Center for Financial Security
UNIVERSITY OF WISCONSIN-MADISON

Retirement & Disability Research Center

Nancy Wong
Professor, UW-Madison

Lydia Ashton
Assistant Professor, UW-Madison

Brett Puetz
PhD Student, UW-Madison

Jaeyoon Choi
PhD Student, UW-Madison

Using SSDI Conversations in Online Forums to Improve Communication and Outreach

Center for Financial Security

University of Wisconsin-Madison

1300 Linden Drive
Madison, WI 53706

608-890-0229
cfs@mailplus.wisc.edu
cfs.wisc.edu

The research reported herein was derived in whole or in part from research activities performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement and Disability Consortium. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA or any agency of the Federal Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof.

Abstract

Text analysis of data collected from online forum conversations, a form of user-generated content (UGC), reveals that Social Security Disability Insurance (SSDI) applicants and recipients, the “customers,” share concerns and confusion about the application and appeal process rules and policies. Extant research suggests that confusions about how SSDI rules are interpreted and applied significantly contribute to high SSDI rejection and appeal rates. This study attempts to provide insights into designing effective communication strategies to reduce confusion and improve customer service experiences and welfare.

Given the size of the data, we first use unsupervised machine learning algorithms to derive topics and model them using epistemic network analysis (ENA) via conversational connections. The resulting ENA provides insights on the structural relationships between different issues surrounding SSDI (e.g., struggles the applicants faced in communicating with and obtaining information from SSA). Taking advantage of the longitudinal nature of the data, we also model trajectory ENAs to investigate how these issues evolve against the backdrop of environmental and policy changes.

To provide deeper contextual value through human judgment, we use the derived topics as seed words in nCoder (an automated classifier). The resulting codes can be used in different applications, from analyzing the efficacy of existing policy to providing practical policy recommendations.

Keywords: SSDI, Online Forums, Text Analysis, Social Support
JEL Codes: **H530** National Government Expenditures and Welfare Programs - Disability Insurance, **H550** Social Security and Public Pensions, **J29** Other - Time Allocation, Work Behavior, and Employment Determination

1. Introduction

The Social Security Disability Insurance (SSDI) program's application process can be complex and time-consuming. Approximately two-thirds of applicants have claims that are initially denied. These claims take an average of two years to be resolved, and nearly two-thirds are eventually awarded benefits (Autor et al. 2015; SSA Annual Statistical Supplement 2019). This high initial rate of denial and subsequent award on appeal suggests the existence of a knowledge gap between applicants and the Social Security Administration (SSA). The extended length of the process produces a strain on applicants by delaying access to benefits; additionally, this adds administrative burden and preventable costs on the SSA. We collected and analyzed a data set of conversations on online forums relating to SSDI to explore this potential knowledge gap.

The study's two primary objectives are to provide insights on effective communication strategies to reduce confusion and improve customer service experiences and welfare. First, the analysis identifies the major areas of confusion about SSA rules and decision criteria using a machine-learning hybrid approach to natural language processing (NLP) and text analytics, and second, to evaluate the impact of how and when SSA customers obtain such information on their interpretation of this information.

We first compare the conversation patterns between the initial application and appeal processes of applicants. We use a text analytics approach called epistemic network analysis (ENA) to model the discussions of individuals participating in online forums related to SSDI, focusing on the difference between conversations of initial applications and appealing one's denial. The results suggest that being denied and going through the appeals process has stronger connections with pain and medical conditions and providing sufficient medical evidence. We next explore the impact of iClaims and field office closures on the conversation patterns among applicants capitalizing on the longitudinal nature of our data (pre-iClaim: 2004—2008, post-iClaim: 2009—2014). Results suggest that conversations shifted from focusing on questions surrounding medical evidence and suggestions (during the pre-iClaim period) to expressions of frustrations, mental health, and pain associated with medical evidence (during the post-iClaim period).

2. Literature Review

SSDI is a social insurance program intended to cover long-term disruptions to employment due to disability. The two main requirements for being awarded benefits are having a sufficient work history and documentation of a qualifying medical condition. Exploring conversations on online forums will provide insight into the challenges faced by SSDI applicants. Prior research of online discussions relating to disability found participants derived benefits such as gaining information, ability to discuss diverse and "taboo" subjects, and help in problem-solving (Finn 1999). Online forums provide information to people having difficulty obtaining services due to disability, overcoming geographic barriers, or with limited socialization opportunities.

Despite the potential issue for self-selection, analyzing UGC from online forums provides an effective way of understanding customers' misinterpretations of information and needs that are not determined by the researchers but grounded in the customers' actual experiences (Timoshenko & Hauser 2019). The self-selection nature of UGC can be an opportunity if the objective is to understand confusions and problems experienced in the application and appeal process, as customers are more likely to turn to such forums to seek help and share their experiences. The discussion threads in these forums will provide in-depth understanding of what information or advice is exchanged between applicants, forum contributors, and advisors (legal or medical). They provide real time progression of the application process and their eventual resolution. We can thus identify areas of confusion and the pathways of how these confusions were resolved.

2.1 Understanding the knowledge gap.

Research has shown that low levels of Social Security literacy, such as what people know about Social Security and their understanding about how the system works, significantly impact their choices and hence, their ability to receive full retirement benefits. In particular, factors such as age, income, and education significantly impact this knowledge gap and level of preparedness preventing applicants from leveraging resources available to them (Godtland et al. 2007; Yoong, Rabinovich and Wah 2015). Identifying specific areas of confusion and overcoming this knowledge gap will provide evidence for intervention pathways in communication and outreach efforts.

Jerit (2009) shows that differences in cognitive ability often lead to disparities in knowledge across low and high SES individuals. However, while expert opinions widen the knowledge gap on issues related to social security and Medicare, providing contextual information such as historical background, causes, and consequences of policies could narrow the knowledge gap between low and high education individuals (Jerit 2009). Therefore, it is critical to equip low SES individuals with the right contextual information to help them understand complex policy issues and enable them to negotiate the application process. It also provides the possibility of overcoming individual determinants of knowledge (e.g., education, income, age, race, and gender) with environmental determinants of knowledge (e.g., information on historical background and causes and consequences of policies). Analyzing online conversations surrounding SSDI applications could help identify where such knowledge gaps exist.

2.2 Community of Support

The socioemotional and task-oriented nature of the support shared in online forum conversations lends itself to our application of textual relationships and sentiment analysis to investigate discussions about problems and difficulties participants encounter in the application process for SSDI. Pharmaceutical companies have used similar applications, monitoring online forums for information related to patients' experiences of adverse drug reactions to reveal consumer reactions to new medications (Netzer et al. 2012). An analysis of relevant forums for the occurrences of specific words, such as "retirement," "PTSD," and "veteran," over time can reveal common and emerging trends and issues. Such observations could serve as a social listening post that can monitor applicants' ongoing discussions on the internet to extract and quantify user discussions to gain insight into the pinch points that lead to rejections and appeals and account for the bulk of applicants' frustrations.

3. Data and Methods

The data used in the analysis was scraped from seven online discussion forums between 2004 and 2020 focusing on SSDI (Federal Soup, FreeAdvice, Hadit, MSWorld, NeuroTalk, Physical Evaluation Board, and SSDFacts). Next, we segmented the data by each post, resulting in a total of 141,728 posts contained in 19,987 unique threads written by 9,015 unique authors. Because these forums are open to applicants seeking help, each thread generally starts with a question from one user. Other users respond by sharing advice or personal experiences and providing explanations or references to alternative resources or even emotional support.

We first used topic modeling, an unsupervised machine learning technique, that is, without predefined tags or training data previously classified by humans, that automatically analyzes text data to determine cluster words and discover abstract topics in a collection of documents, such as our UGC which is a collection of posts. The two most popular topic modeling algorithms are Non-Negative Matrix Factorization (NMF), which uses a linear-algebra based algorithm that performs dimensionality reduction and clustering simultaneously, and Latent Dirichlet Allocation (LDA), which uses a probabilistic approach. We derive topics using both of these algorithms. We find that terms included in the NMF are more contextually meaningful, more consistent with recent findings, and provide more coherent topics (O'Callaghan et al. 2015).

The emergent nature of the discussion topics means that we let the data "speak for itself" and choose a "cut-off" point based on the maximum topic coherence derived between two independent methods (LDA and NMF). Although the study's original intent was in exploring concerns or confusions related to policies on Windfall Elimination Provision (WEP), Government Pension Offset (GPO), the burden of Continuing Disability Reviews (CDRs), and effects of COVID-19 on SSDI. However, the first three topics have little or no mentions in these conversations. WEP is mentioned in only 242 (0.17 percent), and GPO is mentioned in only 20 (0.01 percent), out of 141,728 posts total. Discussions on COVID-19 only occurred toward the end of the data collection period. It would be more appropriate to isolate COVID-19 related conversations as a separate analysis in the future.

The topics derived from machine learning algorithm are next used as seed words in a supervised machine learning algorithm (nCoder) to identify codes for the associative network analysis. Next, we used Epistemic Network Analysis or ENA which extracts the meaningful features in the data operationalized as patterns of connections between codes within conversations (Shaffer 2017). ENA is well-suited to analyzing the data set of SSDI related conversations. It contains a rich array of experiences, and the large nature of this data set precludes manual analysis. We combine ENA with nCoder, a tool that makes it feasible for researchers to efficiently conduct qualitative analysis and transform a large amount of raw data into meaningful information using human judgment.

Shaffer (2013) developed ENA to model theories of cognition, discourse, and culture. However, researchers can apply the same method to different questions by modeling how groups of people frame, investigate, and solve complex problems. Our data fit the three assumptions necessary to apply ENA: (1) topics are meaningful features that can be systematically identified; (2) conversation threads provide a local structure; and (3) topics are connected to one another within threads (Shaffer and Ruis 2017). Therefore, ENA can model the organized knowledge in SSDI online forums and capture the relationships among topics by quantifying their co-currency within threads. The resulting networks can be analyzed by comparing them both visually and statistically.

We apply ENA to our data using the R package rENA (Marquart et al. 2019).[1] The ENA algorithm constructs a network model for each topic in the data by connecting it to topics within the recent temporal context. Our model defines the recent temporal context as each post plus the three previous posts within a given thread. ENA aggregates the resulting networks using a binary summation in which the links for a given post reflect the co-occurrence of each pair of topics. To visualize the network nodes and connections, ENA uses singular value decomposition (SVD) to decompose the structure of the data into a set of uncorrelated components in a high-dimensional space.

Fig. 1 shows a visualization of the ENA network for the last year in our data, 2019. The nodes correspond to the topics, and edges reflect the relative frequency of co-occurrence, or connections, between two topics. The group mean network, which averages points for each group, is plotted as a solid square surrounded by a larger square denoting the confidence interval (mean plotted point). The meaning of each axis in a given ENA space can be constructed by intuitively evaluating the placement of nodes. The position of the nodes is kept identical across plots, which allows the comparison of networks using the group mean of networks.

[1] An R package is a collection of code, data and documentation that users of R, a computing language and environment, can use to conduct specific tasks.

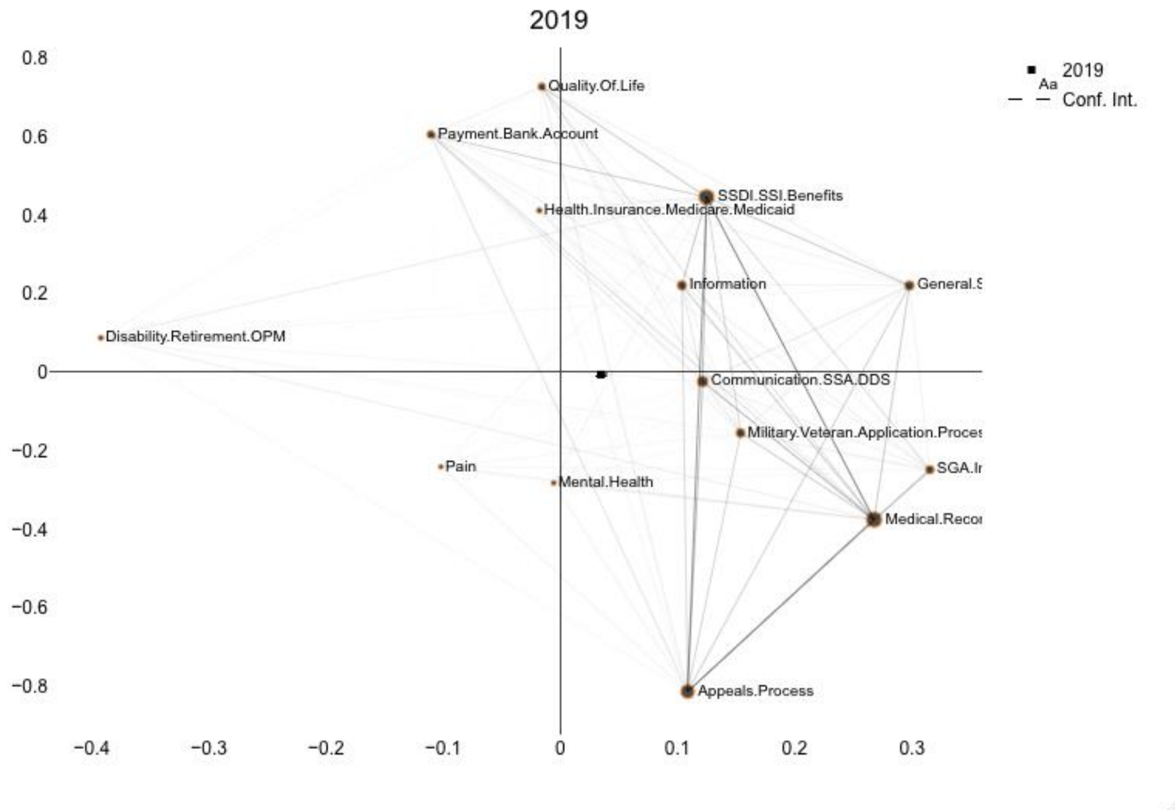

Fig. 1 Epistemic Network Visualization. The ENA network for 2019 shows a rich pattern of connections, however, the most prominent connections are between SSDI SSI Benefits, Appeals Process, and Medical records.

In general, moving from high to low along the y-axis indicates a shift from quality of life, SSDI and SSI benefits, and health insurance, Medicare, and Medicaid conversations toward conversations about the appeals process. Similarly, moving from right to left along the x-axis indicates a shift from general SSDI or SGA conversations toward conversations focused on disability retirement.

4. Results

4.1. Analysis 1: Comparison Between Application and Appeals Process

We segmented posts by sentence and coded them using the automated classifier nCoder (see Table 1). All codes had Cohen's $\kappa > 0.90$ and Shaffer's ρ (0.90) < 0.05 between a human rater and nCoder. This means that we achieve a statistically significant level of reliability between a human rater and their conceptual judgments and the automated classifiers. In other words, the measure of agreement for all codes between the human and nCoder were $\kappa > 0.90$, indicating a very high level of agreement with Shaffer's ρ (0.90) < 0.05 meaning that agreement generalizes to the whole dataset. We define conversations at the thread level and use infinite stanzas because of the permanence of existing posts within a thread. A robustness check using a large moving stanza was qualitatively similar to using the infinite stanza. We classified conversations as relating to initial application or denial appeals based on the more frequent code.

Table 1 SSDI forum-derived codes

Name	Definition	Example
DENIAL APPEALS	Refers to SSDI application denial or traversing the appeals process	*It was her lawyer that placed the paperwork in for her for the Reconsideration*
INITIAL APPLICATION	Refers to a new application; Code is mutually exclusive with DENIAL APPEALS	*My father applied for Social Security disability benefits back in June of this year.*
MEDICAL EVIDENCE	Refers to documentation and records needed to justify the qualifying medical condition	*Have you looked at your medical records to see what kinds of things your doctors are documenting?*
MENTAL HEALTH	Refers to mental health conditions; this may refer to listings in the SSA Blue Book	*I suffer from Major Depression and anxiety now, since my accident that has changed my life.*
NEUROLOGICAL CONDITION	Refers to neurological conditions; this may refer to listings in the SSA Blue Book	*My husband [was] diagnosed with Monomelic Amyotrophy recently.*
PAIN	Refers to the feeling of pain and descriptions of related sensations	*I can sit for an hour or so before my hips, shoulder and elbows start to hurt from arthritis*

Conversations classified under denial appeals exhibit stronger connections between pain, mental health, and medical evidence, while initial application discussions have more connections between neurological condition and medical evidence.

For example, **Error! Reference source not found.** illustrates these connections where User one talks about their denial and their documentation of medical conditions. Another theme of the data is the centrality of medical evidence to prove one's condition, as seen in the excerpt when User one talks about including every medical issue they have, and User two responding with advice to only include evidence relevant to the SSA's determination criteria.

Table 2 Qualitative Example relating to DENIAL APPEALS

User	Excerpt
User 1	Well in the denial letter **[DENIAL APPEALS]** this is what they wrote: "The medical evidence **[MEDICAL EVIDENCE]** shows that you have CFS, CTS, asthma, joint pain **[PAIN]**, memory loss, and visual loss"… I wrote in all of my conditions, symptoms and pains, aches **[PAIN]** and problems … down the side of the application **[INITIAL APPLICATION]** to get it all in.
User 2	The bluebook will give you some idea of what SSA considers to be disabling. Having looked at your recent post, what I would suggest is that you get your neurosurgeon and rheumatologist to write a letter, preferably in your medical records **[MEDICAL EVIDENCE]**, stating your Residual Functional Capacity.

The conversation type mean plotted points (see subtraction plot in Fig. 2) have a statistically significant difference. A two-sample t test assuming unequal variance of the y-coordinate values showed the INITIAL APPLICATION network (mean = 0.10, SD = 0.78, N = 296) was statistically significantly different at the alpha = 0.05 level from the DENIAL APPEALS network (mean = -0.06, SD = 0.75, N = 502; t(597.38) = 2.80, p = 0.01, Cohen's d = 0.21). This means that we found a statistically significant difference in the discourse patterns between the INITIAL APPLICATION threads compared to the DENIAL APPEALS threads.

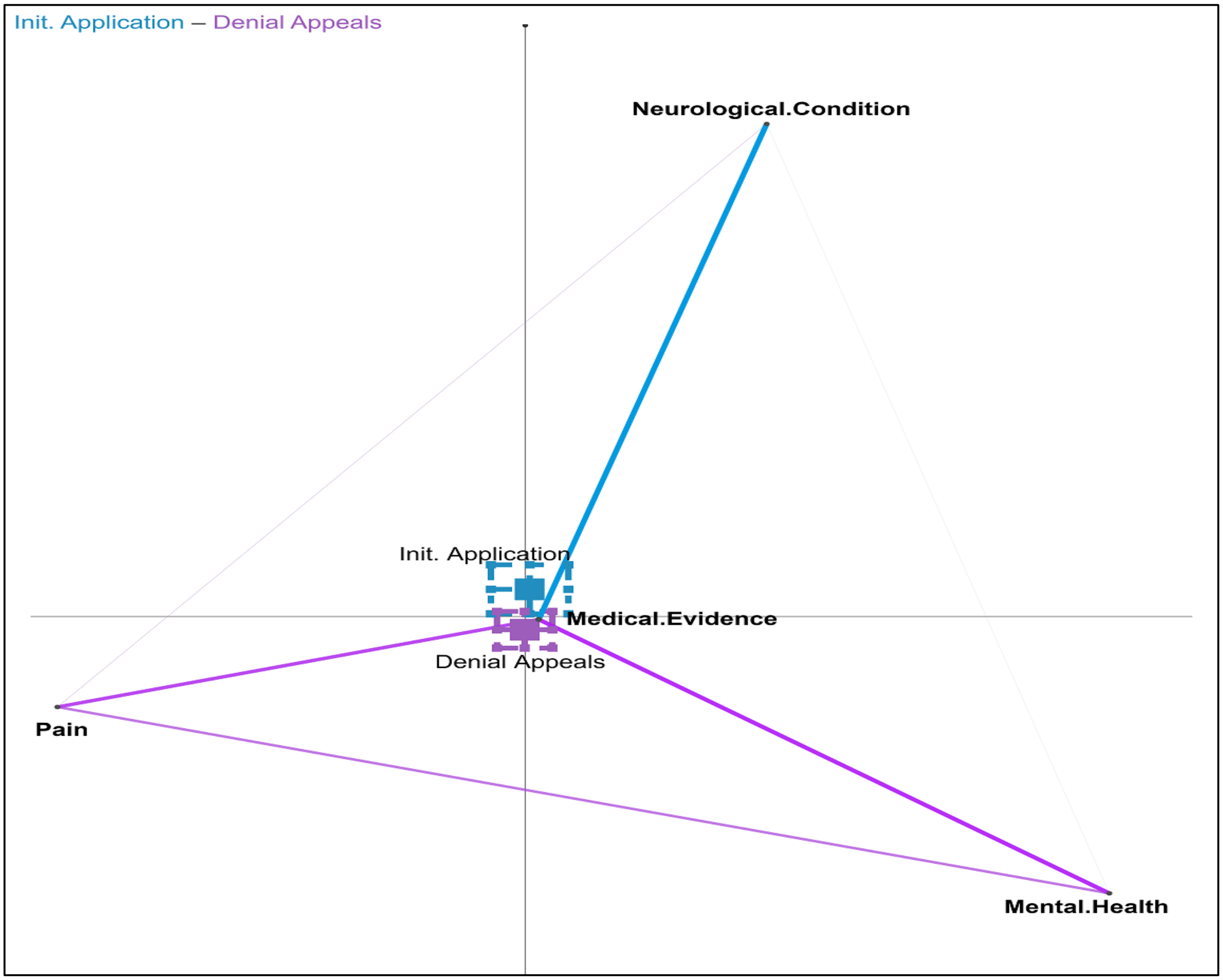

Fig. 2 ENA subtraction graph comparing the differences between INITIAL APPLICATION (blue) and DENIAL APPEALS (purple) classified conversations. INITIAL APPLICATION networks shows strong connections between the neurological condition node and medical evidence, while the INITIAL APPLICATION network shows strong connections between medical evidence, pain, and mental health nodes. Squares are group means; the dashed boxes are 95% confidence intervals (t-distribution).

These study results show systematic differences in conversations between filing an initial application and discussing denial appeals. Conversations containing initial application make stronger connections between neurological condition and medical evidence. In contrast, denial appeals make stronger connections among pain, mental health, and medical evidence. The increased connections between initial application and neurological condition suggest it is relatively easier to document the impacts of one's medical condition.

4. 2. Analysis 2: Introduction of iClaims and Field Offices Closures

In December 2008, the Social Security Administration (SSA) launched iClaim, an improved version of the online application system, intending to reduce application completion time, allowing third parties to access and submit applications, and facilitating access to application status. Foote et al., 2019 estimate that iClaim accounted for 7.5 percent of the 24 percent increase in SSDI applications nationwide from 2008 to 2011. However, it is unclear whether iClaim has streamlined the complex processes and alleviated the applicants' difficulties. To understand the challenges that SSDI applicants face regarding the iClaim program and how they communicate about them, we analyze and compare the discourse patterns in online discussion forums on SSDI during the pre- and post-iClaim periods.

To code the dataset, we first conducted a grounded analysis of the posts to find meaningful elements of the forum discourse, resulting in six codes (see **Error! Reference source not found.**”). Then we developed an automated coding scheme with regular expression matching.

Table 3 Codes from SSDI forums

Name	Definition	Example
SUGGESTION	Suggesting the next possible steps or solutions.	*"…. go to a disability lawyer first before going to your doctor and gathering any further medical documentation."*
RESOURCE	Referring to the external resources such as policy documents and helpful websites.	*"The website of the OPM is http://www.opm.gov/"*
QUESTION	Asking questions or seeking information from other users.	*"Do you have performance deficiencies because of your condition? …."*
FRUSTRATION	Expressing negative emotions, mainly frustration.	*"After being injured for three years, I am exhausted by all this."*
POSITIVE EMOTION	Expressing positive emotions.	*"So glad that I listened and filed for reconsideration and took on a proactive role in making sure that I had everything I needed and then some."*
COMMUNICATION	Referring to communication with SSA, from contacting a	*"I called local SSA last week & set up an appt via phone to make an application for them."*

	local office to submitting an application online.	

We use Epistemic Network Analysis to analyze the users' discourse in collaborative discussion forums. Because we are interested in investigating how individuals expressed their thoughts, we model the co-occurrences only within each post. We define pre-iClaim posts as those written between May 2004 and November 2008 and post-iClaim posts as those written between December 2008 and September 2014, the end of the Fiscal Year (FY) 2014. We excluded the period from October 2014 onward because the number of applications filed through the internet remained virtually unchanged until April 2020, the start of the COVID-19 pandemic. As a result, there were 5,999 Pre-iClaim posts and 82,929 Post-iClaim posts.

Table 4 illustrates a representative pattern of discourse in the Pre-iClaim and Post-iClaim periods. During the Pre-iClaim period, user *BL* first asked many questions to understand the post initiator's situation more clearly. Then, *BL* suggested some possible solutions (e.g., "follow the reconsideration") to solve the post initiator's problems. On the other hand, during the Post-iClaim period, user *bluerinse103* had an issue with their SSA online account, which led them to call the SSA and learn that they may need to drive a long distance to visit their field office, which causes them a lot of stress. The latter illustrates a situation in which the online application system, which was intended to ease the applicants' burden, causes the applicant feelings of frustration.

Table 4 Excerpts of a post written during the Pre-iClaim and Post-iClaim periods

Period	User	Post
Pre-iClaim	*BL*	*I think while your concerns are valid, you are overreacting at this point.* ***Did you file the reconsideration? …. Have you called*** *your Local Office to speak to someone there?* ***Is it possible for you [QUESTION]*** *to visit your Local Office to speak with someone in person? ….* ***Just make sure you follow the reconsideration [SUGGESTION]****.*
Post-iClaim	*bluerinse103*	*…. following their instructions to try again on Monday to give their system a chance to update the new information, I got the same message. So I* ***called them [COMMUNICATION]*** *again and explained everything …. So they verified all of my information and told me that ….* ***I should try again [SUGGESTION],*** *and then if it didn't work then,* ***I would probably need to visit***

	[SUGGESTION] *my local office. I was so* ***hopeful and relieved [POSITIVEEMOTION]*** *over the weekend, and now I have to deal with the prospect of driving a long distance … but this is just so* ***frustrating [FRUSTRATION]***

The mean plotted points for the posts written during Pre-iClaim and Post-iClaim in Fig. 3 suggest that the two groups are different in terms of their positions on the x-axis. A two sample t-test assuming unequal variance of the x-coordinate showed that Pre-iClaim group (mean= -0.21, SD= 2.91, N=5999) was statistically significantly different at the alpha 0.05 level from Post-iClaim group (mean= 0.04, SD=2.55, N=82929; t(6683.3)= -6.6802, p=2.579e-11, Cohen's d= 0.09). This means that there was a statistically significant difference in the discourse patterns between the Pre-iClaim posts compared to the Post-iClaim posts.

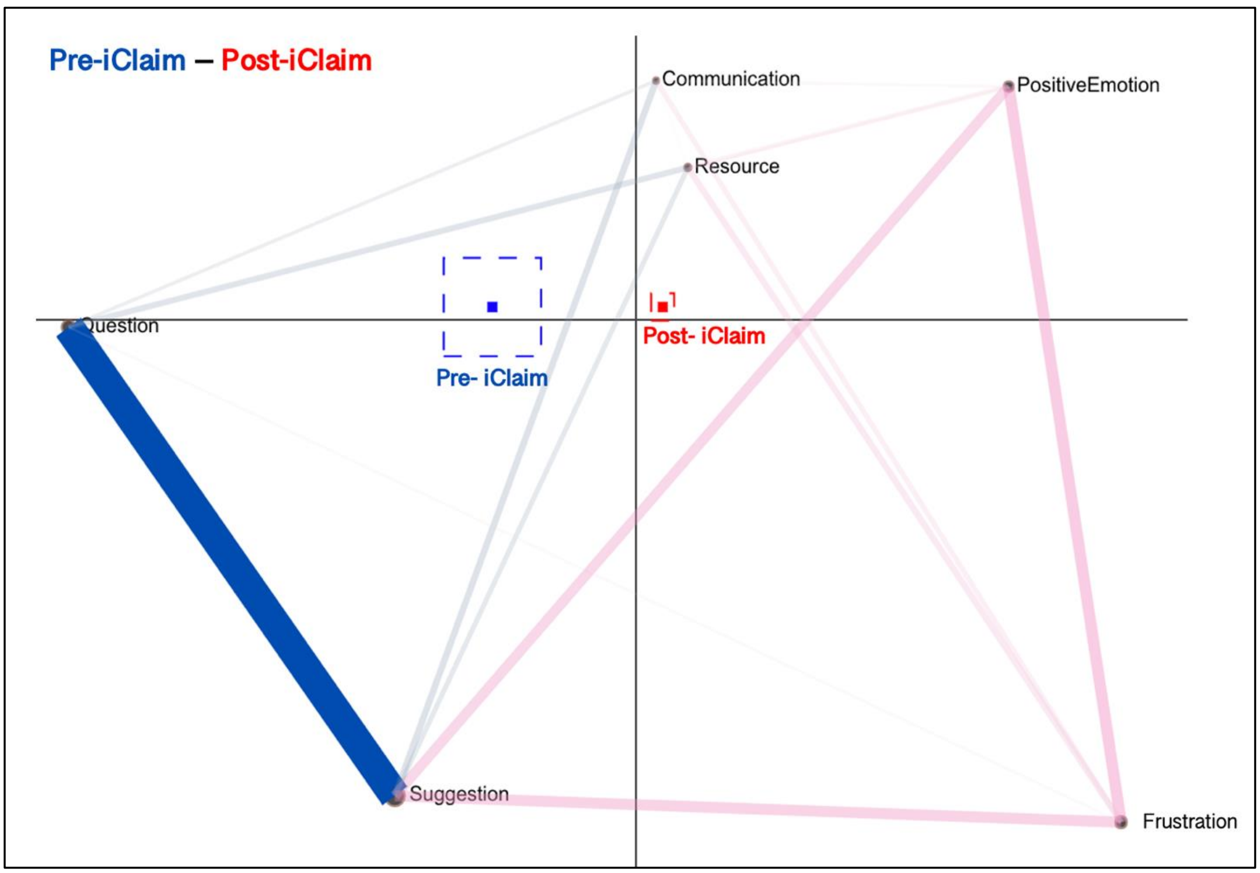

Fig. 3 ENA subtraction graph comparing the differences between Pre-iClaim (blue) and post-iClaim (red) classified conversations. The Pre-iClaim network shows strong connections between the Question and the Suggestion node, while the Post-iClaim network shows strong

connections between Suggestion and Frustration. Squares are group means; the dashed boxes are 95% confidence intervals (t-distribution).

The Pre-iClaim network plot shows a strong connection between QUESTION and SUGGESTION. In contrast, the Post-iClaim network plot shows a strong connection between SUGGESTION and FRUSTRATION. As the mean subtraction plot shows, Pre-iClaim posts tend to mention COMMUNICATIONS and RESOURCES in the context of interactional aspects. In contrast, Post-iClaim posts are more likely to mention COMMUNICATIONS and RESOURCES in the context of emotions.

This study shows how Pre-iClaim posts and Post-iClaim posts reveal different discourse patterns in online discussion forums on SSDI. When referring to communication with the SSA, the Pre-iClaim posts focus more on the interactional aspects, whereas Post-iClaim posts concentrate more on the emotional aspects (e.g., "Yes 4 years, I lost my house and was beyond frustrated" and "I'm about to have back surgery in 10 days and a lot of time has passed because I got discouraged and almost gave up"). The results suggest that applicants face particular challenges when using iClaim that put an emotional strain on them. For example, applicants may expect that using iClaim would eliminate the effort of visiting the field office but end up having to call or visit their field office anyway after encountering some difficulties with the online application system may experience feelings of frustration. Future studies could further examine why and how iClaim could be putting pressure on SSDI applicants. For example, different state Disability Determination Services (DDS) have different technology infrastructures to access the Health Information Technology (HIT) for them to access the electronic medical records of the applicants in order to collect medical evidence in using the iClaim system (SSAB Roundtable on Medical Evidence Collection 2021). These technological disconnects (gaps) contribute to the significant delays (having to resort to mail and paper copies) and frustrations.

5. Discussion

Findings from this study suggest that we can learn much from the socioemotional and task-oriented nature of support that is shared in online forum conversations. In Analysis 1, comparing discussion patterns between initial applications and those negotiating the appeals process indicates significant differences exist between the types of information exchanged. We find that denied applicants struggle to adequately document their underlying medical conditions with higher relation to pain and mental health. These conditions may be more difficult to document and justify due to a lack of a physical diagnosis. Additionally, the individual circumstances of such conditions may place a larger burden on an applicant to thoroughly document their condition. Therefore, individuals may not be fully cognizant of how important documentation is and the potential difficulty of documenting some conditions to provide the necessary medical evidence. Consequently, it may be imperative to give more explicit guidelines regarding the level of evidence and documentation, especially for conditions that may require differing proof. Changes like this could significantly impact the number of initially denied applicants but subsequently awarded benefits upon appeal—the derived program efficiencies helping both the SSA and SSDI applicants. A possible extension of this work would be to explore the medical evidence code further. Currently, this code does not explain much variation along either axis due to its omnipresence. Therefore, useful information is likely to be observed from separating this code to examine different medical conditions and the type of medical evidence needed.

In Analysis 2, the comparison between pre-iClaim and post-iClaim conversations suggest that there is significant increase in negative emotions largely attributed to the burden of negotiating the application and the appeal process. The field office closures which iClaim could help support actually exacerbated the assistance that applicants needed for clarifications and assistance. The different levels of digital access further complicate the process for applicants in submissions and for collection of medical evidence to support their applications, consistent with our findings in Analysis 1. The current findings suggest specific steps of the application (i.e., functional report) and verification process (i.e., consultative examination) in the iClaim system requires significant technological support in each state's electronic records exchange (ERE) system with SSI.

The current findings suggest that the public policy sphere faces barriers to addressing societal challenges, including better serving people who most need support, particularly people with disabilities. Consequently, exploring online conversations among this vulnerable group provides a unique opportunity to identify and understand the specific barriers that cannot be solicited from surveys but coming directly from the customers' experiences that they share with others in seeking help, and point to specific pathways toward creating solutions.

5.1. Implications for Vulnerable Populations

Armour (2018) finds that the availability of online applications increases the application rates of medically marginal applicants. About 18 percent of the increase in DI applications (1992—2004) can be attributed to the benefits statement. This increase suggests the importance of informational costs in DI, suggesting that information provision is an important policy lever among the population covered by DI. However, the DI application process is also plagued by information asymmetry, such that 20 percent of applicants accepted into the program are work-capable while 60 percent of rejections meet disability requirements (Benítez-Silva, Buchinsky, and Rust 2004). The current study aims to reduce the information asymmetry of the 60 percent rejected who meet disability requirements since these applications were subsequently approved.

5.2. Limitations

One limitation of this study is it most likely does not represent the entire population of SSDI applicants despite containing many observations spanning more than a decade. Additionally, this study only examines one aspect of a complex process and may not capture more important or interacting facets of the application and appeals process.

Future work should evaluate the representativeness of the sample of individuals participating in online forums. For example, the share of SDDI applicants who are military or veterans could be compared to the number of users in veteran- and military-oriented forums. Weighing the data to reflect the sample representation would be ideal, but the potential of future applications is not limited since the profile of SSDI applicants on public forums may change. Still, that change is not significant to this analysis, which provides a longitudinal view of posts and unique threads over time, reflecting changes in the environment in terms of economic and demographic trends.

Certain SSDI applicants may not have access to online forums due to lack of a computer, access to high-speed internet, computer skills, or assistive devices such as screen readers. However, some access issues are probably less of a concern given the rise of smartphones, inexpensive notebook and tablet computers, and other devices. Online forums are also relatively undemanding in terms of internet bandwidth compared to, for instance, video streaming and gaming. However, further analyses of the online forum data point to the digital divide driven by computer usage skills and broadband access beyond simple internet access.

Finally, a fundamental trade-off of applying automated text analysis to this research is the size versus representativeness of the observations. Representativeness in this case concerns (1) the sampling of the participants; (2) the sampling of the environments (the forums selected); (3) the kinds of issues participants are exposed to in given environments; and (4) the states of minds and behaviors participants can express in a given environment (Mahmoodi et al. 2017).

5.3. Future Directions

According to Bonfadelli (2002), the knowledge gap is a key consequence of the digital divide (or digital inequality), the gap between advantaged and disadvantaged computer users and nonusers. One of the general arguments is the “Matthew Effect” (the rich get richer); with digital access, demographic differences, access, skills, interests, and infrastructure could represent costs and barriers, with more benefits flowing to those with greater abilities, resources, and information needs. Similar to the Matthew Effect is the knowledge gap hypothesis. Those with the most resources adopt first, have and gain more skills, and use more and different activities more effectively, thereby increasing and not reducing the knowledge gaps in society (Pearce and Rice 2013).

The usage gap divide due to unequal access to technology argues that factors affecting the access gap also affect the usage gap, but at different levels and other usage types. Differential outcomes from this gap create feedback loops that may increase and institutionalize such differences. Existing social inequalities thus both affect and reinforce various digital divides. Mobile internet may overcome many infrastructural differences between urban and rural and developed and less-developed regions, as wireless connectivity requires far less infrastructure. However, PC-based internet allows for a more optimal experience while mobile-based internet access, although more accessible and affordable, requires compromises.

Pearce and Rice (2013) find that PC-based internet access provides health and government interactions, personal development, news, and information considered social capital enhancing. Mobile-based internet access offers music and videos, commercial transactions, and social communication as consumptive or entertaining. Capital enhancing use of the web improves life chances that include seeking political or government information online, exploring career or job opportunities, or financial or health information seeking.

6. Conclusion

Internet forums and other social media platforms facilitate online communication in an open context, allowing users to share their feelings, experiences, and advice in an informal, nonthreatening environment. As a result, participants may provide information about individual experiences with and thoughts about SSDI that is unlikely to be gained from formal surveys. This hypothesis is supported by online social support (OSS) theory, which states that individuals seek social support when they are confronted with acute stressors. Online forums provide access to this support through task-oriented discussions. SSDI applicants and beneficiaries are financially vulnerable and may feel stigmatized.

References

U.S. Social Security Administration. *Annual Statistical Supplement to the Social Security Bulletin*. By Mark J. Warshawsky and Katherine N. Bent. No. 13-1170. Washington, D.C.: Social Security Administration, 2019. https://www.ssa.gov/policy/docs/statcomps/supplement/2019/index.html, accessed February 25, 2020.

Armour, P. (2018). "The Role of Information in Disability Insurance Application: An Analysis of the Social Security Statement Phase-In." *American Economic Journal: Economic Policy*, 10(3): 1—41. https://doi.org/10.1257/pol.20160605.

Autor David, H., Nicole, M., Mullen Kathleen, J., & Alexander, S. (2015). Does delay cause decay? The effect of administrative decision time on the labor force participation and earnings of disability applicants. No. w20840.

Benítez-Silva, Hugo, Moshe Buchinsky, and John Rust. (2004). How Large Are the Classification Errors in the Social Security Disability Award Process?, *National Bureau of Economic Research* (NBER) Working Paper 10219.

Bonfadelli, H. (2002). "The Internet and Knowledge Gaps: A Theoretical and Empirical Investigation." European Journal of Communication, 17: 65—84. https://doi.org/10.1177/0267323102017001607.

Finn, Jerry. (1999). "An Exploration of Helping Processes in an Online Self-Help Group Focusing on Issues of Disability. *Health & Social Work*, 24(3): 220–31.

Foote, A., Grosz, M., & Rennane, S. (2019). The Effect of Transaction Costs on SSDI. Journal of Policy Analysis and Management, 38(1), 99-123.

Godtland, Erin M., Grgich, Michele, Petersen, Carol Dawn, Sloane, Douglas M., and Walker, Ann T. (2007). "Racial Disparities in Federal Disability Benefits." Contemporary Economic Policy, 25(1): 27—45.

Jerit, J. (April 2009). "Understanding the Knowledge Gap: The Role of Experts and Journalists." *The Journal of Politics*, 71(2): 442—456. https://doi.org/10.1017/S0022381609090380.Mahmoodi, J., Leckelt, M., van Zalk, M. W., Geukes, K., and Back, M. D. (2017). "Big Data Approaches in Social and Behavioral Science: Four Key Trade-Offs and a Call for Integration." *Current Opinion in Behavioral Sciences*, 18: 57—62.

Marquart, C. L., Hinojosa, C., Eagan, B., Siebert-Evenstone, A., & Shaffer, D. W. (2020) nCoder (Version0.2.2.0) [Software]. Available from http://app.n-coder.org.

Marquart, Cody L., Swiecki, Zachari, Collier, Wesley, Eagan, Brendan, Woodward, Roman, and Shaffer, David Williamson. (2019). "RENA: Epistemic Network Analysis." https://CRAN.R-project.org/package=rENA.

Netzer, Oded, Feldman, Ronen, Goldenberg, Jacob, and Fresko, Moshe. (2012). "Mine Your Own Business: Market-Structure Surveillance Through Text Mining." *Marketing Science*, 31(3): 521–43.

O'Callaghan, Derek, Greene, Derek, Carthy, Joe, and Cunningham, Pádraig. (2015). "An Analysis of the Coherence of Descriptors in Topic Modeling." *Expert Systems with Applications*, 42(13): 5645–57.

Pearce, K. E., & Rice, R. E. (2013). Digital Divides from Access to Activities: Comparing Mobile and Personal Computer Internet Users. *Journal of Communication, 63*, 721-744.

Shaffer, David Williamson. (2003). Pedagogical Praxis: Using Technology to Build Professional Communities of Practice." *ACM SIGGROUP Bulletin*, 24(3): 39–43.

Shaffer, D. W. and Ruis, A. R. (2017). "Epistemic Network Analysis: A Worked Example of Theory-Based Learning Analytics." In *Handbook of Learning Analytics*, edited by C. Lang, G. Siemens, A. F. Wise, and D. Gasevic, pp. 175—187.

Shaffer, D. W. (2017). *Quantitative Ethnography*. 1st edn. Cathcart Press: Madison, WI.

SSAB (2021) https://www.ssab.gov/research/roundtable-medical-evidence-collection-the-cornerstone-of-social-security-disability-determination/

Timoshenko, A., & Hauser, J. R. (2019). Identifying Customer Needs From User-Generated Content. *Marketing Science*, *38*(1), 1-20.

Yoong, J., Rabinovich, L., and Wah, S. H. (2015). "What Do People Know About Social Security?." CESR-Schaffer Working Paper 2015-022.

Center for Financial Security

School of Human Ecology
University of Wisconsin-Madison

1300 Linden Drive
Madison, WI 53706

608-890-0229
cfs@mailplus.wisc.edu
cfs.wisc.edu