Center for Financial Security
UNIVERSITY OF WISCONSIN-MADISON

Retirement & Disability Research Center

Trends in Conversations about SSDI in Online Forums

Nancy Wong
Professor

Lydia Ashton
Assistant Professor

Brett Puetz
Doctoral Student

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 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

Applications to disability insurance (DI) have declined in recent years but extant research sheds little light on what is driving these trends. Research surveys and interviews based on self-reported data may not reveal more personal situations or include financially vulnerable populations. This study will help address these limitations by using a text analytics and text analysis approach to explore how individuals communicate with each other about DI on internet forums. Online forums and other social media platforms facilitate online communication in an open context. These communication platforms enable users to share their feelings, experiences, and advice in an informal and nonthreatening environment; as a result, they may provide information that is not available from formal surveys.

We collected data on online conversations that mentioned SSDI from seven online forums for the period 2004–2019. We then created a master data set of approximately 150,000 posts in roughly 15,000 unique threads. We conducted text analytics and text analysis to identify term and word frequencies, as well as topics modeling using unsupervised machine learning. We also developed preliminary collective associative networks (CANs) to delve deeper into the data. After describing our methods, we provide a summary of the findings and recommendations based on the outcomes.

1. Introduction

As of December 2016, 11,832,337 people received Social Security Disability Insurance (SSDI) benefits, including disabled workers, disabled widows and widowers, and disabled adults. This number represents a significant increase from 1970, when SSDI supported 1,812,786 recipients; that increase is driven mainly by an increase in the number of disabled workers, representing about 74% of total recipients in 2016. Adult recipients (18–64) represent 4.7% of the US population with Alabama, Arkansas, Kentucky, Maine, Mississippi, and West Virginia having the highest rates of SSDI beneficiaries (7% or more)~~.~~ However, this number seems to be trending downwards from the high of 12,156,191 in 2013 (https://www.ssa.gov/policy/docs/statcomps/di_asr/2016/sect01.html#)~~.~~

To explore why this number is falling, this study gleaned insights on SSDI claim trends by using text analytics and text analysis with the support of machine learning to extract contextual information from conversations in online forums. Online forums and other social media platforms facilitate communication in an open context. These platforms enable users to share feelings, experiences, and advice in an informal and nonthreatening environment, providing access to information that is unlikely to be gained from formal surveys. In addition, informal communication channels such as online forums are also particularly accessible to members of vulnerable communities, a group that can be difficult to survey. Specifically, the study aims are 1) to collect data from online forums where SSDI is discussed and use text analytics and text analysis on that data to generate chronological collective associative networks (CANs), or “organized knowledge,” regarding the DI application process; 2) to examine how changes in CANs coincide with changes in the broader economic environment as well as changes in SSDI policies during the period of observation; and 3) to develop recommendations for leveraging data from online forums to expand applicants’ access to and use of SSDI information.

2. Literature Review

People turn to online forums for many reasons. For people who suffer from disabilities and are homebound, temporary or permanently, online forums can help overcome the social and emotional difficulties of social isolation (Finn 1999), such as depression, loneliness, alienation, lack of social interaction, lack of information, and lack of access to employment. These self-help groups are attempts by people who share a problem to take control over the circumstances affecting their lives. This form of reaching out is particularly common in the health care arena, with online outlets addressing a variety of topics, ranging from chronic diseases, long-term rehabilitation, and terminal illness to addiction and rare diseases.

Online forums also offer a degree of anonymity, which allows members, especially those with stigmatizing issues (for instance, mental illness or AIDS), to more easily and safely explore these topics. The social support, resource sharing, and collaborative problem solving contribute to participants' sense of empowerment and identification with these online communities, which often translates into greater engagements in their offline lives and civic activities (Pendry and Salvatore 2015). Hence, it is not surprising that the past two decades have seen tremendous growth in online forums in the health domain, from the thousands in the late 1990s to hundreds of thousands by 2012 (Wright 2016). According to the National Cancer Institute (2013), one in six adult Americans participated in a health-related peer support community in 2012.

2.1 Online Social Support Theory

Review of the research clearly shows the significant contribution of online forums in participants' development of successful coping behaviors. A key advantage of online support groups is the access it gives participants to larger number of individuals with more diverse experiences and knowledge, compared to face-to-face support groups. This wider exposure is particularly critical in the problem-focused dimensions of online coping, in terms of active coping approaches, skill-building, and enhanced self-efficacy (Wright 2016).

Research has also found that the online disinhibition effect greatly accelerates development of personal and interpersonal dynamics compared to the traditional support group setting. Activities such as writing, expressing emotions, collecting information, improving understanding and knowledge, developing social relationships, and enhancing decision-making skills foster participants' sense of control, empowerment, and self-confidence. Although risks of overdependence, leading to distancing from in-person contacts, and other unpleasant experiences such as stalking or bullying exist, the benefits generally outweigh the risks (Barak, Boniel-Nissim, and Suler 2008).

2.2 Benefits

In an early study of conversations in a national online group on disability, Finn (1999) found that participants derived benefits such as information (facts, resources, ideas), dialogue on diverse issues, discussion of "taboo" subjects, a sense of "being in the same boat," mutual support, mutual sharing of experiences, problem solving, reduced sense of alienation and isolation, a sense of helping others, development of inspiration and hope, and development of social networks. For members of vulnerable populations, online forums are particularly helpful in providing support and information to people who have difficulty obtaining services due to physical or mental disabilities or difficulty in accessing services due to geographic barriers, lack of transportation, verbal communication difficulties, or limited socialization opportunities.

Consequently, exploring conversations on these online forums will likely provide unique insights on the challenges faced by this population that are unavailable elsewhere. While the anonymity of online forums yields clear benefits, that anonymity also brings risks associated, such as an increased potential for deception, manipulation, cyber-surveillance, and other negative behaviors (Barak, Boniel-Nissim, and Suler 2008).

2.3 What Is Shared

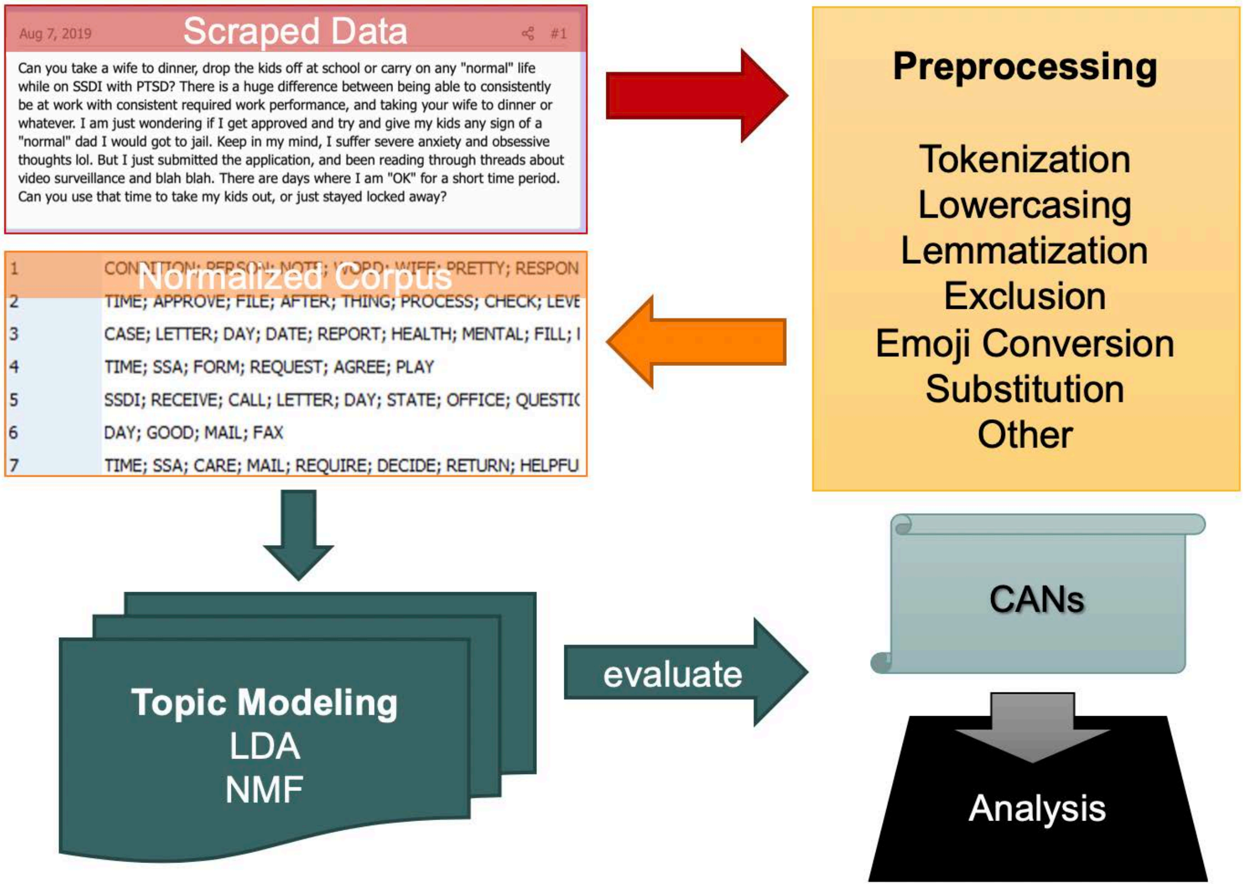
The socioemotional and task-oriented nature of the support shared in online forum conversations lends itself to our application of textual relationships and sentiment analysis to investigate discussions of SSDI in relation to problems and difficulties participants encounter in the application process. Similar to how pharmaceutical companies have monitored online forums for information related to patients' experiences of adverse drug reactions to reveal consumer reactions to new medications (Netzer et al. 2012), analysis of relevant forums for the occurrences of specific words, such as "retirement," "PTSD," "veteran," over time can reveal common and emerging trends and issues. Such observations could serve as asocial listening post that can monitor applicants' ongoing discussions on the internet with the goal of extracting and quantifying 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

In this study, we construct an original data set by gathering user-generated content (UGC) from online forums discussing SSDI. This UGC takes the form of a collection of posts discussing particular topics, our corpus. We then use an untrained machine learning algorithm to extract topics ("topic modeling") from this corpus. These topics are represented by a list of frequent or relevant terms, which can provide a semantically meaningful interpretation of the latent or hidden structure of the corpus. We then use these topics and the terms that connect them to develop chronological CANs, or a map of collective knowledge, regarding the SSDI application process.

We also collect data on economic and environmental determinants, such as unemployment rate and SSDI policy changes, to examine how changes in this collective knowledge map coincide with changes in the broader economic environment, as well as changes in DI policies, during the period of observation. Figure 1 illustrates the data collection process.

Figure 1: Data and Methods



3.1 User-Generated Content

Online forums provide a platform for informal conversations in a nonthreatening environment; they provide an ideal setting to explore discussions surrounding SSDI. Previous work has established that data from UGC, such as that found on forums, is at least as valuable as data derived from more conventional research methods, like interviews and surveys (Timoshenko and Hauser 2019). Additionally, data from online discussion forums has been shown to be useful for describing consumer behavior (Way, Wong, and Gibbons 2011). More than simple repositories of information, online forums often provide a place for individuals to find support (Braithwaite, Waldron, and Finn 1999). Thus, the data collected should provide a holistic view of discussions surrounding SSDI, as opposed to focusing on simple technical questions pertaining to the application process.

a) Forum identification

The data from the online forums is described using specific terms. The smallest unit is a post, which is a message generated by a single user. A thread is composed of a sequence of posts discussing a particular topic. In turn, forums are a collection of an array of related threads covering a general theme. Each forum addresses any number of topics. For the purposes of data collection, topics from each forum were selected based on their relevance to SSDI. This assessment provided a data set broad enough to capture a wide array of discussions but focused enough to minimize irrelevant data.

In order to determine which forums to collect data from, an initial search was followed by a selection process to narrow the field. The research team worked independently to identify candidate forums that discussed SSDI and included a community group. This search yielded a wide list of forums, although some forums appeared on multiple lists. Further information was then collected about each forum, including forum sponsorship, context related to SSDI, top keywords, top outgoing traffic, top referring forums, time visitors spent on forum daily.

b) Forum selection

Based on the additional information, the initial list was reduced from 22 to 7 forums.[1] While the quantity of potential data was important, it was also necessary to choose as wide a range of contexts around SSDI to ensure as rich a data set as possible. To that end, we focused on a diverse set of forums:

- *Federal Soup* (federalsoup.com)

Federal Soup provides general employment information for employees of the federal government. The forum is not part of the government, but it is a registered government vendor that provides services to government agencies. The site hosts a dedicated forum for disability retirement, with discussions centering around SSDI and the interplay between SSDI and the Federal Employees Retirement System (FERS).

---

[1] See Table A.1 for a complete list of sites identified to contain relevant SSDI conversations. We dropped forums where SSDI-related conversations were too spare or scarce.

- *FreeAdvice* (forum.freeadvice.com)

FreeAdvice provides general legal advice covering many aspects of the law. In addition to the forums, the site also offers articles and FAQs on legal issues and resources for finding and accessing legal services. The disability-related forum on FreeAdvice includes threads on both SSDI and Supplemental Security Income (SSI).

- *Hadit* (hadit.com)

Hadit focuses primarily on United States military veterans, with a strong emphasis on those who have already transitioned to civilian life. The site hosts a dedicated SSDI forum where users discuss their experiences with SSDI and offer help to those currently navigating the process. This site also offers other sources of information, including blogs and news.

- *MS World* (www.msworld.org)

MS World is an online community addressing the needs of individuals whose lives have been affected by multiple sclerosis (MS). The forum provides resources addressing all facets of MS, including the disease itself, related news articles, and a creative center where individuals share artistic projects. The site’s dedicated SSDI forum is partially moderated by a volunteer attorney, who answers legal questions.

- *NeuroTalk* (www.neurotalk.org)

NeuroTalk is an online support community that is part of the larger PsychCentral Community Connection. The forums provide information and support for issues around mental illness and neurological conditions. The site has a dedicated SSDI forum where users create threads related to SSDI issues.

- *Physical Evaluation Board* (www.pebforum.com)

The Physical Evaluation Board (PEB) site caters primarily to veterans of the United States Armed Services. In addition to forums regarding policies and procedures for the armed forces, the site hosts a dedicated SSDI forum. This forum is contained within a larger conversation about the transition process from active duty to civilian life. Consequently, it contains many threads concerning the interplay between the Department of Defense, the Department of Veterans Affairs, and the Social Security Administration (SSA).

- *Social Security Disability Facts* (www.ssdfacts.com)

Social Security Disability Facts website is dedicated to Social Security disability programs. The site hosts an expansive set of forums relating to nearly all aspects of the application, appeals, and Continuing Disability Review (CDR) processes, as well as general resources and disability-related issues. The site also offers links to forms and resources offered by the SSA.

c) Data collection

To facilitate the data collection, we use the Rcrawler software package (Khalil and Fakir 2017) within the general R programming framework. This approach provides multiple efficiencies in data collection and analysis. The Rcrawler software package allows simultaneous web crawling and scraping, which means data are collected as the program traverses the forum pages identified as being of interest. This method increases the speed and simplicity of data collection by combining work previously done in separate steps. The efficiency of data collection can be increased further by instructing Rcrawler to collect only specific content from the forums it transverses using the forum URL structure (which indicates how the web addresses for the forum pages are organized) and element patterns (the names given to different web page components).

To create the master data set, the collected data are cleaned and formatted. The first step involves removing duplicate entries from the collected data. Then, special characters used for formatting, such as the tab (\t) and newline (\n) characters, are removed, as is quoted text within the post content where possible, to reduce potential bias within the data set. As the different forums contain different data in differing formats, the data are then harmonized into a consistent format; an important aspect of this formatting is ensuring a common date format. To finalize the master data set, the cleaned and formatted data are merged for further processing and analysis.

d) Summary statistics

Table 1 provides summary statistics for the selected forums, including forum sponsorship type, which refers to how each forum is funded. Ad-revenue forums rely solely on advertising revenue. In commercial forums, expert advice is provided by professionals seeking business, such as lawyers. Institutional forums belong to nonprofit organizations and rely on donations in lieu of advertising revenue. Dates for the first post in each forum range from May 2004 to February 2013. All the forums were active throughout the data collection period; however, we drop any posts after 2019 to eliminate variations arising from COVID-19. The number of threads ranges from 203 to 8,435. The unique number of users per forum ranges from 307 to 3,448. On average, these users participate in 2.9 to 14.0 threads and contribute 5.0 to 26.3 posts.

Table 1: Forum statistics

Forum	Sponsorship Type	Date of First Collected Post	Number of Threads	Unique Users	Mean Thread Participation	Mean Post Contribution
Federal Soup	Ad-revenue	1/18/2006	2063	1457	6.1	11.9
Free Advice	Commercial	5/3/2004	893	1000	2.9	5.0
Hadit	Ad-revenue	9/26/2005	1202	957*	6.0	11.7
MS World	Institutional	8/18/2011	610	559	3.7	5.3
NeuroTalk	Institutional	8/23/2006	1392	1178	5.6	10.6
PEB	Ad-revenue	2/12/2013	203	307	3.1	7.0
SSDFacts	Commercial	9/1/2009	8435	3448	14.0	26.3

*Note: This forum allows guest posts. Guest cannot be identified as unique users.

Figure 2 shows a timeline of the total posts and threads and each forum's first collected post. In addition to the overall number of posts collected from each forum, this visualization provides insight into the usage and application of the data. A total of 149,068 posts were collected across 15,130 unique threads.[2] Figure 3 shows a timeline displaying the number of threads by forum.

[2] See Table A.2 for descriptive statistics for the forums.

Figure 2: Descriptive statistics for posts

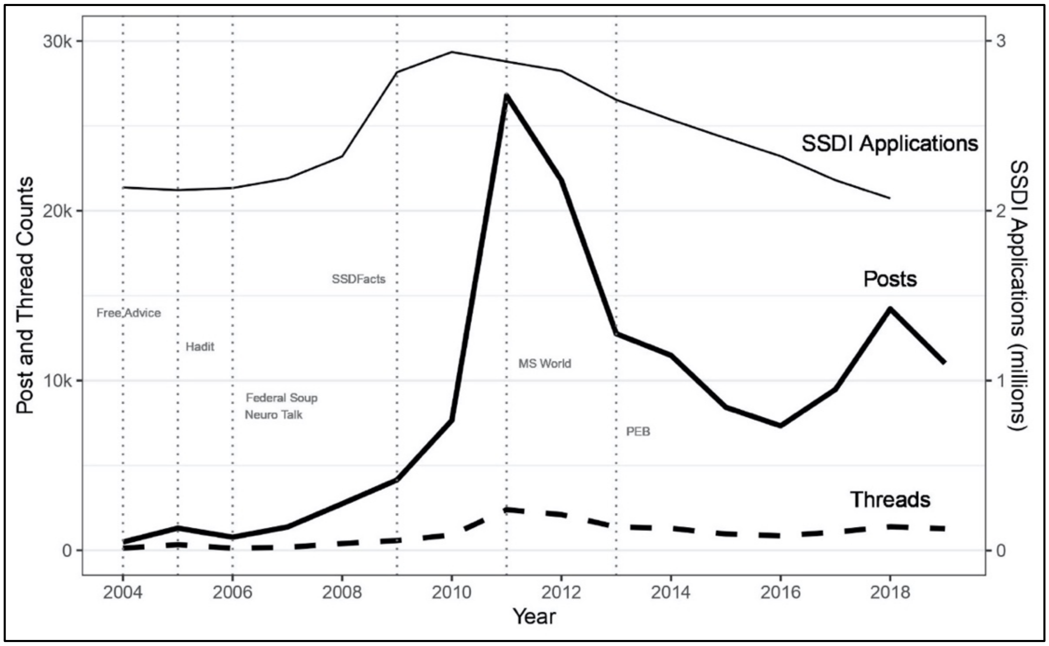
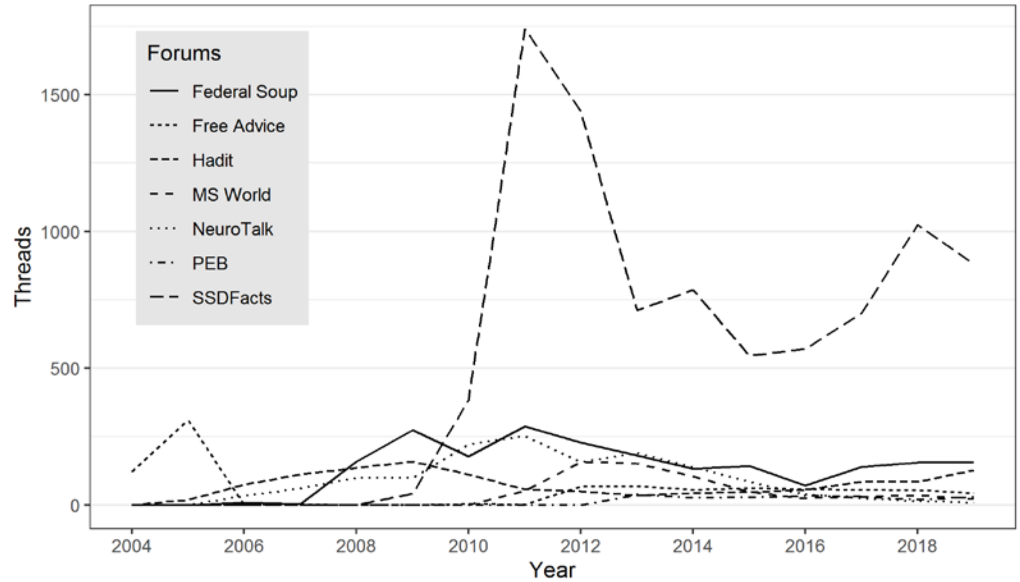


Figure 3: Threads by forum



3.2 Topic Modeling

Topic modeling is an unsupervised machine learning technique that automatically analyzes text data to identify cluster words for a set of documents. In this section, we describe how we use topic modeling to generate text-based descriptions of online forum conversations about SSDI. We preprocess the data, conduct text analytics to identify potential term artifacts and important relationships, and then use a machine learning algorithm to derive the topics.

a) Preprocessing

The first step in topic modeling is preprocessing the corpus, which entails cleaning the text narratives so they can be handled computationally. We first tokenize the corpus (breaking the text into pieces) and eliminate numbers, punctuation, and stop words, which are words used frequently that do not convey useful information. Then we convert all strings to lowercase and use lemmatization to group inflected forms so they can be analyzed as single items. We also convert emojis to text descriptions. Figure 4 demonstrates how the text from a hypothetical post looks before and after preprocessing. The end result is what is known as our "bag of words."

Figure 4: Preprocessing

Before

"Looking for info. Can I apply and receive SSDI while being on active duty? I have been diagnosed with Meniere's disease which makes it impossible for me to work □, or even being on active duty. Thanks"

After

[looking, information, apply, receive, ssdi, active, duty, diagnose, menieres, disease, impossible, work, frown, active, duty, thanks]

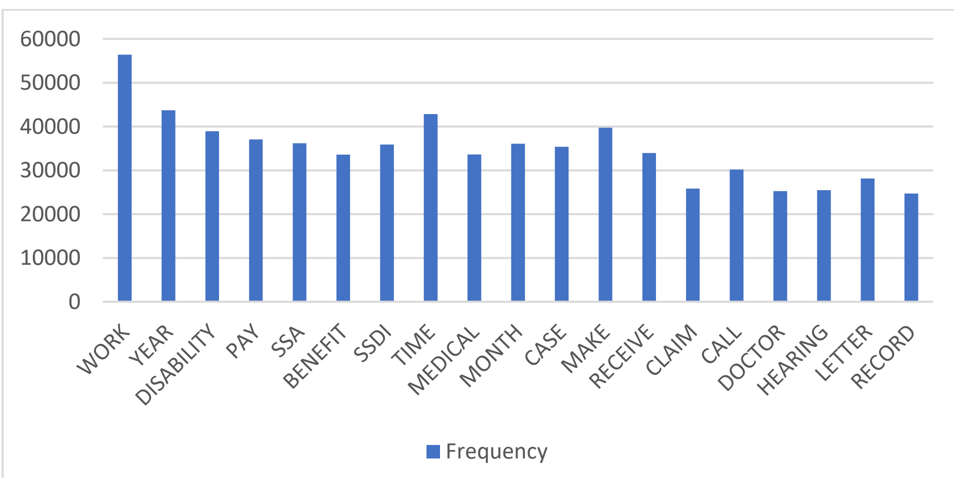
b) Text analytics

We process the data further using data analytics to identify frequent artifacts, misspellings, abbreviation, and entities. This is important because the frequency of these terms may be misinterpreted; for instance, SSDI may be referred to as SSD, Social Security Disability Insurance, or Social Security Disability. We also identify n-grams, which are terms formed by more than one word that have more or different contextual meaning when they occur together that are specifically relevant to our corpus, and connect the terms with an underscore symbol .

The average term (a word, n-gram, acronym or entity) count per post before and after preprocessing is 102 and 101 words, respectively. After preprocessing, we have about 78,000 unique terms with more than 14 million instances of these terms.

Figure 5 shows the top 20 terms by frequency after preprocessing. Some of the most frequent terms are expected, such as work, disability, SSDI, medical, and benefit. Other, less-expected terms of interest, such as information, hope, and check, may provide novel insights. Time-related terms (e.g., time, year, day) appear very frequently, indicating that users are often discussing the complexity and time burden of the application process. However, this figure provides just a first look at the data, so caution should be exercised when making any interpretations.

Figure 5: Top 20 terms by frequency



Text analysis allows exploration of trends in specific terms. For example, Figure 6 shows that the terms *work* and *disability* follow the trends in SSDI applications. Figure 7 shows frequency trends for specific terms. While terms with matching trends may provide an early signal of the types of questions the data may illuminate, further analysis should be performed to investigate how the terms are related to each other and how much they co-occur before drawing conclusions.

Figure 6: Term frequency: Work and Disability

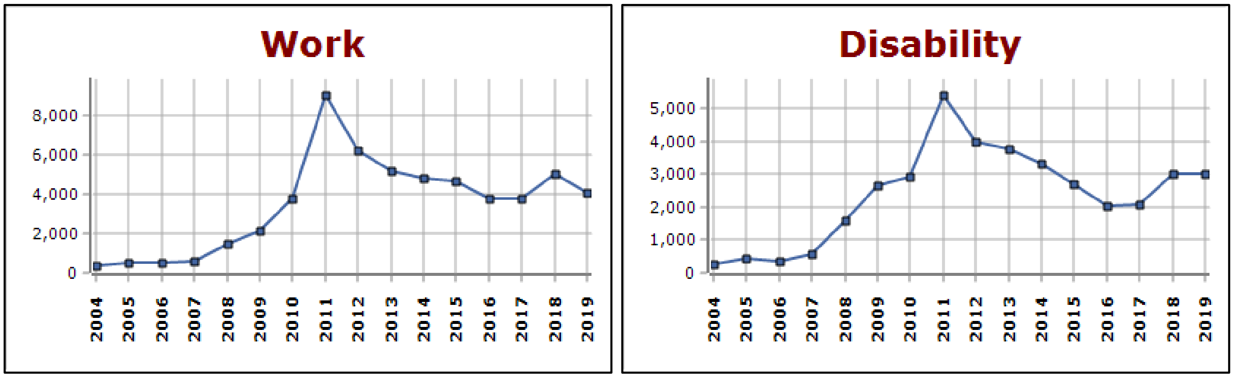
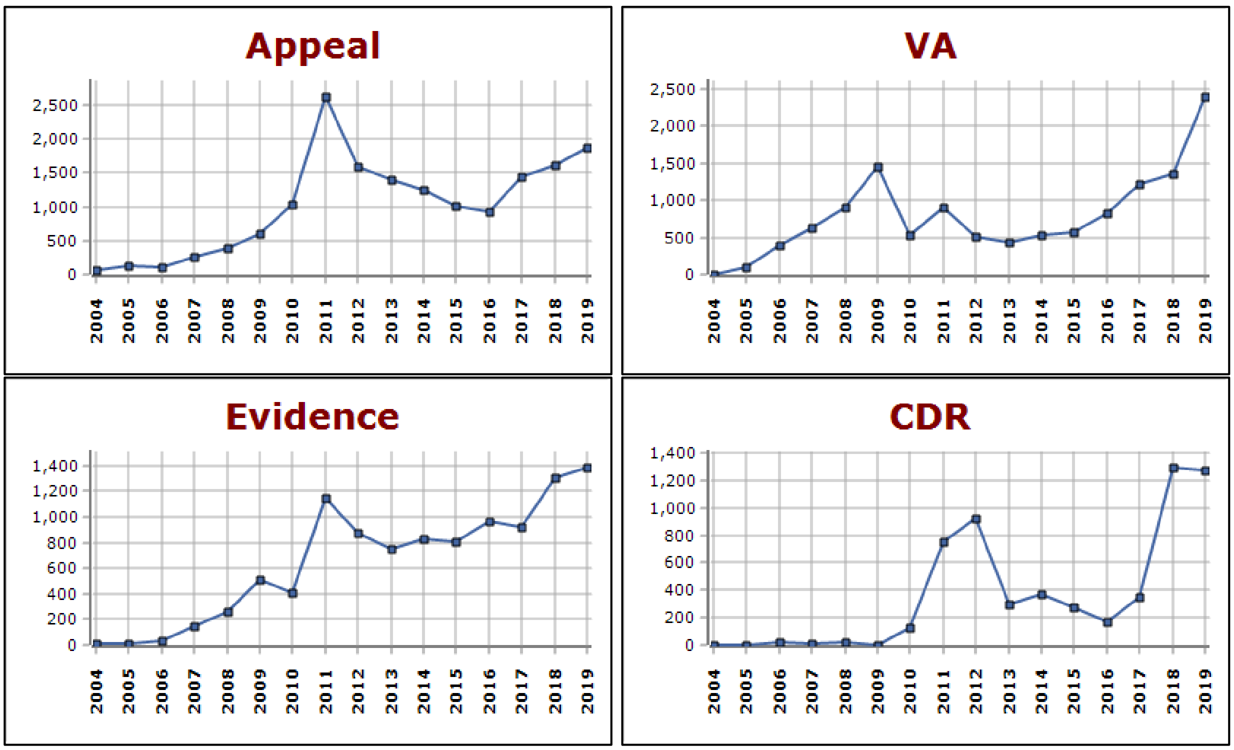
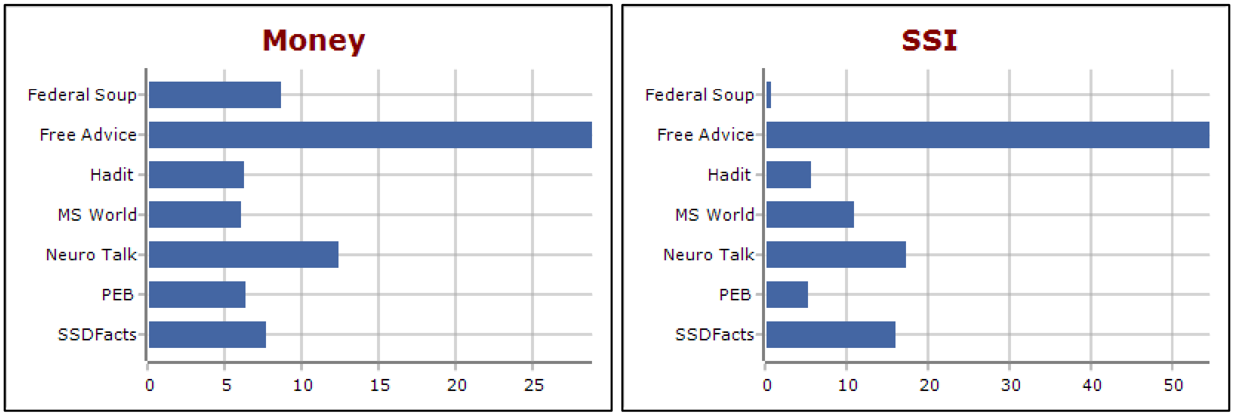


Figure 7: Term frequency: Appeal, Veteran, Evidence, and CDR



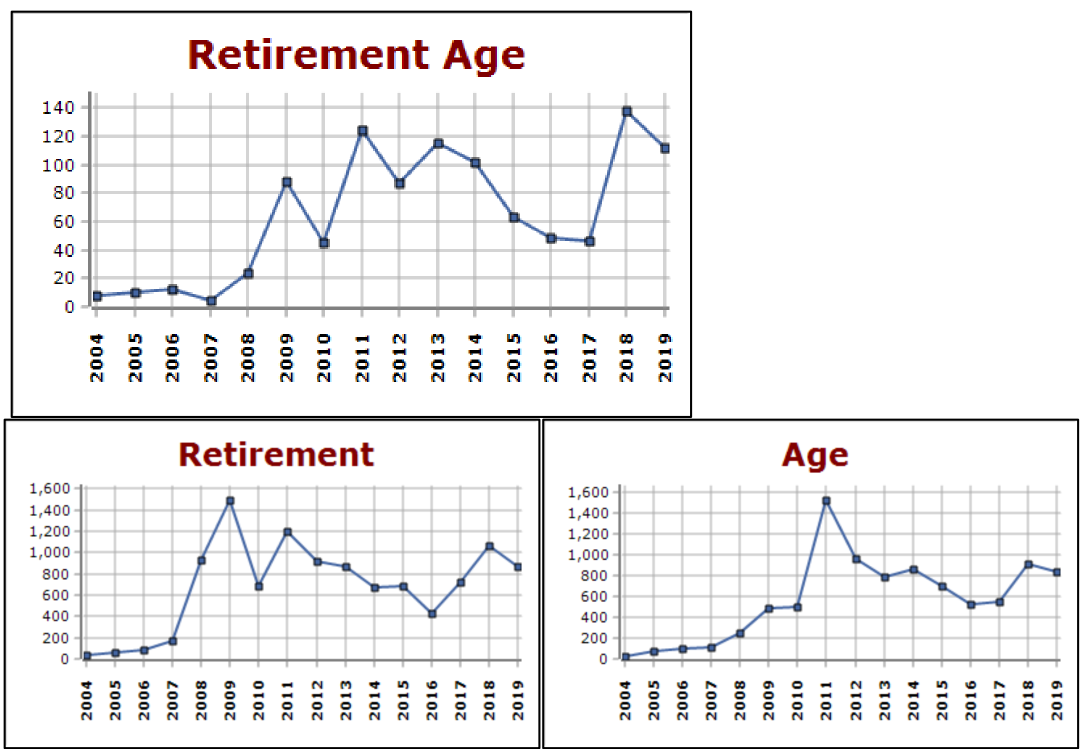
Term occurrence rate by forum can also be determined. For example, Figure 8 maps the term rate for *money* and *SSI* by forum, showing that these terms have a significantly higher frequency in the Free Advice forum.

Figure 8: Term frequency by forum: Money and SSI



Finally, a concurrency analysis identifies combinations of terms that have more contextual meaning together (n-grams). We conduct a concurrency analysis for groups of a minimum of 2 and a maximum of 5 terms. These groups of terms are then replaced by a new single term, composed of the original terms connected by underscores. Figure 9 shows how n-grams work by illustrating the more detailed picture of the trends in SSDI-related conversations provided by the n-gram "retirement age" compared to the individual terms.

Figure 9: Term frequency: Retirement age



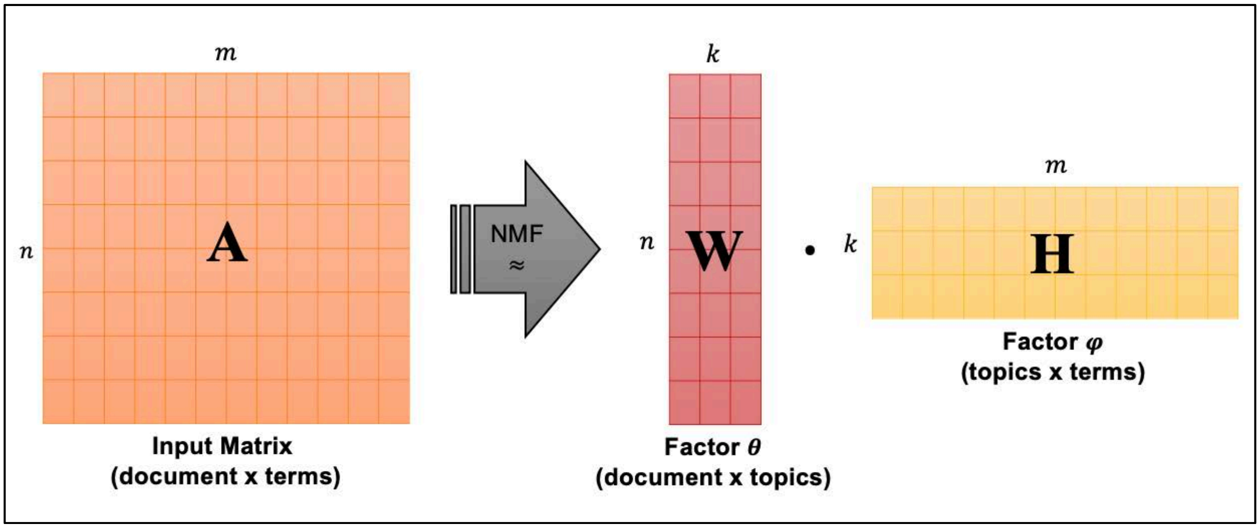
It is nearly impossible to conduct these types of analytics for each term or possible combination of our list of 78,000 relevant terms. Topic modeling helps simplify this process by focusing on the terms that provide the most contextual meaning in the corpus.

c) Topic modeling

Topic modeling is an unsupervised machine learning technique, one that works without predefined tags or training data previously classified by humans, that automatically analyzes text data in a collection of documents to determine cluster words and discover abstract topics. 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. Deriving topics using both of these algorithms, we find that terms included in the NMF are more contextually meaningful, consistent with recent findings that suggest that NMF provides more coherent topics (O'Callaghan et al. 2015). In the following section, we describe the NMF topic modeling process in detail. Appendix B describes LDA topic modeling.

Non-Negative Matrix Factorization. NMF decomposes, or factorizes, high-dimensional vectors into lower-dimensional representations. Figure 10 illustrates how NMF can be applied to our corpus to derive topics. Matrix A is a document-term matrix A, where each row (n) represents a post and each column (m) represents a unique term in the corpus, so that each element in the post represents the weight of a certain term. NMF takes matrix A and modifies the initial values of the factors W and H so that the product approaches A, that is, until either the approximation error converges or the maximum iterations are reached. The elements of factor W contain the weights for each topic (k) in a post, and elements of factor H contain the weights for each term in a topic. The terms with the highest weights are used to textually represent each topic.

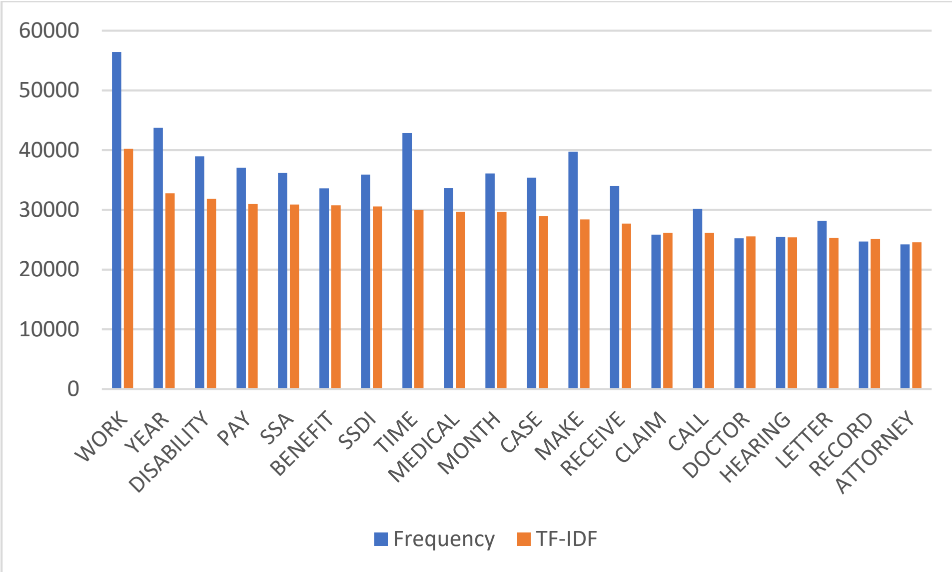
Figure 10: Topic modeling with NMF



NMF requires two inputs, matrix A and k. First, matrix A is constructed by calculating the weight of a certain term in a post using term frequency–inverse document frequency (TF-IDF), which measures how important a term is to a post. Figure 11 overlaps the top 20 terms by TF-IDF with their frequencies. For example, time-related terms appear to be less relevant. This could be because in some forums, users sign posts with a timeline of their application process; as a result, these terms occur frequently in posts but are not necessarily relevant to the entire corpus. Our model includes the top 600 terms by TF-IDF.[3]

[3] Technically, matrix A could include all of the unique terms in the corpus text; however, handling such a large matrix would be computationally difficult. Therefore, it is standard practice to select top terms by TF-IDF.

Figure 11: Top 20 terms by TF-IDF and frequency



Second, to identify the optimal number of latent topics, we iterate the algorithm for different levels of *k* and evaluate how the average topic coherence changes in each model. Topic coherence measures the degree of semantic similarity between high-scoring words in the topic. Coherence ranges from 0 to 1, with 1 representing the highest coherence level. Coherence helps distinguish between topics that are semantically interpretable and topics that are artifacts of statistical inference.

Figure 12 shows that the average topic coherence reaches a maximum in Model 10 and plateaus between Model 10 and Model 15, dropping from there. This indicates that the optimal number of topics occurs between Model 10 and Model 15.

Figure 12: Mean coherence score by number of derived topics

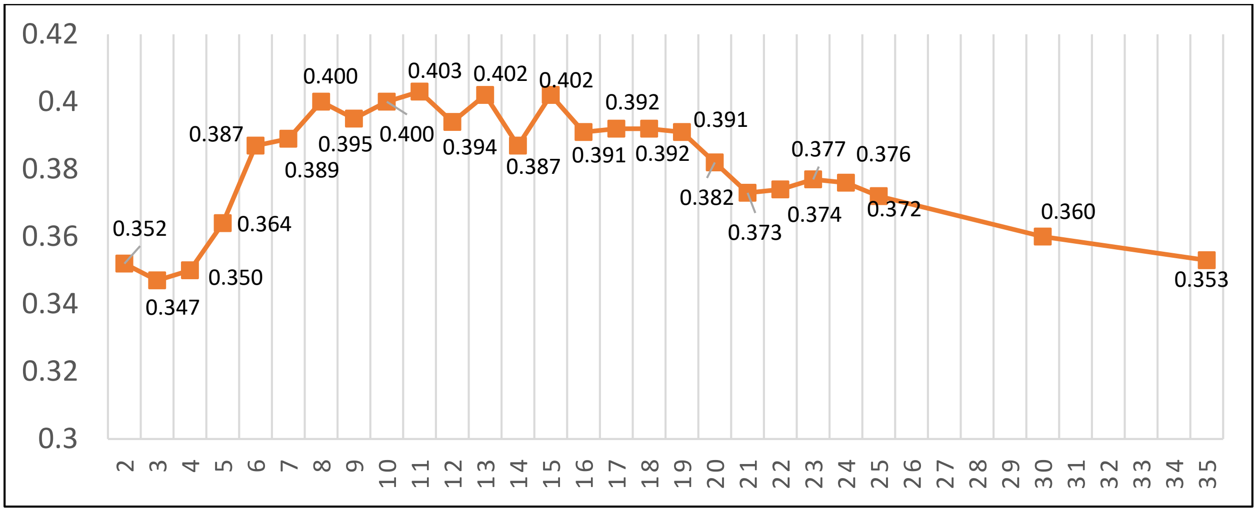


Table 2 charts the topics derived against the number of topics included, which allows the optimal number of topics to be further pinpointed. Topics 1 through 7 are derived independent of the number of topics included in the model. When 15 topics are included, the model identifies all of the topics included in the previous models. This result, combined with the results from the mean topic coherence analysis, suggest that 15 is the optimal number of topics to include in the model.

Table 3 lists the topics derived in Model 15, the most relevant terms for each topic by TF-IDF, topic coherence, and the percentage of cases (the proportion of posts that include this topic). Coherence ranges from 0.347 to 0.523, suggesting the topics included in our model have a good coherence score overall. Table A-3 contains a more extensive list of the most relevant terms for each topic by TF-IDF with topic weights (our factor $\boldsymbol{\varphi}$ or matrix H).

Table 2: Topics derived by model (number of topics included)

Topics	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1. SGA, Impairment	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2. SSDI, SSI, Benefit	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3. Communication with SSA/DDS		X	X	X	X	X	X	X	X	X	X	X	X	X
4. Information			X	X	X	X	X	X	X	X	X	X	X	X
5. Appeals Process				X	X	X	X	X	X	X	X	X	X	X
6. Pain					X	X	X	X	X	X	X	X	X	X
7. Military/Veteran Application Process						X	X	X	X	X	X	X	X	X
8. Disability Retirement, OPM							X	X		X	X	X	X	X
9. Health Insurance, Medicaid, Medicare								X		X	X	X	X	X
10. Medical Records												X	X	X
11. General Social Security									X	X	X	X	X	X
12. Mental Health									X			X	X	X
13. Payment, Bank Account									X		X	X		X
14. Quality of Life										X			X	X
15. Community Support & Engagement											X		X	X

Table 3: Model 15 topics and summary statistics

Topic	Terms with highest ITF-IDF weight	Coherence	% Cases
Forum Referral	http; gov; www; pom; secure; link; app; ssa; nsf lnx;	0.523	6.55%
SGA, Impairment	impairment; perform; activity; ability; functional; meet; medically; prevent; sga;	0.445	4.92%
General Social Security	obtain; individual; claimant; handle; law; june; forum; insurance; social security;	0.443	8.93%
Health Insurance, Medicare, Medicaid	medicare; coverage; premium; insurance; cost; plan; prescription; medicaid; drug;	0.443	3.89%
Military/Veteran Application Process	veteran; rating; service; cue; ptsd; claim; military; service connected;	0.427	4.76%
Disability Retirement, OPM	opm; retirement; fers; annuity; agency; owcp; retire; retirement age; regular retirement;	0.424	4.96%
Appeals Process	alj; appeal; hearing; council; judge; remand; deny; decision; attorney; denial; appeals council;	0.419	14.62%
Payment, Bank Account	account; bank; check; payment; direct deposit; payment center; checking account;	0.413	4.27%
Mental Health	depression; anxiety; disorder; psychiatrist; medication; therapist; psychologist; mental health;	0.406	3.10%
Pain	pain; walk; leg; sit; stand; foot; nerve; knee; surgery; pain meds; chronic pain; nerve damage;	0.393	2.63%
Medical Records	record; medical; form; doctor; copy; fill; send; request; report; ce; medical records;	0.378	14.38%
SSDI, SSI, Benefits	amount; benefit; ssdi; eligible; income; earn; earnings; period; ssi; onset; month; onset date;	0.36	13.67%
Communication with SSA/DDS	office; call; local; phone; mail; local office; ssa office; local ssa office; ss office;	0.359	10.12%
Quality of Life	child; live; parent; food; rent; money; family; income; child support; food stamps; backpay	0.347	3.59%
Community Support & Engagement	favorable, decision, hope; congratulations; happy; good	---	---

3 CANs

We use epistemic network analysis (ENA), which identifies and quantifies connections among elements in coded data and represents them in dynamic network models, to represent the CANs, or organized knowledge, in SSDI online forums. CANs provide visual representations of the chronological trajectory of modeled topics and their connections within the corpus. We then compare the chronological CANs visually.[4]

ENA Methods

ENA Methods. Shaffer (2003) developed ENA to model theories of cognition, discourse, and culture. However, the epistemic frame can be applied to different research 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 be used to 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). The ENA algorithm constructs a network model for each topic in the data by connecting it to topics within the recent temporal context. In our model, the recent temporal context is defined as each post plus the three previous posts within a given thread. The resulting networks are aggregated 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.

Figure 13 shows a visualization of the ENA network for the last year in our data, 2019, in which nodes correspond to the topics and edges reflect the relative frequency of co-

---

[4] Future work could formally quantify the weighted structure of the connections in these networks.

occurrence, or connection, between two topics.[5] The centroid of the mean network, which averages the connection weights, is plotted as a solid square surrounded by a larger square denoting the confidence interval. 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 comparison of networks using the centroids of the mean networks.

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

Figure 14 compiles the centroids of each network by year to represent the CANs. Examining the placement of the nodes shows how the organization of knowledge has changed over the years. For example, clustering around particular connections seems to follow specific environmental changes. For example, conversations move from information seeking in the earlier years of observation toward communication with SSA/DDS and the appeals process during the recession years. Interestingly, the trajectory of the centroids during the recession recovery years, 2010–2013, suggest a shift toward conversations related to pain and mental health.

[5] Appendix C includes ENA networks by year for all years in our data.

Figure 13: Epistemic network visualization

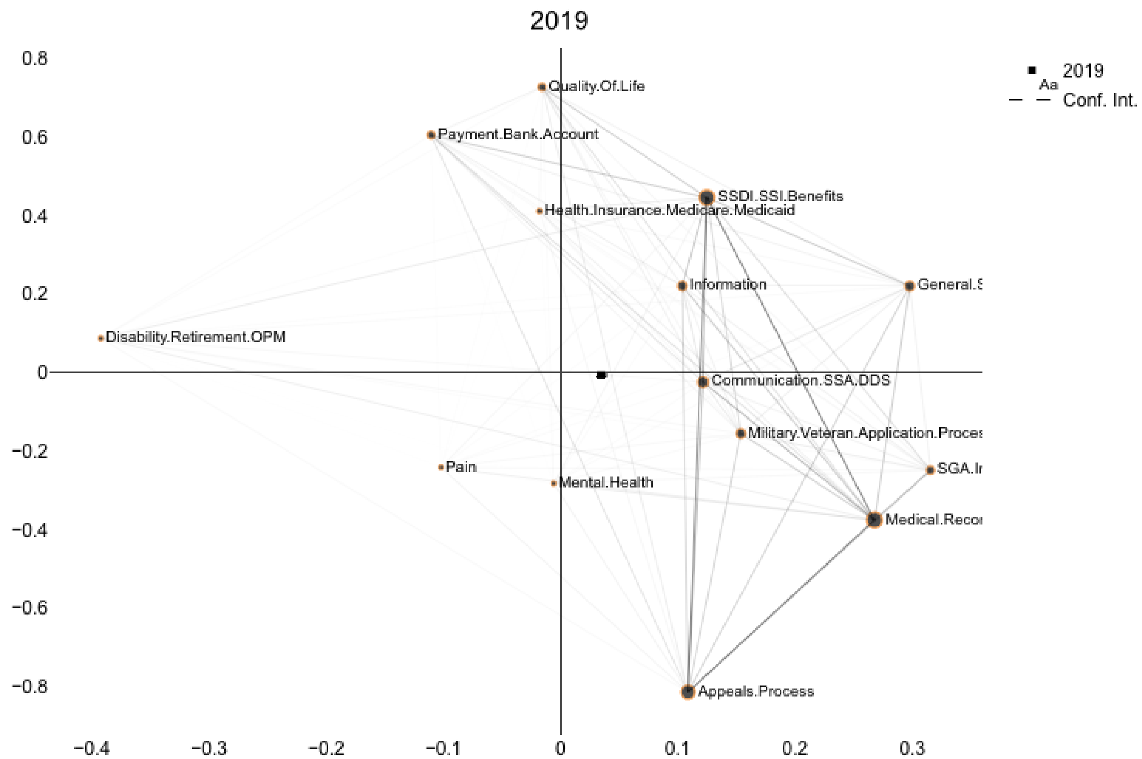
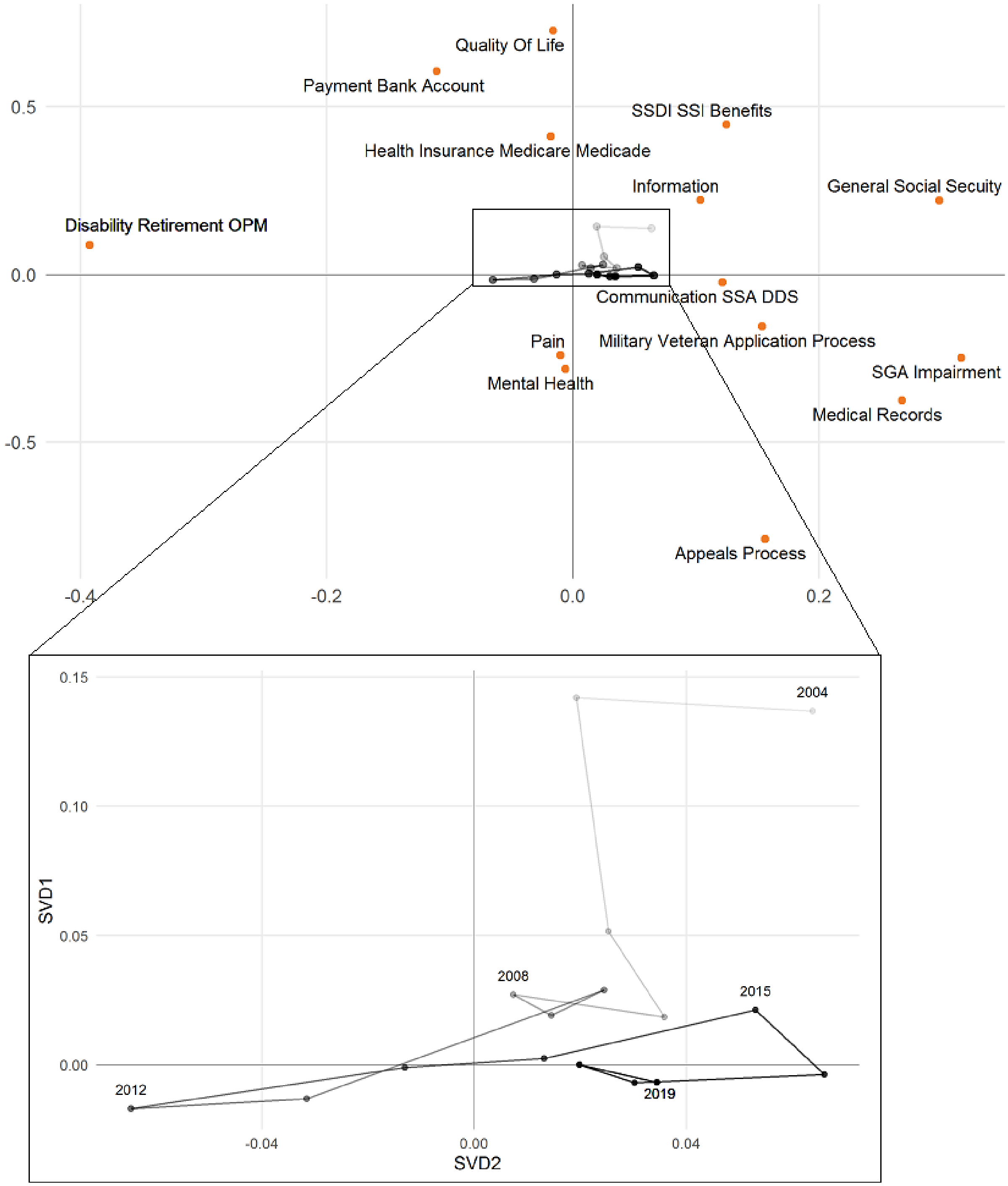


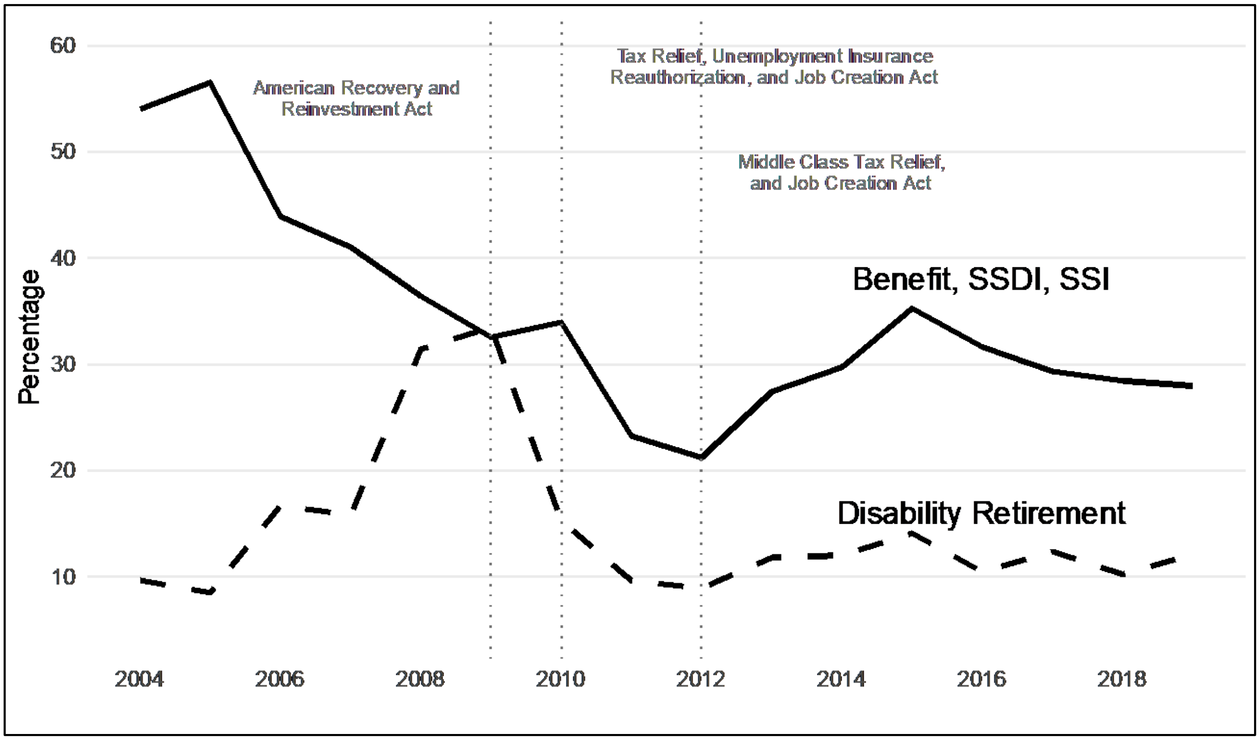
Figure 14: CANs



3.4 Analysis

Figure 15 presents a preliminary graphical analysis of two derived topics, "Disability Retirement" and "Benefit, SSDI, SSI," presented with relevant SSDI policies. For example, conversations on retirement disability arguably should increase during periods of increased unemployment, as they did during the great recession period (2007–2009, the area in gray).

Figure 15: Topic rate by year



4. Discussion

4.1 Implications for vulnerable populations

Research surveys and interviews based on self-reported data may not reveal personal situations or reach financially vulnerable populations. Pendry and Salvatore (2015) find that online forums are positively linked to well-being for stigmatized group members.

A unique contribution of this research is the longitudinal view of the evolution of conversations on issues germane to the communities of users surrounding SSDI applications. In addition to providing insights on the structural relationships between different issues, the longitudinal nature of the data also reveals the trajectory of these issues over time against the backdrop of environmental and policy changes. This information would be particularly useful in predictive model building and hypothesis testing, as well as in developing interpretive or explanatory models surrounding these issues.

4.2 Limitations

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 would not be limited even without this weighting. The profile of SSDI applicants on public forums may change, but that change is not significant to this analysis, which is meant to provide 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 smart phones, 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.

Finally, a key trade-off of applying automated text analysis to this research is bigness 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 are able to express in a given environment (Mahmoodi et al. 2017).

4.3 Future Directions

We believe UGC can have multiple applications in providing insight to SSA focus areas. Statistics derived from CANs can be used to evaluate how changes in the SSDI online community's organized knowledge coincides with changes in the economic environment of SSDI policies. Furthermore, similar to the approach followed in Netzer et al. (2012), which used text mining and semantic analysis of UGC to identify "overwarned" side effects of diabetes drugs, UGC can be used to identify misunderstandings about SSDI rules and policies that may complicate and extend the application and review process. Furthermore, text mining coupled with sentiment analysis has shown to be predictive of pharmaceutical recalls based on user reviews (Batt et al. 2020). This suggests that listings of impairments could be used to train a machine learning model to code disability trends, and sentiment analysis to predict SSDI claims. Finally, UGC presents a unique opportunity to evaluate how COVID-19 has affected SSDI trends, applications, and appeals.

The next step in this research is to develop testable hypotheses about terms and key topics, such as mental health and its associative terms (depression, anxiety, disorder, psychiatrist, medication, therapist, psychologist, mental health) and develop ENA models to explore the relationships and directionality of these terms within each topic area identified in the text analysis.

A further use of the data could be to mine for related terms of interest in the conversations for comparisons. For instance, SSI, Workers Compensation, and Private Disability Insurance could be analyzed to explore differences in attitudes and concerns between applicants to these programs. UGC may also be used as a "social listening" tool to monitor reception to policy changes and sentiments associated with macroeconomic and environmental changes (for instance, COVID-19) in real time.

5. 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 when individuals are confronted with acute stressors, they seek social support. Online forums provide access to this support through task-oriented discussions. SSDI applicants and beneficiaries are financially vulnerable and may feel stigmatized.

References

Arun, R., Suresh, V., Veni Madhavan, C. E., & Narasimha Murthy, M. N. (2010). On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations. In M. J. Zaki, J. X. Yu, B. Ravindran, & V. Pudi (Eds.), Advances in Knowledge Discovery and Data Mining (pp. 391–402). Springer. https://doi.org/10.1007/978-3-642-13657-3_43

Barak, Azy, Meyran Boniel-Nissim, and John Suler. (2008). Fostering empowerment in online support groups. *Computers in Human Behavior 24*(5): 1867–83.

Batt, R., Bavafa, H., & Soltani, M. (2020). Quality improvement spillovers: Evidence from the hospital readmissions reduction program. Available at SSRN 3132770.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022.

Braithwaite, Dawn O., Vincent R. Waldron, and Jerry Finn. (1999). Communication of Social support in computer-mediated groups for people with disabilities. *Health Communication 11*(2): 123–51.

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. Neurocomputing, 72(7), 1775–1781. https://doi.org/10.1016/j.neucom.2008.06.011

Deveaud, R., SanJuan, E., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. Document numerique, Vol. 17(1), 61–84.

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.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. Proceedings of the National Academy of Sciences, 101(suppl 1), 5228–5235. https://doi.org/10.1073/pnas.0307752101

Khalil, S., & Fakir, M. (2017). RCrawler: An R package for parallel web crawling and scraping. SoftwareX, 6, 98-106.

Mahmoodi, J., Leckelt, M., van Zalk, M. W., Geukes, K., & 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, Cody L., Zachari Swiecki, Wesley Collier, Brendan Eagan, Roman Woodward, and David Williamson Shaffer. (2019). RENA: Epistemic network analysis." https://CRAN.R-project.org/package=rENA

National Cancer Institute. (2013). Health Information National Trends Survey. Retrieved from http://hints.cancer.gov

Netzer, Oded, Ronen Feldman, Jacob Goldenberg, and Moshe Fresko. (2012). Mine your own business: Market-structure surveillance through text mining. *Marketing Science 31*(3): 521–43.

O'Callaghan, Derek, Derek Greene, Joe Carthy, and Pádraig Cunningham. (2015). An Analysis of the coherence of descriptors in topic modeling. *Expert Systems with Applications 42*(13): 5645–57.

Pendry, Louise F., and Jessica Salvatore. (2015). Individual and social benefits of online discussion forums. *Computers in Human Behavior 50*: 211–20.

Phan, X. H., Nguyen, L. M., & Horiguchi, S. (2008, April). Learning to classify short and sparse text & web with hidden topics from large-scale data collections. In Proceedings of the 17th international conference on World Wide Web (pp. 91-100).

Shaffer, David Williamson. (2003). Pedagogical praxis: Using technology to build professional communities of practice." *ACM SIGGROUP Bulletin 24*(3): 39–43.

Shaffer, D. W., & Ruis, A. R. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gasevic (Eds.), Handbook of learning analytics (pp. 175–187). Society for Learning Analytics Research.

Way, Wendy, Nancy Wong, and Damiana Gibbons (2011) "Social Network Sites and Internet Forums: An Investigation of Interactions around Personal Finance in the Online Social World" Madison, WI: Center for Financial Security Working Paper.

Wright, Kevin B. (2016). Communication in health-related online social support groups/communities: A review of research on predictors of participation, applications of social support theory and health outcomes." *Review of Communication Research 4*:65–87.

Appendices

Appendix A

Descriptive Statistics

Table A.1: Online forums with SSDI-relevant conversations

Forum	Forum / Subforum Topic
AARP	AARP Forum / Social Security
Bogleheads	Investments / Social Security
City Data	Health and Wellness / Health Insurance
Consumer Affairs	Miscellaneous / Social Security Disability
DIS Boards	DIS-topic / Social Security Disability Insurance
Disability Secrets	Disability Secrets
DSL Report	Open Forum / SSI Disability
MDS Foundation	MDS Patient Message Board / Filing SSDI Claims
Psych Central	Insurance and Finances / Approval for SSI/SSDI
Social Security Intelligence	Social Security
Something Awful	Debate and Discussions / Social Security
The Student Doctor Network	Physicians & Residence Forum / Military Medicine
This is MS	This is MS
VISTA Campus	Individuals with disabilities
Federal Soup	Retirement / Disability Retirement
Free Advice	Law / Social Security Disability & SSI Law
HADit.com	Veteran / Social Security Disability
MS world	MS / Social Security Disability
NeuroTalk	NeuroTalk Support Groups / SSDI
Physical Evaluation Board Forum	Physical Evaluation for Veterans / SSDI
Social Security Disability Facts	Social Security Disability
Disability Benefits Help	Disability Benefits / SSDI

Note: Low-value forums (e.g., forums with few threads containing SSDI-related topics and forums with very sparse conversations) were discarded.

Table A.2: Monthly topic and post statistics

Forums	Topic Mean (Std. Dev.)	Topic Maximum	Topic Minimum	Post Mean (Std. Dev.)	Post Maximum	Post Minimum
Federal Soup	18.8 (10.7)	53	1	105.6 (69.7)	304	1
Free Advice	8.2 (9.3)	45	1	39.9 (36.5)	187	1
Hadit	10.4 (7.4)	79	1	65.4 (64.0)	695	3
MS World	9.8 (8.4)	41	1	29.6 (23.5)	90	1
NeuroTalk	13.6 (10.2)	51	1	80.8 (84.2)	419	1
PEB	4.4 (2.3)	13	1	26.0 (19.7)	102	2
SSDFacts	92.3 (51.8)	285	5	764.8 (536.8)	2599	5

Table A.3: Most relevant terms for each topic by TF-IDF

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15
CONGRATULATIONS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NEWS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HOPE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GLAD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GOOD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HAPPY	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LUCK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HEARING	**0.55**	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.05	0.00	0.00	0.00
ATTORNEY	**0.49**	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
ALJ	**0.48**	0.07	0.04	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.00	0.13	0.00	0.00	0.00
JUDGE	**0.45**	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.17	0.00	0.00	0.04	0.00	0.00	0.00
APPEAL	**0.42**	0.23	0.03	0.07	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00
DENY	**0.36**	0.23	0.00	0.11	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
LAWYER	**0.36**	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
COURT	**0.32**	0.06	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
HIRE	**0.31**	0.05	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.01	0.09	0.01	0.00
DENIAL	0.27	0.23	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
WIN	0.23	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
FEE	0.22	0.01	0.00	0.00	0.00	0.12	0.00	0.01	0.00	0.10	0.00	0.00	0.00	0.00	0.00
LEVEL	0.21	0.20	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.18	0.00	0.01	0.00
CHANCE	0.13	0.06	0.00	0.02	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00
HEAR	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
INITIAL	0.11	**0.36**	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
EVIDENCE	0.06	**0.35**	0.00	0.18	0.00	0.00	0.00	0.00	0.04	0.00	0.03	0.22	0.00	0.00	0.00
DETERMINATION	0.00	**0.35**	0.00	0.01	0.01	0.00	0.00	0.05	0.04	0.00	0.00	0.10	0.00	0.07	0.01
APPLICATION	0.00	**0.32**	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.04	0.00	0.03
FILE	0.09	**0.30**	0.06	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
PROCESS	0.00	0.29	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
RECONSIDERATION	0.22	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
DDS	0.00	0.29	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.15	0.01	0.00	0.00	0.00
REQUEST	0.03	0.29	0.00	0.02	0.10	0.00	0.00	0.00	0.00	0.00	0.23	0.00	0.00	0.00	0.00
MEDICAL	0.00	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.26	0.27	0.00	0.00	0.00
SUBMIT	0.00	0.29	0.02	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.02	0.00	0.00
PROVIDE	0.00	0.27	0.00	0.03	0.00	0.00	0.01	0.08	0.00	0.07	0.00	0.10	0.06	0.08	0.02
DOCUMENT	0.00	0.26	0.00	0.04	0.00	0.00	0.02	0.04	0.00	0.00	0.08	0.12	0.00	0.03	0.00
REVIEW	0.00	0.25	0.00	0.00	0.02	0.00	0.00	0.00	0.09	0.00	0.13	0.00	0.00	0.03	0.00
DETERMINE	0.00	0.24	0.04	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.24	0.00	0.05	0.15
COMPLETE	0.00	0.24	0.00	0.00	0.03	0.00	0.04	0.01	0.00	0.00	0.23	0.04	0.02	0.02	0.00
INCLUDE	0.00	0.22	0.00	0.09	0.00	0.00	0.01	0.07	0.00	0.10	0.05	0.10	0.02	0.04	0.02
CASE	0.22	0.22	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
STATEMENT	0.00	0.21	0.00	0.06	0.00	0.03	0.01	0.00	0.00	0.00	0.06	0.06	0.09	0.01	0.00
REQUIRE	0.00	0.21	0.00	0.00	0.00	0.00	0.01	0.05	0.00	0.08	0.00	0.16	0.06	0.03	0.07
FACT	0.00	0.21	0.00	0.01	0.00	0.00	0.16	0.00	0.00	0.00	0.00	0.06	0.00	0.05	0.00
FOLLOW	0.00	0.20	0.01	0.03	0.02	0.00	0.05	0.07	0.00	0.01	0.01	0.09	0.01	0.08	0.00
REPRESENTATIVE	0.02	0.20	0.00	0.00	0.16	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00
REASON	0.00	0.19	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.06	0.05	0.00	0.00
MAKE	0.00	0.19	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
STEP	0.02	0.19	0.00	0.00	0.00	0.00	0.02	0.03	0.00	0.00	0.00	0.11	0.00	0.00	0.00
DECIDE	0.02	0.18	0.03	0.00	0.00	0.00	0.09	0.00	0.05	0.02	0.00	0.00	0.00	0.02	0.03

POINT	0.00	0.17	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00
INFORMATION	0.00	0.17	0.00	0.00	0.11	0.00	0.00	0.06	0.00	0.00	0.03	0.00	0.00	0.05	0.00
STATE	0.00	0.16	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.05	0.00	0.02	0.02
SUPPORT	0.00	0.16	0.00	0.01	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.05
RULE	0.00	0.15	0.00	0.00	0.00	0.00	0.01	0.08	0.00	0.00	0.00	0.10	0.01	0.05	0.13
CURRENT	0.00	0.15	0.00	0.06	0.00	0.00	0.02	0.02	0.00	0.07	0.01	0.12	0.06	0.02	0.01
IMPORTANT	0.00	0.15	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.01	0.04	0.11	0.00	0.04	0.00
SYSTEM	0.00	0.14	0.00	0.02	0.01	0.00	0.10	0.00	0.00	0.01	0.00	0.00	0.04	0.03	0.01
ORDER	0.00	0.14	0.00	0.03	0.00	0.00	0.04	0.01	0.00	0.03	0.00	0.05	0.00	0.01	0.06
NOTICE	0.00	0.13	0.05	0.00	0.10	0.03	0.01	0.00	0.09	0.00	0.00	0.00	0.00	0.02	0.00
PERSON	0.00	0.13	0.00	0.00	0.07	0.00	0.12	0.00	0.00	0.00	0.00	0.04	0.00	0.04	0.05
EXPLAIN	0.00	0.11	0.00	0.00	0.02	0.00	0.11	0.04	0.02	0.00	0.03	0.06	0.00	0.00	0.00
READ	0.00	0.11	0.00	0.00	0.00	0.00	0.06	0.05	0.00	0.00	0.02	0.04	0.00	0.00	0.00
CHANGE	0.00	0.10	0.08	0.00	0.00	0.00	0.06	0.00	0.02	0.05	0.00	0.03	0.00	0.00	0.00
CORRECT	0.00	0.09	0.00	0.02	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LINE	0.00	0.09	0.00	0.01	0.05	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00
ADD	0.00	0.08	0.00	0.04	0.00	0.00	0.07	0.00	0.00	0.02	0.00	0.02	0.00	0.00	0.00
MEMBER	0.00	0.07	0.00	0.00	0.00	0.00	0.05	0.03	0.00	0.00	0.00	0.00	0.00	0.06	0.00
AGREE	0.04	0.04	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
DATE	0.00	0.06	**0.64**	0.01	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.01	0.00	0.00	0.00
ONSET	0.00	0.02	**0.62**	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.10	0.00	0.00	0.03
MONTH	0.00	0.00	**0.37**	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.11
PERIOD	0.00	0.06	**0.35**	0.00	0.00	0.00	0.00	0.04	0.00	0.07	0.00	0.09	0.00	0.00	0.17
JULY	0.00	0.00	0.26	0.02	0.04	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
AUGUST	0.00	0.00	0.25	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MARCH	0.01	0.00	0.24	0.04	0.03	0.00	0.00	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.00
AFTER	0.00	0.07	0.24	0.00	0.00	0.00	0.11	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
WAIT	0.01	0.00	0.22	0.00	0.04	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00
START	0.00	0.01	0.22	0.00	0.00	0.00	0.15	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00
PRIOR	0.00	0.16	0.22	0.03	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.10	0.00	0.02	0.01
APRIL	0.00	0.00	0.21	0.03	0.05	0.00	0.00	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.00
EARLY	0.00	0.07	0.18	0.02	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.01	0.08	0.00	0.04
APPROVE	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00
RECEIVE	0.00	0.00	0.15	0.00	0.09	0.14	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.01	0.14
APPROVAL	0.02	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00
VA	0.00	0.00	0.00	**0.72**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VETERAN	0.00	0.03	0.00	**0.61**	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.01	0.00
RATING	0.00	0.01	0.00	**0.51**	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.02	0.00	0.00	0.00
PTSD	0.00	0.00	0.00	**0.48**	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00
SERVICE	0.00	0.09	0.00	**0.34**	0.00	0.00	0.00	0.04	0.00	0.09	0.00	0.00	0.30	0.04	0.00
CLAIM	0.00	**0.31**	0.00	**0.31**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00
RATE	0.03	0.01	0.00	0.29	0.00	0.00	0.00	0.02	0.00	0.06	0.00	0.01	0.04	0.00	0.03
AWARD	0.00	0.00	0.12	0.22	0.00	0.14	0.00	0.00	0.16	0.00	0.00	0.00	0.00	0.01	0.01
OPINION	0.04	0.13	0.00	0.14	0.00	0.00	0.07	0.00	0.00	0.00	0.05	0.09	0.00	0.02	0.00
OFFICE	0.00	0.07	0.00	0.00	**0.67**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LOCAL	0.00	0.05	0.00	0.00	**0.64**	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CALL	0.00	0.00	0.00	0.00	**0.59**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PHONE	0.00	0.01	0.00	0.00	**0.43**	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NUMBER	0.00	0.08	0.00	0.00	0.29	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.03	0.05	0.00
MAIL	0.00	0.00	0.01	0.00	0.28	0.04	0.00	0.00	0.12	0.00	0.14	0.00	0.00	0.00	0.00
CONTACT	0.00	0.13	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.01	0.04	0.00
TODAY	0.00	0.00	0.00	0.00	0.23	0.03	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00
WEEK	0.00	0.00	0.07	0.00	0.23	0.00	0.13	0.00	0.03	0.00	0.04	0.00	0.00	0.00	0.00

STATUS	0.00	0.12	0.04	0.00	0.21	0.02	0.00	0.00	0.03	0.00	0.00	0.00	0.09	0.00	0.00
UPDATE	0.00	0.03	0.00	0.00	0.19	0.04	0.00	0.00	0.04	0.00	0.01	0.00	0.00	0.00	0.00
SPEAK	0.01	0.06	0.00	0.00	0.17	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ONLINE	0.00	0.08	0.00	0.00	0.15	0.11	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.00	0.00
WORKER	0.00	0.09	0.00	0.00	0.12	0.00	0.04	0.00	0.00	0.00	0.00	0.01	0.09	0.03	0.02
ACCOUNT	0.00	0.00	0.00	0.00	0.01	**0.68**	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00
BANK	0.00	0.00	0.00	0.00	0.02	**0.65**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DEPOSIT	0.00	0.00	0.00	0.00	0.05	**0.62**	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
CHECK	0.00	0.00	0.00	0.00	0.08	**0.36**	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00
MONEY	0.00	0.00	0.00	0.00	0.00	**0.33**	0.08	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.16
PAYMENT	0.00	0.00	0.19	0.00	0.06	**0.33**	0.00	0.00	0.00	0.03	0.00	0.00	0.04	0.00	0.14
BACKPAY	0.00	0.00	0.23	0.00	0.00	0.26	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
MONTHLY	0.00	0.00	0.15	0.00	0.00	0.26	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.25	0.16
PAYEE	0.00	0.03	0.00	0.00	0.01	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.03
SET	0.00	0.06	0.04	0.00	0.06	0.11	0.06	0.00	0.00	0.01	0.00	0.06	0.00	0.02	0.00
BAD	0.00	0.00	0.00	0.01	0.00	0.00	**0.34**	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00
PAIN	0.00	0.00	0.00	0.00	0.00	0.00	**0.34**	0.00	0.00	0.00	0.09	0.22	0.00	0.00	0.00
LIFE	0.00	0.00	0.00	0.00	0.00	0.00	**0.34**	0.00	0.00	0.08	0.00	0.02	0.00	0.00	0.03
FEEL	0.03	0.00	0.00	0.00	0.00	0.00	**0.33**	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
WALK	0.00	0.00	0.00	0.00	0.00	0.00	**0.33**	0.00	0.00	0.00	0.02	0.15	0.00	0.00	0.00
HOUR	0.00	0.00	0.00	0.04	0.02	0.00	0.29	0.00	0.00	0.00	0.02	0.11	0.00	0.00	0.00
LEAVE	0.00	0.01	0.01	0.01	0.02	0.00	0.29	0.00	0.00	0.00	0.00	0.01	0.12	0.00	0.00
SIT	0.04	0.00	0.00	0.00	0.02	0.00	0.29	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00
HOME	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.00	0.00	0.11	0.00	0.01	0.00	0.00	0.10
PROBLEM	0.00	0.01	0.00	0.05	0.00	0.00	0.28	0.00	0.00	0.01	0.01	0.14	0.00	0.00	0.00
DEPRESSION	0.00	0.00	0.00	0.15	0.00	0.00	0.27	0.00	0.00	0.00	0.03	0.22	0.00	0.01	0.00
ANXIETY	0.00	0.00	0.00	0.13	0.00	0.00	0.26	0.00	0.00	0.00	0.04	0.18	0.00	0.01	0.00
MINUTE	0.06	0.00	0.00	0.04	0.04	0.00	0.25	0.00	0.01	0.00	0.09	0.05	0.00	0.00	0.00
FRIEND	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.02
MEDICATION	0.00	0.00	0.00	0.01	0.00	0.00	0.25	0.00	0.00	0.12	0.11	0.24	0.00	0.00	0.00
HARD	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
THING	0.00	0.01	0.00	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
PEOPLE	0.00	0.06	0.00	0.00	0.00	0.00	0.24	0.00	0.00	0.02	0.00	0.00	0.00	0.03	0.00
FAMILY	0.00	0.00	0.00	0.00	0.00	0.00	0.23	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.12
JOB	0.00	0.00	0.00	0.00	0.00	0.00	0.23	0.00	0.00	0.00	0.00	0.22	0.18	0.00	0.04
PLACE	0.00	0.08	0.00	0.00	0.00	0.00	0.22	0.00	0.00	0.05	0.00	0.01	0.00	0.00	0.00
HAND	0.00	0.06	0.00	0.00	0.00	0.00	0.21	0.00	0.00	0.00	0.03	0.06	0.00	0.00	0.00
SURGERY	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.01	0.03	0.10	0.00	0.00	0.00
LOSE	0.04	0.01	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.05	0.00	0.00	0.01	0.00	0.06
TIME	0.00	0.05	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
LIVE	0.00	0.00	0.00	0.00	0.00	0.01	0.19	0.00	0.00	0.13	0.00	0.05	0.00	0.00	0.16
STAY	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00
DEAL	0.00	0.03	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.02	0.00	0.00	0.00	0.01	0.00
TALK	0.02	0.00	0.00	0.00	0.11	0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
KIND	0.00	0.01	0.00	0.00	0.00	0.00	0.18	0.00	0.00	0.01	0.00	0.02	0.00	0.01	0.00
DAY	0.00	0.00	0.08	0.00	0.12	0.00	0.18	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00
ISSUE	0.00	0.11	0.00	0.08	0.00	0.00	0.17	0.00	0.00	0.01	0.00	0.12	0.00	0.04	0.00
MOVE	0.00	0.05	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00
AGO	0.00	0.00	0.04	0.17	0.04	0.00	0.17	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00
ABLE	0.00	0.03	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.05	0.00	0.16	0.00	0.00	0.03
STOP	0.00	0.00	0.05	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.11
END	0.00	0.05	0.08	0.00	0.00	0.00	0.17	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00
SITUATION	0.00	0.08	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.06	0.00	0.00	0.00	0.05	0.03

SCHOOL	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.16
MIND	0.00	0.02	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
EXPERIENCE	0.00	0.12	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.06	0.00	0.05	0.00
WORD	0.00	0.05	0.00	0.00	0.00	0.00	0.15	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00
TURN	0.02	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.06
BIG	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00
UNDERSTAND	0.00	0.06	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.01	0.00	0.06	0.00	0.00	0.00
HOLD	0.00	0.06	0.00	0.00	0.06	0.03	0.14	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
COUPLE	0.00	0.00	0.04	0.00	0.05	0.00	0.14	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
WRONG	0.00	0.08	0.00	0.01	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FINALLY	0.00	0.00	0.10	0.03	0.04	0.00	0.13	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00
LONG	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.01	0.00	0.08	0.01	0.00	0.00	0.00
FIGHT	0.10	0.00	0.00	0.07	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MENTION	0.00	0.07	0.00	0.05	0.00	0.00	0.13	0.00	0.00	0.00	0.03	0.06	0.00	0.00	0.00
CLOSE	0.00	0.06	0.00	0.00	0.03	0.01	0.12	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
PRETTY	0.00	0.02	0.00	0.01	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HAPPEN	0.00	0.05	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
WORRY	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00
LOVE	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
REMEMBER	0.00	0.01	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.01	0.05	0.00	0.00	0.00
ADVICE	0.03	0.02	0.00	0.02	0.00	0.00	0.10	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00
LONGER	0.00	0.04	0.02	0.00	0.00	0.00	0.10	0.00	0.00	0.02	0.00	0.07	0.04	0.00	0.05
WIFE	0.00	0.00	0.00	0.08	0.00	0.00	0.09	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.08
IDEA	0.00	0.06	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.01	0.00	0.02	0.00	0.00	0.00
MATTER	0.00	0.07	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.06	0.00	0.01	0.04
GUESS	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FIGURE	0.00	0.00	0.02	0.00	0.00	0.04	0.07	0.00	0.00	0.04	0.00	0.00	0.01	0.00	0.04
HUSBAND	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.04
APPRECIATE	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SHARE	0.00	0.03	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00
SOUND	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HTTP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.89**	0.00	0.01	0.00	0.04	0.00	0.10	0.00
GOV	0.00	0.02	0.00	0.00	0.00	0.00	0.00	**0.86**	0.00	0.04	0.00	0.04	0.00	0.08	0.02
LINK	0.00	0.02	0.00	0.00	0.00	0.00	0.00	**0.47**	0.00	0.01	0.00	0.00	0.00	0.03	0.00
SSA	0.00	0.14	0.00	0.00	0.00	0.04	0.00	**0.34**	0.00	0.00	0.00	0.09	0.00	0.00	0.03
FORUM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FAVORABLE	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	**0.85**	0.00	0.00	0.00	0.00	0.00	0.00
FULLY	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	**0.81**	0.00	0.00	0.01	0.00	0.00	0.00
DECISION	0.08	0.23	0.00	0.00	0.01	0.00	0.00	0.00	**0.47**	0.00	0.00	0.00	0.00	0.00	0.00
LETTER	0.00	0.00	0.08	0.01	0.19	0.07	0.00	0.00	0.23	0.00	0.05	0.00	0.00	0.00	0.00
MEDICARE	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.02	0.00	**0.62**	0.00	0.00	0.00	0.00	0.00
COST	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	**0.50**	0.00	0.00	0.00	0.00	0.00
INSURANCE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.50**	0.00	0.02	0.04	0.27	0.00
PLAN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	**0.46**	0.00	0.01	0.01	0.00	0.00
MEDICAID	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.44**	0.00	0.00	0.00	0.00	0.14
COVER	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.42**	0.00	0.01	0.00	0.00	0.00
HEALTH	0.00	0.00	0.00	0.01	0.00	0.00	0.08	0.00	0.00	**0.32**	0.00	0.16	0.05	0.26	0.00
PAY	0.00	0.00	0.06	0.00	0.00	0.26	0.00	0.00	0.00	0.30	0.00	0.00	0.07	0.00	0.17
PART	0.00	0.06	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.27	0.00	0.06	0.00	0.00	0.00
BILL	0.00	0.00	0.00	0.00	0.00	0.18	0.06	0.00	0.00	0.26	0.00	0.00	0.00	0.00	0.00
PROGRAM	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.25	0.00	0.02	0.00	0.04	0.24
CARE	0.00	0.00	0.00	0.01	0.00	0.00	0.19	0.00	0.00	0.22	0.00	0.05	0.00	0.00	0.00
FREE	0.00	0.02	0.00	0.00	0.02	0.00	0.03	0.02	0.00	0.21	0.00	0.00	0.00	0.00	0.00

FORM	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	**0.59**	0.00	0.00	0.01	0.00
FILL	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	**0.53**	0.01	0.00	0.00	0.00
DOCTOR	0.00	0.00	0.00	0.02	0.00	0.00	0.12	0.00	0.00	0.03	**0.42**	0.19	0.00	0.00	0.00
SEND	0.00	0.05	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	**0.38**	0.00	0.00	0.00	0.00
RECORD	0.02	0.24	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	**0.32**	0.18	0.00	0.00	0.00
REPORT	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	**0.31**	0.14	0.00	0.00	0.00
COPY	0.00	0.14	0.00	0.04	0.05	0.00	0.00	0.00	0.00	0.00	**0.31**	0.00	0.00	0.00	0.00
CE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.28	0.11	0.00	0.00	0.00
EXAM	0.00	0.00	0.00	0.23	0.00	0.00	0.01	0.00	0.00	0.00	0.27	0.08	0.00	0.00	0.00
CDR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.26	0.03	0.00	0.01	0.00
SHORT	0.00	0.00	0.01	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.26	0.03	0.00	0.05	0.00
APPOINTMENT	0.00	0.00	0.00	0.00	0.14	0.00	0.13	0.00	0.00	0.04	0.22	0.04	0.00	0.00	0.00
VISIT	0.00	0.00	0.00	0.00	0.04	0.00	0.11	0.00	0.00	0.14	0.21	0.09	0.00	0.01	0.00
PAPERWORK	0.02	0.01	0.01	0.00	0.06	0.00	0.04	0.00	0.00	0.00	0.20	0.00	0.09	0.00	0.00
NOTE	0.00	0.11	0.00	0.07	0.00	0.00	0.08	0.00	0.00	0.00	0.20	0.14	0.00	0.00	0.00
DOC	0.00	0.00	0.00	0.09	0.00	0.00	0.09	0.02	0.00	0.00	0.17	0.06	0.00	0.00	0.00
WRITE	0.00	0.09	0.00	0.00	0.00	0.00	0.05	0.00	0.02	0.00	0.16	0.09	0.06	0.00	0.00
PAGE	0.00	0.11	0.00	0.00	0.01	0.00	0.00	0.12	0.01	0.00	0.15	0.00	0.00	0.03	0.00
SCHEDULE	0.05	0.02	0.07	0.04	0.06	0.00	0.02	0.00	0.01	0.00	0.11	0.03	0.00	0.00	0.00
SIGN	0.03	0.02	0.00	0.00	0.06	0.04	0.00	0.00	0.00	0.08	0.09	0.00	0.00	0.00	0.00
IMPAIRMENT	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	**0.56**	0.00	0.05	0.00
ABILITY	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	**0.54**	0.00	0.01	0.00
ACTIVITY	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	**0.50**	0.00	0.06	0.03
PERFORM	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	**0.50**	0.09	0.02	0.00
FUNCTION	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16	**0.46**	0.00	0.00	0.00
SEVERE	0.00	0.05	0.00	0.04	0.00	0.00	0.16	0.00	0.00	0.00	0.00	**0.43**	0.00	0.01	0.00
MENTAL	0.00	0.00	0.00	0.06	0.00	0.00	0.11	0.00	0.00	0.00	0.09	**0.37**	0.00	0.07	0.00
LISTING	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	**0.36**	0.00	0.00	0.00
SGA	0.00	0.04	0.05	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	**0.35**	0.00	0.00	0.15
PHYSICAL	0.00	0.02	0.00	0.02	0.00	0.00	0.12	0.00	0.00	0.00	0.09	**0.34**	0.00	0.03	0.00
CONDITION	0.00	0.14	0.00	0.13	0.00	0.00	0.07	0.00	0.00	0.00	0.05	**0.34**	0.03	0.01	0.00
AFFECT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.34**	0.00	0.03	0.05
MEET	0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.29	0.00	0.01	0.09
TREATMENT	0.00	0.05	0.00	0.07	0.00	0.00	0.11	0.00	0.00	0.10	0.11	0.28	0.00	0.00	0.00
WORK	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.28	0.05	0.00	0.16
PROVE	0.00	0.19	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.27	0.00	0.00	0.03
DISORDER	0.00	0.00	0.00	0.18	0.00	0.00	0.17	0.00	0.00	0.00	0.02	0.25	0.00	0.04	0.00
DIAGNOSIS	0.00	0.02	0.00	0.16	0.00	0.00	0.04	0.00	0.00	0.00	0.08	0.25	0.00	0.00	0.00
TEST	0.00	0.00	0.00	0.02	0.00	0.00	0.08	0.00	0.00	0.00	0.13	0.23	0.00	0.00	0.00
RESULT	0.00	0.14	0.00	0.09	0.00	0.00	0.05	0.01	0.00	0.00	0.03	0.20	0.00	0.03	0.00
LIST	0.00	0.12	0.00	0.03	0.00	0.00	0.05	0.04	0.00	0.00	0.08	0.20	0.00	0.00	0.00
PAST	0.00	0.13	0.05	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.19	0.00	0.05	0.03
TREAT	0.00	0.11	0.00	0.05	0.00	0.00	0.10	0.00	0.00	0.03	0.14	0.19	0.00	0.02	0.00
LIMIT	0.00	0.12	0.00	0.00	0.00	0.00	0.03	0.07	0.00	0.09	0.00	0.18	0.00	0.00	0.16
SHOW	0.00	0.15	0.00	0.06	0.00	0.08	0.04	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.00
DETAIL	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.12	0.15	0.00	0.00	0.00
TYPE	0.00	0.11	0.00	0.00	0.00	0.00	0.13	0.02	0.00	0.00	0.00	0.14	0.00	0.02	0.00
LIGHT	0.00	0.05	0.00	0.00	0.00	0.00	0.07	0.02	0.00	0.00	0.00	0.13	0.00	0.00	0.00
CONTINUE	0.00	0.08	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.06	0.04	0.11	0.05	0.06	0.06
EXPECT	0.00	0.07	0.04	0.00	0.00	0.00	0.07	0.00	0.03	0.00	0.00	0.09	0.00	0.02	0.00
GREAT	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
OPM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.68**	0.00	0.00
RETIREMENT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.65**	0.04	0.09

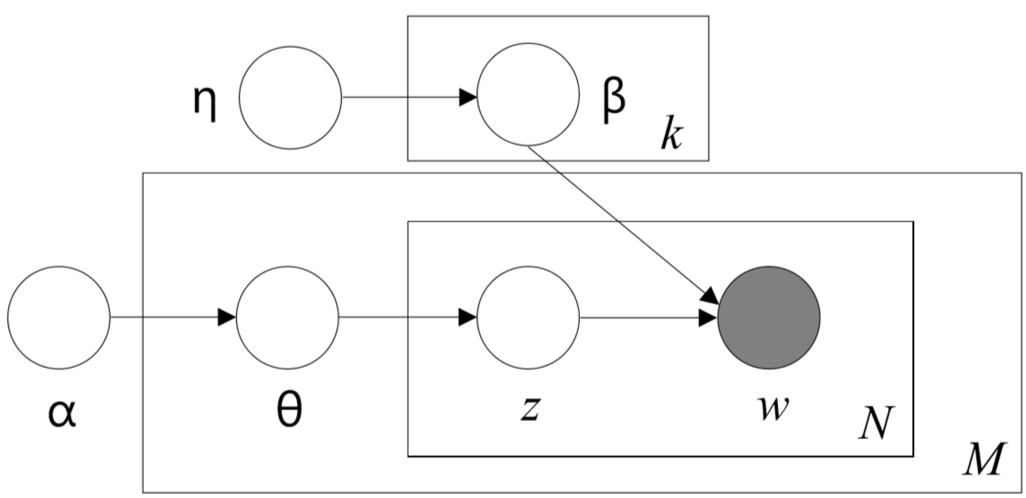
AGENCY	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	**0.48**	0.01	0.00
EMPLOYEE	0.00	0.12	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.07	0.00	0.00	**0.39**	0.02	0.00
FEDERAL	0.10	0.06	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.06	0.00	0.00	**0.37**	0.05	0.04
DR	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.05	**0.34**	0.00	0.00
SECURITY	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.74**	0.14
SOCIAL	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	**0.73**	0.13
INDIVIDUAL	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.07	0.00	0.12	0.00	**0.48**	0.03
CLAIMANT	0.06	0.16	0.00	0.00	0.00	0.00	0.00	0.02	0.03	0.00	0.00	0.13	0.00	**0.42**	0.00
LAW	0.14	0.12	0.00	0.02	0.00	0.00	0.00	0.01	0.00	0.05	0.00	0.00	0.02	**0.36**	0.00
JUNE	0.00	0.00	0.23	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.31**	0.00
FORUM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.30	0.00
DISABILITY	0.00	0.07	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15	0.25	0.30	0.08
BASE	0.00	0.17	0.00	0.08	0.00	0.00	0.00	0.03	0.00	0.03	0.00	0.13	0.00	0.28	0.15
ANSWER	0.00	0.00	0.00	0.00	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.24	0.00
REGARD	0.00	0.13	0.00	0.04	0.00	0.00	0.04	0.03	0.00	0.01	0.02	0.05	0.00	0.14	0.00
QUESTION	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.03	0.08	0.00	0.10	0.00
INCOME	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.02	0.00	0.15	0.00	0.00	0.00	0.00	**0.55**
SSI	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.52**
SSDI	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.45**
EARN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.03	0.09	0.00	**0.40**
QUALIFY	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.14	0.00	0.09	0.00	0.00	**0.36**
AMOUNT	0.00	0.00	0.04	0.00	0.00	0.22	0.00	0.02	0.00	0.04	0.00	0.00	0.01	0.00	**0.36**
BENEFIT	0.00	0.01	0.06	0.00	0.00	0.08	0.00	0.03	0.00	0.02	0.00	0.00	0.00	0.25	**0.36**
ELIGIBLE	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.05	0.00	0.19	0.00	0.00	0.00	0.00	**0.34**
CREDIT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	**0.34**
CHILD	0.00	0.00	0.00	0.00	0.00	0.11	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.31**
DISABLE	0.00	0.08	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.15	0.00	0.26	0.26
TAX	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.01	0.00	0.09	0.00	0.00	0.09	0.00	0.25
AGE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.17	0.15	0.05	0.24
APPLY	0.00	0.06	0.11	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.07	0.01	0.23
FULL	0.00	0.03	0.05	0.00	0.00	0.00	0.08	0.03	0.00	0.01	0.00	0.05	0.10	0.02	0.17
HIGH	0.00	0.00	0.00	0.07	0.00	0.00	0.09	0.00	0.00	0.09	0.00	0.05	0.11	0.00	0.16
SON	0.00	0.00	0.00	0.00	0.00	0.09	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11
SS	0.00	0.00	0.00	0.00	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10
RETURN	0.00	0.08	0.02	0.00	0.07	0.00	0.08	0.00	0.00	0.00	0.06	0.02	0.07	0.01	0.10
DEPEND	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.07	0.00	0.06	0.00	0.00	0.10
CONFUSE	0.00	0.00	0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.06

Appendix B
Latent Dirichlet Allocation

Latent Dirichlet Allocation (LDA) is a generative probabilistic method used to create a model of a corpus (Blei, Ng, and Jordan 2003). The corpus is composed of many documents, with a document defined in this implementation as an individual post. These documents are modeled as being composed of combinations of latent topics, which are in turn composed of word distributions. These probability distributions take the form of a Dirichlet distribution, giving the method its name. Identifying latent topics allows for the discovery of not only the overarching themes present within a corpus but also the distribution of topics through the corpus as well as the words most strongly associated with each topic.

Figure B.1 provides a visualization of the construction method for an LDA model. The corpus is constructed of M documents, each containing an N number of observed words w. There exist k topics with individual topics represented by z. The distribution of topics among documents is specified by θ and parameterized by α; β represents the distribution of words over topics, parameterized by η. While LDA parameters may be numerically estimated in several ways, we used the Gibbs sampling approach to calculate these values (Phan, Nguyen, and Horiguchi 2008).

Figure B.1: Graphic representation of LDA topic modeling



Source: Blei, Ng, and Jordan 2003

One of the assumptions in generating an LDA model is that the number of topics, k, is known and fixed. This is not the case in practice; the optimum number of topics must be ascertained through analysis. An ideal number of topics is large enough to be descriptive without being too general or overly specific, which so noise associated with word frequency to manifest as nonsensical topics. A standardized method for selecting the number of topics within an LDA model does not exist; we employ a combination of four metrics for identification (Arun, Madhavan, and Murthy 2010; Cao et al. 2009; Deveaud, SanJuan, and Bellot 2014; Griffiths and Steyvers 2004). Figure B.2 displays these metrics for models with numbers of topics ranging from 2 to 30. Assessing the local minimum and maximum demanded by the metrics described by Cao et al. (2009) and Deveaud, SanJuan, and Bellot (2014), we identified 15 as an appropriate number of topics. The other two metrics did not convey very useful information. This topic number is also consistent with the number of topics identified in the NMF model.

Figure B.2: LDA topic number analysis

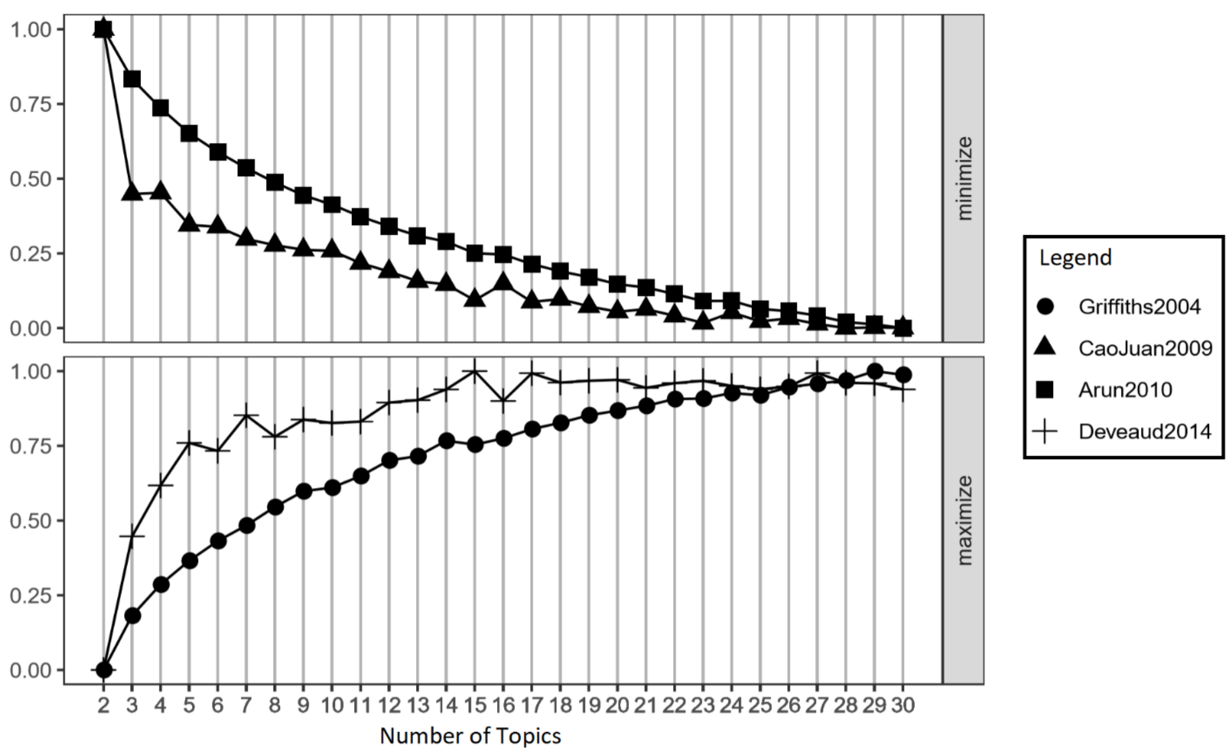


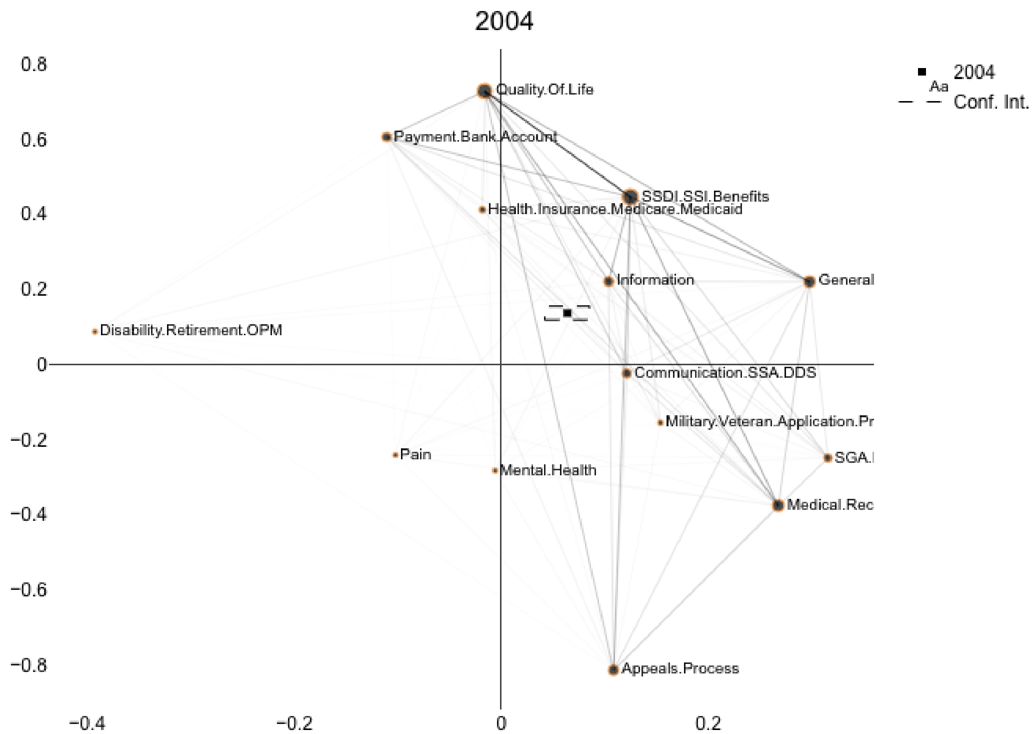
Table B.1 lists the topics derived from the LDA model, including the five words with the highest probability of being associated with each topic. Just as topics are not mutually exclusive to individual documents (i.e., documents exist as combinations of topics), topics may contain words that are present within other topics. The words present are the output of lemmatization performed during the preprocessing of the data. This process converts multiple forms and tenses into a consistent base form. This information is useful when looking at topic composition. For instance, in the "Appeals Process" topic, the word "hear" is very likely converted from the original "hearing."

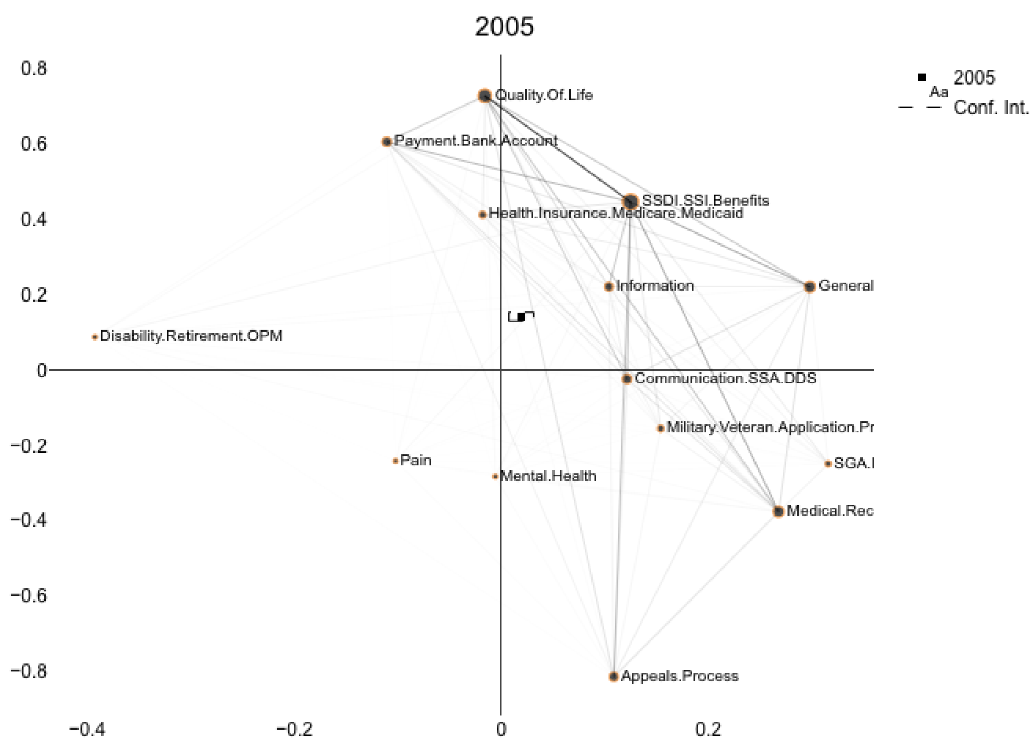
Table B.1: LDA model topics and word composition

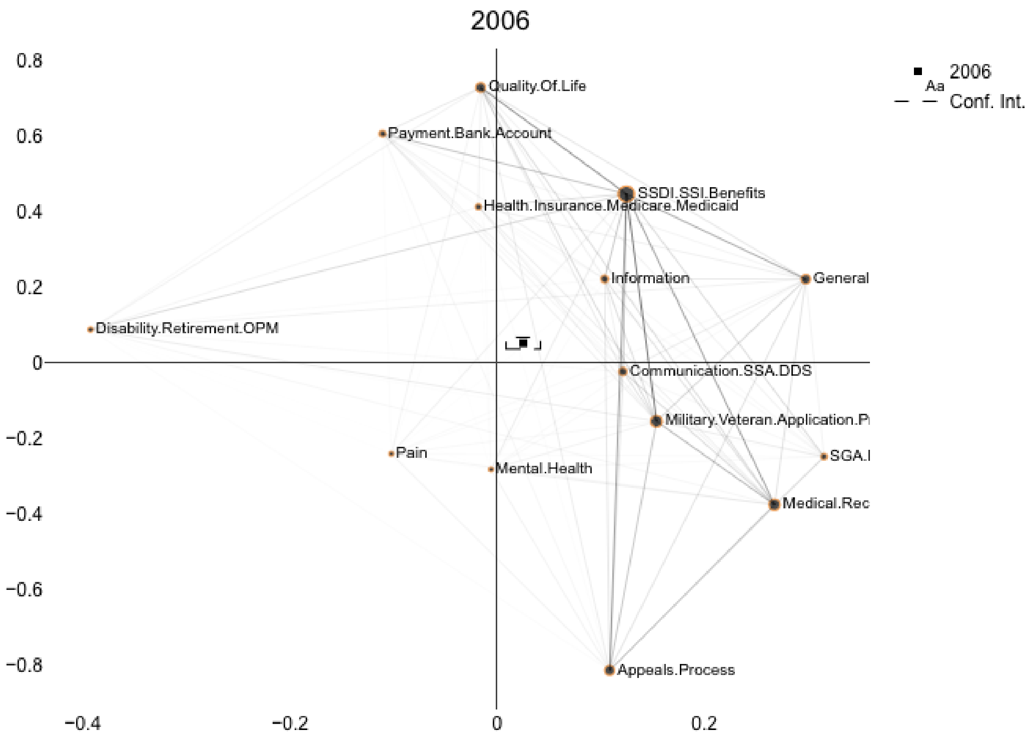
Topic	Terms with Five Highest Probabilities
General Social Security	Security, Social, SSDI, Benefit, SSI
Medical Eligibility	Medical, Meet, Disability, List, Claimant
Appeals Process	Attorney, Appeal, Hear(ing), Judge, Decision
Communication with SSA/DDS	Day, Letter, Call, Office, Tell
Process Anxiety	Time, Bad, Guess, Worry, People
Health Insurance	Check, Pay, Money, Insurance, Medicare
Application Process	Form, SSA, Medical, Record, Review
Medical Examination Process	Mental, Medical, Doctor, Treatment, CE
Approval Decision & Timeframe	Approve, Date, Letter, Receive, Month
Military/Veteran Application Process	File, VA, Service, Claim, Rate
Community Support & Engagement	Hope, Happy, Feel, Luck, News
Forum Utilization	Time, Post, Answer, Question, Read
Pain	Time, Day, Pain, Surgery, Hour
Federal Disability Retirement	Disability, Retirement, Job, Dr., OPM
Quality of Life	Live, People, Family, Money, Life

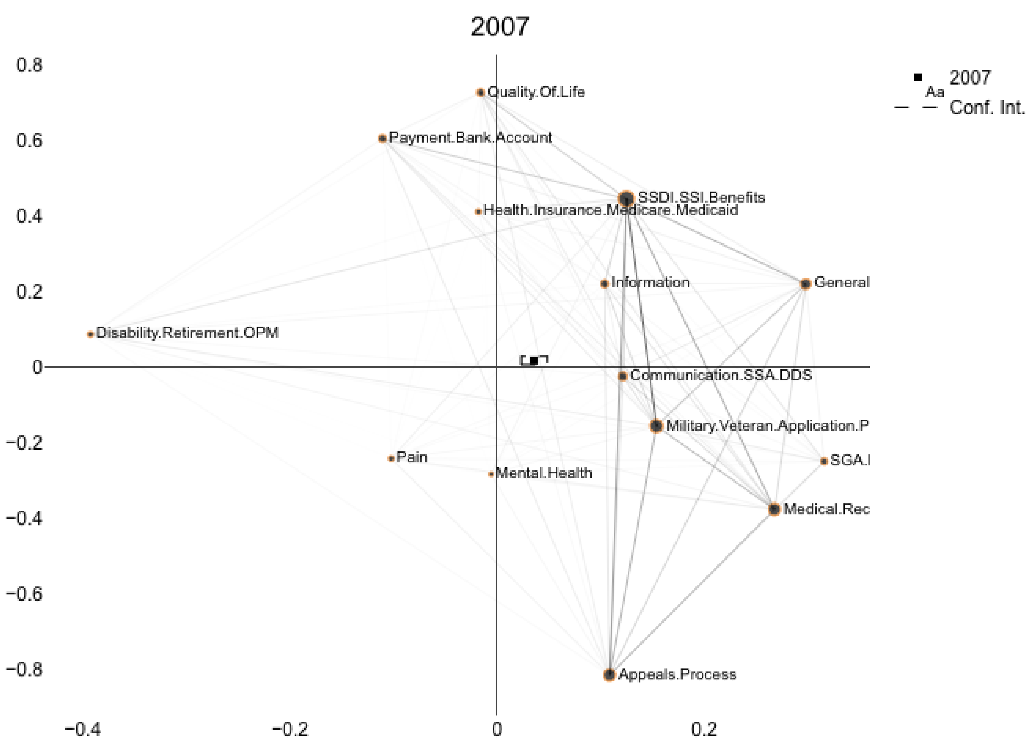
Appendix C

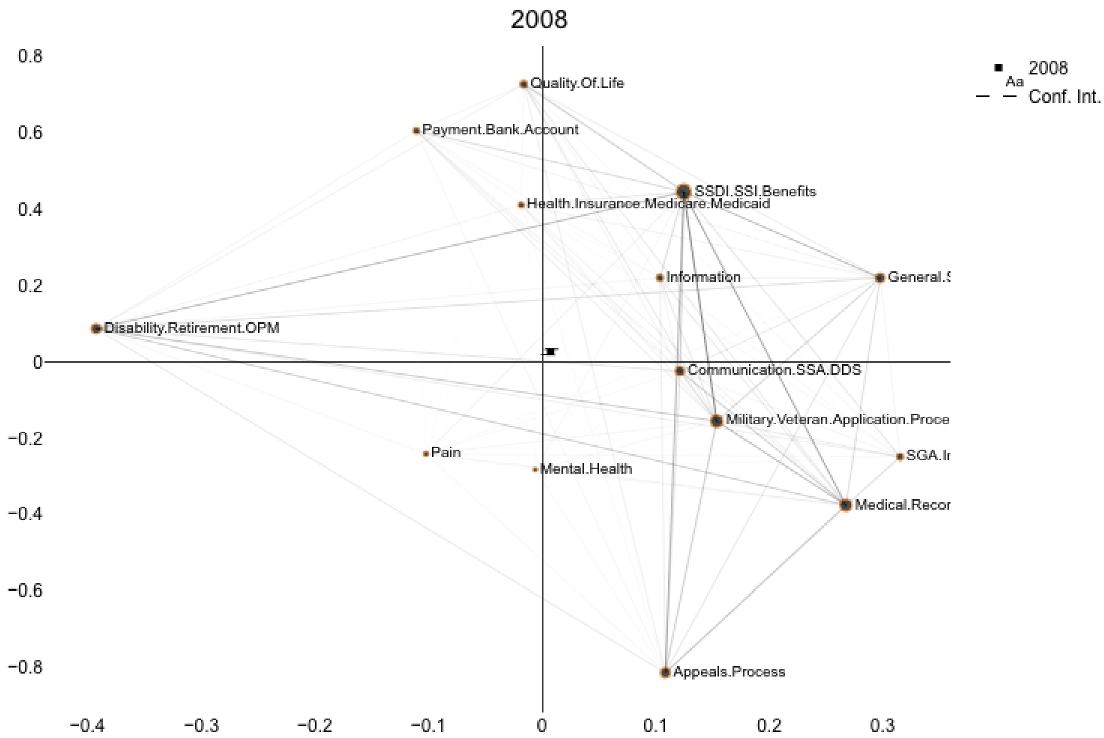
ENAs by Year

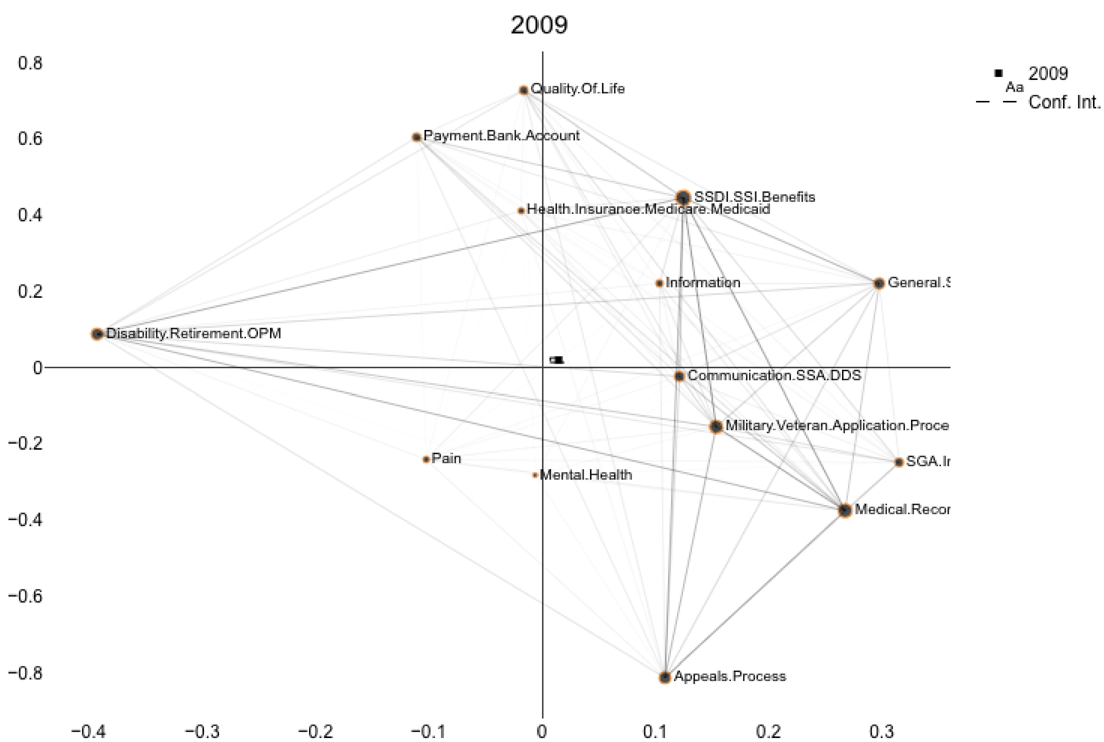


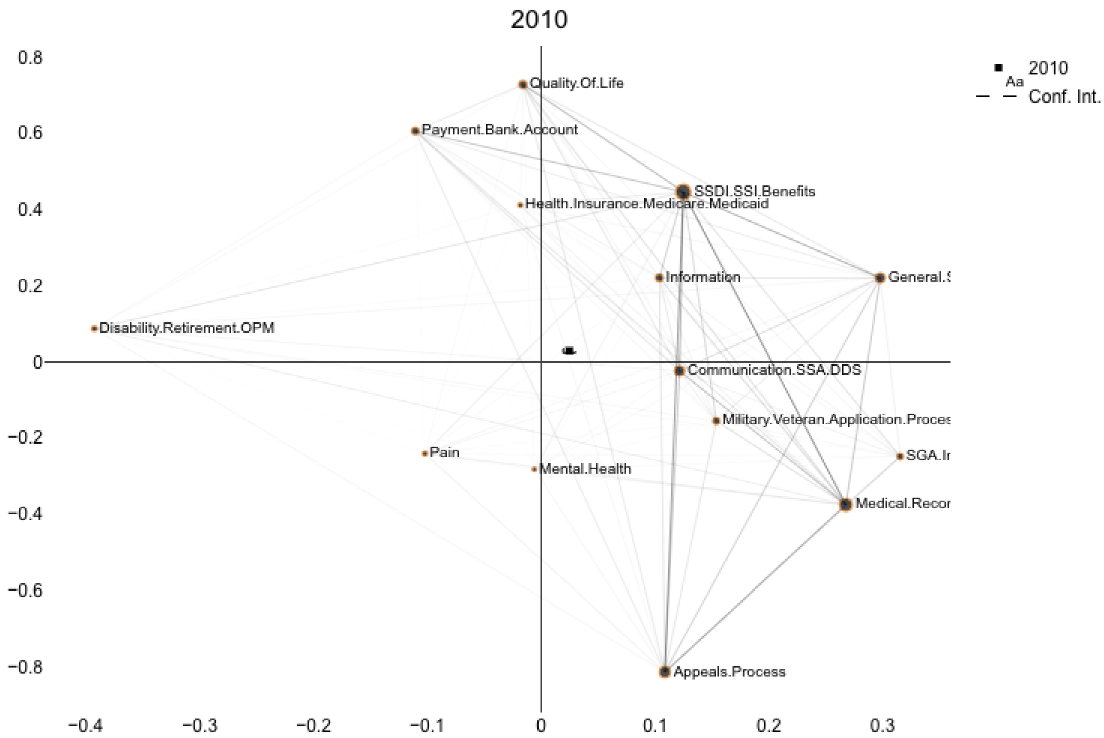


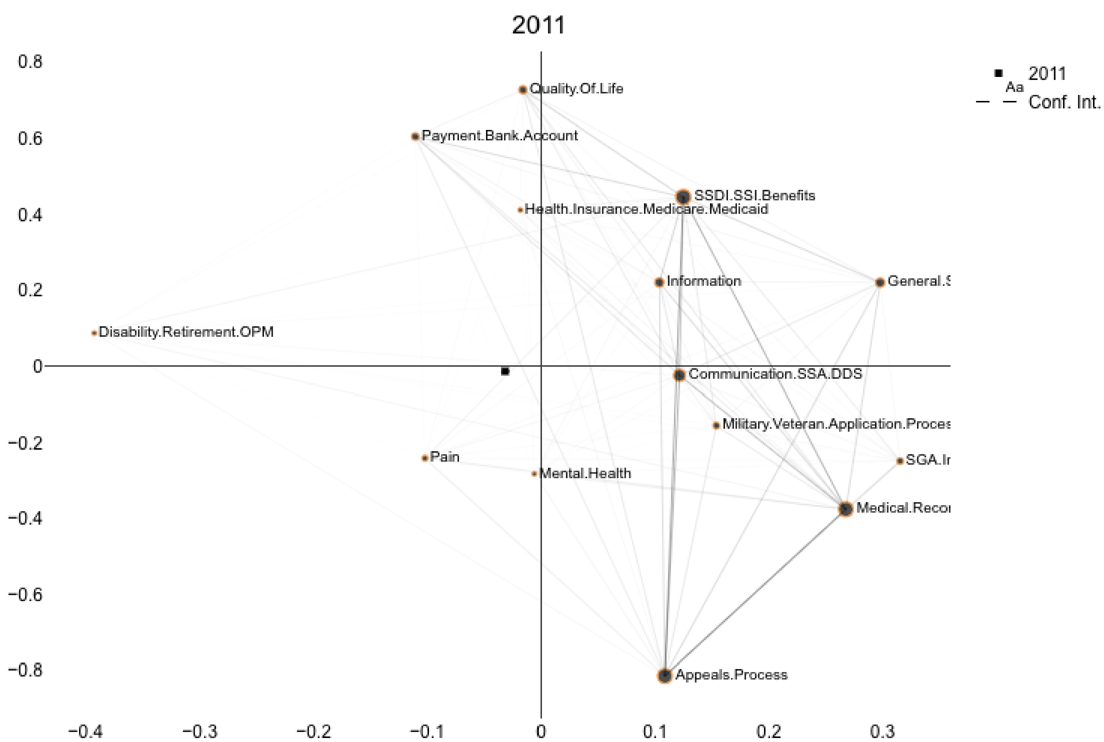

2006
2006
Conf. Int.
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
General
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.P
Pain
SGA.I
Mental.Health
Medical.Rec
Appeals.Process
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.2
0
0.2

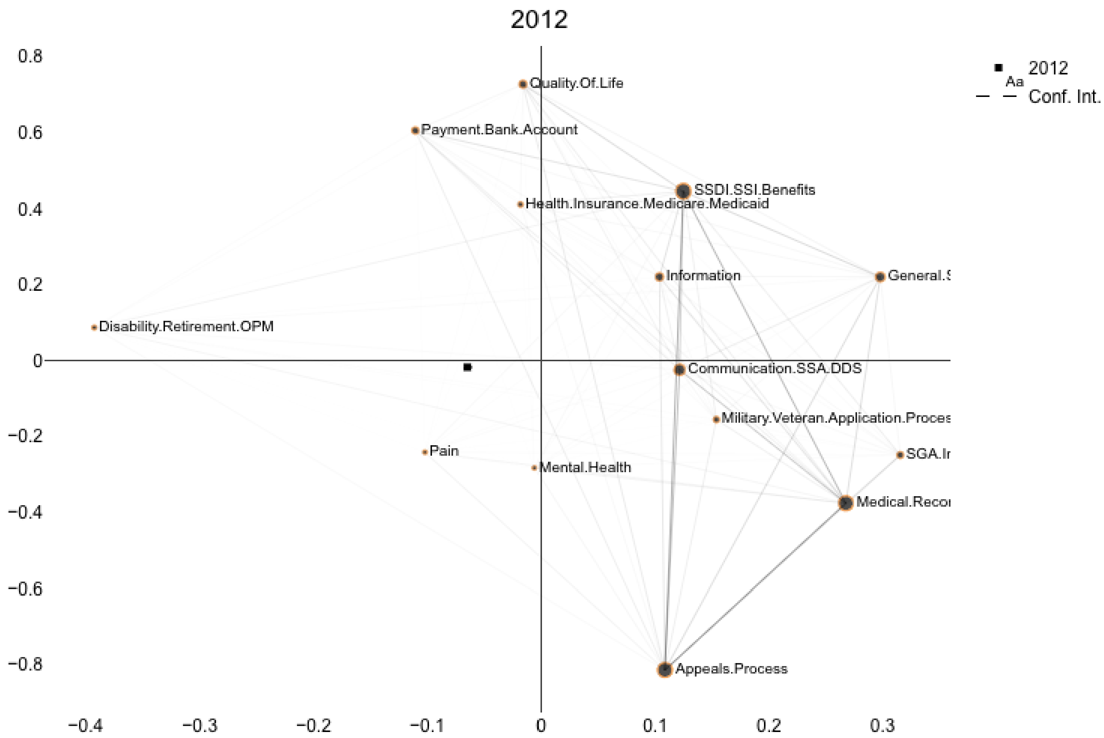

2007
2007
Conf. Int.
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
General
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.P
Pain
SGA.I
Mental.Health
Medical.Rec
Appeals.Process
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.2
0
0.2

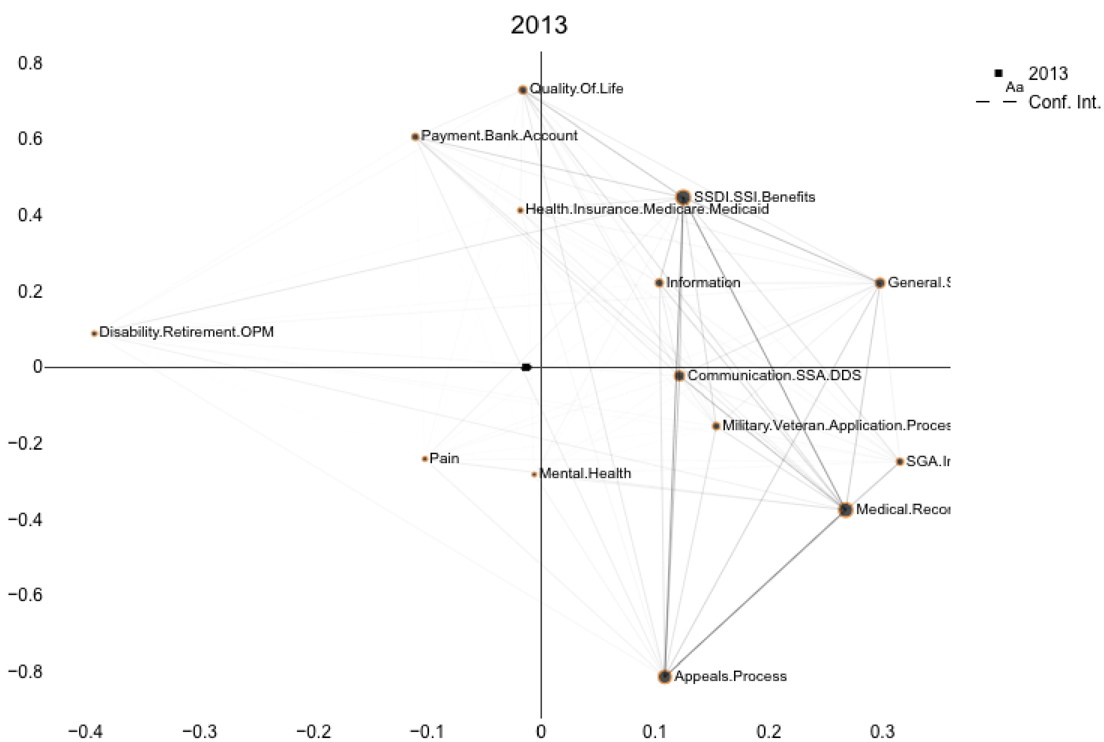

2008
2008
Conf. Int.
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance:Medicare.Medicaid
Information
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proce
Pain
Mental.Health
Medical.Recor
Appeals.Process
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3

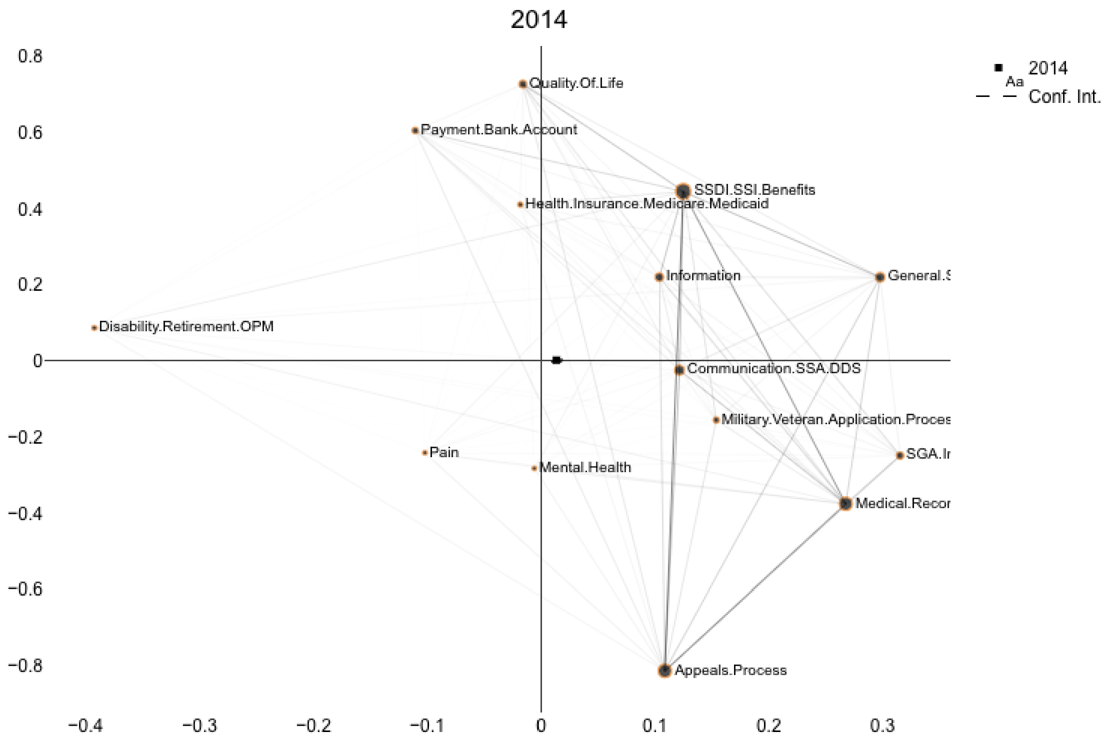

2009
2009
Conf. Int.
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance:Medicare.Medicaid
Information
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proce
Pain
Mental.Health
Medical.Recor
Appeals.Process
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3

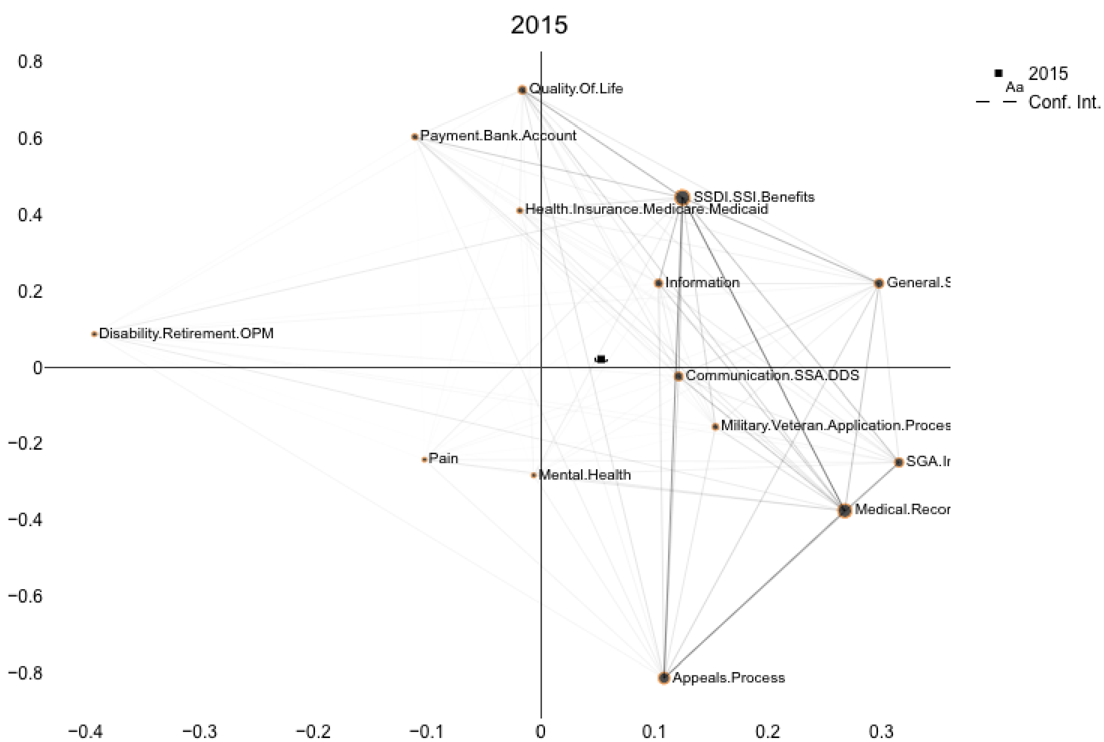

2010
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
Mental.Health
Medical.Recor
Appeals.Process
Aa 2010
Conf. Int.

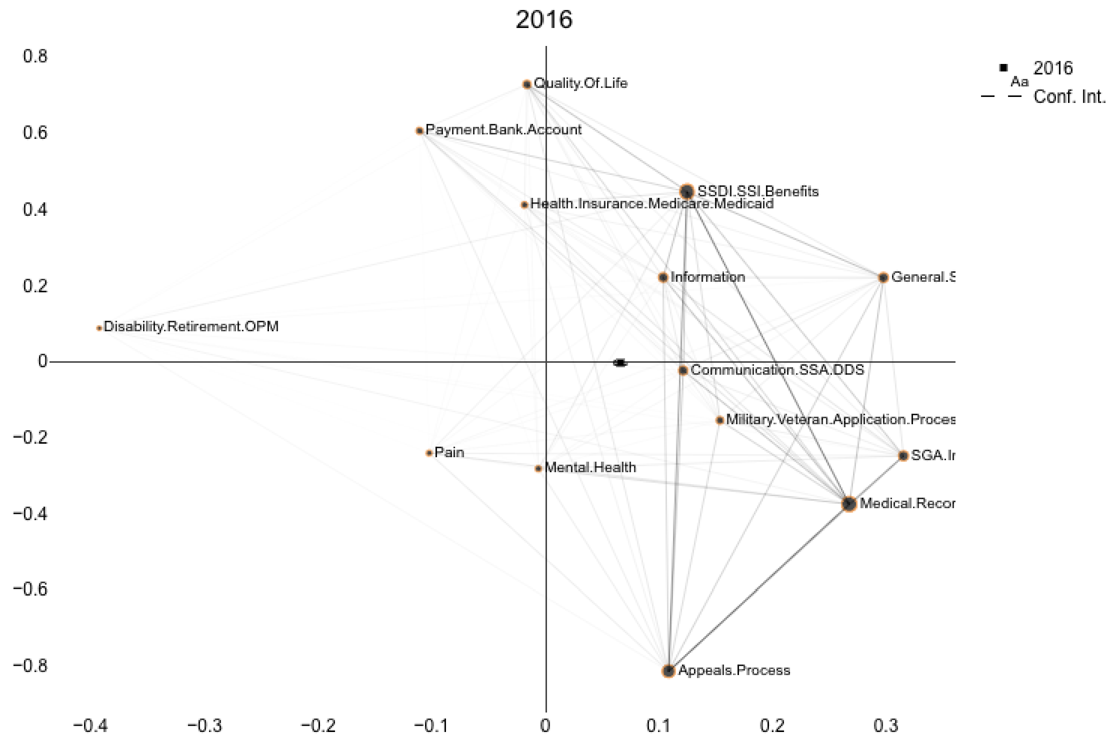

2011
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
Mental.Health
Medical.Recor
Appeals.Process
Aa 2011
Conf. Int.

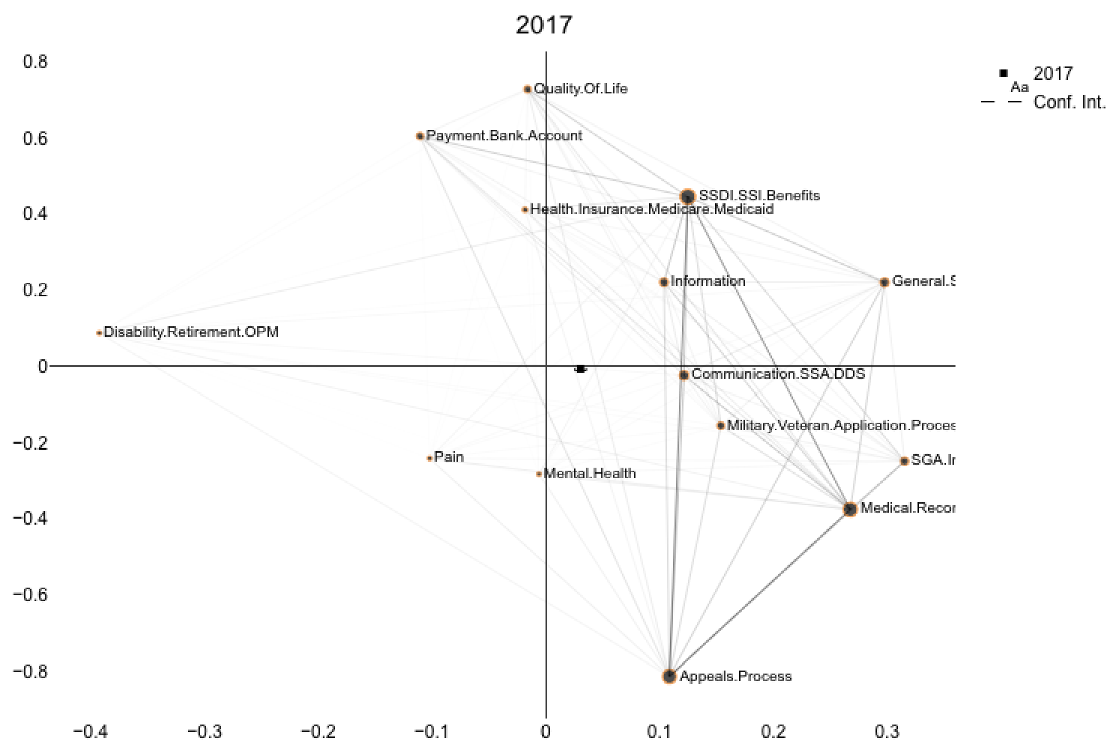

2012
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
Mental.Health
Medical.Recor
Appeals.Process
2012
Aa
Conf. Int.
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3

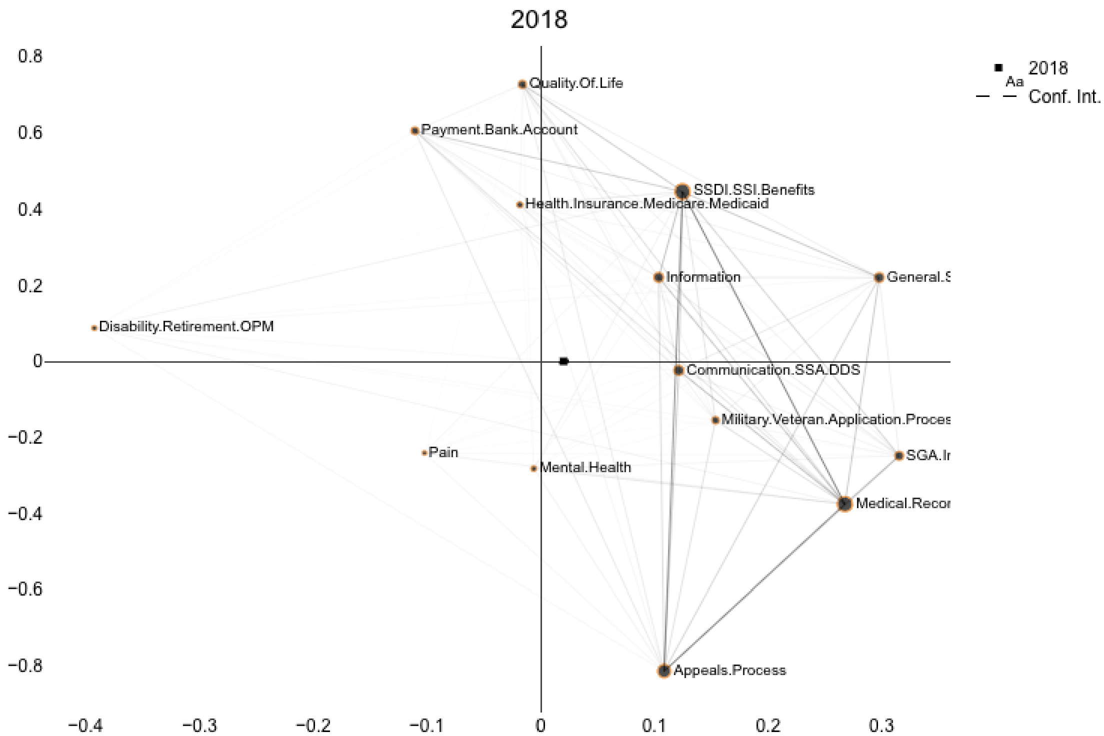

2013
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
Mental.Health
Medical.Recor
Appeals.Process
2013
Aa
Conf. Int.
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3

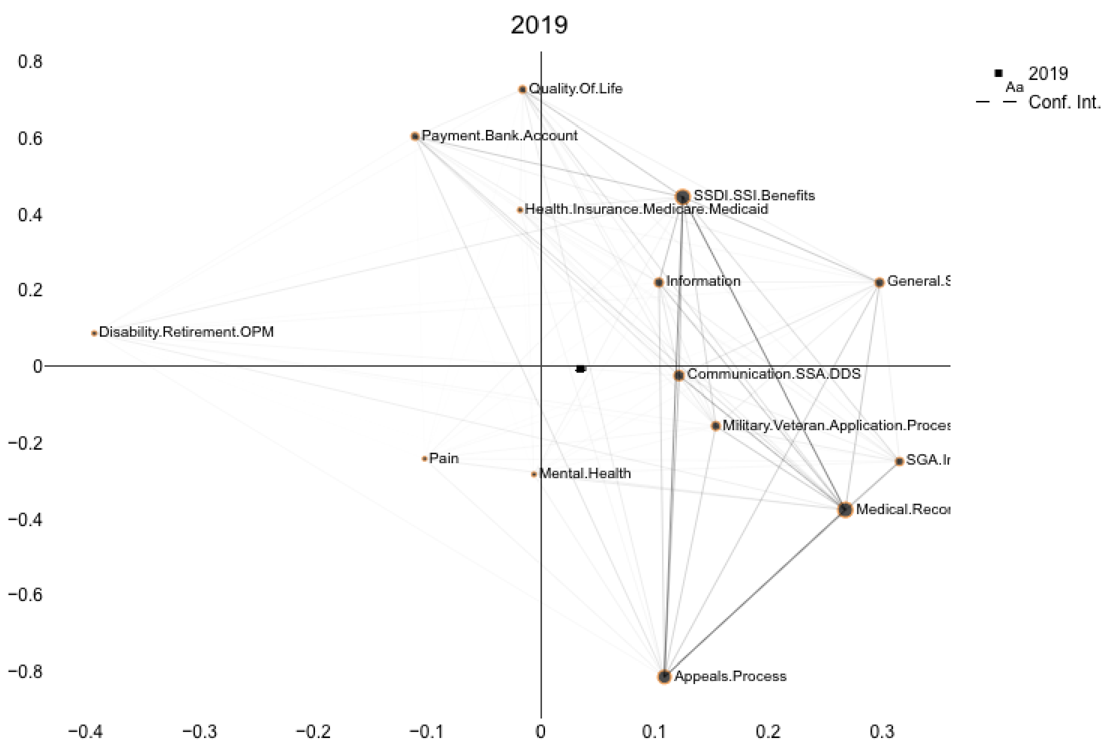

2014
2014
Aa
Conf. Int.
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
General.S
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
SGA.Ir
Mental.Health
Medical.Recor
Appeals.Process
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3


2015
2015
Aa
Conf. Int.
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
General.S
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
SGA.Ir
Mental.Health
Medical.Recor
Appeals.Process
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3


2016
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
General.S
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
Mental.Health
SGA.Ir
Medical.Recor
Appeals.Process
2016
Conf. Int.
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3


2017
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
General.S
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
Mental.Health
SGA.Ir
Medical.Recor
Appeals.Process
2017
Conf. Int.
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3


2018
2018
Conf. Int.
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
Mental.Health
Appeals.Process
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3


2019
2019
Conf. Int.
Quality.Of.Life
Payment.Bank.Account
SSDI.SSI.Benefits
Health.Insurance.Medicare.Medicaid
Information
Disability.Retirement.OPM
Communication.SSA.DDS
Military.Veteran.Application.Proces
Pain
Mental.Health
Appeals.Process
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3

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