Innovation Economics and Natural Language Processing

University of Texas at Dallas

Glen Cooper
(gxc162030@utdallas.edu / glco@sbcglobal.net / 214-914-6756)

for

Dr. Hicks - Fall 2020 PPPE 6365

List of Abbreviations

The following abbreviations appear in this paper:

AI – Artificial Intelligence
ALPAC - Automatic Language Processing Advisory Committee
CAGR - Compound Annual Growth Rate
CPU - Central processing unit
FFNN - Feed-forward neural networks
GPT – General Purpose Technology
GPU - Graphics processing unit
LISP - Locator/Identifier Separation Protocol
LSTM - Long short-term memory
NIH - Department of Energy, National Institutes of Health
NIST - National Institute of Standards and Technology
NLP – Natural Language Processing
NRC - U.S. National Research Council
NSF - National Science Foundation
RED - Real-time early detection
RNN - Recurrent neural nets

Introduction

Natural Language Processing (NLP) is a branch of a larger branch of technology called Artificial Intelligence (AI). AI is considered the fourth industrial revolution. After providing a brief explanation of AI and NLP, this paper will generally demonstrate how NLP has developed and continues to develop in accordance with innovation economic theories. It will also be shown how NLP is or is likely to impact society also along ways predicted by innovation economic. Specifically, this paper will discuss the impact on NLP from individuals, universities, government, convergent technologies, market forces, oligopolies, and entrepreneurs. Next this paper will discuss the specific impacts or likely impacts of NLP on the economy, organizational capital, and human capital. Finally, this paper will review the important public policies issues related to NLP technology and then end with a few concluding remarks.

Artificial Intelligence and Natural Language Processing

Artificial Intelligence (AI) is being called the fourth industrial revolution by many. Note, there is some debate on the revolutionary "count" with some like Atkinson and Wu (2017) following the Schumperterian technology long-wave theory labeling it the "sixth wave" (p. 21). Without regard to this dispute, the first industrial revolution began in the 1800s. This first revolution saw advances in "hard" capital such as railways, heavy industries, and the steam engine. The second revolution began near the end of the 19th century and is identified with electrification and the assembly line "soft" process innovation. The third industrial revolution began in the 1970s and was distinguished by automation through electronics. This revolution saw the innovation of the personal computer and the internet. It allowed global access to information and an increase in data processing capabilities. The newest revolution is the technical integration of cyber and physical systems into physical production processes. The introduction of AI distinguishes this fourth revolution (Lischka, 2011). As with previous revolutionary technology, AI is a General Purpose Technology (GPT) in that it is pervasive, improving over time, and able to spawn new innovations (Brynjolfsson & McAfee, 2014).

The story of AI can begin with Dr. John McCarthy's (1927-2011) from Stanford University. He was the first to use the term Artificial Intelligence. Dr. McCarthy's proposal was that "… every aspect of learning and other properties of intelligence can be described so precisely that a machine can simulate it" (Lischka, 2011). In 1950, Alan Turing (1912 – 1954) wrote an inspiring paper describing a test for a "thinking" machine. He stated that if a machine could be part of a conversation through the use of a teleprinter, and it imitated a human so completely that there were no noticeable differences, then the machine could be considered capable of thinking. Shortly after this, in 1952, the Hodgkin-Huxley model, showed how the brain uses neurons in forming an electrical network. These events helped inspire the idea of AI (Foote, 2019). Consistent with its GPT character, AI has a wide range of applications, including: 1) Machine learning that automates model building and allows machines to operate autonomously; 2) Virtual personal assistants that help users by providing reminders, scheduling appointments, and finding service providers; 3) Machine vision that allows computers to identify objects; and 4) Natural language processing that allows computers to understand language (Chen, Christensen, Gallagher, Mate, & Rafert, 2016).

This last component of AI will be the focus of this paper. NLP is an aspect of AI that helps computers understand, interpret, analyze, generate, and utilize human languages. NLP allows computers to communicate with people in a natural way by allowing computers to read text, hear speech, and interpret both. Generally, NLP attempts to close the gap between human and computer communications (Foote, 2019).

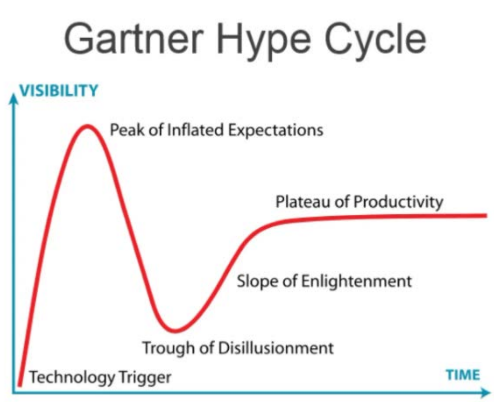
Individuals, Universities & Government

As with many innovations NLP resulted from a spin-off of the efforts occurring within the university system (Block & Keller, 2009). Ferdinand de Saussure (1857 – 1913) was a Swiss linguistics professor that taught courses at the University of Geneva under the concept of "Language as a Science." Professor Saussure developed an approach describing languages as "systems." He proposed that within a language, sounds represent concepts – concepts that shift meaning as the context changes (Foote, 2019). Saussure viewed society as a system of shared norms from which language develops and provides for reasonable thinking, decisions, and actions by individuals. After Saussure's death two of his colleagues, Albert Sechehaye (1870 – 1946) and Charles Bally (1865 – 1947), recognized the importance of his work and collected his and his students' notes. Using that material, they wrote and published, in 1916, the Cours de Linguistique Generale. This book laid the foundation for the structuralist approach to linguistics. The concepts of the structuralist approach could and were translated into computer code (Foote, 2019). One of the first persons to utilize the computer for language processing was an Italian Jesuit priest named Roberto Busa (1913 – 2011). As a pioneer of computational linguistics, he collaborated with IBM founder Thomas J. Watson (1874 – 1956) to create a computer-readable indexing of the complete works of St. Thomas Aquinas, the 13th-century Catholic priest and philosopher. The "Index Thomisticus" project took more than 30 years and eventually was published in 56 volumes based on more than 11 million computer punch cards, one for every word analyzed (Eggers, Malik, & Gracie, 2018).

NLP first began receiving widespread recognition in the 1950s, when researchers and linguistics experts began developing computer programs to automate language translation (Eggers, Malik, & Gracie, 2018). Again the universities planned a key role in the advancement of NLP. These early computer programs were called rationalistic (Church, 2007). The adoption of this approach was mainly due to the widespread acceptance of Noam Chomsky's (1928 -) idea that language was innate. In 1957, Chomsky postulated that key parts of language were hardwired in the brain at birth and were part of the human genetic inheritance (Chomsky, 2002). The rationalist approaches endeavored to design hand-crafted rules to incorporate knowledge and reasoning mechanisms into intelligent NLP systems based upon a complex set of handwritten rules (Deng & Liu, 2018). One of the first attempts at translation was the Georgetown University's experiment conducted in 1954. This experiment involved fully automatic translation of more than sixty Russian sentences into English. While this experiment might be labeled a "toy," it received widespread attention and created excitement around NLP possibilities (Wikipedia, 2020). In 1958, the programming language, LISP (Locator/Identifier Separation Protocol), a computer language for NLP development and still in use today, was released by the aforementioned Dr. McCarthy (Foote, 2019). Then in 1964, ELIZA, a "typewritten" comment and response program was designed by Joseph Weizenbaum (1923 – 2008) at the MIT Artificial Intelligence Laboratory. This program imitated a psychiatrist using Rogerian psychotherapist reflection techniques. The program communicated by rearranging sentences and following relatively simple grammar rules. However, there was no understanding on the computer's part. The program did produce startlingly human-like interaction. Here is an example of an ELIZA conversation (Foote, 2019, Wikipedia, 2020):

This period of development coincided with the early developments of artificial intelligence, characterized by expert knowledge engineering. These expert systems where topic germane expert computer programs with very specific knowledge domains. The experts designed these programs using symbolic logical rules (i.e. if-then combinations), but the programs lacked learning capabilities and could not generalize solutions to new situations (Deng & Liu, 2018).

However, while these technologies were limited, a consequence of the excitement this new technology engendered launched NLP on to Gartner, Inc.'s hype cycle. The Gartner hype cycle theorizes that technologies move through a process of trigger (i.e. the development of a new technology), peak of inflated expectations, trough of disillusionment (i.e. community begins to lose hope), slope of enlightenment (i.e. beginnings of rapid advancements) and plateau of productivity (i.e. mature growth) (Gartner Inc., 2020). Graphically depicted as follows:



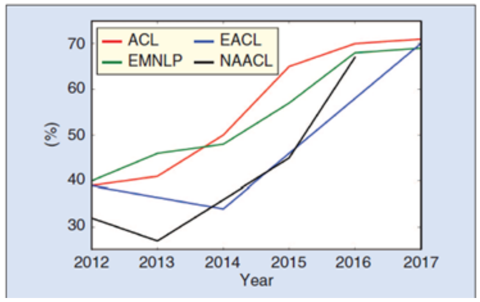
Another source of technology development is government (Block & Keller, 2009) investment. The excitement for this new technology resulted in the U.S. National Research Council (NRC), an organization receiving grants from Federal agencies, created the Automatic Language Processing Advisory Committee (ALPAC) in 1964. This committee was tasked with evaluating the progress of NLP research. However, after twelve years of research, and $20 million dollars of investment, machine translation systems were still more expensive than manual human translations, and no computer came close to carrying out basic human conversations. By 1966, NLP research was considered a dead end by many (Foote, 2019). That is, the technology had entered Gartner's trough of disillusionment. As innovation economics predicts, innovation is expensive, the development of new technology often involves lengthy trial-and-error processes, and most endure many dead ends (Taylor, 2016). The rationalistic approach was one of those dead ends.

It took nearly fourteen years for NLP research to recover. The stoppage had, however, initiated the Gartner slope of enlightenment, with earlier concepts of rationalistic linguistics being replaced with pure statistical approaches (Deng & Liu, 2018). With this realignment we see innovation demonstrating steady improvement but were the past does not necessarily portend what comes next (Brynjolfsson & McAfee, 2014). The new statistical approach was paralleled with corresponding approaches in AI to computer vision. It became clear in both NLP and computer vision that learning and perception capabilities were crucial for complex AI systems and something that was missing in the expert systems (Deng & Liu, 2018). Thus, the 1980s initiated a fundamental reorientation, replacing NPL's handwritten rules systems of simple approximations with more rigorous statistics. This was the result of both the steady increase of computational power and the shift to statistical machine learning algorithms. Decision trees being a good example of a machine learning algorithm (Foote, 2019). These new statistical models were also capable of making soft probabilistic decisions something that the expert systems could not do.

One of the features of the new computer age is the digitization of just about everything (Brynjolfsson & McAfee, 2014). Digitization was also critical to the new approach because without data to operate upon statistics is ineffective. Data initially came from governmental developments of multilingual textual corpora produced by the Parliament of Canada and the European Union as a result of laws calling for the digital translation of all governmental proceedings into all official languages of the government (Wikipedia, 2020) and, of course, the internet began developing and creating textual materials around this period. Note specifically, how foreign language databases can be a type of Rosetta stone giving words references to multiple meanings (Young, Hazarika, Poria & Cambria, 2018) and providing NLP systems with significant referential material. Throughout the 1980s, IBM was responsible for the development of several successful, complicated statistical models. In the 1990s, the popularity of statistical models for NLP analyses rose dramatically as the pure statistics methods became remarkably valuable in keeping pace with the tremendous flow of digitized text (Foote, 2019).

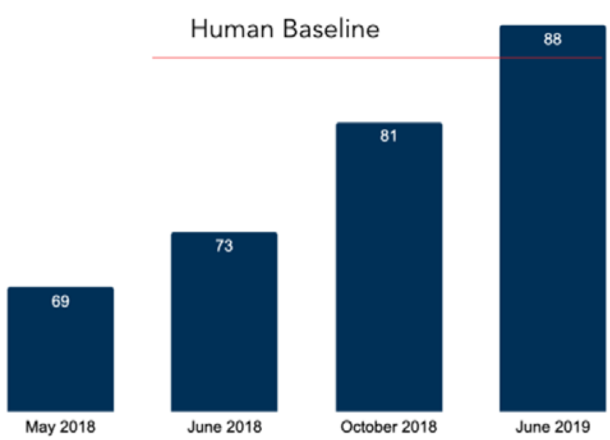
Ironically, while Chomsky's ideas initially advanced NLP, in the end, by his own words, "you do not get discoveries in the sciences by taking huge amounts of data, throwing them into a computer and doing statistical analysis on them: that's not the way you understand things …" (Young, Hazarika, Poria & Cambria, 2018). And indeed, more was needed for NLP to advance than pure statistics. This is where neural nets or deep learning revolutionized NLP. Statistical methods are based on data-intensive machine learning, which is now call "shallow" due to the general lack of abstraction. Neural nets allow for the construction of many-layered or "deep" representations of data (Deng & Liu, 2018).

In a collaboration between the university system and business the California Institute of Technology and Bell Laboratories sponsored in 1986 the NeurIPS an invitation-only meeting on neural networks for computing. This interdisciplinary meeting was one of the first to included researchers exploring artificial neural networks (Conference on Neural Information Processing Systems, 2020). The concepts of neural net deep learning advanced quickly ultimately leading to the development of Long short-term memory (LSTM) recurrent neural nets (RNN) in 1997. In addition to the RNN model, feed-forward neural networks (FFNN) that moves data only in one direction, from input nodes, through hidden nodes, and then to output nodes were developed. The FFNN has no cycles or loops as in the RNN model making it much faster. Currently, these neural net models are considered the cutting edge of NLP (Foote, 2019) as demonstrated from the rapidly increasing level of professional journal publication appearing in this area (Young, Hazarika, Poria & Cambria, 2018):



ACL – Association for Computational Linguistics,
EACL – European Association for Computational Linguistics,
NAACL – North American Association for Computational Linguistics,
EMNLP - Empirical Methods NLP

These neural net approaches have moved NLP to an inflection point with a wide range of practical, including business, applications (Deng & Liu, 2018) or on to the Gartner plateau of productivity phase with striking successes in practical tasks. Not only has machine translation been advanced but speech recognition has become possible and commercially viable. Consider that one current speech recognition program (i.e. BERT) surpassed human baseline General Language Understanding Evaluation (GLUE) index with only about one year of neural net training (Delangue & Nadella, 2020):



https://gluebenchmark.com/leaderboard from https://www.stateof.ai

Speech recognition programs are even robust against noisy environments (Deng & Liu, 2018)!

In addition to machine translation and speech recognition, NLP has now become practical in many other areas, including language dialog, lexical analysis and parsing, knowledge graphing, information retrieval, question answering from text, social computing, language generation, and text sentiment analysis. Today, deep learning is a dominating method applied to practically all NLP tasks (Deng & Liu, 2018). Current NLPs, however, are still not able to pass Alan Turing's test, and currently do not sound like real human beings. At least not yet (Foote, 2019).

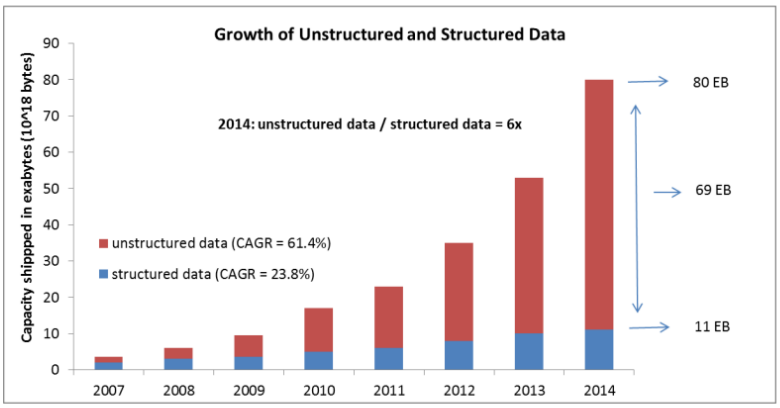
There are two schools around government role in innovation. The first is that government should be a primary driver of innovation while the second school says that government interference only hinders innovation (Taylor, 2016). In regards to NLP, and AI in general, the U.S. government appears to have taken a somewhat limited role to date. However, this may be changing. On February 11, 2019, President Trump signed Executive Order 13859 - "Maintaining American Leadership in Artificial Intelligence" (Exec. Order No. 13859, 2019). This strategy is a concerted effort to promote and protect national AI technology and innovation. The Initiative implements a whole-of-government strategy in collaboration and engagement with the private sector, academia, the public, and like-minded international partners. It directs the Federal government to pursue five pillars for advancing AI: (1) invest in AI research and development (R&D), (2) unleash AI resources, (3) remove barriers to AI innovation, (4) train an AI-ready workforce, and (5) promote an international environment that is supportive of American AI innovation and its responsible use. The U.S. is also actively leveraging AI to help the Federal government work smarter in its own services and missions in trustworthy ways (United States, 2019). Fiscal 2020 also represents the first year in which the White House has released a figure for civilian agencies' combined investments in AI-related R&D. In support of Executive Order 13859, the Department of Energy, National Institutes of Health (NIH), National Institute of Standards and Technology (NIST), and National Science Foundation (NSF) will devote a combined $850 million to AI research, according to the fiscal 2020 budget request. Additionally, the Defense Department plans to allocate $4 billion toward AI and machine learning R&D activities in fiscal 2020, a sizeable increase above fiscal 2019 spending levels, according to a Bloomberg Government analysis. While Executive Order 13859 arrived without specific funding figures attached, it is now clear government agencies recognize the urgency of making AI-related investments (Cornillie, 2019).

Convergent Technologies

The contributions from individual theories and, university and government support were the foundations of NLP technology progress, if it were not for the convergences of two key technology developments NLP would likely not be possible. First, the decades of exponential growth in computing performance, and, second, the digitization of large datasets (Allen & Chan, 2017).

Most are familiar with Moore's law identified in 1965 by Gordon Moore (1929-) the co-founder of Intel Corporation. This law states that central processing unit (CPU) computing power will roughly double every year. A prediction that has remained remarkable true (Brynjolfsson & McAfee, 2014). However, CPU power received a remarkable enhancement when it was coupled with graphics processing unit (GPU) computing. A CPU works together with a GPU to increase the throughput of data and the number of concurrent calculations within an application. GPUs were originally designed to create images for computer graphics and video game consoles, but since the early 2010's, GPUs can also be used to accelerate calculations involving massive amounts of data. A GPU complements the CPU architecture by allowing repetitive calculations within an application to be run in parallel while the main program continues to run on the CPU. The CPU can be thought of as the taskmaster of the entire system, coordinating a wide range of general-purpose computing tasks, with the GPU performing a narrower range of more specialized tasks (usually mathematical). Using the power of parallelism, a GPU can complete more work in the same amount of time as compared to just a CPU (Conference on Neural Information Processing Systems, 2020). This CPU plus GPU has turbo charged the ability of NLP software to perform with credibility.

As mentioned above digitization is the work of turning information and media—text, sounds, photos, video, data from instruments and sensors, and so on — into the language of computers (Brynjolfsson & McAfee, 2014). As shall be discussed below some of this digitization has come from government sources but also as computer technology has proliferative into every aspect of life much of the material is now begin created by business and individuals with much of it being created on the fly. The rise of the web and sites like Wikipedia, Facebook, YouTube, Twitter, Instagram, Pinterest, Reddit, and others provide gigabits of unstructured language text written in natural language, which allowed NLP models to be trained (Delangue & Nadella, 2020, Wikipedia, 2020). Most documents (e.g. email, news groups, news articles, business reports, research papers, blogs, resumes, proposals, academic reports, etc.) are all being created digitally. Today approximately two and a half exabytes (10^{18}) of unstructured data are created daily. In fact, the amount of unstructured data that was created in the past two days is equivalent to the same amount of data that was created from the beginning of humankind through the end of 2003. The graph below shows the exponential growth rate of both unstructured and structured data (Zhao, 2017):



Source: S&P Global Market Intelligence Quantamental Research. IDC. 2015.

EB = Exabytes (10^{18})

All this digital language-based data has been a boom for NLP as it provides the platform of operations for the CPU plus GPU computing power combination to work upon. This convergence has propel NLP forward.

Market Forces

Innovative activity is a cumulative process and once the free market innovation machine was launched its inherent structure leads the machine to grow more powerful and productive with the passage of time (Baumol, 2002). If you consider any organization or individual life you realize that they are built around text. From sales, customer support, customer reviews, comments, internal collaboration, product descriptions, emails, calendars, and personal journals — the use of text is never ending. Every aspect of a business and life is centered around language and text (Delangue & Nadella, 2020). As noted earlier this text has become digitized and machine readable. Now that computing performance has improved, and the techniques have developed to take advantage of this data, the market is being driven to increase commercial investment in the technology (Allen & Chan, 2017). Two fundamental factors are primarily driving this investment: companies seeking competitive advantage and the demand by customers for improved experiences.

Modern growth models estimate final output based on physical capital, hours worked, human capital per person, and the stock of ideas. Traditional growth accounting (Solow, 1956) calculates the stock of ideas as a residual. Modern growth theory explains that residual in terms of economic forces. Embedded in the modern production function is the key insight of Paul Romer (1955-) (Romer, 1986)): that new ideas lead to increasing returns. New ideas come from an idea production function that depends on the number of people looking for new ideas as well as on the existing stock of ideas (Fernald & Jones, 2014). NLP of large databases of scientific and non-scientific papers can cross-reference and summarize these materials for easier, or perhaps automated, idea generation is one of the driving market forces driving NLP technologies (Singh, 2018).

Microeconomic competition is another driving force of NLP technology. Companies are using NLP to conduct market analysis to better understand markets, competitors, and other important industry details. The ability of human beings to effectively make use of the vast amount of digital information is low as this task can be boring, tedious, and time consuming. The explosion of information and need for more sophisticated and efficient information handling tools highlights the need of NLP technology (O'Neil & Paik, 1998). NLP technologies can help to efficiently analyze free text and to discover valuable and relevant knowledge from it in the form of structured information. The goal is to extract salient facts about pre-specified types of events, entities, or relationships, in order to build more meaningful, rich representations of their semantic content, which can be used to populate databases that provide structured information (Singh, 2018). For example, NLP engines are being used to automatically build competitor landscapes by scanning the internet and gathering, listing, and ranking and relating market competitors (Maddipudi, 2020, Ruliputra, 2019). These NLP engines can also automatically incorporate competitors pricing and promotions, and public opinion data from various sources and present the data for better decision making (Singh, 2018).

Another application of NLP is to better understand the customer. Customers now provide feedback and desired requirements through an increasing number of channels including customer generated ratings, emails, chat transcripts, and support forum discussions (Fortune Business Insights, 2020). Understanding this customer feedback and associated requirements is becoming increasing difficult to search and retrieve manually. Therefore, companies are turning to NLP tools to retrieve relevant information from these sources, classify them and form structured database about customer requirements (Singh, 2018). In addition to determining customer requirements, it is equally important to know how customers feel about a company's product or service. Companies are using an NLP technique known as sentiment analysis to gather intelligence on customer's feelings (Sun, Luo & Chen, 2017). Scanning social media platforms such as Facebook, Twitter, Instagram and blogs as well as conventional news (Singh, 2018) companies can review positive and negative customer sentiments and convert it into insightful information (Maddipudi, 2020).

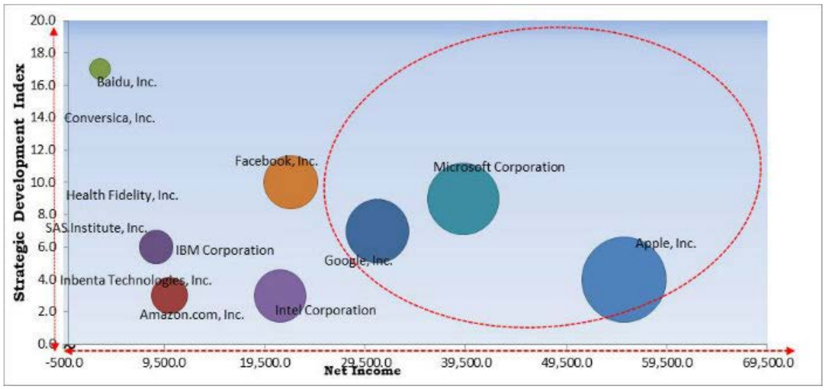
Competitive advantage is also being sought through cost reduction and efficiency. Many employees are burdened with making response to emails and scheduling calendar events. However, most emails have a Yes/No reply, and many contain information about upcoming meetings. Reading, indexing, replying, and calendaring such emails is a time-consuming task. NLP technology can automatically respond to emails and calendar events saving significant time (Singh, 2018). Another example can be found in resume processing. Human resource departments can spend considerable time extracting information from resumes. Manually reading individual resume, checking their prospects as per job requirements, replying to submissions, and schedule interviews is very time consuming. Using NLP technology, resumes can be processed and replied to automatically (Singh, 2018). A very significant application of NLP is found in documenting business activities. Documenting and reporting business activity is among the most time-consuming tasks for most businesses. NLP techniques allow companies to convert unstructured text information into structured data entry and reports. Sophisticated solutions can also identify and request missing data in an automated way (Maddipudi, 2020). Such methods can also complete regulatory forms reducing the compliance burden (Métais, Meziane, Saraee, Sugumaran, & Vadera, 2016) or to evaluate whether regulatory templates have been correctly and accurately completed (Arora, Sabetzadeh, Briand, & Zimmer, 2015).

Governments have also been utilizing NLP to reduce costs and improve services. The European union's horizon 2020 program launched an initiative called Real-time Early Detection (RED) alert, aimed at countering terrorism by using NLP to monitor and analyze social media conversations (Eggers, Malik, & Gracie, 2018). The Police Department of Durham, NC uses NLP in crime fighting by enabling the police to observe patterns and interrelations in criminal activities and identify pockets of high incidents of crime, thus allowing for quicker interventions. The method has contributed to a 39% drop in violent crime in the area from 2007 to 2014 (Eggers, Malik, & Gracie, 2018). The U.S. Defense Department uses NLP to automatically extract operationally relevant information from unstructured text to help defense analysts derive actionable insights (Eggers, Malik, & Gracie, 2018). The US Department of energy's Oak Ridge National Laboratory is leveraging NLP capabilities to extract data on energy egosystem components to rank the top clean energy innovation egosystems in the U.S. They are using NLP to transform text and numerical data into metrics on clean energy innovation activity and geography. This helps investors, researchers, and corporations rapidly identify, quantify, and compare clean energy innovation (Eggers, Malik, & Gracie, 2018). Similarly, researchers at the environmental defense fund are working to develop a system backed by NLP that can analyze applications for oil and gas permits submitted under the national Environmental Protection act. The system will provide a deeper analysis of filed applications, thereby helping local regulators and other stakeholders determine whether a project may pose a threat to wildlife, water, or cultural heritage sites (Eggers, Malik, & Gracie, 2018).

The other large market driver for NLP is the demand for improved customer experience. Innovation typically moves customers to higher levels on the experience chain; from Commodity, Goods and Services, and finally, to Experience (Hicks, 2020, slide 52). Most value to consumers now comes from convenience and efficiency (Brynjolfsson & McAfee, 2014). NLP is a technology that enables both convenience and efficiency and so is in high consumer demand. The growing deployment of web-based business-to-customer applications and the increasing use of computer technology by the ordinary consumer is increasing the demand for efficient human–machine interfaces through natural speech (VynZ Research, 2019). NLP fills that need. For example, customers often find it difficult to find products online. NLP can be used to extract attributes of interest and classify various advertisements into pre-defined classes like cars, shoes, kitchenware, electronic, apparel, and so on (Singh, 2018). Another example of a consumer convenience provided by NLP is the ability of consumers to assemble personally customized magazines through NLP searches (Métais, Meziane, Saraee, Sugumaran, & Vadera, 2016).

Oligopolies

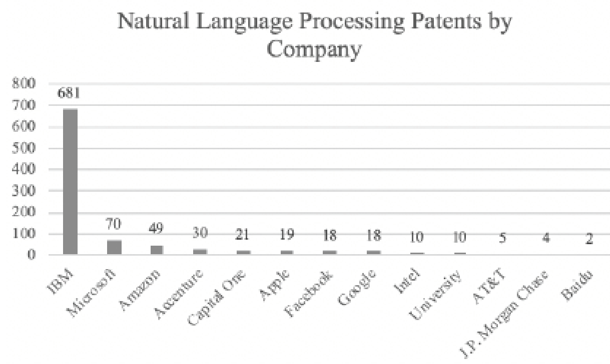
Innovation economics theorizes that oligopolistic competition plays a major role in innovations, especially in the high-tech market. Indeed, Microsoft Corporation, Google, Inc., and Apple, Inc. where the forerunners in the NLP Market (Kbv Research, 2019). William Baumol (1922 – 2017) postulates that for large high-tech business a firm's innovation has replaced prices as the prime competitive weapon. For Baumol, firms must innovate or die (Baumol, 2002). This concept appears to be playing out in NLP development. For example, machine translation using NLP neural nets has quickly been incorporated into the major technology firms. Google announced the first stage of its move to neural machine translation in September 2016 and Microsoft, made a similar announcement two months later. At the same time Facebook had been working on the conversion to neural machine translation, and by August 2017, had their version in full deployment (Deng & Liu, 2018). Additionally, as Kbv Research, a market research firm's study demonstrates, based on net income, most of the NPL market is being captured by a few major companies (Kbv Research, 2019):



Additionally, as newness has become too important to be left to serendipity entire industries build internal innovation capacities (Hicks & Budiman, 2020). This also appears to be a factor in NLP development as a quick internet search suggest that the major industry competitors in NLP have significant research and development expenditures:

Major Solution Providers per Market Analytics	R&D Budget (USD Billions)	Year
Amazon Web Services Inc. (AWS)	23.3	2017
Microsoft Corporation (Microsoft)	19.3	2020
Alphabet Inc. (Google)	16.6	2017
Facebook Inc. (Facebook)	13.6	2019
Intel Corporation (Intel)	13.1	2017
Apple Inc (Apple)	11.6	2017
IBM Corporation (IBM)	5.99	2019
Baidu Inc. (Baidu)	2.7	2018

Additional confirmation of this concept appears in the United States patent records which shows IBM, Microsoft, and Apple own the largest portions of the United States NLP patents:



Interestingly IBM owns a significant 36% portion of the NLP patents market patents. Microsoft and Amazon own the second and third largest portions of the market with 70 and 49 patents, respectively. Microsoft, Amazon, Apple, Facebook, and Google have a combined 174 NLP patents, or only about 9% of the total. Thus, IBM owns more than three times as many NLP patents as the five companies combined (Haney, 2020). In a paper by Salome Baslandze (2016), an assistant professor at Einaudi Institute for Economics and Finance, she theorizes that the degree of a firms openness rests on the two counteracting incentives: on the one hand, firms can more easily learn from each other when knowledge is diffused; on the other hand, the improved access to knowledge increases the scope of business-stealing. In high-tech sectors where firms benefit more from external knowledge, the former – open knowledge diffusion - effect will likely dominate whereas in sectors that do not rely much on external knowledge the latter - business-stealing- effect will dominate. (Aghion, Jones, & Jones, 2017). A potential reason for the disparity in IBM patent filings may be that IBM, being an older firm, may have developed a culture for creating patents and is more influenced by the business-stealing-effect while many of the younger firms may be following an open knowledge diffusion model in which proprietary inventions are considered and anathema (Hicks, 2020, slide 35). Like Nike's approach when it released many of its patents in 2010 (Johnson, 2010) it is possible IBM will also open their patents to the community in the future. The openness model is also suggested by many major technology firms cooperatively funding collaborative efforts by investing in such consortiums as OpenAI that is developing transformative NLP aimed at developing a general language model (Haney, 2020) to be shared in open source software.

Another indication that oligopolistic competition is playing a major role in NLP innovations is the flood of new product launches and product expansions being released by the major technology companies. For example:

- Nov-2019: Microsoft Research's Natural Language Processing Group unveiled a dialogue generative software pre-trained on more than 147M dialogues.
- Nov-2019: Baidu introduced upgrades that include a streamlined NLP toolkits.
- Oct-2019: Google introduced two new dialog datasets for NLP development. These datasets include thousands of conversations as well as annotations and labels for training digital assistants in determining the intentions and preferences of users.
- Aug-2019: Conversica made advancements in its conversational AI platform that powers front-office conversations systems to allow for personalizing interactions.
- Aug-2019: Facebook announced the launch of Misspelling Oblivious (word) Embeddings (MOE), a new model for bolstering the use of NLP (Kbv Research, 2019).

Many innovation and competitiveness scholars point out that there is a networking advantage when innovative firms cluster together in small, relatively well-defined regions. That is, proximity is important. This may relate to the ability to focus skilled labors, markets for inputs, and spillover (Taylor, 2016) within a defined space. This phenomenon appears to be occurring within the group of major NLP companies. According to MarketsandMarkets, a U.S. based market research company, the major NLP market solution providers are comprised of the following companies (Market Analytics, 2019):

Major Solution Providers per Market Analytics	US Headquarters
Alphabet Inc. (Google)	Mountain View, CA
Intel Corporation (Intel)	Santa Clara, CA
Apple Inc (Apple)	Cupertino, CA
Baidu Inc. (Baidu)	Sunnyvale, CA
Veritone Inc. (Veritone)	Costa Mesa, CA
Facebook Inc. (Facebook)	Palo Alto, CA
OpenAI	San Francisco, CA
Bitext Innovations S.L. (Bitext)	Redwood City, CA
Health Fidelity Inc. (Health Fidelity)	San Mateo, CA
Conversica Inc. (Conversica)	Foster City, CA
Inbenta Technologies Inc. (Inbenta Technologies)	Sunnyvale, CA
Narrative Science	Chicago, IL
Linguamatics Ltd. (Linguamatics)	Marlborough, MA
3M Company (3M)	Saint Paul, MN
SAS Institute Inc. (SAS Institute)	Cary, NC
Automated Insights Inc. (Automated Insights)	Durham, NC
IBM Corporation (IBM)	Armonk, NY
Dolbey Systems Inc. (Dolbey)	Concord, OH
SparkCognition	Austin, TX
Amazon Web Services Inc. (AWS)	Seattle, WA
Microsoft Corporation (Microsoft)	Readmond, WA

Notice the clear clustering within the state of California.

Thus, NLP technology development appears to be following the patterns predicted by innovation economics in regard to oligopolistic competition, activities, and behaviors.

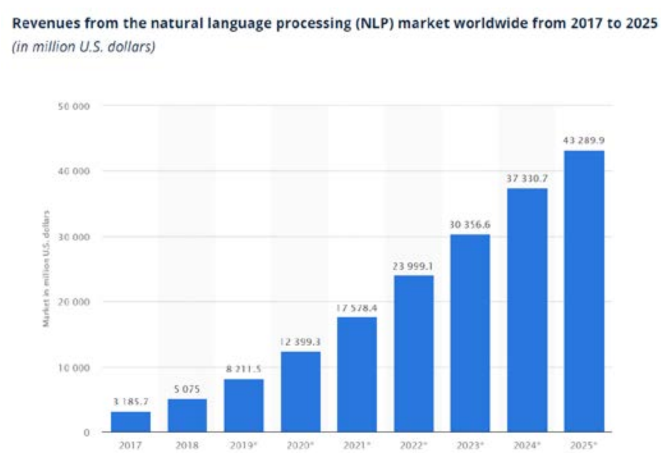
Entrepreneurs

Dr. Phelps (1933-), a professor at Columbia University, and Nobel Laureate in economics, said that "a nation's culture ultimately makes a difference for the nation's economic performance." Dynamism is necessary to encourage entrepreneurial types to initiate their activities (Phelps, 2007). Oligopolistic competition may play a major role in innovations (Baumol, 2002) and most innovation may come from existing firms improving their own products (Hsieh & Klenow, 2017) but as knowledge spillover occurs from the larger firms dynamism picks up and spreads technology through entrepreneurial activity (SBA report, 2002). Evidence suggested that knowledge spillovers might occur predominately between, rather than within, industries, consistent with the theories of Jane Jacobs (1916 - 2006) (Acs & Armington, 2003, Jacobs 1969). NLP entrepreneurialism is being encouraged through the many NLP information extraction tools being offered for public access such as OpenNLP (Java machine learning toolkit), Natural Language Toolkit (Suite of Python libraries for NLP), DBpedia Spotlight (Open source tool for Named Entity Recognition and Named Entity Linking tool), Open-Calais (Automated IE web service from Thomson Reuters) to name but a few. Additionally, commercial tools are also becoming available such as IBM Intelligent Minerand, SAS Text Analytics, and Business Objects (Singh, 2018). However, given the market is relatively new it is difficult to obtain processed information on actual NLP entrepreneurial activity but from limited searches most activity appears to be within the consulting area. This appears reasonable as the effective working of most software requires consulting services to install and increase the efficiency of the implemented processes (Market Analytics, 2019). The availability of NLP software in market provides ample opportunities for consulting services. This appears to bear out as one aggregator, Clutch's website (Clutch, 2020) demonstrates, of the AI service providers many have 10% to 60% of their focus in NLP technologies. Not only can consultants seek their own markets, but they might also partner with the larger firms. In this way, larger organizations can outsource to more entrepreneurial firms to allow for more responsiveness to changes in the environment (Aghion, Jones, & Jones, 2017) and potentially take advantage of adjacent possibilities (Hicks & Budiman, 2020) more quickly. Two examples of this are when, in 2019, Google began collaborating with Grid Dynamics, a provider of engineering and consulting services, to allow Grid to help its customers optimize supply chain management using NPL. And again in 2019, Health Fidelity began collaborating with Change Healthcare, an independent healthcare technology provider, to embedded NLP within medical servicing to increase claim accuracy and compliance (Kbv Research, 2019).

As David Warsh (1944-) suggested innovations and the new "sets of instructions" that arise coupled with the entrepreneurs that put them to use can be the key to growth (Warsh, 2006). Determining exactly how this concept develops within the NLP market is an area of rich opportunity for future research.

Economy

Worldwide revenue from the NLP market is forecast to increase rapidly in the next few years. As show in the graph below, the NLP market is predicted to be almost fourteen times larger in 2025 than it was in 2017, increasing from around three billion U.S. dollars in 2017 to over 43 billion in 2025 (Liu, 2020):

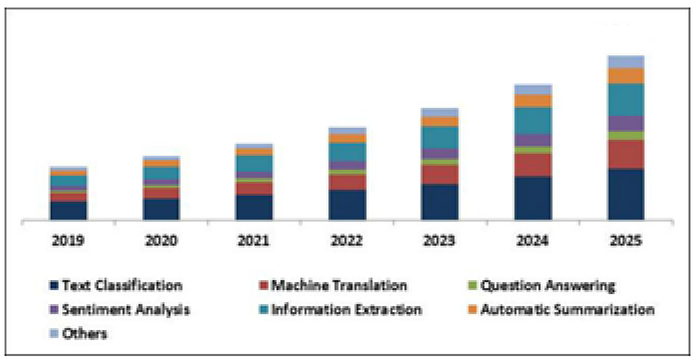


The global NLP market size is projected to have a Compound Annual Growth Rate (CAGR) of between 20% to 30% (Fortune Business Insights, 2020, VynZ Research, 2019, Kbv Research, 2019). Geographically, North America has accounted for the largest share in the NLP (VynZ Research, 2019) growth. Rapid developments in infrastructure and the high adoption of digital technologies are the two major drivers of the NLP market growth in that region (Market Analytics, 2019).

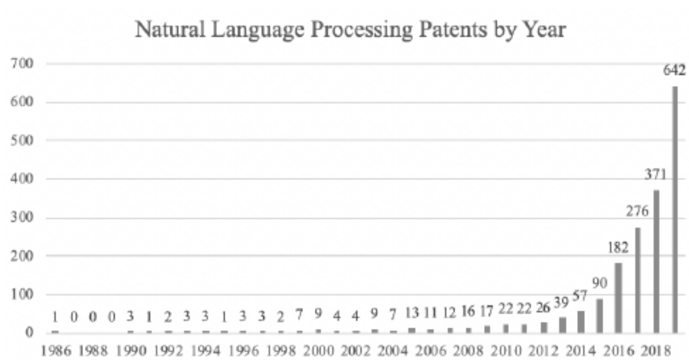
The NLP market is generally segmented as follows (Market Analytics, 2019):

- Banking, Financial Services and Insurance (BFSI)
- Manufacturing
- Healthcare and Life Sciences
- Retail and eCommerce
- Government and Defense
- Media and Entertainment
- Telecommunications and IT
- Travel and Hospitality
- Others (Research and Education, and Energy and Utilities)

Kbv Research, projects the growth by segment as follows (Kbv Research, 2019):



As noted above, patent levels may be understated, but the above growth levels are also observable in the patent data. The first NLP patent was issued to Tokyo Shibaura Denki Kabushiki Kaisha, in the year 1986. The graph below shows the significant growth since than (Haney, 2020):



These market projections and observations leave out the value of operational efficiencies and potential human capital improvements that might result from NLP implementation. Additionally, the value of consumer satisfaction from such items as improved news reading capabilities (Patankar & Bose, 2016) and life experiences are not, as of yet, being measured. Such measures as GDP and gross revenue, even if it were perfectly measured, does not quantify welfare. The trends in the official statistics tend to underestimate organizational and consumer bounty (Brynjolfsson & McAfee, 2014). These additional economic benefits from NLP implementation are an opportunity for further exploration and research.

Organizational Capital

Because economies of scale operate at the plant level, in the traditional Dr. Robert Solow (1924-) (Solow, 1956) model economic growth relied on capital investment in larger plants. However, capital accumulation can explain only a small amount of the variation in economic growth (Ciccone & Hall, 1996, Acs & Armington, 2003). Production in this new age depends less on physical equipment and structures and more on such things as organizational and human capital. Organizational capital is the intangible benefits achieved by such items as new business processes, production techniques, organizational forms, business models, etc. (Brynjolfsson & McAfee, 2014). Organizational changes and associated new skills are often slow to exhibit an impact. This concept is known as the Solow Paradox (Solow, 1987): we see transformative new technologies everywhere but in the productivity statistics. In the development of new technologies, such as NLP, there can be implementation and restructurings lags. Basically, it takes considerable time - often more than is commonly appreciated - to be able to sufficiently harness a new technology. There are two main sources of the delay between recognition of a new technology's potential and its measurable effects. One is that it takes time to build the stock of the new technology to a size sufficient enough to have an aggregate effect. The other is that complementary investments are necessary to obtain the full benefit of the new technology, and it takes time to discover and develop these complements and to implement them (Brynjolfsson, Rock & Syverson, 2017). Consistent with this theory, due to the relatively recent convergence of technologies making NLP possible, there is no indication that organizational capital in being enhanced by NLP. This is yet another area that appears ripe for further observation and research.

Human Capital / Labor

Human capital relates to the level of education and job skills developed by a labor workforce. This is the capital that humans acquire and make the labor force more productive. Automation does indeed substitute for labor as it is typically intended to do (Brynjolfsson & McAfee, 2014). According to a survey by Pew Research Center, 65 percent of US citizens expect that within 50 years a robot or an intelligent algorithm will be doing their work (Smith, 2019). An example of this substitution may be playing out with respect to NLP technology in the area of resume processing. NLP resume processing systems can potentially replace a large portions of a corporate human resource (HR) department's activity (Kopparapu, 2010, Kelkar, Shedbale, Khade, Pol & Damame, 2020). Beyond just parsing resumes for keywords, NLP is being used to judge the videos of job applicants. Consumer goods giant Unilever said that its system has delivered six-figure savings annually. Another area where NLP may replace labor is in credit analysis. NLP systems are routinely being used to assess creditworthiness, even without financial data (Splunk, 2020). However, the full extent of this labor replace aspect of NLP implementation has not yet been fully explored.

Automation also complements labor (Autor, 2015), creates new employment opportunities, and can also enhance human capital. NLP may complement the abilities of HR professionals by amplifying the comparative advantage of the professional in supplying problem-solving and creativity skills (Autor, 2015) when reviewing job applicates. NLP is creating new employment opportunities through such strategies as crowdsourcing the annotation of corpora or lexicons which is a core process to the development of the NLP field (Sabou, Bontcheva & Scharl, 2012). Finally, NLP might provide important enhancements in the area of education through improvements in the way educational courses are conducted.

All these evaluations are another source for future investigation as NLP becomes more fully implemented.

Public Policy

Many of the benefits from NLP have already been discussed. Items such as, increasing economic value, organizational enhancements, increasing employment opportunities, and consumer bounty. These benefits have been achieved within the U.S. because many of the governmental policies necessary to aid innovation are, for the most part, in place. These policies include: defining and enforcing property rights, maximizing freedom of exchange, consumption, and production, assuring quality data is available to market actors, preventing a small number of actors from eliminating competition (i.e. controlling monopolistic power), providing those public goods that private markets do not provide more efficiently, ensuring that investors, producers, and consumers each bear the costs, and capture the benefits, etc. (Taylor, 2016). Therefore, the remainder of the paper will not focus on policies designed to enhance progress but will review some of the potential risks from the continuing implementation of NLP.

Government planning during the development and implementation of disruptive technologies like NLP is always important. However, the potential pervasiveness of NLP and related intelligent machine technologies may make governmental planning more critical than with previous disruptive technologies. This is because the potential for intelligent computer programs to become self-developing resulting in exponential growth will mean that they may have far reaching impacts beyond the ability for societal control (Makridakis, 2017). Intelligent programs have moved beyond prototyping and into execution and implementation (Sander & Wolfgang, 2014), therefore, now is the time for serious contemplation. Below are a few of the more serious concerns brought about from NLP and related intelligent computer technologies, including concerns relating to employment, legal accountability, bias, privacy, and national competitiveness and security.

Employment

One of the largest concerns with the application of new technologies is its impact the employment of labor. As David Ricardo (1772-1823) wrote as far back as 1821, the "substitution of machinery for human labor, is often very injurious to the interests of the class of laborers" (Ricardo, 1821). Many fear that these advances will bring extreme Taylorism. Taylorism is a production efficiency methodology that breaks every action, job, or task into small and simple segments (BusinessDictionary.com, n.d.). Those that fear this outcome suggest that only the simplest tasks will be left to humanity if any task at all. About 47 per cent of total US employment is at risk from AI develops, read the catch line in the report by Frey/Osborne in 2013 (Frey & Osborne, 2017).

In general, although employment in certain industries has been reduced in the past due to technological advancements, the net effect of technological advancement has not appeared to lead to a reduction in long-term total employment (Atkinson, 2013). Clearly, the past two centuries of automation and technological progress have not made human labor obsolete: the employment-to-population ratio rose during the 20th century even as women moved from home to market; and although the unemployment rate fluctuates cyclically, there is no apparent long-run increase (David, 2015). Extrapolating from the impact of the industrial and digital revolutions it seems that technology has created more jobs than it has destroyed; although there may be a transitional period of increased unemployment until new opportunities are created to serve the emerging needs of those with increased incomes (Stewart, Debapratim, & Cole, 2015).

This raises two policy concerns. First, how to manage the public perceptions that occur during the implementation of a disruptive technology and, second, how to manage the periods of labor transition. Regarding the first issue, political leadership should begin developing communication plans to educate the public and better acclimate it to the potential for coming change. Regarding the second issues, policies should be designed which utilize social security or other such systems to help compensate and transition the labor force. Of course, here the financial pressure on social welfare systems will be the central problem. Progressive tax reform and distributive subsidies are one solution. Development of self-funded social security accounts, namely lifelong learning accounts, are another. Additionally, policy makers should review school curricula and, just as importantly, teacher training to better prepare future workers for the soft skills they will need as intelligent machine technology advances (EIU, 2018, Wisskirchen, et al. 2017).

Legal accountability

Intelligent programing so far has developed in a legal/regulatory vacuum. Virtually no courts appear to have developed standards specifically addressing who should be held legally responsible for harm caused. Ex ante regulation may be difficult because NLP development can have different components designed without conscious coordination, programmers widely dispersed geographically, and opaque coding such that outside observers may be unable to detect potentially harmful system features. Additionally, the autonomous nature of NLP neural net programs creates issues of foreseeability and control that might render ex post regulation ineffective (Scherer, 2015) as well. For example, a novel NLP system can help predict suicidal ideation during conversations or through unstructured clinical notes (Cook, 2016). Who becomes responsible if the system fails to identify a suicidal patient that ultimately becomes successful? Some steps along solving these problems are developing. For example, the US Defense Advanced Research Projects Agency (DARPA) launched the eXplainable AI (XAI) challenge in 2016. This project is designed to encourage the development of systems to help explain how neural net systems, such as those used in NLP applications, produce the results they produce. In 2018, the European General Data Protection Regulation (GDPR) became effective. The GDPR gives the right to European citizens to an explanation regarding decisions taken by automatic intelligent systems (Alonso & Bugarín, 2019). In 2020 the Department of Defense released a document called "Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense" (Maddipudi, 2020) to help develop standards in AI development. While these are good beginnings more focus is needed in this area.

Bias

For NLP systems "Training is everything," says Eric Sammer, distinguished engineer at Splunk, a NASDAQ traded machine learning software company. "A lot of these algorithms are being trained on existing human practices that are inherently biased and problematic. It'd be naive to assume we can eliminate that from NLP algorithms at the outset" (Splunk, 2020). NLP is being used to screen job applicants, grade college entrance exams, asses creditworthiness, provide information to police and the justice systems. The training of these systems comes from existing datasets which can contain existing human bias against certain demographic groups and, therefore, produce undesirably unfair results (Boddington, 2017). For example, programs used in criminal justice that influenced sentencing decisions were found to be twice as likely to wrongly flag black defendants as future criminals as they were to inaccurately label white defendants (Splunk, 2020). Policy makers and individual developers should monitor this problem and develop appropriate responses.

Privacy

Some may argue that privacy is a thing of past generations, others disagree. The U.S. has adopted a commercial data integrity and breach reporting philosophy to privacy laws whereas the European Union, with GDPR, has taken privacy as an individual right philosophy. Both approaches have their advantages and disadvantages. However, with ever-increasing personal data being generated by commercial enterprises and by individuals sua sponte, privacy is likely to continue to be a matter of public concern. This will become even more of an issue as NLP begins to process the enormous amount of dark data (i.e. textual data not previously reviewable such as emails and text messaging) generated within corporations and be able to provide management with direct access to employee communications.

National Competitiveness Concerns

While the U.S. is likely to remain a world competitor, NLP technology, and AI more broadly, may develop to become the country's most important resource and largest asset. This could lead to what is known as the Resource Curse. This problem refers to countries where natural resources comprise a large portion of the economy. In this type of situation countries tend to be more unstable than countries with more diversified economies (Stevens, Lahn, & Kooroshy, 2015). Perhaps this a possibility only in the U.S.'s advanced future but policy makers should be on guard for its possibility.

National Security Concerns

National security is likely to become an important state issue as NLP and AI develops. Advances in NLP and AI will affect many issues of military security such as the ability to analyze governmental communications and mine important strategic information. These capabilities will become more affordable and available to a broader range of actors. This can give weak states and non-state actors access to a type of propaganda vehicle that can become increasingly indistinguishable from the truth. By manipulating data, actors can throw off or break down entire learning models, hijacking them or rendering them useless. Adversarial machine learning, the technique employed to fool models through malicious input, has tricked Google's AI into thinking a turtle was a rifle. Funny, at least outside a security checkpoint. Policy makers should expect also to see attempts to poison algorithm with specious data samples specifically designed to throw off the learning process of a machine learning algorithm. It is not just about duping smart technology but making it so that the algorithm appears to work fine — while producing the wrong results (Splunk, 2020). Population size will become less important for national power. Small countries and non-state actors that develop these technology capabilities will punch far above their weight. Governments should take these risks seriously and develop appropriate responses.

Conclusion

This paper has demonstrated how NLP has developed along the lines predicted by innovation economics. Originating from the work of a few individuals which was carried forward through the university systems. As the technology progressed the government began promoting its concepts. Finally, propelled by converging technologies, market forces pushed and pulled the technology forward. Also as predicted by innovation economics a few large companies are dominating the development of the technology while entrepreneurs are taking advantage of the spillover opportunities. The technology has begun to have an impact on the economy, however, once again as predicted by innovation economics, the newness of the technology has not resulted in materially observable changes to organizational or labor practices. Giving clarity to these last potential changes is an area that is ripe for further research. Finally, like all new technologies, NLP has risks which should be analyzed and mitigated by policy makers.

References:

Acs, Z., & Armington, C. (2003). Endogenous growth and entrepreneurial activities in cities. US Bureau of the Census.

Aghion, P., Jones, B. F., & Jones, C. I. (2017). Artificial intelligence and economic growth (No. w23928). National Bureau of Economic Research.

Allen, G., & Chan, T. (2017). Artificial intelligence and national security. Cambridge, MA: Belfer Center for Science and International Affairs.

Alonso, J. M., & Bugarín, A. (2019). ExpliClas: automatic generation of explanations in natural language for WEKA classifiers. In 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (pp. 1-6). IEEE.

Arora, C., Sabetzadeh, M., Briand, L., & Zimmer, F. (2015). Automated checking of conformance to requirements templates using natural language processing. IEEE transactions on Software Engineering, 41(10), 944-968.

Atkinson, R. D. (2013). Stop Saying Robots Are Destroying Jobs–They Aren't.

Atkinson, R. D., & Wu, J. (2017). False alarmism: Technological disruption and the US labor market, 1850–2015.@ Work Series.

Autor, David, H. J. J. O. E. P. (2015). Why are there still so many jobs? The history and future of workplace automation. Journal of economic perspectives, 29(3), 3-30.

Baslandze, Salome, (2016). The Role of the IT Revolution in Knowledge Diffusion, Innovation and Reallocation, EIEF manuscript.

Baumol, W. J. (2002). *The free-market innovation machine: Analyzing the growth miracle of capitalism*. Princeton university press.

Block, F., & Keller, M. R. (2009). Where do innovations come from? Transformations in the US economy, 1970–2006. Socio-Economic Review, 7(3), 459-483.

Boddington, P. (2017). Towards a code of ethics for artificial intelligence. Springer International Publishing.

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. WW Norton & Company.

Brynjolfsson, E., Rock, D., & Syverson, C. (2017). Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics (No. w24001). National Bureau of Economic Research.

BusinessDictionary.com. (n.d.). Taylorism. Retrieved October 15, 2020, from https://www.emerald.com/insight/search?q=Taylorism

Chen, N., Christensen, L., Gallagher, K., Mate, R., & Rafert, G. (2016). Global economic impacts associated with artificial intelligence. Analysis Group.

Chomsky, N. (2002). Syntactic structures. Walter de Gruyter.

Church, K. (2007). A pendulum swung too far. Linguistic Issues in Language Technology, 2(4).

Ciccone, C. and R.E. Hall, (1996), “Productivity and the Density of Economic Activity,” American Economic Review, 86(1), 54-70.

Clutch. (2020). Top NLP Developers - Leaders Matrix. Retrieved November 30, 2020, from https://clutch.co/developers/artificial-intelligence/natural-language-processing/leaders-matrix.

Conference on Neural Information Processing Systems. (2020). Retrieved November 25, 2020, from

Omni sci. (2020). CPU vs GPU: Definition and FAQs. Retrieved November 26, 2020, from https://www.omnisci.com/technical-glossary/cpu-vs-gpu

https://en.wikipedia.org/wiki/Conference_on_Neural_Information_Processing_Systems.

Cook, B. L., Progovac, A. M., Chen, P., Mullin, B., Hou, S., & Baca-Garcia, E. (2016). Novel use of natural language processing (NLP) to predict suicidal ideation and psychiatric symptoms in a text-based mental health intervention in Madrid. Computational and mathematical methods in medicine, 2016.. https://www.hindawi.com/journals/cmmm/2016/8708434/.

Cornillie, C. (2019). Finding Artificial Intelligence Money in the Fiscal 2020 Budget: BGOV. Retrieved November 27, 2020, from https://about.bgov.com/news/finding-artificial-intelligence-money-fiscal-2020-budget/.

David, H. J. J. O. E. P. (2015). Why are there still so many jobs? The history and future of workplace automation. Journal of economic perspectives, 29(3), 3-30.

Delangue, C., & Nadella, S. (2020, April 06). Sponsored post: How the revolution of natural language processing is changing the way companies understand text. Retrieved December 03, 2020, from https://techcrunch.com/sponsor/nvidia/how-the-revolution-of-natural-language-processing-is-changing-the-way-companies-understand-text/?guccounter=1

Deng, L., & Liu, Y. (2018). Deep Learning in Natural Language Processing. In Deep Learning in Natural Language Processing. Springer Singapore Pte. Limited https://doi.org/10.1007/978-981-10-5209-5.

Eggers, W. D., Malik, N., & Gracie, M. (2018). Using AI to unleash the power of unstructured government data. Retrieved November 26, 2020, from https://www2.deloitte.com/content/dam/Deloitte/lu/Documents/public-sector/lu-ai-unstructured-government-data.pdf.

EIU, (2018). Who is Ready for the Coming Wave of Automation? The Economist Intelligence Unit, Retrieved May 6, 2020, from http://www.automationreadiness.eiu.com/whitepaper.

Exec. Order No. 13859, 3 C.F.R. (2019).

Fernald, J. G., & Jones, C. I. (2014). The future of US economic growth. American economic review, 104(5), 44-49.

Foote, K. (2019). A Brief History of Natural Language Processing (NLP). Retrieved September 29, 2020, from https://www.dataversity.net/a-brief-history-of-natural-language-processing-nlp/.

Fortune Business Insights. (2020). NLP Market Size, Share: Natural Language Processing Market Report 2026. Retrieved November 27, 2020, from https://www.fortunebusinessinsights.com/industry-reports/natural-language-processing-nlp-market-101933.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation?. Technological forecasting and social change, 114, 254-280.

Gartner_Inc. (2020). Definition of Hype Cycle - Gartner Information Technology Glossary. Retrieved November 28, 2020, from https://www.gartner.com/en/information-technology/glossary/hype-cycle.

Haney, B. S. (2020). Patents for NLP Software: An Empirical Review. Available at SSRN 3594515.

Hicks, D.A. (2020), Technological Changes - Spontaneous & Strategic Sources [PowerPoint slides]. dahicks@utdallas.edu.

Hicks, D.A.& S. Budiman (2020), Innovation Capital Timescapes: Investment-Mix “Churn” and the ICT “Adjacent Possible,” Draft unpublished manuscript.

Hsieh, C. T., & Klenow, P. J. (2017). The reallocation myth. Fostering a Dynamic Global Economy. https://ieeexplore.ieee.org/abstract/document/618923.

Jacobs, J., (1969), The Economy of Cities, New York: Vintage.

Johnson, S. (2010). The genius of the tinkerer. Wall Street Journal, 25.

Kbv Research. (2019). Natural Language Processing Market Size, Share & Forecast 2025. Retrieved November 25, 2020, from https://www.kbvresearch.com/natural-language-processing-market/.

Kelkar, B., Shedbale, R., Khade, D., Pol, P., & Damame, A. (2020). Resume Analyzer Using Text Processing”. Journal of Engineering Sciences, 353-361. https://jespublication.com/upload/2020-110557.pdf.

Kopparapu, S. K. (2010). Automatic extraction of usable information from unstructured resumes to aid search. In 2010 IEEE International Conference on Progress in Informatics and Computing (Vol. 1, pp. 99-103). IEEE.

Lischka, K. (2011). John McCarthy: Der Vater der Rechner-Cloud ist tot - DER SPIEGEL - Netzwelt. Retrieved May 5, 2020, from https://www.spiegel.de/netzwelt/web/john-mccarthy-der-vater-der-rechner-cloud-ist-tot-a-793795.html.

Liu, S. (2020). Global natural language processing market 2017-2025. Retrieved November 26, 2020, from https://www.statista.com/statistics/607891/worldwide-natural-language-processing-market-revenues/.

Maddipudi, T. (2020). Natural Language Processing- Business use cases. Retrieved November 26, 2020, from https://medium.com/byteridge/natural-language-processing-business-use-cases-512debac8151.

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. Futures, 90, 46-60.

Market Analytics (2019). Natural Language Processing Market. Retrieved November 25, 2020, from https://www.marketsandmarkets.com/Market-Reports/natural-language-processing-nlp-825.html.

Métais, E., Meziane, F., Saraee, M., Sugumaran, V., & Vadera, S. (Eds.). (2016). Natural Language Processing and Information Systems: 21st International Conference on Applications of Natural Language to Information Systems, NLDB 2016, Salford, UK, June 22-24, 2016, Proceedings (Vol. 9612). Springer.

O'Neil, Paul and Paik, Woojin (1998). The Chronological Information Extraction System (CHESS), IEEE.

Patankar, A. A., & Bose, J. (2016). Bias Based Navigation for News Articles and Media. In International Conference on Applications of Natural Language to Information Systems (pp. 465-470). Springer, Cham.

Phelps, E. S. (2007) Entrepreneurial culture: why European economies lag behind the US Wall Street Journal. 2007. Opinion Journal, 4.

Ricardo, David. (1821) "On the Principles of Political Economy and Taxation" John Murray, 3rd. ed., Chapter 31.3.

Romer, P. (1986), 'Increasing returns and economic growth'. American Economic Review 94, 1002–1037.

Ruliputra, R. N., Sucahyo, Y. G., Gandhi, A., & Ruldeviyani, Y. (2019). Why do Enterprises Adopt Natural Language Processing Services? Startups' Landscape and Opportunities in Artificial Intelligence. In 2019 International Conference on Advanced Computer Science and information Systems (ICACSIS) (pp. 313-318). IEEE. https://ieeexplore.ieee.org/abstract/document/8979806.

Sabou, M., Bontcheva, K., & Scharl, A. (2012). Crowdsourcing research opportunities: lessons from natural language processing. In Proceedings of the 12th International Conference on Knowledge Management and Knowledge Technologies (pp. 1-8) https://dl.acm.org/doi/abs/10.1145/2362456.2362479?casa_token=IzliSb5HBwsAAAAA:OitqtIkOlhotr62J32PFcZtV93iurG1HaKgfPP7NbWZRglwUz1Xvv6MruYQ-MssoYsuaJNM1P7s.

Sander, A., & Wolfgang, M. (2014). The rise of robotics. Bcg perspectives.

SBA (Small Business Association). (2002). The Influence of R&D Expenditures on New Firm Formation and Economic Growth. BJK Associates, Maplewood, NJ.

Scherer, M. U. (2015). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. Harv. JL & Tech., 29, 353.

Singh, S. (2018). Natural language processing for information extraction. arXiv preprint arXiv:1807.02383.https://arxiv.org/pdf/1807.02383.pdf.

Smith, A. (2019). Future of Workforce Automation: America's Predictions. Retrieved May 5, 2020, from https://www.pewresearch.org/internet/2016/03/10/public-predictions-for-the-future-of-workforce-automation/.

Solow, R.M., (1956), "A Contribution to the Theory of Economic Growth", Quarterly Journal of Economics 94, 614-623.

Solow, Robert. (1987). "We'd Better Watch Out." New York Times Book Review, July 12: 36.

Splunk. (2020). Emerging Tech Predictions 2020. Retrieved September 29, 2020, from https://www.splunk.com/pdfs/ebooks/emerging-tech-predictions-2020.pdf.

Stevens, P., Lahn, G., & Kooroshy, J. (2015). The resource curse revisited. Chatham House for the Royal Institute of International Affairs.

Stewart, I., Debapratim, D., & Cole, A. (2015). Deloitte. https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/finance/deloitte-uk-technology-and-people.pdf.

Sun, S., Luo, C., & Chen, J. (2017). A review of natural language processing techniques for opinion mining systems. Information fusion, 36, 10-25. https://www.sciencedirect.com/science/article/pii/S1566253516301117?casa_token=DgeSse_bSZwAAAAA:_HB1CNpXF4YFP97M15FmdBzpn4muqQFuv4ROneMoMgdnZyTWqciA5DKeZ3AiO-a1nT_SkFJ_.

Taylor, M. Z. (2016). The politics of innovation: Why some countries are better than others at science and technology. Oxford University Press.

United States. (2019). Artificial Intelligence for the American People. Retrieved November 26, 2020, from https://www.whitehouse.gov/ai/executive-order-ai/.

VynZ Research. (2019). Market Research Company Focuses On Providing Valuable Insights On Various Industries. Retrieved November 26, 2020, from https://www.vynzresearch.com/ict-media/natural-language-processing-nlp-market.

Warsh, D. (2006). Knowledge and the wealth of nations: A story of economic discovery. WW Norton & Company.

Wikipedia. (2020, November 24). Natural language processing. Retrieved October 28, 2020, from https://en.wikipedia.org/wiki/Natural_language_processing

Wisskirchen, G., Biacabe, B. T., Bormann, U., Muntz, A., Niehaus, G., Soler, G. J., & von Brauchitsch, B. (2017). Artificial intelligence and robotics and their impact on the workplace. IBA Global Employment Institute, 2012-2017.

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. ieee Computational intelligenCe magazine, 13(3), 55-75.

Zhao, F. (2017). Natural Language Processing – Part I: Primer, Unveiling the Hidden Information in Earnings Calls. Https://www.spglobal.com/marketintelligence/en/documents/sp-global-market-intelligence-nlp-primer-september-2018.pdf.