DSMVis: Interactive Visual Exploration of the DSM-5 for Mental Health Providers

Ji Won Chung
Smith College

Isha Raut
Mount Holyoke College

Ji Young Yun
Smith College

Kelly Pien
Smith College

Subashini Sridhar
Smith College

Morganne Crouser*
Justice Resource Institute

R. Jordan Crouser†
Smith College

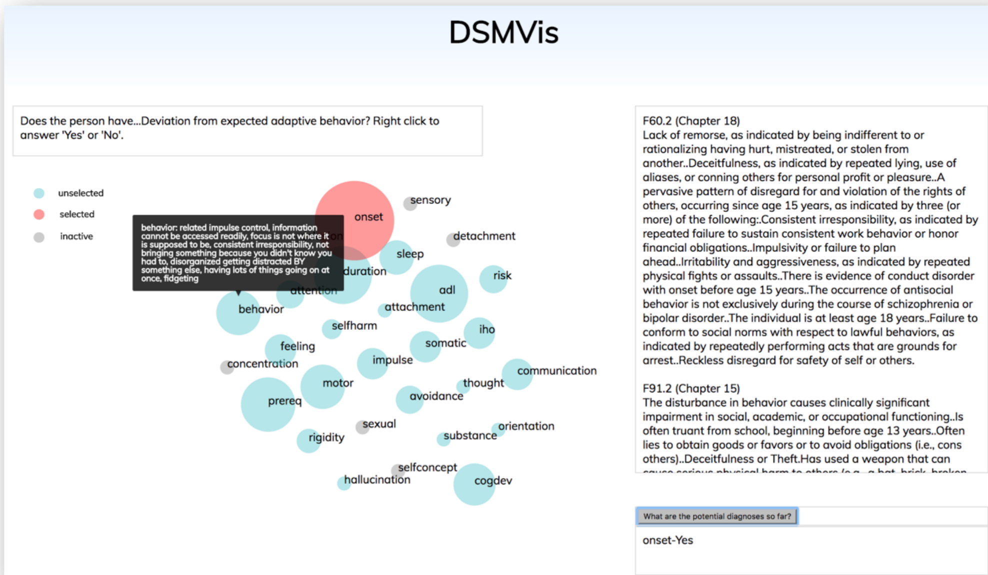

Figure 1: The DSMVis interface uses an interactive bubble chart showing the 29 major categories of symptoms for mental health diagnoses. A clinician can include or exclude symptom categories based on their observation or a client's self-reporting to explore potential diagnoses. For each potential diagnosis, the full text of the diagnostic criteria from the Diagnostic and Statistical Manual (5th ed.) appears on the right-hand side as a reference. To help guide the clinician's reasoning, the circles are dynamically resized to highlight on the proportion of potential diagnoses that include each symptom family.

Abstract

Issues regarding mental illness and substance abuse are some of the most rapidly growing segments of the global disease burden today. Despite this, collaboration between computational scientists and mental health practitioners has been nearly nonexistent. Catalyzed by participating in recent Symposia on Computing and Mental Health[1], the Human Computation and Visualization Laboratory at Smith College has established an ongoing collaboration with community-based clinicians at the Justice Resource Institute (JRI)[2]. In this paper, we present DSMVis, a co-designed interactive visual exploration system designed to help mitigate diagnostic bias.

*e-mail: mcrouser@jri.org
†e-mail: jcrouser@smith.edu, corresponding author

[1]http://mentalhealth.media.mit.edu/
[2]https://www.jri.org/

1 Introduction

Mental illness and substance abuse issues account for more than 189 million disability-adjusted life years (DALYs) every year, a global impact on human life comparable to that of cancer [3]. In the United States, the Diagnostic and Statistical Manual of Mental Disorders (now in its fifth edition, hereafter referred to as the DSM-5) offers a common language and standard criteria for the classification of mental disorders. It is used extensively by both healthcare providers and insurance companies. However, it originally was not created to aid diagnosis but rather to list symptoms of conditions [1]. In addition, 13 separate groups of experts focused developing different parts of the book, resulting in little cross-group collaboration or standardization between chapters [1]. This inherent organizational bias of the DSM-5 is one problem we try to resolve.

Clinicians themselves report that they are also subject to specialization and diagnostic bias. A clinician may have a tendency to look at certain chapters over others depending on their field of expertise. Indeed, the DSM has received significant criticism related to diagnostic reliability – the degree to which different clinicians agree on a diagnosis [2]. The goal of DSMVis is to help mental health clinicians mitigate their diagnostic or specialization bias by facilitating interaction with multiple chapters of the DSM simultaneously.

2 Data Collection

Our current methods of data collection are imperfect and biased. However, the primary goal of the project was to create a working prototype of an interface and we are in the process of making the the techniques more impartial. For this stage of the project, the online version of the DSM-5 was web-scraped using the Python package BeautifulSoup. 207 diagnoses and their associated diagnostic criteria were collected. To better reflect the categorization that practicing clinicians use to diagnose their client, similarly themed symptoms were manually grouped into single categories with the guidance of collaborating clinicians. For instance, fatigue and insomnia were grouped together in a 'sleep' category. The new categories counteract the inherent organizational bias of the DSM-5 and minimize redundant categories. To minimize our own biases, a survey sent to clinicians around the country verified the correct classification of the resulting 29 categories. Each surveyee was asked to correct or suggest a different category if the classification was incorrect. The final data set consisted of 207 rows of diagnoses and 31 columns denoting the chapter the DSM-5 diagnosis was from, the definition of the diagnosis, and 29 columns that represented the 29 categories. A diagnosis containing a category's symptom was marked with a 1, else, a 0.

3 DSMVis

The interface was designed for clinicians, people without a strong background in computer science. Hence, the focus of DSMVis was to be user-friendly and easy to maneuver with little to no training. The bubble chart, created with D3.js, was implemented for the following reasons: (1) it is an intuitive type of visualization, (2) it is a compact way of showing a variety of categorical and quantitative information, (3) it enables different data types to be grouped, compared, and contrasted.

Each bubble represents a symptom and its area represents the frequency that symptom appears over the entire filtered set of diagnoses. To represent the frequency, a form of quantitative data, numerically would provide too much information to the user. The area was used to visually represent frequency because it concisely and intuitively communicates that a larger area corresponds to higher importance. Upon hovering over a bubble, the user can see two features: (1) a black pop-up on the bubble containing the clinician-generated description of the category and (2) a corresponding question that appears on the top left of the screen. In response to the question, the user can right-click a bubble to answer 'yes' or 'no'. The selection of 'yes' turns the circle red and fixes its area to 1, scaled by 10000. The selection of 'no' turns the circle blue and updates the area.

A user can double-click a circle to deselect a selection. The user's selections are shown in a box at the bottom right corner. Upon clicking the 'What are the potential diagnoses so far?' button, the user can see the list of filtered possible diagnoses with their corresponding chapters and descriptions. A legend is located on the top left of the bubble chart to indicate what the three colors represent. Red indicates that the user has selected the bubble, blue indicates an unselected bubble, and gray indicates there is no data to be filtered based on that symptom.

DSMVis revealed insights on the data set made only possible through its visualization. The bubble chart showed the concentration bubble was gray to begin with, indicating there is no diagnosis determined by the concentration category.

4 Evaluation

A clinician from the JRI tested DSMVis by inputting symptoms from test case 8.1 in the DSM-5 Clinical Cases book. The lack of a prescribed starting question and increased autonomy over input symptoms and ordering of questions were appreciated by the test user. The clinician also responded positively to the tooltip's ability to provide additional information on a category in a compact way. The visualization's instant updates regarding color and sizes in response to user input were also well-received.

The test user also indicated a need for a clearer system. The next version of the interface will not include D3's force collision because it overlaps and relocates the bubbles. This makes the bubbles harder to find and bubble labels and tooltip texts harder to read. It was also unclear to the user whether the red bubbles represented the unanswered questions or those answered 'no'. In the future, we plan to refine the legend or add one more bubble mode for clarification. Double-clicking to deselect a bubble is a feature that is easy to learn, but not immediately obvious. From the user we learned that a more intuitive action is to select 'no' from the context menu. However, leaving the option to select 'no' to a question that we previously answered 'yes' confused the user. Therefore, if the double-clicking feature is maintained, the context menu will be deactivated if bubbles are selected.

5 Conclusion and Future Work

Refinement of the interface to increase accuracy and decrease human bias is still in progress. To mitigate our own biases in grouping systems, we will utilize the Natural Language Toolkit to group symptoms based on their cosine similarities. To reduce the number of indistinguishable diagnoses that the system produces, we will include additional filtering options for symptoms such as age of onset, episode and symptom duration. We also plan to address the possibility of comorbidity. Additionally, we will compare overlap rates in chapters of the DSM-5 to those in the International Statistical Classification of Diseases and Related Health Problems (ICD-10) book for medical doctors. We hypothesize the ICD-10's overlap rates should be lower. However, comparable rates could indicate our symptom grouping method is flawed.

In addition, we intend to introduce a confidence rating for the symptoms and diagnoses. A clinician can mark a symptom as 'tentative' if the clinician is uncertain of its presence. The interface will indicate the confidence of each suggested diagnosis. Another major goal is to reduce our own biases and those of a clinician's client. A visual representation of a clinician's proximity to a diagnoses could skew the clinician's reports. Possible solutions include keeping the symptom categories vaguely-worded or computerizing the categorization of symptoms. The interface could text-mine and categorize the input of symptoms made by a clinician. However, this model not only reduces clinician bias, but also a clinician's autonomy over the interface which could reduce a clinician's trust in the interface.

Additionally, a client's exposure to the interface could incline them to answer questions in a biased manner. As part of our future work, we will work with clinicians to evaluate the accessibility of the interface to the public, and critically examine the role and impact of technological diagnostic aides in clinician-client interactions.

Acknowledgments

The Justice Resource Institute was instrumental in providing their domain expertise. D3.tip was created by Justin Palmer and adapted for D3 version four by Constantin Gavrilete and David Gotz. D3.contextmenu was created by Patrick Gillespie.

References

[1] A. P. Association. *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association, Arlington, VA, 5th ed., 2013. doi: 10.1176/appi.books.9780890425596.Introduction

[2] R. Freedman, D. A. Lewis, R. Michels, D. S. Pine, S. K. Schultz, C. A. Tamminga, G. O. Gabbard, S. S.-F. Gau, D. C. Javitt, M. A. Oquendo, et al. The initial field trials of dsm-5: new blooms and old thorns, 2013.

[3] H. A. Whiteford, L. Degenhardt, J. Rehm, A. J. Baxter, A. J. Ferrari, H. E. Erskine, F. J. Charlson, R. E. Norman, A. D. Flaxman, N. Johns, et al. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. *The Lancet*, 382(9904):1575–1586, 2013.