

# Enabling Proactive Quality in Commercial Airplanes using Natural Language Processing



## BUSINESS PROBLEM

Quality management processes typically rely on structured, often numerical, data to measure performance and identify areas for improvement. Recording quality data in free-text forms is also common, but its unstructured format makes drawing insights challenging. Effective exploitation of the information contained in such data becomes a barrier, especially to proactive quality programs, that is hard to overcome without large amounts of manual effort.

## DATA SOURCES

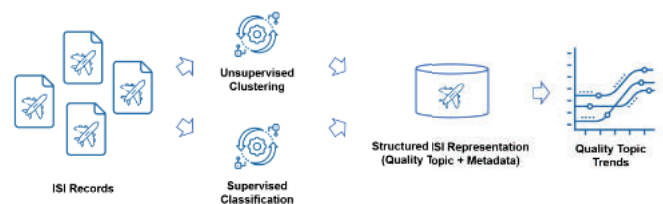
In-Service Investigation data, including completed investigations, associated quality records and metadata, and customer communications. All data is sourced from Boeing's manufacturing execution and quality data systems.

### Data Types and Format

Primarily un-structured text data (investigation reports, customer messages) and some structured metadata (summarization codes)

## APPROACH

Boeing's In-Service Investigation (ISI) process is used as a case study. These investigations share broad quality topics, but the free-text nature of the final reports makes it challenging to find these relationships. A combination of NLP and machine learning methods were investigated to create representations of these investigations that could be more easily combined for trend analysis.



## IMPACT

The ability to draw greater insights from in-service investigations depends on the ability to relate investigations by the type of quality issues they explore. As exactly-repeated issues are rare, this relies on being able to put investigations into groups of broader similarity, called "quality topics". This becomes challenging given the distributed nature of investigations - investigations for different airplane models are handled by different teams, and there is a significant history of historic investigations. To create these groupings, a supervised classification technique was successfully developed. This method uses an NLP transformation of investigation attributes (such as the title and summary metadata) to predict the investigation's quality topic; certain models showed strong performance at this task. Part of the in-service investigation process includes reviewing past investigations for instances of similar quality escapes. The acts of both collecting data and making comparisons is done by hand. An assistive "helper" tool was created to show the potential to automate this process: the tool automatically loads past investigations and uses NLP-based methods to estimate similarity between these documents. Such a tool, if built in production, would improve the speed and consistency of this time-intensive process.

### DRIVERS

Advancements in natural language processing (NLP) and machine learning and the availability of complete libraries reduce the burden for non-expert users to obtain value from these processes; such methods have powerful abilities to extract information from free-text documents.

### BARRIERS

Process improvements where historic data is important requires balance - decisions made on how to approach the problem needed to take into consideration not only future investigations but also how information from past investigations could be incorporated.

### ENABLERS

Boeing has active, high-profile efforts to improve the use of quality data, which include looking at more novel technologies to bring results.

### ACTIONS



This solution was developed iteratively with guidance by internal stakeholders. Major milestones included use case identification (ISI process), detailed planning (group investigations by quality topic), and trials on several types of machine learning methods. Each milestone was punctuated by demonstrations and feedback from teams who would be potential users of the results.

### INNOVATION

This work included developing several helper technologies: - Method to compare the preservation of NLP embedding's semantic similarity after being subject to various dimensionality reduction techniques - Rule-based algorithm to expand common short-hand part number descriptions into full form - Graphical investigation "helper" tool that performs live similarity comparisons on documents using NLP embedding representations

### IMPROVEMENT

This project focused on finding and demonstrating feasible solutions to improve the in-service investigation process, though these changes have not yet been implemented.

### BEST PRACTICES

For process improvement projects, best practice is to identify an ideal state but make progress towards it in small pieces. This way if a larger piece does not end up proving successful, there are still smaller improvements to fall back on.

### OTHER APPLICATIONS

The methods proposed to extract value from in-service investigations are not restricted to only this type of document or to quality management; they should work well with other use cases that seek to represent unstructured free-text data in a more structured form.