Bridging Natural Language Processing AI technique and Corporate Communications

Today's communication channels and media platforms generate a huge amount of data, which - through advanced AI- (Machine Learning) based techniques - can be leveraged to significantly enhance business networking, improve the efficiency of public relations, management, and extend the possible application areas of communication components. As a sub-discipline of AI, Natural Language Processing (NLP) is frequently utilized in the field of corporate communications (CC) to boost target- group satisfaction through information retrieval and automated dialogue services. This paper gives an overview of the use of NLP in different disciplines of CC, discusses general corporational/organizational practices, and identifies promising research topics for the future while pointing out the ethical aspects of user-data handling and customer engagement. The findings of this synthesizing study are based on primer qualitative research building on the methodology of deep interviews and focus group research involving experts practicing in the fields of CC and NLP. Based on the feedbacks of the participants, a refined CC model was developed, as well as a model mapping conventional NLP techniques onto CC disciplines and tasks they are utilized for.

Keywords: *business management, natural language processing, public relations, corporate communication, deep learning, information society, artificial intelligence, AI ethics*

Author Information

Dániel Gergő Pintér
https://www.linkedin.com/in/dániel-gergő-pintér-phd-01532487/
Péter Lajos Ihász
https://www.linkedin.com/in/peter-ihasz/

How to cite this article

I
N
F
O
R
M
Á
C
I
Ó
S

T
Á
R
S
A
D
A
L
O
M

Bridging Natural Language Processing AI techniques and Corporate Communications: towards an integrative model[1]

Dániel Gergő Pintér – Péter Lajos Ihász

Abstract

Today's communication channels and media platforms generate a huge amount of data, which - through advanced AI- (Machine Learning) based techniques - can be leveraged to significantly enhance business networking, improve the efficiency of public relations, management, and extend the possible application areas of communication components. As a sub-discipline of AI, Natural Language Processing (NLP) is frequently utilized in the field of corporate communications (CC) to boost target-group satisfaction through information retrieval and automated dialogue services. This paper gives an overview of the use of NLP in different disciplines of CC, discusses general corporational/organizational practices, and identifies promising research topics for the future while pointing out the ethical aspects of user-data handling and customer engagement. The findings of this synthesizing study are based on primer qualitative research building on the methodology of deep interviews and focus group research involving experts practicing in the fields of CC and NLP. Based on the feedbacks of the participants, a refined CC model was developed, as well as a model mapping conventional NLP techniques onto CC disciplines and tasks they are utilized for.

Keywords: business management, natural language processing, public relations, corporate communication, deep learning, information society, artificial intelligence, AI ethics

1. Introduction

1.1. The era of Information Society

With evolving into an information society, access to the social resources and information has been completely rearranged, changing significantly the technological, economic and cultural aspects of everyday life. (Beniger 1986) While in the industrial era the devices and natural resources were definitive for the economy, nowadays knowledge is considered

as the most valuable product. (Castells 1996, 1998) Labour aimed to process information became more significant than direct physical work: the creation, distribution, and manipulation of information is currently the most significant economic and cultural activity, profoundly changing all aspects of social organization. (Pintér 2016; Castells 1997) Thus, digital experience, network-based interaction and unlimited communication have become a basic experience and daily need. (Pintér 2016; Castells 1996) Accordingly, business management started to rely on automated data mining, data analysis, and automated response generation in order to harvest this novel and profound resource. Companies collect and store a large amount of customer data in order to enable better business decisions and gain an advantage in the global market through performing communication that builds on a better understanding of customer needs. (Castells 1997)

1.2. Problem identification, goal and structure

As one pillar of corporate business management, corporate communications - activities aimed to establish and maintain favorable internal and external reputation of the corporation (Riel 2002) - processes and responds to the input of stakeholders and various target audience groups. Many e-commerce websites, for example, allow customers to express their opinions about the products and services the company offers. The reviews are considered not only by fellow customers but with the right information retrieval techniques, these easily obtainable feedbacks serve as a valuable source of information for the companies as well. As another source of information, social media can also be harvested through Artificial Intelligence (AI) - based information retrieval. (Russel and Norvig 2003) Sentiment analysis (SA), for example, is a way to extract semantic information from feedbacks, where opinions, sentiments, emotions, attitudes toward entities and their attributes are computationally identified. Topic extraction, dialogue act classification or summarization are further examples from a wide palette of information retrieval practices. (Carenini et al. 2011; Goldberg 2016)

Extracting customer intelligence from such user-generated content, however, is a challenging task, as it involves dealing with data requiring natural language processing (NLP) techniques. Nevertheless, various machine learning-based NLP methods exist, making possible the effective extraction of customer intelligence, and thus, indirectly enhancing business networking, improving the efficiency of public relations management, and extending the possible application areas of communication components. (Jozefowicz et al. 2016; Goodfellow et al. 2016)

The goal of this study is to underline the seldom researched (and at a first glance non-trivial) connection between the field of communications and computational intelligence. This paper gives an overview on the use of AI-based NLP information retrieval and answer generating practices (NLP techniques) in different disciplines of corporate communications, discusses general practices in a corporational/organizational environment, identifies promising research topics for the future and elaborates on the ethical aspects of AI-based user-data handling and automated communication. As a result of the presented synthesizing study, two models have been developed 1) a model refining the definitions on the disciplines and tasks of CC, and 2) a model mapping NLP techniques onto CC tasks they are applied for. We believe our findings can be utilized by the experts of both

fields not only on a theoretical, but on a practical level as well, and that our research can motivate further discussion.

The paper is organized as follows. Section 2 introduces the methodology, elaborating on the details of the experiments for validation. Section 3 describes and specifies the disciplines, tasks and goals of corporate communications and synthesizes between contradicting definitions, introducing our refined model for CC. Section 4 elaborates on the role of AI within today's information society, and specifies the mainstream NLP techniques applicable in CC. In section 5, a model developed to bridge NLP techniques and CC tasks is introduced, along the results of the qualitative research it builds upon. Section 6 exemplifies the application of NLP techniques along the three CC disciplines and discusses the ethical, legal and moral considerations elaborated before. Finally, section 7 concludes the main findings of the study, highlights its limitations, and outlines possible future work.

2. Methodology and experimental setup

The conceptual basis of the study has been laid down through:

1. reviewing the basic literature on CC;
2. synthesizing the fundamental academic definitions of the CC disciplines (Management Communication, Organizational Communication and Marketing Communication);
3. analyzing and redefining one of the most accepted integrative model of CC from a task-focused perspective, differentiating between strategic and operational, as well as internal and external dimensions;
4. summarizing AI-based NLP techniques relevant to CC disciplines based on literature review.

As a primer qualitative research to

- validate and refine our literature review-based findings and conclusions,
- support the reliability of the refined CC model,
- construct a novel NLP-CC model – a model, mapping NLP techniques onto the tasks of the refined CC model – helping to organize the tasks of several general practices used in corporational/organizational environment,
- and to complement our findings with AI ethical and moral aspects,

we conducted:

1. <u>structured deep interviews</u> with communications and AI experts separately. The interviews were 45-60 minutes long one-on-one sessions with 6-6 experts from each field. The questions concentrated on a) what tasks and professional challenges emerge during the daily practice in the fields of CC/NLP, b) what are the possible solutions and best practices in terms of used techniques and services, and c) in which task do they feel that there is a need for improvement (see the questions in detail in

Section 5). The interviews were standardized to have 8 identical questions towards each interviewee so comparisons can be made with confidence between sample subgroups or between different interview periods. The interviewees were selected based on their professional background and seniority level: having a minimum 3 years experience in communications / AI (preferably working with NLP), working mainly on tasks of the operational dimension (e.g. coding, content creation, social media management etc.)

2. <u>focus group research</u> between a heterogeneous group of communication and NLP experts (7-7 people from each field, different from the participants of the deep interviews). The participants were asked about their perceptions, opinions, and attitudes towards a preliminary classification mapping NLP techniques on CC tasks based on the results of the deep interviews. The session was 90 minutes long. Questions were asked in an interactive group setting where participants were free to talk with other group members about the possible discrepancies and insufficiencies of our approach. The authors were present as mere mediators, motivating debate and supporting the group to reach consensus. The participants were selected along the same criteria used in the case of the deep interviews, with the addition that we paid special attention to have at least one expert from each disciplines who works on tasks from the strategic dimension (e.g. reputation management, campaign planning, technical leading, project management etc.), has a minimum of 5 years working experience, and has at least moderate insight to the other field.

Thus, as a result of the deep-interviews, we refined the CC model and built an initial version of the NLP-CC model. The models were further refined in the light of the discussion in the focus group research. The finalized model is illustrated and introduced in Section 5.

3. Corporate Communications as a highly interdisciplinary profession

3.1 Literature review: towards an exact definition

In the scientific literature corporate communication is defined as a set of activities and professional techniques involved in planning, managing and orchestrating all internal and external communications, aimed at creating favourable reputation among stakeholders, target groups and business partners on which the company depends (Fombrun and Riel 2007). According to the definition of Goldhaber (1993, 15), corporate communication is the process of "creating and exchanging messages within a network of interdependent relationships to cope with environmental uncertainty". In his research, Riel (2003, 53) summarizes that CC "is the orchestration of all the instruments in the field of organizational identity (communications, symbols and behaviours of organizational members) in such an attractive and realistic manner as to create or maintain a positive reputation for groups with which the organization has an interdependent relationship".

Frandsen and Johansen (2013) have synthesized the common features of some of the prevailing definitions as seen below:

1. CC is a strategic management function that takes a strategic approach to communication activities and is tied the overall strategy of the company.
2. It integrates external and internal communication activities spread among a series of organizational practices to build, maintain, change and/or repair one or more positive images and/or reputations.
3. All these activities take place inside the relationship with the external and internal stakeholders of the company. (Frandsen and Johansen 2013)

To sum up this interdisciplinary approach, the biggest goal of this profession is twofold: on one hand, to help organizations explain their mission, unique selling proposition and social responsibility; and on the other, to combine the vision of the company, business philosophy and its values into a coherent and credible message to stakeholders and larger audiences, such as employees, media, channel partners and the general public. (Goodman and Hirsch 2010) As Christensen and Cornelissen (2011) pointed out, CC encompasses and coordinates all company's communication activities as an integrated whole with the aim of building and maintaining a valuable image across different stakeholder groups, markets and audiences (Christensen and Cornelissen 2011; Cornelissen 2008).

As it can be seen CC has several, often contradicting or only partly-overlapping definitions, making the selection process of a widely-accepted definition with well formulated specifications to build upon highly difficult and subjective. Thus, in order to achieve the primary goal of this paper - defining CC tasks mappable to NLP techniques - we turned to a discipline-segmenting, task-oriented definition instead of deducing operational communication tasks from a normative description selected from above.

3.2 Corporate communications defined through its sub-disciplines

According to Mazzei's research, CC paradigm prefers the sender's, in this case, the company's perspective, assuming for itself an orchestration role, and justifying centralized control of the entire communication function (Christensen and Cornelissen 2011). This is in accordance with Riel's (1995) well-known, task-oriented definition, which states that CC includes three categories of communication defined by the senders of the communication:

1. <u>Management communication</u> implemented by CEOs, executive managers and teamleaders for: developing offline and online CC strategies; planning and implementing internal and external communications directives; systematically organizing the flow of information; supervising, coordinating, disseminating, revising and monitoring of all the formal channels of communication. Its biggest goals are to develop a shared vision within and besides the organization, gain reputation and maintain trust, enable and manage change processes, support the financial stability and development of the corporation and help employees to grow professionally.

2. <u>Organizational communication</u> includes heterogeneous communication activities: public relations, public affairs, investor relations, government relations, employer branding, corporate social responsibility, investor relations, communication with the labour market, environmental communications, corporate advertising and internal communications.

3. <u>Marketing communication</u> gets the bulk of the budgets in most organizations, encompasses commercial and business communication activities developed to support the sale of goods and services. It typically includes: product advertising, personal selling, promotion, direct marketing, branding and sponsorship activities. (Riel 1995)

This discipline-differentiating definition could be seen as an integrative model linking stakeholders and various internal - and external - target audience groups with the communicational tasks of the organization. This approach gives us a detailed view of CC, which can be used as the basis to link it with NLP techniques.

3.3 A refined CC model

Nevertheless, the tasks and goals mentioned in the description of the disciplines by Riel (1995), partially negligate the synthesization of the common features of CC definitions by Frandsen and Johansen (see Section 3.2). To resolve this discrepancy we redefine management communication as management OF communication (ManC), since it stands for monitoring and analysis-based planning process rather than for the conduct of communication processes in practice. In further accordance with Frandsen and Johansen (2013), we altered Riel's original model from a function-based perspective differentiating between strategic and operational, as well as internal and external (within operational) dimensions. Some tasks were re-specified within the disciplines of Organizational Communication (OC) and Marketing Communication (MC) accounting for the possible intersections. Furthermore, we complemented the model with crisis communication as well, as a crucial but neglected task of CC. (An and Cheng 2010; Pintér 2018) According to the communication scholar Timothy Coombs, as a sub-discipline of public relations profession, crisis communication is designed to protect and defend an organization facing an often unpredictable public challenge to its reputation and performance. (Bundy et al. 2016; Coombs 2007)

The re-definition of the model and re-specification of the tasks were verified and complemented by the valuable remarks of CC experts during the deep interviews and the focus group research (see Section 2 and Table 2). Figure 1 below illustrates the revised model of Riel's.

4. AI in corporate communications

In this section NLP within AI is discussed, and conventional NLP techniques are listed, which can be then mapped onto the tasks of the refined CC model (see Section 3.3).
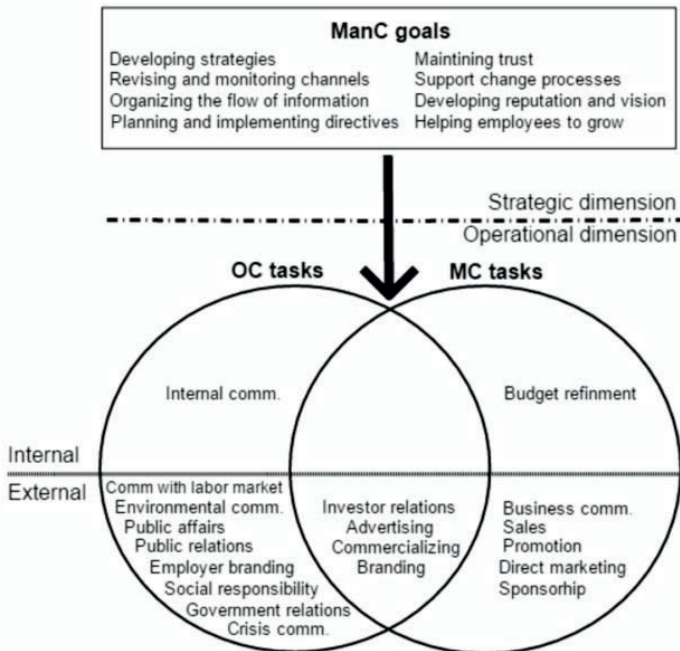
Figure 1: Revised model of CC tasks and disciplines

4.1 Defining NLP within AI

Artificial intelligence (AI) is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans. The use of AI conventionally boils down to machine learning (ML), the implementation of algorithm based programs, that perceive their environment and take actions to achieve their goals, while learning from experience to maximize their chance of success. Typical AI / ML algorithms mimic "cognitive" functions, such as classification, pattern recognition, prediction etc. (Poole 1998)

Natural language processing (NLP) is a subfield of AI concerned with the interactions between computer and human, in particular how to develop computational programs (typically MLs) to process, analyze and/or respond to natural language inputs. The three main subfields of natural language processing is automatic speech recognition and synthesization (ASR), natural language understanding (NLU), and natural language generation (NLG). As ASR is related to the pre-processing of NLU and post-processing of NLG, it is often considered as part of them and not as an individual subfield. (Goldberg 2017)

4.2 NLP in the light of CC tasks

Several of the aforementioned CC tasks rely on the summarization and analysis (in alignment with business goals) of data gathered from the target group in order to define and revise strategic steps. Other tasks require the conduction of customer targeting communi-

Table 1. Conventional NLP techniques

Process	Techniques	Description
Analysis	Intent analysis	Inferring the intents of the locutor in the form of dialog acts usually from text source. (Ihász et al. 2019)
	Sentiment analysis/ emotion recognition	Inferring basic emotions (sadness, joy etc.) or sentiments (grouping emotions along their valence into negative, positive and neutral polarities) from text, audio and/or visual sources. Processing multiple source channels in parallel is conventional for this technique. (Ihász et al. 2019; Liu 2015)
	Topic extraction	Inferring one or multiple topics of given text units. (Bun et al. 2002)
	Entity recognition	Extracting target entities (e.g. objects) or named entities (e.g. actual or legal persons) from a given textual source. Extraction from audio source is also possible but unconventional. (Nadeau and Satoshi 2007)
	Content summarization	Summarizing text content into shorter units. (Mani 1999)
Generation	Response generation	Generating responses in the light of user inputs based on hand-made templates or probability-based, deep learning-driven language models. Response generation is achieved by chatbots and preluded by some sort of meaning inference (analysis techniques), and rule-based reasoning. It often involves communication with external information sources (database, ontology or question-answering system). The generated answers are always in text form which can be further synthesized into audio form. (Sordoni et al. 2015)
	Question-Answering	Systems that retrieve information (based on online, web-based or offline, database search) in accordance with the user input and output them in text or audio form. (Xiong et al. 2016)
	Machine translation	Systems translating a source language into target language based on deep learning-driven language models. Input and output is possible both in text and audio.(Kohen 2009)

cation. This is where NLP techniques can boost CC tasks through quick, automatized solutions, able to process complex and robust data. Accordingly, NLP can be subdivided into content analyzing (former NLU and ASR) and content generating (former NLG and ASR) sub-processes.

Table 1 specifies the typical content analysis and content generation techniques along their brief descriptions. The contents were selected as a result of discussion with both AI and CC experts from participating in the deep interviews and focus group research (see Section 2 and Table 2). It is important to note as well, that from a linguistic perspective, the NLP techniques specified in the table are processing semantic - and pragmatic-level (in other words meaning-level) information, preluded by several pre-processing techniques, handling phonetics and phonology - (e.g transcription), morphology - (e.g. part-of-speech-tagging), and syntactic - (e.g. syntactic parsing) level information. In this paper we limit ourselves to discuss only NLP techniques processing meaning-level data, as techniques for lower dimensional information processing could not be effectively associated with CC tasks.

5. Results: the proposed NLP-CC model

In this section the proposed model is presented, mapping NLP techniques onto CC tasks. The mapping is based on the results of the deep interviews and refined through a focus group research.

Figure 2 illustrates an extended CC model, specifying NLP techniques related to CC tasks and goals (and excluding all unrelated tasks and goals). No NLP techniques to the goals of ManC were mapped, since they represent overarching-goals influencing the whole of the company, rather than specific tasks included in OC and MC. These goals are satisfied through the realization of the tasks of the operational-level disciplines (as indicated by the blue arrows). Thus, ManC is, in reality, uses all NLP methods, only in an indirect way, through the result of the implemented operational-level CC tasks. Such tasks (of OC and MC) then are governed accordingly by the strategic-level communication goals, constituting a highly interactive relationship between ManC and OC; and ManC and MC separately. It is important to note, that although ManC achieves a bi-directional information exchange with MC and OC, the two operational-level disciplines - even though sharing common tasks - are not interacting with each other.

The initials below the specific CC tasks indicate the NLP techniques they involve, while the colour (of the initials and the arrows), indicates the direction of the information-flow. Although NLP techniques for content analyzis and generation were differentiated in Table 1, this specification not necessarily overlap the differentiation of information gathering and outward communication. Machine translation for example, which is a content generating technique, is used for content generation and information gathering as well in several OC and MC tasks (indicated by the duplicated use of the initial M in both read and blue colours). The CC tasks which share resonating goals were grouped (indicated by brackets) since they utilize the same set of NLP techniques.
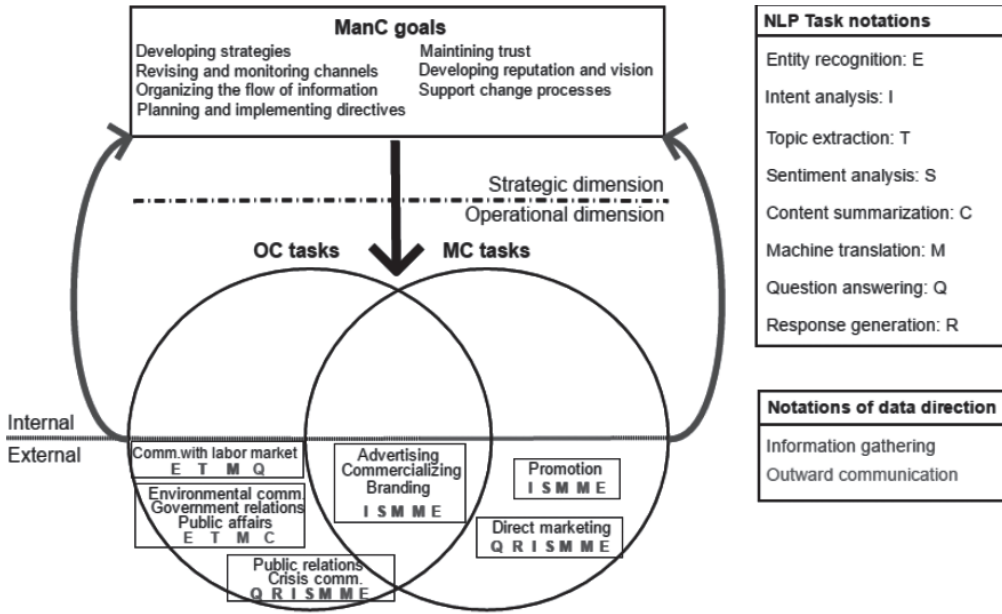
Figure 2: The NLP-CC Model, mapping NLP techniques on CC tasks

Table 2. below specifies the questions and answers of the deep interviews and focus group research the above model was built upon. Ratio of consensus is also indicated, measuring the extent of alignment among a) the 6-6 CC and NLP experts (separated by expertise), participating in the deep interviews, and b) between the 7-7 CC and NLP experts attended the focus group research. The authors concentrated on analyzing the intersection related aspects of CC and NLP as comprehensive as possible, thus even when opinions were highly divided (50% of consensus), they tended to acknowledge the existence of a given problem/aspect. The debate generating questions during the focus group research was mainly identical to the ones asked in the deep interviews (and thus not specified in the table), with a greater emphasis on the questions where the ratio of consensus was initially low, e.g. on the ethical aspects regarding NLP in internal communication. Considering the length limitations of the paper not all concerns were specified in the disagreement section, only one example was provided per question for each group of experts reflecting the most concerning issues. (Over-specifying the separate concerns would shift the balance of the paper which would no longer concentrate on giving a holistic overview on the overlap of NLP and CC.) The results were used to validate and revise the refined CC model, and to support and expanding it to the NLP-CC model.

Table 2. Experimental results: summary of questions and answers

Question	Answer with consent of majority	Ratio of consensus	Main points of disagreement
1. Are the tasks listed (in the refined CC model / Table 1 of NLP techniques) representing your field of expertise to a satisfactory-level?	Yes.	CC experts of deep interviews: 66% (4/6)	Some experts argued that the environmental comm., public affairs and business comm. tasks are fused in some cases. Thus their separation into different tasks and to different disciplines can be inadequate. On the other hand, one expert claimed that the advertising and commercializing tasks are purely MC tasks, and should not be positioned in the intersection of OC and MC.
		NLP experts of deep interviews: 100% (6/6)	All NLP experts agreed on the techniques of Table 1 as conventional NLP techniques.
		Focus group participants:79% (In agreement: 11/14; 6/7 NLP, 5/7 CC experts)	Some CC and NLP experts found the differentiation between the goals of strategic dimension and the tasks of operational dimension ambiguous. CC experts pointed out that the strategic goals and operational tasks are sharing an interactive connection, where the borderline between goal formulation and task implementation is not distinguishable. One NLP expert found it difficult to visualize how the implementation of specific tasks serve several strategic goals in parallel. (E.g. employer branding serves the goals of maintaining trust, supporting change processes and developing reputation as well).

2. What are the CC tasks that are plausible to be solved by any NLP solution? (What CC tasks to keep in the model?)	See the tasks specified in the OC and MC sets of Figure 2.	CC experts of deep interviews: 83% (5/6)	The definition and application of sentiment analysis was debated. (E.g. how it is realized in the commercializing task and is it required at all.)
		NLP experts of deep interviews: 66% (4/6)	Differentiation between activities related to the tasks of MC were debated. (E.g. promotion vs. direct marketing)
		Focus group participants: 57% (In agreement: 8/14; 4/7 NLP, 4/7 CC experts)	There is a lack of mutual understanding between the experts of the two fields. CC experts tend to overestimate the possible use-cases of NLP, while NLP experts having difficulties with recognizing the outcomes of several CC tasks. (E.g. NLP techniques are not applicable on the investor relations task. What is the exact purpose and outcome of environmental comm.)
3. What are the specific NLP techniques that are plausible to apply for the given CC tasks?	See the mapped techniques notated by initials in the OC and MC sets of Figure 2. and a detailed discussion in Section 6.1.	CC experts of deep interviews: 50% (3/6)	The role of machine translation and topic extraction was highly controversial in the case of OC.
		NLP experts of deep interviews: 66% (4/6)	The role of entity recognition was ambiguous in the case of MC. (E.g. how it is realized in the task of promotion.)
		Focus group participants: 50% (In agreement: 7/14; 4/7 NLP, 3/7 CC experts)	There was discrepancy in the NLP-technique- specific grouping of the CC tasks. (E.g. although environmental communication, government relations, and public affairs share partially similar goals, according to some of the participants they might not utilize the same NLP techniques: is machine translation and content summarization required for the inherently outward-directed public affairs task?).

4. What are the NLP techniques that serve the purpose of outward communication / information gathering? (Direction of data-flow.)	See the techniques differentiated by colour in the OC and MC sets of Figure 2.	CC experts of deep interviews: 83% (5/6)	NLP techniques of outward communication was debated to be applicable in the case of some OC tasks. (E.g. since misinformation in crisis communication can threaten personal safety and economical stability, automated response generation may deemed to be dangerous to apply.)
		NLP experts of deep interviews: 66% (4/6)	
		Focus group participants: 86% (In agreement: 12/14; 7/7 NLP, 5/7 CC experts)	
5. Are there any NLP techniques plausible for ManC purposes?	NLP techniques are utilized for the goals of ManC only through the results of OC and MC in an indirect way (indicated by the blue arrows in Figure 2.)	CC experts of deep interviews: 100% (6/6)	It was agreed by all experts that in order for the ManC to successfully fulfill its strategic planning role (towards the operational level) it continually requires input from the operational-level tasks of OC and MC disciplines.
		NLP experts of deep interviews: 50% (3/6)	The realization of strategic and operational dimensions in terms of CC tasks was ambiguous. (E.g. how exactly the use of question-answering for the task of direct marketing serves the realization of ManC strategic goals.)
		Focus group participants: 79% (In agreement: 11/14; 4/7 NLP, 7/7 CC experts)	There was a lack of understanding on the interactive structure of CC disciplines from the NLP experts-side. (Similar to the concerns of the CC experts detailed in the focus group disagreement section of Question 1)

6. Are there any NLP techniques applicable for internal (as opposed to external) OC and MC tasks?	No.	CC experts of deep interviews: 83% (5/6)	The experts from both sides agree on that NLP is not utilized widely (to their knowledge) at the moment for internal communication purposes. However, there were suggestions from both sides on technologies applicable (but currently unexploited) for gathering information from and about the employees of the company. The ethical way of using NLP would be automated and anonymized survey and content analysis (through all the analysis-related NLP techniques mentioned in Table 1) to measure the level of employee engagement; to identify topics of interest; to understand the overall perception about the company. The unethical application would be monitoring and distinguishing certain employees based on the content shared on the official platforms of the company. E.g. fishing for employees planning to leave the company / damages the reputation of the company etc. Both NLP-based information gathering methods could provide inputs for internal communications. Some CC experts also suggested the development of chatbot-like personal assistants to navigate the employees through complex internal platforms / to find specific intra-company information in a centralized way.
		NLP experts of deep interviews: 66% (4/6)	
		Focus group participants: 78% (In agreement: 11/14; 6/7 NLP, 5/7 CC experts)	

7. What are the potential use cases of NLP techniques for CC, that would worth to be implemented in the future?	CC experts of deep interviews: Question - answering could be used for swift information gathering from intra-company sources. E.g. measuring employee-satisfaction via the automated intention/sentiment analyzis of free-input feedback forms.	100% (6/6)	None
	NLP experts of deep interviews: monitoring online behaviour through topic extraction in order to measure employee engagement.	50% (6/3)	E.g. monitoring of private search-history (even on company machines) would raise several ethical and legal issues.
	Focus group participants: internal communication is unexploited (from an NLP utilizing perspective) and should be considered to at least the same extent as external communication, because employees are an equally important target group of CC. The NLP techniques which are implemented for external purposes	100% (14/14)	Although there were some debate related to ethical considerations, nevertheless, all participants agreed on that the power of NLP could be exploited on internal communications as well, if the resources and number of employees render it rational.

8. Are there any ethical problems that can emerge in the application of NLP to CC task in your opinion?	Yes. (See a detailed discussion in Section 6.2.)	CC experts of deep interviews: 83% (5/6)	The artificial agents may be confused to human agents during the official interaction between the customers and the company (e.g. help desk chatbots). In the case of communication with artificial agents legal responsibility is not clarified.
		NLP experts of deep interviews: 50% (3/6)	The process of information gathering and analyzis often induces several ethical problems related to private data management. (See Section 6.2.)
		Focus group participants: 100% (14/14)	Although all experts agreed on that there are several potential ethical problems, opinions from the representatives of the two fields differed on the nature of the possible sources what can endanger the customers. CC experts highlighted the harmfulness of misleading customers, while initially failed to recognise what NLP experts stressed, that information gathering and analysis is often conducted without the full consent and awareness of the customers. NLP experts on the other hand concentrates on the pragmatic side of implementation, and consider less the possible harms the misuse of the application of NLP can cause.

6. Discussion

In this section the results of the primer qualitative research (represented in Table 2 and Figure 2) is discussed with general examples of organizational/corporational practices along two axis: results related to the bridging of NLP and CC; and results related to ethical considerations.

6.1 General practices of applying NLP for CC in organizational/corporational environment

The deep interviews and focus group research revealed several practices implemented in the industry in the relation of CC.

<u>Management of Communication:</u>

Although no NLP techniques can be directly associated with the strategic goals of this discipline, the primer qualitative research revealed that several directives support and motivate the application of computational linguistics in order to improve the overall efficiency of the company. Thus, management must account for AI-related planning and implementation:

- integrating the use of AI into the company's workflow and processes;
- recruiting AI specialists;
- organizing trainings focusing on the best practices of AI-leveraging communications;
- supporting digitalization and the adoption of AI business mindset.

According to the participants of the research, the following NLP-applications of OC and MC give input for strategic planning. The applications listed below are often overarching, utilized in several CC tasks.

<u>Organizational Communication:</u>

- Finding a common thread across heterogenous target groups: mainly utilizing the NLP techniques of entity recognition, sentiment analyzis, and machine translation. It is applied along all of the OC tasks.
- Predict media trends, discrepancies, conflicts and identify deception: mainly utilizing the NLP techniques of entity recognition, text summarization, sentiment analyzis, and machine translation. It is applied mostly in the OC tasks of public affairs, public relations, government relations, social responsibility and crisis communication.
- Issue management, tracking rumors and crisis indicators, determining and avoiding PR crises: mainly utilizing the NLP techniques of entity recognition, text summarization, sentiment analyzis and intent analyzis. It is applied mostly in the OC tasks of public affairs, public relations, government relations, and crisis communication.
- Public Relations and impact measurement, metrics-improvement: mainly utilizing the NLP techniques of entity recognition, sentiment analyzis, and topic extraction. It is applied mostly in the OC tasks of public relations, employer branding, crisis communication and communication with the labour market.
- Creating media lists, corporate-related content, writing data-driven stories, and automatically delivering news: mainly utilizing the NLP techniques of text summarization, entity recognition, response generation and question answering. It is applied mostly in the OC tasks of internal communication, public relations, employer branding and crisis communication.

<u>Marketing Communication:</u>

- Creating more accurate buyer personas, identifying target customers: mainly utilizing the NLP techniques of entity recognition, sentiment analyzis, and topic extraction. It is applied mostly in the MC tasks of business communication, sales, sponsorship and direct marketing.

- Building online chatbots and voice assistants, automating customer service: mainly utilizing the NLP techniques of entity recognition, intent analyzis, response generation and question answering. It is applied mostly in the MC tasks of promotion, sales and direct marketing.
- Marketing measurement, market research and analysis, metrics improvement: mainly utilizing the NLP techniques of entity recognition, sentiment analyzis, content summarization and topic extraction. It is applied mostly in the MC tasks of budget refinement, business communication, sales, direct marketing.

6.2 Ethical and moral considerations of using AI-based NLP techniques in CC

Although this sub-section discusses issues form the receiver's perspective (as opposed to the sender's perspective used to specify the tasks of the developed models), to fully account for the concerns of the participants of the deep interviews and focus group research, and to understand the discussed CC task – related NLP techniques in their full scope, we found it important to elaborate on the main ethical issues that can arise from the use of NLP techniques for CC tasks.

The ever-growing scope of Natural Language Processing AI techniques present new opportunities for corporations to mine customer intelligence and to fine-tune / improve their communications accordingly. On the other hand, the utilization of such techniques often result in handling of confidential data, possibly unwanted by the target group, who is not provided with the means for adequate self-management of data-privacy. As it has been pointed out by the participants during the research conducted, the use of NLP for internal communication purposes, for example, offers several unexploited possibilities, but may lead to unwanted and unauthorized monitoring of the employees. Thus, the unprecedented growth of data is creating not only business opportunities but complex ethical problems as well, with no standardized solutions to deal with them. The rapid pace of development in data science is not met with the understanding of the legal, moral, and cultural issues associated with its collection, storage, and reputation management-oriented analysis. (Mittelstadt and Floridi 2016; Larson 2013)

The aforementioned problems stemming from the careless use of AI technology or insufficient knowledge about how to use them properly can be considered as indicators for organizational crises with unpredictable negative effects. It is thus important to map the ethical aspects, accounting for cases where the interests of the customers' could be potentially harmed, in order to prepare for and possibly avoid such crisis. Without doing so, the mere usage of AI techniques (or any device utilizing such techniques) need to be considered as possible threats not only for the individual customers, but in an indirect way, for the reputation and image of the whole company. (Neuman and Pintér 2019).

Based on the academic literature and the experimental results of this study, we inferred that the most widely discussed issue regarding the usage of NLP techniques for CC purposes is that the requirement of transparency during data-handling is often not satisfied. The main ethical issues are briefly listed below focusing on non-transparent information gathering and interaction practices.

Information gathering and analyzis

- Private data management consent is often given without reading / full understanding of the risks and consequences. The user, however, might not want his data to be analyzed. (Saqr 2017) The failure of proper consent-gathering has been widely discussed and acknowledged. (Mantelero 2014)
- In order to use a given online service, consensus to private data management is obligatory. With no alternative means to use a service, the user is often pushed to accept any terms and conditions and forced to give access to sensitive information.
- There is no time limit on the utilization of gathered customer data. Terms of agreement (and utilization purposes) might change with time. (Saqr 2019)

The users should be adequately and completely informed about the purposes of the data gathering: how it is collected, stored, transferred and processed. The customer should also be informed if a third party will acquire access / participate in processing of her/his data.

- Users should be given clear guarantees that data will not be sold or transferred to other legal entities without their proper consent.
- Although it is not widely implemented, NLP could be used effectively for internal communication purposes, e.g. to measure employee satisfaction and survey attitudes towards certain work-related topics. On the other hand, this can easily lead into the monitoring of the employees (private search-history on company machines etc.) which would raise a new dimension of ethical and legal problems: where is the line between spying and surveying?

Interaction and communication

- Customers may confuse the artificial agents to human agents. When they realize they have been interacting with a chatbot they may feel tricked. As a possible result, the level of trust in the brand decreases which has a negative effect on the company's reputation.
- Portraying chatbots with human-like avatars (with names, images, personal info) may give customers the false impression they are interacting with a real person, which can lead to long lasting emotional / financial harm for both sides.
- Users are often not provided with sufficient information to clarify who is the legal person who can be held responsible in the case they are misinformed by an artificial agent.

To account for the issues listed above, specific cases of data-abuse are need to be identified and analyzed, in order to develop novel approaches which could guarantee the users the necessary knowledge and means for a more effective self-management of data-privacy.

The rapid growth of information necessitates the orchestrated efforts of all parties involved in CC to understand the legal, ethical, and cultural problems we are facing. As conventional practices do not help to solve the issues of our modern information society, we need novel, widely adaptable, but rigorous policies, applicable not just for the current problems but for future challenges as well. (Varley-Winter and Shah 2016) Thus, a general framework should be developed, governing and supervising NLP-based data-handling. It is up for research to further investigate this problem.

7. Concluding remarks, limitations and further research

This study contributes to the fields of communications and computational linguistics with three main findings:

1. A function-based model of CC has been developed building on Riel's original concept. The synthesizing model resolves the discrepancies and contradictions with other definitions / daily practice; differentiates between strategic and operational, as well as internal and external dimensions; and re-specifies several tasks accounting for the possible intersections of disciplines.
2. As an extension of our validated CC model, an NLP-CC model has been developed, mapping NLP techniques on CC tasks. General examples of organizational/corporational practices has also been discussed giving context to the mapping.
3. Our research addressed the ethical aspects of utilizing NLP for CC purposes revealing two potential sources where the customers' interests can be harmed:
   a) There are no policies regulating the use of artificial agents, which would protect the customer from misleading them into the false belief that they are interacting with a real person /grant the customer proper channels to hold someone responsible in case being misinformed by an artificial agent.
   b) The customer is often not properly informed about how, where and when her/his personal data is utilized.

The findings of this study is based on the author's professional experience in the field of communication and computational linguistics supported by primer qualitative research building on the methodology of deep interviews and focus group research involving experts from both fields. Based on the feedbacks of the participants we defined possible associations between CC tasks and NLP techniques. (The qualitative research also revealed several important unexploited possibilities. E.g. NLP is not utilized widely for internal communication purposes, however several technologies applicable for gathering information from and about the employees). However, we outlaid only general examples of typical organizational/corporational practices without mapping them to company specific use-cases which would lead to a deeper understanding about the key factors of the intersection of the two disciplines. Further empirical research is needed to validate and strengthen the representativeness and universality of our findings especially in the cases where the ratio of consent within the collected answers was lower. In addition, surveying the best practices of large, multinational companies typically utilizing customer feedback, such as Amazon, Rakuten etc. and communication and media agencies would provide further insights on the overlap of the discussed fields.

Both fields could further benefit from supporting the results of qualitative research (conducted so far) with quantitative research methodology as well: surveying larger, statistically significant, diverse sub-samples (accounting for demographic, experience-wise etc. variability) of target population (NLP/CC experts) based on and extending the questions of the deep interviews and focus group research (see Table 2) through online questionnaires.

This pioneering study tries to synthesize NLP (a field of sciences supported by quantitative data and exact definitions) and CC (a field of social sciences with less exact, ever-

changing definitions). Mapping between the two fields is a complex task with no conventional solutions to the knowledge of the authors. Accordingly, during the creation of the proposed models, the authors forced to make several subjective decisions: what aspects to consider during the selection of participants, how to resolve disagreement/ contradiction between the interviewed experts and to what extent should they rely on the expertise of the interviewed participants and previous work.

Our study concentrates rather on finding the correspondence of NLP and CC, than solving the definitional problems and providing a rigorous differentiation between the disciplines of corporate communication. Although our models were developed from an AI perspective, they also account for several discrepancies of the well-known but overly-flexible and obsolete (in the context of modern information society) framework of CC by Riel. This approach, however, is just one of the many possible ones, which can also be thought as a debate-generating initiative. Instead of concentrating on a holistic perspective as this study did, it could be beneficial to analyze specific use-cases between given NLP techniques and CC tasks on a more detailed level: e.g. conducting case-studies how entity recognition can be utilized for crisis communication purposes, etc. Modern, deep learning-based NLP solutions could be, for example, utilized in crisis communication to: predict organizational crises, to analyze the possible outcomes, to classify different crisis categories, to speed up response-generation or to build AI bots assisting crisis communication experts. (Coombs 2007)

This is a practice-driven study where theory is criticized and refined considering the daily routine. As it has been pointed out by Pinter (2019) - paraphrasing Gregory-Miller's (1988) models of "Public understanding of science" - in fields with strong pragmatic outcomes and dependencies on ever-improving technical tools such as communications and computational linguistics, regular revision of theory in the light of an equally valuable consideration factor: practice, is crucial. Furthermore, involving experts in scientific research to keep track of the emerging challenges and the development they necessitate in this rapidly changing, modern information society is of outmost importance.

Bibliography

An, S. K. and I. H. Cheng. "Crisis Communication Research in Public Relations Journals: Tracking Research Trends Over Thirty Years", in *The Handbook of Crisis Communication* edited by Coombs, W. Timothy; Holladay, Sherry J., 65-89. Oxford, UK., Wiley-Blackwell, 2010. http://dx.doi.org/10.1002/9781444314885.ch3

Beniger, J. R. *The Control Revolution: Technological and Economic Origins of the Information Society*, Cambridge, Mass.: Harvard University Press. 1986.

Bun, K. K. and M. Ishizuka. "Topic extraction from news archive using TF* PDF algorithm. "*Proceedings of the Third International Conference on Web Information Systems Engineering*, Institute for Electrical and Electronics Engineer (2002) http://dx.doi.org/10.1109/WISE.2002.1181645

Bundy, J; M. D. Pfarrer; C. E. Short and W. T. Coombs. "Crises and crisis management: Integration, interpretation, and research development", *Journal of Management* vol 43, no 6 (2016): 1661–1692. http://dx.doi.org/10.1177/0149206316680030

Carenini, G.; G. Murray and R. Ng. *Methods for Mining and Summarizing Text Conversations: Synthesis Lectures on Data Management*, Morgan & Claypool Publishers, 2011. http://dx.doi.org/10.1145/2348283.2348529

Castells, M. *End of Millennium: The Information Age: Economy, Society and Culture* vol 3. Cambridge, Massachusetts; Oxford, UK: Blackwell. 1998. ISBN 978-0-631-22139-5.

Castells, M. *The Power of Identity, The Information Age: Economy, Society and Culture* vol. 2. Cambridge, Massachusetts; Oxford, UK: Blackwell. 1997. ISBN 978-1-4051-0713-6.

Castells, M. *The Rise of the Network Society, The Information Age: Economy, Society and Culture* vol. I. Cambridge, Massachusetts; Oxford, UK: Blackwell. 1996. ISBN 978-0-631-22140-1.

Christensen, L. T. and J. Cornelissen. "Bridging corporate and organizational communication: review, development and a look to the future", *Management Communication Quarterly* vol 25, no 3 (2011): 383-414. https://doi.org/10.1177/0893318910390194

Coombs, W. T. *Ongoing Crisis Communication: Planning, Managing, and Responding*, Los Angeles: Sage. 2007.

Cornelissen, J. *Corporate Communication: A Guide to Theory and Practice*, Sage, London. 2008.

Frandsen, F. and W. Johansen. "Corporate communication" in *The Routledge Handbook of Language and Professional Communication* edited by Bhatia, Vijay; Bremner, Stephen, Routledge, London. 2013.

Goldberg, Y. "A Primer on Neural Network Models for Natural Language Processing", *Journal of Artificial Intelligence Research* vol 57. (2016):345–420. arXiv:1807.10854

Goldberg, Y. "Neural network methods for natural language processing", *Synthesis Lectures on Human Language Technologies* vol 10, no 1 (2017) https://doi.org/10.2200/S00762ED1V01Y201703HLT037

Goldhaber, G. M. *Organizational Communication*, (6th ed.), Brown & Benchman, Madison, WI. 1993. [originally published in 1974]

Goodfellow, I.; Y. Bengio and A. Courville. *Deep Learning*, MIT Press, 2016.

Goodman, M. B. and P. B. Hirsch. *Corporate Communication*, Peter Lang Inc., International Academic Publishers, New York, 2010. ISBN-13: 978-1433106217

Gregory, J. and S. Miller. *Science in public: communication, culture, and credibility*, NewYork: Plenum.1988.

Ihász, P. L.; M. Kovács and V. V. Kryssanov. "Emotion Recognition through Intentional Context." *International Journal of Affective Engineering* vol 18, no 1 (2019):17-25. https://doi.org/10.5057/ijae.IJAE-D-18-00002

Jozefowicz, R.; O. Vinyals; M. Schuster; N. Shazeer and Y. Wu. *Exploring the Limits of Language Modeling*. 2016. arXiv:1602.02410v2

Koehn, P. *Statistical machine translation*, Cambridge University Press. 2009.

Larson, E. "Building trust in the power of "big data" research to serve the public good." *JAMA: the journal of the American Medical Association* vol 309, no 23 (2013): 2443-2444. https://doi.org/10.1001/jama.2013.5914

Bing, L. *Sentiment analysis: Mining opinions, sentiments, and emotions*, (1st ed.) New York, USA.: Cambridge University Press, 2015.

Mani, I. *Advances in automatic text summarization*. MIT Press, Cambridge, United States. 1999.

Mantelero, A. "The future of consumer data protection in the E.U. Re-thinking the 'notice and consent' paradigm in the new era of predictive analytics." *Computer Law & Security Review*, vol 30 (2014): 643–60. https://doi.org/10.1016/j.clsr.2014.09.004

Mittelstadt, B. D. and L. Floridi. "The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts." *Science and Engineering Ethics* vol 22, no 2 (2015): 303-41. https://doi.org/10.1007/s11948-015-9652-2

Neuman, P. and D. G. Pintér. "Technology-based critical phenomena: a Borgmannian approach of crisis prediction." in *Essays in Post-Critical Philosophy of Technology* edited by Héder, Mihály; Nádasi, Eszter, Ch.13: 125-134. Vernon Press. 2019. ISBN: 978-1-62273-457-3

Pintér, D. G. "Revealing the connections between Situational Crisis Communication Theory and Framing Theory with their obstacles and opportunities for improvement in connection with media analysis of police news conference on the 2016 Budapest explosion." *Doctoral Dissertation*, Eötvös Loránd University, 167-168. (2019)

Pintér, D. G.. "Various Challenges of Science Communication in Teaching Generation Z: an Urgent Need for Paradigm Shift and Embracing Digital Learning." *Opus et Educatio* vol 3, no. 6 (2016): 1-25. http://dx.doi.org/10.3311/ope.146

Pintér, D. G. „Media Bias and the Role of User Generated Contents in Crisis Management: a Case-Study about the Communication of the Hungarian Police Forces after 2016 Budapest Explosion." *Corvinus Journal of Sociology and Social Policy* vol 9, no 1 (2018): 101-125. https://doi.org/10.14267/CJSSP.2018.1.05

Poole, D.; R. Goebel and A. Mackworth. *Computational intelligence: A Logical Approach*, New York: Oxford University Press. 1998.

Riel, C. B. M. van., "Defining corporate communication" in *Corporate Communication: A Strategic Approach to Building Reputation*, edited by Bronn, Peggy Simcic; Wiig, Roberta. Gyldendal Norsk Forlag, Oslo, (2003):21-40. ISBN 9780415328265.

Riel, C. B. M. van and C. J. Fombrun. *Essentials Of Corporate Communication*, Abingdon & New York: Routledge. 2007. ISBN 9780415328265.

Riel, C. B. M. van. *Principles of Corporate Communication*, Prentice Hall Europe, Hemel Hempstead. 1995.

Russell, S. J. and P. Norvig. *Artificial Intelligence: A Modern Approach* (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall. 2003. ISBN 0-13-790395-2

Saqr, M. "Big data and the emerging ethical challenges," *International Journal of Health Sciences* vol 11, no 4 (2017): 1–2.

Sordoni, A; M. Galley; M. Auli; C. Brockett; Y. Ji; M. Mitchell; J. Y. Nie; J. Gao and B. Dolan. "A neural network approach to context-sensitive generation of conversational responses." *Proceedings of NAACL-HLT* (2015) arXiv:1506.06714

Varley-Winter, O. and H. Shah. "The opportunities and ethics of big data: practical priorities for a national Council of Data Ethics," *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences* vol 374, no 2803 (2016) https://doi.org/10.1098/rsta.2016.0116

Xiong, C.; V. Zhong and R. Socher. "Dynamic coattention networks for question answering." *Proceedings of ICLR* (2017) arXiv:1611.01604