Metrics for Validation and Traceability of Project Management Requirements

Daniela López De Luise
Department of Computer Science and Engineering
CI2S Labs & GIBD/ National Technological University, Concepción Del Uruguay Branch

Patricia R. Cristaldo
Department of Computer Science and Engineering
GIBD / National Technological University, Concepción Del Uruguay Branch

Lucas La Pietra
Department of Computer Science and Engineering
GIBD / National Technological University, Concepción Del Uruguay Branch

Anabella De Battista
Department of Computer Science and Engineering
GIBD / National Technological University, Concepción Del Uruguay Branch

Abstract

Requirements are part of a project's scope. It is a text mostly expressed in natural language and usually collects its content from different documents. This paper introduces a subset of requirement metrics that belong to a larger project for validation and traceability tools. The metrics presented here are based on mathematical simple models and certain Natural Language Processing approaches used to evaluate if the requirements document is relevant, precise, succinct, complete, consistent, and testable. As part of the scope its usage and case studies are also covered. Statistical analyses and a global study of applicability are presented as well. The detailed metric derivation and exhaustive listing are described also, along with with a short functional description of an implementation prototype.

Keywords- Software Projects Management, Management Metrics, Data Mining, Natural Language Processing

I. INTRODUCTION

Software project management can be thought as a combination of Science and Management. It covers many topics like strategies for management, scope definition, managing of interested parties, customers, and risks, planning and control of activities and requirements, and the determination of business goals. It depends on the project's director ability to handle problems mainly related to administration and technology.

There are many methodologies and guides for project management in the market. Among others they are PMBOK [1], PRINCE2 [2][3], APM [4], ISO 21500 [5], SCRUM [6] [7], KANBAN [8], CRISP-DM [9] [10], XP [11], DSDM [12], and CRYSTAL [13]. The right project leading looks for the a completion in time and quality [14] [15]. According to the CHAOS report of 2018, barely 29 percent of the projects are carried out in time, budget, characteristics and expected functionality. By contrast, about 37 percentage of them don't succeed in one of them. Circa 52 percent of the cases are behind schedule, are beyond the budget or implement fewer requirements [16] [17]. The number increases by 10 percent according to reports of the year 2010. Furthermore, the project cancellation rate without any product is about 19 percentage.

Among other causes for failing in accomplishing goals in time and form, there are: poor planning [18] [19] [20], deficient requirement determination [21] [22] [23], and lack of skills, organization and management by leaders [24].

Many authors propose a hybrid approach combining agile and traditional methodologies [25] [26] [27] [28][29] [30]. The advantage here is that the manager must focus not only in the final target but also in the timing for opting among one or the other methodology.

Many efforts are oriented to systematically evaluate the relation between quality and its effects. Some valuable efforts have been made to handle the requirements of any type and its impact in the process and final product [31] [32].

This paper extends some of the previous publications presented above, introducing metrics and indicators that aim to assess and track requirements in order to determine the main aspects and components of the project management methodologies that are mandatory for success. The approach takes as a startup point the textual expression of project requirements. The metrics derived here are a measure framework useful for cross-evaluating different methodologies upon every specific business case, by qualitatively comparing the associated indicators. Thus emerges the possibility to make the applicability of certain methodology explicit to any project step depending on its current context. In order to do that, the obtained metrics are related to a specific set of context-based indicators. This, also enables a proper validation and tracking of declared requirements, which is mandatory for a proper control of them.

The rest of this paper shows the state of the art in metrics and methodologies for management control (section II), introduction of the proposed project management scope (section III), a case study (section IV), the discussion (section V), the conclusions and future work (section VI).

II. RELATED WORK

Management for Software Projects has increased interest in the scientific and academic community. This recent trend can be observed in the publication topics as shown in Figure 1.

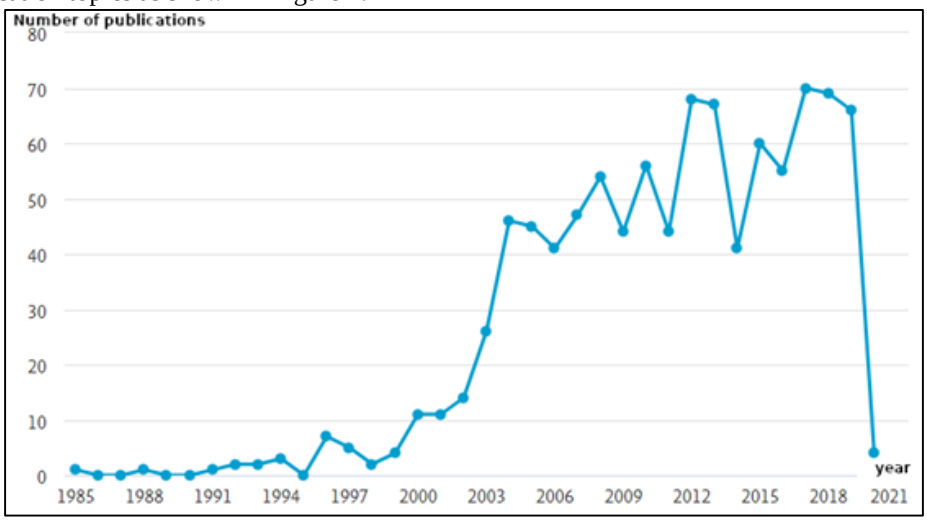


Fig. 1: Number of publications per year

Among the topics, there is a strong emphasis in requirement management and engineering; sometimes with approaches related to NLP (Natural Language Processing) and a few times natural) and occasionally with focus on metrics to evaluate methodologies itself.

In the field of NLP Galinier [31] presents a formal model to reduce incoherence and inconsistency among definitions and requirements of the system. Naziri and col. [35] apply NLP to user history transformation from the requirement specification to design stage. This model can also be extended to many languages [36]. In [37] and [38] a programmatic version of requirements expressed in NL are presented. The proposals in [39] and [40] work in the same line. Extensions of this concept are the automatic document generation in C [41], restricted domain requirement conceptualization [42], and industry modeling extensions with NLP to identify and validate change requirements [43]. Other approaches have been proposed using NLP mainly with semantics [44][45][46][47] and graphical tools [48].

Some metrics are useful to evaluate projects performance. Many of them refer to the product quality model ISO 25010 or to classical quality metrics [49][50][51][52][53]. In [54] there is a step forward with a prediction framework to analyze databases performance. Regarding business process management, in [55][56] there is a proposal of a set of specific quality metrics. Pmbok [1] also defines performance metrics following EVM (Extended Value Management), but they are not suitable for projects with parallel paths [57]. Other approaches determine a performance duration index [58]. For risk management there are specific metrics [59][60], sometimes using agile to track and control [61] best practices related to ISO/IEC/IEEE 12207 and ISO/IEC TR 29110-5-1-2.

This work is centered in the development of metrics for evaluating the implementation process of management models. It is strongly aligned with papers like [61][62]. The proposed set of metrics are suitable for project management and constitute a novel type of tools mainly based on conceptual NLP taken from [44].

III. METRICS FOR PROJECT SCOPE MANAGEMENT

This proposal works on several management cornerstones. Among them is the scope management, which is the main focus of this paper. The remaining will be introduced in future publications.

According to the Best Practices Management Guide [1], scope management is the set of processes to assure all the required tasks needed to successfully complete a project. Considering that, it is important to properly formalize the Business Scope Statement as a proper and complete list of requirements. This way it settles a solid foundation for a quality project. Though this paper is part of a broader project that aims to determine sets of metrics covering the entire management models application process, the main focus here are metrics for evaluating how a model handles the scope of a specific project under certain circumstances. A case study is in the next section (section IV) while the rest of this section details the definitions and elements that constitute the current proposal for working with scope metrics.

A. Parse Trees

Several elements of the metrics' equations consider that a Parse Tree (PT) has been performed. It is a NLP entity for representing the structure of the sentences in a text written in natural language. In this context every word is considered an element that belongs to a grammar G, is part of a dictionary D, and has a linguistic category among a predefined set of categories. Every category

determines the working function of that elements. The grammar follows certain formation rules (frequently called syntactic). For computational linguistics purposes, the tree explains how the categories and rules are working on the text being represented.

The derivation process has two types of elements: non-terminal (those applied to elements of the grammar as categories and rules) and terminal (elements in the dictionary). Every terminal belongs to a category and is part of a rule, and combined with other rules builds a tree that express a structural hierarchy of categories. That is the syntactic structure of a string. Every sentence collapse to a root node or element.

The process is known in the computational linguistic field as phrase marking that leads to Phrase-markers or P-markers. They were first introduced by Noam Chomsky and represents the deep structure of a sentence as generated by applying certain linguistic predefined structure rules. As an example let´s take the Mission and vision statement of MVMT that says: “Style shouldn’t break the bank” [63].

Bottom up from left to right, sentence's words are categorized, labeled and submitted to rules in order from top to bottom.

Here apply the following labels:

Style ← Noun

Shouldn´t ← Auxiliary Verb

break ← Verb

the ← determiner

bank ← Noun

And the nodes are the following:

N: Noun

AV: Auxiliary Verb

V: Verb

D: Determiner

VP: Verbal Phase

NP: Noun Phrase

A simple set of grammar rules could be:

D + N → NP

AV + V → VP

VP + NP → VP

N + VP → S

Figure 2 shows a simple Parse Tree. There, circles are non-terminal elements and rectangles are terminal elements. Observe that terminal elements are the current set of words in the sentence.

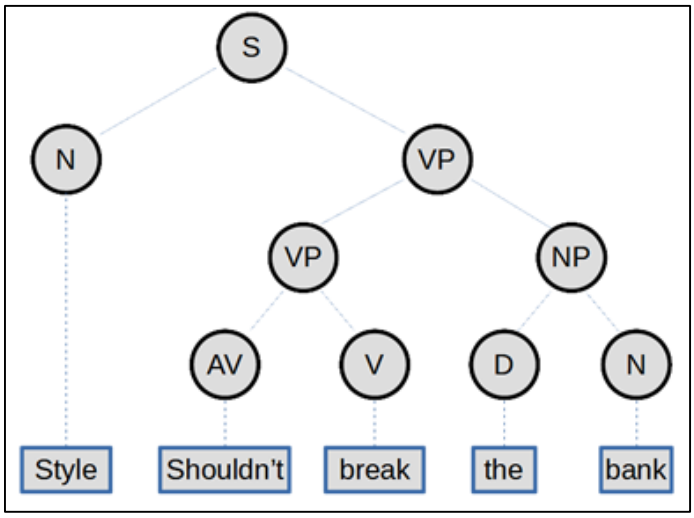


Fig. 2: Sample of Parse Tree for MVMT mission and vision

Phrase markers, or P-markers, were introduced in early transformation generative grammar, as developed by Noam Chomsky and others. A phrase marker representing the deep structure of a sentence is generated by applying phrase structure rules. Then, this application may undergo further transformations.

Phrase markers may be presented in the form of trees (as in the above Parse Tree), but they are often given instead in the form of "bracketed expressions", which occupy less space in the memory and are suitable for automatic processing. Following back in the previous example, the bracketed expression corresponding to the constituency-based tree given above may be something like this:

$$[_S\ [_N\ \textit{Style}][_{VP}\ [_{VP}\ [_{AV}\ \textit{Shouldn't}][_V\ \textit{Break}]][_{NP}\ [_D\ \textit{the}][_N\ \textit{bank}]]]]$$

It is important to note that sometimes the slang of graph theory applies in this context. Thus, leaf or leaf nodes refer to terminal elements, nodes are non-terminal elements and root is a special elements labeled “S”, the starting element that represents

the whole sentence. A child node is an element under other element in the tree structure, which is called father node. Note also that the tree has always the root node on top and leafs on the bottom. A sub-tree is a sub-graph that starts in certain node, and its descendants. The level of a node is the number of nodes to get the root.

B. Proposed Metrics

1) Requirement Validation Criteria (CVR)

CVR is an acronym for "required validation criteria" in Spanish, Criterios de Validación de Requerimientos. Following the literature review in the previous sections, it is strongly determined by the meaningfulness degree of the Business Statement Documentation (mission, vision and scope). The proposed metric to assess this is based on a Parse Tree built from the document as described previously.

$$CVR = p(nec).nec + p(conc).conc + p(comp).comp + p(cons).cons + p(amb).amb + p(verif).verif \quad (1)$$

With:

- p(nec)=1 (if business risk=high), 0.5 (if business risk=neutral), 0.1 (if business risk=low).
- p(conc)= xx/number of child nodes.
- p(comp)= Number of nodes out of the sub-tree/ (number of descendants + number of nodes out of the sub-tree)
- p(cons)= number consistencies/number of requirements
- p(amb)=1
- p(verif)=xx/number of nodes until level n
- xx=number of deployable requirements=number of leaf nodes in the sub-tree of the node

2) Concise Degree (GradConc)

GradConc is the acronym for concise Degree in Spanish, Graduación de Conciso. This is other relevant topic that literature remarks. Eq. .02 interprets it by the proportion of nouns that are in the project scope document, that are also in the requirement PT. This is combined with the same rate calculated for verbs. Therefore GradConc is:

$$GradConc = (num.\ nouns) + (num.\ verbs) / num.\ Words\ in\ project\ scope\ statement \quad (2)$$

3) Completeness Degree (GradCompl)

GradCompl is the acronym for Completeness degree om Spanish, Graduación de Completo. The metric is in eq.03, evaluating how the text of the statement covers explicitly all the required information. It must be projected in a sub-tree:

$$GradCompl = (num.\ Functional\ requirements) + (num.\ Non\text{-}functional\ requirements)/RQN \quad (3)$$

At this time, RQN stands for Business requirements (in Spanish Requerimientos del Negocio, RQN). Determines the needs and opportunities of the enterprise (eq. 04). It is expressed by the requirement list. When it doesn't exist it is valuated as 0.

$$RQN = \sum(r)/\sum (RQN, RQI, RQS, RQP, RQC) \quad (4)$$

r=number of elements (Z>=0)
RQN Ɛ [0..1)

Regarding RQI, it represents the parties involved (in Spanish Requerimientos de los Interesados, RQI): needs of those that participate in the project. It is extracted from the requirement list. When there is no elements, it is valued as 0

$$RQI = \sum(r)/\sum (RQN, RQI, RQS, RQP, RQC) \quad (5)$$

r=number of elements (Z>=0)
RQN Ɛ [0..1)

RQS are the solutions requirements (in Spanish Requerimientos de Soluciones, RQS): the set of characteristics and functionalities for the product or service being the goal of the project. Itis taken from the requirement list. If it doesn't exist it is valied as 0.

$$RQS = \sum(r)/\sum (RQN, RQI, RQS, RQP, RQC) \quad (6)$$

r=number of elements (Z>=0)
RQN Ɛ [0..1)

RQP is the project requirement (in Spanish Requerimientos del Proyecto, RQP): The set of actions and processes that the project must provide. It is taken from the requirement list. If it doesn't exist it is valued as 0.

$$RQP = \sum(r)/\sum (RQN, RQI, RQS, RQP, RQC) \quad (7)$$

r=number of elements (Z>=0)
RQN Ɛ [0..1)

RQC is the quality requirement (in Spanish Requerimientos de Calidad, RQC): The set of decisions and criteria that the product must accomplish. It is tajen from the requirement list. If it doesn't exist it is valued as 0.

$$RQC = \sum(r)/\sum (RQN, RQI, RQS, RQP, RQC) \quad (8)$$

r=number of elements (Z>=0)
RQN Ɛ [0..1)

4) Consistency Degree (GradCons)

GradCons refers to the Consistency (in Spanish Graduación de Consistente, GradCons). Following up the literature recommendations it covers consistency with eq. 09. It watches on contradictions in requirements, if there are negative particles of the language along with positive for the same noun/verb as follows:

GradCons= (num. Contradictions in requirement PT) + (num. Restrictions requirements)/ num. leaf nodes in requirement PT sub-tree (9)

5) Ambiguity Degree (GradAmb)

GradAmb represents the metric for Ambiguity Degree (in Spanish Graduación de Ambigüedad, GradAmb). It works considering the bias interpretation of Descriptive adjectives discarding other types of adjectives:

GradAmb= (num. Nodes Descriptive adjectives) + num. nodes (10)

6) Verification Possibility (GradVerif)

GradVerif is about the possibility of perform certain type of requirement verification. It stands for its name in Spanish: Graduación de Verificable (GradVerif). The requirement statement must allow inspections, analysis, demonstrations and/or tests. The metrics is determined as:

$$GradVerif= \sum(pv)\ (node)/\ num.\ nodes \quad (11)$$

With:

pv(node)=1 (if ∃ a finite process and/or ∃ tests for this requirement), else 0

7) Traceability Degree of the Requirement (GrTraz)

GrTraz is the Degree of requirement's traceability (in Spanish Grado de Trazabilidad del Requerimiento, GrTraz). The Requirement statement must contain status for every element: approved, designed, implemented, tested, etc. The metric is as follows:

$$GrTraz= \sum(attribute: ID + sentence + owner + origin + priority + status)/6 \quad (12)$$

With:

- ID={1 if there is a unique ID, else 0}
- sentence={1 if it can be extracted from scope statement, else 0}
- priority ={1 if its priority is explicit in the requirement list, else 0}
- status= {1 if its status is explicit according status-list, else 0}
- owner = {1 if it is explicit who determined the status, else 0 }
- origin= {1 if it is explicit in the statement, else 0}

The next section is a study case using these metrics in real world.

IV. CASE STUDY

This section tests the proposed metrics with a reduced set of enterprises from the software industry. As can be seen in Figure 3, respondents profile are diverse (project leaders, managers, owners, and others related to the informatics world but with no degree in the field).

The type of enterprise is also diverse, as can be seen in Figure 4.

The database is partitioned in two groups: the first one cover cases when Project Manager (PM) has previously determined the project and product scope. The second is for the rest of the cases, where PM is supported in the process of scope statement writing, using some of the polling data. The polling is part of a testing protocol that was implemented by social media LinkedIn. The metrics proposed in the previous section will be applied in this section to these enterprises as a postmortem tool for analysis. Then a result discussion and context evaluation is performed.

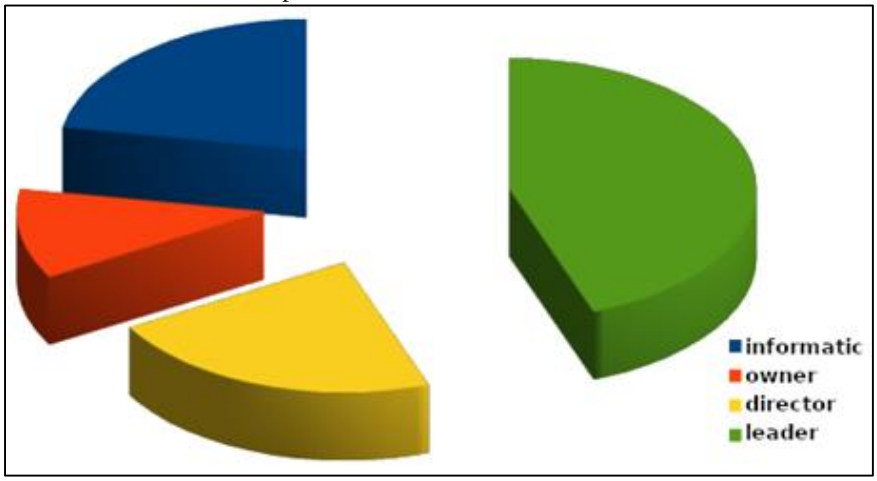


Fig. 3: Respondent profile

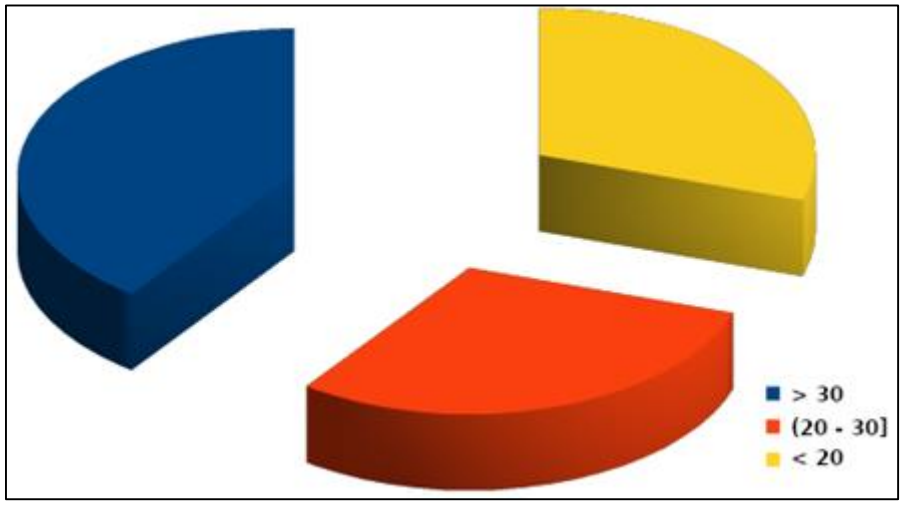


Fig. 4: Enterprise size

V. DEMOGRAPHIC CHARACTERISTICS

Most of the tested volunteers belong to private organizations (80%), as just 20% are from public sector. About 40 % of the enterprises have between 11 and 20 employees at the polling time, 30% more than 30, and 20% less than 10. Circa 65% of volunteers are managers or project directors, and 35% coordinators or staff. Almost 20% of the individuals have accredited knowledge in methodologies and best practices for project management. Just 60% of the respondents have performed the project and product scope statement.

A. Metrics Behavior

This section the behavior of the metrics introduced in previous sections is analyzed in the context of the current test data-set.
As can be seen in Figure 5, the scope statement is not performed in every enterprise. Metrics aim to determine not only the main characteristics of the organization but also the main features of the documentation according to different types of enterprises and business focus.

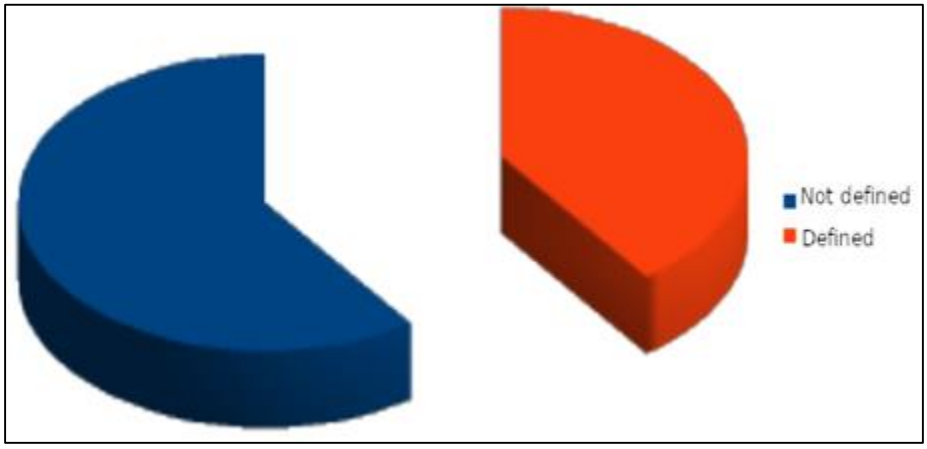


Fig. 5: Scope statement document

In order to consider and interpret the respondent subjectivity in the context of real projects, the previous existence of the statement is remarked. Equations 04 to 08 (see Table 1) work on Scope statement documents and other collected items not included in them. It is important to note that these metrics are just a subset of a larger set to evaluate management key documents like those related to mission and vision.

Table 1: Results from metrics RQN, RQI, RQS, RQP, RQC

	RQN	*RQI*	*RQS*	*RQP*	*RQC*
ID1	*0.3*	*0.2*	*0.25*	*0.05*	*0.05*
ID2	*0.13*	*0.13*	*0.38*	*0.38*	*0.0*
ID3	*0..35*	*0.21*	*0.35*	*0.64*	*0.02*
ID4	*0.03*	*0.0*	*0.01*	*0.00*	*0.02*
ID5	*0.02*	*0.0*	*0.0*	*0.01*	*0.02*
ID6	*0.31*	*0.32*	*0.11*	*0.08*	*0.12*
ID7	*0.25*	*0.29*	*0.12*	*0.24*	*0.17*

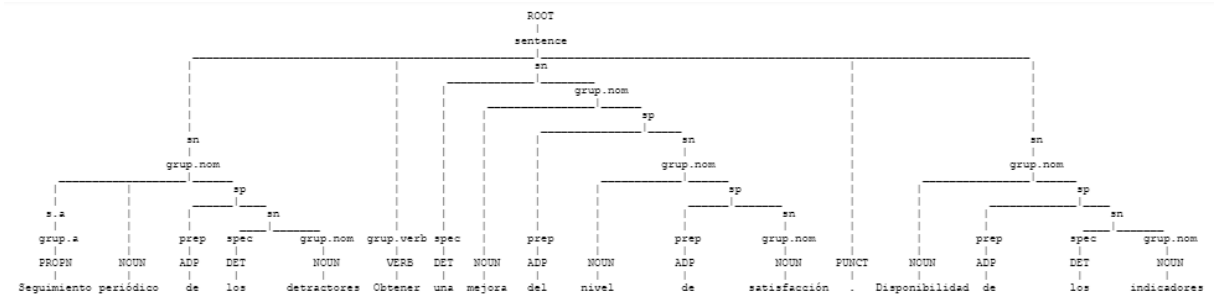
	RQN	*RQI*	*RQS*	*RQP*	*RQC*
ID8	*0.68*	*0.67*	*0.70*	*0.16*	*0.15*
ID9	*0.51*	*0.33*	*0.41*	*0.62*	*0.72*
ID10	*0.41*	*0.55*	*0.52*	*0.41*	*0.75*

The linguistic automatic processing of every scope statement with Octave NLTK(c), starts with a Tagger as in Figure 6. It marks when a word is noun, verb or adjective.

```
from nltk.parse import CoreNLPParser
import nltk.tree
import pandas as pd
from collections import Counter
import json


def obtenerCorpus(numeroEmpresa,numeroPregunta):
    dfnueva=df.drop(columns=['Pregunta'])
    if numeroEmpresa==10:
        dfnueva['UnionEmpresas'] = dfnueva.values.sum(axis=1)
        if numeroPregunta==14:
            celda=dfnueva['UnionEmpresas'].str.cat(sep=', ')
        else:
            celda=dfnueva.iloc[numeroPregunta]['UnionEmpresas']
    else:
        if numeroPregunta==14:
            celda=dfnueva.iloc[:, numeroEmpresa-1].str.cat(sep=', ')
        else:
            celda=dfnueva.iloc[numeroPregunta,numeroEmpresa-1]
    return celda

def parsearTexto(texto):
    listaTagsOraciones=[]
    parse=list(parser.raw_parse(texto))
    for oracion in parse:
        tags=oracion.pos()
        listaTags=[]
        for tupla in tags:
            listaTags.append(tupla[1])
        dictTags = dict(Counter(listaTags))
        listaTagsOraciones.append(dictTags)
    return listaTagsOraciones


df = pd.read_excel('EncuestasFormateadas.xlsx')
parser = CoreNLPParser(url='http://localhost:9003')
parse=list(parser.raw_parse('Yo corro en mi bicicleta roja'))



for numeroPregunta in range(13):
    for numeroEmpresa in range(9):
        #analisis NLP
        corpus=obtenerCorpus(numeroEmpresa,numeroPregunta)
        if corpus!='      ':
            listaTagsOraciones=parsearTexto(corpus)
            for num, oracion in enumerate(listaTagsOraciones, start=1):
                rutadatos=f'Pregunta{numeroPregunta}-Empresa{numeroEmpresa}-Tags-{num}.txt'
                tagsenstr=json.dumps(oracion)
                with open(rutadatos, 'wt') as out:
                    out.write(tagsenstr)
            print(f'Tags para la pregunta {numeroPregunta} de la empresa {numeroEmpresa} guardadas con éxito')
```

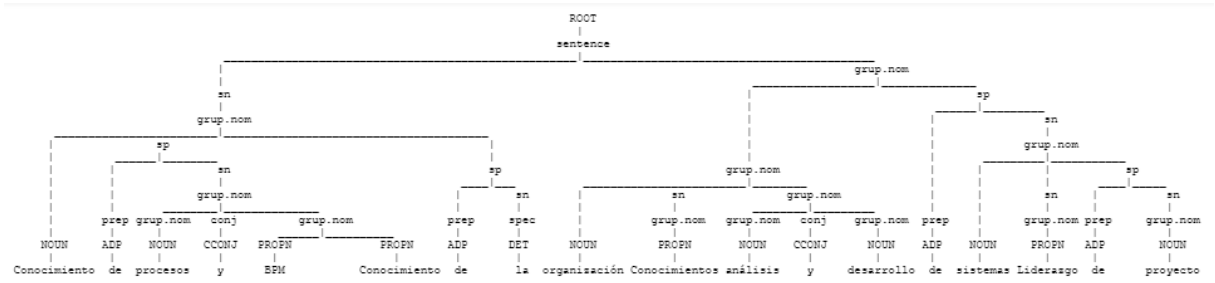
Fig. 6: Tagger

Table 2: Result for metric CVR

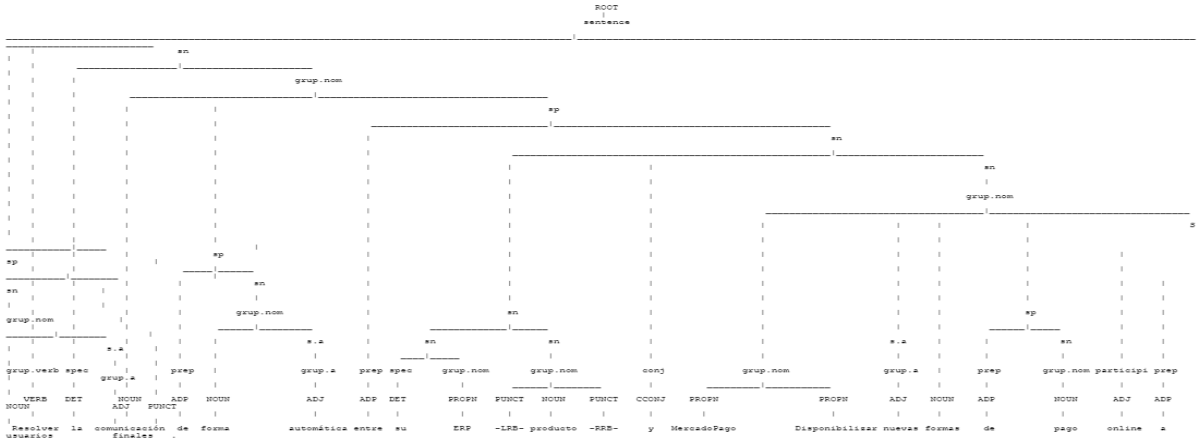
Case	***CVR***
ID1	

Case ***CVR***

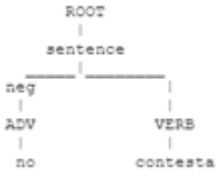
ID2



ID3

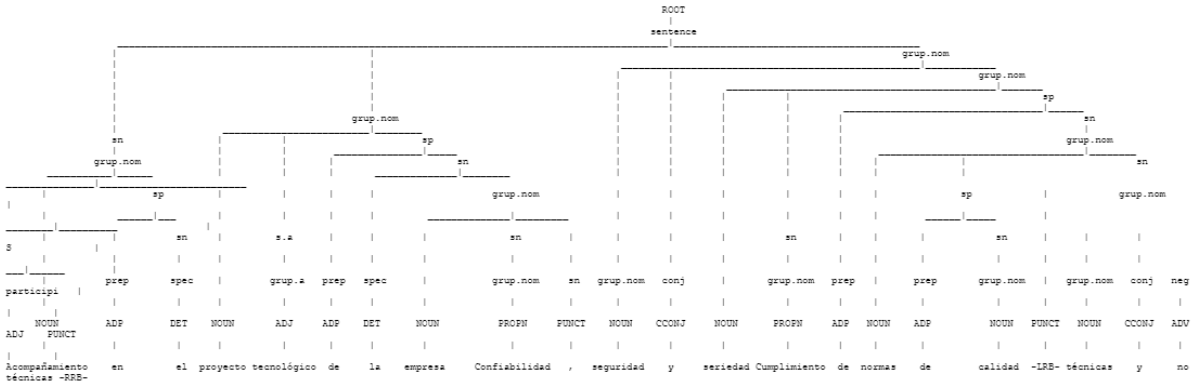


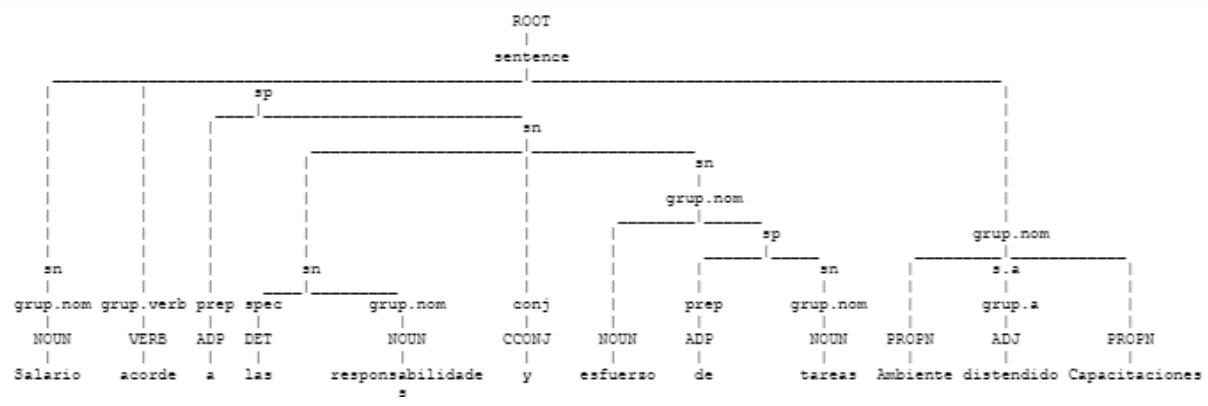
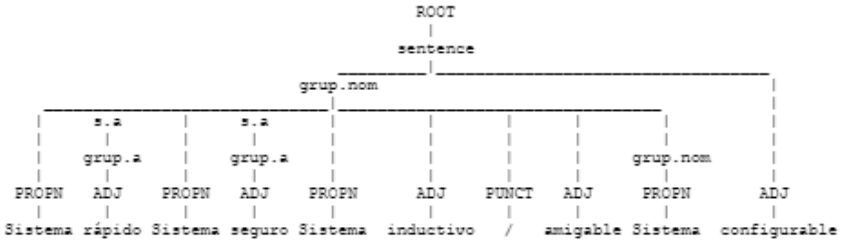
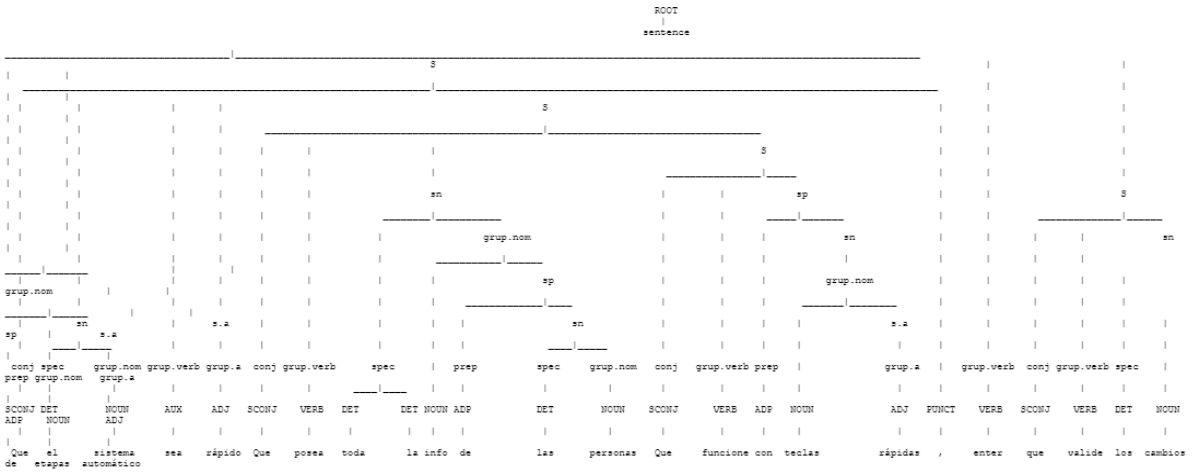
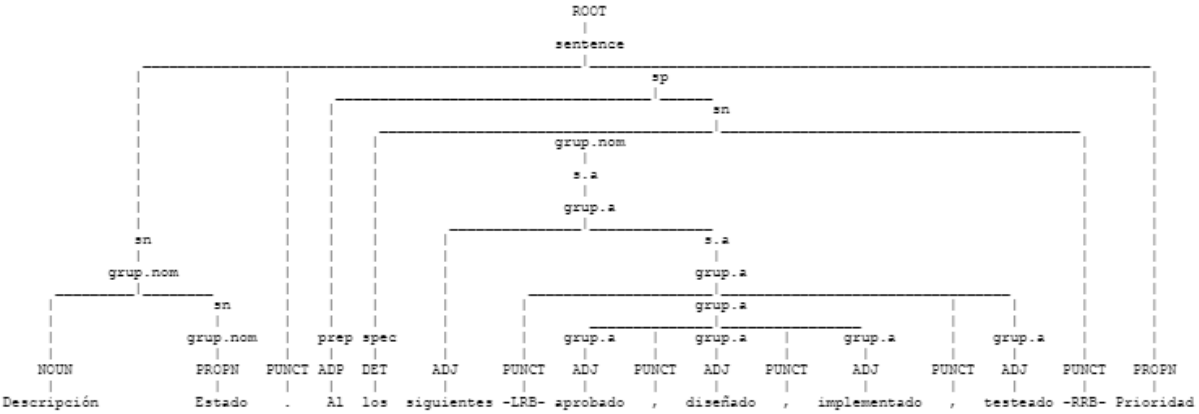
ID4



ID5

ID6



Case	***CVR***
ID7	
ID8	
ID9	
ID10	

To provide a better comparison basis, texts belonging to the same sections in the Scope Statement are evaluated in the same bunch. Table 3 shows a summary of the PT obtained.

Table 3: Summary of tags

	Tags	*CVR*
ID1	*{"PROPN": 1, "NOUN": 7, "ADP": 4, "DET": 3, "VERB": 1, "PUNCT": 1}*	*0.47*
ID2	*{"NOUN": 7, "ADP": 4, "CCONJ": 2, "PROPN": 4, "DET": 1}*	*0.50*
ID3	*{"VERB": 1, "DET": 2, "NOUN": 6, "ADP": 4, "ADJ": 4, "PROPN": 3, "PUNCT": 3, "CCONJ": 1}*	*0.63*
ID4	*{"ADV": 1, "VERB": 1}*	*0.04*
ID5	*{"ADV": 1, "VERB": 1}*	*0.04*
ID6	*{"NOUN": 8, "ADP": 4, "DET": 2, "ADJ": 2, "PROPN": 2, "PUNCT": 3, "CCONJ": 2, "ADV": 1}*	*1.03*
ID7	*{"NOUN": 4, "VERB": 1, "ADP": 2, "DET": 1, "CCONJ": 1, "PROPN": 2, "ADJ": 1}*	*0.21*
ID8	*{"PROPN": 4, "ADJ": 5, "PUNCT": 1}*	*0.19*
ID9	*{"SCONJ": 4, "DET": 5, "NOUN": 6, "AUX": 1, "ADJ": 3, "VERB": 4, "ADP": 3, "PUNCT": 1}*	*1.18*
ID10	*{"NOUN": 7, "ADJ": 2, "CCONJ": 1, "ADP": 5, "DET": 4, "PROPN": 1, "PUNCT": 2}*	*0.89*

Table 4 is the word distribution by ID. L denotes a linear behavior, when words have identical frequency (value 1). Entries with NO, are cases with no answer to the question. The remaining are cases with one or more words, with an average number higher than the rest, and are considered as a special analysis case.

Note that 42% of 50 entries are linear, and 10% have no answer. Both cases summarize 52%.

Table 4: Word distribution frequency by question

ID	*RQN*	*RQI*	*RQS*	*RQP*	*RQC*
1	*L*	*L*	*PAGOS (PAYMENTS)*	*L*	*L*
2	*L*	*CONOCIMIENTO (KNOWLEDGE)*	*SISTEMA (SYSTEM)*	*L*	*L*
3	*PAGOS (PAYMENTS)*	*L*	*ALTA (HIGH)*	*PLATAFORMA (PLATFORM) INTEGRACIÓN (INTEGRATION)*	*NO*
4	*L*	*NO*	*NO*	*L*	*L*
5	*L*	*NO*	*NO*	*NO*	*DEBE (MUST) SER (BE)*
6	*L*	*TÉCNICAS (APPROACHES)*	*L*	*L*	*L*
7	*DISPONIBILIDAD (AVAILABILITY)*	*L*	*L*	*PROYECTO (PROJECT) CUENTA (ACCOUNT) PROCESO (PROCESS)*	*L*
8	*SISTEMA (SYSTEM)*	*SISTEMA (SYSTEM)*	*MÓDULO (MODULE) PRÉSTAMO (LOAN)*	*PROCESO (PROCESS)*	*L*
9	*GESTIÓN (MANAGEMENT)*	*L*	*MÓDULO (MODULE)*	*PROCESO BANCO (BANK PROCESS)*	*DATOS (DATA) DEBEN SER (MUST BE)*
10	*ENCUESTAS (POLLINGS) RESULTADOS (RESULTS)*	*GOBIERNO (GOVERNMENT) SERVIDORES (SERVERS)*	*APLICACIÓN (PROGRAM) DEBE (MUST) SER (BE)*	*CUMPLIMIENTO (FULFILLMENT) NORMAS (REGULATIONS)*	*CRÉDITOS (CREDITS) MÓDULO (MODULE) PRÉSTAMO (LOAN)*

The next section analyzes the relation between these results, linguistics and the type of enterprise.

VI. DISCUSSION

The Scope Statement document must describe the characteristics, functions and requirements for the product or service being provided and for the process itself. In this context, metrics must be able to determine validation criteria for business requirements (RQN), involved individuals (RQI), solutions (RQS), project as a whole (RQP) and quality (RQC). They must also provide information for their traceability as well.

From Table 2, CVR is highest for ID 9, followed by ID 6, ID 10 and ID 3 (all with a value greater than 0,5). This indicates that these enterprises show a higher degree of business scope understanding, which is expressed in the way the documents are

written. The cases that correspond to enterprises were already defined, despite its size. In fact they are big, medium and small. For none of them the respondent is the owner but an expert on project management.

Cases ID 4 and 5 present values under 0.1. These are cases where the tested person is the director. Typically the scope is defined at an organizational level but not in a detailed way. They are not able to answer typical questions about scope statement. Also present lower values of RQN, RQI, RQS, RQP, y RQC. Both cases correspond to state organizations.

Regarding ID 1 and 2, the values are between 0,3 and 0,5. That means a little better value than ID 4 and 5, and represent documents with higher degree of lexical diversity. Documents are more elaborated, expressive and larger. This is related to more information in the text document and therefore to a more detailed scope maturity.

These cases are in the intermediate zone for CVR, related to medium and large enterprises, geographically distributed in several provinces and countries. The field of the industries are supply companies and several types of staffing. The degree of understanding of the business is clearly higher to previous cases but not perfect. By contrast, the scope has even better defined.

CVR for ID 7 and 8 is between 0.1 and 0.3, representing a higher degree of elaboration and more efficacy in expression. Correspond to enterprises that provide software development services. Both of them are small and one is medium size. The volunteers have scarce management knowledge in all cases, but some of them are under training. None of the companies have branches and two of them are located in capital cities of relevant provinces. Here, the degree of definition of the scope is balanced and quite good.

Both ID 6 and 8, present the highest values. They represent large enterprises, and the volunteer is the project leader. None of the companies have branches but they are located in a very relevant capital city in the country. It is important to note that in one of them the scope statement document was not already existent. Both cases are very well established businesses, leaders in the market and with a solid experience in project management methodologies.

Table 5 is a summary of results and its relation with cases ID. Also proposes a qualification indicator for the value ranges of the metric.

Table 5: Indicator and CVR

ID	*GVR*	*Indicator*
4,5	*<0.1*	*LOW*
1,2,7,8	*[0.1 – 0.5]*	*MEDIUM*
3,6,9,10	*>0.5*	*HIGH*

In a similar way, Table 6 presents a summary for GrTraz values and certain indicators.

Table 6: Indicator and GrTraz

ID	*GrTraz*	*Indicator*
4,5	*<0.1*	*VERY LOW*
7,8	*[0.1 – 0.3)*	*LOW*
1, 2	*[0.3 – 0.5)*	*MEDIUM*
2,7,8	*[0.3 – 0.5)*	*HIGH*
3,10	*>=1*	*VERY HIGH*

It is interesting to note that Table 4 denotes the focus of the enterprise on every stage of the business production line. For ID 3, just “NO” results relevant for the text of services, “Pagos” (payment) for its business requirement declaration and "Plataforma" and "integración" (platform and integration respectively) for project requirement text. As the enterprise belongs to software industry, it could express the fact that there is a software product under construction where the most relevant fact is the expected payments to be obtained.

The status of ID 5 is worse since there is relevant word, then there is no relevant concept automatically detected.

When the enterprises have a requirement statement in MEDIUM level (related to performance in RQN, RQI, RQS, RQP, and RQC), there are certain relevant words automatically detected. For example in ID 7 the word "disponibilidad" (availability) arises as a value factor for the company, which is also its main feature presented in the market.

Enterprises with requirement statement in HIGH level, there are more relevant words and are more specific and very close related to its main product characteristics.

VII. CONCLUSIONS AND FUTURE WORK

This paper introduces a set of metrics for project management evaluation in the context of specific projects. They are focused on the Scope Statement document, and a used case for the Requirement Validation Criteria (CVR) and the Requirement traceability Degree (GrTraz). The rest of the metrics are Business Understanding Degree (GCN), Degree of Completeness of Scope (GCA) and many others are out of the scope of this work. The use case explained here is a test set of 10 companies. A number of indicators are proposed for the values obtained.

From the results obtained here it can be said that metrics relate to quality of the scope and requirements specifications, which can be easily contrasted against traditional linguistic and ad-hoc approaches.

Among other pending tasks, there is the definition of the indicators like for time, cost, quality, resources, stakeholders, communications, risks and expenses management.

Also remains to test the rest of the metrics and to validate results with larger sets, evaluate the impact on projects when metrics are applied and considered during the project evolution.

An interesting pending is the relationship between metrics and the implicit knowledge in documents, probably using data mining, to obtain extra hidden information from data.

REFERENCES

[1] Project Management Institute. A Guide to the Project Management Body of Knowledge. 6 Ed. ISBN- 10: 9781628251845, 2017.
[2] PRINCE2. An introduction to PRINCE2: managing and directing successful projects. Office of Government Commerce. Stationery Office, 123 p. ISBN-10: 0113311885, ISBN-13: 978-0113311880. 2009.
[3] Böhm, A. Application of PRINCE2 and the Impact on Project Management. ISBN (eBook) 978-3-640-42634-8, 2009.
[4] Highsmith, J. Agile project management: creating innovative products. 2 Ed. Boston, MA: Addison-Wesley. 432 p. 2010.
[5] ISO. ISO 21500:2012 Guidance on Project Management, Geneva: ISO, 2012.
[6] Sutherland, J. Scrum: The art of doing twice the work in half the time. New York: Crown Business. 256 p. ISBN-10: 038534645X, ISBN-13: 978-0385346450, 2014.
[7] Van Solingen R., Van Lanen R. (Scrum for managers) Scrum voor Managers. Den Haag: Academic Service. EAN: 9789012585903, 2014.
[8] Lei H., Ganjeizadeh F., Jayachandran P., Ozcan, P.A statisc al analysis of the effects of Scrum and Kanban on software development projects. Robotics and Computer-Integrated Manufacturing, http://dx.doi.org/10.1016/j.rcim.2015.12.001, 2015.
[9] Shearer C. The CRISP-DM model: The new blueprint for data mining. Journal of Data Warehousing, no. 4, pp. 13-22, 2000.
[10] Shafique U., Qaiser H. A Comparative Study of Process Models Data Mining (KDD, CRISP-DM and SEMMA).International Journal of Innovation and Scientific Research, pp.217-222, 2014.
[11] Livermore JA Factors that signifcantly Impact the Implementation of an agile software development methodology. Journal of Software, pp. 31–36, 2008.
[12] Richards K. Agile project management: running PRINCE2 projects with DSDM Atern. Norwich: The Stationery Office. 98 p.ISBN-10: 0113310587, ISBN-13: 978-0113310586, 2013.
[13] Nardonea R, Marrone S, Getilea U, Amatoc A, Barberioc G, Benerecetti M, De Guglielmo R, Di Martinoe B, Mazzoccaa N, Perona A, Pisani G, Velardi L, Vittorini V. An OSLC-based environment for system-level functional testing of ERTMS/ETCS controllers. https://doi.org/10.1016/j.jss.2019.110478, 2020.
[14] Varajão, J. Dominguez, C. Ribeiro, P. Paiva, A. Critical success aspects in Project management: similarities and differences between the construction and software industry, Tech. Gazette 21 (3) 583–589, 2014.
[15] Ranjbarfard M, Hatami Z. critical Success Factors for Implementing Business Intelligence Projects (A BI Implementation Methodology Perspective). https://doi.org/10.28945/4607, 2020.
[16] TSG. The CHAOS Report. The Standish Group, 2018. Disponible en https://secure.standishgroup.com/reports/flyers/CM2018- TOC.pdf, 2018.
[17] Ibraigheeth M, Fadzli S. Core Factors for Software Projects Success. e-ISSN : 2549-9904, 2019.
[18] Lehtinen T., Mäntylä M., Vanhanen J., Itkonen J., Lassenius, C. Perceived causes of software project failures – An analysis of their relationships. Information and Software Technology, pp. 623–643, 2014.
[19] Ramos P., Mota C. Perceptions of success and failure factors in information technology projects: a study from Brazilian companies. Procedia - Social and Behavioral Sciences, pp. 349 – 357, 2014.
[20] Montequin S., Fernandez C., Fernandez O., Balsera J. Analysis of the Success Factors and Failure Causes in Projects: Comparison of the Spanish Information y Communication Technology (ICT) Sector. Journal Information Technology Project Management, pp. 18-31, 2016.
[21] Chow T., Chao D. A survey of critical success factors in agile software projects. Journal of systems and software. Science Direct, pp.961– 971, 2008.
[22] Elkadi H. Success and failure factors for e-government projects: A case from Egypt. Egyptian Informatics Journal, pp. 165–173, 2013.
[23] ElEmam K., Koru A. A replicated survey of IT software project failures Software. IEEE Software, pp. 84–90. 2008.
[24] Blaskovics B. The impact of project manager on project success – The case of ICT sector. Assistant Professor, Strategy and Project Management Department, Corvinus University Budapest E-mail: balint.blaskovics@uni-corvinus.hu, 2016.
[25] Esteki M., Gandomani T., Farsani H. A Risk Management Framework for Distributed Scrum Using PRINCE2 Methodology. Bull. Electrical Engineering and Informatics. DOI: 10.11591/eei.v9i3.1905. 2020.
[26] Jabar M., Mohd Ali N., Jusoh Y., Abdullah S., Mohanarajah S. A pilot examination of an improved agile hybrid model in managing software projects success. Test Engineering and Management, pp. 3040-3046, 2019.
[27] Smoczyńska A., Pawlak M., Poniszewska-Marańda, A. Hybrid agile method for management of software creation. Advances in Intelligent Systems and Computing, pp. 101-115. 2019.
[28] Włodzimierz Wysockia W., Orłowski, C. A multi-agent model for planning hybrid software processes. Procedia Computer Science pp. 1688–1697, 2019.
[29] Mousaei M., Gandomani T. A new project risk management model based on Scrum framework and Prince2 methodology. International Journal. 2018.
[30] Neelu L, Kavitha D. Estimation of software quality parameters for hybrid agile process model. https://doi.org/10.1007/s42452-021-04305-0, 2021.
[31] Galinier F. A DSL for Requirements in the Context of a Seamless Approach. https://doi.org/10.1145/3238147.3241538, 2018.
[32] Garousi V, Bauer S, Felderer M. NLP- assisted software testing: A systematic mapping of the literature. https://doi.org/10.1016/j.infsof.2020.106321, 2020.
[33] Cuccurullo C., Aria M., Sarto F. Foundations and trends in performance management. A twenty-five years bibliometric analysis. Scientometrics, pp. 595–611, 2016.
[34] Van Eck N., Waltman L. Software survey: Vosviewer, a computer program for bibliometric mapping. Scientometrics, pp. 523–538, 2010.
[35] Nasiri S, Rhazali Y, Lahmer M, Chenfour N. Towards a generation of class diagram from user stories in agile methods. Procedia Computer Science, pp. 831–837, 2020.
[36] Bruel J, Ebersold S, Galinier F, Naumchev A, Mazzara M, Meyer B. Formality in software requirements, DOI:10.13140/RG.2.2.29708.05762, 2019.
[37] Naumchev A, Meyer B, Mazzara M, Galinier F, Bruel J, Ebersold S. Expressing and verifying embedded software requirements. ArXiv. Preprint arXiv:1710.02801, 2017.
[38] Naumchev A, Meyer B. Complete contracts through specification drivers. arXiv: 1602.04007, 2016.
[39] Garousi V, Bauer S, Felderer M. NLP-assisted software testing: A systematic mapping of the literature. https://doi.org/10.1016/j.infsof.2020.106321, 2020.
[40] Nazir F, Butt W, Anwar M, KhanKhattak M. The applications of natural language processing (NLP) for software requirement engineering - a systematic literature review. International Conference on Information Science and Applications Information Science and Applications, pp 485-493, 2017.
[41] Menaka Pushpa A. Automatic Source Code Documentation using Code Summarization Technique of NLP. Procedia Computer Science, pp.2522–2531, 2020.
[42] Brundagea M, Sexton T, Hodkiewicz M, Dima A, Lukens S. Technical language processing: Unlocking maintenance knowledge. https://doi.org/10.1016/j.mfglet.2020.11.00, 2020.
[43] Dawood H, Siddle J, Dawood N. Integrating IFC and NLP for automating change request validations doi: 10.36680/j.itcon.2019.030. 2020.
[44] Schlutter A, Vogelsang A. Knowledge Extraction from Natural Language Requirements into a Semantic Relation Graph. https://doi.org/10.1145/3387940.3392162, 2020.
[45] Mahmoud A, Niu N. On the role of semantics in automated requirements tracing.https://doi.org/10.1007/s00766-013-0199-y, 2015.
[46] Guo J, Cheng J, Cleland-Huang J. Semantically Enhanced Software Traceability Using Deep Learning Techniques. International Conference on Software Engineering (ICSE). IEEE. https://doi.org/10.1109/ICSE.2017.9, 2017.

[47] Michalke V, Hartig K. Explanation Retrieval in Semantic Networks. https://doi.org/10.14279/depositonce-7136, 2016.
[48] Schlutter A, Vogelsang A. Knowledge Representation of Requirements Documents Using Natural Language Processing. https://doi.org/10.14279/depositonce-7776, 2018.
[49] Arvanitou E., Ampatzoglou A., Chatzigeorgiou A., Galster M., Avgeriou P. A mapping study on design-time quality attributes and metrics. The Journal of Systems and Software, pp. 52–77, 2017.
[50] Tebes G., Peppino D., Becker P., Matturro G., Solari M., Olsina Analyzing and documenting the systematic review results of software Testing ontologies. Information and Software Technology 123, 2020.
[51] Jaleel A., Arshad S., Shoaib M., Awais M. Design Quality Metrics to Determine the Suitability and Cost-Effect of Self-* Capabilities for Autonomic Computing Systems. Digital Object Identifier 10.1109/ACCESS.2019.2944119, 2019.
[52] Friedman A., Flaounas I. The Right Metric for the Right Stakeholder: a Case Study of Improving Product Usability. https://doi.org/10.1145/3292147.3292224, 2018.
[53] Ismail S., Mohd F., Jalil M., Wan Kadir W. Development metrics measurement level for component reusability evaluation approach (CREA). DOI: 10.11591/ijece.v9i6.pp5428-5435, 2019.
[54] Raza B., Aslam A., Sher A., Malik A., Faheem M. Autonomic performance prediction framework for data warehouse queries using lazy learning approach. Applied Soft Computing Journal 91, 2020.
[55] Kbaiera W., Ghannouchib S. Determining the threshold values of quality metrics in BPMN process models using data mining techniques. Procedia Computer Science pp.113–119, 2019.
[56] Ahmad M., Odeh M., Green S. Metrics for Assessing the Basic Alignment between Business Process and Enterprise Information Architectures with reference to the BPAOntoEIA Framework. IEEE 978-1-7281-0385, 2018.
[57] Vandevoorde S., Vanhoucke M. A simulation and evaluation of earned Value metrics to forecast the project duration. Journal Operational Research Society 58, pp. 1361-1374, 2007.
[58] Khamooshi H., Golafshani H. EDM: earned Duration Management, a new approach to schedule performance management and measurement. International Journal Project Management, 1019-1041, 2014.
[59] Vanhoucke, M., 2011. On the dynamic use of project performance and schedule risk information during project tracking. Omega 39, pp. 416-426, 2011.
[60] Wood D. High-level integrated deterministic, stochastic and fuzzy cost-duration analysis aids project planning and monitoring, focusing on uncertainties and earned value metrics. Journal of Natural Gas Science and Engineering pp. 303-326, 2017.
[61] Mas A., Mesquida A., Pacheco M. Supporting the deployment of ISO-based project management processes with agile metrics. https://doi.org/10.1016/j.csi.2019.103405, 2020.
[62] Silva Ferreira M., Almeida Martins L., Júnior P., Costa H. Measuring developer work to support the software project manager: and exploratory study. https://doi.org/10.1145/3364641.3364651, 2019.
[63] https://www.oberlo.com/blog/inspiring-mission-vision-statement-examples. Taken on April 14th, 2021.