

Na computação em nuvem são utilizados dois tipos de virtualização para fomentar a sua infraestrutura: paravirtualização e a virtualização completa. Essas virtualizações são realizadas nos servidores dos *datacenters* e nos servidores distribuídos, que por sua vez atuam fornecendo as camadas para os modelos de serviços. (TAURION, 2009)

5.2.1 Modelos de serviço

Com a aplicação da virtualização nos servidores, geram-se três novos modelos de entrega de serviço para os usuários: *Software* como um Serviço – (SaaS), *Plataforma* como um Serviço – (PaaS) e *Infraestrutura* como um Serviço – (IaaS). Esses modelos foram desenvolvidos, a partir da Arquitetura Orientada a Serviço – SOA¹⁴, e são os que compõem a arquitetura da Computação em Nuvem.

- **Software como um Serviço:** *Software as a Service* – SaaS, é o modelo em que um aplicativo está hospedado como um serviço para que os usuários possam acessá-lo por meio de dispositivos que se conectem a *Internet*. Os *softwares* aplicativos são hospedados nos servidores virtualizados dos provedores, possibilitando o acesso por meio de navegador Web. Como o *software* está na Web, ele pode ser acessado pelos usuários de qualquer lugar e a qualquer momento, permitindo maior integração entre as tecnologias utilizadas pelos agentes dos setores de uma mesma instituição. Um grande exemplo do SaaS são os aplicativos da Google, como: Gmail, Google Docs e entre outros da família *Google Cloud*.
- **Plataforma como um Serviço:** *Platform as a Service*, este serviço é específico para o desenvolvimento de um aplicativo, banco de dados e serviços na Web, ou seja, serviços que auxiliam o SaaS, como troca de

¹⁴ Faz referência a um conceito que leva um conjunto de princípios e metodologias aplicadas por engenheiros de *softwares* para projetar e desenvolver *softwares* que forneçam serviços interoperáveis.

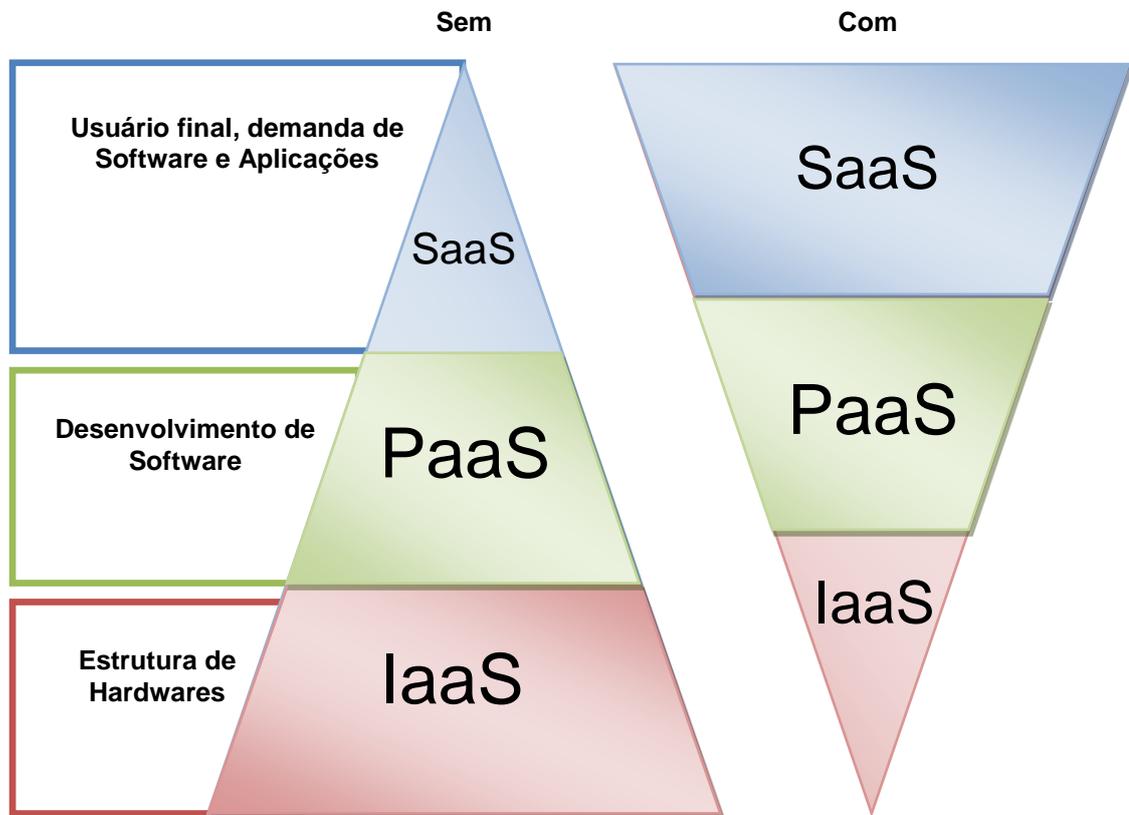
mensagens, serviço de informação, de conectividade, e muitos outros, isto é, os serviços que dão suporte ao funcionamento dos aplicativos. A PaaS oferecem uma plataforma de desenvolvimento onde o software pode ser desenvolvido, testado e implantado, ou seja, o ciclo de vida de um software pode ser operado em um PaaS. Esse modelo fornece todos os recursos necessários para o desenvolvimento de *softwares* aplicativos e conseqüentemente os serviços, sem a necessidade de instalação de *software*, tudo feito pela grande rede, ou seja, nuvem. Um excelente exemplo de PaaS é o Google *Cloud* que provê aplicações na infraestrutura da Google para o desenvolvimento de outros aplicativos baseados em nuvem.

- **Infraestrutura como um Serviço:** *Infrastructure as a Service* - IaaS, fornece recursos de *Hardwares* para o usuário aumentar o poder computacional da sua máquina ou servidor. O processamento, memória e armazenamento são configurados de forma escalável, isto é, o usuário pode configurar os recursos de acordo com a demanda de serviços. Com a customização da infraestrutura tecnológica para gerir os serviços de uma instituição, tudo se torna mais simples para os usuários que não detem o conhecimento aprofundado na TI. Em outras palavras, Poderíamos dizer que a IaaS seria *Hardwares* virtualizados como um serviço que mantém os *softwares*, Sites e Servidores na nuvem. O Google *Cloud*, como exemplo, inaugurou recentemente a sua IaaS, fornecendo esse e outros modelos de infraestrutura.

Esses três serviços compõem a base da computação em nuvem que reformulou os meios de pensar e executar a tecnologia, de forma que venha a facilitar a interação do usuário com as partes técnicas da TI e eliminar a complexidade da logística computacional.

Veremos a seguir uma figura que retrata os três serviços sem a utilização da virtualização, como também os serviços de virtualização junto à internet, proporcionando assim a Computação em Nuvem.

Figura 15 – Representação dos serviços sem e com aplicação de virtualização.



Fonte: Dados da pesquisa, 2013.

Como podemos perceber, o IaaS sem a virtualização necessitava de uma ampla estrutura de *Hardwares* e, logo, de pessoas para gerenciá-los para sustentar as plataformas de desenvolvimento de softwares e aplicativos que, por sua vez, geram uma variedade pequena de *softwares* para os usuários. Entretanto, com a virtualização e o embarque dos serviços em nuvem, a estrutura de *Hardwares* foi potencializada, ocorrendo uma redução no número de pessoas para gerenciamento dos *Hardwares* e na quantidade da estrutura e equipamentos físicos.

A PaaS foi embarcada na estrutura virtualizada da IaaS e proporcionou um modelo mais dinâmico de desenvolvimento, ocasionando um aumento na gama de *softwares* ofertados para os usuários (SaaS) e para mercado tecnológico.

Segundo Hurwitz et al (2010), com a adesão e o desenvolvimento da computação em nuvem no mercado informático, outros três tipos de serviços de ambiente computacional foram ampliados na Web:

- **DaaS - Desenvolvimento como um serviço:** É um serviço que fornece ferramentas de desenvolvimentos de aplicativos de forma simples e compartilhada. Um exemplo deste serviço é o Cloud 9¹⁵ ou Outsystem¹⁶.
- **CaaS – Comunicação como um Serviço:** São serviços de comunicação unificada para ambientes corporativos oferecidos pelos provedores e fabricantes. Exemplos: VOIP, Chat, SMS-Mail e outros.
- **EaaS – Tudo como um Serviço:** É a utilização de todas as tipologias de serviço para desenvolvimento da T.I.C.

5.2.2 Tipos de nuvem

Todas as instituições adotam políticas de restrições, isso é pelo fato de não desejarem que todos os usuários acessem e utilizem determinados recursos nos seus sistemas ou setores. Seguindo este sentido, a computação em nuvem oferece os mais diversificados recursos de segurança que atendem as necessidades mais básicas até as mais restritas, onde somente alguns usuários devidamente autorizados possam utilizar todos os serviços providos.

Essas questões de acesso, disponibilidade, segurança e outros fatores estão imbuídos na escolha do tipo de nuvem a ser utilizada pelo usuário. Neste ambiente computacional há quatro diferentes tipos de nuvem a ser implementados. As restrições ou abertura de acesso ficam a cargo do processo de negócio e políticas realizadas pelo usuário ou instituição e pelo tipo de informação. (TAURION, 2009)

¹⁵ <https://c9.io/>

¹⁶ <https://outsystems.com/>

Taurion (2009) e Veras (2013), entram em acordo sobre a existência de quatro tipos de nuvem e elas podem ser classificadas como: Pública, Privada, Comunitária e Híbrida.

- **Nuvem Pública:** É um modelo de computação em nuvem que distribui para o usuário que deseja processar, armazenar e/ou compartilhar seus dados em um sistema que forneça recursos da TI. Toda infraestrutura é oferecida como um serviço de informação, onde outros usuários poderão ou não ter acesso. As aplicações de todos os usuários ficam reunidas nos mesmos sistemas de armazenamento, o que pode representar uma ineficiente a primeira vista. No entanto, se a implementação de uma nuvem pública considerar questões fundamentais, como desempenho no processamento e segurança, a existência de outras aplicações sendo executadas na mesma nuvem permanece transparente tanto para os prestadores dos serviços como para os usuários.
- **Nuvem Privada:** Este tipo de infraestrutura de nuvem é, geralmente, utilizado exclusivamente por instituições, sendo esta nuvem provida em sistema local ou remoto e gerenciado pela própria empresa ou por terceiros. Isto implica dizer que a instituição é responsável pela gestão, operação, segurança e todas as atividades meio desempenhadas pela instituição. Neste tipo de nuvem são agregadas políticas de acesso aos serviços. As aplicações dessas políticas são realizadas de forma técnica, ou seja, utiliza-se dos recursos de gerenciamento de redes ou de aplicativos que simulam os serviços de rede para configurar os provedores a modo que reforce as tecnologias de autenticação e autorização.
- **Nuvem Comunitária:** Esta infraestrutura de nuvem é compartilhada por diversas instituições, suportando uma comunidade exclusiva que compartilha seus interesses como: a missão, requisitos de segurança, políticas, entre outros. Essa infraestrutura pode ser gerida pelas próprias instituições ou por terceiros que ofereçam o serviço de gerenciamento.
- **Nuvem Híbrida:** Esta nuvem é a junção das Nuvens Pública, Privada e/ou Comunitária. A nuvem híbrida permite que uma nuvem privada tenha seus recursos estendidos a partir da nuvem pública ou comunitária. Em outras

palavras, os usuários ou instituições que utilizam a nuvem privada podem executar as aplicações em uma nuvem pública ou comunitária e realizar o armazenamento dos dados na sua nuvem privada. Esta portabilidade é permitida a partir da configuração dos proprietários da nuvem privada.

Podemos perceber que as nuvens são como ferramentas que podem ser aplicadas para nos auxiliarem nas mais diversificadas atividades do nosso cotidiano. Não obstante, essas ferramentas estão disponibilizadas a todos, isto é, para usuários e instituições que pretendem utilizar os meios terceirizados ou não.

5.3 UTILIZANDO A NUVEM

É notório o crescente uso da computação em nuvem por instituições públicas e, principalmente, por instituições privadas. Isto nos remete a grande adoção e aceitação desta infraestrutura, tanto do mercado tecnológico e empresarial quanto dos usuários comuns que a utilizam para as mais diversas finalidades.

Para aqueles que almejam montar uma nuvem privada, apenas irá depender dos projetos para estruturá-la e das necessidades dos usuários ou instituições. Essa estrutura pode ser implantada e acessada por qualquer instituição ou usuário que esteja devidamente cadastrado. O sistema oferecerá a demanda de *softwares* e aplicativos para as mais variadas soluções que vão de implementações de redes sociais a sistemas de gerenciamentos Governamentais e Militares. O Governo brasileiro já aderiu à tecnologia da computação em nuvem como podemos notar na seguinte matéria da Agencia Brasil da Empresa Brasil de Comunicação - EBC:

De acordo com informações precisas do Serviço Federal de Processamento de Dados (Serpro), a partir de 1º de setembro será lançada a primeira computação em nuvem (cloud computing) do governo federal. O ambiente abrigará, a priori, sistemas para o Programa Cidades Digitais. A tecnologia vai oferecer soluções de educação, atendimento médico hospitalar, gerenciamento e comunicações para cerca de 200 municípios brasileiros. [...] de acordo com Mazoni, presidente do SERPRO, o pacote de serviços inclui sistemas de ouvidoria, gerenciamento de saúde básica com integração ao cartão único de saúde, educação, sistema de gerenciamento escolar e suíte de comunicação. Para o presidente do Serpro, a nuvem brasileira está em um período de estabilidade e plenamente preparada para um acesso três vezes maior do que o previsto. Mazoni disse ainda que o governo brasileiro precisa trabalhar com tipos de software que sejam auditáveis, mas que acima de tudo não tenham compromisso com outros países. (CRISTALDO, EBC, 2013) Acesso em: 20 de Janeiro de 2014.

Com a criação da nuvem privada do SERPRO, os servidores federais irão executar seus serviços em um modo mais dinâmico, pois a interoperabilidade dos serviços será gerida sobre um mesmo ambiente.

Além de fornecer *softwares* e aplicativos que são os elementos de processamento, a computação em nuvem amplia os meios estratégicos para o armazenamento dos dados. Esses dados podem ser armazenados nos servidores virtualizados e nos servidores físicos, podendo ainda ser copiados, como *Backups* nos *Datacenters*, onde apenas os administradores terão acesso a esses arquivos.

Com o sentido de armazenar dados para um acesso futuro, trilharemos o trabalho unificando as duas áreas abordadas, a preservação digital e a computação em nuvem.

6. PRESERVAÇÃO EM NUVEM

Percebemos, em outros capítulos, que para preservar um objeto digital o profissional da informação necessita de conhecimentos, não apenas na sua área, mas obter habilidades essenciais no campo da informática. Também vimos que o impulso tecnológico possibilitou a criação de sistemas que podem ser ajustáveis para a elaboração de estratégias. (ARELLANO, 2008) Entretanto, essas estratégias são soluções que trazem, ao profissional da informação, opções para melhor gerir o sistema de preservação de objetos digitais.

Com o advento da computação em nuvem as possibilidades para preservar o objeto digital foram ampliadas. Isto se deve ao poder computacional oferecido por essa ferramenta. Como vimos anteriormente, ela virtualiza a infraestrutura de *hardwares*, tornando-a *software* e deixando-a escalável de acordo com a necessidade do usuário. Com essa nova tecnologia, o profissional da informação ganha mais possibilidades para elaborar estratégias em seu plano de preservação. (VERAS, 2013)

A computação em nuvem possui múltiplos atributos para o campo da Ciência da Informação, pois com ela podemos potencializar o armazenamento de dados, tráfego de dados em uma rede, ampliar a segurança ao acesso, entre outros. Para que essa ferramenta venha ser um trunfo para o profissional da informação, ele terá que discernir as estratégias e modelos que serão adotados em seu plano de preservação para que utilize todo potencial desta tecnologia.

Diante disso, neste capítulo abordaremos os mecanismos referenciais para a preservação em nuvem. Discorreremos ainda sobre os repositórios digitais que são imprescindíveis para preservar o objeto digital. Por fim, apresentaremos o recurso estratégico para a preservação digital em nuvem.

6.1 MECANISMOS DE REFERÊNCIA PARA PRESERVAÇÃO EM NUVEM

Todos os dias milhares de documentos são gerados devido às atividades realizadas por instituições, órgãos, pessoas, entre outros. Esse volume crescente de documentos a cada dia nos faz pensar sobre os meios com os quais iremos armazená-lo e acessá-lo futuramente. Para Arellano (2004, p. 2004), “muitos dos acervos e coleções digitais estão sendo construídas fora das bibliotecas por diferentes organizações”. Arellano (2008), também afirma que os profissionais da informação encontram-se indecisos em qual estratégia adotar para preservar esses registros ao longo dos anos. As referentes afirmações assentam os profissionais da informação na questão mais vigente da era da informação, como preservar o documento digital?

Em um debate oferecido pela 2ª Jornada Arquivística, em Novembro de 2013, na Universidade Estadual da Paraíba, com o tema: Conservação e Preservação em fotografias, o professor Alberto Banal, convidado para o debate e para ministrar sobre seu projeto nas comunidades quilombolas da Paraíba, indagou os participantes do evento perguntando-lhes: como preservar um acervo com mais de 25 mil fotografias digitais e que a cada semana, decorrente as atividades do projeto, acrescenta-se mais 500 fotografias? A pergunta do professor Alberto Banal lançada aos convidados, que entre os quais estavam professores e alunos do curso de arquivologia da Universidade estadual da Paraíba, apenas fez realçar aquilo que Arellano (2008), afirma no início deste subcapítulo.

Entre afirmações e indagações, elencamos algumas questões que os profissionais da informação refletem diante de situações similares ao que o convidado deixou no debate. Estes são apenas alguns pensamentos sobre os métodos estratégicos mais comuns de armazenamento e, conseqüentemente, de preservação que estamos habituados a executar:

- i. Poderíamos arquivar esses documentos no mesmo dispositivo que o gerou? Isto é, no conjunto de *hardware* e *software*?
- ii. Esse dispositivo arquivaria toda minha produção documental?

- iii. Teríamos que realizar *backup*, isto é, migrar os dados para mídia óptica ou outro tipo de dispositivo de armazenamento para formatar o HD e recomeçar o armazenamento de novos documentos?
- iv. Comprariamos outro HD?
- v. Se fossemos uma instituição, enviaríamos todos os documentos produzidos para o servidor e estaria tudo resolvido?

Embora as perguntas sejam claras e objetivas, as respostas trazem um aspecto subjetivo em cada individuo que as responderem, pois uma afirmação imediata traria consequências drásticas futuramente, pelo fato de não estarem realizando uma prática eficiente para preservação dos documentos.

Entretanto, se as práticas de preservação de documentos estão se tornando obsoletas, como preservar os documentos em um espaço que seja ilimitado e renovar os métodos de preservação? Uma das soluções é o armazenamento em nuvem.

Mesmo que o armazenamento em nuvem traga sua infinidade de possibilidades para preservar os dados, o profissional da informação terá que predefinir, no plano de preservação, mecanismos que o auxiliarão na gestão do sistema de preservação em nuvem e que evitará a perda de qualquer objeto digital. Esses mecanismos são estudados desde 1997, quando a iniciativa *Preserving Access for Digital Access – PADI*, pertencente a biblioteca nacional da Austrália, “decidiu fomentar mecanismos para auxiliar a garantir que a informação em formato digital seja gerenciada para a preservação e o acesso futuro”. (LYALL; BRANDIS, 1998, apud THOMAZ, 2004, p. 48)

Segundo Katia Thomaz (2004, p. 128), os cinco mecanismos imprescindíveis para a montagem de uma infraestrutura de preservação digital são: “sistema de cópia de segurança, sistema de armazenamento redundante, sistema de detecção e recuperação automática de falhas, sistema de segurança de acesso físico e lógico e o sistema hierárquico de armazenamento”.

Todas as ferramentas que proporcionam a preservação digital, possivelmente, possuem algum destes mecanismos que trilha o profissional da informação a realizar

um modelo sistemático ou não de preservação digital. Tendo como base esses mecanismos aplicados em conjunto a determinadas ferramentas, a perspectiva de preservação eleva-se exponencialmente oferecendo uma estrutura de preservação digital segura. Arellano (2008, p. 51), ainda ressalta dizendo que para os “detentores de acervos digitais, é cada vez mais imperiosa a necessidade de contar com mecanismos que garantam a preservação de seus documentos em formato digital”.

Deste modo, percebemos que para preservar em nuvem o sistema deverá possuir estes mecanismos citados acima. No entanto, a própria tecnologia da computação em nuvem se encarrega de acrescentar outros dois pontos que corroboram para uma melhor eficácia na preservação digital e na recuperação. A nuvem abarca para a preservação digital, o sincronismo e a multiplataformas, proporcionando uma interação completa, segura e ágil. (VERAS, 2013) Para compreendermos mais sobre esses mecanismos, descreveremos a seguir cada um deles com o foco dirigido para a nuvem.

- ✓ **Sistema de Cópia de Segurança (Backup):** É a reprodução fidedigna de dados de um dispositivo de armazenamento. A cópia de segurança é necessária para que, por ventura os dados sejam perdidos, apagados acidentalmente, sofreram alguma ação sinistra ou foram corrompidos, venham ser recuperados a modo que substitua os originais. (FARIA, 2009)
- ✓ **Sistema de Armazenamento Redundante (RAID):** É um conjunto redundante de discos de armazenamento independentes. Este sistema cria um subsistema de armazenamento com dois ou mais discos (serviço de armazenamento em nuvem) distintos, mas quando integrados proporcionam um trabalho simultâneo no serviço de backup. Isto é, o sistema sincroniza dois discos rígidos, ou duas ou mais nuvens de armazenamento, proporcionando um espelhamento de dados, pois tudo que for armazenado em disco rígido ou uma nuvem, essa replicará automaticamente para o disco rígido ou a nuvem seguinte, caso ocorra falha em um disco rígido ou nuvem, a outra assumirá a execução dos dados. Sendo assim, as ferramentas dos serviços de armazenamento em nuvem oferecem ao administrador, a escolha de uma redundância selecionada ou sincronizada. O administrador pode

escolher os diretórios de dados que serão espelhados. (RAINER;CEGIELSKI, 2012)

- ✓ **Sistema de Detecção e Recuperação de Falhas:** É o reconhecimento de possíveis falhas no sistema de armazenagem e a sua correção automática. Geralmente, as ferramentas disponíveis em nuvem, ou que gerenciam uma nuvem ou uma rede, possuem este recurso que auxilia o administrador a gerir e corrigir o sistema e os dados armazenados. (RAINER;CEGIELSKI, 2012)

- ✓ **Sistema de Segurança de Acesso:** Pode ser classificado como sistemas que fornecem serviços protocolizados para proteção ou restrição dos dados. Também pode ser compreendido como um conjunto de procedimentos que visam garantir um acesso controlado de entrada, uso e saída dos usuários e administradores do serviço de qualquer sistema de informação. Geralmente, os serviços em nuvem possuem registro de toda atividade do administrador e dos usuários que detém o acesso compartilhado à nuvem, datando desde o horário de entrada, identificação, quais documentos produzidos ou alterados, até a hora da saída do sistema. (PAIVA, 2008)

- ✓ **Sistema Hierárquico de Armazenamento:** É um sistema que divide e define vários níveis de armazenamento, objetivando a proteção, preservação, frequência de uso dos dados, entre outros. Cada uma dessas partes pode ser entendida como classe que contém seus respectivos níveis. Os dados armazenados podem ser alocados para as diferentes classes estipuladas, isto é, com base no seu caráter intrínseco e pela temporalidade. Portanto, os dados com um valor mais corrente devem ser alocados na primeira classe. Os dados que de consulta esporádicas são alocados para a classe intermediária. Os dados que necessitam uma maior proteção e que não estão sendo consultados devem ser alocados na terceira classe. (EMC, 2009)

- ✓ **Sincronização:** É o gerenciamento de múltiplas ações ou processos executáveis do mesmo recurso ou de recursos diferentes, integrando-os para obter um resultado. No que toca a nuvem, a sincronização é utilizada para integrar um sistema ou aplicativo a outros sistemas, possibilitando o compartilhamento dos aplicativos, dos dados, espelhamento dos dados, expansão de espaço de armazenamento, ampliação de ferramentas para execução das atividades. As APIs são um grande exemplo que oferece este recurso de sincronização. Um exemplo mais claro é o Google drive que permite os arquivos serem sincronizados com Microsoft Office, possibilitando o usuário executar o mesmo documento e realizar as tarefas em ambas as plataformas. (TAURION, 2009)

- ✓ **Multiplataformas:** É um aplicativo ou sistema que consiste em executar mais de um sistema operacional. Estes aplicativos emulam outros sistemas que por sua vez executam outros aplicativos ou arquivos. Sem a utilização dos aplicativos emuladores o arquivo não poderia ser executado. Estes aplicativos são chamados de máquinas virtuais. Um exemplo de um sistema multiplataforma é o Java Virtual Machine - JVM. Para o nosso caso, seria um documento ser acessado por plataformas ou sistemas diferentes sem que haja erro. (HUMBLE; FARLEY)

Visto que os cinco mecanismos se encontram presentes ou passíveis de implementação na infraestrutura de computação em nuvem, podemos afirmar que a preservação feita pela nuvem passa a ser uma referência como estratégia de preservação.

6.2 ANALISANDO O ACERVO FOTOGRÁFICO DIGITAL

Sabendo da referente apresentação e da reflexiva e contundente provocação do professor Alberto Banal¹⁷ na 2ª Jornada Arquivística – UEPB e tomando conhecimento de seu acervo digital do projeto social que mantém a memória das comunidades quilombolas da Paraíba, compreendemos que a sua produção documental é vasta, resolvemos conhecer o seu arquivo e questioná-lo sobre os procedimentos adotados por ele para preservar os documentos digitais.

Realizamos a entrevista em seu escritório, onde nos foi apresentado a sua infraestrutura tecnológica de gestão e armazenamento. A primeira vista, vislumbramos um bom ambiente tecnológico composto por um Computador desktop robusto, um Laptop, um Scanner, cinco HD's externos ativos, um Modem e um Roteador. Estes equipamentos fornecem a sua rede local.

O entrevistado nos relatou que dois HD's externos são interligados ao desktop para ampliar o espaço de armazenamento e os outros três HD's externos são *backups* do HD interno do desktop e dos outros dois externos. O Scanner é utilizado para digitalizar o acervo analógico. O Modem fornece a conexão com a internet e esta é distribuída pelo roteador para todo ambiente.

Em meios às perguntas lançadas ao entrevistado, notamos que a sua compreensão sobre a tecnologia é razoável, pois, quando perguntado se teria algum conhecimento sobre computação em nuvem, o mesmo respondeu que utilizava um serviço simples em nuvem para armazenar e disseminar uma pequena parte do seu trabalho e que já pensou armazenar seu acervo em nuvem, mas o valor do aluguel, possivelmente, seria mais oneroso que manter seus dados em seu próprio escritório.

No decorrer da entrevista, percebemos que o gestor realizava algumas estratégias operacionais básicas de preservação como a migração dos dados para outras mídias de armazenamento. O mesmo realizava a migração dos HD's online para os HD's de backup. No entanto, o que nos preocupou foi que o HD's de backup

¹⁷ Fotógrafo Italiano que trabalha com Comunidades quilombolas, desde 2005, na Paraíba. Nesta palestra a provocação para reflexão foi a seguinte: Gostaria de saber o que vocês arquivistas podem fazer para organizar e preservar o meu acervo que possui mais de 150 mil fotos?

permanecem conectados ao computador e isso é um fator de risco extremo. O próprio gestor afirma que o risco é latente, mas que estes são os únicos recursos que ele compreende e os utiliza há anos.

Identificamos que o entrevistado possui uma boa noção em informática, porém não faz uso dos procedimentos metodológico e políticos da preservação digital. Entretanto, encontramos dois problemas que também foram relatados pelo entrevistado. O primeiro se trata da organização e classificação do acervo e o segundo subdivide em três, segurança dos dados, migração dos suportes e uma política de *Disaster Recovery*¹⁸, isto é, preservação digital.

Enxergando a sugestiva problemática do acervo fotográfico digital e levando como um desafio à pergunta lançada pelo professor Banal na palestra realizada na 2ª Jornada Arquivística, materializou-se, neste momento, a referente discussão da pesquisa sobre a preservação digital, até então teórica.

Uma vez materializado o problema desta pesquisa, resolvemos apresentar a estratégia conceitual de preservação em nuvem ao entrevistado como uma solução para os referidos problemas de organização e preservação do acervo fotográfico. Frisamos ainda que o referente modelo atende as políticas básicas de *Disaster Recovery*, tão falado pelo entrevistado.

Partindo da ideia de organizar e preservar os objetos digitais do acervo fotográfico: Quilombos da Paraíba, apresentaremos a estratégia de preservação de documentos arquivísticos digitais em nuvem que garantirá a salvaguarda por longo prazo e o seu acesso. No entanto, antes de prosseguirmos com a descrição e apresentação da preservação em nuvem, abordaremos sobre uma peça fundamental para que ocorra a preservação em nuvem, o repositório digital.

¹⁸ São políticas e procedimentos que estão relacionados à preparação para a recuperação de dados ou manutenção de infraestrutura tecnológica, que são vitais para uma organização após desastres naturais ou induzidos pelo homem.

6.3 REPOSITÓRIOS DIGITAIS: O caminho para a nuvem

Para Arellano (2008), um repositório digital é um serviço de armazenamento de objetos digitais que tem capacidade de manter e gerenciar objetos digitais por longos períodos de tempo, provendo o seu acesso apropriado. Os repositórios digitais estão divididos em três classes: “temáticos, institucionais e centrais”. Os temáticos abrangem determinadas áreas do conhecimento. Os institucionais são utilizados como um sistema de informação para armazenamento, preservação e disseminação, cabendo ainda o acesso à produção intelectual digital de comunidades científicas. Já os repositórios centrais, “seriam provedores de serviço que agrupam os objetos digitais dos repositórios temáticos e institucionais”. (ARELLANO, 2008, p.124)

Os estudos sobre os repositórios para preservação da informação digital é crescente a cada dia nos países que são considerados os precursores da Arquivística contemporânea e da Ciência da informação. Canadá, Portugal, Espanha e Estados Unidos possuem projetos de pesquisa focados para a preservação digital com a utilização de repositórios de informação.

No Brasil, uma das pesquisas que está a frente, no que toca a preservação digital e autenticidade dos documentos digitais, sob a coordenação do Arquivo Nacional é o projeto InterPares¹⁹, que realiza uma ação colaborativa internacional para fomentar procedimentos técnicos e teóricos de preservação de documentos digitais e a sua autenticidade nos sistemas de informação, logo, nos repositórios digitais. As universidades brasileiras e os institutos de pesquisa científica adotaram o Dspace²⁰ como repositório padrão para preservação de suas produções acadêmicas e outros materiais digitais. (DURANTI, 2007)

O IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia descreve um repositório digital como: “uma base de dados online que reúnem de maneira organizada a produção científica de uma instituição ou área temática”.

¹⁹ *International Research on Permanent Authentic Records in Eletronic System* (Pesquisa Internacional sobre Documentos Arquivísticos Autênticos) – Visa estabelecer políticas e procedimentos de preservação e autenticidade.

²⁰ *Software* de código fonte aberto, desenvolvido pelo Instituto de Tecnologia de Massachussets, que fornece facilidades para a preservação e o gerenciamento de acervo digital.

Arellano (2008), afirma que os repositórios digitais são “coleções de informação digital, que podem ser construídas de diferentes formas e com diferentes propósitos”.

Os repositórios digitais podem armazenar os mais variados formatos de dados, possibilitando a instituição preservarem arquivos de multimídia ou qualquer objeto digital. Geralmente, esses repositórios digitais são alimentados com a produção científica de uma determinada instituição, de uma área temática da instituição ou independente de instituições. (ARELLANO, 2008)

Estes tipos de repositórios foram fomentados a partir da Iniciativa de Arquivos Abertos (*Open Archive Initiative*), cujo objetivo é de criar mecanismos tecnológicos para tornar interoperáveis os diferentes repositórios funcionando segundo a proposta do *Open Archive*. Haja vista que estes “tipos de repositórios foram criados para disseminar a produção acadêmica para a sociedade” e realizar a preservação de toda a produção. Portanto, os repositórios nasceram da necessidade de preservar a memória digital e oferecer visibilidade à instituição, contudo, esses repositórios não se limitam apenas as instituições ou organizações científicas, mas para toda sociedade. (MARCONDES; SAYÃO, 2009)

Os repositórios digitais podem ser desenvolvidos e/ou configurados também para as atividades de gerenciamento administrativas, pois, uma vez que os repositórios são *Softwares Open Source* (Código Aberto), são passíveis a uma reprogramação para suprir as necessidades da instituição, organização ou do usuário funcional. Nos repositórios digitais, geralmente, é encontrado o modelo sistêmico referencial de preservação, o OAIS. Com o modelo OAIS integrado aos repositórios, ganha-se uma grande base sólida e eficiente para preservar os objetos digitais, como afirma Arellano (2008, p. 135):

A aplicação dos princípios do modelo OAIS, e particularmente, a implementação de um arquivo de acesso livre em concordância com os modelos de funcionalidade e estrutura da informação do OAIS é o pré-requisito chave para estabelecer repositórios confiáveis e garantir a preservação de longo prazo dos seus atributos.

Arellano (2008), também afirma que para garantir a confiança nos repositórios digitais por parte dos autores e dos usuários, devem-se manter alguns requisitos

mínimos da preservação digital, tais como: autenticidade, integridade, confiança, disponibilidade e capacidade de reuso. Com esses requisitos integrados aos repositórios, teremos a inclusão, não apenas dos objetos digitais, mas de toda estrutura digital do documento, isto é, metadados e identificadores descritivos.

Os repositórios digitais utilizam esses metadados e identificadores descritivos, que são construídos a partir do protocolo OAI-PMH²¹, para facilitar a recuperação dos documentos arquivados nos repositórios que utilizam o referido protocolo da Iniciativa de Arquivos Abertos (OAI) em seus mecanismos de busca, tornando o sistema mais preciso e satisfatório para o usuário. (ARELLANO, 2008)

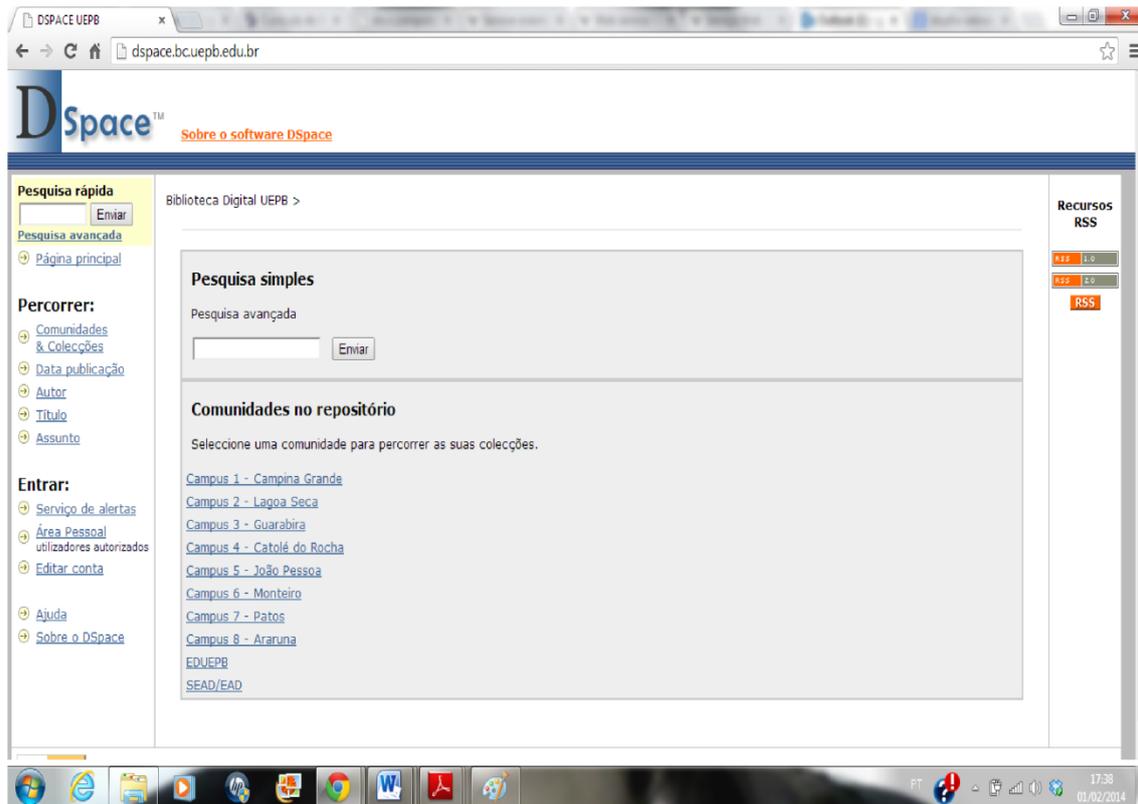
Atualmente, há vários *softwares* que podemos encontrar na internet para criarmos um sistema de preservação digital, tais como: Greenstone, Invenio, Fedora, RODA, OCLC e outros. No entanto, resolvemos abordar o repositório mais utilizado pelas instituições públicas brasileiras que segundo Grácio, (2011) e Arellano (2008), é o Dspace.

O Dspace é um software de código aberto que possui a funcionalidade primária de preservar os objetos digitais em seus variados tipos e formatos. Criado pela Massachusetts Institute of Technology (MIT) *Libraries* em parceria com o laboratório da *Hewlett-Packard* (HP), é utilizado como repositório institucional para Arquivos, Bibliotecas e Centro de pesquisa e informação. Sua funcionalidade secundária está atrelada a configuração adotada pela instituição, pois o Dspace pode ser configurado para atender todo o conjunto das unidades administrativas, proporcionando além da preservação, uma gestão eletrônica de documentos. (ARELLANO, 2008).

A seguir, veremos dois modelos de DSpace em atividade. O DSpace da UEPB utiliza o modelo padrão, isto é, não foi configurado ou editado para oferecer um mais dinamicidade e conforto visual para o usuário. Já o LUME da UFRGS, foi modificado, reconfigurado e ampliado, pois além de realizar a preservação facilita a busca para o usuário, oferecendo um design visual e informacional dinâmico, deixando-o atrativo.

²¹ É um protocolo desenvolvido pela Open *Archival Initiative*-OAI, que define um mecanismo para coleta de registros de metadados em repositórios para uma recuperação futura.

Imagem 3 – DSpace Institucional da Universidade Estadual da Paraíba



Fonte: <http://dspace.bc.uepb.edu.br>

Imagem 4 – DSpace modificado para atender as necessidades de preservação da instituição.



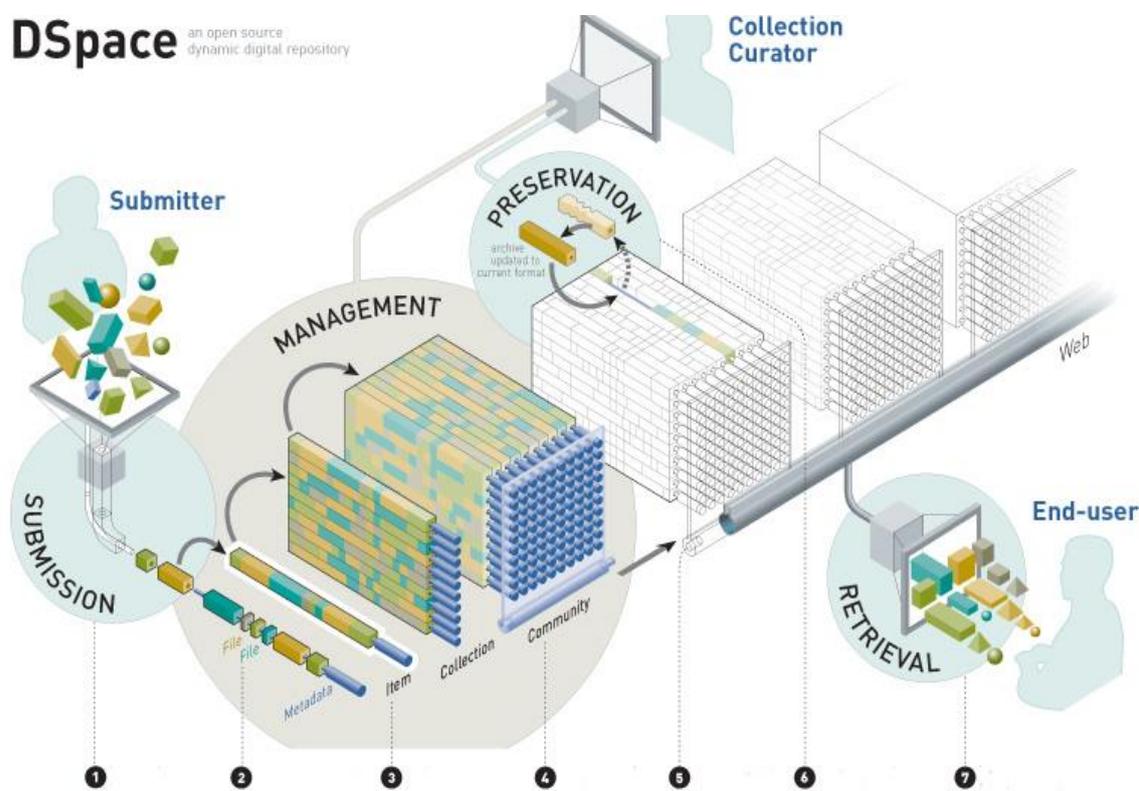
Fonte: <http://www.lume.ufrgs.br>

Assim como a UEPB utiliza o Dspace, o LUME e outras universidades em todo o mundo fazem uso deste *software* que foi implementado com os padrões, *Open Archive Initiative* (OAI), *Open Archival Information System Reference Model* (OAIS) e o *Dublin Core*²², que são aceitos internacionalmente. Com o uso do modelo OAIS integrado no Dspace transforma-o em um sistema de arquivamento digital, devido aos mecanismos referenciais de descrição e preservação digital. (ARELLANO, 2008)

O uso ou a preservação no Dspace é realizado a partir do modelo referencial OAIS, pois este modelo garante a gestão, a preservação e uma recuperação mais eficiente. Vejamos como ocorre o processo sistemático no Dspace.

O uso ou a preservação no Dspace é realizado a partir do modelo referencial OAIS, pois este modelo garante a gestão, a preservação e uma recuperação mais eficiente. Veremos abaixo como ocorre o processo sistemático no Dspace.

Figura 16 – Processo sistemático funcional do Dspace



Fonte: dspace.org

²² É uma organização iniciativa que delimita padrões de metadados que visa descrever os mais variados objetos digitais. O Dublin Core utiliza a linguagem de marcação XML para interoperação e identificação dos documentos no sistema. Também utiliza-se a linguagem RDF para representar os metadados na Web.

Percebe-se que o processo sistêmico do Dspace, uma vez que são imbuídos os padrões da OAI, assemelha-se com o modelo referencial de preservação OAIS. Podemos ver ainda que o processo se divide em sete partes.

A primeira parte é da submissão (*SUBMISSION*), que é realizada por uma interface web, permitindo o autor ou gestor (*SUBMITTER*) inserir os objetos digitais. Feito isso, partimos para a segunda parte que consiste em os objetos digitais converterem-se em um fluxo de bits ou *bitstream*, que serão organizados em conjuntos relacionados. Cada *bitstream* tem um formato técnico e outras informações técnicas serão atreladas nele. Esta informação técnica será mantida no *bitstream* para auxiliar com a preservação ao longo do tempo e facilitar na sua recuperação futura. (ARELLANO, 2008)

A terceira parte consiste na relevância dos metadados. Após o bitstream receber as informações técnicas, este objeto digital será considerado como um item (*ITEM*), ou seja, um “átomo de arquivo” que consiste em agrupar conteúdos relacionados e descrições associadas, isto é, metadados. Esses metadados são adicionados para indexar o objeto digital em seu respectivo acervo ou coleção (*COLLECTION*) relacional lógico.

A quarta parte infere nas questões organizacionais. Segundo Arellano (2008, p.141), “o DSpace faz uso da terminologia de comunidades para definir os espaços dos conteúdos onde os usuários podem localizar os documentos”. Esta é considerada uma das partes mais relevantes do DSpace, pois ela corresponde à organização estrutural de todo o arquivo digital, isto é, condiz a qual comunidade (*COMMUNITY*), departamento, setor, laboratório, está acondicionado o objeto digital. Não obstante, é nesta parte também que o gestor organiza o formato documental pertencente à comunidade.

A quinta parte não refere-se a um processo, mas a uma funcionalidade ou recurso oferecido. A arquitetura modular do DSpace permite a criação de grandes repositórios ou de repositórios distribuídos, isto é, repositórios integrados através dos recursos de redes computacionais, podendo ser uma rede local (LAN), rede metropolitana (MAN) ou rede a longa distância (WAN). Este processo é garantido por conta dos protocolos, tanto de redes quanto dos repositórios, que “permitem a

distribuição destes repositórios e/ou arquivos digitais unificados pelos mecanismos da interoperabilidade”. (ARELLANO, 2008, p.102)

A sexta parte incide em um processo chamado de *workflow*²³. Este processo consiste em aceitação ou rejeição, revisão, alteração, autenticação e todos os procedimentos predefinidos pelas políticas da instituição, empresa ou organização. Todo esse processo é gerenciado por uma entidade (*COLLECTION CURATOR*) que conheça as políticas predefinidas e adotadas para submissão e gestão do conteúdo.

O DSpace se compromete em ir mais além na preservação (*PRESERVATION*) e acesso dos objetos digitais. Também é nesta parte que o objeto digital é atualizado para um formato mais acessível para o dispositivo a qual o usuário esteja utilizando. Por exemplo: Se um documento foi armazenado no DSpace como formato *Word Microsoft Office (.doc)*, o gestor do Dspace irá optar o melhor formato para que o usuário tenha o acesso ao documento. Neste caso seria a atualização de Word para PDF, possibilitando a execução do documento a partir de qualquer dispositivo que realize a busca no repositório. Caso fosse um objeto audiovisual, este poderia ser submetido como um arquivo WMV e seria atualizado para MP4 que é uma extensão midiática mais acessível para ser executado em qualquer dispositivo digital.

Esta sexta parte torna-se o destaque do DSpace devido a este processo de atualização contínua do objeto digital. Isso faz que os documentos submetidos ao DSpace estejam sempre acessíveis e interpretáveis em qualquer dispositivo que o acesse. Se caso um novo formato entrar em uso corrente no ambiente digital e se for subtendido que esse novo formato é acessível por todos os dispositivos, o Dspace atualizará para o referente formato corrente. Esse recurso nos dá a certeza e a segurança de que o objeto digital sempre será interpretado quando recuperado. (ARELLANO, 2008)

Após os processos de submissão, conversão em fluxo de bits, atrelar os metadados descritivos, organizar e classificar o documento digital e atualizar para um formato mais acessível, chega à parte da recuperação (*RETRIEVAL*). Nesta parte, refere-se ao usuário final (*End-user*), que a partir de uma interface web com o

²³ Consiste em um conjunto regras e procedimentos necessários para alcançar a automação dos processos de negócios.

auxílio de um mecanismo de busca, os documentos digitais em seus variados formatos podem ser exibidos no navegador ou descarregados no dispositivo do usuário, isto é, os documentos textuais e iconográficos podem ser visualizados pelo navegador. Já os documentos audiovisuais podem ser descarregados para o dispositivo do usuário, deixando a cargo do usuário escolher qual *software* será utilizado para executar o arquivo descarregado.

Diante deste processo sistêmico do DSpace, percebemos que os repositórios são ferramentas essenciais para a preservação digital e principalmente o DSpace devido a sua versatilidade.

Embora o DSpace seja imponente diante de outros repositório e métodos de preservação, ainda nos faz refletir sobre a sua capacidade de armazenamento, pois o seu limite é a capacidade do servidor físico aonde ele esteja instalado. Se quisermos uma ferramenta que armazene, ou melhor, realize todos os processos que o DSpace executa e tenha um espaço ilimitado, a solução não é um *software*, mas sim a infraestrutura de *hardware* e como ela será gerenciada.

Partindo do princípio de que necessita-se manter os objetos digitais do acervo fotográfico preservados no servidor local e uma cópia fiel em um lugar remoto, de modo que, no momento em que os objetos sejam arquivados no servidor local, uma cópia também é gerada automaticamente no servidor remoto. Não obstante, ainda temos que pensar sobre a recuperação desses objetos digitais e que seja amplamente acessível pelo seu gestor. Para tanto, resolvemos implementar, de forma estratégica, uma rede WAN que irá prover recursos computacionais para sustentar o DSpace numa infraestrutura de rede virtualizada, tendo seu acesso para arquivamento e recuperação via rede local, intranet ou pela internet, gerando assim o caminho para a preservação dos documentos arquivísticos em nuvem.

6.4 O RECURSO ESTRATÉGICO PARA A PRESERVAÇÃO DE DOCUMENTOS ARQUIVÍSTICOS DIGITAIS

Segundo Veras (2011, p. 11), um dos conceitos chaves da transformação da TI é o “conceito de alinhamento estratégico”, respaldado na governança em TI, possibilita a perfeição na execução de projetos. Arellano (2008, p. 51) acrescenta relatando que “o desenvolvimento de padrões e de mecanismos legais para lidar com arquivos eletrônicos precisa de estratégias metodológicas bem definidas”.

Para proceder à estratégia de preservar os documentos em nuvem, devemos construir uma infraestrutura que contenha servidores para sustentar a mesma, *softwares* para gerenciar os processos e fornecer os recursos integrantes, a rede para conectar os servidores que irão prover os recursos para armazenar, preservar por um prazo longo e tornar acessíveis os documentos digitais. A rede também disponibilizará o acesso ao gestor e/ou usuário convidado por meio dos computadores das suas respectivas redes local ou pela internet.

O processo de qualquer projeto na área da tecnologia da informação é longo e complexo. Portanto, faremos a descrição dos processos da infraestrutura para preservação em nuvem de forma clara e objetiva, mas não se desprezando dos detalhes essenciais que são imprescindíveis para o sucesso de todo o trabalho.

Contudo, para iniciarmos a infraestrutura deve-se realizar a escolha do sistema operacional que administrará o servidor, que é fundamental para obtermos uma boa performance na infraestrutura. Este sistema operacional será encarregado por gerenciar os recursos de *hardware*, rede e *softwares* que irão complementar o funcionamento da infraestrutura.

Quanto às questões de *hardware*, uma vez que necessitamos armazenar um número considerável, que no nosso caso são, inicialmente, 25.000(vinte e cinco mil) de documentos digitais, é evidente que precisaremos de um processador com uma frequência de processamento alta, memória equiparada ao processador e um ou alguns discos rígidos (HD) com grande capacidade de armazenamento. (VERAS, 2011)

Baltzan e Phillips (2012, p. 124), dizem que para erguer uma boa infraestrutura devem-se obter algumas características, como “flexibilidade, escalabilidade, confiabilidade, disponibilidade, desempenho”, entre outros.

Então, com a aquisição de *hardware* e *software*, far-se-á a instalação e posteriormente a configuração necessária para que o servidor administre a infraestrutura de forma segura e ágil.

Depois que estabilizarmos o servidor, isto é, certificar de que ele seja gerenciável sob pequenas falhas, atualizável, escalável, realize *backup* de sua configuração, passível para criação de uma ERD²⁴, seguro, processável e suportável a um vasto fluxo de dados e que diagnostique qualquer problema referente à rede a qual gerencie, estabeleceremos as inerentes conexões para firmarmos uma rede privada.

A rede estabelecida será uma “rede longa distância (WAN)”, uma vez que esta tornará responsável por realizar a conectividade entre os continentes sul-americano e europeu, mais precisamente, Brasil e Itália. (ROSS, 2008). Este processo da criação da rede tende a ser minucioso, pois exige um alto conhecimento, no que tange as infraestruturas de redes computacionais e protocolos de interoperação. Qualquer equívoco poderá acarretar em falhas que afetarão na rota de fluxo dos dados e, possivelmente, também permitirá brechas para invasores. Neste ponto iniciam-se as políticas e os procedimentos que a rede irá adotar quanto à segurança da informação.

Se sua empresa está bastante interconectada e considera a TI como estratégia para os negócios, mas ainda não tem políticas de segurança da informação, é importante pensar em criá-las, começando por uma análise de vulnerabilidades para, em seguida, selecionar os controles necessários. [...] **o plano de segurança da informação poderia ser considerado como um adendo do plano de tecnologia da informação.** (ARAGON; ABREU, 2009, p.75 GRIFO NOSSO).

As políticas de segurança da informação, como os autores citam acima, podem ser adicionados no plano de TI, no nosso caso, na rede e ao sistema de preservação digital, isto é, no repositório em nuvem. No entanto, é evidente que a segurança da informação seja imprescindível para o nosso trabalho, mas ela poderá

²⁴ Disco de reparação de emergência. É utilizado para recuperar as configurações de todo o sistema.

ser implementada posteriormente, uma vez que o foco do trabalho não é estabelecer processos, políticas e normalização como a ISO 27000²⁵ e 27001²⁶, mas sim criar uma estratégia de preservação para o acervo Quilombos da Paraíba.

A rede de longa distância será formada a partir dos servidores que gerenciam as redes locais (LAN) dos referidos países. Para criarmos a WAN, far-se-á a utilização da virtualização dos componentes físicos dos servidores locais do Brasil e da Itália. Uma vez aplicada a virtualização, esses servidores virtualizados serão instâncias dos servidores físicos, ou seja, os servidores físicos, em modo sincronizado, gerenciarão os servidores virtuais que irão prover os serviços de armazenamento e acessibilidade dos objetos na web.

Cada servidor virtualizado receberá uma função a qual se dedicará, isto é, a partir da virtualização podemos gerar vários servidores a qual serão distribuídos tarefas. Por exemplo: um servidor para fornecer segurança, um segundo como servidor de e-mail, outro para hospedar aplicações e executá-las em seu ambiente, isto é, criando varias abstrações de servidores, gerando assim uma nuvem de serviços.

No entanto, como o nosso caso é específico em armazenar documentos em um repositório utilizando o ambiente tecnológico virtualizado e criando desta forma a nuvem, possibilitando ainda, o acesso remotamente, utilizar-se-á os servidores virtualizados apenas para instalar o repositório, configurá-lo para prover o seu serviço de armazenar e preservar os documentos e torna-los acessíveis por meio da internet.

Obrigatoriamente, todos os computadores e servidores físicos e virtuais serão roteados para que não ocorram erros no acesso aos serviços disponibilizados ou aos servidores de armazenamento, finalizando desta forma a WAN. Definidas as rotas de toda a rede, implementar-se-á o repositório nos servidores virtualizados possibilitando o seu acesso pela internet. Este processo também é compreendido como web service. (VERAS, 2013)

²⁵ São normas estabelecidas pela ISO sobre políticas de segurança da informação.

²⁶ Normas para sistemas de gestão da segurança da informação.

Uma vez que os servidores, através de configurações, irão disponibilizar o acesso do repositório pela web, o próprio repositório irá prover uma interface que possibilitará o processo de armazenamento e recuperação dos documentos digitais. Isto se deve a programação predefinida do repositório e a um conjunto de *softwares* utilizados para fornecer o DSpace como *Web service*, facilitando, deste modo, o trabalho do gestor e os seus utilizadores. No entanto, a interface do repositório poderá ser customizada para poder atender as necessidades do gestor, no que tange o armazenamento ou a recuperação dos documentos digitais. (ARELLANO, 2008)

O processo de instalação do DSpace torna-se complexo devido ao conjunto de *softwares* utilizados para construção da sua arquitetura. Para uma configuração básica do DSpace, onde sejam permitidos todos os procedimentos relatados no trabalho, o servidor que receberá o DSpace terá que estar configurado com os seguintes *softwares*:

- **Apache Tomcat** – É um servidor de aplicações que tem capacidade para atuar também como um servidor web ou ser parte integrante de um servidor web. O *Apache Tomcat* irá possibilitar que o DSpace seja interpretado como um serviço na web, isto é, fará que ele seja executado a partir de uma página na internet com o protocolo HTTP. (APACHE.ORG, 2013)
- **Apache Maven** – Utilizando o conceito de Projeto Modelo de Objetos – POM, o *Apache Maven* é um gerenciador de projetos que organiza os processos de multitarefas como uma estrutura de diretório. Ele torna todos os processos mais flexíveis evitando, desta forma, erros ou falhas em um programa. Como os documentos no DSpace são “convertido” em formato XML, sua utilização para gerenciar esse tipo de formato é essencial. (APACHE.ORG, 2013)
- **Apache ant** – É uma ferramenta que automatiza a instalação e a construção do DSpace. (APACHE.ORG, 2013)

- **PostGre sql** - É um banco de dados que é utilizado para armazenar os dados inseridos no sistema, isto é, no DSpace. (POSTGRE-SQL.ORG, 2013)
- **JDK** – É uma plataforma JAVA que permite desenvolver ou implementar aplicativos para desktops, servidores, programas e sistemas que necessite do seu ambiente. Uma vez que o DSpace é construído na linguagem JAVA, necessitará desta plataforma para a sua execução. (ORACLE.COM, 2013)
- **Handle Corporation for National Research Initiatives** – É um subsistema de distribuição de informação. O conjunto de protocolos que é contido neste subsistema permite a identificação, o armazenamento e a distribuição. Este subsistema auxilia a organização dos documentos no DSpace, otimizando ainda mais os recursos de preservação e recuperação do DSpace. (CNRI, 2013)

Após a instalação destes *softwares* e as configurações necessárias os servidores estarão prontos para instalar o DSpace. O processo de instalação do DSpace é simples, porém a sua customização se torna exaustiva. O gestor terá que delimitar quantos e quais administradores terão acesso total do sistema, configurar o banco de dados do DSpace, estabelecer os procedimentos de submissão, configurar o motor de busca, definir os parâmetros dos metadados entre outros procedimentos intrínsecos que darão ao DSpace um melhor desempenho.

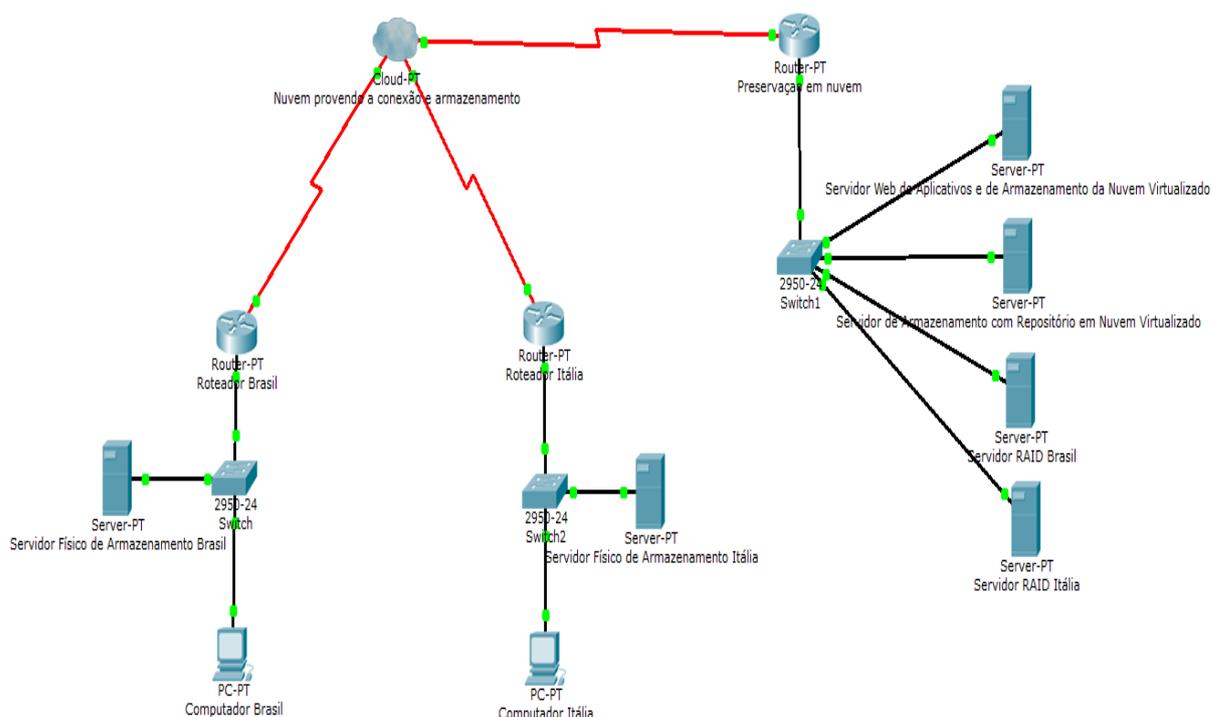
Terminada a instalação e configuração do DSpace, far-se-á testes para observar o funcionamento do referente sistema de preservação. Caso seja encontradas falhas ou mau funcionamento do sistema, o administrador terá que corrigir a partir da configuração ou reinstalação de todo o sistema. Se o teste for positivo, dar-se-á início a aplicação de um servidor DNS²⁷ para o link do DSpace.

²⁷ Domain Name System: É um sistema de resolução de nomes que são atribuídos a número de protocolo de internet (IP), o DNS traduz o número IP em um nome, exemplo: O nome google.com é o DNS de 173.194.45.34.

O DNS, configurado pelo administrador, será responsável por relacionar ou traduzir o endereço IP do servidor onde o DSpace está instalado para um nome que seja mais fácil de lembrar para poder acessar o DSpace pela internet, isto é, ao invés de digitar o IP do computador, o usuário irá digitar um nome no navegador, como por exemplo: Quilombos.com.br. Evidente que se o administrador irá requerer um domínio próprio e para tanto, terá que realizar o registro deste domínio.

Finalizado o processo de instalação do DNS, encerra-se o processo da arquitetura de preservação em nuvem. Para visualizarmos melhor a referente infraestrutura, veremos abaixo a figura que a descreve.

Figura 17 – Infraestrutura do recurso estratégico de preservação digital em nuvem de documentos arquivísticos.



Fonte: Dados da pesquisa, 2013.

Como é visto na figura acima, utilizando o mínimo de recursos, a rede local brasileira e a rede local italiana se conectam, a partir dos seus respectivos roteadores, gerando assim uma rede de longa distância que provê a comunicação entre os computadores e os servidores dos distintos países. Até este ponto, os

procedimentos são simples e de fácil compreensão, pois é desta forma que se estabelece uma conexão a longa distância.

Neste ponto, se quiséssemos apenas armazenar os documentos, no nosso caso o acervo fotográfico, no servidor físico brasileiro, mas que seja também armazenado no servidor físico italiano poderíamos criar a rede WAN e habilitar o compartilhamento entre os computadores e servidores. Esta solução é prática e funcional, porém não é viável aos padrões da preservação digital. Pois, apenas o compartilhamento dos documentos por uma rede e o seu armazenamento em um HD não garante atender os requisitos e mecanismos da preservação digital. (THOMAZ, 2004).

Continuando com a infraestrutura, ao invés de criar uma WAN ponto-a-ponto, a partir da virtualização dos servidores físico brasileiro e italiano foram gerados servidores virtuais que auxiliará no gerenciamento de toda rede e fornecerá suporte para o repositório DSpace, outros softwares e ao armazenamento. A partir dos servidores virtualizados e com auxílio dos roteadores das redes locais, foi criada e configurada a WAN.

Com os servidores virtualizados e a WAN firmada, será dividida a dedicação dos servidores virtualizados. Uma vez que utiliza-se recursos mínimos, resolveu-se virtualizar quatro servidores.

- Servidor virtualizado 1º - O primeiro será dedicado aos aplicativos, serviço de DNS, serviço de autenticação, autorização, auditoria e entre outros serviços que possibilita um melhor desempenho para o sistema de preservação.
- Servidor virtualizado 2º - O segundo servidor será desempenhado para receber o repositório DSpace e o conjunto de *softwares* descritos em seu processo de implementação. Este servidor se dedicará exclusivamente ao repositório DSpace.
- Servidor virtualizado Brasil 3º - O terceiro servidor se dedicará a um processo de redundância RAID nos servidores físicos e virtualizados. Ele será um subsistema que entrará em execução caso qualquer um dos servidores ocorra alguma falha. Este subsistema é possível devido a sincronismo entre

os servidores físicos. Caso ocorra uma eventual falha em qualquer servidor, este irá substituí-lo automaticamente sem que haja perda de dados. A sua instância virtualizada procede no servidor físico italiano para precaver uma possível falha do servidor físico brasileiro.

- Servidor virtualizado Itália 4º - O quarto servidor se dedicará a mesma funcionalidade do terceiro, porém a sua virtualização é realizada no servidor físico brasileiro para oferecer o suporte caso o servidor físico italiano venha falhar.

Neste momento percebe-se que a computação em nuvem é a utilização de memória, processamento e capacidade de armazenamento de computadores e servidores virtualizados e interligados por meio da grande rede, fornecendo recursos e serviços para as mais variadas atividades. (VERAS, 2011) Para o caso específico deste trabalho, a computação em nuvem provê recursos para a utilização de um repositório que manterá os documentos digitais por um longo prazo.

Deste modo, finaliza-se a infraestrutura básica para o recurso estratégico de preservação de documentos digitais em nuvem.

Com a infraestrutura montada, o gestor do acervo poderá preservar as fotografias de três formas. A primeira, ele poderá submeter no repositório as fotos individualmente ou compactadas em grupo, onde, cada foto ou cada grupo de fotos, receberá suas respectivas descrições, isto é, metadados que facilitarão a recuperação do acervo.

Este processo se torna seguro e confiável devido ao objeto digital não permanecer apenas em um simples dispositivo de armazenamento, mas sim incluso em um sistema que realiza a sua segurança, forneça acesso ininterrupto para submissão e recuperação por meio da internet para qualquer dispositivo digital ou plataforma utilizada.

Em segundo lugar, o gestor poderá apenas utilizar os recursos da rede WAN para estabelecer o compartilhamento dos servidores. Isto significa dizer que o gestor poderá armazenar o seu acervo, que se encontra no computador brasileiro, também no servidor italiano. Este processo, assim como o primeiro, atende diretamente uma

das políticas de Disaster Recovery que visa distribuir as cópias dos dados gerados por uma organização para outros lugares remotos do seu local de proveniência.

A terceira forma utiliza a sinergia da primeira com a segunda para, além de proverem os recursos necessários da preservação de todo acervo digital, cria-se um sistema de redundância de dados. Este sistema visa distribuir os dados, no nosso caso o acervo, em todos os servidores de armazenamento, isto é, no servidor do Brasil, Itália e no repositório que se encontra em nuvem. Não devemos confundir sistema de redundância de dados com os servidores RAID que irão fazer a redundância, não só dos dados, mas de toda infraestrutura, caso venha a falhar qualquer servidor de toda infraestrutura do recurso estratégico em nuvem.

Com este recurso estratégico, utilizando o repositório, o gestor poderá submeter um grande fluxo de documentos sem se preocupar com a capacidade do seu armazenamento, pois assim que o repositório consumir, totalmente, o seu espaço de armazenamento virtual, ele irá iniciar a consumação do espaço de armazenamento dos servidores físicos, mas isto só acontecerá se o gestor permitir. Caso os recursos de espaço do armazenamento físico estejam acabando, ao invés do gestor comprar um HD com uma capacidade grande de espaço e desperdiçá-lo migrando os dados do servidor para o HD, com essa infraestrutura de preservação em nuvem poderá instalar o HD no servidor, oferecendo assim mais espaço de armazenamento para o repositório poder consumi-lo.

Para finalizarmos, outra forma para ampliar o espaço de armazenamento do repositório é a instalação de um dispositivo de armazenamento de rede (NAS) que pode ser agregado na infraestrutura e configurado pelo navegador. Este dispositivo direciona todo fluxo de dados do repositório para o os seus discos, evitando a sobrecarga nos discos rígidos dos servidores físicos e/ou virtualizados. Este dispositivo pode conter três, quatro ou cinco discos rígidos com grande capacidade para o armazenamento de todo o fluxo documental, não apenas o acervo fotográfico digital, mas de todos os objetos digitais que forem submetidos na infraestrutura de preservação em nuvem.

7. CONSIDERAÇÕES FINAIS

O referente trabalho buscou, não apenas, apresentar um recurso estratégico conceitual de preservação digital, mas também propiciar a interação do profissional da informação, em especial o arquivista, com as novas tecnologias da informação e comunicação. A preservação digital é um desafio que pode ser vencido se, e somente se, o profissional da informação manter-se atualizado no campo da tecnologia da informação e comunicação e não deixar-se cair em obsolescência neste novo mundo da informação.

Este trabalho, assim, descreveu os processos de preservação de documentos digitais para demonstrar que há um número variado de métodos, técnicas e opções para a preservação dos mais variados documentos digitais. Também demonstrou que esses métodos e técnicas podem se remodelarem, diante de uma nova infraestrutura ou arquitetura tecnológica, para provê a sua funcionalidade em um ambiente atual.

Evidente que as escolhas dos métodos e técnicas de preservação competem ao profissional da informação, mas para que haja êxito em sua aplicação o profissional terá que compreender, não só a ferramenta, mas toda holística do documento, do sistema de informação e da instituição para ampliar seu conhecimento e poder adotar uma estratégia de preservação.

Explicar os processos funcionais da computação em nuvem corroborou para expandir as opções de arquitetura para a aplicação de estratégias. Isto se deve a sua elasticidade, escalabilidade e poder computacional que ganha a partir da virtualização. A virtualização também foi outro agente complementar que impulsionou a alusiva pesquisa, pois sem a potencialização das estâncias virtuais, o referente recurso de preservação em nuvem não poderia ser criado.

A passagem do repositório, mas precisamente o DSpace, foi marcada pela sua versatilidade como uma ferramenta de preservação que se torna indispensável e adaptável para qualquer organização que anseia preservar a sua informação e/ou documentação digital.

Os elementos apresentados no decorrer do trabalho foram substanciais para o desenvolvimento do recurso estratégico de preservação em nuvem. O processo minucioso para erguer infraestrutura em nuvem, implementar o repositório em nuvem e descrever todos os recursos necessários para criar a alusiva estratégia, exigiu de nós um grande esforço para traduzir os processos do campo da área tecnológica e computacional para a área das Ciências Sociais aplicadas a qual se encontra a Arquivologia.

Ao longo da pesquisa, vimos que a preservação digital é vista por muitos autores como um grande problema na era da informação, porém identificamos que a solução está no seu próprio ambiente tecnológico, mas com o auxílio de um profissional que detenha os conhecimentos exatos para serem atrelados às tecnologias, gerando desta forma uma nova técnica, recurso, modelo ou método estratégico de preservação.

Em nosso recurso estratégico, foi percebido que o processo mais exaustivo encontra-se na criação da infraestrutura de rede e nas configurações dos servidores, pois este é um trabalho que necessita de uma grande destreza para que não haja transtornos em um futuro próximo.

Sabemos que outros modelos de arquitetura de redes podem ser formados para estabelecer a comunicação entre os computadores, servidores e *hardwares*, mas a implementação do repositório, neste caso o DSpace, é única para qualquer tipologia ou arquitetura adotada.

Também compreendemos que essa infraestrutura criada, pode se conectar com uma infraestrutura terceirizada. Isto é, conectar a nuvem que criamos com uma empresa que fornece serviços de armazenamento em nuvem. Esse processo, embora elevasse os custos, seria mais uma estratégia complementar, pois além de ter os documentos em sua nuvem privada, haveria outra nuvem (terceirizada) interligada oferecendo o armazenamento dos documentos, ou seja, mais uma alternativa para preservar e recuperar os documentos digitais.

Este trabalho fez uma busca pelas melhores ferramentas que se encontram disponíveis para realizar a preservação digital com o mínimo de recursos. A análise feita sobre o nosso recurso estratégico, aplicado diretamente no acervo fotográfico

digital, prova que a infraestrutura de computação em nuvem unida a um conjunto de mecanismos de preservação digital possui, não apenas uma, mas várias estratégias para preservação de documentos digitais que o profissional da informação pode desenvolver.

O trabalho ainda respondeu às questões erguidas pelo palestrante na 2ª Jornada Arquivística – UEPB e entrevistado. Tais questões referenciadas sobre a organização, a preservação e aplicação de um recurso de *Disaster Recovery* (Plano de recuperação de desastres) no acervo fotográfico foram resolvidas, a partir da estratégia de preservação em nuvem.

Estamos seguros em afirmar que a estratégia de preservação em nuvem é a solução mais viável para manter os dados, informações, documentos e/ou todo tipo de objeto digital. Esperamos que a perspectiva dos profissionais da informação, quanto à preservação digital, mude para uma visão mais ampla sobre seus conceitos, métodos, técnicas e estratégias.

Por fim, sabemos que o profissional da informação, em especial o arquivista, se provê com os conhecimentos necessários para desempenhar as suas complexas atividades cotidianas. Essas referentes atividades aprisionam o profissional da informação em um mundo rotineiro, vendando-o e impossibilitando-o de enxergar um mundo insólito que se encontra ao seu redor. Um mundo capaz de solucionar, do mais simples aos mais complexos, problemas que envolvem a arquivística. Um mundo que fomenta a era da informação proporcionando e possibilitando, a este profissional, a reformulação de antigas práticas, conceitos, métodos, técnicas e estratégias sejam aplicadas nos novos Arquivos. O arquivista, sem dúvida, necessita dos conhecimentos do passado para gerir o presente, mas tem o dever de reformular os conhecimentos do passado para garantir o futuro.

REFERÊNCIAS

- APACHE FOUNDATION SOFTWARE. **Tomcat**. Los Angeles, 1999. Disponível em: <<http://tomcat.apache.org/>>. Acesso em: 10 jan. 2014
- APACHE FOUNDATION SOFTWARE. **Ant**. Los Angeles, 1999. Disponível em: <<http://ant.apache.org/>>. Acesso em: 10 jan. 2014
- APACHE FOUNDATION SOFTWARE. **Maven**. Los Angeles, 1999. Disponível em: <<http://maven.apache.org/>>. Acesso em: 10 jan. 2014
- ARAGON, A.; ABREU, V. F. **Implantando a governança de TI: da estratégia à gestão dos processos e serviços**. 2ª ed. Rio de Janeiro: Brasport, 2009. 439 p.
- ARELLANO, M. A. M. **Critérios para a preservação digital da informação científica**. 2008. 356 f. Tese (Doutorado em Ciência da Informação e Documentação) – Faculdade de Ciência da Informação e Documentação, Universidade de Brasília, 2008. Disponível em: <http://btdt.bce.unb.br/tesesimplificado/tde_busca/arquivo.php?codArquivo=4547>. Acesso em: 20 jan. 2013
- ARELLANO, M. A. M. Preservação de documentos digitais. **Ci. Inf.**, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/305/1452>>. Acesso em: 20 jan. 2014.
- BALTZAN, P.; PHILLIPS, A. **Sistemas de informação**. 1ª ed. Porto Alegre: AMGH, 2012. P. 355.
- BRASIL. CONARQ. **Carta para Preservação do Patrimônio Arquivístico Digital: Preservar para garantir o acesso**. Rio de Janeiro, RJ. CONARQ: 2004. Disponível em: <<http://www.documentoseletronicos.arquivonacional.gov.br/cgi/cgilua.exe/sys/start.htm>>. Acesso em: 01 jan. 2014.
- BRASIL. CONARQ. **Modelo de requisito para sistemas informatizados de gestão arquivística de documentos**. Rio de Janeiro, RJ. CONARQ: 2006. 132 p. Disponível em: http://www.unesp.br/ccad/mostra_arq_multi.php?arquivo=6964. Acesso em: 02 jan. 2014.
- BRASIL. CONARQ. **Recomendações para digitalização de documentos arquivísticos permanentes**. Rio de Janeiro, RJ. CONARQ: 2010. 28 p. Disponível em: <http://www.conarq.arquivonacional.gov.br/media/publicacoes/recomenda/recomendaes_para_digitalizacao.pdf>. Acesso em: 02 jan. 2014.

BRASIL. CONARQ. Resolução nº 20, de 16 de Julho de 2004. **Dispõe sobre a inserção dos documentos digitais em programas de gestão arquivística de documentos dos órgãos e entidades integrantes do sistema Nacional de Arquivos.** Disponível em: < <http://www.conarq.arquivonacional.gov.br/cgi/cgilua.exe/sys/start.htm?infoid=71&sid=46>>. Acesso em: 01 jan. 2014.

CNRI – CORPORATION FOR NATIONAL RESEARCH INITIATIVE. Reston, 2013. Disponível em : < <http://www.cnri.reston.va.us/>>. Acesso em 10 jan. 2014

CREASY, J. R. The origin of the VM/370 Time-Sharing System. **J. Res. Develop.** Palo Alto, v. 25, n. 5, p. 483-490, September, 1981. Disponível em: < ftp://ftp.cis.upenn.edu/pub/cis700-6/public_html/04f/papers/creasy-vm-370.pdf>. Acesso em: 07 jan. 2014.

CRISTALDO, H. Serpro vai lancar o primeiro Sistema em computação em nuvem do governo federal em setembro. **Agência Brasil – EBC**, Brasília, 15, ago. 2013. Disponível em: < <http://memoria.ebc.com.br/agenciabrasil/noticia/2013-08-15/serpro-vai-lancar-primeiro-sistema-em-computacao-em-nuvem-do-governo-federal-em-setembro>>. Acesso em: 14 jan. 2014.

DATE, C. J. **Introdução a sistemas de banco de dados.** Tradução de Daniel Vieira. 8ª ed. Rio de Janeiro: Elsevier, 2003. 851 p.

DECLARAÇÃO UNESCO / UBC VANCOUVER. **A memória do mundo na Era digital:** Digitalização e preservação. In: Conferência Mundial sobre: A memória do mundo na Era digital. UNESCO. Vancouver/Canadá, 2012.

DEMO, P. **Metodologia do conhecimento científico.** 1ª ed. São Paulo: Atlas, 2000. 216 p.

DEMO, P. **Metodologia científica em ciências sócias.** 3ª ed. São Paulo: Atlas, 1995. 296 p.

DIGIRATI. **Guia Prático de Redes de Computadores.** [Equipe Digirati Books]. – São Paulo: Digirati Books, 2009.

DURANTI, L. An Overview of InterPARES 3. **J. Int. Res.**, Columbia, v. 1, n. 1, p. 1-27, September, 2007. Disponível em: < http://archivo.cartagena.es/files/36-180-DOC_FICHERO1/03-duranti_overview.pdf>. Acesso em: 8 jan. 2014.

EMC² - CORPOTARION. **Armazenamento e gerenciamento de informação:** Como armazenar, gerenciar e proteger informações digitais. 1ª ed. Porto Alegre: BOOKMAN, 2009. 461 p.

FARIA, H. M. **Bacula:** ferramenta livre de backup. 1ª ed. Rio de Janeiro: Brasport, 2010. 208 p.

FERREIRA, C. A. S. **Preservação digital**: Uma perspectiva orientada para bibliotecas. 2011. 155 f. Dissertação (Mestrado em Informação, Comunicação e Novas Media) – Faculdade de Letras, Universidade de Coimbra, Coimbra, 2011.

FERREIRA, M. **Introdução à preservação digital**: Conceitos, estratégias e actuais consensos. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006. 88 p. Disponível em: <<https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf>>. Acesso em: 10 jan. 2014.

FERREIRA, M. **Preservação de longa duração de informação digital no contexto de um arquivo histórico**. 2009. 273 f. Tese (Doutorado em Tecnologias e Sistemas de Informação) – Escola de Engenharia, Universidade do Minho, 2009. Disponível em: <<http://repositorium.sdum.uminho.pt/bitstream/1822/9563/4/thesis-mferreira-0.33.pdf>>. Acesso em: 20 jan. 2014.

GIL, A. C. **Como elaborar projetos de pesquisa**. 3ª ed. São Paulo: Atlas, 1996. 159 p.

GOLDEN, B. **Virtualization for Dummies**. 1ª ed. Indianapolis: Wiley Publishing, Inc. 2008. 341 p. (Coleção for Dummies)

GRACIO, J.C.A. **Preservação digital na gestão da informação**: Um modelo processual para as instituições de ensino superior. 2011. 223 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Ciência da Informação, Universidade Estadual Paulista, 2011. Disponível em: <http://www.athena.biblioteca.unesp.br/exlibris/bd/bma/33004110043P4/2011/gracio_jca_dr_mar.pdf>. Acesso em: 15 jan. 2014

HEDSTRON, M.; et al. **Invest Save**: Report and recommendations the NSF-DELOS working group on digital archiving and preservation. Michigan, 2003. Disponível em: <http://eprints.erpanet.org/94/01/NSF_Delos_WG_Pres_final.pdf>. Acesso em: 08 jan. 2014.

HUMBLE, J.; FARLEY, D. **Entrega Contínua**: como entregar software de forma rápida e confiável. 1ª ed. Porto Alegre: Bookman, 2013. 496 p.

HURWITZ, J.; et al. **Cloud Computing for Dummies**. 1ª ed. Indianapolis: Wiley Publishing, Inc, 2010. (Coleção for Dummies).

INNARELLI, H. C. Os dez mandamentos da preservação digital. In: SANTOS, V. B.; INNARELLI, H. C.; SOUSA, T. R. B. **Arquivística**: temas contemporâneos. Brasília: SENAC, 2007. Cap. 1, P. 21-71.

LAVOIE, B; DEMPSE, L. Thirteen Ways of Looking at...Digital Preservation. **D-Lib Mag.** vol. 10, n., 7/8, jul/ago, 2004. Disponível em: <<http://www.dlib.org/dlib/july04/lavoie/07lavoie.html>>. Acesso em: 04 jan. 2014.

LE COADIC, Y. F. **A Ciência da Informação**. 1ª ed. Brasília: Briquet de Lemos, 1996. 119 p.

LE GOFF, J. **História e Memória**. 1ª ed. Campinas, São Paulo: UNICAMP, 1991. 476 p.

LEE, K, et al. The State of The Art and Practice in Digital Preservation. **J. Res. Natl. Inst. Stand. Technol.** Washingto Gaithersburg, v. 107, n. 93, p. 93-106, jan.-fev. 2002.

LÉVY, P. **O que é o virtual?**. Tradução de Paulo Neves. 1. Ed. São Paulo: Ed. 34, 1996. 160 p. (Coleção TRANS)

LÉVY, P. **Cibercultura**. Tradução de Carlos Irineu. 1ª ed. São Paulo: Ed. 34, 1999. 264 p. (Coleção TRANS)

LOBO, A. P. Embratel conduz projeto de computação em nuvem latino-americana. **Convergência Digital**, São Paulo, 15, jun. 2013. Disponível em: < <http://convergenciadigital.uol.com.br/cgi/cgilua.exe/sys/start.htm?infoid=34072&sid=124#UvwjpvldV1Y>>. Acesso em: 20 jan. 2014.

LUZ, C. **Arquivologia 2.0: A informação digital humana – Excertos de um Arquivista 2.0 no mundo digital**. 1ª ed. Florianópolis: Bookess, 2010. 116 p.

MARCONDES, C. H; SAYÃO, L. F. Documentos digitais e novas formas de cooperação entre sistemas de informação em C&T. **Ci. Inf**, Brasília, v. 31, n. 3, p. 42-54, 2002. Disponível em: < <http://www.scielo.br/pdf/ci/v31n3/a05v31n3.pdf>>. Acesso em: 05 jan. 2014.

MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de Metodologia Científica**. 5ª ed. São Paulo: Atlas, 2003. 311 p.

MARCONI, M. A.; LAKATOS, E. M. **Metodologia do trabalho científico: Procedimentos básicos. Pesquisa bibliográfica, projeto e relatório. Publicações e trabalhos científicos**. 7ª ed. São Paulo: Atlas, 2007. 228 p.

MARCONI, M. A.; LAKATOS, E. M. **Técnicas de pesquisa: Planejamento e execução de pesquisas. Amostragens e técnicas de pesquisa. Elaboração, análise e interpretação de dados**. 7ª ed. São Paulo: Atlas, 2008. 296 p.

MILLER, M. **Cloud Computing: Web-based applications that change way you work and collaborate online**. 1ª ed. Indianapolis: QUE Publishing, 2008. 292 p.

MORLEY, D.; PARKER, C. **Understanding Computers: Today and tomorrow, introductory**. 13ª ed. Boston: Cengage Learning, 2010. 450 p.

NEVES, J. L.; **Pesquisa qualitativa: características, usos e possibilidades.** Caderno de Pesquisas em Administração, São Paulo, v. 1, n. 3, 2 sem., 1996.

NIST – National Institute of Standards and Technology. NIST Cloud Computing Reference Architecture: Recommendations of the national Institute of Standards and Technology. J. Natl. Inst, Technol. Spec. Publ., Gaithersburg, v. 500, n. 292, 35 p. September 2011. Special Publication.

O'BRIEN, J. A.; MARAKAS, G. M. **Administração de sistemas de informação.** 15ª ed. Porto Alegre: BOOKMAN, 2010. 589 p.

ORACLE TECHNOLOGY NETWORK. **JDK.** Redwood, 2013. Disponível em: <<http://www.oracle.com/technetwork/pt/java/javase/downloads/jdk7-downloads-1880260.html>>. Acesso em: 10 jan. 2014

PAIVA, M. F. **Boas práticas para projetos, homologação, implantação, utilização e manutenção de sistemas de gestão da informação que armazenam imagens digitais de documentos com fidedignidade e confiabilidade.** 1ª ed. São Paulo: Target, 2008. 99 p.

PINTO, M. M. G. A. Do “efêmero” ao “sistema de informação”: a preservação na era digital. **Flup. Arq & Biblio.** Porto, v. 2, n. 15, p. 53-59. Disponível em: <<http://ler.letras.up.pt/uploads/ficheiros/3083.pdf>>. Acesso em: 02 jan. 2014.

POSTGRESQL.ORG – GLOBAL DEVELOPMENT GROUP. Los Angeles, 2013. Disponível em: <<http://www.postgresql.org/>>. Acesso em: 10 jan. 2014

RAINER, R. K.; CEGIELSKI, C. G. **Introdução a sistemas de informação.** 3ª ed. Rio de Janeiro: Elsevier. 2012. 431 p.

RONDINELLI, R. C. **Gerenciamento arquivístico de documentos: Uma abordagem teórica da diplomática arquivística contemporânea.** 1ª ed. Rio de Janeiro: FGV, 2002. 160 p.

ROSS, J. **Rede de Computadores.** 1ª ed. Rio de Janeiro: Antenna, 2008. 145 p.

ROTHENBERG, J. **Avoiding technological quicksand:** finding a viable technical foundation for digital preservation. Washington, D.C.: Council on Library and Information Resources, 1999. p. 1-35. Disponível em: <<http://www.clir.org/pubs/reports/rothenberg/pub77.PDF>>. Acesso em: 7 jan. 2014.

SAYÃO, L. F.; MARCONDES, C. H. Software livres para repositórios institucionais: alguns subsídios para a seleção. In: SAYÃO, L. F.; et al. **Implantação e gestão de repositórios institucionais:** políticas, memória, livre acesso e preservação. 1ª ed. Salvador: EDUFBA, 2009. cap. 2, p. 23-54. Disponível em: <https://repositorio.ufba.br/ri/bitstream/ufba/473/3/implantacao_repositorio_web.pdf>. Acesso em: 01 fev. 2014

SCHELLENBERG, T. R. **Arquivos Modernos**: Princípios e técnicas. 4ª ed. Rio de Janeiro: FGV, 2004. 388 p.

TAURION, C. **Cloud Computing**: transformando o mundo da tecnologia da informação. 1ª ed. Rio de Janeiro: Brasport, 2009. 201 p.

THOMAZ, K. P.; SOARES, A. J. A preservação digital e o modelo de referência Open Archival Information System (OAIS). *Dat.Gram. Rev. Ci.* v. 5, n. 1, 2004. Disponível em: <http://www.dgz.org.br/fev04/Art_01.htm>. Acesso em 7 jun. 2013.

THOMAZ, K.P. **A preservação de documentos eletrônicos de caráter arquivístico**: Novos desafios, velhos problemas. 2004. 389 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Ciência da Informação, Universidade Federal de Minas Gerais, 2004. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/VALA-68ZRKF/doutorado_katia_de_padua_thomaz.pdf?sequence=1>. Acesso: 20 jan. 2014

TURBAN, E. et al. **Tecnologia da informação para gestão**: Transformando os negócios na economia digital. 6ª ed. São Paulo: BookMan, 2008. 711 p.

TURBAN, E.; VOLONINO, L.; BRODBECK, A. **Tecnologia da informação para gestão**: em busca do melhor desempenho estratégico e operacional. 8. ed. Porto Alegre: Bookman, 2013. 468 p.

VERAS, M. **Virtualização**: componente central do Datacenter. 1ª ed. Rio de Janeiro: Brasport, 2011. 331 p.

VERAS, M. **Arquitetura corporativa de nuvem**: amazon web servisse (AWS). 1ª ed. Rio de Janeiro: Brasport, 2013. 387 p.

VERAS, M. **Cloud Computing**: nova arquitetura da TI. 1ª ed. Rio de Janeiro: Brasport, 2012. 208 p.

VERMAAT, M. E. **Discovery Computers**: Student Success Guide. 1ª ed. Boston: Cengage Learning, 2013. 834 p.

YATES, R. B.; NETO, R. **Recuperação da informação**: Conceitos e tecnologia das máquinas de busca. 2ª ed. Porto Alegre: Bookman, 2013. 581 p.

APÊNDICES

APÊNDICE: COLETA DE DADOS COM O GESTOR DO ACERVO FOTOGRÁFICO DIGITAL DO PROJETO: COMUNIDADES QUILOMBOLAS DA PARAÍBA

A inerente entrevista remete as questões de conhecimento e práticas de preservação digital e a utilização da tecnologia computação em nuvem. Tomando conhecimento sobre acervo digital do projeto social que mantém a memória das comunidades quilombolas da Paraíba e sabendo que a sua produção documental é vasta, resolvemos entrevistar o gestor do acervo. Segue abaixo, a transcrição da entrevista concedida pelo membro, produtor e gestor do acervo digital fotográfico do projeto: Comunidades Quilombolas da Paraíba. A entrevista foi cedida no dia: 11.12.2013, no escritório da sua residência. A entrevista durou, aproximadamente, 60 minutos.

ENTREVISTA: Preservação digital e utilização da computação em nuvem.

1- Quando surgiu o interesse pela arte de fotografar e por quê?

Resposta da questão 1º

Entrevistado	Desde muito cedo! Viajei por muitos lugares ao longo da vida. Sou um cidadão do mundo.
---------------------	--

2- Todos esses anos fotografando, como você arquiva seu acervo fotográfico?

Resposta da Questão 2º

Entrevistado	Eu mantenho arquivado de uma forma simples, mas que até hoje está me servindo. As mais antigas (analógicas), algumas foram digitalizadas, mas tenho muitos álbuns e diapositivos acondicionados com papel neutro em caixas arquivo, eles estão bem conservados. As mais recentes (digitais) estão arquivadas no meu computador e nos HD's externos.
---------------------	---

3º Sabemos que a tecnologia digital trouxe algumas facilidades para o mundo da fotografia, como: a manipulação das imagens, “economia” e facilidade em fotografar, mas em contra ponto, trouxe a problemática do armazenamento. Após a adesão da tecnologia digital, como você vem armazenando seu acervo?

Resposta da Questão 3º

Entrevistado	Eu estou armazenando em meu computador e em cinco HD's externos, sendo três destes HD's externos backups do HD do meu computador e dos outros dois HD's externos. Também tenho alguns backups na Itália, mas percebo que necessito construir uma estratégia de <i>Disaster Recovery</i> *.
---------------------	--

* São políticas e procedimentos que estão relacionados à preparação para a recuperação de dados ou manutenção de infraestrutura tecnológica, que são vitais para uma organização após desastres naturais ou induzidos pelo homem.

4º Onde se encontra o seu acervo analógico? Já pensou em digitaliza-lo?

Resposta da Questão 3º

Entrevistado	Meu acervo analógico encontra-se, uma boa parte, digitalizado. Porém há outras fotografias mais antigas que não digitalizei. Eu mesmo digitalizei. Digitalizei até os diapositivos. Utilizei um Scanner.
---------------------	--

5º Como é feito o arquivamento? O acervo é classificado?

Resposta da Questão 5º

<p>Entrevistado</p>	<p>Eu passo as fotos para o meu computador e organizo nomeando as pastas com as cidades, onde se encontra a comunidade quilombola e as respectivas datas da visita. São milhares de arquivos acondicionados nos HD's e o grande problema é que o mapa de todo acervo está na minha cabeça.</p>
----------------------------	--

6º Possui algum software para arquivar e recuperar as fotos do acervo?

Resposta da questão 6º

<p>Entrevistado</p>	<p>Não. Apenas o sistema operacional.</p>
----------------------------	---

7º Você já tentou armazenar seu acervo em um banco de dados, ou em um servidor?

Resposta da Questão 7º

<p>Entrevistado</p>	<p>Não!</p>
----------------------------	-------------

8º Você tem conhecimento sobre repositórios digitais ou arquivos digitais?

Resposta da Questão 8º

Entrevistado	Não!
---------------------	------

9º Armazenamento em nuvem, você possui conhecimento sobre esta tecnologia?

Resposta da Questão 9º

Entrevistado	<p>Sim! Eu armazeno alguns arquivos em nuvem. Eu tenho uma assinatura no Imgur.com* por um preço incrível. Eu pago US\$ 24,00 (Vinte e quatro dólares), ao ano por um espaço ilimitado de armazenamento. Eu hospedo as fotos e rapidamente tenho acesso a elas. Eu utilizo outros como Google Drive, não para armazenar meu acervo, mas para outras coisas. O Google Drive me oferece 7 GB (Sete Gigas), gratuitamente, mas eu vi que neste momento o Imgur.com é muito mais rápido que o Drive do Google.</p>
---------------------	--

* [Imgur](https://imgur.com). É um servidor para hospedagem de imagens fundada em 2009.

10º Qual a principal dificuldade em relação ao seu acervo digital?

Resposta da Questão 10º

Entrevistado	A preservação e segurança. Pois, não é a primeira vez que um queima um HD ou outras coisas. Já... Acho que já queimei uns quatro ou cinco HD's nestes últimos anos. Felizmente, já tinha começado a fazer as cópias e por isso não perdi todos os dados. Mas eu sei que tem esse perigo.
---------------------	--

11º Você acredita que um arquivista possa gerir seu acervo com os novos recursos tecnológicos que, possivelmente, irá preservar e permitir o seu acesso por longa data?

Resposta da Questão 11º

Entrevistado	Eu acredito que este seja o trabalho do arquivista. O arquivista tem que encontrar as soluções mais adequadas para um "novo tipo de Arquivo". Sabemos que há procedimentos para arquivar as fotografias, mas não é que eu queira preservar apenas 500 ou 1.000 fotografias que tenho em papel. Eu quero preservar as minhas 25.000 fotografias digitais, pois eu tenho que deixar esse trabalho para a sociedade. Hoje, eu estou realizando todo esse trabalho em prol das comunidades quilombolas como um trabalho voluntário. Então, eu tenho sempre que pensar que, se este trabalho é importante, é útil hoje, terá que continuar a ser útil futuramente. Essa também é a função social do Arquivo.
---------------------	---

12º Quanto você pagaria para preservar, não apenas seu acervo, mas toda a memória do seu trabalho que lhe custou anos para produzir?

Resposta da Questão 12º

Entrevistado	Quanto eu pagaria, não sei te dizer. Pois tudo irá depender dos custos, mas se eu estiver em frente de alguém que me ofereça uma solução e diga: - "Olha, para resolver esse teu problema você precisará fazer isso, isso, e tudo sairá por um custo razoável". Sim! Eu pagaria o necessário, mas
---------------------	---

	<p>quem vier falar comigo para realizar a proposta, tem que saber do que está falando e que tenha a visão inclinada para a causa do trabalho social nas comunidades. Quanto à memória do trabalho, eu acredito que o arquivista tem essa obrigação, assim como um padre, um sacerdote ou médico tem a sua obrigação. Um arquivista, se ele desempenha a função de arquivista, ele tem uma missão e essa missão é para a humanidade. Porque a memória, a história, ou seja, tudo isso é uma parte integrante, não só da atualidade, mas também do futuro.</p>
--	--

ANEXO

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado participante:

Sou estudante do curso de graduação em Arquivologia pela Universidade Estadual da Paraíba. Estou realizando uma pesquisa sob supervisão da professora Dr^a. Francinete Fernandes de Sosal, cujo objetivo é desenvolver um recurso estratégico de preservação a partir da utilização da computação em nuvem.

Sua participação envolve-se em uma entrevista, que será gravada se assim você permitir, para um melhor registro e apuração dos dados.

A participação nesse estudo é voluntária e se você decidir não participar ou desistir de continuar em qualquer momento, tem absoluta liberdade de fazê-lo.

Na publicação dos resultados desta pesquisa, você está de acordo em liberar a sua identidade: () Sim ou () Não. Caso não esteja de acordo sua identidade será mantida no mais rigoroso sigilo e serão omitidas todas as informações que permitam identificá-lo(a).

Mesmo não tendo benefícios diretos em participar, indiretamente você estará contribuindo para a compreensão do fenômeno estudado e para a produção de conhecimento científico.

Quaisquer dúvidas relativas à pesquisa poderão ser esclarecidas pelo pesquisador, Glaucco Ranniere de Souza Pontes, fone: (83) 9653-4503.

Atenciosamente,

Assinatura do estudante

Local e data

Matrícula:

Assinatura da orientadora

Consinto em participar deste estudo e declaro ter recebido uma cópia deste termo de consentimento.

Assinatura do participante

Local e data