

SISTEMAS ESPECIALISTAS EM BIBLIOTECAS: DESENVOLVIMENTO DE UM PROTÓTIPO PARA CATALOGAÇÃO*

Sônia Elisa Caregnato**

Nigel Ford***

O artigo discute, sucintamente, o que são sistemas especialistas, que impacto eles estão tendo nas áreas da biblioteconomia e ciência da informação e aponta alguns setores, dentro dessas disciplinas, onde eles estão sendo aplicados. Descreve, mais detalhadamente, o protótipo de um sistema especialista desenvolvido no Department of Information Studies da University of Sheffield, Inglaterra, como parte de um trabalho de mestrado, cujo objetivo é o de auxiliar na aplicação das regras contidas no capítulo 21 do Código de catalogação anglo-americano, segunda edição.

1 INTRODUÇÃO

Este artigo documenta os resultados de um trabalho de pesquisa, desenvolvido no Department of Information Studies da University of Sheffield, como parte dos requisitos necessários para a obtenção do título Master of Science in Information Management¹. O objetivo da pesquisa era o de investigar e explorar as possibilidades de aplicação de sistemas especialistas a bibliotecas e/ou serviços de infor-

* Este artigo é baseado no trabalho de mestrado da autora, financiado pelo CNPq durante o período de outubro de 1991 a setembro de 1992.

.. Bolsista do CNPq e estudante de doutorado no Department of Information Studies, The University of Sheffield, Western Bank, Sheffield S10 2TN, England.

... Professor no mesmo departamento.

mação. Para tanto, além da revisão bibliográfica do assunto, criou-se um sistema especialista para auxiliar na execução de uma das atividades profissionais do bibliotecário. A intenção em criar o sistema era a de demonstrar e testar a aplicabilidade da tecnologia à área da informação, mais do que produzir um sistema que fosse realmente operacional.

Antes de se delinear as características específicas do sistema, uma introdução aos conceitos de inteligência artificial e sistemas especialistas faz-se necessária.

1.1 Inteligência artificial

A inteligência artificial (IA) é definida por BARR e FEIGENBAUM² como "a parte da ciência da computação envolvida com o desenvolvimento de sistemas inteligentes de computadores, isto é, sistemas que exibem características que nós associamos com a inteligência no comportamento humano - compreensão da língua, aprendizado, raciocínio, solução de problemas, etc". A partir dessa definição, fica fácil entender o caráter interdisciplinar da IA e o porquê do interesse que ela desperta em áreas como psicologia, lingüística e filosofia, para citar algumas.

Sua origem é datada da metade dos anos cinqüenta mas o interesse nela aumentou consideravelmente na última década, a partir da divulgação, em 1981, do plano de pesquisa japonês para o desenvolvimento de computadores de quinta geração, os quais deveriam ser capazes de manipular informação usando técnicas da IA. Segundo MORRIS e O'NEIL³, essa iniciativa japonesa fez o interesse na disciplina se difundir, especialmente nos países desenvolvidos, o que, por sua vez, trouxe o desenvolvimento de uma das principais áreas de aplicação da IA: os sistemas especialistas.

1.2 Sistemas especialistas

Os sistemas especialistas (SE) são programas que empregam técnicas da IA para simular um especialista humano em uma área determinada e limitada do conhecimento, isto é, a partir de uma base de conhecimento num domínio específico, eles podem tirar conclusões e dar aconselhamento. Uma definição formal é aquela aceita pela British Computer Society, citada por VICKERY e BROOKS⁴, que considera um sistema especialista como "a inclusão, dentro de um

computador, de um componente baseado em conhecimento a partir de uma habilidade humana especializada, de tal forma que o sistema possa oferecer conselhos inteligentes ou tomar uma decisão inteligente sobre uma função de processamento. Uma característica adicional desejável, que muitos considerariam fundamental, é a capacidade do sistema de, quando requisitado, justificar a sua própria linha de raciocínio numa forma diretamente inteligível pelo usuário".

MORRIS⁵ e BIBB⁶ identificam algumas características dos SEs. Uma síntese das idéias de ambos é mostrada abaixo:

- Os SEs são embutidos de conhecimento organizado, o qual é representado na forma de fatos, relacionamentos e heurísticas⁷ obtidos de especialistas humanos. Como os especialistas humanos, eles devem solucionar problemas complexos em uma área determinada do conhecimento, ao nível especializado.
- O conhecimento é mantido separadamente dos mecanismos de controle e inferência do programa, numa parte chamada de base de conhecimento, o que permite o seu refinamento e melhoramento.
- O conhecimento é representado através de símbolos.
- Os SEs têm a capacidade de explicar e/ou justificar as suas perguntas, decisões e conclusões, quando requeridos, no meio ou no fim de uma consulta.
- Os SEs são desenvolvidos para simular especialistas e também podem cometer erros. Contudo, os programas podem, normalmente, ser modificados quando isso ocorre e alguns sistemas têm a habilidade de aprender com os erros.
- Os SEs podem lidar com informação incompleta ou incerta.

Segundo ALBERICO e MICCO⁷, um SE consiste em não menos do que três componentes: a base de conhecimento, o mecanis-

* Forma de solucionar problemas que não segue etapas seqüenciais e que, nem sempre, produz soluções precisas, mas consegue lidar com problemas complexos, dos quais nem todas as variantes são conhecidas. Tentativa e erro.

mo de inferência e a interface. A base de conhecimento é a parte central do sistema e contém conhecimento de fatos e regras para inferir novo conhecimento. O mecanismo de inferência consiste em regras e princípios para a operação do sistema; ele decide quais regras da base de conhecimento usar e em que ordem, quando um conflito aparece. A interface é o componente que permite a comunicação entre o usuário e o sistema.

O uso que pode ser feito de SEs é bastante diversificado. Dez categorias de aplicações são identificadas por HAYES-ROTH et al.⁸: interpretação, predição, diagnóstico, projeto, planejamento, monitoramento, depuração, reparo, instrução e controle.

Os benefícios da aplicação desses sistemas são particularmente visíveis nos setores industriais e comerciais, onde ganhos financeiros são facilmente mensuráveis, visto que eles adicionam flexibilidade a uma atividade que passa a ser feita por menos pessoas e de uma forma mais barata. Contudo, outras vantagens também podem ser enumeradas: os SEs podem libertar o profissional de atividades rotineiras e torná-lo disponível para outras mais complexas e criativas; tornam disponível 24 horas por dia um conhecimento especializado, raro e de alto custo; proporcionam padronização na tomada de decisões; formalizam o conhecimento de especialistas; podem solucionar certos problemas mais rapidamente do que especialistas humanos; são um recurso atraente para treinamento, já que combinam conhecimento armazenado com a habilidade de explicar a linha de raciocínio e, finalmente, permitem que o indivíduo os use na hora e no ritmo que preferir.

Alguns aspectos desejáveis de problemas que podem ser resolvidos pelos SEs têm sido amplamente debatidos e servem como referência para a análise da apropriação da aplicação de um sistema especialista a uma determinada área. O problema a ser resolvido deve, idealmente, ter as seguintes características: ser específico e bem definido; requerer soluções heurísticas e ser baseado em conhecimento e não em senso comum; não ser trivial; a área de aplicação dever ser identificável com um especialista humano, sendo que este é caro e raro.

Nem todos os SEs desenvolvidos seguem todas estas orientações. Muitas aplicações têm sido desenvolvidas graças à disponibilidade de ferramentas baratas e, ao mesmo tempo, sofisticadas para a criação de sistemas. Essas ferramentas podem ser:

- linguagens de programação como LISP e PROLOG, que foram especialmente desenvolvidas para uso em aplicações da IA. Linguagens convencionais, contudo, também podem ser usadas;
- esqueletos para sistemas especialistas (*shells*), que são estruturas básicas adaptáveis a diferentes áreas de aplicação. Comumente, englobam facilidades para a criação da base de conhecimento, mecanismos de inferência e interface;
- ambientes de engenharia do conhecimento, que são mais sofisticados do que os esqueletos porque integram ferramentas para auxiliar em todo o processo de criação de um SE, desde a aquisição do conhecimento até a sua implementação.

2 SISTEMAS ESPECIALISTAS E BIBLIOTECAS

Os SEs têm sido tradicionalmente aplicados a áreas altamente especializadas, como, por exemplo, diagnóstico médico, configuração de computadores e prospecção geológica. Mais recentemente, a biblioteconomia e a ciência da informação também tornaram-se alvos de pesquisas na aplicação desses sistemas a algumas de suas atividades específicas.

2.1 O papel do profissional da informação

Além do papel que os SEs têm em automatizar algumas das atividades profissionais do bibliotecário e/ou cientista da informação, esses profissionais também têm um papel a desempenhar na criação daqueles sistemas. MORRIS e O'NEILL³ esclarecem que o profissional da informação é responsável pela disseminação da informação e que, como é esperado que os SEs tenham um impacto em todos os aspectos da atividade humana, modificando as necessidades e a maneira como a informação é usada, é lógico que os profissionais da informação devem estar a par dos novos desenvolvimentos e prontos a responder a eles. Na opinião de FORD⁹, o que possivelmente tenderá a acontecer é uma extensão de papéis, de profissional da informação para profissional do conhecimento, sendo que por profissional do conhecimento ele define aquele que administra tanto os recursos informacionais quanto os que armazenam conhecimento de uma forma diretamente aplicável, como o fazem os SEs.

O bibliotecário e o cientista da informação têm muito a oferecer à disciplina de IA. MORRIS e O'NEILL³, numa investigação dos possíveis papéis do profissional da informação na área dos SEs, identificaram sete maneiras pelas quais esses profissionais poderiam contribuir com seu conhecimento:

- 1 como especialistas em fontes de informação, providenciando informação básica sobre o assunto ao engenheiro de conhecimento* ;
- 2 como engenheiros de conhecimento, já que muitas das habilidades requeridas são aquelas que o profissional da informação tem: habilidade de conduzir entrevistas, de comunicação interpessoal e habilidade para elucidar, organizar, estruturar e disseminar o conhecimento, para mencionar algumas;
- 3 colaboração em estudos de usuários, para superar as limitações encontradas em algumas aplicações, as quais provaram ser pouco usadas por falta de prévia avaliação das necessidades de informação;
- 4 como agente da informação, sendo intermediário entre o usuário e o sistema, recomendando o uso do sistema, sugerindo fontes alternativas e interpretando os resultados de consultas;
- 5 atualizando e estendendo os SEs depois de estes estarem em uso;
- 6 ligando ciência da computação à ciência da informação, já que é esperado que a colaboração interdisciplinar será necessária em projetos futuros;
- 7 desenvolvendo SEs para clientes, especialmente aqueles cujas necessidades não estão sendo supridas pelas pesquisas convencionais em IA.

* Função comparável ao do analista de sistema. É responsável por capturar o conhecimento, analisá-lo e incorporá-lo ao SE.

As autoras concluem que há um grande potencial para os profissionais da informação nessa área. Contudo, muito vai depender do interesse pessoal, do entusiasmo e do comprometimento desses profissionais com a tecnologia,.

Para BRITTAIN¹⁰, há duas áreas onde os novos graduados em biblioteconomia/ ciência da informação terão um papel importante a desenvolver e ambas estão completamente fora do enfoque tradicional na sua formação: na fase de aquisição do conhecimento, que fará parte do sistema e na operação, manutenção e melhoramento dos SEs em diversos campos. Ele conclui dizendo que o papel do profissional da informação nessas áreas deve ser explorado antes que outros o façam e, mais tarde, reivindiquem exclusividade sobre o que é hoje uma área interdisciplinar.

2.2 Áreas de aplicação

A recuperação da informação é, provavelmente, a área dentro da ciência da informação na qual os SEs têm sido mais freqüentemente aplicados. Já se criou uma grande variedade de sistemas, que funcionam como intermediários humanos, ao providenciar assistência inteligente para usuários finais de bases de dados *online*, como por exemplo, CANSEARCH e MENUSE e EP-X¹¹.

Sistemas desse tipo normalmente auxiliam o usuário na escolha da base de dados apropriada, na formulação da estratégia de busca e na reformulação dessa estratégia, se ela não produzir resultados satisfatórios. Além disso, técnicas de IA têm sido usadas para desenvolver arquiteturas que exploram melhores maneiras de armazenar, descrever e recuperar documentos.

A referência é outra área na qual os SEs têm sido freqüentemente aplicados. Ela é vista como o meio de providenciar serviços de informação, usando-se a coleção da biblioteca ou qualquer outro recurso disponível¹², o profissional agindo como uma espécie de consultor, faz o papel intermediário entre o usuário e a informação. Sendo assim, aparentemente, há um grande potencial para a utilização de SEs em referência, já que são comumente usados para consultoria, na maior parte das aplicações.

Um exemplo de SE para este fim é PLEXUS¹³, que foi desenvolvido para responder a perguntas de referência, usando fontes de informação na área de jardinagem. O sistema interage com o usuário, fazendo perguntas destinadas a entender o problema do cliente,

usando, para isso, um dicionário e um banco de dados em jardinagem, uma classificação hierárquica dos conceitos nessa área e uma lista de palavras a serem ignoradas (*stopwords*). As respostas do sistema são na forma de direcionamento a recursos, tanto humanos quanto bibliográficos, que satisfarão as necessidades de informação do cliente.

A classificação também tem sido considerada uma área promissora para a aplicação de SEs. Basicamente, ela é uma atividade complexa, que consome bastante tempo e exige do classificador um bom conhecimento do assunto que está sendo classificado e uma sólida compreensão dos princípios e notações do sistema em uso. Além disso, é necessário um controle para manter consistência no uso das notações. De fato, a inconsistência é o principal problema que os SEs poderiam solucionar nessa área.

CLARKE e CRONIN¹⁴ idealizaram um sistema especialista para a classificação de monografias que interagisse com o classificador, em linguagem natural, fazendo perguntas sobre título, assunto e capítulos da obra. O sistema iria, então, examinar, pesar e ordenar os termos, de acordo com o uso e a relevância para, então, compará-los às subdivisões existentes em sua memória e decidir-se pela melhor notação.

Contudo, até o presente momento, nenhum sistema especialista foi de fato desenvolvido com êxito nessa área. Uma tentativa foi feita na University of Strathclyde¹⁵, na Escócia, onde se pretendia criar um sistema para auxiliar em classificação usando o Código Decimal de Dewey. Vários esqueletos para sistemas especialistas foram analisados, mas nenhum se mostrou apropriado e o projeto foi abandonado.

A catalogação, outra atividade que se mostra apropriada para o desenvolvimento de SEs, será considerada em mais detalhe a seguir. A intenção deste artigo é dar apenas alguns exemplos do que vem sendo feito no campo da aplicação de SEs a bibliotecas e serviços de informação. Vários trabalhos abordam o tema de uma forma bem aprofundada, dentre os quais sugere-se consultar FORD⁹, MORRIS⁵ e ALBERICO e MICCO⁷. Entre outros, menos recentes, mas também úteis estão CLARKE e CRONIN¹⁴, VICKERY e BROOKS⁴, GIBB⁶ e DAVIES¹⁶.

3 SISTEMAS ESPECIALISTAS E CATALOGAÇÃO

A catalogação é uma atividade complexa, que consome bastante tempo e que, obviamente, requer conhecimento especializado porque

não consiste, simplesmente, na aplicação de regras, mas também da tomada de decisões intelectuais por parte do catalogador. Por suas características, a catalogação é uma atividade que se torna uma provável candidata para a aplicação do SEs.

DAVIES e JAMES¹⁷, num estudo da aplicabilidade de SEs à catalogação, distinguem os seguintes tipos de conhecimentos necessários ao catalogador: conhecimento dos diferentes tipos de materiais e suas características, das regras de catalogação usadas, das particularidades da entidade catalogadora e conhecimento das regras de arquivamento.

Como já foi dito anteriormente, a catalogação é uma atividade baseada na aplicação de regras. No mundo ocidental moderno, o principal conjunto de regras que governam essa atividade é o Código de Catalogação Anglo-Americano, em sua segunda edição (CCAA2).

3.1 CCAA2

O CCAA2 é dividido em duas partes: a primeira abrange regras para a descrição normalizada de um item e consiste de um capítulo geral (cap. 1), o qual se aplica a todas as formas de materiais de bibliotecas e os capítulos subsequentes, (cap. 2 - 13), que se aplicam a tipos particulares de materiais. A segunda parte consiste em regras para a determinação da entrada principal e das secundárias (cap. 21), forma dos cabeçalhos e títulos uniformes (cap. 22-25) e forma das remissivas (cap. 26).

A forma como as regras são apresentadas, particularmente na parte 2 do código, é semelhante à forma com que as regras de produção são criadas na base de conhecimento de um sistema especialista. Dessa maneira, a automação parece fácil, contudo, a aplicação das regras do CCAA2 não é tarefa simples. As regras são freqüentemente consideradas inconsistentes, complexas e ambíguas, características que tornariam sem efeito os aspectos que, à primeira vista, identificaram essa atividade como apropriada para a aplicação de SEs. No entanto, algumas pesquisas mostraram que o desenvolvimento de SEs para auxiliar o processo de catalogação pode ser produtivo. HJERPPE e OLANDER¹⁸, num estudo sobre a relevância da aplicação do CCAA2 como base de conhecimento para SEs, concluíram que um sistema operacional para catalogação é viável na prática, mas que não seria eficaz para a maioria das bibliotecas, no atual estágio da tecnologia, por causa da extensão do CCAA2.

Há consenso na literatura de que as heurísticas dos catalogadores deveriam ser adicionadas à base de conhecimento como forma de melhorar o desempenho de sistemas desse tipo. DAVIES e JAMES¹⁷ sugerem que sejam adicionadas, também, regras contidas nos códigos de práticas das bibliotecas, os quais estão relacionados às particularidades de cada instituição. Outros autores, porém, acham que o código deve mudar se deseja a sua implementação na forma de sistema especialista. HJERPPE e OLANDER¹⁸ sugerem que um novo código com menos regras poderia trazer mais consistência à atividade de catalogação e, conseqüentemente, na sua implementação como um sistema especialista.

3.2 Sistemas especialistas em catalogação

Pesquisas no desenvolvimento de SEs para uma catalogação baseada no CCAA2 estão ainda em progresso. Três diferentes abordagens podem ser identificadas: uma interessada em desenvolver sistemas para dar orientação ao catalogador na aplicação das regras, outra interessada na aplicação e execução das regras juntamente com o catalogador para automaticamente criar os registros e, finalmente, uma envolvida com a criação de sistemas completamente automatizados, que dispensariam o especialista humano.

a) Sistemas para consultoria

Segundo MORRIS⁵, as primeiras tentativas de se produzir um sistema especialista para catalogação foram as desenvolvidas por BLACK et al e EYRE. Os primeiros, usando diferentes esqueletos para sistemas especialistas, criaram duas versões de um sistema especialista chamado HEADS que se mostrou insatisfatório por causa das limitações nos esqueletos. Eles concluíram que um sistema para consultoria, por si só, não era de muita valia e então propuseram a sua integração com mecanismos que produzissem, automaticamente, os registros catalográficos. O último desenvolveu em PROLOG um sistema constituído de *menus* para orientar na forma dos nomes de pessoa a serem usados como cabeçalhos. O sistema mostrou-se praticável, mas o autor concluiu que um sistema desse tipo não seria útil já que, segundo ele, seria mais fácil usar a versão impressa do código.

CATALYST (GIBB e SHARIFF¹⁹) é um dos sistemas para

consultoria mais conhecidos. Ele foi desenvolvido na University of Strathclyde, com o objetivo de orientar no uso e forma dos pontos de acesso segundo o CCAA2. O sistema foi desenvolvido a partir de um esqueleto para sistemas especialistas e utiliza os capítulos 21-25 do CCAA2 como base de conhecimento, além de fazer uso da estrutura proposta no algoritmo criado por SHAW et al.²⁰.

Um método diferente no desenvolvimento de sistemas para consultoria em catalogação é o relatado por ERCEGOCAC e BORKO²¹, desenvolvido em um trabalho de doutorado na University of Califórnia. Até então, todos os sistemas eram criados tendo como base de conhecimento a parte 2 do CCAA2 e não eram específicos a um tipo de material. O sistema criado por eles, chamado MAPPER, orienta na descrição normalizada de mapas de uma folha, emanados de três editoras dos Estados Unidos. O sistema, além de fazer uso das regras relevantes do CCAA2, também inclui heurísticas obtidas de especialistas em catalogação de mapas. A idéia do sistema é auxiliar estudantes e profissionais iniciantes na aplicação das referidas regras.

b) Sistemas para a criação de registros

Sistemas para a criação de registros são uma tentativa de integrar sistemas para consultoria com mecanismos que permitiriam a produção do registro catalográfico. A primeira tentativa nesse sentido foi a empreendida por DAVIES e JAMES¹⁷ da University of Exeter, Inglaterra.

O sistema desenvolvido usa *menus* como forma de agilizar o processo de comunicação com o usuário. A interação continua até o sistema ter informações suficientes para determinar os pontos de acesso requeridos, quando então ele apresenta na tela um formulário, o qual é dependente da natureza do item que está sendo catalogado, com alguns campos já preenchidos e outros a serem preenchidos pelo catalogador. Contudo, o sistema nunca foi completamente implementado por falta de recursos e restringiu-se à criação de um pequeno protótipo em PROLOG e PASCAL.

Enquanto o trabalho em Exeter estava em progresso, outro projeto tinha início na Universidade de Linkoping, Suécia. Os autores, HJERPPE e OLANDER¹⁸, tinham como principal objetivo avaliar a relevância, as técnicas e as ferramentas disponíveis para o desenvolvimento de SEs no contexto das bibliotecas.

Duas versões do sistema ESSCAPE (*Expert System for Simply Choice of Access Points for Entries*) foram desenvolvidas, ambas baseadas no capítulo 21 do CCAA2. O primeiro sistema interage com o usuário, fazendo perguntas sobre o item que está sendo catalogado até obter informações suficientes para chegar a uma conclusão. Quando isso acontece, a sugestão da forma da entrada principal e secundária do item é apresentada ao usuário. A segunda versão somente indica as regras que devem ser consultadas para a solução do problema.

c) Sistemas completamente automáticos

Outra parte da pesquisa nessa área concentrou-se na produção de sistemas completamente automáticos, ou seja, que lêem e interpretam dados automaticamente, selecionam as regras apropriadas e, finalmente, produzem o correspondente registro catalográfico.

O trabalho mais importante nessa área é o de WEIBEL et al.²² na *Online Computer Library*, em Dublin. O sistema foi desenvolvido para catalogar um item automaticamente, através das informações recuperadas da página de rosto. O sistema limita-se à catalogação descritiva de um item através da identificação dos seguintes elementos: título (e informação suplementar), responsabilidade, editor, lugar e data de publicação e edição. Essas informações são obtidas da página de rosto do documento, fazendo uso da técnica de reconhecimento ótico de caracteres e é processada por um conjunto pequeno de regras. Os autores dizem que 80% dos campos bibliográficos foram identificados corretamente pelo sistema, num conjunto de páginas de rosto escolhidas ao acaso.

No entanto, segundo eles, se um melhor desempenho é desejado desse tipo de sistemas vários problemas precisam ser solucionados, por exemplo, a irregularidade do *layout* das páginas de rosto, problemas de interpretação gráfica dos instrumentos usados e a complexidade das regras de catalogação. Por esses problemas serem, em sua maioria, de difícil solução, os autores acreditam que um sistema completamente automatizado para catalogação não será possível num futuro próximo, embora o método seja promissor.

Depois dessa revisão das pesquisas na aplicação de IA à catalogação, a conclusão a que se chega é a de que muito ainda está por ser feito. Sistemas para consultoria têm sido criados, mas a sua utilidade fora do meio acadêmico é duvidosa. DAVIES et al.¹² lembram

que os catalogadores somente usariam esse tipo de sistema quando conscientes de um problema para o qual eles não sabem a resposta. Tal consciência, contudo, dependeria do conhecimento prévio do código de catalogação e, dessa forma, ainda poderiam ser produzidos registros catalográficos incorretos. Além do mais, o efeito na economia do processo seria pequeno, por causa da pouca utilização.

Para realmente reduzir-se o tempo e o custo consumidos em catalogação, são necessários sistemas completamente automatizados. Contudo, isto não parece praticável no futuro próximo, já que os processos humanos usados em catalogação não foram completamente analisados e o CCAA2 ainda tem muitas limitações. Segundo HJERPPE e OLANDER¹⁸ estas são:

- 1 a estrutura hierárquica das regras não corresponde com exatidão à numeração das mesmas;
- 2 os conceitos básicos não são sempre bem definidos;
- 3 alguns conceitos se sobrepõem;
- 4 as alternativas não são mutuamente exclusivas e há pouca orientação na escolha das mesmas quando as circunstâncias são diferentes;
- 5 as distinções e exceções parecem não ter sido bem elaboradas.

DAVIES et al.¹² observam que seria prematuro embarcar em uma revisão total do código antes que se alcance um consenso quanto ao papel, conteúdo e estrutura do catálogo e do processo pelo qual ele é criado e mantido. Contudo, como expõe FORD⁹, a sua reestruturação é de fundamental importância, se desejamos a preservação dos benefícios que esse tipo de código traz.

4 DESENVOLVIMENTO E IMPLEMENTAÇÃO DO SISTEMA

4.1 Entendendo o problema

A primeira fase no desenvolvimento do sistema foi a de encontrar uma atividade específica, dentro da biblioteconomia ou da ciên-

cia da informação, que fosse bem definida e específica e bem documentada na literatura. Esses requisitos eram importantes por se tratar de um trabalho de mestrado, com claros limites de tempo para conclusão. Escolhemos a catalogação, porque é identificada na literatura como uma área promissora para a aplicação de SE e por ser uma área técnica e bem determinada, com um corpo definido de regras bem documentadas e que requer conhecimento especializado para ser executada com precisão.

O segundo passo foi a escolha de uma ferramenta para o desenvolvimento do sistema. As linguagens de programação PROLOG e LISP foram descartadas devido à falta de experiência no seu uso. Além do mais, o limite de tempo estabelecido era curto e deveria ser rigorosamente seguido. Em vista disto, o uso de um esqueleto para sistemas especialistas foi considerado a melhor opção. Os esqueletos que estavam à disposição naquele momento - KnowledgePro, Crystal e Leonardo - foram devidamente avaliados e o último foi considerado o mais apropriado.

4.2 Leonardo Shell

Leonardo Shell é um programa para a criação de SEs. Ele foi desenvolvido para ser usado em PCs pela *Creative Logic Limited* e requer um microcomputador com um mínimo de 512k de memória e um disco rígido com pelo menos 1M de espaço disponível. Naquela época, existiam três versões diferentes, para este trabalho, sendo usada a versão 3.4.

Os principais componentes do *Leonardo* são:

- a base de conhecimento, que é criada pelo idealizador do sistema e composta de regras e objetos com respectivos quadros (*frames*);
- o editor, que permite a criação e a edição da base de conhecimento;
- uma linguagem de procedimentos própria do *Leonardo*, usada na criação de telas e cálculos complexos, na produção de relatórios e como interface à base de dados externos;
- funções de controle e execução. A primeira converte a base de conhecimento do formato textual para o formato codificado, usado pelo computador e verifica se as regras, objetos e procedimentos são reconhecíveis e logicamente consistentes.

A segunda executa o sistema que foi verificado com êxito;

- facilidades para manusear os arquivos. Uma base de conhecimento é mantida em um arquivo cujo nome tem a extensão “.PKB”. Cada vez que uma base de conhecimento é acessada, o arquivo correspondente deve ser aberto. Arquivos externos do tipo texto também podem ser lidos e escritos;

Como foi dito anteriormente, *Leonardo* representa o domínio do conhecimento através de regras, objetos e quadros. Para desenvolver um sistema no programa é fundamental estar consciente da maneira como os blocos se comportam.

a) Regras

As regras são expressões condicionais que entram no sistema na forma

SE <condição> ENTÃO <ação>

Isto é, se a condição é encontrada, então a ação é invocada. Condição e ação são cláusulas que contêm, pelo menos, um objeto-operador-valor conjunto. Por exemplo, a regra básica em *Leonardo* tem a seguinte estrutura:

SE <objeto 1> <operador> <valor 1>

ENTÃO <objeto2> <operador> <valor 2>

Além das regras de produção, são necessárias também algumas regras afirmativas para controlar a execução das regras de produção.

b) Objetos

Leonardo identifica objetos quando da verificação das regras e, então, cria um quadro para cada um deles, de acordo com o tipo: real, texto, lista ou procedimento.

c) Quadros

Todos os objetos em *Leonardo* têm um quadro a eles associado, que armazena informações sobre os objetos, em campos apro-

priados. Alguns desses campos são preenchidos pelo próprio sistema e outros podem ser editados.

4.3 Criando a base de conhecimento

Como não seria possível usar todo o conhecimento contido no CCAA2 por causa de sua extensão e complexidade, uma parte dele foi escolhida para o protótipo, mais exatamente o capítulo 21, por ser considerado complexo mas, ainda assim, possível de decomposição. Além disso, a disponibilidade do algoritmo para a parte 2 do código, elaborado por SHAW et al.²⁰, veio ajudar na criação do sistema.

Quatro subdivisões foram identificadas no conhecimento contido no capítulo 21, quais sejam: publicações legais, publicações religiosas, obras relacionadas e categorias não especiais. Para cada uma delas, criou-se um conjunto de regras, o qual, juntamente com o conjunto principal, alcançou um total de 203 regras de produção, elaboradas segundo a estrutura em *Leonardo*.

O processo de conversão das regras do capítulo 21 do CCAA2 para as regras de produção, usadas pelo sistema, não foi tarefa simples, como aparentava à primeira vista. As regras do código são bastante complexas e uma única delas pode envolver vários conceitos diferentes. Além do mais, a enumeração hierárquica das regras não corresponde a uma subdivisão hierárquica das informações contidas no código. Para algumas foi necessário criar até oito regras de produção na base de conhecimento.

Além da criação das regras, foram também editados os quadros para cada objeto que *Leonardo* identificou, quando da verificação das mesmas, ou seja, foram estabelecidos os possíveis valores de cada objeto, as perguntas que o sistema deveria formular quando o valor daquele objeto era demandado, etc.

Procedimentos, que também são um tipo de objeto, foram elaborados para criar as telas finais, que são a forma de apresentação da resposta a uma consulta. Todas regras do sistema, exceto o conjunto principal de regras, invocam um procedimento como parte da ação. No total, 161 procedimentos foram criados.

5 O SISTEMA EM OPERAÇÃO

O primeiro passo para se usar a base de conhecimento é acessar

Leonardo através do comando DOS. O usuário carrega o sistema com os arquivos da base de conhecimento e escolhe a opção EXECUTE do *menu* do sistema. Quando isso é feito, aparece a seguinte tela:

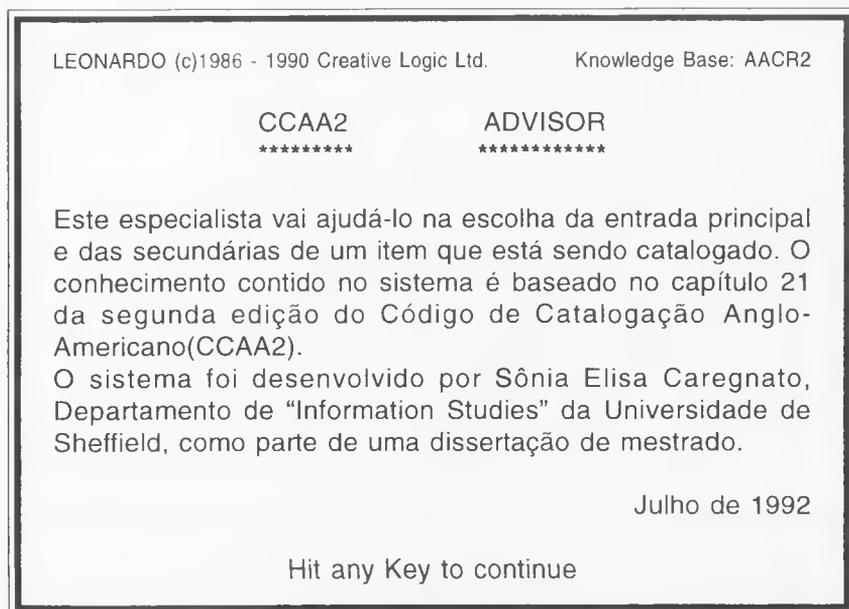


Figura 1: Primeira tela do sistema especialista.

Seguindo-se as instruções que aparecem no final da tela, qualquer tecla é pressionada e isto dá acesso ao primeiro *menu* do sistema.

Este sistema foi originalmente escrito em inglês. Neste artigo, no entanto, serão traduzidos os textos escritos pelo idealizador do sistema, mas não os comandos de *Leonardo*.

Este sistema para consultoria lhe dá quatro opções para a identificação da categoria do trabalho a ser catalogado:

publicações legais, publicações religiosas, obras relacionadas e categorias não especiais. Qualquer informação adicional sobre cada uma dessas opções é mostrada ao lado, quando você seleciona. Selecione a opção apropriada e pressione a tecla ENTER.

Qual das seguintes opções se aplica ao item?

- | | |
|--------------------------|---|
| ⇒ publicações legais | Publicações legais inclui trabalhos tais como leis, tratados, decisões dos tribunais, etc. Não inclui textos sobre leis, edições anotadas de comentários a leis, etc. |
| publicações religiosas | |
| obras relacionadas | |
| categorias não especiais | |

Keys: 1 Help 2 Quit 3 Why? 5 Volunteer 6 Backup 7 Expand 8 Review

Figura 2": Primeiro menu do sistema especialista.

As quatro principais subdivisões do conhecimento contido no capítulo 21 são mostradas. Suponhamos que o usuário do sistema queira catalogar o livro *Intelligent information system: progress and prospects*, editado por Roy Davies, um conjunto de capítulos escritos por autores diferentes e produzido sob uma direção editorial.

O sinal ⇒ é usado para indicar a opção que está selecionada.

O primeiro ponto a ser considerado é a natureza do trabalho. As opções são aquelas mostradas no primeiro *menu*. Explicações complementares para cada opção podem ser obtidas quando a opção é selecionada. Neste caso, selecionando "categorias não especiais", a tela fica assim:

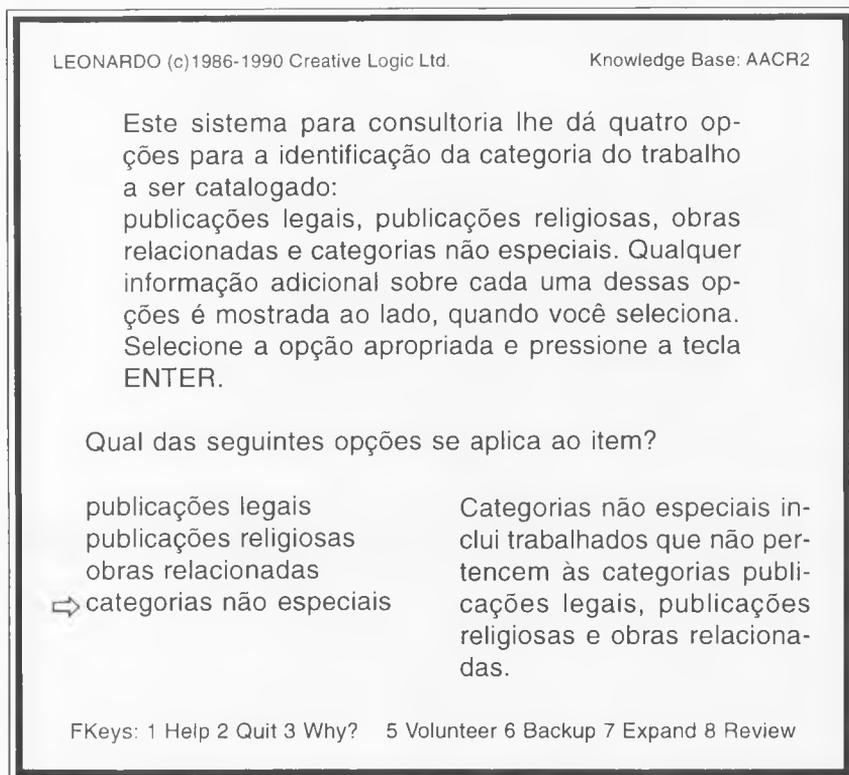


Figura 3": Tela quando a opção "categorias não especiais" é selecionada.

Quando a tecla ENTER é pressionada o *menu* para a opção é apresentado.

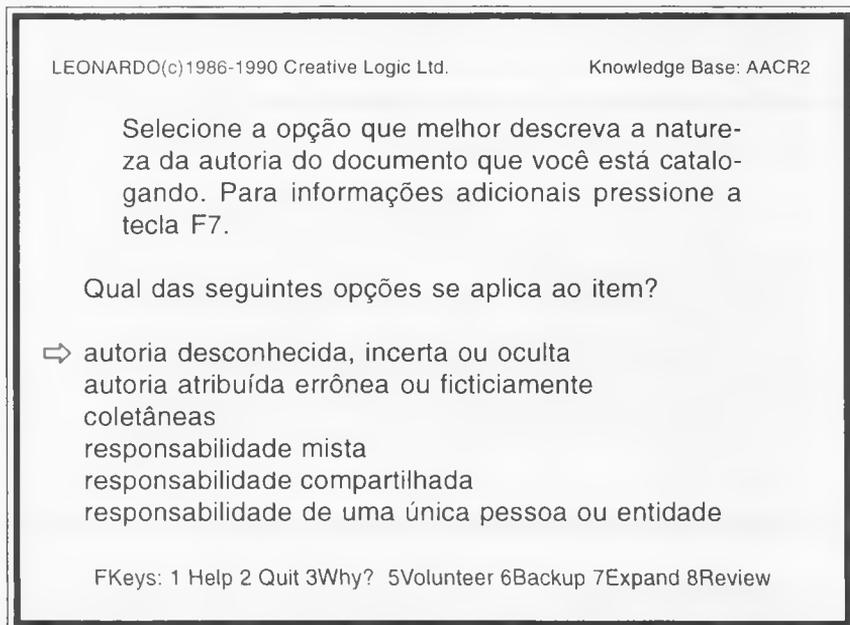


Figura 4: Menu para a opção "categorias não especiais".

Para informações suplementares acerca da aplicabilidade de cada uma das opções mostradas no *menu*, a tecla F7 (EXPAND) pode ser pressionada, como indicado.

LEONARDO(C) 1986 - 1990 Creative Logic Ltd.

Knowledge Base: AACR2

– "autoria desconhecida, incerta ou oculta" se aplica a obras cuja autoria é indicada por iniciais, numerais ou marcas tipográficas, a obras anônimas e a obras cuja autoria é atribuída a alguém.

– "autoria atribuída errônea ou ficticiamente" se aplica a obras cuja autoria é errônea ou ficticiamente atribuída a uma pessoa ou uma entidade.

– "coletâneas" são coletâneas de obras independentes por pessoas e entidades distintas; coletâneas que consistem em extratos de obras independentes por pessoas e entidades distintas; obras que consistem em contribuições de pessoas e entidades, produzidas sob a direção de um coordenador e obras que consistem, em parte em obras independentes

Qual das seguintes opções se aplica ao item?

- ⇒ autoria desconhecida, incerta ou oculta
- autoria atribuída errônea ou ficticiamente
- coletâneas
- responsabilidade mista
- responsabilidade compartilhada
- responsabilidade de uma única pessoa ou entidade

Type any Key to see the next screen

Figura 5: Primeira tela mostrada quando a tecla F7 é pressionada.

Por pessoas e entidades diferentes e, em parte, de contribuições produzidas sob a direção de um coordenador.

– "responsabilidade mista" é aquela cujo conteúdo intelectual ou artístico recebeu contribuições de diversas pessoas ou entidades, desempenhando diferentes tipos de atividades.

– "responsabilidade compartilhada" é aquela produzida pela colaboração entre duas ou mais pessoas ou entidades desenvolvendo o mesmo tipo de atividade na criação do seu conteúdo.

– "responsabilidade de uma única pessoa ou entidade" se aplica a obras de autoria de uma só pessoa ou provenientes de uma única entidade.

Qual das seguintes opções se aplica ao item?

autoria desconhecida, incerta ou oculta
autoria atribuída errônea ou ficticiamente

⇒ coletâneas

responsabilidade mista

responsabilidade compartilhada

responsabilidade de uma única pessoa ou entidade

Type any Key so see the next screen

Figura 6: Segunda tela mostrada quando a tecla F7 é pressionada.

Lendo-se as informações suplementares fica claro que a obra, que é produzida sob direção editorial, deve ser tratada como coleção. O usuário escolhe aquela opção e a consulta prossegue com a seguinte tela:

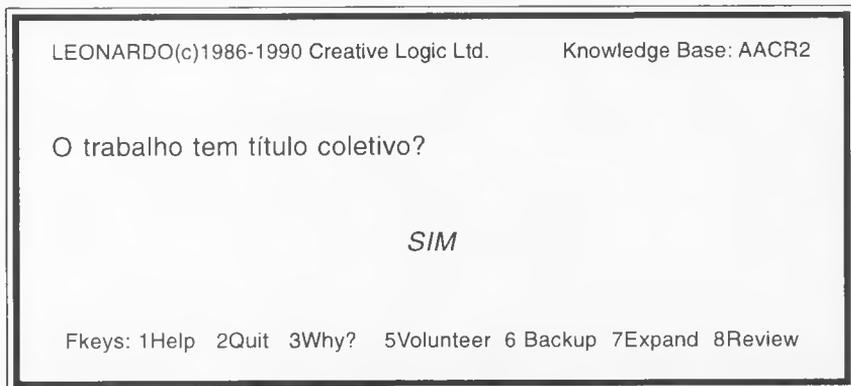


Figura 7: Interação entre sistema e usuário

Como o trabalho tem um título coletivo o usuário digita "sim" e a próxima tela que aparece é:

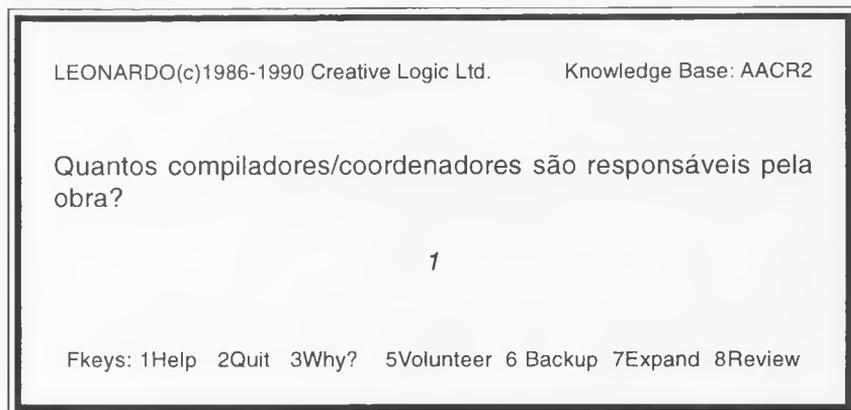


Figura 8: Interação entre sistema e usuário.

O usuário digita o número 1, já que um só editor é responsável pelo trabalho. A tela seguinte é:

LEONARDO(c)1986-1990 Creative Logic Ltd. Knowledge Base: AACR2

Quantas contribuições ou obras independentes estão incluídas na obra?

10

Fkeys: 1Help 2Quit 3Why? 5Volunteer 6 Backup 7Expand 8Review

Figura 9: Interação entre sistema e usuário.

A resposta é 10, correspondendo aos dez capítulos do livro, cada um escrito por um ou mais contribuidores. A tela que é mostrada é:

LEONARDO(c)1986-1990 Creative Logic Ltd. Knowledge Base: AACR2

Quantos colaboradores estão envolvidos na obra?

13

Fkeys: 1Help 2Quit 3Why? 5Volunteer 6 Backup 7Expand 8Review

Figura 10: Interação entre sistema e usuário.

O usuário digita o número 13 o qual corresponde ao número de pessoas responsáveis por alguma parte do trabalho.

O sistema já fez perguntas suficientes para inferir um resultado a partir das regras disponíveis na base de conhecimento. A tela mostrada como resposta à consulta é a seguinte:

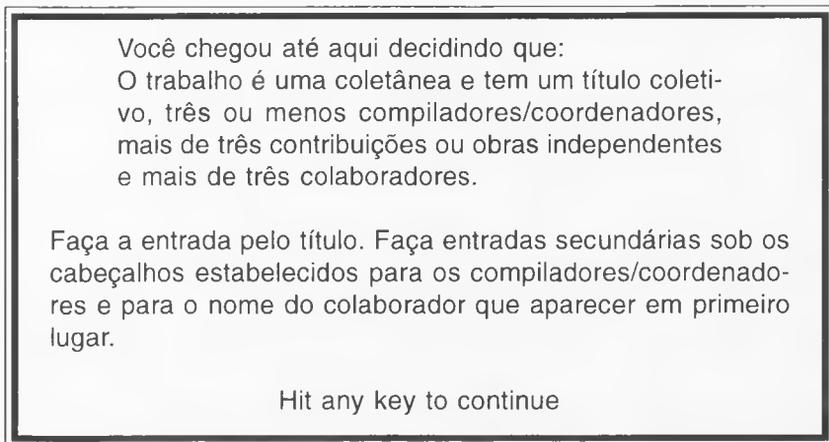


Figura 11: Resposta do sistema à consulta do usuário.

Quando o usuário pressiona alguma tecla a última tela é mostrada:

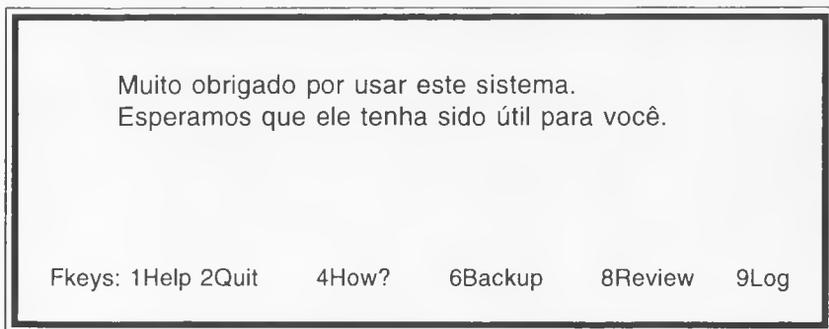


Figura 12: Encerramento da consulta.

No final da consulta o usuário pode pressionar a tecla F4 (HOW?) para obter explicações sobre como aquela solução foi alcançada. Durante a consulta, se o usuário pressiona a tecla F3 (WHY?), *Leonardo* explica porque está fazendo aquela pergunta.

“VOLUNTEER”(F5), “BACKUP” (F6) e “REVIEW”(F8) também estão disponíveis durante ou no final da consulta. A primeira permite que o usuário interrompa a consulta e dê respostas antes de ser questionado, cortando assim o tempo gasto na consulta. A segunda permite a mudança nas respostas que foram dadas anteriormente e a última mostra todo o diálogo entre o usuário e o sistema durante a sessão.

6 CONCLUSÕES

O sistema aqui apresentado de nenhuma forma fez total uso do potencial oferecido pela tecnologia dos sistemas especialistas. Aquilo que FORD⁹ escreveu sobre o sistema CATALYST é também aplicável a este: ele é mais um *smart manual* do que um sistema especialista e poderia ter sido programado usando-se técnicas de programação mais tradicionais. Contudo, como ele mesmo explica, não há nada de errado nisto se o sistema for eficaz. Vale lembrar também que, no caso desta aplicação, o objetivo era mais acadêmico do que prático.

Como trabalho de pesquisa, o exercício foi altamente proveitoso porque, além de avaliar a conveniência do CCAA2 como base de conhecimento para sistemas especialistas, mostrou, também, que há espaço para profissionais da área da biblioteconomia e ciência da informação na aplicação da tecnologia. Estes profissionais estão, de uma forma ou de outra, participando do desenvolvimento de sistemas baseados em conhecimento e do surgimento de ferramentas mais poderosas e baratas que estão abrindo possibilidades para aqueles que, não tendo prévia experiência em programação, desenvolverem seus próprios sistemas. Além do mais, como foi mostrado em estudos citados, vários outros papéis estão se abrindo, nesta área, aos profissionais da informação.

Isso provavelmente vai nos dar a habilidade de estender nosso papel de profissionais da informação para profissionais do conhecimento.

EXPERT SYSTEMS IN LIBRARIES: THE DEVELOPMENT OF A PROTOTYPE FOR CATALOGUING

This paper discusses briefly what expert systems are, the impact they are having on the Library and Information Science (LIS) fields and outlines some areas within these fields where they have been applied. In addition, it describes in detail a system developed in the Department of Information Studies at the University of Sheffield as part of a M. Sc. dissertation to assist in the use of the rules of the Chapter 21 of the Anglo-American Cataloguing Rules (AACR2).

7 REFERÊNCIAS BIBLIOGRÁFICAS

- 1 CAREGNATO, S. E. **The development of a demonstration expert system for cataloguing using Leonardo expert system shell.** Sheffield, 1992. 238p. (Dissertação- Mestrado - University of Sheffield).
- 2 BARR, A., FEIGENBAUM, E. A. (Ed.) **The handbook of artificial intelligence.** London: Pitman, 1981. v.1.
- 3 MORRIS, A. O'NEILL, M. Information professionals - roles in the design and development of expert systems **Information Processing & Management**, v.24, n.2, p.173 - 81, 1988.
- 4 VICKERY, A., BROOKS, H. Expert systems and their applications in LIS. **Online Review**, v. 11, n.31, p. 149-65, 1987.
- 5 MORRIS, A. (Ed.) **The application of expert systems in libraries and information centres.** London: Bowker-Saur, 1992.
- 6 GIBB, F. **Expert systems in libraries.** London, Taylor Graham, 1986.
- 7 ALBERICO, R., MICCO, M. **Expert systems for reference and information retrieval.** Westport: Meckler, 1990. 395p.
- 8 HAYES-ROTH, F., WATERMAN, D. A., LENART. An overview of expert systems. In: _____ (Ed.) **Building expert systems.** London: Addison-Wesley Publishing Co., 1983. v.1, cap. 1, p. 3-29.
- 9 FORD, N. **Expert systems and artificial intelligence: an information manager's guide.** London: Library Association Publishing, 1991.

- 10 BRITTAİN, M. Implications for LIS education of recent developments in expert systems. **Information Processing & Management**, v.23, n.2, p.139-152, 1987.
- 11 POLLITT, A. S. Intelligent interfaces to text retrieval systems. IN: GILMAN, P. (Ed.) **Text retrieval: the state of the art**. London: Taylor Graham, 1990. p. 129-208.
- 12 DAVIES, R., SMITH, A. G., MORRIS, A. Expert systems in reference work. IN: MORRIS, A. (Ed.) **The applications of expert systems in libraries and information centres**. London: Bowker-Saur, 1992. p. 91-132.
- 13 VICKERY, A., BROOKS, H., ROBINSON, B., et al. A reference and referral system using expert systems techniques. **Journal of Documentation**, v. 43, n. 1, p.1-23, Mar. , 1987.
- 14 CLARKE, A., CRONIN, B. Expert systems and library/information work. **Journal of Librarianship**, v.15, n.4, p.227-293.
- 15 SHARİFF, C. **Developing an expert system for classification of books using microbased expert system shells**. London, British Library Research and Development Department, 1988. (British Library Research Paper, 32).
- 16 DAVIES, R. (Ed.) **Intelligent information systems: progress and prospects**. Chichester: Ellis Horwood, 1986.
- 17 DAVIES, R., JAMES, B. Towards an expert system for cataloguing: some experiments based on AACR2. **Program**, v.18, n.4, p.283-297, Oct. 1984.
- 18 HJERPPE, R., OLANDER, B. Cataloging and expert systems: AACR2 as a knowledge base. **Journal of the American Society for Information Science**. v.40, n.1, p. 27-44, 1989.
- 19 GİBB, F., SHARİF, C. CATALYST: an expert assistant for cataloguing. **Program**, v.22, n. 1, p. 62-71, 1988.
- 20 SHAW, M., DENT, B., EVANS, D. et al. **Using AACR2: a step-by-step algorithmic approach to part II of the Anglo-American Cataloguing Rules**. London: The Library Association, 1980.
- 21 ERCEGOVAC, Z., BORKO, H. Design and implementation of an experimental cataloging advisor - mapper. **Information Processing & Management**, v. 28, n.2, p.241-257, 1992.
- 22 WEİBEL, S., OSKİNS, M., VİZİNE-GOETZ, D. Automated title page cataloguing: a feasibility study. **Information Processing and Management**, v. 25, n.2, p. 187-203, 1989.