The state of natural language and applied natural language. Language processing in logistics and supply chain management

Joel Cedric Lengeling * [1]

[1] Chair of Logistics, Technical University Berlin – Fachgebiet Logistik Sekr. H 90 Technische Universität Berlin Straße des 17. Juni 135 10623 Berlin, Germany

Abstract:

Purpose of the communication: As of today, research with regards to natural language processing (NLP) and applied NLP in logistics and supply chain management (LSCM) is limited. Natural languages are the languages that are used for everyday communication by humans and NLP is one of the names given to the field that studies the processing of natural language (NL) by machines; either in written or spoken form. Applied NLP refers to the application of methods or algorithms that have been established trough NLP research to process NL. The information flow in LSCM is usually carried by NL. Any interruption by any incident along that supply chain leads to a call, an email or any other form of communication applying NL. Any request by a customer leads to communication in NL. Any deal between a supplier and a manufacturer leads to a contract written in NL. Thus NL is necessary to almost all forms of information flow in LSCM. Nevertheless limited research towards NL, NLP and applied NLP in LSCM has been conducted thus far. This contribution aims to provide an overview to a sector of that still limited research by conducting a survey on NL in LSCM while assessing the current state of applied NLP.

Research design, methodological approach: Starting point of this contribution is a digital questionnaire. Experts in the field are asked to rate the amount of NL that accumulates along the customer order process within their enterprises. Those experts are recruited from the retail, manufacturing and service industry. After rating the amount of NL that accumulates, the experts are asked to specify further with regards to applied NLP.

Results obtained: The outcome of this contribution provides an overview regarding the amount of NL that occurs along the customer order process while also indicating how the current state of applied NLP in LSCM is.

Theoretical contributions: The LSCM research community benefits from this research by receiving a better indication regarding the areas in which research regarding NLP and applied NLP in LSCM might be useful and even necessary.

Managerial contributions: Industry benefits from this contribution by receiving an indication regarding areas in LSCM that might benefit from applying NLP, while also receiving an overview regarding the current state of applied NLP in the industry.

Limitations: The limitations for digital survey research apply.

*Speaker

Keywords: natural language, natural language processing, NLP, applied natural language processing, logistics and supply chain management

THE STATE OF NATURAL LANGUAGE AND APPLIED NATURAL LANGUAGE PROCESSING IN LOGISTICS AND SUPPLY CHAIN MANAGEMENT

Abstract:

Purpose of the communication: As of today, research with regards to natural language processing (NLP) and applied NLP in logistics and supply chain management (LSCM) is limited. Natural languages are the languages that are used for everyday communication between humans and NLP is one of the names given to the field that studies the processing of natural language (NL) by machines; either in written or spoken form. Applied NLP refers to the application of methods or algorithms that have been established trough NLP research to process NL. The information flow in LSCM is usually carried by NL. Any interruption by any incident along that supply chain leads to a call, an e-mail or any other form of communication applying NL. Any request by a customer leads to communication in NL. Any deal between a supplier and a manufacturer leads to a contract written in NL. Thus NL is necessary to almost all forms of information flow in LSCM. Nevertheless limited research towards NL, NLP, and applied NLP in LSCM has been conducted thus far. This contribution aims to provide an overview to a sector of that still limited research by conducting a survey on NL in LSCM while assessing the current state of applied NLP

Research design, methodological approach: Starting point of this contribution is a digital questionnaire. Experts in the field are asked to rate the amount of NL that accumulates along the customer order process within their enterprises. Those experts are recruited from the retail, manufacturing and service industry. After rating the amount of NL that accumulates, the experts are asked to specify further with regards to applied NLP.

Results obtained: The outcome of this contribution provides an overview regarding the amount of NL that occurs along the customer order process while also indicating how the current state of applied NLP in LSCM is.

Theoretical contributions: The LSCM research community benefits from this research by receiving a better indication regarding the areas in which research regarding NLP and applied NLP in LSCM might be useful and even necessary.

Managerial contributions: Industry benefits from this contribution by receiving an indication regarding areas in LSCM that might benefit from applying NLP, while also receiving an overview with regards to the current state of applied NLP in the industry.

Limitations: The limitations for digital survey research apply.

Keywords: "natural language processing”; "applied natural language processing”; "logistics and supply chain management"; “survey”;
Paper type: full paper

1. INTRODUCTION

Natural languages are the languages that are used for everyday communication between humans (Bird, Klein, and Loper 2009). Within the logistics and supply chain management (LSCM) industry, communication between entities is usually carried out applying natural language. A number of very likely scenarios are: Whenever an urgent incident along the supply chain occurs, the manager of the supply chain gets a phone call from the shipper. Whenever a customer has an issue with a product he bought, he sends an e-mail to the customer service department of the retailer. Whenever a procurement agent has decided to source a certain component from a new supplier, a contract is signed linking the two businesses. Phone calls, e-mails, and contracts are all examples of applied natural language, as the language applied by them, is the language for everyday communication between humans.

The processing of natural language by a machine, requires the application of rather specific methods and algorithms. Those specific methods and algorithms have been subject to research and developed by experts from a variety of fields. Major contributions have been achieved by computer scientists, linguists, electrical engineers, cognitive science scientists, and data scientists. One result of that scientific melting pot is a rather diverse number of technical terms. Occasionally different terms are applied to describe almost identical subjects. The scientific background of the researcher and the decade of the publication seem to impact which terms are applied by the researcher. Natural Language Processing (NLP) and Computational Linguistics (CL) are two dominating terms within the English speaking scientific community dealing with those matters. There seems to be a prevalence for NLP by computer scientists and for CL by linguists.

In 1999 Cunningham summarized CL to "CL is a part of the science of language that uses computers as investigative tools" and NLP to "NLP is part of the science of computation whose subject matter is computer systems that process human language"(Cunningham 1999). In 2009, Bird et. al defined NLP as "to cover any kind of computer manipulation of natural language. At one extreme, it could be as simple as counting word frequencies to compare different writing styles. At the other extreme, NLP involves “understanding” complete human utterances, at least to the extent of being able to give useful responses to them" (Bird, Klein, and Loper 2009). In 2010, Carstensen et. al defined the research subject of CL as the processing of natural language, both in written or spoken form, by a computer (Carstensen, Ebert, Endriss, Jekat, Klabunde, and

Langer 2010). Rather subjectively, throughout this text and the survey, NLP is used to refer to the interdisciplinary field that studies algorithms and methods that enable machines to process natural language. Cunningham defined (Natural) Language Engineering as "the application of NLP to the construction of computer systems that process language for some task usually other than modelling language itself" (Cunningham 1999). Nonetheless, Natural Language Engineering as a name for applied NLP does not seem to have prevailed. In order to maximize the audience for this contribution, the application of a method or an algorithm, that has been established trough NLP research, as part of an overlaying software product, is referred to as applied NLP. Two well-known classical examples of software products that rely on NLP internally are spell checkers and spam detection systems (Jurafsky and Martin 2009). Newer examples are personal digital assistants and online translation services. Generally speaking, any product that executes any algorithm or method that has natural language as an input is considered an example of applied NLP. Specific examples from the domain of LSCM are rule based voice assisted picking systems and customer services chat bots.

No scientific contribution in which LSCM processes or LSCM activities with a high natural language relevance are identified, has been found in the preparation of this contribution. It seems that as of today, research regarding natural language, NLP, and applied NLP in LSCM is still limited. Hellingrath and Lechtenberg noted in their review on applications of artificial intelligence in LSCM that “surprisingly, the application of approaches such as natural language processing or image recognition to SCM tasks has not been well researched so far” (Hellingrath and Lechtenberg 2018). By conducting a survey among industry experts within the LSCM industry, this contribution aims to identify LSCM processes and LSCM activities in which natural language has a high relevance. Additionally, this contribution aims to identify the current state of applied NLP within the LSCM industry.

In the next section, the framework applied throughout the survey is presented. In section 3 the methodology of the survey is described. In section 4 the results and implications of the survey are presented. In section 5 limitations are described, an outlook is given and some final remarks are made.

2. THE CUSTOMER ORDER PROCESS FRAMEWORK

As previously described, this contribution aims to identify processes and activities with a high natural language relevancy within LSCM, while additionally identifying the current state of applied NLP within the LSCM industry. Straube defines process oriented logistics from the viewpoint of an enterprise, as the planning, controlling, execution, and verification of all flows of materials and all flows of information between enterprises, from customer to suppliers and between all the other value adding participants (Straube 2004). This definition shows how vast the domain of LSCM is. In order to support the participating industry experts dealing with this vastness, a scientific framework describing LSCM has been applied. The applied framework is called “Kundenauftragsprozess”, or translated into English, “Customer order process”. That customer order process has been developed over time by Baumgarten. He initially introduced it in 1999 and further detailed it in 2000 (Baumgarten 1999, Baumgarten and Inga-Lena 2000). As an international survey audience was chosen, the original German description of the framework was translated into English by the author of this contribution. Throughout the rest of this contribution that English translation of the framework is referred to.

Baumgarten develops the customer order process applying methods defined within process chain management. Those methods allow the definition of processes in an iterative way. A process is made up out of (sub-)processes, and activities and thus, multiple concatenated or chained activities and processes create a process. One difference between activities and processes is, that a process might be further divided while an activity may not be further divided. It is important to note, that activities need to be target driven, and that activities need both a measurable input and a measurable output. Activities should generate a measurable increase in value. Activities and processes are either directly related to the product (primary) or related to an administrative task (secondary). Processes enable the flow of material or the flow of information, they are repetitive and deterministic. Additionally, processes should aim to be customer driven. All activities and processes might be executed in sequence, in parallel, or iteratively.

Baumgarten defined the customer order process itself applying four sub-processes: *Development*, *Supply*, *Order Execution,* and *Disposal*. According to Baumgarten, the *Development* process incorporates all activities and processes that take place between the initial product idea and the market entry of the product. The *Supply* process incorporates all activities and processes that need to be undertaken in order to supply the production process. The *Order Execution* process incorporates all activities and processes that need to be undertaken to fulfill a customer order. The *Disposal* process incorporates all activities and processes that need to be undertaken to recycle the product.

Baumgarten subdivides the *Supply* process into the (sub-)processes and activities: *Assessment of Demand*, *Purchasing,* and *Inbound Logistics*. Baumgarten further subdivides the *Order Execution* process into the (sub-)processes and activities: *Order Processing*, *Production Planning and Controlling*, *Production*, *Distribution,* and *After-Sales-Services*. Although Baumgarten subdivides *Disposal* into three further (sub-)processes and activities, the author of this contribution decided to not subdivide *Disposal* further due to the survey setting. Throughout the survey, participants were asked questions in relation to the different processes. Without subdividing *Disposal* further, participants are asked questions related to ten different processes. With subdivision of *Disposal*, that number would have risen to twelve. Such a large number was deemed not fitting for a survey setting. The resulting framework is illustrated in Figure 1.

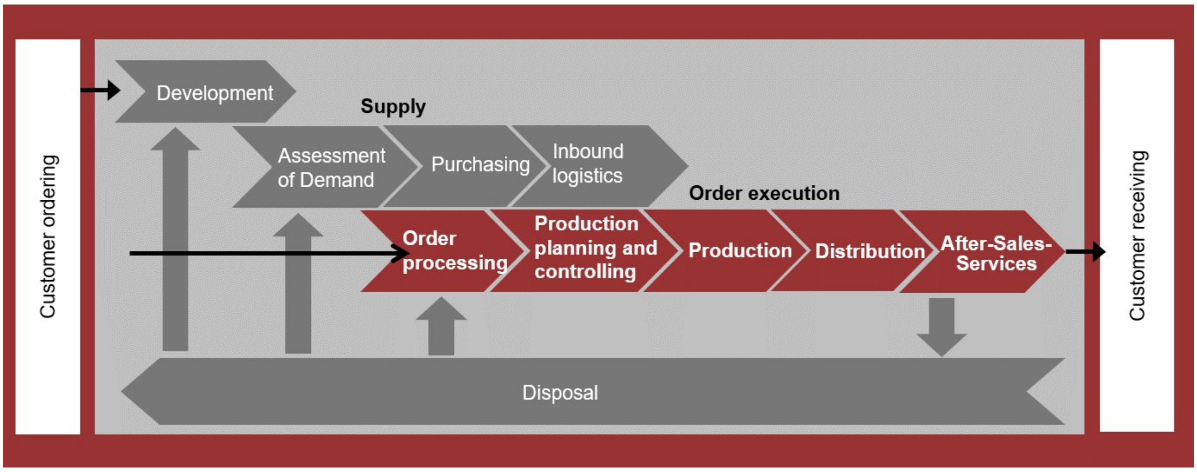

Figure 1 – Customer order process, illustration based on Baumgarten (Baumgarten 1999, Baumgarten and Inga-Lena 2000)

3. RESEARCH DESIGN

According to Saunders et. al. a survey strategy tends to be used for exploratory and descriptive research. (Saunders, Lewis, and Thornhill 2009). Saunders et. al. define an exploratory study in their glossary as “research that aims to seek new insights into phenomena, to ask questions, and to assess the phenomena in a new light". Due to the still limited available research regarding natural language, NLP, and applied NLP in LSCM such a study seems appropriate. According to Saunders et. al., a questionnaire might be used as an instrument when applying a survey strategy (Saunders, Lewis, and Thornhill 2009). The methodology applied throughout the survey is thus exploratory and qualitative.

In order to identify LSCM processes and LSCM activities with a high natural language relevance while also identifying the current state of applied NLP within the LSCM industry, a digital questionnaire was applied. During two months in the autumn of 2019, roughly 700 industry experts from the manufacturing, retail, and logistics service industries have been given the opportunity to participate. In total 13 members from the logistics service industry – including one participant that offers software services –, 17 members from the manufacturing industry, and 3 members from the retail industry, thus a total of 33 invitees, accepted the invitation.

The European Union has defined small and medium-sized enterprises (SMEs), as enterprises that employ less than 250 employees while having a turnover below 50 million euros. The participants were asked to identify, the type of enterprise they are employed by, applying that same definition. Of the 33 participants, 4 are employed by a SME, 24 are employed by a bigger enterprise and 5 noted, that they rather would prefer, not to answer that question.

Since the LSCM domain is rather vast, participants were first asked to rate their own expertise within the different processes. Throughout the rest of the survey, participants were only asked questions related to the processes for which they indicated that they "have some expertise" or that they are "an expert in the field". As a result, 31 of the 33 participants have been asked questions related to the *Distribution* process while ten participants have been asked questions related to the *Development* process. Only one participant indicated to have at least some expertise in two processes, while ten participants indicated to have expertise in five processes, and four participants indicated expertise in all ten processes.

The participants have been asked to rank the processes they specified they had expertise in, regarding the relevancy of natural language, the relevancy of applied NLP, and the current state of NLP. The definition provided in the Introduction Chapter of this contribution for natural language, NLP, and applied NLP was also provided to the participants. The participants were only aske regarding the general application of computers to process natural language in the domain and not regarding the application of a specific method or algorithm. In addition, optional questions allowing participants to further elaborate on their answers, were incorporated into the questionnaire.

As previously mentioned, the participants have been only questioned about processes for which they indicated at least some expertise. Thus, for example, seven participants indicated to have at least some expertise in both *Development* and *Assessment of Demand*, thus, seven pairwise comparisons with both those processes – see Table 1 – do exist. Throughout the rankings a lower rank indicates a higher relevancy. The number of pairwise comparisons between the different processes are displayed in Table 2. In total, 45 process comparisons with each containing a variable number of pairwise comparisons within, do exist.

Table 1 – Natural language relevancy ranking between Development and Assessment of Demand.

Participant Id	**Development rank**	**Assessment of Demand rank**
1	**3**	**4**
2	**3**	**5**
3	**7**	**1**
4	**8**	**4**
5	**1**	**2**
6	**2**	**3**
7	**2**	**1**

Due to the fact, that the number of participants does not allow to draw conclusions solely relying on Bernoulli's Theorem, the methodology applied does not rely on a certain statistical distribution. Additionally enterprise and industry background was not taken further into consideration due to the number of participants. Nevertheless, results are presented in Table 3, Table 4, and Table 5 of Section 4.1. Those tables contain the number of pairwise comparisons, which rank the process of the row with a higher relevancy compared to the process of the column. Thus, for instance, in Table 2, the matrix entry with row P1 and column P2 contains the value four, since four participants (id 1, 2, 5, and 6; see Table 1) ranked the *Development* process to have a higher relevancy for natural language compared to the *Assessment of demand* process.

Table 2 – Matrix displaying the number of pairwise comparisons between the processes. P1 aka Development, P2 aka Assessment of Demand, P3 aka Purchasing, P4 aka Inbound Logistics, P5 aka Order Processing, P6 aka Production Planning and Controlling, P7 aka Production, P8 aka Distribution, P9 aka After-Sales-Services and P10 aka Disposal.

	Development	Assessment of Demand	Purchasing	Inbound Logistics	Order Processing	Order Processing	Production	Distribution	After-Sales-Services	Disposal
	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**
P1	X	7	8	9	9	7	8	10	6	6
P2	7	X	15	18	17	15	13	19	12	11
P3	8	15	X	23	21	17	16	22	14	12
P4	9	18	23	X	28	22	19	30	17	14
P5	9	17	21	28	X	23	19	28	16	13
P6	7	15	17	22	23	X	17	22	13	11
P7	8	13	16	19	19	17	X	19	12	15
P8	10	19	22	30	28	22	19	X	17	14
P9	6	12	14	17	16	13	12	17	X	10
P10	6	11	12	14	13	11	15	14	10	X

4. RESULTS AND IMPLICATIONS

Throughout the first subsection of this section, the results of the different rankings are presented. Throughout the second subsection of this section, some implications, with additional anecdotal answers provided by the questionnaire, are given.

4.1. Results

Table 3 displays the number of pairwise comparisons that rank the process of the row with a higher natural language relevancy compared to the process of the column. The distribution, of the difference between the number of pairwise comparisons between two processes – the difference, for any two processes x and y where $x \neq y$, between the values stored in (Px, Py) and (Py, Px) – is displayed in Figure 2. Thus, there are six process comparisons – P1 and P3, P1 and

P8, P1 and P9, P3 and P7, P4 and P5 as well as P4 and P8 – which have the exact same number of pairwise comparisons rating either process with a higher relevancy compared to the other process. Taking the process comparison between P1 and P9 as well as Bernoulli's Theorem into consideration, the fact that according to three experts, the *Development* process has a higher natural language relevancy compared to the *After-Sales-Service* process, and that according to three other experts, the *Development* process has a lower natural language relevancy compared to the *After-Sales-Service* process, is not that expressive. At the same time, the fact that in the comparison between the *Inbound Logistics* process (P4) and the *Order Processing* process (P5), 14 experts rated that the *Inbound Logistics* process has a higher natural language relevancy compared to the *Order Processing* process and that 14 other exports rated, that the *Inbound Logistics* process has a lower natural language relevancy compared to the *Order processing* process, seems more expressive. Figure 2 shows that there are 27 process comparisons, with a smaller (0, 1 or 2) difference in number of pairwise comparisons between two processes. Of the 18 remaining process comparisons, nine were caused by two particular processes. These two processes are the *Production* process (P7) and the *Disposal* process (P10). Of the 15 experts that had experience in both process areas, 10 ranked the *Production* process higher than the *Disposal* process and thus 5 ranked the *Production* process lower than the *Disposal* process. Besides this process ranking, the *Production* process was involved in three other rankings were the difference in number of pairwise comparisons between the two involved processes was larger than two. The *Disposal* process was involved in five other rankings were the difference in number of pairwise comparisons between the two involved processes was larger than two. Additionally, the *Disposal* process stood out with regards to process comparisons. It is the only process that was consistently ranked lower in the overall number of pairwise comparisons between processes. This is especially recognizable when comparing the row and the column of P10 in Table 3.

Table 3 – Matrix displaying the natural language relevancy rankings between processes. P1 aka Development, P2 aka Assessment of Demand, P3 aka Purchasing, P4 aka Inbound Logistics, P5 aka Order Processing, P6 aka Production Planning and Controlling, P7 aka Production, P8 aka Distribution, P9 aka After-Sales-Services and P10 aka Disposal.

	Development	**Assessment of Demand**	**Purchasing**	**Inbound Logistics**	**Order Processing**	**Order Processing**	**Production**	**Distribution**	**After-Sales-Services**	**Disposal**
	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**
P1	X	4	4	5	5	3	7	5	3	4
P2	3	X	7	11	9	11	10	10	8	7
P3	4	8	X	12	11	8	8	13	8	7
P4	4	7	11	X	14	12	8	15	10	10
P5	4	8	10	14	X	11	10	13	10	10
P6	4	4	9	10	12	X	12	10	9	8
P7	1	3	8	11	9	5	X	10	7	10
P8	5	9	9	15	15	12	9	X	8	10
P9	3	4	6	7	6	4	5	9	X	8
P10	2	4	5	4	3	3	5	4	2	X

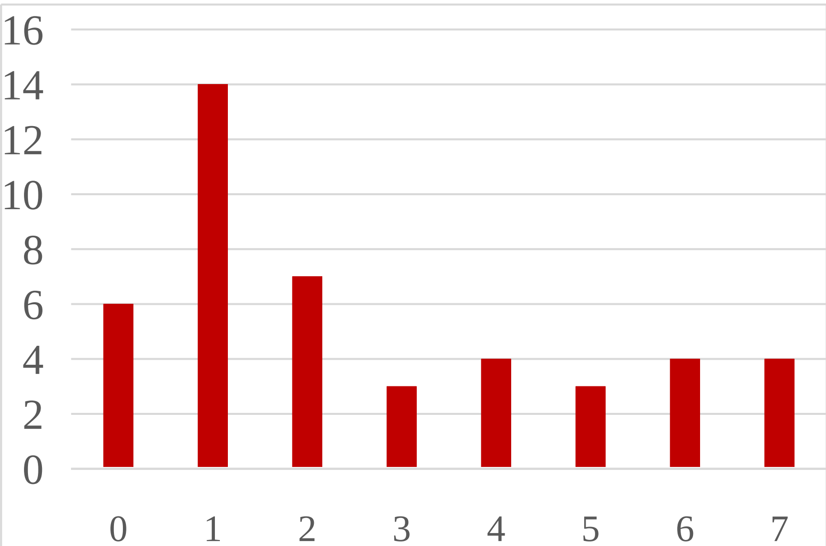

Figure 2 – The distribution of the difference between (PX,PY) and (PY,PX) of Table 5

Table 6 displays the number of pairwise comparisons that rank the process of the row with a higher NLP relevancy compared to the process of the column. The distribution, of the difference between the number of pairwise comparisons between two processes – the difference, for any two processes x and y where x ≠ y, between the values stored in (Px, Py) and (Py, Px) – is displayed in Figure 3. Compared to the natural language relevance comparisons, the NLP relevance comparisons is more pronounced. Due to the results, it is possible to create a top five ranking. The first ranked process was consistently ranked higher in the overall number of pairwise comparisons between the processes. The second ranked process was ranked higher in the pairwise comparisons between the processes, except in the ranking with the first ranked process. The third ranked process was ranked higher in the pairwise comparisons between the processes, except in the ranking with the first ranked and second ranked process and so on. In this top five, the first rank belongs to the *Assessment of Demand* process, the second rank belongs to the *Order processing* process, the third rank belongs to the *Purchasing* process, the fourth rank belongs to the *Inbound logistics* process, and the fifth rank belongs to the *Development* process. This is recognizable when comparing the rows and the columns of P1, P2, P3, P4, and P5 in Table 6. It has to be noted though, that in the top three of this top five, the difference between the number of pairwise comparisons between the processes is minimal. In the pairwise process comparison between the first ranked process, the *Assessment of Demand* process, and the second ranked process, the *Order processing* process, nine participants ranked *Assessment of Demand* to have a higher NLP relevancy and eight participants ranked *Order Processing* to have a higher NLP relevancy.

Table 4 – Matrix displaying the natural language processing relevancy ranking between processes. P1 aka Development, P2 aka Assessment of Demand, P3 aka Purchasing, P4 aka Inbound Logistics, P5 aka Order Processing, P6 aka Production Planning and Controlling, P7 aka Production, P8 aka Distribution, P9 aka After-Sales-Services and P10 aka Disposal.

	Development	**Assessment of Demand**	**Purchasing**	**Inbound Logistics**	**Order Processing**	**Order Processing**	**Production**	**Distribution**	**After-Sales-Services**	**Disposal**
	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**
P1	X	3	3	4	4	4	5	5	4	4
P2	4	X	8	11	9	10	10	12	8	8
P3	5	7	X	14	10	13	13	15	12	10
P4	4	7	9	X	6	13	12	20	10	11
P5	5	8	11	22	X	19	15	22	12	13
P6	3	5	4	9	4	X	9	12	7	5
P7	2	3	3	7	4	8	X	12	7	10
P8	5	7	7	10	6	10	7	X	11	10
P9	2	4	2	7	4	6	5	4	X	8
P10	2	3	2	3	0	6	5	4	2	X

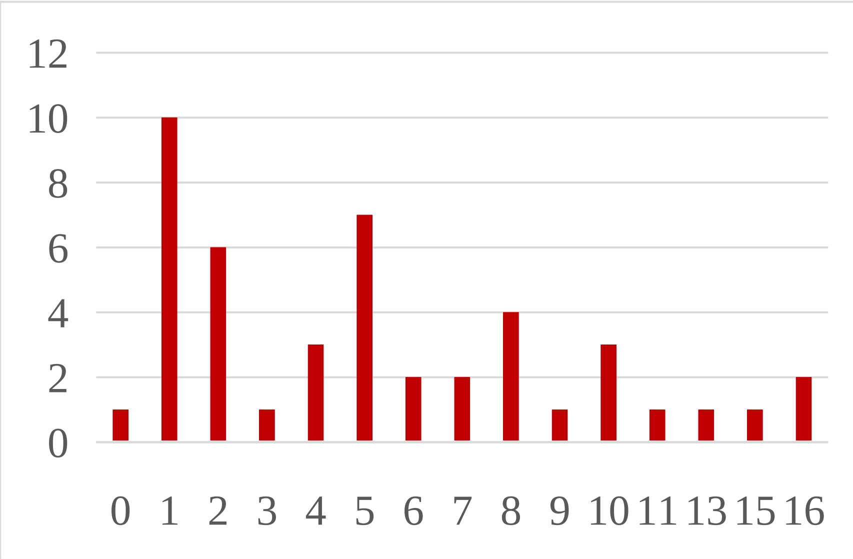

Figure 3 – The distribution of the difference between (PX,PY) and (PY,PX) of Table 6

Table 5 displays the number of pairwise comparisons that rank the current state of NLP higher in the process of the row compared to the process of the column. The distribution, of the difference between the number of pairwise comparisons between two processes – the difference, for any two processes x and y where x ≠ y, between the values stored in (Px, Py) and (Py, Px) – is displayed in Figure 4. There are two processes that draw attention among the currents state of NLP pairwise process comparisons. The first process is the *Assessment of Demand* process, the second process is the *Disposal* process. The *Assessment of Demand* process is, similar to the relevancy for NLP ranking, the only process that has been consistently ranked higher in the overall number of pairwise comparisons between processes. At the same time, the *Disposal* process is, similar to the relevancy for natural language ranking, the only process that has been consistently ranked lower in the overall number of pairwise comparisons between processes. This is recognizable when comparing the two rows and columns of P2 and P10 in Table 5. Otherwise, the ranking regarding the current state of NLP is similar to the ranking for NLP relevancy. The *Order processing process*, the *Purchasing* process, and the *Inbound logistics* process where are all ranked rather high.

Table 5 – Matrix displaying the current state of natural language processing ranking between processes. P1 aka Development, P2 aka Assessment of Demand, P3 aka Purchasing, P4 aka Inbound Logistics, P5 aka Order Processing, P6 aka Production Planning and Controlling, P7 aka Production, P8 aka Distribution, P9 aka After-Sales-Services and P10 aka Disposal

	Development	**Assessment of Demand**	**Purchasing**	**Inbound Logistics**	**Order Processing**	**Order Processing**	**Production**	**Distribution**	**After-Sales-Services**	**Disposal**
	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**
P1	**X**	**2**	**3**	**4**	**3**	**2**	**6**	**6**	**5**	**4**
P2	**5**	**X**	**11**	**12**	**10**	**10**	**9**	**14**	**10**	**9**
P3	**5**	**4**	**X**	**12**	**8**	**8**	**11**	**14**	**12**	**11**
P4	**5**	**6**	**11**	**X**	**11**	**13**	**10**	**15**	**12**	**12**
P5	**6**	**8**	**13**	**17**	**X**	**20**	**15**	**19**	**12**	**13**
P6	**5**	**5**	**9**	**9**	**3**	**X**	**12**	**13**	**10**	**9**
P7	**2**	**4**	**5**	**9**	**4**	**5**	**X**	**11**	**9**	**11**
P8	**4**	**5**	**8**	**15**	**9**	**9**	**8**	**X**	**13**	**12**
P9	**1**	**2**	**2**	**5**	**4**	**3**	**3**	**4**	**X**	**9**
P10	**2**	**2**	**1**	**2**	**0**	**2**	**4**	**2**	**1**	**X**

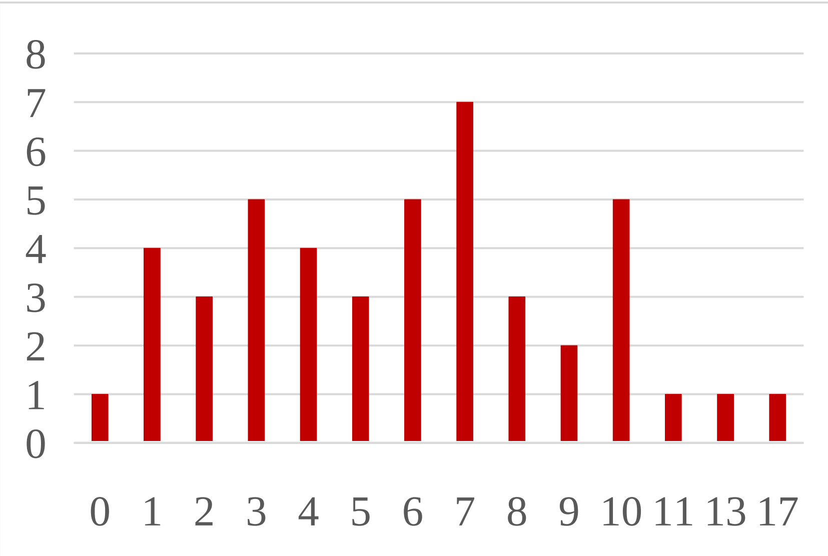

Figure 4 – The distribution of the difference between (PX,PY) and (PY,PX) of Table 5

4.2. Implications

As previously noted, participants were asked optional questions in which they could further elaborate on their answers. Throughout this section the results will be discussed while also incorporating some anecdotal answers provided by participants. Additionally, one participant noted in an open question, that he wasn't confident about some answers he provided. Thus, his results have to be taken with a grain of salt.

The natural language relevancy ranking indicates, that there are no processes that draw extremely high attention, but indicates a more or less equally distributed ranking of processes. This does make sense taking the earlier given definition by Straube into consideration. Straube defined process oriented logistics from the viewpoint of an enterprise, as the planning, controlling, execution, and verification of all flows of materials and all flows of information between enterprises, from customer to suppliers and between all the other value adding participants (Straube 2004). The flow of information within LSCM is usually carried out applying natural language. A LSCM example which does not necessarily rely on applying natural language in order to communicate information, is an automatic replenishment systems. Such systems rely on electronic data interchange which might be implemented without applying natural language (Daugherty, Myers, and Autry 1999).

In order to identify NLP use cases within the LSCM domain, the author is performing a Systematic Literature Review (SLR) in parallel to this contribution. Throughout that SLR, multiple potential use cases have already been identified. It has to be noted though, that most identified use cases are either just indirectly related to LSCM or that they need to be translated into a LSCM use cases. Multiple scientific contributions in which authors perform opinion mining and sentiment analysis have been identified. Throughout this survey, the *Assessment of Demand* process has been ranked high in the NLP relevancy ranking and in the current state of NLP ranking. Conducting opinion mining and sentiment analysis as part of the *Assessment of Demand* process does seem to make sense. As a matter of fact some participants noted, that NLP might be applied to improve forecasting. Multiple participants also mentioned conducting surveys and the application of NLP as part of the After-Sales-Service process. Chen et al.

suggest combining user reviews with geographical data of the user in order to improve marketing (Chen, Wan, and Xu 2016).

Some other use cases that have been identified through the SLR and that might be related to process participants have been questioned about are: web intelligence gathering and handwritten text recognition. In their review on web intelligence gathering, Berkan and Trubatch suggest applying web intelligence gathering for supply chain reconnaissance (Berkan and Trubatch 2002). Although Berkan and Trubatch did not define supply chain reconnaissance in their contribution, applying web intelligence gathering in order to improve the supplier selection process seems sensible. An example of a NLP use case that might translated into a LSCM use case is handwritten text recognition. Giménez et. al. studied handwritten digit recognition (Giménez, Andrés-Ferrer, Juan, and Serrano 2011). In a world that is still not fully digital, handwritten digit recognition might be useful when processing the order of a customer. Handwritten text recognition might also be useful during the *Inbound Logistics* process, or in the *Distribution* process.

Multiple participants suggested applying NLP in a general sense to manage communication during the negotiation process with potential suppliers. One participant noted that his employee applies rule based NLP in order to classify incoming orders containing natural language. Applying NLP as part of change management for the *Development* process was also mentioned by a participant.

Nevertheless, there might be additional NLP use cases that are translatable into the LSCM domain. Considering the natural language relevancy rankings, there have to be some not yet identified NLP use cases in the industry. Anecdotally, it seems that industry experts don't really think about natural language, while they are consistently confronted by natural language. Generally speaking, any process or activity during which natural language occurs in some form or another, might benefit from applying NLP and in almost any process or activity in the domain of LSCM natural language occurs. More research is necessary in order to identify NLP use cases.

5. FINAL REMARKS

Limitations related to the research design do apply in this case. The survey does not allow to draw conclusions solely relying on Bernoulli's Theorem as a result, the analysis of the results of the survey has only been qualitative. Additionally, the research framework describing LSCM might have been a limitation. It might be more successful to do a deep dive on one specific process with experts in that process instead of staying on a high level. Surveying industry experts about a – at least for some of them – new technology, is also a potential limitation. Anecdotally, it seems that industry experts don't really think about natural language, while they are consistently confronted by natural language. It might thus be hard to be confronted with an aspect of work life that has been taken for granted. More quantitative and qualitative research is still necessary in the domain. There are potentially many not yet identified NLP use cases within LSCM. The author plans to continue researching the subject in the coming years.

6. ACKNOWLEDGEMENTS

The authors would like to thank the Kühne Foundation for the financial support of this contribution. Additionally the author would like to thank the test users of his digital questionnaire.

7. BIBLIOGRAPHY

[Baumgarten 1999] Baumgarten, H. (1999). Prozeßkettenmanagement in der logistik. In H. Baumgarten and J. Weber (Eds.), *Handbuch Logistik Management von Material- und Warenflußprozessen*, pp. 226–238. Schäffer-Poeschel Verlag.

[Baumgarten and Inga-Lena 2000] Baumgarten, H. and D. Inga-Lena (2000). Prozesskettenmanagement als basis für logistik-management. In H. Baumgarten, H.-P. Wiendahl, and J. Zentis (Eds.), *Logistik-Management Strategien Konzepte Praxisbeispiele*, Chapter 2.3.1. Springer-Verlag.

[Berkan and Trubatch 2002] Berkan, R. C. and S. L. Trubatch (2002). Fuzzy logic and hybrid approaches to web intelligence gathering and information management. In *2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings (Cat. No.02CH37291)*, Volume 2, pp. 1033–1038 vol.2.

[Bird, Klein, and Loper 2009] Bird, S., E. Klein, and E. Loper (2009). *Natural Language Processing with Python*. Sebsdtopol, CA, USA: O'Reilly Media, Inc.

[Carstensen, Ebert, Endriss, Jekat, Klabunde, and Langer 2010] Carstensen, K.-U., C. Ebert, C. Endriss, S. Jekat, R. Klabunde, and H. Langer (2010). *Computerlinguistik und Sprachtechnologie – Eine Einführung* (Dritte Auflage). Heidelberg, Germany: Spektrum Akademischer Verlag.

[Chen, Wan, and Xu 2016] Chen, G., Y. Wan, and X. Xu (2016). An analysis of the sales and consumer preferences of e-cigarettes based on text mining of online reviews. In *2016 3rd International Conference on Systems and Informatics (ICSAI)*, pp. 1045–1049.

[Cunningham 1999] Cunningham, H. (1999). A definition and short history of language engineering. *Natural Language Engineering 5*(1), 1–16.

[Daugherty, Myers, and Autry 1999] Daugherty, P. J., M. B. Myers, and C. W. Autry (1999). Automatic replenishment programs: An empirical examination. *Journal of Business Logistics 20*(2), 63–82.

[Giménez, Andrés-Ferrer, Juan, and Serrano 2011] Giménez, A., J. Andrés-Ferrer, A. Juan, and N. Serrano (2011). Discriminative bernoulli mixture models for handwritten digit recognition. In *2011 International Conference on Document Analysis and Recognition*, pp. 558–562.

[Hellingrath and Lechtenberg 2018] Hellingrath, B. and S. Lechtenberg (2018). Applications of artificial intelligence in logistics. In K. Furmans and T. Wimmer (Eds.), *Understanding Future Logistics –Models, Applications, Insights*, Number 9 in International Scientific Symposium on Logistics, Bremen, Germany, pp. 9–24. Bundesvereinigung Logistik International.

[Jurafsky and Martin 2009] Jurafsky, D. and J. H. Martin (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (Second Edition). Upper Saddle River, NJ, USA: Prentice Hall.

[Saunders, Lewis, and Thornhill 2009] Saunders, M., P. Lewis, and A. Thornhill (2009). *Research methods for business students* (Fifth Edition). Prentice Hall.

[Straube 2004] Straube, F. (2004). *e-Logistik: Ganzheitliches Logistikmanagement* (First Edition). Berlin, Heidelberg, Germany: Springer-Verlag.