Beyond Traditional SMS – Can Resilience Engineering and Deep Learning Neural Networks Be Used to Anticipate Disruptions in the NAS?

Richard Abbott
Objectstream, Inc.
7725 W Reno Ave. #307, Oklahoma City, OK 73127
richard.abbott@objectstream.com

ABSTRACT

Air Traffic Management organizations are part of a complex socio-technical system. In a tightly coupled, intractable system like ATM, it is not possible to identify many incident and accident causes or system risks by using the safety management methodologies in place today. This paper explores how using resilience engineering and deep learning neural networks together could help an air navigation service provider develop its resilient performance potential to recognize and manage unexpected situations when they present themselves.

1 INTRODUCTION

Commercial aviation is a very complex socio-technical system, yet it is defined as ultra-safe, where the risk of disaster is below one accident per one million operations, i.e. 10^{-6} (Amalberti, 2001, p. 111). This paper will explore how today's traditional approach to managing safety may not be enough in the complex, highly dynamic modern air transportation system.

Since the early 2000s, the concept of resilience engineering (RE) has been studied, researched, and applied as a complement to today's safety management systems. The premise of this concept is the notion that how an organization performs depends on its capability to perform and that its *actual* performance depends on potential performance. Resilience engineering allows an organization to develop and manage what Dr. Eric Hollnagel refers to as "resilience potentials" (Hollnagel, 2017, p. 23). While this paper is not a tutorial on resilience engineering or the tools available to assess resilience potentials, it will discuss how this concept could be used to supplement, or move beyond, a traditional Safety Management System (SMS).

Artificial intelligence – we've all heard about it. It's used in video games, pattern recognition to identify trends in social media posts and online buyers' purchasing habits, intelligence and cybersecurity, and many other applications. The aviation industry uses forms of artificial intelligence to develop models that optimize airline seat pricing; develop predictive analytics used to determine aircraft maintenance needs; collect and analyze aircraft performance data on specific flight routes to predict optimum fuel loading; and manage the complexity of airline crew scheduling. The Air Traffic Management industry, with its large amounts of unstructured heterogeneous data sets, audio, video, images, text, etc., lends itself well to artificial intelligence solutions. This paper will investigate the application of Deep Learning Artificial Neural Networks to enhance an air navigation service provider's resilience performance potentials.

2 RESILIENCE ENGINEERING – The Evolution of Safety Management

The FAA Air Traffic Organization's Safety Management Systems manual defines Safety Risk Management as follows:

> A process within the SMS composed of describing the system; identifying the hazards; and analyzing, assessing, and controlling risk. SRM includes processes to define strategies for monitoring the safety risk of the National Airspace System (NAS).

This is the "traditional" approach to managing safety risk. It has worked well for decades and continues to do so. The process typically flows as follows:

- Describe the system by decomposing it to understand the role of each system component.
- Identify as many system hazards as you can.
- If the risk any of those hazards pose is unacceptable, develop and implement a risk control (such as a rule, process, or other action to mitigate the risk), making sure prior to implementing the control that the control itself doesn't impose another risk (substitution risk).
- Use a safety assurance process to determine whether the controls are working as intended.

This process is based on a linear, barrier model approach such as depicted in James Reason's Swiss Cheese model. These models rely on what has been described as the *causality credo*, which states that every cause has an effect. The reverse of this idea (every effect has a cause) is prominent in most organizations' safety management systems. The thinking is, if accidents happen as a combination of active failures and latent conditions (as implied by the Swiss Cheese model), then risks must be the likelihood of weakened defenses in combination with active failures (Hollnagel, 2014).

Resilience engineering, as an approach to managing safety in complex systems, is said to have had its genesis in the early 2000s when safety scientists such as Eric Hollnagel and David Woods conducted studies and research. Currently there is an abundance of literature exploring resilience engineering in detail, hence this paper will not delve into the details. There is, however, one fundamental concept that Woods, Hollnagel, and their colleagues identified, and it serves at the core of RE: Safety II.

2.1 Holistic safety – aka Safety II

Before discussing Safety II (sometimes referred to as the New View of safety), a review of how most organizations think about safety (usually called Safety I) is important. Safety in high-risk industries today primarily focuses on maintaining a system where as few things as possible go wrong. When things do go wrong, it is assumed that they do so because of a system component failure – whether it be procedures, technology, workers, or the organization itself. In most cases, the human is viewed as the most likely culprit, as humans are the most variable component in the system. Safety management systems respond to this breakdown by attempting to eliminate causes or improve barriers. Procedures are revised, additional training is provided, or new rules are put in place to keep the same thing from happening again (Hollnagel, Wears, & Braithwaite, 2015).

This Safety I view came into practice in the 1960s when systems were thought to be less complex and could be completely understood (tractable), therefore easily decomposed. System components either worked or they didn't; in other words they were bimodal. These assumed characteristics led to improving safety by looking for causes and fixing them. Almost every safety management system today is built around these assumptions. As an example, the *FAA Air Traffic Organization 2015 Safety Report* states, "Collect, Find, Fix describes how we improve the safety of our airspace" (FAA, 2015). To summarize, the starting point for traditional safety management (Safety I) is that either something has gone wrong or a risk has been identified. In either case, the find and fix approach is used.

In the latter part of the 20th century, as safety professionals and the safety science research community studied the safety landscape in complex socio-technical industries such as nuclear power plant management, medicine, and aviation, it became apparent that while learning from accidents and identifying system risks were still viable and important endeavors, something more was needed. The core Safety I assumptions discussed above were no longer valid in complex systems such as air traffic management. These systems (and their components) were tightly coupled and could not be completely decomposed – they were intractable. They were also *not* bimodal – components didn't either succeed or fail. Yes, sometimes there were acceptable outcomes and sometimes the outcomes were unacceptable, but neither outcome could be attributed to components working or not working.

Two concepts will help illustrate why this is the case. The first is *work-as-imagined*. In the case of ATM, for example, facilities, technical requirements, and operational work requirements are designed and engineered by a diverse range of people, including engineers, technicians, technical writers, policy makers, and numerous others. Their thinking and subsequent work products are a result of how they "imagine" the system should work to produce the most efficient and safest operations in the NAS. While these teams will always include controllers – those who do the work – their work products will never be in complete alignment with how the work is done. This is not due to a lack of effort, but happens because the system is intractable (not completely understood) and works in different ways depending on the interactions of all the various components, i.e. complexity. There will always be a gap between *work-as-imagined* and *work-as-done*. The more intractable the system is, the larger the gap is likely to be.

Safety II recognizes this gap. Safety II (unlike Safety I) sees the human component as the reason almost every outcome is successful because of humans' ability to adapt their actions to the situation at hand. The question became; if successful outcomes occur so much more often than unsuccessful ones do, why do we measure safety by focusing on "the lack of safety," i.e. when things don't go well? Safety II advocates suggested that "we should avoid treating failures as unique, individual events, and rather see them as an expression of everyday performance variability" (Eurocontrol, 2013). The focus of safety management is moving away from preventing something from going wrong to instead trying to learn why so much goes right and how to support activities that contribute to those many successful outcomes. Managing safety with this in mind is the focus of resilience engineering.

2.2 Resilience engineering – the basics

Webster's Dictionary defines resilience as "the quality of being resilient"; that is, the ability to recover or bounce back. This definition is problematic from a safety management perspective – the problem being, how do you know if a person or an organization is resilient? How do you measure it (resilience)? In many ways, it's like safety culture – a monolithic concept that everyone seems to understand yet which cannot be measured or assessed in any scientific way. And as the saying goes, to manage something, you must be able to measure it. To do that, one cannot think of resilience as something an organization has but rather as a characteristic of how an organization performs – in other words, *what* it does and *how* it does it. The most recent definition of resilience in this context (it has evolved over the last 20 years) is as follows:

> Resilience is an expression of how people, alone or together, cope with everyday situations – large and small – by adjusting their performance to the condition. An organization's performance is resilient if it can function as required under expected and unexpected conditions alike (changes/disturbances/opportunities). (Hollnagel, 2017)

Resilience engineering, in the context of the above definition, looks at everything an organization does, not just safety. It looks at what goes wrong *and* what goes right. The objective is to focus on how an organization performs. To do that, RE focuses on work-as-done. Understanding how the work is done (versus how it is imagined) is critical to ensure that as many good outcomes as possible emerge from the organization's activities. Dr. Hollnagel and his colleagues in the resilience engineering community suggest that all organizations have a potential for resilient performance and have identified four domain-independent abilities (they refer to them as "potentials") that are necessary (for resilient performance) and can be assessed, monitored, developed, and continuously improved. The four potentials and brief definitions[1] are as follows (Hollnagel, 2017):

- The potential to **respond**. This potential refers to knowing what to do when regular and irregular changes, disturbances, and opportunities are recognized. Note that this isn't about responding just when things go wrong but also when opportunities to create new ways of doing things present themselves. Responding includes activating prepared actions such as, in the case of ATM, contingency plans.
- The potential to **monitor**. This potential is about knowing what to look for that could affect an organization's performance (positively or negatively) in the near term. Monitoring includes not just looking at the external environment but looking at the organization's internal performance as well.
- The potential to **learn**. This includes knowing what has happened and learning from experiences.
- The potential to **anticipate**. An organization that exhibits resilient performance can anticipate potential disruptions or new opportunities. This potential looks further into the future.

The most important thing to know about these four potentials is that they are dependent on each other. Most organizations look at these functions and say, "We do this already." And in most cases, especially in an air traffic management organization, they most emphatically do. However, organizations must ask themselves: Can we do more? What are we capable of doing? What is our *resilient potential*? Tools and methods (such as the Resilience Analysis Grid [RAG] and the Functional Resonance Analysis Method [FRAM]) are available to help answer these questions. Numerous research papers, books, tutorials, and training classes exist that delve much deeper into resilience engineering and the use of the tools and methods for assessing and developing resilient potentials. The author will leave it up to you, the reader, to search out what's most appropriate for your organization. But perhaps there's more that can be done. Can artificial intelligence help

[1] For a detailed explanation of resilient potentials, see Safety II in Practice (Hollnagel, 2017).

organizations learn more about themselves and the socio-technical systems they are a part of? The remainder of this paper will explore these questions.

3 ARTIFICIAL INTELLIGENCE

Today almost every industry is using artificial intelligence to help it operate more efficiently, increase revenue growth, and target the right customers for its products and services. Automobiles, homes, commercial aircraft, and millions of other physical devices can connect to the internet where they collect and share data. This concept, the Internet of Things, or IoT, generates vast amounts of data that can in turn be analyzed and help industries and companies make rapid improvements in their products. We are all familiar with intelligent voice-enabled assistants like Siri and Alexa. Mobile phones use facial recognition in place of passwords that allow users to access their phones as well as many of the applications on their phones. Most of these AI applications are based on a subset of artificial intelligence known as machine learning, which enables systems to predict outcomes by learning patterns in data. The learning evolves as the data changes without a need for rules-based programming. Arthur Samuel, a machine learning pioneer, defined machine learning in 1959 as "a field of study that gives computers the ability to learn without being explicitly programmed" (Bell, 2015). Deep learning is a subset of machine learning that makes use of multi-layered artificial neural networks that learn to recognize specific patterns in data through a training process (Rueckert, 2017). Deep learning requires significant processing power (now available through dedicated graphics processing units or GPUs) and excels at classification and prediction tasks. Using computers to understand and process human language, whether written or spoken, is very difficult. This is the focus of Natural Language Processing or NLP. Deep learning artificial neural networks have made significant advancements in the development and implementation of NLP solutions over the last few years. This technology could be helpful in the future to help Air Traffic Management organizations assess and develop their resilient performance potentials.

3.1 Artificial neural networks – an overview

Artificial neural networks (ANN) are sets of computer algorithms modeled loosely after the human brain that are designed to recognize patterns and make predictions. They are capable of processing large amounts of diverse data and grouping that data into specific categories. Neural networks can classify data when they have been "trained" with labeled data sets. During the training process, input data is compared to the labeled data and are processed through the neurons (nodes) of the ANN input layer using a variety of algorithms that are chosen based on the objective of the network. Errors that arise based on classifications that are different from the labeled data are fed back to the input layer, where small adjustments in how the nodes are weighted (which cause them to fire or not) are made. Through this process the weights are continuously adjusted until the incoming data is classified to a specified degree of accuracy. This type of learning (classifying data based on labeled data sets) is known as supervised learning.

Data can also be grouped based on the input data's similarities. This is known as clustering and is used when there are no labeled data sets to train on. This is the case when the input data is unstructured. Sound, text, images, and video are examples of unstructured data. Most of the data in the world is unstructured, i.e. not labeled. Learning about the similarities in unstructured data is called unsupervised learning. Different types of algorithms (than those used in supervised learning) are used to find clusters or subgroups in the input data set. The goal is to learn interesting things about the data that are useful and could help an organization meet its objectives.

There is a middle ground that makes use of both labeled and unlabeled data. It is referred to as semi-supervised learning. Labeled data can be difficult and expensive to obtain. Using unlabeled data is subjective and may require a significant effort to adjust the network's weightings so the groups of clusters make sense and are useful for the intended application. With semi-supervised learning, a small labeled data set can be used to train a larger unstructured data set. With this high-level understanding of artificial neural networks and the three common types of learning methodologies they use, we can explore deep learning and how deep learning artificial networks can be used to analyze human language in both spoken and written forms.

3.2 Deep learning and natural language processing

Deep learning neural networks are differentiated from the neural networks described above by the number of hidden layers between the input and output layer. If there is more than one hidden layer, the network is referred to as a "deep" neural network. In deep learning networks, each layer processes data using distinct features that are based on the previous layer's output. The more layers that are present, the more the nodes are able to recognize more complex features since they aggregate and recombine features from the previous layer (Skymind, 2015).

Natural Language Processing, or NLP, uses deep learning artificial neural networks to process, analyze, and understand human language. In the past few years, there have been substantial advancements in this subfield of artificial intelligence. Much of this advancement involves new techniques to process unlabeled data using unsupervised or semi-supervised learning. One such technique, or feature set, is word embeddings. Word embeddings attempt to capture similarity between words and are typically pre-trained and used in the first hidden layer of a deep learning ANN. Traditional word embeddings tried to consider all the sentences in a document or other source where a word is present and then create a global representation of that word. That didn't always work – for example, the word "bank." Is it a financial institution or land next to a body of water? This dilemma spawned the more recent development of "contextualized" word embeddings that produce word embeddings based on the context rather than using a global perspective. In short, artificial intelligence's rapid technological advancements are making significant strides in improving a computer's ability to process human language. The industry expects that the rapid change seen over the last few years will be a stepping stone from NLP to natural language understanding (Young, Hazarika, Poria, & Cambria, 2018).

3.3 The use of AI in air traffic management

Several research and technical papers have been published over the last few years that have recognized possible applications of AI in aviation systems such as air traffic management for both manned and unmanned aircraft. At the 2018 Civil Air Navigation Services Organization (CANSO) Global ATM Summit and 22nd Annual General Meeting, CANSO Director General Poole noted, "Digitisation … has the potential to transform Global ATM performance, bringing huge benefits … in terms of increased efficiency and *enhanced safety*" (CANSO, 2018). Similarly, a working paper published for use at the 2018 ICAO Thirteenth Air Navigation Conference recognized how AI can contribute to ATM, as evidenced by this statement:

> Faced with the challenges of growing traffic, resource demands, increasing uncertainties, and operational complexity, ATM can exploit the power of AI to

> empower current operators and boost productivity through the capability of decision making under uncertainties of optimized situational strategies that procedures or simple algorithms cannot provide. (ICAO, 2018)

Using Natural Language Processing deep learning solutions to classify and analyze aviation safety reports is the basis for a paper published in 2016. This paper notes how data contained in numerous aviation safety databases (such as the Aviation Safety Report System [ASRS] and the European Coordination Center for Accident and Incident Reporting Systems [ECCAIRS]) "is extremely valuable *for learning lessons* from past incidents and accidents" (Tanguy, Tulechki, Urieli, Hermann, & Raynal, 2016). Very important and meaningful revelations come from this paper. First, the authors noted that a human-based manual analysis of the diverse reports in many data sets requires numerous resources and is very complex. Their solution was to investigate how NLP tools based on supervised machine learning techniques could help to automatically classify the text-based data. Second, this work took place four to five years ago. As noted in section 3.2 of this paper, NLP processing tools have evolved substantially (mostly in 2018) since this work was done. While this work was based on machine learning; today's NLP initiatives focus on using deep learning artificial neural networks with algorithms that are much better suited to identifying patterns and data anomalies that could help ATM service providers provide more efficient, safer operations in the airspace they manage.

4 CONCLUSIONS

Safety in the aviation industry has always focused on identifying what can or does go wrong. This approach to managing safety is characterized in two important ways. First, incidents and accidents are investigated in order to identify the cause, whether it be an "active" failure or a "latent" condition. Considerable effort is usually undertaken to identify that single (usually human component failure) root cause most likely responsible. And second, to minimize the probability of incidents and accidents occurring, organizations implement their risk management process. This process looks at the entire system, attempts to identify any possible hazards, and then analyzes these hazards to determine risk. If the risk assessment results in a particular risk exceeding a predetermined threshold whereby it is deemed "unacceptable", risk controls are put in place to reduce the risk to an "acceptable" level. The safety assurance process then takes over and through audits and other monitoring processes, the system is continuously evaluated to verify whether the risk controls are working. If the number of incidents is reduced or audits don't raise any concerns, safety is assumed. Safety, then, is when nothing happens. It is measured by the lack of things going wrong. This approach to safety management has worked well and continues to do so, as the fact that commercial aviation is ultra-safe appears to imply.

However, over a period of several decades, the aviation system has become very complex. Because of the system's tight coupling and intractability, it is no longer possible to identify all the potential hazards and risks. Components (human and technical) of the larger socio-technical system can (and will) interact with the environment, each other, and daily operations in ways that cannot be identified in advance. In the late 1990s and early 2000s, safety professionals and safety scientists began to research ways in which organizations could be better prepared for the sometimes unexpected and unplanned-for events that occurred. The outcome of this research was the concept of safety-based resilient engineering. The research concluded that since organizations cannot identify all the hazards and risks in their system, then they must increase their ability to manage unexpected situations. It made more sense, to this new resilient engineering community, to look at safety in a more holistic way – not from the concept of being unsafe when things go wrong, but from the concept of being safe when so many more things go right – and focus on learning why things go right most of the time and ensure the organization does more of those things.

The reason things go right most of the time is because of the human's ability to adapt to a situation. This normal human behavior of adjusting one's performance to meet the demands of the situation became the foundation of resilience engineering. It seemed to the RE community that developing an organization's ability to adapt (increasing its resilient performance) would increase its ability to better function under both expected and unexpected conditions. Over the last two decades, tools and methods have been developed to assess, develop, and manage an organization's resilient performance potentials. These potentials, *respond*, *monitor*, *learn*, and *anticipate*, are dependent on each other and are key to an organization's ability to be prepared for unexpected developments it might encounter.

While tools (such as the Resilience Analysis Grid and FRAM – both referenced in this paper) exist, the information used by an organization's safety department (or consultants) to populate and use these tools is mostly a manual process. There are many databases that contain performance and safety information of interest, including the Bureau of Transportation Statistics, the FAA's operational contingency plan database, the Air Traffic Safety Action Program, the Technical Operations Safety Action Program, and the Aviation Safety Report System.

This paper posits that the organization could learn more about the environment they are part of (the ATM community, for example) by collecting data from the abovementioned databases and others and processing this diverse structured and unstructured data through a deep learning artificial neural network using natural language processing algorithms. The ability of an NLP neural network to identify patterns quickly, automatically, and with high levels of correlation to specific safety and performance categories and clusters could contribute to an organization's ability to increase its resilient performance and therefore be better prepared to react and perform when an unexpected situation presents itself.

Richard Abbott is Director of Safety for Objectstream, Inc. He serves as a member of the ANSI accredited U.S. Unmanned Aircraft Systems Technical Advisory Group to ISO responsible for developing world-wide UAS standards. In that role, he is currently working with the UAS Traffic Management (UTM) workgroup tasked to develop international UTM guidelines and standards. He is a long-time member of the International System Safety Society and the International Society of Air Safety Investigators. He is a former FAA Aviation Safety Inspector and a former Program Manager for the FAA System Approach to Safety Oversight program office. Since retiring from the FAA, he has provided safety management consulting and training services to airlines, maintenance repair organizations, and civil aviation authorities world-wide. The views expressed in this paper are the author's and do not reflect official positions of Objectstream, Inc. or its clients.

5 REFERENCES

Amalberti, R. (2001). The paradoxes of almost totally safe transportation systems. *Safety Science, 37*(2–3), 109–126. doi:http://dx.doi.org/10.1016/S0925-7535(00)00045-X.

Bell, J. (2015). *Machine Learning: Hands-On for Developers and Technical Professionals*. Indianapolis, IN: John Wiley and Sons.

CANSO (2018). *CANSO Global ATM Summit and 22nd AGM: Air Traffic Management in the Age of Digitisation and Data*. Retrieved from https://www.canso.org/canso-global-atm-summit-and-22nd-agm-air-trafficmanagement-age-digitisation-and-data.

Eurocontrol (2013). *From safety-I to safety-II*: *A white paper*, DNM Safety, Brussels. Retrieved from www.eurocontrol.int/sites/default/files/content/documents/nm/safety/safety_whitepaper_sept_2013-web.pdf.

FAA. (2015). *Air Traffic Organization 2015 Safety Report*. Retrieved from www.faa.gov/about/office_org/headquarters_offices/ato/service_units/safety/.

Hollnagel, E. (2014) *Safety-I and safety-II: the past and the future of safety management*. Burlington, VT: Ashgate Publishing Company.

Hollnagel, E., Wears, R., Braithwaite, J. (2015). *From safety-I to safety-II: A white paper*. Published simultaneously by the University of Southern Denmark, University of Florida, and Macquarie University, Australia. The Resilient Health Care Net. Retrieved from https://www.england.nhs.uk/signuptosafety/wp-content/uploads/sites/16/2015/10/safety-1-safety-2-whte-papr.pdf.

Hollnagel, E. (2017). *Safety-II in Practice: Developing the Resilience Potentials*. New York: Routledge.

ICAO (2018). Thirteenth Air Navigation Conference, Working Paper 232. Retrieved from https://www.icao.int/Meetings/anconf13/Pages/default.aspx.

Rueckert, Marco (2017). *Artificial intelligence in ATM*. Paper presented at 2017 ATCA Technical Symposium, Atlantic City, NJ.

Skymind (2015). A beginner's guide to neural networks and deep learning. Retrieved from https://skymind.ai/wiki/neural-network.

Tanguy, L., Tulechki, N., Urieli, A., Hermann, E., Raynal, C. (2016). Natural Language processing for aviation safety reports: from classification to interactive analysis. *Computers in Industry*, Elsevier, 2016, 78, pp. 80–95. doi:http://dx.doi.org/10.1016/j.compind.2015.09.005.

Young, C., Hazarika, D., Poria, S., Cambria, E. (2018). Recent trends in deep learning based natural language processing v8. Retrieved from https://arxiv.org/abs/1708.02709.