

CADERNO DE RESPOSTA - PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSC 27

Questão 1

O conceito de indexação de acordo com autores brasileiros da Biblioteconomia e Ciência de Informação refere-se sobre a atividade de atribuir termos, palavras-chave ou frases que expressem o conteúdo semântico do documento a ser analisado. Cabe ressaltar que o processo de indexação está pautado em 2 etapas, nos dois: análise do conteúdo, onde é decidido de que se trata o documento, ou seja, os assuntos que abarcam o conteúdo e esta atribuição por sua vez pode ser feita por meio de termos retirados do próprio documento, neste caso por meio de linguagem natural utilizada pelo próprio autor, ou por meio de atribuições de termos feito pelo indexador através de vocabulários controlados e externos aos documentos, tais como taxonomias e tesouros; a segunda etapa denomina-se tradução, onde os termos escolhidos não expressam o documento de forma padronizada, visando atender as necessidades informacionais do usuário, assim como o nível de especificidade desejado.

Quanto a tipologia pode-se inferir que a indexação pode ser manual, onde lhe são atribuídos termos extraídos de vocabulários controlados, taxonomias, tesouros ou ontologias. Ou de forma automática



CADERNO DE RESPOSTA – PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e Informação
CÓDIGO CANDIDATO(A) RSC27

onde por meio de algoritmos e ferramentas computacionais são atribuídos descritores, considerando e comparando os termos presentes no documento, em oposição a termos "vazios", onde o resultado será apenas de termos relevantes para sua recuperação.

Uma das tendências teóricas para análise de assunto é a Colton Classification, criada por Ranganathan. Neste caso, a atribuição de descritores obedece a Teoria Facetada de Ranganathan. A teoria facetada tem por objetivo indexar documentos por meio de facetas, que sob perspectivas vistas de composição de diferentes atributos e propriedades ao documento. Para tanto Ranganathan idealizou o PMEST. O "P" de propriedade que identifica "o que é" o documento, "M" para o material ou substância do qual o documento é composto. A letra "E" referente ao estado ou o que aconteceu ao documento. E por fim, o "S" quanto ao espaço relativo a referência do documento, e o "T" quanto ao tempo.

A teoria de Ranganathan atribuiu aos documentos indexador uma perspectiva analítico e sintético, garantindo maior flexibilidade e complexidade na recuperação de informação, ao combinar diferentes facetas na descrição temática dos itens do acervo

CADERNO DE RESPOSTA – PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSC27

Outra tendência técnica quanto a análise de assunto e indexação de documentos é a perspectiva socio-cultural desenvolvida por Hyland. Quanto a esta abordagem, Hyland considerava 3 etapas durante a análise de assunto.

A contextualização, ou seja, quais ~~são~~ são as necessidades informacionais do usuário em relação ao ambiente no qual está inserido. Então, por exemplo, para um público especializado, a contextualização deveria, durante a indexação pré-coordenada, ou seja, aquela realizada no ato da indexação (entrada), ser a baixo nível de relevância. Pois durante a indexação pré-coordenada, seria necessário alto nível de precisão, considerando a utilização de operadores booleanos pelo público-alvo.

A próxima etapa é a perspectiva epistemológica, onde diferentes perspectivas e estilos de pensamento são considerados na atribuição de termos para descrever o documento. E por fim, o contexto social e cultural no qual o usuário está inserido, e que deve refletir os valores da sociedade, garantindo assim uma integração mais holística na recuperação de informações.

Deve ressaltar que não há um modelo "correto" de análise de assunto e indexação de documentos.

CADERNO DE RESPOSTA – PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSC27

mas sim, a visões e atuações por parte dos profissionais, visando as necessidades informacionais, de forma mais eficaz e eficiente na recuperação. Uma possibilidade pedagógica e/ou didática que pode ser problematizada em sala de aula é o uso de linked data e aplicações de ontologias em representações temáticas complexas, ganhando não só a qualidade das informações recuperadas, assim como o tratamento efetivo dos descritores. Deste modo, observa-se que as aplicações de Inteligência Artificial e deep learning, nos tendem e desafiam impulsionar a uma sociedade imersa em análise e atribuições de dados para eficaz recuperação da informações.

Tal possibilidade pedagógica visa não só apresentar possíveis soluções em meio a integrações em ambiente digital, mas também evidenciar alternativas de interoperabilidade, comunicações e linguagem, utilizando tanto a análise de assunto quanto o próprio ato de indexação, como mediador entre o conhecimento inserido nos documentos e a necessidade dos usuários.

CADERNO DE RESPOSTA - PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSC 27

Questão 2

Os sistemas de recuperação da informação (SRI), são estruturas lógicas que tem por objetivo a composição de subsistemas cuja finalidade trata-se de evidenciar os melhores mecanismos de busca, consulta, e recuperação da informação a seus usuários. Os SRI estão presentes em uma pluralidade de unidades de informação, tais como, bibliotecas físicas ou digitais, repositórios e bases de dados.

Os subsistemas que compõem um SRI podem ser definidos por: ~~sistema~~ subsistema de indexação, que tem por finalidade a atribuição de descritores usando a caracterização temática dos documentos, e por consequência, a recuperação da informação; subsistema de consulta, onde a aplicação de operadores booleanos, podem maximizar o grau de recuperação dos motores de busca; subsistema de recuperação da informação, este objetiva trazer e apresentar os resultados estabelecidos nos subsistemas anteriores, e garantir maior precisão em um sistema de recuperação da informação; e por fim, o subsistema de feedback, onde o usuário refaz a busca, seja por uso de instrumentos de indexação, como vocabulário controlado, taxonomia, tesauro, etc, ou analisa o resultado dos documentos apresen-

CADERNO DE RESPOSTA – PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSE24

fator e verifica se tal amortiz representa uma
ter relevante a sua pesquisa

No que tange a evolução histórica de sistemas
de recuperação de informações Talamo, aborda quatro
momentos distintos.

O primeiro chamado de Era pré-digital, onde
por meio de catálogos e fichas impressas eram
feitos as buscas no acervo. Porém tratava-se de um
sistema pouco flexível, e que impedia a combinação
de termos, bem como exigia do usuário certo conhe-
cimento dos mecanismos de busca utilizados pela
biblioteca.

Em 1980, atinge-se o primeiro passo de auto-
matização do acervo com uso de ferramentas
computacionais e o emprego de instrumentos que
objetivaram a rápida recuperação de informações.

Até o fim do século 20, a partir da década de 1990,
temos a implantação de motores de busca, como
o AltaVista e o Google, que por meio da World Wide
Web viabilizou a ampliação das buscas não somente
de nos acervos das bibliotecas, mas também nos
principais repositórios e bases de dados.

Atualmente, a utilização de sistemas moder-
nos com adoção de inteligência artificial
trouxe aos usuários uma exponencial capacidade de

CADERNO DE RESPOSTA - PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSE27

de buscar aos usuários, considerando a aplicação de ferramentas como machine e deep learning com enfoque em resultados mais precisos na utilização de algoritmos.

Dentre os critérios de avaliações de um SRI, destaco 7 requisitos fundamentais para a constituição de um sistema eficiente.

O critério de precisão, tem por finalidade indicar quais documentos são úteis a pesquisa dentre os documentos recuperados. O critério de relevância, que objetiva identificar o volume de documentos recuperados em relação a totalidade do acervo. O critério de relevância, busca indicar, dentre os documentos recuperados, quais atendem os critérios de busca.

Também há de se considerar a usabilidade de do sistema, tanto em buscas simples como avançadas. Ou seja, se a interface é intuitiva e didática proporcionando ao usuário facilidade de busca.

Outro critério é a cobertura temática. Um SRI eficiente deve ter uma boa cobertura de assunto no que concerne ao acervo do conhecimento requerido pelos usuários.

O tempo de resposta e a interoperabilidade de são outros critérios essenciais para a composição de um sistema de recuperação de



CADERNO DE RESPOSTA – PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSC 27

informações, pois objetivam, respectivamente, uma rápida resposta ao exterior de busca solicitadas pelo usuário durante a consulta, assim como a integração e conectividade entre diferentes sistemas e instituições.

As redes de informações podem ser vistas como estruturas colaborativas, que visam atender as necessidades informacionais dos usuários. Um bom exemplo é a Library of Congress (LC), por meio de sua coleção centralizada. A LC realiza e distribui a todas as bibliotecas participantes do EUA fichas catalográficas padronizadas de todos os itens de seu acervo, garantindo a uniformidade das informações entre as bibliotecas, o controle bibliográfico do que é produzido no país, e também, o uso por parte dos usuários das melhores meta-bases de busca disponibilizadas pela LC.

Os sistemas de informação, por sua vez, são tipificados quanto às suas atividades, são eles: quanto à interação, ou seja a facilidade que o usuário dispõem no uso de interfaces de busca; quanto ao conteúdo, as evidências em conjunto de metadados bibliográficos, ou de texto completo; e quanto às buscas e o uso de combinações de descritores, que pode ser manual; ou automática no que se refere o uso de IA para a recuperação.

CADERNO DE RESPOSTA – PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSC27

As principais possibilidades de investigações quanto a sistemas de recuperação de informações evidenciam uma tendência no que tange o uso de inteligência artificial visando a tender de forma eficaz os sistemas de avaliação de um SRI, ora mencionados.

Deve ressaltar que o uso massivo de IA em sistemas de recuperação pode trazer benefícios ao buscar e consultar informações realizadas pelos usuários, mas se faz necessário devida atenção e acompanhamento as atividades abarcadas por esta ferramenta sobretudo no que tange os subsistemas de SRI.

A inteligência artificial faz uso de algoritmos alimentados pela prática humana, neste caso, os mecanismos de busca. Para tanto o uso de nuanças como folksonomia e instrumentos de controle terminológico ~~em~~ carecem de maiores pesquisas que possam objetivar a recuperação eficiente de informação aos usuários.



CADERNO DE RESPOSTA – PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSC27

Questão 3

A organização e representação do conhecimento (ORC) pode ser vista como o conjunto de processos e práticas que visam a organização, representação, disseminação e acesso aos documentos. Desta forma, o principal objetivo da ORC está pautado em dar visibilidade às informações, de modo a atender as necessidades informacionais de seus usuários.

Quando se fala na aplicação da ORC no contexto da web semântica, a abordagem metodológica visa ambientar tais processos sob a perspectiva dos recursos digitais. Desta forma, o principal componente da web semântica, ou seja, a expressividade de semântica tem por finalidade estabelecer uma conexão entre os processos dispostos pela organização e representação do conhecimento atrelados à acessibilidade dos documentos, como por exemplo, os níveis de relevância e prioridades.

Nessa vista o desenvolvimento de ORC no contexto da web semântica, a aplicação de ontologias torna-se viável, uma vez que determina quais os conceitos que delimitam a domínio do conhecimento a ser representado, bem como o estabelecimento

CADERNO DE RESPOSTA - PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSC 27

das relações evidenciadas pelos conceitos do domínio. Para tanto, alguns mecanismos podem auxiliar na implantações de ontologias na web semântica.

O RDF (Resource Description Framework) são recursos padronizados para a descrição na web. O RDF, por sua vez, se utiliza de SKOS (Simple Knowledge Organization System), que se trata de um modelo RDF, utilizando linguagens documentarior, tais como taxonomias, tesouros e esquemas de classificações amplamente reconhecidos, tais como a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal, idealizadas respectivamente por Dewey; e Paul Otlet e Henry de la Fontaine.

O FRBR trata-se de um modelo conceitual para a organização e representação do conhecimento baseado em entidade-relacionamento que visa, por meio dos tarefas de usuário, permitir a recuperação de informação dos recursos bibliográficos. A família FRBR, composta por FRAD para registro de autoridades, e FRSAO para assunto, ~~foram~~ atualmente foram aprimorados para o modelo LRM (Library Reference Model), o que evidencia melhor atuação e recuperações dos registros bibliográficos na web.

CADERNO DE RESPOSTA - PROVA ESCRITA - 25/11/2024

Setor: MC-071 Área: Gestão, Organização e Representação do Conhecimento e
Informação
CÓDIGO CANDIDATO(A) RSC27

Em contrapartida o modelo OWL, objective estabelecer uma linguagem padronizada para ontologias complexas e que necessitam de maior especificidade na recuperação da informações.

Para tanto atributos de interoperabilidade, ou seja, a capacidade de diferentes sistemas e instituições interagirem entre si sem perda significativa de informações, é ideal para o contexto de web semântica da ORC. A adoção de linked data, na publicação e conectividade dos registros bibliográficos evidenciam a necessidade de informacionais no contexto da sociedade da informações.

Assim como a adoção do Dublin Core (DC), que são conjuntos de metadados idealizados para a representação de recursos bibliográficos, amplamente utilizados na representação documental e similar ao padrão de metadados MARC21.

Portanto, conclui-se que a organização e representação do conhecimento no contexto de web semântica pode maximizar as potencialidades do acesso e uso dos recursos na web, e trazer melhorias quanto ao avanço de educação, ciência e cultura.