

Predictive analysis for aviation systems configuration

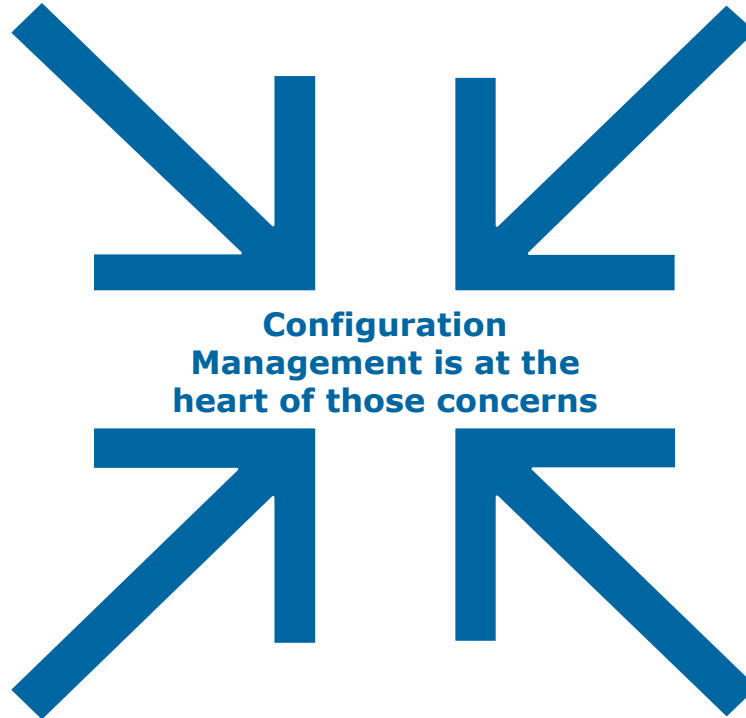
Return on experience – focus on Design Office domain

29-09-2017

The Design Office of main aerospace industrials are facing the following challenges

- Challenge to retain knowledge while there is no longer development programme and people/experts are moving to other activities

- Market is pushing industrials to decrease their development lead-times to improve product time to market



- By integrating new technologies and sticking to demanding customers, configuration management is getting more and more complex over the time

- Internal customers (manufacturing, customer service) are relying on DO data to perform their activities

Few words on Configuration Management in aviation industry

► Assumptions:

- Each time there is a need to modify the design of the product, a demand for change is created
- Design Office experts must identify all the functional and physical parts of the product impacted by the demand of change
- For each impacted part, the Design Office being in charge will have to propose a new design
- The lead-time to identify all the parts being impacted can be very long for complex modifications



Can Analytics be a game changer and help Design Office experts ?

Proof of Value objectives

- ▶ This company aims at delivering state of the art tools to their Design Office expert so they can:



Reduce the lead-time between the demand for change and the end of the investigations



Improve the quality of the information contained in the Engineering documents, which are used by internal customers



Provide Design Office with a support to identify impacts of the related change

Our conviction: Analytics is a game changer for Configuration Management



- ▶ Our experience on a similar projects demonstrated the following improvements:

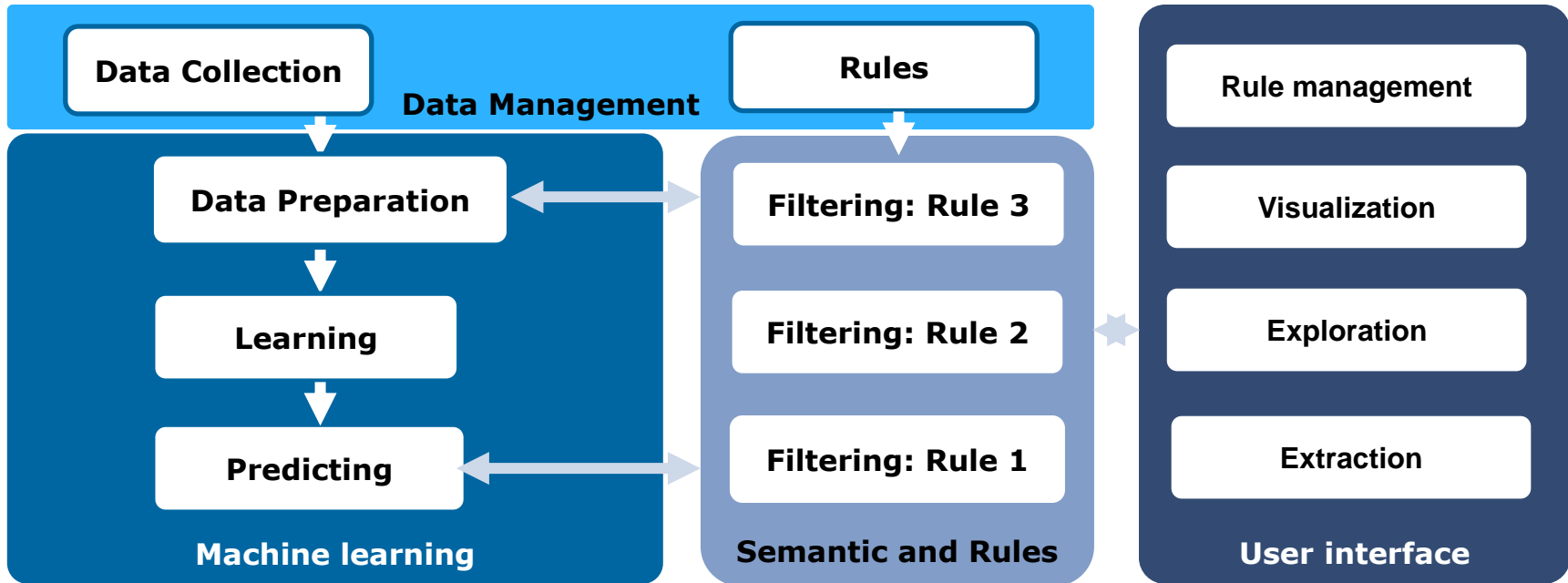
Reduction of the necessary loops to validate a technical subject

Technical investigations do not only rely on current experts but on the history of all previous subjects

Workload reduction for experts, focusing on complex topics

Use of correlated data (gathered from other processes) brings added value to users

The modules implement the complete set of functional requirements



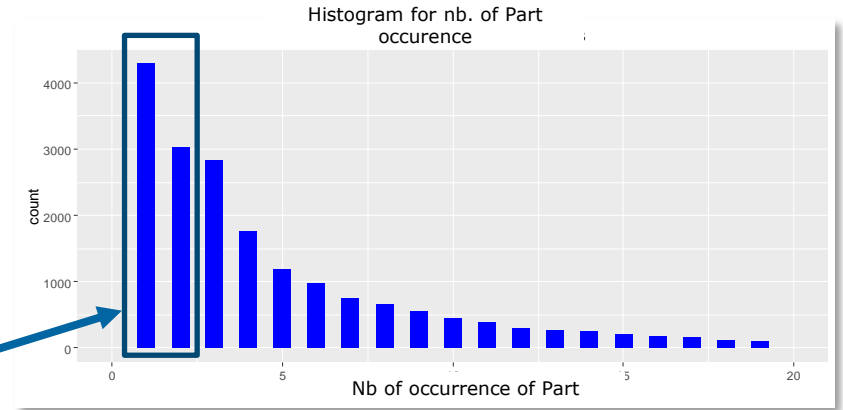
Data Science approach: Data exploration

► Basic Intuition

- **Parts mentioned together in a same DO document in the past are somehow correlated and may appear again together**
- The more frequently they appear together, the more probably they will appear again together
- No clear order relation : Part1 => Part2 or Part2 => Part1

► Focus on the file of Part attachments to DO Document

- 21% of Part s occur just once in data set => **prediction is not possible**
 - 14% of Part s occur twice => prediction is possible in some cases, but without strong confidence
- => **Max sensitivity is limited by the data set ~ 75%**
- Prediction space has the size of Part space = 22.531 items, which is too big to use some "natural" approaches



► Text fields

- Simple text mining is possible, on some variable
- Harvesting full information needs a more complete semantic framework in order to understand correctly the concepts and relationships

=> **Max sensitivity might be limited by the data set ~ 75%**

Data approaches: Mathematical models

► We used a number of various strategies

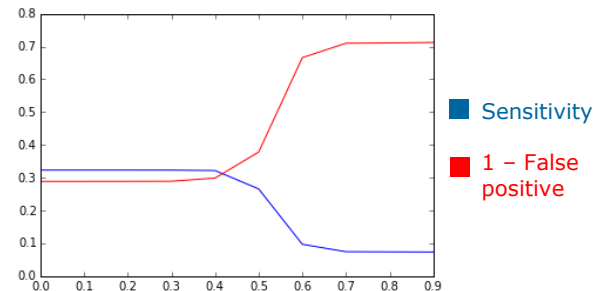
- Recommendations : direct use of correlations between Parts
- Regressions : modelling one Part as an expression of the others
- Bayesian networks : modelling probability of one Part given the appearance of another Part
- Forests / decision trees : modelling more complex structures in the Parts representing choices or rules
- Clustering : finding groups of Parts easier to model in a homogeneous manner
- Text mining : as an extension to another approach, based on text fields

► And found hurdles and some discoveries

- Confirmed Partorder independence intuition through Bayesian network model
- We obtained contrasted results : good sensitivity with very bad false positive or reverse
- Trying to solve, we identified that large DO Document might contain false correlations between Parts (eg when DO Document title mentions MISCELL.)
- We improved false positives by :
 - ⇒ Removing DO Document larger than 65 Parts from learning set
 - ⇒ Adding a negative bias on DO Document size, i.e. when Part1 appears with Part2 in a document, the strength of the signal between Part1 and Part2 decreases with DO Document size

	Part 1	Part 2	Part 3	Part n
DO Doc1	0	0	1	0	0	0	1	0	0
DO Doc2	0	0	0	0	0	1	0	0	0
DO Doc3	0	0	0	0	0	0	0	0	0
...	0	0	1	0	0	0	0	0	0
...	0	0	0	1	0	0	0	0	0
...	0	0	0	0	0	0	0	0	0
DO Doc6797	0	0	0	0	0	1	0	0	1

The central object is the large matrix that links Parts together through the DO Document



Example of the compromise between sensitivity (green) and false positive (red) for regression models

Results of machine learning activities and improvement proposals

► Results

- The best compromises sensitivity / false positive are obtained with 2 algorithms :
- **Success :**
 - The sensitivity is in line with what can be expected considering the intrinsic limitations of data used and problem complexity
- **Limitations :**
 - The false positive remain high and needs to be improved
 - A significant ratio of DO Document return no predictions, due to the high number of low rate Parts

RESULTS	Sensitivity	False positive
Classification model	40%	78%
Direct custom correlation	45-65%	75%

► To solve limitations

1. Expert review

- Review with experts the cases of bad prediction
- Identify root causes and model workarounds : add rules, filter out data, ...

2. Using additional data

- Using other data about Parts (e.g. Product structure elements) to improve the knowledge of the structures uniting the Parts

3. Semantic analysis

- Extract **Named Entities** from DO Document titles and Parts description files
- Add the **Named Entities** as additional features for machine learning

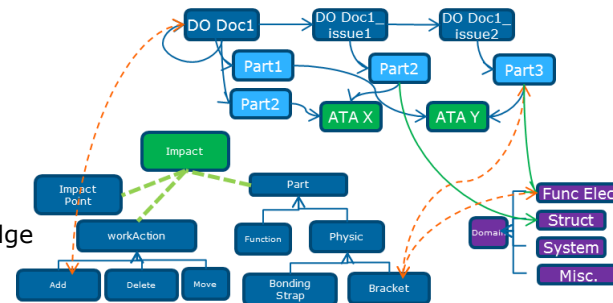


Illustration of a simple semantic graph to model concepts used in Parts and DO Document

About Semantic Analysis

Understand DO Document with Ontologies and Text Mining

– Model DO Document with Graphs and Ontologies

- Support for Knowledge processing
 - Rules for pattern matching, feature additions, ...
 - Graph processing for ML, features extraction, Search Engine, ..

– Use ontologies to recognize entities from texts

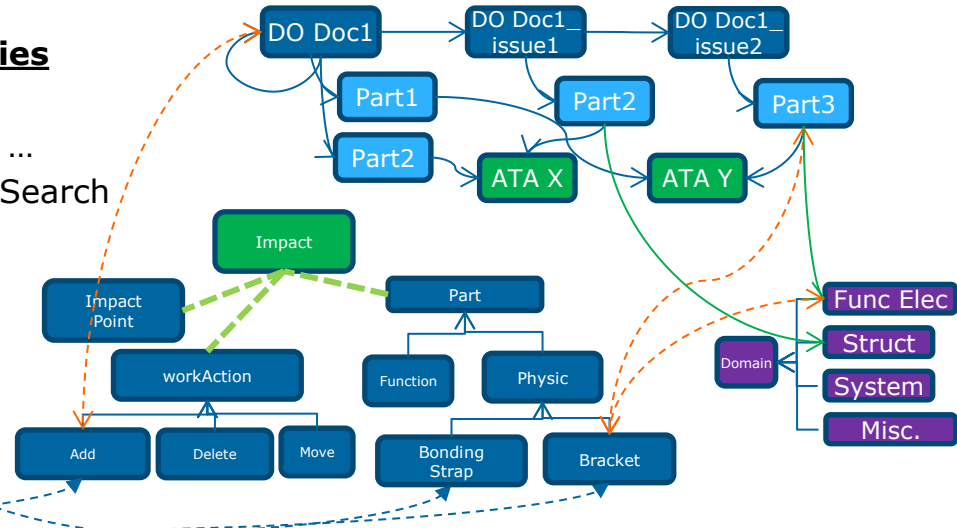
- example from a demand for change Title:

INSTALL ATTACHMENT FITTINGS FOR BONDING STRAPS

- Add extracted entities into semantic graph
 - New features for ML ; Better clustering

– Machine Learning on Semantic Graphs

- example algorithm: Tensor Factorization [Nickel 2015]



Benefits of the Proof of Value approach implemented

1

Increase business value (cost & lead-time reduction, quality improvement) using analytics

2

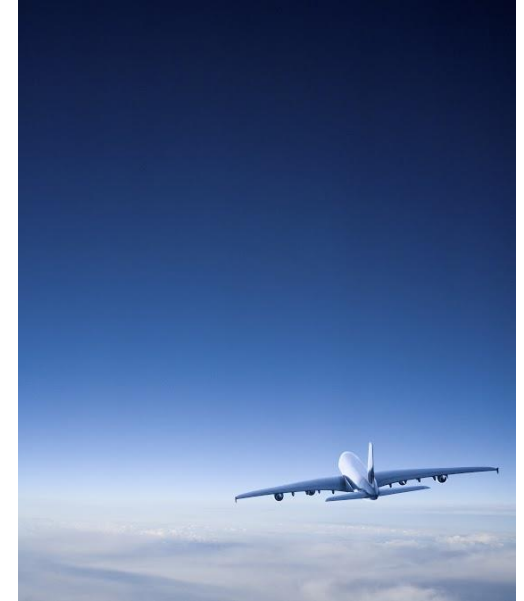
Demonstration of the added-value of analytics and data visualisation in this specific perimeter

3

Get a prototype (in weeks), that will be used by DO experts (for months) on a real data set

4

Prepare the next steps for further deployment



We bring agility in Business Analytics ... and industrialize Big Data

Innovation Booster

Agile Analytics Factory

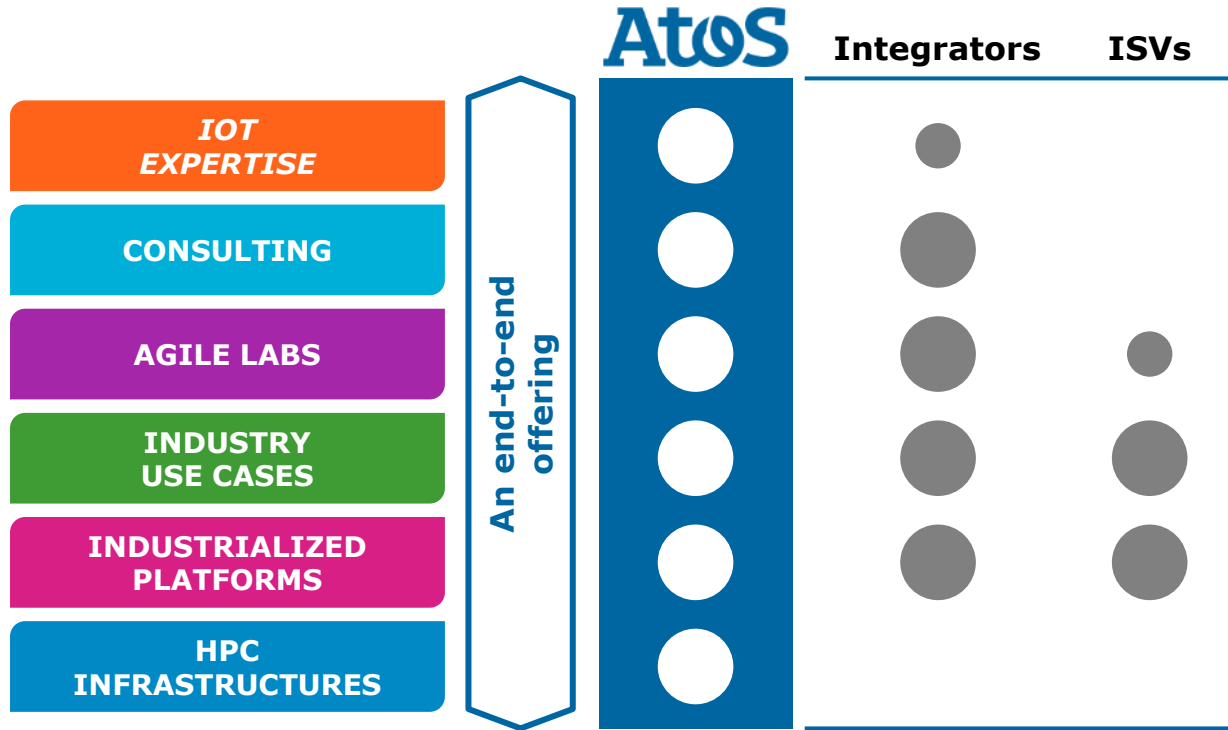
Your business opportunities confirmed in weeks

- ▶ Fast B-Cases
- ▶ Co-innovate on best-in-class cases
- ▶ Bootstrap analytics applications design

Industrial Big Data

- ▶ Data Management & Knowledge Integration
- ▶ Secured platforms for massive IoT management and Analytics
- ▶ Business models (aaS, BPO, ...)

Our end-to-end Atos Codex capabilities help our customers be one step ahead



Our difference:

- ▶ **Open innovation:** we collect the world's intelligence and make it work for our clients
- ▶ **Made to measure:** we help our clients differentiate by crafting the best fitted platform to their unique business context
- ▶ **Sovereign:** get secure, avoid lock-in

Thanks

For more information please contact:
SGHIOUER, KAOUTAR <kaoutar.sghiouer@atos.net>
JARRIGE, PIERRE <pierre.jarrige@atos.net>

Atos, the Atos logo, Atos Codex, Atos Consulting, Atos Worldgrid, Worldline, BlueKiwi, Bull, Canopy the Open Cloud Company, Unify, Yunano, Zero Email, Zero Email Certified and The Zero Email Company are registered trademarks of the Atos group. September 2017. © 2017 Atos. Confidential information owned by Atos, to be used by the recipient only. This document, or any part of it, may not be reproduced, copied, circulated and/or distributed nor quoted without prior written approval from Atos.

The Atos logo is displayed in a white, bold, sans-serif font. The letter 'o' is stylized with a white outline and a blue fill, and the letter 'S' is also white with a blue outline.