Speech Analytics for Actionable Insights: Current Status, Recommendation, and Guidance

Other titles in Foundations and Trends® in Information Systems

Structured Conflict Approaches used in Strategic Decision Making: from Mason's Initial Study to Virtual Teams
Jerry Fjermestad
ISBN: 978-1-68083-616-5

Qualitative and Critical Research in Information Systems and Human-Computer Interaction: Divergent and Convergent Paths
Eleanor Wynn and Helena Vallo Hult
ISBN: 978-1-68083-556-4

Enterprise Personal Analytics: The Next Frontier in Individual Information Systems Research
T. Clohessy, T. Acton, E. Whelan and W. Golden
ISBN: 978-1-68083-460-4

Empirical Research in Information Systems: 2001–2015
S. Shuraida, H. Barki and A. Luong
ISBN: 978-1-68083-430-7

Speech Analytics for Actionable Insights: Current Status, Recommendation, and Guidance

J. P. Shim
Department of Computer Information Systems
J. Mack Robinson College of Business
Georgia State University
Atlanta, GA 30302
USA

Aaron M. French
Anderson School of Management
The University of New Mexico
Albuquerque, NM 87131
USA

Bikesh Upreti
Department of Information and Service Economy
Aalto University School of Economics
Espoo
Finland

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Information Systems

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

J. P. Shim, A. M. French and B. Upreti. *Speech Analytics for Actionable Insights: Current Status, Recommendation, and Guidance*. Foundations and Trends® in Information Systems, vol. 4, no. 3, pp. 213–274, 2020.

ISBN: 978-1-68083-697-4

Foundations and Trends® in Information Systems
Volume 4, Issue 3, 2020

Editorial Board

Editorial Scope

Topics

Foundations and Trends® in Information Systems publishes survey and tutorial articles in the following topics:

- IS and Individuals
- IS and Groups
- IS and Organizations
- IS and Industries
- IS and Society
- IS Development
- IS Economics
- IS Management
- IS Research Methods

Information for Librarians

Foundations and Trends® in Information Systems, 2020, Volume 4, 4 issues. ISSN paper version 2331-1231. ISSN online version 2331-124X. Also available as a combined paper and online subscription.

Contents

1 Introduction **3**

2 The Background of Speech Analytics **7**
2.1 History of Speech Analytics 8
2.2 The Need for Speech Analytics 9

3 Approaches and Methods of Speech Analytics **14**
3.1 Speech Analytics Methods 14
3.2 Real-Time versus Post-Call Capabilities 17
3.3 Text Mining Approaches 19

4 Major Vendors and Real-World Applications in Speech Analytics **24**
4.1 Major Vendors in Speech Analytics 24
4.2 Real-World Applications in Speech Analytics 30
4.3 Case Study #1: Morgan Stanley 31
4.4 Case Study #2: Bluegreen Vacations 34
4.5 Case Study #3: Bank of America 35
4.6 Case Study #4: HomeServe UK 38

5 Speech Analytics Issues and Recommendations **40**
5.1 Issues in Speech Analytics 40
5.2 Recommendations and Guidance 48

6 Concluding Comments **50**

Acknowledgments **52**

Appendix **53**

References **59**

Speech Analytics for Actionable Insights: Current Status, Recommendation, and Guidance

J. P. Shim[1], Aaron M. French[2] and Bikesh Upreti[3]

[1]*Department of Computer Information Systems, J. Mack Robinson College of Business, Georgia State University, Atlanta, GA 30302, USA; jpshim@gsu.edu*
[2]*Anderson School of Management, The University of New Mexico, Albuquerque, NM 87131, USA; afrench@unm.edu*
[3]*Department of Information and Service Economy, Aalto University School of Economics, Espoo, Finland; bikesh.upreti@aalto.fi*

ABSTRACT

In today's hypercompetitive customer-centric marketplace, every enterprise strives to gain competitive advantage through customer loyalty, high customer satisfaction and low customer turnover. Through the use of a combination of tools such as analytic technologies, and data-mining techniques and access to real-time data, companies are now able to place a greater emphasis on customer engagement and satisfaction. Today's increased enforcement of standards and stricter legal compliance rules have led call centers to take proactive steps to ensure that enforcement is in compliance with regulations through the use of speech analytics. In the realm of analytic technologies, speech analytics is quickly becoming one of the most demanded technologies in customer engagement optimization and the fastest growing technology in call centers. Organizations have been searching for

J. P. Shim, Aaron M. French and Bikesh Upreti (2020), "Speech Analytics for Actionable Insights: Current Status, Recommendation, and Guidance", Foundations and Trends® in Information Systems: Vol. 4, No. 3, pp 213–274. DOI: 10.1561/2900000020.

ways to translate this wealth of information into holistic, accurate, and actionable insights. The nascent technology is increasingly in more demand as its features become more relevant for call centers, and as organizations seek to capture the voice of the customer (VoC), or the customers' wants and needs, and improve first call resolution (FCR) through post-call and real-time solutions. Speech analytics is a complementary approach for organizations driving optimization, personalization and targeting across their digital channels. Since the latest real-time voice/speech analytics technology can mine 100% of the company's voice contacts, organizations can now be much more successful in capturing and tagging the reasons for the customers' call. Given today's ubiquitous computing trends, such as mobile voice and digital data stream, post-call and real-time applications are solutions to capture the voice of the customer, and improve operations and efficiency, customer experiences, and improve customer loyalty.

Keywords: speech analytics; text mining; data insights; call analytics; customer engagement optimization; use cases; artificial intelligence.

1

Introduction

In today's hypercompetitive business environment, enterprises may strive to gain a sustainable competitive advantage by achieving the highest marks on the most influential elements of customer service: establishing brand and customer loyalty, high customer satisfaction, and a high customer retention rate together with low customer attrition rate. With access to real-time data, companies continually pursue opportunities to achieve this competitive edge and increase the likelihood of success. Tapping into such opportunities requires improving the understanding of customers' perceptions toward the company's brand image, products, and services through a variety of tools such as big data, analytics technologies, and data-mining techniques. For example, call center management has taken steps to implement integrated phone systems designed to improve the quality of its call center agents through insights identified from speech analytics.

In the early days of speech analytics, the adopters were unclear on how to interpret the results from speech analytics technology as there were no actionable insights or results from reports of repeated words or phrases. Around the early 2000s, speech analytics quickly became more widespread as organizations in a variety of industries demanded its

implementation within call centers to maximize customer engagement and customer relationships. Thanks to technological advancements in speech analytics software and an ever-growing data repository, today, companies can uncover business insights and analyze customer perceptions, emotions, and tones as never before. These valuable insights are discovered within a wide range of customer interactions that occur in a variety of customer service channels such as telephone conversations, in-person communication, or customer support channels (e.g., chat, app, text, social media, or email). In Aberdeen Group's benchmark report, "Speech Analytics: Listen to your Customers", companies reported that nearly two-thirds of interactions with customers in call centers involved voice (www.aberdeen.com). The combination of speech analytics technology and big data has garnered the attention of practitioners and academicians alike. Big data components allow call centers and corporations to process and analyze massive amounts of data and translate them into valuable and actionable goals (Kappelman *et al.*, 2014; Moore, 2014).

Over the past decades, numerous research studies have been conducted in the areas of customer engagement and customer relationship management, illustrating the roles and activities performed by customers at various touchpoints. This approach has led firms to strive for a better understanding of the context shaped by the cultural, social, and relational dimensions (Brodie *et al.*, 2011; Chan, 2005). In the realm of analytics technologies, speech analytics is quickly becoming one of the most in-demand technologies in customer engagement optimization and the fastest-growing technology in call centers. Despite being a nascent technology, speech analytics is already increasingly in high demand as its features become more relevant and crucial not only for call centers but also for organizations that seek to capture the voice of the customer (VoC), also known as the customers' wants and needs. The VoC describes the in-depth process of capturing the customer's expectations, preferences, and aversions. Similar to VoC, first call resolution (FCR) is another important facet in achieving customer satisfaction. This is where call center agents successfully listen to customer needs and provide a resolution to solve the customer's issue on the first call. The speech analytics technology is instrumental for successful FCR, through

post-call and real-time solutions by conducting a root-cause analysis, as it investigates the reasons why customers would contact a call center and utilize that information to remedy specific problems at the source.

Given today's ubiquitous computing trends, such as mobile voice and digital data stream, the voice of the customer drives continuous improvement cycles in operational improvement and efficiency, and customer experiences. As identified in the previous researches, voice and speech analytics fall under the umbrella of digital data streams (DDS), which harnesses real-time or near real-time data (Piccoli *et al.*, 2015). Piccoli and Pigni's survey of CIOs demonstrated that there exist exponential untapped opportunities to exploit DDS (Piccoli and Pigni, 2013). While currently used within organizations for monitoring purposes, the combination of a DDS with predictive analytics has the potential to revolutionize customer service and also efficiently improve call center agents' performance.

Speech analytics utilizes speech recognition, predictive analytics, and authentication of the data streams (in this case, the voice of the customer) while assessing customers' complaints in real-time. Assessment occurs through the collection and analysis of current data mixed with historical facts to determine patterns and to predict trends. In the current research, the authors have chosen to focus primarily on speech analytics, serving as an umbrella term encompassing speech analytics, audio-mining technologies. The use of speech analytics typically refers to a broader range of speech products, such as analyzing voice identification, emotion detection, and phonetics/speech analysis.[1] Due to the complexity of vocal elements of a speaker's behavior, the current research excludes in-depth analysis of speech analytics (i.e., syllable emphasis, tone, pitch, tempo, etc.) and focuses primarily on speech analytics. This paper proceeds with the discussion of an overview of enterprise needs for speech analytics, a brief history of the speech recognition, the infrastructure of phonetic versus transcription approaches and real-time versus post-call solutions, major speech analytics vendors and their features, applications found within case studies, and recommendations and guidance. The primary goal of this paper is to help

[1] www.rankminer.com/post/voice-analytics-vs-speech-analytics-difference.

business decision-makers educate themselves on the burgeoning field of speech analytics as well as to understand how it impacts the broader enterprise landscape.

References

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). "Latent Dirichlet allocation". *Journal of Machine Learning Research*. 3: 993–1022.

Bloehdorn, S. and A. Hotho (2004). "Boosting for text classification with semantic features". In: *International Workshop on Knowledge Discovery on the Web*. Berlin, Heidelberg: Springer. 149–166.

Brodie, R. J., L. D. Hollebeek, B. Juric, and A. Ilic (2011). "Customer engagement: Conceptual domain, fundamental propositions, and implications for research". *Journal of Service Research*. 14(3): 252–271.

Chan, J. O. (2005). "Toward a unified view of customer relationship management". *The Journal of American Academy of Business*. March: 32–38.

Ciarelli, P. M., E. Oliveira, C. Badue, and A. F. De Souza (2009). "Multi-label text categorization using a probabilistic neural network". *International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM)*. 1: 133–144.

Davenport, T. H. and J. Kirby (2016). "Just how smart are smart machines?" *MIT Sloan Management Review*. 57(3): 21.

Davies, J. (2015). "Are you ready for real-time speech analytics?" *Gartner*. June 19.

Deerwester, S., S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman (1990). "Indexing by latent semantic analysis". *Journal of the American Society for Information Science*. 41(6): 391–407.

Dev Sarma, B. and S. R. M. Prasanna (2018). "Acoustic–phonetic analysis for speech recognition: A review". *IETE Technical Review*. 35(3): 305–327.

Devlin, J., M. W. Chang, K. Lee, and K. Toutanova (2018). "BERT: Pre-training of deep bidirectional transformers for language understanding". arXiv: 1810.04805.

DMG Consulting LLC (2015). "2015–2016 speech and text analytics market report and consulting services". June.

Envision Intelligence LLP (2019). "Global speech analytics market 2020–2026".

Fluss, D. (2014a). "Speech analytics is starting to make a difference". URL: www.destinationcrm.com, November.

Fluss, D. (2014b). "Speech analytics is starting to make a difference". *Customer Relationship Management*. November: 58–59.

Fluss, D. (2015). "Analytics revolutionizes the QA process". *Speech Technology*. Spring: 48.

Garland, J., C. Blake, M. Finlay, and D. Lanham (2014). *Phonetic-Based Dialogue Search: The Key to Unlocking an Archive's Potential*. Atlanta, GA: Nexidia, Inc.

Hartigan, J. A. and M. A. Wong (1979). "Algorithm AS 136: A K-means clustering algorithm". *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 28(1): 100–108.

Hoel, T. and W. Chen (2016). "Implications of the European data protection regulations for learning analytics design". In: *Workshop Paper Presented at the International Workshop on Learning Analytics and Educational Data Mining (LAEDM 2016) in Conjunction with the International Conference on Collaboration Technologies (CollabTech 2016)*. Kanazawa, Japan. 14–16.

Jeffs, C. (2015). "Phonetic vs. transcription". Presented on behalf of Verint at Georgia State University, September.

Joachims, T. (1998). "Text categorization with support vector machines: Learning with many relevant features". In: *European Conference on Machine Learning*. Berlin, Heidelberg: Springer. 137–142.

Kappelman, L., E. McLean, V. Johnson, and N. Gerhart (2014). "The 2014 SIM IT key issues and trends study". *MIS Quarterly Executive*. 13(4): 237–263.

Le, Q. and T. Mikolov (2014). "Distributed representations of sentences and documents". In: *International Conference on Machine Learning*. 1188–1196.

Le Prell, C. and O. Clavier (2017). "Effects of noise on speech recognition: Challenges for communication by service members". *Hearing Research*. 349: 76–89.

Lee, J. Y., J. T. Lee, H. J. Heo, C. H. Choi, S. H. Choi, and K. J. Lee (2015). "Speech recognition in real-life background noise by young and middle-aged adults with normal hearing". *Journal of Audiology & Otology*. 19(1): 39–44.

Meteer, M. (2011). "Choosing the right technology for your speech analytics project". Unpublished paper (retrieved March 1, 2016, from URL: www.callminer.com/wp-content/whitepapers/The-Right-Technology-for-Speech-Analytics.pdf).

Mikolov, T., K. Chen, G. Corrado, and J. Dean (2013). "Efficient estimation of word representations in vector space". arXiv: 1301.3781.

Minkara, O. (2013). "Speech analytics: Making effective use of voice in a multi-channel world". Aberdeen Group. September.

Moore, B. (2014). "Speech analytics: It's all about statistics". *Pipeline Articles*: 2.

Mordor Intelligence (2019). "Speech analytics market—Growth, trends and forecast (2019–2024)". URL: https://www.mordorintelligence.com/industry-reports/global-speech-analytics-market-industry.

Owens, M. (2016). "Sentiment analytics". Presented on behalf of LP Insights at Georgia State University, January.

Pennington, J., R. Socher, and C. Manning (2014). "Glove: Global vectors for word representation". In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 1532–1543.

Piccoli, G. and F. Pigni (2013). "Harvesting external data: The potential of digital data streams". *MIS Quarterly Executive*. 12(1): 143–154.

Piccoli, G., J. Rodriguez, and R. Watson (2015). "Leveraging digital data streams: The development and validation of a business confidence index". In: *Proceedings of the Forty-Eighth Hawaii International Conference on System Sciences,* January.

Pigni, F., G. Piccoli, and R. Watson (2016). "Digital data streams: Creating value from the real-time flow of big data". *California Management Review.* 58(3): 5–25.

Shankar, V. (2018). "How artificial intelligence (AI) is reshaping retailing". *Journal of Retailing.* 94(4): vi–xi.

Shim, J. P., J. Koh, S. Fister, and H. Y. Seo (2016). "Phonetic analytics technology and big data: Real-world cases". *Communications of the ACM.* 59(2): 84–90.

Turney, P. D. and P. Pantel (2010). "From frequency to meaning: Vector space models of semantics". *Journal of Artificial Intelligence Research.* 37: 141–188.