Conversational Playlist Creation Dataset

Dataset: storage.googleapis.com/gresearch/cpcd/index.html

The Conversational Playlist Creation Dataset (CPCD) contains 917 music-seeking conversations paired with song ratings collected using a wizard-of-oz methodology. This dataset is intended for research on how users express their preferences in consumption domains, like music, and to evaluate conversational recommendation systems. The dataset also includes questions asked by the wizard to elicit preferences from the user.

Data Card

DATASET TEAM(S)

Conversational Playlist Creation

DATASET CONTACT

- Group Email: conversational-playlist-team@google.com
- Website: www.google.com

DATASET AUTHORS

- Arun Tejasvi Chaganty
- Megan Eileen Leszczynski
- Ravi Ganti
- Shu Zhang
- Filip Radlinski
- Krisztian Balog

PRIMARY DATA MODALITY

Image Data
Text Data
Tabular Data
Audio Data
Video Data
Time Series
Graph Data
Geospatial Data
Multimodal (Please specify)
Others (please specify)
Unknown

DATASET SNAPSHOT

Size of dataset	9MB
Number of Conversations	917
Number of Dialog Turns	4,921
Number of Tracks	106,736
Labeled Classes	Positive / Negative
Number of Labels	8,550
Average labels per instance	N/A
Algorithmic Labels	N/A
Human Labels	N/A
Other	N/A

DESCRIPTION OF CONTENT

Each datapoint is a conversation between two speakers, a “user” and a music recommendation “system” (which is simulated by a human “wizard”). Each turn of the conversation includes a music-seeking query from the user, a list of search queries executed by the wizard, the wizard’s response and which of the shown results were ultimately liked or disliked by the user. The datapoint also includes metadata for all the tracks referenced in the conversation.

DATASET SUBJECT

EXAMPLE: DATA POINT

DATA FIELDS

Sensitive Data about people

Non-Sensitive Data about people

Data about natural phenomena

Data about places and objects

Synthetically generated data

Data about systems or products and their behaviors

Unknown

Others*

(*please specify)

This is what a typical data point looks like.

Example Data Point:


```
{
  "id": "00079c9c8dd7b55a",
  "turns": [
    {
      "user_query": "I'd like to create a
playlist that is perfect for a dance party. I'd
like to create a playlist that I can utilize for
a dance party.",
      "system_response": "Hello! I can help with
that! What kind of genre or artist would you like
to add to your playlist?",
      "search_queries": [],
      "search_results": [],
      "liked_results": [],
      "disliked_results": []
    },
    {
      "user_query": "Any genre works as long as
it is upbeat and something people can jump around
and sing to!",
      "system_response": "Here is a selection of
songs that match your preferences. Any other
artist or genre?",
      "search_queries": [
        "Party Rock Anthem",
        "Don't Stop Believin"
      ],
      "search_results": [
        [
          "3JDHlCcn-UY",
          "9acyYo1JdRg"
        ],
        [
          "0cBShyKk-w0",
          "14d5a5pxiV8"
        ],
      ],
      "liked_results": [
        "Jx_O6PHdWww",
        "OIPmhkzN2ug"
      ],
      "disliked_results": [
        "tYvFa2ARD24"
      ]
    }
  ],
  "tracks": {
    "4z-bOdAdias": {
      "track_artists": [
        "Beyoncé"
      ],
      "track_release_titles": "I AM...SASHA
FIERCE",
```


Each datapoint is a conversation consisting of the following fields:

- **id (string):** a unique identifier for this conversation.
- **turns (list[Turn]):** A list of turns in the conversation.
- **tracks (dict[string, Track]):** The metadata associated with each track referenced in turns above.
- **goal_playlist (list[string]):** The list of track ids for the final target list of “liked” tracks in this conversation.

}

Each turn consists of:

- **user_query (string):** user query for this turn.
- **system_response (string):** wizard response for this turn.
- **search_queries (list[string]):** list of queries entered by the wizard this turn.
- **search_results (list[string]):** list of track ids retrieved for each query in “search_queries”.
- **liked_results (list[string]):** list of track ids shown to users and liked by them.
- **disliked_results (list[string]):** list of track ids shown to users and disliked by them.

Each track consists of:

- **track_ids (string):** identifier for the track. It corresponds to a YouTube video id.
- **track_titles (string):** title of the track.
- **track_artists (list[string]):** names of the artist on this track.
- **track_release_titles (string):** title of the release or album.

```
      "track_titles": "Single Ladies (Put a Ring
on It)",
      "track_ids": "4z-bOdAdias"
    },
    "7YQESUr8Cxc": {
      "track_artists": [
        "Gwen Stefani"
      ],
      "track_release_titles": "Love Angel Music
Baby (Deluxe Version)",
      "track_titles": "Hollaback Girl",
      "track_ids": "7YQESUr8Cxc"
    }
  },
  "goal_playlist": [
    "4z-bOdAdias",
    "7YQESUr8Cxc",
    "BciS5krYL80",
    "Jx_O6PHdWww"
  ]
}
```


DATASET PURPOSE(S)	KEY DOMAINS OR APPLICATION(S)	PRIMARY MOTIVATION(S)
Monitoring **Research** Production Others (please specify)	Domains Natural Language Processing, Recommendation Systems, Conversational AI Problem Space Conversational recommender systems	• Provide a dataset of music-seeking conversations paired with item ratings. • Evaluate conversational recommendation systems.
DATASET USAGE	**INTENDED AND/OR SUITABLE USE CASE(S)**	**UNSUITABLE USE CASE(S)**
Safe for production use **Safe for research use** Conditional use- some unsafe applications Only approved use Others (please specify)	Qualitative analysis of music-seeking user queries in a conversational system. Training conversational recommendation systems.	The dataset was created in accordance with Google's AI Principles and is not intended to be used in a way that would cause or likely to cause overall harm.
SAFETY OF USE WITH OTHER DATA	**ACCEPTABLE TRANSFORMATIONS**	**BEST PRACTICES FOR JOINING OR AGGREGATING WITH DATASET**

Safe to use with other data Conditionally safe to use with other data Should not be used with other data Unknown Others* (Please specify)	**Joining with other datasets** **Subsampling and splitting** **Filtering** **Joining input sources** Cleaning missing values **Anomaly detection** **Grouping and summarizing** **Scaling and reducing** **Statistical transformations** Redaction or Anonymization Others (please specify)	The dataset includes the video ids for songs on YouTube. For example, the YouTube video for a track id `7YQESUr8Cxc` can be accessed at `https://music.youtube.com/watch?v=7YQESUr8Cxc`.

VERSION STATUS	DATASET VERSION	MAINTENANCE PLAN
Regularly Updated New versions of the dataset have been or will continue to be made available. Actively Maintained No new versions will be made available, but this dataset will be actively maintained, including but not limited to updates to the data. **Limited Maintenance** The data will not be updated, but any technical issues will be addressed. Deprecated This dataset is obsolete or is no longer being maintained.	**Current Version** 1.0.0 **Last Updated** 01/2023 **Release Date** TBD	• **Versioning:** N/A - CPCD is a static dataset. Minor releases correspond to any errors fixed in the dataset. • **Update:** CPCD is not updated. • **Errors:** Please contact conversational-playlist-team@google.com. • **Feedback:** Please contact conversational-playlist-team@google.com.

ACCESS POLICY	RETENTION POLICY	WIPEOUT POLICY
CPCD is an open-access public dataset.	N/A (Public data exemption)	N/A (Public data exemption)

DATA COLLECTION METHODS	DATA SOURCES	DATA COLLECTION

API Artificially Generated **Crowdsourced - Paid** Crowdsourced - Volunteer Vendor Collection Efforts Scraped or Crawled Survey, forms or polls Taken from other existing datasets Unknown To be determined Others (please specify)	N/A	Data was collected through a platform where two human annotators (a user and a wizard) interacted. Users initiated conversations by requesting for a playlist for a specific purpose (e.g. “music to lift me up when I’m sad”). Wizards could search and add songs on YouTube through the interface, and were encouraged to use Google web search to research recommendations. Wizards could also ask users questions (“did you have a particular genre in mind?”). Users can rate songs added by the wizard. Users submitted their conversations at the end of 5 rounds and after rating at least 15 songs.
INCLUSION CRITERIA	**EXCLUSION CRITERIA**	**DATA PROCESSING**
N/A	All turns where the user or wizard coordinated on the task (e.g, “Hello, are you there?”, “Thanks!”) were filtered.	N/A (no data processing was applied)
SENSITIVE DATA	**FIELDS WITH SENSITIVE DATA**	**SECURITY AND PRIVACY HANDLING**
User Content User Metadata User Activity Data Identifiable Data S/PII Business Data Employee Data Pseudonymous Data Anonymous Data Health Data Children’s Data **None** Others* (*please specify)	N/A	N/A
SENSITIVE HUMAN ATTRIBUTES	**SOURCE(S) OF HUMAN ATTRIBUTES**	**RATIONALE FOR COLLECTING HUMAN ATTRIBUTES**

Race Gender Ethnicity Socio-economic status Geography Language Sexual Orientation Religion Age Culture Disability Experience or Seniority **None** Others (please specify)	N/A	N/A
ANNOTATION WORKFORCE TYPE	**ANNOTATION CHARACTERISTICS**	**ANNOTATION DESCRIPTION**
Annotation Target in Data Machine-generated Annotations **Human Annotations - Expert** **Human Annotations - Non-expert** Human Annotations - Employees Human Annotations - Contractors Human Annotations - Crowdsourcing Human Annotations - Outsourced / Managed Teams Unlabeled Others* (*Please specify)	N/A (dataset was created through annotator interactions)	N/A
	ANNOTATOR BREAKDOWN	**ANNOTATOR DESCRIPTION**
	Users Annotator type: Paid - Non-Expert Total unique annotators: 111 Expertise of annotators: Trained for task Wizards Annotator type: Paid - Expert Total unique annotators: 11 Expertise of annotators: Trained for task	Users Annotators were recruited from an online crowdworking marketplace. Annotators were required to be fluent English speakers from the United States, and to be regular music listeners; they did not require a music background. Annotators were provided slides as training material: the slides included instructions on how to use the interface and some broad guidelines on what types of queries are preferred (conversational, detailed and stating broad preferences) and not preferred (short “commands” or fixating on specific songs). Users were surveyed for which

		genres of music they are familiar with; the survey results were used to match users and wizards. Wizards Annotators were recruited from a trusted vendor supplier of full-time crowdworkers. Annotators were required to be fluent English speakers from the United States, and to be extremely familiar with at least one music genre. Annotators were provided slides as training material and several rounds of feedback in early pilot iterations of the task. The slides included instructions on how to use the interface and some broad guidelines on how to search for relevant recommendations and to elicit user preferences. Wizards were surveyed for which genres of music they are familiar with; the survey results were used to match users and wizards.