A Sentiment Analysis of Patient Safety Event Reports

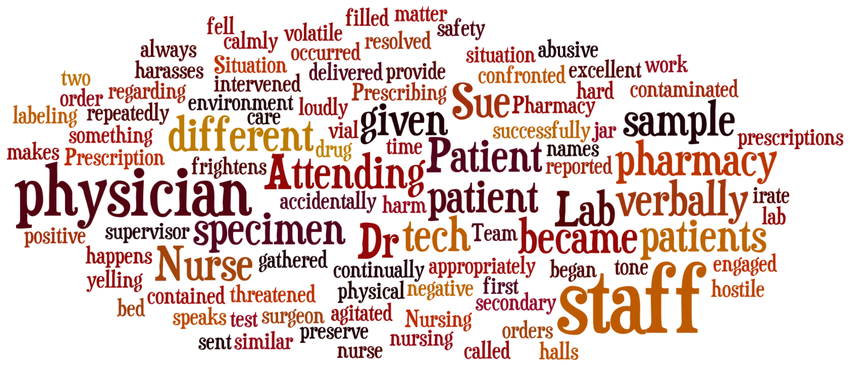
Leatt Gilboa[1]; Daniel Hoffman[2]; Allan Fong[2]; Raj Ratwani[2]
[1]MedStar Institute for Innovation, [2]National Center for Human Factors in Healthcare

MedStar Health National Center for Human Factors in Healthcare

MedStar Institute for Innovation

Introduction

- **Patient Safety Event** (PSE) reporting **systems** are **widely implemented across the country**
- Reporting systems generally **contain large free text description fields**
- Analyzing the free text to understand the event and culture is critical
- Natural Language Processing (NLP) techniques, such as sentiment analysis (SA) can be used better understand these data
- **The objective** of this is to describe an NLP driven approach to **better understand sentiment** in the **free text of PSE reports**.

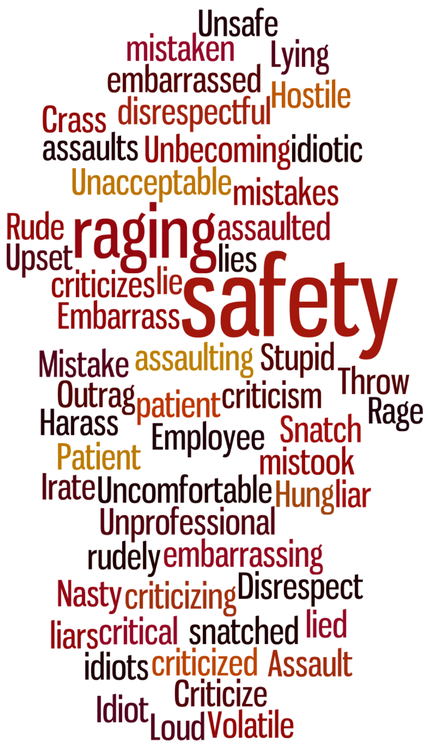


*Word cloud generated from synthesized PSE reports

Methods

Investigation Process

1. **Prepare Data** - *Clean text by removing numbers, punctuations, and extra spaces*
2. **Initial Analysis** - *Using the AFINN library rank the PSE reports from most "emotive" to least.*
3. **Manual Review** -*Two investigators manually review the 60 highest ranking reports for sentiment . Compile list of sentiment keywords found in these reports (that may or may not correspond to AFINN).*
4. **Inter-Researcher Reliability** - *Compare individually compiled keyword lists and reach consensus.*
5. **Test Keywords** – *Search all 40,000 PSE reports with keywords from Step 3. Manually review the pulled reports for sentiment.*
6. **Weighting Scale** - *Gather keyword frequencies and assess each keyword's correlation to sentiment.*



*Word cloud generated from censored library

Example process flow (Investigation process steps 3 – 5)

Report Text

Dr. Kyle is always hostile to my staff and repeatedly harasses me. No matter what happens, something positive or negative, Dr. Smith is verbally abusive. This makes the work environment hard to provide excellent care for the patient.
*This is not a real report, but is representative of the qualities found in highly sentimental reports.

Initial Feeling: Negative report
Initial Scan: Located 3 keywords that contribute to the report being negative

String	Positive usage	Neutral	Negative usage
Harass			1
Hostile			1
…	…	…	…

Searched for "hostile" and "harass" in rest of 40,000 reports. If a report has a single instance of the keyword it is counted

of affected reports
9
51
…

Did these reports convey an emotional or sentimental writer?

Sentiment Correlation
yes
yes
…

How many reports that contain these keywords describe patient behavior?

About the patient?
~50%
…

Results & Discussion

- Uncovered **36 keywords** with varying degree of sentiment reliability.
- **Context** appears **relevant** to sentiment.

1. **Expression** - Keywords that portray the author's tone (for example, "the hostile environment made me feel unsafe").
2. **Quoting** - Keywords that are used to quote patients or to describe patient behavior, but are independent of the author's tone (for example, "The patient was hostile, so we called security").
3. **Nuance** – Combination and location of keywords with others appears influential.

Keywords that are not utilized in general sentiment libraries were **uncovered** (such as "hung up", "assault", "volatile", and "throw")

Conclusion

- Sentiment analysis may be a promising way to understand culture based on patient safety reports
- Next steps are to correlate SA data with other indicators of culture

Acknowledgments:

This work was supported by MedStar Health Quality and Safety