

## ***Chatterbot Tira-Dúvidas do Curso de Informática do IFMS***

**Giulia Faustini Milan<sup>1</sup>, Bárbara Santos Munhão<sup>1</sup>, Antonio Miguel Faustini Zarth<sup>1</sup>**

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul (IFMS)  
Caixa Postal 79.113-000 – Campo Grande – MS – Brasil

{giulia-milan, barbara.munhao}@hotmail.com, antonio.zarth@ifms.edu.br

**Abstract.** *This paper presents the creation of a chatterbot with principles in Artificial Intelligence and Natural Language Processing, created by students from IFMS. It was created from the observation of the lack of information that the population of Mato Grosso do Sul presents relative of the federal institution and the Technical Course Mid-Level Integrated Informatics. Thus, the goal is to create an alternative tool for the dissemination of the course and institution, whose anthropomorphic aspects make it more attractive to prospective students and those interested in the Informatics Course.*

**Resumo.** *Este artigo apresenta a criação de um chatterbot com princípios em Inteligência Artificial e Processamento de Linguagem Natural, produzido por estudantes do IFMS. Sua elaboração se deve à observação da falta de informações que a população do Mato Grosso do Sul apresenta em relação à instituição federal e ao Curso Técnico de Nível Médio Integrado em Informática. Desta forma, o objetivo é a criação de uma ferramenta alternativa para a divulgação do curso e da instituição, cujos aspectos antropomórficos a tornam mais atrativa para aos interessados e futuros alunos do técnico em informática.*

### **1. Introdução**

Segundo [Russell and Norvig 2004], *chatterbots*, do inglês *chat*, que significa conversa e *bot*, que significa robô (robô de conversação), são programas que simulam conversas com humanos, através de padrões semânticos ou sintáticos de uma base de dados pré-definida.

Devido à falta de informação que fora percebido sobre o IFMS (Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul) e sobre o Curso Técnico de Nível Médio Integrado em Informática na população local, dada sua recente implantação no Estado, foi proposta a criação de um *chatterbot*, para que o público pudesse ter informações sobre a instituição e sobre o curso, de uma maneira mais simples e rápida: através da internet e em linguagem natural.

### **2. Metodologia**

Para a criação de uma base de dados inicial com as dúvidas gerais da população, foram realizadas entrevistas informais com alunos, professores e pais de alunos da instituição. Em sequência, a base de dados foi transformada e ampliada em códigos da linguagem de formatação AIML (*Artificial Intelligence Markup Language*) [AIM ], que faz possível criar ou customizar um bot em linguagem natural. Para isso foi utilizado o Interpretador AIML *ProgramD*, uma plataforma *bot opensource* amplamente utilizada e gratuita.

O *chatbot*, denominado Alana em homenagem ao matemático Alan Turing, foi disponibilizado em um blog (<http://alana.blog.br>). Neste, estão disponíveis textos relacionados ao projeto além do próprio *bot*, possibilitando a visita de todos que têm acesso à internet.

À medida que os códigos foram sendo criados e aperfeiçoados, selecionou-se grupos de testes para conversar de forma espontânea com o *bot*. Os grupos foram compostos principalmente por adolescentes com idade média de 14 anos, os quais são o foco do curso. Essas conversas foram gravadas em um *log*, possibilitando a análise posterior. Em sequência, os códigos foram corrigidos e ampliados. Então, outros grupos de testes eram escolhidos, gerando uma retroalimentação.

### 3. Discussão dos Resultados e Conclusões

Com as análises do log, pôde-se analisar o uso de siglas, “vícios de linguagem” e erros ortográficos não reconhecidos pelo bot. Para isso foram utilizadas estratégias semânticas e sintáticas. Estas facilitaram a expansão do entendimento nas conversas, pois foram englobados diversos sinônimos e erros de ortografia nos códigos.

No total foram efetuados quatro testes, sendo estes realizados à medida que os códigos fossem corrigidos e ampliados. Fora observado que a cada novo teste, o *chatbot* respondia uma maior quantidade de perguntas de forma mais clara e objetiva. No último teste foram analisadas 90 perguntas, nas quais 71 foram respondidas de forma coerente. Resultando em aproximadamente 80% das questões respondidas corretamente. Finalizados os quatro testes, realizou-se a contagem das linhas de AIML escritas, abrangendo mais de 15.000.

Com as análises mais recentes, constatou-se que a maioria das perguntas feitas ao *bot* não são relacionadas com seus objetivos. Isso ocorre pois o blog foi divulgado apenas para os estudantes e servidores da instituição. Estes já tem conhecimento das informações dadas pela Alana, não se interessando por perguntas que se enquadram no objetivo do robô.

A principal contribuição desse trabalho é proporcionar uma forma alternativa de comunicação com a sociedade e o público alvo, explorando aspectos antropomórficos que tornam as informações dispostas mais atraentes que as em formato de editais. Sendo assim, este trabalho poderá servir de guia para implementações semelhantes.

O objetivo de produzir uma ferramenta acessível para cativar futuros estudantes da instituição foi cumprido, pois ao ser criado por estudantes do Curso de Informática gera interesse indiretamente. Já o propósito de ser utilizado pela população em geral refletindo no aumento da procura pelo IFMS, só será observado de médio a longo prazo.

### Referências

Artificial Intelligence Markup Language. <http://alicebot.org/aiml.html>.  
Site acessado em 20 de agosto de 2012.

Russell, S. and Norvig, P. (2004). Inteligência Artificial. *Editora Campus*.