

# Descrevendo a plataforma de Serviços de Recomendação RecS-DL utilizando o Framework 5S

Daniel Carlos Guimarães Pedronette<sup>1</sup>, Ricardo da Silva Torres<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Estadual de Campinas (Unicamp)  
CEP 13081-970 – Campinas – SP – Brasil

daniel.pedronette@students.ic.unicamp.br, rtorres@ic.unicamp.br

**Abstract.** *This paper presents a formal description of the platform RecS-DL using the framework 5S.*

**Resumo.** *Este artigo apresenta um conjunto de definições e descrições formais da plataforma de recomendação RecS-DL, utilizando o framework 5S.*

## 1. Introdução

Teorias e *frameworks* formais são cruciais para especificar e entender de forma clara e não ambígua as características, estruturas e o comportamento de sistemas de informação complexos, como os sistemas de bibliotecas digitais.

*Frameworks* formais abstraem as características gerais e comuns de um conjunto de sistemas desenvolvidos para problema similares, explicando suas estruturas, processos e práticas comuns. Além disso, podem ser utilizados no projeto de sistemas reais, provendo especificações precisas de requisitos. Estas especificações, por sua vez podem ser comparados à implementação para verificações de corretude [Gonçalves 2004].

O *framework* 5S provê formalismos, uma linguagem própria e uma unificação teórica e prática de componentes, interações e serviços aplicados às bibliotecas digitais. Este framework baseia-se nas definições de:

- *Streams*: tipos de informações multimídia suportados pela biblioteca digital;
- *Structures*: como as informações são estruturadas e organizadas;
- *Spaces*: propriedades e operações dos componentes da biblioteca digital;
- *Scenarios*: serviços e comportamento da biblioteca digital;
- *Societies*: conjunto de atores e gerenciadores que atuam nos serviços.

O framework 5S define formalmente todos os principais conceitos relacionados a Bibliotecas Digitais. Permite ainda extensões de conceitos abordados e a descrição formal de sistemas de Bibliotecas Digitais. Este artigo tem como objetivo estender algumas definições do framework 5S e utilizá-lo para definir formalmente a plataforma RecS-DL.

## 2. Estendendo o conceito de Recomendação através do Framework 5S

O framework 5S define o conceito principal de um Serviço de Recomendação. Neste artigo, apresentaremos definições que estendem esse conceito e apresentam uma abordagem mais detalhada dos conceitos relacionados a *Recomendação*. Algumas dessas definições serão utilizadas para definição da plataforma RecS-DL.

**Recomendação colaborativa:** Dada uma coleção, um ator e um conjunto de ratings para objetos dessa coleção, produzir um subconjunto da coleção para esse ator particular.

**Recomendação baseada no conteúdo:** Dada uma coleção e um subconjunto de objetos dessa coleção, produzir outro subconjunto da coleção que se assemelhe ao subconjunto dado.

**Definição 1.** *Um serviço de recomendação baseado no conteúdo  $S_{Rec}$  é um conjunto de cenários  $S_{Rec} = \{srec_1, srec_2, ..., srec_t\}$  onde cada cenário  $srec_i$  corresponde a uma sequência de eventos onde cada evento  $e_i$  é associado a uma função  $OP_{Rec}$  definida como segue:*

$OP_{rec} : (C \times H_{rec}) \times Sim_s \rightarrow 2^{Conteudo}$ , onde:

- $H_{rec} : A \rightarrow H$ , é uma função que mapeia um Ator  $A$  a um conjunto de dados de histórico  $H = \{h_1, h_2, ..., h_n\}$ , onde  $h_i$  identifica um objeto acessado pelo ator  $A$ ;
- $Sim_s = OP_{rec}(h, ObjDig) \mid h \in H, OP_{rec} : H \times C \rightarrow R$  é uma função que associa um número real a um objeto histórico  $h \in H$  e a um objeto digital  $ObjDig$ .

A função  $OP_{rec}$  define o tipo de Recomendação provida pelo serviço. Na Recomendação Colaborativa, o cálculo de  $OP_{rec}$  é dado pela aplicação de algoritmos colaborativos ao de histórico de comportamento do usuário. Já na Recomendação baseada em conteúdo, o cálculo de  $OP_{rec}$  consiste no uso de técnicas recuperação de informações aplicadas aos objetos contidos no histórico do usuário. Em ambos os casos, o número real resultante representa a adequação de dado objeto da coleção aos dados de histórico, onde as que obtiverem os maiores valores compõem o subconjunto  $2^{Conteudo}$  recomendado.

### 3. Estendendo conceitos de Software através do Framework 5S

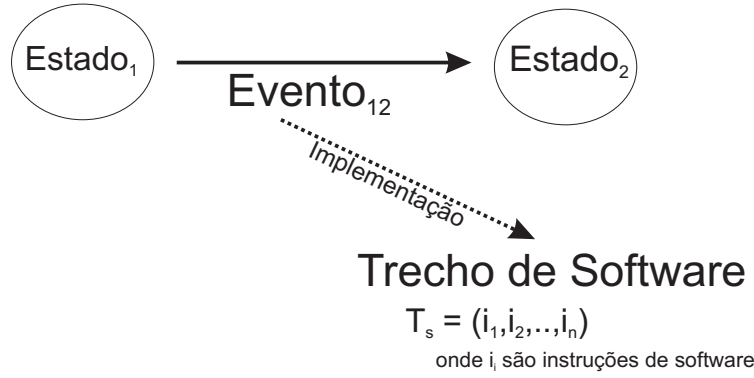
Serão apresentadas a seguir, as definições propostas relacionadas ao conceito de software. Essas definições, por sua vez, serão utilizadas adiante para compor outras definições mais complexas que constituirão a descrição da plataforma RecS-DL.

**Definição 1.** *Um trecho de software executável  $T_s = (i_1, i_2, ..., i_n)$  é um conjunto de instruções de software através dos quais é implementado um evento  $e_s$ . (ver Def. 6 [Gonçalves 2004]).*

A Figura 1 apresenta uma ilustração de um evento de transição e como esse evento relaciona-se com o conceito de Trecho de Software, dado pela Definição 1.

**Definição 2.** *Um método de software é uma tupla  $M_s = (T, I_e, I_s, I_{me}, F_s, c_s)$  responsável pela implementação de um cenário  $c_s$  (ver Def. 7 [Gonçalves 2004]), onde:*

- $T = (t_1, t_2, ..., t_n)$  é um conjunto de trechos de software executáveis;
- $I_e$  é uma interface de dados de entrada;
- $I_s$  é uma interface de dados de saída;
- $I_{me} = (i_{me1}, i_{me2}, ..., i_{men})$  é um conjunto de interfaces de acesso a métodos externos;



**Figura 1. Trecho de Software.**

- $F_s: \{T \times I_e \times I_{me}\} \rightarrow I_s$  é uma função que mapeia os trechos de software executáveis, as interfaces de dados de entrada e métodos externos aos dados de saída.

**Definição 3.** Um *objeto digital de software* é um objeto digital (para maiores detalhes, ver Def. 16 [Gonçalves 2004]),  $ObjDigSoft = (h_{desc}, C_{Mt}, C_{Md}, C_{Rep}, C_{Cn})$ , que atende às seguintes extensões e restrições:

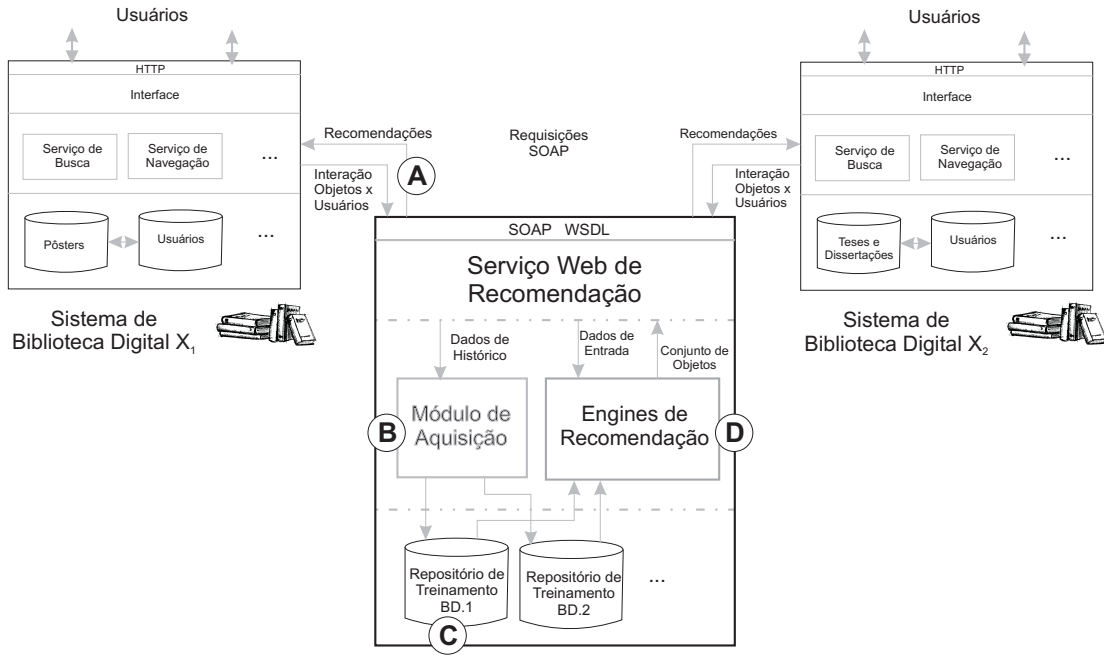
- $h_{desc} \in H$ , representa um conjunto de identificadores únicos (rótulos);
- $C_{Mt} = (mt_1, mt_2, \dots, mt_n)$  é um conjunto de métodos de software;
- $C_{Md} = (md_1, md_2, \dots, md_n)$  é um conjunto de especificações de metadados descritivos (ver Def. 12 [Gonçalves 2004]), responsáveis por armazenar informações sobre o software, interfaces e parâmetros de configuração;
- $C_{Rep} = (R_1, R_2, \dots, R_n)$  é um conjunto de repositórios acessados pelos métodos de software do conjunto  $C_{Mt}$ , para armazenamento ou recuperação de informações;
- $C_{Cn} = (cn_1, cn_2, \dots, cn_n)$  é um conjunto de cenários implementado pelos métodos de software do elemento  $C_{Mt}$ , e  $C_{Cn} \subset Serv$ , onde  $Serv$  é um serviço definido total ou parcialmente pelo conjunto  $C_{Cn}$ .

**Definição 4.** Uma *plataforma de software* é um ambiente para oferecimento de serviços através da execução de objetos de digitais de software, definida por uma tupla  $P = (C_{Od}, C_{Serv}, C_{Rep}, C_{Md})$ , onde:

- $C_{Od}$  conjunto de objetos digitais de software;
- $C_{Serv}$  conjunto de serviços providos pela plataforma e implementados pelo conjunto  $C_{Od}$ ;
- $C_{Rep}$  é um conjunto repositórios  $R$ , manipulados pela plataforma de software;
- $C_{Md}$  é um conjunto de especificações de metadados descritivos sobre os serviços, objetos digitais de software e repositórios manipulados.

#### 4. Definição da Plataforma RecS-DL

Serão descritos a seguir todos os principais componentes da plataforma RecS-DL, desde os principais módulos até os serviços e métodos oferecidos. Uma visão geral da plataforma é ilustrada na Figura2, cujos elementos serão referenciados e explicados nas definições seguintes.



**Figura 2. Arquitetura Geral da Plataforma RecS-DL.**

**Definição 1.** Seja uma plataforma de software definida como  $P = (C_{Od}, C_{Serv}, C_{Rep}, C_{Md})$ , então existe uma plataforma **RecS-DL**, onde:

- $C_{Od}$  é um conjunto de objetos digitais dado por  $C_{Od} = (Módulo\ de\ Aquisição, Módulo\ de\ Configuração, E_{Rec_1}, E_{Rec_2}, \dots, E_{Rec_n})$ , onde cada  $E_{Rec_i}$  representa um *Engine de Recomendação*;
- $C_{Serv}$  é formado pelo conjunto serviços  $S = (Serviço\ de\ Recomendação, Serviço\ de\ Configuração)$  providos pela plataforma.
- $C_{Rep}$  representa um conjunto de repositórios configuráveis através do Serviço de Configuração.
- $C_{Md}$  é dado pelo conjunto  $M = (Arquivo\ de\ Configurações\ Gerais\ da\ Plataforma, Arquivos\ de\ Configuração\ de\ Conexões\ com\ Repositórios)$ .

**Definição 2.** Seja um *serviço* um conjunto de *cenários*, então um **Serviço de Recomendação** da plataforma **RecS-DL**, consiste no conjunto de cenários implementados pelo *Módulo de Aquisição* e pelos *Engines de Recomendação* da plataforma.

**Definição 3.** Seja um objeto digital de software definido como  $ObjDigSoft = (h_{desc}, C_{Mt}, C_{Md}, C_{Rep}, C_{Cn})$ , então existe na plataforma **RecS-DL**, o objeto digital de software **Módulo de Aquisição** (ilustrado na parte B da Figura2), onde:

- $h_{desc}$  é Módulo de Aquisição;
- $C_{Md}$  é dado por um arquivo XML contendo o nome e parâmetros dos métodos. É ilustrado na Figura3 um trecho desse arquivo;
- $C_{Mt}$  representa a implementação do conjunto  $C_{Cn}$ , e é dado pelo conjunto de métodos  $C_{Mt} = (insertItem, insertRating, getUserById, getAllUsers, getItemById, getAllItems, getRatedItems)$ ;

- $C_{Cn}$  representa parte do Serviço de Recomendação (ilustrado na parte A da Figura2) da plataforma RecS-DL, onde  $C_{Cn}$  é dado pelo seguinte conjunto de cenários:
  - **insertItem (item,dbName).** **Cenário:**  $\langle e_1 : p = insertItem, e_2 : execute(insertItem), e_3 : p = response(true,false) \rangle$ , onde  $e_1$  é um evento gerado por uma biblioteca digital invocando uma ação em RecS-DL,  $p$  especifica a ação correspondente que está sendo invocada;  $e_2$  representa a execução do método de software que consiste na inserção do elemento *item* no repositório *dbName*; e  $e_3$  representa o evento de retorno da execução onde  $p = response(true, false)$  define se a execução foi concluída corretamente.
  - **insertRating (rating,dbName).**  
**Cenário:**
  - **getUserById (userId,dbName).**  
**Cenário:**
  - **getAllUsers (dbName).**  
**Cenário:**
  - **getItemById (itemId,dbName).**  
**Cenário:**
  - **getAllItems (dbName).**  
**Cenário:**
  - **getRatedItems (userId,dbName).**  
**Cenário:**
- $C_{Rep}$  é dado por  $C_{Rep} = (R_{Rel})$ , onde  $R_{Rel}$  representa um SGBD relacional (ilustrados na parte C da Figura2)) e cuja conexão é configurada pelo Serviço de Configuração da plataforma RecS-DL;

```

<serviceGroup>
  <service name="RecommenderWS">
    <description>Serviço Web de Recomendação Desenvolvido no Instituto de Computação
      (IC) da Universidade Estadual de Campinas (Unicamp)</description>
    <parameter name="ServiceClass"
      locked="false">br.unicamp.ic.recommender.webservice.RecommenderWS</parameter>
    <!-- ===== -->
    <!-- Métodos de Acesso a Dados -->
    <!-- ===== -->
    <operation name="insertUser">
      <messageReceiver class="org.apache.axis2.rpc.receivers.RPCMessageReceiver" />
      <parameter name="user" locked="false" />
      <parameter name="dbName" locked="false" />
    </operation>
    <operation name="insertItem">
      <messageReceiver class="org.apache.axis2.rpc.receivers.RPCMessageReceiver" />
      <parameter name="item" locked="false" />
      <parameter name="dbName" locked="false" />
    </operation>
    ...
  </service>
</serviceGroup>

```

Figura 3. Metadados dos métodos do objeto digital Módulo de Aquisição

**Definição 4.** Seja um objeto digital de software definido como  $\text{ObjDigSoft} = (h_{desc}, C_{Mt}, C_{Md}, C_{Rep}, C_{Cn}, R, Serv)$ , então existe na plataforma **RecS-DL**, um conjunto de objetos digitais de software  $E = (e_{rec1}, e_{rec2}, \dots, e_{recn})$  (ilustrados na parte D da Figura2), onde cada  $e_{rec}$  representa um **Engine de Recomendação**, onde:

- $h_{desc}$  é nome do *Engine de Recomendação*;
- $C_{Md}$  é um descritor composto por metadados em XML, seguindo o padrão *Dublin Core*. São armazenadas informações sobre o software, versão, autor e parâmetros de configuração;
- $C_{Mt}$  e  $C_{Cn}$  são dados pelos seguinte conjunto de métodos e respectivos cenários:
  - **makeModel (engineClassName, dbName)**  
Cenário:
  - **getRecommendationByUser (user, dbName, engineClassName)**  
Cenário:
  - **getRecommendationByItem (item, dbName, engineClassName)**  
Cenário:
- $R$  é um repositório de treinamento, definido por cada *Engine de Recomendação*. Os repositórios variam de acordo com o engine, desde Sistema de Arquivos a SGBD relacionais.
- $Serv$  é parte do Serviço de Recomendação da plataforma **RecS-DL**.

**Definição 5.** Seja um *serviço* um conjunto de *cenários*, então um **Serviço de Configuração** da plataforma **RecS-DL**, consiste no conjunto de cenários implementados pelo *Módulo Configuração* da plataforma.

**Definição 6.** Seja um objeto digital de software definido como  $\text{ObjDigSoft} = (h_{desc}, C_{Mt}, C_{Md}, C_{Rep}, C_{Cn}, R, Serv)$ , então existe na plataforma **RecS-DL**, um objeto digital de software **Módulo de Configuração**, onde:

- $h_{desc}$  é Módulo de Configuração;
- $C_{Md}$  é dado por um arquivo XML contendo o nome, descrição e parâmetros dos métodos;
- $C_{Mt}$  e  $C_{Cn}$  são dados pelos seguinte conjunto de métodos e respectivos cenários:
  - **installEngine (engineName, xmlFile, jarEncodedFile)**  
Cenário:
  - **getParametersList (engineName)**  
Cenário:
  - **installedEngines**  
Cenário:
  - **getEngineInformation (engineName)**  
Cenário:
  - **activateEngine (engineName, dbName)**  
Cenário:
  - **getDefaultEngine (dbName)**  
Cenário:
  - **getActivatedEngines (dbName)**  
Cenário:

- **setDefaultEngine (dbName,engineClassName)**  
Cenário:
- **getParameter (dbName,engineClassName,parameter)**  
Cenário:
- **setParameter (dbName,engineClassName,parameter,value)**  
Cenário:
- **getParameterValues (dbName,engineClassName)**  
Cenário:
- **createDigitalLibrary (dbName)**  
Cenário:
- **configureDigitalLibrary (dbName,urlDB,driver,user,password)**  
Cenário:
- *R* é o repositório de configuração para cada Biblioteca Digital usuária da plataforma **RecS-DL**.
- *Serv* é o Serviço de Configuração da plataforma **RecS-DL**.

## Referências

Gonçalves, M. A. (2004). *Streams, Structures, Spaces, Scenarios, and Societies (5S): A Formal Digital Library Framework and Its Applications*. PhD thesis.