Funded by the Horizon 2020 Framework
Programme of the European Union
SocialTruth - Grant Agreement 825477

Deliverable D3.2
Title: SocialTruth Semantic Analyzer

Dissemination Level:	**PU**
Nature of the Deliverable:	**R**
Date:	**30/09/2019**
Distribution:	**WP3**
Editors:	**Expert System**
Reviewers:	**ICCS, Qwant, LSBU**
Contributors:	**ALL**

Abstract:

The main aim of this deliverable is to define and set-up the technical methodology for semantic content analysis and to provide the tools to the partners of the consortium. The semantic services will be primarily used to analyse all the textual content acquired from the sources retrieved by the crawling and social mining mechanisms.

**** Dissemination Level:*** *PU= Public, RE= Restricted to a group specified by the Consortium, PP= Restricted to other program participants (including the Commission services), CO= Confidential, only for members of the Consortium (including the Commission services)*

***** Nature of the Deliverable:*** *P= Prototype, R= Report, S= Specification, T= Tool, O= Other*

Disclaimer

Revision History

Date	Rev.	Description	Partner
05/07/2019	0.1	First version of DDP	UTP
12/07/2019	0.3	Initial deliverable structure (TOC)	ESF
29/07/2019	0.4	Initial deliverable content	ESF
2/09/2019	0.5	Aggregation of partners feedback and contributions	ESF, QWANT, UTP, ICCS
11/09/2019	0.6	Updated version after periodic technical meeting (remotely)	ESF, QWANT, UTP, ICCS, LSBU
18/09/2019	0.8	Updated version after periodic technical meeting (remotely)	ESF, QWANT, UTP, ICCS, LSBU
24/09/2019	1.0	Updated version after periodic technical meeting (remotely)	ESF, QWANT, UTP, ICCS, LSBU
28/09/2019	1.2	Updated version, based on peer-review feedback	ESF
29/09/2019	1.3	Final version for submission	ESF, ICCS

List of Authors

Partner	Author
ESF	Sonia Collada, Nahid Oulmi, Patrick Chable, Margot Avella
ICCS	Kostas Demestichas, George Koutalieris
UTP	Michal Choras, Rafal Kozik, Pawel Ksieniewicz
QWANT	Stan ASSIER, Noel Martin
LSBU	Manik Gupta, Chathura Galkandage

Table of Contents

Index of figures

Executive Summary

Social Truth's goal is to produce concrete results on fake news detection with significant technical and strategic impact. Cogito is cognitive technology that is able to understand unstructured text and to categorized and/or extract information from it as a human being would do.

This document aims at highlighting the information that will be provided by Expert System solution. We will provide a tool and support regarding the textual and semantic analysis; we will develop a meta-verification system on story classification and ranking. The global aim of the approach we have worked on is, based on a golden corpus, comparing an untargeted document to our base of "qualified documents". The hypothesis applied lies in the fact that "true news" all have a pattern that our tools will highlight. Expert System will be able to go from a large group of articles to automatically select a relevant group to compare to a possible "fake news".

The first milestone is the categorization. We have the untargeted document; we analyze it using the categorization process and reduce the list to compare it to. The second milestone corresponds to narrowing the search by using the clustering. The third milestone is the similarity analysis: how similar is it to equivalent documents? We are talking about the sentiment present or not in the text and the writing style. We end up by extracting the feelings and the writeprint of the document.

Different hypotheses were established about the pertinence of the information provided by the different milestones. Two of them have been verified (the ones related to the tone and the vocabulary use), and one has been rejected (regarding the language level). However, those hypotheses will be tested on bigger corpora to ensure the accuracy of the semantic analyzer. Focusing on the source or on the author might be another signal of relevant contribution to the project. It would thus reduce the risk of the possible mistakes being committed. The use of a thesaurus and ontologies of labialized sources could be considered in future analysis.

The outcome of this deliverable is a semantic analyzer that aims at providing information that will be the input of the expert meta-verification system.

1. Introduction

1.1 Forewords

Social Truth's goal is to produce concrete results on fake news detection with significant technical and strategic impact. To achieve this, eleven partners have been chosen. Five of them are industrial and commercial partners, three of them are researchers and academic partners and the three lasts are end-users and will act as facilitators. The whole consortium covers a large spectrum of abilities, from the blockchain to the data protection and privacy and the web services integration. Expert System, on this spectrum, is in charge and brings his primary expertise in the field of the semantic analysis as well as a significant expertise in machine learning algorithms.

Cogito is cognitive technology that enables human comprehension and insight at scale. Cogito's core algorithms, which are based on a human-like comprehension of text and an embedded knowledge graph, are made more effective and trained to work in different and/or very specific domain, by combining them. This means that Cogito is able to understand unstructured text and to categorized and/or extract information from it as a human being would do.

The two main engines inside Cogito are:

- *Disambiguator,* evaluates and understands a sentence in his context. Thanks to Disambiguator, Cogito can make the difference between the meanings a word can have by understanding the sentence as a whole and part of the text.
- *Sensigrafo*, is the Expert System proprietary knowledge graph, a representation of knowledge where concepts are connected to one other by semantic relationships. Sensigrafo is perpetually evolving and can also be expanded through the acquisition of new knowledge from subject matter experts. Sensigrafo is designed to interact with Disambiguator to resolve the ambiguity in the meaning of each word, a fundamental step in the text analytics process.

By means of these engines and using Expert System's tool Cogito Studio a human being can teach Cogito on how to analyze information and provide signals able to help in the identification of a "fake news".

1.2 Contribution of this deliverable to the SocialTruth solution

It is important to understand the scope of our action; Expert System will not be responsible to decide and to settle whether it is a "fake news" or not. We will provide accurate information and relevant elements of comparison to other members of the consortium. This document aims at highlighting the information that will be provided by Expert System solution. We will provide a tool and support regarding the textual and semantic analysis; we will develop a meta-verification system on story classification and ranking.

Figure 1- Expert System in the consortium

The beginning of the chain is a document and the question of its nature ("fake" or "true" news). The approach shall consist of progressive steps to be followed, and consortium members' work to be applied. We are only a part of this chain and we will not decide whether the document is or not a "true" news.
In order to provide relevant information, we will conduct our research on 1,627 documents.

The whole WP3 is summarized at the left of this chart presented on deliverable D2.2:

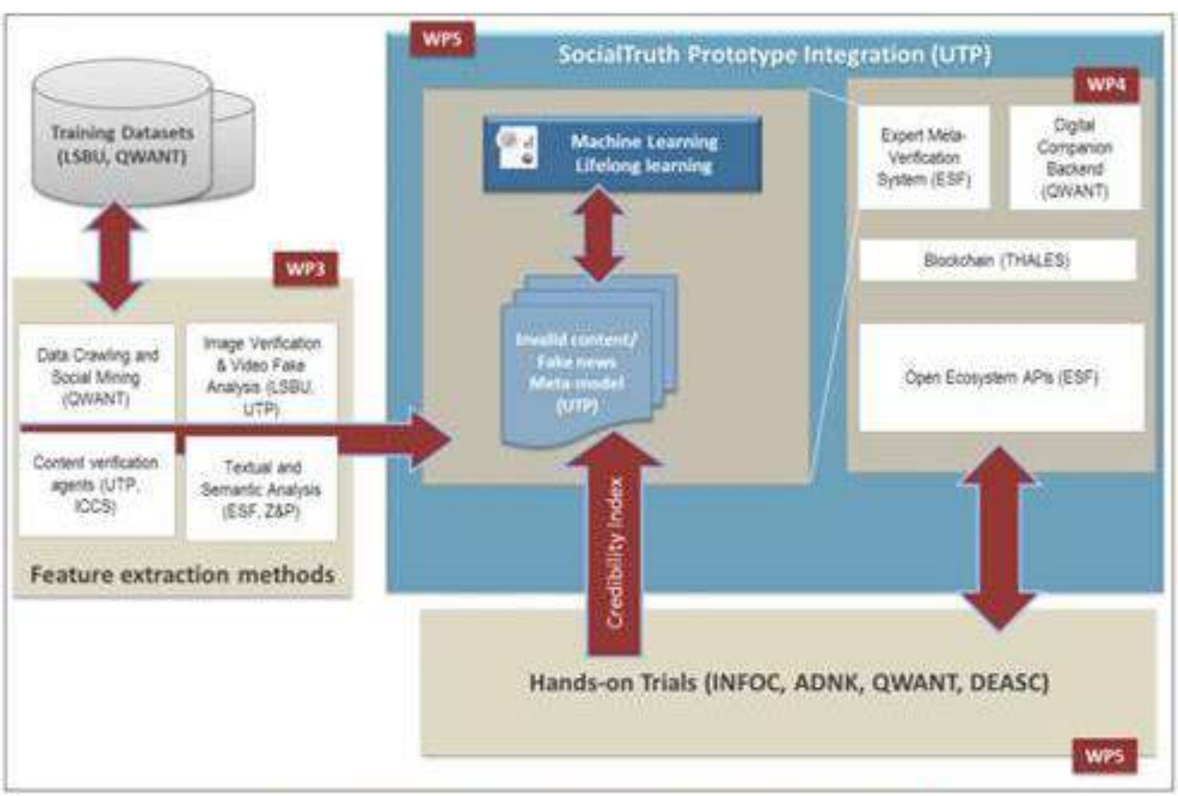


Figure 2- Conceptual workflow and interaction of SocialTruth development activities

As seen on this chart, Expert System will integrate in a larger feature extraction pipeline. Expert System handles Natural Language Processing (NLP), but other types of features such as sources (medias, websites), images, videos, social content (likes, shares...), and features extracted by verification agents (human content verification).

1.3 Presentation of the data: buzzfeed – webis dataset

Social Truth being a project related to data, we need a dataset in order to provide accurate information, furthermore we need this corpus to test our model and our hypothesis. LSBU provided a list of relevant articles about fake news detection and among them we found a resource online giving us an access to 1,627 articles that were fact-checked by professional journalists at Buzzfeed. All 1 627 articles, that we call "qualified documents", were labelled by these journalists as 4 target categories: "mostly true", "mostly false", "mixture of true and false", "no factual content". The topic of the corpus available is: the US presidential election of 2012.

All the articles are from 9 different publishers and they all have been published a week close to the US election. Three out of nine publishers were identified as from the right-wing (Eagle-rising, Freedom-daily, Right-wing-news), three from the left-wing (Addicting-info, The-other-98, Occupy-democrats), three mainstream (ABC, CNN, Politico) publishers. Six out of nine were identified as prolific hyper-partisan ones (either right-wing or left-wing), and all the publishers have earned Facebook's blue checkmark (Facebook's proof of authenticity and elevated status within the network).

1.4 Dataset analysis

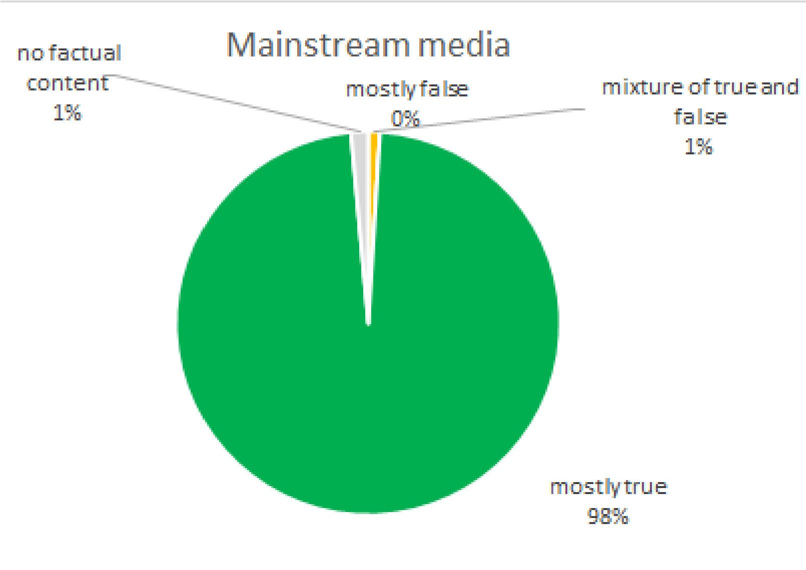
1.4.1 Presentation of our dataset

From this corpus of 1,627 documents:

- 1264 have been labelled as "mostly true"
- 87 have been labelled as "mostly false"
- 212 have been labelled as "mixture of true and false"
- 64 have been labelled as "no factual content"

- 826 are known as coming from mainstream publishers: ABC, CNN, Politico
- 256 are categorized as being from the left-wing: Addicting-info, The-other-98, Occupy-democrats
- 545 are categorized as being from the right-wing: Eagle-rising, Freedom-daily, Right-wing-news

We can see that the truthfulness of a source can be correlated to its political orientation on this dataset, as showed by these graphics:



In the mainstream media, 98% are mostly true, 1% is a mixture of true and false, none are mostly false and 1% is made of no factual content.

Figure 3- Content of mainstream media

In the right-wing media, 51% are mostly true, 8% are no factual content, 28% are a mixture of true and false, and 13% are mostly identified as false.

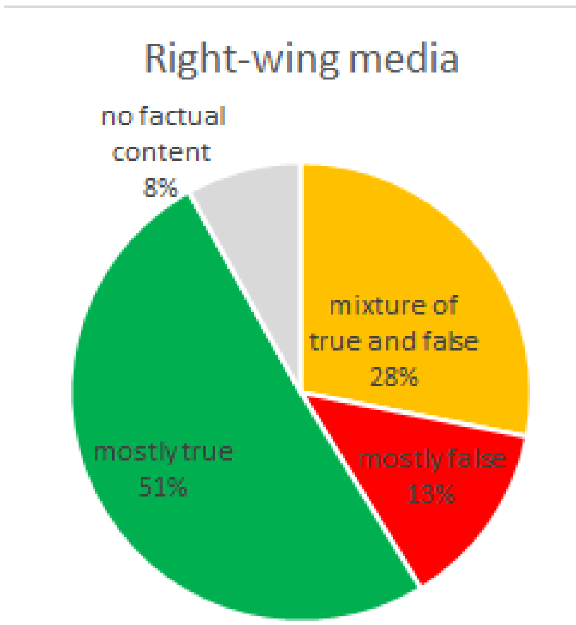
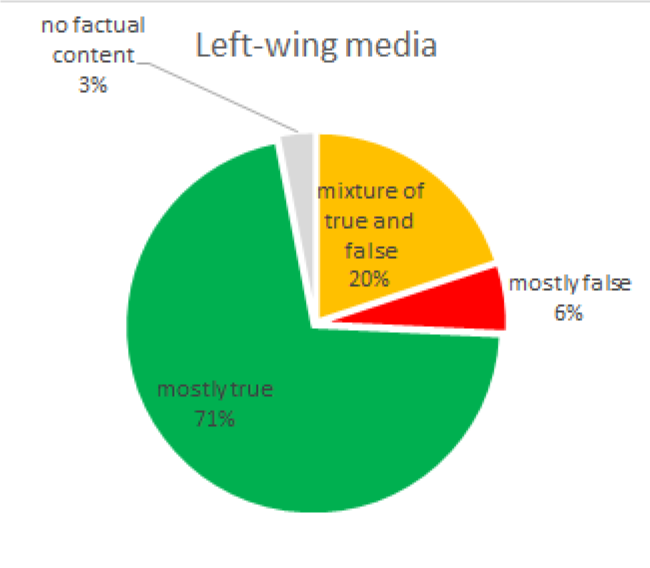


Figure 4- Content of right-wing media



In the left-wing media, 71% are mostly true, 3% are no factual content, 20% are a mixture of true and false and 6% are mostly false.

Figure 5- Content of left-wing media

To summarize, according to the graphs we can say that mainstream media are more reliable sources when it comes to veracity than right and left wings articles. Inside those right and left wings, we can find some differences as well, the left one provides more accurate information (71%) than the right wing (only 51%), and the left wing, although publishing 20% of articles with a mixture of true and false, it remains more accurate than the right-wing articles analyzed for this study (28%).

1.4.2 Sources

A French newspaper, Le Monde, has created a special section to communicate on sources, it is called "Decodex" and has been launched at the beginning of 2017. It aims to judge the credibility of public information. This sources directory has evolved over the years and provide to the readers tips to identify themselves and forge themselves their opinions:

They give advices on how to recognize a reliable website:
- Consult the page "About us"
- Check if it is not a parodist website
- Identify the authors, most likely this information will be on the first page of the website
- Identify the sources of the website
- Is the information presented in a neutral way? Is the title a reflection of the content of the article? Is the website communication factual content or opinions? Is the tone of the article moderate or does it seem inappropriate?
- Is this website known to publish parodic/fake content?

How to judge the reliability of a source:
- Identify the author of the message (is he an authentic journalist?)

- Bu default, an information given on a website by an unknown written might be more false than true
- If several different media give the same information quoting different sources, it might be true
- Try to identify the first source that published the news
- The more an information is surprising, the more it has to be detailed and precise

We can see that our methodology is using the same idea: identify the tone, the linguistic style and the emotion the content aims to produce to the reader to identify a potential “fake news”.

2. Hypothesis and verification

2.1 Hypothesis

Our hypothesis are the following:

1. **A “fake news” must be compared to “qualified” documents to establish its value.** For this hypothesis, we have a dataset of documents, mentioned above, labeled with the 4 targets. The “true news” documents will be used to compare the results of the analysis of “Fake news” documents and understand their stylistic differences.

2. **A “fake news” may show strong sentiments as they aim at influencing the reader.** According to us, “fake news” documents do not provide the same stylistic writing and content that “true news” ones.

3. **A “fake news” may go through the “standard”, polish publishing process and be written by less skilled people.** With this hypothesis we imply that accredited journalists working for official (mostly mainstream) media hence to provide accurate information, in opposition with a casual writer copying the journalistic style and trying to disguise the information to make it look like it was written by a professional.

2.2 Methodology

The methodology is divided in several milestones. The global aim of the approach we have worked on is comparing an untargeted document to our base of “qualified documents”. To avoid wasting time, we must reduce the search perimeter: the main metric is about analyzing the similarity between two documents and the elements to analyze are the degree of emotion, anger, hatred, basically everything that differs from a neutral point of view. Another element is the stylistic writing, whether it is professional or not, and the last element is the question of the source: who is writing this? It is a professional journalist? What is the message he tried to put together? Is the shape of the message appropriate?

The main questions to focus on, are: Who is the writer? What is the message? How can I validate this source? Are there emotions in the document? How am I reacting to this document? By using a logic based after similarity we can proceed to elimination. The aim is to reduce the search perimeter to a few examples only.

The document will be classified as follow:

- True
- False
- Half true
- Can’t be tested (lack of factual content)

Our hypothesis lies in the fact that “true news” all have a pattern that our tools will highlight. Expert System will be able to go from a large group of articles to automatically select a relevant group to compare to a possible “fake news”.

As explained in the figure below, the first milestone is the categorization. We have the untargeted document, we put it through the categorization process and reduce the list to compare it to. As to narrow the search we are using the clustering (second milestone). The third milestone is the similarity analysis: how similar is it to equivalent documents? We are talking about the sentiment present or not in the text and the writing style. We end up by extracting the feelings and the writeprint of the document.

The methodology is divided in several milestones. The global aim of the approach we have worked on is comparing an untargeted document to our base of “qualified documents”. To avoid wasting time, we must reduce the search perimeter: the main metric is about analyzing the similarity between two documents and the elements to analyze are the degree of emotion, anger, hatred, basically everything that differs from a neutral point of view. Another element is the stylistic writing, whether it is professional or not, and the last element is the question of the source: who is writing this? It is a professional journalist? What is the message he tried to put together? Is the shape of the message appropriate?

The main questions to focus on, are: Who is the writer? What is the message? How can I validate this source? Are there emotions in the document? How am I reacting to this document? By using a logic based after similarity we can proceed to elimination. The aim is to reduce the search perimeter to a few examples only.

The document will be classified as follow:

- True
- False
- Half true
- Can’t be tested (lack of factual content)

Our hypothesis lies in the fact that “true news” all have a pattern that our tools will highlight. Expert System will be able to go from a large group of articles to automatically select a relevant group to compare to a possible “fake news”.

As explained in the figure below, the first milestone is the categorization. We have the untargeted document, we put it through the categorization process and reduce the list to compare it to. As to narrow the search we are using the clustering (second milestone). The third milestone is the similarity analysis: how similar is it to equivalent documents? We are talking about the sentiment present or not in the text and the writing style. We end up by extracting the feelings and the writeprint of the document.

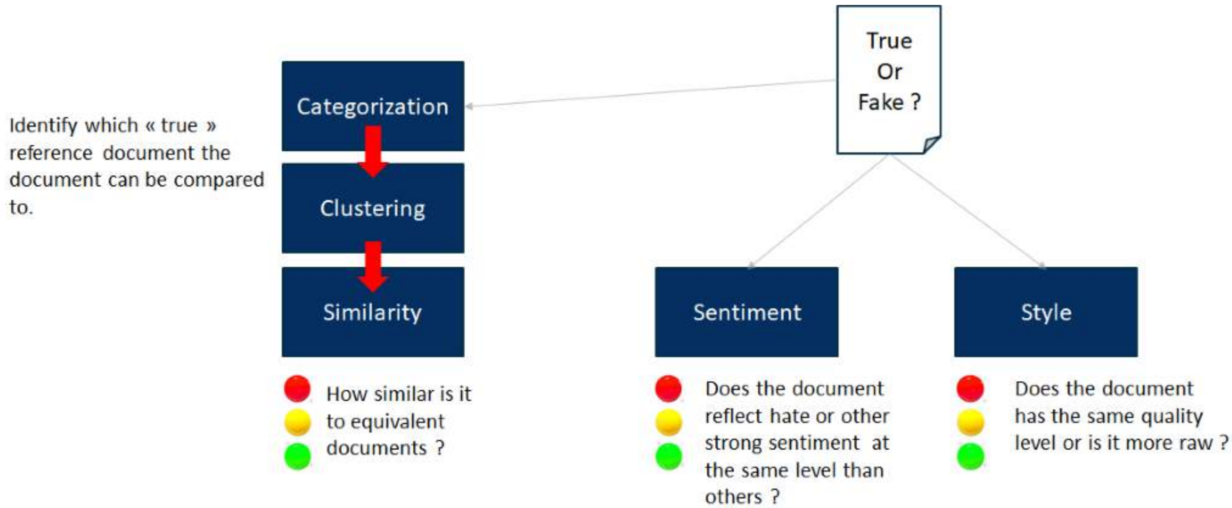


Figure 6- Expert System methodology

2.2.1 Part 1: Reduce the complexity of the data using the categorization

The first part of the methodology is to categorize documents. We will categorize documents based on their topic to analyze documents topic-wise, because we cannot compare oranges with apples. This part aims to reduce the complexity of the data and to compare new documents (untargeted ones) that are about the same topic than our "qualified documents" corpus.

- We first use the **categorization tool to understand in which group it belongs**.
  The categorization tool is divided in two substeps, the frist one is using the Cogito's default taxonomy, which will yield a general categorization of the documents and the second one is the use of an external taxonomy (such as the Mediatopics taxonomy), which will yield a finer categorization of the documents.
  Example: the use of this capacity, to be able to categorize millions of documents on the US President for example.

- We use the **clustering, that will allow us to identify the right sub-group** and therefor reduce the domain of comparison.
  Example: In all the articles talking about the US President, the clustering will allow us to only select the ones talking about his birthplace (or any other topic selected).

- Finally, we will use **similarity comparison** to select the documents that are the closest to one another, to provide the most accurate and precise comparison.

In other words, the approach is to go from general to specific. The figure below shows to identify the fact-checking set:

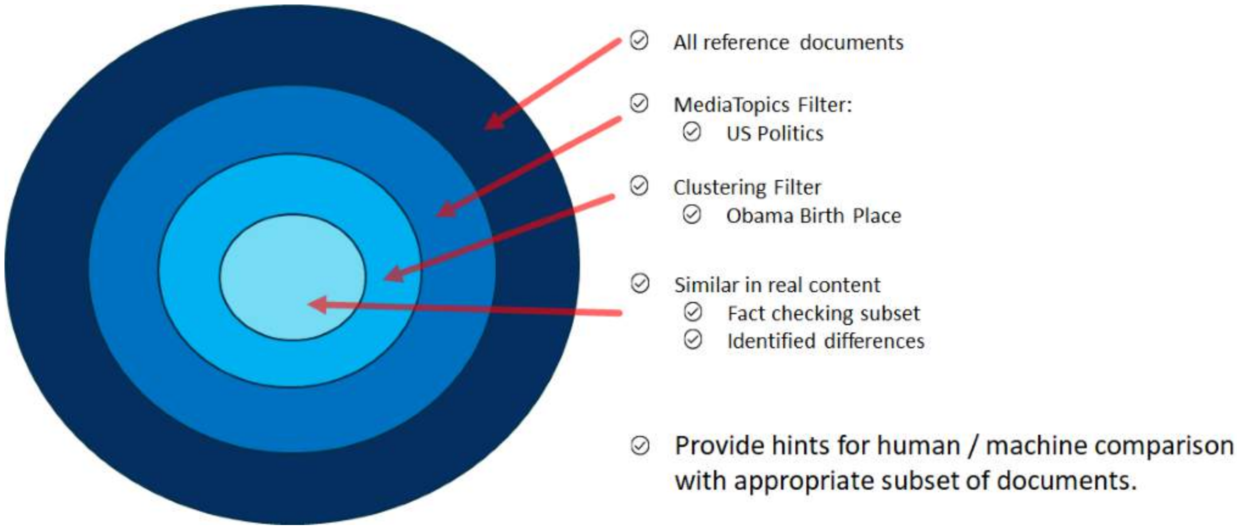


Figure 7- From general to specific

It goes from all the reference documents to the ones filtered by Mediatopics (categorization, step 1), to the clustering filter (step 2), to ending with the similarity analysis (step 3). This approach aims only to provide hints for humans to be able to make a fully informed decision, with appropriate subset of documents.

2.2.2 Part 2: How to establish the appropriate and necessary signals to the analysis and detection process

- **Signal 1**: indication of **the tone**
  The tone is important to consider and will be used to analyze the documents. The tone is linked to the second hypothesis *"A "fake news" may show strong sentiments as they aim at influencing the reader."*. According to us, a "fake news" will convey stronger feelings and will aim to create a strong impulse to the reader, whereas an authentic article will be neutral.

- **Signal 2**: **the language level of the author**
  Our hypothesis is that fake news editors do not provide the same linguistic style as true news editors. This is linked to the third hypothesis *"A "fake news" may go through the "standard", polish publishing process and be written by less skilled people"*.

- **Signal 3**: **Similarity analysis** to detect outliers
  The similarity allows us to identify the documents that are distinctly different from the "qualified" group of documents. It gives us a signal on homogeneity with the "true news" that are constitutive of our reference corpus. The process of identification comes after a deep search and the selected documents remaining are specific: they must be from the same topic. Then the reader can witness the discrepancies between the corpus on the one side and that needs to be targeted as "fake news" or not on the other side.

This milestone is constitutive of our first hypothesis *"A "fake news" must be compared to "qualified" documents to establish its value"*.

Expert System do not pretend to bring a unique solution, but we do bring evidences (signal 1, 2 and 3). The output of this methodology should give the operator the tools to take advantage of the anomalies that COGITO has highlighted. We are in a processing chain with the other consortium's members.

3. Implementation of the expert system tools

In this part, we will explain in detail how the expert system tools will be implemented to the tailor-made solution elaborated for the project Social Truth. We have chosen to introduce this part with a funnel figure to give a graphic idea of our method.

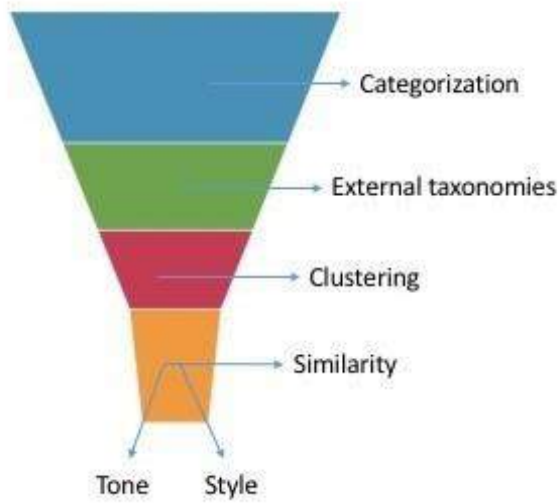


Figure 8- Implementation of the ES tools, step 1

The categorization is the first part of a long process that involves the use of external taxonomies, after this step comes the clustering, to focus on more specific articles on which, as for the third and last step, the similarity process will be applied.

3.1 Categorization of the corpus

3.1.1 Cogito Standard Domains

The Cogito Standard Domains panel is a taxonomy developed by Cogito in intern. It provides the list of all domains, chosen from a closed list of predefined domains, that Cogito Studio automatically identifies in the analyzed text. This information is available regardless of the type or number of linguistics rules developed. in a project. The information is produced automatically each time. The taxonomy is available in thirteen languages (English, Italian, French, Spanish, German, Dutch, Portuguese, Russian, Japanese, Korean, Chinese, Arabic and Hebrew).

For instance, let’s take the following "true news” document:

“Following the shooting death of an unarmed black man by a police officer in Tulsa, Oklahoma, last Friday, Democratic presidential nominee Hillary Clinton spoke out today against police violence. “This horrible shooting again. How many times do we have to see this in our country? In Tulsa, an unarmed man with his hands in the air,” Clinton said, calling into "The Steve Harvey Morning Show." Forty-year-old Terence Crutcher was fatally shot by a white police officer after his SUV stalled on the road. Video of the incident released after the shooting appears to show

Crutcher with his hands raised in the air. Calling the act "unbearable" and saying officer-involved shootings "need to be intolerable," Clinton appealed to a white audience to address unconscious discrimination. "Maybe I can, by speaking directly to white people, say look, this is not who we are. We have got to do everything possible to improve policing, to go right at implicit bias," she said. Another unarmed Black man was shot in a police incident. This should be intolerable. We have so much work to do. #TerenceCrutcher -H Clinton's policy proposal for criminal justice reform includes developing national standards on police officers' use of force, supporting legislation to end racial profiling, and committing $1 billion in funding to training programs and research to "tackle" implicit bias. ABC News' Julia Jacobo and Josh Haskell contributed to this report."

And here are the extraction results by Cogito (Standard domains):

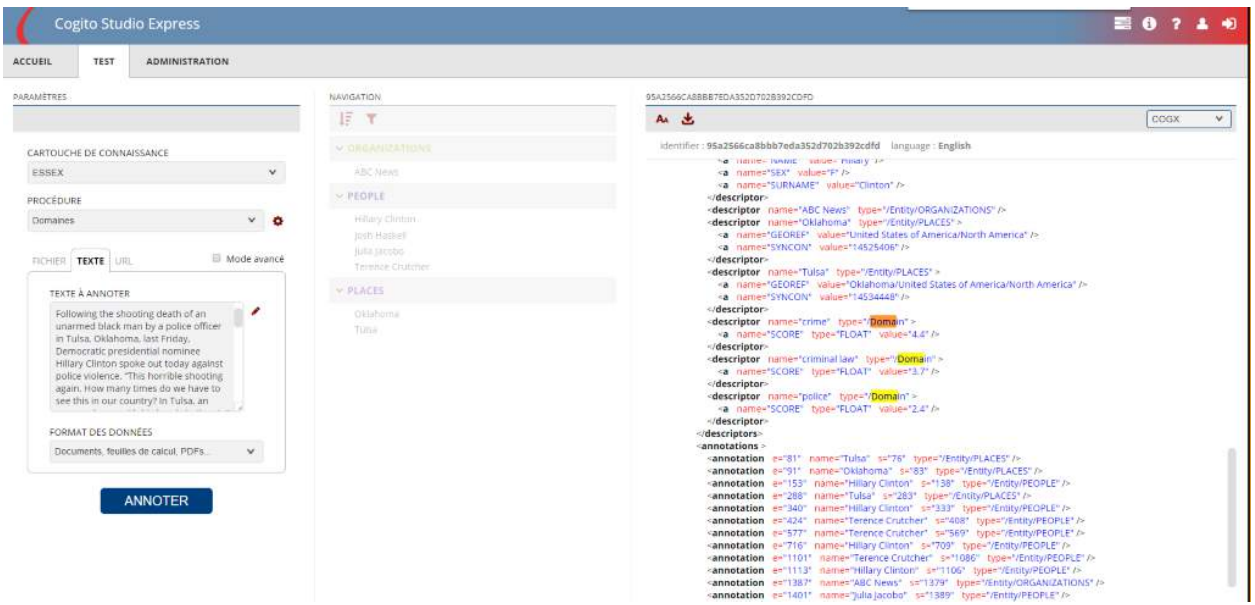


Figure 9- Cogito Standard Domains

In this example, you can see matching domains are in correlation with the text (the text is talking about a mass shooting in the US), the domains that cogito has recognized are: crime, criminal law and police. Those results are not incorrect, but we think we can go deeper in the accuracy, with the help of external taxonomies.

3.1.2 The use of external technologies

Our methodology has been thought with the use of the external taxonomy Mediatopics. Mediatopic is indeed an external taxonomy, available online and downloadable in Cogito. Mediatopics is a taxonomy that is specific to media, thus has more known concepts that our tools that can give a more generic overview.

If we take the document illustrated in paragraph above, with domains being crime, criminal law and police, the domains extracted via the Mediatopics taxonomy are finer: "Assault, Discrimination, Police, Political candidates, Racism".

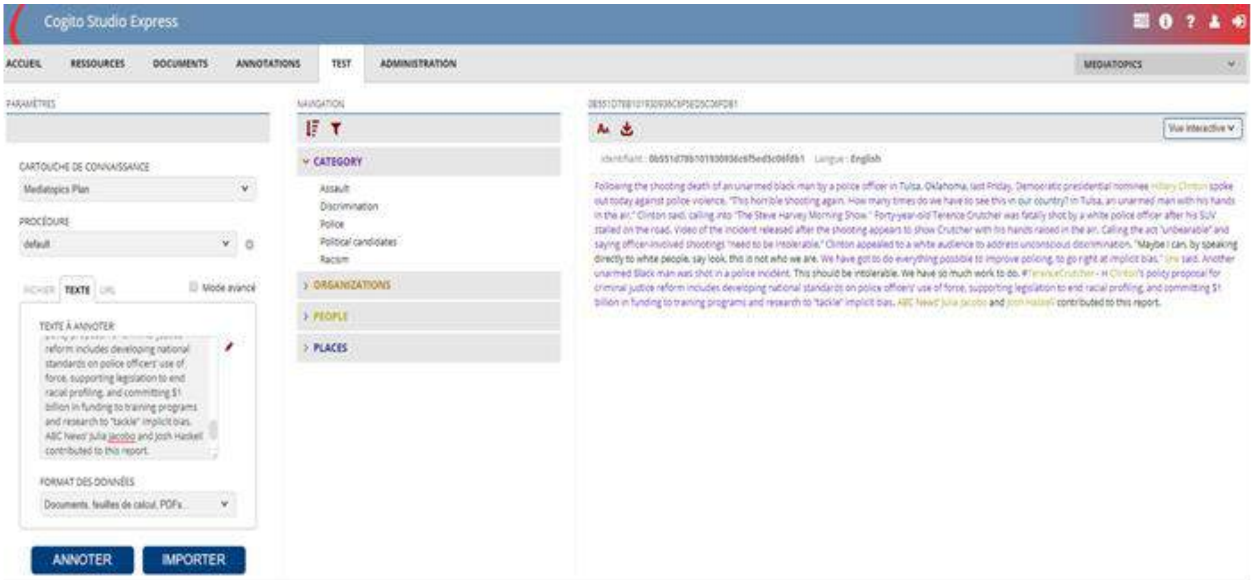


Figure 10- Mediatopics taxonomy results 1

With Mediatopics in the website itself this is the result we have: we can see their taxonomy, on the same keyword, have more details:

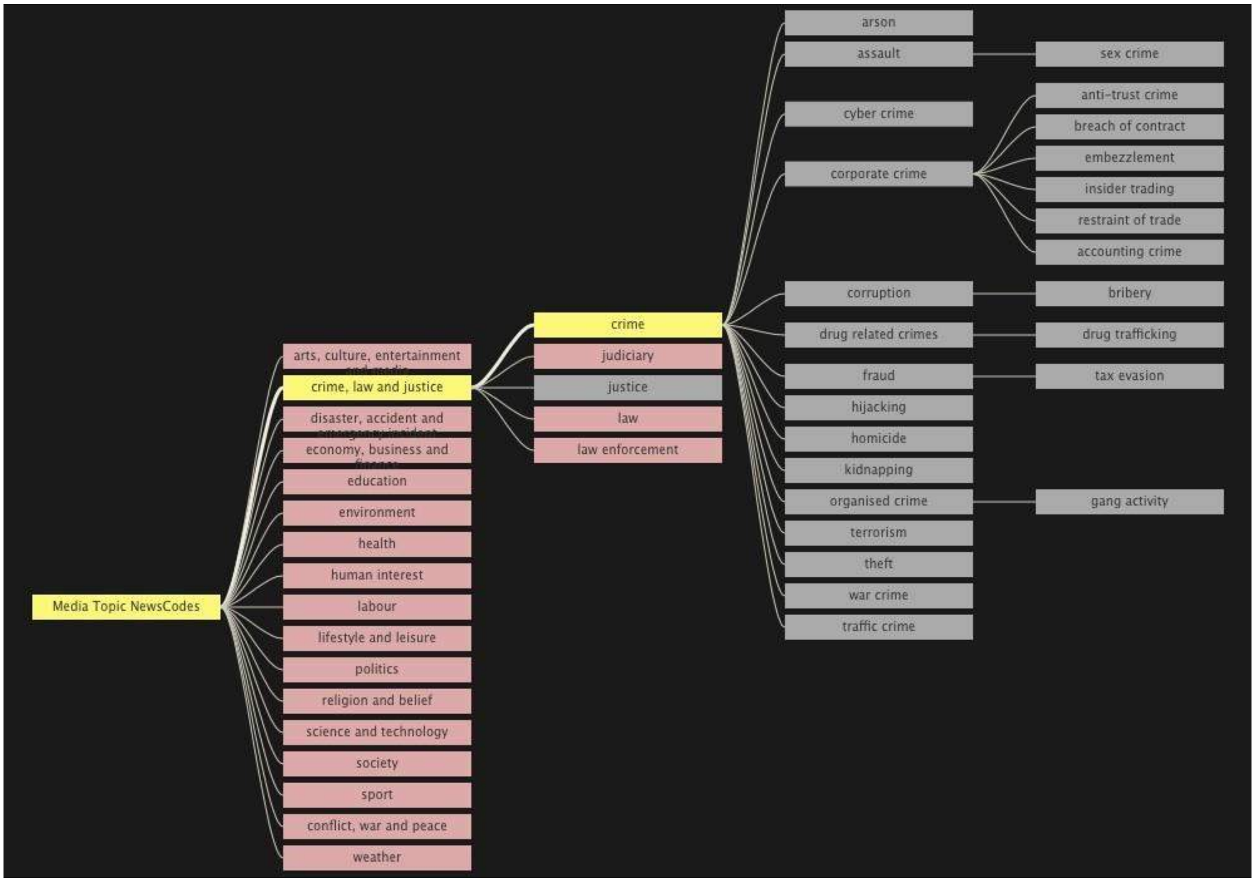


Figure 11- Mediatopics taxonomy results 2

3.1.3 Automatic categorization through Clustering

The clustering is a user-friendly tool, very graphic and playful. In an instant, the main categories are visible and recognizable. This approach is useful to treat subject that are unknown to the operator.

The clustering allows the categorization process from a set of documents without any supervision of an external taxonomy. It allows as well to detect subjects that would not have been detected via an extern taxonomy.

Figure 12- Clustering visual

To summarize the steps we have described, see the below figure that also links to the part to come:

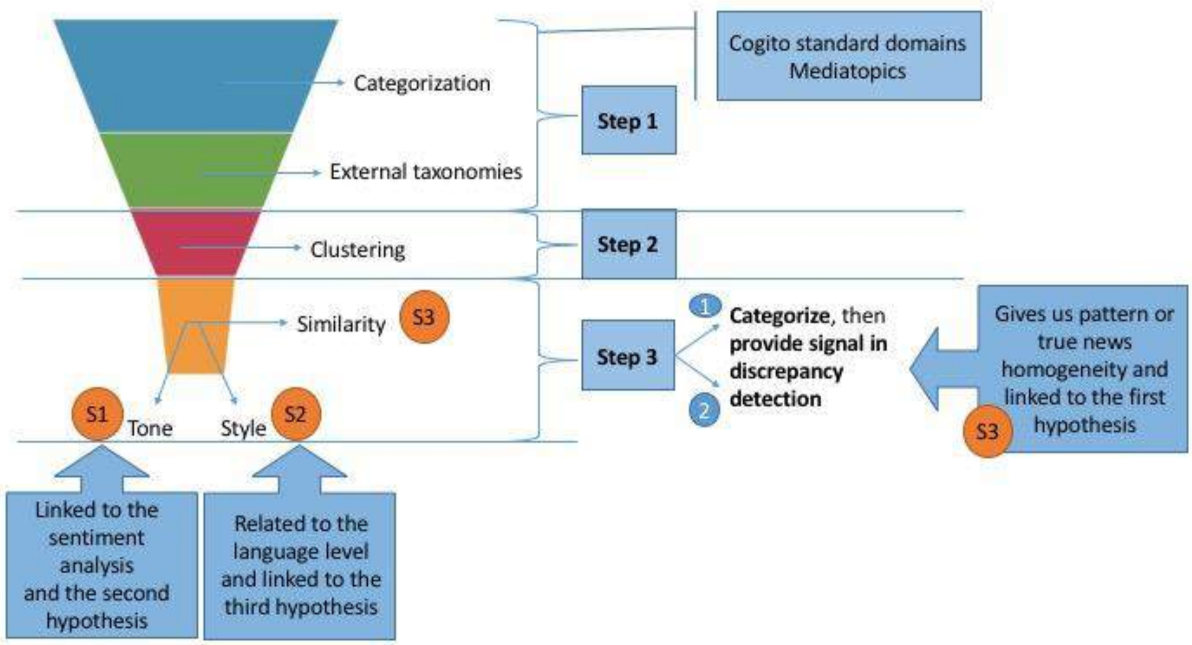


Figure 13- Implementation of the Expert System tools, step 2

3.2 Semantic Analysis

The semantic analysis is part of the signal 1, the detection of the tone. According to the previous figure, this part will focus more on the third step, the similarity and thus explore how the signals emerge from the method, and at the same time are linked to the hypothesis and our tools.

The semantic analysis is able to understand how the writer feels towards the topic he is writing about. It understands the language he uses. Expert System has thought a system that takes a word and give it a score according to the meaning it is carrying, on how it is expressed. It shows how the writer likes and dislikes and more importantly on how much he likes and dislikes - to highlight the question of intensity.

3.2.1 Automatic extraction thanks to ESSEX

Thanks to the help of the ESSEX skill cartridge, some information will, automatically and for all documents, be extracted. It will be the same nature of information in all the documents, which gives us a faire point of comparison.

Expert System's ESSEX (Expert System Semantic Engine eXtended server) is Expert System's software platform for document analysis. ESSEX exposes a set of simplified interfaces which allow for coordinated access to the base analysis capabilities, categorization and extraction. ESSEX is the semantic heart of every Expert System products. A Skill Cartridge is an annotation resource with a set of customizable knowledge components that define the information to be extracted from documents.

Essex is the core engine of the Expert System semantic platform; its main functionalities are:

- Semantic Analysis of texts;
- Conversion of documents from binary to text format;
- Recognition of document idiom and information extraction.

ESSEX works with the standard HTTP protocol, using an API paradigm via POST requests. The main scope of this instrument is the automatic management of the download and deployment of linguistic resources, in order to avoid a static installation on the machine. This way it is possible to dynamically load Expert System LPKs, making them particularly suited to clusters of machines in cloud.

Below an example of all the extracted fields thanks to the work of the ESSEX skill cartridge:

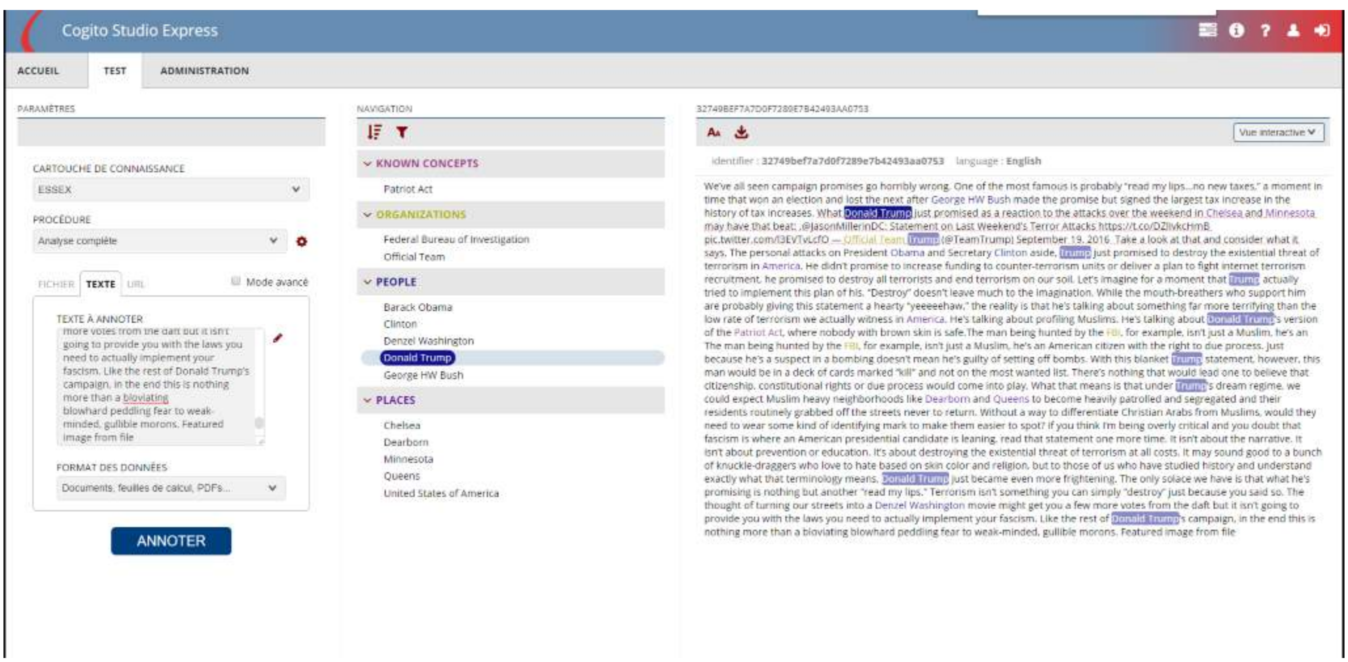


Figure 14- Essex skill cartridge

We can see in this picture all the places, people, organization and known concepts (in the *Sensigrafo*) and extracted from an unstructured text sample.

3.2.2 Sentiment analysis (Signal 1, the tone)

The ESSEX skill cartridge helps the sentiment analysis: how to detect the extreme values of a feeling transcription.

3.2.2.1 *Useful to detect extreme point of view*

Emotions are linked to each extracted entity, as the examples shown below:

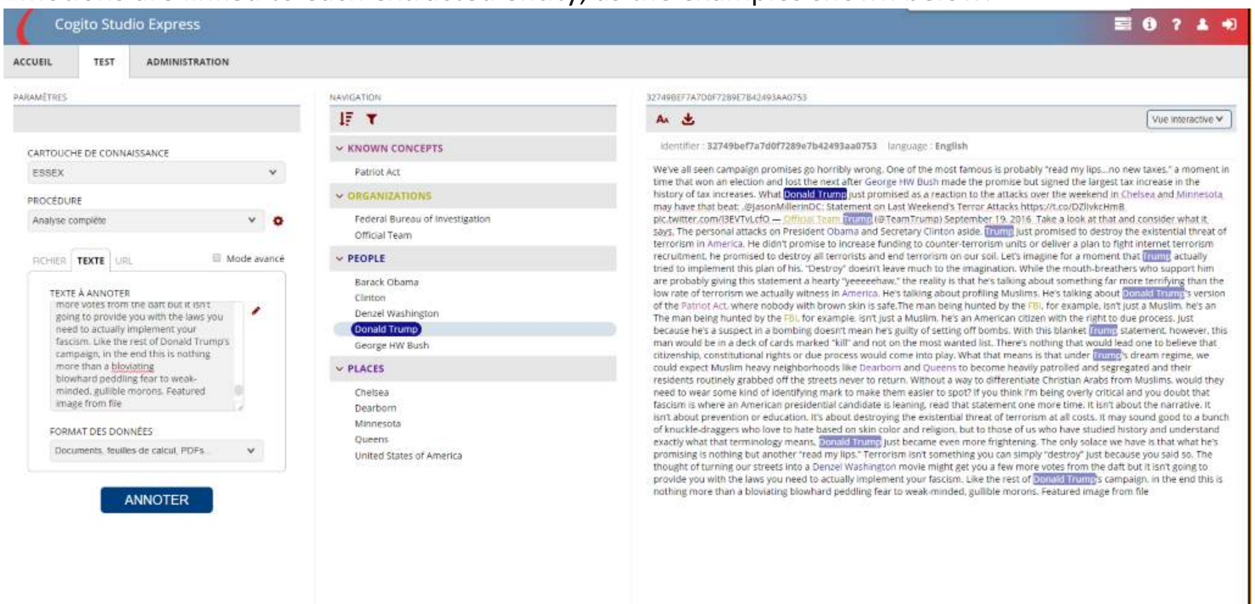


Figure 15- Emotions linked to ESSEX

On the above shot, you can see that Cogito automatically extracted "Donald Trump" from the text. It also extracted entities such as the Patriotic Act or the FBI.

On the below screenshot, you can see that Cogito attaches a Sentiment direction and magnitude to each of the entities extracted. Cogito will also attach a sentiment index to the whole document, giving the general tone of the document.

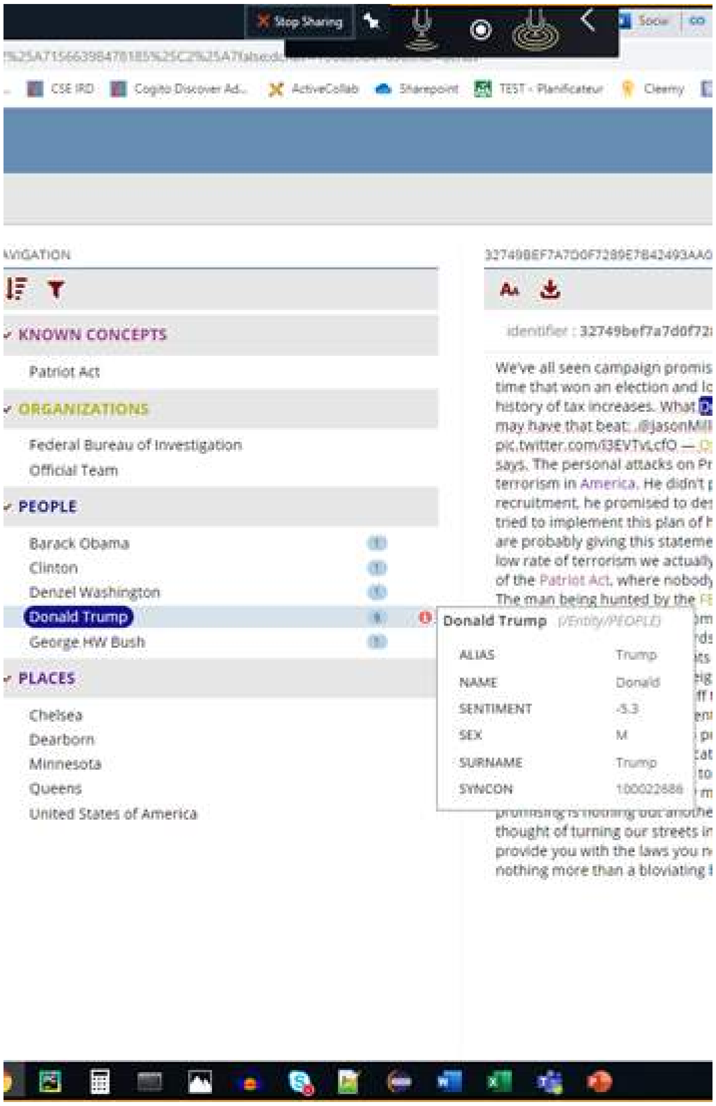


Figure 16- Sentiment attached to each entity

The sentiment attached to Donald Trump is negative, it means in the text the author expresses a negative point of view towards the actual US president.

3.2.2.2 *Useful to compare feelings in text based after pattern FAKE NEWS vs TRUE NEWS*

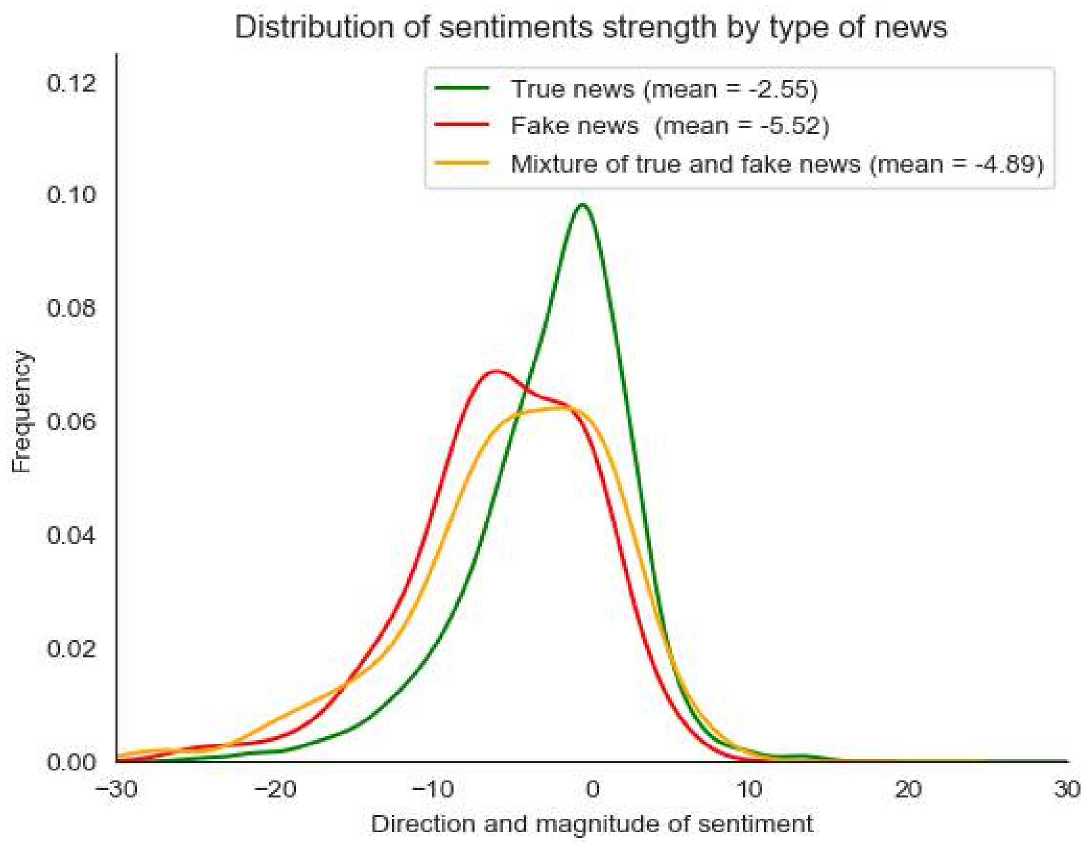


Figure 17- Distribution of sentiments strength by type of news

The above graph is showing us that the "fake news" documents have an average of deeper emotional moments than "true news". However, we still need to put things in perspective because we only have 87 documents targeted as "fake news" in our corpus, and we also have to take in consideration that the document distribution may vary.

3.2.2.3 *The problem of quotations un the document*

Some articles of the dataset can quote a large part of a public figure's speech. In this case, we have to make a difference between the wording of the author of the article and the public figure that is being quoted. For instance, the paragraph:

Democratic presidential nominee Hillary Clinton spoke out today against police violence. "This horrible shooting again. How many times do we have to see this in our country?"

Would lead to a misinterpretation of the author's intentions by our tools.

With the Expert System tools, we can create personalized rules which will be able to understand when someone is quoted and when someone actually speaks his mind.

3.2.2.4 *Objective: checking if a hypothesis saying that the semantic analysis contributes to the "fake news" detection process*

Our analysis found a mean sentiment of –5.52 for Fake News vs –2.55 for True News with a p-value of 0.000017. $P < 0.05$, so we can our hypothesis stating that emotions conveyed in Fake News differ from those conveyed in True News is valid.

To summarize this part, the below figure explicit the metric behind the signal 2 and how it is related to the signal 3. In the remaining part, we will go further in the explanation of the signal 3 and how its value add sense to our global method.

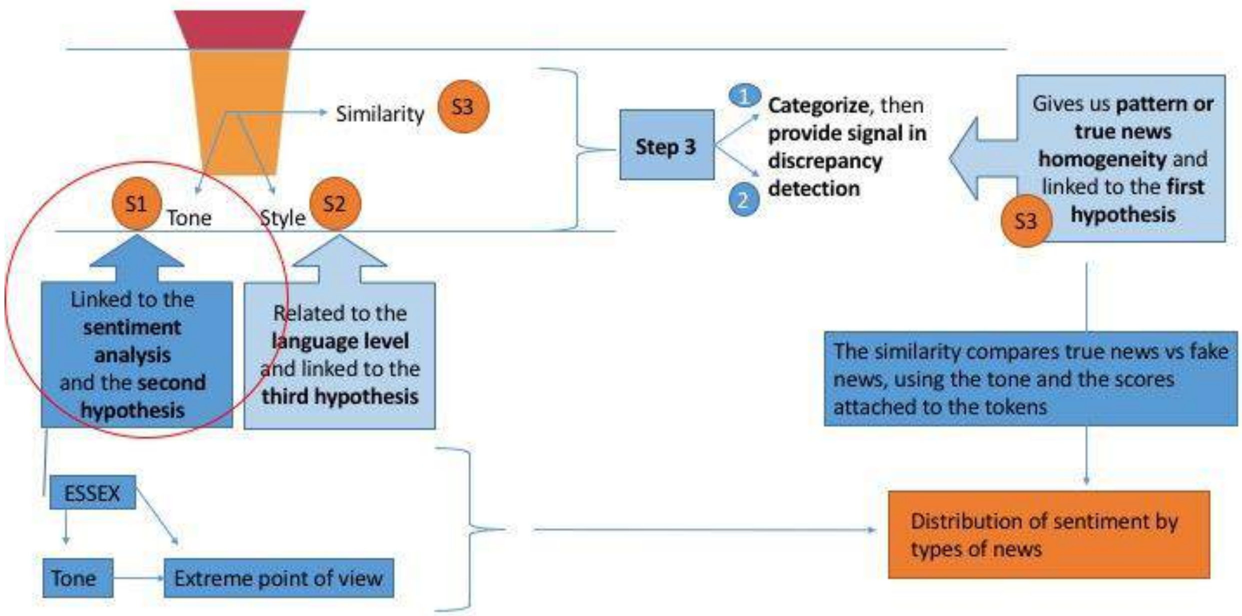
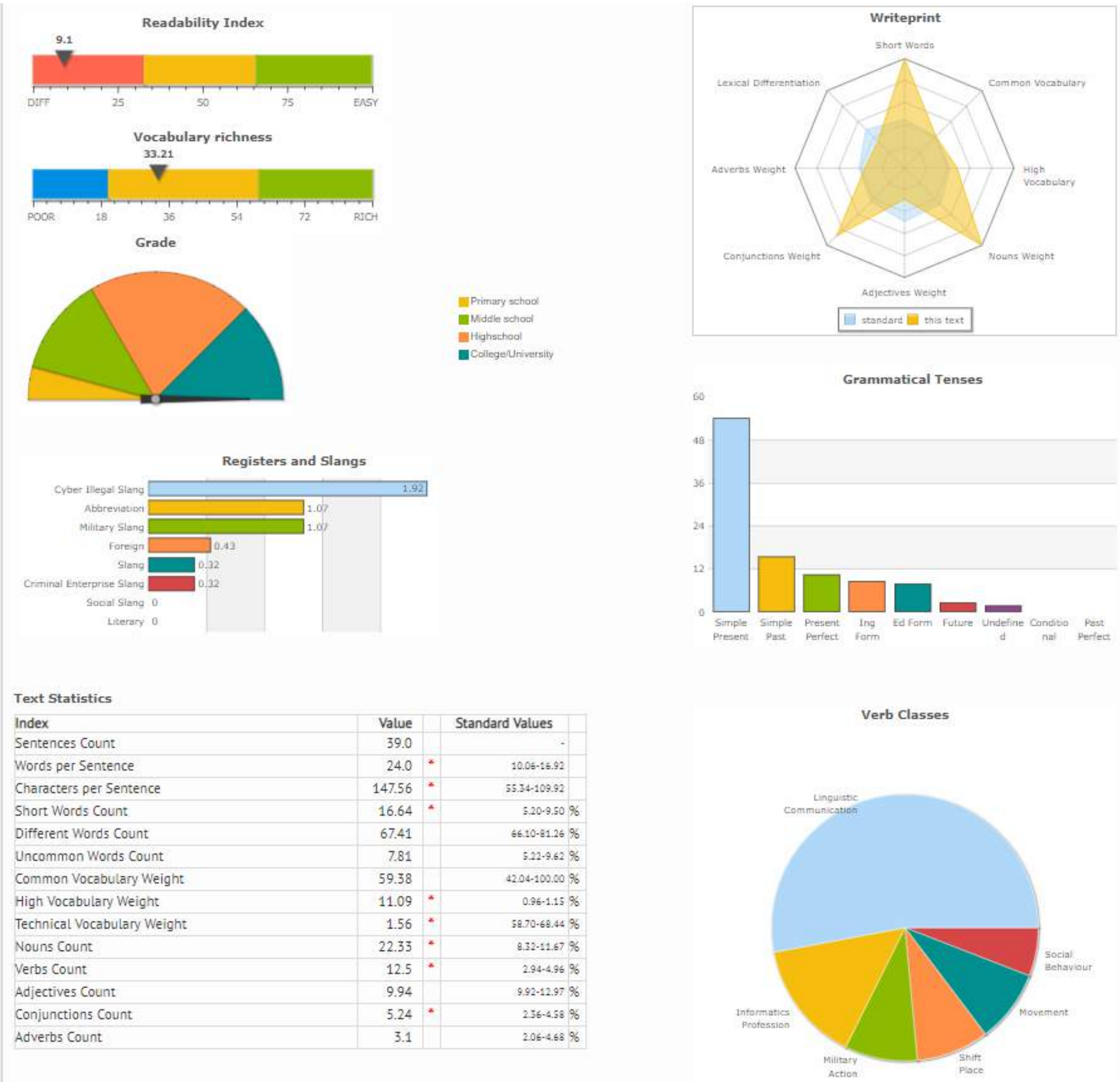


Figure 18- Implementation of the Expert System tools, step 3

3.3 CI API

3.3.1 Presentation of the different stylistic mark: language register, emotions, linguistic level (Signal 2)

The Writeprint service performs is a stylometric analysis of the document, which ranges from readability and vocabulary richness, to verb types and tenses, registers, document structure and grammar. Stylometric data is provided in the shape of indices which, as a whole, make up for a complete fingerprint of the document – that is, a "writeprint". By comparing a number of documents on the basis of their writeprint, author invariants are highlighted by this powerful tool for authorship analysis.



Index	Value		Standard Values	
Sentences Count	39.0		-	
Words per Sentence	24.0	*	10.06-16.92	
Characters per Sentence	147.56	*	55.34-109.92	
Short Words Count	16.64	*	5.20-9.50	%
Different Words Count	67.41		66.10-81.26	%
Uncommon Words Count	7.81		5.22-9.62	%
Common Vocabulary Weight	59.38		42.04-100.00	%
High Vocabulary Weight	11.09	*	0.96-1.15	%
Technical Vocabulary Weight	1.56	*	58.70-68.44	%
Nouns Count	22.33	*	8.32-11.67	%
Verbs Count	12.5	*	2.94-4.96	%
Adjectives Count	9.94		9.92-12.97	%
Conjunctions Count	5.24	*	2.36-4.58	%
Adverbs Count	3.1		2.06-4.68	%

Figure 19- Different stylistics marks analyzed with CI API

As shown in the figure, in order to understand authorship analysis, the Writeprint platform looks for specific information: the readability index, the vocabulary richness, the grade, the use of slang, the verb classes and the grammatical tenses.

3.3.2 Fact-mining

The Fact Mining service provides extraction of Entities, Tags and Domain-specific Entities within the sentences related to a specific fact's taxonomy. In other words, the Fact Mining service performs Text Mining of Entities and Tags within specific text sections, thus extracting connections between relevant entities and specific contexts/domains.

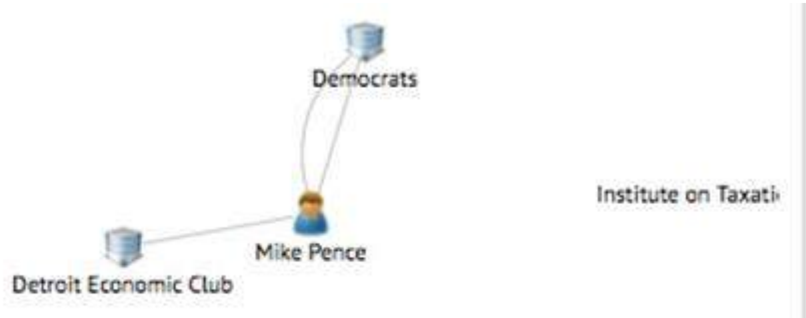


Figure 20- Extraction of fact-mining

As shown in the document above, an article about the actual US President Donald Trump, the tool CI API is able to distinguish the main characters mentioned. If we look for more details:

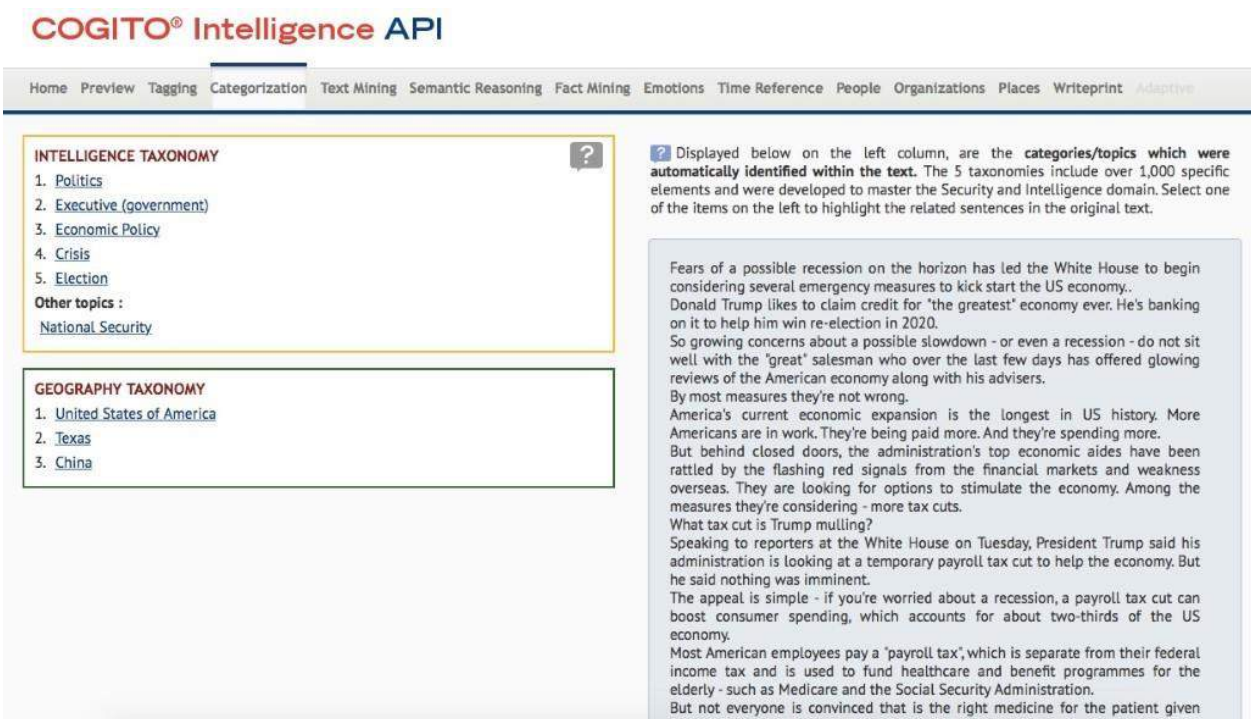


Figure 21- CI API platform

As mentioned in the document, displayed above on the left column, are the **categories/topics which were automatically identified within the text.** The 5 taxonomies include over 1,000 specific elements and were developed to master the Security and Intelligence domain. This tool goes deeper and gives a more specific output.

3.3.3 Stylographic analysis (Signal 2, the writing style)

3.3.3.1 *Vocabulary use*

We can see significant patterns in the vocabulary use within the dataset; the vocabulary of “Hatred” and “Offence” are dominant in Fake and mixed News whereas “Success”, “Hatred”, and “No Emotions” are dominant in True News:

Below are graphs summarizing the findings:

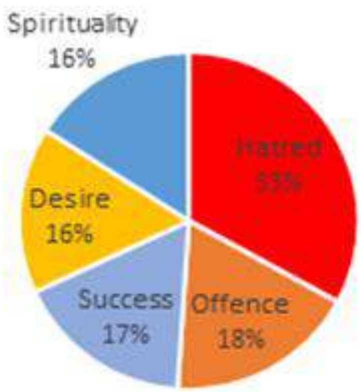


Figure 22- Vocabulary of fake news

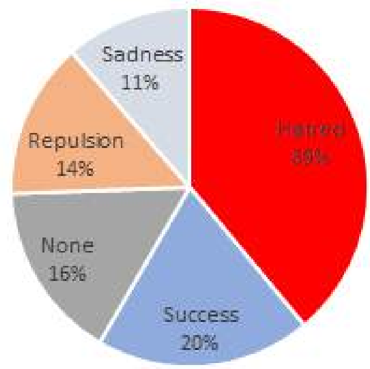


Figure 23- Vocabulary of mixed true and false news

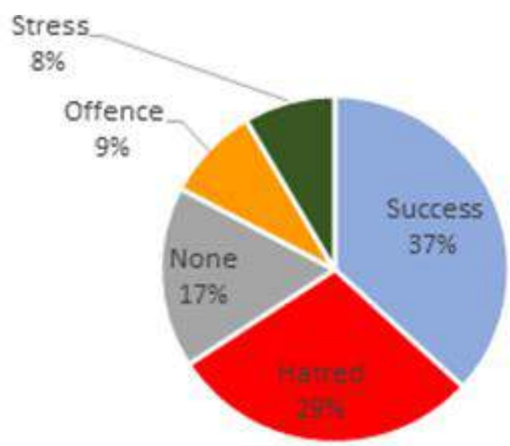


Figure 24- Vocabulary of true news

We can see that none of the topics such as spirituality or desire appear in the true news pattern, however the vocabulary used in the true news do appear in the fake news, like the one related to success, hatred and offence. This is linked to one of our hypothesis, even if the writer is borrowing some linguistic style and thus making his article look like a professional one, another level of analyze must be provided.

3.3.3.2 *Readability index*

Readability is the ease with which a reader can understand a written text. In natural language, the readability of text depends on its content (the complexity of its syntax and vocabulary), its presentation (font size, line height and line length). It is and we are using it because the readability Index is an indicator of the complexity of a document. It takes the indicators mentioned to build a more general readability value. Readability also has an impact on the reader. For readers with poor reading comprehension, raising the readability from mediocre to good can make the difference between success and failure of its communication goal.

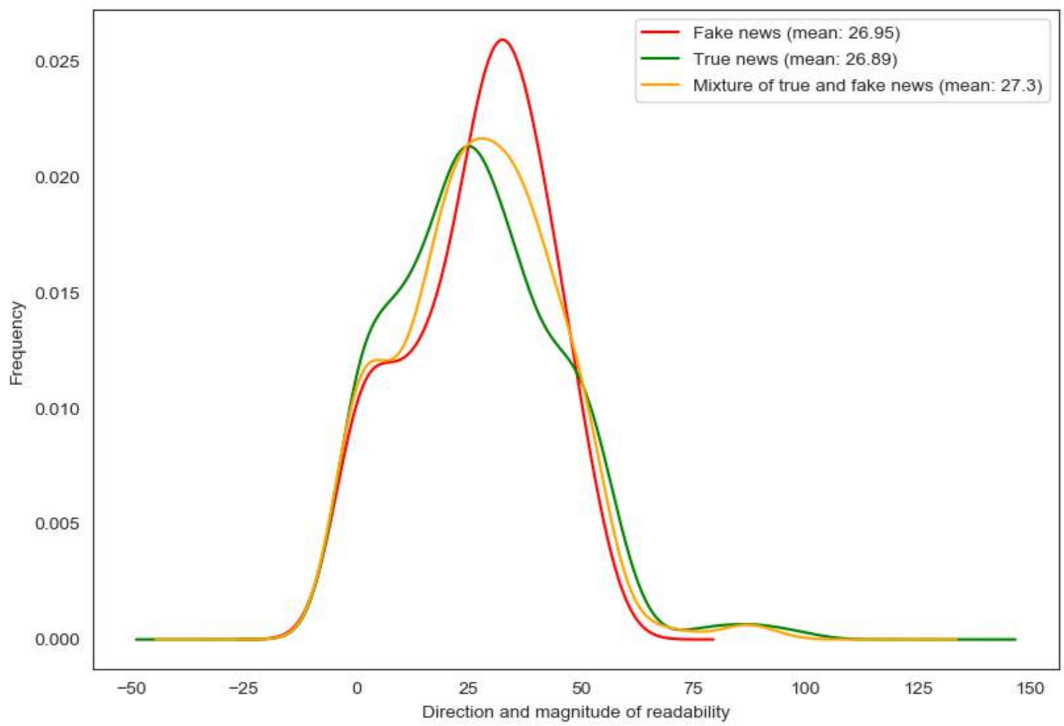


Figure 25- Readability graph

For Y axis, we can see the frequency of the data and for X axis, we can see the direction and magnitude of readability. From what we can see, there is almost no difference between the types analyzed.

The correlation between veracity and readability of the information is not significant on this dataset. It means that the readability slightly differs, but the difference is not statistically significant (p-value = 0.9). Our hypothesis is that Fake News differ from True News in the way that they utilize a different level of language (easier to read, more straightforward) is not valid on this dataset.

3.3.3.3 *Vocabulary richness*

The vocabulary richness, on the other hand, is slightly different from fake news (especially mixed true/false news) to true news.

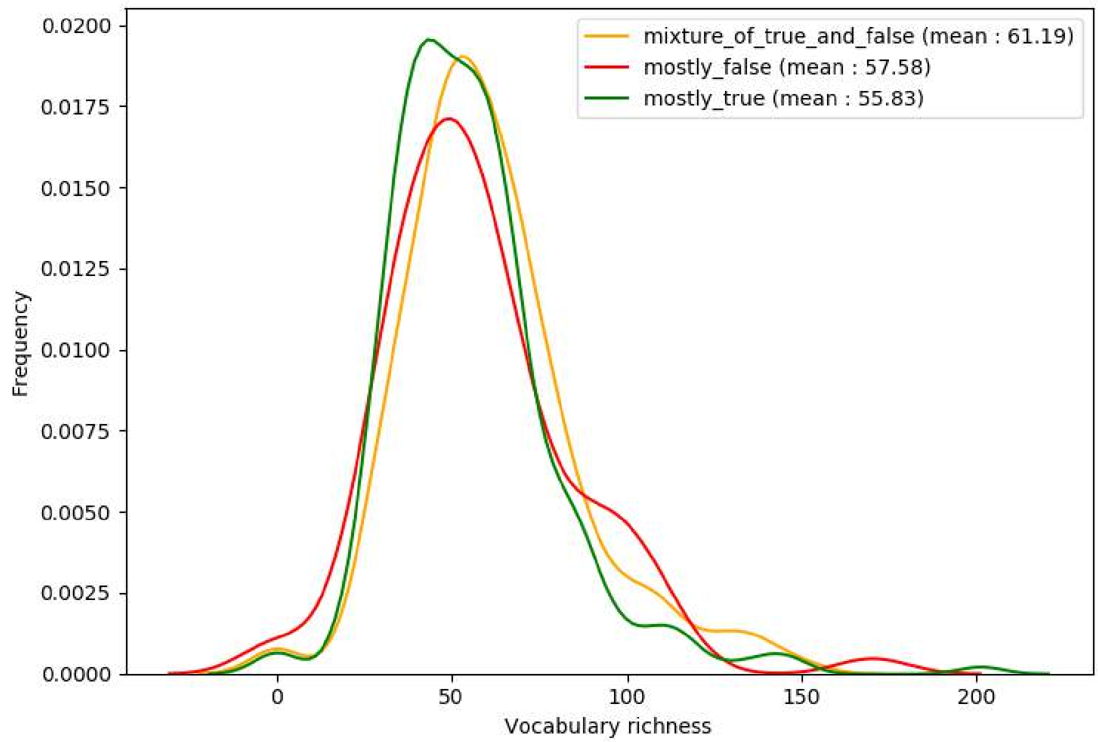


Figure 26- Vocabulary richness graph

Finally, on the 2 variables plot you can see that the difference between vocabulary richness (Y axis) is higher than the difference between readability index (X axis).

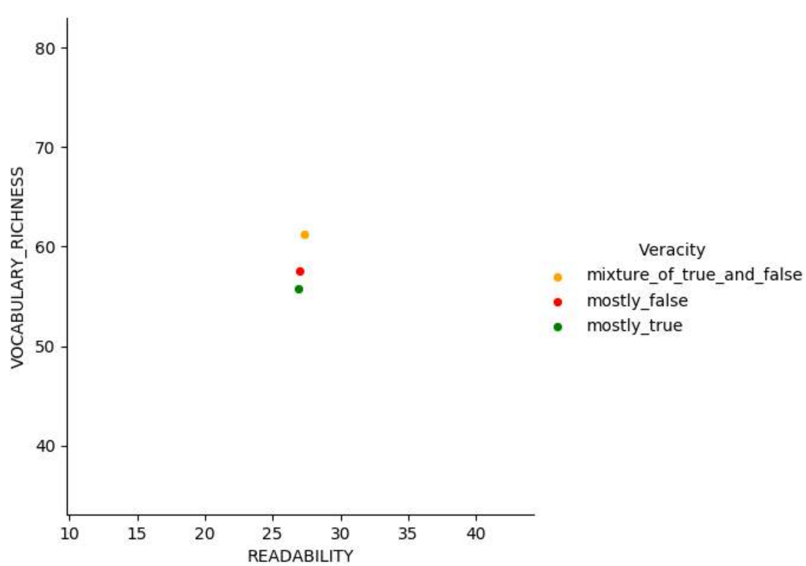


Figure 27- Correlation between vocabulary richness and readibility

To summarize the last part of the document we have recapitulated all the steps with the tools combined mentioned in the third part, in the figure below:

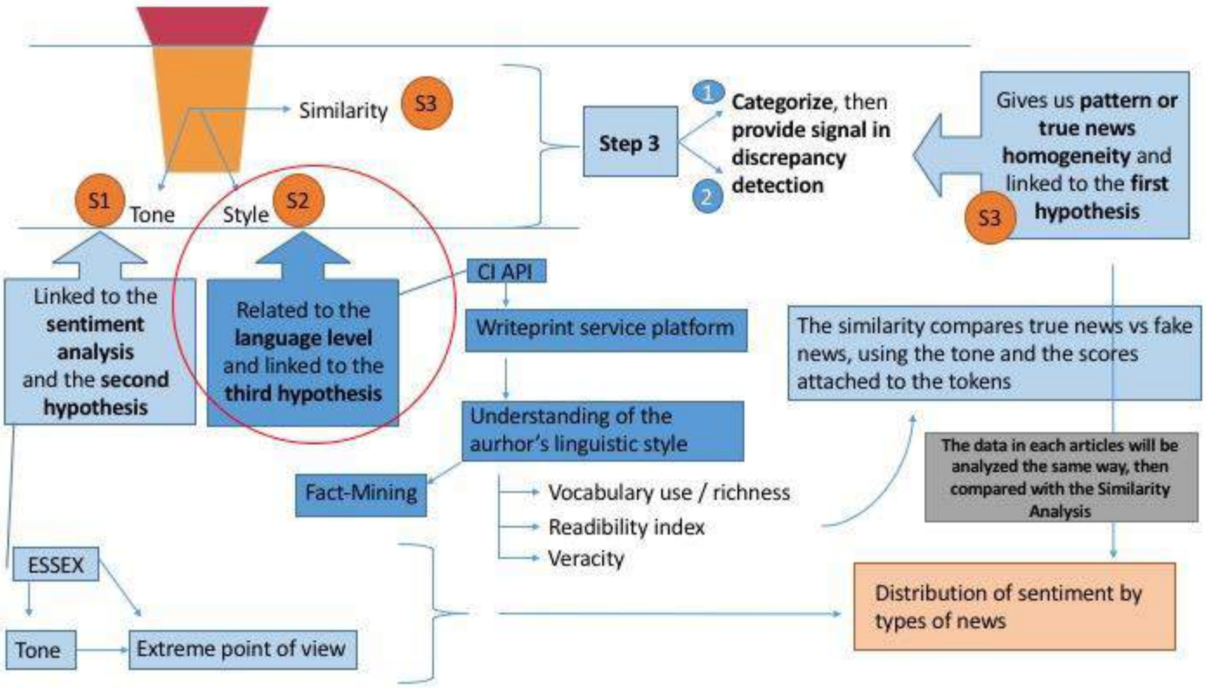


Figure 28- Implementation of the Expert System tools, step 4

3.4 Similarity analysis between same category documents

3.4.1 Use of similarity (Signal 3)

The Similarity analysis of Cogito can detect documents that are very different from others.

For example, this screenshot below:

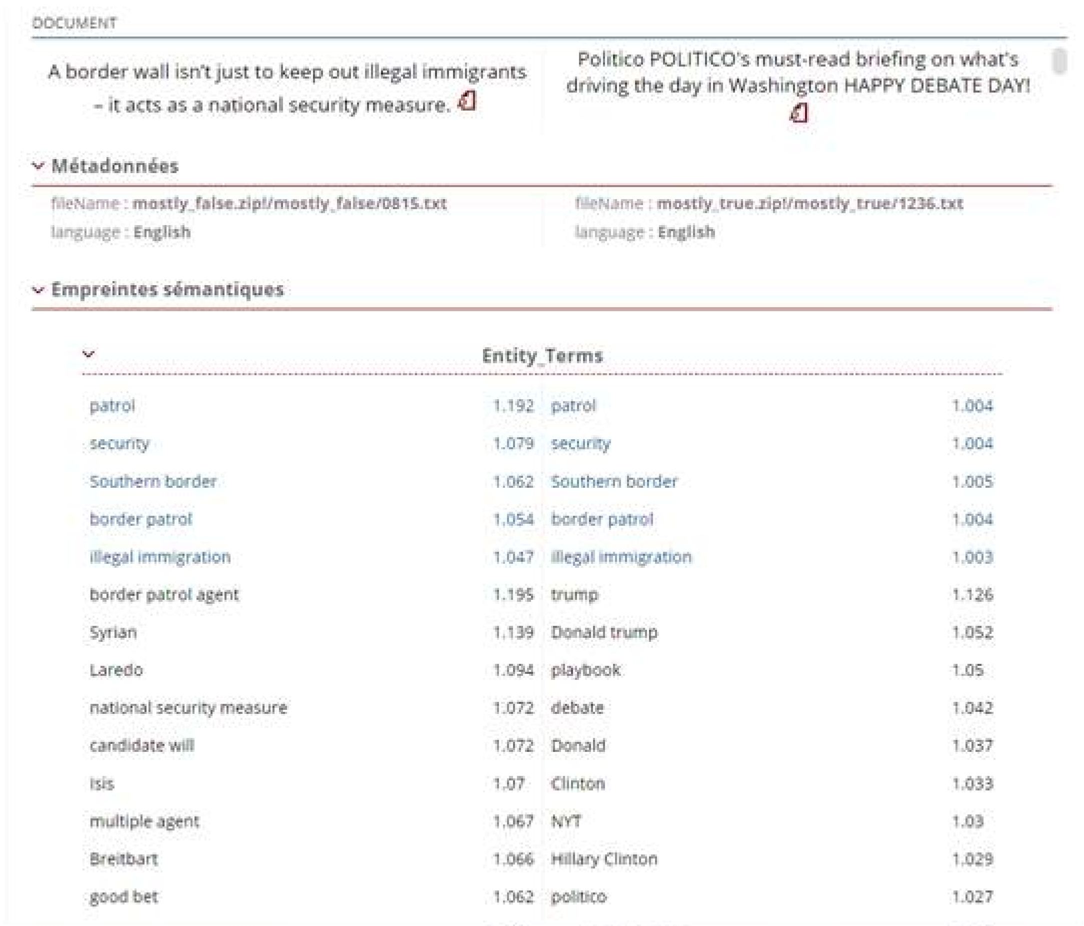


Figure 29- Cogito similarity analysis

It shows a similarity analysis between a True document and a Fake document.

The similarity between these two documents is only 5%, although they speak about the same topic. We can see that the documents talk about the same subject in general ("border patrol", "southern border", "illegal immigration" ...). However, we can see semantic discrepancies between the expressions used in one document compared to the other. For example: "Laredo", "Isis", "Breitbart" ... for the Fake document vs. "playbook", "NYT", "Trump" ... For the True document.

On this shot of a comparison between two True news:

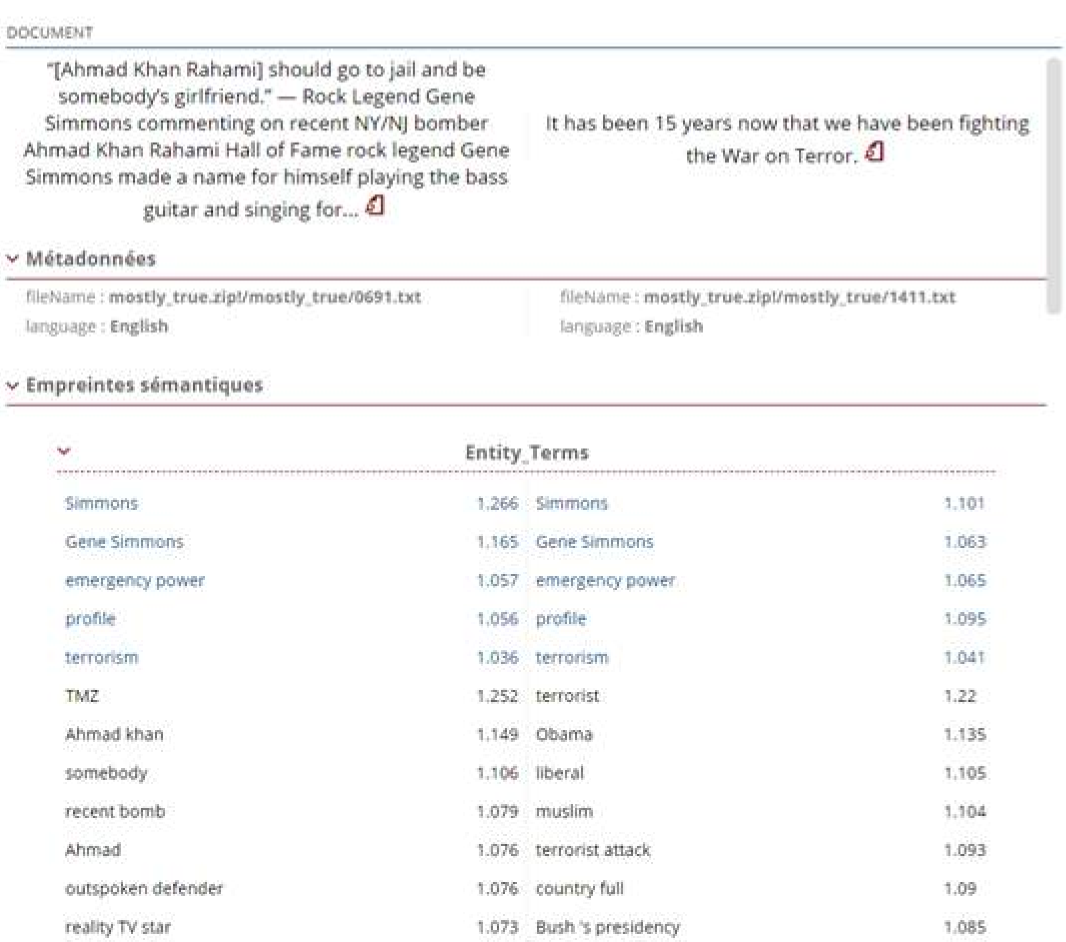


Figure 30- Comparison between two "true news"

We can see that the similarity between the two documents is 13%.

3.4.2 Elaboration of statistical pattern to detect outliers on different signals

In the context of this project, outliers are documents that deviate far from the expected value in terms of Writeprint and emotions. Our model would raise flags on documents that deviate too far from the mean emotions value of True News.

We have summarized the project in its globality in the figure below:

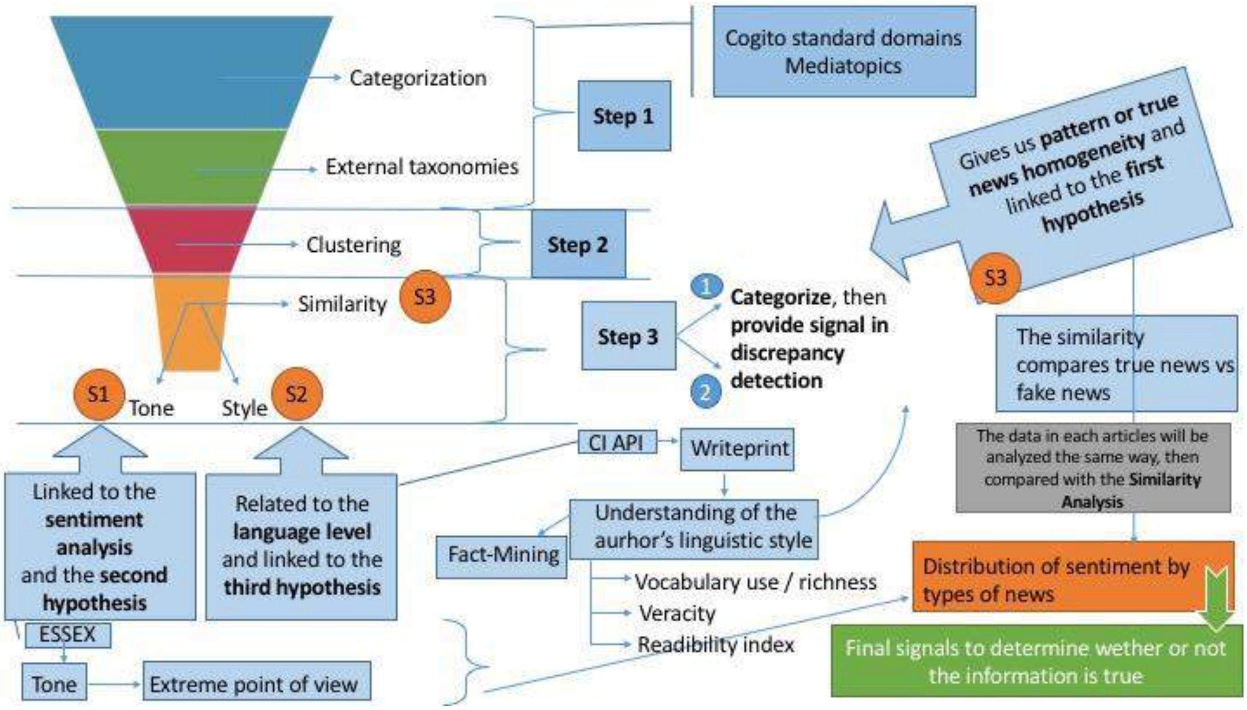


Figure 31- Implementation of the Expert System tools, step 5

4. Access to the Cogito tools and methods

Expert System will use the RESTful model. As said before, the tools used by Expert System are:

- ESSEX (accessible via the Cogito Discover REST API) for semantic analysis
- Similarity (accessible via the Cogito Discover REST API) for similarity analysis
- Cogito Intelligence API (accessible via its own REST API) for stylometric and writeprint analysis

Each of these tools will be available via REST APIs. They can be used either via on-premise installations or via Expert System's webservers.

4.1 Cogito Discover REST API

ESSEX via Cogito Discover can be queried either via the REST API. The ESSEX semantic analysis, the sentiment analysis, as well as the similarity analysis will be launched via the Discover REST API. Expert System will provide the complete documentation of the Discover REST API as part of this WP. The documentation of the Cogito Discover REST API consists of a Swagger online Documentation, a Postman collection & Postman environment as well as a pdf file that will allow other technical teams to effectively use the REST API.

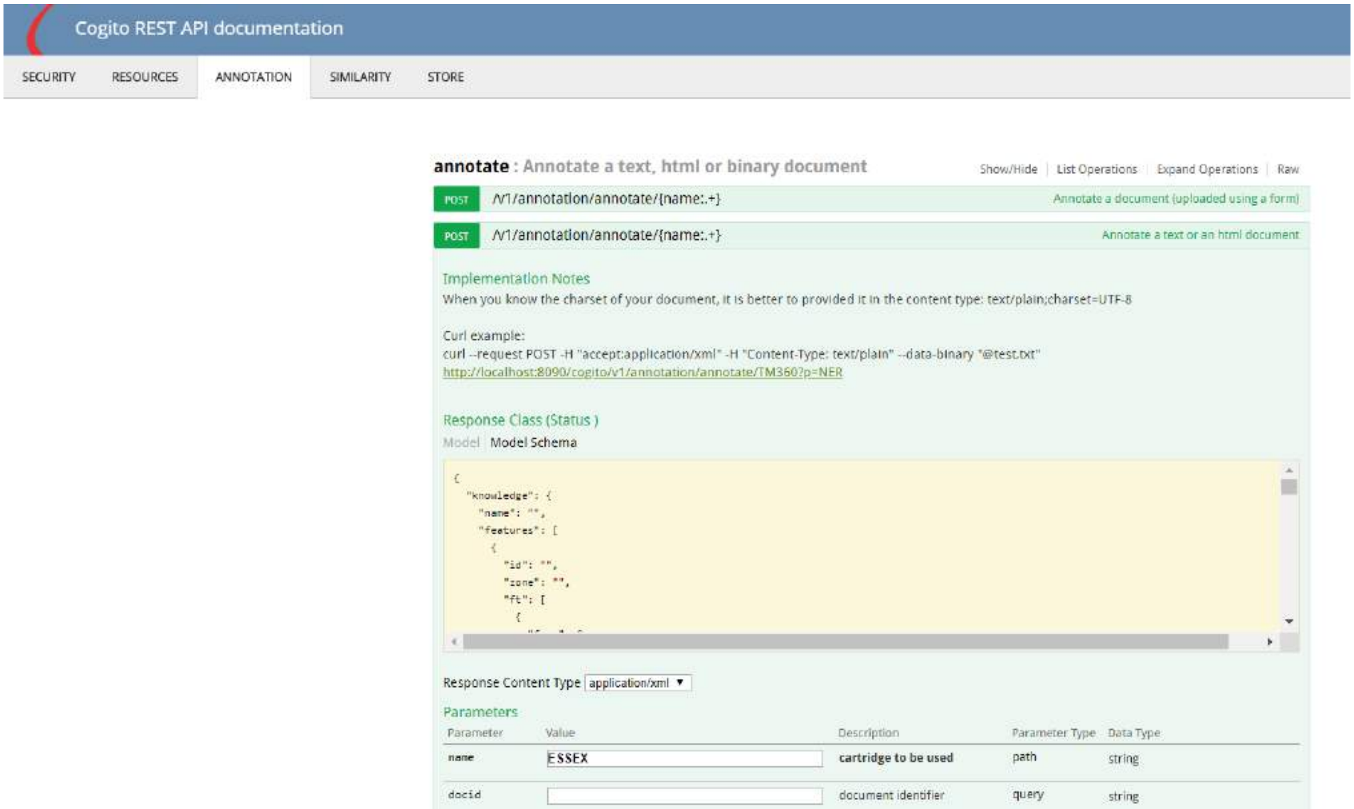


Figure 32- Cogito Discover REST API Swagger documentation

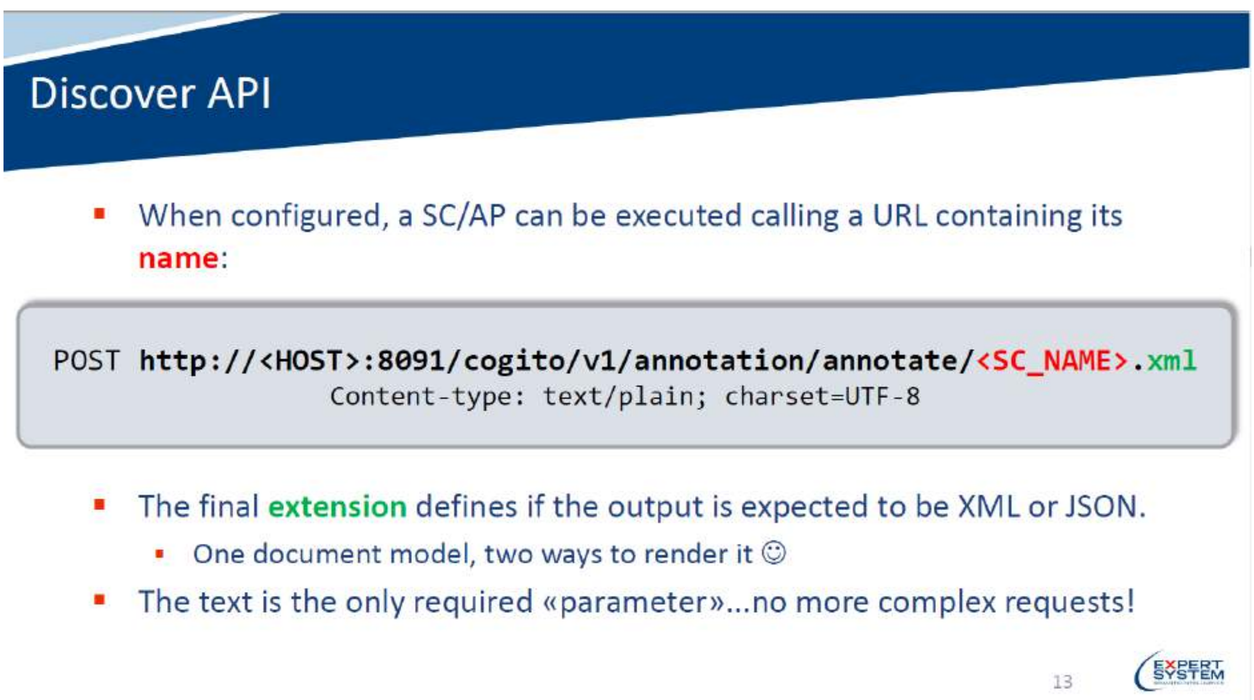


Figure 33- Cogito Discover REST API documentation

4.2 Cogito Intelligence REST API

The Cogito Intelligence API endpoints for stylometric and writeprint analysis differ slightly from the Cogito Discover REST API endpoints and need a different documentation that will be provided to the consortium as well. This documentation will consist in two pdf files as well as a Postman collection & Postman environment.

Cogito Intelligence API
Developers' Guidelines

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

WEB SERVICE - REQUEST

Requests to Cogito Intelligence API are nothing but HTTP-POST requests, containing a String payload in JSON format. The JSON-formatted payload must contain two required String parameters: the text to be analyzed and the API key received upon registration. An example follows of a typical analysis request to Cogito Intelligence API in its RESTful version, request headers and payload content for a typical analysis request:

- HTTP headers (example):

```
POST http://www.intelligenceapi.com/CogitoIntelligenceAPIrest/services/tm_dme HTTP/1.1
Content-Type: application/json;charset=UTF-8
Accept: application/json
Host: www.intelligenceapi.com
```

- JSON Payload = request content (example):

The request content is made up of the DOCUMENT, the API key provided upon subscription, and an optional METADATA structured field. The request must be included within the HTTP message payload.


```
{
"DOCUMENT" :
        "Regardless of what President Goodluck Jonathan's government would have us believe, we are losing the war against Boko Haram and its leader Abubakar Shekau. We're losing it to the original Boko Haram and to its various franchises, including those in Jonathan's government\nDays after the chief of defence staff, Air Marshal Alex Badeh, took over in January, he vowed to end the Boko Haram onslaught by April 2014. He had barely finished speaking when gunmen struck, killing 74 people in separate attacks in Borno and Adamawa states. Badeh ate the humble pie and promptly disavowed setting any deadline to end the killings.\nIn the last two weeks, gunmen presumed to be Boko Haram have killed over 200 people, including children and teenagers, with over 30 killed on Tuesday.",
"KEY" :
        "ba7ed852f700000288e22b20000007ed0000000852f74dfcb288e00000",
"METADATA" : {
        "author" : "Azubuike Ishiekwene",
        "source" : "http://www.leadership.ng"
```


Figure 34- Cogito Intelligence API documentation

The consortium partners will be able to launch stylometric analysis of any document thanks to the Cogito Intelligence API. Expert System also provided a link to a visual demonstration and test platform : https://www.intelligenceapi.com/demo/ that is already accessible by the members of the consortium.

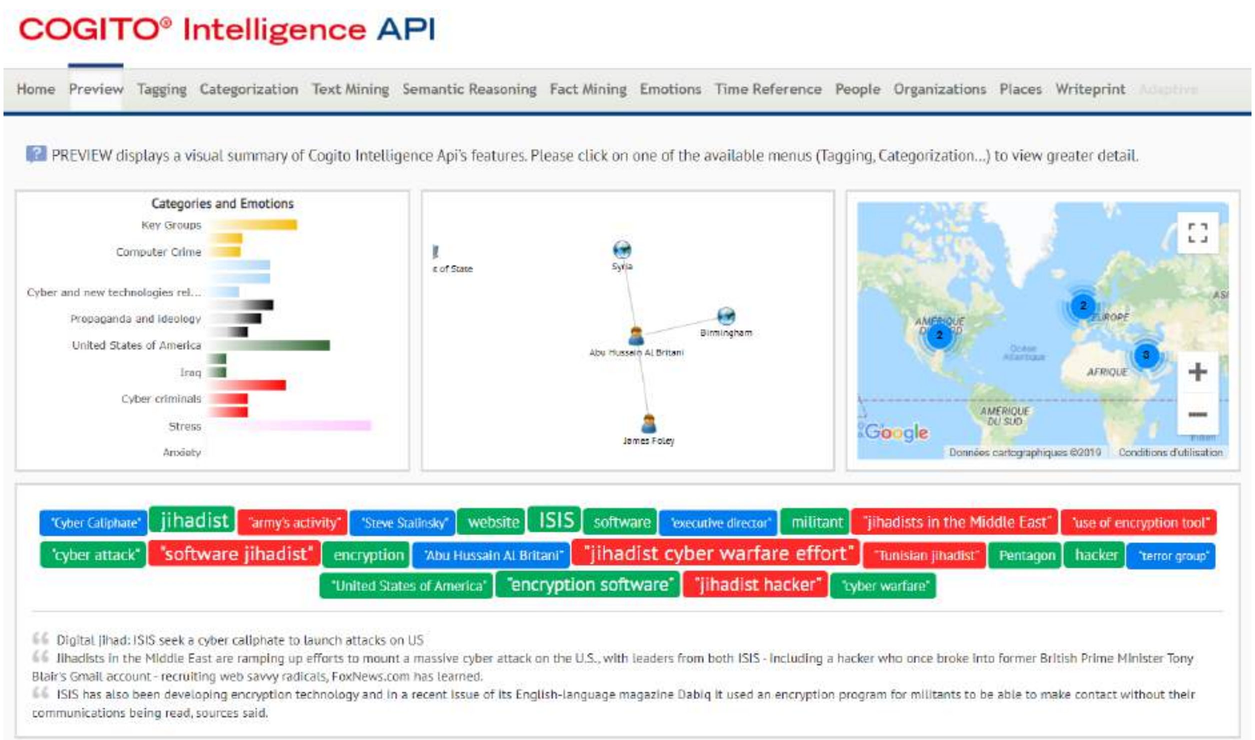


Figure 35- Demonstration interface for Cogito Intelligence API

4.3 A common data format for transfer and manipulation: Cogx

A data format for all features extracted by Cogito is needed. Cogx is an XML-like format for Cogito data manipulation. For a given document, the Cogx records all of its features : categories, sentiments, writeprint, etc…

Cogx comes with a Java library for accessing its different branches. It can also be accessed via Xpath queries.

A complete documentation of the Cogx format will also be provided to the partners in the consortium. This documentation will consist in a pdf file and a Javadoc.

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

Expert System Cogito eXchange Format

DATA MODEL

The documents to be analyzed with Expert System solutions may be in different formats and have different content. Text mining solutions should have a plain text representation of these documents available before applying automatic information extraction and category analysis. Furthermore, some descriptive information - such as the document metadata, which might for example state the source, author and date - can be useful when you want to present these documents or to update those that are likely to change often.

Some Expert System solutions add information to documents in order to provide them with more analytical descriptions and to prepare them for other applications. For example, when an extraction process is applied to a corpus, each document is associated with a set of features. Storing these features along with the document representation enables future classification.

For these reasons, Expert System has chosen to represent documents as sets of fields. These fields directly give birth to tags of the XML format.

DOCUMENT FIELDS

Documents are described through three types of fields: data fields, mining fields and operational fields. They contain data extracted from or related to the content of the document, as well as operational information related to the processing chain. Some of these fields can be empty.

- Data fields: contain information directly extracted from the original document. They are normalized to facilitate the comparison of documents from different sources. These data fields include Publication date, Authors, Title, Source, etc. data field values are found between the `<metadata>` tags.

- Mining fields: contain information generated by text mining operations. They are created by Expert System text mining processes and store various types of features available for further online analysis.

- Operational fields: contain information related to processing and to the management of the document within the system, for example, the last update date, lists of warnings and error messages, information of whether documents were properly annotated or not, and so on. These values are to be found in the `<status>` tag.

Documents can actually be viewed as wrappers for the fields mentioned above. Each document always has an identifier attribute and can also have additional attributes like a URI (Uniform Resource Identifier) that references the source.

Documents may have been logically divided into zones. Some fields may then refer to a particular zone. For example, a patent reference can have zones like an abstract, a body or a bibliography. Abstract and body have a plain text representation.

Note: *zone* attributes are only provided for convenience as all text content will be concatenated into one and zones will only be annotations.

The Following sections describe the fields that can be associated to a document. Each field is presented in a table that is associated to its field type (data, mining or operational field). The name of the field gives birth to the corresponding element name in Cogito eXchange format documents.

Note: The fields used or that can be used in Cogito eXchange format documents are not restricted to recommendations of any particular organism.

Data fields

Field name	Description
language	Language of the "*intellectual*" content of the resource. This is the main language of the document. This is a special name for an attribute and will be used by Cogito®. **Example**:

EXPERT SYSTEM

4

Figure 36- Documentation of the Cogx format

5. Conclusion and Next Steps

The main achievement of this deliverable has been to define, to set and to test -with success- the technical methodology for semantic content analysis, and to provide the tools to the partners of the consortium. It also has been to define the common format for all Cogito data, the Cogx. The partners of the project will be able to use and to manipulate these Cogx for further analysis detailed in WP4 and WP5.

Two of our hypotheses have been verified (the ones related to the tone and the vocabulary use), and one has been rejected (regarding the language level). However, those hypotheses have to be tested again because our dataset only contains 9 different publishers and it might not be enough to please the level of accuracy this project aims to achieve. Focusing on the source or on the author might be another signal of relevant contribution to the project. It would thus reduce the risk of the possible mistakes being committed. The use of a thesaurus and ontologies of labialized sources (like for instance Russia Today of Mediapart) could be a good start.

The outcome of this deliverable is a semantic analyzer that aims at providing information that will be the input of the expert meta-verification system. The expert meta-verification system will combine the verification results from the content verification services created in WP3, social, semantic and multimedia content, in order to compute a meta-score that accurately depicts the credibility of the digital content under consideration.