RICI Revista Ibero-Americana de Ciência da Informação

ISSN 1983-5213

Auditoria de repositórios digitais preserváveis

Diego Bil Silva Barros
Universidade Federal do Pará, Biblioteca Central, Belém, PA, Brasil
diegobilbarros@gmail.com

Igor Dias Ferrer
Instituto Brasileiro de Informação em Ciência e Tecnologia, Brasília, DF, Brasil
igorferrer@ibict.br

Cleusa Maria de Souza Maia
Instituto Brasileiro de Informação em Ciência e Tecnologia, Brasília, DF, Brasil
cleusamaia@ibict.br

Resumo: Este trabalho tem por finalidade demonstrar os atributos e responsabilidades de um repositório digital arquivístico confiável. A confiabilidade dos repositórios digitais arquivísticos dá-se por meio da sua certificação, que deve ser com base na norma ISO 16363/2012, fundamentada no modelo conceitual OAIS (Open Archival Information System), e na aplicação de ferramentas tais como o TRAC (Trustworthy Repository Audit & Certification). Será abordado o uso de metodologias de auditoria e a adoção de estratégias que visem evitar a obsolescência tecnológica, permitindo o acesso aos documentos a longo prazo, mantendo sua autenticidade.

Palavras-chave: Auditoria digital; Obsolescência tecnológica; Preservação digital; Repositório digital.

ARTIGOS

Audit of preservable digital repositories

Abstract: This paper aims to demonstrate the attributes and responsibilities of a reliable archival digital repository. The reliability of archival digital repositories is achieved through its certification, which must be based on ISO 16363/2012, based on the Open Archival Information System (OAIS) conceptual model, and on the application of tools such as TRAC (Trustworthy Repository Audit & Certification). It will be approached the use of auditing methodologies and the adoption of strategies that aim to avoid the technological absoluteness, allowing the access to the documents in the long term, maintaining its authenticity.

Keywords: Digital audit; Digital preservation; Digital repository; Technological obsolescence.

Auditoría de repositorios digitales preservables

Resumen: Este trabajo tiene por finalidad demostrar los atributos y responsabilidades de un repositorio digital archivístico confiable. La confiabilidad de los repositorios digitales archivísticos se da por medio de su certificación, que debe basarse en la norma ISO 16363/2012, basada en el modelo conceptual OAIS (Open Archival Information System), y en la aplicación de herramientas tales como el TRAC (Trustworthy Repository Audit & Certification). Se abordará el uso de metodologías de auditoría y la adopción de estrategias para evitar la obsolescencia tecnológica, permitiendo el acceso a los documentos a largo plazo, manteniendo su autenticidad.

Palabras clave: Auditoría digital; Obsolescencia tecnológica; Preservación digital; Repositorio digital.

1 Introdução

A quantidade de informações que surgem a todo instante faz com que a humanidade seja capaz de desenvolver novos conhecimentos e gerar diversas formas de contribuir para o avanço científico. À medida que essa quantidade aumenta outra preocupação também se eleva exponencialmente: a necessidade da preservação digital.

Desde a antiguidade o homem busca meios de preservar e conservar sua memória. A evolução dos processos de meio de comunicação aliado ao desenvolvimento da ciência foram fatores que corroboraram para a compreensão de fatores ligados à integridade e recuperabilidade da informação com o passar dos tempos. Assim, independentemente do tipo de suporte de informação, a necessidade de preservar a informação sempre foi um sentimento comum ao indivíduo (VIDAL, 2010).

A importância da salvaguarda das informações nos dias de hoje está mais evidente à medida em que, com o passar do tempo, os suportes informacionais surgem em diferentes formatos. A adoção de políticas e estratégias de preservação digital corroboram de forma positiva para que os problemas com a perda de informação sejam minimizados. Além de estratégias, é interessante que haja conscientização e boa vontade dos gestores para investir nessas ferramentas. Dessa forma, as ações que visam a preservação, manutenção e recuperação da informação em meio digital se caracterizam como formas de estratégias e políticas de preservação (SCHÄFER; CONSTANTE, 2012).

Dentro dessas estratégias de preservação digital, temos um importante fator que deve ser levado em consideração: a auditoria dos repositórios digitais. Portanto, auditar os repositórios digitais significa torná-los confiáveis e, na medida do possível, mais seguros para garantir que as informações ali dispostas estejam preservadas ao longo do tempo.

Em razão do avanço tecnológico ser uma constante no mundo digital, percebe-se que a implantação de políticas, de estratégias e de normas não conseguem acompanhar as alterações necessárias para que um repositório digital possa tonar-se confiável. Corroborando com este pensamento, Thomas *apud* Santos (2016, p. 68) destaca que "a crescente proliferação de documentos digitais, […], tem ameaçado a capacidade humana de continuar utilizando os arquivos como fontes de informação em virtude dos novos desafios impostos pela preservação".

A necessidade de preservação de documentos digitais é fundamental para garantir seu acesso, e com isso novos desafios para evitar o comprometimento de sua autenticidade e arquivamento a longo prazo foram surgindo. Neste sentido Innarelli (2006, *apud* Santos 2016, p. 76) afirma que:

> "O entendimento da complexidade e fragilidade dos documentos digitais deixa claro que a preservação digital não é resolvida pela própria tecnologia e nunca será, é resolvida com o estabelecimento de políticas e agendas de trabalho que, quando levadas a sério e incorporadas no dia a dia, permitirão a perpetuação dos acervos digitais, mesmo que estes deixem de ser digitais para serem atômicos, biológicos, futurológicos, etc."

As atividades de preservação digital não dependem exclusivamente da tecnologia utilizada; é fundamental a implantação de políticas de preservação, estratégias e normas, uma vez que a obsolescência tecnológica é uma consequência motivada por uma evolução desenfreada de *softwares*, formatos de arquivos e sistemas operacionais. Diante desta realidade o planejamento da gestão de arquivos permitirá que eles continuem autênticos a longo prazo.

Neste contexto os planos e políticas de preservação digital têm um papel importante para as organizações, pois possibilitam o gerenciamento dos riscos críticos, das atividades de pessoal responsável, baseado em diretrizes e requisitos internacionalmente adotados. Para garantir a confiabilidade, o acesso e a integridade dos documentos a longo prazo é necessário a implantação de estratégias de preservação para documentos digitais, juntamente com as políticas de preservação digital, pois com sua ausência torna o patrimônio digital vulnerável à obsolescência.

Diante disso, com a aplicação de estratégias e ferramentas de preservação digital em longo prazo, torna-se viável desenvolver um processo de auditoria interna segundo a norma ISO 16363 (Audit and certification of trustworthy digital repositories), padrões internacionais e modelos de referência como OAIS (Open Archival Information System) para aferir e reportar a maturidade do repositório.

Há muitos desafios para as áreas da Ciência da Informação, Biblioteconomia e Ciência da Computação em torno dos repositórios digitais de dados de pesquisa, principalmente em termos de prática biblioteconômica e de aplicação de tecnologias. As bibliotecas especializadas – cumprindo sua missão - começam a expandir e renovar suas habilidades e conhecimentos, centradas em documentos, para estabelecer as bases de uma biblioteconomia de dados capaz de lidar com os estoques crescentes desses ativos informacionais (SALES; SAYÃO, 2015).

As políticas de preservação digital constituem um papel estratégico para o acervo, pensando no longo prazo, na autenticidade, no acesso, desta forma, observa-se que partes dos esforços da preservação digital podem ser minimizadas através de uma produção que vislumbre os padrões de preservação. Com isto, torna-se fundamental o desenvolvimento de manuais, normas, sistemas informatizados próprios para gestão e preservação. Além da definição de uma política de preservação e do estabelecimento de estratégias de preservação,

a implementação de um repositório digital torna-se fundamental para executar estas estratégias de preservação bem como as políticas.

Nesse território da pesquisa contemporânea, reordenado pela geração e uso intensivo de dados, os meios de compartilhamento e o amplo acesso aos dados de pesquisa criam pontos de inflexões nas metodologias de identificação de novos fenômenos, na validação e na reprodutibilidade das pesquisas e nas formas de socialização dos pesquisadores. Este cenário de grandes novidades abre perspectivas inéditas para descobertas em todas as áreas do conhecimento.

O termo "dado de pesquisa" tem uma amplitude de significados que vão se transformando de acordo com domínios científicos específicos, objetos de pesquisas, metodologias de geração e coleta de dados e muitas outras variáveis. Pode ser o resultado de um experimento realizado num ambiente controlado de laboratório, um estudo empírico na área de ciências sociais ou a observação de um fenômeno cultural ou da erupção de um vulcão num determinado momento e lugar. Dados digitais de pesquisa ocorrem na forma de diferentes tipos de dados, como números, figuras, vídeos, *softwares*; com diferentes níveis de agregação e de processamento, como dados crus ou primários, dados intermediários e dados processados e integrados; e em diferentes formatos de arquivos. Essa diversidade vai sendo delineada pelas especificidades de cada disciplina, suas condicionantes metodológicas, protocolos e seus objetivos, e se torna um desafio, mesmo para o pesquisador, pelo alto grau de contextualização necessário, definir precisamente o que é dado de pesquisa de uma forma transversal aos diversos domínios disciplinares (BORGMAN, 2010; PAMPEL *et al.,* 2013).

Além de oferecer uma base tecnológica para a execução dos processos de contextualização dos dados, os repositórios têm um papel importante nas interações que envolvem a validação do trabalho de pesquisa e na própria dinâmica social da comunicação científica. A possibilidade de se ter os dados de pesquisa disponíveis *online*, indexados, documentados e anotados relativos a uma pesquisa publicada ou pré-publicada num artigo acadêmico, redimensiona a revisão por pares, estendendo-a a uma comunidade mais ampla e conectada em rede. "Um repositório permite exame, prova, revisão, transparência de resultados de pesquisa por outros especialistas que vão além da revisão por pares do artigo acadêmico publicado" (UZWYSHYN, 2016, p. 1).

O repositório deve fornecer evidências para mostrar que ele opera um sistema de gerenciamento de dados e metadados adequado para garantir integridade e autenticidade durante os processos de ingerir, arquivar e armazenar o acesso aos dados. A integridade garante que as alterações nos dados e metadados sejam documentadas e possam ser

rastreadas. A autenticidade cobre o grau de confiabilidade dos dados depositados originais e sua proveniência, incluindo a relação entre os dados originais e os divulgados, e se as relações existentes entre conjuntos de dados e metadados estão ou não existentes.

Os repositórios devem funcionar em conjunto com os depositantes para garantir que haja informações disponíveis suficientes sobre os dados, de modo que possibilite a avaliação substancial dos dados. Essa avaliação da qualidade torna-se cada vez mais relevante quando a equipe é multidisciplinar, onde os pesquisadores podem não ter a experiência pessoal para fazer uma avaliação da qualidade apenas a partir dos dados. Os repositórios também devem ser capazes de avaliar a qualidade técnica dos depósitos de dados em termos de integridade e qualidade dos materiais fornecidos e da qualidade dos metadados.

Os dados ou metadados associados podem ter problemas de qualidade relevantes para o valor de pesquisa, mas isso não impede seu uso na ciência se um usuário puder tomar uma decisão bem informada sobre sua adequação por meio da documentação fornecida. Os repositórios devem garantir que os dados possam ser entendidos e utilizados de forma eficaz no futuro, apesar das mudanças na tecnologia. Este requisito avalia as medidas tomadas para garantir que os dados sejam reutilizáveis.

Os repositórios precisam operar em infraestruturas básicas confiáveis e estáveis que maximizem a disponibilidade do serviço. Deve analisar ameaças potenciais, avaliar riscos e criar um sistema de segurança consistente. E ele deve descrever cenários de danos com base em ações mal-intencionadas, erros humanos ou falhas técnicas que representam uma ameaça para o repositório e seus dados, produtos, serviços e usuários.

2 Auditoria de Repositórios Digitais

Os repositórios digitais são importantes instrumentos de divulgação e disseminação da produção científica de uma instituição. Dessa forma, para que o processo de comunicação científica seja evidenciado, é interessante que os trabalhos sejam disponibilizados em plataformas de livre acesso à produção científica. Assim, Weitzel (2006) explicita que o repositório digital é uma coleção de documentos digitais, no qual aqueles que adotam o modelo OAI-PMH (*Open Archive Initiative – Protocol for Metadata Harvesting*) se tornam interoperáveis a partir do compartilhamento de seus metadados.

De modo geral, os repositórios podem ser temáticos ou institucionais. Os repositórios temáticos são aqueles que lidam com a produção científica institucional, porém com alguma área do conhecimento em particular. Já os repositórios institucionais podem ser considerados aqueles que armazenam a produção de uma instituição, sem distinguir área temática (LEITE *et*

al, 2012; WEITZEL, 2006). De qualquer forma, ambos trabalham com a produção científica e corroboram para o processo de comunicação científica.

Certamente, a criação de um repositório digital perpassa por uma série de requisitos que devem ser observados no momento do planejamento. Logo, podemos lançar mão de alguns questionamentos básicos importantes no momento de planejar a implementação desses repositórios, tais como: *Que tipo de documentos serão disponibilizados? Quais os formatos desses objetos digitais? Como pensar em um planejamento de preservação digital para o repositório? O meu repositório é confiável? Será possível adotar auditorias internas ou externas para avaliar o repositório? Podemos adotar algum parâmetro de auditória para Repositórios Digitais?*

O processo de preservação digital é algo que serve como base para o início do processo de garantia de acesso aos documentos no futuro. Assim, Márdero Arellano (2004) relata que a preservação dos documentos digitais é determinada pela capacidade de um determinado objeto digital se manter utilizável e acessível para as próximas gerações. Sendo assim, é importante que as soluções tecnológicas sejam adotadas de forma a corroborar com o contexto da preservação digital em repositórios.

Além da preservação digital, um fator que se revela importante no contexto é a confiabilidade das informações disponíveis. Sobre isso, Thomaz (2007) afirma que um repositório digital pode ser considerado confiável a partir de no mínimo três níveis de aplicação: *i)* a confiança de que os produtores estão enviando as informações corretamente; *ii)* a confiança de que os consumidores estão recebendo as informações corretas; *iii)* a confiança de que os fornecedores estão prestando os serviços adequados. Assim, para que que um arquivo digital seja considerado confiável, foram identificados os seguintes atributos: Conformidade com o modelo de referência SAAI, Reponsabilidade Administrativa, Viabilidade Organizacional, Sustentação Financeira, Adequação Tecnológica, Sistema de segurança, Responsabilidade (*accountability)* de procedimentos (RLG/OCLC, 2002 *apud* THOMAZ, 2007).

Outro ponto importante a ser analisado é a questão da certificação. Sobre isso, Thomaz (2007, p.84) afirma que "a certificação se tornou um componente-chave para repositórios digitais contemporâneos". Assim, é a partir desse ponto que um repositório digital poderá, ao longo do tempo, obter a confiança dos pesquisadores que o acessam. Conscientes sobre o problema, Thomaz (2007) afirma que em janeiro de 2007, o Consultative Commitee of Space Data System (CCSDS) reuniu um grupo de trabalho com a finalidade de desenvolver uma norma de âmbito internacional para auditoria e certificação de repositórios digitais, tendo como principais fontes de referência desse trabalho o *Trustworthy Repositories Audit &*

Certification: Criteria and Checklist (TRAC), o *Catalogue of Criteria for Trusted Digital Repositories*, o *Digital Repository Audit Method Based on Risk Assessment* (DRAMBORA) e extratos da norma ISO/IEC 27001 *Information Technology – Security techniques – Information security management systems – Requirements* com relação à segurança.

A auditoria de repositórios digitais se constitui como uma importante ferramenta para que seja possível auxiliar os repositórios no caminho da preservação digital de seus conteúdos. Para que se alcance um efetivo processo de preservação digital, torna-se necessário perpassar pelos parâmetros de certificação digital e, consequentemente, trabalhar a curadoria digital nesses repositórios.

Santos e Flores (2015, p. 210) afirmam que "o processo de auditoria consiste em verificar e avaliar as metodologias adotadas pela instituição", corroborando então para averiguar a conformidade do repositório digital com relação as normas e o comprometimento com as ações de preservação digital no que tange a estrutura tecnológica, física e técnica. Assim, a partir dos resultados da auditoria, é possível avaliar o grau de confiabilidade que o repositório possui, sendo possível obter ou não a certificação de um repositório digital confiável (SANTOS; FLORES, 2015).

Térmens e Leija (2017) explicam que os sistemas de auditoria permitem saber se o repositório digital é seguro e se é possível confiar nele. De forma geral, os métodos de auditoria corroboram para certificar se as ações de preservação digital estão sendo aplicadas de forma adequada, podendo então constatar pontos positivos ou negativos de um repositório. As auditorias tradicionais podem se apresentar de forma onerosa para uma instituição, cabendo então conhecer as diversas formas de auditoria e suas aplicações – além de que essas auditorias necessitam de profissionais capacitados para que sejam realizadas.

Existem algumas iniciativas importantes no processo de auditoria de repositórios digitais – inclusive algumas já listadas anteriormente. Diante disso, é importante salientar que cada uma contribui de forma específica e, ao mesmo tempo, podem ser utilizadas de maneira geral para todos os repositórios digitais. A metodologia TRAC, por exemplo, apresenta uma série de critérios usados como referência para a certificação dos repositórios digitais, disponibilizando ferramentas para que seja possível realizar auditoria, avaliação e certificação potencial de repositórios. De forma geral, o TRAC possibilita desenvolver critérios de identificação dos parâmetros que vem a promover o acesso ao documento digital ao longo dos anos, contribuindo assim para que seja alcançada a certificação digital desses repositórios digitais (SANTOS; FLORES, 2015).

O DRAMBORA é constituído de ferramentas que possibilitam a auditoria interna de repositórios digitais. Assim, essa metodologia possibilita que um administrador de repositórios digitais conheça e avalie seus pontos positivos e negativos com relação aos aspectos da preservação digital, identificando e avaliando os riscos que o repositório possui a fim de sanar e se adequar às boas práticas mais confiáveis (SANTOS; FLORES, 2015).

A norma ISO 16363 deriva do TRAC (Trustworthy Repositories Audit and Certification Checklist), publicado em 2007 pelo Research Library Group (RLG) e o National Archives and Records Administration (NARA), tendo-se configurado também em norma ISO no ano de 2012.O TRAC tem como objetivo fornecer uma ferramenta que permita auditar, avaliar, e certificar os repositórios digitais, identificando para isso a documentação necessária para realizar uma auditoria e estabelecendo metodologias apropriadas para determinar a robustez e a sustentabilidade do repositório digital.

Frente a isso, Carvalho (2015) relata que a norma ISO 16363 não possui uma indicação referente ao nível de conformidade de determinado repositório com relação aos parâmetros de auditoria, ou seja, à luz da norma é possível apenas saber se o repositório se enquadra ou não em determinado aspecto, cabendo ao auditor determinar se as características do repositório são suficientes ou não para assegurar o cumprimento da norma. Assim, observando essa problemática, o próprio autor adotou níveis de maturidade em repositórios digitais – a partir das recomendações e avalições de auditoria da norma ISO 16363.

Segundo Carvalho (2015, p. 5), os respectivos níveis de maturidade de repositórios digitais, constante do quadro 1, podem ser evidenciados da seguinte forma: 1 – Inexistente; 2- Incipiente; 3- Em formação; 4- Operacional e 5- Proativo. No nível *inexistente*, "o repositório não implementa quaisquer processos que poderão ir de encontro às exigências do requisito normativo". No nível *incipiente* "o repositório está consciente da necessidade de existirem processos para suprir o requisito, porém estes não se encontram devidamente formalizados ou são realizados de forma *ad-hoc".* No terceiro nível – *Em formação* – pode verificar que "o repositório possui processos definidos que satisfazem o requisito normativo, porém estes ainda não se encontram totalmente implementados e/ou disseminados". No nível *Operacional* "existem políticas, procedimentos e processos implementados que satisfazem as exigências do requisito normativo". Por fim, no nível *Proativo*, "existem políticas, procedimentos e processos devidamente enquadrados num sistema de gestão que visa a monitorização e a melhoria contínua tendo por base um plano estratégico assente em factos, i.e. objetivos, metas e indicadores".

Quadro 1 –Níveis de maturidade no cumprimento dos requisitos normativos

Nível	Designação	Descrição
1	Inexistente	O repositório não implementa quaisquer processos que poderão ir de encontro às exigências do requisito normativo.
2	Incipiente	O repositório está consciente da necessidade de existirem processos para suprir o requisito, porém estes não se encontram devidamente formalizados.
3	Em formação	O repositório possui processos definidos que satisfazem o requisito normativo, porém estes ainda não se encontram totalmente implementados ou disseminados.
4	Operacional	Existem políticas, procedimentos e processos implementados que satisfazem as exigências do requisito normativo.
5	Proativo	Existem políticas, procedimentos e processos devidamente enquadrados num sistema de gestão que visa a monitorização e a melhoria contínua tendo por base um plano estratégico, objetivos, metas e indicadores.

Fonte: http://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf

No quadro 1 podem ser visualizados os níveis de maturidade baseados na norma adotada pelo TRAC, que permite à equipe de auditoria identificar um nível de conformidade de cada repositório em relação do referencial normativo adotado e com base nessa maturidade elaborar um plano de ações de melhoria específico para cada repositório.

3 Preservação de Documentos Arquivísticos Digitais Confiáveis

A preservação é fundamental para garantir o acesso a documentos digitais autênticos a longo prazo, mas envolve alguns desafios, como os que envolve o avanço tecnológico, a adoção de práticas e políticas efetivas para evitar o comprometimento de sua autenticidade e acesso a longo prazo.

Para estabelecer um repositório digital confiável deve-se basear no modelo de referência Open Archival Information System (OAIS). De acordo com o Conarq é necessário o uso de parâmetros para que os repositórios digitais sejam confiáveis, para garantir a autenticidade, a confiabilidade, o acesso e preservação. Com esse entendimento percebe-se que a preocupação não está somente no avanço tecnológico, mas também em acompanhar e fazer manutenção dos acervos documentais por longos períodos, de acordo com as *Diretrizes para a implementação de repositórios arquivísticos digitais confiáveis*, do Conarq (2015, p. 5).

Entretanto, os documentos digitais possuem complexidade e especificidades, tais como, a vulnerabilidade e a facilidade de alterar, reformatar e falsificar, tanto o seu conteúdo, quanto o seu formato e suporte; sem deixar qualquer vestígio, comprometendo sua autenticidade e o seu acesso futuro (CONARQ, 2004; 2011; 2012; CORRÊA, 2010; INNARELLI, 2006; FERREIRA, 2006; INTERPARES, 2007a). É dentro deste contexto que existe uma

preocupação em relação a vulnerabilidade dos documentos digitais dentro de uma visão mais ampla em relação as estratégias de preservação digital, as políticas organizacionais adotadas, a obsolescência tecnológica e a capacitação de pessoal qualificado.

4 Repositório Digital Confiável

Para que o repositório digital seja considerado confiável poderá ser avaliado por meio de um processo de auditoria interna segundo a norma ISO 16363 (Audit and certification of trustworthy digital repositories) que é usado como mecanismo e instrumento de medida, em que será demonstrado a confiabilidade e autenticidade do repositório digital, assim como a política institucional, segurança e estratégias de preservação digital. Esta norma apresenta três dimensões a saber: Infraestrutura Organizacional, Gestão de Objetos Digitais e Infraestrutura e Gestão de Segurança. Sendo cada uma delas com sua funcionalidade:

1) Infraestrutura organizacional: abrange a estrutura governativa do repositório e a sua viabilidade organizacional, analisando as questões relacionadas com a gestão do repositório, processos e recursos humanos afetos. Inclui ainda alguma documentação como a política de preservação, documentação dos processos relacionados com o repositório, a sustentabilidade financeira das instituições que gerem os repositórios, os contratos associados ao serviço, licenças e a missão e responsabilidades do serviço de repositório prestado;

2) Gestão de objetos digitais: analisa o processo de ingestão e gestão de objetos digitais do repositório, ou seja, a forma como incorpora a informação digital assim como a criação e gestão dos Pacotes de Informação de Arquivo (AIP). Engloba ainda o planejamento da preservação e a forma como os AIP são preservados. Finalmente, observa as componentes de gestão de informação do serviço e a gestão de acessos; e

3) Infraestrutura e gestão da segurança: engloba as questões técnicas relacionadas com a gestão e controle de riscos inerentes à infraestrutura e a gestão da segurança. Este componente relaciona-se diretamente com as infraestruturas de alojamento disponibilizados pelo serviço.

Essa metodologia apresenta-se sob a forma de um *checklist* que contém os requisitos que um repositório digital deve possuir para que se possa tornar confiável e seguro. O *checklist* serve, numa primeira instância, como ferramenta de autodiagnostico e, posteriormente, de auto avaliação. Este instrumento permite identificar potenciais omissões e pontos de falha nos sistemas e organizações responsáveis pela preservação da informação digital (RLG *et al*., 2007).

Por meio da sincronia de estratégias, políticas e sistemas informatizados, definindo, executando e gerenciando as atividades, respectivamente será possível minimizar os riscos de perda, e chegar a níveis seguros.

5 Auditoria de repositórios institucionais de dados científicos

Existem outras recomendações para realizar auditoria, são elas:

- Audit and Certification of Trustworthy Digital Repositories (ACTDR): é um documento de recomendação técnica utilizado como base para fornecer auditoria e certificação de confiabilidade aos repositórios digitais. Ele fornece uma especificação detalhada de critérios pelos quais os repositórios digitais devem ser auditados.
- Catalogue of Criteria for Trusted Digital Repositories da Network of Expertise in long-term STORage(NESTOR): um catálogo de critérios para repositórios digitais confiáveis para preservação de longo prazo, seguindo uma abordagem baseada na comunidade alemã.
- Data Seal of Approval (DAS): garante que os dados arquivados ainda possam ser encontrados, entendidos e usados no futuro, realizando uma auto avaliação para repositórios digitais de dados de pesquisa.

Os cumprimentos dos padrões de auditoria de repositório têm sido frequentemente citados como uma barreira à participação em iniciativas de preservação colaborativa. Um exemplo desse impedimento acontece na Comissão Europeia que organizou uma série de reuniões para discutir uma abordagem a nível europeu, e estabeleceu um *Memorando de Entendimento que define o Quadro Europeu de Auditoria e Certificação de Repositórios Digitais*. Este memorando cria uma abordagem em camadas para a certificação, permitindo uma auto avaliação de nível de entrada e revisão de pares com base no Selo de Aprovação de Dados, uma auto avaliação mais ampla (com base nas normas DIN 31644 ou ISO 16363) e uma escala externa de Auditoria baseada na norma ISO 16363.

Figura 1 - Selo de Aprovação de Dados

Fonte: https://www.datasealofapproval.org/en/

O Selo de Aprovação de Dados (DSA, 2008), constante na figura 1, é um processo de auto avaliação para repositórios digitais de dados de pesquisa. Embora seja necessária uma

despesa de tempo para solicitar o DSA, é muito menos oneroso que a norma ISO 16363, com apenas dezesseis diretrizes sobre as quais a organização é avaliada. As diretrizes são baseadas nos seguintes cinco critérios:

- Os dados podem ser encontrados na Internet;
- Os dados são acessíveis (direitos e licenças claros);
- Os dados estão em um formato utilizável;
- Os dados são confiáveis;
- Os dados são identificados de forma única e persistente para que possam ser encaminhados.

Embora a DSA esteja na superfície de uma auto auditoria, essa auto auditoria é então revisada por pares antes de atribuir um selo, aumentando assim o nível de autoridade para o processo. A abertura e a transparência são encorajadas e as instituições são convidadas a disponibilizar gratuitamente suas evidências (essencialmente documentação, políticas e procedimentos). Ao contrário de uma auditoria de acordo com a norma ISO 16363, o revisor de pares não é obrigado a visitar a instituição para ver que as políticas e procedimentos estão funcionando na prática, então este processo é muito baseado na confiança.

6 Conclusão

A auditoria é uma etapa que vem garantir ao repositório a confiabilidade que é necessária para que as pesquisas científicas sejam efetivadas de forma consistente e recuperável ao longo do tempo. Por conseguinte, de nada adiantaria a manutenção de repositórios que não buscam se manter certificados de que os procedimentos de preservação adotados são pertinentes.

Existem diversas formas de auditoria em repositórios digitais, com a finalidade de torna-los confiáveis. Vale ressaltar que todas as formas são válidas de aplicação e podem ser utilizadas por todos os repositórios institucionais, visando a salvaguarda das informações científicas produzidas ao longo do tempo.

Devido a vulnerabilidade dos documentos digitais em repositórios arquivísticos digitais é fundamental um sistema que gerencie as tendências dos padrões, crie ambientes de armazenamentos seguros e implementação de estratégias que possuam mecanismos para registrar toda e qualquer alteração ocorrida aos objetos digitais.

Um dos requisitos para avaliar um repositório digital está relacionado a sua funcionalidade oferecida pelo *software*, que diz respeito à gestão de objetos digitais, onde será avaliado a maturidade. Sendo esta evidência fornecida pelo gestor de repositório. A avaliação

faz parte de uma das dimensões da norma ISO16363/2012, Infraestrutura Organizacional, referente a missão e responsabilidades do serviço de repositório prestado.

A preservação de documentos digitais, como é o caso dos repositórios arquivísticos digitais implica constantemente em acompanhar, implementar e adotar requisitos para auditoria, como por exemplo o TRAC ou outros modelos em paralelo, que são fundamentais para a preservação em longo prazo. É necessária a implementação de políticas para os repositórios de dados de pesquisa, pois só a partir da política implementada poderá ser realizada uma auditoria de qualidade, onde será contemplada todas as exigências normativas. Existem alguns processos e requisitos que são obrigatórios para que se tenha um mínimo de aceitação em relação à auditoria do repositório. Esses processos, políticas e procedimentos devem se enquadrar em um sistema de gestão para monitorização e cumprimento dos objetivos do repositório de dados de pesquisa, estando sempre em constante melhoria.

Referências

BORGMAN C. L. **Research data: who will share what, with whom, when, and why?** In: CHINA-NORTH AMERICAN LIBRARY CONFERENCE. Beijing: CALA, 2016. Disponível em: http://works.bepress.com/borgman/238/ Acesso em: 15 set. 2017.

CARVALHO, J. Auditoria ISO 16363 a Repositórios Institucionais. IN: CONGRESSO NACIONAL BAD. 12., 2015, Portugal. **Anais eletrônico**. Évora: CNBAD, 2015. Disponível em: https://www.bad.pt/publicacoes/index.php/congressosbad/article/viewFile/1459/pdf_91 Acesso em: 25 out. 2017.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). **Requisitos Para a Implementação de Repositórios Arquivísticos Digitais Confiáveis - RDC-Arq.** 2015. Disponível em: http://www.conarq.gov.br/images/publicacoes_textos/diretrizes_rdc_arq.pdf Acesso em: 28 set. 2017.

DATA SEAL OF APPROVAL. **Selo de aprovação de dados e diretrizes**. Versão 2017-2019. Disponível em: https://assessment.datasealofapproval.org/guidelines_54/html Acesso em: 10 nov. 2016.

FONTANA, F.F. ARCHIVEMATICA como ferramenta para acesso e preservação digital à longo prazo**. Ágora**, Florianópolis, v.24, p. 62-82, 21 abril 2014. ISSN 0103-3557. Disponível em: https://agora.emnuvens.com.br/ra/article/view/457 Acesso em: 13 out. 2017.

LEITE, F. *et al.* **Boas práticas para a construção de repositórios institucionais da produção científica**. Brasília: IBICT, 2012. 34 p.

MÁRDERO ARELLANO, M. A. Preservação de documentos digitais. **Ciência da Informação**, Brasília, v. 33, n.2, p. 15-27, maio/ago. 2004. Disponível em: http://revista.ibict.br/ciinf/article/view/1043 Acesso em: 15 out. 2017.

SANTOS, H. M.; FLORES, D. Repositórios digitais confiáveis para documentos arquivísticos: ponderações sobre a preservação em longo prazo. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 20, n. 2, p. 198-218, abr./jun. 2015. Disponível em: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2341 Acesso em: 11 out. 2017.

SANTOS, Ângela Sofia de Sousa. **Repositório Digital Seguro: um modelo de requisitos para um provedor de serviços de certificação.** Mestrado em Ciência da Informação. Faculdade de Engenharia Universidade do Porto), p. 32, 2015.Disponível em: https://repositorio-aberto.up.pt/handle/10216/79978 Acesso em: 09 out. 2017

SANTOS H. M. Flores D. Os fundamentos da diplomática contemporânea na preservação de documentos arquivísticos digitais. **Biblos,** Revista do Instituto de Ciências Humanas e da Informação, Rio Grande, v. 30, n. 2, 2016. Disponível em: https://www.seer.furg.br/biblos/article/view/4825/4440 Acesso em: 11 out. 2017

SAYÃO L. F.; SALES L. F. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação,** Londrina, v. 21, n. 2, p. 90-115. Disponível em: http://www.uel.br/revistas/uel/index.php/informacao/article/view/27939 Acesso em: 2 out. 2017.

SCHÄFER, M. B.; CONSTANTE, S. N. E. Políticas e estratégias para a preservação da informação digital. **Ponto de Acesso**, Salvador, v. 6, n. 3, 2012. Disponível em: https://portalseer.ufba.br/index.php/revistaici/article/view/6449 Acesso em: 19 out. 2017.

TÉRMENS, M. LEIJA, D. Auditoría de preservación digital con *NDSA Levels*. **El profesional de la información**, Barcelona, v. 26, n. 3, p. 447-456, 2017. Disponível em: https://recyt.fecyt.es/index.php/EPI/article/view/epi.2017.may.11 Acesso em: 8 nov. 2017.

THOMAZ, K. P. Repositórios Digitais confiáveis e certificação. **Arquivística.net**, Rio de Janeiro, v. 3, n.1, p. 80-89, jan./jun. 2007. Disponível em: www.brapci.inf.br/index.php/article/download/10726 Acesso em: 10 out. 2017.

UZWYSHYN, R. Research data repositories: the what, when, why, and how. **Computers in Libraries**, Westport, v. 36, n. 3, Apr. 2016. Disponível em: http://www.infotoday.com/cilmag/apr16/Uzwyshyn--Research-Data-Repositories.shtml Acesso em: 16 set. 2017.

VIDAL, A. A conservação e a preservação de documentos digitais: um desafio na era da sociedade de informação. **Revista da Faculdade de Ciências Humanas e Sociais**, Porto, v. 7, p. 144-154, 2010. Disponível em: http://hdl.handle.net/10284/2809 Acesso em: 17 out. 2017.

WEITZEL, S. R. O papel dos repositórios institucionais e temáticos na estrutura da produção científica. **Em questão**, Porto Alegre, v. 12, n.1, p. 51-71, jan./jun. 2006. Disponível em: http://www.redalyc.org/html/4656/465645954004/ Acesso em: 28 out. 2017.

Recebido/Recibido/Received: 2017-11-21
Aceitado/Aceptado/Accepted: 2017-12-29