Despite advances in hardware, training, regulations, and operations, commercial aviation incidents attributable to human error remain consistent. A review of the literature was conducted to determine what factors are contributing to these anomalous incidents, as well as what research exists on human error, its causes, and its management.

Although aviation accident databases exist, only the NASA Aviation Safety Reporting System (ASRS) utilizes a completely voluntary reporting scheme for all incidents. This database is also unique that it provides a rich narrative text section, and therefore was chosen for this study. Reports from the ASRS database were analyzed with advance text mining tools and methods that afforded large scale processing of unstructured text.

Data was first analyzed using traditional statistical methods such as frequencies and multinomial logistic regression. Relevant and recent approaches in text mining such as Knowledge Based Discovery (KBD) and Literature Based Discovery (LBD) were employed to create associations between factors and anomalies, as well as generate predictive models. Finally, advances in dimensional reduction techniques identified concepts or keywords within records, thus creating a framework for an unsupervised document classification system.

Findings from this study reinforced traditional views on what factors contribute to civil aviation anomalies, however, new associations between previously unrelated factors and conditions were also found. Dimensionality reduction also demonstrated the possibility of identifying salient factors from unstructured text records.