NATURAL LANGUAGE PROCESSING

ENABLING THE POTENTIAL OF A DIGITAL HEALTHCARE ERA

MARKET SCAN REPORT
ABOUT CHILMARK RESEARCH

Chilmark Research is a global research and advisory firm whose sole focus is the market for healthcare IT solutions. This focus allows us to provide our clients with the most in-depth and accurate research on the critical technology and adoption trends occurring throughout the healthcare sector. Areas of current research focus include among others: Clinician Network Management, Cloud-computing Models for Healthcare, IT-enabled Accountable Care Organizations, Care Coordination, Adoption of Mobile Technology and Consumer-facing Health and Wellness Applications and Services.

Using a pragmatic, evidence-based research methodology with a strong emphasis on primary research, Chilmark Research structures its research reports to serve the needs of technology adopters, consultants, investors and technology vendors. In addition to reports for the general market, Chilmark Research performs research for clients based on their specific needs. Such research has included competitive analyses, market opportunity assessments, strategic assessment of market and vendors for partnership and/or acquisition.

In 2012, Chilmark Research launched its newest service, the Chilmark Advisory Service (CAS). The CAS was created in direct response to clients’ request for a continuous feed of research on the most pertinent trends in the adoption and use of healthcare IT. This is an annual subscription that provides not only access to a number of re-search reports throughout the year, but also direct access to Chilmark Research analysts to answer specific client needs. Please contact us directly for further information about CAS.

Chilmark Research is proud of the clients it has had the pleasure to serve including Abbott Labs, Bluetooth Special Interest Group, Catholic Healthcare East, Cerner, HCA, Highmark, IBM, Kaiser-Permanente, McKesson, McKinsey, Microsoft, and Thomson Reuters to name a few. It is our hope that at some future date we will have the pleasure to serve you as well.

Chilmark Research LLC
One Beacon Street
Boston, MA 02108
www.ChilmarkResearch.com
info@chilmarkresearch.com
Ph. 617.615.9344

The information in this report is proprietary to and copyrighted by Chilmark Research. No part of this report may be reproduced or distributed without prior permission of Chilmark Research. The information contained within the report is not intended as a solicitation of an offer to buy or sell any investment or other specific product. All information and opinions expressed in this report were obtained from sources believed to be reliable and in good faith. No representations or warranty expressed or implied is made as to its accuracy or completeness. Trademarked and service marked names appear throughout this report. Rather than use a trademark or service mark symbol with every occurrence, names are used in an editorial fashion, with no intention of infringement of the respective owner’s trademark or service mark.
# TABLE OF CONTENTS

## EXECUTIVE SUMMARY

9

## MARKET DYNAMICS

11

- Drivers for NLP in Healthcare ................................................................. 12
- An Evolving Market .................................................................................. 13
- NLP's Contribution to Intelligent Systems .................................................. 14
- Measuring NLP Performance ..................................................................... 14
  - NLP ROI ................................................................................................... 15

## TYPES OF NLP SYSTEMS AND METHODOLOGIES

17

- Rule-based Systems .................................................................................... 18
- Machine Learning and Statistics-based Systems ........................................... 18
  - Supervised Learning .................................................................................. 19
  - Gold Standard Corpus (GSC) .................................................................... 19
  - How much of the training data should be withheld for testing? ...................... 20
  - Are open source gold standard datasets available? ........................................ 20
- Unsupervised Learning ................................................................................. 20
- Deep Learning .............................................................................................. 20
  - Transfer Learning ...................................................................................... 21
- Hybrid Systems ............................................................................................ 21

## IMPLEMENTATION CHALLENGES OF NLP SOLUTIONS

22

## CUSTOMER TYPES AND END USERS

24

## PRIMARY NLP USE CASES IN HEALTHCARE

25

- Proven (Mainstay) NLP Healthcare Use Cases ............................................. 26
  - Speech Recognition ................................................................................... 27
  - Clinical Documentation Improvement (CDI) ............................................... 27
  - Data Mining Research ............................................................................... 27
  - Computer-Assisted Coding ....................................................................... 28
Emerging NLP Healthcare Use Cases ................................................................. 29
  Clinical Trial Matching .................................................................................. 30
  Prior Authorization ......................................................................................... 30
  Clinical Decision Support .............................................................................. 30
  Risk Adjustment and Hierarchical Condition Categories ............................. 31
Next-Generation NLP Healthcare Use Cases ................................................ 32
  Ambient Virtual Scribe .................................................................................. 33
  Computational Phenotyping and Biomarker Discovery ............................... 33
  Population Surveillance .................................................................................. 33
Notable Use Case – Chatbots .......................................................................... 34

BARRIERS TO ADOPTION 35

Human Challenges ......................................................................................... 35
  Skills Gap ...................................................................................................... 35
  System Usability ............................................................................................ 35
  Cyber Attacks ................................................................................................ 35
Data Challenges ............................................................................................. 35
  Quality, Completeness, and Bias ................................................................. 35
  Metadata (or lack thereof) ............................................................................ 36
  Adversarial Data ............................................................................................ 36
  Data Warehouse → Data Lake → Data Operating System ....................... 36

NLP VENDOR LANDSCAPE 37

NLP Vendor Use Case Analysis ..................................................................... 37

NLP VENDOR PROFILES 39

  3M ................................................................................................................. 39
  Artificial Intelligence in Medicine (an Inspirata Company) ......................... 40
  Clinithink ....................................................................................................... 41
  Digital Reasoning .......................................................................................... 41
  Health Catalyst .............................................................................................. 44
  Health Fidelity ............................................................................................... 45
  IBM Watson Health ........................................................................................ 46

NLP: ENABLING THE POTENTIAL OF A DIGITAL HEALTHCARE ERA
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>JULY 2018</th>
</tr>
</thead>
</table>

Linguamatics Health.................................................................................................................. 47
M*Modal................................................................................................................................ 48
Nuance ..................................................................................................................................... 49
Optum ........................................................................................................................................ 50
SyTrue ....................................................................................................................................... 51

FUTURE NLP MARKET PLAYERS .......................................................................................... 52

Cloud Computing Vendors and Technology-Enabled Services........................................... 52
  Alphabet – Google & Verily................................................................................................. 52
  Microsoft .............................................................................................................................. 53
  Amazon .................................................................................................................................. 53

EHR Vendors ......................................................................................................................... 53

CONCLUSIONS AND RECOMMENDATIONS ......................................................................... 55

APPENDIX A: SCOPE AND METHODOLOGY .................................................................. 56

APPENDIX B: MEASURING NLP PERFORMANCE .......................................................... 57
  Auto Summarization............................................................................................................. 57
  Automatic Speech Recognition............................................................................................ 58
  Reading Comprehension ...................................................................................................... 58
  Transcribing Conversations ................................................................................................. 58

APPENDIX C: ACRONYMS USED IN THIS REPORT ...................................................... 59

ENDNOTES .......................................................................................................................... 60
TABLES AND FIGURES

EXECUTIVE SUMMARY

MARKET DYNAMICS

- Figure 1: Ambient Virtual Assistant adoption rate relative to previous consumer technology paradigms
- Figure 2: Projected growth in Alexa Skills
- Table 1: Current State of Big Data in Healthcare vs. the Emerging Paradigm
- Figure 3: The Three Phases of AI Integration in Healthcare
- Figure 4: Formulas for Calculation Common NLP Performance Metrics

TYPES OF NLP SYSTEMS AND METHODOLOGIES

- Table 2: Advantages and Challenges of NLP Methods
- Figure 5: Basic processes required for implementing an NLP system
- Figure 6: The Three Levels of Linguistic Analysis
- Table 3: Syntax-related data preprocessing tasks

IMPLEMENTATION CHALLENGES OF NLP SOLUTIONS

- Figure 7: Example of preprocessed NLP
- Table 4: Customer Types and End Users

CUSTOMER TYPES AND END USERS

PRIMARY NLP USE CASES IN HEALTHCARE

- Figure 8: Current Standard NLP Use Cases in Healthcare and Key Vendors
- Figure 9: Emerging Use Cases for NLP in Healthcare and Key Vendors
- Figure 10: Next-Generation Use Cases of NLP in Healthcare and Key Vendors

BARRIERS TO ADOPTION

- Table 5: Comparison of Batch and Streaming Data Processing
NLP VENDOR LANDSCAPE .................................................. 37
  Table 6: NLP Vendors and Use Cases Supported ............................................... 38

NLP VENDOR PROFILES .......................................................... 39

FUTURE NLP MARKET PLAYERS .................................................. 52

CONCLUSIONS AND RECOMMENDATIONS ........................................... 55

APPENDIX A: SCOPE AND METHODOLOGY ......................................... 56

APPENDIX B: MEASURING NLP PERFORMANCE ....................................... 57
  Table 7: The Confusion Matrix: A Standard Method of Visualizing NLP Performance .................................................. 57
  Figure 11: Algorithms That Outperform Humans at Reading .......................... 57
  Figure 12: Speech Recognition vs. Human on Switchboard HUB500 Dataset 58

APPENDIX C: ACRONYMS USED IN THIS REPORT .................................... 59

ENDNOTES ................................................................................. 60
Executive Summary

The great struggle to digitize the business of healthcare and practice of medicine is over; however, the war to wrangle and analyze the data collected will rage on for years. In the rush to digitize and not unreasonably disrupt established clinical workflows and documentation practices, as much as 80 percent of the data captured by IT systems is unstructured, and the text is often poor quality. In its current format and given the high cost in human time and effort it would take to read, the extensive library of health data is effectively unusable. Thus, the quest to drive better healthcare decision making and analytics remains an unfulfilled promise.

Yet the need to leverage unstructured data is growing in importance as the business model for reimbursement of care shifts from fee-for-service to value-based care (VBC). Natural Language Processing (NLP), a subcategory of artificial intelligence (AI), has the ability to augment and automate human behaviors and skills. It can be used to parse and abstract key information from a variety of sources, such as clinician notes, thereby unlocking unstructured data and helping ease payers and providers through the transition to a new reimbursement model. This augmentation - leading to automation of such mundane tasks as quality reporting and creating patient registries among others - has the potential to save healthcare organizations (HCOs) significant money and time.

The current market for NLP technology in healthcare is nascent, dominated by a few legacy vendors that are focusing on front-end speech recognition (for computer assisted physician documentation) and back-end coding (to optimize billing). While a number of tech giants continue to advance NLP technology for more general use, there are only a handful of niche solutions from highly specialized healthcare vendors pursuing additional use cases. We have outlined several of these vendors in this report. Many academic institutions are also developing their own solutions, often using open-source software. Such academic solutions offer the potential that joint ventures with vendors could lead to commercialization and potentially broader acceptance of NLP.

The best-performing NLP engines are able to combine the precision of traditional rule-based methods with advanced machine learning methods. State-of-the-art NLP systems also exploit the latest in deep learning methods – they excel by utilizing painstakingly developed gold-standard (i.e., expert-annotated) training datasets to learn how to classify new cases based on the accumulated knowledge of all historical cases.

This report describes a dozen significant NLP healthcare use cases, including computer assisted coding, speech recognition, and data mining. Five of these solutions have proven ROI and are commercially available from numerous well-established vendors. Another four are going through the initial phase of the adoption cycle and are primed to have an immediate impact under the new value-based care paradigm. These solutions focus on identifying the highest-cost patients earlier, tracking basic quality metrics related to annual follow-up, and reducing readmissions. The last group of solutions will mature over the next three-plus years and includes computational phenotyping for precision medicine, ambient virtual scribes to improve the electronic health record (EHR) user experience, and digital biomarker discovery using advanced voice-based diagnostic techniques.

There are many challenges to developing sophisticated NLP applications; these include the complexity of natural language, multiple technology approaches, and choice of metrics to measure success. Implementing NLP brings yet more challenges, including hiring people with specialized skills and achieving data liquidity. While these challenges are not insurmountable, they require a full appreciation of what it will take to succeed.

This report provides insights and advice for HCOs that are considering, implementing, and deploying NLP technologies. Following months of market research and interviews with healthcare providers and NLP vendors, we have compiled advice regarding newer technologies, proven methods, and the most impactful use cases. We also identify and analyse key performance metrics that users should expect to see cited by NLP vendors as key differentiators. This report concludes with profiles and analysis of a dozen significant NLP-focused vendors that Chilmark Research considers representative of the stack of technologies, development platforms, use cases, and services.
Market Dynamics

Natural Language Processing (NLP) is a subfield of computer science, computational linguistics, artificial intelligence (AI) and machine learning (ML). It has several sub-disciplines, including Natural Language Understanding (NLU), Natural Language Generation (NLG), and Natural Language Query (NLQ). NLP can be defined as the automatic processing of human natural language.

NLP has two large, overarching and related use cases:

1. Understanding human speech and extracting meaning.

2. Unlocking unstructured data in documents and databases by abstracting out key concepts and values and making this information available for decision support and analytics.

Nearly everyone uses NLP technologies every day without even realizing it. Spellcheck is the original NLP feature and one of its most common use cases. Voice assistants such as Apple’s Siri are routinely used to improve a user’s smartphone experience. Search engine use of NLP is also commonplace. In 2016, Google changed the algorithm powering its search engine from its original keyword-driven algorithm to a natural language query approach. Google Translate also uses NLP for machine translation to present any webpage in the native language of the browser.

Figure 1: Ambient Virtual Assistant adoption rate relative to previous consumer technology paradigms
The most significant advancement in consumer adoption of NLP is the rapid adoption of ambient voice assistants built into smart speakers. Consumers have already filled their homes and lives with this newest category of consumer device faster than any other in history – a few years rather than a few decades. (See Figure 1.) Since the beginning of 2017, 400 million Android users were introduced to Google Assistant. Amazon’s Alexa, with 39 million such devices sold in the US as of the end of 2017, is on track to reach 55-percent of households by 2022. Since 2017, the number of Alexa Skills went from 7,000 to 40,000.

We project Alexa’s Skills to start growing exponentially as Amazon continues offering tools that make it easier for non-technical users to create their own customized skills. (See Figure 2.)

**DRivers for NLP in Healthcare**

The long-promised digitization of the business of healthcare and practice of medicine is nearly complete. This major accomplishment in the quest to modernize the industry has given rise to a population-scale dataset with unprecedented potential to generate insights that provide opportunities to improve quality of care and decrease costs. The next stage of this digitization journey will require advances in the two key areas of NLP, speech interfaces and analyzing unstructured data.

Three main imperatives are driving the need for healthcare-specific NLP:

1. Supporting the needs of value-based care (VBC) and population health management (PHM).
2. Coding and analyzing encounters more effectively.
3. Decreasing physician workload and burnout.

**Support for VBC and PHM.** The shift in business models and outcomes is driving the need for better use of unstructured data. Legacy health information systems have focused almost exclusively on extracting value from the 20 percent of clinical data that is captured in structured, computable formats. Next-generation management care and PHM applications, patient engagement portals, and advanced predictive/prescriptive analytics algorithms will be limited in their impact unless they can also tap into the information locked up in the 80 percent of clinical data that is unstructured. NLP offers a viable technical solution to this problem.

**Coding.** Because payers could arbitrarily approve or deny claims without auditing the EHR, this has led to widespread deficiencies in clinical documentation practices. Clinical documentation was traditionally used to inform billing/coding teams who would in turn use their specialized knowledge to choose the code with the highest reimbursement rate. With the shift from patient-level to population-level metrics for determining reimbursement demands, coders have had to shift their workflow from accurately documenting procedures and therapies to accurately documenting all diagnoses and related complications/comorbidities.

**Physician Workload and Burnout.** One of the greatest complaints of physicians today is the soul-crushing user experience of current-generation EHRs. Research has determined that for every hour spent in direct contact with patients, physicians now spend two hours doing EHR-based clerical work. As a direct result, physician burnout...
Appendix A: Scope and Methodology

For this report, Chilmark Research identified more than 250 vendors that claimed NLP capabilities in healthcare. From this long list, Chilmark combined secondary and primary research methodologies to identify more than two dozen healthcare NLP companies that had a level of “market mind-share.” These meetings and calls were typically 30 to 60 minutes including discussions of market drivers, technology, use cases, clients, adoption, pricing, and competition.

Chilmark also had discussions with many academic medical centers, consultants, and some government entities involved in reporting registries and NLP research. These interviews provided further perspectives on NLP, core use cases, vendors of note, and specific vendor experiences.

From this work, a subset of a dozen NLP companies were chosen based on multiple factors, including their experience/ traction in healthcare, the breadth (or uniqueness) of their solution, their technology approach, and their results. Vendors all had an opportunity to review their profile and provide comments, which were considered and, where relevant, incorporated into the profile. In the end, the final vendor assessments are from the Chilmark Research analysts who authored this report. Only one vendor (Optum) opted to not provide feedback/review of their profile.

In compiling this extensive report, Chilmark Research maintained absolute objectivity throughout the entire research process, and it is our sincere hope that this report brings greater clarity to this rapidly developing and potentially high-value market.
Appendix B: Measuring NLP Performance

Developing useful and powerful NLP applications is all about optimizing methods to improve performance using annotated training datasets. The performance of NLP systems can be crucial in determining whether to apply those systems alongside humans or replace humans and calculating ROI. The two most important measures are precision and recall. The F1 score is calculated by averaging the precision and recall and represents the . The goal of any NLP system is to achieve optimal performance on both measures. (See Table 7.)

<table>
<thead>
<tr>
<th>Actual</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pred.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>True Pos.</td>
<td>TRUE POSITIVE</td>
<td>FALSE POSITIVE</td>
</tr>
<tr>
<td>True Neg.</td>
<td>FALSE NEGATIVE</td>
<td>TRUE NEGATIVE</td>
</tr>
</tbody>
</table>

Table 7: The Confusion Matrix: A Standard Method of Visualizing NLP Performance

Examples of metrics for different types of NLP processing include the following:

Auto Summarization

When comparing auto-summarization applications to human performance, BLEU and ROUGE scores are variations of precision and recall respectively.

> BLEU scores measure how often the words that appeared in the machine-generated summary appeared in the human-reference summaries.

> ROUGE scores measure how often the words that appeared in the human-reference summaries appeared in the machine-generated summary.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Model</th>
<th>EM</th>
<th>F1</th>
</tr>
</thead>
</table>
| 1    | Hybrid AoA Reader (ensemble)  
Jan 22, 2018  
Joint Laboratory of HIT and iFLYTEK Research | 82.482 | 89.281 |
| 1    | QANet (ensemble)  
Mar 06, 2018  
Google Brain CMU | 82.744 | 89.045 |
| 1    | Reinforced Mnemonic Reader + A2D (ensemble model)  
Feb 19, 2018  
Microsoft Research Asia & NUDT | 82.849 | 88.764 |
| 2    | SLQA+ (ensemble)  
Jan 5, 2018  
Alibaba iDST NLP | 82.440 | 82.607 |

Figure 11: Algorithms That Outperform Humans at Reading
Automatic Speech Recognition

When assessing automatic speech recognition (ASR) systems, Word Error Rate (WER) is the standard measure. To determine WER, the ASR output is dynamically aligned to a reference transcript, which then identifies three different types of error: Substitution errors, insertion and deletion errors.

Reading Comprehension

Stanford Question Answering Dataset (SQuAD) is a test of reading comprehension based on a dataset of more than 500 Wikipedia articles and 500,000 question-answer pairs associated with the articles. At the end of 2017, humans remained unequaled in their question answering skills, but in the five months since the start of 2018 at least four separate research groups from Microsoft, Alibaba, Google, and China’s Joint Laboratory of HIT and iFLYTEK Research have bested humans on the SQuAD dataset.

Transcribing Conversations

Algorithms have also now matched humans at transcribing conversations, achieving 95-percent Word Accuracy (WAcc) using the Switchboard HUB5’00 data set of 40 audio files and the corresponding human transcriptions. (See Figure 12.)

---

**Speech Recognition, Switchboard HUB5’00**

<table>
<thead>
<tr>
<th>Year</th>
<th>Best AI System</th>
<th>Human Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>85%</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 12: Speech Recognition vs. Human on Switchboard HUB5’00 Dataset*
# Appendix C: Acronyms Used in this Report

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADE</td>
<td>Adverse Event</td>
</tr>
<tr>
<td>AMCs</td>
<td>Academic Medical Centers</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>ASR</td>
<td>Automatic Speech Recognition</td>
</tr>
<tr>
<td>BI</td>
<td>Business Intelligence</td>
</tr>
<tr>
<td>CAC</td>
<td>Computer Assisted Coding</td>
</tr>
<tr>
<td>CAPD</td>
<td>Computer Assisted Physician Documentation</td>
</tr>
<tr>
<td>CCD</td>
<td>Continuity of Care Document</td>
</tr>
<tr>
<td>CCE</td>
<td>Clinical Concept Extraction</td>
</tr>
<tr>
<td>CDI</td>
<td>Clinical Documentation Improvement</td>
</tr>
<tr>
<td>cTAKES</td>
<td>Text Analysis and Knowledge Extraction</td>
</tr>
<tr>
<td>CTM</td>
<td>Clinical Trial Monitoring</td>
</tr>
<tr>
<td>CUI</td>
<td>Concept Unique Identifier</td>
</tr>
<tr>
<td>DBN</td>
<td>Deep Belief Network</td>
</tr>
<tr>
<td>DL</td>
<td>Deep Learning</td>
</tr>
<tr>
<td>DNR</td>
<td>Drug Name Recognition</td>
</tr>
<tr>
<td>DOS</td>
<td>Data Operating System</td>
</tr>
<tr>
<td>EDW</td>
<td>Enterprise Data Warehouse</td>
</tr>
<tr>
<td>EHR</td>
<td>Electronic Health Record</td>
</tr>
<tr>
<td>ELT</td>
<td>Extract Load Transform</td>
</tr>
<tr>
<td>ETL</td>
<td>Extract Transform Load</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>GSC</td>
<td>Gold Standard Corpora</td>
</tr>
<tr>
<td>HCC</td>
<td>Hierarchical Condition Category</td>
</tr>
<tr>
<td>HCOs</td>
<td>Healthcare Organizations</td>
</tr>
<tr>
<td>HIE</td>
<td>Health Information Exchange</td>
</tr>
<tr>
<td>HIT</td>
<td>Health IT</td>
</tr>
<tr>
<td>HMM</td>
<td>Hidden Markov Models</td>
</tr>
<tr>
<td>LHS</td>
<td>Learning Health System</td>
</tr>
<tr>
<td>LSA</td>
<td>Latent Semantic Analysis</td>
</tr>
<tr>
<td>LSTM</td>
<td>Long Short-Term Memory</td>
</tr>
<tr>
<td>MedLEE</td>
<td>Medical Extraction and Encoding</td>
</tr>
<tr>
<td>MIPS</td>
<td>Merit-based Incentive Payment System</td>
</tr>
<tr>
<td>ML</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>NER</td>
<td>Named Entity Recognition</td>
</tr>
<tr>
<td>NLPaaS</td>
<td>Natural Language Processing as a Service</td>
</tr>
<tr>
<td>NLPOS</td>
<td>Natural Language Processing Operating System</td>
</tr>
<tr>
<td>NLQ</td>
<td>Natural Language Query</td>
</tr>
<tr>
<td>NLU</td>
<td>Natural Language Understanding</td>
</tr>
<tr>
<td>NN</td>
<td>Neural Network</td>
</tr>
<tr>
<td>PHM</td>
<td>Population Health Management</td>
</tr>
<tr>
<td>PPC</td>
<td>Payer Provider Convergence</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive Predictive Value</td>
</tr>
<tr>
<td>RA</td>
<td>Risk Adjustment</td>
</tr>
<tr>
<td>RNN</td>
<td>Recurrent Neural Network</td>
</tr>
<tr>
<td>ROI</td>
<td>Return on Investment</td>
</tr>
<tr>
<td>SaaS</td>
<td>Software as a Service</td>
</tr>
<tr>
<td>SL</td>
<td>Supervised Learning</td>
</tr>
<tr>
<td>SNOMED-CT</td>
<td>Systematized Nomenclature of Medicine–Clinical Terms</td>
</tr>
<tr>
<td>STS</td>
<td>Semantic Textual Similarity</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>UIMA</td>
<td>Unstructured Information Management System</td>
</tr>
<tr>
<td>UMLS</td>
<td>Unified Medical Language System</td>
</tr>
<tr>
<td>VBC</td>
<td>Value-Based Care</td>
</tr>
<tr>
<td>VPAs</td>
<td>Virtual Personal Assistants</td>
</tr>
<tr>
<td>WAcc</td>
<td>Word Accuracy</td>
</tr>
<tr>
<td>WER</td>
<td>Word Error Rate</td>
</tr>
</tbody>
</table>
Endnotes


28. Advanced Medical Coding Solutions & Coding Automation. M*Modal Available at: mmodal.com/coding-solu-
NLP: ENABLING THE POTENTIAL OF A DIGITAL HEALTHCARE ERA


Ray Sipherd, S. C. C. The third-leading cause of death in America most doctors don’t want you to know about. (2018).


AI for saving lives. – Corti is a digital assistant that augments medical professionals. Available at: www.corti.ai/. (Accessed: 1st May 2018)


Pharmacovigilance is defined by the WHO as the marketing, distribution and use of drugs in a society, with special emphasis on the resulting medical, social and economic consequences. [Introduction to Drug Utilization Research: Chapter 1: What is drug utilization research and why is it needed?: 1.1. Definition and domains. Available at: apps.who.int/medicinedocs/en/d/js4876e/2.html. (Accessed: 19th February 2018)


Learn with Google AI. Google AI Available at: ai.google/education/. (Accessed: 27th April 2018)


The Health Catalyst Data Operating System (DOSTM) Solution. Health Catalyst Available at: www.healthcata-
D’Onfro, J. Google is hiring people to work on improving visits to the doctor's office with voice and touch technology. CNBC (2013). Available at: www.cnbc.com. (Accessed: 17th June 2018)
Microsoft Research. EmpowerMD - Medical Conversations to Medical Intelligence. YouTuBe (Youtube, 2018).


