

# **CONIC-SEMESP**

## 13º Congresso Nacional de Iniciação Científica

Anais do Conic-Semesp. Volume 1, 2013 - Faculdade Anhanguera de Campinas - Unidade 3. ISSN 2357-8904

**TÍTULO:** IMPLEMENTAÇÃO DE MOTORES ROBÓTICOS INTELIGENTES INTEGRADOS COM O REPOSITÓRIO DSPACE

**CATEGORIA:** CONCLUÍDO

**ÁREA:** ENGENHARIAS E TECNOLOGIAS

**SUBÁREA:** COMPUTAÇÃO E INFORMÁTICA

**INSTITUIÇÃO:** CENTRO UNIVERSITÁRIO EURÍPEDES DE MARÍLIA

**AUTOR(ES):** RAFAEL LUIZ DE MACEDO

**ORIENTADOR(ES):** ELVIS FUSCO

Realização:



Apoio:



## 1. RESUMO

O AIML (*Artificial Intelligence Markup Language*) é um conjunto de elementos (TAG) XML (*eXtensible Markup Language*) capaz de representar e relacionar expressões<sup>2</sup> em linguagem natural permitindo a criação de motores robóticos inteligentes capazes de manter um diálogo simples, porém, os motores robóticos são limitados na quantidade de perguntas em que possam responder, assim, não conseguindo manter um diálogo simples por muito tempo. Esse trabalho faz a integração dos motores robóticos inteligentes desenvolvidos na linguagem AIML com o repositório digital da plataforma DSpace, para assim, auxiliar os usuários nas buscas de arquivos digitais contidos nesses repositórios por meios de perguntas e respostas, além, de fazer o uso do software analisador de metadados que é desenvolvido nesse trabalho, para fins, de aplicar um tratamento nas informações armazenadas nos metadados do repositório e gerando novas bases de conhecimento em forma de perguntas e respostas para esses motores robóticos inteligentes. Com essas novas bases de conhecimentos geradas pelo software analisador, os motores robóticos inteligentes poderão realizar perguntas sobre os assuntos relacionados aos arquivos que os usuários estão buscando no repositório.

## 2. INTRODUÇÃO

As tecnologias atuais têm permitido cada vez mais a percepção que é possível fazer uma máquina ser inteligente a ponto de responder perguntas feitas por meio de motores robóticos construídos com inteligência artificial.

A linguagem destina-se à comunicação sobre o mundo. Quando se estuda a linguagem, passa-se a conhecer mais o mundo. Podem-se testar teorias sobre o mundo se conhecem que sustentação elas dão à tentativa de compreender a linguagem. E, se conseguirmos criar um modelo computacional de linguagem, pode-se ter uma ferramenta poderosa para a comunicação sobre o mundo. [RICH, Elaine e KNIGHT, Kevin, 1993, p.433].

Por um lado, as formas em que os seres humanos e computadores normalmente comunicam são muito diferentes. Os seres humanos tendem a gastar um monte de tempo em chit chat e diálogo informal com pouco ou nenhum efeito. Computadores são conhecidos por dar respostas precisas, verdadeiras e lógicas. A taxa de troca de informação de mais diálogo humano é muito baixa, não mais do que 1kbit por segundo, mas a comunicação de computador é mais rápida. Alicebot / AIML

é uma tentativa de colmatar esta divisão [WALLACE, Dr. Richard S., 2003, “The Element of AIML Style”].

Uma das linguagens que possibilita a criação e desenvolvimento desses motores robóticos inteligentes capazes de se manter em um diálogo é a linguagem de marcação AIML.

Os motores robóticos desenvolvidos na linguagem AIML têm uma grande limitação na quantidade de perguntas que possam responder, por principal motivo de ser desenvolvido em cima de um ou mais contextos e não podendo abranger dinamicamente a quantidade de perguntas e respostas em seu banco de informação.

Com o surgimento da *Internet*, as informações estão cada vez mais acessíveis, de forma que, apenas com uma máquina que tenha acesso a rede *Internet* é possível realizar buscas para obter novas informações.

A quantidade de informação que são postadas diariamente na Web tem facilidade cada vez mais o acesso para pessoas acessarem essas informações. Conseqüentemente tem crescido o número de repositórios online que possuem artigos, monografias, jornais, revistas e outros tipos de arquivos digitais e uma das ferramentas mais utilizadas é conhecido como DSpace.

Esse trabalho tem o objeto de integrar os motores robóticos inteligentes com a ferramenta DSpace, para auxiliar nas buscas de arquivos digitais. A integração facilitará os usuários em conseguirem encontrar os arquivos que procuram, por meio de perguntas e respostas que serão trocados entre o motor robótico e o usuário.

Para que seja feita interpretação de informações, serão necessários conceitos de interpretação de linguagem natural. Em particular, o Processamento de Linguagem Natural impõem esses conceitos.

O principal objetivo desse trabalho é integrar um motor robótico com o repositório digital da plataforma DSpace e desenvolver um software analisador de metadados, que junto com o motor robótico realizarão buscas semânticas nos metadados do repositório, para que encontre arquivos relacionados ao assunto do que está sendo buscando e com o software analisador gerará novas bases de conhecimentos de perguntas e respostas para os motores robóticos em cima desses metadados do repositório.

Para a integração do motor robótico com o repositório Dspace e o desenvolvimento do software analisador, será utilizado o repositório do Centro Universitário Eurípides de Marília – UNIVEM que é um repositório da plataforma

Dspace. O motor robótico desenvolvido nesse trabalho, auxiliará os usuários que realizarem buscas de arquivos digitais nesses repositórios.

## 2.1. AIML

A linguagem AIML é um conjunto de elementos (TAG) XML (eXtensible Markup Language) capaz de representar e relacionar expressões em linguagem natural, permitindo a criação de motores capazes de manter um diálogo simples. Cada conjunto de elementos AIML possui um ou mais elementos chamadas de categoria, sendo que as categorias são desenvolvidas em cima de um contexto; uma categoria é formada pelo conjunto dos elementos *template* e *pattern*, que são os elementos responsáveis por interpretar a mensagem pelo usuário e enviar uma mensagem de resposta ao usuário.

O *Chatterbots* (*Chat* = Conversa e *bot* = Robô) como conhecidos atualmente são motores robóticos inteligentes que interage com usuários por meio de perguntas e respostas. Atualmente vem crescendo a quantidade de motores robóticos inteligentes utilizados em atendimentos online e por meio de dispositivos de comunicação por voz.

As empresas tem utilizado esses motores robóticos, para realizarem um auto-atendimento ao cliente e assim agilizando os atendimentos e economizando tempo, por motivo desses motores robóticos fazerem o reconhecimento do assunto que o cliente quer tratar com a empresa, para assim, redirecionar o atendimento ao setor competente.

Um motor robótico bastante conhecido e é utilizado para conscientizar as pessoas em cuidarem do meio ambiente é o robô Ed, da empresa CONCEPT junto com o apoio da Petrobras. [CONCEPT, Agosto 2013.]

O surgimento da linguagem AIML foi resultado do desenvolvimento do software de processamento de linguagem natural (PLN) conhecida A.L.I.C.E. (*Artificial Linguistic Internet Computer Entity*), a licença do AIML está sob a GNU GPL.

Os motores robóticos desenvolvidos na linguagem AIML, possuem dois módulos, linguagem e motor. A linguagem é todo o conhecimento que esses motores possuem, da linguagem natural e das informações desenvolvidas na linguagem AIML. O motor realiza a comunicação entre as duas linguagens, natural e AIML, ocorrendo o reconhecimento das informações contidas nas duas linguagens, para

assim, os motores robóticos poderem reconhecer perguntas escritas feitas pelos humanos e responder essas perguntas.

Existe uma limitação nos módulos dos motores robóticos, os mesmos não conseguem reconhecer perguntas realizadas fora do contexto em que foram desenvolvidos, por não poder reconhecer perguntas de outros contextos que não tenha em seu banco de conhecimento, às conversas não durão muito tempo.

Alan Turing (1912-1954), em seu famoso ensaio “*Computing Machinery and Intelligence*” [Turing 1950], sugeriu que, em vez de perguntar se as máquinas podem pensar, devemos perguntar se máquinas podem passar por um teste de inteligência comportamental, que veio a ser chamado de teste de Turing [Stuart Russell e Peter Norvig, 2004, p.916].

O Teste de Turing é realizado por meio de perguntas e respostas feitas para uma máquina com motor robótico inteligente e para um humano, assim como, a máquina e o humano, cada um é alocado em uma sala separadamente. Um avaliador realizará as perguntas para essas duas salas, mas é feito as perguntas sem o avaliador saber em qual das salas está a máquina com o motor robótico.

Conforme as respostas que o avaliador vai recebendo, o mesmo poderá descobrir em qual das salas se encontra a máquina e na outra o humano. Os motores robóticos inteligentes que reconhecem a linguagem natural, não conseguem criar novas perguntas e respostas de forma autônoma, fazendo com que ocorra uma perda de não conseguirem reconhecer uma pergunta que está fora do seu conhecimento ou repetirem uma resposta já utilizada em perguntas diferentes.

A máquina que conseguir passar no teste de Turing é considerada inteligente, atualmente não existem máquinas capazes de passar por esse teste.

A linguagem AIML é uma linguagem livre (*open source*), assim, possibilitando o uso em pesquisas de melhoria nas criações dos motores robóticos inteligentes capazes de reconhecer informações escrita na linguagem natural.

Uma linguagem de marcação possui várias plataformas gratuitas e pagas para a utilização da linguagem e plataformas desenvolvidas em linguagens de programação como Java, Python, C, C++ e etc.

A plataforma ProgramD é a mais utilizada quando se fala em desenvolvimento de motores robóticos na linguagem AIML e a mais completa em recurso da linguagem, por motivo da plataforma ser desenvolvida na linguagem Java e ser *open source*.

Esse trabalho desenvolverá um motor robótico inteligente voltado para os repositórios digitais da plataforma Dspace, para auxiliar os usuários nas buscas de arquivos digitais contidos no repositório.

## **2.2. DSPACE**

Atualmente tem crescido a quantidade de matérias digitalizadas, como livros, artigos científicos, jornais, monografias e etc. Com os materiais digitalizados, também houve a necessidade de disponibilizar os materiais de modo que qualquer pessoa possa ter acesso, ou seja, porque não disponibilizar esses arquivos de forma online, para assim, qualquer parte do mundo possa-se ter acesso a esses materiais.

Na disponibilização dos arquivos digitais por meio da rede *Internet* foram surgindo algumas ferramentas de repositórios digitais e uma dessas ferramentas mais conhecidas é o DSpace.

O DSpace é uma plataforma de software Open Source de organização de arquivos digitais, criado pelo MIT (*Massachusetts Institute of Technology*) e *Hewlett-Packard*. O MIT e o *Hewlett-Packard* trabalharam durante dois anos para o desenvolvimento da plataforma de repositório digital DSpace [TANSLEY, Robert, et al.].

A plataforma DSpace é voltado para a área acadêmica, com fins de servir como base para o desenvolvimento futuro de abordar a preservação a longo prazo dos arquivos e problemas de acesso [TANSLEY, Robert, et al.].

Quando um arquivo é registrado no repositório DSpace, são necessárias algumas informações do arquivo para ser preenchido no repositório, por se tratar de uma plataforma de organização de arquivos digitais é necessário criar um rótulo de informação do arquivo registrado.

As informações que são registradas no banco de dados do repositório DSpace, são carregadas para as páginas do repositório em formato de metadados.

Para integrar os motores robóticos inteligentes desenvolvidos na linguagem de marcação AIML, será desenvolvido nesse trabalho novos elementos na linguagem AIML para realizarem buscas semânticas em cima dos metadados. Os novos elementos integrado com o software analisador, realizaram um tratamento nas informações contidas nos metadados para geração de novos conhecimentos para os motores robóticos no formato de perguntas e respostas.

Uma busca de informações em metadados para um específico contexto, é necessário a utilização dos conceitos de um domínio da Ontologia, para assim, obter um resultado mais preciso na realização de uma busca e compreender de forma clara as informações que estão sendo buscadas no repositório DSpace.

### **3. OBJETIVOS**

O principal objetivo neste trabalho é abranger e potencializar a linguagem de marcação AIML com um software analisador de metadados, o mesmo realizará requisições aos repositórios digitais da plataforma DSpace, que com os metadados de resposta vindo do repositório será analisado e tratado pelo software analisador e, assim, gerando nova base de conhecimento em forma de pergunta e resposta para os motores robóticos inteligentes.

Os motores robóticos inteligentes desenvolvidos neste trabalho, serão integrados ao repositório da plataforma DSpace, para auxiliar nas buscas de arquivos digitais toda vez que um usuário acessar o repositório e realizar uma busca.

### **4. METODOLOGIA**

A metodologia utilizada para o desenvolvimento do trabalho foi dividida em etapas:

- O levantamento bibliográfico e estudo do AIML e XML.
- Pesquisa de correlatos.
- Definição do modelo de integração.
- Implementação de protótipo.
- Teste e validação.
- Comparação com correlatos.

Para que o software analisador possa interpretar as informações contidas nos metadados, é necessário o uso de algumas regras de Processamento da Linguagem Natural.

### **5. DESENVOLVIMENTO**

O desenvolvimento do software analisador de metadados que aplica uma análise e tratamento de informações contidas na resposta da requisição realizada ao repositório da plataforma DSpace, foi iniciado o estudo bibliográfico da linguagem de marcação AIML e do XML.

É necessário ter o conhecimento de como é criado um arquivo de conhecimento da linguagem AIML, para desenvolver um motor robótico inteligente da linguagem AIML.

Com o estudo bibliográfico das linguagens AIML e XML concluídas, foram levantados conhecimentos bibliográficos do repositório digital da plataforma DSpace, para assim, poder ser implantando no software analisador de metadados uma função que realiza requisições aos servidores dos repositórios. Todas requisições são realizadas por meio do serviço de HTTP Cliente/Servidor, que recebe como resposta do servidor do repositório um conjunto de metadados em XML.

Para o desenvolvimento do software analisador foi utilizado o repositório aberto do Centro Universitário Eurípides de Marília – UNIVEM [Repositório Institucional UNIVEM, 2012].

O software analisador está integrado a um motor robótico desenvolvido na linguagem AIML e quando um usuário acessa o repositório para busca de um arquivo digital, o mesmo poderá realizar perguntas para o motor robótico descrevendo qual o assunto, o autor, a data de publicação e outros dados do arquivo que está buscando.

O motor robótico acionará no software analisador uma função que dispara uma requisição ao servidor do repositório por meio de uma url HTTP, que contém as informações passadas pelo usuário. A resposta que é recebida do servidor é estruturada no formato da linguagem XML, contendo metadados que são específicos do repositório DSpace.

O software analisador tem pré-definido alguns elementos de metadados, para que possa analisar e reconhecer todo o arquivo XML vindo de resposta da requisição feita ao repositório Aberto do UNIVEM.

O software reconhecerá aonde estão as informações que procura no arquivo XML para assim já gerar um novo arquivo de perguntas e respostas estruturados na linguagem AIML. Assim como o software possui pré-definido os elementos XML, também tem pré-definido alguns elementos padrões da linguagem AIML, que são necessários no desenvolvimento de um arquivo de perguntas e respostas da linguagem.

Na saída do software analisador é apresentado um novo arquivo AIML contendo novas perguntas e respostas referente ao arquivo digital que o usuário está buscando. Esse novo arquivo AIML é automaticamente utilizado pelo motor robótico



para que possa responder as perguntas sobre o assunto que o usuário está buscando no repositório.

## 6. RESULTADOS

Neste trabalho foi desenvolvido um software analisador de metadados que analisa e gerar novos conhecimentos em formato de perguntas e respostas, para assim, os motores robóticos inteligentes possam fazer o uso dessas novas perguntas e respostas referente aos metadados dos arquivos que estão contidos nos repositórios da plataforma DSpace.

O software já realiza requisição aos servidores do repositório, recebe as informações na estrutura XML vindo como resposta dessas requisições, faz uma análise e tratamento nos metadados contidos nas informações XML, para assim, gerando um novo arquivo AIML com as novas perguntas e respostas. Na figura 1, mostra uma resposta vinda de uma requisição do repositório do institucional aberto do UNIVEM e na figura 2, mostra um arquivo AIML gerado pelo software analisador de metadados após ser analisado toda as informações do arquivo XML.

```
<?xml version="1.0" encoding="UTF-8" ?>
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
  http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2013-08-30T18:44:36Z</responseDate>
  <request verb="ListRecords" from="2012-11-22T00:00:01Z"
  metadataPrefix="oai_dc" until="2012-11-23T08:00:01Z">http://aberto.univem.edu.br/oai/request</request>
  <ListRecords>
  <record>
  <header>
  <identifier>oai:aberto.univem.edu.br:11077/804</identifier>
  <timestamp>2012-11-22T05:00:06Z</timestamp>
  <setSpec>hdl_11077_180</setSpec>
  </header>
  <metadata>
  <oai_dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/" xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
  http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
  <dc:title>Interpretador AIML Alimentado com TAGS HTML5 - Manual de Comandos do AIML</dc:title>
  <dc:creator>Macedo, Rafael Luiz de</dc:creator>
  <dc:description>O AIML (Artificial Intelligence Markup Language, Linguagem de Marcação da Inteligência Artificial) é um conjunto de
  tags XML (eXtensible Markup Language, Linguagem de Marcação Extensível) capaz de representar e relacionar expressões em linguagem natural
  permitindo a criação de motores capazes de manter um diálogo simples em linguagem natural. </dc:description>
  <dc:date>2012-11-21T11:56:02Z</dc:date>
  <dc:date>2012-11-21T11:56:02Z</dc:date>
  <dc:date>2012-11-21</dc:date>
  <dc:type>Dissertação</dc:type>
  <dc:identifier>http://hdl.handle.net/11077/804</dc:identifier>
  <dc:language>pt_BR</dc:language>
  </oai_dc:dc>
  </metadata>
  </record>
  </ListRecords>
</OAI-PMH>
```

Figura 1. Arquivo XML recebido da requisição realizada ao servidor do repositório digital.

```

<?xml version='1.0' encoding='ISO-8859-1'?>
<aiml version='1.0.1'
  xmlns='http://alicebot.org/2001/AIML-1.0.1'
  xmlns:html='http://www.w3.org/1999/xhtml'
  xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'
  xsi:schemaLocation='http://alicebot.org/2001/AIML-1.0.1 http://aitools.org/aiml/schema/AIML.xsd'>
  <category>
    <pattern>Qual artigo você tem *</pattern>
    <template>Interpretador AIML Alimentado com TAGS HTML5 -
      Manual de Comandos do AIML</template>
  </category>
  <category>
    <pattern>Quem * o autor *</pattern>
    <template>Macedo, Rafael Luiz de</template>
  </category>
</aiml>

```

Figura 2. Arquivo AIML com as novas perguntas e respostas geradas pelo Software Analisador de Metadados.

Com o novo arquivo que foi gerado pelo software analisador, os motores robóticos inteligentes automaticamente fazem o uso desse arquivos, para que possam realizar perguntas e respostas com o usuário que está buscando os arquivos digitais que são relacionados ao contexto dessa nova base de conhecimento.

## 7. CONSIDERAÇÕES FINAIS

A linguagem de marcação AIML possui uma limitação na criação de motores robóticos de conversação, fazendo com que seja limitado o número de perguntas que possam ser feitas a esses motores robóticos. Devido a essa limitação, os motores robóticos não conseguem manter um diálogo simples por muito tempo com um humano, pois as perguntas que são realizadas fora do contexto de conhecimento em que esses motores robóticos foram desenvolvidos, não conseguem ser respondidas.

Para tal problema, propõe-se a extensão da AIML com a incorporação de um software analisador de metadados que, analisará e aplicará um tratamento, nas informações contidas nos metadados encontrados nos repositórios digitais da plataforma Dspace. Transformando essas informações encontradas em novas perguntas e respostas para abranger o conhecimento da linguagem de marcação AIML e aumentar o número de perguntas que possam ser realizadas para esses motores robóticos inteligentes.

## 8. FONTES CONSULTADAS

ALICE AI Foundation “AIML: The Artificial Intelligence Markup Language”  
<http://www.alicebot.org/aiml.html>, Agosto 2013.

ALVES, Rachel C. V. (2005) “Web Semântica: Uma Análise focada no uso de metadados”, Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências – Universidade Estadual Paulista – UNESP, Marília/SP.

BUSH, Noel. “ProgramD” <http://aitools.org/Programd>, Agosto 2013.

CONCEPT, “Converse com o Robô Ed!”  
<http://www.ed.conpet.gov.br/br/converse.php>, Agosto 2013.

MENEZES, Paulo, B. (1998), “Linguagens Formais e Autômatos” Segunda Edição. Porto Alegre.

RICH, Elaine e KNIGHT, Kevin. (1993) “Inteligência Artificial” – Segunda Edição. São Paulo/SP.

RUSSELL, Stuart e NORVIG, Peter. (2004) “Inteligência Artificial” – Tradução da Segunda Edição. Rio de Janeiro/RJ.

School of Mathematics and Statistics University of St. Andres. “História do Matemático Alan Turing”, <http://apprendre-math.info/portugal/historyDetail.htm?id=Turing>, Escócia, Agosto 2013.

TANSLEY, Robert, et al.. “The DSpace Institucional Digital Repository System: Current Functionality”, Joint Conference on Digital Libraries, pág. 87-97, EUA, 2003

UNIVEM, Centro Universitário Eurípides de Marília, 2012 - “Repositório Institucional - UNIVEM Aberto”, <http://aberto.univem.edu.br/>, Agosto 2013.

WALLACE, Dr. Richard S. (2003) “The Element of AIML Style”, © ALICE A. I. Foundation, Inc., Outubro.

WALLACE, Dr. Richard; TAYLOR, Anthony. “AIML Reference Manual”, <http://www.alicebot.org/documentation/aiml-reference.html>, Agosto 2013.