**Twitter Query Methodology: Discussions of Eric**

In designing the queries for Twitter’s API, I wanted to strike a balance between the volume of data returned and the usefulness of that data – ie. we want a broad enough dataset to do interesting analysis on, but not at the cost of diluting the signal-to-noise ratio with junk tweets. I divided my queries into three primary domains:

1. Tweets directly referencing Eric, his exhibit, and/or the Kickstarter project to rebuild him
2. Tweets adjacent to references of Eric, eg. retweets & replies of tweets from category 1
3. Tweets from the geographic area of the London Science Museum during the time period Eric’s exhibit was displayed

The categories aren’t strictly defined, and each has some overlap with the others. They were more designed to let me visualize different areas to search in than to categorize the results; once duplicate results are discarded they can all be collated and analyzed as part of one dataset (although the data determining which set of tweets came from which query/which category will be preserved if we ever do want to use the categories for analysis).

Within each category, I aimed to start with a focused query at the center and then “branch out” from there into broader and broader queries. The first category was fairly simple in this regard – I started with eg. “Tweets from 2016 – 2018 that contain the phrase ‘Eric the Robot’”, before branching out to more and more general queries such as “Tweets from 2016 – 2018 that contain the words ‘Eric’ and ‘Robot’” or “Tweets from 2016-2018 that contain the words ‘Kickstarter’ and ‘Eric’”. The second category was also fairly simple, and could be largely built off of the first, as it’s trivial to search for eg. “Retweets and replies to tweets from 2016 – 2018 that contain the phrase ‘Eric the Robot’” instead of the tweets themselves.

The third was slightly more complicated, as it contained both geospatial and relational factors. I designed a set of core queries for this category (eg. “Tweets from the time period of Eric’s exhibit that contain the word ‘robotics’”) to be repeated, and then performed each of them three times, each time with a broader geographic area – expanding from a 200m radius around the Science Museum (effectively covering only the Science Museum itself) to 750m and 1500m radii, respectively. This category could be further explored via examining retweets and replies to tweets captured in this category.

**Twitter Query Methodology: General Sentiment Towards Robots (Real & Fictional)**

As robots, automation, and AI occupy an enormous space in the modern consciousness, there are an order of magnitude or two more tweets discussing them than Eric specifically; as such, designing the categories for the broad-topic search required a slightly different approach than the previous more focused categories.

Instead of several smaller queries, I constructed a single, monolithic query for each subcategory I came up with, and raised the API’s response limit to 100,000 tweets each to ensure a robust dataset. The two main divisions I worked with were between

1. Discussions of robots in fiction/pop culture vs. robots in real life
2. Discussions of broad concepts vs. specific models of robot (or robotic characters)

Following the above divisions resulted in four categories:

1. Discussions of fictional robots in a general sense (eg. mentions of droids or cyborgs)
2. Discussions of fictional robots as characters (eg. C-3PO or R2-D2)
3. Discussions of real robots in a general sense (eg. mentions of automation or drones)
4. Discussions of specific models of real robots (eg. Boston Dynamics’ Spot or NASA’s Curiosity)

Additionally, although not strictly related to the topic of robotics, I thought that the popular conception of AI & machine learning in the current zeitgeist was at least robotics-adjacent, and so included a fifth category:

1. Discussions of AI/Machine Learning, both in a general sense (eg. mentions of deep learning or artificial neural networks) and any discussions of specific AI models (eg. mentions of the GPT or DALL-E series).

Of note is that when a search term might have returned ambiguous or unrelated results due to non-specific naming (eg. Star Trek’s Commander Data returning results about data science), I searched for the relevant term being used with at least one contextual keyword at the same time (eg. “Star Trek” or “Commander”). Although not perfect, this method at least allows for some collection of data where it would otherwise be obscured by unrelated topics.

**Twitter Queries (Plaintext)**

**Category 1 (time period: 2016-03-01 – 2018-03-01)**

Tweets containing the words “Eric” and “Robot”

Tweets containing the phrase “Eric the Robot”

Tweets containing the words “Eric”, “Robot”, and either “Museum” or “Restoration”

**[control]** Tweets containing the words “Eric” and “Museum”

**[control]** Tweets containing the words “Eric” and “Kickstarter”

**[control]** Tweets containing the words “Robot” and “Kickstarter”

Tweets containing the words “Robot”, “Kickstarter”, and “Museum”

Tweets containing the words “Kickstarter”, “Restoration”, and either “Eric” or “Robot”

Tweets containing the words “Kickstarter”, “Museum”, and “Restoration”

Tweets containing the phrases “London Science Museum” and “Eric”

Tweets containing the phrases “London Science Museum” and “Robot”

Tweets containing the phrases “London Science Museum”, “Robot”, and “Kickstarter”

Tweets containing the words “Russell”, “Richards”, or “Reffell”, and “Eric” or “Robot”

Tweets containing the words “Russell”, “Richards”, or “Reffell”, and “Robot”

Tweets by the Science Museum’s account containing the words “Eric” and “Robot”

Tweets by the Science Museum’s account containing the words “Eric”

**[control]** Tweets by the Science Museum’s account containing the words “Robot”

Tweets by the Science Museum’s account containing “Kickstarter” and “Eric” or “Robot”

**Category 2 (time period: 2016-03-01 – 2018-03-01)** **(potential edge case not caught: main tweet mentions “Eric” and “Robot”, replies only refer to him as “Eric”)**

Tweets replying to the Science Museum’s account containing the words “Eric” or “Robot”

Retweets of the Science Museum’s account containing the words “Eric” or “Robot”

Retweets of the Science Museum’s account containing the words “Eric” and “Robot”

Retweets of the Science Museum’s account containing the word “Eric”

Tweets that are replies containing the words “Eric” and “Robot”

Tweets that are retweets containing the words “Eric” and “Robot”

Tweets that are quotes containing the words “Eric” and “Robot”

**Category 3 (time period: 2016-08-01 – 2017-10-01)**

Tweets containing the words “Eric” and “Robot” that have geotagging enabled

Tweets from within a 200m radius of the Science Museum that contain the word “Eric”

**[control]** Tweets from within a 200m radius of the Science Museum that contain the word “Robot”

**[control]** Tweets from within a 200m radius of the Science Museum

Tweets from within a 750m radius of the Science Museum that contain the word “Eric”

**[control]** Tweets from within a 750m radius of the Science Museum that contain the word “Robot”

**[control]** Tweets from within a 750m radius of the Science Museum

Tweets from within a 1500m radius of the Science Museum that contain the words “Eric” and “Robot” (returned error)

**[control]** Tweets from within a 1500m radius of the Science Museum that contain the word “Robot”

Tweets from within London that contain the words “Eric” and “Robot”

Tweets from within Great Britain that contain the words “Eric” and “Robot”

Tweets from within Great Britain or Ireland that contain the words “Eric” and “Robot” (no tweets from Ireland, same results as previous query – discarded)

**Category 4 (time period: 2006-03-22 – 2023-04-01)**

Fictional Robots (General Concepts): Cyborg, Android, Droid, Cybernetics, Mecha, Power Armour, Von Neumann Probe, Laws of Robotics, Nanobot, Nanomachine, Replicant

Fictional Robots (Specific/Character-Based): RoboCop, TARS, CASE, T-800, T-1000, Johnny Five, Robbie the Robot, Cylons, Cybermen, Daleks, Decepticons/Autobots, Data, C-3PO, R2-D2, BB-8, Ultron

Real Robots (General Concepts): Bot, robot, Automata, Bionics, Boston Dynamics, Powered Exoskeleton, Automation, Industrial Robot, Surgical Robot, Drone, Teleoperation, Telepresence, Space Probes, Rovers

Real Robots (Specific Models): ~~Mechanical Turk~~, Spot, Atlas, BigDog, Roombas, iRobot, ASIMO, iCub, NASA Valkyrie, Spirit (rover), Opportunity (rover), Curiosity (rover), Voyagers 1 & 2 (probes)

Machine Learning/AI: Machine Learning, AI, Artificial Intelligence, Deep Learning, ANN/Artificial Neural Network, GAN/Generative Adversarial Network, DALL-E, GPT, LLM/Large Language Model, Generative model, OCR/Optical Character Recognition, CAPTCHA/reCAPTCHA, Turing Test, Chinese Room, Self-Driving, Siri, Alexa

**Twitter Queries (Verbatim)**

**Fictional Robots (General Concepts):** "cyborg" OR "android" OR "droid" OR "cybernetics" OR "mecha" OR "mech" OR ("power armour" OR "power armor") OR "laws of robotics" OR "nanobot" OR "nanomachine" OR ("replicant" ("blade runner" OR "robot" OR "cyborg" OR "android"))

**Fictional Robots (Specific/Character-Based):** "RoboCop" OR "TARS" OR ("CASE" ("INTERSTELLAR" OR "TARS" OR "ROBOT")) OR ("T-800" OR "T800") OR ("T-1000" OR "T1000") OR ("Johnny Five" OR "Johnny 5") OR "Robbie the Robot" OR "Cylon" OR ("Cybermen" OR "Cyberman") OR "Dalek" OR "Decepticon" OR "Autobot" OR ("Data" ("Star Trek" OR "Enterprise" OR "android" OR "artificial")) OR ("C-3PO" OR "C3PO" OR "Threepio") OR ("R2-D2" OR "R2D2" OR "Artoo") OR ("BB-8" OR "BB8") OR "Ultron"

**Real Robots (General Concepts):** "bot" OR "robot" OR "automaton" OR "bionic" OR "boston dynamics" OR "powered exoskeleton" OR "automation" OR "industrial robot" OR "surgical robot" OR "drone" OR "teleoperation" OR "telepresence" OR ("probe" ("space" OR "NASA" OR "ESA" OR "Exploration")) OR "rover"

**Real Robots (Specific Models):** ("Spot" ("Boston Dynamics" OR "Robot")) OR ("Atlas" ("Boston Dynamics" OR "Robot")) OR "BigDog" OR "Roomba" OR "iRobot" OR "ASIMO" OR "iCub" OR ("valkyrie" ("NASA" OR "robot")) OR ("Spirit" ("Mars" OR "NASA" OR "science" OR "rover")) OR ("Opportunity" ("Mars" OR "NASA" OR "science" OR "rover")) OR ("Curiosity" ("Mars" OR "NASA" OR "science" OR "rover")) OR ("Voyager" ("1" OR "2" OR "PROBE" OR "NASA" OR "space" OR "probe"))

**Machine Learning/AI:** "machine learning" OR "machine vision" OR "AI" OR "artificial intelligence" OR "deep learning" OR ("ANN" OR "artificial neural network") OR ("GAN" OR "Generative Adversarial Network") OR "DALL-E" OR ("GPT" OR "ChatGPT") OR ("LLM" OR "Large Language Model") OR "generative model" OR ("OCR" OR "Optical Character Recognition") OR ("CAPTCHA" or "reCAPTCHA") OR "Turing Test" OR "Chinese Room" OR ("self-driving" OR "self driving") OR "siri" OR "alexa"