

[Open in app](#)[Follow](#)

520K Followers



# Counsel Chat: Bootstrapping High-Quality Therapy Data

There is a dearth of good psychotherapy data. Let's work on that.



Nicolas Bertagnolli May 11 · 10 min read



Photo by [Polina Zimmerman](#) from [Pexels](#)

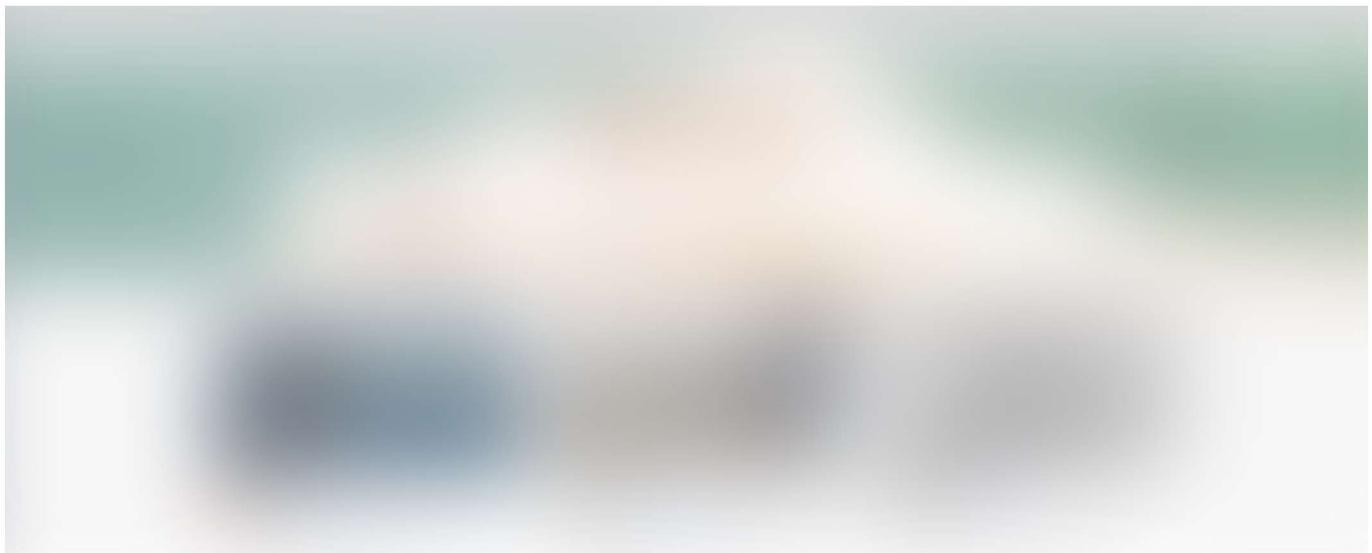
[Open in app](#)

thing I found particularly difficult in this domain is the lack of high-quality data. Sure, you can go scrape Reddit and get some interesting therapeutic interactions between individuals, but in the author's opinion, this is a poor substitute for what an actual interaction between a client and a therapist is like. Don't get me wrong, there are datasets. They are just, more often than not, proprietary or pay to play.

My hope with this post is to introduce a data set of reasonably high-quality therapist responses to mental health questions from real patients. I will discuss the data source, basic information about what is in the data set, and show some simple models we can train using this data culminating with training a chatbot! I am unaffiliated with [counselchat.com](https://counselchat.com), but I think they are doing good work and you should check them out. All code for this project is available [here](#).

An unfortunate fact about Medium is that it doesn't allow you to coauthor pieces. This was a joint project with [Dr. Grin Lord](#) who both suggested this project in the first place and helped with all of the analysis. Also with the help of the CounselChat Co-Founders, Eric Ström and Phil Lee. Phil is a serial entrepreneur focused on innovative machine learning and data. Eric is an attorney and licensed Mental Health counselor.

## Counsel Chat



Screenshot of Counselchat.com's landing page.

[Open in app](#)

On the site, therapists respond to questions posed by clients, and users can like responses that they find most helpful. It's a nice idea and lends itself to some interesting data.

The coolest thing about this data is that there are verified therapists posting the responses. Not every reply is excellent, but we know that it comes from a domain expert. If you were using Reddit data the person providing advice could be anyone. Here we know that the individuals providing the advice are qualified counselors. Pretty neat! It is important to keep in mind that in-person interactions with a therapist are often very different from what we see publicly online. Another thing, this is not a dialogue between a therapist and a patient. It only involves a single talk turn.

## The Data

Initially, we scraped the data from [www.counselchat.com](http://www.counselchat.com). But after reaching out to the founders of counselchat.com for comment they provided us with all of their data for this article! That data dump of both the scraped data and true data is available here as a [CSV](#). So special thanks to Philip and Eric for being so kind and willing to share what they've built with the community. All analysis is done using my original scraped data.

There are 31 topics on the forum, with the number of posted responses ranging from 317 for the topic of “depression” to 3 for “military issues” (Figure 1–3). There are 307 therapist contributors on the site, most of whom are located on the West Coast of the US (Washington, Oregon, California). They range in licensing from Ph.D. level psychologists, social workers, and licensed mental health counselors.

The site had some duplication of questions and answers that appeared to be generated by the users themselves (e.g., where a therapist copy-pasted a response from one question to another similar question); we did not clean the data for this kind of duplication.

The dataset is presented as a CSV with 10 columns described below:

- **questionID** — A unique question identifier which is distinct for every question
- **questionTitle** — The title of the question on counsel chat

[Open in app](#)

- **topic** — The topic the question was listed under
- **therapistInfo** — The summary of each therapist, usually a name and specialty
- **therapistURL** — a link to the therapist's bio on counselchat
- **answerText** — The therapist response to the question
- **upvotes** — The number of upvotes the answerText received
- **split** — The data split for training, validation, and testing.

In general, most questions have only a few responses with 75% of questions having two or fewer total responses. However, many questions have a lot of therapist engagement. The most commented-on question is Do I have too many issues for counseling? It's nice to see such wonderful participation from therapists on questions like this. As a minor aside, one thing which I really enjoy about working with therapy data is that you get to see humanity's capacity for kindness and understanding.



Figure 1 — Visualizing the number of responses

[Open in app](#)

intimacy.



Figure 2 — Visualizing the count of questions by topic

But, funny enough, counseling-fundamentals jumped 12 places to take the number three spot in the number of responses. It looks like therapists have lots of opinions about therapy : ).



[Open in app](#)

In general, most questions are pretty short. The therapists seem to be providing much longer responses. The average question length is 54 words but the average response is 170 words long.

Counselchat.com, like any good social website these days, has the ability to upvote a therapist's response to a question. If we take a look at the number of responses that have upvotes we can see that about 30% of responses get upvoted. Most responses did not receive upvotes. The range of upvotes for a single counselor response to a question was from 0 to 8; with the median response receiving 1 upvote.

[Open in app](#)

## Upvote Prediction

To see what might contribute to an upvote I trained a simple classifier using TF-IDF on n-grams, one using BERT features, and one that combined the two. Unfortunately predicting upvotes appears kind of difficult. Our first pass TF-IDF model's performance isn't great. By using BERT we can squeak out a little bit higher precision but still not good overall. For the BERT model, I used BERT as a feature extractor as I did in this [other post](#).



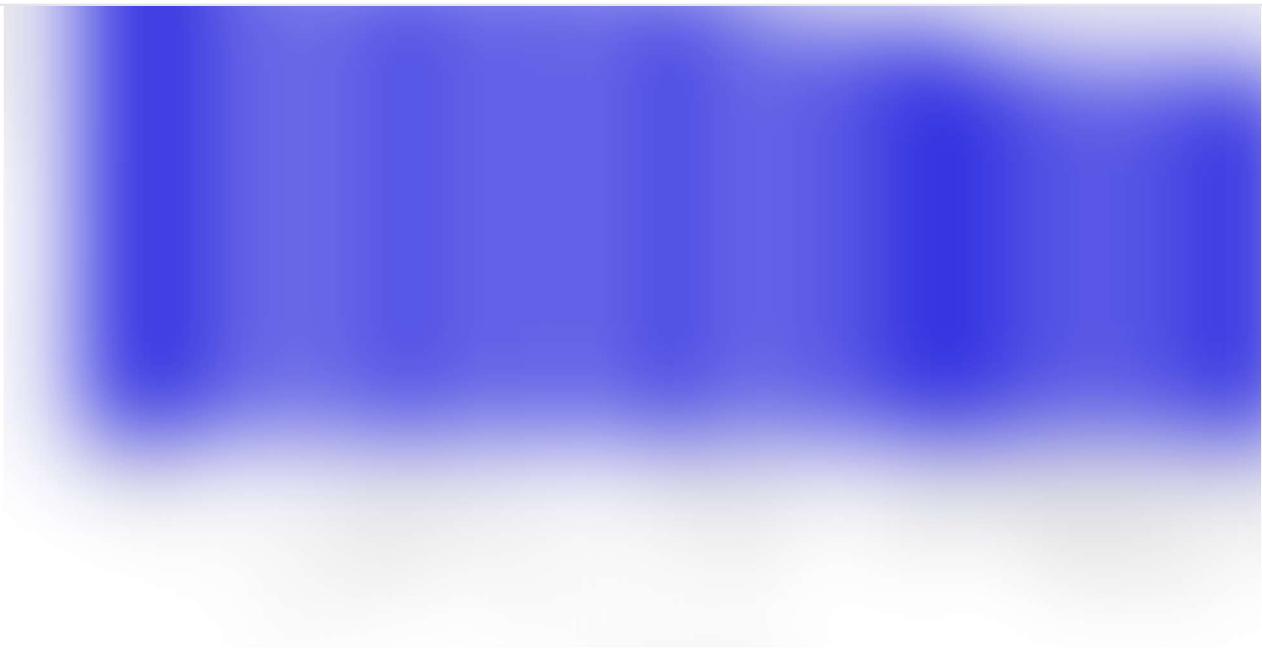
## Training a Topic Classifier

Looking beyond upvotes, classifying therapist responses into different categories is also interesting. It's sometimes useful to know if people are talking about depression, or maybe intimacy. To that end, I trained an SVM on TF-IDF features. Since BERT features didn't seem to be giving us all that much here I went with a more interpretable model.

[Open in app](#)

Unfortunately, performance on the validation set doesn't look great. If we look at support in the validation set we see that many topics only show up one or two times, so it's no surprise that we have basically no ability to predict these. Over time these results should improve with more data, especially as more people use the platform.

I took a look at the top features for this model using the code I describe in [this post](#). It looks like words related to drug use, duty, and loss pop up. It seems somewhat reasonable given that anxiety, depression, grief-and-loss could all be related to "loss." "Substance-abuse" is also a category that I'd bet "12 step" is pretty strongly correlated with. Anyway, it's good to spot check these models and make sure they are producing words that make some intuitive sense.

[Open in app](#)

## Visualizing Topics

I thought it might also be interesting to see how well BERT features divide the space so I created some simple UMAP 2D embeddings. We don't see a strong separation between the classes in general. However, different groups of topics do appear closer together in some cases and further apart in others. Take workplace relationships (purple) for example, it's very very close to relationship-dissolution (black), but completely separate from counseling fundamentals (bright green).

[Open in app](#)

## Training a Therapist Chat Bot

Recently there has been an explosion of apps trying to make mental health more accessible using conversational agents, see [woebot.io](https://woebot.io), or [wysa.com](https://wysa.com) to get an idea of what's out there. Since mental health bots are so hot right now I figured we should train one with our new data.

Training reasonably good chatbots has become surprisingly easy. I don't want to do a tutorial on training chatbots here when Hugging Face has such an [exceptional walkthrough](#). Though it's worth touching on some of my experiences. To train this chatbot I used their code almost verbatim. I made two tiny modifications to the code and had to parse the counselchat.com data into the correct form for their transformer based model. Ultimately, I had to write and modify less than 50 lines of code.

### Code Modifications

The Hugging Face model is awesome for getting a pretty decent chatbot up and running without much data. I didn't want my bot to have any personality so I had to comment out line 93 inside the **get\_data\_loaders** function. This line makes some permutations of the existing persona, but I wasn't passing in any personas to the model so this is empty and needs to be commented out.

```
persona = [persona[-1]] + persona[:-1] # permuted personalities
```

[Open in app](#)

```
assert max_l <= 512
```

inside the **pad\_dataset** function. BERT and GPT2 have some input size restrictions and when these are exceeded you can get some cryptic errors in the middle of training like:

```
RuntimeError: index out of range: Tried to access index 512 out of a  
table with 511
```

Adding this line helped me find out if any of my data was too big before running a model. It doesn't help the model at all but was super helpful when I was debugging : ).

## Converting the Data for Chat Bot Training

The data format for the Hugging Face transformer model seemed a bit confusing at first, but it was fairly easy to generate. The training data needs to be a JSON file with the following signature:

```
{
  "train": [
    {
      "personality": [
        "sentence",
        "sentence"
      ],
      "utterances": [
        {
          "candidates": [
            "candidate 1",
            "candidate 2",
            "true response"
          ],
          "history": [
            "response 1",
            "response 2",
            "etc..."
          ]
        }
      ]
    }
  ]
}
```

[Open in app](#)

Let's break this down a little bit. The larger JSON object has two main keys. "train" and "valid". Train is the training data and is a list of personality, utterances pairs. Valid is the same but for the validation set. The personality is a list of sentences defining the personality of the speaker. See the Hugging Face tutorial for more details on this. For my model, I simply left this blank to indicate no personality information. The candidates section holds a list of candidate responses. This list contains some non-optimal responses to the history of the conversation where the last sentence is the ground truth response. Lastly, we have to define the history. In general, this is a list of strings where each position holds a new talk turn. For our bot, we only have the two talk turns, one from the person asking the question and one as the therapist's response. So we just fill this in with the question. We do this because we are looking for good single responses to problems, not training a general therapist bot.

If you want another example of how to format the data, Hugging Face has a good one in the repository you can find [here](#).

With the data properly formatted we can pass it directly to the Hugging Face model. I trained mine with the following specification:

```
python3 train.py --dataset_path counsel_chat_250-tokens.json --  
gradient_accumulation_steps=4 --lm_coef=2.0 --max_history=1 --  
n_epochs=3 --num_candidates=4 --train_batch_size=2
```

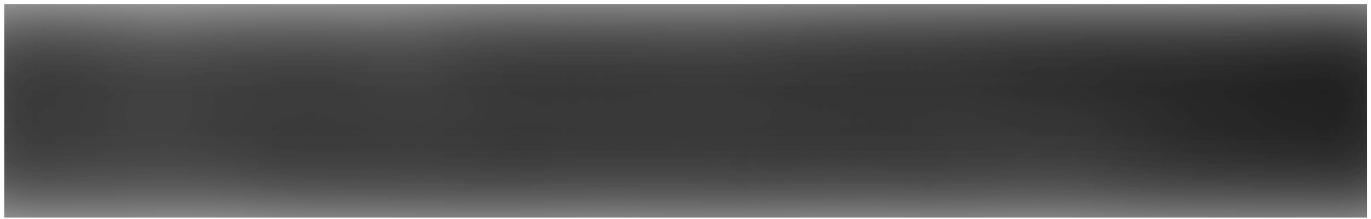
Voila! We have a chatbot trained on real therapist responses. If you'd like to play with the model please download it [here](#). Then all you need to do is unzip the file and run

```
make build  
make interact CHECKPOINT_DIR=counselchat_convai
```

[Open in app](#)

and increase the memory allowed for your container in Docker -> Preferences to something like 5GB. This seems to be a common problem when folks try and build this container.

Here's an example of the dialogue! We can see that the model is a good listener. Providing advice, and some empathy and compassion. It also seems to be handy in apocalyptic scenarios offering to bring me tools.



Some decent responses

## Conclusion

It's hard to get access to good therapist-patient interactions, but there is good data out there if you look around. Counselchat is an excellent source of limited quality therapist interactions. I hope you find some cool applications of this psychotherapy data in your field.

---

## Sign up for The Daily Pick

By Towards Data Science

Hands-on real-world examples, research, tutorials, and cutting-edge techniques delivered Monday to Thursday. Make learning your daily ritual. [Take a look](#)

[Get this newsletter](#)

Emails will be sent to thoshithakaduru.dev@gmail.com.

[Not you?](#)

[Open in app](#)[About](#) [Help](#) [Legal](#)

Get the Medium app

