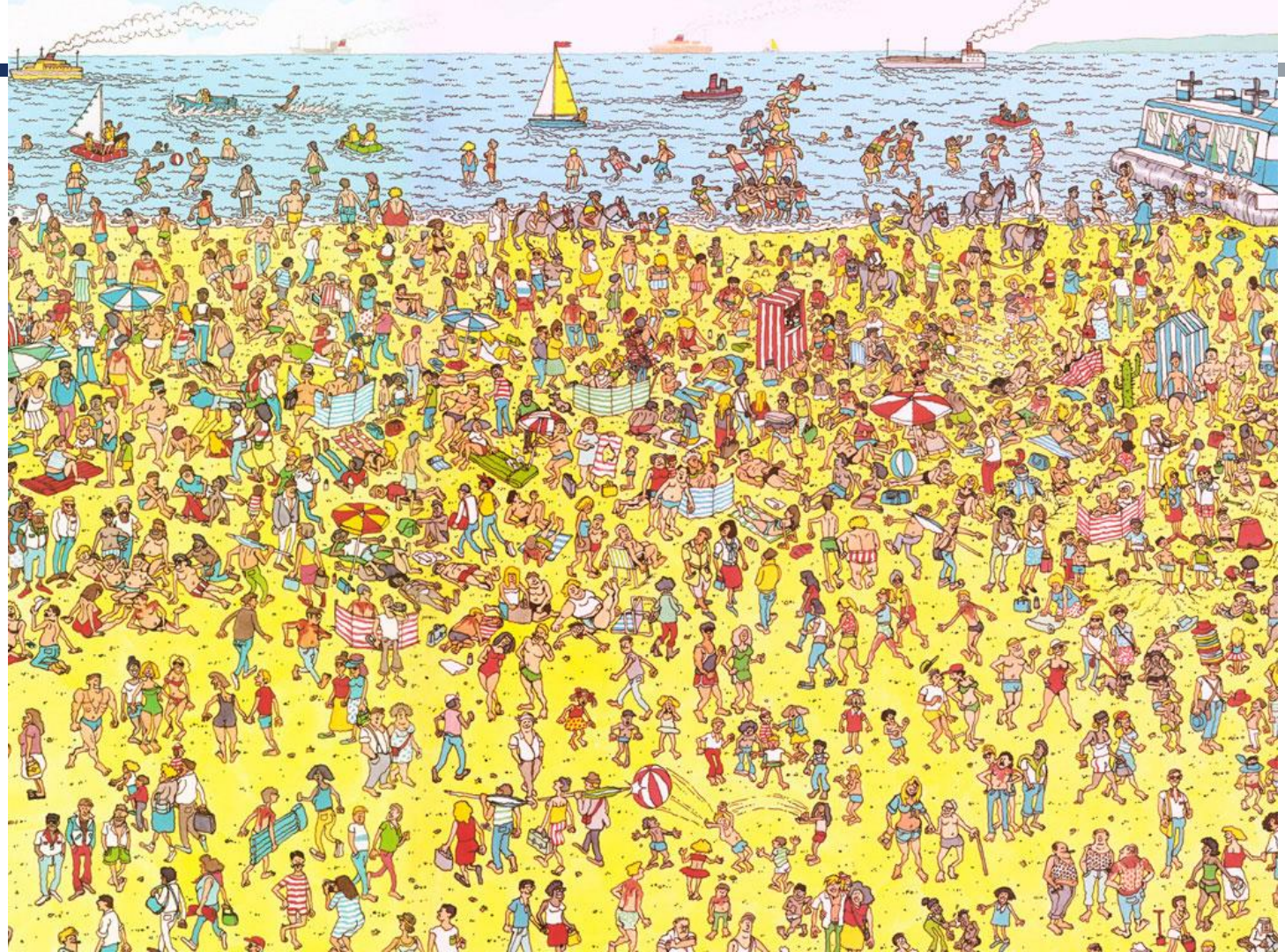
# REVISÃO SISTEMÁTICA DE LITERATURA

Profa. Lucila Ishitani
Introdução à Pesquisa em Informática
ICEI – PUC Minas

# POR QUE SISTEMÁTICA?

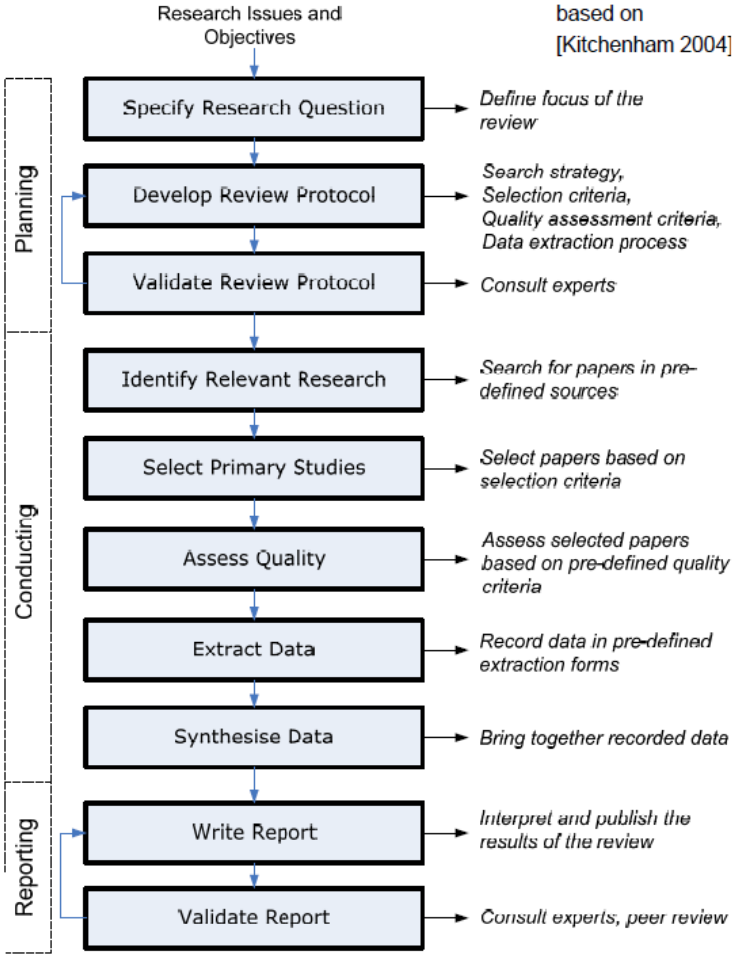
| Resultados da pesquisa por "Software Metrics“ em jan/2018 | |
|---|---|
| Google | 384.000 |
| Scholar Google | 46.700 |
| IEEE | 5.031 |
| Science Direct | 1.834 |
| ACM | 494 |



# SLR Process

Key elements:

- *3 phases*
  - Planning
  - Conducting
  - Reporting
- *Explicit definition of a review protocol!*
- *Continuous validation and quality assessment!*




Fonte: http://seaa2011.oulu.fi/tutorials/RR_SEAA2011ResearchMethods_SLR.pdf

# EXEMPLO 1

- Hall, T.; Beecham, S.; Bowes, D.; Gray, D.; Counsell, S., "A Systematic Literature Review on Fault Prediction Performance in Software Engineering," *Software Engineering, IEEE Transactions on* , vol.38, no.6, pp.1276,1304, Nov.-Dec. 2012

- Objective: to “investigate how the context of models, the independent variables used and the modelling techniques applied, influence the performance of fault prediction models.”

# EXEMPLO 1

| Research Questions | | Motivation |
|---|---|---|
| RQ1 | How does context affect fault prediction? | Context has been shown to be a key factor in the comparative use of software metrics in general [5]. Context is important in fault prediction modelling as it can affect the performance of models in a particular context and the transferability of models between contexts. Currently the impact context variables have on the transferability of models is not clear. This makes it difficult for potential model users to select models that will perform well in a particular context. We aim to present a synthesis of current knowledge on the impact of context on models and the transferability of models. |
| RQ2 | Which independent variables should be included in fault prediction models? | There are a range of independent variables that have been used in fault prediction models. Currently the impact individual independent variables have on model performance is not clear. Although the performance of independent variables has been investigated within individual studies, no comparison of performance across studies has been done. This makes it difficult for model builders to make informed decisions about the independent variables on which to base their models. We aim to present a synthesis of current knowledge on the impact independent variables have on models. |
| RQ3 | Which modeling techniques perform best when used in fault prediction? | Fault prediction models are based on a wide variety of both machine learning and regression modelling techniques. Currently the impact modelling technique has on model performance is not clear. Again, the performance of modelling techniques has been investigated within individual studies, but no comparison of performance across studies has been done. This makes it difficult for model builders to make effective technique selections. We aim to present a synthesis of current knowledge on the impact of modelling technique on model performance. |

# EXEMPLO 1

- Inclusion criteria (a paper must be. . . )
  - An empirical study
  - Focused on predicting faults in units of a software system
  - Faults in code is the main output (dependent variable)
- Exclusion criteria (a paper must not be. . . )
  - Focused on: testing, fault injection, inspections, reliability modelling, aspects, effort estimation, debugging, faults relating to memory leakage, nano-computing, fault tolerance.
  - About the detection or localisation of existing individual known faults.

# EXEMPLO 1

- Pesquisa:
  - Included papers were published between January 2000 and December 2010.
  - There were four elements to our searches:
    - Key word searching using the search engines:
      - ACM Digital Library
      - IEEExplore
      - ISI Web of Science
    - An issue-by-issue manual reading of paper titles in relevant journals and conferences. [...] These were chosen as highly relevant software engineering publications found previously to be good sources of software engineering research.
    - A manual search for publications from key authors using DBLP3. These authors were selected as appearing most frequently in our list of papers: Khoshgoftaar, Menzies, Nagappan, Ostrand and Weyuker.
    - The identification of papers using references from included studies.

# EXEMPLO 1

Table 10
Conferences and journals manually searched

| Conference manually searched | Journals manually searched |
|---|---|
| International Conference on Software Engineering (ICSE) | IEEE Transactions of Software Engineering |
| International Conference on Software Maintenance (ICSM) | Journal of Systems and Software |
| IEEE Int'l Working Conference on Source Code Analysis and Manipulation (SCAM) | Journal of Empirical Software Engineering |
| International Conference on Automated Software Engineering | Software Quality Journal |
| IEEE Int'l Symposium and Workshop on Engineering of Computer Based Systems | Information & Software Technology |
| International Symposium on Automated Analysis-driven Debugging | |
| International Symposium on Software Testing and Analysis (ISSTA) | |
| International Symposium on Software Reliability Engineering | |
| ACM SIGPLAN Conference on Programming language Design and Implementation | |
| Int'l Workshop on Mining Software Repositories | |
| Empirical Software Engineering & Measurement | |
| PROMISE | |
| Foundations of Software Engineering | |

# EXEMPLO 1

Table 3
Paper selection and validation process

| Selection Process | # of papers | Validation |
|---|---|---|
| Papers extracted from databases, conferences and journals | 2,073 | 80 random papers independently classified by 3 researchers |
| Sift based on title and abstract | -1,762 rejected | Fair/good inter-rater agreement on first sift (k statistic test) |
| Full papers considered for review | 311 primary<br>80 secondary | Each paper is read in full and 80 secondary papers are identified from references |
| Rejected on full reading | -185 rejected | Papers are rejected on the basis that they do not answer our research questions |
| Comparison to Catal and Diri's review | 2<br>(our searches missed) | |
| Papers accepted for the review | 208 papers | |

# EXEMPLO 1

## Assessment criteria:

- Phase 1: Establishing that the study is a prediction study.
- Phase 2: Ensuring sufficient contextual information is reported.
- Phase 3: Establishing that sufficient model building information is reported.
- Phase 4: Checking the model building data.

Table 8
Results of applying assessment criteria

| Number of papers passed | Number of papers failed | | | | | |
|---|---|---|---|---|---|---|
| | Phase 1: Prediction | Phase 2: Context | Phase 3: Model | Phase 4: Data | Other reasons | Total failed |
| 36 | 41 | 114 | 2 | 13 | 2 | 173 |

# EXEMPLO 2

- Kitchenham, B., Mendes, E., Travassos, G.H. (2007) A Systematic Review of Cross- vs. Within-Company Cost Estimation Studies, IEEE TSE, 33 (5), 316-329.
- Objective: “to determine under what circumstances individual organisations would be able to rely on cross-company based estimation models.”

# EXEMPLO 2

- Questões:
  - RQ1: What evidence is there that crosscompany estimation models are not significantly worse than within-company estimation models for predicting effort for software/Web projects?
  - RQ2: Do the characteristics of the study data sets and the data analysis methods used in the study affect the outcome of within company and cross-company effort estimation accuracy studies?
  - RQ3: Which estimation method(s)/process(es) were best for constructing cross-company effort estimation models?

- Search string:

(software OR application OR product OR Web OR WWW OR Internet OR World-Wide Web OR project OR development)

**AND**

(method OR process OR system OR technique OR methodology OR procedure)

**AND**

(cross company OR cross organisation OR cross organization OR cross organizational OR cross organisational OR cross-company OR cross-organisation OR cross-organization OR cross-organizational OR cross-organisational OR multi company OR multi organisation OR multi organization OR multi organizational OR multi organisational OR multi-company OR multi-organisation OR multi-organization OR multi-organizational OR multi-organisational OR multiple company OR multiple organisation OR multiple organization OR multiple organizational OR multiple organisational OR multiple-company OR multiple-organisation OR multiple-organization OR multiple-organizational OR multiple-organisational OR within company OR within organisation OR within organization OR within organizational OR within organisational OR within-company OR within-organisation OR within-organization OR within-organizational OR within-organisational OR single company OR single organisation OR single organization OR single organizational OR single organisational OR single-company OR single-organisation OR single-organization OR single-organizational OR single-organisational OR company-specific)

**AND**

(model OR modeling OR modelling)

**AND**

(effort OR cost OR resource)

**AND**

(estimation OR prediction OR assessment)

# EXEMPLO 2

- Fontes:
  - Compendex,
  - IEEE Xplore,
  - Springer Link,
  - ISI Web of Knowledge,
  - ScienceDirect,
  - Wiley Inter Science Journal Finder,
  - ACM Digital Library
  - Individual journals (J) and conference proceedings (C)

- Empirical Software Engineering (J)
- Information and Software Technology (J)
- Software Process Improvement and Practice (J)
- Management Science (J)
- International Software Metrics Symposium (C)
- Int. Conference on Software Engineering (C)
- Empirical Assessment in Software Engineering (manual search) (C)

# EXEMPLO 2

- Inclusion criteria:
  - any study that compared predictions of cross-company models with within-company models based on analysis of single company project data.
- Exclusion criteria:
  - studies where projects were only collected from a small number of different sources (e.g. 2 or 3 companies),
  - studies where models derived from a within-company data set were compared with predictions from a general cost estimation model.

# EXEMPLO 2

- Avaliação de qualidade
  - Is the analysis process description complete?
    - Was the data investigated to identify outliers and to assess distributional properties before analysis?
    - Was the result of the investigation used appropriately?
    - Were the resulting estimation models subject to sensitivity or residual analysis?
    - Was the result of the sensitivity or residual analysis used appropriately?
    - Were accuracy statistics based on the raw data scale?
  - Is it clear what projects were used to construct each model?
  - Is it clear how accuracy was measured?
  - Is it clear what cross-validation method was used?
  - Were all model construction methods fully defined (tools and methods used)?

# EXEMPLO 2

- Avaliação de qualidade
  - How good was the study comparison method?
    - Was the single company selected at random (not selected for convenience) from several different companies?
    - Was the comparison based on a completely independent hold out sample or on n-fold cross-validation for the within-company model?
  - The size of the within-company data set, measured according to the criteria presented below. Whenever a study used more than one within-company data set, the average score was used:
    - Less than 10 projects: Poor quality (score = 0)
    - Between 10 and 20 projects: Fair quality (score = 0.33)
    - Between 21 and 40 projects: Good quality (score = 0.67)
    - More than 40 projects: Excellent quality (score = 1)

# EXEMPLO 3

- Darja Šmite , Claes Wohlin,Tony Gorschek, Robert Feldt. (2010). Empirical evidence in global software engineering: a systematic review. Empirical Software Engineering, 15, 91–118.
- Objective: “to capture the current empirically evaluated knowledge in the area and to identify needs and opportunities for future research.”

# EXEMPLO 3

- Questões:
  - RQ1: What is the state-of-the-art in empirical studies of GSE?
    - RQ1.1: Who is Involved in GSE?
    - RQ1.2: Where are the Development Sites Located?
    - RQ1.3: What is Studied in GSE?
    - RQ1.4: How Successful are the Cases Reported in Literature?
    - RQ1.5: Why are Companies Involved in GSE?
  - RQ2: What is the strength of the empirical evidence reflected in the empirical GSE?

# EXEMPLO 3

- Search string:
- The final search strings were based on the experience from the pilot searches and consisted of a Boolean expression:
- (A1 or A2 or A3 or A4)
- AND (B1 or B2 or B3 or B4)

A1—global software development
A2—global software engineering
A3—distributed software development
A4—distributed software engineering
B1—empirical
B2—industrial
B3—experiment
B4—case study

# EXEMPLO 3

- Fontes:
  - Compendex,
  - IEEE Xplore,
  - Springer Link,
  - ISI Web of Knowledge,
  - ScienceDirect,
  - Wiley Inter Science Journal Finder,
  - ACM Digital Library

# EXEMPLO 4

- Marcela Genero, Ana M. Fernández, H. James Nelson, Geert Poels, Mario Piattini. (2011). A Systematic Literature Review on the Quality of UML Models. Journal of Database Management, 22(3), 46-70.
- Objetivo: “to explore the history of the field and to determine its current state of the art by locating, evaluating, and interpreting relevant research to date that is related to model quality with a focus on UML.”

# EXEMPLO 4

- Questões:
  - RQ1. Which type of UML model quality has been investigated by researchers?
  - RQ2. Which research methods are used in research on UML model quality?
  - RQ3. What is the nature of the research results on UML model quality?
  - RQ4. Which research goals are aimed at in research on UML model quality?
  - RQ5. Which type of UML diagrams is the focus of the research on UML model quality?

# EXEMPLO 4

The search terms used were constructed using the following steps (Brereton, Kitchenham, Budgen, Turner, & Khalil, 2007):

- Define the major terms.
- Identify alternative spellings, synonyms or related terms for major terms.
- Check the keywords in any relevant papers we already had.
- Use the Boolean OR and AND to link the major terms.

| Major terms | Alternative terms |
|---|---|
| Quality | quality OR consistency OR maintainability OR understandability OR completeness OR comprehension OR comprehensibility OR testability OR defect OR effectiveness OR complexity OR readability OR metric OR measure OR efficiency OR validation OR verification OR layout |
| UML | UML OR Unified Modeling Language |
| Representation | Representation OR diagram OR model |

# EXEMPLO 4

Fontes:

- SCOPUS database,
- ScienceDirect with the subject Computer Science,
- Wiley InterScience with the subject of Computer Science,
- IEEE Digital Library,
- ACM Digital Library,
- SPRINGER database.

# EXEMPLO 4

Inclusion criteria:

- Papers which dealt with UML and the tangible results of the modelling process (the UML diagram),
- were written in English,
- and were published between 1997 and 2009.

# EXEMPLO 4

- Exclusion criteria:
  - pure discussion and opinion papers,
  - studies available only in the form of abstracts or presentations,
  - duplicates,
  - research focusing issues other than UML model quality (for example, functional size measurement),
  - or where quality is mentioned only as a general introductory term in the paper's abstract and an approach or other type of proposal related to quality is not amongst the paper's contributions.
  - Papers were also excluded if they dealt with the quality and complexity of UML as a language (for example, how to make UML the language simpler) rather than on the quality and complexity of the models produced by UML,
  - if the paper was a summary of a workshop.

*Table 4. Chronology of the development of activities in the SLR*

| Time | Planning | Conducting | Reporting | Outcomes |
|---|---|---|---|---|
| **First phase** | | | | |
| July 2007 | Protocol development | | | Review protocol. |
| Sept 2007 | | Data retrieval (until Sept 2007) | | Form with the general information of the papers (1500 papers). |
| | | Study selection upon abstracts and titles | | Form with the general information of the selected papers (483 papers). |
| Mar 2008 | | Retrieval of the files of the primary studies | | Repository of papers (483 papers). |
| Apr 2008 | | Remove duplicates | | Form with the general information of the papers (399 papers). |
| Jul 2008 | Protocol improvement | Pilot data extraction | | Data extraction form with the classification scheme refined. |
| Aug 2008 | | Study selection and Data extraction upon the full text | | Data extraction form completed with the classification of 215 primary studies. |
| Feb 2009 | | Resolution of doubts in classification of primary studies in group | | Revisited data extraction form with classification of the primary studies **(193).** |
| Mar 2009 | | Data synthesis | | |
| July 2009 | | | Report the results of the SLR | Pilot report |

| Second phase | | | | |
|---|---|---|---|---|
| Mar 2010 | | Update of searches<br>Data retrieval<br>(until Dec 2009) | | Form with the general information of the papers (979). |
| Mar 2010 | | Study selection upon abstracts and titles | | Form with the general information of the selected papers (140). |
| | | Retrieval of the files of the primary studies | | Repository of papers 140). |
| | | Remove duplicates | | Form with the general information of the papers (103). |
| Feb 2010 | | Study selection and Data extraction upon the full text | | Data extraction form completed with the classification of primary studies (103) |
| Mar 2010 | | Resolution of doubts in classification of primary studies in group | | Revisited data extraction with the classification of primary studies **(73)** |
| Apr 2010 | | Data synthesis | | |
| Jul 2010 | | | Report the results of the SLR | Final report |

# EXEMPLO 5

- B. Kitchenham, P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering - A systematic literature review," Information and Software Technology, vol. 51(1), pp. 7-15, 2009.
- Objetivo: “to review the current status of Evidence-based Software Engineering (EBSE) since 2004 using a tertiary study to review articles related to EBSE and, in particular, we concentrate on articles describing systematic literature reviews (SLRs).

# EXEMPLO 5

- Questões:
  - RQ1. How much SLR activity has there been since 2004?
  - RQ2. What research topics are being addressed?
  - RQ3. Who is leading SLR research?
  - RQ4. What are the limitations of current research?
    - RQ4.1. Were the research topics limited?
    - RQ4.2. Is there evidence that the use of SLRs is limited due to lack of primary studies?
    - RQ4.3. Is the quality of SLRs appropriate, if not, is it improving?
    - RQ4.4. Are SLRs contributing to practice by defining practice guidelines?

**Table A1**
Sources searched for years 2004–2007 (including articles up to June 30 2007).

| Year | 2004 | 2005 | 2006 | 2007 | Total |
|---|---|---|---|---|---|
| IST (Total) | 85 | 95 | 72 | 47 | 299 |
| IST (Relevant) | 0 | 2 | 2 | 0 | 4 |
| IST (Selected) | 0 | 0 | 2 | 0 | 2 |
| JSS (Total) | 139 | 122 | 124 | 43 | 428 |
| JSS (Relevant) | 4 | 0 | 0 | 0 | 4 |
| JSS (Selected) | 3 | 0 | 0 | 0 | 3 |
| IEEE SW (Total) | 51 | 52 | 48 | 24 | 175 |
| IEEE SW (Relevant) | 1 | 0 | 5 | 2 | 9 |
| IEEE SW (Selected) | 1 | 0 | 3 | 0 | 4 |
| TSE (Total) | 69 | 66 | 56 | 25 | 216 |
| TSE (Relevant) | 2 | 1 | 0 | 3 | 7 |
| TSE (Selected) | 0 | 1 | 0 | 3 | 4 |
| CACM (Total) | 148 | 141 | 158 | 64 | 511 |
| CACM (Relevant) | 1 | 0 | 0 | 0 | 1 |
| CACM (Selected) | 1 | 0 | 0 | 0 | 1 |
| ACM Sur (Total) | 12 | 11 | 13 | 3 | 39 |
| ACM Sur (Relevant) | 0 | 0 | 1 | 0 | 1 |
| ACM Sur (Selected) | 0 | 0 | 0 | 0 | 0 |
| TOSEM (Total) | 10 | 12 | 12 | 6 | 40 |
| TOSEM (Relevant) | 0 | 2 | 0 | 0 | 2 |
| TOSEM (Selected) | 0 | 0 | 0 | 0 | 0 |
| SPE (Total) | 64 | 59 | 68 | 29 | 220 |
| SPE (Relevant) | 0 | 0 | 0 | 0 | 0 |
| SPE (Selected) | 0 | 0 | 0 | 0 | 0 |
| ICSE (Total) | 58 | 58 | 36 | 64 | 216 |
| ICSE (Relevant) | 0 | 0 | 1 | 0 | 1 |
| ICSE (Selected) | 0 | 0 | 1 | 0 | 1 |
| ISESE (Total) | 26 | 50 | 56 | n/a | 132 |
| ISESE (Relevant) | 0 | 2 | 1 | n/a | 3 |
| ISESE (Selected) | 0 | 2 | 0 | n/a | 2 |
| IET SW (Total) | 22 | 28 | 22 | 9 | 81 |
| IET SW (Relevant) | 0 | 0 | 0 | 1 | 1 |
| IET SW (Selected) | 0 | 0 | 0 | 0 | 0 |
| EMSE (Total) | 14 | 19 | 20 | 12 | 61 |
| EMSE (Relevant) | 1 | 0 | 0 | 0 | 1 |
| EMSE (Selected) | 1 | 0 | 0 | 0 | 1 |
| Metrics (Total) | 36 | 48 | n/a | n/a | |
| Metrics (Relevant) | 1 | 0 | n/a | n/a | 1 |
| Metrics (Selected) | 1 | 0 | n/a | n/a | 1 |
| Total | 734 | 761 | 685 | 326 | 2506 |
| Total relevant | 10 | 7 | 10 | 6 | 33 |
| Total selected | 7 | 3 | 6 | 3 | 19 |

**Table 1**
Selected journals and conference proceedings.

| Source | Acronym |
|---|---|
| Information and Software Technology | IST |
| Journal of Systems and Software | JSS |
| IEEE Transactions on Software Engineering | TSE |
| IEEE Software | IEEE SW |
| Communications of the ACM | CACM |
| ACM Computer Surveys | ACM Sur |
| ACM Transactions on Software Engineering Methodologies | TOSEM |
| Software Practice and Experience | SPE |
| Empirical Software Engineering Journal | EMSE |
| IEE Proceedings Software (now IET Software) | IET SW |
| Proceedings International Conference on Software Engineering | ICSE |
| Proceedings International Symposium of Software Metrics | Metrics |
| Proceedings International Symposium on Empirical Software Engineering | ISESE |

# EXEMPLO 5

- Critérios de inclusão:
  - Systematic Literature Reviews (SLRs) i.e. literature surveys with defined research questions, search process, data extraction and data presentation, whether or not the researchers referred to their study as a systematic literature review.
  - We included articles where the literature review was only one element of the articles as well as articles for which the literature review was the main purpose of the article.
  - Meta-analyses (MA).

# EXEMPLO 5

- Critérios de exclusão:
  - Informal literature surveys (no defined research questions; no defined search process; no defined data extraction process).
  - Papers discussing the procedures used for EBSE or SLRs.
  - Duplicate reports of the same study (when several reports of a study exist in different journals the most complete version of the study was included in the review).

**Table A2**
Candidate articles not selected.

| Source | Authors | Reference | Year | Title | Reason for rejection |
|---|---|---|---|---|---|
| TSE | T. Mens and T. Tourwé | 30(2), pp 126–139 | 2004 | A survey of software refactoring | Informal literature survey |
| TSE | S. Balsamo, A. Di Marco, P. Inverardi | 30(5), pp. 295–309 | 2004 | Model-based performance prediction in software development | Informal literature survey |
| IET Software | S. Mahmood, R. Lai and Y.S. Kim | 1(2), pp 57–66 | 2007 | Survey of component-based software development | Informal literature survey |
| IEEE Software | D.C. Gumm | 23(5) pp. 45–51 | 2006 | Distribution dimensions in software development | Literature survey referenced but not described in article |
| IEEE Software | M. Shaw and P Clements | 23(2) pp. 31–39 | 2006 | The golden age of software Architecture | Informal literature survey |
| IEEE Software | M. Aberdour | 24(1), pp. 58–64 | 2007 | Achieving quality in open source software | Informal literature survey |
| IEEE Software | D. Damian | 24(2), pp. 21–27 | 2007 | Stakeholders in global requirements engineering: lessons learnt from practice | Informal literature survey |
| JSS | E. Folmer and J. Bosch | 70, pp. 61–78 | 2004 | Architecting for usability: a survey | Informal literature survey |
| IST | Hochstein and Lindvall | 47, pp. 643–656 | 2005 | Combating architectural degeneration: a survey | Informal literature survey |
| IST | S. Mahmood, R. Lai, Y.S. Kim, J.H. Kim, S.C. Park, H.S. h | 47, pp. 693–707 | 2005 | A survey of component-based system quality assurance and assessment | Informal literature survey |
| TOSEM | J. Estublier, D. Leblang, A. van der Hoek, R. Conradi, G. Clemm, W. Tichy, D. Wiborg-Weber | pp. 383–430 | 2005 | Impact of software engineering research on the practice of software configuration management | Informal literature survey |
| TOSEM | Barbara G. Ryder, Mary Lou Soffa, Margaret Burnett | pp. 431–477 | 2005 | The impact of software engineering research on modern programming languages | Informal literature survey. No clear search criteria, no data extraction process. |
| ACM Surv | J. Ma and J. V. Nickerson | 38(3), pp. 1–24 | 2006 | Hands-on, simulated and remote laboratories: a comparative literature review | Not a software engineering topic |
| ISESE | S. Wagner | | 2006 | A literature survey of the quality economics of defect-detection techniques | Informal literature survey although quantitative data tabulated for different testing techniques. |

# EXEMPLO 5

## Quality assessment

- Each SLR was evaluated using the York University, Centre for Reviews and Dissemination (CDR) Database of Abstracts of Reviews of Effects (DARE) criteria [3]. The criteria are based on four quality assessment (QA) questions:
  - QA1. Are the review's inclusion and exclusion criteria described and appropriate?
  - QA2. Is the literature search likely to have covered all relevant studies?
  - QA3. Did the reviewers assess the quality/validity of the included studies?
  - QA4. Were the basic data/studies adequately described?

[3] Centre for Reviews and Dissemination, What are the criteria for the inclusion of reviews on DARE? 2007. Available at <http://www.york.ac.uk/inst/crd/faq4.htm, 2007 (accessed 24.07.07>).

# EXEMPLO 5

- Quality assessment
  - QA1:Y (yes), the inclusion criteria are explicitly defined in the study, P (Partly), the inclusion criteria are implicit; N (no), the inclusion criteria are not defined.
  - QA2:Y, the authors have either searched 4 or more digital libraries and included additional search strategies or identified and referenced all journals addressing the topic of interest; P, the authors have searched 3 or 4 digital libraries with no extra search strategies, or searched a defined but restricted set of journals and conference proceedings; N, the authors have search up to 2 digital libraries or an extremely restricted set of journals.
  - QA3:Y, the authors have explicitly defined quality criteria and extracted them from each primary study; P, the research question involves quality issues that are addressed by the study; N no explicit quality assessment.
  - QA4:Y Information is presented about each study; P only summary information about primary studies is presented; N the results are not specified.
  - The scoring procedure was Y = 1, P = 0.5, N = 0, or Unknown (i.e. the information is not specified). [...] When a question was scored as unknown we e-mailed the authors of the paper and asked them to provide the relevant information and the question re-scored appropriately.

**Table 3**
Quality evaluation of SLRs.

| Study | Article type | QA1 | QA2 | QA3 | QA4 | Total score | Initial rater agreement |
|---|---|---|---|---|---|---|---|
| S1 | SLR | Y | P | N | Y | 2.5 | 4 |
| S2 | SLR | Y | P | P | P | 2.5 | 4 |
| S3 | MA | Y | P* | P | P | 2.5 | 4 |
| S4 | SLR | Y | P | N | P | 2 | 4 |
| S5 | SLR | Y | Y | N | Y | 3 | 4 |
| S6 | SLR | Y | P | N | Y | 2.5 | 4 |
| S7 | SLR | Y | Y* | Y | Y | 4 | 4 |
| S8 | SLR | Y | Y | P | Y | 3.5 | 4 |
| S9 | SLR | Y | Y | N | Y | 3 | 4 |
| S10 | SLR | P | N | P | P | 1.5 | 4 |
| S11 | SLR | Y | Y | Y | Y | 4 | 4 |
| S12 | SLR | Y | P* | N | Y | 2.5 | 4 |
| S13 | SLR | Y | N | P | P | 2 | 4 |
| S14 | SLR | Y | Y* | N | Y | 3 | 4 |
| S15 | SLR | P | Y | N | Y | 2.5 | 3 |
| S16 | SLR | P | P | N | P | 1.5 | 3 |
| S17 | SLR | Y | N | N | Y | 2 | 2 |
| S18 | SLR | Y | N | N | N | 1 | 4 |
| S19 | SLR | Y | P | N | P | 2 | 3 |
| S20 | SLR | Y | Y | Y | P | 3.5 | 3 |

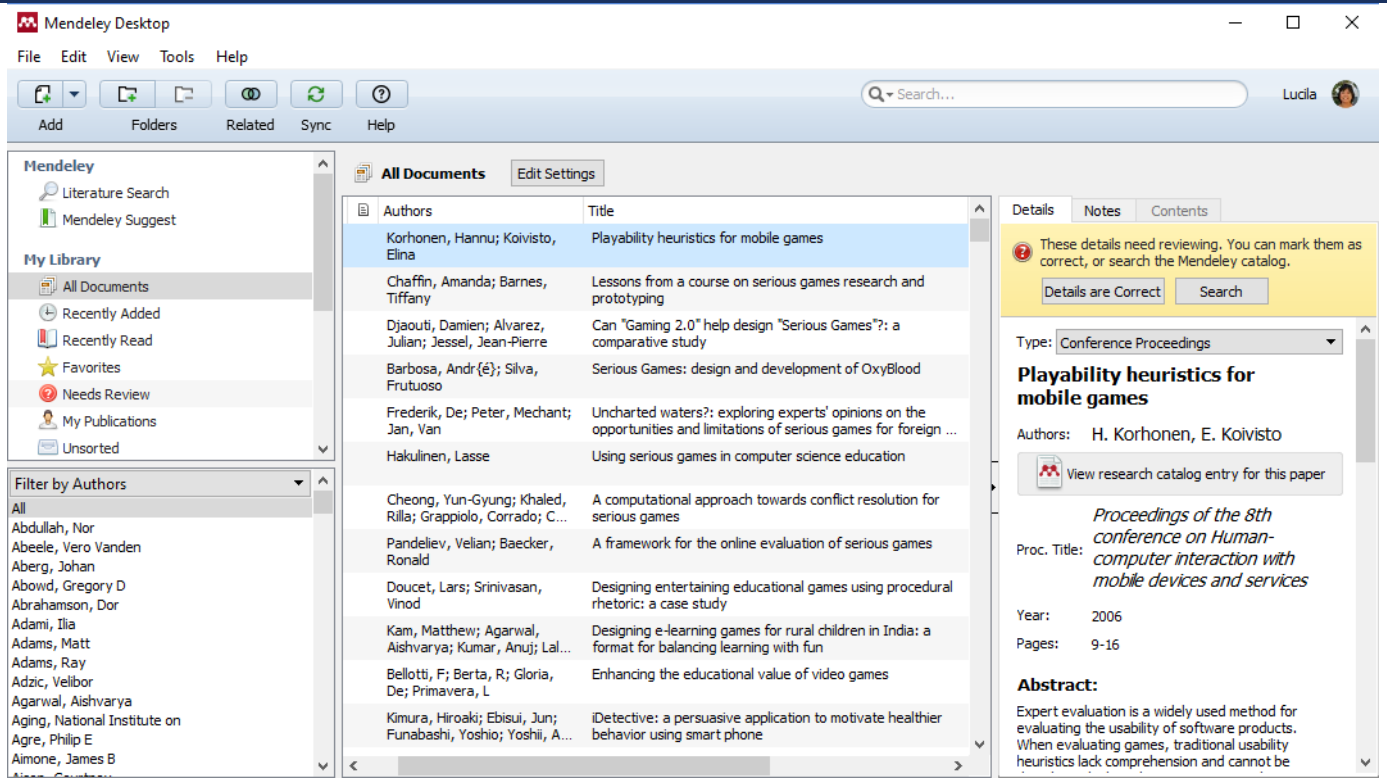
# SOFTWARE DE GESTÃO DE BIBLIOGRAFIA

| Softwares | EndNote Web | Mendeley | Zotero | F1000Workspace |
|---|---|---|---|---|
| **Guias de uso** | Sim | Sim | Sim | Sim |
| **Como pode ser usado** | Na web | No computador e na web | No computador e na web | Na web. No computador (em breve) |
| **Custo** | Grátis* | Grátis | Grátis | Grátis |
| **Identifica registros duplicados** | Sim | Sim | Sim | Sim |
| **Estilos e padrões de Citações** | Sim | Sim | Sim | Sim |
| **Importa registros de bases de dados** | Sim | Sim | Sim | Sim |
| **Compartilha dados e bibliotecas** | Sim | Sim | Sim | Sim |
| **Permite anotações** | Sim | Sim | Não | Sim |
| **Integra processadores de texto** | Microsoft Word | Microsoft Word LibreOffice | Microsoft Word LibreOffice | Microsoft Word Google Docs |

*A versão Desktop do EndNote é paga.

Fonte: http://www.sibi.usp.br/apoio-pesquisador/gerenciadores-referencias-citacoes/

# SOFTWARE DE GESTÃO DE BIBLIOGRAFIA



# REFERÊNCIAS

B. Kitchenham, "Procedures for Undertaking Systematic Reviews," Joint Technical Report, Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd (0400011T.1) 2004.

P. Brereton, B. A. Kitchenham, D. Budgen, M. Turner, and M. Khalil, "Lessons from applying the systematic literature review process within the software engineering domain," Journal of Systems and Software, vol. 80(4), pp. 571-583, 2007.

B. A. Kitchenham, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Version 2.3, EBSE Technical Report, EBSE-2007-01, Software Engineering Group, School of Computer Science and Mathematics, Keele University, Keele, Staffs, ST5 5BG, UK and Department of Computer Science, University of Durham, UK, 2007.

B. Kitchenham, P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering - A systematic literature review," Information and Software Technology, vol. 51(1), pp. 7-15, 2009.

# REFERÊNCIAS


Kitchenham, B., Mendes, E., Travassos, G.H. (2007) A Systematic Review of Cross- vs. within-company cost estimation studies, IEEE TSE, 33 (5), 316-329.

Darja Šmite , Claes Wohlin,Tony Gorschek, Robert Feldt. (2010). Empirical evidence in global software engineering: a systematic review. Empirical Software Engineering, 15, 91–118.

Marcela Genero, Ana M. Fernández, H. James Nelson, Geert Poels, Mario Piattini. (2011). A Systematic Literature Review on the Quality of UML Models. Journal of Database Management, 22(3), 46-70.

Sérgio A. Loureiro. Revisão Sistemática da Literatura. Disponível em: https://goo.gl/7zUWgP. Acesso em: 06 abr. 2013.

Rick Rabiser. Systematic Literature Reviews: Introduction and Personal Experiences. Euromicro SEAA 2011 – Research Methods Tutorial. Disponível em: https://goo.gl/HGtsHg. Acesso em: 06 abr. 2013.

Hall, T. et al. "A Systematic Literature Review on Fault Prediction Performance in Software Engineering," Software Engineering, IEEE Transactions on , vol.38, no.6, pp.1276-1304, Nov.-Dec. 2012