

**INDEXAÇÃO NA PLATAFORMA FamilySearch**

***INDEXING IN THE FamilySearch PLATFORM***

***Deise Maria Antonio Sabbag***

*Doutora em Ciência da Informação*

*Professora da USP*

*Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto*

*Docente do PPGCI- UNESP/MARILA*

*deisesabbag@usp.br*

***Inara Veronez do Nascimento***

*inaraveronez@usp.br*

Recebido em: 30/05/2022

Aprovado em: 20/06/2022

**Resumo**

O presente trabalho tem como objetivo apresentar um estudo sobre a indexação dentro da plataforma genealógica FamilySearch, criada pela Igreja de Jesus Cristo dos Santos dos Últimos Dias e considerada a maior do segmento. O objetivo consistiu em analisar quais os recursos e métodos são empregados na plataforma para indexar. Para isso foi necessário analisar as etapas do processo, os recursos necessários, os objetivos da instituição e suas dificuldades. Foi possível identificar que o FamilySearch realiza um tratamento prévio dos documentos antes de sua disponibilização para indexação, além de dar suporte aos voluntários neste processo. A iniciativa do trabalho desenvolvido pela igreja partiu do interesse em realizar pesquisas genealógicas para concretização de práticas relacionadas aos princípios de salvação estabelecidos pela crença. Foram identificadas dificuldades relacionadas ao volume de materiais para indexação, a idade dos documentos trabalhados, a qualidade do software de indexação e o preparo dos voluntários.

**Palavras Chaves:** Indexação. Indexação via web. Folksonomia. Linguagem documentária. Genealogia.

**Abstract**

*The present work aims to present a study on indexing within the FamilySearch genealogical platform, created by the Church of Jesus Christ of Latter-day Saints and considered the largest in the segment. The objective was to analyze which resources and methods are used in the platform to index. For this, it was necessary to analyze the steps of the process, the necessary resources, the objectives of the institution and its difficulties. It was possible to identify that FamilySearch performs a pre-treatment of documents before they are made available for indexing, in addition to supporting volunteers in this process. The initiative of the work developed by the church came from the interest in carrying out genealogical research to implement practices related to the principles of salvation established by belief. Difficulties related to the volume of materials for indexing, the age of the documents worked, the quality of the indexing software and the preparation of volunteers were identified.*

**Keywords:** Indexing. Web Indexing. Folksonomy. Documentary language. Genealogy.



## 1 INTRODUÇÃO

Por muito tempo as atividades de organização do conhecimento ficaram restritas aos profissionais da informação, no entanto, com o avanço da era digital e da globalização, a produção e o acesso à informação transformou-se, criando um cenário no qual o usuário produz e classifica as informações.

Partindo desta perspectiva, indexar tornou-se intuitivo para os usuários da internet em virtude do seu aspecto interativo, onde tudo e todos estão conectados, por isso, classificar o conhecimento transfigura-se como uma prática inerente do internauta mesmo que não intencionalmente.

Logo, as plataformas colaborativas ganharam relevância, como os fóruns de discussão e blogs, por exemplo. Neste momento, cada usuário pode compartilhar sua experiência pessoal e conhecer as vivências de outras pessoas, por isso, as páginas colaborativas são marcadas pela interatividade dos usuários, onde a produção e classificação da informação não pertencem mais apenas a um grupo específico.

Por esse motivo, algumas organizações, como a plataforma FamilySearch, se apoiam em um sistema colaborativo. Basicamente, este site disponibiliza documentos como certidões de casamento, nascimento, batismo, óbito e demais registros com a parceria de igrejas, cartórios e prefeituras (Introdução a História da Família - Manual do Aluno, 2021). Após a disponibilização destes documentos é necessária a indexação em sua página na web, para isso, conta com o apoio de voluntários.

A indexação via web é executada com apoio de um programa de indexação online, onde o voluntário irá analisar documentos digitalizados e realizar a indexação preenchendo o máximo de informações possíveis. Os resultados obtidos apontam que o FamilySearch atualmente é a maior plataforma de genealogia do mundo (FamilySearch, 2021).

O objetivo geral deste artigo consiste em analisar o processo de indexação via web realizado pelo site FamilySearch da Igreja de Jesus Cristo dos Santos dos Últimos Dias

Para realização deste trabalho foi feita uma pesquisa bibliográfica exploratória na plataforma Brapci (Base de Dados em Ciência da Informação), em seguida, foi criado um perfil na plataforma a fim de explorar o software de indexação. A avaliação levou em consideração os seguintes aspectos: os parâmetros utilizados para indexação, nível de dificuldade e a qualidade final da indexação; bem como as questões tecnológicas que envolvem, usabilidade do programa, velocidade do processamento dos documentos, ferramentas disponíveis para trabalhar com a imagem do documento.

## 2 INDEXAÇÃO

A palavra “indexar”, significa “organizar em forma de índice”, do Latim *index*, “dedo indicador, apontador, lista”, que gerou **indicare**, “apontar” (ORIGEM DA PALAVRA, 2021).

Encontramos em registros antigos ato de indicar a localização de um documento, na Mesopotâmia no século II a.C., eram grafadas em tábuas de argila um resumo de cada livro (FUJITA, SILVA, 2004). Em vista disso, é nítida a existência de uma preocupação da humanidade diante a organização e recuperação de documentos, mesmo em tempos remotos e utilizando-se métodos rudimentares. Somente no século XIX iniciou-se o aprimoramento das técnicas de indexação em decorrência da massa documental. Um grande exemplo deste feito é a Encyclopaedia Britânica, considerada o maior aperfeiçoamento do século. Neste contexto, é perceptível que o século XIX operou grandes mudanças para a indexação, o que conseqüentemente impulsionou a criação de novos métodos de organização da informação já no século XX, como a Classificação Decimal Universal (CDU), Classificação Decimal de Dewey (CDD), os princípios de pré-coordenação de La Fontaine e Paul Otlet, adicionando na CDU com o acréscimo de dois pontos para relacionar assuntos de classes diferentes. (FUJITA; SILVA, 2004).

Sabe-se que a indexação envolve duas importantes etapas: análise conceitual e tradução. A análise conceitual, resumidamente, define qual é o assunto abordado em um determinado documento, para isso, o indexador precisa questionar sobre o que se trata, o motivo da inclusão daquele material no acervo, e qual o interesse dos usuários; a tradução transforma a análise documental em linguagem de indexação (LANCASTER, 1933).

Lancaster (1933) reforça sobre a atenção para a recuperação posterior de documentos, estando relacionado um item à sua relevância em uma determinada enunciação. Em outras palavras, consiste em estabelecer quais resultados aparecerão após a pesquisa de um termo, contudo, é válido ressaltar que

esse fator está diretamente ligado com a precisão da busca dependendo da exaustividade e especificidade do termo utilizado.

Desta forma, a indexação extrai os termos capazes de condensar a ideia principal do documento, podendo ser uma seleção mais precisa, onde reduz-se a quantidade de termos, ou, mais exaustiva, gerando um resultado mais genérico. Para este processo, é imprescindível que o bibliotecário conheça as políticas da instituição, pois estas são responsáveis por criar diretrizes para indexação.

Em decorrência da necessidade de indexar cada vez mais documentos em menos tempo, as atividades pertinentes à organização do conhecimento, incluindo os métodos de indexação, ganharam novos recursos. No Brasil, de acordo com Santos (2009) citado por Corrêa e Lapa (2014), a indexação automática teve início no final da década de 1960, com a utilização de um programa chamado KWIC utilizado no desenvolvimento de índices publicados pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).

Sabe-se que este processo se iniciou em decorrência do volume de materiais a serem indexados. De acordo com Robredo (1982), citado por Celerino e Tavares (2018, p.10):

O processo de indexação deve comparar cada palavra presente no documento com um conjunto de palavras de significado vazio (previamente selecionadas), e isso irá ajudar o processo de eliminação, já que o programa de computador irá considerar as palavras restantes como palavras significativas.

Dessa forma, o sistema irá compreender quais terminologias devem estar associadas entre si ou não, comparando os seus significados dentro de uma espécie de tesouro.

Percebe-se que essa nova ferramenta foi capaz de conferir um novo caminho para a indexação, pois, no momento atual, as ferramentas de classificação e recuperação da informação já são utilizadas nas redes sociais, mesmo que com propósitos diferentes, como o uso das “tags”, por exemplo.

O ato de taggear um post na internet tem como objetivo impulsionar um determinado conteúdo, sendo que, posteriormente, ele poderá ser recuperado por alguém quando pesquisado. Como exemplo, uma empresa que vende artigos esportivos ao compartilhar a foto de uma nova campanha pode utilizar as “tags” para impulsionar sua nova propaganda, como: esportes, saúde, fitness, exercícios físicos, etc. Ou seja, a empresa está criando essas classificações, indexando o seu produto nas redes sociais objetivando sua venda.

Esta atividade, nos remete a indexação colaborativa ou baseada no usuário deve ser usada em um sistema de recuperação dinâmica que permita aos usuários contribuírem com o processo de indexação e recuperação. (BROWN *et al.*, 1996). A classificação social, segundo Matusiak (2006), citado por Marques (2018, p. 273) traz a linguagem do usuário, sua perspectiva, experiência, e oferece oportunidades para o envolvimento do usuário.

Este processo já é alvo de estudos, os especialistas o chamam de folksonomia. De acordo com o site Biblioteconomia Digital (2011), folksonomia “[...] é uma forma relacional de categorizar e classificar informações disponíveis na web, sejam elas representadas por meio de textos, imagens, áudio, vídeo, ou qualquer outro formato.”

Portanto, a folksonomia é um fenômeno que surgiu na internet com objetivo de encontrar um modo de organização e recuperação da informação. De acordo com Santos (2017, p.109):

Folksonomia trata-se do resultado do processo de etiquetagem dos recursos disponíveis na web, ou seja, os próprios usuários são capazes de classificar os documentos, [...]. Desta forma, pode-se dizer que o “recuperar” da informação está intimamente ligado à forma de organização. E que a Folksonomia se trata de uma ferramenta de classificação da informação orientada pelos usuários das informações e documentos.

Essas mudanças trazidas pela internet e as redes sociais é um fenômeno observado por muitas instituições, estas, perceberam que o usuário pode colaborar de alguma forma com a plataforma, como no caso do FamilySearch.

A partir desta ideia é plausível considerar que possa existir benefícios do trabalho colaborativo para biblioteconomia e ciência da informação, e, em especial, no processo de indexação, uma vez que este é um trabalho que carece mais do que processos técnicos, mas também um olhar meticuloso de

quem o executa capaz de compreender qual a melhor forma de o fazê-lo levando em consideração diversos fatores, como as coleções, o propósito da instituição, o perfil do usuário, etc.

Dada esta questão, não se trata apenas do papel do bibliotecário em executar um trabalho de indexação corretamente, mas também, de toda uma estrutura social responsável pela forma como as informações são dispostas, podendo induzir o profissional da informação tomar decisões favoráveis a este sistema excludente. Ou seja, os próprios recursos dispostos para indexação podem induzir o bibliotecário a executar o seu trabalho inadequadamente.

### **3 A HISTÓRIA DA IGREJA JESUS CRISTO DOS SANTOS DOS ÚLTIMOS DIAS**

Segundo site institucional, a Igreja de Jesus Cristo dos Santos dos Últimos Dias possui atualmente 16 milhões de membros, tendo mais de 30 mil congregações localizadas em aproximadamente de 160 países (FAMILYSEARCH, 2021).

A igreja mantém suas atividades consolidando uma mentalidade corporativa, onde parte de seu sustento vem de empresas como supermercados e padarias onde os colaboradores são voluntários. Além disso, investem na tradução do evangelho, disponível em mais de 110 línguas, e, mais de 400 mil jovens e adultos estão inscritos em escolas de educação religiosa.

De acordo com o Livro dos Mórmons (2006), a jornada da igreja iniciou-se com Joseph Smith Junior, em 1820, quando tinha apenas 14 anos, estava em um bosque localizado em Nova York nos Estados Unidos fazendo uma oração pela primeira vez e recebeu uma visita de Jesus Cristo. Durante essa visita, Joseph teria perguntado a qual igreja Jesus Cristo se filiou, e ele teria respondido que nenhuma delas é a verdadeira igreja de Cristo, foi então, que ele ficou encarregado de restaurar a fé cristã.

O interesse genealógico da Igreja iniciou-se pela prática de batismos<sup>1</sup> dos entes queridos já falecidos. Segundo a doutrina essa atividade permite que aqueles que já morreram e não conheceram o mormonismo, tenham suas almas salvas. É principalmente essa causa que move a igreja ao trabalho genealógico.

Acredita-se que as pessoas que não foram batizadas em vida, podem ser batizadas pelos seus descendentes, o que é chamado de batismo por procuração, isso significa que alguém vivo recebe ordenanças em nome do ente falecido, o que consideram uma grande responsabilidade familiar.

Com isso, a igreja fundou a maior **biblioteca genealógica** do mundo localizada em Salt Lake City, a Biblioteca de História da Família. Atualmente, esta biblioteca colabora não somente com os membros da igreja, mas com todos aqueles que possuem algum tipo de interesse na pesquisa genealógica para obter acesso a documentos de seus antepassados. De acordo com Art Johnson, gerente de recrutamento do departamento de história da família, o trabalho executado pelo FamilySearch é uma espécie de corrida contra o tempo, pois, diariamente, registros genealógicos ao redor do mundo estão sendo perdidos. Para ele, os missionários realizam um trabalho essencial para a preservação dos documentos de maneira efetiva (A IGREJA DE JESUS CRISTO, 2021).

Para intensificar a proteção dos documentos, no ano de 1964 a igreja fundou o primeiro centro de história da família local (o que nos dias de hoje desdobrou-se em mais de 5 mil centros espalhados ao redor do mundo), e em 1966, construiu-se um cofre em uma montanha para proteção dos registros, localizado na Cordilheira Wasatch, nos Estados Unidos. Esta construção entrou para a lista dos dez cofres mais seguros do mundo de acordo com a revista *Hypescience* (2013). Toda esta estrutura foi construída pensando-se no crescimento constante dos registros, permitindo que posteriormente estes ficassem disponíveis da plataforma FamilySearch, fundada no ano de 1999, tornando-se o maior site genealógico do mundo atingindo mais de 1 bilhão de registros no ano de 2014.

A igreja ainda julga a plataforma FamilySearch como um “milagre moderno”, porque a partir de pequenas informações, como o nome do antepassado, é possível localizar informações completas. Ainda os dados fornecidos pela instituição, apontam que um dos avanços mais significativos para a plataforma ocorreu no ano de 2005, quando foram adquiridos 15 *scanners* de alta velocidade.

Em notícia apresentada pelo Church News, em 2020, o FamilySearch é capaz de acrescentar um novo registro digital em um período de 24 horas, o que demorava em média de seis a nove meses no ano

---

<sup>1</sup> Para essa religião o batismo deve ser feito por quem tem autoridade do sacerdócio realizando a “imersão”, neste processo significa que a pessoa encerrará uma vida pecadora e iniciará uma nova etapa dedicada à nova religião (Sempre Fiéis: Tópicos do Evangelho, 2021).

de 2019. Como resultado, mais documentos estão sendo indexados em menos tempo, permitindo com que cada vez mais pessoas encontrem informações de seus antepassados, independentemente se há ou não um propósito religioso.

#### 4 INDEXAÇÃO DE INFORMAÇÕES VIA WEB NO FAMILYSEARCH

Inicialmente o usuário precisa fazer um cadastro gratuitamente utilizando um endereço e-mail já existente para criar um usuário e senha. Ao clicar no menu superior direito da tela “criar conta” onde ele será direcionado para uma página onde deverá preencher seus dados pessoais e concordar com os termos de uso do site.(Figura 1).

**Figura 1** - Criação de um perfil no FamilySearch

Fonte: FamilySearch, 2021.

Ao selecionar opção de indexação logo no topo da tela, o usuário encontrará no primeiro menu a Visão Geral da indexação, onde é possível realizar um tour e receber as instruções para iniciar um projeto de indexação (Figura 2).

**Figura 2** - Página Inicial

Fonte: FamilySearch, 2021.

Ao clicar no botão **experimente** localizado no canto inferior direito da tela, o site redireciona o usuário para uma tela de demonstração com o objetivo de compreender como funciona o processo de indexação via web. Logo, algumas orientações básicas de como começar a usar o sistema aparecem em destaque na tela. (Figura 3).



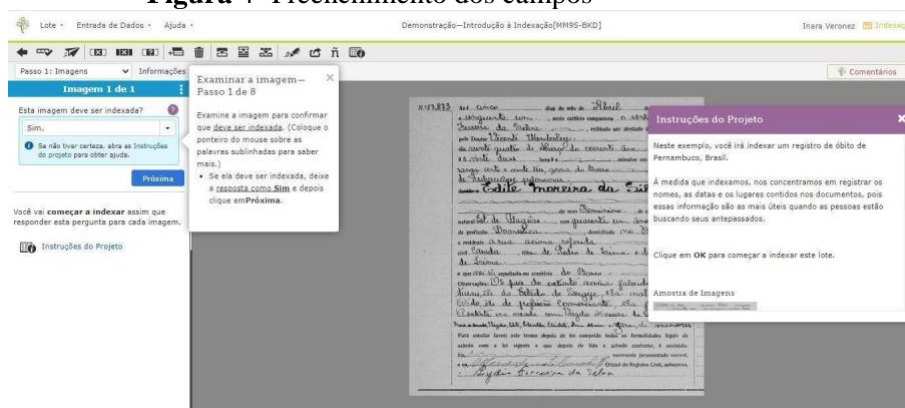
Figura 3 - Software de Indexação



Fonte: FamilySearch, 2021.

Quando selecionada a opção “Começar Demonstração” o usuário terá acesso ao sistema, onde ele passará por um total de oito etapas, o primeiro passo consiste em uma análise do documento. Antes de iniciar este primeiro passo, é orientado que o voluntário conheça o projeto, para isso do lado esquerdo da tela, é possível visualizar as instruções do projeto (Figura 4)

Figura 4- Preenchimento dos campos



Fonte: FamilySearch

A primeira pergunta é "Esta imagem deve ser indexada?", e as opções de resposta são “Sim” e "Não. Sem dados a serem extraídos". Caso exista dúvida na resposta. Ao clicar no ponto de interrogação, há orientações para responder a pergunta. Ao selecionar a resposta correta, clicando em “próxima” é possível seguir para o passo de número dois (Figura 5).

Figura 5 - Informações de apoio



Fonte: FamilySearch

Nos próximos passos, o usuário preencherá o campo com todas as informações do documento, no caso acima, trata-se de uma certidão de óbito, portanto, deve-se preencher obrigatoriamente todos os campos marcados com asterisco, como: nome, sobrenome, data da morte, idade. Caso haja outros dados, como nome dos pais, cônjuge, etc., é possível também incluir (Figura 6)

**Figura 6 - Detalhamento passo a passo**

The screenshot shows the FamilySearch interface during the data entry phase. On the left, a sidebar contains a form with fields for: 'Nomes do(a) Falecido(a) \*', 'Sobrenomes do(a) Falecido(a) \*', 'Sexo do(a) Falecido(a) \*', 'Dia da Morte \*', 'Mês da Morte \*', 'Ano da Morte \*', 'Idade do(a) Falecido(a) \*', and 'Nomes do Pai'. The main area displays a handwritten document with a pop-up window titled 'A indexação de informações - Passo 2 de 8'. The pop-up contains the following text: 'Indexar as informações da imagem, conforme solicitado em cada campo. Se você tiver dúvidas, consulte o ajuda.com o campo acima.' and a bullet point: '• Digite o nome da pessoa. Para continuar, clique em Tab.'

Fonte: FamilySearch, 2021.

Finalizando esta etapa, é possível enviar o registro clicando no símbolo do avião de papel na parte superior da tela, a seguir, uma mensagem de verificação de qualidade do trabalho aparecerá na tela. A verificação automática realiza apontamentos sobre o resultado da indexação realizada, definindo a necessidade de acréscimos ou correções. (Figura 7)

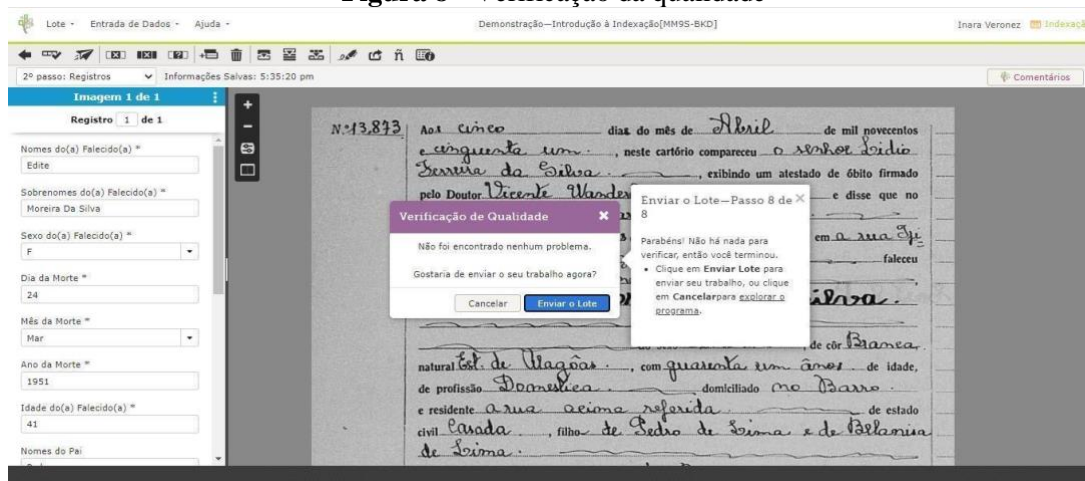
**Figura 7- Finalização do trabalho**

The screenshot shows the FamilySearch interface with a confirmation dialog box in the center. The dialog box has a yellow background and contains the text: 'Uma Verificação de Qualidade tem que ser feita x antes do envio deste lote. Deseja iniciar a Verificação de Qualidade agora?'. Below the text are two buttons: 'Iniciar verificação de Qualidade' and 'Cancelar'. The background shows the same data entry form and document as in Figure 6.

Fonte: FamilySearch, 2021.

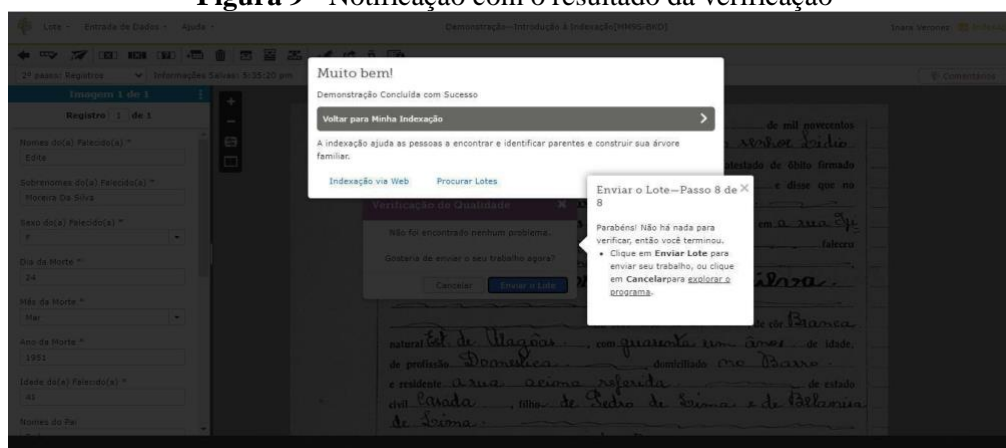
Após a iniciação da verificação de qualidade realizada automaticamente pelo programa, caso não haja erros, será possível prosseguir e enviar o lote, seguido de confirmação (Figura 8).

Figura 8 - Verificação da qualidade



Fonte: FamilySearch, 2021.

Figura 9 - Notificação com o resultado da verificação



Fonte: FamilySearch, 2021.

A imagem final informa que a indexação foi feita com sucesso, sendo assim o usuário pode passar para a próxima etapa.

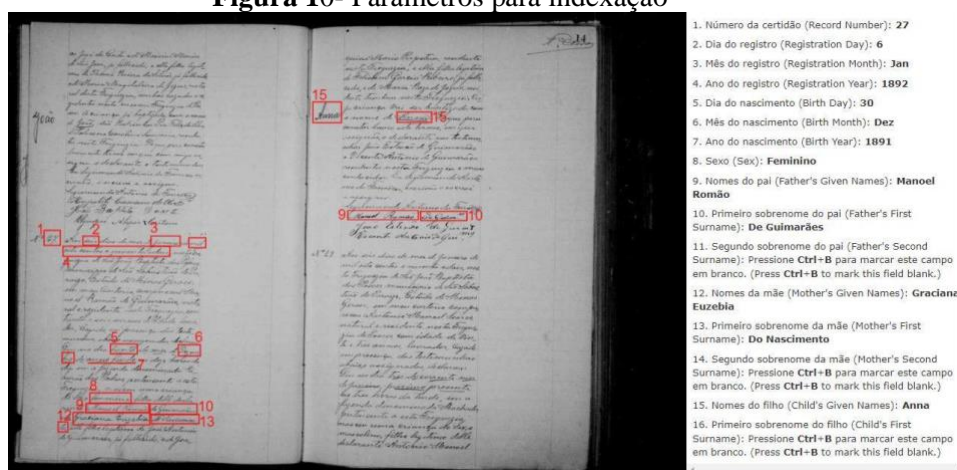
A partir da experiência com a plataforma, foi possível identificar alguns aspectos e constatar os seguintes pontos positivos a respeito da ferramenta: o software é simples, permitindo que o usuário saiba com facilidade manuseá-lo; o sistema possui um rápido processamento e carregamento de imagens em alta qualidade; possui ferramentas para “manuseio” das imagens, sendo possível ampliá-las, mudá-las de posição e configurar o brilho; é possível compartilhar o documento que está sendo trabalhado a fim de solicitar ajuda para indexá-lo; o sistema possui opção para filtrar o tipo de documento que o usuário deseja trabalhar, sendo possível selecionar o nível de dificuldade.

A respeito dos pontos negativos identificados na ferramenta: as lacunas para preenchimento das informações são restritas, pois, há um limite no tipo de informações a serem inseridas; os menus disponibilizados na parte superior esquerda da tela são pouco esclarecedores, os botões de ações são representados com imagens e não textos, o que em primeiro contato pode parecer confuso.

A respeito dos parâmetros utilizados pela plataforma para indexação, antes de iniciar qualquer projeto, o usuário recebe um passo a passo após escolher o material que deseja trabalhar, onde há links de artigos para ajudar o voluntário a conhecer melhor o documento. Como na imagem a seguir, estão selecionados os campos onde aparecem os dados utilizados para indexação do documento com uma explicação à direita da tela, especificando sobre o que se trata cada lacuna (Figura 10).



Figura 10- Parâmetros para indexação



Fonte: FamilySearch, 2021.

As instruções ainda abordam sobre o uso adequado de caracteres internacionais (aqueles não existentes no teclado), ajuda com caligrafia, como trabalhar com informações corrigidas, riscadas ou ilegíveis.

Para analisar os recursos da plataforma utilizou-se um estudo realizado por Borges e Lima (2015) que avaliou o desempenho de softwares de indexação automática pontuando os principais critérios para o desenvolvimento de um software de indexação. Os critérios utilizados por eles, levaram em consideração os seguintes aspectos: “a semântica e a sintaxe; a Linguística computacional e o Tratamento de documentos textuais para fins de recuperação da informação”.

A metodologia utilizada pelos pesquisadores dividiu o trabalho em duas etapas, a primeira consistiu em leitura de artigos técnicos-científicos de indexação automática que continham resultados de experimentos com softwares de indexação automática, a amostra selecionada foi composta por 103 pesquisas, entre elas, artigos nacionais e internacionais entre os anos de 1950 e 2008. Os pesquisadores selecionaram os seguintes critérios elencados pelos autores dos artigos: (1) Nome do critério conforme definido pelo autor; (2) Objetivo do critério; (3) Descrição Do critério; (4) Fontes de identificação; (5) Análise do critério mediante elaboração de parágrafo síntese indicando aplicações, vantagens e desvantagens observadas pelo autor.

Na segunda etapa, houve uma análise de combinação dos critérios, fazendo-se um recorte de 12 textos a partir da primeira etapa, e novos critérios de análise foram incluídos (BORGES; LIMA, p. 53, 2015): (1) Título da pesquisa; (2) Objetivos da pesquisa; (3) Nome do pesquisador autor responsável; (4) Período De realização do trabalho; (5) Localidade de realização/aplicação da pesquisa; (6) Listagem dos critérios utilizados; (7) Software Utilizados/ou desenvolvidos na pesquisa; (8) Comparação com indexação manual; (9) Registros de métodos comparativos entre os critérios mencionados; (10) Tipo de documento analisado; (11) Área do conhecimento em que se contextualiza o trabalho; (12) Identificação dos resultados como satisfatórios ou insatisfatórios de acordo com análise dos próprios autores pesquisadores e (13) Numeração do texto conforme amostra de estudo nº 1 (Borges e Lima, p. 53, 2015).

Ao final deste estudo, os pesquisadores concluíram há existência de oito critérios necessários para o desenvolvimento de softwares de indexação automática, sendo eles: Formatação de frases-termo (critério que permite a formação de novos termos a partir de termos muito abrangentes); Frequência absoluta de ocorrência da palavra no texto (quantas vezes um termo se repete em um texto); Identificação de palavras (análise da sequência de caracteres que formam a palavra); Identificação de radicais de palavras (palavras que são da mesma família); Lista de palavras proibidas / Palavras proibidas (pertinente aos termos que não devem ser selecionados para indexação do documento); Peso numérico (quantas vezes o termo aparece no documento); Vocabulário semântico / vocabulário de cabeçalhos conceituais / Tesouro (análise de cabeçalhos conceituais que aparecem nos títulos de artigos que possuem vocabulário semântico).

Comparando os critérios desta pesquisa à ferramenta de indexação do FamilySearch, percebe-se que o processo de indexação realizado na plataforma se aproxima muito mais da **indexação manual** do que automática, pois o voluntário precisa analisar o conteúdo e preencher os campos disponíveis.

Para Lancaster (1968) existem alguns problemas que a indexação manual pode trazer, como perguntas exaustivas ou específicas demais, falta de especificidade, exclusão de termos importantes, falta de vínculo entre o usuário e sistema.

Diante deste pensamento, é intrigante como a plataforma FamilySearch é uma referência em indexação apesar de não utilizar recursos automáticos, portanto, a questão levantada é: por qual/ quais motivo(s) a indexação é realizada manualmente? A primeira possibilidade está atrelada a idade dos documentos trabalhados, considerando que determinados documentos possuem séculos de idade, o que compromete a qualidade dos materiais dispostos para leitura, além da caligrafia contida nos mesmos, pertinente a época em que o documento foi produzido. A segunda, poderia ser a dificuldade atrelada ao desenvolvimento de um software capaz de identificar caligrafias (não mais utilizadas) em documentos, visto que as letras sofrem alteração com o passar dos anos, além da questão da condição física do material, onde o documento pode apresentar um estado de deterioração mais avançado.

Outro aspecto observado, é que esta tecnologia de reconhecimento de caracteres em imagens é conhecida como OCR (Optical Character Recognition), em português Reconhecimento Ótico de Caracteres. Essa ferramenta é capaz de executar a leitura de arquivos de imagem, datilografados, escritos a mão ou impressos. (Wikipédia, 2021). Para utilizar deste recurso, o documento precisa estar escaneado com qualidade mínima de 300dpi<sup>2</sup>, e a qualidade do documento original precisa ser boa (Sesín, 2017).

**Figura 11** - Alfabeto

A		a	
B		b	
C		c	
D		d	
E		e	
F		f	
G		g	
H		h	
I		i	
J		j	
L		l	
M		m	
N		n	
O		o	
P		p	
Q		q	
R		r	
S		s	
T		t	
U		u	
V		v	
X		x	
Y	Basically the same as the lower case y.	y	
Z	Basically the same as the lower case z.	z	

**Fonte:** FamilySearch, 2021.

Na plataforma FamilySearch, é como se os usuários voluntários fizessem o trabalho deste programa. Contudo, se para os próprios softwares não é fácil o reconhecimento de determinadas caligrafias, os voluntários possivelmente terão dificuldades de análise. Justamente por este motivo, há uma estrutura de apoio aos usuários, que inclui solicitação de ajuda com a caligrafia, além da disponibilização de um documento com um alfabeto de letras antigas<sup>3</sup>. Sabe-se que o trabalho de indexar atravessa questões de caligrafia e tecnologia, pois, o conhecimento prévio, os vieses ideológicos, a posição de poder que o indivíduo ocupa na sociedade, sua formação, treinamento, compreensão e interpretação individual de cada texto irá interferir no processo final. Entende-se então, que além do

<sup>2</sup> DPI (*Dots per Inch*) em português, pontos por polegada é uma medida que indica o número de pontos individuais que existem dentro de uma polegada na superfície da imagem. Ou seja, quanto mais pontos individuais por polegada, maior será a qualidade de resolução da imagem (WIKIPÉDIA, 2013).

<sup>3</sup> Ver figura 14.

desenvolvimento de um sistema para a realização da indexação, o trabalho da plataforma pode ser demorado e complexo dependendo da quantidade de lotes, qualidade da conservação do documento, resolução da imagem, quantidade de informações disponíveis e grafia utilizada.

Atualmente o FamilySearch oferece três tipos de suporte para indexação, o tutorial interativo realizado acima, ajuda e treinamento, que consiste em um menu com as perguntas mais frequentes, e de ajuda com a caligrafia. Mas, caso necessário, é possível entrar em contato com algum voluntário para pedir ajuda via chat ou telefone.

Em entrevista com Beatriz Clemente, líder do setor de indexação, diz que há todo um cuidado com as informações colhidas para indexação, pois é necessária auditoria dos documentos, como nos casos de catálogos com muitos volumes ou uma cidade que desapareceu. Paralelamente a este trabalho, existe uma revisão dos municípios, para verificação de todo o histórico da cidade.

Essa associação de termos colabora com todo o processo de pesquisa, por definição da Associação Brasileira de Normas Técnicas (1992), indexar é o “[...] ato de identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos e que constituem uma linguagem de indexação”. Dessa forma, o vocabulário escolhido poderá aumentar ou diminuir as chances de recuperação do documento.

Podemos levar em consideração que este processo de estudo e análise dos documentos e de suas respectivas origens, abordando o contexto histórico de cada município, contribui com as atividades de indexação, pois estabelecem uma direção a ser seguida. De acordo com Sousa e Fujita (2014, p. 32) “[...] o processo de indexação é uma atividade complexa e as normas e políticas são suportes, com os quais o profissional pode ter amparo na execução da tarefa, ao invés de adotar a perspectiva do bom senso”.

Ainda no relato apresentado pela líder de indexação, ela apontou que atualmente não há bibliotecários trabalhando em seu grupo e não soube informar se em demais setores existem bibliotecários voluntários. Sem dúvidas, um profissional da informação poderia orientar com prontidão os grupos de indexação, acelerando o processo de disponibilização dos materiais digitalizados e consequentemente reduzindo o volume da massa documental recebida.

Por este motivo, é perceptível que a auditoria dos documentos e a estrutura do software de indexação do FamilySearch colabora com o processo realizado pelo voluntário, pois ele recebe um material previamente lapidado, e este tratamento prévio promove uma direção a ser seguida. Portanto, mesmo que o voluntário não possua formação voltada para as áreas de Biblioteconomia e Ciência da Informação é possível indexar.

Após a indexação do documento, outro voluntário mais experiente fica responsável pela verificação do mesmo, analisando se as informações adicionadas na indexação são fiéis ao documento. Em entrevista, foi apontado que este tipo de função é destinado somente aos voluntários que possuem mais de mil indexações corretas. Por este motivo, a familiaridade entre o voluntário e o documento é um quesito importante.

## **5 CONSIDERAÇÕES FINAIS**

A partir dos dados colhidos e analisados neste trabalho, é possível afirmar que o FamilySearch ocupa um papel de destaque na área de genealogia, sobretudo, acerca do trabalho de indexação. Por este motivo, é enriquecedor conhecer e explorar projetos bem sucedidos a fim de promover o aprendizado por meio de observação e reflexão.

Os objetivos do trabalho se pautaram em uma análise dos serviços oferecidos pela plataforma viabilizando a identificação de pontos fortes e fracos, juntamente com os seus propósitos. Em suma, para esta etapa foi preciso a realização de testes da ferramenta de indexação a partir da criação de um perfil, bem como pesquisa bibliográfica e entrevista com um profissional da equipe de indexação. A partir deste processo foi possível agrupar material suficiente para uma análise da plataforma, levando em consideração o aspecto tecnológico e os pontos importantes para a indexação de documentos a partir de critérios estabelecidos por estudiosos e pesquisadores da área. Portanto, foi plausível estabelecer uma comparação entre os estudos pertinentes a indexação e a plataforma.

Das etapas pertinentes a indexação, há um processo de auditoria realizado pelos funcionários onde identifica-se a origem dos documentos e, em seguida é feita uma análise das páginas que não são necessárias para indexação, garantindo que o material seja digitalizado e disponibilizado o mais rápido possível. Em seguida, os projetos de indexação ficam disponíveis e os voluntários podem começar a

escolher em quantos desejam trabalhar, selecionar o nível de dificuldade e caso necessário, podem acessar artigos informativos ou entrar em contato com o suporte oferecido. Após análise foi possível concluir que a plataforma possui funcionamento otimizado e o software de indexação pode ser facilmente utilizado por todos.

Outro aspecto analisado levou em consideração os números significativos de indexação, buscando compreender como o FamilySearch tornou-se o maior site do ramo de genealogia mesmo contando com o apoio de terceiros. Não há uma resposta concreta para esta pergunta, mas as possibilidades mais convincentes apontam que isto ocorre por duas razões: o pensamento corporativo da igreja e a permuta de informações entre a plataforma e os usuários.

A estrutura similar de uma empresa, marcada principalmente pela divisão de grupos e tarefas, além dos líderes de cada setor, permitem a identificação de metas e resultados. Isso é perceptível principalmente pelos dados comparativos do trabalho executado ao longo dos anos, eles incluem as mudanças ocorridas na organização associada de seus resultados, como os números de indexação, os recursos aplicados, como máquinas de digitalização e melhorias no programa de indexação, por exemplo.

Outro ponto a ser citado é a troca de benefícios entre plataforma e os usuários, pois, ao navegar é possível realizar uma tarefa de indexação, e, independente do usuário se voluntariar ou não, ele pode acessar os documentos. Por isso, apesar do site oferecer um serviço para estruturar árvores genealógicas, o usuário pode realizá-lo sozinho, podendo acessar o que deseja na plataforma, assim, o processo torna-se mais atrativo para quem está realizando alguma pesquisa.

Contudo, esta pesquisa pode colaborar com a área da biblioteconomia e ciência da informação, pois com ela foi possível analisar as etapas do processo de indexação e os recursos necessários para que sejam produzidas, especialmente, no modo via web, visto que esta prática torna-se mais frequente em consequência das novas tecnologias. Em consequência, é esperado que a partir dos levantamentos feitos esta pesquisa possa levantar interesse nas áreas de pesquisa pertinentes a indexação via web e o bom aproveitamento do trabalho voluntário para atividades pertinentes à organização do conhecimento.

## REFERÊNCIAS

ACHILLES, D.; SOUSA, B. P.; SABBAG, D. M. A. Interação pelo catálogo on-line dos processos de seleção e representação temática: exploração da literatura lésbica enquanto assunto. In: ENCONTRO NACIONAL DE PESQUISA EM CIENCIA DA INFORMACAO, 20., 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/123829> . Acesso em: 28 set. 2021.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 12676: **Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação**. Rio de Janeiro, 1992.

BORGES, G. S. B.; LIMA, G. N. B. O. Desenvolvimento de softwares de indexação automática: breve avaliação dos principais critérios. **Informação & Tecnologia**, v. 2, n. 2, p. 49-70, 2015. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/41528>

BROWN, Pauline et al. (1996) - The democratic indexing of images. **New Review of Hypermedia and Multimedia** [Em linha]. Vol. 2, Nº. 1, p. 107-120. [Consult. 8 set. 2018]. Disponível na Internet: .

CHURCH OF JESUS CHRIST. Os Computadores e a Pesquisa de História da Família. In: **Introdução a História da Família: Manual do Aluno**. [S. l.], 2012. Disponível em: <https://referenciabibliografica.net/a/pt-br/ref/abnt> Acesso em: 13 set. 2021.

DICIONARIO Online de Português. [S. l.: s. n.], 2021. Disponível em: <https://www.dicio.com.br/cognicao/> Acesso em: 3 jul. 2021.

D'ORNELAS, Stephanie. **10 cofres muito bem guardados**. Disponível em: <https://hypescience.com/10-cofres-muito-bem-guardados/> . Acesso em: 18 out. 2021.



- FERREIRA, Luciana; MARQUES, Luana Farias Sales. Indexação colaborativa de acervo de imagens em acesso aberto: a experiência do IBGE. *Cadernos BAD* (Portugual), n. 1, p. 270-284, 2018. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/110070> . Acesso em: 29 set. 2021.
- FINO, C. N. Vygotsky e a Zona de Desenvolvimento Proximal (ZDP): três implicações pedagógicas. **Revista Portuguesa de Educação**, v. 14, nº 2, p. 273-291, 2001.
- FUJITA, Mariângela Spotti Lopes. A leitura documentária e o processo de compreensão do indexador: Memorial de investigação científica. In: FUJITA, Mariângela Spotti Lopes; NEVES, Dulce Amélia de Brito; DAL'EVEDOVE, Paula Regina. **Leitura documentária: estudos avançados para indexação**. Marília: Cultura Acadêmica, 2017. p. 15-50. ISBN 978-85-7983-917-7. Disponível em: <https://www.marilia.unesp.br/Home/Publicacoes/leitura-documentaria---ebook.pdf> Acesso em: 22 jun. 2021
- FUJITA, Mariângela Spotti Lopes. A política de indexação para representação e recuperação da informação. In: GIL-LEIVA, Isidoro; FUJITA, Mariângela Spotti Lopes (Eds.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. p.17-28.
- FUJITA, Mariângela Spotti Lopes; SILVA, Maria dos Remédios da. A prática de indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, Campinas, p. 133-161, maio/ago. 2004. Disponível em: <https://www.scielo.br/j/tinf/a/cNngvqQdWfBGrJtLSdLRKnP/?lang=pt&format=pdf> . Acesso em: 14 jun. 2021.
- IPE, M. Knowledge sharing in organizations: a conceptual framework. **Human Resource Development Review**, v. 2, n. 4, p. 337-359, Dec. 2003.
- LAPA, R. C.; CORRÊA, R. F. A indexação automática no brasil no âmbito da ciência da informação (1973-2012): indicadores bibliométricos de autoria. In: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 4, 2014. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/45382> . Acesso em: 28 set. 2021.
- MATHES, A. **Folksonomies**: cooperative classification and communication through shared metadata. 2004. Disponível em: <http://www.adammathes.com/academic/computer-mediatedcommunication/folksonomies.pdf> .
- PAGSEGURO. **Hashtags**: saiba como usar para impulsionar o perfil da sua empresa. [S. l.], 2021. Disponível em: <https://blog.pageseguro.uol.com.br/hashtags-saiba-como-usar-para-impulsionar-o-perfil-da-sua-empresa/> . Acesso em: 15 nov. 2021.
- PONTIFÍCIA UNIVERSIDADE CATÓLICA. **A Web dos significados**. Rio de Janeiro: PUC. p. 105-151, 2009. Disponível em: [https://www.maxwell.vrac.puc-rio.br/15543/15543\\_6.PDF](https://www.maxwell.vrac.puc-rio.br/15543/15543_6.PDF) . Acesso em: 16 ago. 2021.
- PORFÍRIO, Francisco. "O que é ética?"; Brasil Escola. Disponível em: <https://brasilecola.uol.com.br/o-que-e/o-que-e-sociologia/o-que-e-etica.htm> . Acesso em 10 de novembro de 2021.
- RICHARD, Ivan. **Biblioteca da Presidência trata golpe de 1964 como "vitória da revolução"**: Instituição diz que usa o termo por razão histórica. **Brasília**: Agência do Brasil, p. 1, 6 abr. 2014. Disponível em: <https://agenciabrasil.ebc.com.br/politica/noticia/2014-04/biblioteca-da-presidencia-trata-golpe-de-1964-como-vitoria-da-revolucao> . Acesso em: 10 out. 2021.
- ROY, Ludovic Le. **O que é OCR** (Reconhecimento Ótico de Caracteres) e como funciona?. 2017. Disponível em: <http://sesin.com.br/o-que-e-ocr-reconhecimento-otico-de-caracteres-e-como-funciona/>
- SANTOS, H. S.; OLIVEIRA, J. R.; LIMA, J. S. Folksonomia: representação da informação na web. **Bibliomar**, v. 16, n. 1, p. 105-114, 2017. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/126361> Acesso em: 30 set. 2021.



SANTOS, Raimunda Fernanda dos; FERNANDES CORRÊA, Renato. A Folksonomia e a Representação Colaborativa da Informação em Ambientes Digitais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, [s. l.], v.8, ed.1, jan./jun. 2015. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/369/369> . Acesso em: 25 ago. 2021.

FERREIRA, L.; MARQUES, L. F. S. Indexação colaborativa de acervo de imagens em acesso aberto:

SANTOS, Wigvan Junior Pereira dos. **Ética**. [S. l.], [2020]. Disponível em: <https://www.preparaenem.com/filosofia/etica.htm> Acesso em: 14 set. 2021.

SILVA, I. A. O; SANTOS, C. A. C. M. Indexação Colaborativa na Web. **CRB-8 Digital**, São Paulo, v. 4, n. 1, p. 90-98, abr. 2011. Disponível em: <http://revista.crb8.org.br>

TAVARES, W. Q.; CELERINO, V. G. A importância da bibliometria para a indexação automática. **Folha de Rosto**, v. 4, n. 2, p. 7-15, 2018. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/109164> Acesso em: 28 set. 2021.

THE CHURCH OF JESUS CHRIST OF LATTER-DAY SAINTS. **Overview**. [S. l.], 2021. Disponível em: <https://www.churchofjesuschrist.org/learn/about-us?lang=eng> Acesso em: 13 jul. 2021.

WALKER, Sydney. Como o FamilySearch está usando a tecnologia da computação para acelerar o processo de publicação de imagens de registros. **Church News**, [S. l.], p. 1, 30 mar. 2020. Disponível em: <https://www.thechurchnews.com/pt/membros/2020-03-30/historia-da-familia-familysearch-explorar-imagens-historicas-5432/> Acesso em: 15 nov. 2021.

WHATCOTT, Jace. **Missões de História da Família**. Disponível em: <https://br.igrejadejesuscristo.org/informacoes-basicas-sobre-as-missoes-de-historia-da-familia> Acesso em: 8 nov. 2021.

WIKIPEDIA. Alexander Cruden. [S. l.], 2021. Disponível em: [https://en.wikipedia.org/wiki/Alexander\\_Cruden](https://en.wikipedia.org/wiki/Alexander_Cruden) . Acesso em: 3 out. 2021.

WIKIPEDIA. **Folksonomia**. [S. l.], 2019. Disponível em: <https://pt.wikipedia.org/wiki/Folksonomia> . Acesso em: 30 nov. 2021.

WIKIPEDIA. **Lei de Zipf**. [S. l.], 2021. Disponível em: [https://pt.wikipedia.org/wiki/Lei\\_de\\_Zipf#cite\\_note-2](https://pt.wikipedia.org/wiki/Lei_de_Zipf#cite_note-2) .Acesso em: 30 nov. 2021.

WIKIPEDIA. **Reconhecimento ótico de caracteres**. [S. l.], 2021. Disponível em: [https://pt.wikipedia.org/wiki/Reconhecimento\\_%C3%B3tico\\_de\\_caracteres](https://pt.wikipedia.org/wiki/Reconhecimento_%C3%B3tico_de_caracteres) Acesso em: 23 nov. 2021.

ZAMBONI, Rita Costa Veiga; FRANCELIN, Marivalde Moacir. Garantia Cultural, Garantia Ética e Hospitalidade na Organização e Representação do Conhecimento. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17; 2016, Salvador. 2016. Disponível em: <http://www.ufpb.br/evento/index.php/enancib2016/enancib2016/paper/viewFile/4113/2382>.