Improving Air Traffic Management operations with machine learning collaboration on private data sets: discussion of use cases of interest for the ATM stakeholders

(The SESAR AICHAIN Solution)

AICHAIN team: Javier Busto (SITA FOR AIRCRAFT), Sergio Ruiz (EUROCONTROL), Andre Rungger (SWISS), Oliva Garcia-Cantu (Nommon), Salman Toor (Scaleout)

Project site: www.aichain-h2020.eu

Coordinator contact: Javier.busto@sita.aero

Project partners: SITA FOR AIRCRAFT, EUROCONTROL, NOMMON, SCALEOUT, SWISS

SESAR 2020 Exploratory Research project addressing call topic SESAR-ER4-2019 - Digital Information Management (DIM).
Full project title: ***A platform for privacy-preserving Federated Machine Learning using Blockchain to enable Operational Improvements in ATM***

This project has received funding from the SESAR Joint Undertaking (JU) under grant agreement No 894162. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the SESAR JU members other than the Union.

Purpose of the meeting

1. To present a solution that enables **privacy-preserving machine learning collaboration on private data sets to enable operational improvements in ATM** (based on the results from SESAR ER4 project AICHAIN).
2. To **identify use cases of your interest** that could be enabled with AICHAIN (the solution can be applied in many use cases)

The AICHAIN Solution

AICHAIN Solution context and motivation

There is a need to enable **access to private data** to enable higher levels of **automation** and **performance** in ATM

AICHAIN

Privacy-preserving machine learning collaboration on private data sets to enable operational improvements in ATM

Digitalisation *in Aviation & Air Traffic Management (ATM)*

Enables

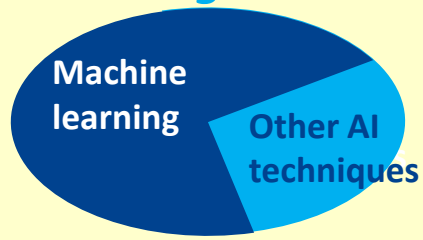
More Operational Data

Sharable Data

Private Data (non-sharable)

Enables

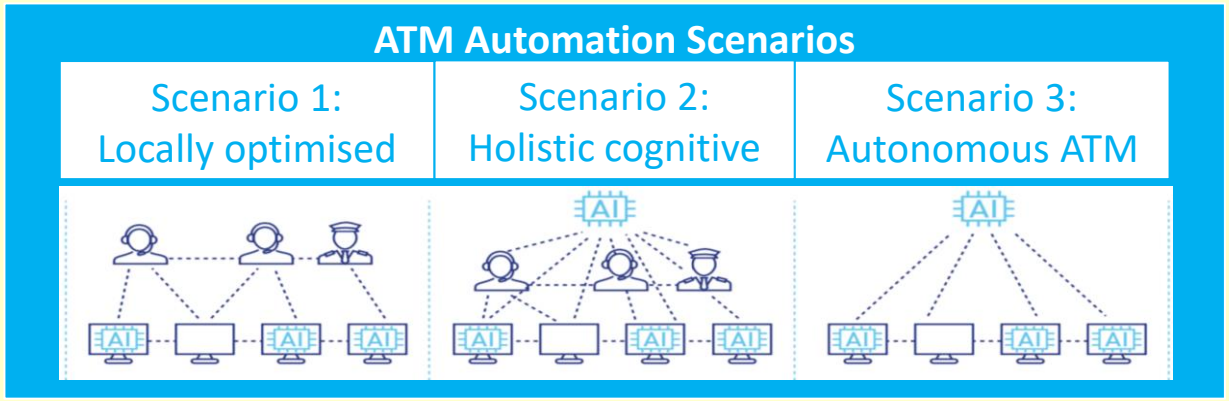
Artificial Intelligence *in Aviation*



Machine learning *will play a key role to foster the needed levels of automation*

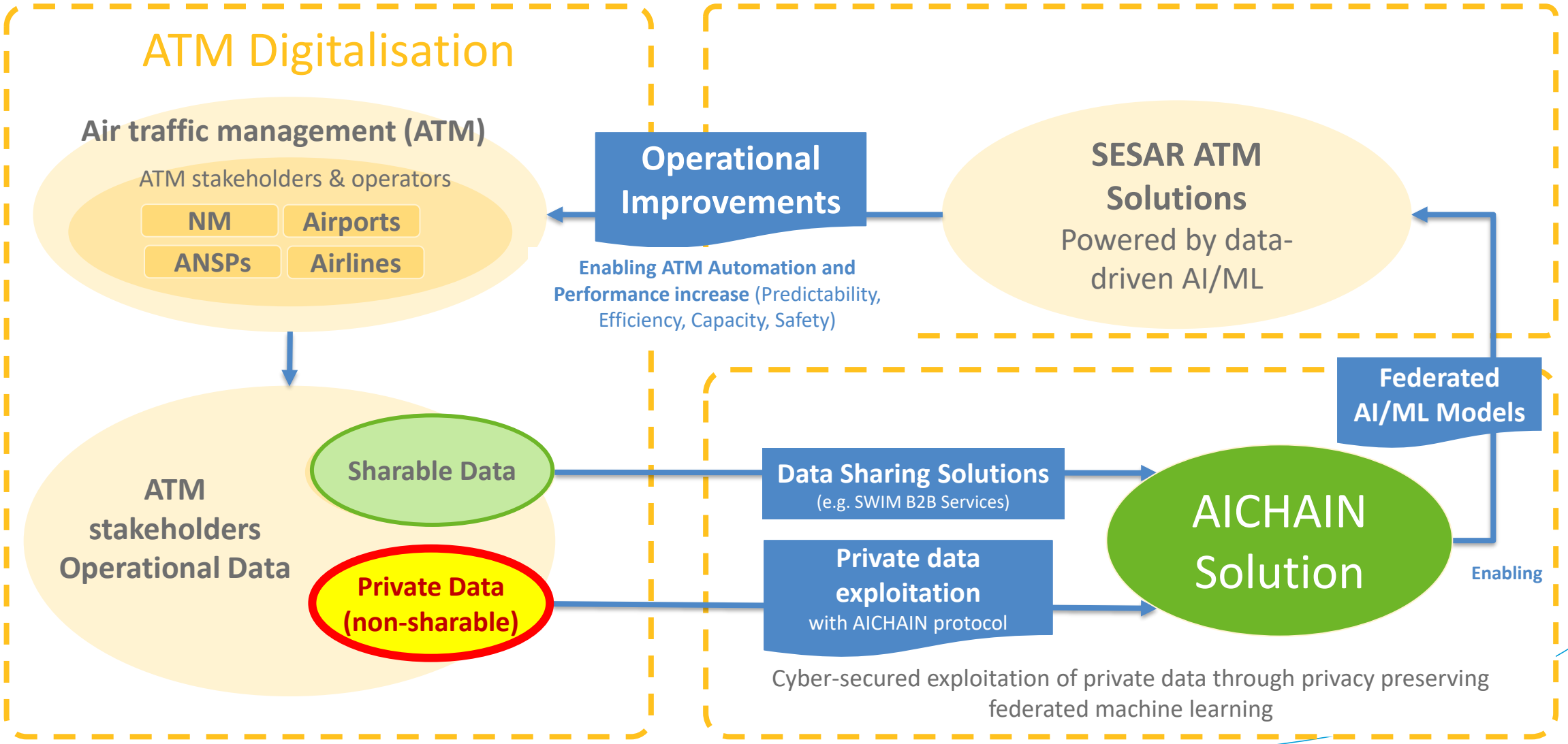
Enables

Automation *in future ATM: three foreseen scenarios*



ATM Operational Improvements

The AICHAIN Solution as a new SESAR technology enabler for ATM operational improvements



The AICHAIN Solution:

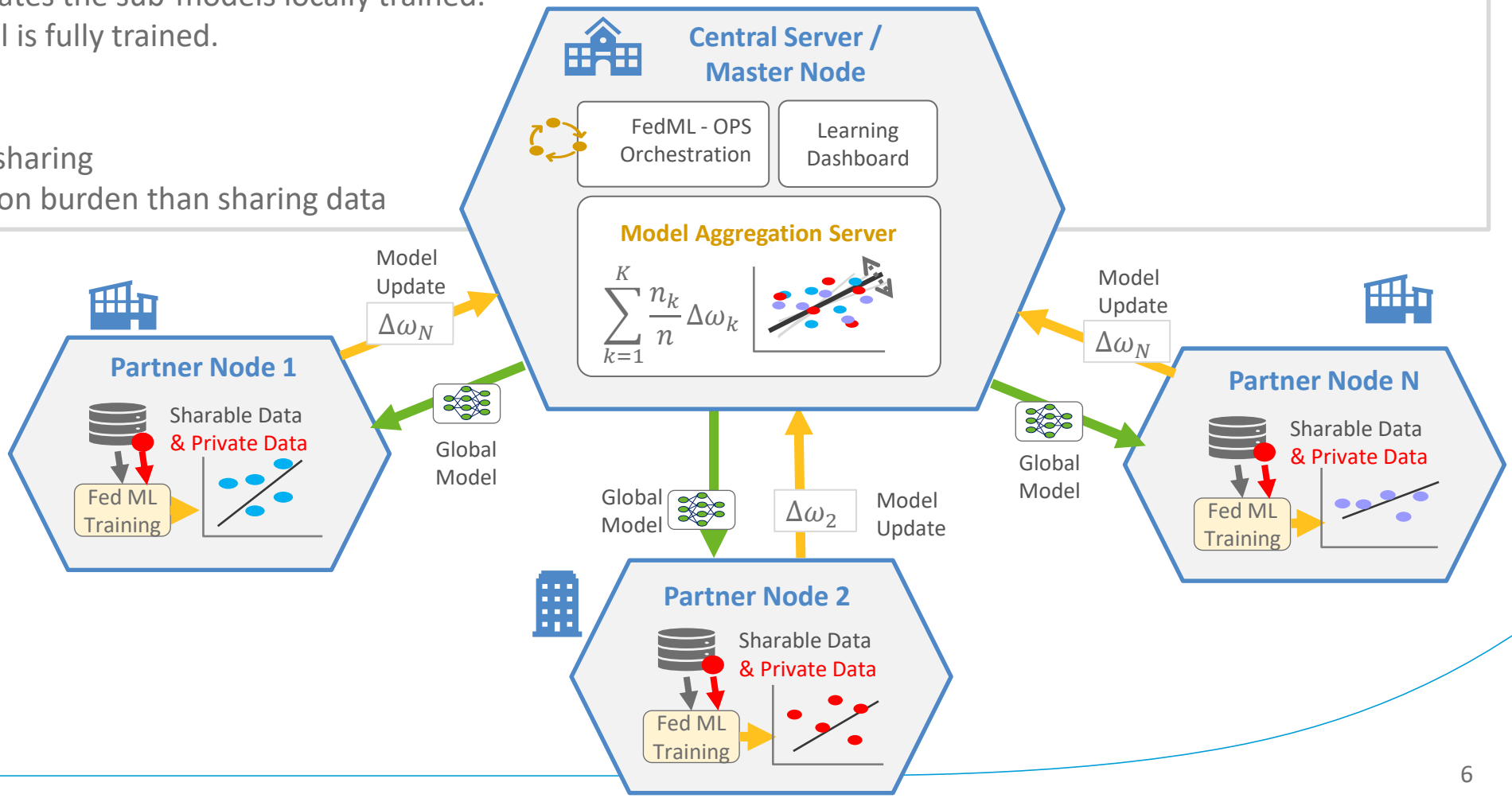
The Federated Learning concept

Collaborative training of ML models on private data sets:

- Each node train a single shared model locally with private data.
- The server aggregates the sub-models locally trained.
- Iterate until model is fully trained.

Key benefits:

- No need for data sharing
- Less communication burden than sharing data



The AICHAIN Solution:

Federated Learning enhanced with Blockchain

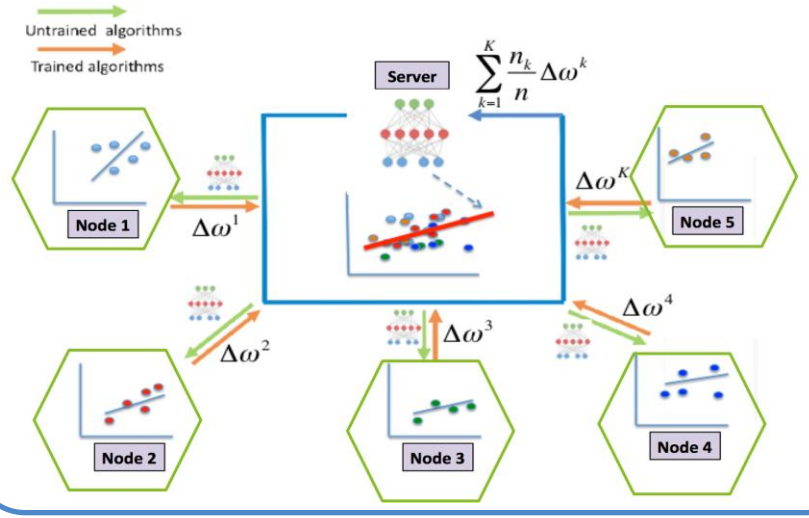
A Platform for Privacy-preserving Federated Machine Learning enhanced with Blockchain to enable Operational Improvements in ATM

AICHAIN Solution

=

Federated Machine Learning

(Non-shareable data exploitation for AI with privacy-preserving ML techniques that do not involve data sharing)



+

Blockchain-based Governance & Incentives

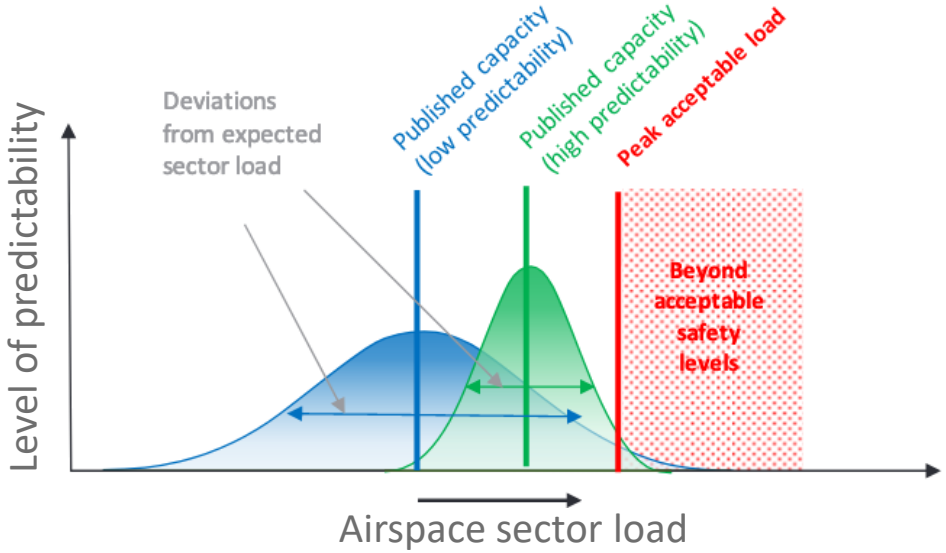
(with distributed ledgers, smart contracts and tokens)

Experimental results from two ATM use cases

Experimental results from two use cases

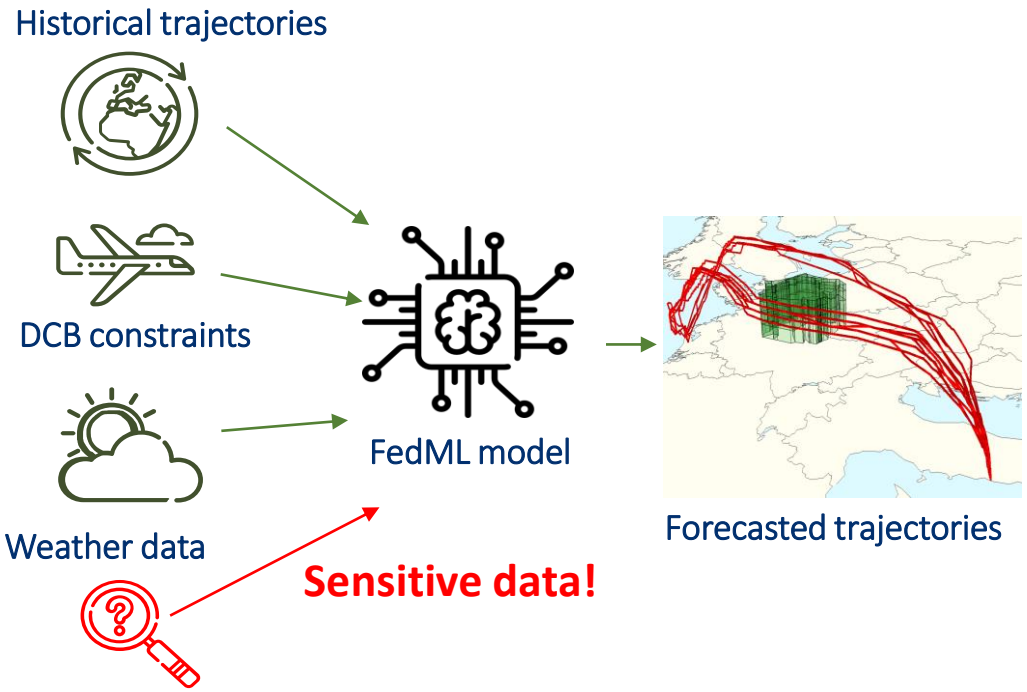
Two baseline models have been enhanced with private data

UC1: Improving Take-off time prediction



Uncertainty reduction in departure time → Reduction of capacity buffers

UC2: Improving 2D Route-selection prediction



Throughput at sectors can be increased if traffic is more predictable

Benefits: less congestion and less delay costs

UC1: Improving Take-off time prediction

Features:

Simplified version of the model at MUAC

NM features

EVENT
EVENTCLASS
FLTSTATE
ADEP
ADES
ARCTYP
IRULES
RDYSTATE
TAXITIME
AOARCID
FLTTYP
DEPAPTYPE
CDMSTATUS
AOOPR
ATFMDELAY
IFPSDISCREPANCY_REG
IFPSDISCREPANCY_ARCTYP
IFPSDISCREPANCY_OBT
DEPSTATUS
ADESOLD
RWY
FLIGHT_DURATION
EOBT_IFP_TO_EOBT
ADEPETO_IFP_TO_ADEPETO
IOBT_TO_EOBT
TIMESTAMP_IFP_TO_TIMESTAMP
TIMESTAMP_TO_ADEPETO
TIMESTAMP_TO_EOBT
TIMESTAMP_TO_TSAT
TIMESTAMP_TO_TOBT
TIME_FROM_REG_CHANGE
TURNAROUND_LEG
FLIGHT_DURATION_LEG
TIMESTAMP_TO_TA_LEG
ADEPETO_IFP_TO_ADEPETO_LEG
EOBT_IFP_TO_EOBT_LEG
TIMESTAMP_LEG_TO_TIMESTAMP
EVENT_LEG
AOOPR_LEG
FLTSTATE_LEG
ADEP_LEG
AOARCID_LEG
FLTTYP_LEG
HOUR
MONTH
DAY

SWISS features

- PAX_BOARDING_STATUS
- DEPARTURE_GATE
- SWISS_EXIT
- SWISS_EXOT
- SWISS_TURNAROUND_LEG
- DEPARTURE_GATE_ASSIGNED
- CREWHADTAILCHANGEPREVIOUS
- CREWCONNECTIONTIMEPREVIOUSFLIGHTSCHEDULED
- CREWCONNECTIONTIMEPREVIOUSFLIGHTACTUAL
- OCCUPATION
- PREDINPAX_GX
- SCDOUTPAX_GX
- SCDINPAX_GX
- SWISS_RWYNUM
- SWISS_RWYSPEC
- NUMPAXBOOKED
- NUMPAXFLOWN
- AIRCRAFTCAPACITY
- SWISS_TIMESTAMP_TO_TIMESTAMP
- SWISS_EOBT_TO_EOBT
- SWISS_ETOT_TO_ETOT
- SWISS_EOBT_TO_SOBT
- SWISS_TIMESTAMP_TO_ALL_DOORS_CLOSED

SWISS airline kindly provided a dataset with the private features needed for a subset of their flights

- **Objective of the regression model**: to optimise the ETOT calculated by the current system without ML models (the ML model computes the difference between the current ETOT and the actual take-off time)

UC2: Improving 2D Route-selection prediction Features:

Private features available in the experiments

Private feature approximated

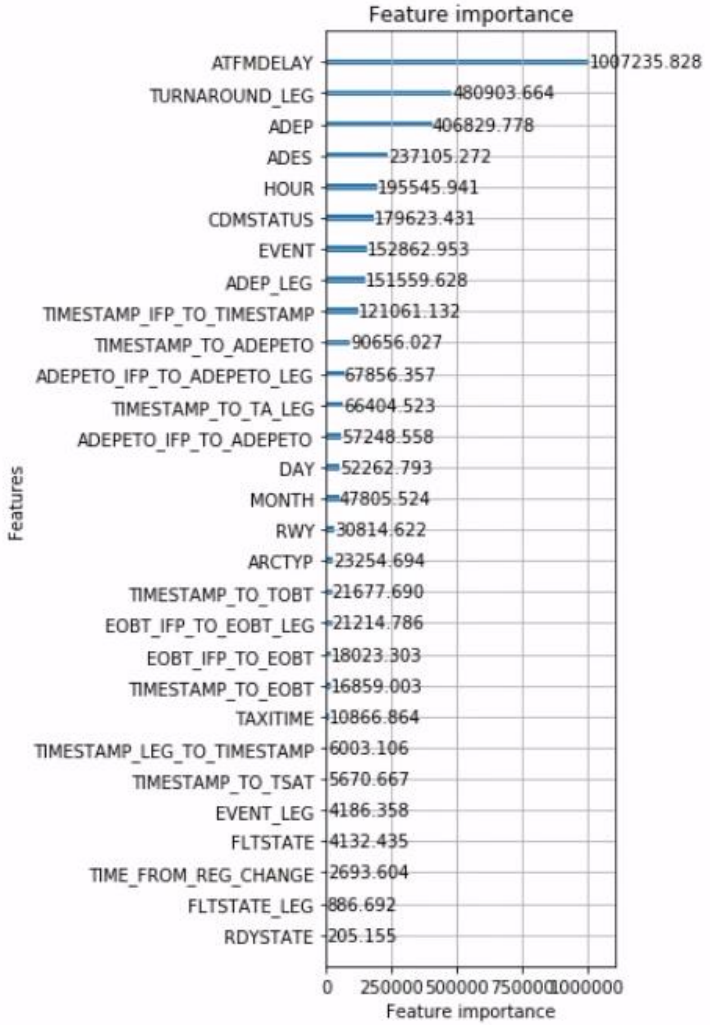
Variable Name	Description
Airline TOW	**Measured Take-Off weight by the airline of each flight**
Connecting passengers	**Number of passengers that have a flight connection in the destination airport**
DoW	The day of week of the flight codified accordingly
Flight Time	The ETOT hour of the flight
DoY	The day of year the flight takes place in
Longitude diff	Geodesic longitudinal separation between origin and destination
Latitude diff	Geodesic latitudinal separation between origin and destination
Airport population	Population density of the Origin/Destination surroundings areas
Airport GDP	Gross Domestic product of the Origin/Destination surroundings areas
Daily flights	Number of flights for each od pair and day
Airline market share	Airline's flight share for each od pair and day

Variable Name	Description
Route length	The length in kilometres of a given route
Wind length	Length of the route in kilometres adjusting the effect of the along wind
Charges	The charges paid for the current route for a given aircraft
Fuel cost	**Estimation of the cost of fuel for each given route**
Direct costs	Sum of the fuel and charges costs
CAPE	Used as a storm proxy
K-index	Weather metric that approximates the probability of a thunderstorm to happen
Humidity	The relative humidity observed along the route, that is a requisite for thunderstorms to occur
Local wind at origin/ destination	Variable that measures how aligned and in what value local wind at the airport is
Military zones	The route crosses a closed military zone, not use as a feature but to discard routes
Regulations	The duration of the regulation affecting the route

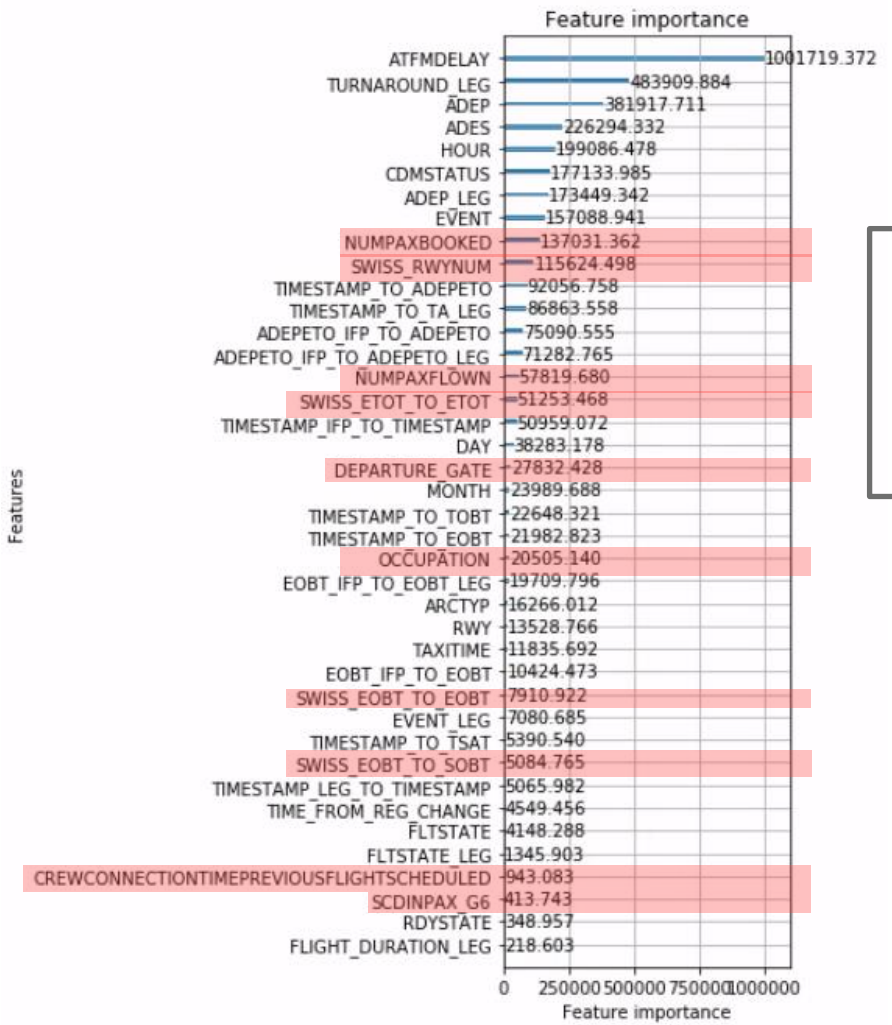
UC1: feature importance analysis

Non-augmented vs augmented models

Non-augmented model (NM data)



Augmented model (NM+SWISS data)

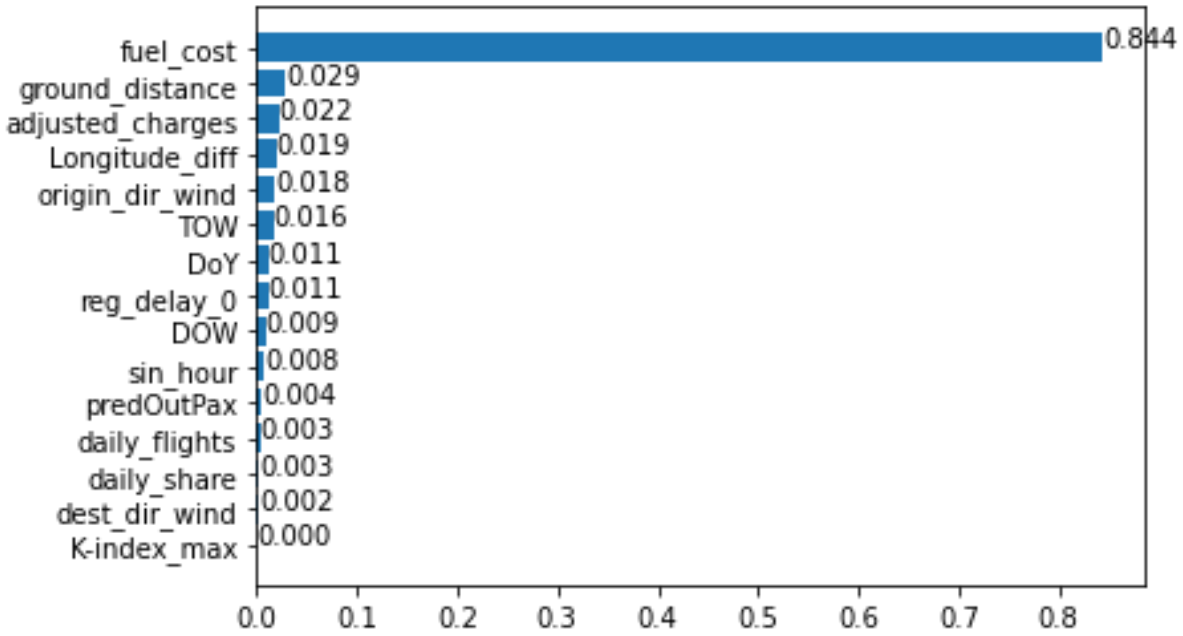


Private features (in red) do improve the model performance

UC2: feature importance analysis

Augmented model (NM+SWISS data)

Feature importance



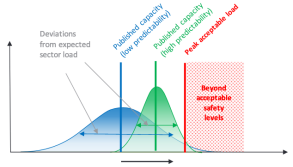
The most important feature in this use case is the fuel cost, which is considered private data.

Some models' performance can be highly dependent on private features

Note: in our experiments the fuel cost was approximated because the actual fuel cost was considered too private by the airline owner! It is expected that using the actual cost through federation the model performance could improve significantly, due to the importance of such feature in this model.

Experimental results from the two use cases
Model performance improvements observed:

UC1: Improving Take-off time prediction



Absolute error of the predictions on the test set

	ETFMS (legacy)	NM data (baseline)	NM+SWISS (solution)
Mean	10,5	9,3	9,0

+11.4% (ETFMS → NM data)

+14.3% (ETFMS → NM+SWISS)

Relative improvement: 14.3/11.4 = +25%

UC2: Improving 2D Route-selection prediction

Accuracy of the predictions on the test set

	Most flown (non-ML)	NM data (baseline)	NM+SWISS (solution)
Mean	0,87	0.95	0.954

+9.2% (Most flown → NM data)

+9.6% (Most flown → NM+SWISS)

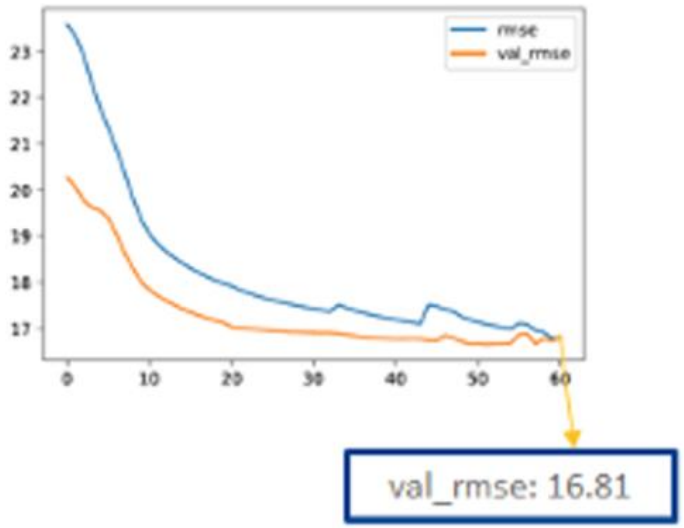
Relative improvement: 9.6/9.2 = +4%

Note: there is still room for improvement, either fine-tuning the model and/or adding more data

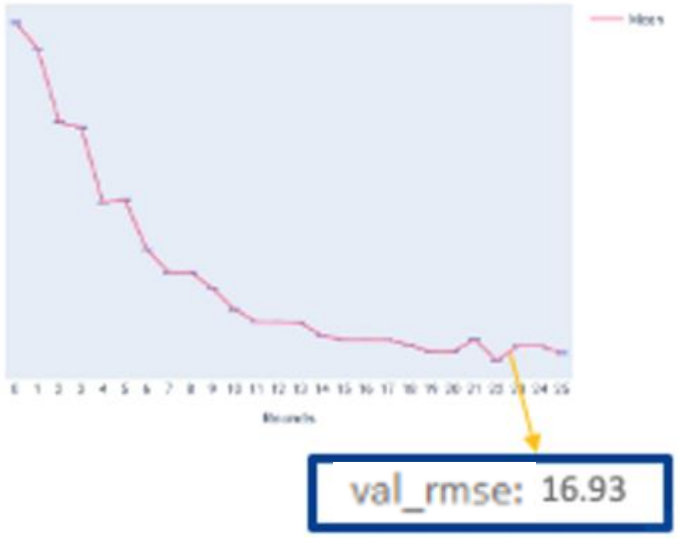
Federated machine learning exploits the value of private datasets

Non-federated vs federated models

Augmented model (V1)



Federated augmented model (V2)



Both augmented models (non-federated and federated) yielded the same model performance (see root mean square error metric below).

Conclusion: federated learning can exploit all the value from the private datasets while privacy is preserved (the federated augmented model could be built without sharing the private data).

Discussion of use cases of your interest

which could be enabled with AICHAIN Solution

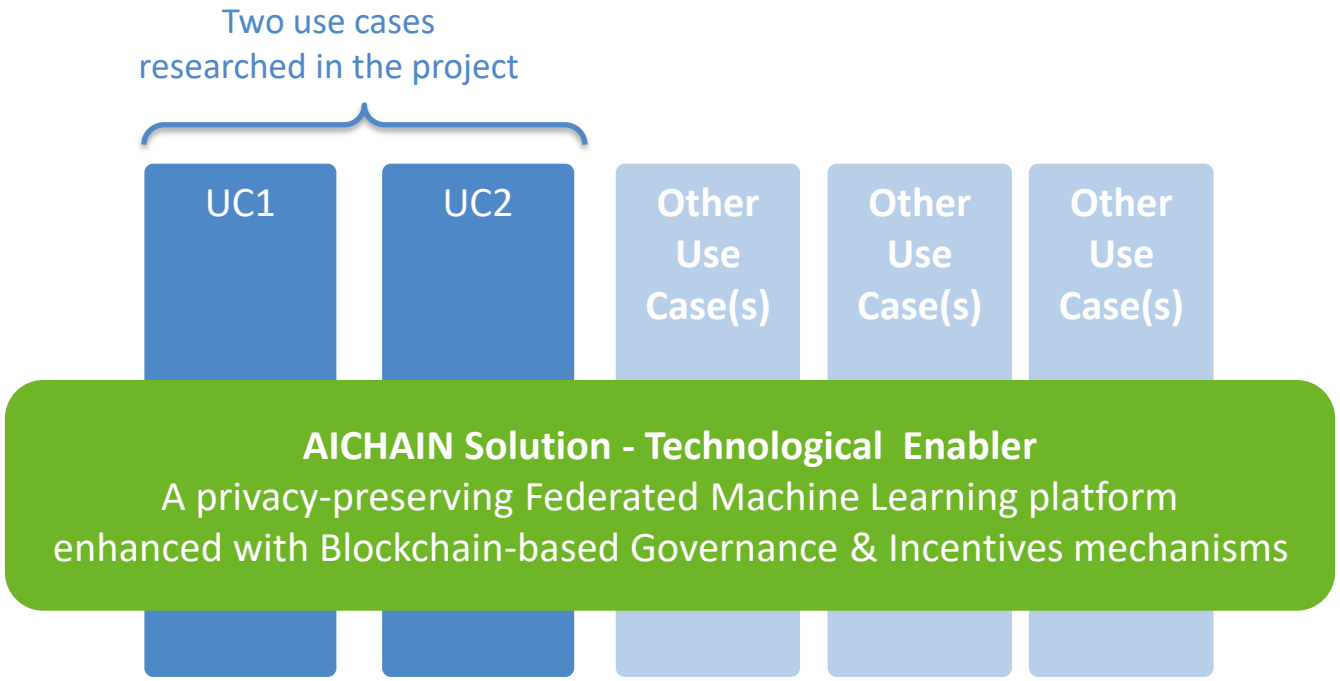
(it can be applied in many use cases)

AICHAIN solution: many potential use cases

The AICHAIN Solution is proposed as a new SESAR Solution of type technology enabler.

It can be applied transversally, i.e. in many use cases

1. **VERTICAL dimension** : Multiple applications and use cases
2. **HORIZONTAL dimension:** New SESAR Technology Enabler Proposal



Explored in the project:

- **UC1: Improving Estimated Take off Time**
- **UC2: Improving ATM AU´s 2D Route prediction**

Other use-cases possible:

- **Curfew management**
- **Flight efficiency indicators for ATC and ATFM**
- **Inter modality – Cargo-Drone Hub operation improvements**
- **Inter modality – End-to-end passenger journey operational improvements**
- **Other**

Is there any use case of your interest?
Let's talk about that!

The AICHAIN Solution:

Improving Air Traffic Management with machine learning collaboration on private data sets

Project site: www.aichain-h2020.eu

Coordinator contact: Javier.busto@sita.aero

Project partners: SITA FOR AIRCRAFT, EUROCONTROL, NOMMON, SCALEOUT, SWISS

SESAR 2020 Exploratory Research project addressing call topic SESAR-ER4-2019 - Digital Information Management (DIM)

Full project title: ***A platform for privacy-preserving Federated Machine Learning using Blockchain to enable Operational Improvements in ATM***

This project has received funding from the SESAR Joint Undertaking (JU) under grant agreement No 894162. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the SESAR JU members other than the Union.