that matter most to service: (1) self-learning and (2) connectivity. Self-learning in artificial intelligence means a machine that can automatically evolve with experience, like some machine learning models. Networked AI is creating the emerging phenomenon of collective intelligence, like the Internet of Things (IoT), that enables smart service-based simple sensors.

To counter the combined impact of the increased workload, especially regarding text processing and heightened expectations in customer service, there is a growing emphasis on using AI, more specifically text analytics (TA) and natural language processing (NLP) solutions, so we focus on this AI area in this article. Recently, the rapid development of TA and NLP has been transforming not only the field of scientific research but also the services companies provide and, indeed,

\* Corresponding author.

E-mail addresses: [szabina.fodor@uni-corvinus.hu](mailto:szabina.fodor@uni-corvinus.hu) (S. Fodor), [csaki.csaba@uni-corvinus.hu](mailto:csaki.csaba@uni-corvinus.hu) (C. Csáki), [andrea.ko@uni-corvinus.hu](mailto:andrea.ko@uni-corvinus.hu) (A. Kó).<https://doi.org/10.1016/j.jjime.2026.100395>

Available online 11 February 2026

2667-0968/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

all aspects of our lives (Chin et al., 2023). Generative artificial intelligence tools, in particular, Large Language Models (LLM) such as OpenAI's ChatGPT (Ray, 2023) promise to provide effective solutions in customer service as well.

Given the nature of the work of customer helplines, like the need for written and spoken communication and the pressure to process vast amounts of text, various forms and techniques of NLP, text mining (TM) and text analytics enjoy special attention (Borg et al., 2021). Text mining is a semiautomated process of extracting patterns from large amounts of unstructured data sources, while text analytics is a broader term; it is text mining extended with information retrieval (Sharda et al., 2021). Text mining and text analytics are often used interchangeably, so we use text analytics in this paper. However, the application of these can bring additional burdens, especially in the case of a failed AI solution. The most common option in customer service context is the use of chatbots (Klopfenstein et al., 2017), which come in three forms: rule-based (question-and-answer or keyword search), NLU-based (limited) open-text chatbots/voicebots/phonebots, and the most recent LLM-based pretrained AI models. NLP is an essential component of TA and is a field of AI and computational linguistics. NLP addresses the problem of "understanding" human language, formalising it to make it easier for computer programs to process (Sharda et al., 2021). In customer service fields, it could be used in many ways for the interpretation of telephone conversations, the analysis of social media posts, or the categorization, preparation or even answering of incoming messages based on their content (Belk et al., 2023; Hese, 2022; Kumar et al., 2021).

While the latest AI-driven NLP solutions are widely discussed in the literature, these works are often not connected to the daily tasks of customer service. Managers and decision-makers of customer service units are not always able to utilize the latest results of AI research, especially in the case of small languages and in those countries where digitalization is not in a mature phase (Kim, 2023). The most advanced NLP and TA solutions are developed for the English language, while solutions available for small languages are fewer, typically due to the size of the market (Joshi et al., 2020). Another barrier to spreading these solutions is the level and status of digitalization in a given country, especially when the adoption of relevant digital technologies is low for industry and the public.

This paper assesses the utilization of text analytics and NLP solutions in customer service based on literature feedback from customer service experts, and using an example country, the case of Hungary. It details the suitable and typical NLP solutions in a particular customer service domain and the possibilities for an extension. The rapid growth of NLP and text analytics can only be fully exploited in customer service if we understand the current state of their application. This research helps to assess such a situation, and beyond this snapshot, it discusses the challenges, barriers, and opportunities of using TA and NLP in customer service.

The main contribution is that the paper provides an overview of the utilization of NLP solutions in customer service through a unique methodological approach. This approach includes a systematic literature review, BERT-based topic modelling, and validation of the results with domain experts. Finally, trends and challenges for customer service in TA and NLP utilization are identified based on the exploration.

## 2. AI in customer care - emerging technologies and linguistic challenges

In recent years, customer service has become heavily reliant on technology and has experienced significant changes, primarily due to the advancements in artificial intelligence technologies (Huang & Rust, 2018). On the analytical side of customer care, the application of TA, information extraction, and various knowledge-representation techniques, while from a communication point of view, speech recognition, voice generation, question/answering, and chatbot-based techniques

are being discussed in the literature (Fu et al., 2021; Saberi et al., 2017) as potential AI-based language technologies and tools that are becoming mature and ready for deployment in customer service contexts. The most popular solutions are based on various question-answer techniques, while the fastest-growing segment is chatbots (Mnasri, 2019; Nguyen, 2019). The underlying models are typically based on machine learning (ML) approaches, which are usually utilized as supervised learning, which requires tagged training data (often created due to laborious manual work). In simple cases, a smaller training set might be adequate, but performance can be improved using more extensive data sets and longer training processes. While "over-training" could be an issue, another challenge relates to the length of the text, which is often limited by the size of the model's input layer.

Beyond the traditional machine learning techniques, growing interest is gained by the more complex solutions such as GPT and BERT models and their spin-offs. These massive models utilize deep learning based on a so-called "transformer" architecture using "attention-based" model elements and build a language model using an auto-regressive approach. They have potentially hundreds of millions (or even billions) of parameters, use massive datasets of Terabytes in size for training, and run on dedicated, expensive hardware. Building such a language model thus requires a lot of resources: time, money, technology, and human labour, including specialized expertise (Kim, 2023). The resulting solution, however, can extract information related to the context of text pieces and be able to interpret words and sentences in relation to their immediate or even broader textual environment (depending on the number of input parameters). A pre-trained generic model can be fine-tuned to a specific goal. Although they are available online for limited free use and there is a lot of research focusing on investigating their abilities, their application in the context of customer care is yet to be explored (Senese et al., 2020).

Since most techniques and models mentioned above are resource intensive (including a massive input corpus to be prepared), one issue considers the limited availability of solutions for small languages. An emerging research area at the intersection of NLP applications and customer care addresses the issue of small languages (Joshi et al., 2020). The problem is rooted in the size of the market, that is, the population speaking a given language. In the context of digital presence, "small" is understood as a language with less than ten million speakers. Such small languages are disadvantaged both from a technical and an economic point of view. In a small market, it is not immediately economically feasible to develop NLP technologies as they require considerable time and effort to develop a sufficient language-specific base (the language corpus).

Even with an acceptable starting corpus, domain or profession-specific frameworks (such as a legal, health, or telecom corpus) would also be needed before the final development of a particular application or solution. Considering the European Union, only 5 of its 27 members have more than 20 million inhabitants (speaking the same language), and 15 have a population below 10 m. This issue also impacts several Asian and African countries as well.

The size of the target population also impacts the cost of development compared to the potential number of users. In addition, small languages would more likely face problems rooted in the lack of general NLP solution tools that special industries could rely on. Small languages often do not hit the break-even point of the economies of scale for manufacturers and service providers. The result is that the choice and the quality of applications for small languages are vastly poorer than that of large ones (Akdemir et al., 2022; Chiang & Trimi, 2020). There is no place here to compare commercial products, but a quick look at the number of articles in the English and Latvian Wikipedia will be convincing enough (2567,509 versus 17,527).

In addition, in small countries and the context of smaller languages, smaller organizations face extra hurdles. This does not mean that big firms do not have difficulties, but they, especially those backed by multinationals, will have more resources to at least try to utilize the

advantages of AI and NLP. Still, skill shortages further add to the difficulties. One additional challenge considers linguistic accuracy: although beyond a certain point, which is not necessarily a high priority to all applications of NLP. Suppose the task is summarising the content of many documents or extracting their elements, and the client accepts the risk in exchange for a quicker or cheaper survey. In that case, a recall rate of 80 % might be sufficient for many uses (but, for example, 70 % may not).

To demonstrate the point, one may consider the specific context of Hungary, where the challenge of supporting customer service digitalization discussed so far is clearly present, making it a suitable example case. Most NLP solutions are developed for English, and Hungarian ones are rare. As is typical in such countries, there are only a few working AI language models, even with the emergence of attention-based generative LLMs (and even reports about them do not reach the general scientific community, as papers are typically written in their native tongue). In Hungary, there are only two widely known models, and they are indeed reported in the local language (Nemeskey, 2022; Yang, Dodé, Héja, et al., 2024). Even if reports are published in English (Yang, Dodé, Ferenczi, et al., 2024), they get little attention in the mainstream. Additionally, text analytics and NLP solutions for the Hungarian language raise some computer linguistics issues, which must be managed. The Hungarian language is morphologically rich and agglutinative, like Finnish, Estonian, and Turkish. Words with 3-4 suffixes are quite common, and morphological units include essential information rather than syntax (Akdemir et al., 2022). The additional constraint of spreading these solutions is the level and status of digitalization. Hungary was positioned 23rd out of 27 EU Member States in the Digital Economy and Society Index (DESI) 2021, while the rank was 22nd in 2022. Although these challenges are specific to the Hungarian customer service sector, they can be generalised and interpreted broadly, as will be shown in the findings and discussion sections.

### 3. Research design

Based on the previous theoretical review, our main research objective was to identify the key challenges and opportunities of using text analytics to support customer services as part of digitalization efforts. To this end, the research questions put forward by the work reported here to address the objective are:

**RQ1:** In what key industrial domains are text analytics techniques utilized during customer service digitalization?

**RQ2:** What text analytics techniques can support customer service operations?

**RQ3:** What main customer service channels are relevant when using text analytics techniques?

**RQ4:** How do customer service managers in a small-language country see the applicability of text analytics techniques?

The research used a hybrid approach to achieve the objective: it mixed qualitative and quantitative techniques. The research method consisted of three main phases (see Fig. 1). RQ1 and RQ2 are answered

during Phases 1–2, while RQ3 and RQ4 are answered in the discussion of Phase 3.

#### 3.1. Phase 1: domain specification

To ascertain pertinent study areas, we conducted an AI-assisted content mining analysis of scholarly findings published over the past decade pertaining to the field of customer service. We collected the author, title, year of publication, abstract of refereed journal articles, and conference proceedings from the Scopus database. The following search term (Search-term#1) was used to locate articles:

**TITLE-ABS-KEY ("customer support") OR TITLE-ABS-KEY ("contact center") OR TITLE-ABS-KEY ("call center") OR TITLE-ABS-KEY ("customer helpline") OR TITLE-ABS-KEY ("customer service")**

A list of 12,140 records published between January 2013 and February 2024 was extracted. The collected data were pre-processed, duplicate (n=27) and non-English records (n=1667) were removed, and records without abstracts (n=333) were deleted. This left a total of 10,113 records. We concatenated the title with the abstract for each record and considered it our corpus. (The specific steps are detailed in Appendix Fig. A3).

This work utilizes BERTopic (Grootendorst, 2020a, 2022), a cutting-edge topic modelling technique to extract topics. Our topic modeling approach involved four steps.

- First, the corpus was converted to its embedding representation a 384-dimensional vector space using a pre-trained sentence-transformers language model "all-MiniLM-L12-v2".
- Second the dimensionality of the resulting embedding is reduced to optimize the clustering process. Although principal component analysis (PCA) is perhaps the most popular method for reducing dimensionality, the uniform manifold approximation and projection (UMAP) method was used to maintain the local and global structure of the data sets. Based on UMAP, our research selects the 15 nearest neighbours to emphasize local structures preferentially.
- Third, the reduced embedding is clustered using hierarchical density-based spatial clustering of applications with Noise (HDBSCAN), which allows noise to be modelled as outliers. This prevents unrelated documents from being assigned to any group. The minimum size of the clusters is set at 55 to identify a sufficiently large number of articles to increase the relevance of our results.
- As a fourth step to characterize the resulting clusters, we merged the abstracts belonging to the clusters. Using quantized LLM technology, including KeyBERT extraction technique (Grootendorst, 2020b; Sammet & Krestel, 2023), and the class-based frequency-inverse document frequency approach (c-TF-IDF) to characterize the forty-three resulting clusters. Quantized LLM was used for fine-tuning and enhancing the accuracy of topic representation. Our study utilized the pre-trained LLMs "Open-Hermes-2.5-Mistral-7B-GGUF". A prompt has been established for

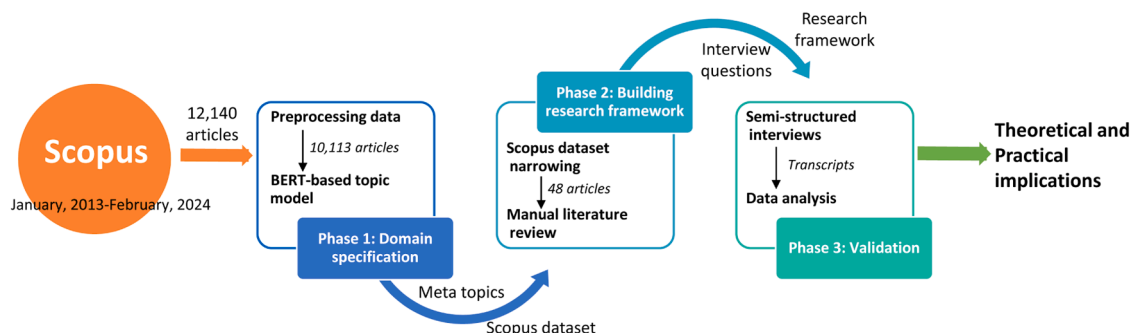


Fig. 1. Research phases (Source: prepared by the authors).

LLM to utilize in generating topic labels derived from clustering keywords and abstracts associated with the respective topic.

### 3.2. Phase 2: building research framework

Following the topic clusters resulting from the BERTopic analysis based on a comprehensive set of papers, a more focused search was executed on the Scopus database to identify the text analytics techniques that support customer services. The selection of articles was made in four steps as follows (The specific steps are detailed in Appendix Fig. A3):

1. *STEP*: the search string used in the first phase (Search-term#1) was augmented with adding text analytics techniques related to keywords and the publication type restricted to articles,

(Search-term#1) AND ((TITLE-ABS-KEY ("nlp") OR TITLE-ABS-KEY ("natural language processing") OR TITLE-ABS-KEY ("text mining") OR TITLE-ABS-KEY ("text analytics"))

This resulted in a total of 168 records, from which duplicates (n=3) and non-English articles (m=17) have been removed.

2. *STEP*: the title and the abstract of the resulting reference set of 148 were manually reviewed; those not relevant to the research were excluded (k=33),

3. *STEP*: for the remaining 115 articles, full texts were available for 97, which were subsequently downloaded and examined.

4. *STEP*: upon reviewing the complete text, publications featuring literature reviews or solely comparing various technologies were excluded from the list. The remaining 48 pertinent articles were manually classified.

To augment the literature categorization, interviews were conducted with key representatives of three leading Hungarian consulting firms specialized in providing customer care improvement services. These leading experts were engaged to identify and define critical areas relevant to customer service digitalization and text analytics in this language context. Interviewees were selected with the support of the Hungarian Artificial Intelligence (AI) Coalition. We participated in the "AI in customer service" workgroup research, where we dealt with the challenges and barriers of NLP applications at the Hungarian Customer Service. We had an expert team (dedicated representatives from the main Hungarian customer service centres and their NLP providers) in this workgroup. Interview questions were generated together with these experts.

### 3.3. Phase 3: validation

Based on the research framework (see Fig. 1), the related literature review and the experts' feedback, we put together a data collection method including a more profound survey instrument (see questions in Appendix Table A1) and options to validate and expand on the findings

**Table 1**  
Facts of the expert interviewees.

Position	Relation with customer service	Industry	Size	Owner
Sr. manager	Supervisor	Government	Very large	State
Sr. manager	Supervisor	Retail	Very large	Foreign
Sr. manager	Supervisor	Retail	Very large	Foreign
Sr. manager	Supervisor	Transportation	Large	Municipality
Sr. manager	Supervisor	Finance, insurance	Large	Foreign
Sr. manager	Supervisor	Finance, insurance	Very large	Foreign
Sr. manager	Advisor	Consultancy	Small	Domestic
Sr. manager	Advisor	IT consultancy	Medium	Domestic
Manager	Supervisor	Finance, insurance	Medium	Foreign
Manager	Supervisor	Finance, insurance	Very large	Foreign
Manager	Customer service	Finance, insurance	Very large	Domestic
Manager	Customer service	Finance, insurance	Very large	State
Manager	Advisor	Consultancy	Medium	Domestic
Employee	Supporting	Telecommunications	Very large	Foreign
Sci. advisor	Advisor	Telecommunications	Very large	Foreign

of the first two phases. This validation was done in a selected country, Hungary (as justified in Section 2). Participating organizations were determined to be approached based on representativeness in the Hungarian economy. Data were collected from participants through semi-structured interviews (see Table 1). The interviewees were international experts, and the companies surveyed were Hungarian, but more than half of them were foreign-owned.

## 4. Findings

### 4.1. Findings of phase 1

Running BERTopic on the Scopus data collected resulted in 43 topics being extracted. The two-dimensional visualization of the resulting topics (see Fig. 2) shows their topological structure. The topics created are succinctly outlined in Table A2 in the Appendix. Each topic is accompanied by a label assigned by the LLM and a set of 10 keywords defined by KeyBERT.

The top 5 topics account for more than 50 % of all abstracts considered, and the top 10 topics account for more than two-thirds of all scientific papers (Fig. 3 Panel A). If we look at the change in the number of abstracts for the top 5 topics over time (see Fig. 3 Panel B), we see that the number of topics has been significantly affected by the COVID-19 pandemic that hit in 2019. This could explain the surge in articles on health (Topic 4) and chatbots (Topic 0), as well as the steep decline in tourism-related topics (Topic 1). Service quality related to waiting times (Topic 2) has become less interesting in recent years, while service quality related to online shopping and customer satisfaction (Topic 3) has increased interest.

The 43 topics were manually categorized into seven meta-topics, namely *Technology*, *Communication channel*, *Industry*, *TA task*, *Service quality*, *Service characteristics*, and *Well-being of customer service employees*. The research domain was defined based on four meta-topics (*Technology*, *Communication channel*, *Industry*, *TA task*), as the topics of *Service quality*, *Service characteristics*, and *Well-being of customer service employees* were irrelevant to this research. Within these four meta-topics of interest for our research, we have identified five technology categories, three communication channels, thirteen industry categories and six NLP task categories.

### 4.2. Findings of phase 2

Executing the Scopus search string returned 148 items after removing non-English and duplicates, which were previewed by title and abstract, and found that 97 articles were relevant to the research. These 97 articles were evaluated on the basis of their full content. In the manual assessment, 11 articles were literature reviews, while 38 articles did not directly contribute to customer service, leaving the remaining 48

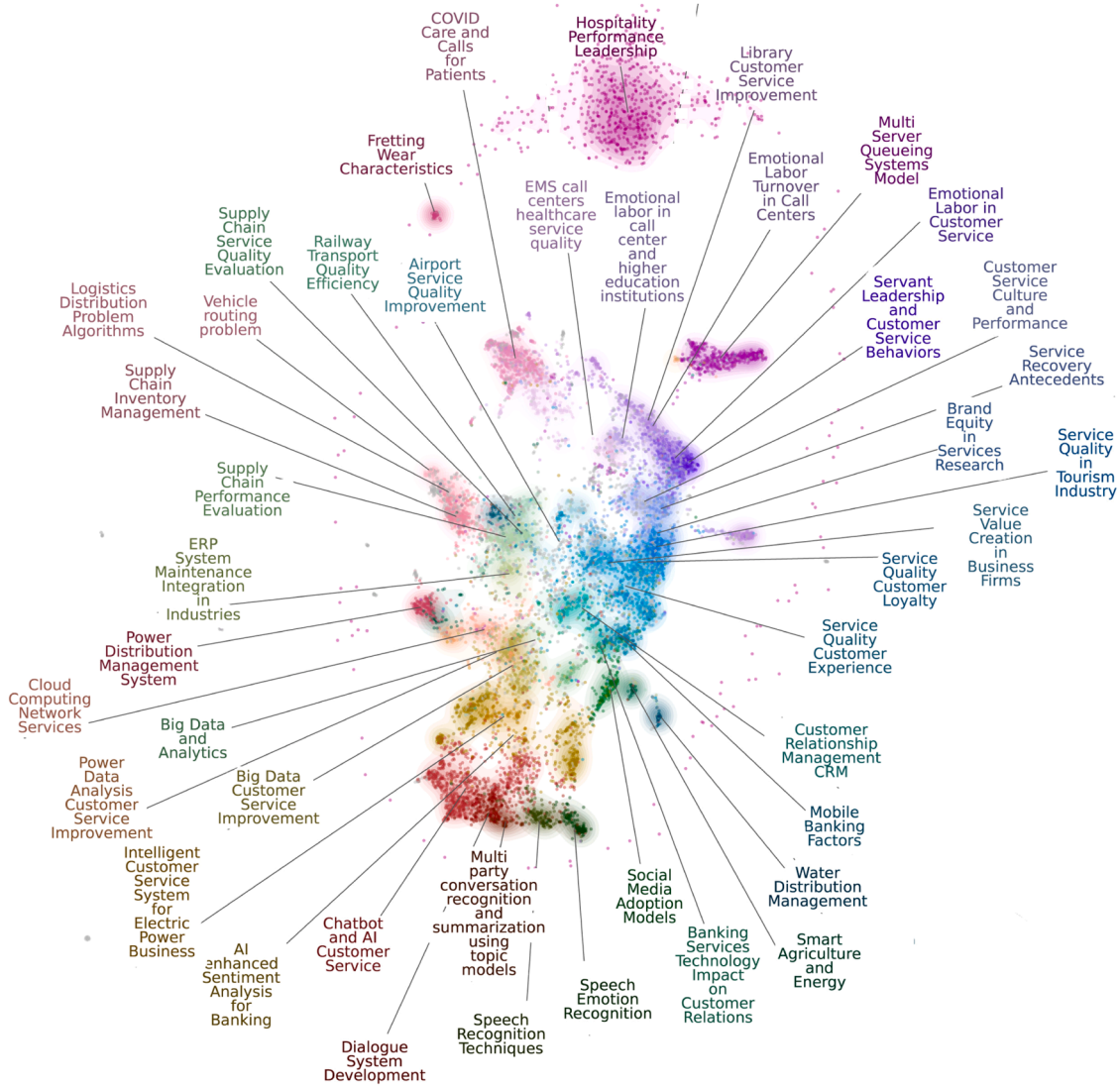


Fig. 2. Distribution of 43 topics in a two-dimensional space. (Source: prepared by the authors).

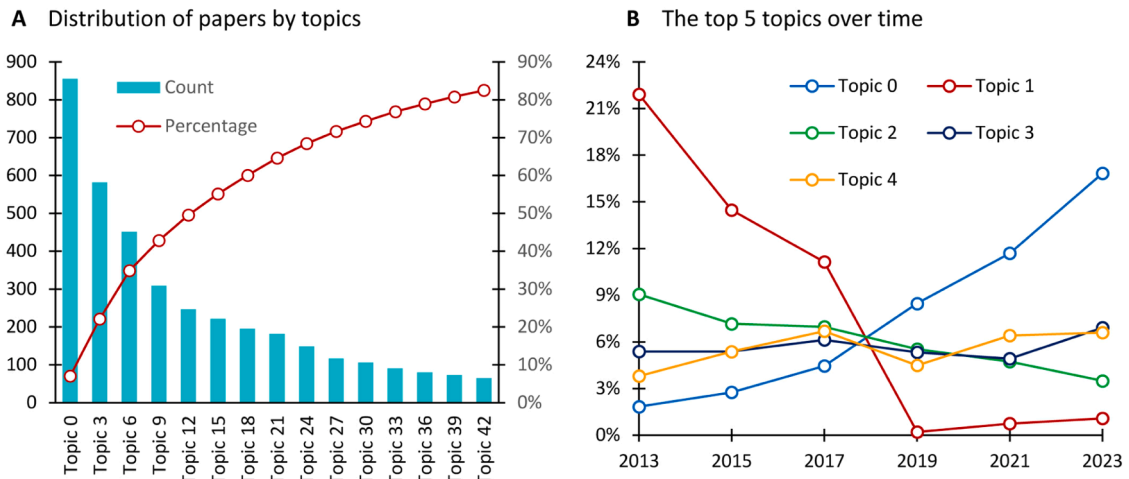


Fig. 3. The distribution of papers by topic. The top 5 topics Chronological change in publishing in the top 5 topics (Source: prepared by the authors).

relevant journal articles to work with (see also in Appendix Fig. A3). All 48 papers were reviewed and categorized using all three dimensions. BERTopic topics were used as categories when possible; otherwise, new categories were established as needed. An appropriate name was created for the category (within its corresponding dimension) in both cases. As a result:

- The *Industry* dimension had 14 categories (10 from BERTopic analysis covering 19 BERTopics, while six new ones were manually identified).
- The *Text Analytics Task* dimension ended up having 11 categories (5 of which were from BERTopic covering 7 from BERTopics, and 6 new, manually identified).
- The *Channel* dimension ended up having six categories (3 of which were from BERTopic and 3 new, manually identified).

Overall, the manual analysis covered 27 from BERTopics (out of 43), while 17 new categories were identified manually. Table 2 shows the final three-dimensional paper groupings after manually categorizing all articles. Among the industries, finance was cited highest with 7, followed by IT, retail, and aviation with 6, 5, and 4. Manufacturing, logistics, and engineering had 3–3 each. Nearly 20 % of the articles did not specify a domain or industry from where they acquired their data or used broad customer care data, such as Twitter comments from numerous companies. IT services, engineering, logistics, and aviation were separated and picked up by the preceding phase. However, retail, public sector, and e-commerce were not. Text analytics seems to be used for a broad spectrum of domains and in many different industries. None of the fourteen industries listed in Table 2 seems to dominate (although Logistics was well covered by the BERTopic analysis).

Classification, topic modeling, sentiment analysis are the most frequent TA tasks, with 13, 11, 9, and 8 articles, respectively (41 in 67 appearances, over 60 %). Knowledge management and Q&A follow with 5 (7–7 %). Four other activities occurred 1–4 times, while 6 worked on less popular tasks like voice mining or machine reading. The tasks are much more unequal than Industry. Tagging and information extraction seem pointless, too. Conversely, sentiment analysis and opinion mining "hype" impacted the list.

It is not surprising that chat channel is a popular means of allowing customers to access helpline services, as 13 of the 48 papers (more than 25 %) reported using this. It might be somewhat surprising, however, that Chat is as popular as another well-investigated channel, Social media (also with 13). This is because many NLP research uses Twitter data as a starting point for analysis. Phone access, Online feedback, and Email appeared to have high importance in technology-based Customer support, with these channels being mentioned 7, 5, and 4 times respectively. Quite a few (6) AI-text research covered several (or even all) channels. These works took an integrated view and typically investigated a mix of tasks too.

After analysing the findings of the systematic literature review (outlined in Table 2), informal discussions were conducted with three Hungarian experts involved in the customer service industry to validate further data collection. They confirmed that

- there is a perception that core activities can function without customer service;
- CS is often seen as a hygiene factor;
- innovation in this area has been absent in the last few years;
- helplines in Hungarian enterprises rely heavily on personal contact rather than sophisticated AI solutions like chatbots;
- Hungarian firms do not formally utilize online feedback and social media for customer service helplines but rather for marketing and informational purposes;
- while chat (such as online chat started from corporate web pages) is growing, chatbot use is not customary, and penetration is low (from the consulting experts' clientele of 250 leading Hungarian

**Table 2**

Manual sorting of 48 papers into categories along the three dimensions (BT ID is the ID of the BERTopic from Phase 1).

MANUAL CATEGORIES	BT ID	REFERENCES
<b>INDUSTRY</b>		
Aviation	28	(Gunarathne et al., 2022; Liao & Tan, 2014; Lucini et al., 2020; Zhou et al., 2022)
e-Commerce	None	(Palese & Usai, 2018)
Utility	20;21;26;36	(Papadia et al., 2022; Rebelo et al., 2022)
Manufacturing and Engineering	40	(Akella et al., 2017; Ittoo & Bouma, 2013; Ohata et al., 2022)
Finance, banking, insurance	19; 33;42	(Aattouri et al., 2023; Asano et al., 2017; Chakrabarti & Luger, 2015; Fan & Ilk, 2020; Liao et al., 2020; Quattrocchi et al., 2023)
Healthcare, fitness	4;6	(Kelahan et al., 2016; Lal & Neduncheliyan, 2023; Ram Kumar et al., 2023)
IT and telecommunication	None	(Ali Zaidi et al., 2022; Borg et al., 2021; Ilk et al., 2020; Jabr et al., 2014; Kecht et al., 2023; Tanaka et al., 2019)
Marketing	38	(Mahr et al., 2019; Ramsey & Bapna, 2016)
Public sector	None	(Bruni et al., 2023; Nirala et al., 2022)
Retail	None	(Ashton et al., 2013; Manchiaiah et al., 2019; Mao, 2021; Ngai et al., 2021; Zhan et al., 2021)
Logistics and Supply chain	9;17;25;35;39;41	(Aman et al., 2021; Luo & He, 2021; Wang et al., 2023)
Tourism, hospitality	1;13	(Chen & Tussyadiah, 2021; Kar et al., 2021)
Fast food	None	(Rose et al., 2019)
General	None	(Hardalov et al., 2019; Lin et al., 2023; Müller et al., 2016; Nordheim et al., 2019; Przegalinska et al., 2019; Raeesi Vanani, 2019; Shah et al., 2023)
<b>TEXT ANALYTICS TASK</b>		
Categorization	None	(Akella et al., 2017; Ali Zaidi et al., 2022; Borg et al., 2021; Bruni et al., 2023; Gunarathne et al., 2022; Kelahan et al., 2016; Lucini et al., 2020; Mao, 2021; Ohata et al., 2022; Przegalinska et al., 2019; Quattrocchi et al., 2023; Ramsey & Bapna, 2016; Rose et al., 2019)
Clustering	36	(Chen & Tussyadiah, 2021; Kar et al., 2021; Liao & Tan, 2014; Mahr et al., 2019; Manchiaiah et al., 2019; Palese & Usai, 2018; Rebelo et al., 2022; Teng & Khong, 2021)
Concept linking	None	(Fan & Ilk, 2020; Mahr et al., 2019; Manchiaiah et al., 2019)
Information extraction	None	(Ashton et al., 2013; Chen & Tussyadiah, 2021; Ilk et al., 2020; Ittoo & Bouma, 2013)
Knowledge management	21	(Asano et al., 2017; Chakrabarti & Luger, 2015; Mahr et al., 2019; Ngai et al., 2021; Zhou et al., 2022)
Question answering	None	(Hardalov et al., 2019; Kecht et al., 2023; Lal & Neduncheliyan, 2023; Rebelo et al., 2022)
Speech recognition	22	(Aattouri et al., 2023; Asano et al., 2017)
Sentiment analysis, Emotion recognition and Opinion mining	5;24	(Aman et al., 2021; Kar et al., 2021; Liao et al., 2020; Liao & Tan, 2014; Lucini et al., 2020; Raeesi Vanani, 2019; Ram Kumar et al., 2023; Shah et al., 2023; Zhan et al., 2021)
Tagging	None	(Mao, 2021)
Topic modelling	37	(Aman et al., 2021; Kar et al., 2021; Liao et al., 2020; Lin et al., 2023;

(continued on next page)

Table 2 (continued)

MANUAL CATEGORIES	BT ID	REFERENCES
Other	None	Lucini et al., 2020; Luo & He, 2021; Mahr et al., 2019; Müller et al., 2016; Palese & Usai, 2018; Papadia et al., 2022; Zhan et al., 2021) (Jabr et al., 2014; Nordheim et al., 2019; Tanaka et al., 2019; Teng & Khong, 2021; Wang et al., 2023)
COMMUNICATION CHANNEL		
Chat	0;12	(Asano et al., 2017; Chakrabarti & Luger, 2015; Hardalov et al., 2019; Ilk et al., 2020; Lal & Neduncheliyan, 2023; Ngai et al., 2021; Nirala et al., 2022; Nordheim et al., 2019; Przegalinska et al., 2019; Rebelo et al., 2022; Shah et al., 2023; Wang et al., 2023; Zhou et al., 2022)
Comment, review, online survey	None	(Ashton et al., 2013; Jabr et al., 2014; Lucini et al., 2020; Manchaiah et al., 2019; Palese & Usai, 2018)
Email	None	(Ali Zaidi et al., 2022; Borg et al., 2021; Ittoo & Bouma, 2013; Ohata et al., 2022)
Phone	2;4; 6;24; 37	(Aattouri et al., 2023; Fan & Ilk, 2020; Kelahan et al., 2016; Liao et al., 2020; Müller et al., 2016; Papadia et al., 2022; Ramsey & Bapna, 2016)
Social media	14	(Aman et al., 2021; Chen & Tussyadiah, 2021; Gunarathne et al., 2022; Kar et al., 2021; Kecht et al., 2023; Liao & Tan, 2014; Luo & He, 2021; Raeesi Vanani, 2019; Ram Kumar et al., 2023; Rose et al., 2019; Teng & Khong, 2021; Zhan et al., 2021)
All (or most – e.g. any text in general)	None	(Akella et al., 2017; Bruni et al., 2023; Lin et al., 2023; Mahr et al., 2019; Quattrocchi et al., 2023; Tanaka et al., 2019)

companies, 1/3 (78) had chat function on their website, but only 10 % (24) had chatbots (only 3 used AI).

They recommended three more typical communication channels relevant to this context that were not apparent in the literature: in-person, video chat, and mobile app.

#### 4.3. Findings of phase 3

A survey instrument was developed and implemented using 15 semi-structured interviews (see Appendix A2). The answers showed that telephone customer services are dominant, with almost half of respondents' workplaces, 43 %, using them predominantly. This is followed closely by two channels: face-to-face customer service (22.5 %) and email (19.6 %). These are followed by the mobile app channel (8.7 %), which is well behind. Chat is only 3.3 % but is making a noticeable appearance as a potential future growth area.

Significant industry differences could be noticed as telecommunication is far ahead of other industries and is ahead of other digitalization solutions in general. In the financial sector, it is not necessarily true that companies wish to reduce personal contacts as they consider secondary customer effects and private banking prefers face-to-face encounters with lead customers. Although voice analytics solutions are on the market, few firms are moving in this direction. The primary argument is that there is no capacity and no time to train such solutions from scratch, and no pre-trained, "standard" systems are available.

However, respondents expect significant changes and innovations in customer services soon. In addition to assessing the current state, the

most significant changes are expected to be in the rise of the chat channel (appearing in various chat interfaces and applications) along with the decline of the in-person and phone channels – while they expect the constancy of email. Although professionals predict significant changes and progress, there is also considerable concern about the task size ahead. Interviewees unanimously perceive any development as a critical but complex step. A reassuring response to this concern is required, considering the need for training to support the development and use of AI-based solutions. Responding quickly to customer requests has become a critical challenge, and knowledge bases have been developed to support this. To speed up communication, some companies use predefined templates.

On the positive side, there is no risk of not finding suitable suppliers in the domestic market. From an end-to-end operational point of view, respondents prefer systems developed by external suppliers that are then operated in-house.

## 5. Discussion

Based on the data collection and reflecting on the literature, we have identified the following theoretical and practical implications (including challenges, open issues, and active research areas) for NLP in the context of CS.

### 5.1. Practical implications - challenges

The field of NLP has significantly improved recently, mainly because of the appearance and spread of LLMs, especially ChatGPT, but most companies are not satisfied or ready to use them because of their limitations and potential but often hidden biases (Hadi et al., 2023; Ray, 2023; Tamkin et al., 2021). This observation in the literature is further emphasized by our experts, who confirmed that several issues and limitations could be identified concerning these solutions. These include unavailability of training data, quality issues of the data source (incomplete or outdated knowledge), lack of domain knowledge, difficulties with managing factual accuracy, lack of handling the context, moral and ethical challenges, inability to incorporate users' expertise, lack of customization regarding feedback, and issues with meeting resource requirements for training and deployment.

#### 5.1.1. Training data issues

Morphologically rich and agglutinative languages have special requirements regarding training data, and they need about four times more data during the training to have equal results with the English language (Akdemir et al., 2022). Training data for particular domains is usually unavailable, and it is time- and resource-intensive to produce them. Hungarian customer services are also faced with these challenges if they plan to prepare their own LLM solutions. Supervised learning requires tagged training data, where tagging is usually provided manually. LLM is not always a perfect solution in this case because its training data may contain outliers, as it learns from unusual or extreme examples that are not representative, which can potentially distort overall modelling results. Most datasets are private, which makes reproducibility, credibility, and transparency of results difficult (Winata et al., 2022). Because of these issues, training is usually time- and resource-consuming and requires special hardware and software environments (Hadi et al., 2023).

#### 5.1.2. Data source quality issues

New language models such as ChatGPT may provide content on various topics. However, they can be incomplete or outdated and may lack the depth of expertise and knowledge in specific fields (Chiang & Lee, 2023; Liu et al., 2023). This can limit their usefulness in specialized fields or applications requiring domain-specific knowledge. Semantic approaches and applications require context management, which LLMs usually do not provide. There is also no guarantee of factual accuracy.

The interviewees also highlighted the problem of terminology quality, primarily referring to the management of context dependency (the meaning of specific terms may vary depending on context).

### 5.1.3. Customization issues

Customization is essential in customer services but requires unique development approaches in specific business situations. Customization needs structured domain knowledge (usually in the form of semantic solutions and ontologies), context management, reasoning, communication possibilities with the users, and the incorporation of the possibility of user knowledge. All these functions are not adequately available in standard TA solutions: semantic understanding is still poor, even in GPT-4 (Chang et al., 2023). Customization is also a critical factor in the context of this study. According to interviewees, it is challenging for companies to develop their LLM model or use the available solutions.

### 5.1.4. Moral and ethical issues

While NLP solutions are often poor at addressing moral and ethical considerations, recent language models do not appear to solve this problem either (Chiang & Lee, 2023; Hadi et al., 2023; Tamkin et al., 2021). Traditional ML models often fail to meet equality and fairness standards (primarily due to biases in the data used). Such problems were present in the interviews and confirmed by our experts. At the same time, the new wave of LLMs has been shown to provide content that is morally ambiguous or does not comply with ethical standards. After proper human supervision, such content can only be used in customer service applications. This reminds us that with the advent of LLMs, data cleansing did not go away; indeed, it may even be amplified due to limited resources and weaker training corpora in the case of non-central languages (studied in this research). Another concern to be addressed in the context of customer helplines using sophisticated chatbots is the potential violation of data privacy, especially in the context of the European GDPR regulation and the new AI Act. A further challenge relates to anthropomorphic chatbots, in other words, whether they should mimic human customer service agents' verbal, vocal and visual cues as much as possible or retain more "robotic" features. Indeed, it is a decision with some moral considerations whether customers should always be informed when communicating with a bot or when communicating with a human person. Pushing this line of thinking further leads to the customer service decision whether customers should be allowed the option of choosing between a human or a bot – and this should not be decided on costs only.

### 5.1.5. Actionable recommendations for AI-driven customer service in Hungary

Although a significant number of the companies we surveyed have launched or are planning to launch TA/NLP projects to support customer services, the solutions available cannot be applied to Hungarian companies on a one-to-one basis. These solutions are too general, do not adequately address the business context (lack of terminologies and taxonomies describing domain specialities), and no appropriate data sets are available for model training purposes. Accordingly, when designing TA/NLP projects, we recommend that priority is given to supporting the context-dependency of models, including the development of terminology, taxonomies and ontologies for the specific domain. The development of data sets for teaching models is also necessary. All this will increase the resource requirements, and duration of projects and will require specific expertise. Similar considerations can be made for other countries that face the same challenges, although in each case, the specific details need to be carefully investigated before drawing conclusive recommendations.

## 5.2. Theoretical implications - open issues and active research areas

The literature and the practice under study showed a diverse picture of using NLP and TA in customer service. Drawing on our assessment

and the previously discussed challenges, further open issues, opportunities, and research areas can be identified.

### 5.2.1. Ethical and moral issues

LLMs usually need vast amounts of data for training, which raises data protection issues. For the protection of user privacy and the prevention of the compromise or misuse of confidential information and data, it is necessary to implement protection measures. Responsible and ethical data anonymization techniques, data retention policies, and data consent mechanisms have to be incorporated into the development and implementation of LLMs.

### 5.2.2. Interpretability and explainability

Interpretation of LLM outputs is challenging and complex because their architectures are usually black-box models, often with billions of parameters. However, interpretability and explainability are necessary for customer services, not only in financial and healthcare domains. Trust and transparency support users to understand the reasoning behind a given model response. LLMs must be more transparent and accountable for broader application in practice. Additionally, interpretability contributes to the debugging and improvement of the model. AI interpretability and explainability are still active research areas.

### 5.2.3. Human-related functions

It is still an open question how LLMs should deal with human-like features such as emotion management, empathy, creativity, and critical thinking. These human interactions are valuable in several application fields, particularly customer service.

### 5.2.4. Training data, data quality and customization

According to Winata et al. (Winata et al., 2022) more than 60 % of the datasets used in publications are private, which impairs modelling reproducibility. Domain-specific training data sets are rare in many cases, and sometimes it is very challenging to prepare them. Considering, for example, fraud detection in a financial context, it is difficult to get proper quality data for training. Synthetic data generation can provide a solution, but its usage raises additional questions. More public datasets, especially domain-specific ones, are required to enable a wider-availability of LLMs. Tests should be developed for data quality checking and management. Developing a system of comparative indicators for training data sets and LLMs would also be helpful. Improvements in the integration of semantic layers would facilitate customization.

### 5.2.5. Management of the resource and the time-consuming nature of LLMs

Operational issues, like the efforts required by LLMs, are providing additional research directions. It would be beneficial to investigate the resource and effort requirements of various specific architectures, including ways of optimization. Is there any part of text processing worth automatizing, and what are the consequences of resulting organizational and process changes regarding resource and time management? According to respondents, the technology is available, and companies are enthusiastic, but are not exploiting the opportunities yet. Implementation is slow due to the known issues. In addition, customer service managers are not necessarily aware of the new NLP developments and have limited power to initiate projects. Respondents also highlighted the importance of training and the need to change the corporate culture.

### 5.2.6. Theoretical advancements for AI-driven customer support

From the theoretical point of view, based on the literature review and the discussion with the experts, we set up a three-dimensional (industry- task- communication channel) model of the applications of text mining in services management (see Table 2). We didn't find a similar approach in the literature. Compared with the previous studies (e.g. Kumar et al. (2021)) our approach provides a more comprehensive and

detailed view for the investigation of the applications of text mining in services management. According to the discussion with domain experts, the theoretical advances in AI-driven customer service for small languages aim to overcome the obstacles presented by limited linguistic resources and data scarcity. Traditional AI models, such as those based on large-scale neural networks, frequently necessitate massive volumes of annotated data to attain excellent performance, which is uncommon for tiny or underrepresented languages. Recent theoretical advances stress using transfer learning, multilingual models, and low-resource machine learning techniques to close this gap. For example, pre-trained multilingual models such as mBERT (Pires, 2019) and XLM-R (XLM-RoBERTa) (Conneau, 2019) use cross-lingual transfer to apply information from high-resource languages to smaller ones. Furthermore, advances in few-shot and zero-shot learning paradigms enable AI systems to generalize efficiently with less input, making them better suited to small languages.

Another important theoretical area is the integration of unsupervised and semi-supervised learning methods, which reduces the reliance on labelled datasets. Self-supervised pretraining and data augmentation are being investigated as methods for producing synthetic data or extracting relevant patterns from constrained corpora. Furthermore, researchers are looking into the usefulness of linguistically informed models, which incorporate grammatical and syntactic rules unique to tiny languages, improving accuracy and contextual comprehension. These theoretical advances not only increase the feasibility of AI-driven customer support for small languages but also promote linguistic inclusivity, ensuring that AI technologies can serve diverse global populations equitably.

### 5.3. Limitations and future work

Limitations of this research are related to the specialities of the Hungarian language (small and morphologically rich). However, the research method and the phases of our research can be repeated for other languages. Another limitation is the sample size of the interviews. Future work can include the same investigation for other languages and the investigation of the possibility of generalization. It is worth repeating the same research in Hungary and monitoring the progress of customer service digitalization in terms of using text analytics.

## 6. Conclusion

Our research shows that advances in AI, especially in TA and NLP, have fundamentally changed the way customer service works. Regarding the Industry dimension of our analysis (see Table 2), there is a good spread of different domains, and while IT, as well as retail or airlines, are apparent areas where one would expect the use of advanced technologies, the area of energy or engineering might be a bit surprising and less evident. No sophisticated NLP functionality seemed to be widely used, such as speech recognition or semantic nets. While opinion mining is prevalent, this can be considered a simple technical solution, as most papers reflected on positive-negative judgements without looking deeper and providing any elaborate analysis of emotions. They search for words and expressions that relate to a narrow set of opinion indicators but do not look at the emotions behind those words, at least in CS applications.

Regarding the use of various channels, human chat – either online or as video chat – appears to replace basic phone conversions. Chatbots are not widely used and are not expected to grow significantly. This is due partly to COVID, since during the pandemic, video chat became an option to replace customers visiting the branch, as it allowed for proper identification.

Confirming the validity of our multi-stage methodology, the discussion with the experts yielded several new results that were not apparent from either the BERTopic-assisted or the manual literature review. The complex methodology we follow is a unique approach compared to traditional literature analysis. In describing the findings of Phase 3, it

was mentioned that three additional channels were identified on top of the channels that appear in the literature. This is understandable, as the focus of the literature research was to investigate the role of NLP in customer service. Having these channels put forward by experts, on the other hand, indicates that many customer services in a country with a small language still rely on traditional channels - primarily due to the lack of advanced technologies. They may only be supported if communication through those channels is available in a format that is accessible by advanced TA solutions. These channels might be missing from international literature because they are not that relevant anymore, but they are still part of the practice of less digitally advanced countries.

Furthermore, employees working in customer services and managers of such departments also reported that customer services are often a "hygiene" factor, a compulsory constraint imposed by law. So, companies are often content to settle for a "good enough", acceptable solution. Customer helpines are easily seen as cost factors to be minimized and not being considered to directly affect the bottom line. It is not always clear what its role is and how it impacts customer retention, churn, or even cross- and up-sales opportunities. There may also be perceptual and attitudinal problems behind the lack of clunky automation and AI solutions. Overall, the CS area is often not a priority in the digitalization projects of companies. Customer Care units do not have the departmental budgets to initiate digitalization projects. Customer services depend on other core departmental digitalization initiatives and financial resources for technological improvements in the back-end work. This is in line with the observation by Hryhorak et al. (Hryhorak et al., 2020) mentioned in the Introduction section.

Another interesting finding is that companies do not even plan to automate specific customer service processes. Typically, these are processes where the sector prefers a personalized customer relationship. Large companies are observed to operate two types of customer service, one for retail and SME, where they are more likely to use automated, AI-supported solutions, and the other for so-called "accounts" (highlight customers), which are managed by account managers and which they do not even intend to automate.

### Data & code availability

All data and code supporting the findings of this study are publicly available in the Zenodo repository at <https://zenodo.org/records/18374192>. The repository includes topic-modeling artifacts (sentence embeddings, UMAP and HDBSCAN outputs), as well as complete parameter settings and random seed configurations, enabling deterministic re-runs and full reproducibility of the results.

### CRedit authorship contribution statement

**Szabina Fodor:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Csaba Csáki:** Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Andrea Kó:** Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Funding acknowledgments

This research was supported by the National Research, Development and Innovation Office (NKFIH) under grant number 2024-1.2.3-HU-RIZONT-2024-00030.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.jjimei.2026.100395](https://doi.org/10.1016/j.jjimei.2026.100395).

## References

- Aattouri, I., Mouncif, H., & Rida, M. (2023). Modeling of an artificial intelligence based enterprise callbot with natural language processing and machine learning algorithms. *IAES International Journal of Artificial Intelligence*, 12(2), 943. <https://doi.org/10.11591/ijai.v12.i2.pp943-955>
- Akdemir, A., Shibuya, T., & Güngör, T. (2022). A comprehensive analysis of subword contextual embeddings for languages with rich morphology. *Deep Learning Applications* (Volume 3., 31–72. [https://doi.org/10.1007/978-981-16-3357-7\\_2](https://doi.org/10.1007/978-981-16-3357-7_2)
- Akella, K., Venkatachalam, N., Gokul, K., Choi, K., & Tykal, R. (2017). Gain customer insights using NLP techniques. *SAE International Journal of Materials and Manufacturing*, 10(3), 333–337. <https://doi.org/10.4271/2017-01-0245>
- Ali Zaidi, S. S., Fraz, M. M., Shahzad, M., & Khan, S. (2022). A multiapproach generalized framework for automated solution suggestion of support tickets. *International Journal of Intelligent Systems*, 37(6), 3654–3681. <https://doi.org/10.1002/int.22701>
- Aman, J. J., Smith-Colin, J., & Zhang, W. (2021). Listen to E-scooter riders: mining rider satisfaction factors from app store reviews. *Transportation Research Part D: Transport and Environment*, 95, Article 102856. <https://doi.org/10.1016/j.trd.2021.102856>
- Asano, H., Bessho, K., Sadamitsu, K., Nishida, K., & Saito, K. (2017). Natural language processing technology for agent services. *NTT Technical Review*, 15(8), 1–5. <https://doi.org/10.53829/ntr201708fa3>
- Ashton, T., Evangelopoulos, N., & Prybutok, V. R. (2013). Exponentially weighted moving average control charts for monitoring customer service quality comments. *International Journal of Services and Standards*, 8(3), 230–246. <https://doi.org/10.1504/IJSS.2013.057237>
- Belk, R. W., Belanche, D., & Flavián, C. (2023). Key concepts in artificial intelligence and technologies 4.0 in services. *Service Business*, 17(1), 1–9. <https://doi.org/10.1007/s11628-023-00528-w>
- Borg, A., Boldt, M., Rosander, O., & Ahlstrand, J. (2021). E-mail classification with machine learning and word embeddings for improved customer support. *Neural Computing and Applications*, 33(6), 1881–1902. <https://doi.org/10.1007/s00521-020-05058-4>
- Bruni, R., Bianchi, G., & Papa, P. (2023). Hyperparameter black-box optimization to improve the automatic classification of support tickets. *Algorithms*, 16(1), 46. <https://doi.org/10.3390/a16010046>
- Cassandra, C., Hartono, S., & Karsen, M. (2019). Online helpdesk support system for handling complaints and service. *Chakrabarti, C., & Luger, G. F. (2015). Artificial conversations for customer service chatter bots: architecture, algorithms, and evaluation metrics. Expert Systems with Applications*, 42(20), 6878–6897. <https://doi.org/10.1016/j.eswa.2015.04.067>
- Chang, Y., Wang, X., Wang, J., Wu, Y., Zhu, K., Chen, H., Yang, L., Yi, X., Wang, C., & Wang, Y. (2023). A survey on evaluation of large language models. *arXiv preprint arXiv:2307.03109*.
- Chen, Y., & Tussyadiah, I. P. (2021). Service failure in peer-to-peer accommodation. *Annals of Tourism Research*, 88, Article 103156. <https://doi.org/10.1016/j.annals.2021.103156>
- Chiang, A.-H., & Trimi, S. (2020). Impacts of service robots on service quality. *Service Business*, 14(3), 439–459. <https://doi.org/10.1007/s11628-020-00423-8>
- Chiang, C.-H., & Lee, H.-y. (2023). Can large language models be an alternative to human evaluations? *arXiv preprint arXiv:2305.01937*.
- Chin, H., Marasini, D. P., & Lee, D. (2023). Digital transformation trends in service industries. *Service Business*, 17(1), 11–36. <https://doi.org/10.1007/s11628-022-00516-6>
- Chung, M., Ko, E., Joung, H., & Kim, S. J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*, 117, 587–595. <https://doi.org/10.1016/j.jbusres.2018.10.004>
- Conneau, A. (2019). Unsupervised cross-lingual representation learning at scale. *arXiv preprint arXiv:1911.02116*.
- Correani, A., De Massis, A., Frattini, F., Petruzzelli, A. M., & Natalicchio, A. (2020). Implementing a digital strategy: learning from the experience of three digital transformation projects. *California Management Review*, 62(4), 37–56. <https://doi.org/10.1177/0008125620934864>
- Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48, 24–42. <https://doi.org/10.1007/s11747-019-00696-0>
- Fan, S., & Ilk, N. (2020). A text analytics framework for automated communication pattern analysis. *Information & Management*, 57(4), Article 103219. <https://doi.org/10.1016/j.im.2019.103219>
- Fu, T., Zampieri, G., Hodgson, D., Angione, C., & Zeng, Y. (2021). Modeling customer experience in a contact center through process log mining. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 12(4), 1–21. <https://doi.org/10.1145/3468269>
- Grootendorst, M. (2020a). BERTopic: leveraging BERT and c-TF-IDF to create easily interpretable topics. *Zenodo*. <https://doi.org/10.5281/zenodo.4430182>
- Grootendorst, M. (2020b). *KeyBERT: minimal keyword extraction with BERT*.
- Grootendorst, M. (2022). BERTopic: neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.
- Gunarathne, P., Rui, H., & Seidmann, A. (2022). Racial bias in customer service: evidence from Twitter. *Information Systems Research*, 33(1), 43–54. <https://doi.org/10.1287/isre.2021.1058>
- Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Wu, J., & Mirjalili, S. (2023). A survey on large language models: applications, challenges, limitations, and practical usage. *TechRxiv*.
- Hardalov, M., Koychev, I., & Nakov, P. (2019). Machine reading comprehension for answer re-ranking in customer support chatbots. *Information*, 10(3), 82. <https://doi.org/10.3390/info10030082>
- Hese, G. V. (2022). Bridging the gap between social media marketing and customer care. *Journal of Brand Strategy*, 11(2), 128–136.
- Hryhorak, M., Trushkina, N., Popkowski, T., & Molchanova, K. (2020). Digital transformations of logistics customer service business models. *Intellectualization of Logistics and Supply Chain Management*, (1), 57–75. <https://doi.org/10.46783/smart-scm/2020-1-6>
- Huang, M.-H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, 21(2), 155–172.
- Ilk, N., Shang, G., & Goes, P. (2020). Improving customer routing in contact centers: an automated triage design based on text analytics. *Journal of Operations Management*, 66(5), 553–577. <https://doi.org/10.1002/joom.1084>
- Inavolu, S. M. (2024). Exploring AI-driven customer service: evolution, architectures, opportunities, challenges and future directions. *Preprint*.
- Ittoo, A., & Bouma, G. (2013). Term extraction from sparse, ungrammatical domain-specific documents. *Expert Systems with Applications*, 40(7), 2530–2540. <https://doi.org/10.1016/j.eswa.2012.10.067>
- Jabr, W., Mookerjee, R., Tan, Y., & Mookerjee, V. S. (2014). Leveraging philanthropic behavior for customer support: the case of user support forums. *MIS Quarterly*, 38(1), 187–208. <https://doi.org/10.25300/MISQ/2014/38.1.09>
- Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. *arXiv preprint arXiv:2004.09095*.
- Kar, A. K., Kumar, S., & Ilavarasan, P. V. (2021). Modelling the service experience encounters using user-generated content: A text mining approach. *Global Journal of Flexible Systems Management*, 22(4), 267–288. <https://doi.org/10.1007/s40171-021-00279-5>
- Keck, C., Egger, A., Kratsch, W., & Röglinger, M. (2023). Quantifying chatbots' ability to learn business processes. *Information Systems*, 113, Article 102176. <https://doi.org/10.1016/j.is.2023.102176>
- Kelahan, L. C., Fong, A., Ratwani, R. M., & Filice, R. W. (2016). Call case dashboard: tracking R1 exposure to high-acuity cases using natural language processing. *Journal of the American College of Radiology*, 13(8), 988–991. <https://doi.org/10.1016/j.jacr.2016.03.012>
- Kim, C., & Lee, J. (2024). Discovering patterns and trends in customer service technologies patents using large language model. *Heliyon*, 10(14). <https://doi.org/10.1016/j.heliyon.2024.e34701>
- Kim, M. (2023). Connecting artificial intelligence to value creation in services: mechanism and implications. *Service Business*, 17(4), 851–878. <https://doi.org/10.1007/s11628-023-00547-7>
- Klopfenstein, L. C., Delpriori, S., Malatini, S., & Bogliolo, A. (2017). The rise of bots: A survey of conversational interfaces, patterns, and paradigms.
- Kumar, S., Kar, A. K., & Ilavarasan, P. V. (2021). Applications of text mining in services management: A systematic literature review. *International Journal of Information Management Data Insights*, 1(1), Article 100008. <https://doi.org/10.1016/j.jjimei.2021.100008>
- Lal, M., & Neduncheliyan, S. (2023). An optimal deep feature-based AI chat conversation system for smart medical application. *Personal and Ubiquitous Computing*, 1–12. <https://doi.org/10.1007/s00779-023-01713-4>
- Lazo, M., & Ebarido, R. (2023). Artificial intelligence adoption in the banking industry: current state and future prospect. *Journal of Innovation Management*, 11(3), 54–74. <https://doi.org/10.24840/2183-0606>
- Liao, X., Chen, G., Ku, B., Narula, R., & Duncan, J. (2020). Text mining methods applied to insurance company customer calls: a case study. *North American Actuarial Journal*, 24(1), 153–163. <https://doi.org/10.1080/10920277.2019.1649155>
- Liau, B. Y., & Tan, P. (2014). Gaining customer knowledge in low cost airlines through text mining. *Industrial Management & Data Systems*, 114(9), 1344–1359. <https://doi.org/10.1108/IMDS-07-2014-0225>
- Lin, H., Zhu, J., Xiang, L., Zhai, F., Zhou, Y., Zhang, J., & Zong, C. (2023). Topic-oriented dialogue summarization. In *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. <https://doi.org/10.1109/TASLP.2023.3271118>
- Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., & Liu, Z. (2023). Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-Radiology*, Article 100017. <https://doi.org/10.1016/j.metrad.2023.100017>
- Lucini, F. R., Tonetto, L. M., Fogliatto, F. S., & Anzanello, M. J. (2020). Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. *Journal of Air Transport Management*, 83, Article 101760. <https://doi.org/10.1016/j.jairtraman.2019.101760>
- Luo, S., & He, S. Y. (2021). Understanding gender difference in perceptions toward transit services across space and time: A social media mining approach. *Transport policy*, 111, 63–73. <https://doi.org/10.1016/j.tranpol.2021.07.018>
- Mahr, D., Stead, S., & Odekerken-Schröder, G. (2019). Making sense of customer service experiences: a text mining review. *Journal of Services Marketing*. <https://doi.org/10.1108/JSM-10-2018-0295>
- Manchaiah, V., Amlani, A. M., Bricker, C. M., Whitfield, C. T., & Ratinaud, P. (2019). Benefits and shortcomings of direct-to-consumer hearing devices: analysis of large secondary data generated from Amazon customer reviews. *Journal of Speech*,

- Language, and Hearing Research, 62(5), 1506–1516. [https://doi.org/10.1044/2018\\_JSLHR-H-18-0370](https://doi.org/10.1044/2018_JSLHR-H-18-0370)
- Mao, L. L. (2021). Understanding retail quality of sporting goods stores: a text mining approach. *International Journal of Sports Marketing and Sponsorship*, 22(2), 330–352. <https://doi.org/10.1108/IJSM-03-2020-0029>
- Mnasri, M. (2019). Recent advances in conversational NLP: towards the standardization of Chatbot building. *arXiv preprint arXiv:1903.09025*.
- Müller, O., Junglas, I., Debortoli, S., & Vom Brocke, J. (2016). Using text analytics to derive customer service management benefits from unstructured data. *MIS Quarterly Executive*, 15(4).
- Nemeskey, D. M. (2022). Introducing huBERT. XVII. Magyar Számítógépes Nyelvészeti Konferencia, 3–14.
- Ngai, E. W., Lee, M. C., Luo, M., Chan, P. S., & Liang, T. (2021). An intelligent knowledge-based chatbot for customer service. *Electronic Commerce Research and Applications*, 50, Article 101098. <https://doi.org/10.1016/j.elerap.2021.101098>
- Nguyen, T. (2019). Potential effects of chatbot technology on customer support: A case study.
- Nirala, K. K., Singh, N. K., & Purani, V. S. (2022). A survey on providing customer and public administration based services using AI: chatbot. *Multimedia Tools and Applications*, 1–32. <https://doi.org/10.1007/s11042-021-11458-y>
- Nordheim, C. B., Følstad, A., & Bjørkli, C. A. (2019). An initial model of trust in chatbots for customer service—Findings from a questionnaire study. *Interacting with Computers*, 31(3), 317–335. <https://doi.org/10.1093/iwc/iwz022>
- Ohata, E. F., Mattos, C. L. C., Gomes, S. L., de Souza Rebouças, E., & Rego, P. A. L. (2022). A text classification methodology to assist a large technical support system. *IEEE Access: Practical Innovations, Open Solutions*. <https://doi.org/10.1109/ACCESS.2022.3213033>
- Palese, B., & Usai, A. (2018). The relative importance of service quality dimensions in E-commerce experiences. *International Journal of Information Management*, 40, 132–140. <https://doi.org/10.1016/j.ijinfomgt.2018.02.001>
- Papadia, G., Pacella, M., & Giliberti, V. (2022). Topic modeling for automatic analysis of natural language: a case study in an Italian customer support center. *Algorithms*, 15(6), 204. <https://doi.org/10.3390/a15060204>
- Pires, T. (2019). How multilingual is multilingual BERT. *arXiv preprint arXiv:1906.01502*.
- Przegalinska, A., Ciecchanowski, L., Stroz, A., Gloor, P., & Mazurek, G. (2019). In bot we trust: A new methodology of chatbot performance measures. *Business Horizons*, 62(6), 785–797. <https://doi.org/10.1016/j.bushor.2019.08.005>
- Quattrocchi, G., Tamburri, D. A., & Van Den Heuvel, W.-J. (2023). Making service continuity smarter with artificial intelligence: an approach and its evaluation. *Software: Practice and Experience*, 53(10), 1887–1901. <https://doi.org/10.1002/spe.3230>
- Raeesi Vanani, I. (2019). Text analytics of customers on twitter: brand sentiments in customer support. *Journal of Information Technology Management*, 11(2), 43–58. <https://doi.org/10.22059/JITM.2019.291087.2410>
- Ram Kumar, T. N., Shidaganti, G., Anand, P., Singh, S., & Salil, S. (2023). Analyzing and automating customer service queries on Twitter using Robotic Process Automation. *Journal of Computer Science*, 19(4), 514–525. <https://doi.org/10.3844/jcssp.2023.514.525>
- Ramsey, G., & Bapna, S. (2016). Text mining to identify customers likely to respond to cross-selling campaigns: reading notes from your customers. *International Journal of Business Analytics (IJBAN)*, 3(2), 33–49. <https://doi.org/10.4018/IJBAN.2016040102>
- Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Rebello, H. D., de Oliveira, L. A., Almeida, G. M., Sotomayor, C. A., Magalhães, V. S., & Rochocz, G. L. (2022). Automatic update strategy for real-time discovery of hidden customer intents in chatbot systems. *Knowledge-Based Systems*, 243, Article 108529. <https://doi.org/10.1016/j.knosys.2022.108529>
- Rose, S., Sreejith, R., & Senthil, S. (2019). Social media data analytics to improve the customer services: the case of fast-food companies. *International Journal of Recent Technology and Engineering*, 8(2), 6359–6366. <https://doi.org/10.35940/ijrte.B2205.078219>
- Saberi, M., Hussain, O. K., & Chang, E. (2017). Past, present and future of contact centers: a literature review. *Business Process Management Journal*. <https://doi.org/10.1108/bpmj-02-2015-0018>
- Sammet, J., & Krestel, R. (2023). Domain-specific keyword extraction using BERT.
- Senese, M. A., Rizzo, G., Dragoni, M., & Morisio, M. (2020, 2020). MTSI-BERT: a session-aware knowledge-based conversational agent.
- Shah, S., Ghomeshi, H., Vakaj, E., Cooper, E., & Fouad, S. (2023). A review of natural language processing in contact centre automation. *Pattern Analysis and Applications*, 26(3), 823–846. <https://doi.org/10.1007/s10044-023-01182-8>
- Sharda, R., Delen, D., & Turban, E. (2021). *Analytics, data science, & artificial intelligence: Systems for decision support*. Pearson.
- Tamkin, A., Brundage, M., Clark, J., & Ganguli, D. (2021). Understanding the capabilities, limitations, and societal impact of large language models. *arXiv preprint arXiv:2102.02503*.
- Tanaka, K., Yagi, T., & Tetsuya, I. (2019). Advanced initiatives for Contact Center AI. *NTT Technical Review*, 17(9), 1–4. <https://doi.org/10.53829/ntr201909fa1>
- Teng, S., & Khong, K. W. (2021). Examining actual consumer usage of E-wallet: A case study of big data analytics. *Computers in Human Behavior*, 121, Article 106778. <https://doi.org/10.1016/j.chb.2021.106778>
- Wang, S.-M., Huang, Y.-K., Shih, C.-W., & Li, P.-C. (2023). Evaluation of service quality on natural language processing service: a case on train station AI service. *Review of Integrative Business and Economics Research*, 12(4), 71–87.
- Winata, G. I., Aji, A. F., Yong, Z.-X., & Solorio, T. (2022). The decades progress on code-switching research in NLP: A systematic survey on trends and challenges. *arXiv preprint arXiv:2212.09660*.
- Xu, Y., Shieh, C.-H., van Esch, P., & Ling, I.-L. (2020). AI customer service: task complexity, problem-solving ability, and usage intention. *Australasian marketing journal*, 28(4), 189–199. <https://doi.org/10.1016/j.ausmj.2020.03.005>
- Yang, Z. G., Dodé, R., Ferenczi, G., Hatvani, P., Héja, E., Madarász, G., Ligeti-Nagy, N., Sárosy, B., Szaniszló, Z., & Váradi, T. (2024, 2024). The first instruct-following large language models for Hungarian.
- Zhan, Y., Han, R., Tse, M., Ali, M. H., & Hu, J. (2021). A social media analytic framework for improving operations and service management: A study of the retail pharmacy industry. *Technological Forecasting and Social Change*, 163, Article 120504. <https://doi.org/10.1016/j.techfore.2020.120504>
- Zhou, X., Wu, O., & Jiang, C. (2022). Increasing naturalness of human-machine dialogue: the users' choices inference of options in machine-raised questions. *Knowledge-Based Systems*, 243, Article 108485. <https://doi.org/10.1016/j.knosys.2022.108485>