

This document created by Karen A. Monsen, Sasank Maganti, Christie Martin, and Mary Jo Kreitzer is licensed under the Creative Commons Attribution Non-Commercial Share Alike 4.0 International License in January, 2020. To view a summary of the license, go to <https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode>

Nursing Insights From CaringBridge Notes

2018 Nursing Knowledge: Big Data Science Conference
Pre-Conference Track 2: Social Media Analytics and Mobile Technology

June 13, 2018

Karen A. Monsen
Sasank Maganti
Christie Martin
Mary Jo Kreitzer

Acknowledgments

- CaringBridge funding for the study: Understanding the impact of social technology on wellbeing, altruism & health outcomes (Kreitzer, M. J., PI).
 - Center for Spirituality and Healing Team
 - College of Computer Science and Engineering Team
 - Omaha System Partnership-Center for Nursing Informatics

Objectives

- Describe CaringBridge, a compassionate technology social media platform
- Describe ontologically-based text mining
- Discuss methodological approaches for use in social media text mining
- Identify applications and implications for social media text mining methods

CaringBridge

[Donate to CaringBridge](#)



[Sign In or Register](#)

[About Us](#)

[How It Works](#)

[Start A Site](#)

[Resources](#)

[Search](#)

STRONGER TOGETHER

Share, Connect and Rally Support
Personal, Protected Websites for Every Health Journey

Visit a Friend's Website

Person's Name

NOLAN L., CaringBridge User

KRISTEL S., CaringBridge Donor

START YOUR FREE WEBSITE

[Start a Site](#)

Save Time with One Update

No more repeating the story over and over. Connect with all of your family and friends at once, giving you time to focus on what matters.

Nursing Knowledge:
Big Data Science Conference



SCHOOL OF NURSING
UNIVERSITY OF MINNESOTA

CaringBridge.org

- Nonprofit social network
 - Facilitate communication with friends and family
 - Support for loved ones during a health journey
- Free personal website
- Place for healing and comfort
 - 740,000 web sites
 - 235 countries



Sona Mehring (founder) with JoAnn Hardegger and Darrin Swanson holding a photo of their daughter Brigid, the baby who inspired CaringBridge in 1997. (Photo: May 2011. <https://www.caringbridge.org/about-us>)

UMN-CaringBridge Collaboration

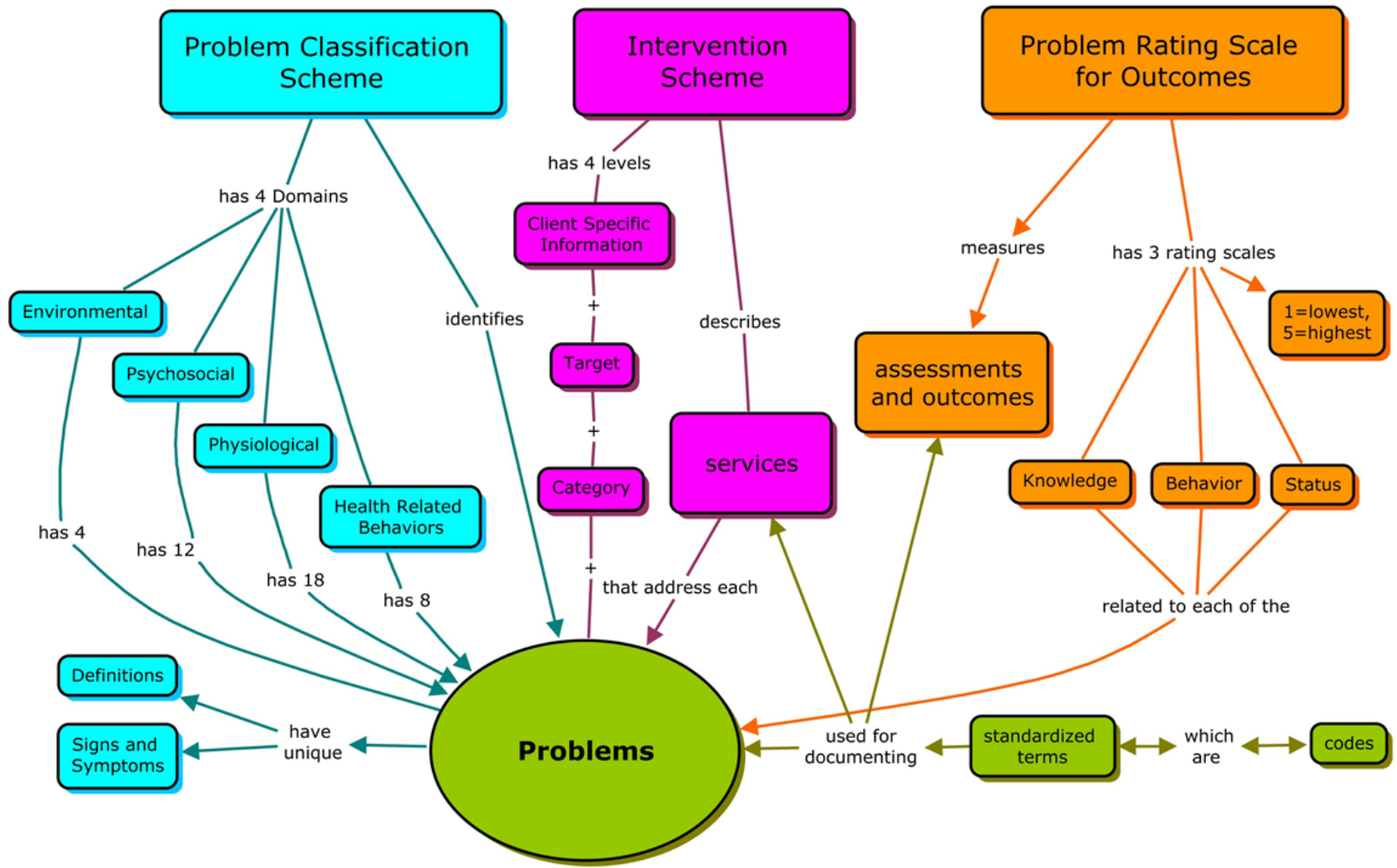
- Liwanag Q. Ojala, Chief Executive Officer, CaringBridge
- Mary Jo Kreitzer, Director, Bakken Center for Spirituality and Healing, University of Minnesota
 - Karen A. Monsen, School of Nursing
 - Arindam Banerjee, Computer Science and Engineering
 - Lana Yarosh, Computer Science and Engineering

The Ontology Question Guiding the Text-mining Approach

- *“What kinds of things exist or can exist in the world, and what manner of relations can those things have to each other?”*

<http://semanticweb.org/wiki/Ontology>

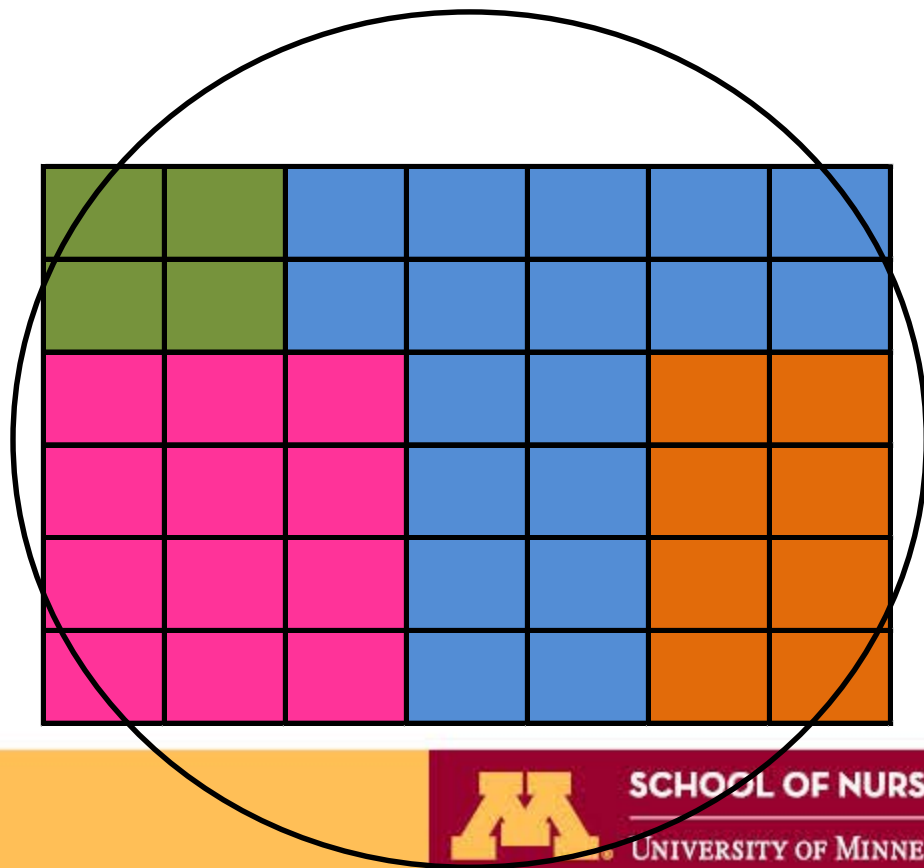
The Omaha System (Martin, 2005)



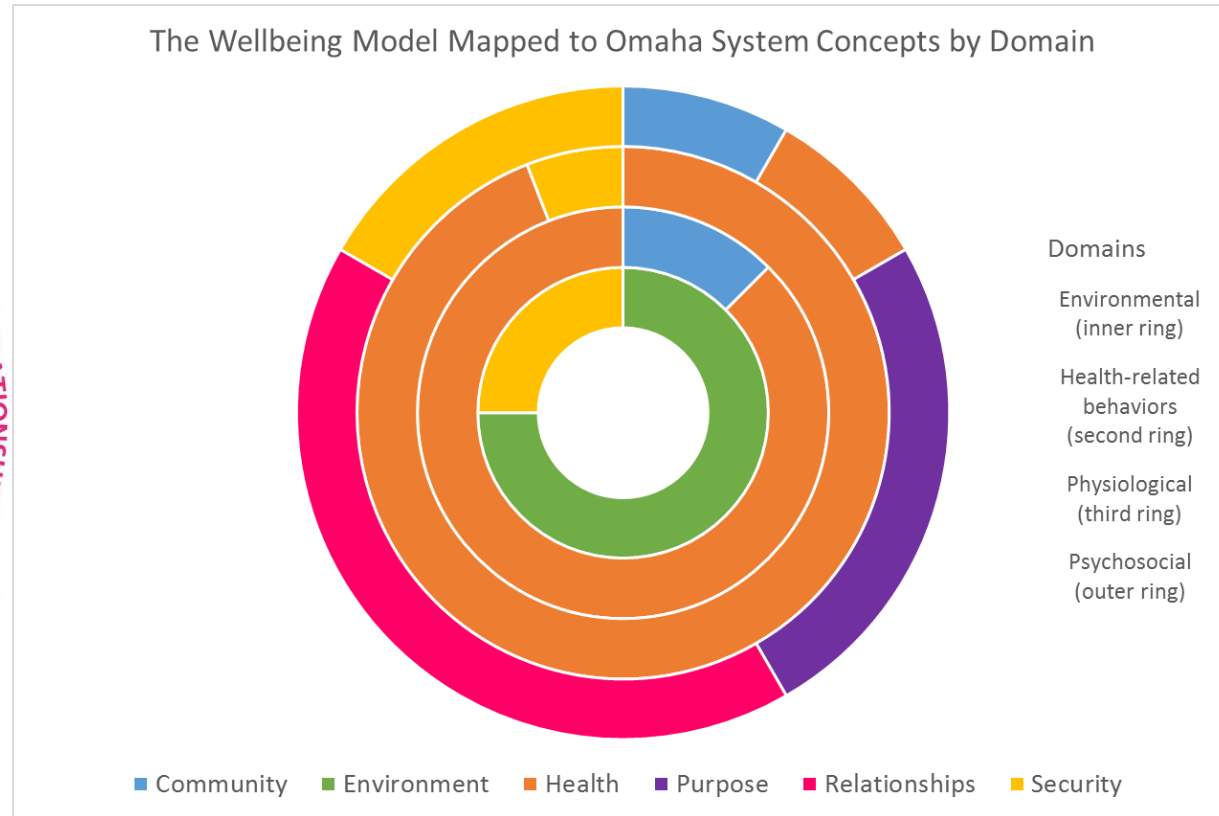
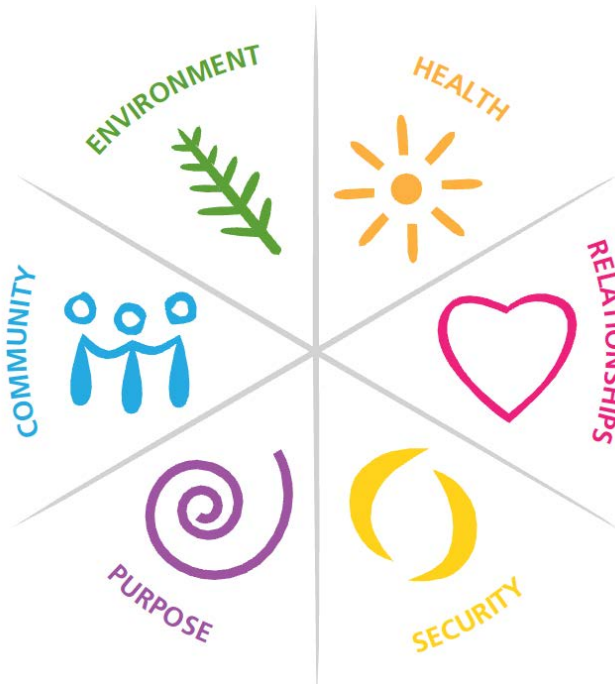
42 Neutral Concepts

Describe all of health and health care

Called “Problems” but may refer to strengths as well



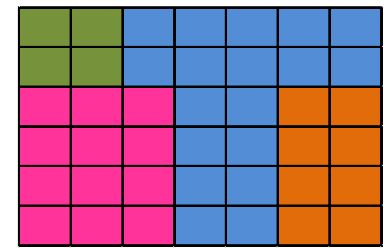
Comprehensive Whole



MJ Kreitzer & M. Koithan (Eds.), Integrative nursing. Cary, NC: Oxford University Press.

Problem Classification Scheme

- Describes defined health concepts in four domains
 - Environmental (4 problems)
 - Psychosocial (12 problems)
 - Physiological (18 problems)
 - Health-related Behaviors (8 problems)
- Every problem has a definition and a unique list of signs/symptoms
 - See <http://omahasystem.org/problemclassificationscheme.html>



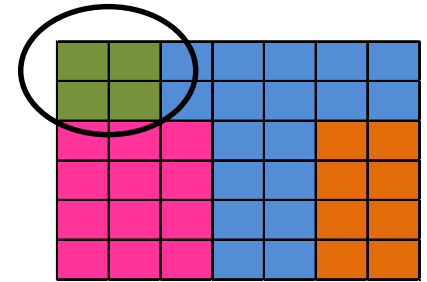
Environmental Domain

Income

Sanitation

Residence

Neighborhood/workplace safety



Psychosocial Domain

Communication with community resources

Social contact

Role change

Interpersonal relationship

Spirituality

Grief

Mental health

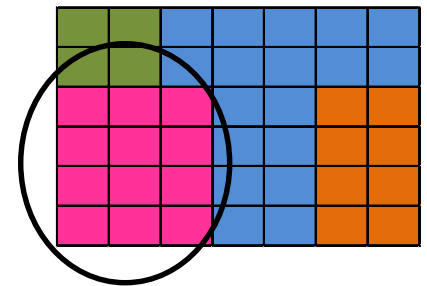
Sexuality

Caretaking/parenting

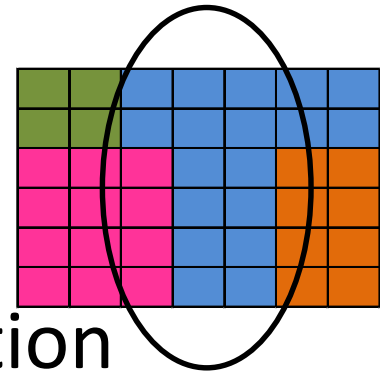
Neglect

Abuse

Growth and development



Physiological Domain



Hearing

Vision

Speech and language

Oral health

Cognition

Pain

Consciousness

Skin

Neuro-musculo-skeletal
function

Respiration

Circulation

Digestion-hydration

Bowel function

Urinary function

Reproductive function

Pregnancy

Postpartum

Communicable/infectious
condition

Health-related Behaviors Domain

Nutrition

Sleep and rest patterns

Physical activity

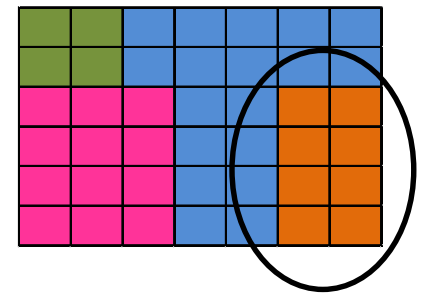
Personal care

Substance use

Family planning

Health care supervision

Medication regimen



Specific Aims

1. To develop ontology-based text mining methods for application in a CaringBridge social media corpus
2. To examine the use of the Omaha System as a basis for understanding whole-person health/wellbeing in CaringBridge journals
3. To describe CaringBridge journal content from a whole-person perspective

Methods

- Sample consisted of free text from 13,757,900 CaringBridge journal entries
- The text dataset was prepared by removing stop words and html text using standard NLP procedures by Giaquinto, Banerjee, and team
- Word counts of the 42 Omaha System problem concepts were obtained using shell scripts and python programming on Minnesota Supercomputing Institute High Performance Computing systems in four steps

Text Mining Approach

- Step 1: Problem concept stems
- Step 2: Combining problem concept stems
- Step 3: Adding S/sx stems
- Step 4: Adding synonyms and related words

Sleep and rest patterns

- Periods of suspended motor and sensory activity and periods of inactivity, repose, or mental calm
 - Sleep/rest pattern disrupts family
 - Frequently wakes during night
 - Sleepwalking
 - Insomnia
 - Nightmares
 - Insufficient sleep/rest for age/physical condition
 - Sleep apnea
 - Snoring

Sleep and Rest Patterns

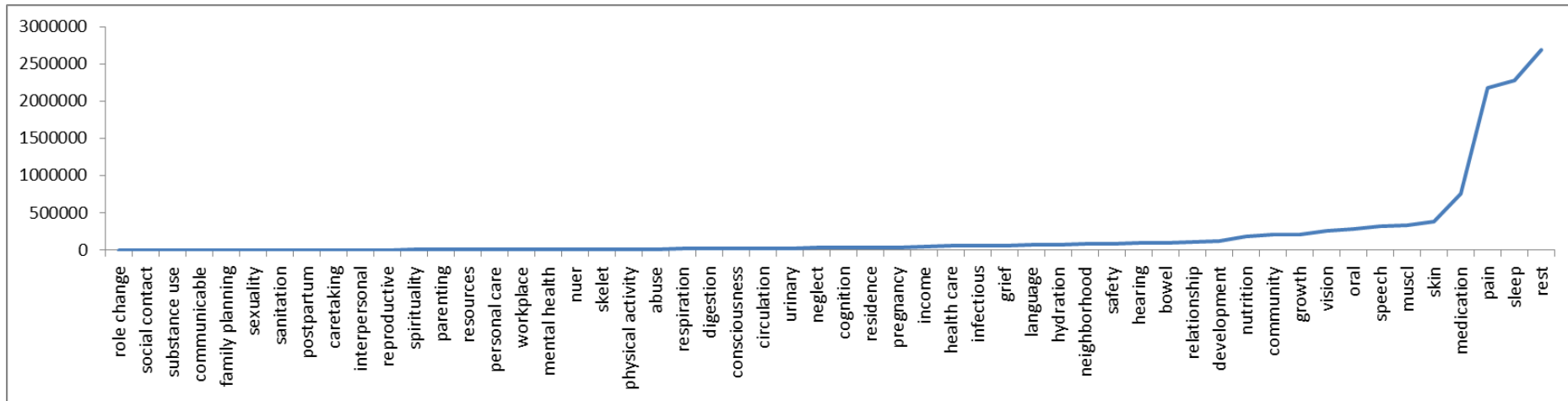
Stems, S/sx, Synonyms

- Sleep
- Rest
 - awake during night
 - sleepwalk
 - insomnia
 - Nightmare
 - snoring
- repose
- snore
- awake at night
- awake all night
- up all night
- nap
- doze
- shut eye
- shuteye

Results

Step 1: Problem concept stems

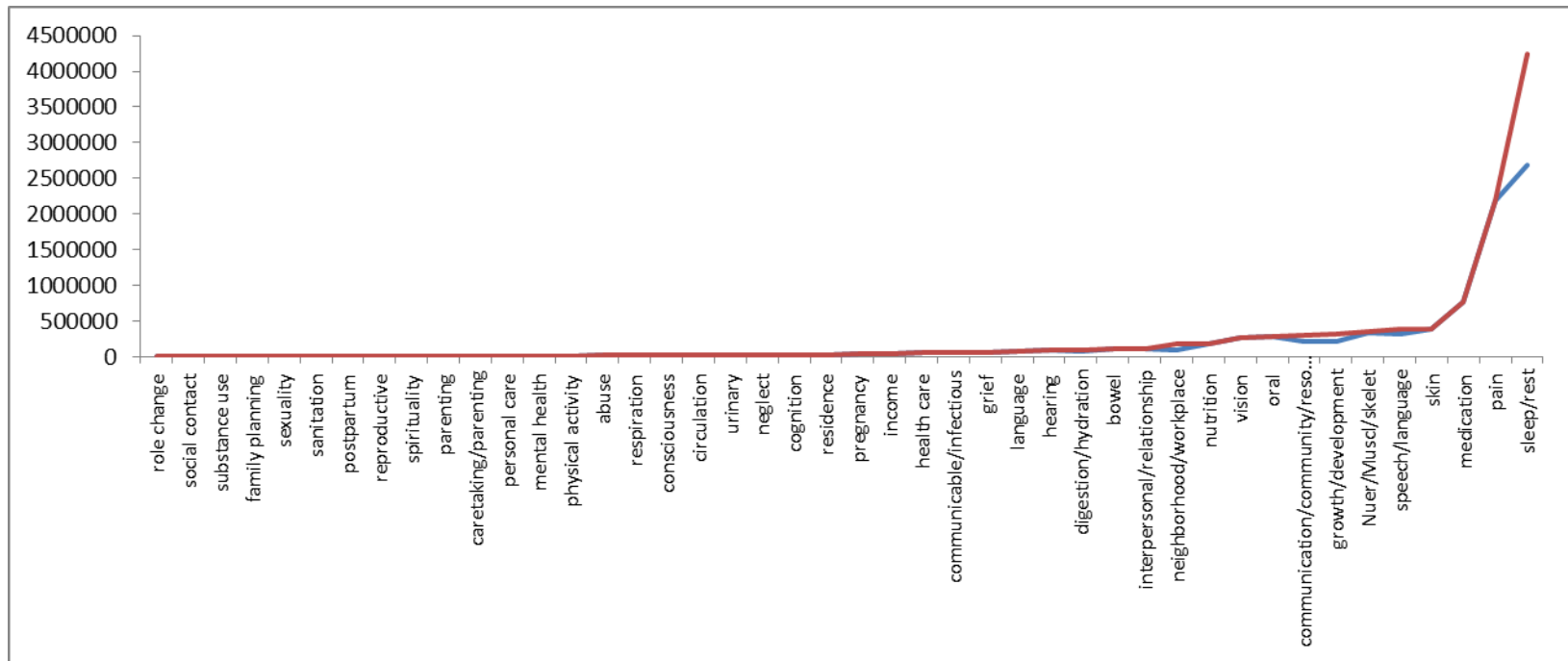
- All 55 terms and stems were present



- Range from 336 (Role change) - 2,685,494 (Rest)

Step 2: Combine problem concepts from terms and stems

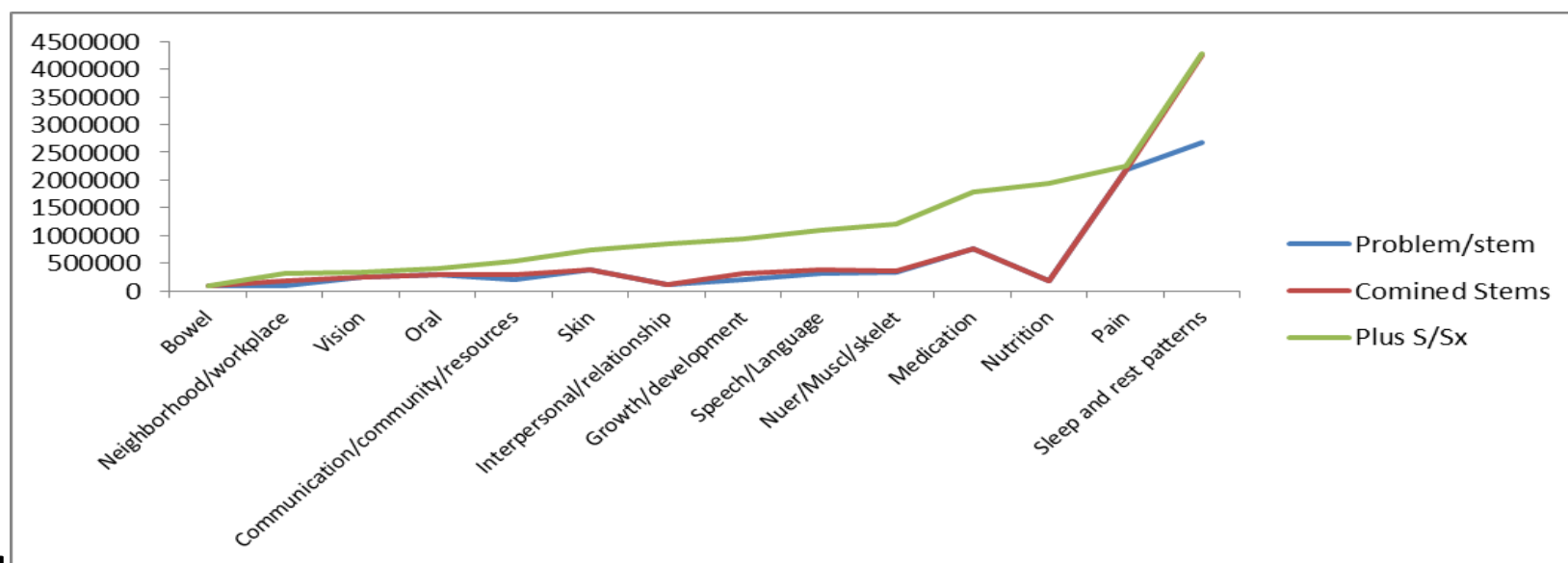
- Showing 42 concepts



- Maximum = Sleep/rest (>4M)

Step 3: Adding signs/symptoms

- showing problem terms and stems with $N > 100,000$

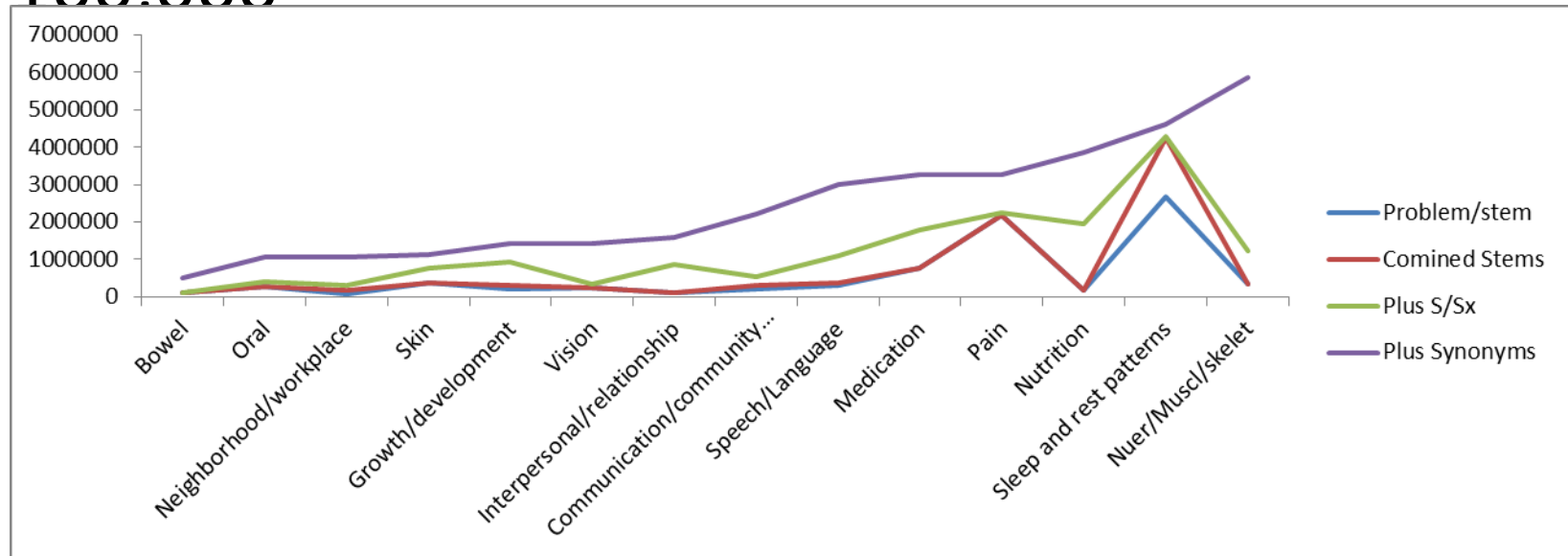


- Increases ranged from 2.5% to 957.3%

- Maximum = 4.5M

Step 4: Adding related words

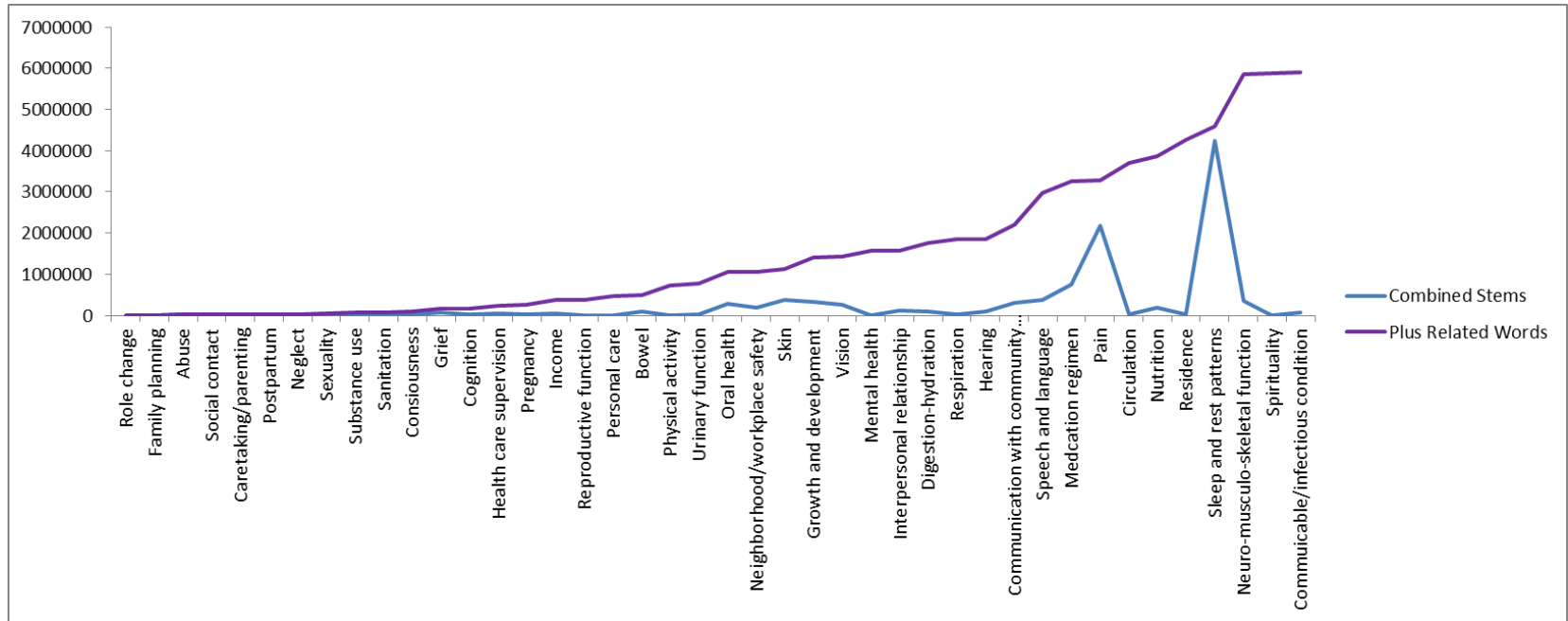
- showing problem terms and stems with $N > 100,000$



- Increases ranged from 7.4% to 381.2%
- Maximum = 6M

Step 4: Adding related words

- showing all problem concepts



- Increases ranged from 0.01% (Abuse) % to 1548.5% (Spirituality)

- Maximum = 6M

Preliminary Validation

- “Things looked fine far as the **skin** [*Skin*] color, oozing, swelling, and blisters, it’s all there, just in the right amounts, nothing to worry about.”
- “They will run some tests to see if **parainfluenza** [*Communicable/infectious condition*] is still cause of his **cough** [*Respiration*].”
- “He does not **walk** unless necessary like getting to car so he’s quite **weak** [*Neuro-musculo-skeletal function*].”

Other Important Observations

- Some CaringBridge authors express extreme emotions

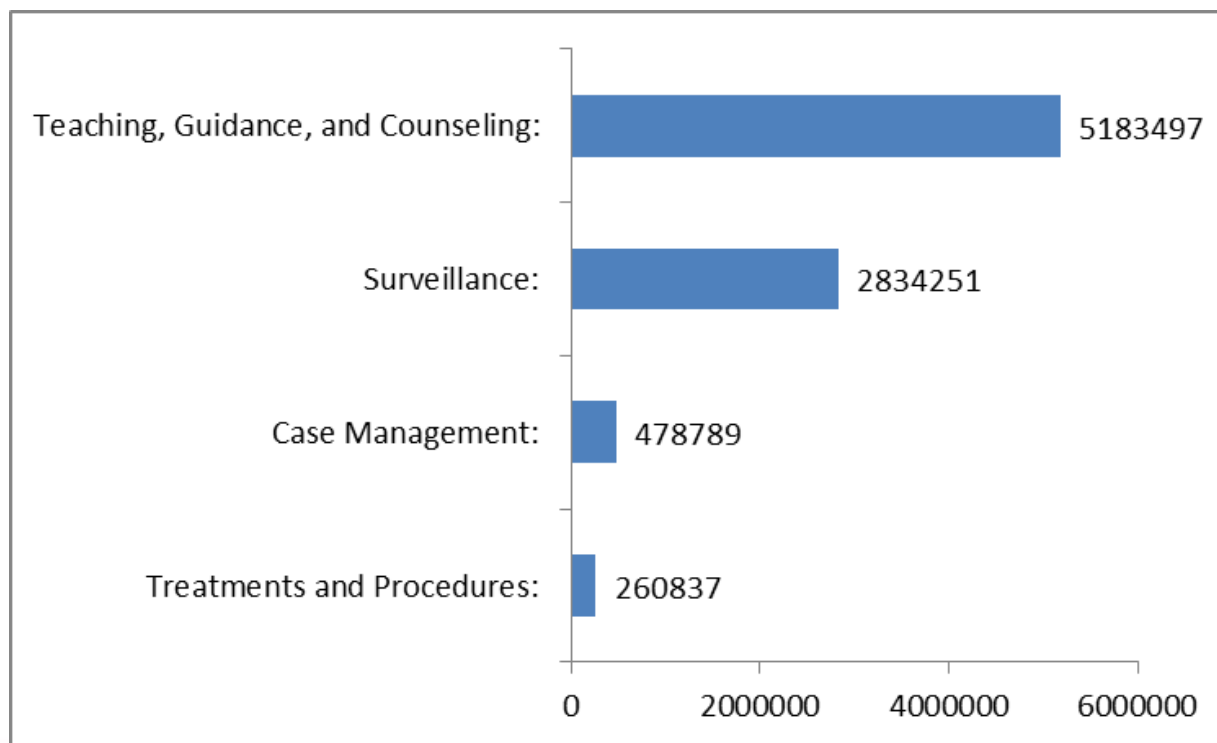
```
aaaaaaaaaaaaaaaaahhhh  
aaaaaaaaaaaaaaaaahhhbye  
aaaaaaaaaaaaaaaaaagony  
aaaaaaaaaaaaaaaaaaggggggggggh  
aaaaaaaaaaaaaaaaaada  
aaaaaaaaaaaaaaaaarrrrrrrrrrrrrrrrggggggggggggghhhhhhh  
aaaaaaaaaaaaaaaaarrrrrrrrrrrrrrrrggggggggggggghhh  
aaaaaaaaaaaaaaaaarrrgggggggg  
aaaaaaaaaaaaaaaaannnnnnny  
aaaaaaaaaaaaaaaaaallllllllllllllllllllll
```

Possible Intervention Topics

- Many people talk about
 - Sleep and rest
 - Pain
 - Nutrition
 - Medications

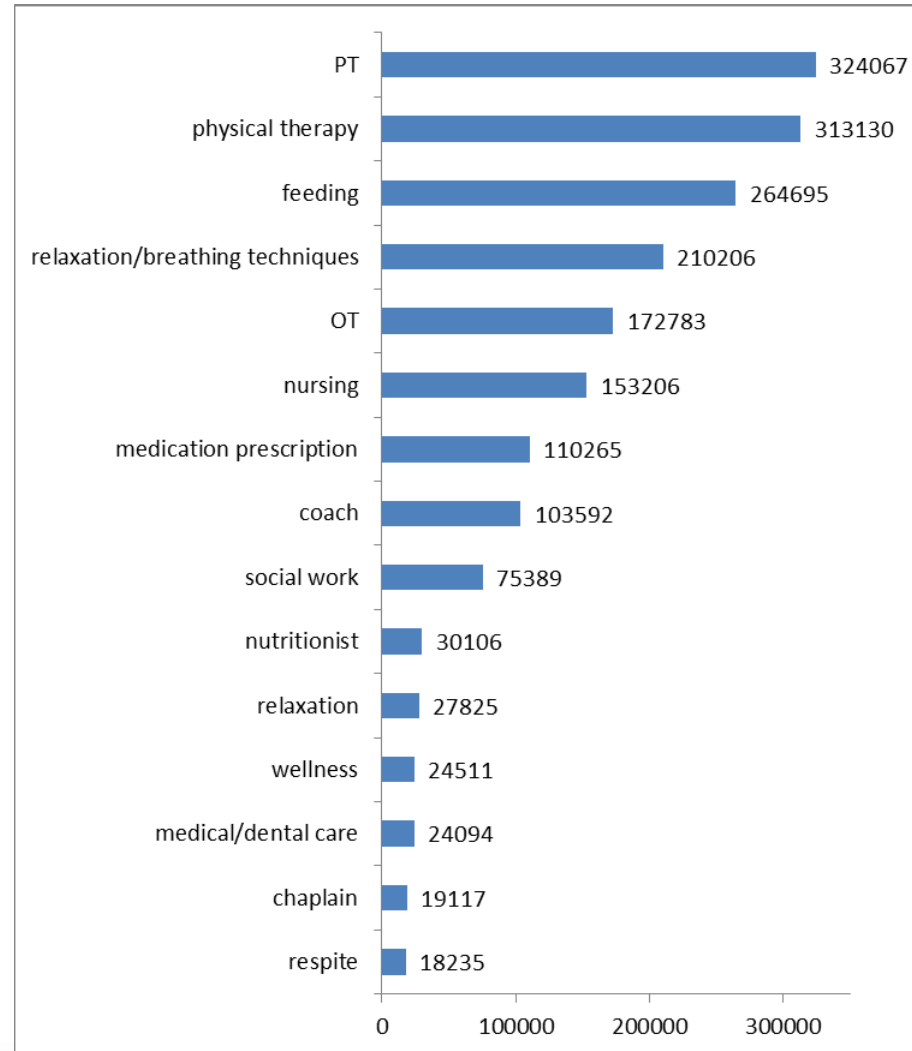
Intervention Terms: Categories

- Many people talk about



Intervention Terms: Target

- Many people talk about
 - PT
 - Feeding
 - Breathing techniques
 - OT
 - Nursing
 - Prescription
 - Coach
 - Social work



Limitations

- Challenges of interpretation:
 - Concepts may be positive or negative
 - Idiomatic bias: need to take words at face value, and not include some words: e.g. “See” could relate to vision or be used in many other idiomatic expressions
 - Semantic equivalence: many ways to say the same thing, difficult to capture all synonyms
 - Overlap: Despite the Omaha System’s taxonomic structure, there may be areas of overlap in some concepts expressed in natural language

Next Steps

- Validate meaning and concept saturation through review of randomly selected journals
- Automate model selection for each concept
- Tag each journal with problem concepts
- Cluster authors to identify meaningful subgroups
- Apply sentiment analysis techniques to understand whether concepts are positive or negative
- Identify intervention needs and outcomes

Model Selection Tool

74 Text Classification

Select the file (only .csv files) - Use 'Browse' option --> C:/Users/mons0122/Downloads/sleep_aggregated.csv

Algorithm -->

Feature Selection

Use Bag of Words as features

Use Tf-idf features

Use list of words(separated by comma) as features

Enter sentence to lemmatize :

Lemmatized sentence :

Test Options

K-fold K =

Train test % train split =

Supply test set

Output:

```
Accuracies for each fold:  
0.8888888888888888  
0.9629629629629629  
0.9259259259259259  
0.8888888888888888  
0.8846153846153846  
0.8846153846153846  
0.7692307692307693  
0.9615384615384616  
  
Average Accuracy: 0.895833333333
```

Data already cleaned

Preliminary Findings from Use of Model Selection Tool

- TBD

Applications for Use in Other Social Media Platforms

- Word2vec technique will identify words in any corpus associated with the standardized words
- Specialized vocabularies may need to be built for each platform
- Model selection tools can help identify best models to be used by concept, with each corpus and platform

Spinning Words into Data

- Implications for our future work
 - We hope to continue the process of tagging journals with 42 concepts, 4 intervention categories, and 75 intervention targets in order to apply typical clustering and analytics methods to text data as we would to quantitative data
 - These methods should be tested with other datasets from diverse platforms and populations

Conclusions

- The Omaha System text mining approach revealed differential representation of CaringBridge content from a whole-person perspective
- Further research is needed to extend this approach to inform important clinical questions and intervention opportunities

Thank you!

- Questions?
- mons0122@umn.edu

