

Cooperation between SNCF and UIC for the use of Machine Learning tool on Safety event databases

Guillaume FOEILLET, SNCF DSG on behalf of SNCF RESEAU DGNUM



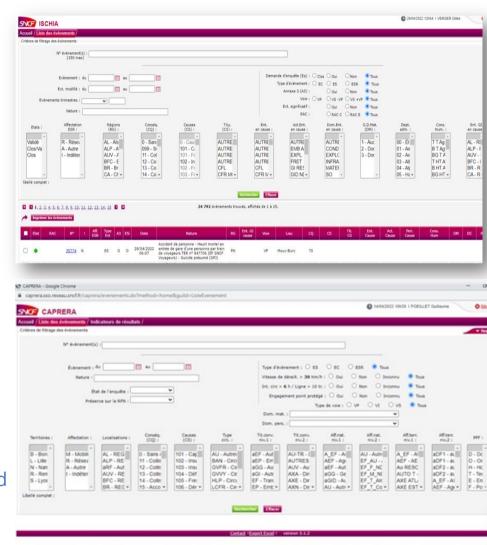




SNCF SAFETY EVENT DATABASES:

ISCHIA & CAPRERA

- -More than 15 years of railway safety events for [SNCF RESEAU] Infrastructure Manager side & [SNCF VOYAGEURS] Passenger Train Operator side
- Currently, information can only be accessed via limited predefined filtering criteria or via the safety event tag number
- Free-chosen key words can not be provided as filtering seed
- → **Stakes**: optimise the use of data, switching from a customised research mode to a multi-purpose / multi-criteria request mode



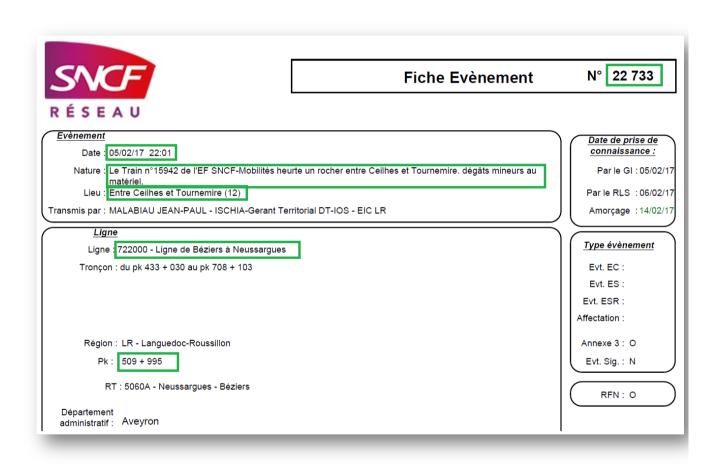
as presented during IRSC 2022



Detailed view of extracted structured fields



- Document number
- Date of safety event
- Where event took place : region, department, kilometric point
- Train operation mode
- Train number
- Track type
- Type of power car/engine
- Use
- Line number / Kilometric point



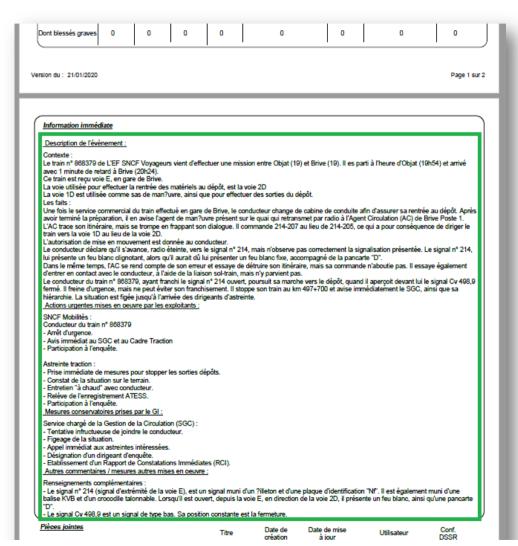
as presented during IRSC 2022



Detailed view of extracted non-structured fields

Acquisition of non-structured

- Safety event context (circumstances)
- Causes, root causes
- Consequences
- Type of risk(s)
- Decisions made, actions to be taken





Current SNCF Use

Words provided by SA SNCF in the query bar of Browser

Parameters for the SA SNCF dedicated to the query bar of ML browser

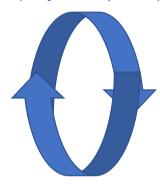
- Brackets
- Number of acceptable levels
- Logical operators

<u>Links between words + Lists of</u> <u>types, according to SA SNCF</u>

- Pure synonyms (FR)
- Railway words related
- Companies of Railway world
- Rolling stock types
- Rolling Components of Rolling stock
- Types de infrastructure components

••••

(subject to updates)



DGNUM
TOOL FOR
MACHINE
LEARNING

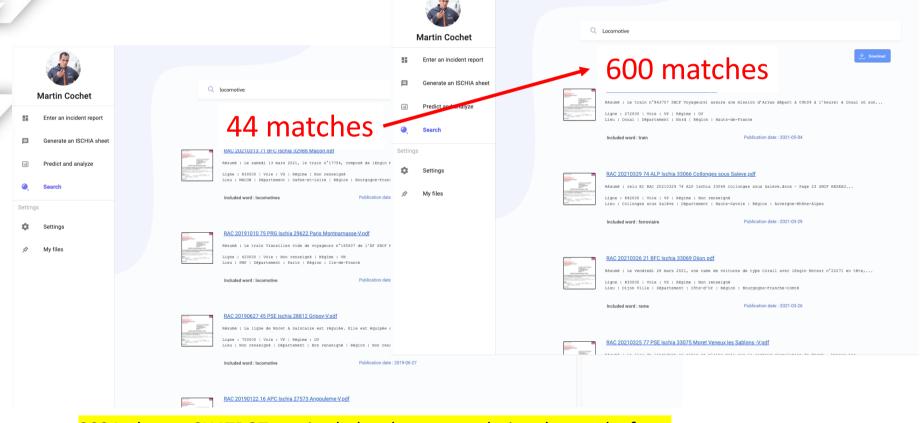
DataBase containing safety events of an SA within SNCF Group

Results for SA SNCF from query: on screen and in files 2024: includes CHATBOT service

Parameters for SA SNCF Relative to output files



Machine Learning's enhanced efficiency for parsing



2024: the ew CHATBOT service helps the user analysing the results from the parsing and can build on its own synthesis document given the collected results

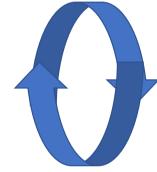


For your company -without -for free

-without disclosure
-for free

YOUR Company provides
keywords in the query bar
of Browser

(decision taken within YOUR company to accept updates)



YOUR
TOOL FOR
MACHINE
LEARNING

dedicated parameters
for the guery bar of ML browser

YOUR company defines

- Brackets
- Number of acceptable levels
- Logical operators

YOUR COMPANY sets

- + links between words
- + Lists of types,
- Pure synonyms (your idioma)
- Railway words related
- Companies of Railway world
- Rolling stock types
- Components of Rolling stock
- Types of infrastructure components

.....

DataBase containing safety events of YOUR company including

Results for YOU from query:
on screen and in files
2024: includes CHATBOT service

UIC itself

Parameters for YOU as user relative to output files

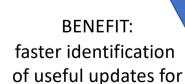


SNCF (DGNUM)

MOC



Numerous members from railway sector



our common ML tool

Identify some railway companies which would be interested in using the ML tool from DGNUM



Some expected results from this cooperation with/via UIC...

- Effortless promotion and without expenditure via UIC towards all its members.
- The share of this ML tool will inevitably induce share of Safety Best Practices.
- Since AI is moving on very fast, it helps us all start for free, then decide what AI is worth using for.
- SNCF RESEAU DGNUM can help for the the initial technical setting of the ML / AI topic.
- No threat to your internal stakes / your on-going projects, no IP issues to be considered.
- For all parties: no obligation of results, no obligation of means.
- Can contribute to speed up the use of Machine learning in the railway domain.
- Can contribute to a faster harmonisation of what is the railway safety.
- Useful application topics for the railway sector can be identified and shared between companies.
- Will help improve faster the DGNUM ML tool with enhanced functionalities.
- Results obtained by other companies will be a positive feedback, promote ML within SNCF group.



Understand better data contents, identify potential safety strategy improvements

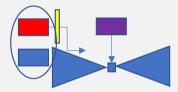
[Non-structured data] are now also taken into account

- Safety event context (circumstances)
- Root causes
- Consequences
- Risks
- Risk Control Measures (RCM) decided and set

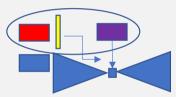
It offers some prediction opportunities

- type of event (precise set of root causes)
- Identification of cumulative consequences
- Risk control measures to prevent or mitigate root cause
- RCM to prevent or mitigate consequence
- Type of rolling stock with type of cause / of consequence
- Location of incident (regional, local) with root cause
- Impact of safety event given the root cause
- Type of signal passed at danger given the root cause

Find out [cause & cause]

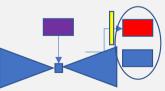


Find out [cause & safety event]

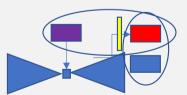


Improve existing Bow-ties Improve existing Bow-ties Create new Bow-ties

> Find out [consequence & consequence]



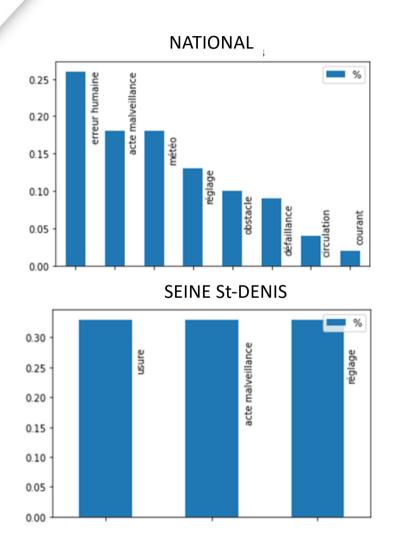
Find out [safety event & consequence

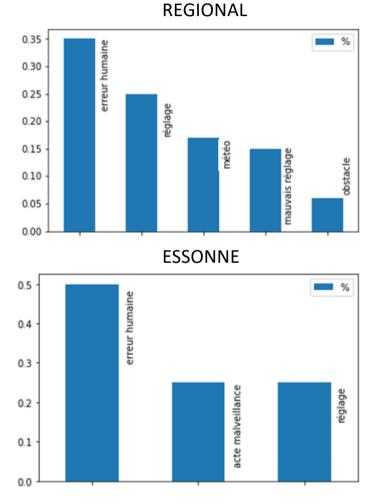


Improve existing Bow-ties Improve existing Bow-ties Create new Bow-ties



Example of compared analysis obtained







Thank you for your attention!