

Um Mapeamento Sistemático Preliminar sobre Frameworks de Avaliação de Sistemas Legados

Jonnathan R. Lopes, Lukas F. Gaedicke, Andréa S. Bordin

¹Universidade Federal do Pampa (UNIPAMPA)
Av. Tiarajú, 810, Ibirapuitã – Alegrete, RS – Brasil

{jonnathan.riquelmo, lukasfgaedicke}@gmail.com, andreabordin@unipampa.edu.br

Abstract. *Legacy systems need to be evaluated in order to know the best evolution strategy, it's not an easy task due to the inherent complexity that this type of system can have. In this way, it's important to know what aspects should be observed in this evaluation process. This paper presents a preliminary systematic mapping of literature which presents the dimensions and criteria for the evaluation of this type of system. The data collected in 17 studies show take into account technical and business aspects comprised of several criteria.*

Resumo. *Sistemas legados necessitam ser avaliados com o propósito de se saber qual a melhor estratégia de evolução, o que não é uma tarefa fácil devido à complexidade inerente que esse tipo de sistema pode possuir. Assim, é importante saber quais os aspectos que devem ser observados nesse tipo de avaliação. Esse estudo apresenta um mapeamento sistemático preliminar da literatura, no qual são apresentadas as dimensões e os critérios para a avaliação desse tipo de sistema. Os dados coletados em 17 estudos mostram que a avaliação deve levar em conta aspectos técnicos e de negócio compostos por diversos critérios.*

1. Introdução

O termo sistema legado possui definições na literatura que descrevem seus diversos aspectos. Para [Bennett 1995] são sistemas que foram desenvolvidos em algum momento do passado, mas que ainda mantêm um alto valor agregado e que muitas vezes são considerados críticos para as organizações que os possuem. Esses sistemas habitualmente foram desenvolvidas sem os métodos de engenharia de software adequados e frequentemente são mantidos para acomodar novas funcionalidades com o passar do tempo [Ransom et al. 1998]. Mais recentemente, [Johann 2016] corroborou com a ideia de que sistemas legados são importantes para o negócio porque geram receitas e que, em muitos casos, são o principal sistema de suporte para o negócio.

Por muitas vezes possuírem um alto valor de negócio para as organizações, os sistemas legados podem ser um grande problema. Isso acontece devido às constantes manutenções realizadas ao longo dos anos que acabam deteriorando o sistema como um todo, tornando a manutenibilidade um processo cada vez mais oneroso para as organizações.

Exposto isso, é notável que evoluir sistemas legados não é uma tarefa trivial. A tomada de decisão em relação ao que fazer com sistemas desse tipo exige uma análise ponderada e criteriosa de vários aspectos. Dessa forma, se faz necessário saber quais as

dimensões e critérios que devem ser analisados para uma correta avaliação do sistema e também quais as possíveis decisões a serem tomadas em relação a esses sistemas.

Na literatura existem algumas proposições de *frameworks* para o auxílio na tomada de decisão. A maioria propõe dimensões e conjuntos de critérios que devem ser observados quando um procedimento de avaliação de sistemas legados é realizado.

Assim, o objetivo deste artigo é realizar um mapeamento sistemático (MS) preliminar da literatura buscando investigar as propostas para avaliação de sistemas legados e as respectivas características que especificam as dimensões e critérios que devem ser levados em conta nesse tipo de processo. O trabalho relacionado que mais se aproxima desta pesquisa é o de [Agilar et al. 2016], um MS sobre processos, técnicas e ferramentas para apoiar a modernização de sistemas legados. Contudo, esse trabalho não aborda especificamente questões relacionadas às dimensões e critérios de avaliação de sistemas, as quais antecedem o processo de modernização em si.

A estrutura do artigo segue conforme especificado: na Seção 2 é descrita a metodologia do protocolo executado, a Seção 3 apresenta os resultados obtidos através do mapeamento e, finalmente, a Seção 4 expõe as considerações finais do estudo elaborado.

2. Metodologia

2.1. Protocolo do Mapeamento Sistemático

Para a obtenção de resultados confiáveis e reproduzíveis é vital que um processo bem definido e estruturado seja seguido. O protocolo de MS utilizado para este estudo foi o de Petersen *et al.* 2008 [Petersen et al. 2008], cujas etapas são exibidas na Figura 1.

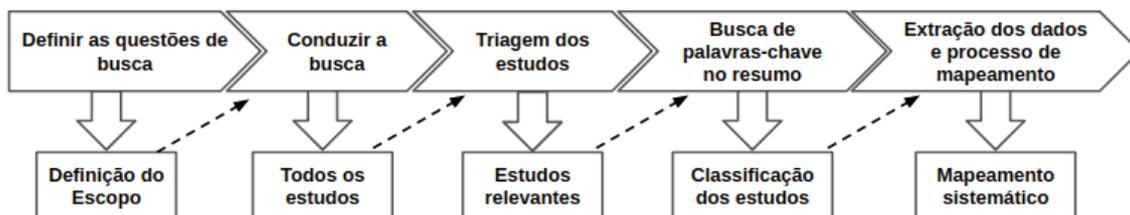


Figura 1. Processo de Mapeamento Sistemático [Petersen et al. 2008]

Em linha gerais, o protocolo visa definir todas as etapas e atividades necessárias para conduzir o MS, como as questões de pesquisa (QP), critérios de seleção, bases a serem utilizadas, a estratégia de leitura dos estudos recuperados e a síntese dos resultados.

2.2. Bases de Busca

Para a realização desse MS utilizou-se bases de dados que possuem: (I) mecanismo de pesquisa baseado na web; (II) mecanismo de busca capaz de usar palavras-chave; e (III) documentos da grande área da Ciência da Computação. As bases selecionadas para a realização do trabalho foram: *ACM Digital library*, *Engineering Village (Compendex)*, *IEEE Xplore* e *Scopus*. Todas as bases são amplamente reconhecidas no meio acadêmico.

2.3. String de Busca

Foi definido um conjunto de palavras referentes ao tema de pesquisa, bem como sinônimos considerados expressivos. Para a construção da *string* foram investigados os termos

que representam o tema investigado, sendo eles, *e.g.* Legacy System, Legacy Application, Framework, Assessment, Management, entre outros.

```
(legacy system OR legacy software OR legacy information system OR  
legacy software system OR software modernization OR Legacy Code OR  
Legacy Application) AND (framework OR decision making OR  
decision-making OR assessment OR evaluation OR management OR model)
```

Figura 2. String Genérica de Busca

Após o levantamento de termos centrais e seus sinônimos houve a construção de uma *string* padrão, exibida na Figura 2. Contudo, é importante ressaltar que, para cada uma das bases de dados utilizadas, a *string* padrão sofreu modificações. Isso se fez necessário para ter a adequação quanto à parâmetros exigidos pelos mecanismos de busca. Para a validação da *string* foram realizadas buscas-piloto. Com um conjunto de artigos predefinidos que deveriam retornar, foi efetuado a pesquisa e então verificado quantos destes trabalhos eram recuperados. Para tanto, mediante os resultados finais desse processo, as *strings* foram consideradas satisfatórias para o prosseguimento da MS.

2.4. Questões de Pesquisa

Após a definição do contexto e objetivo foram elaboradas três questões de pesquisa (QP) que guiaram o restante do MS: **QP1** - *Quando e onde os estudos têm sido publicados?* ; **QP2** - *Quais dimensões são abordadas no processo de avaliação de sistemas legados?* ; **QP3** - *Em cada dimensão quantos critérios são utilizados no processo de avaliação de sistemas legados?*.

2.5. Critérios de Seleção de Estudos

Os critérios de inclusão indicam por qual ou quais critérios um estudo é incluído no MS, ou seja, considerado relevante. Da mesma forma, os de exclusão indicam por qual ou quais critérios um estudo é excluído, ou seja, considerado não relevante [Nakagawa et al. 2017]. Nesse MS os critérios da Tabela 1 foram aplicados.

Tipo	Critério
Inclusão	CII. <i>O estudo deve abordar um framework ou processo para avaliação de sistemas legados.</i>
Exclusão	CE1. <i>O estudo está duplicado.</i>
Exclusão	CE2. <i>O estudo não está escrito em inglês.</i>
Exclusão	CE3. <i>O estudo não fornece acesso completo ao seu conteúdo.</i>
Exclusão	CE4. <i>O estudo não atende ao CII.</i>

Tabela 1: Critérios de Seleção

2.6. Condução do Mapeamento Sistemático

A condução prática do MS iniciou-se com as buscas sendo realizadas nas bases de dados selecionadas. Na Tabela 2 são apresentados os resultados em relação a cada base, bem como o total de estudos primários recuperados. Após, deu-se início a etapa de seleção.

A avaliação dos trabalhos foi realizada em *peer review* (revisão por pares), por meio de quatro iterações, no qual diferentes atividades foram realizadas para alcançar o

Base de Dados Pesquisada	Estudos	% do Total
ACM Digital library	101	9%
Engineering Village (Compendex)	335	31%
IEEE Xplore	136	12%
Scopus	562	48%
TOTAL	1134	100%

Tabela 2: Estudos Recuperados por Bases de Dados

subconjunto final de trabalhos que foram analisados neste MS. A Figura 3 apresenta os ciclos e os artigos restantes em cada iteração.

É importante salientar que houve a inclusão de quatro trabalhos sob a orientação de uma especialista com experiência na área de sistemas legados. Os estudos incluídos somaram-se aos 13 trabalhos resultantes do último ciclo de seleção, o que implicou em um total de 17 trabalhos base para este estudo.

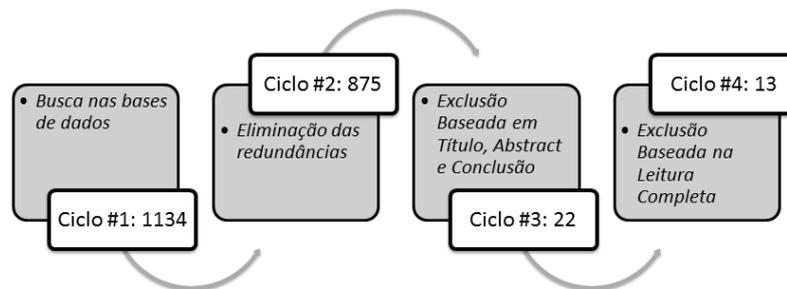


Figura 3. Ciclos de Avaliações dos Trabalhos

2.7. Estratégia de Extração dos Dados

Com o objetivo de registrar as informações necessárias para responder as QPs estabelecidas, foram definidos os dados pertinentes relacionados, sendo eles: **Ano**; **Autor(es)**; **Filiação**; **Dimensões**, ou espectro de análise de *sistemas legados*; **Crítérios**, ou subconjunto de variáveis observadas nas dimensões de análise.

Esse conjunto de informações foi extraído novamente com uma revisão em par e então tabulado. O resultado auxiliou na formulação da síntese dos dados. Após, foi realizada a discussão dos resultados frente às QPs do estudo.

3. Resultados e Discussão

Após o fim do processo de seleção, extração e tabulação dos dados, as QPs foram discutidas. Em relação a **QP1** ("Quando e onde os estudos têm sido publicados?"), foi identificado que os estudos selecionados foram publicados a partir do ano de 1995 até o ano de 2013, estando distribuídos de maneira uniforme no período. No entanto, observa-se que nos últimos 4.5 anos não foram publicados estudos que atendessem os critérios de inclusão do protocolo estabelecido. Destaca-se ainda que o estudo de [Sneed 1995] é o primeiro e mais relevante, sendo citado 296 vezes.

Ainda em relação a **QP1**, 5 dos 17 estudos foram publicados na *Euromicro Conference on Software Maintenance and Reengineering [CSMR]*, como pode ser observado

Referência	Evento [Sigla]	h-index
[Sneed 1995]	IEEE Software	96
[Bennett et al. 1999]	IEE Proceedings - Software (Atual IET Software)	96
[Brooke and Ramage 2001]	International Journal of Information Management [IJIM]	82
[De Lucia et al. 2001]	IEEE International Conference on Software Maintenance [ICSM]	50
[Battaglia et al. 1998] [Ransom et al. 1998] [Koskinen et al. 2005] [Aversano et al. 2005] [Bergmayr et al. 2013]	Euromicro Conference on Software Maintenance and Reengineering [CSMR]	41
[Aversano and Tortorella 2004]	Journal of Software Maintenance and Evolution [JSME]	39
[Rajavat and Tokekar 2011]	Communications in Computer and Information Science [CCIS]	35
[O'Byrne and Wu 2000]	International Workshop on Principles of Software Evolution [IWPSE]	28
[Alkazemi 2014]	Journal of Software [JSW]	24
[Jain and Chana 2015]	International Conference on Communication Technology Proceedings [ICCT]	21
[Kankaanpää et al. 2007]	International Conference on Enterprise Information Systems [ICEIS]	11
[Froncowskiakand and Zandoli 2013]	Proceedings of Student/Faculty Research Day, [CSIS] (Pace University - United States)	0
[Alkazemi et al. 2013]	IEEE International Conference on Systems, Man, and Cybernetics [SMC]	N/A

Tabela 3: Estudos Analisados

na Tabela 3. Para atestar a qualidade dos estudos encontrados, foi pesquisado o *h-index*¹ para todos os trabalhos selecionados. Este fator foi proposto em 2005 por Jorge E. Hirsch e serve como indicativo de qualidade e notoriedade de autores, revistas científicas e eventos. Desta forma, os estudos foram agrupados por evento e publicação, e então ordenados de maneira crescente utilizando por base o ano de publicação.

Para a **QP2** ("*Quais dimensões são abordadas no processo de avaliação de sistemas legados?*"), após a análise dos estudos foram constatadas evidências que colaboraram para o arranjo dos dados, conforme a Tabela 4. A grande maioria dos estudos apontam que os sistemas legados são avaliados segundo duas ou, em alguns casos, até três dimensões, sendo a de Negócio e a Técnica as dimensões mais expressivas.

Existem ainda algumas propostas que abordam a dimensão técnica subdividindo-a em dois tópicos: um que aborda a infraestrutura de software e hardware necessária para o sistema legado, e outra que se refere ao próprio sistema em avaliação. Essa categorização foi adotada na síntese dos dados e pode ser observado na Tabela 4, onde os critérios técnicos foram divididos em dois tipos denotados por *CR-Tec/Sof* e *CR-Tec/Apoio*.

Apesar de alguns estudos apontarem uma quantidade maior de dimensões, foi inferido que as mesmas podem ser classificadas nas duas dimensões já mencionadas. Ainda assim, cabe destacar o estudo detalhado de [Koskinen et al. 2005] que propôs 13 dimensões, as quais foram categorizadas neste MS nessas duas dimensões citadas, *e.g.* *11.Processos de Negócio* é equivalente a dimensão de negócio.

Finalmente para a **QP3** ("*Em cada dimensão quantos critérios são utilizados no processo de avaliação de sistemas legados?*"), observa-se na Tabela 4 que a maioria dos estudos sugeriram critérios de avaliação, com exceção de [Battaglia et al. 1998]. Por outro lado, o estudo de [Koskinen et al. 2005] novamente destaca-se como o que mais cita critérios de avaliação.

É importante ressaltar que o maior número de critérios situa-se na dimensão técnica relacionada ao próprio sistema em avaliação, perfazendo 52% de todos os critérios.

¹<https://www.scimagojr.com/>

REFERÊNCIA	DIMENSÃO	CR Negócio	CR-Tec/Soft	CR-Tec/Apoio	TOTAL
[Sneed 1995]	1.Qualidade Técnica; 2.Valor de negócio	3	25	0	28
[Battaglia et al. 1998]	1.Qualidade Técnica; 2.Valor de negócio	0	0	0	0
[Ransom et al. 1998]	1.Avaliação de valor de negócios; 2.Avaliação de Ambiente Externo; 3.Avaliação de Aplicação	4	9	20	33
[Bennett et al. 1999]	Organização; Técnico	0	6	1	7
[O'Byrne and Wu 2000]	1.Adequação do sistema; 2.Adequação da Plataforma Subjacente; 3.Qualidade do sistema	3	3	5	11
[Brooke and Ramage 2001]	Valor de negócio; Valor técnico	3	6	0	9
[De Lucia et al. 2001]	1.Valor de negócio; 2.Valor técnico	9	12	3	24
[Aversano and Tortorella 2004]	1.Tecnologias 2.Organização 3.Processo 4.Sistemas Legados	7	8	12	27
[Koskinen et al. 2005]	1.Características gerais da aplicação; 2.Requisitos do usuário; 3.Fatores de manutenção; 4.Mantenedores; 5.Processo de manutenção; 6.Qualidades técnicas do sistema; 7.Ferramentas básicas de desenvolvimento; 8.Documentação do sistema; 9.Projeto do sistema; 10.Mantenibilidade do sistema; 11.Processo de negócio; 12.Mudanças de tecnologia; 13.Qualidade do sistema;	2	35	12	49
[Aversano et al. 2005]	1.Valor de negócio; 2.Valor técnico; 3.Sistema legado	3	7	0	10
[Kankaanpää et al. 2007]	1.Valor técnico; 2.Valor de negócio	10	6	0	16
[Rajavat and Tokekar 2011]	1.Domínio do sistema; 2.Domínio gerencial; 3.Domínio técnico	2	2	2	6
[Bergmayr et al. 2013]	1.Negócio; 2.Técnico	1	1	0	2
[Fronckowiak and Zandoli 2013]	1.Negócio; 2.Técnico; 3.Processo de decisão	9	7	4	20
[Alkazemi et al. 2013]	1.Suporte; 2.Negócio; 3.Arquitetura; 4.Tecnologia	3	6	8	17
[Alkazemi 2014]	1.Suporte; 2.Negócio; 3.Arquitetura; 4.Tecnologia	3	6	8	17
[Jain and Chana 2015]	1.Avaliação da aplicação; 2.Qualidade de Serviço; 3. Avaliação da nuvem de destino	0	6	3	9
TOTAL		62	149	74	285

Legenda - **CR**: Critérios, **Tec/Soft**: Técnico/Software, **Tec/Apoio**: Técnico/Apoio.

Tabela 4: Dados Extraídos

Por outro lado, a soma com os 26% que representam os critérios técnicos de apoio ao software totalizam 78%, fornecendo um número expressivo de proposições na dimensão técnica. Os critérios relacionados à dimensão de negócio representam apenas 22%.

4. Considerações Finais

Por meio da análise dos dados do MS realizado foi possível evidenciar que existem diversas proposições de *frameworks* de avaliação de sistemas legados na literatura. Os estudos, em geral, descrevem diversas dimensões e critérios que precisam ser considerados e analisados cautelosamente antes da tomada de decisão por conta do alto valor que os sistemas legados podem representar para as organizações que mantêm seu funcionamento.

A partir dos resultados obtidos é possível observar uma maior frequência de critérios relacionados à dimensão técnica, visto que os dados expostos na seção 3, demonstram um total de 293 critérios técnicos contra apenas 62 critérios de negócio. Um dos motivos dessa tendência pode ser atribuída ao fato dos pesquisadores serem especialistas da área da Computação, e por este fato, acabarem não caracterizando alguns aspectos mais específicos e de difícil identificação na dimensão de negócio. Os resultados obtidos através desse estudo poderão ser utilizados por mantenedores e pesquisadores da área como base para a geração de um arcabouço de conhecimento das dimensões e critérios indispensáveis na análise de sistemas legados.

Como trabalhos futuros, será realizado o *snowballing* dos trabalhos já incluídos nesse estudo e, posteriormente o mapeamento será complementado com a identificação dos critérios mais utilizados nas dimensões, assim como a identificação das alternativas de evolução decorrentes dos processos de avaliação, descritas nos *frameworks* dos estudos analisados. Essa atividade complementar a qualidade do presente trabalho, uma vez que será possível realizar associações entre as dimensões e critérios com as decisões de evolução mais adequadas a serem tomadas, *e.g.* reengenharia, migração ou substituição, em relação a sistemas legados nas organizações.

Referências

- Agilar, E., Almeida, R., and Canedo, E. (2016). A systematic mapping study on legacy system modernization. In *SEKE*, pages 345–350.
- Alkazemi, B. Y. (2014). A framework to assess legacy software systems. *Journal of Software*, 9(1):111–115.
- Alkazemi, B. Y., Nour, M. K., and Meelud, A. Q. (2013). Towards a Framework to Assess Legacy Systems. *2013 IEEE International Conference on Systems, Man, and Cybernetics*, pages 924–928.
- Aversano, L., Esposito, R., Mallardo, T., and Tortorella, M. (2005). Evolving legacy system toward elegacy system in ebusiness context. In *CSMR, CSMR '07*, pages 201–, Washington, DC, USA. IEEE Computer Society.
- Aversano, L. and Tortorella, M. (2004). An assessment strategy for identifying legacy system evolution requirements in eBusiness context. *Journal Of Software Maintenance And Evolution*, pages 255–276.
- Battaglia, M., Savoia, G., and Favaro, J. (1998). Renaissance: A method to migrate from legacy to immortal software systems. *CSMR*, 1998-March(D):197–200.
- Bennett, K. (1995). Legacy systems: coping with success. *IEEE Software*, 12(1):19–23.
- Bennett, K. H., Ramage, M., and Munro, M. (1999). Decision model for legacy systems. *IEE Proceedings - Software*, 146(3):153–159.

- Bergmayr, A., Brunelière, H., Izquierdo, J. L. C., Gorroñoigoitia, J., Kousiouris, G., Kyriazis, D., Langer, P., Menychtas, A., Orue-Echevarria, L., Pezuela, C., and Wimmer, M. (2013). Migrating legacy software to the cloud with ARTIST. *CSMR*, pages 465–468.
- Brooke, C. and Ramage, M. (2001). Organisational scenarios and legacy systems. *International Journal of Information Management*, 21(5):365–384.
- De Lucia, A., Fasolino, A. R., and Pompella, E. (2001). A decisional framework for legacy system management. *IEEE International Conference on Software Maintenance, ICSM*, pages 642–653.
- Fronckowiakand, D. and Zandoli, R. (2013). A new algorithmic approach to the it modernization problem. *Proceedings of Student/Faculty Research Day, CSIS, Pace University*, pages 280–287.
- Jain, S. and Chana, I. (2015). Modernization of Legacy Systems: A Generalised Roadmap. *Proceedings of the Sixth International Conference on Computer and Communication Technology 2015*, pages 62–67.
- Johann, S. (2016). Dave thomas on innovating legacy systems. *IEEE Software*, (2):105–108.
- Kankaanpää, I., Tiihonen, P., Ahonen, J., Koskinen, J., Tilus, T., and & Sivula, H. (2007). Legacy system evolution - a comparative study of modernisation and replacement initiation factors. *ICEIS 2007*, pages 280–287.
- Koskinen, J., Ahonen, J. J., Sivula, H., Tilus, T., Lintinen, H., and Kankaanpaa, I. (2005). Software modernization decision criteria: an empirical study. In *CSMR*, pages 324–331.
- Nakagawa, E., Scannavino, K., Fabbri, S., and Ferrari, F. (2017). *Revisão Sistemática da Literatura em Engenharia de Software: Teoria e Prática*. Elsevier Editora Ltda.
- O’Byrne, P. and Wu, B. (2000). Lace frameworks and technique-identifying the legacy status of a business information system from the perspectives of its causes and effects. In *Proceedings International Symposium on Principles of Software Evolution*, pages 170–174.
- Petersen, K., Feldt, R., Mujtaba, S., and Mattsson, M. (2008). Systematic mapping studies in software engineering. In *EASE*, volume 8, pages 68–77.
- Rajavat, A. and Tokekar, V. (2011). ReeRisk - A decisional risk engineering framework for legacy system rejuvenation through reengineering. *CCIS*, pages 152–158.
- Ransom, J., Somerville, I., and Warren, I. (1998). A method for assessing legacy systems for evolution. *CSMR*, pages 128–134.
- Sneed, H. M. (1995). Planning the Reengineering of Legacy Systems. *IEEE Software*, 12(1):24–34.