

AI CHALLENGE



CHALLENGE COME FROM

MAIA Team has been selected
For The Web Marketing Festival
Hackathon in Rimini.

After that we decided to
propose it as a challenge
to the DSS



2019 | Web
Marketing
Festival

- # THE CHALLENGE
- **The Data Set:**
<https://drive.google.com/drive/folders/1lij4ZShxV5MVxgba-oLkUrYcJwKoNuHw>
 - The Data set consists of a set of RTF documents named in the format ID.rtf (for example: 1000.rtf)
 - The Challenge consists of 3 task

FIRST TASK

- The Teams must indicate for each document the main category to which it belongs

- The Possible categories:

1	Economy
2	Regulation
3	Enviroment
4	Health related Issue
5	Industry
6	Cultures and customs

- The teams must provide a csv file containing the categorization performed for each document.
- Example csv file: each line → ID document, ID category
(1000, 1)
(1001, 3)
(1002, 6)



Labelling

SECOND TASK

- The Teams will have to group / relate / structure the documents analyzing the contents.
- As advice and starting point let us remember that the documents deal with recurring themes and focus on certain topics
- The visualization of the result obtained in this task is part of the third and last task

FINAL TASK

- The Teams will have create the prototype of a web or desktop application that:
 - displaying the relations and groups of documents.
 - Each relation/group of documents must be displayed by one or more keywords that clearly identify the relationship itself and the criteria used (for example, the topic).
 - For each relation/group it will be possible to see the documents that compose it (for convenience their ID and maybe the content)
 - Each Document can be part of more than one relationship/group.
 - Have a prototype search tool by which the user can search in groups found by means of keywords or phrases



LET'S TRY



DIGITAL TREE

Innovation Habitat



DataScienceSeed

Coltiviamo competenze in Data Science
con esperienze dirette basate su sfide e dati a chilometro zero

