Article

Hit or Miss? Evaluating the Potential of a Research Niche: A Case Study in the Field of Virtual Quality Management

Albert Weckenmann ¹, Ștefan Bodi ^{2,*} , Sorin Popescu ² , Mihai Dragomir ² , Dan Hurgoiu ² and Radu Comes ²

¹ Friedrich-Alexander University Erlangen-Nürnberg; 4 Schloßplatz, 91054 Erlangen, Germany; aweckenmann@web.de

² Technical University of Cluj-Napoca, 103-105 Muncii Blvd., 400641 Cluj-Napoca, Romania; sorin.popescu@muri.utcluj.ro (S.P.); mihai.dragomir@muri.utcluj.ro (M.D.); dan.hurgoiu@muri.utcluj.ro (D.H.); radu.comes@muri.utcluj.ro (R.C.)

* Correspondence: stefan.bodi@muri.utcluj.ro; Tel.: +40-740-221-872

Received: 3 February 2019; Accepted: 4 March 2019; Published: 8 March 2019

Abstract: When knowledge is developed fast, as it is the case so often nowadays, one of the main difficulties in initiating new research in any field is to identify the domain's specific state-of-the-art and trends. In this context, to evaluate the potential of a research niche by assisting the literature review process and to add a new and modern large-scale and automated dimension to it, the paper proposes a methodology that uses “Latent Semantic Analysis” (LSA) for identifying trends, focused within the knowledge space created at the intersection of three sustainability-related methodologies/concepts: “virtual Quality Management” (vQM), “Industry 4.0”, and “Product Life-Cycle” (PLC). The LSA was applied to a significant number of scientific papers published around these concepts to generate ontology charts that describe the knowledge structure of each by the frequency, position, and causal relation of associated notions. These notions are combined for defining the common high-density knowledge zone from where new technological solutions are expected to emerge throughout the PLC. The authors propose the concept of the knowledge space, which is characterized through specific descriptors with their own evaluation scales, obtained by processing the emerging information as identified by a combination of classic and innovative techniques. The results are validated through an investigation that surveys a relevant number of general managers, specialists, and consultants in the field of quality in the automotive sector from Romania. This practical demonstration follows each step of the theoretical approach and yields results that prove the capability of the method to contribute to the understanding and elucidation of the scientific area to which it is applied. Once validated, the method could be transferred to fields with similar characteristics.

Keywords: latent semantic analysis; virtual quality management; concept investigation; concept disambiguation; knowledge discovery; sustainable methodologies

1. Introduction

In the paradigm of the knowledge-based society, knowledge becomes an indispensable resource in the further development of the contemporary social and economic status. Hence, in recent years, for supporting the required advancements of modern society, the speed through which knowledge is generated increased exponentially. Yang Lu in Reference [1] associates this phenomenon with the Fourth Industrial Revolution and the emergence of the “Internet-of-Things”, “Cyber-Physical Systems”, and “Enterprise Integration”; however, the increasing computational capacity and connectivity of computers could be mentioned as facilitators. Often, this high dynamic in research (knowledge

generation), especially in interdisciplinary fields, brings the need to express new specific contents, meanings or trends under a linguistic form. In this way, new syntagmas or concepts are brought to light and adopted by professionals as part of the specific jargon in the field. Even if their creators endowed them with a clear meaning at an incipient stage, when they become more popular in an emerging area, these concepts are quickly surrounded by a large amount of new knowledge that is developed with an amazing speed, enriching and enlarging their initial sphere.

The “virtual Quality Management” (vQM) concept could be a significant example for the circumstances described previously. It is born through a semantic operation, joining two established and mature concepts: “virtual” and “QM”, thus it is representative for an area which is in a period of high dynamic development and of interest for companies preoccupied with sustainability from the perspective of operations management and organizational culture.

In this context in which the amount of information relating to new concepts quickly reaches unmanageable levels, regardless of the field, solutions that can analyze extended documentation with the purpose of disambiguating information and capturing the essentials, thus creating knowledge, become the focus of attention and gain in importance. Traditional solutions for that purpose lay in the literature review process, trying to collect, select, filter, and structure the existing and relevant information in a synthetic way. However, this solution has limits: on one hand the large amounts of data can lead to a decrease in the ability of a research team to handle that information and on the other, the selection and interpretation of the chosen references is exposed to strong subjectivity.

Natural Language Processing (NLP), a field of computational linguistics, can be considered as a possible alternative for the above stated issues making users capable of going through large amounts of information fast, retaining only important data. It deals with translating human language to computers, enabling them to summarize and essentialize inputted information from an unlimited number of references. The computer programs deployed for this task use various algorithms, each with their own advantages and disadvantages. Among these worth mentioning are: Part-of-Speech Tagging; Named Entity Recognition; Semantic Role Labeling; and Latent Semantic Analysis (LSA).

The current paper proposes a framework for identifying trends, based on LSA by analyzing the knowledge space resulting from the intersection of two knowledge structures given by the vQM and Industry 4.0 concepts.

The investigatory process for these two concepts was done by applying LSA to a significant number of scientific papers published around them, selected after they underwent an initial raw filtration process proposed by the authors. As a result of this analysis, the generated ontology charts defined the knowledge structures of the two concepts by the frequency, position in text, and causal relation of associated notions. Combining these notions with one another resulted in senseful descriptors. The common ones were retained and further used for identifying new technological solutions and industrial applications, which are considered to emerge from this particular intersection of concepts. They were finally sorted and ranked within the PLC stages using the affinity diagram (KJ diagram) and the matrix diagram.

The intended added value of this approach is to supersede the classic literature review in the state-of-the-art analysis with automated information processing tools, both classic (affinity diagramming, conceptual mapping and analysis—based on ontology charts) and modern (Latent Semantic Analysis) for investigating complex knowledge spaces born at the intersection of new concepts. The illustration of this endeavor was done with the help of the vQM and Industry 4.0 concepts, and its end result is the current and emerging practical Industry 4.0 applications related to vQM throughout the entire PLC stages, particularly in the automotive industry.

2. Objective of the Research

The main objective of the research presented in this paper is to develop an innovative and expedient approach to evaluate the result potential and dissemination outlook of a research topic at the beginning of an investigation endeavor, especially in emerging scientific domains that are

characterized by a high degree of entropy and uncertainty. The presented methodology relies on the use of information processing tools, techniques, and methods (combining both classic and new) to address the most common concerns of research teams and help them arrive at a timely answer with an acceptable degree of validity to the issue concerning the relation between invested resources (time, staff, equipment, etc.) and foreseeable impact.

In this regard, the research approached in the current paper proposes, firstly, to analyze the knowledge density zone obtained by intersecting two knowledge structures, represented here by the vQM and Industry 4.0 axis, throughout the PLC stages (Figure 1).

Figure 1. Knowledge density zone between vQM and Industry 4.0 along the PLC stages.

Secondly, in addition to concept investigation and disambiguation, it aims to identify and hierarchize industrial applications (or new technological solutions) that are expected to emerge.

Finally, the paper intends to validate the proposed methodology through an online survey.

The need to develop this algorithm appeared to the authors due to their involvement in high dynamic manufacturing research demarches, where they were compelled to generate practical solutions for companies before the theories were settled. As such, the validation included in this paper was also based on vQM and Industry 4.0 concepts but can be transferred to any domain.

Contributing to the achievement of these endeavors the following stages were established:

- In the first stage an identification of relevant papers and scientific references was carried out for each of the two main concepts individually: vQM and Industry 4.0. After that, a filtration process was performed for limiting as much as possible the amount of irrelevant information. The ontology charts, which were both similar in nature as two knowledge structures, both for vQM and Industry 4.0, were determined by applying LSA (using a dedicated software program) on the final sets of bibliographic references.
- The second stage was reserved for weighting notions that were in relation with the main concepts. Based on the ontology charts and other information provided by the software each notion relating to vQM and Industry 4.0 was ranked by calculating their CN (correlation number). Next, associations were made between them with the scope of obtaining syntagmas (descriptors) that characterized the two main concepts. A score was also calculated for each of them signifying their importance to both concepts.
- The third stage began with extracting the common meaningful descriptors between the two main concepts. Based on these expressions, applications were identified using search engines.

This way the focus was kept only on those applications that were in close reference to vQM and Industry 4.0.

- The fourth stage, firstly, was focused on identifying the product life cycle stages and secondly, on the construction and completion of the matrix diagram. It had, as inputs on the left, the product life cycle stages together with identified applications and above the common descriptors between vQM and Industry 4.0 (associated by their weighted importance). It illustrates not only the hierarchy of the main categories of applications throughout the product life cycle stages, but also the rank of every sub-category application within the main categories.
- Within the fifth stage, an online questionnaire was applied for checking results obtained from the matrix diagram against an average rating coming from experts from Romania's automotive industry.

The research methodology is summarized in the following figure (Figure 2) showing also inputs and outputs for all its phases in the form of an extended flowchart:

Figure 2. Proposed methodology overview.

3. The Review of the Scientific Literature

Considering the objectives set forth by the paper (detailed previously), in this section the authors assembled the elements relating to the background and literature review into two categories: aspects relating to the need to assist the literature review process with modern instruments and techniques; and background and aspects which related to the main concepts forming the analyzed knowledge space.

3.1. The Need for New Approaches

If the statement that the development of the current society is strongly linked to that of knowledge is already a truism, what seemed fascinating to scientists was and is its speed of growth. The study of knowledge growth has enjoyed attention starting with Price's theory on the exponential growth of science [2], passing through Buckminster Fuller's knowledge doubling curve [3], and arriving at recent works such as Reference [4] which summarizes mathematical models for knowledge growth or References [5,6] which focus on modern digital knowledge spreading channels, to mention just some of the countless publications on this topic in these last years.

At the communication level, new knowledge is asking usually for new concepts to formalize its significance. The emergence of new concepts or concept groups generate in their semantic proximity domain-specific or interdisciplinary knowledge spaces. Any connected new research pass by an exploratory phase when a preliminary investigation of the mentioned spaces is necessary to identify the state-of-the-art, directions or the trending elements in its area. This exploratory investigation is often carried out with the help of literature reviews, which are complex, laborious, and time-consuming tasks, whose results are strongly dependent on the author's critical thinking ability, experience, and competence in the domain. The unprecedented increase in the amount of information that needs

to be processed to capture the relevant aspects of the knowledge space in question leads to either the reduction of the reference sample with the risk of losing important inferences from the analysis result, or the use as support of modern computer-assisted tools able to deliver a primary structure of the conceptual framework for the targeted knowledge spaces.

3.2. Literature Review Methods and Tools

A scientific literature review “has become the most common way of acquiring knowledge and oftentimes sets the direction for a study” [7]. Alongside the methods used and inputted constraints, the type of review can differ. The authors [8] identified a typology set of 14 main reviews for investigating and evaluating specialized literature, out of which worth mentioning are: “critical review”; “mapping review”; “rapid review”; “state-of-the-art review”; and “systematic search and review”.

From the tools and techniques’ perspective, two categories are identifiable: classic tools and those that use computer-assisted instruments. One might argue that, currently, literature reviews are carried out using tools from the latter category; however, some that have more of a classic background are still promoted: concept mapping (structuring the knowledge from a certain area based on main concepts in the form of highly-visual impact maps) [9,10]; affinity diagramming (structuring main ideas extracted from an initial literature search) [11]; snowball sampling (finding a network of professionals specific to a certain area by conducting interviews and asking them for other references) [12]; and literature tables used both for managing and ensuring the traceability of identified references, and for summarizing information and extracting main ideas.

The computer assisted tools are comprised of software programs, that help the user in making this review more efficient considering the following points:

- Researching and identifying relevant literature: Google Scholar; Distiller; arXiv; Citavi; Delve Health; LitAssist;
- Managing references: EndNote; Aigaion; Mendeley; Zotero; LiteRat; Bookends;
- Completing an automated text analysis and recognition (known as natural language processing (NLP)) for assisting the reviewer in classifying, tagging, parsing, semantic reasoning, and reducing the amount of information to a manageable size [13–15], but at the same time keeping its coherence, intelligibility, and meaning of the analyzed text: WordStat; QDA Miner; Carrot²; RapidMiner; Tropes.

3.3. Natural Language Processing Approaches

Natural Language Processing (NLP), a field of computational linguistics, deals with translating human language to computers enabling them to understand, summarize, and essentialize inputted information from an unlimited number of references. An analysis of the specialty literature categorized mainly used models and algorithms (incorporated into software solutions) into ones that complete a basic text processing (part-of-speech tagging; named entity recognition), ones that use statistical language models (bag-of-words model; n-gram models; neural network language model), and ones that discern the text based on the semantical meaning of words within (semantic role labeling; latent semantic analysis). Based on these, the software assists the user by increasing the quality of the review and the amount of information subjected to the analysis, while it reduces the time in which the user can process individually that same amount of text (literature).

Part-Of-Speech (POS) Tagging assigns parts of speech to words from a text body, allowing words having the same written form, but different meanings in different contexts to be understood accordingly by the software. Developed by Reference [16] and later advanced by other contributors [17–19], the method is enjoying recent applications regarding the analysis of written information on social media websites [20–22], dealing with grammar related aspects [23,24], identifying plagiarism [25], or even acting as a human–machine interface [26].

Named Entity Recognition enables the localization and categorization of “important and proper nouns in a text” [27], helping the reviewer to focus on important concepts that characterize the text. According to Reference [28], “this is an important task because its performance directly affects the quality of many succeeding NLP applications such as information extraction”. Its application recently gained popularity for processing semi-structured knowledge bases regarding entity disambiguation/mapping [29–31] and extracting/retrieving information [32] or for analyzing content generated on social media [33–35].

The Bag-of-Words Model identifies and represents by histograms the frequency of words appearing in the text [36] enabling the reviewer to conduct an initial text categorization. Although its application is grounded in NLP (e.g., opinion mining on information generated by social media users [37,38]), recent approaches prove its utility in other fields as well, such as image processing [39–41].

n-Gram Models help determine the probability of a sequence of words in a sentence or in a text. Their application varies from identifying patterns in text [42] to data extraction [43], automatic speech recognition, machine translation, and spell checking [44,45]. Neural Network Language models offer an improved version [46], both having the potential to be integrated into computer-assisted tools for supporting text reviewers.

Semantic Role Labeling traditionally uses a shallow parsing algorithm that identifies “arguments within the local context of a predicate” [47] enabling the software solutions, to a certain degree, to summarize information [48]. However, recent advancements, presented by References [49] and [50] outline the possibility for using this technique without syntactic parsing.

Latent Semantic Analysis (LSA) analyzes the relationship between terms contained in several documents and outputs a set of concepts that are semantically related to the terms and documents subjected to the analysis, and has multiple applications in the fields of information processing and knowledge discovery [51–53]. It uses Single Value Decomposition (SVD), a mathematical tool that allows the reduction of matrix rows, without seriously compromising the structure of columns. The matrix rows m (constructed out of the body of text) are the individual word types and the columns n represent the meaning of that word encoded into sentences and paragraphs. The cells that result contain information about the reoccurrence of a word in a paragraph [54]. After applying the SVD, and based on information obtained from it, an additional step is carried out, “sentence selection”, in which the “most characteristic parts of text” are selected to illustrate ideas that are essential in the body of text [55].

3.4. Virtual Quality Management (vQM)

The expansion of applications in the virtual domain opened new horizons also in the case of “Quality Management”, which evolved alongside the trends from modern industry. A new concept was thus born called “virtual Quality Management” (vQM).

The concept of vQM is based on, but not limited to, simulation studies, which are efficiently deployed for the sole purpose of “generating resilient knowledge and dimensioning quality techniques” [56] that can be applied either to products or processes, before they physically exist. By doing so, products and processes reach a certain level of maturity in the planning stage, so they can be introduced straight into production, having an increased level of performance compared to ones developed using conventional methods.

As identified from several references approaching the virtual area of quality management in manufacturing, simulation also plays an important role for vQM in developing models capable to foresee, adapt to, and optimize different scenarios by analyzing existing data [57], to increase efficiency and implement sustainable development [58] or even to reproduce measurement uncertainty due to temperature, humidity or other external influences [59]. Moreover, the approach proposed by Reference [60], regarding stochastic simulations for tolerance analysis “contributes to the improvement of virtual quality assessment”, which in this case, takes place already in the design stage of a product’s life cycle.

The combination between “simulation” and “virtual reality” allows operators to interact virtually with the fabrication process [61], which in turn enables virtual control and even virtual inspection of the manufacturing flow [62]. Immersive virtual reality applications can also be used efficiently in problem-solving activities, as demonstrated by Reference [63].

The nature of vQM is highly innovative as it is capable of providing the necessary information for deploying quality management techniques before the actual start of the manufacturing processes [59]. Quality and process parameters are obtained with the help of advanced modeling and simulation tools, all contributing to the process design of a manufacturing system. The architecture of vQM, presented in Figure 3, illustrates precisely the above stated particularities.

Figure 3. Architecture of vQM. Adapted from Reference [65].

Another important side of vQM is communication. The development of virtual channels (supported by the internet) brought several advantages in this field, such as improving the way we interact with each other or even with machines. They also increased the speed of information transfer, narrowed the gap between supplier and customer, and enabled product developers to provide their contributions remotely by forming “virtual teams” [64].

Although the definition of the concept is quite clear, the impact vQM has on current production engineering and quality management related concepts and how it interacts with them is yet to be found.

A possible research instrument that can exactly capture and contribute to the disambiguation of the vQM concept is LSA. This type of instrument is proposed because the clarification process should begin right at the semantic value of the expression, by analyzing it in the already used or possible contexts, put forth by the specialty literature of the latest years.

3.5. Industry 4.0

The forth industrial revolution (smart automation) is on the verge of unfolding and the driver of it will be the integration of physical objects in the information network. The main promoter of this leap in manufacturing industry is Germany, who introduced its own term “Industry 4.0” (or Industrie 4.0 in German) in 2011 at the Hannover Messe trade fair.

The year 2013 brought a clear definition of what the requirements are for the next industrial leap. In this sense, in Germany it was published the “Recommendations for Implementing the Strategic Initiative Industry 4.0” report, written by the German Communication Promoters Group of the Industry-Science Research Alliance, in collaboration with the National Academy of Science and Engineering and sponsored by the Federal Ministry of Education and Research. According to this report (also referred to as the Industry 4.0 National Working Group report), every physical object that is connected with the manufacturing process will be interlinked into a single network through the IoT. This network will incorporate everything starting from the factory floor to the delivery process, i.e., multiple systems that are outlaid as “Cyber-Physical Systems”, forming the so-called “Smart

Factory". This sort of system has the capability to be self-aware, thus it can actively intervene in the manufacturing processes to prevent potential faults [66]. Its self-awareness is given by the sensory data collected within the system. Actions that are precisely ordered are based on previously stored information, this way making the system not only self-aware, but self-maintained and self-learning. Due to all these technological advancements, Industry 4.0 is viewed as one of the fastest developing trends in production engineering.

4. Research Questions, Methods, and Results

4.1. Research Questions and Rationale

Within this context, the current research demarche begins with the following set of questions that have been developed starting from the main premise that has been stated, that under severe time constraints complex and resource intensive scientific investigations should only be initiated after a preliminary analysis of available information arrives at the conclusion that there will be sufficient return on investment in the end. The authors laid out in a concentrated form the possible formulation of these questions:

- What is it and how can an ontology chart that depicts the layout of a concept be obtained?
- How can one evaluate the relationships between the main studied concept and associated notions?
- How can a knowledge density zone be defined by intersecting two or more highly volatile (new and ambiguous) concept domains over time?
- Which are the main areas in which this type of information processing can be capitalized upon?

4.2. Creation of Ontology Charts Using Latent Semantic Analysis

4.2.1. Theoretical Aspects (Identifying and Filtering the Documentation)

The input selection in the case of the two main concepts was made in the same way. Mostly, the authors focused on gathering references with the help of search engines that have access to large databases of scientific research that offer the possibility to access full-text articles (Elsevier's Scopus, Elsevier's Science Direct and Google Scholar). However, if additional documents were found with the help of traditional search engines (like Google or Yahoo) they were not disregarded but retained only if they were considered to be relevant after the first filtration process. This type of documentation was manually converted to a format supported by the LSA software (typically .pdf or .txt).

As search engines work with keywords, the challenge was to determine those words that accurately represent the studied concepts. In the first instance, the search was limited to identifying references in relation with the name of the concept, then it was extended to include certain keywords obtained from the top ten most cited scientific papers. These were restricted to six at the most for avoiding the display of insignificant documentation. The timeframe was also preset to the last 7 years for limiting outdated information. It was considered that documentation outside this framework cannot accurately capture state-of-the-art advancements in the studied fields. After a brief revision (the focus was kept on the abstract of the paper, but not limited to it) of displayed papers, a total number of 35 references were considered to be sufficient for illustration purposes for each main concept. They were downloaded and saved in separate folders. The brief review was chosen instead of an extensive literature browsing due to two reasons: the latter one would not serve to the proposed objective, which is quick identification of trending elements and concepts from a specific field of interest (meaning it was much more time consuming); and secondly, there were no guarantees that by reading the paper entirely it could be used in the semantic analysis, so there was the chance that after the review it would be disregarded. We point out the fact that the choices made here represent one of the limitations of the proposed methodology, that for making these decisions relating to the scope of the research it must rely on the experience and expertise of the reviewers, the quantity and quality

of available literature, and on the relative positioning of the studied notions within the processed material when performing the first quantitative evaluation.

Making sure that every selected paper was indeed relevant for the analysis, the methodology included an additional filtration process performed by the review team. Each key word (mentioned above) was analyzed from its distribution perspective, throughout all the 35 papers. If the maximum number out of all keyword frequencies for each paper was not greater than five (viewed as a minimal threshold considering that all of the studied papers were more than five pages in length), the paper was considered to be irrelevant and it was removed from the analysis folder, where all were saved in the convened format. For this, the following condition had to be fulfilled:

$$\max_i(x_i) \leq 5 \quad (1)$$

where x represents the frequency of the i keyword ($i = \{1,8\}$)

For illustration purposes, Figure 4 contains the frequency of mentioned keywords in a certain paper:

Figure 4. Occurrence of keywords in one identified paper.

The above figure contains the occurrence number of each keyword for one of the 35 papers. As it can be seen, the keywords “management” and “organization” are the two least frequent ones. Although “management” appears only two times it is not sufficient to consider this paper irrelevant; for a certain scientific reference to be entirely removed from the analysis it must be scarce in all key words (their maximum must be lower than or equal to five). It can happen that a certain paper is scarce in one key word, but rich in others; in this case, the paper is taken into account, as it is the case of this example, again appealing to the reviewers’ competence, but to a reduced degree compared to the classic literature review. After this final filtration process a set of 30 papers remained, which were subject to the LSA.

The semantic analysis was carried out using a software program called Tropes v8.4 (English version) which is available free of charge and is capable of processing and perceiving, to a sufficiently advanced level, relationships between words, identifying or disregarding words specific to a domain of interest (based on input settings), counting the reoccurrence of words within different contexts, and retrieving neighboring concepts for certain notions or words.

4.2.2. Practical Implications (Defining the Ontology Charts)

As a direct result of the semantic analysis, an ontology chart of vQM was obtained and is presented in the following figure (Figure 5):

Figure 5. vQM ontology chart.

The vQM ontology chart (Figure 5) contains notions that are strongly related to quality management (measurement, control, quality, innovation, SPC). Some make references to instruments deployed through virtual means (such as simulation, modeling and optimization) and introduces “communication” and “collaboration” as notions linked with vQM (although having a low frequency; in sentences they were used closely together with vQM).

For “Industry 4.0”, there appeared considerably more related notions on the ontology chart (Figure 6) than in the case of vQM, which can be explained by the fact that although the two concepts are considered to be relatively new, “Industry 4.0” is more developed and established in the scientific community as it is promoted to be the next generation of production engineering. Also, the known information on this subject reached a more mature level due to a higher number of publications.

Figure 6. “Industry 4.0” ontology chart

The related notions in this chart include the main pillars of Industry 4.0, such as IoT (Internet of Things), cyber (making reference to “Cyber-Physical Systems”), “Big Data”, and “Cloud Computing”; some design principles of Industry 4.0 as stated in Reference [67] (interoperability, decentralization, virtualization), and some notions which are also common to vQM.

The correlation degree of each current notion (denominated k) with regards to the main concept is reflected in its position on the chart. The position can be expressed as a function of the causal relation (the x -axis) and the concentration of relations—or frequency of association (the y -axis) [68]. Thus, the graph supports a visual comparison of the weight of relations between all concepts present on the chart with the main one.

On the x -axis, the number of words separating one concept from another in sentences can be considered as a direct reference for establishing their causal relation. Taking this into account, for quantifying the causal relation for both the “actant” (predecessor, enabler or source) and “acted” (successor, consequence or attribute) concepts, a scale from 0 to 1 was attributed from the center to both ends of the chart. As a certain k notion is further apart (closer to the value 1) on the chart from the main concept, it means that in sentences a significant number of words are placed between them. This value, called the “Causal Relationship Value” (CRV) takes values from 0 to 1:

$$CRV_k = [0 \dots 1] \quad (2)$$

On the y -axis, a fraction is represented between the total number of relations (or association) that each k concept has with the main one and the number of different relations. This value is called the “Concentration Value” or CV. The chart shows the minimum (bottom) and maximum (top) values, as they were calculated by the software program used to complete the semantic analysis. The formula for expressing the CV is:

$$CV_k = \frac{A}{A - B_k} \quad (3)$$

where A is the total number of relations and B_k is the number of relations that each concept has with the main one.

4.3. Weighting Related Notions Based on Ontology Charts

4.3.1. Theoretical Aspects

As a consequence of attributing the two scales for the “actant” and “acted” concepts their strength of the causal relation with the main concept is inversely proportional.

The CV and the inverse of CRV having completely different calculation rules but comparable in terms of magnitude, a correlation number (CN), obtained by the product of the two, was considered to be significant for expressing the importance of each concept in relation with the main one. The formula used for determining the CN is presented as follows:

$$CN_k = CV_k \times \frac{1}{CRV_k} \quad (4)$$

4.3.2. Practical Implications

The CN of each notion from the ontology charts are centralized in Tables 1 and 2, starting from the highest and descending to the lowest:

Table 1. Correlation number of “actant” and “acted” notions regarding vQM.

Actant	Score	Acted	Score
Virtual	37	Model/Modeling	24
Simulation	20	Procedure (referring to process)	20
Information	9	Control	17.4
Quality	8.3	Optimization	12.4
Innovation	5.8	Communication	12.1
Statistical Process Control (SPC)	4.4	Knowledge	9.4
Knowledge management	3	Measurement	5.8
		Collaboration	5.3
		Product	4.8
		Development	3.6

Table 2. Correlation number of “actant” and “acted” notions regarding “Industry 4.0”.

Actant	Score	Acted	Score
IoT (Internet of Things)	31.4	Cloud_computing	17.9
Smart	27.6	Big_data	17.8
Virtualization	19.5	Optimization	17.0
Cyber	13.6	Measurement	12.7
Interoperability	11.5	Procedure	8.3
Simulation	10.0	Model	7.6
Digital	8.2	Collaborative	6.4
Information	5.3	Product	5.8
Innovation	5.0	Development	4.6
Decentralization	4.6	Control	4.0
Quality	3.6		

In the next step, combinations between the “actant” and “acted” notions were made for assuring increased homogeneity between all the terms of each concept. This action was taken for bringing the specific notions of each concept more closely together with the scope of finding common grounds between them.

A cumulative score can be calculated for each resulting association by multiplying the individual correlation values of the two forming notions, thus focusing on their conceptual intersections. This score can be used to determine the hierarchy between the obtained associations. However, only the meaningful associations were retained (thus becoming descriptors for each concept) and ordered in a descending manner in Tables 3 and 4:

Table 3. Cumulative scores of descriptors regarding vQM.

Descriptor	Cumulative Score	Descriptor	Cumulative Score
Virtual model/modeling	888.0	Innovation model	139.2
Virtual control	634.8	Virtual development	133.2
Simulation and modeling	480.0	Process innovation	116.0
Virtual optimization	458.8	Measurement through sim.	116.0
Virtual communication	447.7	Optimization information	111.6
Process simulation	400.0	Quality optimization	102.9
Virtual knowledge	347.8	Quality communication	100.4
Virtual measurement	214.0	Quality knowledge	78.0
Virtual collaboration	196.1	Innovation knowledge	54.5
Simulation knowledge	188.0	Measurement information	52.2
Information control	156.6	Product innovation	27.8
Quality control	144.4	Innovative development	20.9

Table 4. Cumulative scores of descriptors regarding “Industry 4.0”.

Descriptor	Cumulative Score	Descriptor	Cumulative Score
IoT cloud	560.1	Cyber collaboration	87.0
Big_data from IoT	559.1	Process simulation	82.5
Smart cloud (computing)	494.0	Decentralized cloud	82.2
Smart measurement	352.1	Virtual(ized) control	78.4
Big_data virtualization	347.1	Decentralized optimization	78.3
Virtual(ized) optimization	331.1	Simulation & modeling	76.7
Virtual(ized) measurement	247.8	Measurement information	66.8
Cloud interoperability	205.0	Cyber development	62.3
Collaboration through IoT	199.6	Quality optimization	60.7
Interoperability optimization	195.2	Product simulation	58.2
Simulations on big_data	178.3	Cyber control	55.0
Smart collaboration	176.1	Interoperable develop.	52.3

Table 4. Cont.

Descriptor	Cumulative Score	Descriptor	Cumulative Score
Virtual(ized) modeling	149.3	Digital collaboration	51.8
Digital (big_)data	145.2	Development through sim.	45.6
Development through IoT	143.0	Process innovation	40.5
Digital optimization	138.4	Innovation model	37.6
Measurement through sim.	127.3	Digital development	37.1
Control through IoT	126.3	Digital control	32.8
Smart development	126.1	Collaborative innovation	31.2
Virtual(ized) collaboration	123.9	Decentralized collaboration	29.3
Smart control	111.4	Product innovation	28.5
Cyber model	104.8	Quality model	27.4
Digital measurement	103.6	Innovative development	22.4
Process interoperability	94.7	Information control	21.1
Cloud information	93.8	Decentralized develop.	21.0
Optimization information	89.3	Decentralized control	18.5
Virtual(ized) develop.	88.8	Quality control	14.4

In the case of the two concepts, the expressions obtained by pair-wise associations were filtered for redundancies and were verified if they have real-life applicability with the help of search engines. In some cases, connection words were added to further refine the expressions and to assure their coherence related to the main concept.

4.4. Features of the vQM x Industry 4.0 Knowledge Space

4.4.1. Theoretical Aspects

Based on the above steps, we can now characterize the previously ambiguous knowledge space, found at the intersection of the two notions, that are connected through a common goal, namely, sustainability. Overlapping the tables containing the descriptors of vQM and Industry 4.0 (Tables 3 and 4) it can be observed that there are common ones between them. These are used for determining common ground between two apposed concepts. Ranking them was done by calculating their weighted average (thus, still respecting their importance to each main concept) and consequently their weighted importance obtained in percentages.

4.4.2. Practical Implications

As seen in the table below, it can be the case that there is a considerable difference between the total scores of vQM and Industry 4.0. The weighted average provides an aggregate that accurately reflects the importance of each descriptor to both main concepts, regardless of their score difference. Although each common notion has its own LSA score depending on the main concept (vQM or Industry 4.0), the ranking presented in the above table was made according to the calculated weighted average of the score pairs.

These theoretical results can be further expanded for the use of practitioners by mapping onto the knowledge space the adequate concrete implementations that were found in industry. The proposed solution for identifying industrial applications uses the authors' expertise and is similar to identifying relevant papers for the LSA. Firstly, the top five most important common descriptors (obtained from the LSA—see Table 5) were entered into search engines followed by keywords extracted from the top 10 most cited "Industry 4.0" references together with the "application" keyword that was added to limit the amount of retrieved information. This way it is assured that the search was focused on practical applications and were related to "Industry 4.0" as its representative keywords were used. As "Industry 4.0" is considered to be the new trend in production engineering, applications related to it are also categorized as emergent. The retained applications were focused on those that had connections with—or can be used in relationship to vQM. A list of applications is obtained, but they are not sorted in any way.

Table 5. Common notions between vQM and Industry 4.0.

Common Descriptors	LSA Score		Weighted Average	Weighted Importance %
	vQM	Industry 4.0		
Virtual modeling	888.00	149.30	271.87	40.45%
Virtual optimization	458.80	331.10	78.95	11.75%
Simulation and modeling	480.00	76.70	152.74	22.73%
Process simulation	400.00	82.50	56.93	8.47%
Virtual measurement	214.00	247.80	62.27	9.27%
Quality optimization	102.92	60.70	7.35	1.09%
Measurement through simulation	116.00	127.30	11.83	1.76%
Virtual development	133.20	88.80	5.61	0.84%
Process innovation	116.00	40.50	16.88	2.51%
Innovation model	139.20	37.60	6.27	0.93%
Product innovation	27.84	28.50	0.88	0.13%
Innovative development	20.88	22.40	0.53	0.08%
TOTAL	3096.84	1293.2	672.10	100%

4.5. Grouping and Ranking Applications Along Time Stages (e.g., the PLC Stages)

4.5.1. Theoretical Aspects

Focusing on the theoretical connections uncovered before, this last step of the research methodology deploys the new structure of the two-fold common knowledge space into workable results that can be applied in companies. In other words, by using the proper applications in the proper manner, based on the guidelines below, firms can employ vQM tools to rip the competitive benefits of Industry 4.0. Taking into account the development (the third) dimension, to the two axes (vQM and Industry 4.0) another axis is added, the PLC, along which identified applications are grouped into its stages (see Figure 1 for an overview).

There are many views when it comes to defining precisely the product life cycle stages. For this reason, the authors synthesized information from an internationally accepted standard which inclusively covers all stages of a product life cycle: Reference [69]. The stages were grouped into main stages and sub-stages accordingly. These are presented in Table 6:

Table 6. Product life cycle stages, according to ISO/IEC/IEEE 15288:2015.

Main Stage	Sub-Stage
I. Planning, design, and development	Stakeholder requirement definition/analysis Product design and validation
II. Realization	Procurement Manufacturing Measurement and control
III. Utilization	Delivery Customer assistance Maintenance
IV. Disposal	Retirement Recycling

4.5.2. Practical Implications

The affinity diagram (also known as a KJ diagram) was applied for grouping the set of applications into the corresponding life cycle stages, defining both categories and sub-categories of applications. The matrix diagram (presented in Table A1 in the Appendix A Section) contains both these and the common notions with their weighted importance (expressed in percentages).

The cells resulted at the intersection between main or sub-category applications and the common notions were filled out according to the correlation between them, with a number on a scale of 0 to 3 (0 means that there is no relation between the two and 3 signifies a strong relation). The sum of products between these correlation numbers and the weighted importance of the common notions was calculated on each line, thus obtaining an aggregate number that can be viewed as the importance of each application. If a main category application contains more sub-category items, its score is provided by the average of them (indicated in bold in Table A1).

Having the importance scores for every sub- or main-category application, they can be hierarchized for all the stages of the product life cycle (in Table A1 the applications are already ranked). Thus, the matrix diagram contains a ranking of identified vQM applications in the context of Industry 4.0, that are or will become essential in each stage of the product life cycle. Where blank squares exist corresponding to a life cycle stage it means that there were no suitable applications identified that fit the previously mentioned search conditions (see Section 4.3). All identified applications grouped into the PLC stages are presented in Figure 7:

Figure 7. Industrial applications along the PLC stages.

1—Social networking data analysis	14—Automated process optimization	25—Real-time inventory monitoring
2—Google analytics for customer targeting	15—Automated process robustness prediction	26—Automated raw material order and purchase
3—Google Trends analysis used for product benchmarking	16—Preventive actions taken based on collected data	27—Automated raw material approval
4—Plant layout design used for production flow optimization	17—Automated real-time process optimization	28—Smart supplier evaluation and selection
5—Manufacturability analysis	18—Mobile device applications used for process monitoring and control	29—Virtual CMM used for evaluation of uncertainty estimates
6—Work instructions design	19—Smart glasses employing AR used for displaying product/process data	30—Inline metrology used for fault, defect and deviation detection
7—Worker ergonomics planning and design	20—Remote process monitoring and control	31—Absolute multiline technology for online precision monitoring
8—Work time study analysis	21—Edge computing	32—Knowledge base used for problem solving for known issues
9—CAD used for product modeling	22—Remote equipment control and assistance	33—Remote access used for diagnostics and trouble ticket resolution
10—Kinematics simulation, analysis and optimization	23—Vision system used for part inspection	34—FAQ web page used for fault/defect identification and self-help
11—Assembly and part finite element structural analysis	24—Smart glasses employing AR used for operator empowerment	35—Automated online assistant used for tech support through live chat
12—Simulation software used for process planning and flow analysis		
13—Virtual commissioning used for equipment automation		

The above representation combines PLC stages with the identified significant notions and their potential applicability degree to provide an overview for potential academic and industrial users. The circles represent the main categories of applications and their size is proportional with their importance as calculated in the matrix diagram (Table A1). The squares correspond to actual applications and their size also represents their priority in reference to the category they belong to. The denomination of each application is presented below the figure.

5. Research Implications

The main implication of the current research is uncovered by carrying out a validation stage of the proposed LSA-based methodology. This was done through an online questionnaire, applied to groups of professionals with relevant knowledge connected to the actual technological developments from the automotive industry, which is considered in Romania to have the highest degree of quality management and industrial practices proficiency and is also subjected to the most intense competition on a European scale.

The validation demarche raised the following challenges:

- covering the important companies from the automotive sector such that even if statistical relevance is not reached, this will be reflected by including within the survey all/most of the major actors from this field;
- selecting the target group such that it has access to relevant information;
- designing and constructing the questionnaire.

Statistically speaking, there are 510 automotive companies (NACE code 29) in Romania (according to the online database of Reference [70]), but in terms of our analysis not all of them present the same relevance. An analysis of the technological level of these companies clearly leads to the conclusion of an inhomogeneity of their relevance in relation to our analysis. One can expect not to obtain similar information on new technological developments regarding Industry 4.0/vQM from Renault or Ford (big players, global auto makers) and from a small firm that supplies low complexity parts. For this reason, the focus of this study was aimed at those few major manufacturers from this industry branch and questioning consultants, specialists, quality auditors or managers that draw their expertise from not just one, but several important companies and can provide answers with a high degree of relevance. This way, the answers of one individual can represent feedback from more than one company.

Considering the above-mentioned aspects, the companies were grouped into three main categories: the first two categories (A. automotive manufacturers and B. manufacturers of major components—e.g., Dacia–Renault, Ford, Bosch, Continental, Leoni, Rombat, Daimler, Pirelli, RAAL) represent manufacturers with international notoriety; all other companies were left for a third (C) category. The first category had 10 representatives, the second had around 150, and in the third category 350 companies were included.

The investigation was guided towards the first two categories, but there was also a concern about obtaining some samples as well from the third category companies, mainly from those which were preoccupied with new technologies. The reason behind this approach is that SME-type companies (from the C category) are mostly focused on concrete measures of physical QM due to resource constraints, lack of maturity, and limited openness towards adopting new technologies and due to the need to obtain quick and measurable results, with a noticeable impact on their process performance and bottom line.

The surveyed target group consisted of consultants, specialists, quality auditors or even managers that drew their expertise from not just one, but several important companies and could provide answers with a high degree of relevance. An important issue addressed was information confidentiality, as the research team made sure that the expertise and feedback of the focus group was captured but that company specific information was filtered out.

Within the survey, the abovementioned groups were provided with a list of 41 applications determined through the LSA (presented in the Matrix Diagram—Tables A1 and A2) and they had to rate them considering their implementation or future implementation into the companies whose manufacturing processes they were familiar with. Responders also had the possibility to name and rate additional applications that were not included, but were considered to be important by them (centralized in Table A3).

The scale for rating each application was kept the same as in the case of the matrix diagram, from 0 to 3, thus enabling an objective and direct comparison between the score calculated from the matrix diagram and the average score obtained from the survey (0—“it’s not likely to be used by manufacturing companies”; 1—“only some manufacturing companies have or will implement it”; 2—“most of the manufacturing companies have or will implement it”; 3—“all manufacturing companies have or will implement it”).

Table 7 includes the number of responses, the number of companies covered from each type and the percentage of coverage:

Table 7. Number of responses and percentage of coverage.

Target Group	Responses	Type A Companies Covered	Type B Companies Covered	Type C Companies Covered
Top management	6	0	10	9
Technical consultants	8	7	55	28
Specialists	5	2	35	11
Quality auditors	6	1	25	15
TOTAL	25	10 out of 10 (100% coverage)	125 out of 150 (83.3% coverage)	63 out of 350 (18% coverage)

As it can be seen, not all of the 510 companies were covered; however, all of the type A and most of the type B companies were reached within the study.

For analyzing the responses, the chosen reference values were the ones obtained from the questionnaire because they are considered to be the “voice” of the industry. A margin of error was calculated for each score obtained from the matrix diagram and findings showed that it was not more than $\pm 8\%$ (representing the deviation from the reference). The average margin of error was even better, being just under 4% (representing the mean of the deviation modulus values).

The list of applications alongside their rating and margins of error can be consulted in Table A2. It presents not only the accuracy of the research methodology approached within this paper, but also prioritizes trending applications that companies will need for their own development, thus defining the emerging research directions.

6. Conclusions

The current paper focused on analyzing the common knowledge space in specific areas by intersecting the knowledge structures obtained through a concept disambiguation process. The results of this analysis were further used to determine the technological solutions that are expected to emerge from that concept intersection space. Particularly, it was analyzed the possibility of identifying emergent industrial applications by intersecting the knowledge structures of two concepts: Industry 4.0 and vQM.

The investigation and disambiguation process were achieved with the help of a fairly new knowledge extraction method called latent semantic analysis (applied by a specialized software program), which is capable of analyzing large amounts of text and quickly identify the relationship, causality, and frequency of words, thus providing an interpretation of the state-of-the-art literature (in close connection with the two main concepts—vQM and Industry 4.0). By applying this technique,

the common areas of interest (common knowledge space) were identified in the form of descriptors that were further used for identifying applications that are defined by or relate to them.

The set of applications was grouped into the stages of the product life cycle according to the ISO/IEC/IEEE 15288:2015 standard (with the help of a classic sorting tool—Affinity Diagram) and ranked considering their importance scores resulting from applying a matrix diagram. These calculated scores reflect the weighted importance of common descriptors between vQM and Industry 4.0, as they also served as inputs into the matrix diagram.

The research loop was closed with a survey that focused on obtaining information from experts (from the field of quality) with close ties to the automotive industry, that confirms the reliability of the presented methodology with an average margin of error of just under 4%.

As illustrated through the course of this paper, one can note the important advantages of this LSA-based methodology:

- Emerging technological trends can be identified from the common knowledge space of two or more concepts;
- The state-of-the-art knowledge discovery process (concept disambiguation) is based upon an exponentially greater number of scientific references and it is reduced to a few seconds;
- The outputs of the LSA can be further analyzed, sorted, and grouped by deploying various tools and techniques, thus obtaining more useful information.

Considering the abovementioned key points of the methodology, it can be stated that the resulting applications are the product of current research concerns in the two analyzed fields, vQM and Industry 4.0. The constructed matrix diagram also provides a useful preview of industrial applications throughout the product life cycle and ranks them to depict those that are more critical to quality based on their importance scores.

The current framework has a high degree of portability and it can be applied to other concept associations as well for determining common areas or emergent aspects in a specific context.

In a broader understanding, the model can make forecasts for applications that are more likely to become popular in the next period of time. Organizations that are capable to foresee them and focus on their development and implementation have a clear competitive advantage towards achieving sustainable growth and advancement. The applicability of the methodology and the results described in this paper are two-fold. From an academic point of view, they can provide the clarification of the concepts for the use in training and instruction as well as an approach to investigate the two domains in the future to gauge their development. On the other hand, industrial users could benefit from the identified ranking of potential applications to simplify and make more efficient their implementation projects, allowing them to keep the focus on the bottom line and on the concrete results that are needed for market success.

Author Contributions: Conceptualization, Ș.B., A.W., and S.P.; methodology, Ș.B., A.W., and S.P.; literature survey, Ș.B. and R.C.; investigation, Ș.B., D.H., R.C. and M.D.; validation, Ș.B., D.H., M.D.; visualization, R.C. and D.H.; writing—original draft, Ș.B., A.W., S.P. and M.D.; writing—review and editing, Ș.B., M.D. and R.C.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank Dan Marius, the C.E.O. of S.C. Calitop S.R.L. and Călin Neamțu from the Technical University of Cluj-Napoca for their valuable expertise, discussions, and for facilitating connections with the industry.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Matrix diagram for ranking the technological solutions.

				Virtual Modeling	Simulation & Modeling	Virtual Optimization	Process Simulation	Virtual Measurement	Innovation Model	Virtual Development	Process Innovation	Measurement through Simulation	Optimization Quality	Product Innovation	Innovative Development	Sub-Category Score	Category Score
Life cycle stage	Sub-stage	Category	Sub-category	0.40	0.12	0.23	0.08	0.09	0.01	0.02	0.01	0.03	0.01	0.001	0.001		
Planning, design, and development	Stakeholder requirement definition/analysis	E-communication (e-mail, e-conference, IM, file transfers)	-	3	3	3	3	2	1	3	1	2	1	2	2	2.823	2.823
		Online surveys and questionnaires	-	3	3	3	0	1	1	3	1	0	2	0	1	2.432	2.432
		Customer Relationship Management (CRM) systems	-	3	3	1	0	0	0	3	0	0	1	0	1	1.856	1.856
		Data mining	Social networking data analysis	2	2	2	0	0	0	3	0	0	0	0	1	1.552	1.419
	Google analytics for customer targeting		2	1	2	0	0	1	3	0	0	0	3	3	1.451		
	Google Trends analysis used for product benchmarking		2	1	1	0	0	1	3	1	0	3	0	1	1.255		
	Virtual prototyping	CAD used for product modeling	3	3	2	1	2	2	3	1	1	3	3	2	2.432	2.313	
		Kinematics simulation, analysis and optimization through software.	3	3	3	0	1	0	3	0	1	1	1	1	2.430		
		Assembly and part finite element structural analysis	3	3	1	0	2	0	3	0	1	2	0	1	2.076		
	Product design and validation	Virtual process modeling and simulation	Simulation software used for process planning and flow analysis	3	3	2	3	1	1	3	1	3	2	1	2	2.536	2.059
			Virtual commissioning used for equipment automation	2	2	3	1	1	0	3	2	1	2	0	2	2.018	
			Automated process optimization	1	3	3	3	0	0	3	2	2	3	0	1	1.841	
			Automated process robustness prediction	1	3	3	3	0	0	3	2	2	3	0	1	1.841	
	Virtual manufacturing		Plant layout design used for production flow optimization	3	3	3	2	1	1	2	1	2	3	1	1	2.645	2.417
			Manufacturability analysis	3	3	2	3	1	1	3	2	3	3	2	3	2.556	
			Work instructions design	2	3	3	3	2	0	3	2	2	3	0	1	2.431	
			Worker ergonomics planning and design	2	3	3	3	1	0	3	1	2	0	0	0	2.301	
			Work time study analysis	1	3	3	3	3	1	3	1	3	3	1	3	2.150	

Table A1. Cont.

				Virtual Modeling	Simulation & Modeling	Virtual Optimization	Process Simulation	Virtual Measurement	Innovation Model	Virtual Development	Process Innovation	Measurement through Simulation	Optimization	Quality	Product Innovation	Innovative Development	Sub-Category Score	Category Score
Life cycle stage	Sub-stage	Category	Sub-category	0.40	0.12	0.23	0.08	0.09	0.01	0.02	0.01	0.03	0.01	0.001	0.001			
Realization	Procurement	Smart supply chain	Real-time inventory monitoring	1	2	3	3	2	1	0	3	1	3	0	3	1.852	1.465	
			Automated raw material order and purchase	1	1	3	0	1	0	1	3	2	3	0	3	1.420		
			Automated raw material approval	1	1	3	0	1	0	0	0	1	2	0	1	1.341		
			Smart supplier evaluation and selection	1	1	2	1	1	1	2	2	0	3	0	1	1.245		
	Manufacturing	Big data analysis	Preventive actions taken based on collected data	3	2	3	3	3	0	1	3	1	3	0	0	2.758	2.366	
			Automated real-time process optimization	1	2	3	3	3	0	1	3	2	3	0	0	1.974		
		Rapid prototyping (Additive manufacturing)	3D printing used for product manufacturing	3	3	2	0	2	0	3	0	0	1	3	0	2.272	2.272	
		Smart and preventive maintenance	-	2	1	3	2	2	1	0	2	2	3	0	1	2.070	2.070	
		Digital assistance	Remote equipment control and assistance	2	1	2	1	2	1	1	2	0	3	0	1	1.725	1.509	
			Vision system used for part inspection	1	2	2	1	2	0	0	3	0	3	0	0	1.417		
			Smart glasses employing AR for operator empowerment (displaying instructions, assigning tasks and facilitating information access)	2	1	1	0	2	0	1	1	0	2	0	0	1.384		
		Measurement and control	Virtual measuring	Virtual CMMs used for evaluation of uncertainty estimates	2	2	3	1	3	1	2	1	3	2	0	2	2.238	1.907
				Inline metrology used for fault, defect and deviation detection.	2	1	3	1	3	0	0	3	3	3	0	1	2.100	
				Absolute multiline techn. for online equipment precision monitoring	1	1	2	0	3	0	1	2	3	2	0	0	1.383	
			Digital support	Mobile device applications used for process monitoring & control	2	1	1	3	3	0	0	1	1	1	0	0	1.729	1.690
				Smart glasses employing AR used for displaying product/process data	2	2	1	1	3	0	0	1	0	1	0	0	1.652	
	Cloud computing		Remote process monitoring and control	1	2	2	0	2	0	1	0	1	1	0	0	1.331	1.315	
			Edge computing	1	1	2	1	2	0	0	1	1	2	0	0	1.299		

Table A1. Cont.

				Virtual Modeling	Simulation & Modeling	Virtual Optimization	Process Simulation	Virtual Measurement	Innovation Model	Virtual Development	Process Innovation	Measurement through Simulation	Optimization Quality	Product Innovation	Innovative Development	Sub-Category Score	Category Score
Life cycle stage	Sub-stage	Category	Sub-category	0.40	0.12	0.23	0.08	0.09	0.01	0.02	0.01	0.03	0.01	0.001	0.001		
	Delivery	-	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Utilization	Customer assistance	Internet based customer communication support	Knowledge base used for problem solving for known issues	2	2	3	1	0	3	3	3	0	3	2	1	1.952	1.250
			Remote access used for diagnostics and trouble ticket resolution	1	1	2	1	2	1	1	3	0	3	0	0	1.328	
			FAQ page used for fault/defect identification and self-help	0	0	3	0	1	2	0	3	1	3	1	1	0.877	
			Automated online assistant used for tech support through live chat	0	0	3	0	1	2	0	2	0	3	0	0	0.841	
	Maintenance	-	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Retirement	-	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Disposal	Recycling	Recycling assistance platform providing info for customers for recycling	-	1	1	3	0	0	2	1	3	0	3	1	1	1.298	1.298

Table A2. List of rated applications.

Applications	Average of Responses	Calculated from LSA	Error (max. 8%)
Absolute multiline technology used for online precision monitoring	1.33	1.38	3.57%
Assembly and part finite element structural analysis	2.13	2.08	−2.70%
Automated online assistant used for tech support through live chat	0.87	0.84	−2.96%
Automated process optimization	2.00	1.84	−7.94%
Automated process robustness prediction	2.00	1.84	−7.94%
Automated raw material approval	1.33	1.34	0.57%
Automated raw material order and purchase	1.53	1.42	−7.42%
Automated real-time process optimization	1.87	1.97	5.44%
CAD used for product modeling	2.47	2.43	−1.40%
CRM systems used for requirement analysis and sorting	1.80	1.86	3.02%
E-communication: e-mail, e-conference, instant messaging, file transfers	2.87	2.82	−1.53%
Edge computing	1.33	1.30	−2.60%
FAQ page used for fault/defect self-identification and self-help	0.93	0.88	−6.08%
Google analytics for consumer targeting	1.47	1.45	−1.07%
Google trends analysis for product benchmarking	1.33	1.25	−5.91%
Inline metrology used for fault, defect and deviation detection	2.13	2.10	−1.56%
Kinematics simulation, analysis and optimization through software	2.27	2.43	6.71%
Knowledge base used for problem solving for known issues	2.07	1.95	−5.53%
Manufacturability analysis	2.53	2.56	0.88%
Mobile devices used for process monitoring and control	1.87	1.73	−7.40%
Smart glasses employing AR used for displaying product/process data	1.60	1.65	3.12%
Online surveys and questionnaires	2.47	2.43	−1.41%
Plant layout design used for production flow optimization	2.53	2.64	4.20%
Preventive actions taken based on collected data	2.67	2.76	3.31%
Rapid prototyping (3D printing)	2.27	2.27	0.22%
Real-time inventory monitoring	1.93	1.85	−4.20%
Recycling assistance platform providing info for customers for recycling	1.40	1.30	−7.26%
Remote access used for diagnostics and trouble ticket resolution	1.40	1.33	−5.14%
Remote equipment control and assistance	1.87	1.73	−7.59%
Remote process monitoring and control	1.33	1.33	−0.15%
Simulation software used for process planning and flow analysis	2.47	2.54	2.74%
Smart and preventive maintenance for equipment	2.20	2.07	−5.93%
Smart glasses used for operator empowerment	1.47	1.38	−5.66%
Smart supplier evaluation and selection	1.33	1.25	−6.59%
Social networking data analysis for requirement and stakeholder identification	1.47	1.55	5.50%
Virtual commissioning used for reliable equipment automation	2.00	2.02	0.89%
Virtual coordinate measuring machines	2.20	2.24	1.72%
Vision system used for part inspection	1.53	1.42	−7.59%
Work instruction design	2.40	2.43	1.28%
Work time study analysis	2.27	2.15	−5.16%
Worker ergonomics planning and design	2.27	2.30	1.51%
Average error			3.98%

Table A3. List of applications nominated by survey responders.

Applications	Average of Responses
Risk analysis applications	2.5
SAP Smart Data Access	2
Materials Management Operations Guideline/Logistics	1.5
Evaluation—MMOG/LE	1
Design for Lean Six Sigma (DFLSS) used for product development	1
Eight Disciplines (8D) problems solving	2

References

1. Lu, Y. Industry 4.0: A survey on technologies, applications and open research issues. *J. Ind. Inf. Integr.* **2017**, *6*, 1–10. [\[CrossRef\]](#)
2. De Solla Price, D.J. The exponential curve of science. *Discovery* **1956**, *17*, 240–243.
3. Fuller, R.B. *Critical Path*, 1st ed.; St Martins Press: New York, NY, USA, 1981.
4. Vitanov, N.K. *Science Dynamics and Research Production—Indicators, Indexes, Statistical Laws and Mathematical Models*; Springer: Basel, Switzerland, 2016.
5. Schmitt, U. Personal Knowledge Management Devices: The next Co-evolutionary Driver of Human Development. In Proceedings of the International Conference on Education and Social Sciences Abstracts & Proceedings (INTCESS14), Istanbul, Turkey, 3–5 February 2014; pp. 1081–1091. [\[CrossRef\]](#)
6. Petzold, T. *Global Knowledge Dynamics and Social Technology*; Palgrave Macmillan: Basingstoke, UK, 2017.
7. Onwuegbuzie, A.J.; Weinbaum, R. A framework for using qualitative comparative analysis for the review of the literature. *Qual. Rep.* **2017**, *22*, 359–372.
8. Grant, M.J.; Booth, A. A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Inf. Libr. J.* **2009**, *26*, 91–108. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Rosas, R.S.; Ridings, J.W. The use of concept mapping in measurement development and evaluation: Application and future directions. *Eval. Program Plan.* **2017**, *60*, 265–276. [\[CrossRef\]](#)
10. Wang, M.; Cheng, B.; Chen, J.; Mercer, N.; Kirschner, P.A. The use of web-based collaborative concept mapping to support group learning and interaction in an online environment. *Internet High. Educ.* **2017**, *34*, 28–40. [\[CrossRef\]](#)
11. Collins, H. *Creative Research: The Theory and Practice of Research for the Creative Industries*; AVA Publishing SA.: Lausanne, Switzerland, 2017.
12. Varela, A.R.; Pratt, M.; Harris, J.; Lecy, J.; Salvo, D.; Brownson, R.C.; Hallal, P.C. Mapping the historical development of physical activity and health research: A structured literature review and citation network analysis. *Prev. Med.* **2018**, *111*, 466–472. [\[CrossRef\]](#)
13. Bird, S.; Klein, E.; Loper, E. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2009.
14. Sun, S.; Luo, C.; Chen, J. A review of natural language processing techniques for opinion mining systems. *Inf. Fusion* **2017**, *36*, 10–25. [\[CrossRef\]](#)
15. Hansen, S.; McMahon, M.; Prat, A. Transparency and deliberation within the FOMC: A computational linguistics approach. *Q. J. Econ.* **2018**, *133*, 801–870. [\[CrossRef\]](#)
16. Toutanova, K.; Manning, C.D. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In Proceedings of the 2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora: Held in Conjunction with the 38th Annual Meeting of the Association for Computational Linguistics, Hong Kong, China, 7–8 October 2000; Association for Computational Linguistics: Stroudsburg, PA, USA, 2000; pp. 63–70. [\[CrossRef\]](#)
17. Shen, L.; Satta, G.; Joshi, A. Guided learning for bidirectional sequence classification. In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague, Czech Republic, 25–27 June 2007; Association for Computational Linguistics: Stroudsburg, PA, USA, 2007; pp. 760–767.
18. Søgaard, A. Simple semi-supervised training of part-of-speech taggers. In Proceedings of the ACL 2010 Conference Short Papers, Uppsala, Sweden, 11–16 July 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010; pp. 205–208.
19. Manning, C.D. Part-of-speech tagging from 97% to 100%: Is it time for some linguistics? In Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics, Tokyo, Japan, 20–26 February 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 171–189. [\[CrossRef\]](#)
20. Owoputi, O.; O'Connor, B.; Dyer, C.; Gimpel, K.; Schneider, N.; Smith, N.A. Improved part-of-speech tagging for online conversational text with word clusters. In Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Atlanta, GA, USA, 9–14 June 2013; Association for Computational Linguistics: Stroudsburg, PA, USA, 2013; pp. 380–390.

21. Gui, T.; Zhang, Q.; Huang, H.; Peng, M.; Huang, X. Part-of-speech tagging for twitter with adversarial neural networks. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 7–11 September 2017; Association for Computational Linguistics: Stroudsburg, PA, USA, 2017; pp. 2411–2420. [\[CrossRef\]](#)
22. Balusu, R.M.; Merghani, T.; Eisenstein, J. Stylistic variation in social media Part-of-Speech Tagging. In Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics, New Orleans, LA, USA, 1–6 June 2018; Association for Computational Linguistics: Stroudsburg, PA, USA, 2018; pp. 11–19. [\[CrossRef\]](#)
23. Fang, L.; Tuan, L.A.; Hui, S.C.; Wu, L. Syntactic based approach for grammar question retrieval. *Inf. Process. Manag.* **2018**, *54*, 184–202. [\[CrossRef\]](#)
24. Ganesh, B.R.; Gupta, D.; Sasikala, T. Grammar Error Detection Tool for Medical Transcription using Stop Words Parts-of-Speech Tags Ngram Based Model. *APTİKOM J. Comput. Sci. Inf. Technol.* **2017**, *2*, 47–56. [\[CrossRef\]](#)
25. Vani, K.; Deepa, G. Unmasking text plagiarism using syntactic-semantic based natural language processing techniques: Comparisons, analysis and challenges. *Inf. Process. Manag.* **2018**, *54*, 408–432. [\[CrossRef\]](#)
26. Aly, A.; Taniguchi, A.; Taniguchi, T. A generative framework for multimodal learning of spatial concepts and object categories: An unsupervised part-of-speech tagging and 3D visual perception based approach. In Proceedings of the Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob), Lisbon, Portugal, 18–21 September 2017; pp. 376–383. [\[CrossRef\]](#)
27. Mohit, B. Named Entity Recognition. In *Natural Language Processing of Semitic Languages. Theory and Applications of Natural Language Processing*; Zitouni, I., Ed.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 221–245.
28. Cho, H.-C.; Okazaki, N.; Miwa, M.; Tsujii, J. Named entity recognition with multiple segment representations. *Inf. Process. Manag.* **2013**, *49*, 954–965. [\[CrossRef\]](#)
29. Zhao, G.; Wu, J.; Wang, D.; Li, T. Entity disambiguation to Wikipedia using collective ranking. *Inf. Process. Manag.* **2016**, *52*, 1247–1257. [\[CrossRef\]](#)
30. Ni, J.; Florian, R. Improving Multilingual Named Entity Recognition with Wikipedia Entity Type Mapping. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, TX, USA, 1–5 November 2016; Association for Computational Linguistics: Stroudsburg, PA, USA, 2016; pp. 1275–1284.
31. Das, A.; Ganguly, D.; Garain, U. Named entity recognition with word embeddings and Wikipedia categories for a low-resource language. *ACM Trans. Asian Low-Resour. Lang. Inf. Process. (Tallip)* **2017**, *16*, 18. [\[CrossRef\]](#)
32. Mehdi, M.; Okoli, C.; Mesgari, M.; Nielsen, F.Å.; Lanamäki, A. Excavating the mother lode of human-generated text: A systematic review of research that uses the wikipedia corpus. *Inf. Process. Manag.* **2017**, *53*, 505–529. [\[CrossRef\]](#)
33. Liu, X.; Wei, F.; Zhang, S.; Zhou, M. Named entity recognition for tweets. *ACM Trans. Intell. Syst. Technol.* **2013**, *4*, 3. [\[CrossRef\]](#)
34. Derczynski, L.; Maynard, D.; Rizzo, G.; van Erp, M.G.J.; Gorrell, G.; Troncy, R.; Petrak, K.; Bontcheva, K. Analysis of named entity recognition and linking for tweets. *Inf. Process. Manag.* **2015**, *51*, 32–49. [\[CrossRef\]](#)
35. Bellot, P.; Moriceau, V.; Mothe, J.; SanJuan, E.; Tannier, X. INEX Tweet Contextualization task: Evaluation, results and lesson learned. *Inf. Process. Manag.* **2016**, *52*, 801–819. [\[CrossRef\]](#)
36. Zhang, Y.; Jin, R.; Zhou, Z.H. Understanding bag-of-words model: A statistical framework. *Int. J. Mach. Learn. Cybern.* **2010**, *1*, 43–52. [\[CrossRef\]](#)
37. Baccianella, S.; Esuli, A.; Sebastiani, F. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010), Valletta, Malta, 19–21 May 2010; European Language Resources Association: Paris, France, 2010; pp. 2200–2204.
38. Mukherjee, S.; Bhattacharyya, P. Sentiment analysis in Twitter with lightweight discourse analysis. In Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012): Technical Papers, Mumbai, India, 8–15 December 2012; Indian Institute of Technology Bombay: Mumbai, India, 2012; pp. 1847–1864.
39. Song, J.; Song, X.; Liu, Y. Object recognition algorithm based on improved k-means bag-of-words model. *Univ. Politeh. Buchar. Sci. Bull. Ser. C Electr. Eng. Comput. Sci.* **2017**, *79*, 43–56.

40. Durou, A.; Aref, I.A.; Al-Maadeed, S.; Bouridane, A.; Benkhelifa, E. Writer identification approach based on bag of words with OBI features. *Inf. Process. Manag.* **2017**, *56*, 354–366. [[CrossRef](#)]
41. Sarwar, A.; Mehmood, Z.; Mehmood, T.; Qazi, K.A.; Adnan, A.; Jamal, H. A novel method for content-based image retrieval to improve the effectiveness of the bag-of-words model using a support vector machine. *J. Inf. Sci.* **2018**, *45*, 117–135. [[CrossRef](#)]
42. Gernot, A.F. *Markov Models for Pattern Recognition—From Theory to Applications*; Springer: London, UK, 2014.
43. Figueiredo, L.N.; de Assis, G.T.; Ferreira, A.A. DERIN: A data extraction method based on rendering information and n-gram. *Inf. Process. Manag.* **2017**, *53*, 1120–1138. [[CrossRef](#)]
44. Chelba, C.; Mikolov, T.; Schuster, M.; Ge, Q.; Brants, T.; Koehn, P.; Robinson, T. One billion word benchmark for measuring progress in statistical language modeling. In Proceedings of the 15th Annual Conference of the International Speech Communication Association, Singapore, 14–18 September 2014; Causal Productions Pty. Ltd.: Adelaide, Australia, 2014; pp. 2635–2639.
45. Vilares, J.; Alonso, M.A.; Doval, Y.; Vilares, M. Studying the effect and treatment of misspelled queries in Cross-Language Information Retrieval. *Inf. Process. Manag.* **2016**, *52*, 646–657. [[CrossRef](#)]
46. Xu, H.; Li, K.; Wang, Y.; Wang, J.; Kang, S.; Chen, X.; Povey, D.; Khudanpur, S. Neural network language modeling with letter-based features and importance sampling. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing, Calgary, AB, Canada, 15–20 April 2018; pp. 6109–6113.
47. Osman, A.H.; Salim, N.; Binwahlan, M.S.; Altee, R.; Abuobieda, A. An improved plagiarism detection scheme based on semantic role labeling. *Appl. Soft Comput.* **2012**, *12*, 1493–1502. [[CrossRef](#)]
48. Christensen, J.; Soderland, S.; Etzioni, O. Semantic Role Labeling for Open Information Extraction. In Proceedings of the 1st International Workshop on Formalisms and Methodology for Learning by Reading (NAACL HLT 2010), Los Angeles, CA, USA, 5–6 June 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010; pp. 52–60.
49. Zhou, J.; Xu, W. End-to-end learning of semantic role labeling using recurrent neural networks. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China, 27–31 July 2015; Association for Computational Linguistics: Stroudsburg, PA, USA, 2015; pp. 1127–1137.
50. Marcheggiani, D.; Frolov, A.; Titov, I. A Simple and Accurate Syntax-Agnostic Neural Model for Dependency-based Semantic Role Labeling. In Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017), Vancouver, BC, Canada, 3–4 August 2017; Association for Computational Linguistics: Stroudsburg, PA, USA, 2017; pp. 411–420.
51. Della Rocca, P.; Senatore, S.; Loia, V. A semantic-grained perspective of latent knowledge modeling. *Inf. Fusion* **2017**, *36*, 52–67. [[CrossRef](#)]
52. Zhang, P.; Essaid, A.; Zanni-Merk, C.; Cavallucci, D. Case-based Reasoning for Knowledge Capitalization in Inventive Design Using Latent Semantic Analysis. *Procedia Comput. Sci.* **2017**, *112*, 323–332. [[CrossRef](#)]
53. Zhu, G.; Iglesias, C.A. Exploiting semantic similarity for named entity disambiguation in knowledge graphs. *Expert Syst. Appl.* **2018**, *101*, 8–24. [[CrossRef](#)]
54. Babar, S.A.; Patil, P.D. Improving Performance of Text Summarization. *Procedia Comput. Sci.* **2015**, *46*, 354–363. [[CrossRef](#)]
55. Badry, R.M.; Eldin, A.S.; Elzanfally, D.S. Text Summarization within the Latent Semantic Analysis Framework: Comparative Study. *Int. J. Comput. Appl.* **2013**, *81*, 40–45. [[CrossRef](#)]
56. Bookjans, M.; Weckenmann, A. Virtual Quality Management: A holistic approach for the set-up of true-to-life process models and according virtual quality management tools. In Proceedings of the 1st International Conference on Quality and Innovation in Engineering and Management, Cluj-Napoca, Romania, 17–19 March 2011; pp. 17–22.
57. Mertins, K.; Rabe, M.; Gocev, P. Integration of Factory Planning and ERP/MES Systems: Adaptive Simulation Models. In *Lean Business Systems and Beyond*; Koch, T., Ed.; Springer: Boston, MA, USA, 2008; pp. 185–193.
58. Heilala, J.; Vatanen, S.; Tonteri, H.; Montonen, J.; Lind, S.; Johansson, B.; Stahre, J. Simulation-based sustainable manufacturing system design. In Proceedings of the 2008 Winter Simulation Conference, Miami, FL, USA, 7–10 December 2008; pp. 1922–1930. [[CrossRef](#)]
59. Bookjans, M.; Weckenmann, A. Virtual Quality Management—Validation of measurement systems by the use of simulation technologies. *Phys. Procedia Laser Assist. Net Shape Eng.* **2010**, *5*, 745–752. [[CrossRef](#)]

60. Stockinger, A.; Wittmann, S.; Martinek, M.; Meerkamm, H.; Wartzack, S. Virtual Assembly Analysis: Standard tolerance analysis compared to manufacturing simulation and relative positioning. In Proceedings of the International Design Conference (DESIGN 2010), Dubrovnik, Croatia, 17–20 May 2010; The Design Society: Glasgow, Scotland, 2010; pp. 1421–1430.
61. Dix, A. Human–computer interaction, foundations and new paradigms. *J. Vis. Lang. Comput.* **2017**, *42*, 122–134. [[CrossRef](#)]
62. Eschen, H.; Kötter, T.; Rodeck, R.; Harnisch, M.; Schüppstuhl, T. Augmented and Virtual Reality for Inspection and Maintenance Processes in the Aviation Industry. *Procedia Manuf.* **2018**, *19*, 156–163. [[CrossRef](#)]
63. Milella, F. Problem-Solving by Immersive Virtual Reality: Towards a More Efficient Product Emergence Process in Automotive. *J. Multidiscip. Eng. Sci. Technol.* **2015**, *2*, 860–867.
64. Dulebohn, J.H.; Hoch, J.E. Virtual teams in organizations. *Hum. Resour. Manag. Rev.* **2017**, *27*, 569–574. [[CrossRef](#)]
65. Weckenmann, A.; Bookjans, M. Zuverlässigere Prozesse durch Virtuelles Qualitätsmanagement. *Zeitschrift für wirtschaftlichen Fabrikbetrieb* **2008**, *103*, 448–451. [[CrossRef](#)]
66. Lee, J.; Kao, H.-A.; Yang, S. Service innovation and smart analytics for Industry 4.0 and big data environment. *Procedia CIRP* **2014**, *16*, 3–8. [[CrossRef](#)]
67. Kagermann, H.; Wahlster, W.; Helbig, J. *Recommendations for Implementing the Strategic Initiative Industrie 4.0—Final Report of the Industrie 4.0 Working Group*; Acatech: Munich, Germany, 2013.
68. Tropes US. Help Index: Classes and Relations graphs, Version 8.4.2 (build 0019). Released June 2014. Available online: <https://www.semantic-knowledge.com/download.htm> (accessed on 7 June 2016).
69. International Organization for Standardization: ISO/IEC/IEEE 15288:2015. Available online: <https://www.iso.org/standard/63711.html> (accessed on 5 January 2018).
70. List of Romanian Companies. Available online: <https://www.romanian-companies.eu/> (accessed on 12 May 2018).

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).