

Natural Language Processing of Safety Reports in Nuclear Power Plants

University of Southern California

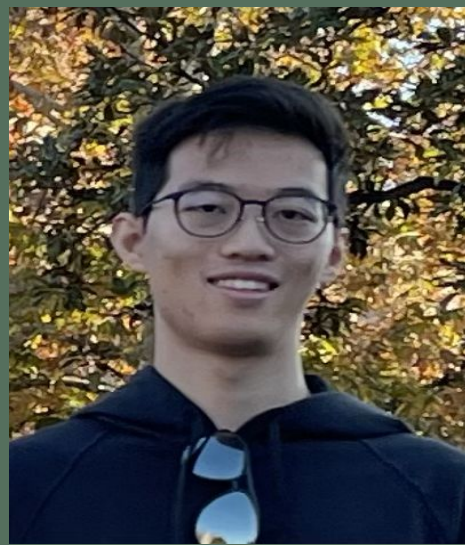
GitHub Link: <https://github.com/ckids-datafirst/2023-fall-nuclear-safety>

The Team



Tejaswi Chaudhari

First Year MSCS
Graduate Student



Xiang Li

First Year ADS
Graduate Student



Rahul Nair

First Year MSCS-AI
Graduate Student



Adrianne Nguyen

First Year CMDS
Graduate Student



Shourya Kothari

First Year MSCS
Graduate Student

**Under the
Guidance of**



Prof. Najmedin Meshkati

USC Viterbi



Prof. Yolanda Gil

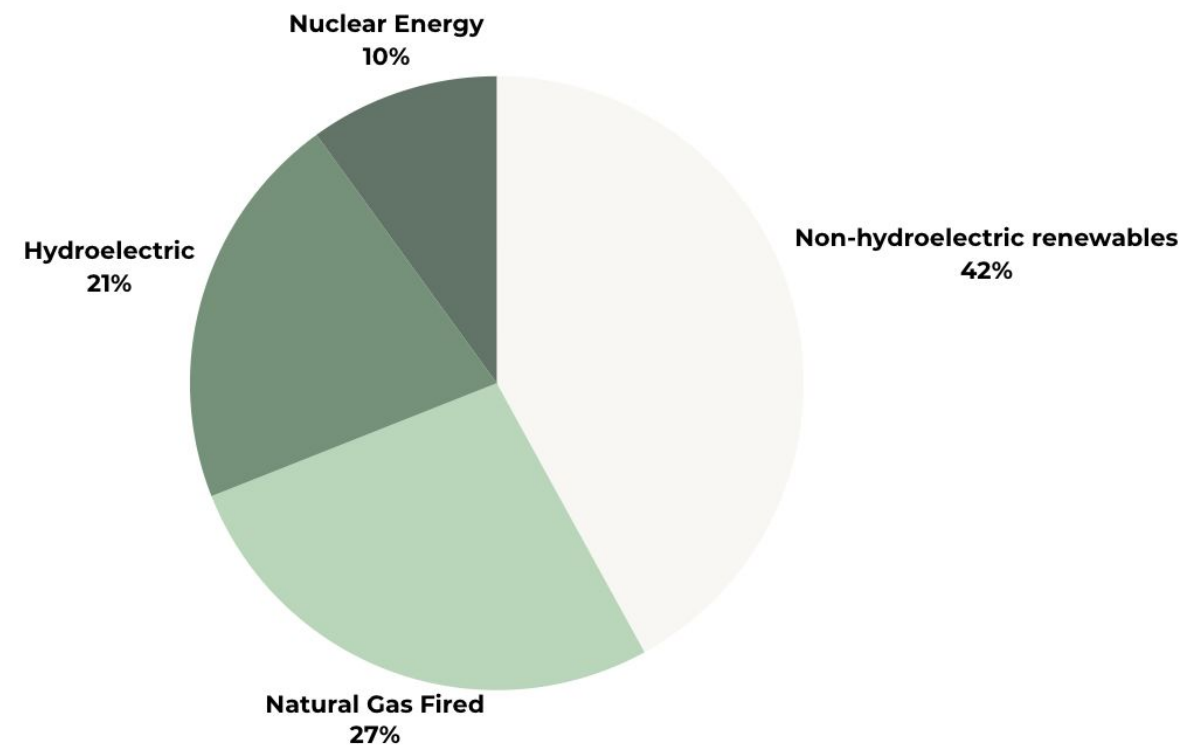
USC Viterbi



Prof. Ulf Hermjakob

USC ISI

Motivation



Diablo Canyon is the **only** nuclear power plant supplying nuclear energy to **California**, which contributes to almost **10%** of the total energy supply of California

Motivation

- Chernobyl and Fukushima disasters resulted in severe public health consequences and environmental hazards
- Following the Nuclear Safety Culture can help us to avoid such disasters in future

Problem

To map nuclear safety issues with nuclear safety traits to improve the safety operations of the power plants and ensure that nuclear safety is maintained according to the safety culture

Data and Resources

A guide by Institute of
Nuclear Power Operations
(INPO), representing 10
Traits of a Healthy Nuclear
Culture

*Pocket Guide
to INPO 12-012:*
**Traits of a
Healthy Nuclear
Safety Culture**

INPO®

Data and Resources

Leadership Safety Values and Actions		Problem Identification and Resolution		Effective Safety Communication	
Respectful Work Environment			Decision-Making		
			Work Processes		
Questioning Attitude	Continuous Learning	Personal Accountability		Environment for Raising Concerns	

Data and Resources

- ~900 page annual reports generated by Diablo Canyon Independent Safety Culture
- 33 annual reports present out of which 13 are available

[32nd Annual Report by the Diablo Canyon Independent Safety Committee, July 1, 2021—June 30, 2022](#)
[Preface](#) | [Executive Summary](#)
[Volume I TOC](#) | [Volume II TOC](#) | [PG&E Response](#) | [Contact the DCISC](#)

Diablo Canyon Independent Safety Committee
Thirty-Second Annual Report on the Safety of Diablo Canyon Nuclear
Power Plant Operations
July 1, 2021—June 30, 2022



Robert J. Budnitz, Chair
Per F. Peterson, Vice-Chair
Peter Lam, Member

Approved: September 28, 2022

A Snapshot of the Dataset

Section	Page	Issue Statement	Safety trait(s)
Volume I 3.1.1 Discussion and Required LERs	38	<i>Unit 2 LER 2021-002-00 was submitted on 12/14/2021, which documented that the Unit 2 reactor was manually tripped on October 15, 2021, in response to increasing water level in a secondary-side feedwater heater. Following shutdown of Unit 2, the feedwater heater was subsequently repaired, and Unit 2 was returned to service.</i>	Problem Identification and Resolution; ▼

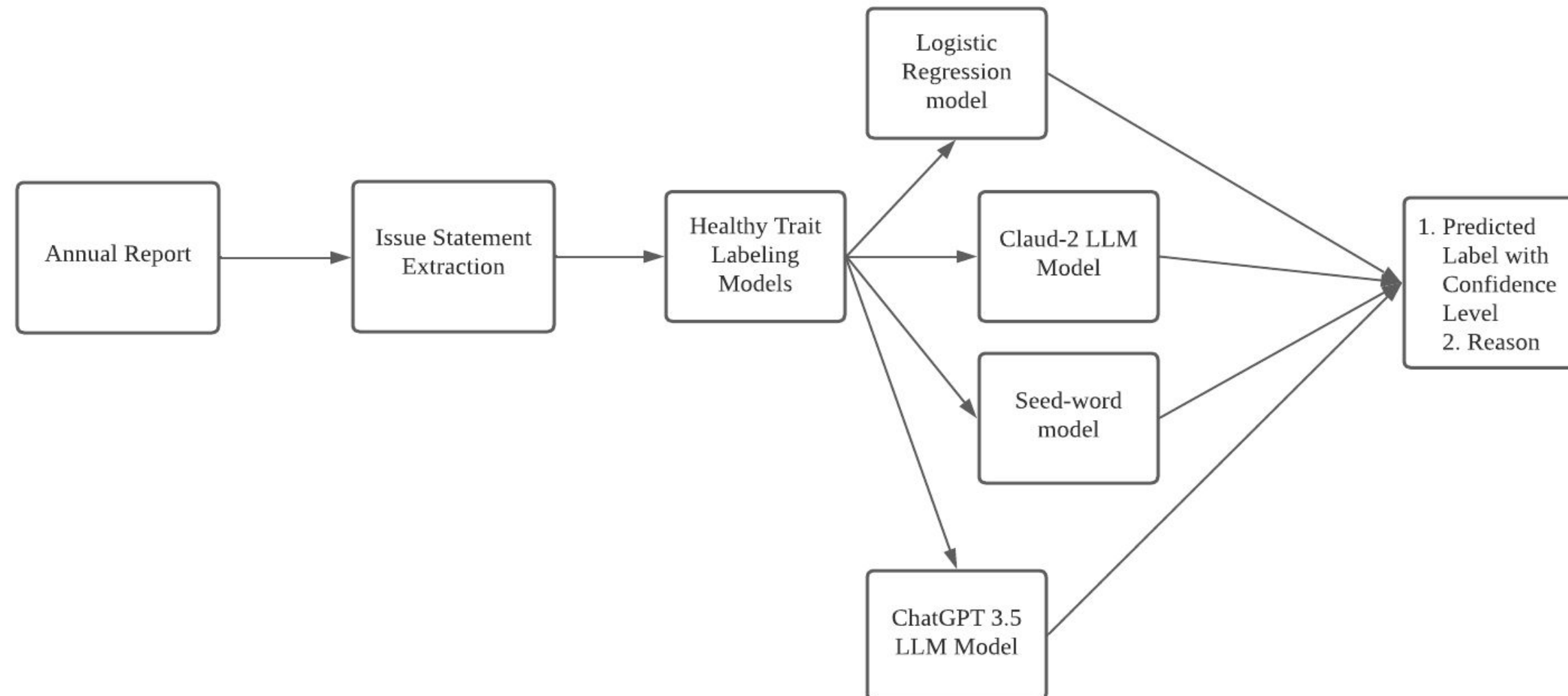
Approach

Step 1: Data Extraction and Pre-processing

Step 2: Baseline Model Selections

Step 3: LLM and Advanced Model Selections

Step 4: Dashboard Design



Approach – Seed Word Model

```
# Define primary and secondary seed words for each safety trait
seed_words = {
    "Personal Accountability": [
        "responsibility",
        "accountability",
        "help",
        "support",
        "trained",
        "qualified",
        "understand",
        "complete",
        "involvement",
    ],

```

```
secondary_seed_words = {
    "Personal Accountability_secondary": [
        "ownership",
        "dedication",
        "commitment",
        "reliability",
        "dependability",
        "self-discipline",
        "initiative",
        "reponsiveness",
        "proactive",
    ],

```

```
# Combine primary and secondary seed words
combined_seed_words = {**seed_words, **secondary_seed_words}
```

Approach – LLMs

GPT3.5 Model



You

Unit 2 LER 2021-002-00 was submitted on 12/14/2021, which documented that the Unit 2 reactor was manually tripped on October 15, 2021, in response to increasing water level in a secondary-side feedwater heater. Following shutdown of Unit 2, the feedwater heater was subsequently repaired, and Unit 2 was returned to service.

Feedwater heater forced outage

ODM Held on EI on Unit1 due to main turbine governor valve 1 (FCV-142) Servo

Positioner Controller (SPC1) Failing

====

what about these three



ChatGPT

Based on the provided information, the issue statements can be classified as follows:

For Unit 2 LER 2021-002-00:

1. **PI.1 Identification:** The issue involves the identification of increasing water levels in a secondary-side feedwater heater, leading to the manual trip of the Unit 2 reactor for corrective action.
2. **PI.3 Resolution:** The issue required the reactor to be manually tripped, followed by the repair of the feedwater heater, emphasizing the need for a timely resolution and corrective action process.

Claude-2 Model

Unit 2 LER 2021-002-00 was submitted on 12/14/2021, which documented that the Unit 2 reactor was manually tripped on October 15, 2021, in response to increasing water level in a secondary-side feedwater heater. Following shutdown of Unit 2, the feedwater heater was subsequently repaired, and Unit 2 was returned to service.

Feedwater heater forced outage

====

please do the traits classification based on the INPO file in the required table format.

Sources

Traits of a Healthy Safety Culture

File • 1

Answer

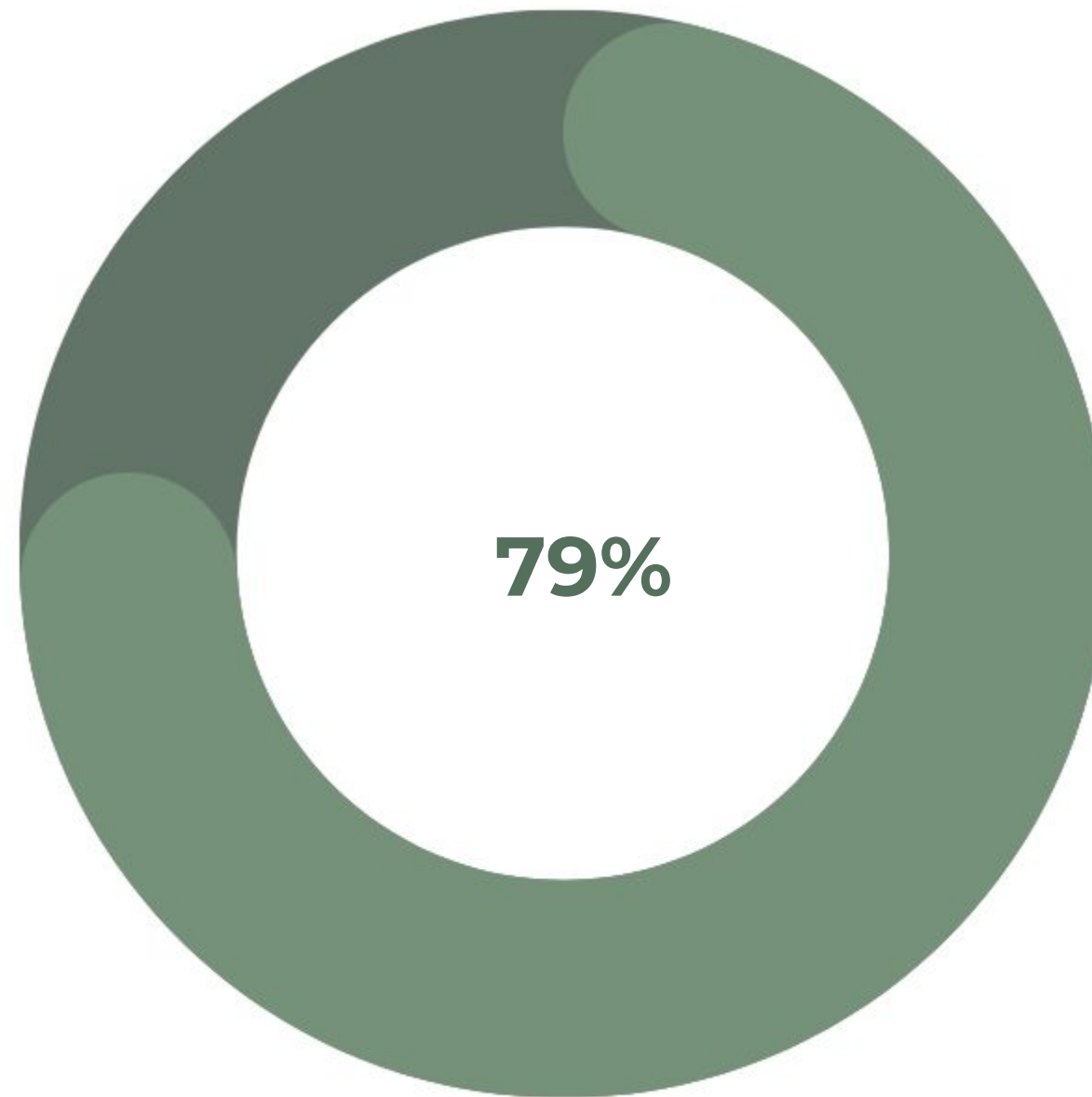
Based on the information provided and the INPO safety culture traits document ¹, here is my assessment of the relevant safety culture traits:

Trait	Confidence
Problem Identification and Resolution	High
Work Processes	High
Personal Accountability	Medium

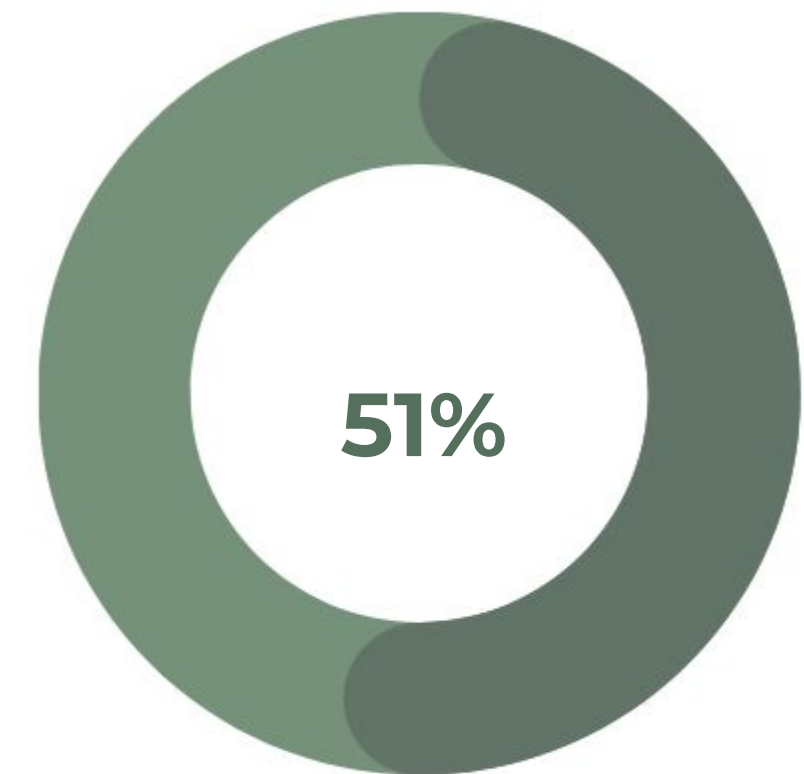
Model Performance: Accuracy



**Seed Word
Model**



**Claude-2 Large
Language Model**



**Logistic Regression
Model**

Discussion and Future Goals

- Test different large language models to improve our baseline model
- Expand to secondary and tertiary seed words for multi-class classification of issues
- Create a dashboard with integrated model
- Build model for issues statement extraction from raw report

Thank You!