Natural Language Processing and Data-Driven Methods for Aviation Safety and Resilience:

from Extant Knowledge to Potential Precursors

08.12.2021

Presenter: Irene Buselli

Irene Buselli, Luca Oneto, Carlo Dambra,
Christian Verdonk Gallego, Miguel García Martínez,
Anthony Smoker, Nnenna Ike,
Patricia Ruiz Martino,
Tamara Pejovic

FARO

saFety And Resilience guidelines for aviatiOn

The FARO project

FARO is a *SESAR Exploratory Research 4*-funded project.

The project mainly addresses the evaluation of the impact of automation solutions on resilient performance and safety. In particular, in this work we focused on a particular manifestation of ATM safety, the loss of separation (LoS).

CRIDA I+D+i - ATM

ENAIRE

EUROCONTROL

LUND UNIVERSITY

POLITÉCNICA

UNIVERSITY OF BELGRADE FACULTY OF TRANSPORT AND TRAFFIC ENGINEERING

ARTIFICIAL INTELLIGENCE & DATA ANALYTICS [ZenaByte]

Scope of the work

A threefold approach:

1. An Exploratory Data Analysis (EDA) on **safety reports,** combining state-of-the-art techniques like **topic modelling** and **clustering.**
2. An algorithm able to extract **TOKAI taxonomy factors** from the free text of the reports, based on **syntactic analysis**.
3. A **Machine Learning** model able to **automatically assess contribution** (ATCo/Pilots/both) few minutes after the LoS.

NLP and statistics on the reports

Data-driven model to be used **before** investigation

The data

A) 89 CEANITA **LoS reports**, written in Spanish and published by Spanish Safety Aviation Agency (AESA), covering safety-related occurrences that happened in the Spanish airspace between **January 2018 and July 2019**.
These incidents reported by CEANITA are just a subset of the total amount of losses of separation, where high-severity incidents are over-represented.

B) Contextual information arranged in structured form, provided by ENAIRE-CRIDA. In particular, high-granularity **ATM data** such as flight tracks and ATM-processed information about the Spanish airspace. More precisely:

- **flight tracks** and related contextual flight information (e.g., type, speed, and heading);
- **ATC events** of the interactions between ATCos and the Controller Working Position (CWP).

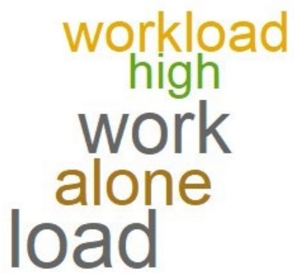
Topic Modelling

Topic modelling is a Natural Language Processing technique able to **extract recurrent topics** from a collection of documents in a completely **unsupervised** way.

015/18 – 03.03.2018
"In a situation of high workload, ACC Canarias Sector F03 authorised Aircraft1 to descend at 6000 ft."

027/18 – 06.04.2018
"Indeed he had a high work load due to a change in configuration"

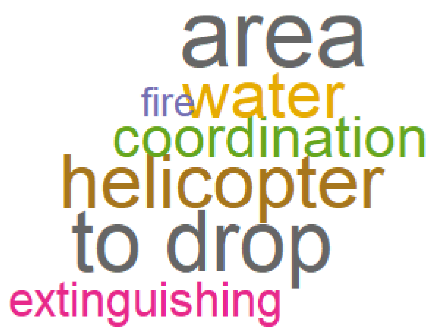
021/18 – 18.03.2018
"APP GCTS during the incident faced a high workload due to the number of communications.



110/18 - 09.08.2018
"Aircraft 2, after dropping water over the hot spot, turned suddenly"

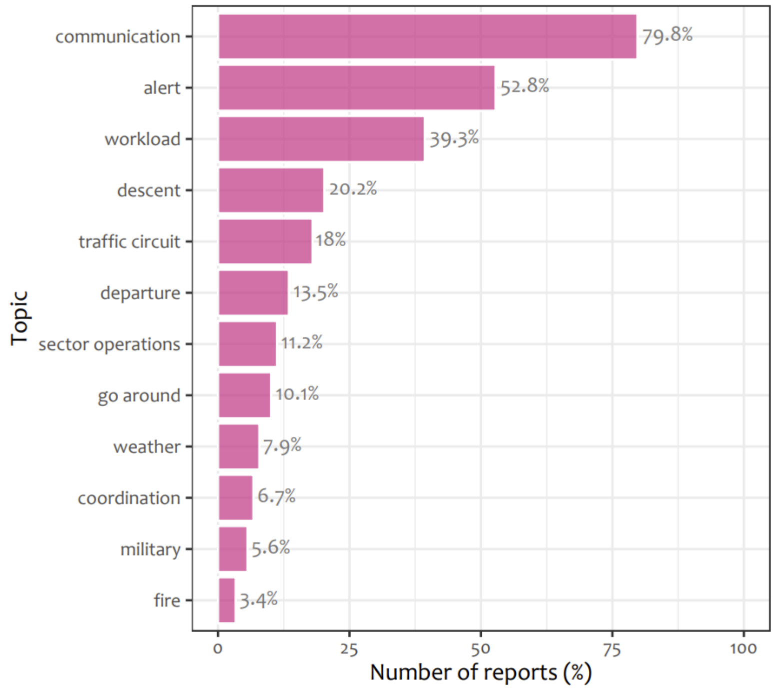
112/18 - 08.08.2018
"The incident happened during the extinguishing of the fire in Llutxent"

111/18 - 06.08.2018
"[…] the helicopter notified his presence there to scoop up water."



Words/Bigrams							Topic
helicopter	drop	water	fires	extinguishing	coordination	drop area	fire
load	work	high	alone	workload	instructions	previous	workload
departure	to take off	aircraft climb	runway	to take off aircraft	rate	they are	departure
wind	tail	down-wind	leg	wind leg	right tail	runway	traffic circuit
weather	adverse	adverse weather	detours	meteorologic conditions	due to weather	thunderstorm	weather
runway	go around	go	around	to take off	to land	aircraft established	go around
sectors	sector aircraft	frequency sector	high	coordination	transfer	limit	sector operations
answer	received	finally	decided	they saw	communication	visual contact	communication
clearance	course descent	aircraft to descend	descent rate	sector to descend	aircraft to maintain	rate	descent
received	coordinating	confirming	to confirm receipt	maintaining formation	sector informs	receipt	coordination
alert	early	early alert	activation function	activation	function	alert function	alert
military	military formation	formation	military aircraft	defence	air defence	main centre	military

Each topic is defined as a **list of words** with different weights (in the table, ordered from the most important to the least). Topics can then be **labelled** and their **prevalence in the reports** can be computed through a probabilistic method.

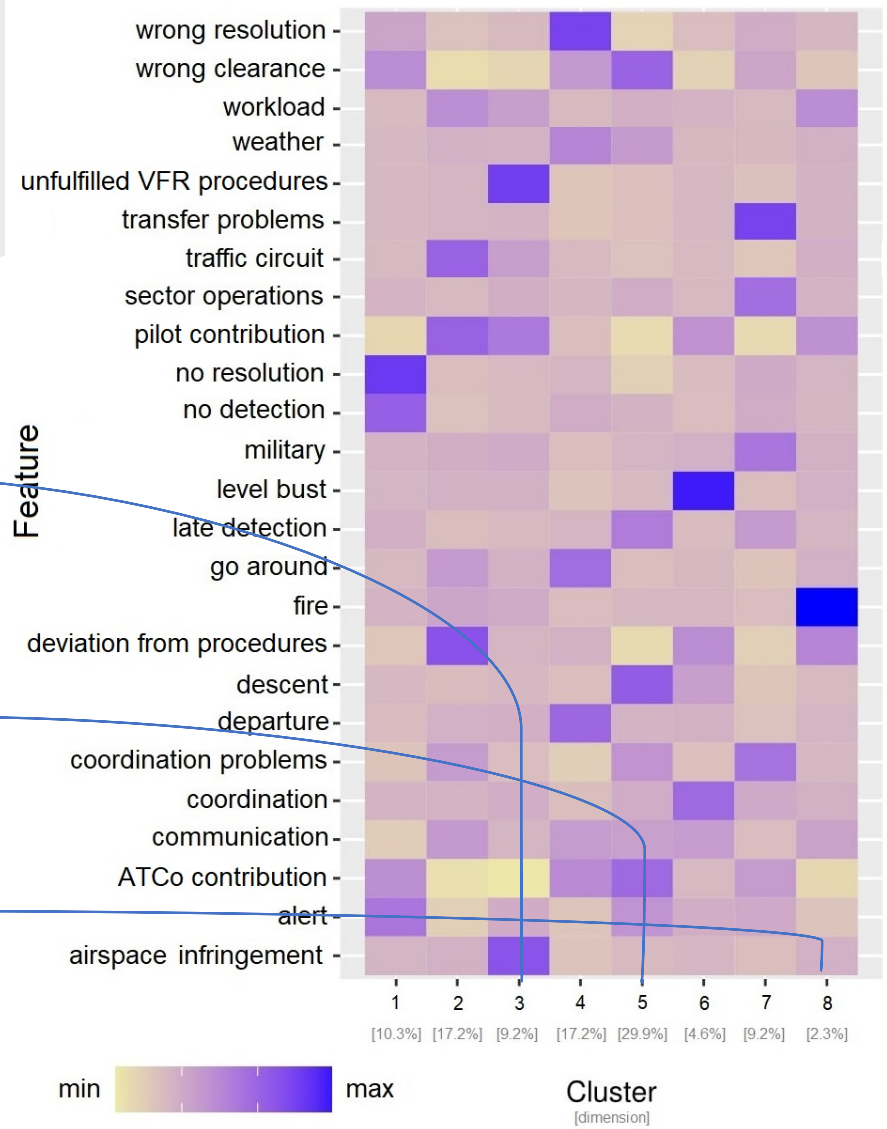


Clustering

Incidents in Cluster 3 appears to be essentially due to **Pilots' errors**, in particular to **airspace infringement** and unfulfillment of the **Visual Flight Rules** (VFR)

The largest cluster (Cluster 5) seems mainly composed of **wrong-clearance and late-detection** incidents, with clearly the highest frequency of **ATCo contribution** and an interesting high prevalence of **"descent" topic**.

Cluster 8 is composed of three incidents where the **main topic is "fire"** (indeed, they are the reports referred to Llutxent fire in summer 2018)



The presence of these topics in each report can be **linked with the other available information** (e.g. main cause and contribution).
In particular, some **clusters** can be identified through unsupervised statistical techniques.

Syntactic-Analysis Algorithm

"Sector CAO authorised Aricraft 2 to descend at FL 340 from FL 380, without detecting that Aircraft 1 was flying at FL 370 […]
When Aircraft 2 started descending, was authorised by Sector CAO at FL 350 without remembering the presence of Aircraft 1.
[…]
Aircraft 1 did not inform ATC of this maneuver.
[…]
Sector CAO did not provide traffic information to Aircraft 2.
On the other hand, Sector ESS provided traffic information to Aircraft 1".

ALGORITHM

Taxonomy factor	Positive/ Negative	Subject
A-1. Perception	Negative	Sector CAO
A-2. Memory	Negative	Sector CAO
A-4. Action	Negative	Aircraft 1
A-4. Action	Negative	Sector CAO
A-4. Action	Positive	Sector ESS

Every CEANITA report contains a free-text description of the main actions performed by ATCo and Pilots. The aim of our algorithm is to **automatically** extract these actions from the free text and to classify them according to the **TOKAI taxonomy** (Part A, so far).

How the algorithm works

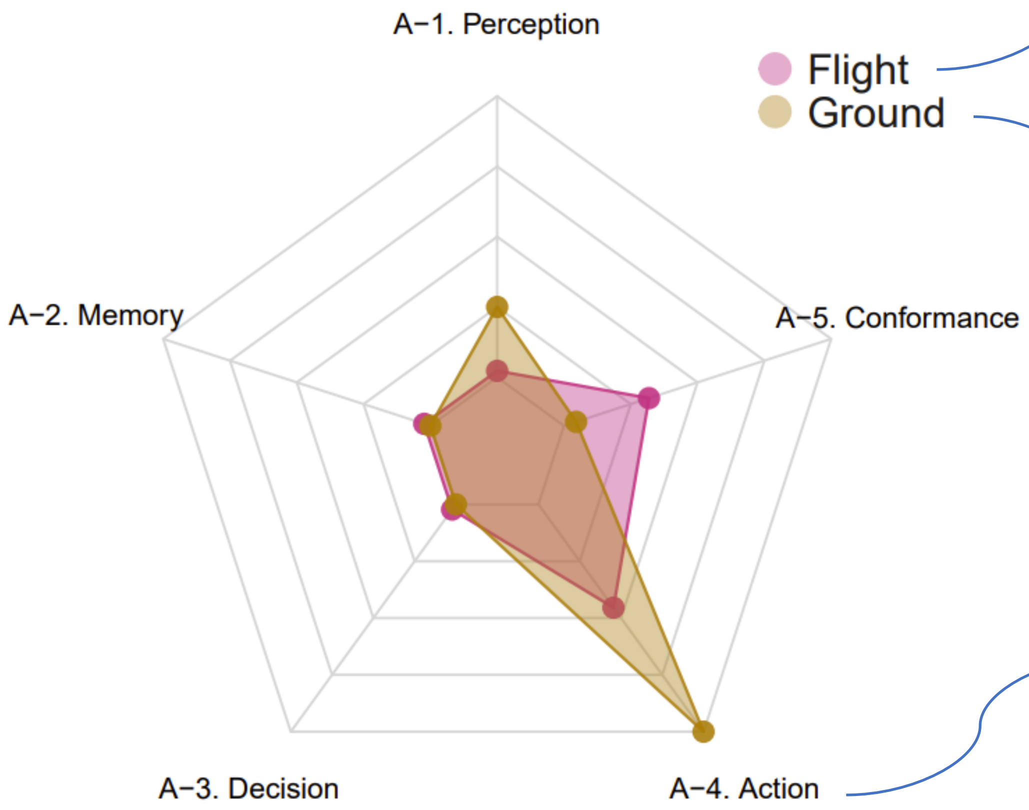
1. It takes the report as input and performs a **syntactic analysis** for every sentence:

	Sentence	Lemma	Part of Speech	Dependency
Sector	Sector	sector	noun	nsubj
SAU	SAU	SAU	propn	appos
instructed	instruye	instruir	verb	root
	a	a	adp	case
	la	el	det	det
Aircraft	aeronave	aeronave	noun	obj
2	2	2	num	nummod
to	a	a	adp	mark
proceed	proceder	proceder	verb	advcl
direct	directo	directo	adj	advmod:lmod
to	a	a	adp	case
	el	el	det	det
point	punto	punto	noun	obl
LOTEE	LOTEE	LOTEE	propn	appos

2. It takes as input the **lists of words defining taxonomy factors** (e.g., for A-1. Perception: “see", "detect", "identify", "recognise“, etc.), which should be provided by the user, and checks if there are corresponding lemmas in the decomposed text of the report.

3. It checks whether the verb of interest is in **negative form** (e.g. «without» + verb, or «not» between subject and verb, etc.) or not, if it is passive form (e.g., «Aircraft was not detected by the Sector», where it recognizes that the active subject is the Sector), and it retrieves the **subject**.

The graph shows the proportion of **negative occurrences** of the different factors in the **entire collection** of reports.



when the subject is Pilot, Aircraft, etc.

when the subject is Controller, Sector, etc.

Interestingly, the majority of these Action occurrences are relative to (lack of) conveyance of information.

Indirect validation

In order to validate our algorithm (since **no ground truth** was available) an **indirect validation** was performed.

A simple model was developed to predict whether the main contribution to the incident was from the ATCo or the Pilots or both, based on:

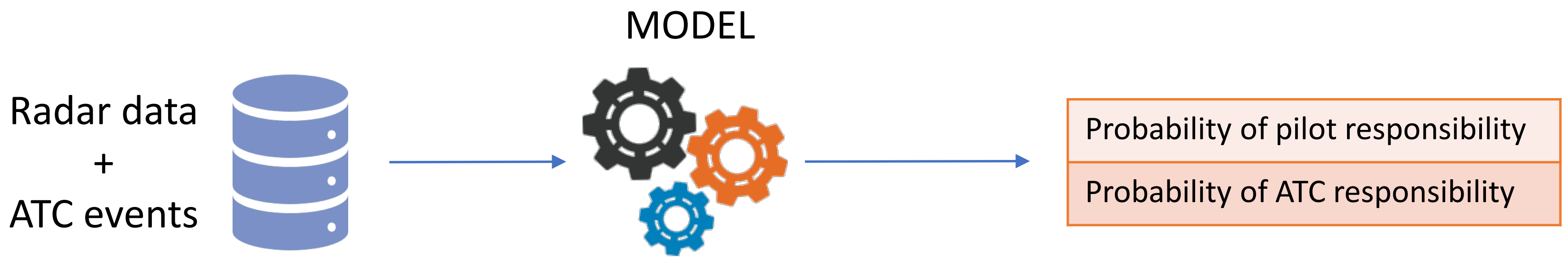
- the number of **positive and negative occurrences** of each taxonomy factor;
- the differences in prevalence between **Flight and Ground elements** for each taxonomy factor;
- the **airspace class**.

Since the factors resulted highly predictive (about 85% of accuracy), we can assess the extracted information is quite reliable.

Contribution assessment model

This model aims at **estimating contribution assessment** (ATCo/Pilot/both) based on the **ATC events** registered in temporal proximity to the LoS and **contextual information**.

This could facilitate the safety practitioners in **prioritising** the investigations and in understanding potential precursors of these LoS events, since the model also returns which events were the most relevant in its contribution assessment.



This model aims at **estimating contribution assessment** (ATCo/Pilot/both) based on the **ATC events** registered in temporal proximity to the LoS and contextual information.

This could facilitate the safety practitioners in **prioritising** the investigations and in understanding **potential precursors** of these LoS events, since the model also returns which events were the most relevant in its contribution assessment.

(A) PILOTS CONTRIBUTION (CONFIDENCE ≥60%)

		Pred. No	Pred. Yes
Truth	No	58.1±0.3	9.3±0.3
	Yes	4.7±0.3	27.9±0.3

(B) ATCO CONTRIBUTION (CONFIDENCE ≥60%)

		Pred. No	Pred. Yes
Truth	No	20.4±0.3	6.1±0.3
	Yes	8.2±0.3	65.3±0.3

(C) PILOTS CONTRIBUTION (CONFIDENCE ≥75%)

		Pred. No	Pred. Yes
Truth	No	60.0±0.0	00.0±0.0
	Yes	3.3±0.1	36.7±0.1

(D) ATCO CONTRIBUTION (CONFIDENCE ≥75%)

		Pred. No	Pred. Yes
Truth	No	29.0±0.1	3.2±0.1
	Yes	3.2±0.2	64.6±0.2

Variable	Importance
Flight type	14.71
Flight rule	11.38
ETO over FIX	9.56
Flight level	8.65
Airspace class	8.60
Radar Contact	7.76
Assume Communications	6.95
Action on Flight Level	5.98
CFL Reached	4.53
Flight Level Validation	3.68
Local Modification of FP	2.65
Modification of Manual Message on Label	1.50
Request to Abandon Hold	0.89
Aircraft Transference Notice	0.70
Authorise Level	0.42
Cancel Notice of Transference	0.37
Undo Authorisation	0
Radar Restriction on Heading or Speed	0
Free Frequency	0

Results are promising and the model presents much room for improvement.

Conclusions

Results are promising, even if some limitations are evident and should be addressed in future developments.

Future work could validate these techniques on **other databases of reports** (e.g., UKAB AirProx Board) and, moreover, these techniques could be **tailored to identify factors to be included in safety taxonomies** or hidden sources of resilient performance (e.g., when not fulfilling a procedure was opportune), based on their presence on the reports.

Finally, integrating **other sources of structured data** (e.g., about weather phenomena, STCA or TCAS activation, or traffic load) to develop richer models could lead to further insights in the estimation of contributors and precursors.

Thank you
for your attention