

Proposta de arquitetura para a integração de sistemas de informação baseados na web com a tecnologia *wiki*¹

Bruno Franciscan Mazzotti²
Anderson Carlos Bueno dos Santos²
Tiago Augusto Pimenta²
Carlos Alberto Alves Meira³
Marcos Cezar Visoli³
Isaque Vacari³

Sistemas de Informação (SI) baseados na *web* proveem serviços para usuários finais publicando dados em formato hipertexto. A arquitetura por meio da qual estes SI's são implementados baseia-se em um modelo cliente/servidor no qual o *back-end* é, quase sempre, um Sistema Gerenciador de Banco de Dados Relacional (SGBDR) (ELMASRI; NAVATHE, 2004). A estrutura rígida do modelo relacional utilizado traz como consequência imediata um baixo grau de flexibilidade para as informações nele contidas, bem como para as possíveis associações entre estas informações.

Contrapondo-se aos SGBDR's, no que diz respeito à flexibilidade da estrutura organizacional dos dados, encontram-se os *wikis*. Um *wiki* é um sistema *web* de caráter colaborativo no qual os usuários elaboram e sugerem a estrutura de seu próprio conteúdo (PRIMO; RECUERO,

¹ *Agradecimentos ao CNPq pelo auxílio financeiro. Projeto Banco de Dados sobre Pragas Quarentenárias Associadas às Culturas Relevantes para o Agronegócio Brasileiro e a Defesa Fitossanitária no País (número 558949/2009-6, chamada 2 do edital 32/2009).*

² *Bolsista do CNPq; brunofm@cnptia.embrapa.br; andersoncbs@cnptia.embrapa.br; tiagoap@cnptia.embrapa.br*

³ *Embrapa Informática Agropecuária; carlos@cnptia.embrapa.br; visoli@cnptia.embrapa.br; isaque@cnptia.embrapa.br*

2003). A forma de edição de conteúdo inerente ao conceito de *wiki* se adequa muito bem a dados de caráter descritivo livre, porém limita a capacidade de cruzamento automático das informações contidas em suas páginas.

Este trabalho propõe uma arquitetura para SI's baseados na *web* que visa integrar o conceito de conteúdo descritivo e colaborativo proveniente dos *wikis* à tradicional visão de conteúdo indexado dos SGBDR's. A metodologia de integração vale-se de uma plataforma de *wiki* já implementada e a utiliza como um módulo interno ao SI *web* tradicional.

Encapsular o *wiki* proporciona a visão de um sistema computacional único e o conteúdo, do ponto de vista do usuário, deve ser fornecido ao sistema e posteriormente recuperado de uma maneira uniforme. Em nenhum momento a situação do *wiki* como um módulo interno de um sistema maior deve transparecer ao mundo externo. Cabe ao próprio sistema fragmentar internamente o conteúdo de entrada e direcionar dados descritivos para o *wiki* e dados indexados para o BD, como ilustrado pela Figura 1.

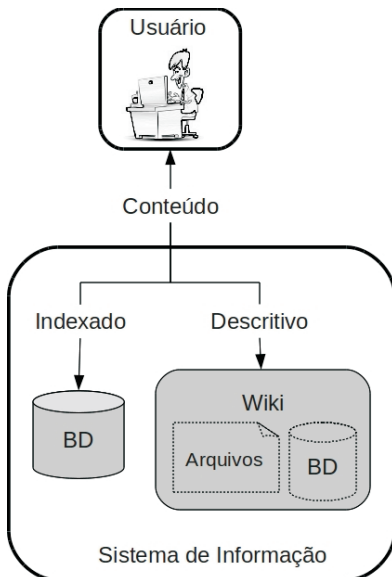


Figura 1. Fluxo de conteúdo no sistema integrado.

Contudo, dados armazenados com o propósito de indexação são de extrema importância para agregar valor semântico às descrições atreladas. O sistema, quando o usuário requisita conteúdo descritivo, deve recuperar também o conteúdo indexado associado e elaborar uma visão unificada dos dados. A metodologia mais indicada para tanto consiste em interceptar o *datapath* de processamento interno do wiki por meio de *hooks* ou *callbacks* (FERG, 2006) e introduzir os dados indexados conforme for conveniente. A Figura 2 resume a sequência de eventos que descreve o referido procedimento, denominado no escopo do projeto como injeção de *wikitexto* (JUNGHANS et al., 2008).

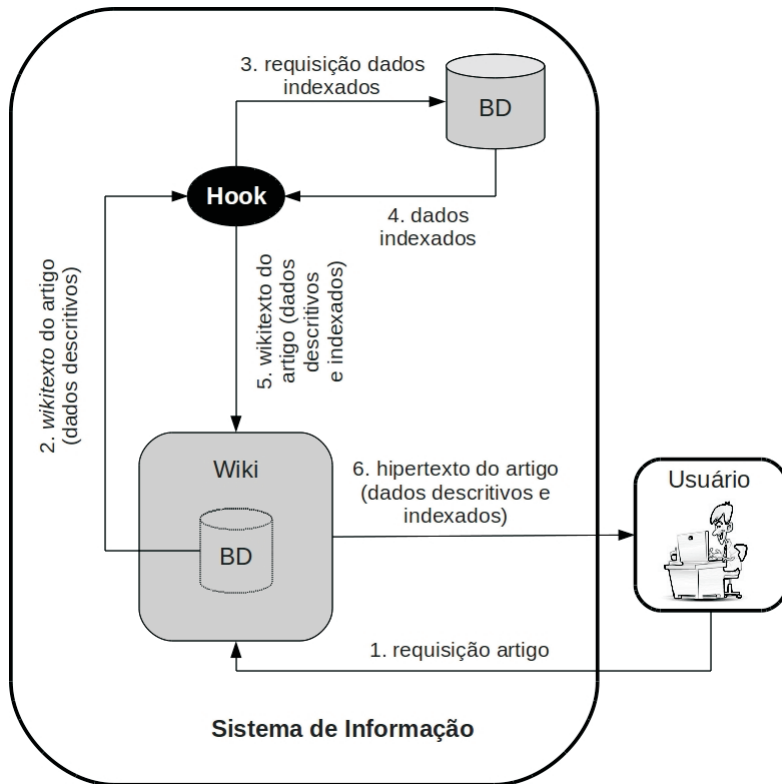


Figura 2. Procedimento de injeção de *wikitexto*.

Assim sendo, o BD que armazena o conteúdo indexado do SI abastece dois clientes que requisitam dados ativamente: o usuário (agente externo) e o próprio wiki (agente interno). As requisições do wiki são necessariamente disparadas pelas requisições do usuário e estas podem ou não desencadear pedidos do wiki para o BD. Se o usuário requisitar somente conteúdo indexado não é preciso recorrer ao conteúdo descritivo do wiki. Caso o usuário necessite de conteúdo descritivo, este será recuperado do wiki e incrementado com conteúdo indexado do BD para prover semântica consistente ao conjunto de dados. Um modelo de aplicação único é representado por duas visões diferentes: hipertexto para o usuário e *wikitexto* para o *wiki*.

Normalmente, um *wiki* oferece funcionalidades de agrupamento de artigos para facilitar os processos de busca e navegação. Em um *wiki* toda a categorização é realizada manualmente pelos usuários, estando portanto sujeita a julgamentos subjetivos individuais, e pode ser tediosa caso o volume de artigos seja considerável. Esta atividade de categorização relaciona-se intrinsecamente com os dados indexados e portanto pode ser automatizada por um componente sincronizador do SI. Este módulo seria responsável por editar as páginas de categoria do *wiki* para que estas reflitam a indexação do BD e consequentemente isentaria os usuários desta tarefa.

Optou-se por utilizar uma solução baseada em *software* livre como plataforma de desenvolvimento para a arquitetura apresentada. Um protótipo para prova de conceito da arquitetura está sendo construído nos laboratórios de Inteligência Computacional e de *Software* Livre. As atividades ligadas ao desenvolvimento deste protótipo se enquadram em um projeto de pesquisa que visa projetar um SI de catalogação de pragas agrícolas quarentenárias, o *BD-Pragas*.

Um BD relacional *PostgreSQL* foi modelado e encontra-se plenamente acoplado a um SI *web* implementado com tecnologias MVC Java. Também foi implementado, em PHP, um módulo que executa o procedimento de injeção de *wikitexto* na plataforma *MediaWiki*. Pretende-se futuramente promover a comunicação deste módulo com o BD, integrar o *MediaWiki* com o SI e também implementar o componente de sin-

cronização. Desta maneira, o protótipo de realização da arquitetura integrada estaria completo e o seu conceito poderia ser submetido a testes de aceitação sistemáticos com os usuários finais.

É possível concluir apenas que a arquitetura apresentada é promissora, uma vez que o trabalho ainda encontra-se em fase inicial. Cabe também ressaltar que após revisão da literatura sobre o tema não foi encontrado nenhum trabalho similar ao proposto.

Referências

ELMASRI, R.; NAVATHE, S. B. XML and Internet Databases. In: ELMASRI, R.; NAVATHE, S. B. **Fundamentals of Database Systems**. 4th ed. Boston: Pearson Education, 2004. p. 841-866.

FERG, S. **Event-driven Programming: Introduction, Tutorial, History**. 2006, p. 42. Disponível em: <<http://eventdrivenpgm.sourceforge.net/>>. Acesso em: 23 ago. 2010.

JUNGHANS, M.; RIEHLE, D.; GURRAM, R.; KAISER, M.; LOPES, M.; YALCINALP, U. A Grammar for Standardized Wiki Markup. In: INTERNATIONAL SYMPOSIUM ON WIKIS, 4., 2008, Porto, Portugal. **Proceedings...** New York: Association for Computing Machinery, 2008. Article n. 21.

PRIMO, A. F. T.; RECUERO, R. C. Hipertexto cooperativo: uma análise da escrita coletiva a partir dos blogs e da Wikipédia. **Revista FAMECOS: Mídia, Cultura e Tecnologia**, v. 1, n. 22, p. 54-65, 2003.