Comparison of COBIT 5 and ITIL V3 using Semantic Analysis
Bruno Sony Lopes Borges
bruno.lopes.borges@tecnico.ulisboa.pt
Instituto Superior Técnico, Lisboa, Portugal
Outubro 2018

Abstract

Enterprise Governance of Information Technology (EGIT) can be defined as a means to achieve business/IT alignment. Organizations can benefit from this alignment by implementing processes defined by EGIT frameworks to become more competitive and produce high-quality products and services. COBIT and ITIL are two well-known Process Reference Models (PRMs) that are widely adopted and adapted by organizations. PRMs are always related to a process assessment model (PAM) which holds all details to determine the capability of the processes of the reference model. However, in multi-framework environments, organizations struggle in assessing these frameworks processes implementation since each framework defines its own scope, structure, definitions, and terminology. To overcome this issue, the authors propose an approach that through semantic similarity compares COBIT PAM and TIPA for ITIL - the PAM for COBIT 5 and ITIL - core concepts to highlight their overlapping.

1. Introduction

The awareness that business involvement is crucial in the governance of Information Technology (IT) initiated a shift in the definition of IT Governance toward Enterprise Governance of IT (EGIT)[1].

Enterprise Governance of IT (EGIT) can be defined as "an integral part of corporate governance which addresses the definition and implementation of processes, structures and relational mechanisms in the organization that enable both business and Information Technology (IT) people to execute their responsibilities in support of business/IT alignment and the creation of business value from IT-enabled business investments" [2].

Several authors argue that organizations should implement EGIT over the use of frameworks, best practices, and ISO standards (due to readability reasons, hereafter called EGIT Frameworks) [3]. Moore [4] identified approximately 315 EGIT Frameworks, and the number of EGIT Frameworks has now increased, as have their application areas, since some of these EGIT frameworks only cover a specific aspect of IT, such as information security, service management, quality, among others.

This situation allows organizations to select and complement their processes from the EGIT frameworks which better fit their contexts [5]. Researchers agree that COBIT, ITIL and ISO 27000 family are the most valuable and popular EGIT frameworks currently being adopted [6], [7]. Research also indicates that organizations are widely adopting COBIT in practice [1], [7], [8].

COBIT, now in its fifth edition, is a process reference model (PRM) that provides a comprehensive practice that assists enterprises in achieving their objectives for the governance and management of enterprise IT [9]. In turn, ITIL is a PRM that provides descriptive guidance on the management of IT processes, functions, roles, and responsibilities related to IT service management (ITSM) [10].

Because of its high-level, broad coverage, COBIT can act as an integrator and can be mapped to other EGIT frameworks that cover specific areas in more detail such as ITIL [11].

For COBIT and ITIL, as PRMs, process management requires each process to be controlled to remain compliant with the objectives of both IT and business [12]. Therefore, PRMs are always related to a Process Assessment Model (PAM) which holds all details to determine the capability of the processes of the reference model [13]. A process assessment is conducted to get a clear view of the current practices in an organization in a particular domain. The goal is to compare these EGIT frameworks to a renowned
reference so that the current status of the processes can be measured and appropriate suggestions for process improvement can be made.

TIPA for ITIL PAM and COBIT PAM are two well-known PAMs. They are based on ISO/IEC 15504 [14], [15], which is a global reference for conducting process capability assessments. From an assessment perspective, both TIPA for ITIL and COBIT 5 PAM break down each process into Base Practices specific to each process and take into account generic practices, which are not restricted to any particular process.

Nowadays, the increasing demands of the different industries coupled with compliance requirements, have forced organizations to adopt multiple EGIT frameworks [16], which add even more complexity to the field, since organizations struggle with the perceived complexity and difficulty of understanding and adopting several frameworks at the same time [17].

Complexity, in general, can be defined as “property of a language [representation] expression which makes it difficult to formulate its overall behavior, even when given almost complete information about its atomic components and their inter-relations.” [18]. Choosing this generic and overarching definition enables the adoption to various fields of application and incorporates specific complexity theories, views, and paradigms [19].

Complex systems can be understood as a heterogeneous amount of elements with diverse interrelations and dependencies [20], [21], [22]. In this context, complexity can be divided into multiple dimensions. Namely, these are task related complexity, structural complexity and time-related complexity/dynamics [16], [23]. While structural complexity can be quantitatively measured [24], the remaining dimensions often depend on the subjective reception of an object by the observer [25], [26]. This observation leads to differentiation between real or structural complexity and perceived complexity [27], [28], which is the type related to EGIT frameworks implementation.

At a time when organizations strive to be efficient and effective, it seems counterintuitive to be wasting resources by having different organizational departments handling both approaches independently [29]. However, the literature lacks a proposal that can assist and guide organizations in assessing different frameworks in a multi-framework environment [30]. Identifying differences, similarities and reducing issues related to the overlapping [31] can be a good start to optimize the costs and risks in the process assessment.

Semantic similarity defines the relatedness between texts based on the likeness of their meaning [32]. Therefore, Semantic Similarity Measures can calculate and present the distance between terms and concepts so that stakeholders can be aware of the overlaps that may exist between different EGIT frameworks.

To sum up, the problem this research intends to help solve is: ‘The terminological disparities between ITIL V3 2011 and COBIT 5 makes best professional judgment the only approach currently available to understand the overlapping between COBIT 5 and ITIL, hindering their simultaneous assessment’.

Therefore, the primary objective of this research is to perform an integration, based on a semantic evaluation, of COBIT and ITIL process assessment core concepts (base practices, inputs/outputs, outcomes and expected results) in order to highlight the overlapping between these concepts and, in a certain way, facilitate the simultaneous assessment of these EGIT frameworks by identifying common aspects of both frameworks.

2. Research Methodology

This research follows the Design Science Research Methodology (DSRM). DSRM involves a rigorous process to design artifacts to solve observed problems, to make research contributions, to evaluate the designs, and to communicate the results to appropriate audiences [33]. Walls et al. [34] argue that the widespread adoption of DSRM within the field of Information Systems research will have more impact on management practice as a result of closer ties between researchers and practitioners. Hence, [35] posits academics should conduct design science research more regularly. Unfortunately, the application of design science research within the context of EGIT remains relatively low [36].

DSRM is appropriated for research that seeks to extend the boundaries of human and organizational capabilities by creating new and innovative artifact [37]. DSRM differentiates from other research paradigms because it tries to develop and reach artifacts that can be proven effective in real-world scenarios [33].

A methodology can be seen as a system of principles, practices, and procedures applied to a specific branch of knowledge. According to [33], a methodology for design science research would include three elements: conceptual principles to define what is meant by design science research, practice rules, and a process for carrying out and presenting the research. The process model consisting of six iterative activities is described next:

- **Problem Identification and Motivation:** The terminological disparities between ITIL V3 2011 and COBIT 5 makes best professional judgment the only approach currently available to understand the overlapping between COBIT 5 and ITIL, hindering their simultaneous assessment.
• **Definition of the objectives for the solution:**
  To perform an integration, based on a semantic evaluation, of COBIT and ITIL process assessment core concepts (base practices, inputs/outputs, outcomes and expected results) in order to highlight the overlapping between these concepts and, in a certain way, facilitate the simultaneous assessment of these EGIT frameworks by identifying common aspects of both frameworks.

• **Design and Development:** Use of the spaCy library for evaluating COBIT 5 PAM and TIPA for ITIL similarity.

• **Demonstration:** Semantic Similarity between Tipa for ITIL and COBIT PAM.

• **Evaluation:** Compare the accuracy of the results against a benchmark proposed by a group of human experts.

• **Communication:** Submission to the 52nd Hawaii International Conference on System Sciences (HICSS) conference.

3. Theoretical Background

In this Section, topics concerning this research scope are introduced and discussed. Existing literature focused on COBIT 5, ITIL, Process Assessment models, and Semantic Similarity techniques are highlighted.

3.1. COBIT 5

COBIT, now in its fifth edition, provides a comprehensive framework that assists enterprises in achieving their objectives for the governance and management of enterprise IT [9].

COBIT 5 is based on five principles: meeting stakeholder needs; covering the enterprise end-to-end; applying a single, integrated framework; enabling a holistic approach; and separating governance from management [10]. Together these principles would allow enterprises to assemble and deploy an effective EGIT and management practice and thus support an outstanding balance between benefits realization, risk management and resources [10].

COBIT 5 unified ISACA’s three frameworks: Val IT, a value delivery focused framework; Risk IT, a risk management focused framework and previous COBIT versions. Hence, this allowed COBIT 5 to cover the lifecycle of governance and management within the scope of enterprise IT [1].

3.2 ITIL

ITIL is a set of comprehensive publications providing detailed guidance on the management of IT processes, functions, roles, and responsibilities related to ITSM [11]. ITIL focuses cycle of life, renewal and decommissioning of services, with a greater business-focused perspective [38]. Its benefits have been addressed from a few relevant academic researchers, that frequently evidenced the following benefits: improvement of Service Quality, improvement of Customer Satisfaction, improvement of Return on Investment [39], [40].

However, according to Strahonja [41], ITIL has also some weaknesses such as the lack of holistic visibility and traceability from the theory (specifications, glossary, guidelines, manuals, amongst others) to its implementations and software applications; its focus on the logical level of processes, instructing what should be done but not how; and its poorly definition of the information models corresponding to process description.

3.3. Process Assessment Models

A process assessment is conducted to get a clear view of the current EGIT frameworks in an organization in a particular domain. The goal is to compare these EGIT frameworks to a renowned reference so that the current status of the processes can be measured and appropriate suggestions for process improvement can be made.

COBIT 5 PAM is a model that aims at assessing the capability of a COBIT 5 process. It scales six process capability levels defined on an ordinal scale, which starts from incomplete to optimizing processes. Process maturity has been a core component of COBIT for more than a decade [42].

TIPA is the result of more than ten years of research work, including experimentation on how to combine ITIL with the ISO/IEC 15504 [43]. TIPA uses the generic approach for process assessment published by the ISO in ISO/IEC 15504-2 – Process Assessment (now ISO/IEC 33000) (ISO/IEC 15504-1/2, ISO/IEC 33000). TIPA is a standards-based approach to ITIL (v2, v3 and v3 2011) assessment that can address challenges (posed by improving the quality of product manufacture or IT processes) in several important ways by providing a repeatable, consistent method for conducting process assessment [44].

TIPA for ITIL PAM and COBIT PAM are based on ISO/IEC 15504 (ISO/IEC 15504-1, 2004; ISO/IEC 15504-2, 2003). It means that they both rely on the same foundation (ISO/IEC 15504), which is a global reference for conducting process capability assessments. From an assessment perspective, both TIPA for ITIL and COBIT 5 PAM break down each
process into base practices specific to each process and take into account generic practices, which are not restricted to any particular process.

COBIT 5 PAM base practices are specific to COBIT processes and ensure proper governance and management of Enterprise IT while TIPA for ITIL base practices are specific to ITIL and guarantee the proper execution of the process to support the service delivery in line with customer needs.

3.4. Semantic Similarity

There is extensive literature regarding the measure of the similarity between documents and sentences. It is broadly used in different fields such as natural language processing, artificial intelligence, and data mining [45].

An effective similarity measure should be able to determine whether two sentences are semantically equivalent, considering the variability of natural language expression [46].

Several similarity measures could be considered when measuring sentences and words:

- **String similarity measures** operate on string sequences and character composition. A string metric is a metric that measures similarity or dissimilarity (distance) between two text strings for approximate string matching or comparison [47].
- **Knowledge-Based measures** determine the degree of similarity between words using information derived from semantic networks [47].
- **Corpus-Based measures** determine the similarity between words according to information gained from large corpora. A Corpus is an extensive collection of written or spoken texts that are used for language research [47].

3.4.1 String-Based Measures. String-Based measures can be partitioned into two types of measures: character-based and term-based measures [47]. Character-based measures operate on the chars of the strings being measured [47]. There are several character-based measures as, for example, Longest Common SubString (LCS), N-gram, Damerau-Levenshtein, Jaro, among others. On the other hand, Term-based similarity measures operate over concepts. There are also several term-based similarity measures such as Cosine distance, Euclidean Measures, Jaccard similarity, among others [47].

3.4.2 Knowledge-Based Measures. Knowledge-based similarity measures can be divided into two groups, more specifically, measures of semantic similarity and measures of semantic relatedness. Semantically similar concepts are deemed to be related based on their likeness. Semantic relatedness covers a broad range of relatedness between words such as is-a-kind-of, is-a-part-of kind or is-the-opposite-of [47].

3.4.3 Corpus-Based Measures. Corpus-based semantic representations exploit statistical properties of textual structure to embed words in a vector space. In this space, terms with similar meanings tend to be located closer to each other. These methods rely on the idea that words with similar meanings tend to occur in similar contexts [48].

There are a large set of Corpus-Based Similarity measures, as, for example, the Latent Semantic Analysis (LSA) and the Hyperspace Analogue to Language (HAL) measures [47].

Word2vec models have become especially popular in embedding generations [48]. These models are trained to reconstruct linguistic contexts of words. Word2vec takes as input a large corpus of text and produces a vector space with each unique word in the corpus being assigned to a corresponding vector in the space. Word2vec methods have an advantage in handling large datasets since they do not consume as much memory as some classic methods like LSA.

4. Proposal

In this section, the authors explain how they used semantic similarity techniques to compare COBIT 5 and TIPA for ITIL v3 PAMs.

The authors started by implementing and testing String-based measures and applying a range of computational techniques that are normally used for analyzing and representing naturally occurring texts such as Sentence Splitting and Tokenization, Removal of Stop words, Stemming, Parts-of-Speech Tagging and Lowercased tokens.

Sentence Splitting and Tokenization is the process of chopping the sentences into pieces, called tokens, and at the same time throw away certain pieces, such as the punctuation. Stop words are common words which would appear to be of little value as the words ‘the’, ‘is’, ‘at’ and ‘on’. Stemming is a technique for transforming words in a way that if they represent the same meaning, they are captured by the same token. Parts-of-speech Tagging relates to the large amount of information that tags give about a word and its neighbors [49].

The authors tested several String-based algorithms such as the Longest Common SubString (LCS), the Cosine and the Jaccard Distance algorithms. The Wordnet was also used as a lexical reference system.
However, these similarity measures and techniques have some limitations when used in certain contexts. For example, the variability of natural language expression and word context in sentences are ignored. Therefore, more advanced techniques and measures were considered to increase the robustness of this proposal. This means that the authors then used some Corpus-Based measures. Due to their strong dependence on the availability of an ontology [50], which falls out of the scope of this research, the authors did not apply any Knowledge-based measures.

Recently, neural-network language embeddings have received increasing attention [48]. So, the authors searched for embedding models to apply in this research. From the research the authors discovered spaCy. spaCy is an open source software library for advanced Natural Language Processing. The library is published under the MIT license and currently offers statistical neural network models for several languages such as English, German, French, and various others.

The similarity between sentences is calculated by comparing word vectors or “word embeddings” that are multi-dimensional meaning representations of a word. When processing natural language, spaCy starts by tokenizing the text to produce a Doc object. The Doc is then processed in several different steps called ‘processing pipeline’.

The pipeline used mainly consists of a tagger, a parser, and an entity recognizer.

- Tagger assigns part-of-speech tags, which is the process of marking up a word in a text as corresponding to a particular part-of-speech (noun, verb, etc.) based on its definition and context.
- Parser assigns dependency labels.
- Named Entity Recognition detects and labels named entities, that seeks to locate and classify named entities in text into pre-defined categories such as the names of persons, organizations, locations, and so forth.

Figure 1 describes the spaCy processing pipeline.

![Spacy Processing Pipeline](image)

Figure 1 Spacy Processing Pipeline

The sentences to be evaluated are all written in the English language. Therefore, the authors used spaCy English models to calculate the semantic similarity between sentences. Several English models are offering different techniques for natural language processing. spaCy en_core_web_sm is the simplest and the smallest model. It uses context vectors in the pipeline processing to generate vectors by locally clustering the context for each word. It is very efficient and fast. By using this model, the authors can calculate the similarity between processes in 15 minutes while with larger models it takes 4 to 5 hours.

However, according to the spaCy documentation, the results are not as good as larger models results. Therefore, despite these issues, the authors also used a larger model (en_core_web_md) in the sentence semantic similarity calculation to improve the results.

As stated before, the primary objective of this research is to perform an integration, based on a semantic evaluation, of COBIT and ITIL process assessment core concepts (base practices, inputs/outputs, and outcomes) since the individual processes are described regarding the process name, purpose, and outcomes. Also, the process dimension of the process assessment model provides information in the form of:

- Base practices (BPs) for the process, defining the tasks and activities needed to accomplish the process purpose and fulfill the process outcomes. Each BP is explicitly associated with a process outcome.
- Input and output work products (WPs) associated with each process and related to one or more of its outcomes.
- Characteristics associated with each WP.

In this research, the authors decided to follow a bottom-up approach as presented in Figure 2.

![Bottom-up approach to the calculation of the processes overall similarity](image)

Figure 2 Bottom-up approach to the calculation of the processes overall similarity.

In the calculation of the similarity scores between Base Practices, Outcomes, and Work Products, the authors adopted a pairwise similarity comparison between sentences. Pairwise comparison is defined as a
process of comparing entities in pairs. Therefore, the authors compared these three core concepts of COBIT 5 PAM and TIPA for ITIL.

The similarity score for each core concept (Base Practice, Outcomes, and Work Product) is calculated by the average of the similarity scores between the highest similarity score of each concept instantiation. An example of this situation is illustrated in Figure 3:

This means that, in a hypothetical case, the similarity score for the Base Practices concept can be the average of the similarity between Base Practice A -> Base Practice’ A + Base Practice A -> Base Practice’ B + Base Practice A -> Base Practice’ C + Base Practice A -> Base Practice’ D + Base Practice A -> Base Practice’ E since Base Practice A is the one that has the highest similarity scores with all the Base Practices’. Therefore, the overall similarity score for the processes is then calculated as follows, in which PS is the process similarity score, WPS is the work product similarity score; BPS is the base practice similarity score and OS is the outcome similarity score. In this research, the authors considered that all the core concepts have the same weight (1/3).

\[ PS = WPS \times \frac{1}{3} + BPS \times \frac{1}{3} + OS \times \frac{1}{3} \]

Sentences with similar meanings should score near 100 percent, and sentences with different meanings should score near 0 percent.

5. Demonstration

The initial focus on any process assessment should be the core (sometimes called primary) processes, which are primarily part of the (in COBIT 5) Build, Acquire and Implement (BAI) and Deliver, Service and Support (DSS) domains [42].

In this section, the authors demonstrate the proposal by assessing the semantic similarity between the COBIT 5 DSS02 Manage Service Requests and Incidents process and its related (according to the ISACA publication) ITIL Incident Management process. Figure 4 presents the base practices with higher semantic similarity scores.

![Figure 4 Base practices results](image1)

Figure 4 Base practices results

Figure 5 presents the work products with higher semantic similarity scores.

![Figure 5 Work Products results](image2)

Figure 5 Work Products results

Figure 6 presents the outcomes with higher semantic similarity scores.

![Figure 6 Outcomes results](image3)

Figure 6 Outcomes results

6. Evaluation

Reasoning about the semantic relatedness of natural language expressions is regularly performed by humans but remains a challenge for computers [51]. When evaluating the relatedness between texts, humans do not
judge relatedness merely at the level of text words, but at a much deeper level that manipulates concepts. Thus, human’s interpretation of a word in a document is far from their background knowledge and experience [33]. Humans have an innate ability to judge semantic relatedness of texts. Therefore, human judgment on a set of text pairs can thus be considered correct by definition, a kind of “gold standard” against which computer algorithms are evaluated [51].

In this specific research, the authors compared the results obtained in this research with a benchmark proposed by ISACA, as one can see in Figure 7.

![Figure 7 Mapping of COBIT and ITIL processes. Adapted from [53].](image)

According to the spaCy documentation, the results are better when using larger models. Therefore, in this section, the authors just present the results achieved by using the larger model. The main issue that exists when using a larger model is that the average computational time takes 4 to 5 hours, whiles small models only take 10 to 15 minutes. Therefore, one should carefully analyze the performance vs. precision of the results.

Figure 8 presents the results using the spaCy library taking into account the following formula:

\[ PS = WPS \times \frac{1}{3} + BPS \times \frac{1}{3} + OS \times \frac{1}{3} \]

As one can observe, the correlation with the human mappings presented in Figure 7 is: 79.60% (Manage Operation -> Access Management); 81.80% (Manage Service Requests and Incidents -> Incident Management); 82.50% (Manage Service Requests and Incidents -> Request Fulfilment); 83.20% (Manage Problems -> Problem Management); 83.80% (Manage Security Services -> Access Management) and 84.00% (Manage Business Process Controls -> Access Management).

### 6.1 Lessons learned

As previously stated, human judgment is considered by definition a kind of “gold standard”. Therefore, one can argue that the results are quite good since 4/6 of the mappings are in line with the experts’ opinion (Figure 7). This means that just the Manage Operation -> Access Management mapping and the Manage Service Requests and Incidents -> Incident Management mapping are not in line with the experts’ opinion.

The results are quite similar, even when talking about different processes. This is not strange since we are talking about processes of the same domain that use similar language to describe their base practices, outcomes, and work products. This means that just a few words change in the descriptions (however, these changes have a clear impact on the meaning but are not enough when using semantic similarity techniques).

Therefore, the language used to describe and explain these concepts should be thoroughly reviewed by experts to facilitate the understanding and, in that way, the differentiation of the processes by the practitioners.

Also, a larger dataset is needed to refine the results. Therefore, one should create an “IT Governance” dictionary that will take into account some particularities of the EGIT field in order to improve the results.
7. Conclusion

In this research, the authors proposed a novel approach to the EGIT field by proposing the use of semantic similarity techniques to address the terminological disparities between ITIL V3 2011 and COBIT 5 that makes best professional judgment the only approach currently available to understand the overlapping between them.

The semantic similarity techniques are broadly used in different fields such as natural language processing, artificial intelligence, and data mining. However, it is rather unusual to use these techniques in the EGIT field. The results were then compared against a benchmark proposed by human experts. The results are quite good since 4/6 of the mappings are in line with the experts’ opinion. This means that the techniques used during this research hit 2/3 of the mappings proposed by the specialists. Taking into account the subjectivity of the field, one can argue that the results are encouraging signs that can allow us to, in the near future, map different EGIT frameworks using semantic similarity techniques rather than human working groups.

This research also has some limitations. First of all, just one domain was analyzed. Therefore, more work should be done regarding other processes. Also, since the approach proposed in this paper is relatively new in the EGIT field, it is difficult to compare it to other proposals.

Regarding future work, it should include several directions such as (a) improving and extending the dataset, this means, analyze other COBIT 5 domains, and (b) improving the proposed approach via investigating other techniques and algorithms.

Also, it would be interesting to analyze the evolution of the results by giving different weights to the core concepts. For example, one can assign a ½ weight for the outcome concept and ¼ for the base practices and work products concepts.

Finally, forming a focus group to extract qualitative information about the obtained results would be interesting.

8. References
