

Uma Ferramenta para Configuração do Processo de Aquisição de Conhecimento no Contexto de Análise de Domínio

Mônica Zopelari Roseti *Leonardo G.P. Murta* *Cláudia M. L. Werner*
{zopelari, murta, werner}@cos.ufrj.br

COPPE - Programa de Engenharia de Sistemas e Computação
Universidade Federal do Rio de Janeiro
Caixa Postal 68511
21945-970 Rio de Janeiro - RJ - Brasil

1. Objetivos

A ferramenta apresentada neste artigo tem como objetivo auxiliar o analista de domínio¹, na configuração de uma sistemática para a execução da aquisição de conhecimento² no contexto de análise de domínio³ (AD), levando em consideração os diferentes contextos de projetos de AD que podem ser realizados. Para isto, a ferramenta conta com uma base de informações sobre uma proposta de sistemática, contendo um processo, um conjunto de técnicas e diretrizes para seleção destas. A sistemática proposta parte do princípio que um processo deve disponibilizar um conjunto de técnicas para aquisição de conhecimento, cada uma com uma forma de elicitação e representação adequada para um determinado tipo de conhecimento do domínio a ser explorado (HARBISON-BRIGGS *et al.*, 1989). Tal princípio está melhor descrito na seção 4 deste artigo. A ferramenta possibilita, ainda, a incorporação de outras propostas de sistemática.

2. Plataformas

A ferramenta foi implementada na linguagem JAVA (NIEMEYER, PECK, 1997), devido aos benefícios advindos da característica de ser uma linguagem portátil entre várias plataformas.

3. Descrição da Funcionalidade

A ferramenta possui, na sua base de informações, uma proposta de processo para aquisição de conhecimento no contexto de AD, definido em termos de etapas e atividades, um conjunto de técnicas para aquisição de conhecimento no contexto de AD associadas às atividades do processo, e um conjunto de diretrizes para seleção de uma configuração apropriada do processo proposto ao contexto do projeto de AD em questão.

O processo proposto foi obtido combinando as atividades do processo genérico de Arango (ARANGO *et al.*, 1994) com a filosofia proveniente do processo de aquisição de conhecimento de SCOTT *et al.* (1991) (ver seção 4).

O conjunto de técnicas disponível é composto pelas seguintes técnicas (MIRADOR, 1976) (HARBISON-BRIGGS *et al.*, 1989) (SCOTT *et al.*, 1991) (GAUSE *et al.*, 1991) (JACOBSON, 1992) (AUGUST, 1993) (LEITE, 1993) (CHI *et al.*, 1994) (HAYES, 1994)

¹ Responsável pela realização de uma análise de domínio.

² O processo de aquisição de conhecimento envolve não apenas a identificação e a coleta da informação, o que seria o caso de elicitação, mas também a representação, a organização e o armazenamento da informação obtida, sempre considerando a natureza evolutiva e contínua deste processo.

³ É o processo de identificar e organizar o conhecimento sobre um domínio de aplicação.

(MCKEEN et al., 1994) (LEITE, 1995) (JACOBSON et al., 1997) (GRISS et al., 1998): Entrevista; Análise de Protocolo; Cenários; Ordenação Conceitual; LEL; Técnica de Delphi; Brainstorming; JAD; Psicometria; Pesquisa Estatística de Comportamento e Auto-Explicação.

As diretrizes disponíveis estão relacionadas às seguintes questões: ótica de análise de domínio adotada; objetivos da análise de domínio; tipo de conhecimento explorado; tempo de duração estimado; pessoas disponíveis; fontes de informações disponíveis; interdependência entre as técnicas.

A ferramenta, inicialmente, apresenta ao analista de domínio várias perguntas a respeito do contexto do projeto, no qual será realizada a análise de domínio, visando identificar o máximo possível de características do mesmo (figura 1). Estas perguntas são, na verdade, identificadores das diretrizes disponíveis.

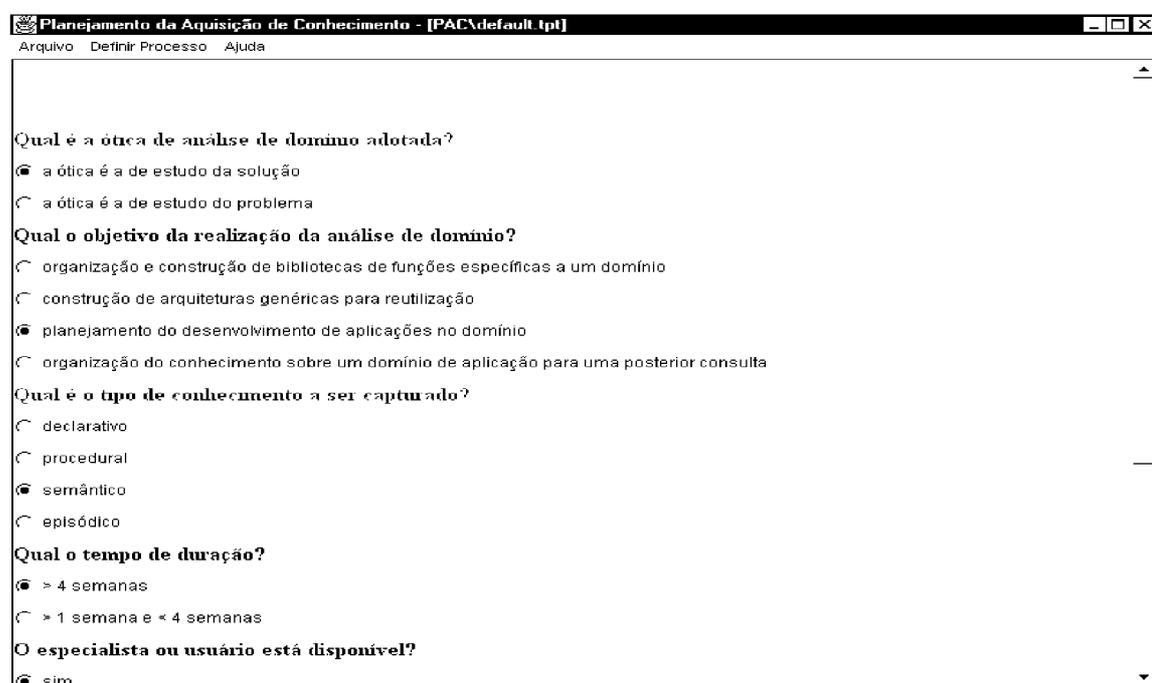


Figura 1 - Tela inicial

De posse das respostas fornecidas pelo analista de domínio, a ferramenta é capaz de configurar o processo. Esta configuração prevê que este processo é composto por um conjunto de etapas, cada etapa é composta por um conjunto de atividades, e que, para cada atividade, existe sempre uma ou mais técnicas relacionadas. A princípio, todas as técnicas, atividades e etapas, estão selecionadas. Cada resposta relacionada à uma determinada diretriz irá excluir uma ou mais técnicas e, possivelmente, uma ou mais atividades e etapas, configurando, desta forma, o processo.

A configuração resultante é apresentada através de uma visão genérica, que descreve apenas as etapas, ou através de uma visão detalhada (figura 2), que descreve, além das etapas, as atividades de cada uma destas com suas respectivas técnicas.

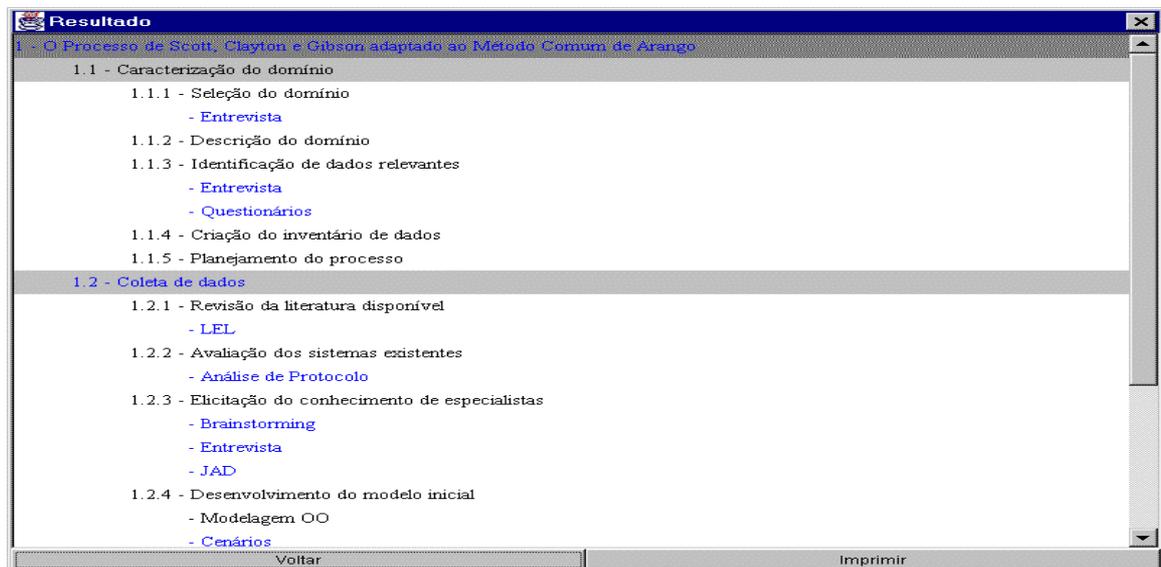


Figura 2 - Visão detalhada de uma configuração resultante

Além disto, a ferramenta apresenta as opções de edição de diretrizes (figura 3), de processos (incluindo etapas e atividades) e de técnicas. Tais opções permitem que a ferramenta incorpore outras propostas de sistemática, possibilitando incluir, modificar e deletar diretrizes, processos e técnicas, além da forma como as diretrizes editadas interferem na configuração do processo resultante.

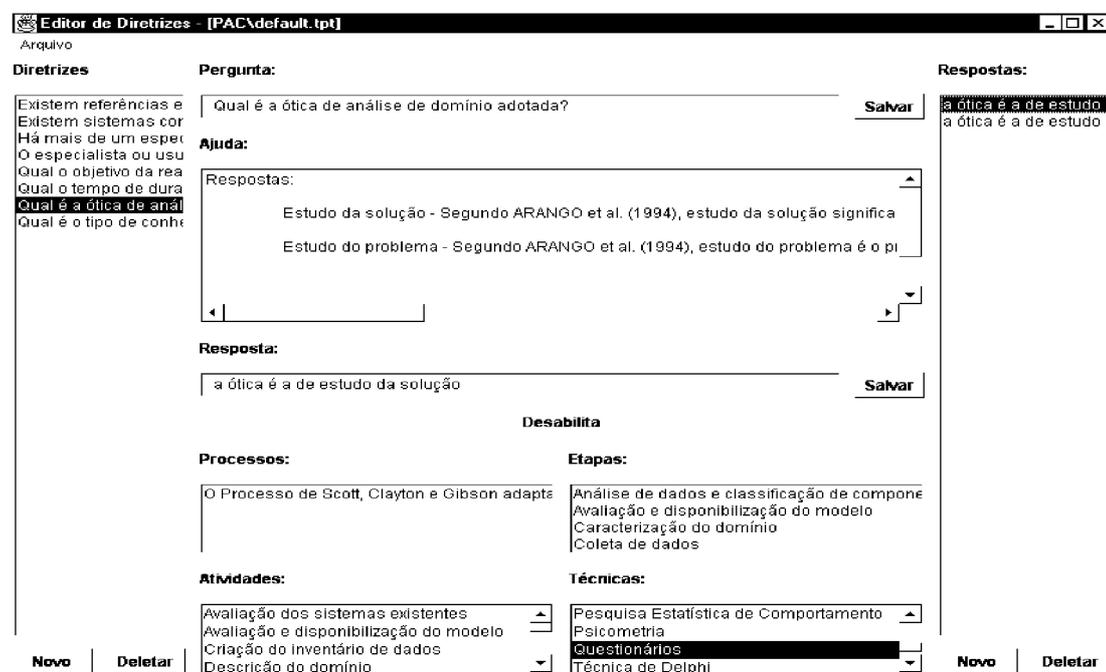


Figura 3 - Tela do Editor de Diretrizes

Ao incluirmos novas diretrizes, podemos caracterizar, de forma cada vez mais precisa, o contexto do projeto em questão. Ao incluirmos novos processo e técnicas, oferecemos ao analista de domínio, propostas de sistemática cada vez mais flexíveis e adaptáveis ao contexto do seu projeto.

Existe, ainda, uma ajuda sensível ao contexto da ferramenta, que informa sobre o processo (suas etapas e atividades), sobre as técnicas e sobre as diretrizes, conforme pode ser visto na figura 4.

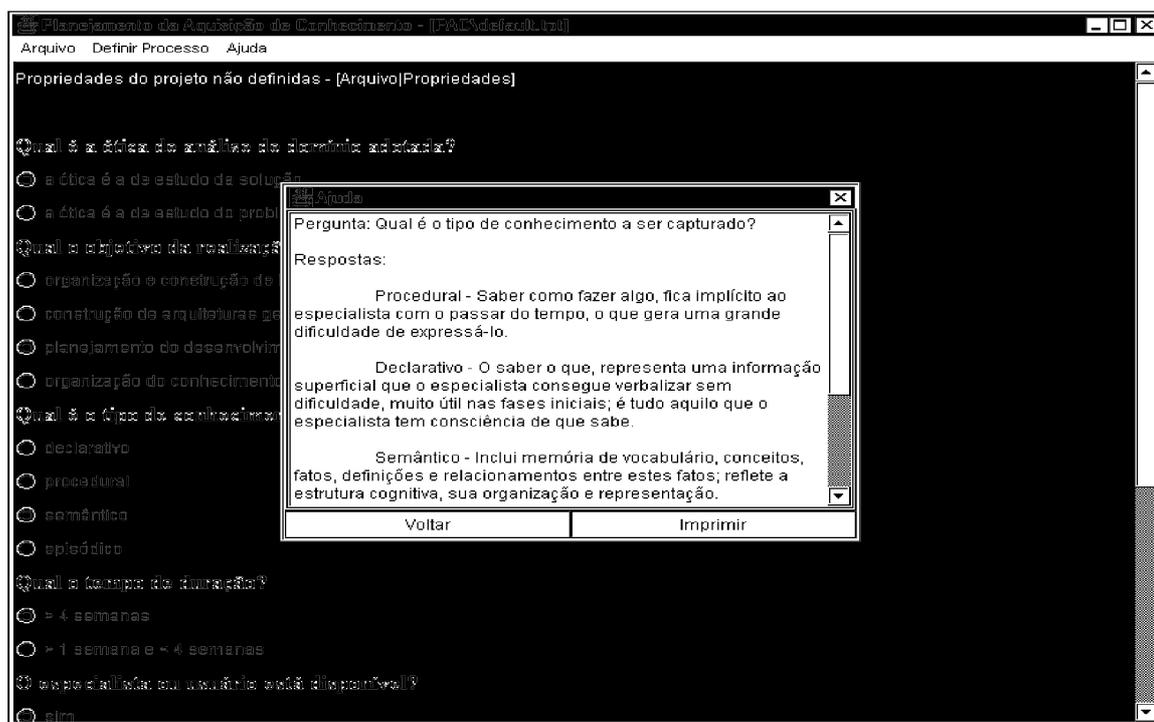


Figura 4 - Janela de Ajuda sobre Diretrizes

4. Embasamento Teórico

Ao realizar estudos sobre várias abordagens para realização de AD, verifica-se a existência de uma fase de aquisição de conhecimento, com a sua importância devidamente reconhecida, porém, sem uma definição explícita da sua sistemática dentro do contexto de uma AD.

Partindo do princípio que, ao realizarmos uma AD, estamos interessados em explorar o máximo possível do conhecimento existente sobre o domínio de interesse, iremos nos deparar com vários tipos de conhecimento (procedural, lógico, etc.), que necessitarão de abstrações diferentes para representá-los. Assume-se, portanto, que para cada tipo de conhecimento deve haver uma representação mais adequada, capaz de registrar da melhor forma possível a sua semântica. Assim, pretende-se fornecer um processo que disponibilize um conjunto de técnicas para aquisição de conhecimento, cada uma com uma forma de elicitación e representação adequada para um determinado tipo de conhecimento do domínio a ser explorado.

Para disponibilizar tal sistemática foi necessário explicitar um processo para a execução da aquisição de conhecimento em AD, bem definido em termos de etapas, atividades e técnicas associadas à estas atividades. A proposta de uma sistemática para aquisição de conhecimento no contexto da AD, baseou-se nas abordagens existentes nas áreas de sistemas convencionais e sistemas baseados no conhecimento (SBCs), sem negligenciar as diferenças destas áreas, sugerindo, então, adequações nos processos e técnicas estudados para o contexto de AD.

Quanto ao processo, em termos de etapas e atividades, foi proposto combinar as

atividades do processo genérico de Arango (ARANGO *et al.*, 1994) com a filosofia proveniente do processo de aquisição de conhecimento de SCOTT *et al.* (1991), devidamente adequado ao contexto de AD. Este último sugere identificar uma fase inicial, na qual o esforço deve ser bem direcionado para entender o máximo possível sobre o domínio de interesse, usando o mínimo possível do tempo do especialista, seguida de uma fase detalhada. De posse do conhecimento previamente capturado na fase inicial, parte-se para uma investigação mais aprofundada para definição do modelo, preocupando-se, constantemente, em revisar e refinar o conhecimento adquirido a cada etapa do processo.

Para resolver a questão das técnicas, partiu-se de técnicas já existentes em outras áreas (SBCs, sistemas convencionais e psicologia), sugerindo adequações para o contexto de AD, de forma a dispor de um conjunto mínimo de técnicas capazes de explorar, de forma satisfatória, os diferentes tipos de conhecimento e diferentes fontes de informações do domínio de interesse.

De posse deste conjunto de técnicas, o analista de domínio deve selecionar quais técnicas ele utilizará num determinado projeto. Para ajudar na seleção das técnicas, baseado no trabalho de MAINDEN *et al.* (1996), sugeriu-se diretrizes de escolha para este conjunto de técnicas.

5. Processo de Desenvolvimento

Inicialmente, elaborou-se uma especificação de requisitos, a partir da qual o projeto da ferramenta foi desenvolvido, utilizando-se a Linguagem de Modelagem Unificada (UML) (FOWLER, SCOTT, 1997) (figura 5).

Foram identificadas as seguintes classes de objetos: *Processo*, *Etapa*, *Atividade*, *Técnica*, *Diretriz* e *Resposta*. Cada processo é composto de várias etapas que, por sua vez, são compostas de diversas atividades. A ferramenta foi projetada para que, a princípio, todas as técnicas e, por consequência, atividades e etapas, fossem selecionadas. Cada resposta relacionada à uma determinada diretriz irá excluir uma ou mais técnicas e, possivelmente, uma ou mais atividades e etapas, configurando, desta forma, o processo.

As classes *Processo*, *Etapa*, *Atividade*, *Técnica* e *Diretriz* apresentam o atributo *info*, cujo objetivo é guardar informações a respeito dos objetos, tais como: descrição, referências bibliográficas, etc. O atributo *selecionado* das classes *Processo*, *Etapa*, *Atividade* e *Técnica* informa se o objeto em questão está selecionado. Conforme explicado anteriormente, todos estão inicialmente selecionados, sendo os objetos da classe *Resposta* que determinam esta seleção, posteriormente através do método *seleciona()*. Isto permite que a ferramenta configure o processo na medida que as perguntas são respondidas.

6. Experiência de Uso e Resultados

A ferramenta descrita neste artigo foi utilizada no contexto de um projeto de pesquisa envolvendo uma universidade e um órgão legislativo municipal. Por um lado (universidade), tinha-se o objetivo de analisar as dificuldades práticas da implantação de um processo adaptado para a aquisição de conhecimento no contexto de AD e por outro (órgão legislativo) a verificação da possibilidade de estabelecer projetos cooperados com outros órgãos legislativos, através de um modelo de domínio passível de ser compartilhado.

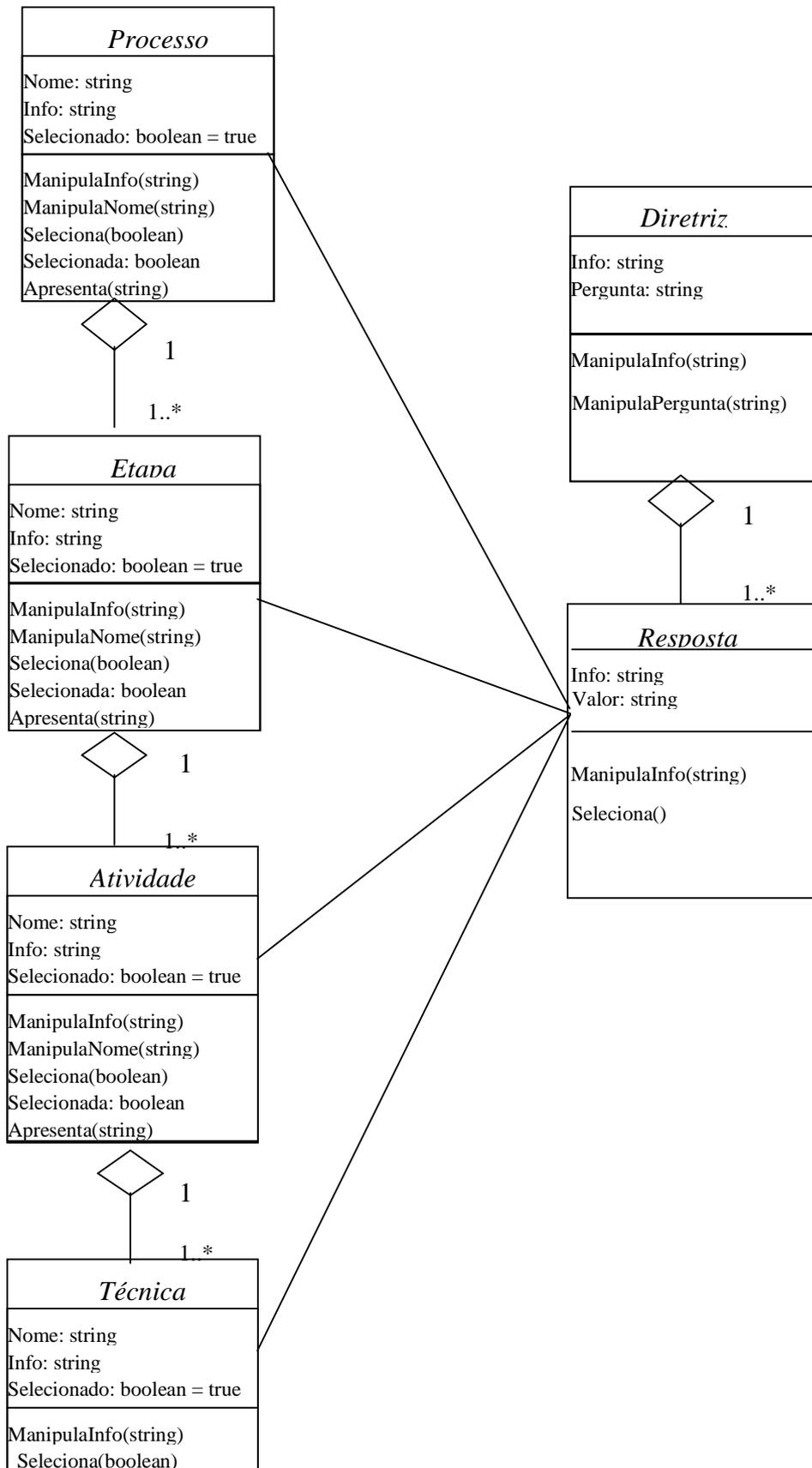


Figura 5 - Modelo da Ferramenta

Para a realização deste projeto, foi selecionado o domínio de tramitação de projetos, devido ao fato de haver um grande interesse por parte dos órgãos legislativos em informatizar esta atividade e de todos os órgãos legislativos executarem esta mesma atividade, existindo, como consequência, uma experiência muito grande neste domínio. O projeto envolveu o órgão legislativo municipal e um órgão legislativo estadual. A ferramenta foi utilizada para configurar o processo proposto, de forma a atender às características específicas deste projeto.

Para efeito de avaliação dos benefícios obtidos a partir desta utilização, comparou-se os resultados deste estudo com os obtidos em um estudo de caso previamente realizado no domínio de química de polímeros, onde não havia uma sistemática específica para realização de aquisição de conhecimento. Constatou-se que, através da utilização de um processo sistemático para a aquisição de conhecimento no contexto de análise de domínio, disponibilizado pela ferramenta, bem definido em termos de etapas, atividades e técnicas, configurado para o projeto em questão, é possível superar grande parte das dificuldades ressaltadas no decorrer do estudo de caso realizado anteriormente, dentre elas:

- Dificuldade de organização do pensamento na execução de suas atividades de aquisição de conhecimento em AD;
- Melhor direcionamento de esforços durante a execução desta atividade;
- Tendência de inércia inicial frente à um domínio desconhecido e/ou complexo;
- Não fluência no vocabulário do domínio;
- Dificuldade em extrair as informações das fontes detectadas e torná-las disponíveis em uma forma adequada;
- Dificuldades para discernir as informações mais relevantes do domínio;
- Desconhecimento sobre a técnica mais adequada num determinado momento;
- Inexistência de uma sequência organizada de passos a serem executados;
- Pouca disponibilidade dos especialistas;
- Não familiaridade dos especialistas com as representações apresentadas;
- Pouca capacidade de inferência por parte do analista de domínio no seu aprendizado sobre o domínio.

7. Conclusão

A ferramenta apresentada permite a disponibilização de uma sistemática para a execução da atividade de aquisição de conhecimento em AD, através de um processo que conta com um conjunto de técnicas, cada uma com uma forma de elicitação e representação adequada para um determinado tipo de conhecimento a ser explorado. Assume-se que, na realização de uma AD, estamos interessados em explorar o máximo possível do conhecimento existente sobre o domínio de interesse e, como consequência, iremos nos deparar com vários tipos de conhecimento. Esta sistemática é, ainda, configurada de acordo com diretrizes que identificam o contexto do projeto de AD a ser realizado.

Para avaliação dos benefícios da proposta e sua disponibilização através da ferramenta foram realizados dois estudos de caso, descritos na seção 6 deste artigo, nos quais pudemos constatar que é viável e benéfico disponibilizar uma ferramenta que armazene informações sobre uma sistemática e que permita configurá-la de acordo com o contexto do projeto em questão. Como a ferramenta permite a incorporação de novos processos, técnicas e diretrizes de seleção, é possível evoluir o auxílio por ela proposto, conforme se evolua com a sistemática proposta.

8. Bibliografia

- ARANGO, G., BASILI, V.R., BORSTLER, J., *et al.*, 1994, *Software Reusability*, 1ª ed. Great Britain, Ellis Horwood.
- AUGUST, J.H., 1993, *JAD - Joint Application Design*, 1ª. ed., Makron Books.
- CHI, M.T.H., LEEUW, N., CHIU, M.H., *et al.*, 1994, "Eliciting Self-Explanations Improves Understanding", *Cognitive Science*, vol. 18, no. 3 (Sept), pp. 439-475.
- FOWLER, M., SCOTT, K., 1997 *UML Distilled - Applying the Standard Object Modeling Language*, 1ª ed. Massachusetts, Addison-Wesley.
- GAUSE, D. C., WEINBERG, G.M., 1991, *Explorando Requerimentos de Sistema*, 1ª. ed., Rio de Janeiro, Makron Books.
- GRISS, M.L., FAVARO, J., d'ALESSANDRO, M., 1998, "Integrating Feature Modeling with the RSEB", in: *Proceedings of 5th International Conference on Software Reuse*, pp. 76-85, Victoria, Canadá, Jun.
- HARBISON-BRIGGS K., MCGRAW, K. L., 1989, *Knowledge Acquisition: Principles and Guidelines*, Prentice Hall.
- HAYES, B. E., 1994, "How to Measure Empowerment", *Quality Progress*, Feb., pp. 41-46.
- JACOBSON, I., 1992, *Object-Oriented Software Engineering: A Use Case Driven Approach*, 1ª. ed., Addison-Wesley.
- LEITE, J.C.S.P., 1993, *Eliciting Requirements Using a Natural Language Based Approach: The Case of the Meeting Scheduler Problem*, Monografias em Ciência da Computação, Departamento de informática, PUC-Rio, Rio de Janeiro, RJ.
- LEITE, J.C.S.P., 1995, "Recovering Business Rules from Structured Analysis Specifications", in: *Proceedings of the Second Working Conference on Reverse Engineering*, pp. 13-21, Ontario, Canada, Jul.
- MAIDEN, N.A.M., RUGG, G., 1996, "ACRE: Selecting Methods for Requirements Acquisition", *Software Engineering Journal*, May, pp. 183-192.
- MCKEEN, J.D., GUIMARAES, T., WETHERBE, J.C., 1994, "The Relationship Between User Participation and User Satisfaction: A Investigation of Four Contingency Factors", *MIS Quarterly*, vol. 3, Dec., pp. 427-451.
- MIRADOR, 1976, *Enciclopédia Mirador Internacional*, Editora Enciclopédia Britânica do Brasil Publicações Ltda., vol 17, pgs 9438-9443, SP - Brasil.
- NIEMEYER, P., PECK, J., 1997, *Exploring Java*, 2ª. edição, O´Reilly.
- PFLEEGER, S. L., 1995, "Experimental Design and Analysis in Software Engineering: How to Set Up an Experiment", *Software Engineering Notes*, vol. 20, nº1(Jan), pp. 22-26.
- SCOTT, A.C., CLAYTON, J. E., GIBSON, E., 1991, *A Practical Guide to Knowledge Acquisition*, 1ª. ed., Addison-Wesley.