

## Scenario

Machine learning and artificial intelligence has gained notoriety with public missteps for bias, racially insensitive outcomes or over-hyped potential. For example, credit algorithms have been found to extend more credit to males or that natural language models associate Muslims with violent terms. In this case you are asked to review the text of all incidents, understanding frequent terms, associations of interest, sentiment analysis or other relevant findings such as topic modeling. The case is nuanced as it is applying machine learning techniques to descriptions of machine learning missteps. The findings may inform practitioners for common “issues” or help regulators spot trends with misbehaving algorithms.

## Contextual Information:

The database is a curated database of 1200+ articles covering a number of specific incidents. It is robust though not likely complete. Multiple industries, algorithm types, and applications are covered. For example, resume and text processing issues are contained in the database alongside scheduling conflicts causing humanistic hardships for low-wage workers.

## Technical Considerations:

* You may subset the data by channel, or time if that is of interest or you may operate on all incidents.
* You may use a subset of the data provided based on your analysis and/or compute constraints. If so, you must show the manner the data was reduced in a script.

## Example Data

The incident date table

Table

Description automatically generated

The incident summary text table has only the first few hundred characters. This is purposeful as a brief synopsis or abridged version of the data and is NOT meant to be a complete academically sound research project.

Graphical user interface, table

Description automatically generated

## Citation Information:

McGregor, S. (2021) Preventing Repeated Real World AI Failures by Cataloging Incidents: The AI Incident Database. In Proceedings of the Thirty-Third Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-21). Virtual Conference.

Original Data: <https://incidentdatabase.ai/>

## Criteria for Success

Apply at least 2 sentiment analysis lexicons.

Cluster the documents using 1 method from among spherical, kmeans or kmediod to identify a number of clusters by words and/or docs

Build a comparative word cloud.