iKeyCriteria: A Qualitative and Quantitative Analysis Method to Infer Key Criteria since a Systematic Literature Review for the Computing Domain

Mayra Carrión-Toro [1], Jose Aguilar [2,3], Marco Santórum [1], María Pérez [1], Boris Astudillo [1], Cindy-Pamela Lopez [1], Marcelo Nieto [1] and Patricia Acosta-Vargas [4,*]

1 Departamento de Informática y Ciencias de la Computación, Escuela Politécnica Nacional, Quito 170525, Ecuador; mayra.carrion@epn.edu.ec (M.C.-T.); marco.santorum@epn.edu.ec (M.S.); maria.perez@epn.edu.ec (M.P.); boris.astudillo@epn.edu.ec (B.A.); cindy.lopez@epn.edu.ec (C.-P.L.); ronnie.nieto01@epn.edu.ec (M.N.)
2 Centro de Estudios en Microelectrónica y Sistemas Distribuidos, Universidad de Los Andes, Mérida 5101, Venezuela; aguilar@ula.ve
3 GIDITIC, Universidad EAFIT, Medellín 050022, Colombia
4 Intelligent and Interactive Systems Laboratory, Universidad de Las Américas, Quito 170125, Ecuador
* Correspondence: patricia.acosta@udla.edu.ec

Abstract: A systematic literature review is a synthesis of the available evidence, in which a review of quantitative and qualitative aspects of primary studies is carried out, to summarize the existing information regarding a particular topic. The researchers extract key criteria from papers collected about their study area, answering research questions and conducting document analysis. Nonetheless, in some cases, these criteria are improperly justified, unknowing their true level of importance in the study subject. Hence, an additional study is necessary to explain the criteria relevance in the papers studied using qualitative and quantitative premises. The correct identification of these key criteria is a critical factor in prioritizing and achieving appropriate results in any scientific research work. In our paper, a new method to determine key criteria from a literature review is proposed, composed of three components: input-process-output. First, the inputs are a set of criteria to evaluate and a set of documents to analyze. Next, the process component examines the document set to indicate whether the criteria to be analyzed are found. The process component produces a Boolean matrix, which is the input of the mathematical logic process that will get the key criteria considered necessary and sufficient as the output component. The iKeyCriteria method has been applied in different computing domains, particularly for serious games design and virtual organizations, giving positive results in each context. Finally, we developed an online tool that provides global support to the execution of our method.

Keywords: data; information retrieval; systematic literature review; key criteria; meta-analysis

1. Introduction

A systematic literature review (SLR) is a research methodology designed to answer a focused research question. This methodology is becoming an important resource for researchers to help in the information search and solve problems in different research areas. Their findings enable them to learn about advancements, characteristics and challenges and thus create or improve existing proposals in specific fields. Currently, we find its application in all study areas [1].

An SLR is a way of evaluating and interpreting all the available research relevant to a particular research domain, thematic area, or phenomenon of interest [2]. For example, Cochrane reviews [3] summarize the results of available and carefully designed studies for controlled clinical trials and provide a high level of evidence on the efficacy of health

interventions. According to [4], an SLR allows the synthesis of research findings to discover areas where further research is needed. In general, an SLR must define research questions, search strings, inclusion and exclusion criteria, quality criteria, among others, and then retrieve literature that helps answer the research questions.

SLRs are carried out in the scientific field, as sufficient analysis to extract key criteria in their study field through answers to research questions and document analysis. Nonetheless, in some cases, these criteria are not adequately justified, and as a result, their relevance is revealed in the papers examined. These facts do not allow us to contrast the level of importance of the study subject. For this reason, an additional study is necessary to explain the criteria relevance in the papers studied using qualitative and quantitative foundations. This method would significantly speed up the research process by generating new proposals or solutions in various study fields. The lack of methods or processes to simplify the relevance of a criteria selection delays the research results and adequate justifications.

We consider it critical to review the literature early in the research process to identify gaps in a specific study area and, based on these findings, to propose new supported knowledge.

In our work, the final documents of the SLR are the input to carry out the qualitative and quantitative analysis to obtain key criteria in a specific domain. This article aims to be a beginning point for researchers interested in applying an SLR methodology, which helps to highlight and identify key criteria in each field of research. Namely, a novel qualitative and quantitative analysis method is proposed here, which allows us to know key criteria related to a specific area of knowledge, which has been tested, particularly in the computing domain.

This proposal is an extension of a regular SLR, since the TF-IDF (term frequency-inverse document frequency) metric is applied, which is related to the frequency of a term in a set of papers, with a qualitative analysis considering researchers' opinion, to determine key patterns using a set of logical axioms, of which key elements are obtained to be considered in each area. The quantitative method based on the TF-IDF metric was automated based on the collected papers from the SLR. The automation process can manage papers in many languages.

The qualitative and quantitative analysis method was applied in different case studies, such as the following, to obtain key criteria in the areas of serious games [5], user-centered design [6], and virtual organizations area for defining a characterization under the Industry 4.0 context [7].

The main contribution of this paper is to help researchers with the selection and verification of the relevance of key criteria in different areas of knowledge, allowing them to obtain the relevance of a criterion in a document as well as in a collection of documents. With these results obtained from key criteria, the researcher can know what elements are important and useful for the generating of new knowledge in his study area. The paper structure is as follows; Section 2 presents related work; Section 3 describes our proposal to obtain the necessary criteria for an SLR; Section 4 demonstrates the automation of our proposal's quantitative analysis process; Section 5 demonstrates a case study in the domain of methodologies for serious games design, and Section 6 reports conclusions and future work.

2. Related Works

The impact of emerging technology is essential [8] when integrating communication through different interactive media. For this reason, a summary of different articles related to the key criteria selection to improve decision-making in different research areas is presented.

In [9], the authors propose a criteria identification to evaluate the health literature critically, which is obtained from an SLR as well as the experts' knowledge consideration, to define a priority.

The authors present a study to contribute to the selection of environmentally friendly ecological suppliers that seek to provide efficient and simple alternatives to take advantage of organic waste from the community in general based on key ecological selection criteria [10]. In this document, the importance of a set of criteria is observed, which allows for making decisions about suppliers that meet the ecological characteristics. However, in this study, the selected criteria are based on literature analysis of general environmental

factors, allowing for the development of common ecological standards. Also, the experts' experience gives weight to the criteria to obtain the most important in the ecological area. Our methodological proposal based on qualitative and quantitative analysis is helpful, for this research since the authors could define procedures to determine basic criteria from an SLR to select green suppliers.

Other authors [11] present criteria selection using text mining approaches. As a result of an SLR on text mining techniques, papers were published. In this article, the criteria definition begins from an SLR observed to avoid the subjectivity presented in some cases.

In contrast, [12] describes a process to identify criteria that aid in the selection of a production system compatible with manufacturing companies. The authors explain the procedure for choosing these criteria: a literature search is carried out first, identifying the criteria. Also, they add criteria obtained from the experts' opinions, and the criteria are validated using a Delphi method. In this work, the method to determine criteria for a specific topic is the main subject, using a Delhi method to solve ambiguities. Meanwhile, we consider a similar approach to criteria validation using logical axioms to obtain the criteria pattern necessary in a specific context.

In earlier research, the authors [13], proposed a method to find selection criteria for Enterprise Resource Planning (ERP)/Customer Relationship Management (CRM) systems used by Pymes in Poland. The method collects information from online surveys and analyses using Spearman's Rank Order Correlation. To find the key selection criteria to ensure the optimal purchase decision of the correct ERP either or both CRM system. An SLR on the selection criteria of ERP either or both CRM systems was conducted, and the opinions of 83 respondents with work experience in Pymes who had previously implemented such systems were considered. According to the authors, the study results can be applied in a similar context related to decision-making in Pymes for the selection of the CRM either or both ERP class system.

In summary, researchers always begin with an SLR when selecting criteria for their works, and the researchers' opinion is used to prioritize, determine relevance, and more. Thus, experts' opinions or mathematical methods are used to prioritize the criteria. However, each of the articles analyzed presents a general criteria selection process, starting from an SLR, without presenting a complete and detailed process that visualizes each activity to be generated to find the selection criteria. Thereby, researchers could replicate their proposals for the criteria selection processes, helping in different areas of specific knowledge.

An SLR answers a defined research question by collecting and summarizing all empirical evidence that fits pre-specified eligibility criteria. A meta-analysis is the use of statistical methods to summarize the results of these studies.

In our case, we applied a rigorous meta-analysis using a statistical method to summarize the results of these studies. Our proposal presents some characteristics to consider for relevant criteria selection in a specific study area, which can be replicated in any context of computer science.

Table 1 shows what aspects are considered with the proposed method. As we see, not all the studies analyzed contemplate these elements when selecting or presenting relevant criteria. They always present their results via SLR.

Table 1. iKeyCriteria method characteristics.

	SLR	Criterion Relevance in a Document	Criterion Relevance in a Collection of Documents	Criteria Categorization According to Relevance- Use Truth Tables	Selection or Inference of Key Criteria
			Characteristics iKeyCriteria Methodology		
[10]	x				x
[12]	x				x
[11]	x				x
[9]	x				x
[13]	x				x

3. Methodological Proposal to Infer Key Criteria

Our method is composed of three components: input-process-output. The process illustrated in Figure 1 presents the activities to be carried out to identify key criteria.

First, the inputs are a set of criteria to evaluate and a document set to analyze. Then, the process offers two paths: one is based on the researcher's opinion who analyzes the document sets to indicate if the criteria to be analyzed are found; the other is based on statistical analysis, which considers the frequency of occurrence of the criteria in the document sets.

The two paths allow the generation of a Boolean matrix, which serves as the input of the mathematical logic process, and finally will allow obtaining the output related to the key criteria considered necessary and sufficient.
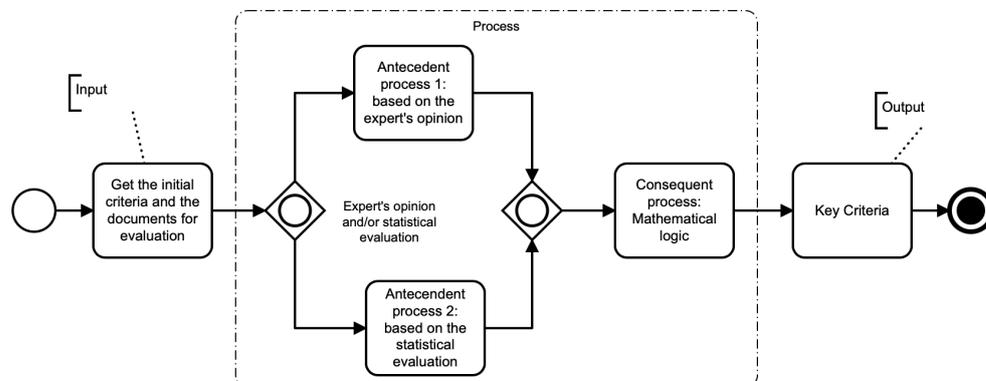
Figure 1. The input-process-output approach.

The following section describes in detail each part of the process of the Figure 1.

3.1. Inputs

The base documents often include identical results and certain differential characteristics that we seek to identify and classify through statistical analysis. The initial criteria are the basic arguments or keywords from independent primary studies obtained from a review of several bibliographical sources of information (documents of the context to be studied) focused on the same question. When the study context is clear, documents or books serve as a base with a set of concepts or key terms, that enclose that domain, and each of the key criteria arises as input to the process. The initial criteria selection can be based on the experience of the researcher, or based on a source of scientific information.

For example, if we want to validate the characteristics that a software requirement must fulfill to be well-formulated, we can start from the individual and group characteristics of a requirement proposed by the ISO/IEC/IEEE 29148 standard [14]. In this case,

the criteria to validate would be appropriate, unambiguous, complete, unique, feasible, verifiable, correct, and compliant. Moreover, the set of base documents would be all software engineering standards related to requirements.

3.1.1. Set of Initial Criteria

The initial criteria must be organized in an analysis matrix. This matrix is the input of the methodological proposal and allows for defining the initial criteria, synonyms, and understanding the context of the term. The columns of the analysis matrix, as shown in Table 2, are as follows.

- Criteria. The initial criteria to validate.
- Definition. A description of the exact meaning of the criterion.
- Context. It provides additional information to allow readers to understand a criterion.
- Synonyms. Terms related to the criteria that allow eliminating ambiguities and identifying the number of occurrences in the collection of documents.

Table 2. Initial criteria matrix.

Criteria	Definition	Context	Synonyms
Term 1	Description of the term	What to look for? Term explanation	Related terms and synonyms
...	...	...	...
Term n	Description of the term n	What to look for? Term explanation	Related terms and synonyms

3.1.2. Set of Documents to Evaluation

The set of documents for evaluation is obtained through an SLR that provides a complete and comprehensive summary of relevant research studies related to the research questions. Our method requires making two groups of documents (P, Q).

- The set of documents named P should group works related to the study context.
- The set of documents named Q that should group works related to the opposite study context.

3.2. Processes

The methodological process is composed of two preceding processes or antecedent processes and one consequent process (seen Figure 1). One of the two antecedent processes can be chosen. The antecedent process 1 is based on expert opinion assumes objectivity or scientific impartiality; the role of the researcher becomes instrumental since anyone would reach the same conclusions if they followed the same steps. However, it is impossible to deny intuition and individual discernment, so the antecedent process 2 is based on the TF-IDF metric, which is based on the number of times a term appears in the document.

3.2.1. Antecedent Process 1 Based on the Expert's Opinion

The first activity carried out in this process is to obtain the justification matrix (see Table 3) that argues the presence, or lack thereof, to finally obtain the Boolean matrix (see Table 4), whose entries are either $0s$ or $1s$, which indicates the absence or presence of the initial criteria on the selected documents from the SLR, based on the expert's opinion. It is made for the groups of documents P and Q.

Table 3. Justification matrix.

	P		Q	
CRITERIA	Document 1P	Document 3P	Document nQ	Document nQ
Term 1	The term appears in the document (Yes or No justification)	Justification	Justification	Justification
...	...	...	...	...
Term n	(Yes or No justification)	Justification	Justification	Justification

Table 4. Boolean matrix.

	P		Q	
CRITERIA	Document 1P	Document 2P	Document 3Q	Document nQ
Term1	$0 \vee 1$	$0 \vee 1$	$0 \vee 1$	$0 \vee 1$
...	...	...	...	...
Termn	$0 \vee 1$	$0 \vee 1$	$0 \vee 1$	$0 \vee 1$

3.2.2. Antecedent Process Based on the `TF-IDF` Metric

`TF-IDF` is a metric to determine the importance of a word in a collection of papers. In our case, this metric will be necessary to identify the importance of the initial criteria to the set of documents. `TF-IDF` value increases proportionally to the number of times a term appears in the document and is offset by the number of documents in the corpus that contain the term, which helps to adjust for the fact that some terms appear more frequently in general. This process starts with the determination of the term-frequency matrix. It aims to obtain the number of times (f) that a term (t) occurs in a document (d) (it is stored in the term-frequency matrix (see Table 5)). It counts each term and its synonyms. However, in the beginning, we use a reduced analysis matrix (see Table 6), which considers the criteria and their synonyms.

Table 5. Term frequency matrix.

	P		Q	
CRITERIA	Document 1P	Document Pn	Document 1Q	Document Qn
Term 1	$tf_{(t,d)}$	$tf_{(t,d)}$	$tf_{(t,d)}$	$tf_{(t,d)}$
...	...	...	...	...
Term n	$tf_{(t,d)}$	$tf_{(t,d)}$	$tf_{(t,d)}$	$tf_{(t,d)}$

Table 6. Reduced analysis matrix.

Criteria	**Synonyms**
Term 1	Related terms and synonyms
...	...

After obtaining the term frequency matrix, the normalized and the inverse frequency are calculated to obtain the `TF-IDF` value. Thus, the normalized frequency matrix is obtained (see Table 7), which is calculated using Equation (1).

Normalized frequency. The normalized frequency adjusts the frequency of the term or relevance score to normalize the effect of document length on the document ranking. It is obtained by dividing the absolute frequency value by the value of the maximum frequency of the term contained in the document.

Equation (1) allows obtaining the content of the normalized frequency table for each term in the document.

$$tf(t_i, d_j, a) = a + (1 - a) * \left(\frac{tf_{(t_i,d_j)}}{tf_{max}(d_j)} \right) \tag{1}$$

where,

- $tf(t_i, d_j, a)$, represents the value obtained from normalized frequency.
- The constant $a = 0.5$ is a value that softens the frequency function of the term and leads to its normalization, recommended by Christopher D. Manning [15] .
- tf_{t_i,d_j}, represents the absolute frequency of the term (*t*) in the document (*d*).
- $tf_{max}(d_j)$, represents the raw frequency of the most frequently occurring term in the document.

Table 7. Normalized frequency matrix.

	P		Q	
CRITERIA	Document 1P	Document Pn	Document 1Q	Document Qn
Term 1	$tf(t_i, d_j, a)$	$tf(t_i, d_j, a)$	$tf(t_i, d_j, a)$	$tf(t_i, d_j, a)$
...	...	...	...	...
Term n	$tf(t_i, d_j, a)$	$tf(t_i, d_j, a)$	$tf(t_i, d_j, a)$	$tf(t_i, d_j, a)$

Also, the inverse frequency matrix is obtained (see Table 8).

Inverse frequency. The inverse document frequency is a measure of whether the term is common or not in the collection of documents. It is obtained by dividing the total documents number by the documents number that contains the term, and the logarithm of this quotient is taken:

The following Equation (2) will be used.

$$idf(t_i, D) = log\left(\frac{n(D)}{df_{t_i}} \right) \tag{2}$$

where

- $n(D)$, refers to the total documents in the collection.
- df_t, refers to the total documents where the term appears.
- The $idf(t_i, D)$ is performed for each term.

Table 8. Inverse document frequency matrix.

Criteria	Valor idf
Term 1	$idf(t_i, D)$
...	...
Term n	$idf(t_i, D)$

Finally, the TF-IDF matrix is obtained (see Table 9).

TF-IDF matrix. It is the product of two measures, multiplying the values obtained in the normalized frequency $tf(t_i, d_j, a)$ by the values of the inverse frequency $idf(t_i, D)$ (see Equation (3)).

$$tfidf(t_i, d_j, a, D) = tf(t_i, d_j, a) * idf(t_i, D) \tag{3}$$

Table 9. TF-IDF matrix.

	P	Q
CRITERIA	Document 1P	Document 1Q
Term 1	$tfidf(t_i, d_j, a, D)$	$tfidf(t_i, d_j, a, D)$
...	...	...
Term n	$tfidf(t_i, d_j, a, D)$	$tfidf(t_i, d_j, a, D)$

With the values obtained in the TF-IDF Matrix, we proceed to create a Boolean matrix (0*s* or 1*s*). For this matrix, the variable k (k = average or variance) is used that represents the average of the frequency of the TF-IDF value. Then, if the value obtained $tfidf(t_i, d_j, a, D) \geq k$, then the value in this matrix will be (1), otherwise the value of (0) is placed (see Table 10).

It is important to mention that the mode, mean, median, and variance can be considered more than the average, which can be selected according to the end-user needs; the option is chosen to finally generate the Boolean matrix.

Table 10. TF-IDF Boolean matrix.

	P		Q	
CRITERIA	Document 1P	Document nP	Document 1Q	Document nQ
Term 1	$0 \vee 1$	$0 \vee 1$	$0 \vee 1$	$0 \vee 1$
...	...	...	...	...
Term n	$0 \vee 1$	$0 \vee 1$	$0 \vee 1$	$0 \vee 1$

3.2.3. Consequent Process

It is the inference process and aims to obtain the matrices of instantiation, behavior, and pairing. It is based on logical axioms to extract the criteria.

The objective of this process is to infer the key criteria. The boolean matrices obtained through the antecedent processes are used as inputs. Next, the process of obtaining the final key criteria is described.

Instantiation Matrix

This matrix is used to specify whether the criterion is in the document being analyzed, and with the pair of values identified by rows of the matrix, it will help to identify through the truth table if it is in any category of the logical pattern matrix.

The instance matrix is made up of two columns (initial criteria and documents *P* and *Q*) (see Table 11). This matrix is made for each term o criteria to be evaluated.

Criteria Column: The values of the row of the boolean matrix of each criterion are placed (antecedent processes 1 or 2). It is necessary to build a matrix for each term evaluated.

Documents Column: In this column, 0 or 1 is placed depending on the study context; 1 (one) when they belong to the set *P* related to the study context, and 0 (zero) for the set of documents *Q* related to the opposite study context.

Table 11. Instantiation Matrix.

Criteria "Term 1"	Document $(P-1)(Q-0)$
$0 \vee 1$	1 ($P1$)
$0 \vee 1$	1 ($P2$)
...	...
$0 \vee 1$	0 ($Q1$)
$0 \vee 1$	0 ($Q2$)
...	...

3.3. Output

Behavior Matrix

This matrix serves to categorize and determine which criteria are relevant in the study context and is validated with the help of truth tables.

The behavior matrix is made up of two sections. (See Table 12). The first section corresponds to a truth table of mathematical logic composed of a column for each input variable (for example, p and q). The second section shows all the possible results of the comparison of the pair of values (criteria, document), of the instantiation matrix. Here in the truth table that we use, p and q mean the following.

p : are the criteria which need their existence (or lack thereof) to be verified in the documents

q : are documents which are either from the study context (P) or the opposite context (Q).

These (p, q) can take $1's$ or $0's$ as appropriate; this pair of values is verified in the instantiation matrix (criteria, documents). If a pair of values from the truth table is found in the instantiation matrix it will be true (1), otherwise it is false, and the value of (0) is placed.

Then, the result of the second section is compared with the matrix of logical patterns. (see Table 12).

Table 12. Behavior matrix.

Truth Table		Result-Criteria
p	q	
1	1	1 true, 0 false
1	0	1 true, 0 false
0	1	1 true, 0 false
0	0	1 true, 0 false

For example, to better understand its logic we have the following.

$p \to q$. If p then q.

The researcher specifies, "If I found the key criterion in my study document, this could be a document from my study context." Such a statement is known as conditional.

p: criterion

q: study context

In such a way that the statement can be expressed as $p \to q$.

Its truth table is as follows (see Table 13):

Table 13. True table.

p	q	$p \to q$
1	1	1
1	0	0
0	1	1
0	0	1

The interpretation of the results of the truth table is as follows:

Analyzing if the researcher lied with the statement of the previous statement: when $p = 1$ means that the key criterion was found in the study document and $q = 1$ that this document is from the study context, therefore $p \to q = 1$ (the researcher told the truth).

When $p = 1$ and $q = 0$, it means that $p \to q = 0$ the researcher lied, since I found the key criterion in the study document, but this document is not from my study context.

When $p = 0$ and $q = 1$, it means that although he did not find the key criterion in the study document, the document does belong to his study context, so he did not lie, such that $p \to q = 1$.

When $p = 0$ and $q = 0$, it means that although the criterion was not found in the study document, that document is not from his study context, therefore $p \rightarrow q = 1$ since he did not lie either.

Considering the previous statements, to achieve the result, we apply mathematical logic that will allow us to classify the criterion into a category.

Table 14 presents the logical axioms that will allow for classifying the criteria. It comprises two sections; the first presents the truth table, and the second shows logical patterns useful for specifying the necessary and sufficient conditions.

Table 14. Logic patterns matrix.

Logic Patterns						
Truth Table		**Necessary and Not Sufficient Condition**	**Not Necessary and Sufficient Condition**	**Necessary and Sufficient Condition**	**Not Necessary and Not Sufficient Condition**	
p	q	$q \rightarrow p$	$p \rightarrow q$	$p \leftrightarrow q$	$(p \wedge q)$	$(p \wedge q) \rightarrow p$
1	1	1	1	1	1	1
1	0	1	0	0	0	1
0	1	0	1	0	0	1
0	0	1	1	1	0	1

Next, we describe each of the categories in which a criterion can be categorized:

1. Necessary and not sufficient (N-NS). $q \rightarrow p$: Criterion appears in all documents of the study context and criterion appears at least once in the opposite study context.
2. Not necessary and sufficient (NN-S) $p \rightarrow q$: Criterion appears at least once in the study context and does not appear in any document from the opposite study context.
3. Necessary and sufficient (NS) $p \leftrightarrow q$: Criterion appears in all documents of the study context and does not appear in any document of the opposite study context.
4. Not necessary and not sufficient (NN-NS) $(p \wedge q) \rightarrow q$: Criterion appears at least once in the study context and also appears at least once in all documents of the opposite study context.
5. None: The criterion was not in any of the categorized groups in the logical pattern matrix.

Pairing Matrix. This matrix is used to categorize all the previously analyzed criteria according to a logical pattern of behavior. See Table 15.

Table 15. Pairing matrix.

	Criterion 1	**Criterion 2**	**Criterion ... n**
	1	1	1
	1	0	0
	0	1	0
	1	1	1
Logical Pattern	Necessary and not sufficient	Not necessary and sufficient	Necessary and sufficient

This method was applied in different study cases helping researchers applying `SLR` to determine important criteria in different case studies [5–7].

4. Automation of the Method

In order to simplify the use of our methodological proposal, we developed a software tool that supports the method's execution and saves time. The tool uses both the $tf - idf$ metric to obtain the Boolean matrix and the researcher's analysis to go directly to the mathematical logic and obtain the categorized results. The following section describes the tool in depth.

4.1. Description of the Tool

Our web application is divided into three modules: information, administration, and the process itself.

1. The informative module has two interfaces: the first at the beginning of the application, which contains a description of the method, and the second contains the tool manual.
2. The user administration module has an interface where the user's information is displayed, but its main task is to record each of the processes carried out by the researcher. It focuses on two states of the process, one when it is completed, in which case the application saves the PDF files, the criteria, and the results, and the other when it is running. This module automatically saves each change of the current state of the process, for example, when a file or criterion is added or deleted, or when the variable k changes the calculation method.
3. The application module or the process itself consists of four sub-modules described as follows:

   - Uploading PDF files: this module consists of four interfaces that allow the researcher to upload to the application the files that will be used in the process. The first two interfaces allow files to be uploaded to each of the two research groups established by the researcher. The following interface displays the files uploaded in each group, and the last one is an informative interface for the correct upload of the files to the application.
   - Uploading criteria and synonyms: this module consists of two text inputs that allow for entering the criteria and all synonyms related to that criterion. All the criteria and synonyms entered by the researcher are shown in a table.
   - Selects the type of process that allows for selecting process 1 or process 2. Process 1 is based on the researcher's opinion; here, the Boolean matrix that the researcher obtained from his analysis is used. Process 2 will use the $tf - idf$ metric and then select the type of measure k with which the comparison will work in the next step of the process; otherwise, when selecting process 1, it will go to the next step directly to obtaining the result.
   - Select type of measurement: this module allows for selecting between the median, mode, mean, and variance to calculate the variable "k" used to obtain the Boolean matrix.
   - Results: The last module of the application shows the results obtained, which are shown in a pie chart with criteria categorized and all the matrices obtained in the process. Additionally, the results obtained can be exported in an excel file.

4.2. Implementation

The application was implemented in the Django framework, a high-level Python Web Framework that encourages rapid development.

Django was chosen as the tool's implementation framework since it is a high-level Python web framework that enables the building of secure and maintained websites. It is a high-level language and well suited for scientific and engineering environments. Python is widely used in statistical data analysis, automation, and the development of dependable, scalable systems using modules that make programming easier and have a low learning curve. Python is a high-level language and well suited for scientific and engineering environments [16,17].

The result of the processes carried out in the method are matrices; Python has libraries such as Numpy that provide high-performance multidimensional array objects that significantly enhance performance and speed up the execution time correspondingly.

The user administration module was developed using Django. It provides a framework for sessions and user management as well as the use of `SQLite` as a database.
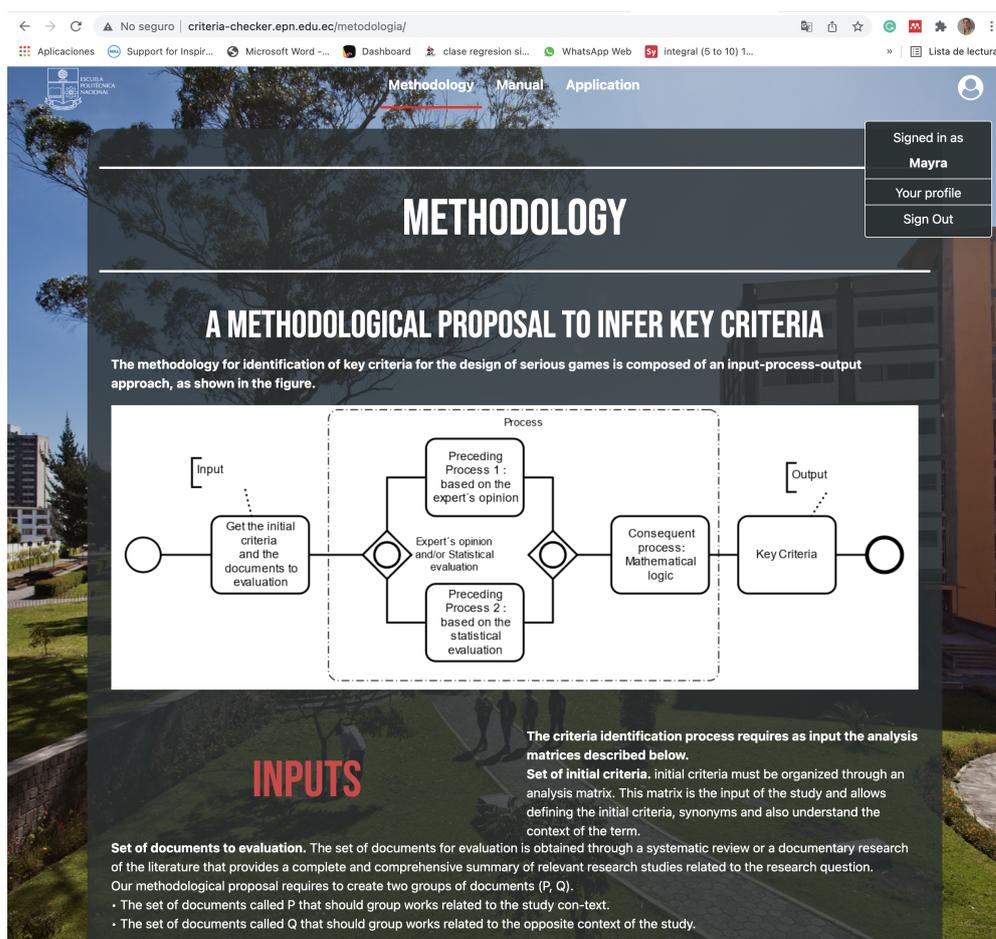
Figure 2. Criteria checker application.

Our tool, called, Criteria checker application (see Figure 2), is installed on the servers of the research laboratory and is available at the web address http://criteria-checker.epn.edu.ec, accessed on 24 April 2022.

5. Application in a Case of Study

5.1. Application Context

The general purpose of this study case is to determine the key criteria for creating educational serious games from the study of software development methodologies and serious games. In this research, the main study context is identified, such as the study of serious games development methodologies, and the opposite context is the study of the traditional software development methodologies, so that the research will be based on the use of these contexts and the application of the Antecedent Process based on the `tf-idf` metric.

This case study begins by following the steps proposed by the Kitchenham methodology [2], in which an `SLR` is carried out whose specific purpose is to obtain the set of documents that meet the search criteria, and that goes through the review process using inclusion and exclusion criteria.

Once the review is executed, the documents found as a result are 26 related papers. This resulting group of documents constitutes the input of our process.

5.2. Initial Criteria

The initial documentary review focuses on the areas of interest, where, within the literature, terminologies referring to the main contexts are identified, which can even be synonymous.

Based on our application context, if the research needs to find information regarding "the key criteria for the creation of educational serious games", several criteria can be defined for each of the main search terms, which are "methodology", "lifecycle", "agile process", "pedagogical objectives", "gamification", and "GamePlay", among others; these initial criteria can be related to synonymous terms such as methodologies, methods, processes, principles, tools, mechanisms, strategies, designs, and more.

The final criteria selection is strongly based on the experience, knowledge, and skills of the researcher, which will define whether or not the words apply to the context, and in turn, will discard those that are not or whose combination of these is not used, traditionally or professionally, in the study field. For example, a combination of words unlikely to be found would be information referring to a "design for the manufacture of serious games".

For the context of the application, the selected criteria are those set out in Table 16.

Table 16. Selected criteria for the application of the method.

Criteria	Synonyms	Criteria	Synonyms
methodology	method	tests	evaluation
lifecycle	stages	roles	character
agile process	scrum	pedagogical objectives	educational aims
resources	digital resources	gamification	ludification
story	narrative	product quality	property
prototyping	mock-up	quality of use	usability
modeling language	machine language	gamePlay	actions

Our study case determined that, of the 26 resulting documents, 12 papers were classified for group *P* and the remaining 14 papers for group *Q*.

5.3. Using the Support Tool

1. The 12 documents related to group *P* are loaded, as shown in Figures 3 and 4.
   The documents of the main context used in this study case are [18–28].
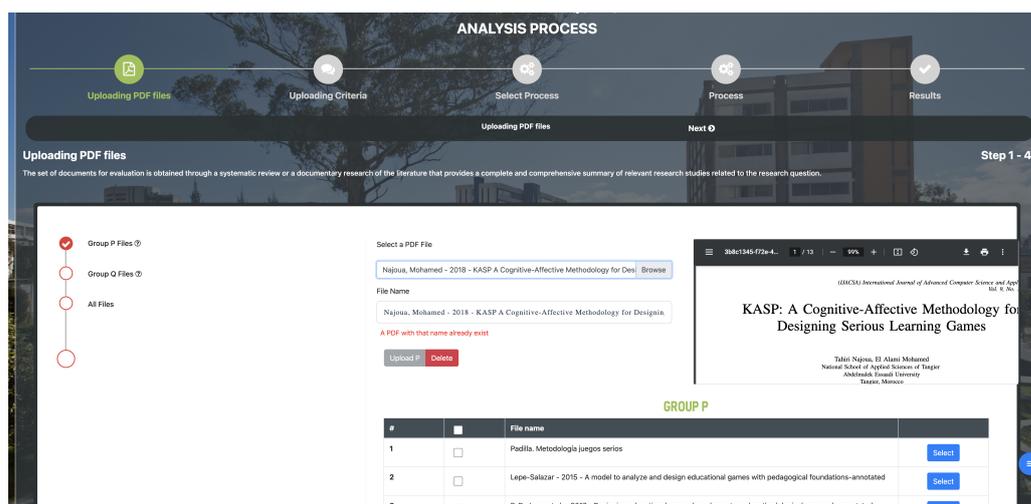
Figure 3. Uploading group P files.
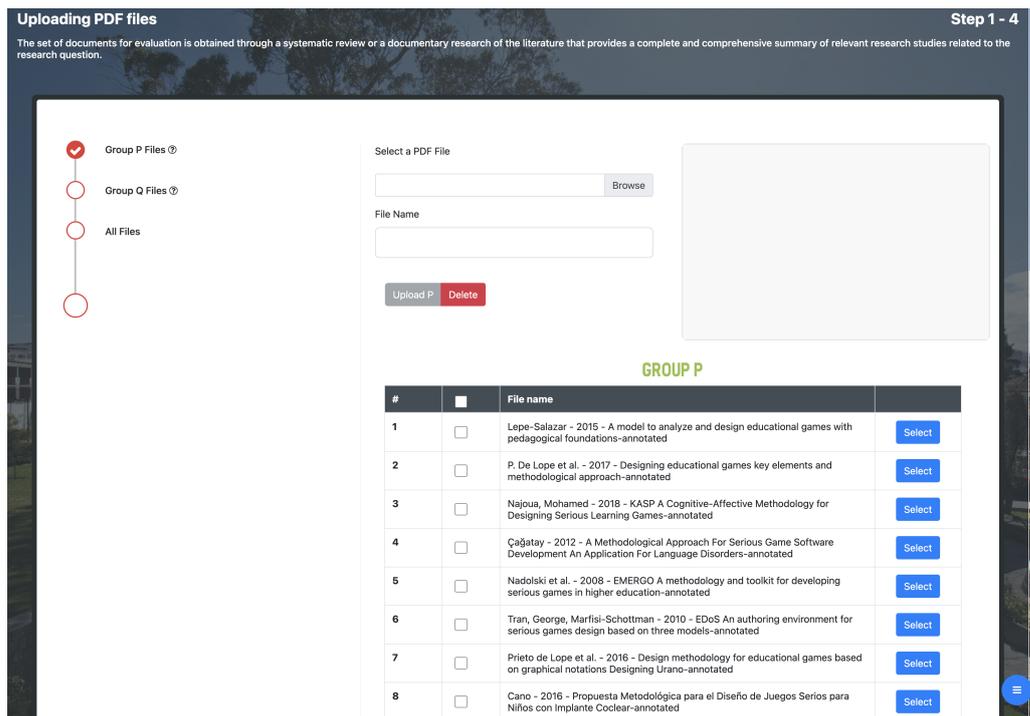
Figure 4. Group P files uploaded.

2. The 14 documents related to group *Q* are loaded, as shown in Figure 5. The documents of the opposite context are [29–42].
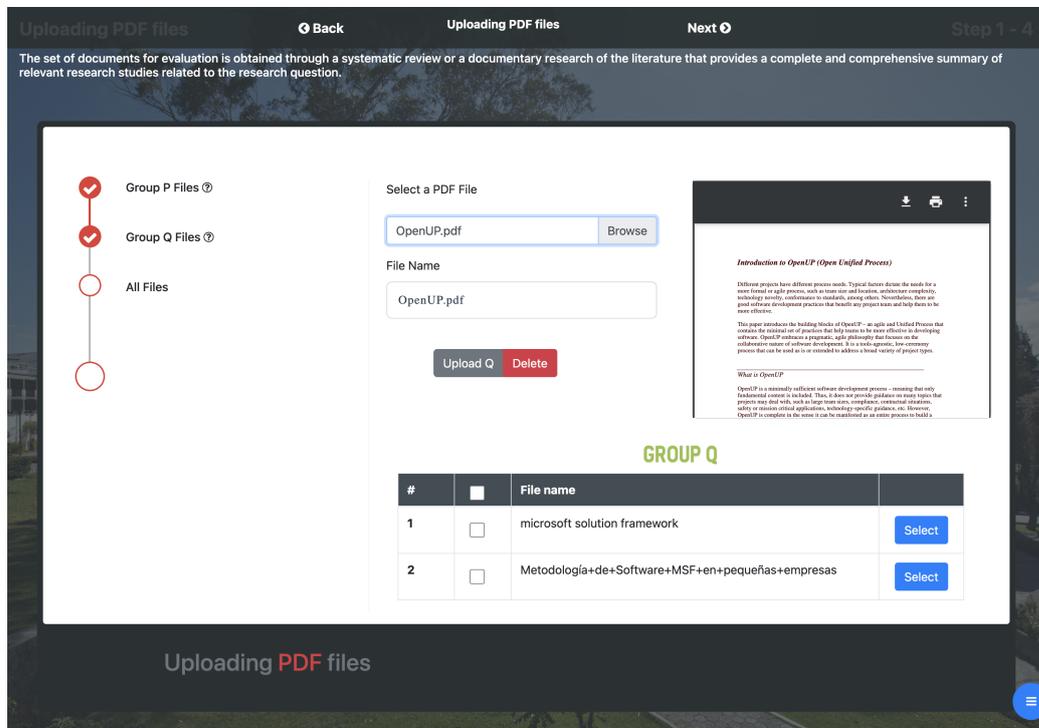
Figure 5. Group Q files uploaded.

3. Is necessary to verify that all the 26 papers have been uploaded without problems, as shown in Figure 6.
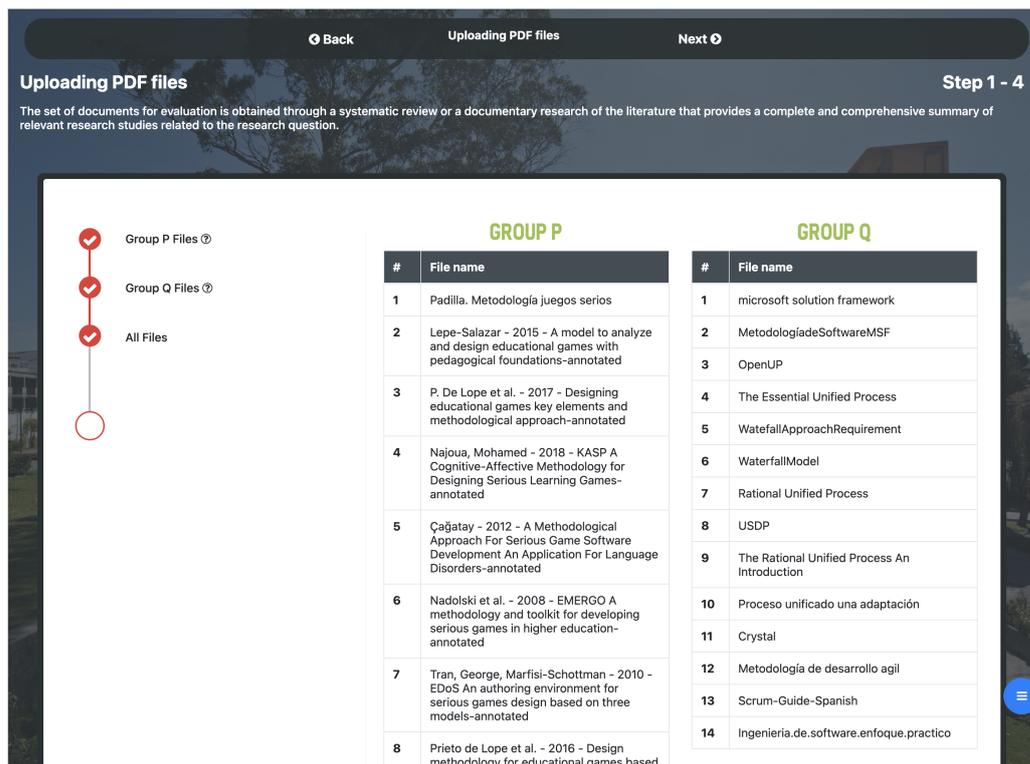
Figure 6. Group P and Group Q files successfully uploaded.

4. Each of the defined criteria with their respective synonyms is entered, as shown in Figure 7.
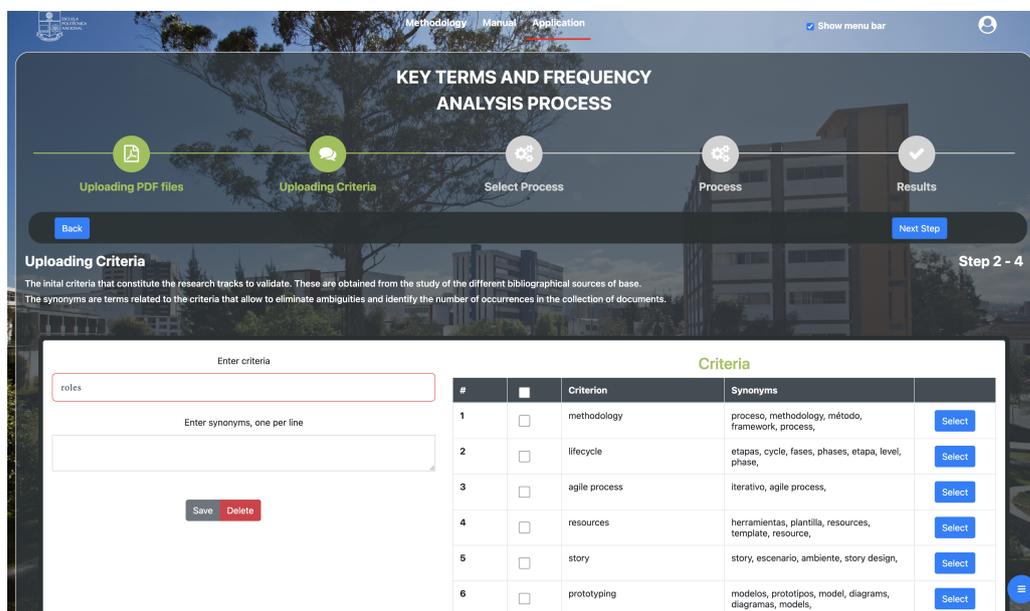
Figure 7. Uploading criteria.

5. Next, is necessary to choose between the antecedent process 1, to measure the expert's opinion, or the antecedent process 2, to apply the $tf - idf$ metric.

We select the `TF-IDF` antecedent process 2. If the obtained value of $tfidf(t_i, d_j, a, D) \geq k$, then the value in this matrix will be (1), otherwise the value of (0) is placed. k represents the (median, mean, mode and variance) of the frequency of the `TF-IDF` value.

Process 2, or `TF-IDF` metric, allows selecting the type of measurement to obtain the Boolean table; for our study case, the "mode" measurement was chosen (Figure 8),

once the app starts finish the calculation of the results, the app shows all the results explained in Section 3.2.2 related to the execution of the TF-IDF method, as shown as an example in Figures 9–11.
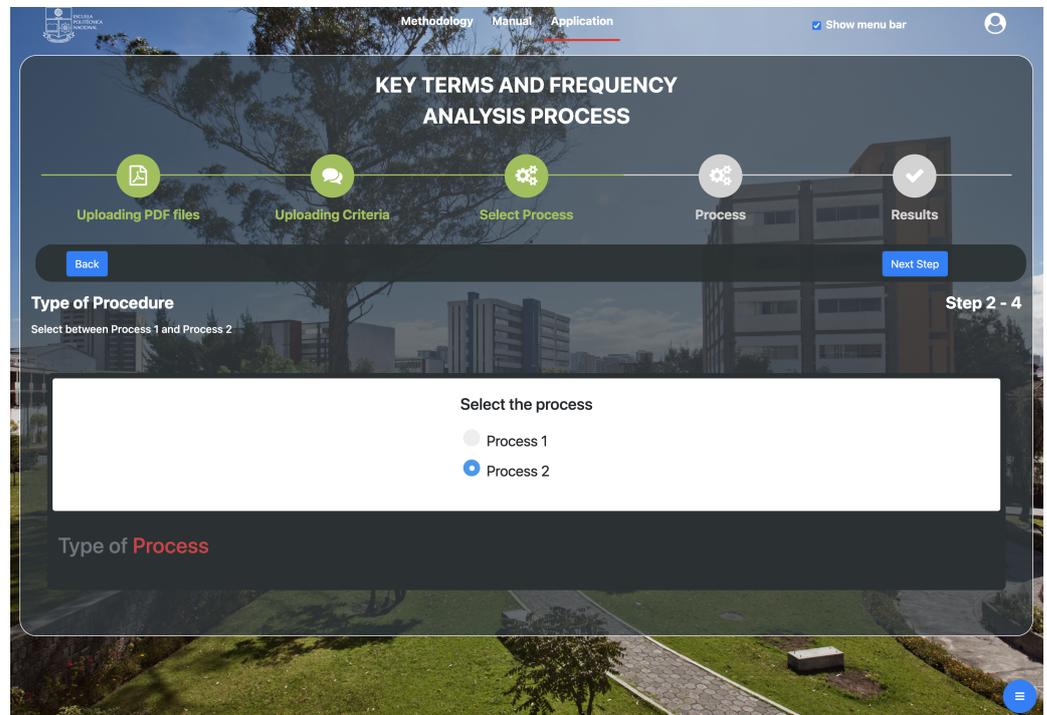
Figure 8. Choice of the antecedent process.

Figure 9. Result tables.

For details of result tables check the Mendeley dataset repository in [43].

Result Tables

Absolute Term Frequency Matrix　Hide

								GRUPO P		
#	Criterios/Documents	Lepe-Salazar - 2015 - A model to analyze and design educational games with pedagogical foundations-annotated	P. De Lope et al. - 2017 - Designing educational games key elements and methodological approach-annotated	Najoua, Mohamed - 2018 - KASP A Cognitive-Affective Methodology for Designing Serious Learning Games-annotated	Çağatay - 2012 - A Methodological Approach For Serious Game Software Development An Application For Language Disorders-annotated	Nadolski et al. - 2008 - EMERGO A methodology and toolkit for developing serious games in higher education-annotated	Tran, George, Marfisi-Schottman - 2010 - EDoS An authoring environment for serious games design based on three models-annotated	Prieto de Lope et al. - 2016 - Design methodology for educational games based on graphical notations Designing Urano-annotated	Cano - 2016 - Propuesta Metodológica para el Diseño de Juegos Serios para Niños con Implante Coclear-annotated	ProcesoJu
1	methodology	52	61	81	323	75	22	107	661	6
2	lifecycle	14	78	26	87	18	30	81	205	1
3	agile process	0	0	0	0	0	0	0	5	
4	resources	4	1	3	0	32	31	6	25	7
5	story	34	18	6	24	2	20	215	75	1
6	prototyping	23	39	51	30	4	63	39	372	1
7	modeling language	1	6	5	105	2	5	1	157	
8	tests	1	13	4	4	0	0	2	298	
9	roles	18	51	17	36	27	21	40	421	1
10	pedagogical objectives	317	20	317	85	70	160	43	81	8
11	gamification	1	2	7	0	0	0	2	0	
12	product quality	0	0	0	0	0	0	0	38	
13	quality of use	1	6	0	5	3	0	6	15	2

Figure 10. Absolute term frequency matrix (see Table 5).

Normalized frequency (TF) Matrix　Hide

| #	Criterios/Documents	Lepe-Salazar - 2015 - A model to analyze and design educational games with pedagogical foundations-annotated	P. De Lope et al. - 2017 - Designing educational games key elements and methodological approach-annotated	Najoua, Mohamed - 2018 - KASP A Cognitive-Affective Methodology for Designing Serious Learning Games-annotated	Çağatay - 2012 - A Methodological Approach For Serious Game Software Development An Application For Language Disorders-annotated	Nadolski et al. - 2008 - EMERGO A methodology and toolkit for developing serious games in higher education-annotated	Tran, George, Marfisi-Schottman - 2010 - EDoS An authoring environment for serious games design based on three models-annotated	Priet 2 me edu bas nota Ura
1	methodology	0.582018927444795	0.891025641025641	0.6277602523659306	1.0	1.0	0.56875	0.748
2	lifecycle	0.5220820189274448	1.0	0.5410094637223974	0.6346749226006192	0.62	0.59375	0.688
3	agile process	0.5	0.5	0.5	0.5	0.5	0.5	
4	resources	0.5063091482649842	0.5064102564102564	0.5047318611987381	0.5	0.7133333333333334	0.596875	0.513
5	story	0.5536277602523659	0.6153846153846154	0.5094637223974764	0.5371517027863777	0.5133333333333333	0.5625	
6	prototyping	0.5362776025236593	0.75	0.5804416403785488	0.5464396284829721	0.5266666666666666	0.696875	0.590
7	modeling language	0.501577287066246	0.5384615384615384	0.5078864353312302	0.6625386996904025	0.5133333333333333	0.515625	0.502
8	tests	0.501577287066246	0.5833333333333334	0.5063091482649842	0.5061919504643962	0.5	0.5	0.504
9	roles	0.5283911671924291	0.8269230769230769	0.526813880126183	0.5557275541795665	0.6799999999999999	0.565625	0.593
10	pedagogical objectives	1.0	0.6282051282051282	1.0	0.631578947368421	0.9666666666666667	1.0	
11	gamification	0.501577287066246	0.5128205128205128	0.5110410094637224	0.5	0.5	0.5	0.504
12	product quality	0.5	0.5	0.5	0.5	0.5	0.5	
13	quality of use	0.501577287066246	0.5384615384615384	0.5	0.5077399380804953	0.52	0.5	0.513
14	gameplay	0.5331230283911672	0.5064102564102564	0.5	0.5046439628482973	0.5	0.5	0.504

Figure 11. Normalized frequency matrix (see Table 7).

The results of the analysis process using *k* (median) are detailed in Table 17 and shown in Figure 12.

Table 17. Key Criteria Matrix based on TF-IDF metric.

Criteria	Logical Pattern
Methodology	None
lifecycle	None
agile process	None
resources	None
story	Sufficient and not Necessary
prototyping	None
modeling language	None
tests	None
roles	None
pedagogical objectives	Sufficient and not Necessary
gamification	None
product quality	None
quality of use	None
gameplay	None
display devices	None
interaction devices	None
feedback	Not Necessary and Not Sufficient
difficulty settings	None
usability	None
accesability	None
visual elements	Not Necessary and Not Sufficient
sound elements	None
initial test	None
context help	None
adaptability	None
learning techniques	Sufficient and not Necessary
avatar	None
context centered design	None
user centered design	None
end users	None
experts	None
linguisic game	None
creative techniques	None
consensus mechanisms	None
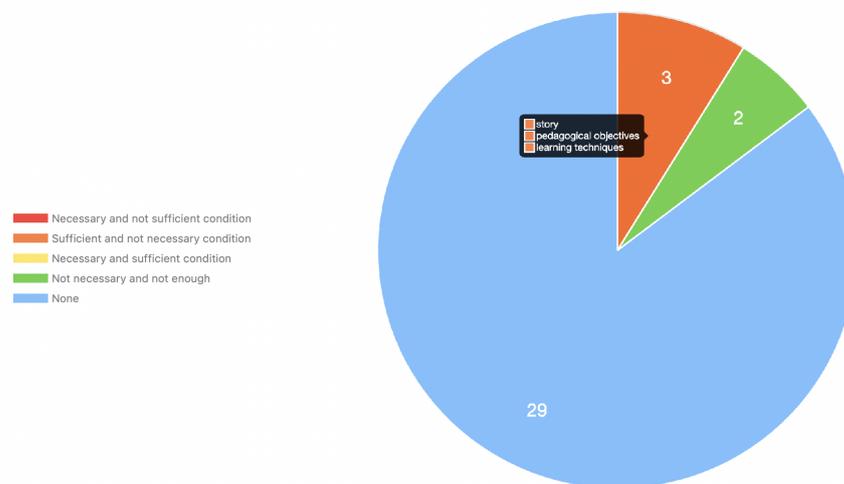
Figure 12. Key Criteria Results.

6. Conclusions

This work proposes a new method that allows the selection of key criteria of a study field, which helps decision-making in that research field.

In most of the studies analyzed, the criteria selection starts from the literature review, allowing researchers to obtain grounded criteria. However, these are not contrasted and validated in the same information set, which would help in determining whether those criteria are really important.

In this study context, we propose an add-on to a systematic literature review, giving researchers an alternative way to find the relevant criteria, characteristics, or congruent patterns to create and propose new ideas.

Our method offers a rigorous meta-analysis using a statistical method to consider relevant criteria selection in a specific study area, which can be replicated in any context. The proposed `TF-IDF` metric is based on a qualitative and quantitative analysis that results in the same frequency analysis table, which leads to a mathematical process to obtain necessary and sufficient criteria as a result, or categorize them in different states, as was presented in the method.

Different study cases have been carried out using the qualitative and quantitative analysis methods, allowing criteria selection in areas like serious games, user-centered design, serious games methodologies, and requirements.

In order to simplify the use of our method, we developed a software tool that supports the method's execution and saves time. However, a tool limitation is that the documents must be unlocked for the analysis.

The web application allows us to validate the criteria defined by the expert using mathematical logic criteria and determine if there is a need to include new criteria to satisfy the systematic literature review performed.

Automation of this method helps researchers obtain a validated criterion in papers that study a specific research domain.

The method applied in this study can serve as a basis for application to any topic that requires validation of the systematic literature review of previous studies. In future work, we suggest further validation of this application in case studies to assess the needs of other researchers, improve the needs of other researchers, and improve based on the feedback.

This study did have some limitations. The first issue is associated with the possible ambiguities of the terms, so it was necessary to include synonymous words and terms in another language. Another limitation, as indicated, in the case of path 1, is that it is based on the subjective opinion of the researcher, which in some cases is advantageous since it can break an ambiguity of terms and analyze whether the document effectively refers to the study context; the other path is based on a statistical process that does not give rise to subjectivity, and is based on the occurrence of the terms, but given possible ambiguity, it might not be as precise. Moreover, it is important to ensure that the documents were studied and did not have security blocks.

In future works, we plan to continue testing this method to obtain feedback and improve the process and its automation, applying it in different case studies to obtain key criteria for both gamification elements in software development related to data analytics and requirement validation in software development.

Finally, this methodology can serve as a starting point for future work in order to show relevant results and trends of research on the subject studied.

Author Contributions: Conceptualization, M.C.-T., J.A., M.S., B.A., C.-P.L., P.A.-V. and M.P.; methodology, M.C.-T., J.A. and M.S.; validation, M.C.-T., M.S. and M.N.; formal analysis, M.C.-T. and M.S.; investigation, M.C.-T. and M.S.; writing—original draft preparation, M.C.-T., M.S., B.A. and P.A.-V.; writing—review and editing, M.C.-T., M.S., J.A., P.A.-V., B.A. and M.P.; supervision, J.A.; M.P. project administration, P.A.-V. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lame, G. Systematic literature reviews: An introduction. In Proceedings of the International Conference on Engineering Design, ICED, Bordeaux, France, 24–28 July 2019; Volume 2019, pp. 1633–1642. [CrossRef]
2. Kitchenham, B.; Charters, S. Guidelines for performing Systematic Literature reviews in Software Engineering. *Engineering* **2007**, *45*, 1051. [CrossRef]
3. Ibrahim, A.; Ahmed, M.; Conway, R.; Carey, J.J. Risk of infection with methotrexate therapy in inflammatory diseases: A systematic review and meta-analysis. *J. Clin. Med.* **2019**, *8*, 15. [CrossRef] [PubMed]
4. Snyder, H. Literature review as a research methodology: An overview and guidelines. *J. Bus. Res.* **2019**, *104*, 333–339. [CrossRef]
5. Carrión, M.; Santorum, M.; Pinaida, A.; Aguilar, J. Estudio para inferir criterios clave para el diseño de Juegos Serios. In Proceedings of the 2019 International Conference on Information Systems and Software Technologies (ICI2ST), Quito, Ecuador, 13–15 November 2019; pp. 63–70.
6. Quimbita, L.; Santorum, M. Estudio de Metodologías Participativas y de Enfoques Centrados en el Usuario para la definición de una Metodoogía de Diseño de Juegos Serios Educativos. Ph.D. Thesis, Escuela Politécnica Nacional, Quito, Ecuador, 2020.
7. Guamushig, T.; Lopez, C.; Santorum, M. Análisis y Caracterización de las Organizaciones Virtuales para la Colaboración en el Contexto de la Industria 4.0. Ph.D. Thesis, Escuela Politécnica Nacional, Quito, Ecuador, 2020.
8. Pavlič, J.; Tomažič, T.; Kožuh, I. The impact of emerging technology influences product placement effectiveness: A scoping study from interactive marketing perspective. *J. Res. Interact. Mark.* **2021** doi: 10.1108/JRIM-02-2021-0041 [CrossRef]
9. Sale, J.E.; Brazil, K. A strategy to identify critical appraisal criteria for primary mixed-method studies. *Qual. Quant.* **2004**, *38*, 351–365. [CrossRef]
10. Banaeian, N.; Mobli, H.; Nielsen, I.E.; Omid, M. Criteria definition and approaches in green supplier selection—A case study for raw material and packaging of food industry. *Prod. Manuf. Res.* **2015**, *3*, 149–168. [CrossRef]
11. Hashimi, H.; Hafez, A.; Mathkour, H. Selection criteria for text mining approaches. *Comput. Hum. Behav.* **2015**, *51*, 729–733. [CrossRef]
12. Dohale, V.D.; Akarte, M.M.; Verma, P. Determining the Process Choice Criteria for Selecting a Production System in a Manufacturing Firm Using a Delphi Technique. In Proceedings of the International Conference on Industrial Engineering and Engineering Management (IEEM), Macao, China, 15–18 December 2019; pp. 1265–1269.
13. Cieciora, M.; Bołkunow, W.; Pietrzak, P.; Gago, P. Online Journal of Applied Knowledge Management Key criteria of ERP/CRM systems selection in SMEs in Poland. *Online J. Appl. Knowl. Manag.* **2020**, *8*, 85–98. [CrossRef]
14. International Organization for Standardization. *Systems and Software Engineering—Life Cycle Processes—Requirements Eengineering Inginierie*; International Organization for Standardizatio: Geneva, Switzerland, 2018. [CrossRef]
15. Manning, C.D.; Raghavan, P.; Schutze, H. *An Introduction to Information Retrieval*; Cambridge University Press, Cambridge, England, 2009; p. 569. [CrossRef]
16. Esteves, R.A.; Wang, C.; Kraft, M. Python-based open-source electro-mechanical co-optimization system for MEMS inertial sensors. *Micromachines* **2022**, *13*, 1 [CrossRef]
17. Django. Django Documentation. Available online: https://docs.djangoproject.com/en/4.0/ (accessed on 16 May 2022).
18. Bousbia, S.; Miladi, E.; Kooli, Z.; Boumaiza, W. Proposal of a methodology of serious games' demystification for the teaching of technical modules. *IEEE Glob. Eng. Educ. Conf. EDUCON* **2016**, *10–13*, 1155–1159. [CrossRef]
19. Mehmet, C. A methodological approach for serious game software development: An application for language disorders. Master Thesis, Atilim University, Ankara, Turkey, 2012.
20. Cano, S.P.; González, C.S.; Collazos, C.A.; Arteaga, J.M.; Zapata, S. Agile software development process applied to the serious games development for children from 7 to 10 years old. *Int. J. Inf. Technol. Syst. Approach* **2015**, *8*, 64–79. [CrossRef]
21. Cano, S.P. Propuesta Metodológica para el Diseño de Juegos Serios para Niños con Implante Coclear. Ph.D. Thesis, Universidad del Cauca, Cauca, Colombia, 2016.
22. Lepe-Salazar, F. A model to analyze and design educational games with pedagogical foundations. In Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology—ACE '15, Iskandar, Malaysia, 16–19 November 2015; pp. 1–14. [CrossRef]
23. Nadolski, R.; Hummel, H.; Van den Brink, H.; Hoefakker, R.; Slootmaker, A.; Kurvers, H.; Storm, J. EMERGO: A methodology and toolkit for developing serious games in higher education. *Simul. Gaming* **2008**, *39*, 338–352. [CrossRef]
24. Najoua, T.; Mohamed, E.A. KASP: A Cognitive-Affective Methodology for Designing Serious Learning Games. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*, 2–11. [CrossRef]
25. De Lope, R.P.; Medina-medina, N.; Soldado Montes, R.; Mora García, A.; Gutiérrez-vela, F.L. Designing educational games: Key elements and methodological approach. In Proceedings of the 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games), Athens, Greece, 6–8 September 2017; pp. 63–70. [CrossRef]
26. Padilla, N. Metodología Para El Diseño De Videojuegos Educativos Sobre Una Arquitectura Para El Análisis Del Aprendizaje Colaborativo. Ph.D. Thesis, Universidad de Granada, Granada, Spain, 2011.
27. De Lope, R.P.; Medina-Medina, N.; Paderewski, P.; Gutierrez-Vela, F.L. Design methodology for educational games based on interactive screenplays. *CEUR Workshop Proc.* **2015**, *1394*, 90–101.

28. Tran, C.D.; George, S.; Marfisi-Schottman, I. EDoS: An authoring environment for serious games design based on three models. In Proceedings of the 4th European Conference on Games Based Learning 2010, ECGBL, Copenhagen, Denmark, 21–22 October 2010; pp. 393–402.

29. Kuhrmann, M.; Ternité, T. Including the Microsoft Solution Framework as an agile method into the V-Modell XT. *Eval. Nov. Approaches Softw. Eng.* **2006**,

30. Arévalo, W.; Atehortúa, A. Metodología de Software MSF en pequeñas empresas MSF software methodology in small businesses. *Cuad. Act.* **2012**, *4*, 83–90.

31. Balduino, R. Introduction to OpenUP (Open Unified Process). *Organization* **2007**, 1–9.

32. Jacobson, I.; Ng, P.W.; Spence, I. The essential unified process. *Dr. Dobb's J.* **2006**, *31*, 40–45.

33. Ru'el, H.J.; Bondarouk, T.; Smink, S. The Waterfall Approach and Requirement Uncertainty. *Int. J. Inf. Technol. Proj. Manag.* **2010**, *1*, 43–60. [CrossRef]

34. Petersen, K.; Wohlin, C.; Baca, D. The waterfall model in large-scale development. *Lect. Notes Bus. Inf. Process.* **2009**, *32*, 386–400. [CrossRef]

35. Company, R. *Rational Unified Process for Systems Engineering*; Company, R.: New York, NY, USA, 2005.

36. Zuehlke Engineering. *Unified Software Development Process*; Technical Report; Zuehlke Engineering. Available online: http://www0.cs.ucl.ac.uk/staff/ucacwxe/lectures/3C05-01-02/aswe5.pdf (accessed on 28 April 2022)

37. Kruchten, P. *The Rational Unified Process An Introduction,* 2nd ed.; Addison-Wesley Professional: Boston, MA, USA, 2000; p. 320.

38. Leyva, E.; González, M. An Unified Process adaptation for the development of Joomla based web applications; In *Ciencias Holguín*; Centro de Información y Gestión Tecnológica de Santiago de Cuba, Holguín, Cuba, 2006; Volume XII, pp. 1–13.

39. Abreu, R.B.; Guntín, Y.O.; Alfonso, Y.Á.; Mena, J.C. Metodología ágil Crystal Clear. Un caso de estudio. *Ser. Cient.* **2009**, *2*, 3.

40. Duarte, A.O. Las Metodologías de Desarrollo Ágil como una Oportunidad para la Ingeniería del Software Educativo. *Rev. Av. Sist. Inform.* **2008**, *5*, 159–171.

41. Schwaber, K.; Sutherland, J. La Guía definitiva de Scrum: Las Reglas del Juego, Scrum Guides, 2013. Available online: https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-Spanish-Latin-South-American.pdf (accessed on 24 April 2022).

42. Tobergte, D.R.; Curtis, S. Ingenieria de Software un enfoque practico. *arXiv* **2013**, arXiv:1011.1669v3.

43. Carrión-Toro, M. "tf-idf metric iKeyCriteria (dataset)", Mendeley Data, V1. Available online: https://data.mendeley.com/datasets/jc8k4cnr2c/1 (accessed on 16 May 2022). [CrossRef]