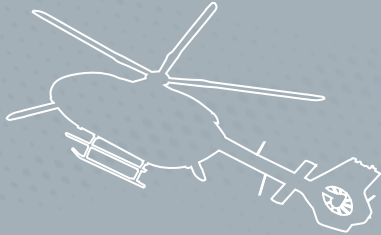# AI Hackathon - Kick-off

HELICOPTERS

January 2024, 10
Digital Acceleration Office
Sebastien Delecraz & Lwena Even

AIRBUS

Agenda

- About us
- Problematic
- Data
- Deliverables
- Evaluation
- Jury
- Awards ceremony

AIRBUS

About us

Helicopters

Design and development

Industry and production

Support and services

Digital
Acceleration **O**ffice

DAO contributes to ensure that Airbus Helicopters take full benefit of the 4th industrial revolution thus providing best products & services to our Customers.

How?
- ❖ **Accelerate** AH digital transformation
- ❖ **Support** the company objectives through digital
- ❖ **Create value** out of data
- ❖ **Prepare** the future

AIRBUS

Problematic

Context

Our department, the Digital Acceleration Office, supports the digital transformation of Airbus Helicopters' Legal department in order to improve:

❖ Information retrieval processes

❖ Automatic creation of legal documents

Task

❖ Natural Language Processing (NLP)

❖ **Input**: sections of legal texts (eg. general terms and conditions of sale)

❖ **Expected output**: summary of input text

AIRBUS

Data

Two datasets available

❖ Opensource dataset
   446 text sections with summaries from ~175 legal documents.

❖ Internal dataset
   413 text sections with summaries produced by Airbus Helicopters' legal department from two different original documents.

Language of texts and summaries: **english**

Example

Raw text: The Contract shall be binding when signed by duly authorised representatives of both Parties and the delivery schedule shall become effective upon receipt by the Seller of the initial down-payment and subject to compliance by the Customer to article 4.1. Helicopter serial number mentioned in the Contract (if any), shall be purely indicative, and shall not create any obligation upon the Seller.

Summary: The delivery schedule shall become effective upon receipt of the first down-payment. The helicopter serial number indicated in the contract is purely indicative.

AIRBUS

Deliverables

Evaluation summaries

Summaries produced by the model from the test corpus which will be sent to you at the end of competition day 6.

+ Source code for reproducibility

Solution presentation (slides)

- General one pager
- Data used
- Data preparation
- Model/algorithm description
- Training description
- Split train/validation

Information on resources used

- Resources used (training + test)
- Number of parameters/equiv.
- Average execution time for inference on test corpus
- Specification of machine used for inference

AIRBUS

Evaluation

Originality/relevance of approach:
Consultation with the jury based on the presentation of the solution.

Frugality of the model:
Evaluation based on resource usage information.

Expert assessment :
On a few predefined summaries from the test corpus (not known to the participants), quality assessment of the summaries by our legal expert (Airbus Helicopters).
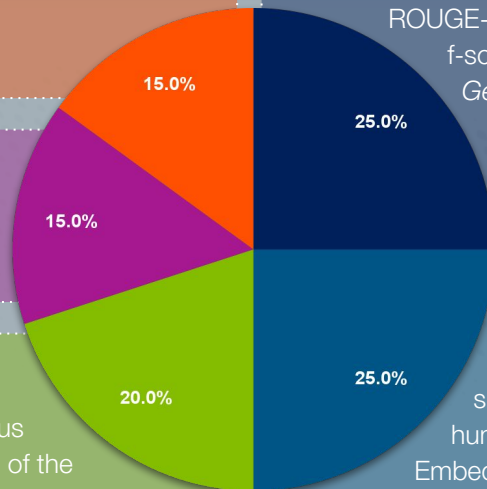
ROUGE metric:
Evaluation of the summaries from the test corpus with the metric ROUGE-1, ROUGE-2, and ROUGE-L (recall, precision and f-score).
General principle: evaluate the overlap between n-grams (sequences of n words) in the automatic summary and those in the human reference.

Similarity score:
Evaluation of the similarity between automatic summaries (1) and the original test corpus (2), as well as human references (3).
Embedding of texts, then calculation of cosine similarity between (1)-(2) and (1)-(3).

15.0%
25.0%
15.0%
20.0%
25.0%

AIRBUS

Jury

Centrale Méditerranée
Ronan Sicre

Centrale Digital Lab
Anne-Laure Mealier

AMU (MAS)
Grégory Maillard

DAO
Sébastien Delecraz

DAO
Lwena Even

Legal
Stéphanie Alex

AIRBUS

Awards ceremony

When & where ?

Between April 4th and 18th

19:00 pm - 22:00 pm

Marseille

What for ?

Celebrate your work

Connect and network

Discover some of our innovative activities

https://forms.gle/CybJAsJVFMdKBdkm9

AIRBUS

Thank you

AIRBUS