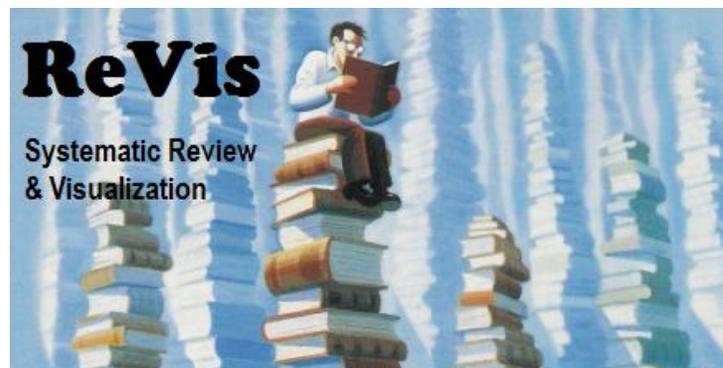




UNIVERSITY OF SÃO PAULO
ICMC - *INSTITUTE OF MATHEMATICS AND COMPUTER SCIENCE*
DEPARTMENT OF COMPUTER SCIENCE
P.O. Box 668 - 13560-970 – São Carlos, SP, Brazil

Revis: User's Manual

Systematic Literature Review Supported by
Visual Analytics



April / 2012

Summary

Systematic Literature Review Supported by Visual Analytics (Revis)	3
Installation	3
How create your set of documents (primary studies) - Input	3
Projects	2
2.1 Create a new Project.....	2
2.1.1 Quality Criteria.....	5
2.1.2 Inclusion / Exclusion Criteria.....	5
2.1.3 PDF's Files.....	5
2.2 Open a Project.....	6
VTM Selection	7
3.1 The Views	7
Stages of Selection Activity	8
4.1 Classification Activity	8
Features of Revis Tool	12
4.2 Neighborhood relationship (called RN connection).....	12
4.3 Topic	13
4.4 Occurrence frequency of an expression.....	14
4.5 Coordination.....	15

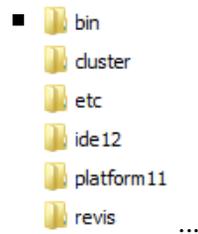
User's Manual

Systematic Literature Review Supported by Visual Analytics (Revis)

The Revis is a flexible visualization tool that has several text handling facilities, which allows for a VTM (Visual Text Mining) exploration of a collection of documents. It is an open source tool that has been developed at the University of São Paulo. The Revis supports the selection of primary studies in systematic literature reviews.

Installation

Firstly, you must downloading the program “Revis” in <http://infoserver.lcad.icmc.usp.br/infovis2/Revis>
After, you must decompress the file. The content will be:



Next, open the folder “bin” and run the file  `revis.exe`

How create your set of documents (primary studies) - Input

There is one input file which can be used on the Revis. This file can be imported by the tool. Following the format of this file is described, detailing its purpose and format.

```
@ARTICLE{1, (a)
author      = {R. Martins, K. Felizardo, ...}, (b)
title       = {Title of article 1}, (c)
year        = {2012}, (d)
abstract    = {Abstract of article 1}, (e)
references  = {Title of reference 1 (f)
              Title of reference 2
              ...
              }
}
```

```

@ARTICLE{2,                                     (a)
author      = {R. Minghim, J. C. Maldonado, ... }, (b)
title       = {Title of article 2},              (c)
year        = {2012},                            (d)
abstract    = {Abstract of article 2},          (e)
references  = {Title of reference 1              (f)
              Title of reference 2
              ...
              }
}

```

Corpus.bib

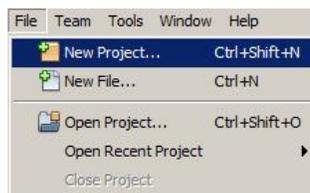
The input file, with extension .bib, is used to create the projection views. In this file, the first line (a) contains the names of the authors of the paper. The authors' names can be written as the first letter of the author's name following by the final station more the last name. One author of other is separated by comma. The second part (b) describes the title of the paper. The third parte (c) is the year of the paper. Then, there is the abstract (d) and the remaining lines (e) are the references of the paper. Each reference is described only by its title, one per line. Above, an example of the input file.

Clue: You should use the **JabRef Tool** to organize your set of documents (primary studies). It is an open source bibliography reference manager.

Projects

Before you start (or continue) selecting the primary studies, it is necessary to create (or open) a project. In this section we discuss how to do this.

2.1 Create a new Project



- **New Project:** This option allows the user to create a new project (systematic review) in the application in three steps:

- The first step is to select the "New Projection" option in file menu (Figure 1 – a).

Figure 1: Menu Option - New Project (Systematic Review)

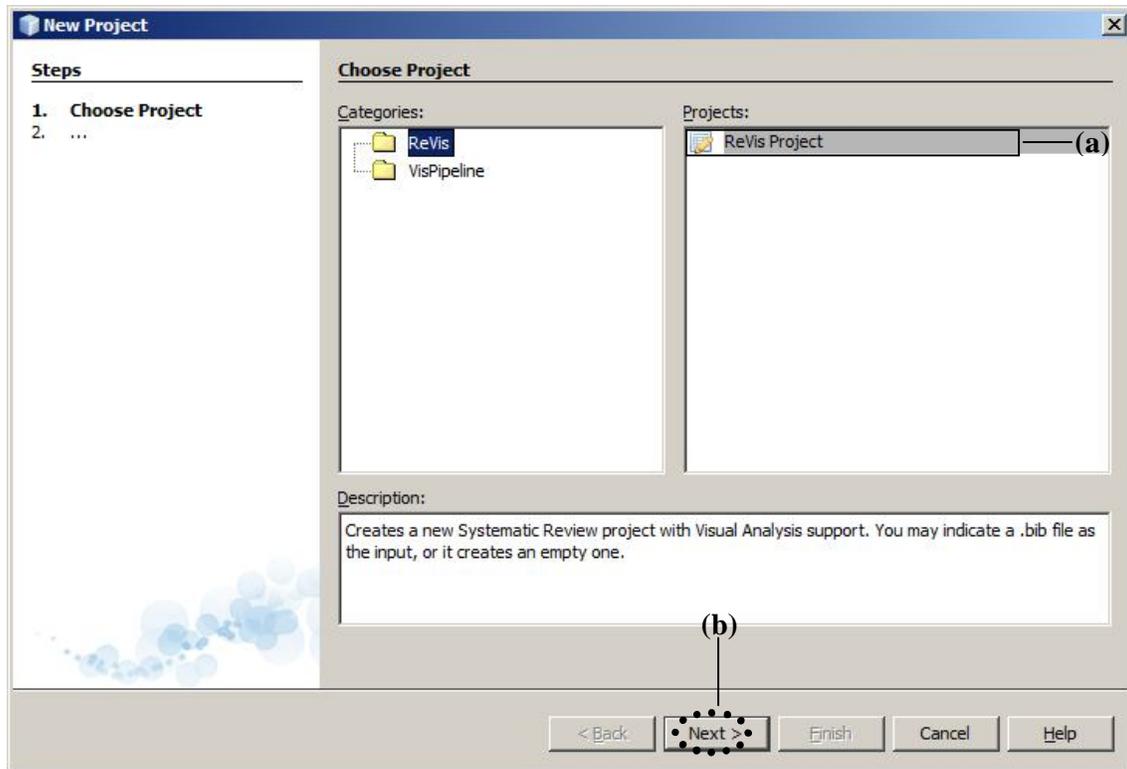


Figure 2: Second step to create a new project

- **Second Step:** The file "*Revis Project*" (Figure 2 – a) is a template of pipeline of systematic revision selection and should be use by user. For this, select the "*Revis Project*" option and click in "Next" button (Figure 2 – b).

If the user wishes, he/she can elaborate your own pipeline. For this, he/she can to select the option "*VisPipeline/Empty pipeline process*", but we recommend the use of the template, it is easier.

- **Third Step:** It is also possible the user to choose the name's of your project (Figure 3 – a) and the location/folder to save it. The primary studies recovered must to be organized as explained in "*How create your set of documents (primary studies) – Input*" section and they must be saved as a BibTeX file. The localization of this file is indicated in the "*BibTeX File*" (Figure 3 – b). After fill out these information, click in "*Finish*" (Figure 3 – c).

Finally, click in *Run Main Project* button (, in the left side of superior toolbar) to visualize the views (to more information about the views, see Section 3.1).

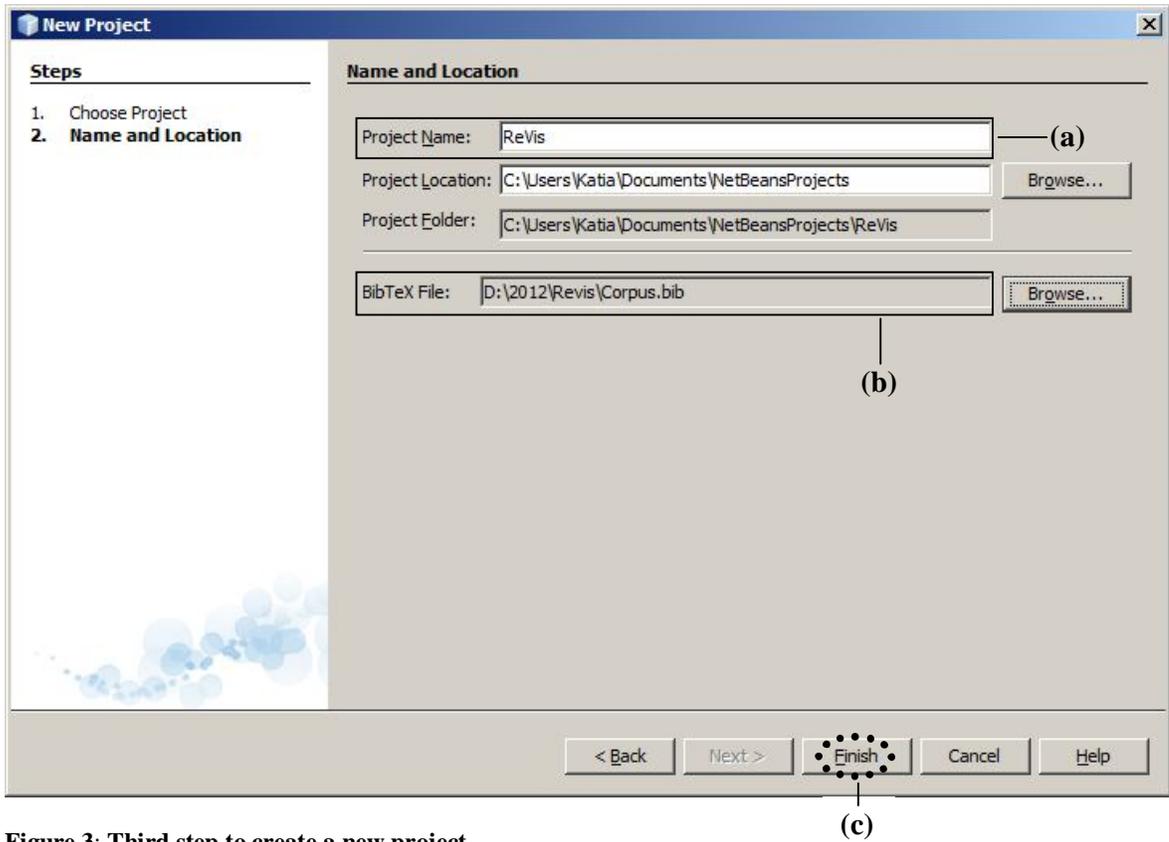


Figure 3: Third step to create a new project

A created project is arranged in two main folder: “  config ” and “  input ”, as shown in Figure 4 and thoroughly described following.

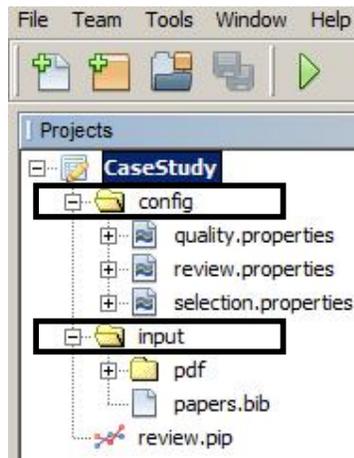


Figure 4: Arranging of a created project

2.1.1 Quality Criteria

In the “  config ” folder is possible to configure the quality criteria, review properties and inclusion/ exclusion criteria. The quality criteria can be defined by clicking onto the “*Quality properties*” file. Then it will show an empty *txt* file where the user writes the quality criteria of your review. Figure 5 exemplifies a template of “*quality.properties*” file filled out hypothetically.

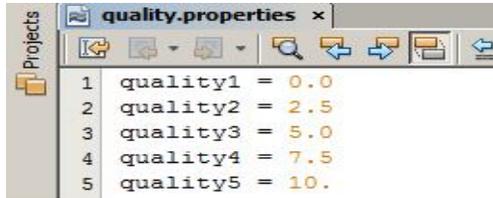


Figure 5: “*Quality.properties*” file

2.1.2 Inclusion / Exclusion Criteria

The inclusion / exclusion criteria can be defined by clicking onto the “*selection.properties*” file. Then it will show an empty *txt* file where the user writes the inclusion / exclusion criteria of your review. Figure 6 exemplifies a template of “*selection.properties*” file filled out hypothetically.

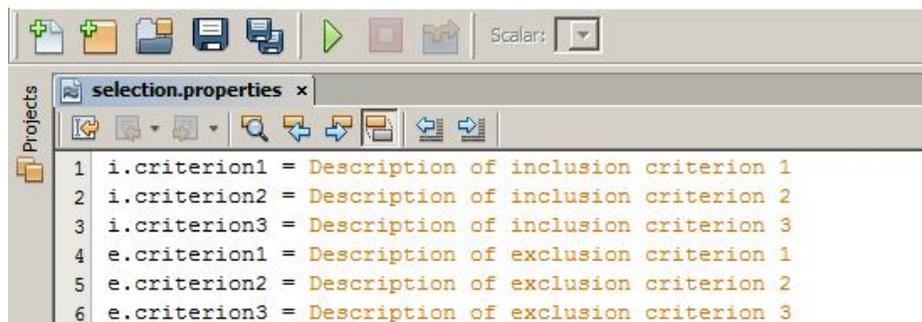


Figure 6: “*Selection.properties*” file

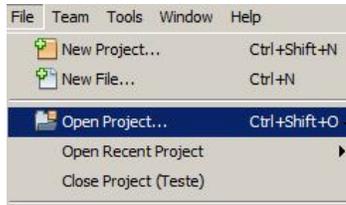
2.1.3 PDF's Files

At this point it is necessary save the pdfs' files will be used in the systematic review selection. To do this, click in “  input ”, folder.



This will open up a new window (named pdf) where you must to save the documents.

2.2 Open a Project



- **Open Project:** This menu option allows the user to access a systematic review created previously.

- The first step is to select the "Open Project" option in file menu (Figure 7 – a).

Figure 7: Menu Option - Open Project

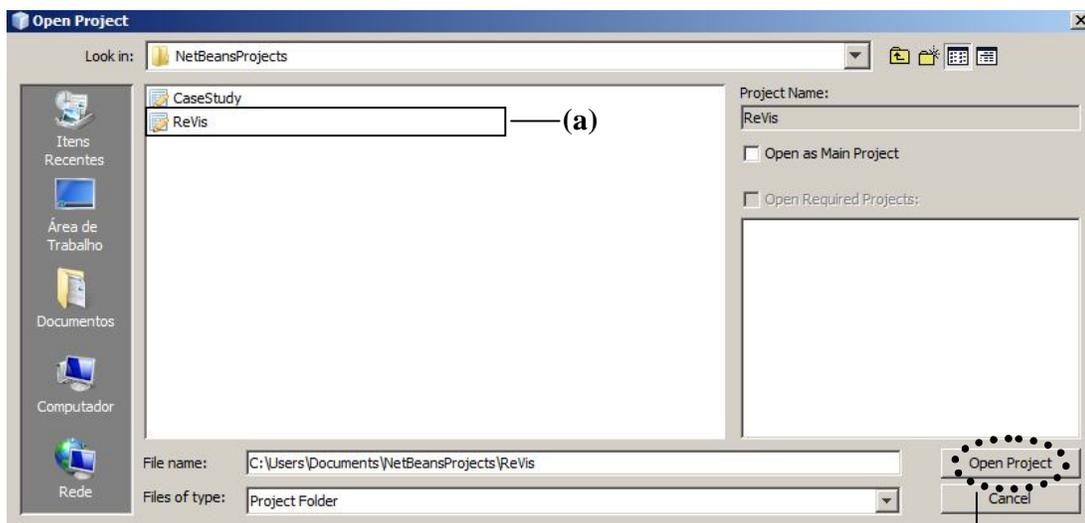


Figure 8: Opening a created project

Then, the user chooses the name of your project (Figure 8 – a) and click in "Open Project" button (Figure 8 – b). As with the "Create a new projection" option, click in *Run Main Project* button () to visualize the views (to more information about the views, see Section 3.1).

VTM Selection

3.1 The Views

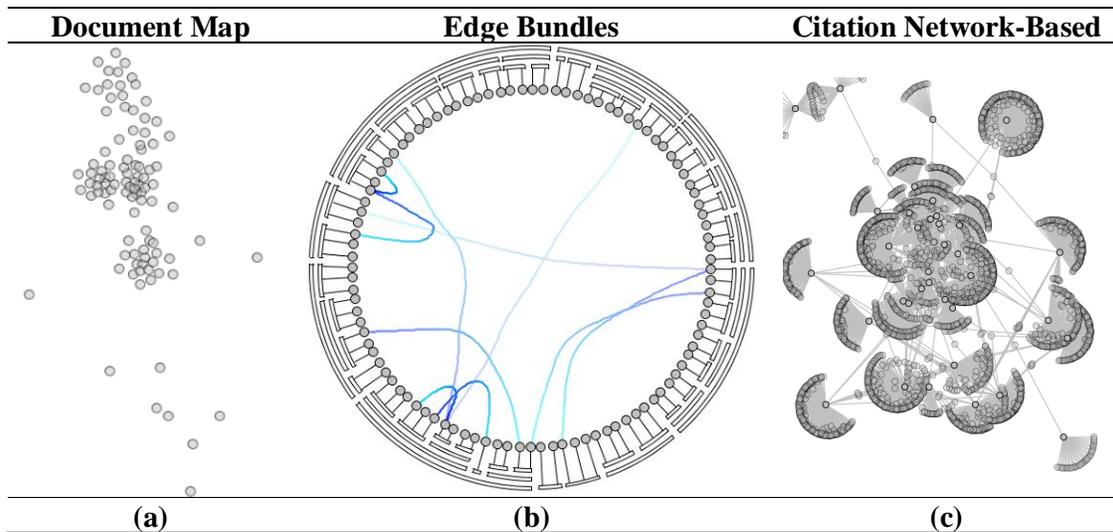


Figure 9: Different visualizations to support the primary studies selection activity

Figure 9 – a presents the “Document Map”, where each point in this map represents a primary study. We can observe regions that concentrate primary studies, indicating content similarity (title, abstract and keywords) among them.

Figure 9 – b shows the “Edge Bundles”, a hierarchical tree visualization technique that shows both nodes (primary studies) and node-links (relationships between nodes - citations) at the same time.

Figure 9 – c shows the “Citation Network-Based”. In this view the points represent the primary studies and the gray points the referenced papers, which are not part of the initial set of considered primary studies. Edges reflect the citations between documents (primary studies).

Stages of Selection Activity

Once a review has been created is initiated the Stage 1 of the selection. In this first stage, the primary studies are selected based on the abstract reading. It is necessary finish the Stage 1 (Current Stage) to start the Stage 2.



To visualize the abstract of any primary study is suffice one double click over it. A new window will appear and the abstract is shown as illustrate in the Figure 10.

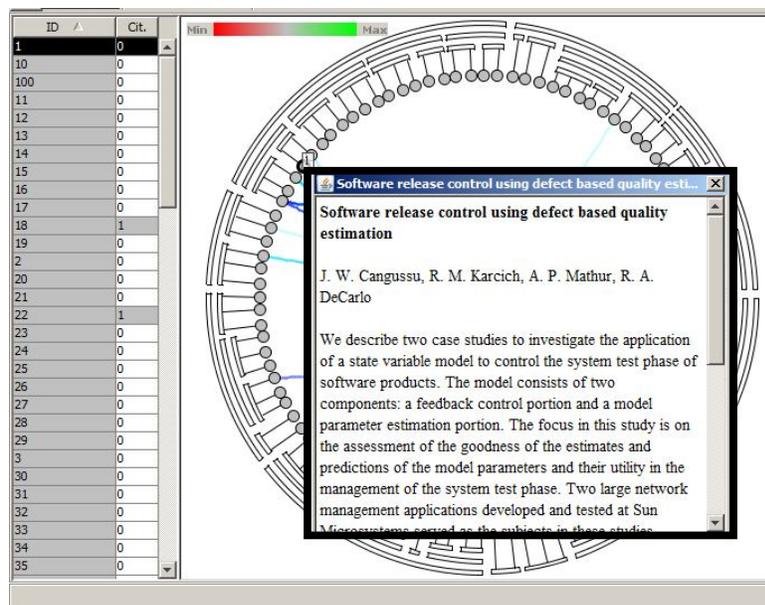


Figure 10: Visualization of abstract's primary study

4.1 Classification Activity

To classify the primary studies in included or excluded is necessary to click with the right-hand of the mouse over the "ID" column of the documents. A new window will be appear, as illustrated in Figure 11.

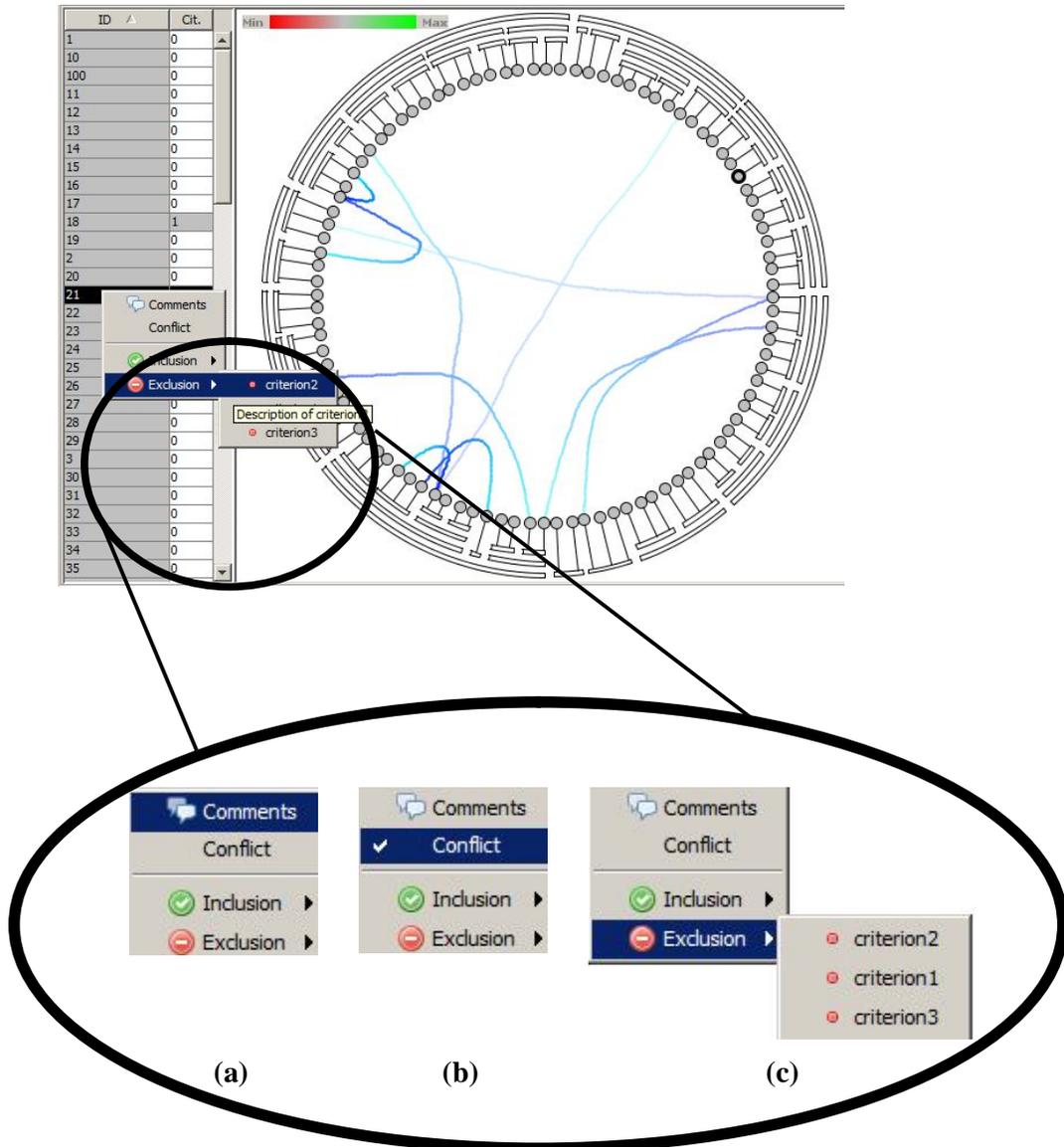


Figure 11: Options of classification of the primary studies

If you wish you can insert your own comment about the primary study (Figure 11 – a). Furthermore, you can to define the document as a “*conflict document*”, i.e., one document that you do not have sure about its the correct classification (included/excluded) (Figure 11 – b). More, you can classify the document as included or excluded and choose the criteria which you used to based your decision (Figure 11 – c).

Figure 12 shows the document map and the edges bundles after that all primary studies were classified by the user. The red points are the primary studies classified as excluded and the green points are the primary studies classified as included.

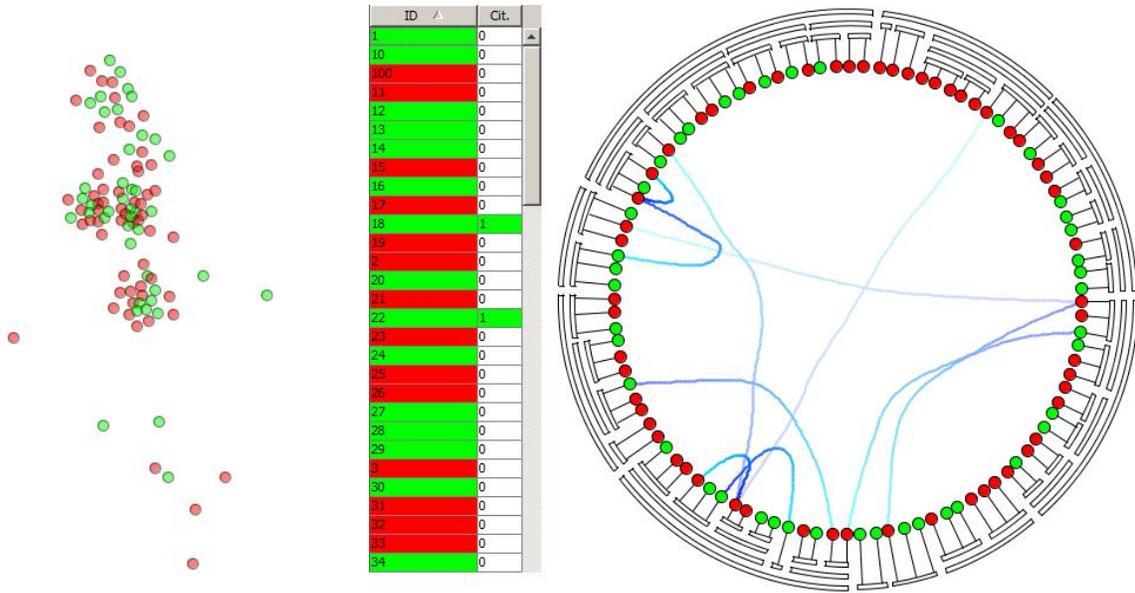


Figure 12: Two views of the three possible, after finished the stage 1

In the second stage, the primary studies are selected based on the full text reading (pdf format). To visualize the full text of any primary study is suffice one double click over it. A new window will appear and the pdf of the document will be shown. Make sure that you saved the pdfs' documents in the folder "input/pdf", as explained in Section 2.1.3. It is important to make it clear that in both stages all VTM strategies can be used to support the selection activity. It is necessary finish the Stage 2 to start the Stage 3.

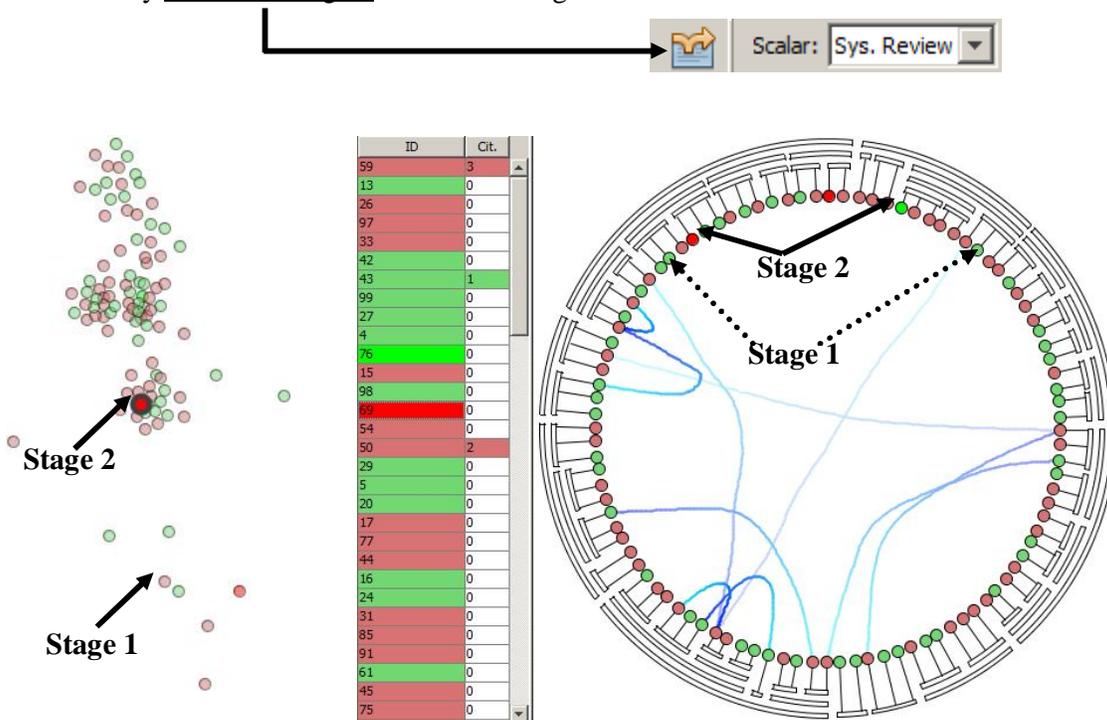


Figure 13: Two views of the three possible, after finished the stage 2

Figure 13 shows the document map and the edges bundles after that all primary studies were classified by the user based on the full text reading. In this case, the red points are the primary studies classified as excluded and the green points are the primary studies classified as included. The dark red points are the primary studies classified as excluded in the second stage (full text reading) and the dark green points are the primary studies classified as included in the second stage (full text reading).

At the end of Stage 2, all papers were already classified in included or excluded. Then, in the Stage 3, only the included studies have their quality measured. To classify the quality of the primary is necessary to click with the right-hand of the mouse over the “ID” column of the documents. A new window will be appear, as illustrated in Figure 14. The scale of values were defined by the user as explained in Section 2.1.1.



Figure 14: Options of quality of the primary studies

Features of Revis Tool

The Revis offers various strategies to support the systematic review selection activity, as “*Neighborhood relationship (called RN connection)*”, “*Topic*” and “*Coordination*”, among other. These strategies will be explained and exemplified following.

4.2 Neighborhood relationship (called RN connection)

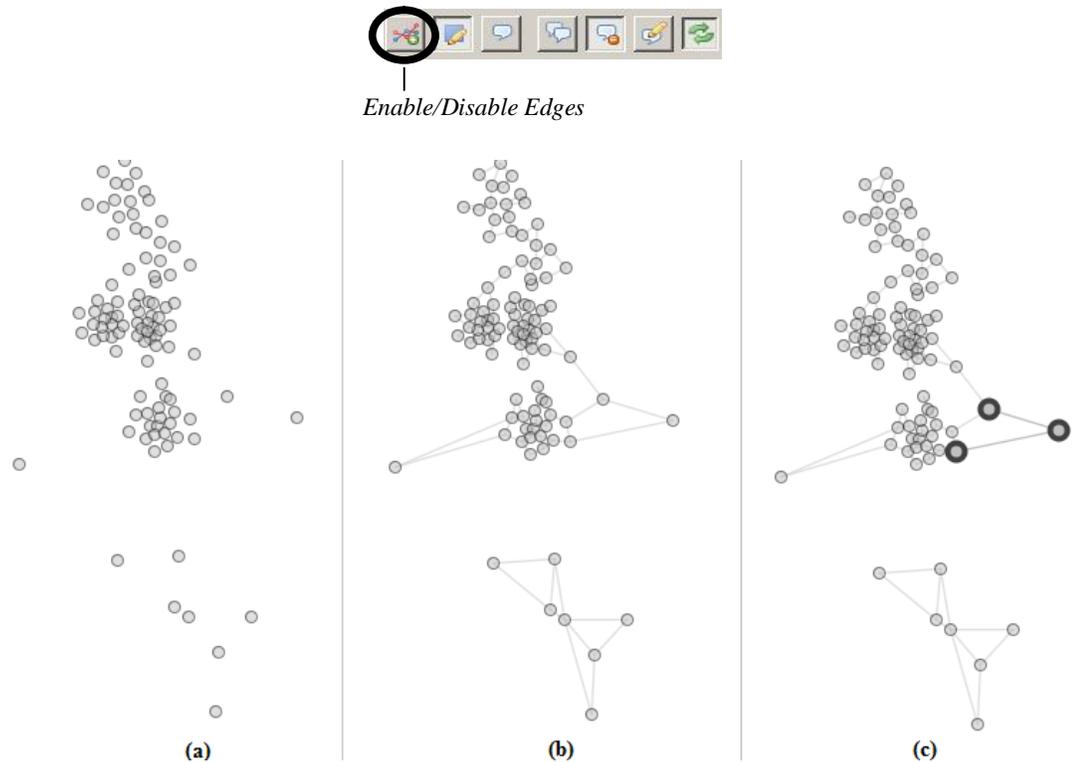


Figure 15: (a) Original document map; (b) Document map after the application of the “*KNN-R2-2*” strategy; (c) One document highlighted and its neighbors (similar documents)

Primary studies around a relevant document are probably relevant, as well as, primary studies around a not relevant document are probably not relevant. One interesting strategy is to identify the neighborhood of these documents. To do this, choose the “*Enable/Disable Edges*” in Menu File (Figure 15 – a). In the original document map will be created edges among the documents, indicating their nearest neighbors (Figure 15 – b). If you wish, you can select one specific document and automatically this point and its neighbors will be highlighted (Figure 15 – c).

4.3 Topic

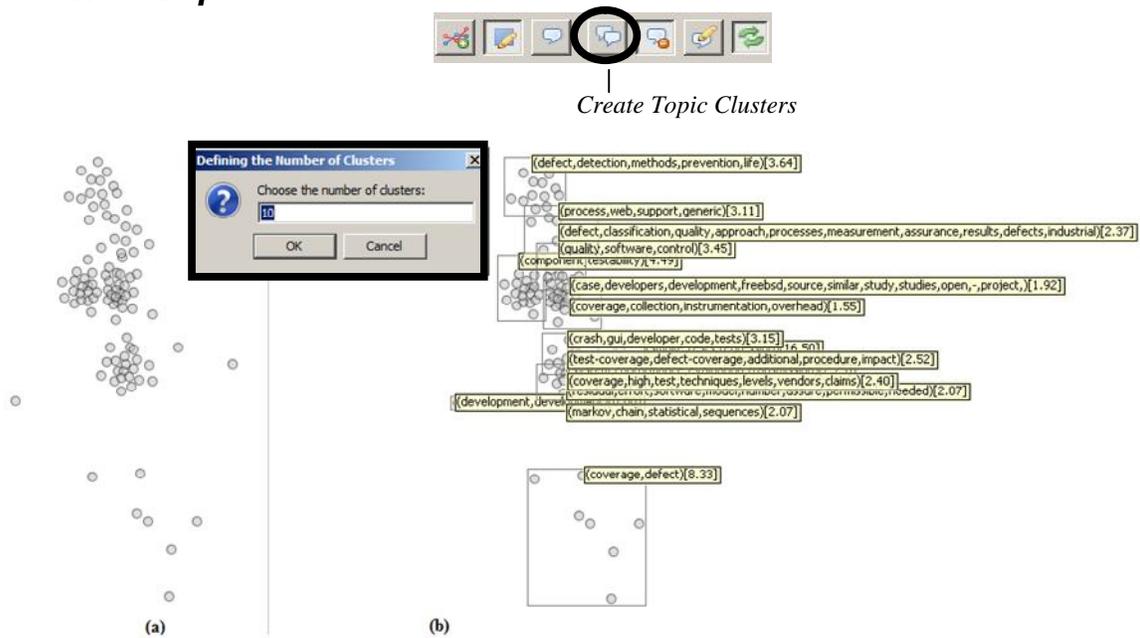


Figure 16: (a) Original document Map; (b) Document map after the application of the "Topic Cluster" strategy

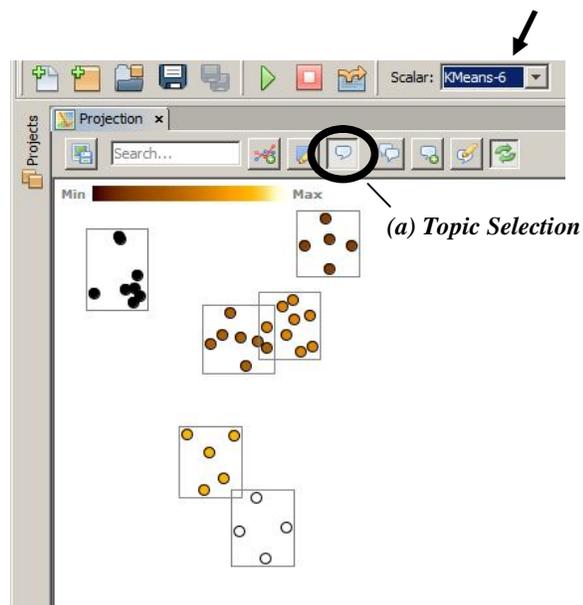


Figure 17: Document Map coloured with K-Means

To identify a region of similar documents, i.e., document that have similar contents in terms of title, abstract and keywords, is possible to create clusters and extract topics – labels that represent the content of the documents contained in the clusters. To do this, first select the "*Create Topic Clusters*" option in menu file. A new window will appear which will request the number of clusters. Write your choose (number of clusters that you wish) and click in the "*OK*" button (Figure 16 – a). The result is the document map marked with the clusters (square delimitation regions) and theirs respective topics (Figure 16 – b).

If you wish, you can colour the points to identify visually the clusters. For this, first select the "*KMeans*" option in "*Scalar*" menu (In front of the word *KMeans* will be appear the number of clusters that you chose anteriorly, 6 in this example). Now, the result is the document map marked with the clusters, being added to each cluster one specific colour to facilitate theirs identification (Figure 17).

More, if you even wish, you can visualize the cluters coloured with K-Means scalar, however, without the square delimitation. For this, press the "*Topic Selection*" button, as illustrated in Figure 17 – a.

4.4 Occurrence frequency of an expression

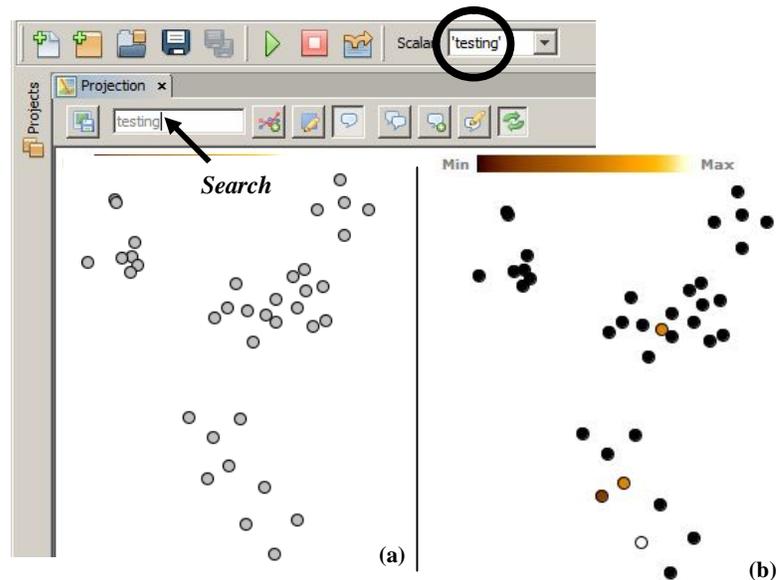


Figure 18: (a) Original document Map; (b) Document map coloured based on the number of times that one expression occurs in the documents

The document map can be coloured based on the occurrence of specific in the document (title, abstract and keywords). In this case, the colour scale varies from black (no occurrence) to brown (many occurrences). The terms are freely chosen by users and they are written in "*search*" the text box (Figure 18 – b). To visualize the coloration, you need choose in the "*scalar*" menu the same expression that you wrote in the "*search*" text box.

4.5 Coordination

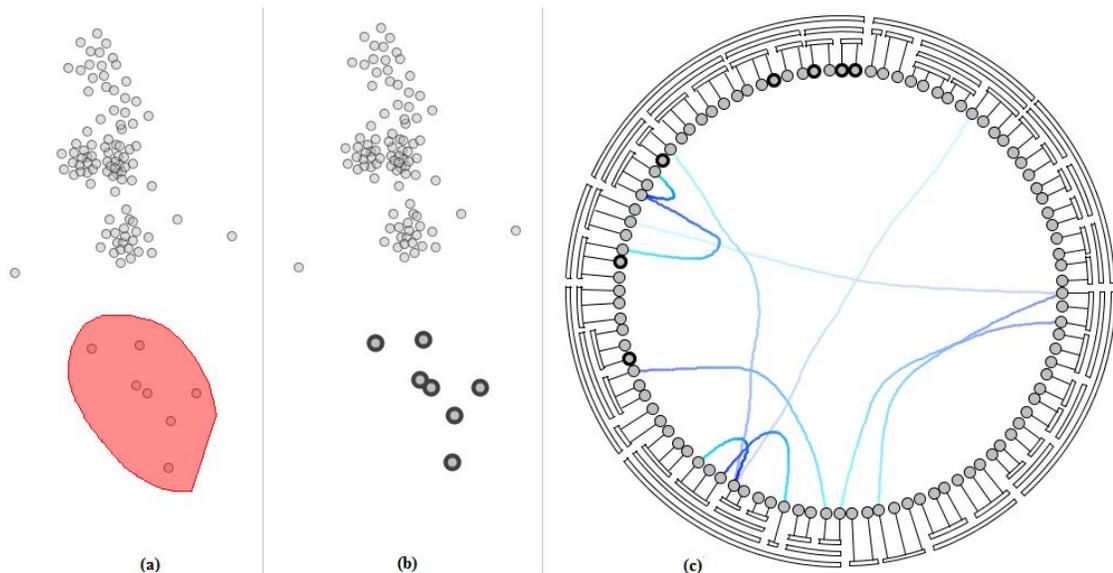


Figure 19: Coordination using a selection points chosen by user

The coordination is the interaction among the different views from the same data and it implies that user selection of a point or a group of documents in a view highlights the corresponding point in the other views. To have the coordination is necessary keep on pressed the “” button (Figure 19).

If you wish, you can select more than one point to coordinate. Thus clicking the right-hand button of the mouse you can select your points (Figure 19 – a). The results are the same points selected by the user highlighted in different views, as illustrated in Figure 19 – b and Figure 19 – c.

More tips about the tool

■ System Requirements

This tool has been designed to run in Windows XP or Windows Vista operating system. It is necessary to be installed on the machine the JDK 1.5.0 (<http://www.java.sun.com>) or an earlier version.

Contacts

For questions, suggestions or problems in the use of the tool, contact:

Rafael Messias Martins	<i>rafaelmm@icmc.usp.br</i>
Katia Romero Felizardo	<i>katiarf@icmc.usp.br</i>