PRESENTATION
SEMIOS SUITE

OneLight

Semios
relecture automatisée des exigences

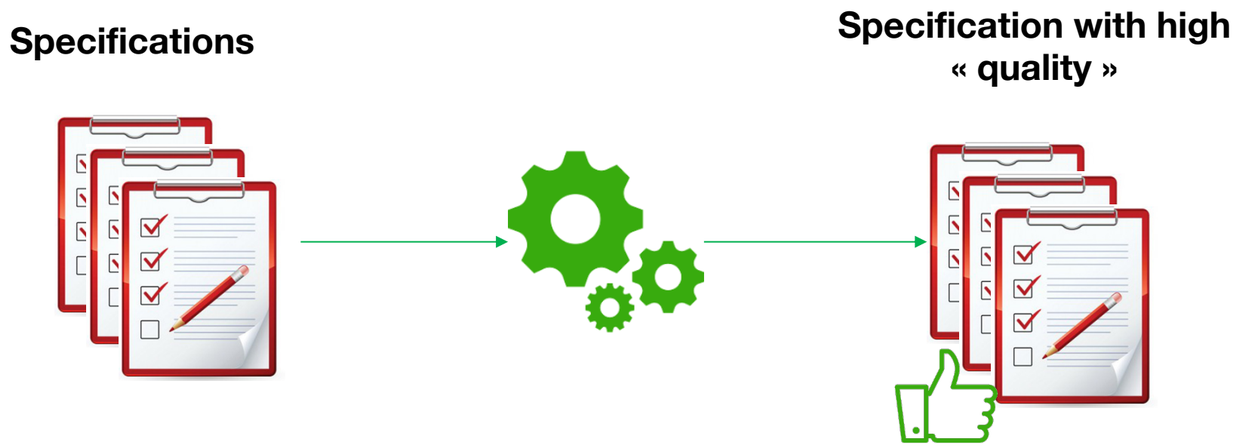
Manel MEZGHANNI
PhD, Data Scientist
Semios Team

Collaboration with: Pr. Florence SEDES

IRIT

GDR GPL- Session GT IE 11 Juin 2019

Industrial Context



- Data from different domains
  - Aeronautic, Automobile, Spatial, Finance, Energy
- How to deal with each domain specification ?
  - Acronyms, Business terms, normes

SUITE SEMIOS

SEMIOS FOR REQUIREMENTS

- A tool for detecting errors in specifications from the conception phase

SEMIOS FOR ACRONYMS

- A tool to detect acronyms based on regular expression

SEMIOS FOR SIMILARITY

- A tool based on IA technology to detect redundancy

Semios

SEMIOS FOR REQUIREMENTS

Presentation of the tool

Based on **Natural Language Processing** techniques (NLP) – **Semantic** analysis

Detects anomalies in your specifications documents at the design stage

Helps to control **requirements quality**

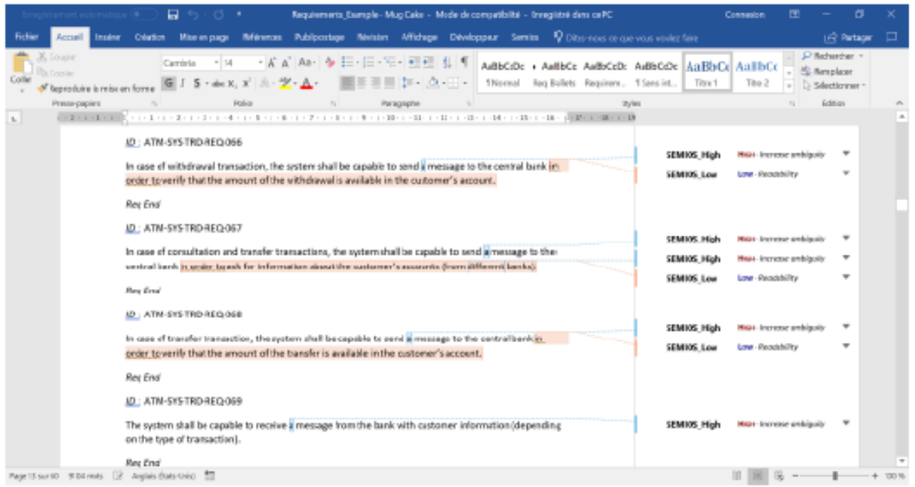
Add-on included in existing tools : Word, Excel, Doors, Doors NG

Reduction of development time

Cost rationalization

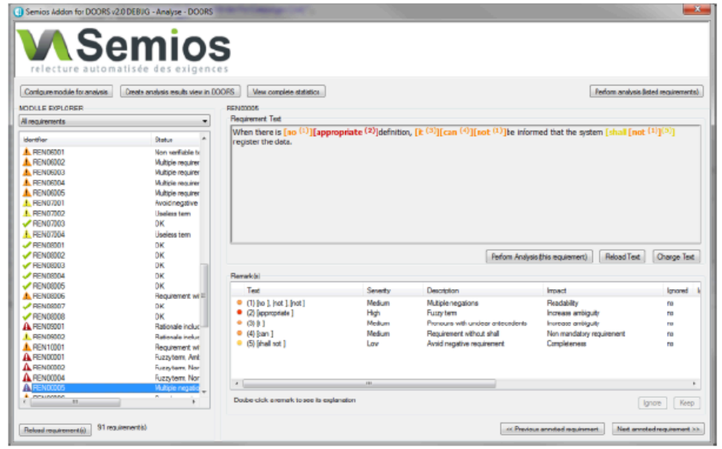
An integrated tool

Word



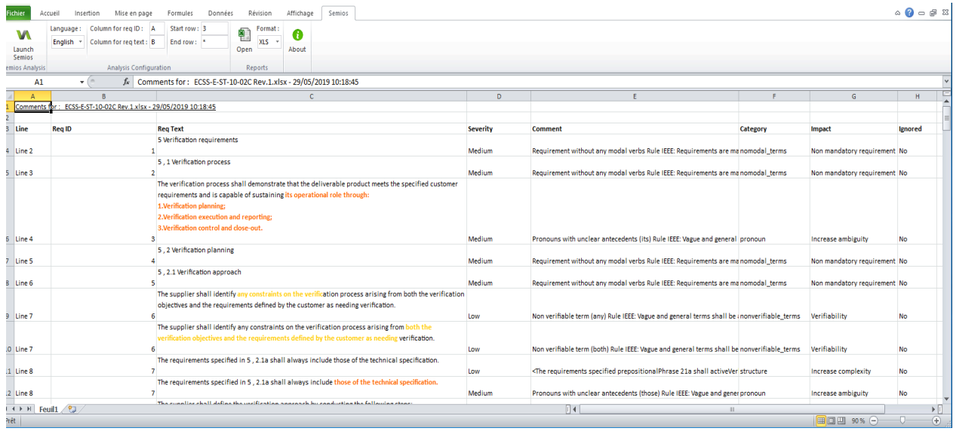
Semios

IBM DOORS



IBM Rational
DOORS Next Generation

X



Semios
an intelligent requirements proofreader

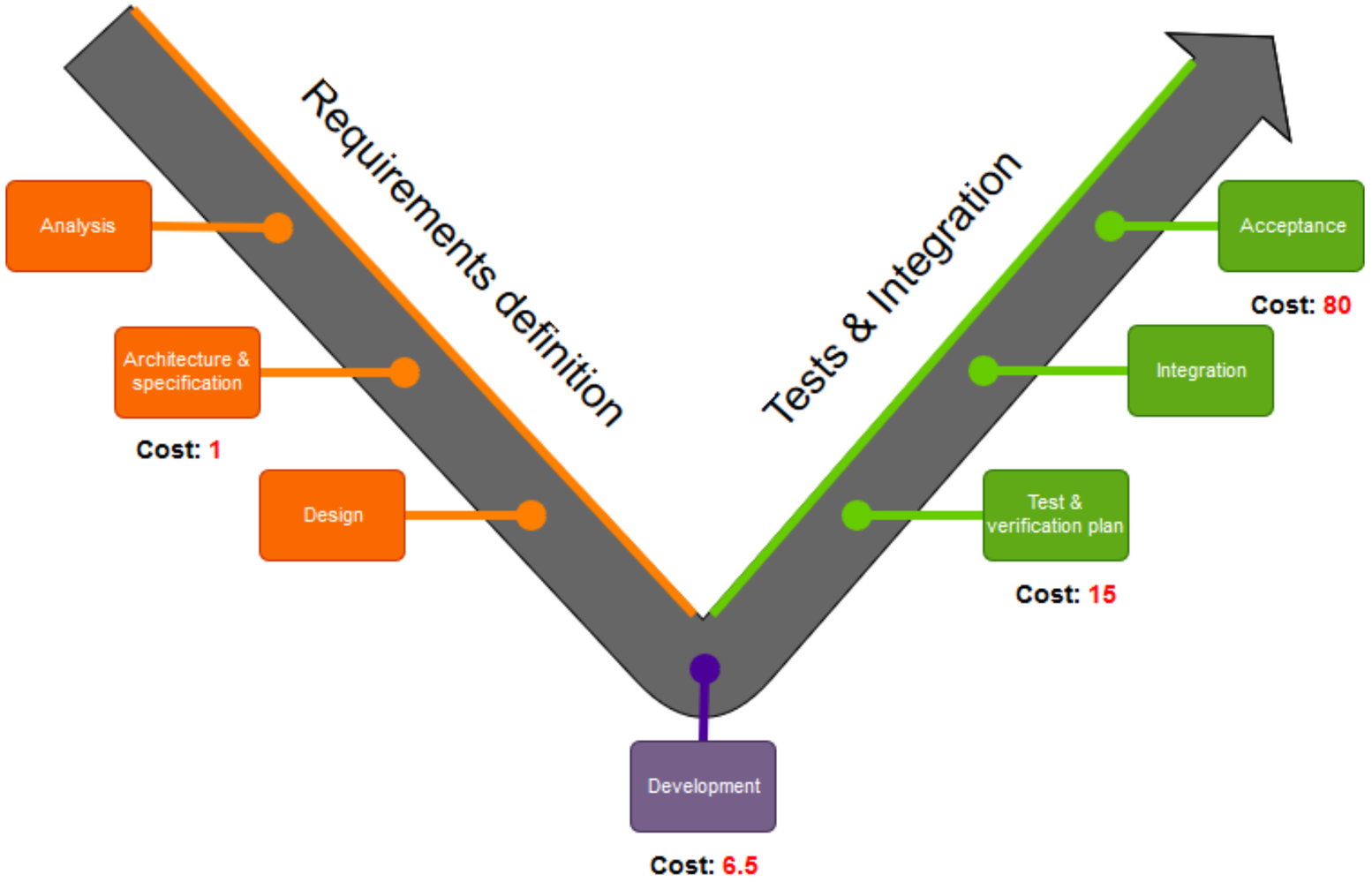
Semios analysis results

The system **[shall [not ²]¹]**accept an envelope if **[it ²]**is **[not ²]**in the right format.

Index	Severity	Impact	Description	
1	Low	Completeness	Avoid negative requirement Rule IEEE: Use positive statements and avoid negative requirements such as shall not.	Ignore
2	High	Increase complexity	Complex negations Rule IEEE: Vague and general terms shall be avoided - Negative statements	Ignore
3	Medium	Increase ambiguity	Pronouns with unclear antecedents Rule IEEE: Vague and general terms shall be avoided - Vague pronouns	Ignore

SEMIOS FOR REQUIREMENTS

How much does a specification error cost depending of the project phase ?



SEMIOS FOR REQUIREMENTS

Quality criteria

IEEE830

Correct
Unambiguous
Complete
Verifiable
Traceable
Consistent
Modifiable
Ranked

ISO29148

Unambiguous
Singular
Consistent
Complete
Feasible
Traceable
Verifiable

AUTRES...

Incose
IREB ARP4754
ASD-STE100

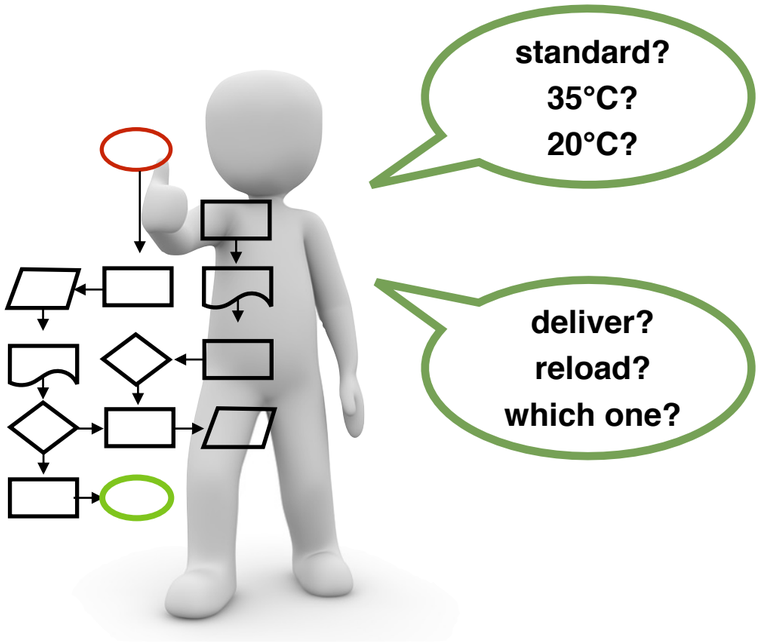
SEMIOS FOR REQUIREMENTS

Non-ambiguity

Risk of misinterpreting during development and test phases

> The maximum pressure loads at the **[standard]** operating temperature shall be 3 bars.

> The system shall deliver data **[or]** reload the configuration checks performed.



SEMIOS FOR REQUIREMENTS

Verifiability

Several tests are needed to cover the requirement, with an additional risk of wrong interpretation when using conjunctions (and, or)

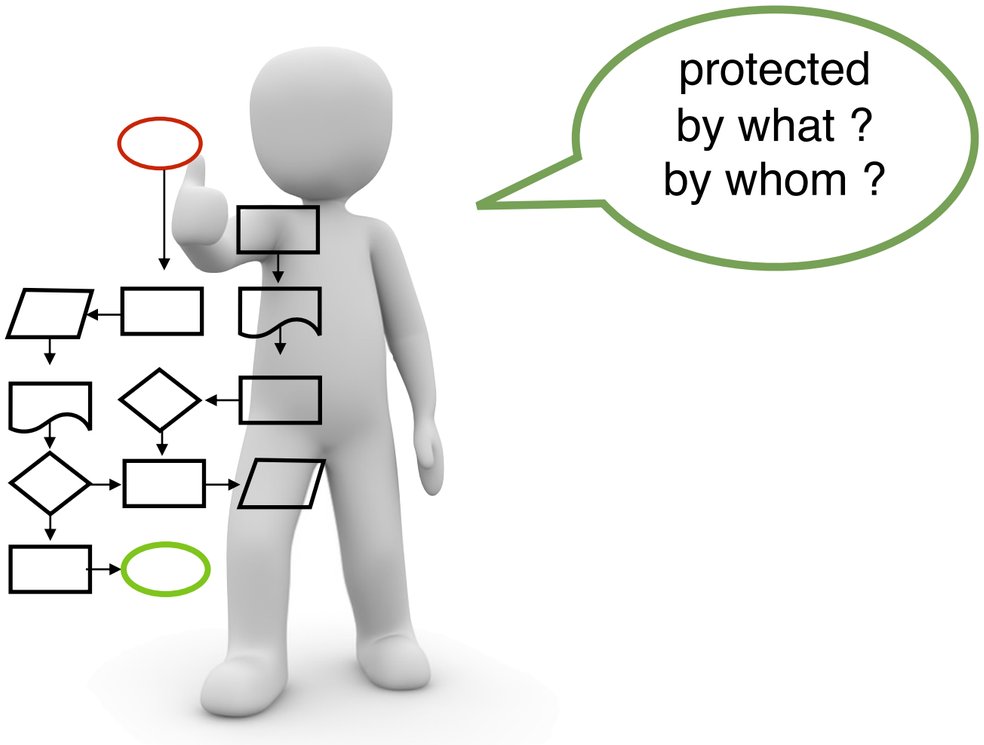
The CPU system **[shall set]** these signals in output or **[shall send]** them directly to the platform and the different supervisor hardware designs **[shall identify]** the software for its adaptation behavior.

SEMIOS FOR REQUIREMENTS

Completeness

Lack of essential information

The system [shall be protected] against over-voltage.

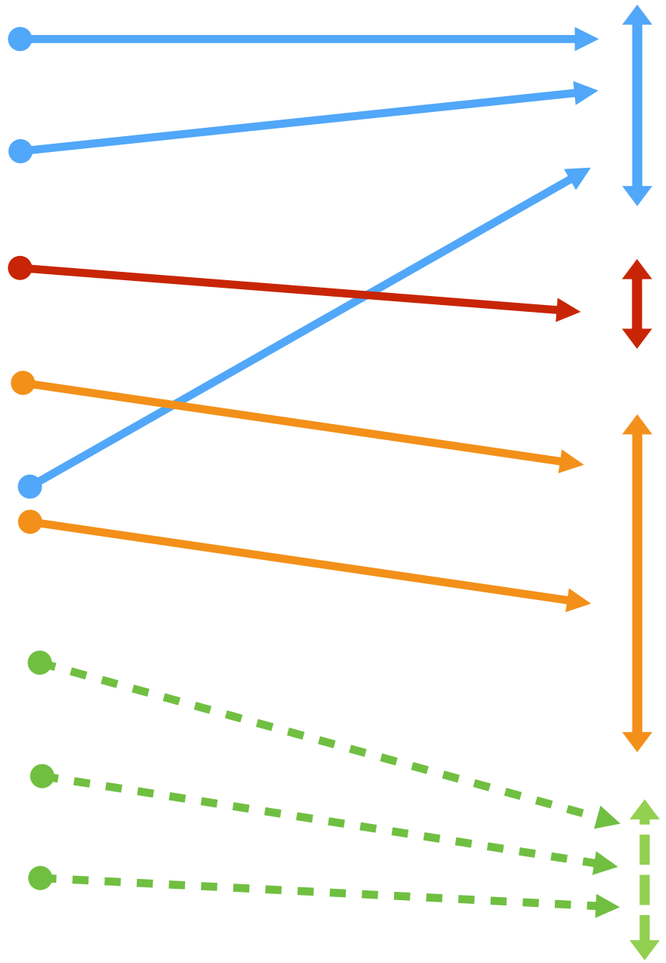


SEMIOS FOR REQUIREMENTS

Summary

Business impacts

- **Ambiguity**
- **Verifiability**
- **Completeness**
- **Readability**
- **Complexity**
- **Under-specification**
- **Necessity**
- **Tolerance**



16 categories of alert →
95 rules for detection and contextualization

Fuzzy term
Unauthorized expression
Complex/ambiguous coordination
Multiple requirements

Passive construction
Incomplete sentence

Justification included in a requirement
Solution/Design sentence
Useless term
Complex/ambiguous negation
Inappropriate structure
Requirements too long
Complex construction

Underspecified term
Non mandatory requirements
Quantities without range

Semios
for Similarity

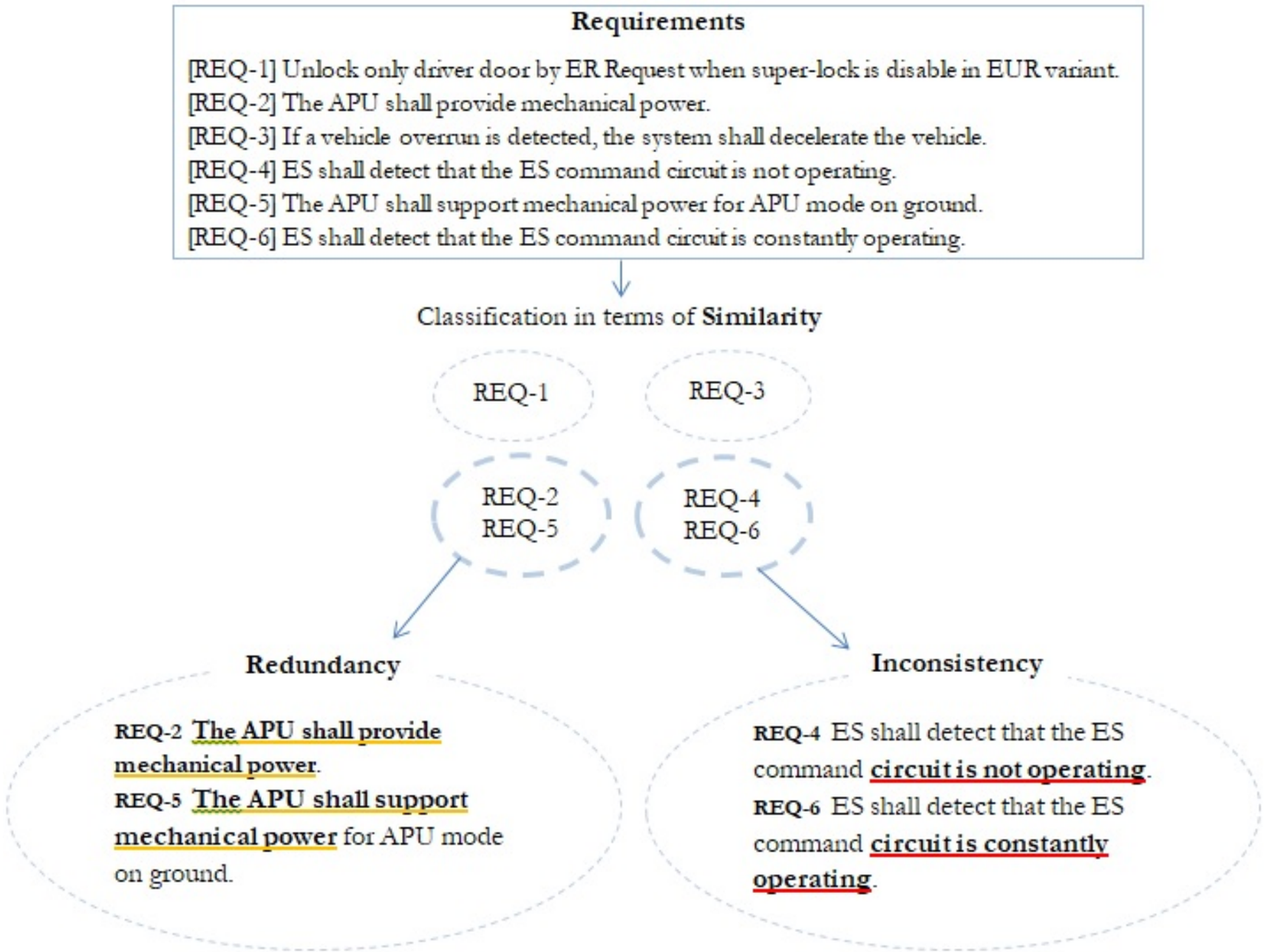
SEMIOS FOR SIMILARITY

Presentation of the tool

- Result from the Project CLE-ELENAA
- Tool from the Semios suite
- Analyses a list of requirements in order to group similar elements together
- Using an AI approach & a distance metric

SEMIOS FOR SIMILARITY

Purpose of the tool



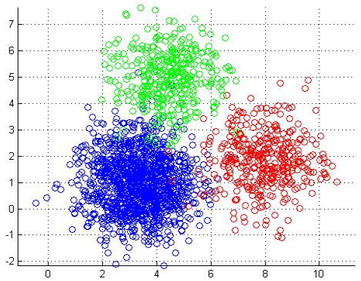
Control the quality of requirements in terms of :

- **Consistency**
- **Redundancy**
- **Contradiction**

SEMIOS FOR SIMILARITY

AI technology in Semios For Similarity : **Clustering**

- **Clustering** : statistic approach
  - ✓ Unsupervised machine learning algorithm : useful for new data
  - ✓ Language independent
- Clustering according to business terms
  - ✓ NLP technology based on PosTagging
- Duplicates generation
  - ✓ Avoid clusters calculation errors



python NumPy pandas $y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$ scikit learn spaCy

spaCy

Bussiness Terms detection

Spacy: neural models for tagging, parsing and entity recognition

- **POS tagging**: Annotate part-of-speech tags on documents

Example: `Unlock only driver door by DR request SW when superlock is disable in EUR variant.`

Postagged exemple: `Unlock` **PROPN** `only` **ADV** `driver` **NOUN** `door` **NOUN** `by` **ADP** `DR` **PROPN** `request` **NOUN** `SW` **PROPN** `when` **ADV** `superlock` **NOUN** `is` **VERB** `disable` **ADJ** `in` **ADP** `EUR` **PROPN** `variant` **NOUN**

- The most used combination patterns in business terms are selected and validated in collaboration with our RE expert:
  - noun-noun (e.g. runway overrun)
  - adjective-noun (e.g. normal mode)
  - proper noun- noun (e.g.BSP data)
  - adjective-adjective-noun (e.g. amber visual indication)
  - noun-noun-noun (e.g. output voltage value)

SEMIOS FOR SIMILARITY

Report output

Report for clustering

3.7 Cluster 7 (2 requirements)

ID	Text
AMR-SR-117	The handheld unit shall function in environments with 99% ambient humidity.
AMR-SR-116	The handheld unit shall function in environments from - 5 degree C through + 50 degree C.

3.5 Cluster 5 (2 requirements)

ID	Text
AMR-SR-115	The handheld unit shall have a mass no larger than 2 , 25kg.
AMR-SR-114	The handheld unit shall be no larger than 25cm x 18cm x 1cm.

Report for distance metric

5.2 Group 2 (6 requirements suggested)

Similarity: Closely similar

ID	Text
AMR-SR-103	**The system shall have a permanently installed network to capture meter readings**
AMR-SR-110	The system shall consist of a series of antennas, towers, collectors, repeaters, or other permanently installed infrastructure to collect transmissions of meter readings from AMR capable meters and get the data to a central computer without a person in the field to collect it.
AMR-SR-104	The system shall collect transmissions of meter readings from AMR capable meters and get the data to a central computer without a person in the field to collect it.

5.6 Group 7 (6 requirements suggested)

Similarity: Likely similar

ID	Text
AMR-SR-110	**The system shall consist of a series of antennas, towers, collectors, repeaters, or other permanently installed infrastructure to collect transmissions of meter readings from AMR capable meters and get the data to a central computer without a person in the field to collect it.**
AMR-SR-104	The system shall collect transmissions of meter readings from AMR capable meters and get the data to a central computer without a person in the field to collect it.
AMR-SR-49	The meter interface unit shall support all data collection functions (data reading, time-triggered operation, and management) of the AMR system.
AMR-SR-103	The system shall have a permanently installed network to capture meter readings

Reports also group together duplicated requirements

Thank you

https://www.semiosapp.com/

Contact : m.mezghanni@semiosapp.com

Partners: