

Banco de dados em *software* livre: uma solução para acesso e recuperação do patrimônio documental

Fabiana Fagundes Fontana

Centro de Excelência em Tecnologia Eletrônica Avançada, Porto Alegre, RS, Brasil
fabifagundesfontana@gmail.com

André Zanki Cordenonsi

Universidade Federal de Santa Maria, Centro de Ciências Sociais e Humanas, Departamento de Documentação, Santa Maria, RS, Brasil
andrezc@gmail.com

DOI: <http://10.26512/rici.v12.n1.2019.12438>

Recebido/Recibido/Received: 2017-10-06

Aceitado/Aceptado/Accepted: 2018-08-28

Resumo: Este trabalho apresenta um estudo sobre a criação de um banco de dados em *software* livre como solução para acesso e recuperação da informação do programa Pacto Nacional pela Alfabetização na Idade Certa, especificamente a documentação produzida e recebida que está sob custódia da coordenação geral do Pacto pela Universidade Federal de Santa Maria. Para a construção do banco de dados, utilizou-se o *software* livre BASE disponível tanto na suíte de *software* Libre Office, quanto no Open Office. O diagnóstico foi realizado por meio da observação direta da situação do acervo analógico e de parte do acervo digital, pois os documentos digitais estavam dispersos e muitos foram perdidos. Para os documentos em suporte papel, foi necessário criar representantes digitais para acessá-los no banco de dados, visando a facilitar o acesso e a recuperação da informação aos usuários. A digitalização foi feita de acordo com as recomendações do Conselho Nacional de Arquivos e com adoção de um vocabulário controlado para nomeação dos arquivos digitais. Foi aplicado um formulário *online* para o levantamento das necessidades informacionais dos usuários visando a construir um banco de dados que suprisse as suas necessidades informacionais. A partir dos resultados obtidos, foi possível construir um banco de dados em *software* livre, que reuniu, em um único local, a maior parte dos documentos, relatórios e apresentações elaboradas para as capacitações dos professores alfabetizadores integrantes do Pacto.

Palavras-chave: banco de dados. educação. recuperação da informação. *software* livre.

Database for free software: a solution to access and retrieval of documentary heritage

Abstract: This paper presents a study on the creation of a database on free software as a solution to access and retrieval of the National Pact for Literacy in the Age One, specifically produced and received documentation that is in the custody of the overall coordination of the Pact Universidade Federal de Santa Maria. To build the database we used the free software available BASE both in the LibreOffice office suite and in the Open Office. A diagnosis was made through direct observation of the situation of analog assets and part of the digital collection, for the digital documents were scattered and many were lost. For paper documents was necessary to create digital representatives to access them in the database, to facilitate access and retrieval of information to users, so the scan was done according to the Conselho Nacional de Arquivos recommendations and adoption of a vocabulary controlled for the indexing of digital files. A survey the information needs of users to build a database that met their information needs has been applied. From the results it was possible to build a database on free

software, which brought together in one location most of the documents, reports and presentations prepared for the training of literacy teachers of Pacto members.

Keywords: database. Education. free software. information retrieval.

Banco de datos en software libre: una solución para el acceso y la recuperación del patrimonio documental

Resumen: Este trabajo presenta un estudio sobre la creación de un banco de datos en software libre como solución para el acceso y recuperación de la información del programa Pacto Nacional por la Alfabetización en la Edad Cierta, específicamente la documentación producida y recibida que está bajo custodia de la coordinación general del Pacto por la Universidad Federal de Santa María. Para la construcción de la base de datos, se utilizó el software libre BASE disponible tanto en la suite de software Libre Office, como en el Open Office. El diagnóstico fue realizado por medio de la observación directa de la situación del acervo analógico y de parte del acervo digital, pues los documentos digitales estaban dispersos y muchos fueron perdidos. Para los documentos en soporte papel, fue necesario crear representantes digitales para acceder a ellos en la base de datos, para facilitar el acceso y la recuperación de la información a los usuarios. La digitalización fue hecha de acuerdo con las recomendaciones del Consejo Nacional de Archivos y con la adopción de un vocabulario controlado para el nombramiento de los archivos digitales. Se aplicó un formulario en línea para el levantamiento de las necesidades informativas de los usuarios con el fin de construir una base de datos que supediera sus necesidades informativas. A partir de los resultados obtenidos, fue posible construir un banco de datos en software libre, que reunió, en un único lugar, la mayor parte de los documentos, informes y presentaciones elaboradas para las capacitaciones de los profesores alfabetizadores integrantes del Pacto.

Palabras clave: base de datos. Educación. software libre. recuperación de la información.

1 Introdução

Esta pesquisa disserta sobre a utilização do *software* livre em código aberto - Base - para criação de banco de dados como instrumento para facilitar o acesso e a recuperação do patrimônio documental de um programa governamental, denominado Pacto Nacional pela Alfabetização na Idade Certa (PNAIC), especificamente a documentação produzida e recebida que está sob custódia da coordenação geral do Pacto pela Universidade Federal de Santa Maria.

A construção e a implementação deste banco de dados com *software* livre visa a dar longevidade ao acesso das informações depositadas no banco, pois o *software* não proprietário permite atualizações gratuitas e periódicas evitando, assim, a obsolescência tecnológica, já que as atualizações periódicas de versões compatíveis permitirão ao usuário acessar as informações ao longo dos anos.

As informações produzidas e recebidas por uma instituição ou pessoa no decorrer de suas atividades, são premissas fundamentais da arquivologia e, portanto, devem ser tratadas e organizadas para facilitar o acesso aos usuários e, em alguns casos, servir como fonte de pesquisa primária a usuários interessados em seu conteúdo informacional e de cunho científico.

Nessa perspectiva, essa pesquisa resulta de uma dissertação que estudou e criou um banco de dados com *software* Base, visando a facilitar o acesso e a recuperação da informação relevante à pesquisa e à gestão administrativa. A documentação analisada é custodiada por

uma coordenadoria geral sediada no Centro de Educação da UFSM (CE) e resulta de um programa governamental nacional denominado Pró-Letramento. Esse programa integra o Pacto Nacional pela Alfabetização na Idade Certa (PNAIC) e a documentação que será armazenada no banco de dados foi produzida e recebida para realizar as ações do Pacto através das ações do Pró-Letramento.

1.1 Patrimônio cultural e documental

Patrimônio cultural pode ser considerado tudo aquilo que é representativo para uma comunidade, que gera uma memória afetiva, que é considerado como legado de uma sociedade ou grupo social. O patrimônio pode ser material: como um objeto, um monumento, um conjunto documental, um prédio, um sítio arqueológico ou imaterial: uma música, uma festa popular, uma dança, uma comida enfim, algo que gere sentimento de pertencimento em uma sociedade como legado cultural, que retrate ou identifique aquele grupo social. Para Grunberg

Patrimônio Cultural não são somente aqueles bens que se herdam dos nossos antepassados. São também os que se produzem no presente como expressão de cada geração, nosso “Patrimônio Vivo”: artesanatos, utilização de plantas como alimentos e remédios, formas de trabalhar, plantar, cultivar e colher, pescar, construir moradias, meios de transporte, culinária, folguedos, expressões artísticas e religiosas, jogos etc. É com todo esse Patrimônio, material, imaterial, consagrado e não consagrado que podemos trabalhar num processo constante de conhecimento e descoberta. (GRUNBERG 2007, p .5).

Quanto aos bens patrimoniais consagrados, ou seja, protegidos por lei, temos em nossa lei maior, a Constituição Federal de 1988, que em seu artigo 216 determina:

Constituem patrimônio cultural brasileiro os bens de natureza material e imaterial, tomados individualmente ou em conjunto, portadores de referência à identidade, à ação, à memória dos diferentes grupos formadores da sociedade brasileira, nos quais incluem: [...] IV – as obras, objetos, documentos, edificações e demais espaços destinados às manifestações artístico-culturais. § 1º O Poder Público, com a colaboração da comunidade, promoverá e protegerá o patrimônio cultural brasileiro por meio de inventários, registros, vigilância, tombamento e desapropriação, e de outras formas de acautelamento e preservação. § 2º Cabem à administração pública, na forma da lei, a gestão da documentação governamental e as providências para franquear sua consulta a quantos dela necessitem. (BRASIL, *Constituição da República Federativa do Brasil*, 1998, p. 111).

Para proteger o patrimônio documental brasileiro temos a lei n. 8.159/1991, que dispõe sobre a política nacional de arquivos públicos e privados. Essa norma, em seu primeiro artigo, determina: "É dever do Poder Público a gestão documental e a proteção especial a

documentos de arquivos, como instrumento de apoio à administração, à cultura, ao desenvolvimento científico e como elementos de prova e informação".

El objetivo de organizar y conservar adecuadamente los archivos adquiere toda su dimensión ciudadana si asumimos que estos centros colaboran muy directamente em la Asunción, entre otros, de los valores de patrimonio público, memoria, identidad y conocimiento, los cuales, y no por casualidad, se asocian normalmente a las finalidades irrenunciables del moderno concepto de cultura. [...]. *Los archivos son patrimonio* a partir del momento en que el ser humano empieza a dejar constancia escrita em soportes duraderos de todo aquello que entiende relevante para llevar a cabo su actividad política o económica, em realidad no hace más que crear de manera espontánea los archivos, aunque inicialmente con la voluntad inequívoca de justificar sus derechos y propiedades y convertirlos em arsenales legitimadores del poder¹. (ALBERCH FUGUERAS *et al.*, 2001, p. 13).

Assim, os acervos documentais custodiados por arquivos, centros de documentação, instituições ou pessoas, são fontes de pesquisa primária e auxiliam a resguardar a memória da sociedade, permitindo contar a história brasileira baseada em provas documentais e, portanto, possuem valor inestimável para a sociedade.

Para compreender o conceito de patrimônio documental como fonte de pesquisa, primeiramente torna-se necessário compreender o conceito de documento arquivístico, aquele que possui potencial após completar seu ciclo de vida em tornar-se um patrimônio documental para a sociedade, além da instituição.

Documentos arquivísticos são aqueles produzidos e recebidos no decorrer das atividades de um órgão ou entidade, independentemente do suporte em que se apresentam, registram suas políticas, funções, procedimentos e decisões. Nesse sentido, constituem-se em documentos arquivísticos, que conferem aos órgãos e entidades a capacidade de: - conduzir as atividades de forma transparente, possibilitando a governança e o controle social das informações; - apoiar e documentar a elaboração de políticas e o processo de tomada de decisão; [...] – assegurar e documentar as atividades de pesquisa, desenvolvimento e inovação, bem como a pesquisa histórica; - manter a memória corporativa e coletiva. (CONARQ/CTDE, 2011, p. 17.)

Os documentos arquivísticos, servem como forma de lembrar, de perpetuar por meio de registro, ações e atividades sejam administrativas, de negócios ou pessoais, para Duranti e Thibodeau (2008, p. 449 *apud* Rondinelli, 2013, p. 253) o “documento arquivístico como um

¹ Tradução nossa: Com o objetivo de organizar e conservar adequadamente os arquivos adquirem toda sua dimensão de cidadania se nós assumimos que estes centros funcionam muito diretamente na ascensão, entre outros, dos valores de patrimônio público, memória, Identidade e conhecimento, os quais, não por casualidade, estão normalmente associadas com os objetivos indispensáveis da moderna concepção de cultura. [...]. Os arquivos são patrimônio a partir do momento que o ser humano começa a deixar escrito em suporte duradouro tudo o que entende relevante para exercer a sua atividade política ou econômica, em realidade os arquivos não nascem de forma espontânea, embora inicialmente com uma vontade clara de justificar os seus direitos e propriedades e convertidos em arsenais legitimadores do poder.

“meio de lembrança” deve ser alguma coisa que o produtor possa de fato manter, associar com outros documentos arquivísticos e, depois, chamar de volta”.

Esse “chamar de volta” tem a ver com o que os autores chamam de “função memorial” do documento arquivístico, a qual deve ser entendida como uma função inerente não somente ao documento de valor permanente, mas ao documento de arquivo em qualquer fase do seu ciclo de vida. (RONDINELLI, 2013, p. 253).

Nessa perspectiva, durante seu ciclo de vida, os documentos arquivísticos passam por três fases distintas: a fase corrente, a intermediária e a permanente. Esses três períodos são determinados por sua frequência e pela natureza de sua utilização. É o denominado Ciclo de Vida dos Documentos, também conhecido como Teoria das Três Idades do arquivo.

Nesse sentido, alguns documentos possuem um valor permanente embutido já desde seu nascimento como é o caso de documentos constituintes de instituições ou com alto potencial para pesquisa, como os dados resultantes de programas institucionais e projetos de pesquisa. Estes constituem-se em patrimônio documental, pois o valor secundário, tanto de prova como de pesquisa estão arraigados nesta documentação, assim, segundo Bellotto:

É nesta dualidade administração/história - ou melhor, valor administrativo/valor histórico - que se deve considerar a documentação estática ou inativa. Não há arquivos que sejam em essência, históricos, e todo papel administrativo, desde sua criação, tem, em potencial, um valor histórico. (BELLOTTO, 2006, p. 117).

Mesmo os documentos permanentes, chamados “inativos” tornam-se “ativos” novamente, mas com outro fim e não mais aqueles para o qual foram gerados, “ativam-se” como fonte primária de pesquisa.

Portanto, os documentos oriundos de pesquisa e de programas acadêmicos constituem-se fontes de pesquisa, mesmo na idade corrente. Assim, a gestão documental de dados de pesquisa, além de auxiliar na tomada de decisão, possibilitará que a documentação, que já possui as características de caráter permanente e de fonte de pesquisa, valorize-se como patrimônio documental, permitindo a outros pesquisadores “beberem em sua fonte” e ampliem sua pesquisa.

O reconhecimento do potencial informacional dos dados de pesquisa para a ciência contemporânea transforma a visão que os caracterizava como simples subprodutos dos processos de pesquisa [...]. Os pesquisadores, as instituições acadêmicas e as agências de fomento à pesquisa começam a compreender que esses dados, se devidamente tratados, preservados e gerenciados, podem constituir uma fonte inestimável de recursos informacionais para a pesquisa científica e para o ensino da ciência. (SAYÃO; SALES, 2015, p. 5).

Desse modo, projetos e programas governamentais, institucionais geram uma produção documental, bibliográfica e resultados científicos consideráveis como fonte de

pesquisa, pois podem contribuir como fonte informacional a muitos pesquisadores. Nesse giro, o compartilhamento de saberes é fundamental para a construção de novos conhecimentos. Como ilustração, citam-se outras iniciativas de utilização de banco de dados para preservação e acesso a informação que já foram construídas nas diferentes esferas governamentais, como, por exemplo, a criação do Centro Regional de Informação Documental da Justiça Federal da 2ª Região, citado por Rodrigues (2018) e a Memória da Justiça do Trabalho do Rio Grande do Sul, citado por Fausto (2013).

1.2 Gestão arquivística de documentos

O objeto de estudo da arquivística é a informação registrada, seja em que suporte for, como, por exemplo, o documento analógico em papel, o microfilme, o arquivo digital de áudio ou vídeo. Assim, a gestão da informação registrada só pode ser realizada se os documentos forem acessíveis e compreensíveis. Isso pode ser alcançado por meio da gestão arquivística de documentos que, como define o Conarq, é:

O conjunto de procedimentos e operações técnicas referentes à produção, tramitação, uso, avaliação e arquivamento de documentos arquivísticos em fase corrente e intermediária, visando sua eliminação ou recolhimento para guarda permanente. (Conarq/CTDE, 2011, p. 17).

Com a nova demanda por administrar dados resultantes de pesquisa científica, a gestão arquivística de documentos propicia aos pesquisadores uma fonte primária de informação de valor inestimável. A gestão agiliza estudos, possibilitando inclusive, dependendo do acervo, o acesso a pesquisas já realizadas e que estão disponíveis e organizadas em um conjunto documental.

O reconhecimento do potencial informacional dos dados de pesquisa para a ciência contemporânea transforma a visão que os caracterizava como simples subprodutos dos processos de pesquisa. Naquele contexto, os dados eram considerados somente na sua configuração final, sem considerar os seus ciclos de vida, versões e linhagens e, via de regra, eram descartados ou armazenados em mídias ou em servidores sem a devida gestão quando os projetos eram concluídos. Quase sempre eram tragados silenciosamente pelo tempo: pela obsolescência tecnológica e pela fragilidade das mídias digitais [...]. Nessa direção, uma gestão eficiente dos dados é fundamental para o desenvolvimento de pesquisas de alta qualidade e excelência. A gestão de dados cobre todos os aspectos relativos à manipulação, organização, documentação e agregação de valor, e tem um papel crucial como facilitador nos processos de compartilhamento dos dados, na garantia da sustentabilidade e acessibilidade dos dados em longo prazo. As ações e comprometimentos promovidos pela gestão, coletivamente, permitem que os dados de valor possam ser reusados em outros projetos ao longo do tempo e do espaço. (SAYÃO; SALES, 2015, p. 1).

Sayão e Sales mencionam a gestão de dados de pesquisa como imprescindíveis para organizar dados de pesquisa. Porém, como essa documentação é orgânica, ou seja, possui uma

relação com outras, e é gerada segundo uma função e para atender a certa necessidade, ela constitui-se, portanto, documentação arquivística. Desse modo, deve ser gerida de acordo com as premissas da arquivologia, ou seja, tem de ser gerida arquivisticamente.

A gestão arquivística de documentos agrega capacidade administrativa para a organização, entidade ou instituição, pois, quando o acervo está tratado e organizado, resulta em economia e eficácia, permitindo que a informação arquivística fixada, independente do suporte, e que seja recuperada mais rapidamente, subsidiando as ações da organização com decisões claras e seguras, reduzindo, assim, a incerteza como fator negativo neste processo.

A gestão arquivística de documentos pode incluir o uso de ferramentas tecnológicas e automatizadas disponíveis atualmente, desde um banco de dados até um sistema de informação complexo, cujo objetivo é armazenar, tratar, manipular e fornecer informações de forma ágil, facilitando as funções ou monitorando o trâmite de processos dentro de uma organização, desde sua criação até a destinação final.

Atualmente, a maioria das instituições possuem documentos em formatos híbridos, ou seja, tanto em formato em papel como no formato digital. Sendo assim, torna-se necessário digitalizar o que ainda está em papel para facilitar o acesso. A digitalização criará representantes digitais, tanto para acesso, como para preservação, visando torná-los acessíveis em logo prazo e permitindo seu acesso por meio de um banco de dados ou sistema informatizado que armazene e que dê acesso a esses documentos. Documentos que não estão sistematizados dificultam o acesso e conseqüentemente a geração de conhecimento.

Para Duranti e MacNeil (*apud* Rondinelli, 2005, p. 63) “um sistema de gerenciamento arquivístico de documentos integra um todo chamado sistema de documentos arquivísticos. Do qual também faz parte o sistema de preservação de documentos” que será regido por um conjunto de regras diferenciadas e consistentes visando a manutenção física e intelectual dos documentos arquivísticos de uso intermediário e permanente para que se tornem acessíveis ao longo do tempo.

Nessa perspectiva, é possível dar continuidade à cadeia de custódia, unindo um repositório digital a um banco de dados ou a um sistema informacional computadorizado de gestão. Isso permite que o patrimônio documental fique protegido durante todo seu ciclo de vida, inclusive após sua destinação final. Lembrando que farão parte dessa nova etapa somente os documentos que possuírem valor secundário e, portanto, que devam ser guardados permanentemente. Nessa etapa, extrapola-se a gestão e adentra-se em outro patamar: a preservação dos documentos permanentes, pela via de sua inclusão nos repositórios digitais confiáveis.

1.3 Digitalização de documentos

A digitalização de documentos é uma das estratégias mais eficazes quando se pretende viabilizar o acesso e a difusão de informações, além de contribuir para a preservação dos originais. O manuseio indevido, a poluição, os contaminantes biológicos (insetos, fungos, bactérias, entre outros) contribuem para a deterioração dos documentos, especialmente os documentos analógicos e fotografias, em suporte papel.

A digitalização, ao criar representantes digitais, possibilita restringir o manuseio dos documentos originais, além de possibilitar o acesso remoto a esses documentos, via sistemas informatizados *online* na web. Segundo Casanovas os documentos digitais podem ter duas origens: nato digitais ou convertidos pela digitalização, sendo assim:

Fueron creados em formato digital y están almacenados em algún soporte óptico o magnético, que está guardado em el centro de documentación o biblioteca. Fueron creados em formato digital y se encuentran disponibles em internet. El centro de documentación, la biblioteca o directamente el usuario no entra “em posesión” de ellos, sino que se tiene derecho de acceso mediante licencias o porque son de uso libre y gratuito. - Fueron creados em formato analógico y deben ser convertidos a formato digital mediante proceso de digitalización². (CASANOVAS, 2008, p. 85).

O Conselho Nacional de Arquivos (CONARQ) em sua publicação denominada – Recomendações para digitalização de documentos arquivísticos permanentes define digitalização como:

Um processo de conversão dos documentos arquivísticos em formato digital, que consiste em unidades de dados binários, denominadas de *bits* - que são 0 (zero) e 1 (um), agrupadas em conjuntos de 8 *bits* (*binary digit*) formando um *byte*, e com os quais os computadores criam, recebem, processam, transmitem e armazenam dados. [...]. No entanto, o produto dessa conversão não será igual ao original e não substitui o original que deve ser preservado. A digitalização, portanto, é dirigida ao acesso, difusão e preservação do acervo documental. (CONARQ, 2010, p. 6).

O termo conversão pode ser compreendido como a mudança de formato, do analógico para o digital, alterando, no caso dos documentos, o suporte que antes era o papel e agora passa a ser digital, com o uso dos dados binários. Mas cabe lembrar que a digitalização pode ocorrer em outros formatos analógicos, como as fitas magnéticas registradas em áudios, vídeos, entre outros.

² Tradução nossa: - Eles foram criados em formato digital e armazenados em algum suporte óptico ou magnético, que está guardado no centro de documentação ou biblioteca. - Eles foram criados digitalmente e estão disponíveis na internet. O centro de documentação, a biblioteca ou o usuário não entram diretamente “em posse” deles, mas, tem direito de acesso por meio de licenças ou porque são de utilização gratuita. - Foram criados em formato analógico e devem ser convertidos para o formato digital pelo processo de digitalização.

A principal dificuldade que os documentos analógicos enfrentam atualmente, na era das Tecnologias Digitais de Informação e Comunicação (TDIC) é a restrição de acesso. Se o usuário necessita de um documento em suporte físico, terá de ir até o local em que tal documento se encontra armazenado para acessá-lo. Além disso, há sempre o risco de que o documento sofra danos pelo manuseio incorreto ou excessivo. Dessa sorte, a própria recomendação³ ressalta uma série de fatores positivos para a digitalização de documentos:

- a. Contribuir para o amplo acesso e disseminação dos documentos arquivísticos por meio da Tecnologia da Informação e Comunicação;
- b. Permitir o intercâmbio de acervos documentais e de seus instrumentos de pesquisa por meio de redes informatizadas;
- c. Promover a difusão e reprodução dos acervos arquivísticos não digitais, em formatos e apresentações diferenciados do formato original;
- d. Incrementar a preservação e segurança dos documentos arquivísticos originais que estão em outros suportes não digitais, por restringir seu manuseio.

O CONARQ, em sua recomendação, ainda indica a execução de um projeto e a adoção de metodologia específica para cada acervo respeitando suas especificidades, além de tratamento arquivístico prévio, como avaliação e seleção dos documentos prioritários a serem digitalizados.

Outro fator fortemente recomendado, é que os documentos sejam digitalizados no local onde estão armazenados, ou seja, nas instalações onde o acervo está custodiado, evitando serem movimentados para outro local, pois este movimento pode acarretar perda e prejuízo à documentação.

1.3.1 Formato de Arquivos para os Representantes Digitais

A digitalização de acervos, por meio da criação de representantes digitais para os documentos analógicos, visa a ampliar o acesso e a facilitar a recuperação da informação.

De acordo com a natureza do documento arquivístico original, diversos dispositivos tecnológicos (*hardware*) e programas de computadores (*software*) serão utilizados para converter em dados binários o documento original para diferentes formatos digitais. No entanto, o produto dessa conversão não será igual ao original e não substitui o original que deve ser preservado. A digitalização, portanto, é dirigida ao acesso, difusão e preservação do acervo documental. (CONARQ, 2010, p. 5).

Para tanto, torna-se necessário estudar o formato que mais bem se adequa às necessidades de acesso desses representantes. Sendo assim, é imprescindível verificar as

³ *Opus citatum*

características dos documentos, sua natureza, forma ou gênero, pois assim será possível traçar estratégias para a digitalização.

Para cada forma ou gênero documental, será adotado um formato específico, de acordo com a intenção ao digitalizar: o acesso e/ou a preservação em longo prazo.

A criação do representante digital se dará por meio da captura digital que, para o CONARQ (2010, p. 5), significa “a conversão em imagem por dispositivo eletrônico (Escâner)”, para o formato digital de um documento originalmente não digital.

O processo de captura digital da imagem deverá ser realizado com o objetivo de garantir o máximo de fidelidade entre o representante digital gerado e o documento original, levando em consideração suas características físicas, estado de conservação e finalidade de uso do representante digital. Recomenda-se a digitalização das capas, contracapas e envoltórios, bem como de páginas sem impressão (frente e verso) especialmente quando contiverem sinalização gráfica de numeração e outras informações. [...] Os metadados técnicos a respeito do ambiente tecnológico (do documento original, da captura digital, do formato de arquivo digital gerado) e as características físicas dos documentos originais devem ser registrados em planilha e sempre que possível, devem ser encapsulados ao próprio objeto digital ou armazenados em um banco de dados. (CONARQ 2010, p. 7).

No presente estudo, adotou-se a nomenclatura “Matriz”, para designar o formato destinado à preservação e “Derivada de Acesso”, para os formatos que vislumbram dar acesso ao arquivo digital gerado, normalmente num tamanho mais leve que as matrizes digitais. Tal escolha justifica-se por serem as mesmas terminologias utilizadas pelo CONARQ.

Existem formatos de arquivos de imagem que são independentes de *software*, ou seja, “rodam”, abrem em diversos programas e, portanto, são ideais para serem eleitos na escolha de formatos dos representantes digitais. Tais formatos são: “TIFF, JPEG e PNG”, a adoção desses formatos para geração de matrizes e derivadas de acesso permitirão melhores condições de acesso em longo prazo, porque acarretam menor dependência de *software* e *hardware*. Além disso, demoram mais para se tornarem obsoletos, pois possuem mais atualizações gratuitas disponíveis para *download*, o que facilita o refrescamento e a atualização de versões.

1.3.1.1 Gerando Matrizes Digitais e Derivadas de Acesso

No momento de capturar imagens para criação de representantes digitais, impende lembrar que esses representantes servirão como matrizes para outras derivadas de acesso. Nesse sentido, há algumas considerações a serem levadas em conta no momento da escolha do formato adotado para a digitalização. A escolha dependerá, por exemplo, do quanto de perda por compressão será admitido e do destino que o representante digital terá: se sua utilização será para dar acesso ou para preservação. Se a intenção é criar uma matriz digital de alta qualidade, de modo a permitir a criação de derivadas de acesso a partir dela, então a

perda por compressão deverá ser menor e, por consequência, o tamanho do arquivo será maior e precisará de mais espaço para seu armazenamento.

No momento de escolha de formatos para criação de matrizes digitais, é de extrema relevância a adoção de formatos que permitam o acesso em longo prazo aos arquivos capturados. Para Ferreira (2006, p. 39) “sempre que possível, deverão ser escolhidos formatos conhecidos pela comunidade de interesse e baseados em normas internacionais abertas”. Isto é, normas *estandarizadas*, as chamadas *International Organization for Standardization (ISO)* (organização internacional para padronização).

Nessa perspectiva, o CONARQ (2010, p. 13) recomenda a adoção dos seguintes formatos para a criação dos representantes digitais matrizes:

1. TIFF – Tagged Image File Format;
2. PNG29 - Portable Network Graphics
3. JPEG 2000

Se a finalidade da criação de representantes digitais é para acesso, o CONARQ utiliza a nomenclatura “Derivada de Acesso”. Ou seja, são derivados de outro formato, visando ao acesso. Esses representantes digitais servem para visualização em tela seja *online* ou *offline*, facilitam o *download* e servem para impressão de imagens em pequenas escalas. Dessa forma, a melhor escolha será por formatos abertos e mais comprimidos, pois estes são mais leves, permitindo um acesso rápido, principalmente *online*, pois não se sabe a velocidade de conexão à Internet, tampouco a capacidade dos computadores de que dispõem os usuários para acessar os arquivos digitalizados.

Nesse caso, há de se pensar que os usuários precisarão carregar a imagem de forma rápida, seja em um *blog*, *site* ou outra plataforma, porém, a perda com a compressão será maior e a definição das cores na imagem será menor.

As derivadas de acesso podem receber tratamento de imagem a fim de permitir melhor visualização ou impressão. Entretanto, critérios éticos devem pautar esse tipo de intervenção para que elas não se tornem dissociadas e não representem corretamente o documento original que as gerou. Sempre que possível, deverão ser utilizados preferencialmente formatos abertos para a geração dessas derivadas, e recomendamos os formatos **JPEG35** e **PNG**. No entanto, pode ser necessária a utilização de outros formatos de arquivo digitais, até mesmo proprietários, em virtude de concepção de acesso por parte dos provedores e custodiadores dos representantes digitais. [...]. O formato de arquivo digital *Portable Document Format* - PDF⁴ ou PDF/A também é recomendado, embora possua uma taxa de compressão menor. Esse formato digital permite dar acesso ao usuário final uma representação fiel do documento original, em um único

⁴ PDF - *Portable Document Format*. - Formato digital desenvolvido pela Adobe em 1993, e em 2008 adotado como padrão ISO-3200, e “que permite capturar e visualizar informações consistentes - a partir de quase todo aplicativo ou sistema operacional - e compartilhá-las praticamente com qualquer pessoa, em qualquer lugar [...]”. Em 2005, a ISO adotou o padrão PDF/*Archive* - PDF/A - ISO 19005-1:2005, que assegura o acesso a longo prazo. Disponível em: <http://www.adobe.com/br/products/acrobat/adobepdf.html>.

arquivo digital, especialmente quando esse é formado por múltiplas páginas e contiver também imagens fixas. (CONARQ, 2010, p. 16, grifo nosso).

A escolha por formato PDF, que também é considerado um formato aberto, torna-se viável para casos em que é necessário digitalizar documentos com várias páginas em sequência, com um tamanho de arquivo relativamente pequeno. Os demais formatos, com exceção do TIFF (que deixará o arquivo num tamanho bem pesado), não permitem digitalizar e comprimir em um único arquivo diversas páginas de um documento ou imagem. Já o formato PDF permite que a sequência lógica da informação não seja perdida, porém dificulta a edição.

Finalmente, cabe ressaltar que no momento de nomear os arquivos do representante digital, seja ele uma matriz ou uma derivada de acesso, alguns cuidados deverão ser tomados na adoção de um controle de vocabulário, visando a favorecer a recuperação do arquivo e o acesso à informação desejada.

1.3.2 Linguagem Documentária para a Construção de um Vocabulário Controlado

Os estudos sobre a linguagem documentária datam da década de 1950. Segundo Cintra (2002, p. 33), “foi quando se iniciou na área da documentação uma preocupação com o armazenamento e a recuperação da informação”. Essa preocupação decorre do crescimento da massa documental e do conhecimento científico adquirido por parte dos profissionais. A solução foi buscar a construção de linguagens próprias que padronizassem o vocabulário, com o intuito de facilitar a recuperação da informação.

Vem desta época a utilização de Linguagens Documentárias – LD, para a recuperação da informação. Essas linguagens são, pois, construídas para a indexação, armazenamento e recuperação da informação e correspondem a sistemas de símbolos destinados a “traduzir” os conteúdos dos documentos. (CINTRA, 2002, p. 33).

Atualmente o conceito de Linguagem Documentária (LD) é visto por diversos autores de forma divergente. Por exemplo, para Lancaster (2004, p.19), LD é sinônimo de Vocabulário Controlado ⁵(VC), sendo este dividido em: “esquemas de classificação bibliográfica (como a classificação Decimal de Dewey), listas de cabeçalhos de assuntos e tesouros”. Para Gomes apud Cintra (2002, p. 40), a LD possui significado mais amplo, sendo o VC, Tesouros e Sistemas de Classificação Bibliográfica instrumentos decorrentes de seu estudo. A Arquivologia compartilha de segunda concepção: a construção de Tesouros e Planos de Classificação de Documentos.

As linguagens documentárias são consideradas metarrepresentações ou representações documentárias, ao lado de outras formas de representação da informação como resumos, catálogos, índices, repertórios e inventários. [...] As LDs operam no quadro das representações documentárias, vistas como linguagens de comunicação entre a informação documentária e o usuário que dela necessita. Para Maria Cristina Pinto (1985, p.169-186), as

⁵ Vocabulário Controlado: conjunto normalizado de termos que serve à indexação e à recuperação da informação. (DBTA - *Dicionário Brasileiro de Terminologia Arquivística*, 2005, p. 164).

LDs atuam nos sistemas de recuperação de informações em dois níveis: orientando o analista sobre quais os melhores termos para representar o assunto de um documento, e orientando o pesquisador sobre a escolha dos termos que corresponderiam à representação do assunto por ele procurado. (DOBEDEI, 2002, p. 56-57).

A LD é mais ampla e poderá abordar vários vocabulários, enquanto a construção de um VC será específica a uma realidade ou fundo documental, com o objetivo de padronizar termos e agilizar a recuperação da informação. Segundo Cunha e Cavalcanti (2008, p. 307) recuperação da informação é a “área que estuda as técnicas de tratamento, armazenamento e recuperação da informação num arquivo ou acervo de documentos, com o objetivo de atender às necessidades de informação do usuário - pesquisa documentária”.

As LDs devem considerar o contexto documental que será retratado e o universo em que se encontra o usuário. A construção de uma LD por meio da adoção de um VC deve usar como base a Linguagem Natural (LN) do sistema nocional desses usuários, ou seja, sempre levar em conta os termos já utilizados por eles e, a partir daí padronizar e conceber o instrumento.

O sistema nocional constitui-se a base de sustentação da LD e, conseqüentemente, do VC. Segundo Cintra (2002, p. 50), “um sistema nocional define-se como um conjunto estruturado de noções que o compõem e no qual cada noção é determinada pela sua posição no sistema”.

Sendo assim, a partir de noções gerais para mais específicas e a partir de análises das relações hierárquicas, utilizando a noção genérica para conjuntos de características comuns e específicas para quando se verifica a existência de diferenças dentro das características comuns, identifica-se pelo menos uma que a diferencie da noção genérica.

Por exemplo, um memorando solicitando transporte é genérico, mas é possível especificá-lo, subdividindo-o em transporte: aéreo e terrestre e, ainda, o terrestre em ônibus e automóvel. Sendo assim, a construção de um VC deve prever a redução de termos do geral para o mais específico possível, de modo que não deixe dúvidas no momento da adoção.

O vocabulário controlado deve ser elaborado de maneira coerente, espelhando de forma precisa o termo que representa. Com a finalidade de agilizar a recuperação da informação, torna-se necessário também adotar uma política de indexação. Segundo Lancaster:

Um vocabulário controlado é essencialmente uma lista de termos autorizados. Em geral, o indexador somente pode atribuir a um documento termos que constem da lista adotada pela instituição para a qual trabalha. Comumente, no entanto, o vocabulário controlado é mais do que uma mera lista. Inclui, em geral, uma forma de estrutura semântica. Essa estrutura destina-se, especialmente, a: 1) controlar sinônimos, optando por uma única

forma padronizada, com remissivas de todas as outras; 2) diferenciar homógrafos. [...] 3) reunir ou ligar termos cujos significados apresentem uma relação mais estreita entre si. (LANCASTER, 2004, p. 14).

Para obtenção de sucesso na recuperação da informação é necessária uma coerente indexação que relacione os termos no momento em que o usuário busca determinado assunto em um universo informacional. Segundo Cunha e Cavalcanti a indexação é:

Descrição do conteúdo de um documento por meio de uma linguagem documentária a fim de facilitar a memorização da informação em arquivos, fichários, bases e banco de dados - análise do conteúdo, análise documentária, descritor, palavra-chave, termo de indexação. (CUNHA; CAVALCANTI, 2008, p. 193).

Na arquivologia, a aplicação mais utilizada para o Vocabulário Controlado é a tradicional lista de cabeçalhos de assuntos que, segundo Lancaster (2004, p.20), “é similar ao tesouro por ser de base alfabética, mas difere dele porque incorpora uma estrutura hierárquica imperfeita e por não distinguir claramente as relações hierárquicas das associativas”.

Sendo assim, o Vocabulário Controlado intenciona controlar sinônimos e ligar termos que se relacionam, para evitar ambiguidades na representação de um assunto. Desse modo, o VC compreende uma linguagem artificial que padroniza e controla os termos, em busca de maior precisão na recuperação da informação.

1.4 Acesso à informação por meio de um sistema

Para Stair e Reynolds (2006, p. 12), um sistema de informação é um tipo especializado de sistema e pode ser definido como um “conjunto de elementos ou componentes inter-relacionados que coletam (entrada) manipulam (processo) e disseminam (saída) dados e informações e oferecem um mecanismo de realimentação para atingir um objetivo”. Esses sistemas de informação podem ser manuais ou computadorizados. No entanto, os sistemas computadorizados facilitam o acesso e abrangem um número maior de usuários, que podem utilizar o sistema de forma simultânea e em mais de um local.

Para que o acesso aos representantes e demais arquivos digitais fique disponível para os usuários, sejam internos ou externos, torna-se necessária a criação de um sistema computadorizado que dê acesso e que viabilize a recuperação da informação desejada por parte desses usuários, para que possam utilizar ou manipular essas informações, seja para gerar conhecimento, ou para auxiliar na tomada de decisão.

Esses sistemas devem permitir a interoperabilidade com outros sistemas informatizados da organização visando ao seu acesso, integração, manutenção e segurança. O objetivo desses sistemas é fazer o gerenciamento técnico e administrativo dos representantes digitais, por meio do controle intelectual (lógico) e da integridade (física) desses representantes digitais, sejam as Matrizes ou as Derivadas. Esse sistema é

distinto de um sistema eletrônico de descrição arquivística e acesso a usuários finais. (CONARQ 2010, p. 21).

O ideal é que seja desenvolvida uma solução em *software* livre, pois assim a interoperabilidade com outros sistemas torna-se flexível e viável, além de permitir que o sistema seja atualizado gratuitamente para novas versões, evitando a obsolescência tecnológica e a dependência de *softwares* proprietários que podem sair de linha a qualquer momento. Assim, ao optar pela implementação de um sistema de informação em software livre, gratuito e em código aberto, a instituição permite ao usuário acessar as informações desejadas por longo prazo.

1.4.1 Banco de Dados em *Software* Livre

O *software* de código aberto, para Stair e Reynolds (2006, p. 145), é aquele “disponível livremente para quem quiser e em forma fácil de alterar” de acordo com as necessidades do usuário, como o Sistema operacional Linux e o servidor de Internet Apache. Os autores ainda esclarecem:

Porque uma organização desenvolveria seus negócios usando software livre? Como pode algo disponível livremente pela internet ser estável, confiável ou ter suporte suficiente para ser colocado no coração das operações diárias de uma companhia? A resposta é surpreendente o software de código aberto é frequentemente *mais* confiável do que o software comercial. Como isso pode ocorrer? Primeiro, ao disponibilizar o código-fonte, de um programa os usuários podem corrigir problemas que encontrarem. Uma correção está frequentemente disponível algumas horas após a descoberta do problema. Segundo, com o código-fonte de um programa acessível a milhares de pessoas, a probabilidade de um defeito ser descoberto e corrigido antes de provocar danos é muito maior do que os pacotes de software tradicionais. (STAIR e REYNOLDS, 2006, p. 145).

Cabe ressaltar que nem todo código aberto terá sua licença de uso gratuita, alguns *softwares* cobram pela licença, no entanto muito outros são totalmente livres e gratuitos.

Assim, ao optar pela implementação de um sistema de informação em software livre, gratuito e em código aberto, a instituição permite ao usuário acessar, por longo prazo, as informações desejadas.

Nessa perspectiva, um banco de dados é um repositório em que são armazenadas informações pertinentes – e que, portanto, precisam ser preservadas –, que estão em formato digital e que possuem potencial de consulta para auxiliarem na tomada de decisão ou gerar novos conhecimentos.

Para Garcia-Molina (2001), banco de dados é uma coleção de dados organizados para facilitar o acesso à informação que será preservada durante um longo período. Esse banco precisa permitir a consulta por parte do usuário, por meio de uma operação que extraia dados especificados e que permita uma relação através da organização dos dados em uma tabela.

Nessa tabela as linhas representam entidades básicas ou fatos de algum tipo, e colunas (atributos) que representam propriedades dessas entidades.

Os relacionamentos são fundamentais para que o resultado da consulta retorne com informações estruturadas, refinadas e sem redundância. Ou seja, devem retornar as informações relevantes e pertinentes ao usuário.

Para garantir que uma informação armazenada possa ser recuperada devemos nos assegurar da eficiência do processo de armazenamento e recuperação. A abordagem relacional (entidade-relacionamento) apresentada pelo Dr. Peter Pin-Chan Chen (1976) é hoje uma das formas comprovadamente segura e eficaz de armazenamento e recuperação da informação. Neste modelo as informações são armazenadas em tabelas que podem ser inter-relacionadas através de operações de conjunto. (CANDIDO, 2005, p. 17).

Para Casanovas (2008), há ainda o modelo lógico global, aquele em que a base de dados dentro do banco estará disponível *online* em rede para diversos usuários, visando a transformar o dado em informação, ultrapassando a noção de coleção de dados. Casanovas destaca que, ao serem acessados por diversos e múltiplos usuários, cada um com sua ótica particular, com sua necessidade informacional, darão um novo sentido aos dados registrados transformando-os em informação relevante, a qual servirá de base para a geração de novos conhecimentos.

Una base de datos tiene tres niveles: - un nivel interno encargado del almacenamiento de datos a nivel físico, con indicadores al siguiente registro, y organizado en páginas; - un nivel lógico, donde se encuentran registradas las entidades, atributos y relaciones que conforman la estructura de los datos; - un nivel externo, para las vistas de usuario de las consultas y de la información recuperada. (CASANOVAS, 2008, p. 39).

Sendo assim, para Casanovas, esse modelo lógico deverá se transformar em um desenho físico, que considerará a forma de armazenamento, seus relacionamentos e seus mecanismos de acesso, resultando em um Sistema de Gerenciamento do Banco de Dados. Esse sistema permitirá controlar a criação, a atualização, a recuperação e a inclusão de novos dados na base de dados, com a função de assegurar que os dados sejam compartilhados pelos usuários com uma variedade de aplicações e ainda se mantenham íntegros, precisos e consistentes, permitindo que a base evolua e se adapte à necessidade crescente dos usuários.

2 Metodologia

O caminho percorrido nesta a pesquisa para alcançar os objetivos traçados é o método indutivo e comparativo, aplicado com a finalidade de verificar as semelhanças e explicar as divergências constatadas durante os estudos no que tange ao *software* livre em código aberto e ao software proprietário para a construção do banco de dados.

Para o levantamento de dados, foram utilizados três procedimentos: pesquisa documental no acervo do PNAIC/UFSM, para conhecer o universo documental para a realização do estudo; revisão bibliográfica, para embasamento teórico da pesquisa; e contatos

diretos e por *e-mail* com usuários internos da documentação e com a coordenadora do PNAIC/UFSM.

Para a coleta de dados, foi elaborado um questionário estruturado e aplicado por meio do formulário *online* - Google Forms. A intenção foi conhecer as necessidades informacionais dos usuários, para que a construção do banco de dados ajudasse a suprir tais necessidades. Essa opção tornou-se viável, pois os usuários encontravam-se dispersos em várias cidades do interior do Estado.

Foi elaborada uma metodologia para digitalização e nomeação dos representantes digitais gerados. Produziu-se um tutorial, contendo o passo a passo de cada etapa da digitalização, desde a escolha do formato do arquivo, até a resolução da imagem em pontos por polegadas (ou dpi, do inglês *dots per inches*) até a nomeação dos representantes digitais.

Para a construção do banco de dados, foram analisados diversos *softwares* proprietários e livres em código aberto. Realizou-se o comparativo dos benefícios que cada programa oferece, além de se analisar os malefícios que a dependência ao *software* proprietário pode resultar. A maior preocupação foi a questão da obsolescência a que todas as tecnologias digitais estão sujeitas.

Finalmente, para a análise dos resultados obtidos, demonstra-se a modelagem com o *software* BRmodelo, que permite demonstrar os relacionamentos das tabelas no banco de dados. Além disso, o BR modelo permite demonstrar a especificação adotada na construção do banco de dados e as vantagens e desvantagens de sua implementação em *software* livre.

3 Análise e discussão dos resultados

Nesta etapa são apresentadas a análise e a discussão dos resultados do estudo.

3.1 Pacto Nacional pela Alfabetização na Idade Certa: Coordenação Santa Maria/UFSM

Este estudo contribui com a elaboração de um banco de dados em *software* livre como instrumento facilitador do acesso e da recuperação da informação arquivística produzida e acumulada pelo PNAIC (Santa Maria/UFSM). Tal informação encontrava-se dispersa e possuía alto potencial para a pesquisa e, portanto, para a geração de novos conhecimentos.

Assim, torna-se necessário conhecer o PNAIC, um Programa do Governo Federal que visa alfabetizar crianças matriculadas em escolas públicas brasileiras. Segundo informações disponíveis no próprio sítio do Ministério da Educação e Cultura (MEC)⁶, o Programa Nacional pela Alfabetização na Idade Certa (PNAIC) é um "compromisso formal assumido pelos governos federal, do Distrito Federal, dos estados e municípios de assegurar que todas as crianças estejam alfabetizadas até os oito anos de idade, ao final do 3º ano do ensino fundamental".

Nesse contexto, em todas as esferas, os governos que desejam aderir ao pacto devem se comprometer em alfabetizar todas as crianças em língua portuguesa e matemática.

As ações do Pacto apoiam-se em três eixos de atuação:

⁶ Disponível em: < <http://pacto.mec.gov.br/o-pacto>>. Acesso em 28 set. 2015

- a) Promoção de formação continuada para os professores alfabetizadores e seus orientadores de estudo em cada estado, por meio de um curso presencial de 2 anos com carga horária de 120 horas por ano, baseado no Programa Pró-Letramento⁷;
- b) Disponibilização de materiais didáticos, obras literárias, obras de apoio pedagógico, jogos e tecnologias educacionais, visando à criação de uma biblioteca acessível e professores e estudantes na própria sala de aula;
- c) Avaliações sistemáticas desenvolvidas e custodiadas pelo MEC, debatidas pelos professores durante o curso de formação, com a possibilidade de serem desenvolvidas e realizadas continuamente pelo professor junto aos alunos; Inserção dos resultados da Provinha Brasil⁸.

O Rio Grande do Sul possui dois coordenadores gerais para promover a formação continuada dos professores alfabetizadores, que serão os multiplicadores desta ação do Governo Federal. Para a realização deste estudo, o foco foi apenas o PNAIC sob coordenação geral de uma docente efetiva da UFSM e atuante no Centro de Educação da referida universidade.

Nessa perspectiva, o pró-Letramento é o nome dado ao programa de formação continuada dos professores. O pró-Letramento objetiva capacitar professores para alfabetizar as crianças na idade certa. Objetiva, também, no âmbito de cada seu município, formar multiplicadores que auxiliem a capacitação de outros alfabetizadores, nas escolas onde atuam.

Esse programa é realizado por meio de parcerias com universidades federais e conta com a adesão dos estados e dos municípios que enviam seus professores para a capacitação. A estrutura organizacional do programa consiste em: MEC – Ministério da Educação (Secretaria de Educação Básica) > Universidades Formadoras > Redes Públicas de Ensino (Secretarias de Educação).

Os participantes do programa são os detentores de grande parte do material produzido para ser utilizado durante a realização das formações. Esse material foi gerado a partir dos cadernos distribuídos pelo governo, os quais são fundamentais para serem preservados, têm caráter único e possuem características de valor permanente. Ora, trata-se de informações registradas, que servem de base para futuras pesquisas em educação e que

⁷ O Pró-Letramento - Mobilização pela Qualidade da Educação - é um programa de formação continuada de professores para a melhoria da qualidade de aprendizagem da leitura/escrita e matemática nos anos/séries iniciais do ensino fundamental. O programa é realizado pelo MEC, em parceria com universidades que integram a Rede Nacional de Formação Continuada e com adesão dos estados e municípios. Disponível em: < <http://portal.mec.gov.br/pro-letramento> >. Acesso em 28 set. 2015.

⁸ A Avaliação da Alfabetização Infantil – Provinha Brasil visa investigar o desenvolvimento das habilidades relativas à alfabetização e ao letramento em Língua Portuguesa e Matemática, aplicadas no 2º ano do ensino fundamental das escolas públicas brasileiras duas vezes ao ano (no início e no final), Disponível em < <http://provinhabrasil.inep.gov.br/> >. Acesso em 28 set. 2015.

serão auxiliares na tomada de decisões administrativas do programa pela coordenação de Santa Maria/UFSM.

Sendo assim, os responsáveis⁹ pelas ações do programa nos Estados são:

- A. Coordenador Geral do Programa:** deve possuir vínculo com a Universidade formadora e fica responsável pela implementação do programa;
- B. Formador:** deve, preferencialmente, ser vinculado à Universidade formadora e responsável pela formação dos orientadores de estudo;
- C. Coordenador Administrativo do Programa:** deve ser um profissional da Secretaria de Educação, responsável pela organização do Programa no município e pela articulação entre a IES e a Secretaria de Educação;
- D. Orientador de Estudo:** deve ser o profissional efetivo do município e receber a formação da Universidade. Sua indicação é feita pela Secretaria de Educação e deverá ser pautada em sua experiência profissional e em sua formação acadêmica. Esse ator é peça-chave no projeto, pois ele será o articulador entre a universidade e os cursistas;
- E. Professor Cursista:** também chamado pelos usuários PNAIC/UFSM de **Professor alfabetizador**, deve ser professor de séries ou anos iniciais do ensino fundamental (1ª a 4ª série ou 1º ao 5º ano), estar atuando em sala de aula e ter feito a inscrição.

Cabe ressaltar que todos os participantes do programa, que possuem responsabilidades e competências a serem desenvolvidas, recebem bolsa e estabelecem compromisso no processo formativo, assinando termos que são enviados para a coordenação geral/UFSM, para serem arquivados.

Esses termos encontram-se todos registrados em suporte papel e são armazenados em pastas plásticas dispostas em estantes, ainda, sem nenhum tratamento arquivístico. Alguns exemplos dos tipos documentais gerados para este programa integrante do PNAIC e que estão em formatos híbridos, tanto em suporte papel como digitais, são:

- a) Termos de compromisso;
- b) Certificados de participação;
- c) *Slides* para formação dos orientadores de estudo e para os professores alfabetizadores baseados no material recebido nos cadernos do Pró-Letramento distribuído pelo MEC;
- d) Avaliação dos orientadores de estudo e dos cursistas – professores alfabetizadores;
- e) Relatórios de acompanhamento e execução dos seminários de formação.

A partir do levantamento dos tipos documentais produzidos e recebidos pelo programa PNAIC/UFSM, foi possível traçar as estratégias para alcançar o próximo objetivo deste estudo. As próximas etapas incluíram a escolha do formato ideal para os arquivos dos representantes digitais dos documentos selecionados para a digitalização, a higienização e a remoção de cliques metálicos dos papéis e a adoção do vocabulário controlado. O acervo escolhido levou em consideração apenas os documentos

⁹ De acordo com o Guia do Pró-Letramento 2012.

considerados relevantes para pesquisa e que auxiliam na tomada de decisão no âmbito do PNAIC/UFSM.

3.2 Higienização dos documentos selecionados para digitalização, formato de arquivo digital e vocabulário controlado

Na sequência, discorre-se sobre a identificação e a seleção dos tipos documentais nas pastas físicas, a escolha do formato dos arquivos dos representantes digitais.

3.2.1 Identificação e seleção dos tipos documentais nas pastas físicas

Os documentos selecionados para digitalização estão atualmente acondicionados em pastas de plástico polionda, dispostas em ordem alfabética, identificando as cidades atendidas pelo Programa e armazenadas em uma estante de madeira. Abre-se, neste momento, um parêntese para ressaltar que esse não é o material indicado para uma estante que armazenará documentos. O material indicado nesse caso é o aço, pois impedirá a infestação por cupins.

Os documentos selecionados são, em grande maioria, Termos de Compromisso de Bolsistas. Foram identificadas em todas as pastas analisadas: uma folha de rosto com a identificação do coordenador local e a do orientador de estudo, seguidas dos tipos documentais acondicionados na pasta.

Como exemplo, apresenta-se, a seguir, o teor do conteúdo de uma pasta digitalizada. A pasta é referente à cidade de Agudo, no ano de 2013. Este é o padrão identificado em todas as pastas armazenadas sob a custódia do PNAIC/UFSM:

- A. Coordenador local: XXXXXXXXXXX
- B. Orientador de estudo: XXXXXXXXXXXXXXXXX

Tipos documentais identificados na pasta:

- a. Atestado de atuação de professora alfabetizadora;
- b. Listagem de professores alfabetizadores aptos para receber certificação;
- c. Ofício encaminhando informações da coordenadora do pacto;
- d. Ofício encaminhando informações do orientador de estudo;
- e. Ofício informando desligamento de professora alfabetizadora do programa;
- f. Termo de compromisso do participante bolsista.

3.2.2 Recomendações preliminares à digitalização

Antes de iniciar a digitalização, foi confeccionado um roteiro contendo as instruções que deveriam ser seguidas no processamento dos documentos. Esse roteiro ficou exposto em local visível, na sala de digitalizações. Essa padronização de método facilitou a digitalização, já que o processamento do acervo foi realizado por pessoas distintas. Assim, o tutorial trouxe segurança aos operadores.

1. Observar se o documento possui alguma sujidade, grampo ou amassado, caso possua, como proceder:
 - a. Sujidade: remover passando a borracha com cuidado e remover o pó de borracha completamente do papel antes de colocar no escâner;

- b. Amassados ou dobrados, desamassar ou desdobrar a folha alisando-a antes de colocar no escâner;
- c. Grampos metálicos: remover os grampos com cuidado antes de colocar no escâner;
- d. Cuidar para que o papel fique bem alinhado, ou seja, não fique torto dentro do escâner antes da digitalização.

Essas recomendações preliminares foram de suma importância, pois a sujeira e a ferrugem advinda do metal comprometem a fibra do papel, provocam amarelamento, manchas e até furos nesse suporte, com o passar do tempo. A remoção do metal também evitou danos ao equipamento, pois o metal e a sujeira poderiam danificar o escâner.

Sendo assim, a adoção dessas recomendações permitiu que os documentos digitalizados fossem, posteriormente, armazenados livres de sujidades e sem metais. Essas medidas são necessárias para se manter os documentos em melhores condições, possibilitando maior longevidade ao acervo físico.

3.2.3 Escolha do formato eletrônico para digitalização

Após estas orientações preliminares para remoção de sujidades e objetos metálicos do papel, seguiu-se para as instruções acerca da adoção do formato de digitalização e sobre como configurar o escâner para digitalizar os documentos nos padrões adotados. O modelo do escâner utilizado foi HP Officejet J4660 All-in-one.

A escolha do formato dos representantes digitais foi o pdf, pois muitos dos documentos são compostos por várias páginas. Assim, poder-se-ia manter a sequência orgânica do documento para leitura, tal como no documento original. A adoção desse formato permite que o documento fique legível por diversos tipos de *software*, o que facilita seu acesso em longo prazo. Como esses documentos digitalizados não precisarão ser editados futuramente, não é necessário mantê-los em formato editável de texto, como o .odt por exemplo.

Nessa perspectiva, seguem as instruções para digitalização dos documentos do PNAIC, disponibilizadas em local visível para os digitalizadores:

1. Abrir a central de soluções HP Officejet J4660 Series que está com atalho na área de trabalho, escolher opção de digitalizar documentos.
2. Escolher atalho e configurar equipamento para digitalização: A próxima etapa será a escolha por formato e configurações de digitalização que já deverão estar salvas no equipamento, como Atalho de digitalização e tipo de arquivo, no caso optou-se por pdf.
3. Confirmando formato de arquivo (pdf) e local de arquivamento: A próxima etapa é a confirmação do tipo de arquivo no formato pdf para digitalização e escolha de local de armazenamento (gravação) do arquivo que será digitalizado e defina o nome do arquivo base que nesse caso será necessário criar uma pasta sempre que iniciar-se uma nova cidade. Primeiro crie a pasta com o nome da cidade (meus documentos>minhas imagens> Pasta “nome da cidade”, e posteriormente clique em procurar e localize. Clique uma vez na pasta referente a cidade que está digitalizando).

4. Digitalizar: O próximo passo será a digitalização, clique em ok e aguarde a próxima instrução.

5. Continuidade de documentos para digitalizar no mesmo arquivo ou conclusão: Na última etapa será necessário confirmar a inserção de mais um documento para ser digitalizado ou finalizar a digitalização. Atenção esta etapa só será realizada se o documento possuir mais de uma página, caso possua, troque o documento dentro do escâner antes de clicar em sim, caso contrário, clique em não e finalize.

6. Nomear o arquivo do Representante Digital: Agora que o documento já foi digitalizado, o arquivo no formato pdf precisará ser nomeado visando facilitar a recuperação das informações. Assim, o arquivo receberá a identificação de forma padronizada através do controle de vocabulário. Ex.: agudo_lista_professores_aptos_certificacao. Não utilize, espaço, vírgulas ou acentos, pois pode dar problemas posteriores no banco de dados. Exemplos de controle de vocabulário para nomear os arquivos dos representantes digitais:

A.Ofícios:nomedacidade_tipodocumentalnº_assunto_documento.

Ex: ajuricaba_oficio252_informando_orientadora_estudo.

B.Termos de compromisso: nomedacidade_tipodocumental_iniciaisdafuncao
desempenhada nome do participante.

Ex: ajuricaba_termo_compromisso_participante[pa]rubia_denise_brivio

1. [pa]: professor alfabetizador

2. [oe]: orientador de estudo;

3. [sp]: supervisora pedagógica

4. [cl]: coordenadora local

Este é um exemplo constante no tutorial criado para auxiliar na atividade digitalização. Nesse documento, que contém ilustrações, é possível visualizar melhor o procedimento padrão adotado. Concluídas essas etapas, foi possível avançar para a próxima: a construção do banco de dados em *software* livre.

3.2.3 Construindo a Modelagem ER e o Banco de Dados em *Software* Livre – BASE

Este momento consistiu nas etapas necessárias para a construção do banco de dados em *software* livre. Para este estudo, foram utilizados os *softwares* gratuitos BRModelo, para a modelagem das entidades e seus relacionamentos; e o *software* livre Base, para a implementação propriamente dita.

O Base integra o pacote do *LibreOffice*, sem depender de *softwares* proprietários. Com essa escolha, busca-se alcançar longevidade de acesso às informações inseridas no banco, além de ser mais simples de manipular e construir.

3.2.3.1 Elaboração da Modelagem Entidade e Relacionamento

Para construir um banco de dados é necessário planejamento. Para tanto, a primeira etapa realizada é a modelagem ER. Essa modelagem consiste na criação de um diagrama contendo as entidades e respectivos atributos, bem como a especificação dos relacionamentos entre tais entidades. Esse modelo favorece a normalização do banco de dados, o que, por sua vez, poderá facilitar a recuperação da informação, pois um modelo bem planejado torna o retorno das buscas mais preciso para o usuário.

Segundo Candido (2005, p. 13) “um modelo de dados é uma descrição dos tipos de informações que estão armazenadas em um banco de dados”. Para a modelagem do diagrama de Entidades X Relacionamentos foi utilizado o *software* gratuito BRModelo¹⁰, ferramenta indicada para a elaboração de modelos de relacionamentos em bancos de dados.

Este modelo, quando apresentado graficamente por meio de um diagrama, como se utilizou neste estudo, recebe o nome de Diagrama de Entidade e Relacionamento (DER)¹¹, conforme ilustra figura 9. Com o modelo finalizado, foi possível vislumbrar o potencial informacional do Banco de Dados.

As informações registradas, oriundas do PNAIC/UFSC, estavam dispersas e foram centralizadas em repositório único. As informações foram salvas em um HD externo para, posteriormente, serem inseridas no Banco de Dados por bolsistas do PNAIC/UFSC. Além desse suporte, foram salvas em um HD Virtual, no Google Drive. Conhecendo o universo informacional que seria inserido no Banco de dados, foi possível construir o Modelo ER que atendesse às necessidades da coordenação do PNAIC/UFSC.

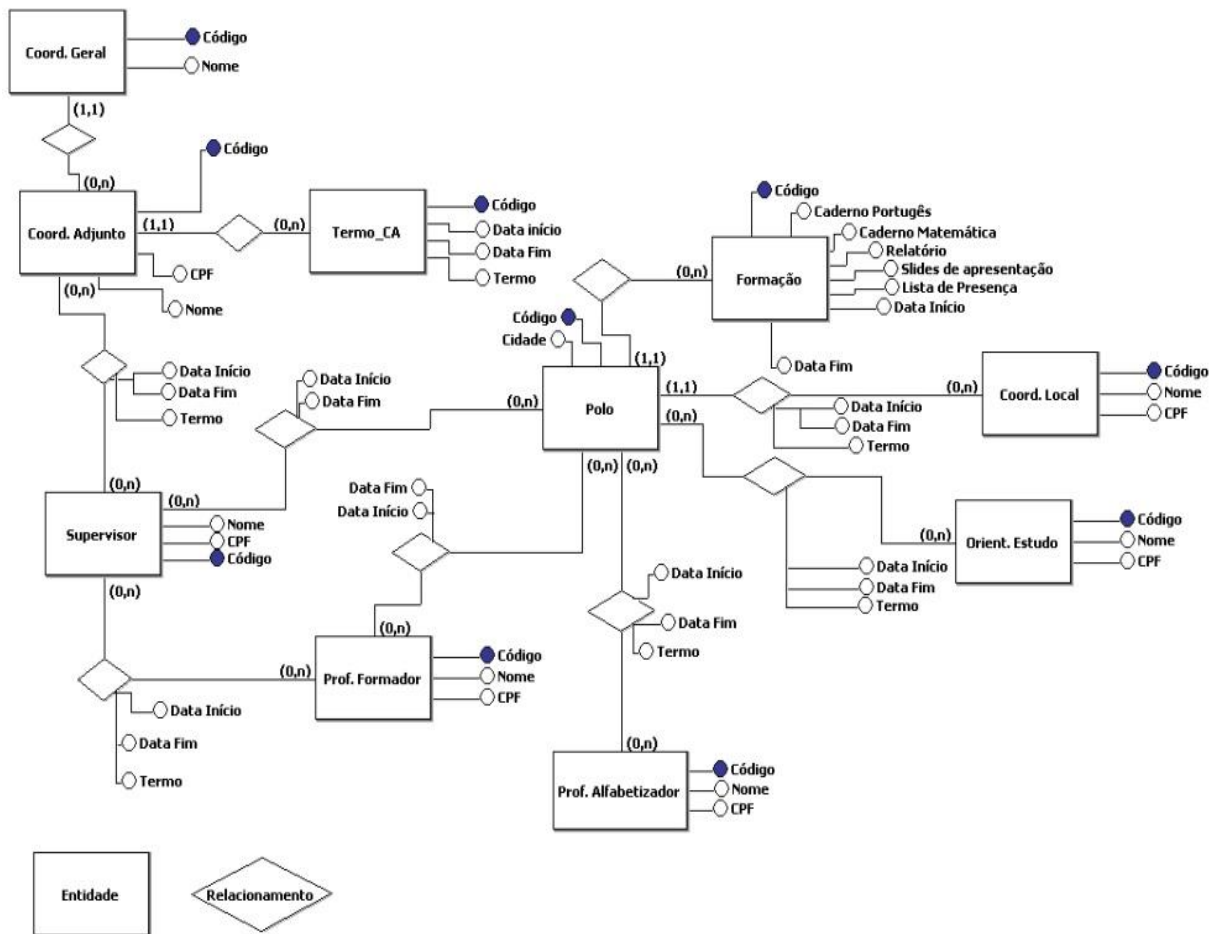
Observa-se, na figura 1, que as chaves primárias representadas pelas bolinhas azuis, são os valores que nunca se repetem na mesma tabela. Assim, podem ser usadas como referência na criação dos relacionamentos com as demais tabelas do banco de dados. Nessa perspectiva, as chaves primárias nunca podem ter valor nulo, tampouco, podem ter informações repetidas.

As entidades encontram-se representadas pelos retângulos que, por sua vez, constituem as tabelas. Os relacionamentos entre as tabelas estão representados pelos losangos. Assim pode-se vislumbrar como as tabelas se inter-relacionam. Mais tarde, os relacionamentos que estão ilustrados na figura 1 serão implementados no banco, mediante o uso de tabelas que contêm basicamente as chaves estrangeiras e, se necessário, outros atributos que qualificam o relacionamento. Importa lembrar que, embora implementados na forma de tabelas, não são entidades. São relacionamentos.

¹⁰ Esta ferramenta foi desenvolvida por Carlos Henrique Cândido sob a orientação do Prof. Dr. Ronaldo dos Santos Mello (UFSC), como trabalho de conclusão do curso de pós-graduação em banco de dados (UNVAG - MT e UFSC). Disponível em:< <http://sis4.com/brmodelo/download.aspx>>. Acesso em 13 de dez. 2015.

¹¹ Esse diagrama é também conhecido como Modelo de Entidade x Relacionamento (MER)

Figura 1 - Modelo de Diagrama de Entidade e Relacionamento - DER utilizado no Banco de Dados
PNAIC- UFSM



Fonte: Captura de tela de modelagem elaborada pela autora

Sendo assim, pode-se avançar para a próxima etapa: conhecer a construção do banco de dados do Pacto, utilizando o *software* Base, um banco de dados flexível por integrar-se completamente com outras aplicações do *LibreOffice*.

3.2.3.2 Construção do Banco de Dados em *Software* Livre

Ao se optar pela utilização do BASE levou-se em consideração a facilidade de manuseio por parte do autor do banco e dos usuários que usufruem das facilidades oriundas de sua utilização. A escolha pelo BASE foi possível, pois o volume de informações armazenadas não será muito grande e, portanto, este *software* é capaz de suportá-lo. Além, claro, de ser uma ferramenta integrante do *LibreOffice*, cuja facilidade de acesso para *download* de qualquer usuário é notória.

Segundo informações encontradas no *site*¹² do fornecedor, o *Base* vem configurado com o completo motor de banco de dados HSQL relacional. Este banco é uma “solução ideal para as necessidades sem complicações, e para as pessoas que necessitam de um sistema simples de usar e fácil de entender; os dados são armazenados dentro do formato ODF (extensão ODB)”.

O *Base* disponibiliza, por meio da opção Banco de Dados do *Base*, um assistente para a criação de bancos de dados. A criação é relativamente simples. Basta que o usuário observe com atenção o passo a passo de cada etapa, a fim de não cometer equívocos.

Cada banco de dados do *Base* é armazenado no HD em forma de arquivo. O passo seguinte à criação desse arquivo é a criação das tabelas. Nessa etapa, o desenvolvedor precisa ter em mãos o DER, para criar os campos que deverão constar nas tabelas. Para isso, pode utilizar o assistente de tabelas do *Base*.

A primeira tabela criada serviu para armazenar os dados da Coordenadora Geral do PNAIC/UFSM. O primeiro campo criado foi o código de identificação, utilizado como chave primária. Segundo o *site* DEVMédia, 2017¹³ a chave primária, ou *primary key*, é o “conceito mais básico relacionado à organização em um banco de dados. Toda tabela irá possuir uma, e somente uma, chave primária. Essa chave é utilizada como o **identificador único** da tabela”. Reforçando o que já foi dito, este campo específico não poderá receber valores repetidos.

Todas as tabelas criadas no banco devem ter uma chave primária, pois ela é essencial para o funcionamento do banco de dados, facilitando as buscas e garantindo que cada valor dentro da tabela seja diferente do outro.

A próxima tabela criada serviu para guardar os dados referentes às Formações. Essa tabela foi relacionada às informações atinentes aos polos, por meio de chaves estrangeiras. Na tabela das formações foram criados os campos: CodFormação; Caderno de Português; Caderno de Matemática; Relatório; Presença; Data de Início; Data de fim e Código do Polo (chave estrangeira para ligação entre a formação e o polo).

Além das chaves primárias também foi preciso criar chaves estrangeiras. As chaves estrangeiras servem para criar o relacionamento entre tabelas.

A chave estrangeira, ou *foreign key*, é um conceito ligeiramente diferente. Ela não diz respeito, especificamente, a uma tabela, mas sim a um relacionamento entre tabelas. De forma sucinta, a chave estrangeira é uma referência, em uma tabela, a uma chave primária de outra tabela. (DEVMEDIA, 2016).

No caso do banco do PNAIC, as chaves estrangeiras relacionaram diversas entidades. Por exemplo, a tabela dos Polos (TabPolo) foi ligada à tabela dos Orientadores de Estudo (TabOE), com a criação de um relacionamento TabOE_Polo, o qual contém as chaves estrangeiras CodOE e CodPolo. Ainda de acordo com DEVMédia diferentemente da chave primária, a chave estrangeira: “pode ser nula (NOT NULL); é um campo em uma tabela que faz referência a um campo que é chave primária em outra tabela; e é possível ter mais de uma (ou nenhuma) em uma tabela”.

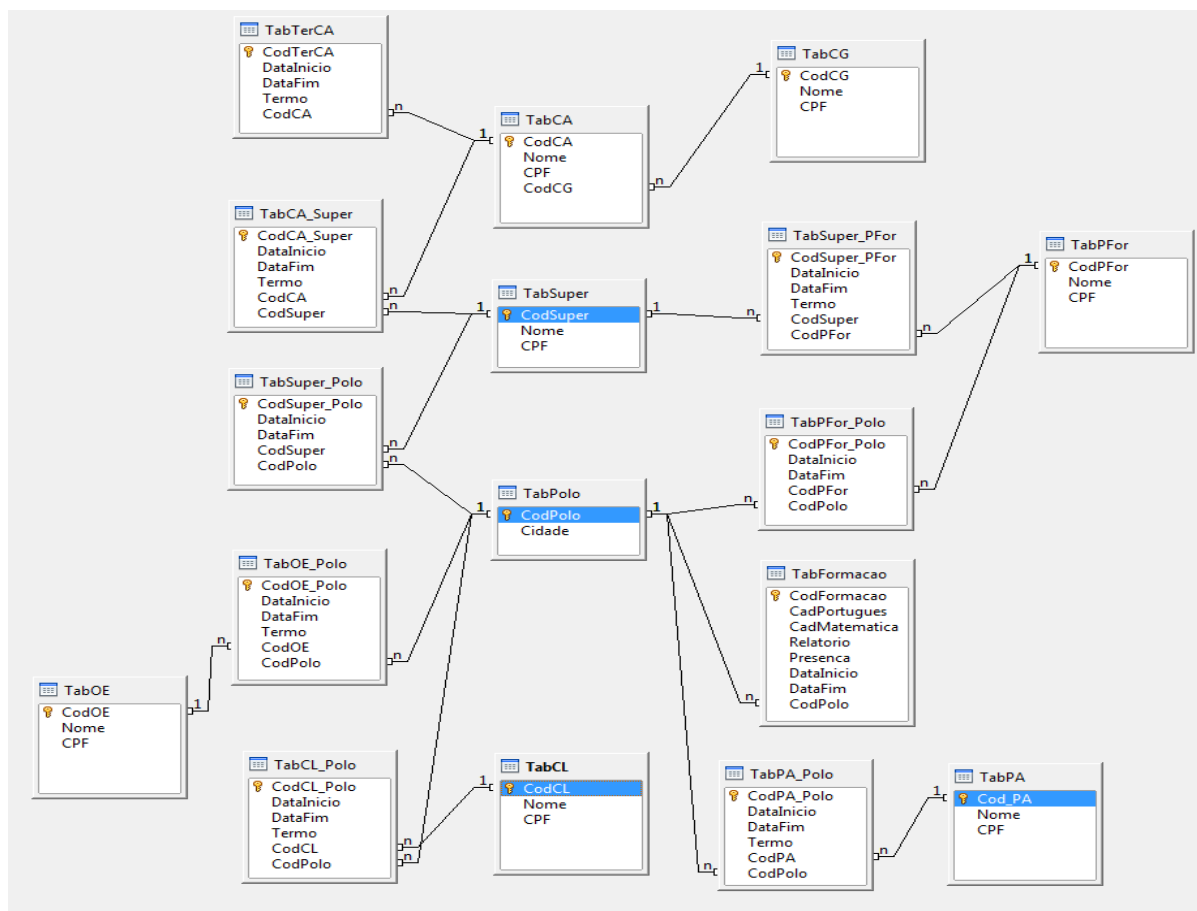
¹² Disponível em: <<https://pt-br.libreoffice.org/descubra/base/>>. Acesso em 13 março 2017.

¹³ Disponível em: <<http://www.devmedia.com.br/sql-aprenda-a-utilizar-a-chave-primaria-e-a-chave-estrangeira/37636>>. Acesso em 08 dez. 2016.

A tabela TabOE_Polo que, como dito, vincula os polos aos orientadores, além da chave primária (CodOE_Polo), possui data de início; data fim; Termo; e as chaves estrangeiras CodOE (código do Orientador de Estudo) e CodPolo (Código do Polo).

A figura 2 ilustra, de forma clara e objetiva, todas as tabelas e relacionamentos necessários para recuperar de forma pontual e precisa as informações depositadas no banco.

Figura 2 - Construção do Banco – Visão Geral das Tabelas das Entidades e seus Relacionamentos



Fonte: Captura de tela elaborada pela autora

Após gerar todas as tabelas, foi necessário criar os respectivos formulários e relatórios de consulta. Os formulários que serão criados devem corresponder às tabelas já criadas, pois os constituem-se a forma de entrada de dados nessas tabelas.

Durante a realização dos testes desta pesquisa, constatou-se que a versão original do Base não suporta arquivos no formato pdf. Assim, procedeu-se a um estudo acerca das possibilidades de alterar essa configuração, de modo a atender às necessidades do PNAIC.

Verificou-se que o Base utiliza um conjunto de blocos de código de programação denominados macros de dados. Essas macros servem para que as tabelas realizem tarefas específicas, tais como adicionar dados, excluir dados, ou aceitar dados em formatos específicos. Após a realização das alterações e dos respectivos testes, formulou-se um quadro com as codificações incluídas nas macros do Base. Essas modificações foram realizadas de modo que o software passe a aceitar o formato pdf, dando suporte às características do acervo do PNAIC.

Alguns documentos confeccionados pela equipe de formação do PNAIC/UFSM foram originalmente produzidos em formatos proprietários. É o caso, por exemplo, dos relatórios e das listas de presença, que estão em formato de texto .doc, bem como dos slides de formação, que se encontram em formato pptx.

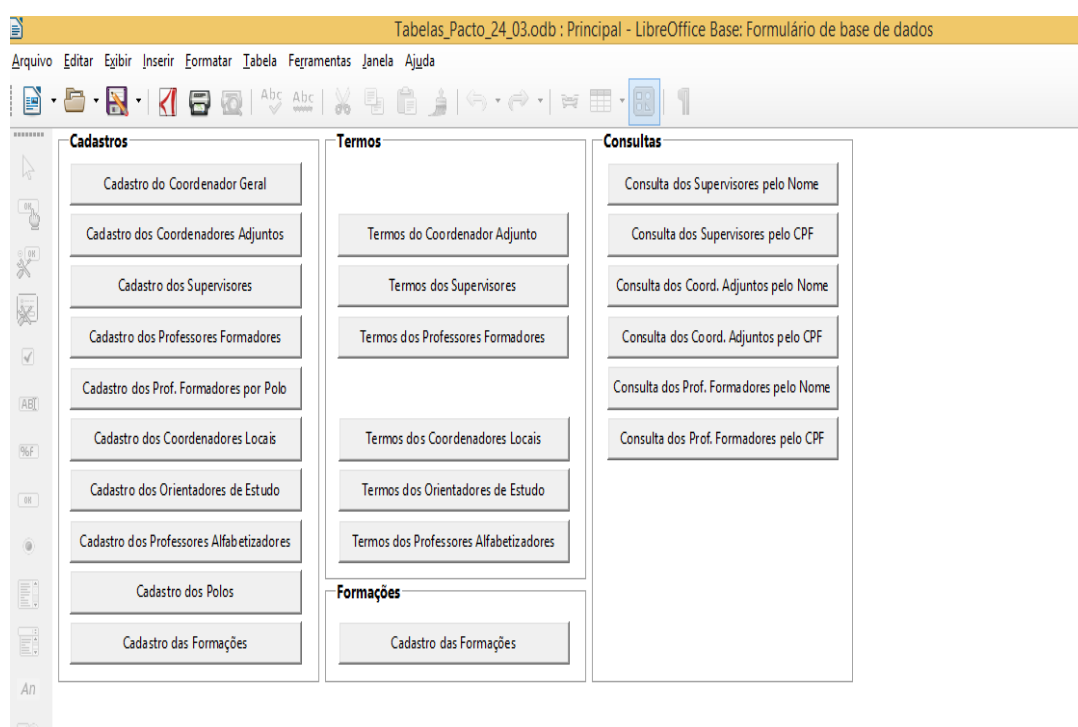
Nesses casos, sugere-se a conversão de todos os arquivos para o equivalente em formato aberto, visando à longevidade que tal formato proporciona, dadas as atualizações de versões gratuitas ao longo do tempo.

Na finalização do banco de dados, é imprescindível se criar uma espécie de menu, por meio do qual o usuário poderá vislumbrar todas as opções de gestão das informações, como ilustra a figura 3.

Nesse módulo possível realizar consultas aos termos do professor formador e a pesquisa pode ser feita em diversos campos ou por somente um deles, dependendo da quantidade de informação prévia que o usuário tiver.

Sendo assim, o banco foi finalizado e está apto para receber arquivos em diversos formatos, tanto abertos, como proprietários. Permite, inclusive, a inserção de imagens e slides de apresentação, além de planilhas, textos editáveis e arquivos em pdf.

Figura 3 - Tela inicial do Banco de Dados finalizado



Fonte: Captura de tela elaborada pela autora

Ressalta-se que, embora o banco permita a inclusão de arquivos proprietários, como .doc, .docx, .ppt, .pptx, .xls, .xlsx, entre outros, a adoção desses formatos é desaconselhável, a fim de se evitar, no futuro, a dependência de fornecedores específicos.

O *design* do banco ficou bem simples e intuitivo, o que facilitará a utilização por usuários com pouca fluência tecnológica. Embora esteja simples, possui funcionalidades suficientes para suprir a contento as necessidades de armazenamento e consulta dos usuários do PNAIC/UFSM.

4 Conclusão

As informações armazenadas no banco de dados proposto são essenciais para a continuidade do programa Alfabetização na Idade Certa, por vários motivos. Primeiramente, trata-se de um programa abrangente, com centenas de professores e escolas envolvidas, o que evidencia a necessidade da organização administrativa e documental do acervo produzido semestralmente pelo programa.

Além disso, esse programa passa por fiscalizações de diversos órgãos federais e é necessário manter uma base informacional fidedigna e correta, visando responder a quaisquer diligências apresentadas. Dessa forma, as atas, folhas de chamada e os cadernos de relatórios produzidos precisam ser acessados para verificação, o que corresponde a boa parte da demanda encontrada.

Por outro lado, essa mesma abrangência revela o esforço governamental que precisa ser corretamente documentado e preservado para gerações futuras, sobretudo estudantes do Centro de Educação. As necessidades informacionais de documentos preservados de forma permanente podem adquirir as mais diferentes formas em tempos futuros. A correta armazenagem e preservação dos documentos garante que pesquisadores e gestores tenham acesso aos resultados do programa, analisando de forma mais fidedigna o alcance da gestão.

Observou-se, neste estudo, a fundamental importância de se conhecer o universo documental antes de se realizar qualquer intervenção arquivística. Com o conhecimento do acervo e das necessidades informacionais dos usuários, é possível traçar as estratégias que serão implementadas de forma assertiva e pontual, necessitando apenas de pequenos ajustes ao longo do percurso.

Constatou-se também a necessidade de investigar e conhecer a tecnologia a ser adotada, pois todo *software* tem limitações. É necessário conhecê-las e planejar o que será feito para resolver os problemas encontrados. Basta ver, como exemplo, o caso das alterações necessárias para aceitação do formato pdf no BASE.

Conforme mencionando, escolheu-se o formato pdf como padrão para os arquivos textuais e, no momento dos testes no banco, constatou-se que esse formato não era suportado. Assim, foi necessário alterar as macros do programa, para que tal formato fosse aceito. Nesse caso, confirmou-se um dos grandes benefícios do *software* livre com licença aberta: permitir a reprogramação do código, para que atenda às necessidades do usuário.

Após a seleção dos documentos que seriam digitalizados, verificou-se a importância de se confeccionar um tutorial com os passos a serem seguidos pelo digitalizador. Esse tutorial trouxe segurança para a adoção dos procedimentos e mais qualidade ao produto dos representantes digitais.

Outro aspecto que se confirmou necessário, foi a adoção do vocabulário controlado para nomear os arquivos digitalizados. Essa escolha foi fundamental, pois padronizou os títulos dos arquivos, de modo a se evitar, por exemplo, o uso de acentuações e espaços que poderiam prejudicar o retorno das consultas no banco.

Outra decisão assertiva, foi a opção pelo *software* livre para a construção do banco de dados. Primeiro, porque, tratando-se de um programa do MEC, é recomendável que se evite o uso de *software* proprietário, em face da política de dados abertos¹⁴ do Governo Federal.

Outro ponto foi a escolha da ferramenta. Ora como a massa documental inserida não é muito extensa, foi possível utilizar um *software* como o Base, que já vem pré-programado. Essa característica é ideal para quem não possui fluência em programação. Afinal, os comandos preveem e permitem o uso de assistente de configuração para a construção do banco, em suas tabelas, formulários, relatórios e consultas.

Outro aspecto nevrálgico foi a possibilidade de alterar a programação do banco, por tratar-se de um *software* em código aberto, a permissão de alterar as macros para que aceitassem outro formato de arquivo que não aquele previamente estipulado foi de grande relevância. Não fosse essa característica do programa, outras medidas mais complexas haveriam de ter sido tomadas para resolver essa questão importante no estudo realizado.

Enfim, pode-se dizer que a adoção do *software* livre foi uma boa solução para a gestão do patrimônio documental do PNAIC/UFSC, pois resolveu a contento uma forma de concentrar todas as suas necessidades informacionais em um único repositório, com o uso de um sistema informatizado que permitiu o acesso e recuperação de forma gratuita, ágil e pontual.

Referências

ALBERCH, Ramon Figueras; BOIX, Lurdes; NAVARRO, Natálys *et al.* **Archivos Y Cultura: manual de dinamización**. Gijón: Ediciones Trea, 2001. 173p.

BRASIL. **Constituição da República Federativa do Brasil**: promulgada em 5 de outubro de 1988. 18. Ed. São Paulo: Saraiva, 1998. (Coleção Saraiva Legislação).

BRASIL. MINISTÉRIO DA EDUCAÇÃO E CULTURA. **O Pacto**. Disponível em: <<http://pacto.mec.gov.br/o-pacto>>. Acesso em 28 set. 2015.

CANDIDO, Carlos. Aprendizagem em Banco de Dados, Implementação de Ferramenta de Modelagem E. R. Disponível em: < <http://www.inf.ufsc.br/~r.mello/bdnc/Especializacao-CarlosCandido-FerramentaModelagemER-2005.pdf>>. Acesso em 15 out. 2016.

BELLOTTO, Heloísa Liberalli. **Arquivos permanentes: tratamento documental**. 4. ed. Rio de Janeiro: Editora FGV, 2006. 320p.

CANDIDO, Carlos. Aprendizagem em banco de dados, Implementação de Ferramenta de Modelagem E.R. Disponível em: < <http://www.inf.ufsc.br/~r.mello/bdnc/Especializacao-CarlosCandido-FerramentaModelagemER-2005.pdf>>. Acesso em 15 out. 2016.

CASANOVAS, Inés. **Gestión de documentos electrónicos**. Buenos Aires: Alfagrama, 2008. 256p.

¹⁴ Decreto Federal Nº 8.777/ 2016, disponível em:< http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8777.htm>. Acesso em 18 nov. 2016.

CINTRA, Anna Maria Marques; TÁLAMO, Maria de Fátima G. Moreira; LARA, Marilda Lopes Ginez de; *et al.* **Para entender as linguagens documentárias**. 2. ed. São Paulo: Polis, 2002. 96p. (Coleção Palavra-Chave, 4).

CONSELHO NACIONAL DE ARQUIVOS (Brasil). Câmara Técnica de Documentos Eletrônicos. **E-Arq Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão arquivística de Documentos**. Rio de Janeiro: Arquivo Nacional, 2011. 136p.

CUNHA, Murilo Bastos da; CAVALCANTI, Cordélia Robalinho de Oliveira. **Dicionário de biblioteconomia e arquivologia**. Brasília, DF: Briquet de Lemos/Livros, 2008.

DEVMEDIA. Chave Primária e Chave Estrangeira. Disponível em: <<http://www.devmedia.com.br/sql-aprenda-a-utilizar-a-chave-primaria-e-a-chave-estrangeira/37636>>. Acesso em 08 dez. 2016.

DOBEDEI, Vera Lúcia Doyle. **Tesauro: linguagem de representação da memória documentária**. Niterói: Intertexto; Rio de Janeiro: Interciência, 2002. 120 p.

GRUNBERG, Evelina. **Manual de atividades práticas de educação patrimonial**. Brasília, DF: IPHAN, 2007. Disponível em: <http://portal.iphan.gov.br/uploads/publicacao/EduPat_ManualAtividadesPraticas_m.pdf> Acesso em 11 jan. 2016.

RONDINELLI, Rosely Curi. **O documento arquivístico ante a realidade digital: uma revisão conceitual necessária**. Rio de Janeiro: Editora FGV, 2013. 280 p.

SAYÃO, Luís Fernando; SALES, Luana Farias. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN/ IEN, 2015. Disponível em: <<http://www.cnen.gov.br/component/content/article?id=160>>. Acesso em 10 fev. 2016.

STAIR, Ralph M; REYNOLDS, George W. **Princípios de sistemas de informação: uma abordagem gerencial**. São Paulo: Pioneira Thomson Learning, 2006. 646 p.