



Hannah DeGraff
www.goShadow.org

The Pain Points in Healthcare: White Paper on Surveying in Healthcare Organizations

INTRODUCTION

The advancement in healthcare technology coupled with a push for value-based care has rapidly increased efforts to utilize data for fact-based solutions. goShadow has co-designed numerous surveys for healthcare organizations to discover and improve low hanging fruit, organizational culture, and outcomes. The surveys consist of likert scale, multiple choice, and open response questions to provide a comprehensive understanding through quantitative and qualitative data. Although quantitative data provides objective answers, qualitative data reveals the feelings, attitudes, and behaviors behind the response. Thus, qualitative data is integral for overcoming the quality chasm that the American health system suffers from (Richard, 2001).

goShadow surveys created for patients render insights on clinical best practices and how to improve processes and experiences using patient-centered care (PCC) principles. Focusing on PCC increases patient satisfaction as well as outcomes. The greater positive experience patients have the more likely they are to return for preventative care before an illness becomes acute. Surveys created for healthcare professionals render insights into how to reduce burnout and increase satisfaction. Burnout rates have been increasing with the constant changes in systems from advancing technology as well as long, sleepless hours working. Discovering ways to ease the burden and make the provider feel supported will go a long way. The rest of this article will focus on surveys created for patients.

Bar charts and word clouds are created to visualize these insights for upper management. The Likert scale and multiple-choice questions are easily placed into a bar chart, but the open response questions require more manipulation. Currently, goShadow analysts manually categorize the text based on their perception of the sentiment and predetermined topics. Predetermined topics were created by goShadow analysts based on the AHRQ TeamSTEPPS (Baker et al., 2008). Developing a data-driven categorization process would save hours of manual labor and expand upon the possibility of standardizing all open-text questions.

BACKGROUND

History of Patient-Centered Care

Interest in patient satisfaction began in 1985 when Notre Dame professor Irwin Press and Medical anthropologist Rod Ganey created the Press Ganey satisfaction survey (Siegrist, 2013). The federal government did not catch onto the trend until 2002 when the Centers for Medicare and Medicaid (CMS) and the Agency for Healthcare Research and Quality (AHRQ) developed the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey. As the first national, standardized, publicly reported survey, HCAHPS has assisted in producing objective comparisons of healthcare organizations, incentives to improve quality of care, and increase transparency (HCAHPS, 2020). The Affordable Care Act of 2010 introduced the pay for performance model to further incentivize health care organizations to improve HCAHPS scores and overall quality patient care.

Although the healthcare system has already begun reinventing itself around PCC, patient-centeredness does not have a clear definition. Pluut has identified three discourses that makeup patient-centeredness: “caring for patients”, “empowering patients”, and “being responsive” (Pluut, 2016). Caring for those who are vulnerable lessens suffering and improves the quality of diagnosis. Empowering and educating patients on self-management provides them with autonomy to make knowledgeable health decisions. Responding appropriately to the patient’s needs, values, and preferences personalizes care to best suit the patient. Focusing on PCC will improve clinical outcomes and be cost-effective for patients and healthcare organizations.

Key Tools

Various tools have been made available to perform Natural Language Processing (NLP) on open text. Excel provides add-ons such as Azure Machine Learning and 3rd party sites such as MeaningCloud provide various text analysis tools as well. However, goShadow has chosen to use the Python programming language as their solution. Python is a powerful, dynamic tool that uses multiple libraries and add-ons to complete NLP tasks.

Key tools used within Python for NLP include the Natural Language Toolkit (NLTK) platform and TextBlob library. NLTK is a platform compatible with Python programs that work with human language data (Natural, n.d.). It uses over 50 corpora and lexical resources to compare the text to. TextBlob is a Python library that provides a simple API for analyzing Natural Language Processing (NLP) tasks (Textblob, n.d.). It performs sentiment analysis as well as many other speech tasks such as parsing, noun phrase extraction, tokenization, etc. A combination of these tools is needed to complete the analysis.

SOLUTION

Pre-Processing

Open text responses often create responses with spelling errors, incorrect punctuation, symbols, and abbreviations. Pre-processing is a data mining technique that strips the raw data into a format that the program can understand. An algorithm can be written to make all the text in lowercase letters and remove all stop words. Stop words are commonly used words such as the, is, and a. Removing filler words focuses the analysis on the important text that gives meaningful information. All characters such as %, #, @, etc. were also removed. Stemming changes all words to their root form to further reduce dimensionality. For example, running becomes run. Once the responses are cleaned, they are downloaded into a CSV file to be manipulated by Python.

Python

The Python algorithm performs a sentiment analysis and categorizes each response into predetermined topics. A sentiment analysis returns a polarity and subjectivity score. The polarity score is a number in the range of $[-1, 1]$ where -1 is negative and 1 is positive. The closer the polarity score is to 0 the more neutral the response is. The subjectivity score reveals how likely the response is based on opinions or emotions. Subjectivity is a number in the range of $[0, 1]$ where 1 is highly opinionated.

Creating a solution to categorize open-text responses into predetermined topics can be attempted in numerous ways. goShadow has decided on a solution that matches the words in a response to a word list. The word list is separated into the predetermined categories with specific words relating to each of the categories. For instance, if the word “help” was read in a response, it would match that response to support because “help” is under the category support in the wordlist. Since the algorithm is simply matching words, the response can fall within more than one category if more words are found in other categories. A confidence interval can assist in determining whether the response truly fits within that category. Some manual analysis must occur to ensure the algorithm’s accuracy.

CONCLUSION

Overall, the Python programming language is a powerful tool for completing NLP tasks for surveys. Further research must be done to ensure the accuracy of the algorithm, but it is a promising start to increasing productivity. NLP will not replace the ability of what humans can do, but it can assist in reducing manual labor. Sentiment analysis and topic modeling are popular ways to utilize NLP for qualitative data. Utilizing a well-trained algorithm to understand the opinions, values, and experience behind open text responses increases productivity and reduces bias.

Works Cited

Baker, D., Krokos, K., & Amodeo, A. (2008, October 7). TeamSTEPPS Teamwork Attitudes Questionnaire Manual. AHRQ.

HCAHPS: Patients' Perspectives of Care Survey. (2020, February 11). Retrieved from <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalHCAHPS>

Natural language toolkit. (n.d.). Retrieved February 27, 2021, from <http://www.nltk.org/>

Textblob. (n.d.). Retrieved February 27, 2021, from <https://pypi.org/project/textblob/>

Pluut B. (2016). Differences that matter: developing critical insights into discourses of patient-centeredness. *Medicine, health care, and philosophy*, 19(4), 501–515. <https://doi.org/10.1007/s11019-016-9712-7>

Richard and hinda rosenthal lectures spring 2001 : Crossing the quality chasm. (2003). ProQuest Ebook Central <https://ebookcentral.proquest.com>