

SEÇÃO 6

MODELO DE REFERÊNCIA OAIS E NORMAS ASSOCIADAS

MÓDULO 7

Gestão e preservação de documentos digitais

SEÇÃO 5

Modelo de referência OAIS e normas associadas

Adaptação do Arquivo Nacional da Costa Rica

Versão 1, 2024

Este curso foi traduzido e adaptado pela Direção Geral do Arquivo Nacional da Costa Rica, em colaboração com a Seção de Arquivologia da Universidade da Costa Rica, a partir do material original de 2011 da Associação Internacional de Arquivos Francófonos, disponível online no Portal Internacional Arquivístico Francófono. Esclarece-se que podem existir variações em relação ao conteúdo original. Para acessar o material em francês, visite <https://www.piaf-archives.org/se-former/module-7-gestion-et-archivage-des-documents-numeriques>.

Conteúdo

Capítulo 1. Objetivo da seção	5
Capítulo 2. Primeiros elementos	8
2.1. Vocabulário.....	8
2.2. O Arquivo em seu ambiente	9
Capítulo 3. O modelo de informação.....	10
3.1. Elementos de modelagem UML.....	11
3.2. Informação de representação.....	13
3.3. Informação de Descrição de Conservação	17
3.4. Pacotes de informação.....	19
Capítulo 4. O modelo funcional	21
4.1. Entidade de Entrada	22
4.2. Entidade de Armazenamento	23
4.3. Entidade de Gestão de dados	24
4.4. Entidade de Administração	25
4.5. Entidade de Planejamento da Conservação	26
4.6. Entidade de Acesso	28
Capítulo 5. Impacto do modelo OAIS.....	29

Capítulo 6. Interface produtor-arquivo	30
Capítulo 7. Norma UNE-ISO 14641:2015 (Norma Afnor NF Z 42-013).....	34
Capítulo 8. Certificação dos Arquivos Digitais.....	36
8.1. O que é a certificação dos Arquivos Digitais?	36
8.2. Projetos sobre certificação e avaliação.....	38
8.3. Outras abordagens previstas.....	43
Capítulo 9. Conclusões.....	43
Bibliografia.....	44

Capítulo 1. Objetivo da seção

O modelo de *referência para um sistema aberto de arquivo de informações* (Reference Model for an Open Archival Information System) é uma iniciativa que surgiu na década de 1990, sendo definido pelo Consultative Committee for Space Data System (CCSDS). Este desenvolveu um modelo conceitual que contempla as características necessárias para a gestão, arquivamento e preservação a longo prazo da informação. Em 2002, tornou-se um padrão para a CCSDS, em 2003, o OAIS se transformou na norma ISO 14721, em 2012, foi realizada uma nova revisão, que levou à criação de sua segunda edição, e em 2015, a norma OAIS foi adotada como uma norma UNE.

Para maior simplicidade, chama-se modelo OAIS. É um modelo abstrato que nos permite compreender em profundidade todas as especificidades do arquivo digital a longo prazo. Seu caráter abstrato às vezes o torna difícil de compreender, mas é essencial.

Por que é essencial?

O modelo é essencial para o surgimento de soluções práticas aplicáveis no campo do arquivamento digital, pois:

- define os conceitos essenciais para a compreensão e a análise do problema,
- propõe uma terminologia que permite às diversas comunidades dialogar e se compreenderem independentemente dos vocabulários específicos dos seus respectivos domínios,
- Contribui para a elaboração de metodologias.

Portanto:

- Constitui uma base absolutamente imprescindível para a definição, elaboração e aplicação de soluções eficazes e confiáveis.
- É uma referência para o desenvolvimento de ferramentas informáticas necessárias para a automatização dos processos e da preservação da informação ao longo do tempo.



ATENÇÃO

- O modelo OAIS é um esquema de referência, não define uma implementação, não é uma aplicação, nem, um programa de computador. A implementação do modelo dependerá em parte das tecnologias, recursos e instrumentos disponíveis.
- Também é importante ressaltar que o modelo OAIS não é destinado especificamente aos profissionais de informática, mas sim a todas as pessoas envolvidas em um processo de arquivamento digital, portanto, os arquivistas devem compreender a lógica geral desse processo e guiá-lo ou participar dele de maneira eficaz.

Objetivo da seção

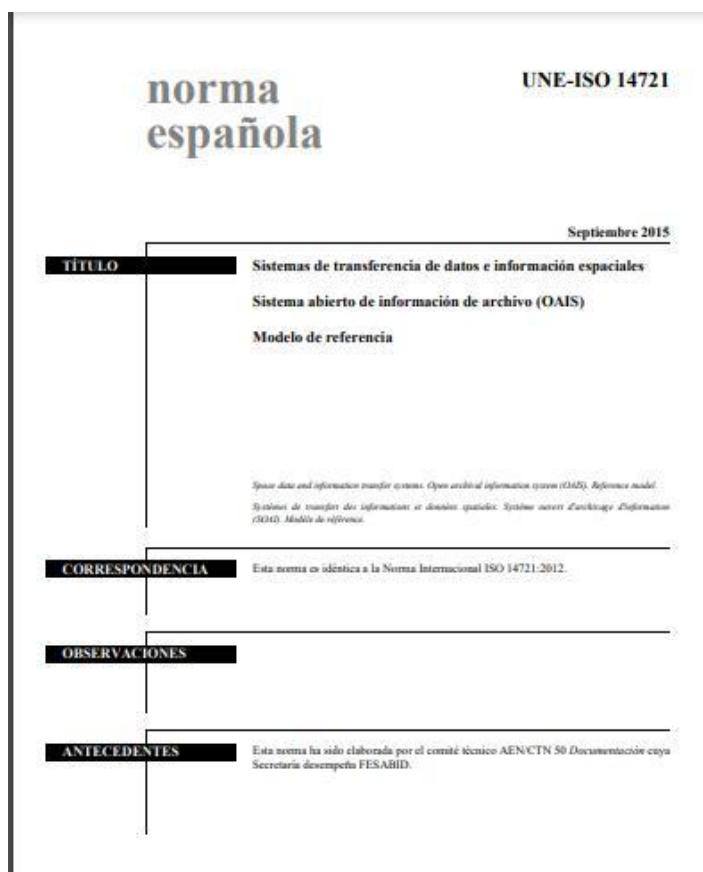
O objetivo principal desta seção 5 é apresentar o modelo OAIS tão simples quanto possível.

No entanto, a apresentação se limita ao fundamental.

Cabe ao leitor ir além, conhecendo diretamente a norma.

O que vamos tratar especificamente?

- Todos os conceitos próprios do modelo OAIS que serão utilizados constantemente neste curso; portanto, é essencial entendê-los bem.
- Várias normas derivadas do modelo OAIS, que tratam do tema da interface entre o serviço de arquivos digitais e os produtores de documentos, assim como da certificação.
- Outra norma geral de alcance mais limitado é a norma Afnor NF Z 42-013 que foi traduzida para o inglês e originou a norma ISO 14641:2015, com o nome de Especificações relativas ao design e funcionamento de sistemas informáticos para garantir a preservação e a integridade dos documentos armazenados nesses sistemas, a qual também será analisada.
- A norma ISO 15489 relativa à gestão de documentos, também deve ser consideradas uma norma geral, mesmo que não seja específica do âmbito digital.



ISO 14721(2015). Imagem de referência da capa da Norma



COMPLEMENTO

Em 1996, foi solicitado à Junta dos chefes executivos do Sistema das Nações Unidas para a Coordenação que elaborasse uma norma para especificar as condições em que seria possível manter o arquivamento a longo prazo dos dados obtidos por meio de experimentos espaciais.

A análise preliminar do CCSDS sobre este tema chegou a duas conclusões:

1. Não era possível estabelecer uma norma desse tipo porque não havia consenso sobre o tema. Em primeiro lugar, deveria ser elaborado um modelo que permitisse definir os conceitos e o vocabulário indispensáveis para a caracterização do problema que precisava ser resolvido.
2. O problema apresentado tinha poucas especificidades próprias do âmbito espacial, por isso é conveniente refletir sobre a elaboração desse modelo da maneira mais ampla e aberta possível.

Assim, o modelo OAIS foi elaborado por representantes dos principais organismos espaciais internacionais, mas com a colaboração e a contribuição contínua de representantes de arquivos nacionais e de grandes bibliotecas.

A decisão do CCSDS de abrir o grupo de trabalho encarregado do modelo, a especialistas de outros campos, de outras áreas, enfrentando o mesmo problema, realmente levou a uma reflexão transversal rica e frutífera. Essa riqueza é a causa do sucesso internacional deste modelo.



GLOSSÁRIO

Arquivo: Organização responsável por conservar a informação para que uma comunidade de usuários alvo possa acessá-la e utilizá-la (glossário do modelo de referência OAIS).

Arquivo de gestão: Documentos e expedientes abertos ou recentemente fechados mantidos nos escritórios para a tramitação dos assuntos.

Arquivo histórico: Documentos que, uma vez avaliados, serão conservados sem limite de duração.

Arquivo central: Conjunto de documentos que, ao terem deixado de ser de uso corrente, devem ser conservados temporariamente por necessidades administrativas ou jurídicas (incluindo os documentos que, uma vez classificados, serão conservados como arquivos definitivos).

Informação: A informação é definida como um conhecimento que pode ser trocado. Na prática, a informação é um elemento de conhecimento que possa ser codificada para sua conservação, tratamento ou comunicação.

Modelo de referência: Marco para a compreensão das relações essenciais entre entidades em um ambiente determinado, assim como o desenvolvimento de normas ou especificações coerentes relacionadas com esse ambiente. Um Modelo de referência baseia-se em um número limitado de conceitos unificadores e pode servir de apoio à formação e à sensibilização dos não especialistas sobre essas normas (glossário do modelo de referência OAIS).

Sistema aberto de arquivo de informação (Open Archival Information System - OAIS): Arquivo, com sua equipe e seus sistemas, encarregado oficialmente de conservar a informação e colocá-la à disposição de uma comunidade de usuários alvo.

O termo aberto em OAIS significa que esta recomendação e as recomendações e normas posteriores relacionadas com a OAIS são elaborados em fóruns abertos, e em nenhum caso o acesso ao Arquivo é livre.

Capítulo 2. Primeiros elementos

Para começar, é necessário definir alguns termos essenciais para evitar qualquer ambiguidade, identificar os envolvidos e delimitar bem o arquivo digital em seu contexto, antes de avançar mais na análise.

2.1. Vocabulário

Na seção 3 deste curso propusemos as definições dos termos informação e dados. Serão apresentadas definições propostas pelo modelo OAIS, além de duas definições novas:



DEFINIÇÃO: ARQUIVO

Organização responsável por conservar a informação para que uma comunidade de usuários alvo possa acessá-la e utilizá-la.

Este termo é equivalente a serviço de arquivo ou serviços de arquivo, abrangendo o conjunto de atividades que vão desde a assistência para arquivar até a conservação e o acesso aos documentos arquivados.

Portanto, o Arquivo garante a sustentabilidade dos dados, o acesso duradouro aos mesmos e a preservação, junto com os dados, de toda a informação necessária para sua compreensão e utilização.

O modelo de referência aplica-se a todas as disciplinas e instituições que mantêm documentos em formato digital. Os termos definidos no modelo podem não corresponder ao vocabulário de uma disciplina particular (por exemplo: arquivos tradicionais, bibliotecas digitais, centros de dados científicos).

Este vocabulário definido na estrutura OAIS, não tem como objetivo substituir as terminologias existentes. Será necessário que as disciplinas e instituições interessadas encontrem a correspondência entre alguns dos seus termos mais comuns e os termos do modelo de referência OAIS. Por exemplo: o termo frequentemente utilizado na literatura arquivística de registros não é utilizada no modelo de referência OAIS, mas pode estar vinculada globalmente ao Conteúdo informativo dentro de um Pacote de informação arquivística. Veremos esses conceitos um pouco mais adiante.



DEFINIÇÃO: LONGO PRAZO (específico para este módulo)

Prazo suficiente para levar em conta as mudanças tecnológicas e, em particular, a gestão dos novos suportes e formatos de dados, assim como a evolução da comunidade de usuários. Este período não está limitado no tempo.

Quanto aos documentos digitais, pode ocorrer que o seu prazo de vida útil seja mais curto do que os documentos em formato papel (por exemplo, no caso dos registros de pessoal, que são mantidos enquanto o agente não cessar suas funções). Parece imperativo aplicar um conjunto de disposições comuns quando se tratar de conservar documentos por um período superior a 10 anos, ou até inferior.

2.2. O Arquivo em seu ambiente

Esse primeiro enfoque é simples, permite identificar os principais envolvidos nas situações e tem como objetivo delimitar claramente qual é a responsabilidade do Arquivo e qual não é:

- O produtor cria objetos digitais que são transferidos ao Arquivo. O produtor, como é definido aqui, não é necessariamente o criador da informação. No âmbito dos arquivos institucionais, fala-se em serviços vertentes e faz-se uma distinção entre o serviço criativo - e, portanto, proprietário - da informação e o serviço vertente. O produtor facilita a informação (em forma de dados digitais) ao Arquivo no âmbito de um acordo denominado Convênio de transferência, também conhecido como Protocolo de transferência, que define as modalidades de apresentação, o modelo de dados que deve ser utilizado e os distintos componentes lógicos que devem ser fornecidos. Podem ser pessoas, sistemas ou agentes os que proporcionam a informação que será conservada.
- O usuário busca, seleciona, controla e recupera os objetos digitais que lhe interessam. São as pessoas ou sistemas que interagem com os serviços de OAIS para poder acessar à informação.
- A direção/entidade responsável (gestão) representa os responsáveis que definem a carta ou o mandato, o âmbito de competência, as prioridades e as orientações gerais do Arquivo, em coerência com a política global da instituição ou empresa e de acordo com as normas vigentes. A direção é frequentemente a origem das fontes de financiamento. Pode proporcionar diretrizes sobre o uso dos recursos (humanos e materiais), arbitra os conflitos entre os atores e avalia os resultados do Arquivo em forma de auditorias.

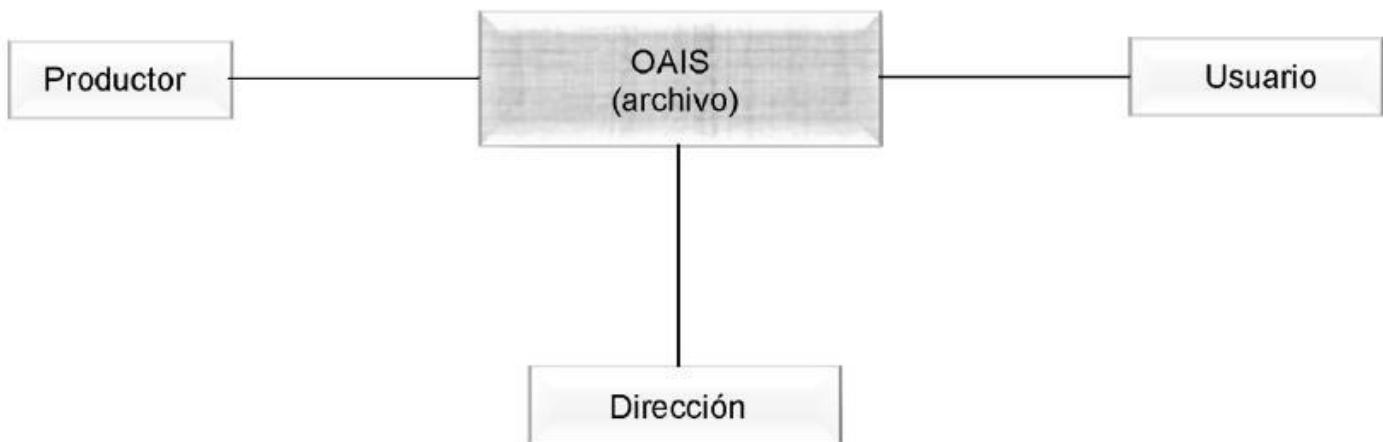


Figura 2-1 – Modelo de entorno de un OAIS

ISO 14721 (2015) p.29 Ambiente do arquivo.

Neste enfoque, o modelo OAIS posiciona o Arquivo como uma entidade com certa independência e responsabilidades próprias bem definidas. É um ponto de vista que diferencia este modelo da norma ISO 15489 dedicada à Records Management (gestão de documentos). De fato, na norma ISO 15489, analisa-se o ciclo de vida do documento (digital ou não digital) desde o momento em que é capturado ou criado até o seu arquivamento. A função de arquivo é somente uma das funções do sistema de informação da empresa.



DEFINIÇÃO: GESTÃO (segundo OAIS)

Responsáveis pela implementação de OAIS em coerência com a política geral do organismo ou da empresa.

É necessário compreender a lógica própria do modelo OAIS:

- que é a entidade funcional “Administração” que se encarrega do endereço do serviço e que assume todas as decisões operacionais
- que é o “endereço” a que define o mandato, a política global e aquela que geralmente também fornece recursos.

Capítulo 3. O modelo de informação

O modelo de informação constitui uma das originalidades da norma OAIS. Isso permite perceber realmente todas as consequências e toda a especificidade que envolve a própria natureza do documento na forma digital. Este modelo descreve os tipos de informação que são gerenciados ou administrados dentro do OAIS.

3.1. Elementos de modelagem UML

O modelo de informação OAIS baseia-se no formalismo UML (Unified Modeling Language) para representação dos objetos do domínio e mais especificamente nos chamados diagramas de classe.

Uma classe descreve o comportamento e o tipo de um conjunto de objetos que compartilham propriedades. Uma classe é um conceito abstrato que representa objetos concretos. Esses objetos concretos são chamados instâncias da classe. Para designar o conceito de automóvel, se falará da classe automotiva ou até mesmo do objeto automotivo. O objeto automóvel branco constituirá uma subclasse da anterior e herdará suas propriedades. O automóvel branco que está realmente em sua garagem será uma instância desta subclasse.

A seguir, são apresentados os principais elementos deste formalismo:

OBJETO

O objeto é representado por um simples retângulo. Elaboração própria a partir de PIAF

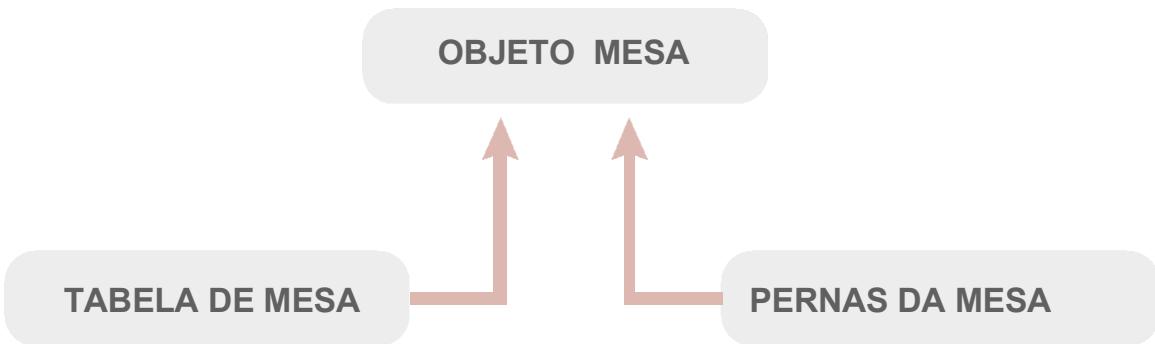
As relações são divididas em três categorias:

- as associações materializadas como acima por um traço simples entre dois objetos. Será possível explicitar, em forma de texto, a natureza desta associação.
- a especialização,
- a composição.



Relação de especialização. Elaboração própria a partir de PIAF

Um objeto (uma classe) pode estar especializado em vários objetos diferentes (subclasses) que possuem todas as propriedades do objeto pai, às quais se acrescentam propriedades adicionais.



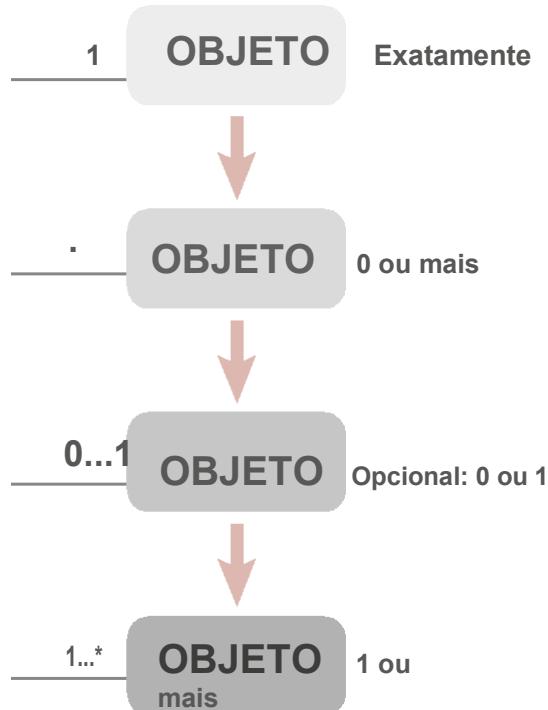
A relação de composição. Elaboração própria a partir de PIAF

Legenda:

Uma classe pode ser composta por várias subclasses. O último elemento que é preciso conhecer refere-se ao que se chama cardinalidade:

Legenda:

Essas regras de cardinalidade permitem definir o caráter opcional ou obrigatório de uma relação, assim como seu caráter único ou múltiplo.



Cardinalidade. Elaboração própria a partir de PIAF

3.2. Informação de representação



EXEMPLO

Suponha que seja necessário conservar a informação contida em um simples objeto digital, um documento digital apresentado como um único arquivo. Essa informação, que constitui o objetivo essencial, será representada em forma de um objeto denominado Conteúdo informativo. O documento é uma sequência de bits

Se só se tem uma sequência de bits, não se tem muito! Elaborado por PIAF

Seria possível pensar que basta conhecer o formato do arquivo para acessar seu conteúdo.

Legenda:

Na prática, hoje em dia existem mais de mil formatos de arquivo, cada um dos quais pode existir em múltiplas versões. Alguns desaparecem, outros são inventados.

3ds, sa, sac, abw, ace, ac3, ag, ai, aif, aifc, aiff, arc, ark, art, alim, asf, asx, au, avi, blend, bmp, cab, caf, cda, cdf, cdrl, cgm, cmt, cob, cpt, css, dat, csv, dif, DivX, diz, DjVu, doc, docm, docx, doct, dom, dob, dpa, ds, dts, dvcam, dvcpro, dwg, dxf, eAAC+, emf, eps, fits, flac, flash, flv, gif, hdf, html, h264, ifc, iges, iklax, jpg, lwo, mhtml, mng, mkv, mpeg, mp2, mp3, mp4, png, odp, ods, odt, egg, ogm, OpenEXR, ppm, pdf, pdfa, ppt, pptx, pxr, ps, psw, raw, rtf, slk, sxw, swt, sue, svg, lga, theora, tif, txd, uot, vorbis, vqf, wml, wav, wma, wmf, wv, xcf, xhtml, xis, xlsx, xslx, xtl, xml, xps, XviD, x3d, zip.....et des centaines d'autres

Quantos formatos de arquivos existem? Elaborado por PIAF

No exemplo, pode-se adicionar que o arquivo contém caracteres codificados de acordo com a norma ISO 646 (idêntica à chamada codificação ASCII). Esta norma define a codificação das letras do alfabeto sem acentuar, dos números e

de alguns caracteres especiais. Para que no futuro seja possível ler este texto, é necessário ter a certeza de que essa norma de codificação está preservada em algum lugar, ou então prever sua preservação no Arquivo.

É possível examinar o conteúdo do arquivo com um simples editor de texto que transforma cada byte do arquivo em uma representação gráfica correspondente ao caractere codificado nesse byte.

Legenda

Visualiza-se o início do arquivo com o Bloco de Notas, mas não se percebe seu significado.

Primeira revisão do arquivo com Bloc Notes. Elaborado por PIAF

Então, deve-se especificar que o arquivo consiste em uma repetição de sequências de comprimento fixo igual a 187 bytes. Cada sequência termina com o caractere de retorno de carro que significa passar para a próxima linha. Todas essas sequências estão estruturadas da mesma maneira e contêm uma sucessão de campos de informação de uma longitude definida:

- 3 números inteiros codificados cada um em 6 bytes
- uma sequência de 24 bytes com codificação interna específica,
- 3 números decimais com uma longitude de 10 bytes cada um, etc.

Com esta informação, vê-se o arquivo com um editor como o textedit, que leva em consideração os saltos de linha. De fato, nos arquivos de texto, a informação “salto de linha” apresenta uma dificuldade, pois não é codificada da mesma maneira por todos os sistemas operacionais:

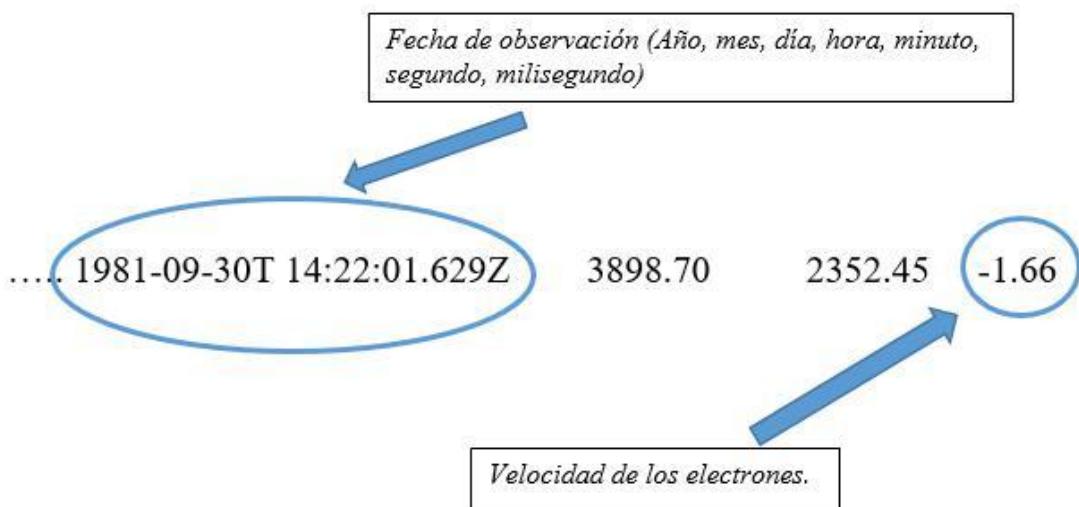
- no Windows, o salto de linha está codificado em dois caracteres (o caractere “Carriage return” + o caractere “line feed”,
- no Linux, um desses dois caracteres é suficiente para desencadear um salto de linha,
- se for lido, com o bloco de notas do Windows, um arquivo de texto criado por outro sistema, os saltos de linha não serão interpretados,
- serão interpretados se utilizarmos um editor de texto compatível com ambos os sistemas.

Legenda

Esta representação é mais clara, mas ainda falta informação essencial, o significado de cada campo.

Visualização de um arquivo levando em conta sua natureza repetitiva. Elaborado por PIAF

Portanto, é necessário definir, um por um e de maneira exaustiva, o significado de cada um dos campos contínuos na estrutura repetitiva.



Definir o significado de cada campo de informação. Elaboração própria a partir de PIAF

Velocidade dos elétrons.

Aparece através desse exemplo, **uma primeira equação fundamental**:

1. Se se deseja manter a informação contida em um objeto digital, a informação que será modelada por um objeto **Conteúdo informativo**, não basta conservar este objeto;
2. É indispensável conservar com este objeto um conjunto de informações, denominadas **Informação de representação**, que permitirá passar dos bits que constituem o objeto digital para o conteúdo informativo desse objeto.

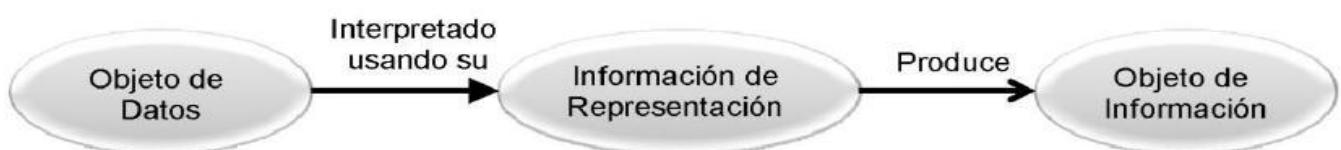


Figura 2-2 – Obtener Información a partir de los Datos

ISO 14721 (2015) p.30. Uma equação fundamental.

Esta equação é traduzida pelo seguinte diagrama:

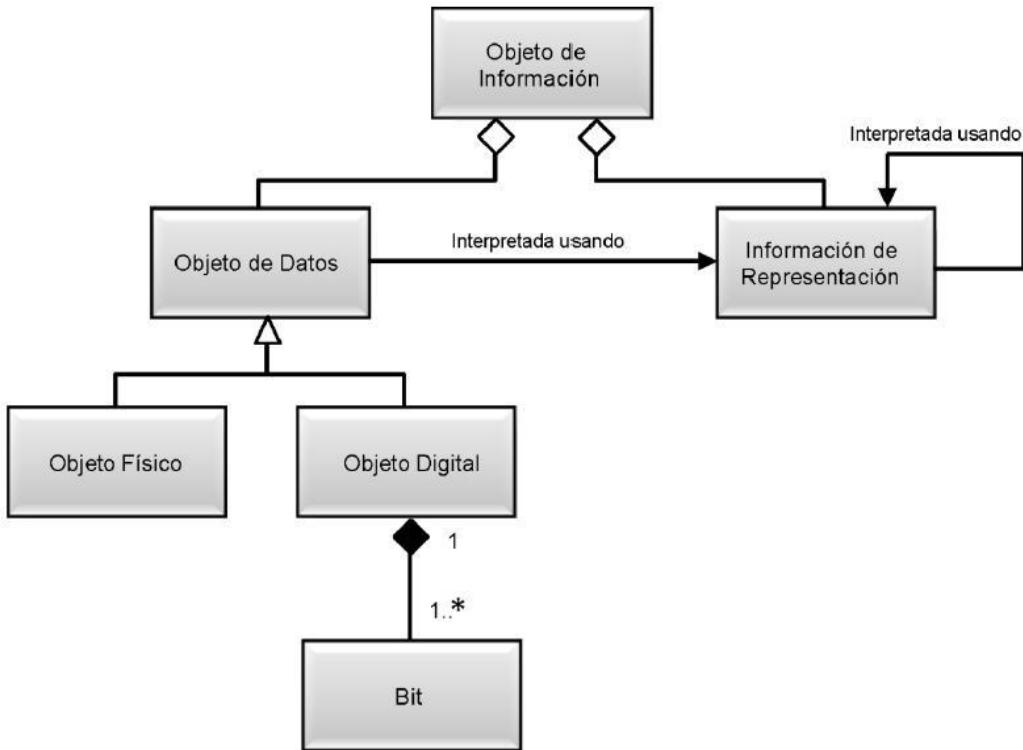


Figura 4-10 – Objeto de Información

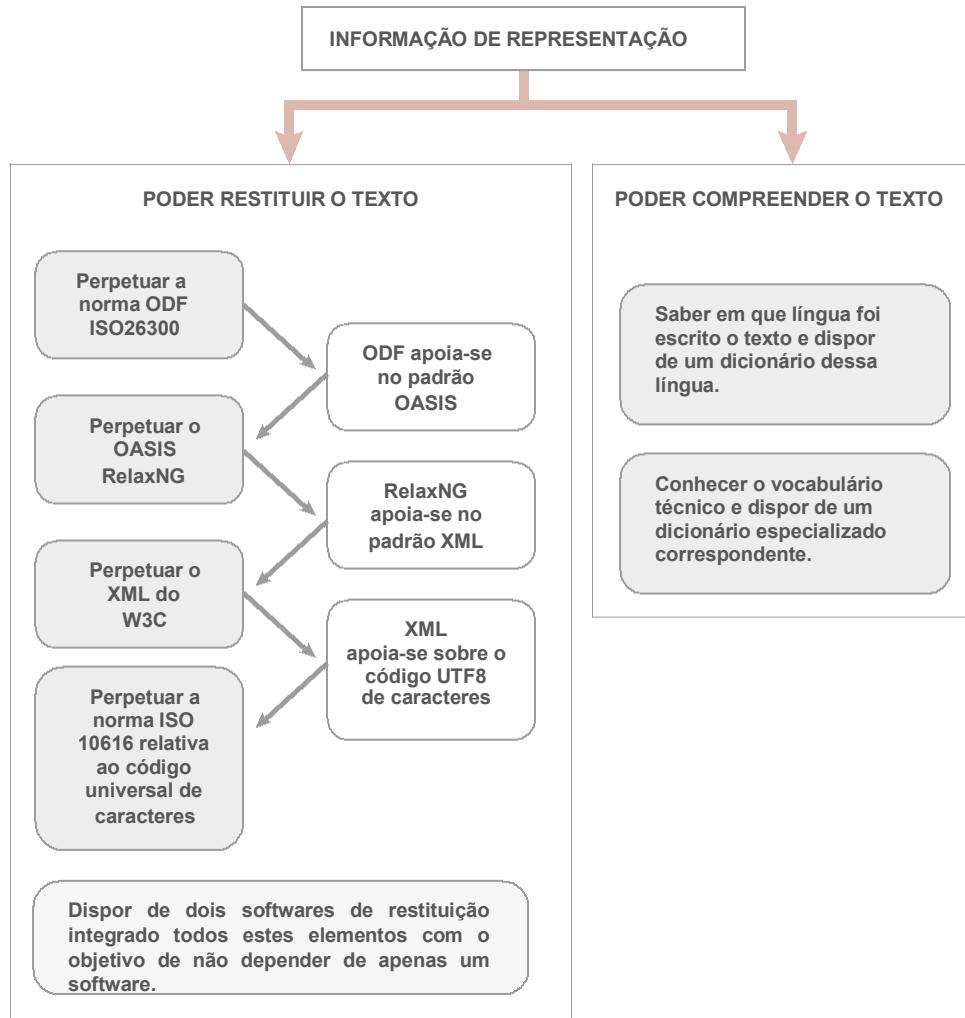
ISO 14721 (2015) p.58. O objeto dados, interpretado com informação de representação, conduz ao conteúdo informativo.

Este diagrama leva aos seguintes comentários:

- Dado que a informação de representação se baseia em elementos que também existem de forma digital (por exemplo, as tabelas de codificação de caracteres alfabéticos), ela é em si mesma um objeto digital que contém informações que devem ser preservadas, portanto, deve dispor de sua própria informação de representação. Isso é o que expressa a seta que sai e depois entra no retângulo Informação de representação,
- O objeto dados pode ser tanto numérico quanto físico (por exemplo, um livro em papel). Neste módulo de formação o foco será nos documentos digitais, mas o modelo OAIS continua válido para os documentos físicos. Isso é importante porque nos próximos anos a maioria dos serviços de arquivo terão que gerenciar simultaneamente documentos em formato papel e digitais.

A informação de representação deve em certo modo refletir a informação que se deseja conservar. Essa preocupação não é nova, quando o arquivista cria um dossiê, ele garante que seja coerente e completo. No entanto, essa preocupação adquire aqui um caráter essencial e sistemático, relacionado aos múltiplos riscos inerentes à informação em formato digital, mas também à crescente instabilidade do contexto em que vivemos.

A informação de representação poderá ser concretizada da seguinte maneira no caso de um documento textual redigido por um serviço técnico:



Caso de um documento técnico armazenado em um arquivo de escritório em formato ODF (Open Document Format). Elaboração própria a partir de PIAF

3.3. Informação de Descrição de Conservação

A informação de representação não é suficiente para que o Arquivo assuma todas as suas responsabilidades. É necessária mais informação, a qual é denominada “Informação de Descrição de Conservação”, que inclui:

Informação de contexto que descreve

- os vínculos entre a informação de conteúdo arquivada e seu ambiente,
- as razões da criação deste conteúdo informativo,
- sua relação com outros objetos de conteúdo informativo. É a informação de contexto que permitirá relacionar um objeto com um ou vários conjuntos (um fundo, um subfundo, série...).

Informação de procedência que indica a origem ou a fonte do conteúdo de informação, que traça todos as mudanças que ocorreram desde sua criação e que identifica os envolvidos nessas mudanças. As mudanças importantes podem estar relacionadas com a modificação do formato dos dados, que será ditada por razões de obsolescência de formatos antigos, para reduzir os custos de utilização de formatos pouco usados ou para responder às necessidades dos usuários cujas exigências requerem um novo formato.

Informação de referência que permite atribuir um identificador único a cada objeto digital. Um exemplo é a identificação existente para livros em formato papel (ISBN...), que adquire uma importância muito particular no digital: os identificadores vinculados a uma infraestrutura de computador ou à arquitetura de um sistema informático - como ocorre frequentemente com os endereços de recursos na Internet - têm uma vida útil limitada devido às mudanças regulares dessas infraestruturas e arquiteturas. A exigência de dispor de identificadores que sejam ao mesmo tempo duradouros e compatíveis (sem recuperação) com os identificadores utilizados pelas instituições associadas é reforçada pelo desenvolvimento de serviços interoperáveis que permitam buscar simultaneamente documentos de arquivo em várias instituições.

Informação de fixação: protege a informação de conteúdo contra alterações não documentadas. Registra os mecanismos de garantia da integridade dos objetos digitais e memoriza a pegada ou as pegadas de integridade que permitem garantir que nenhum objeto digital arquivado tenha sofrido uma modificação sem que esta tenha sido rastreada.

A essas categorias de informação, a nova versão do modelo OAIS adiciona a

Informação de direitos de acesso: oferece as condições de acesso, incluindo conservação, consulta e utilização da informação de conteúdo, também define os regimes de direitos vinculados aos dados (licenças, direitos autorais, ...), precisa as restrições e os controles de acesso aos objetos arquivados.

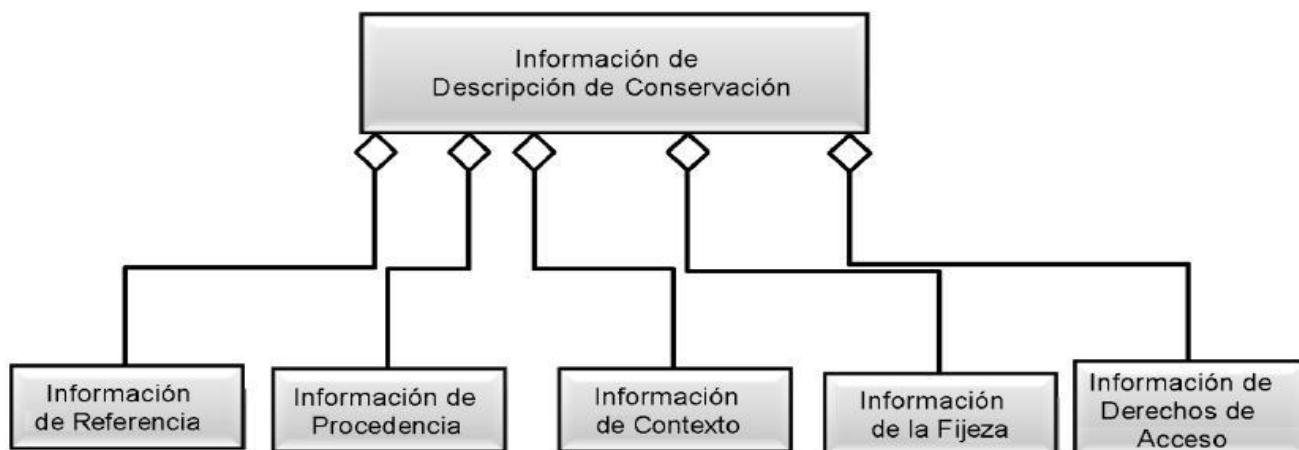


Figura 4-16 – Información de Descripción de Conservación

ISO 14721 (2015) p.71. A *Informação de Descrição de Conservação* e suas cinco subclasses.

Tabla 4-1 – Ejemplos de IDC

Tipo de Información de Contenido	Referencia	Procedencia	Contexto	Fijezas	Derechos de acceso
Datos de Ciencias del Espacio	<ul style="list-style-type: none"> – Identificador del objeto – Referencia de la revista – Misión, instrumento, título, conjunto de atributos 	<ul style="list-style-type: none"> – Descripción del instrumento – Investigador principal – Historia del procesamiento – Historia del almacenamiento y manejo – Descripción del sensor – Instrumento – Modo del instrumento – Mapa de descomutación – Especificación de la interficie de software – Descripción de la información de propiedad 	<ul style="list-style-type: none"> – Historial de la calibración – Conjuntos de datos relacionados – Misión – Historial de financiación 	<ul style="list-style-type: none"> – CRC – Suma de controles – Codificación Reed-Solomon 	<ul style="list-style-type: none"> – Identificación de la Comunidad Específica debidamente autorizada (Control de Acceso) – Concesión de permiso para la conservación y consulta – Enlaces para la Información de la Fijezas y Procedencia (por ejemplo firmas digitales o titulares de derechos)

ISO 14721 (2015) p.66. Exemplo de Informação de descrição de conservação para o caso de dados de ciências do espaço

3.4. Pacotes de informação

Um pacote de informação é um contêiner conceitual no qual são reunidas a informação de conteúdo e a informação de descrição da conservação desse conteúdo. São definidos três tipos de pacotes:

- O Pacote de Informação de Arquivo (Archival Information Package, AIP) reúne um objeto de dados (um objeto digital) e o conjunto de informações necessárias para a permanência das informações contidas neste objeto de dados, a saber, a informação de representação e os cinco componentes da Informação de Descrição de Conservação (contexto, procedência, referência, fixação, direitos de acesso). O AIP é o objeto fundamental manipulado pelo Arquivo Digital.
- O Pacote de Informação de Transferência (Submission Information Package, SIP) é aquele que é fornecido ao Arquivo pelo Produtor. Em algumas ocasiões, pode ser necessário usar vários SIP para obter o conjunto de Informação de Conteúdo e a Informação de Descrição de Conservação, ou seja, um AIP.
- O Pacote de Informação de Consulta (Dissemination Information Package, DIP) é fornecido pelo Arquivo aos consumidores de dados.

Esta noção de pacote permite, antes de tudo, dar uma realidade conceitual ao que o Arquivo deve conservar. De fato, o Arquivo deverá garantir que, qualquer que seja a solução técnica escolhida, ele disponha efetivamente das diferentes categorias de informação numérica que constituem os AIP.

A organização e os formatos escolhidos para os objetos digitais de um SIP podem ser diferentes dos escolhidos para os AIP. Da mesma maneira, não há razão para que correspondam aos formatos e à organização que pedem os usuários. Portanto, os objetos de informação transformam-se para satisfazer diferentes necessidades. Então, um DIP poderá ser criado a partir de vários AIP ou somente a partir de uma parte de um AIP.

Nas implementações, os pacotes podem ter uma realidade física (mas o modelo OAIS não exige isso, pois ele se situa no plano conceitual). Algumas instituições optaram por agrupar em uma mesma estrutura física padronizada todos os elementos dos SIPs, AIPs e, às vezes, dos DIPs. Na seção 9, sobre metadados, falaremos das normas para a criação e gestão desses pacotes de informação.

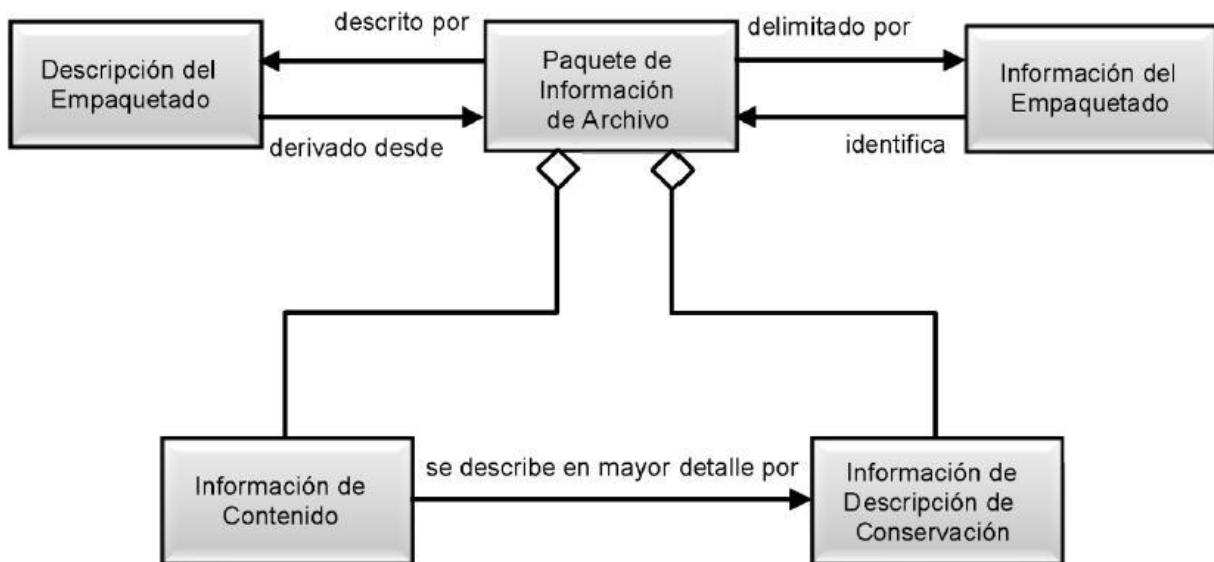


Figura 4-15 – Paquete de Información de Archivo (PIA)

ISO 14721 (2015) p.70. A representação modelada do pacote de informação de arquivo (AIP)

Este diagrama mostra objetos já conhecidos como a Informação de Conteúdo e a Informação de Descrição de Conservação, mas também mostra novas categorias de informação.

Do que se trata?

- A descrição dos links reais ou lógicos dos diferentes componentes de um pacote de informação de arquivo é o que se chama Informação de Empacotamento. Na prática, trata-se de saber como foram organizados os diferentes tipos de informação que constituem um Pacote e devemos conservar essa informação enquanto for útil: por exemplo, a distribuição dos objetos Conteúdo de dados, Informação de representação, Informação de Descrição de Conservação em um conjunto de arquivos e diretórios.

Toda a análise que precede está ditada fundamentalmente pela preocupação do armazenamento. Se deseja que os usuários possam buscar, encontrar, avaliar, selecionar, recuperar e analisar os objetos de Conteúdo de informação que atendam às suas necessidades, é recomendável disponibilizar a esses usuários uma informação denominada Informação de Descrição, que permita realizar as operações mencionadas. Essa Informação de descrição não é nova, ela é extraída da informação anteriormente descrita, constituindo o subconjunto útil para os processos de interação com os usuários. A Informação de descrição pode ser reconstruída, regenerada ou recalculada ao longo do tempo para atender às novas necessidades dos usuários. Por exemplo, corresponde ao que se denominará metadados descritivos. Em termos mais gerais, qualquer informação extraída do AIP e que desempenhe um papel no processo de busca de informações será considerada informação de descrição.

Capítulo 4. O modelo funcional

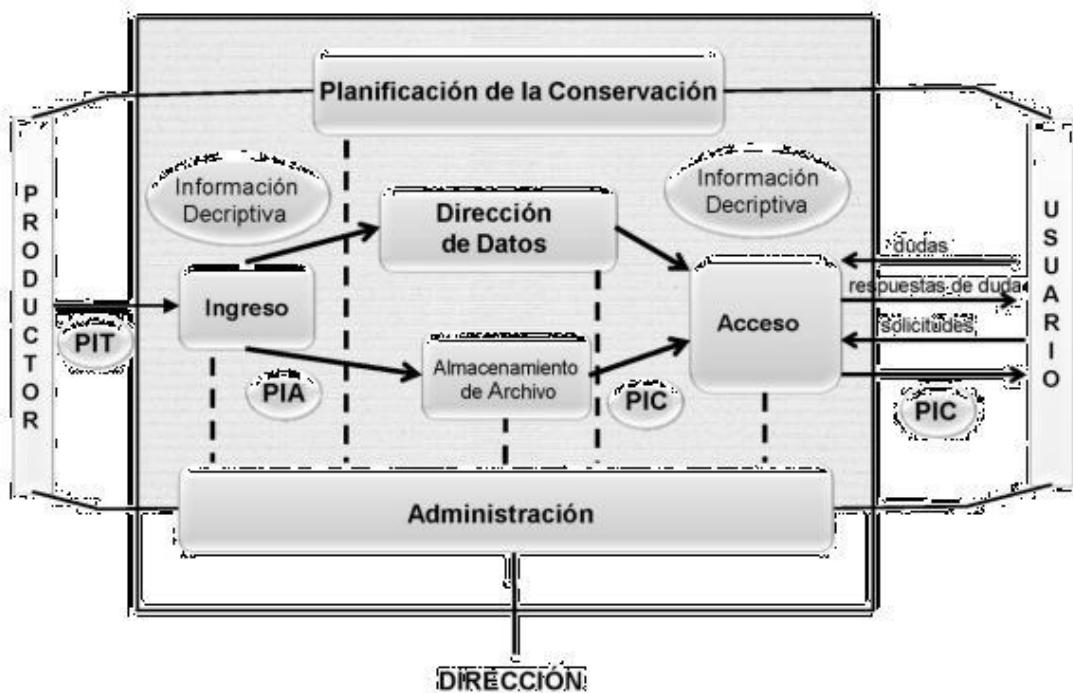
O modelo funcional da OAIS é um pouco mais fácil de entender que o modelo de informação, já que está mais perto das tarefas concretas que deve ser realizado.

O modelo é composto por seis entidades funcionais.

- A entidade Ingestão insere ou transfere para o OAIS os objetos digitais provenientes do produtor, ou seja, os SIP, para usá-los na criação dos AIP.
- A entidade Acesso transmite os objetos digitais para os usuários.
- A entidade Gestão de dados organiza a descrição do conteúdo dos fundos do OAIS.
- A entidade Armazenamento de arquivo prove os serviços e funções para o armazenamento, manutenção e recuperação dos AIP.
- A entidade Administração garante o funcionamento e a coordenação geral do Sistema de Arquivo, esta entidade inclui aspectos como negociar convênios de transferência, auditar as transferências, entre outras.
- A entidade Planejamento da Conservação é responsável pela vigilância e pelo monitoramento para garantir que o Arquivo consiga conservar os AIP em um contexto em constante mudança.

O modelo funcional também identifica:

- interfaces entre entidades funcionais
- interfaces com produtores e usuários.



ISO 14721 (2015) p.41. Esquema geral do modelo funcional.

O modelo tenta cobrir todas as atividades essenciais de um arquivo. Entretanto, o modelo não especifica nem um design, nem uma implementação particular. As implementações reais podem agrupar de maneira diferente as funções apresentadas. Qualquer implementação, para cumprir com a norma, deve ser compatível com o modelo funcional e com o modelo de informação. Isso será o que será analisado.



COMPLEMENTO

Atualmente existem soluções digitais proprietárias e baseadas em software livre que asseguram basear-se no modelo de OAIS, por exemplo a solução Preservica e Archivematica.

Leaders in active digital preservation

Ensuring the accessibility and authenticity of digital information over successive technology cycles and custodians requires a different approach to traditional backup, archiving, storage and content management.

Preservica's standards-based (OAIS ISO 14721) active preservation software combines all the critical capabilities of successful long-term digital preservation into a single integrated platform. It keeps content safely stored, makes sure it can be found and trusted, provides secure immediate access, and automatically updates files to future-friendly formats.

Oferta de serviços da Preservica <https://preservica.com/>

Basado en estándares

Archivematica es un conjunto de herramientas de software libre que permiten al usuario procesar objetos digitales desde que son introducidos en el sistema hasta su publicación acorde al modelo funcional ISO-OAIS. El usuario puede monitorear y controlar los **micro-servicios** de ingestión y preservación a través del panel de control. Archivematica utiliza estándares como METS, PREMIS, Dublin Core, la especificación BagIt (Library of Congress) entre otros estándares reconocidos internacionalmente con el objetivo de generar fiablemente paquetes AIPs (Archival Information Package) para ser grabado en su sistema de almacenamiento preferido.

Oferta de servicios da Archivemática <https://www.archivematica.org/es/>

4.1. Entidade de Entrada

Esta entidade recebe os objetos digitais transferidos pelo produtor. Gerencia o mecanismo de depósito dos SIP, os controles de acesso associados e o conjunto de interações entre o produtor e o repositório durante o processo de transferência. Verificará a conformidade dos SIPs recebidos com o que foi acordado entre o repositório e o produtor, sob a forma de um protocolo de transferência. Se necessário, a administração controlará a integridade dos objetos transferidos.

Quando os SIP forem considerados válidos, a entidade de entrada criará os Pacotes de Informação de Arquivo (AIP), que constituirão a informação de referência a ser preservada. Durante esta tarefa, deverá completar os metadados do Produtor com metadados próprios no Arquivo.

Os AIP serão confiados à entidade de armazenamento. A entidade de entrada também extrairá, a partir dos AIP, a informação descritiva que será utilizada para pesquisa e comunicação.

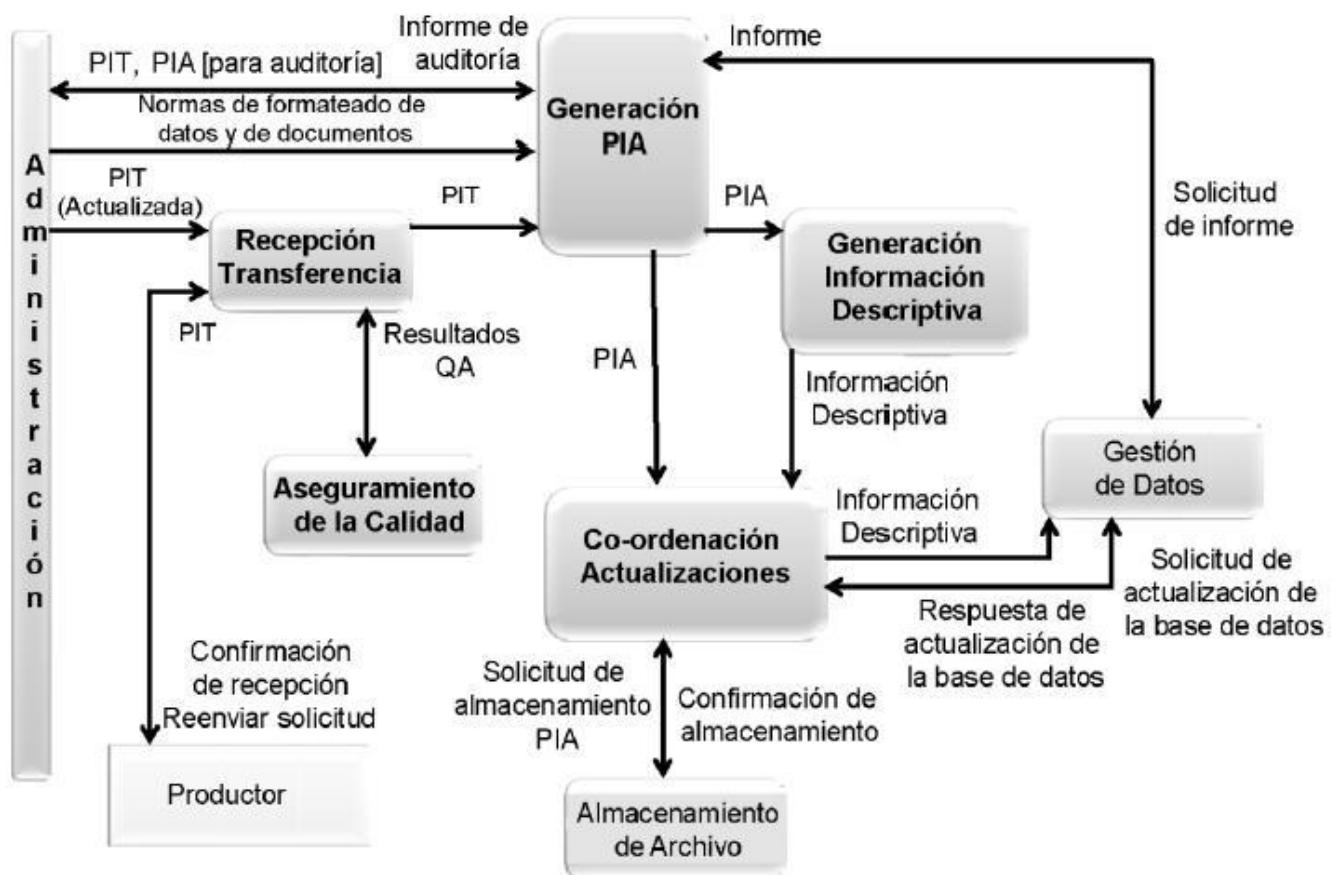


Figura 4-2 – Funciones de la Entidad Funcional de Ingreso

ISO 14721 (2015) p.44. Entidade de Entrada

4.2. Entidade de Armazenamento

Essa entidade encarrega-se esquematicamente de conservar os bits.

- suporta os AIP proporcionados pela entidade de entrada,
- e devolverá esses AIP à solicitação da entidade de acesso.

A entidade de armazenamento gerencia a hierarquia de armazenamento (armazenamento online, armazenamento com acesso restrito) para atender aos requisitos de qualidade de serviço das entidades usuárias.

Garante a migração dos suportes sempre que for necessário.

Realizará controles periódicos e contínuos da integridade da informação que lhe tenha sido confiada. Duplica os objetos digitais e armazena a cópia em uma instalação fisicamente separada para assegurar sua restituição em caso de incidente ou sinistro.

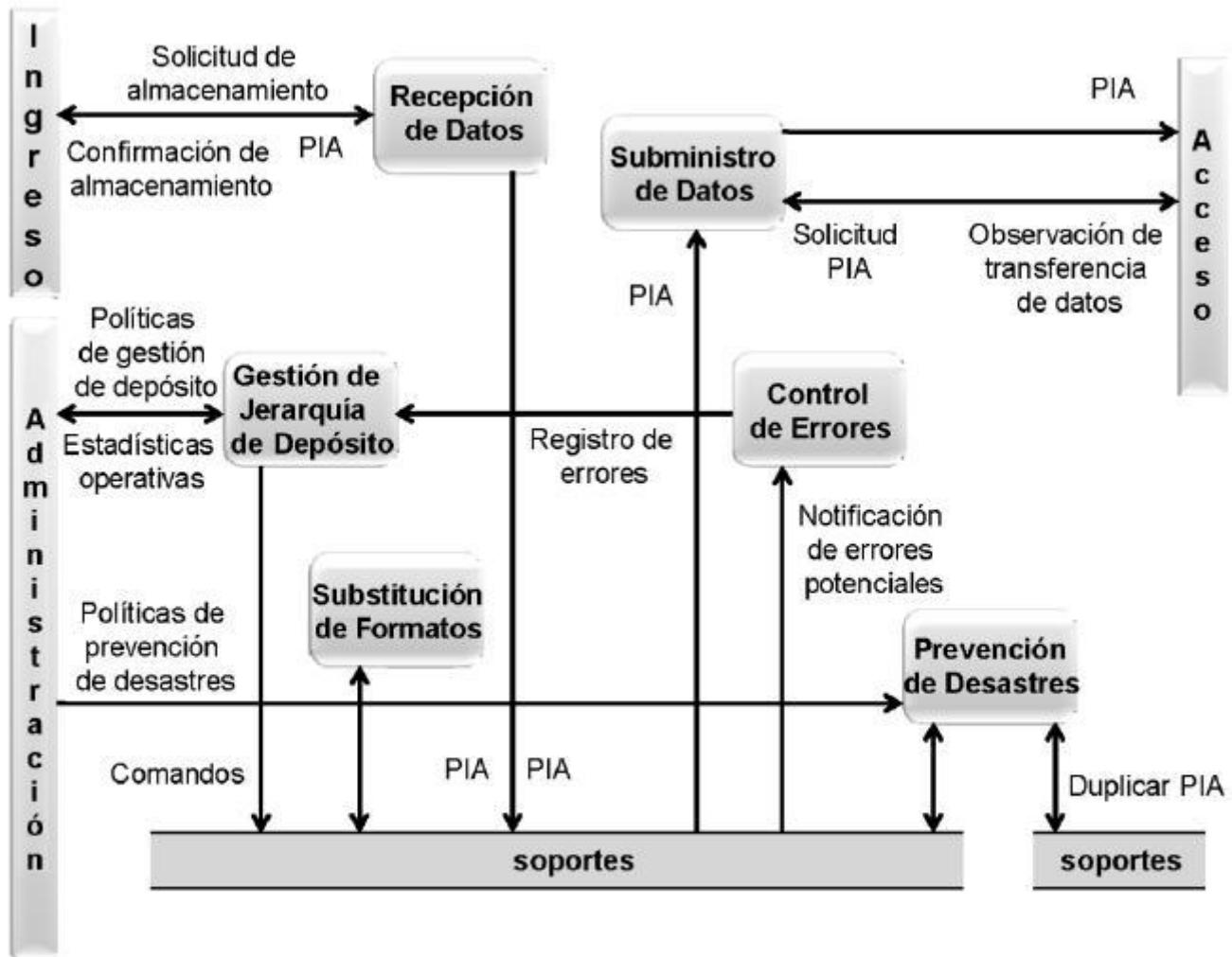


Figura 4-3 – Funciones de la Entidad Funcional Almacenamiento de Archivo

ISO 14721 (2015) p.46. Entidade de Armazenamento

4.3. Entidade de Gestão de dados

Responsabiliza-se pela manutenção e administração do banco de dados que contém a informação descritiva, esse banco de dados contém uma descrição completa do acervo do arquivo.

Esta entidade recebe consultas dos usuários através da entidade de acesso e devolve as respostas. É neste nível que entram em jogo as diferentes técnicas de busca de informação: busca por palavras-chave ou combinações lógicas de palavras-chave, busca de texto completo, indexação. Também gera relatórios sobre a situação dos fundos de arquivo, sobre a evolução e utilização desses fundos.

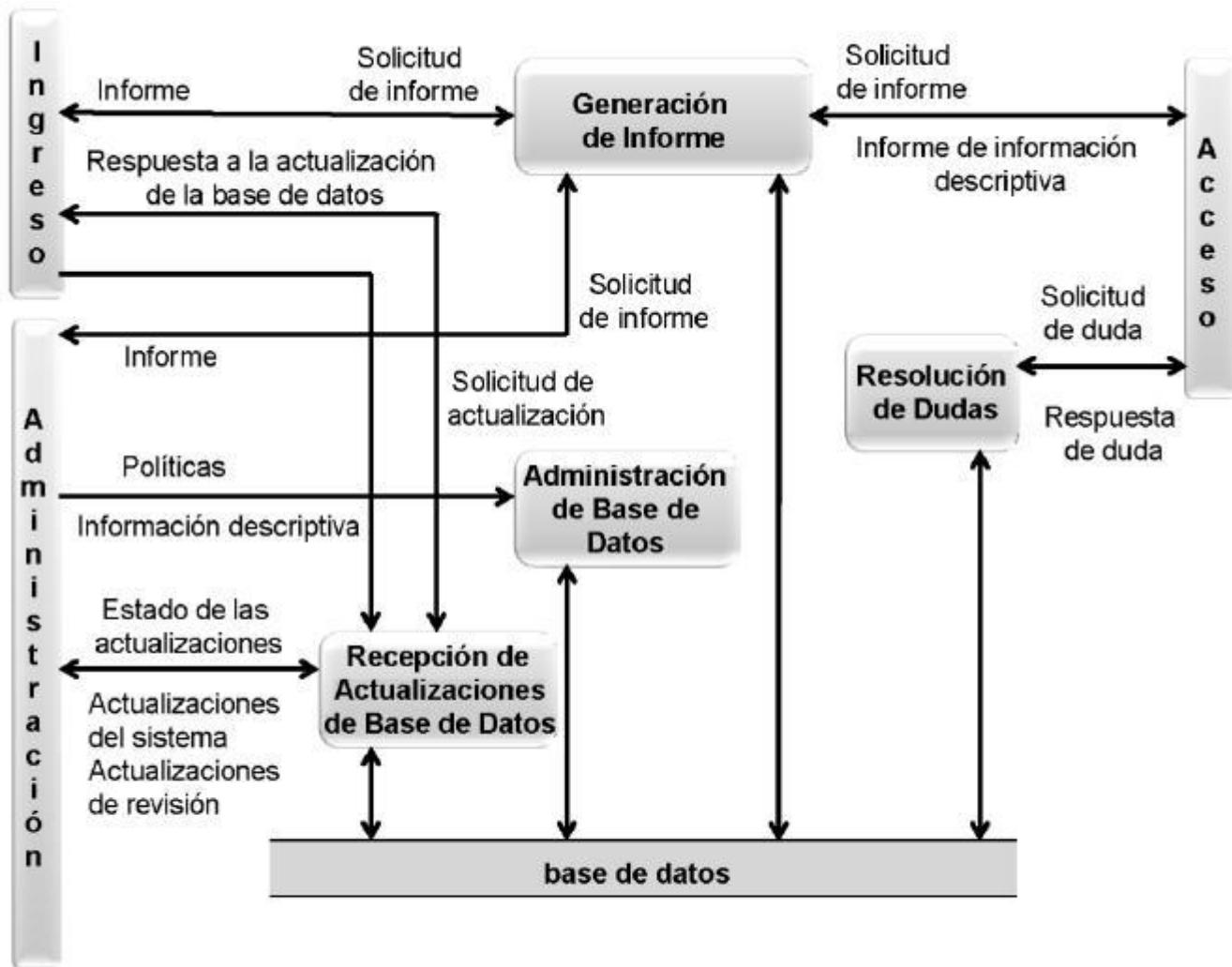


Figura 4-4 – Funciones de la Entidad Funcional Gestión de Datos

ISO 14721 (2015) p.48. Entidade de Gestão de dados

4.4. Entidade de Administração

A entidade de Administração tem uma função global. Garante o funcionamento e a coordenação geral, e toma todas as decisões relacionadas com o funcionamento do Arquivo e realiza uma série de atividades como as seguintes:

- a escolha das regras e normas aplicáveis em matéria de transferência (formatos e documentação que devem ser fornecidos), segurança (controle de acesso), migração (definição do reprocessamento dos suportes), etc., as decisões sobre a aplicação das diferentes migrações.
- a gestão do controle de acesso físico ao Arquivo,
- a administração do sistema informático.

A entidade de Administração está em contato com os agentes externos:

- negocia com os Produtores os protocolos de transferência e zela pelo respeito desse protocolo,
 - com a Direção a qual presta contas das atividades do Arquivo e da qual recebe eventualmente diretrizes.
 - com os usuários para resolver as possíveis dificuldades de acesso ao fundo.

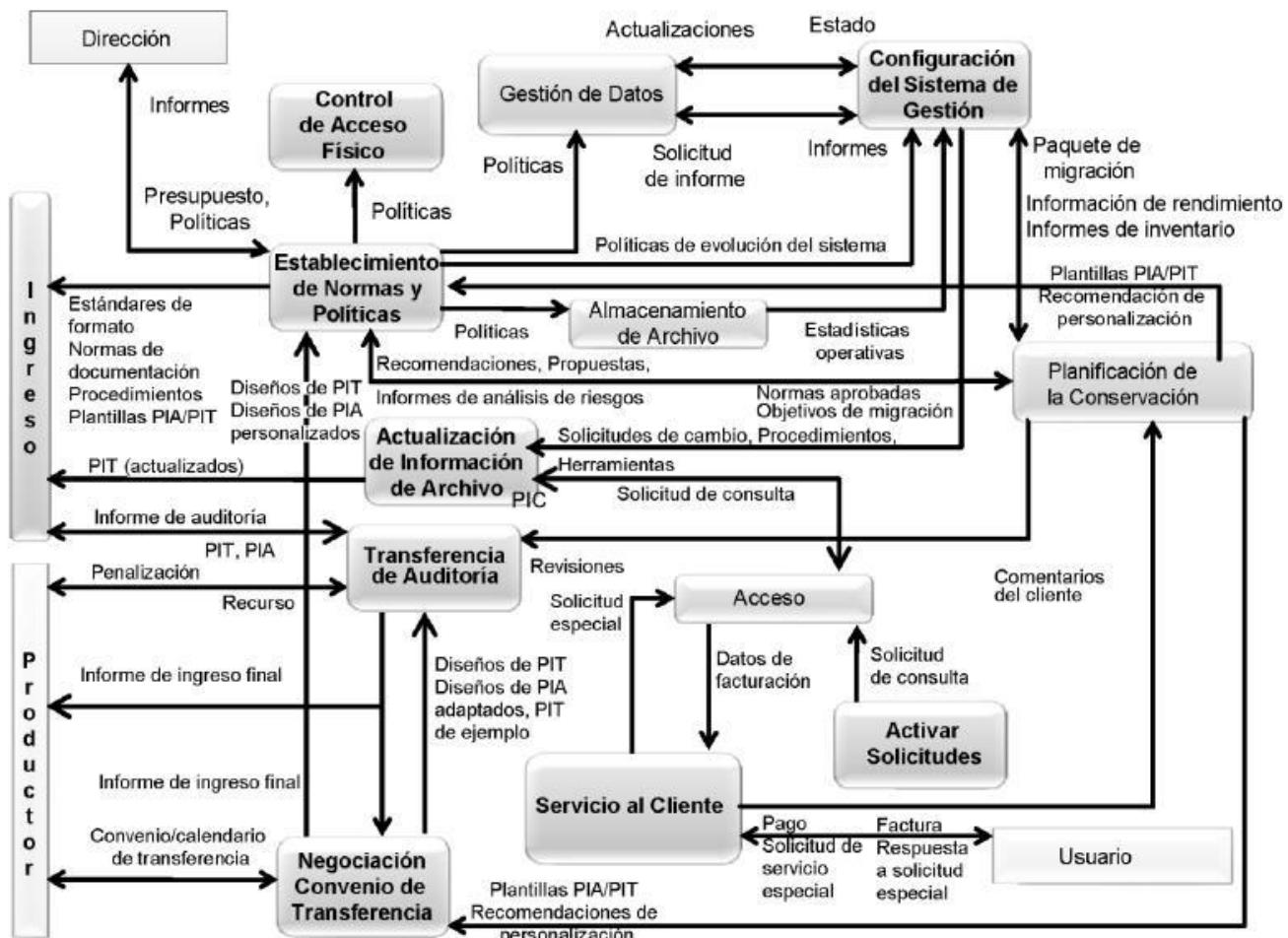


Figura 4-5 – Funciones de la Entidad Funcional de Administración

ISO 14721 (2015) p.49. Entidade de Administração

4.5. Entidade de Planejamento da conservação

Essa entidade é especial na medida em que coloca o Arquivo no contexto da obsolescência das tecnologias e da evolução das necessidades.

De fato, tem como missão analisar as mudanças em curso e futuras, assim como seus impactos sobre a capacidade do

Arquivo de conservar a informação arquivada em um formato utilizável e compreensível:

- o acompanhamento das interrupções tecnológicas envolve um trabalho de vigilância, exame e prototipagem das tecnologias para elaborar e propor planos de evolução do sistema informático do Arquivo e de migração dos dados,
- a análise das evoluções da comunidade de usuários pode levar a propor que os serviços oferecidos pelo Arquivo evoluam na direção desejada por essa comunidade, por exemplo, a implementação de novas interfaces gráficas; por exemplo, na era da Internet, qualquer serviço que permanecesse apenas com acesso à informação online, sem interface gráfica, estaria condenado a desaparecer.

Além do modelo de informação e do modelo funcional, o OAIS aborda outros temas: define uma tipologia das migrações e os modos de cooperação entre os Arquivos Digitais, que serão desenvolvidos na seção 8 sobre as estratégias de armazenamento.

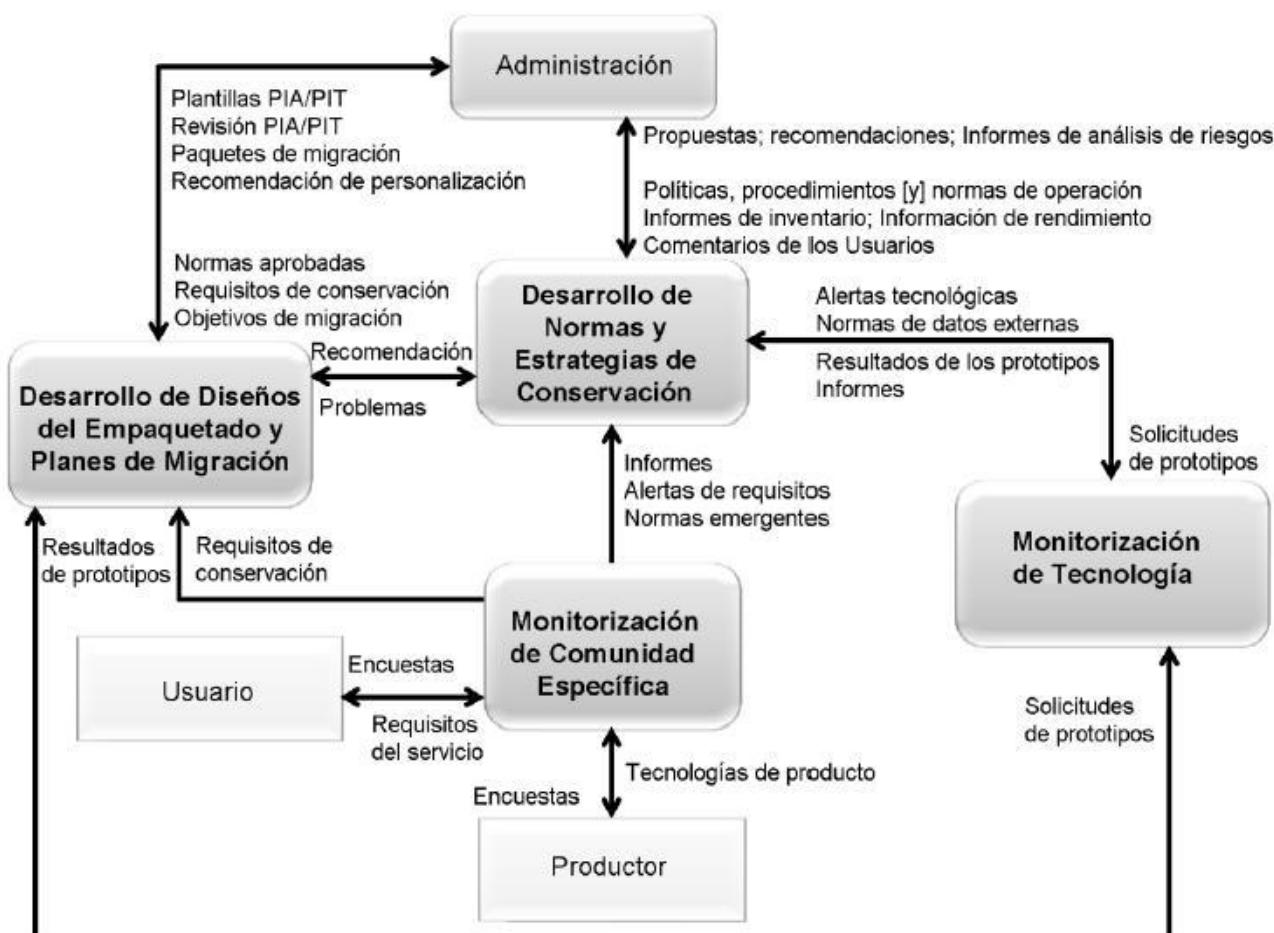


Figura 4-6 – Funciones de la Entidad Funcional de Planificación de la Conservación

ISO 14721 (2015) p.51. Entidade de Planejamento da conservação

4.6. Entidade de Acesso

Implementa a interface com os usuários.

Portanto, deve controlar o acesso dos usuários ao sistema de arquivo, assim como as permissões desses usuários para acessar ou não determinado objeto digital arquivado.

Gerencia as diferentes interações entre o sistema e o usuário: perguntas e respostas a essas consultas, pedidos de objetos arquivados, geração de Pacotes de Informação de Consulta (DIP) a partir dos Pacotes de Informação de Arquivo (AIP) recebidos da entidade de armazenamento e transferência desses DIP ao usuário.

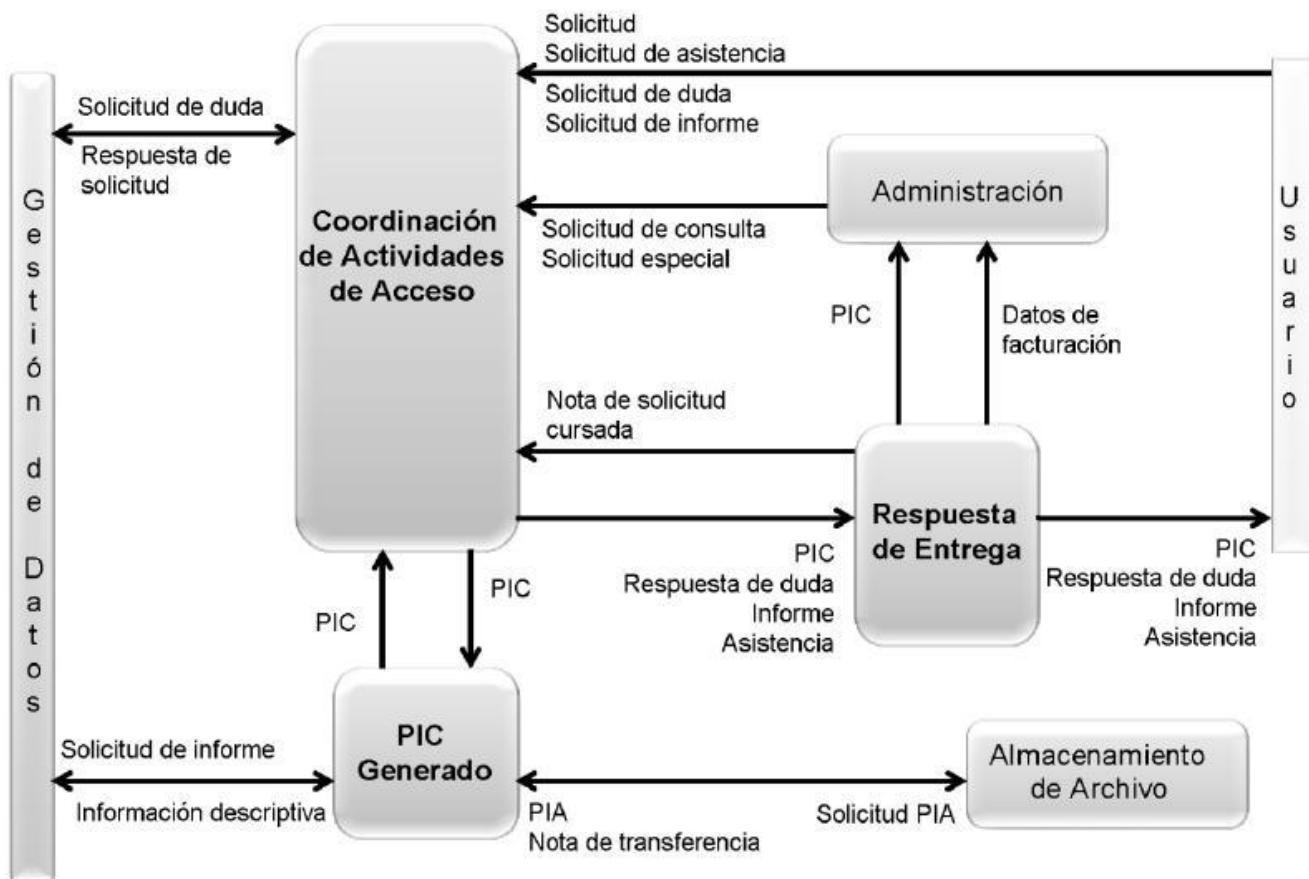


Figura 4-7 – Funciones de la Entidad Funcional de Acceso

ISO 14721 (2015) p.53. Entidade de Acesso



COMPLEMENTO SERVIÇOS DE VALOR AGREGADO

Aqui pode-se observar que o DIP não é necessariamente idêntico ao AIP. O Arquivo pode, de fato, oferecer diferentes serviços que atendam às necessidades dos usuários: fornecer os documentos de arquivo em um formato desejado pelo usuário e não necessariamente no formato em que o documento foi arquivado, extrair um subconjunto de um objeto correspondente à solicitação do usuário. Isso é particularmente interessante no caso de dados geográficos, para os quais apenas será transmitida a área solicitada pelo usuário.

Capítulo 5. Impacto do modelo OAIS

O modelo é utilizado amplamente como ponto de partida para as atividades de preservação da informação:

- por bibliotecas digitais (BnF...),
- pelos arquivos institucionais (NARA, DAF...),
- pelos centros de dados científicos (NASA nos Estados Unidos, CNES na França, ESA para Europa),
- pelas empresas industriais (por exemplo, as indústrias aeronáuticas).



COMPLEMENTO IMPLEMENTAÇÕES CADA VEZ MAIS SIGNIFICATIVAS

O modelo OAIS não especifica uma implementação, mas muitas implementações se baseiam nos conceitos do OAIS, atualmente, a utilização do OAIS é um requisito indispesável em qualquer esforço ou iniciativa de preservação digital e no design ou desenvolvimento de repositórios. Alguns casos de sucesso em que o modelo OAIS foi implementado são a Biblioteca do Congresso, o Centro de Conservação Digital do Reino Unido, no Brasil com a RDC-Arq, o arquivo de revistas acadêmicas JSTOR, a Direção de Arquivos da França, entre outros. O modelo OAIS constitui a referência fundamental aplicável.

O Modelo OAIS é uma base indispesável para a reflexão e a análise das necessidades e limitações específicas de cada empresa ou instituição na aplicação de um Arquivo Digital.

Naturalmente, o caractere universal do modelo e sua abstração lhe conferem limites. Sua aplicação concreta em um campo específico implica que se levem em consideração as especificidades dos objetos digitais desse campo, as práticas, a terminologia e as regulamentações específicas. Nenhuma organização destinada à criação de um Arquivo digital poderá prescindir dessa análise.

Capítulo 6. Interface produtor-arquivo

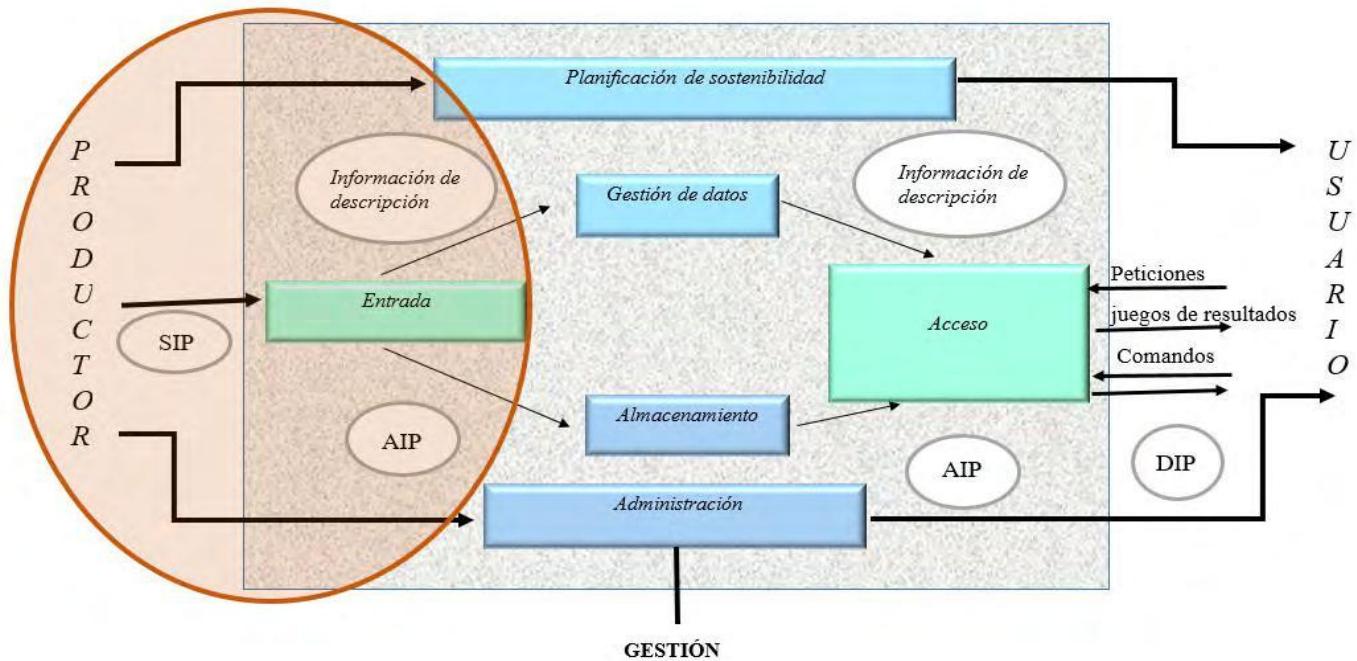
Em um processo de arquivo digital, a fase de transferência é uma das mais críticas.

Quando o Arquivo receber documentos digitais que estejam em conformidade com o previsto e devidamente descritos, grande parte dos obstáculos terá sido superada. É no momento da transferência dos objetos digitais para o Arquivo que este assume realmente a responsabilidade pela permanência da informação contida nesses objetos. Neste momento as dificuldades podem acumular-se:

- os objetos digitais recebidos não se ajustam ao que o Arquivo esperava;
- não foi definido com suficiente precisão o que o produtor deveria transferir;
- descobrimento tardio de anomalias nos objetos digitais entregues (dificuldades particulares de resolver).

Estas dificuldades prejudicam a qualidade da informação arquivada e aumentam o custo da operação. Uma série de normas intervém na fase de transferência. As normas de empacotamento são descritas com mais detalhes na seção 9 Metadados.

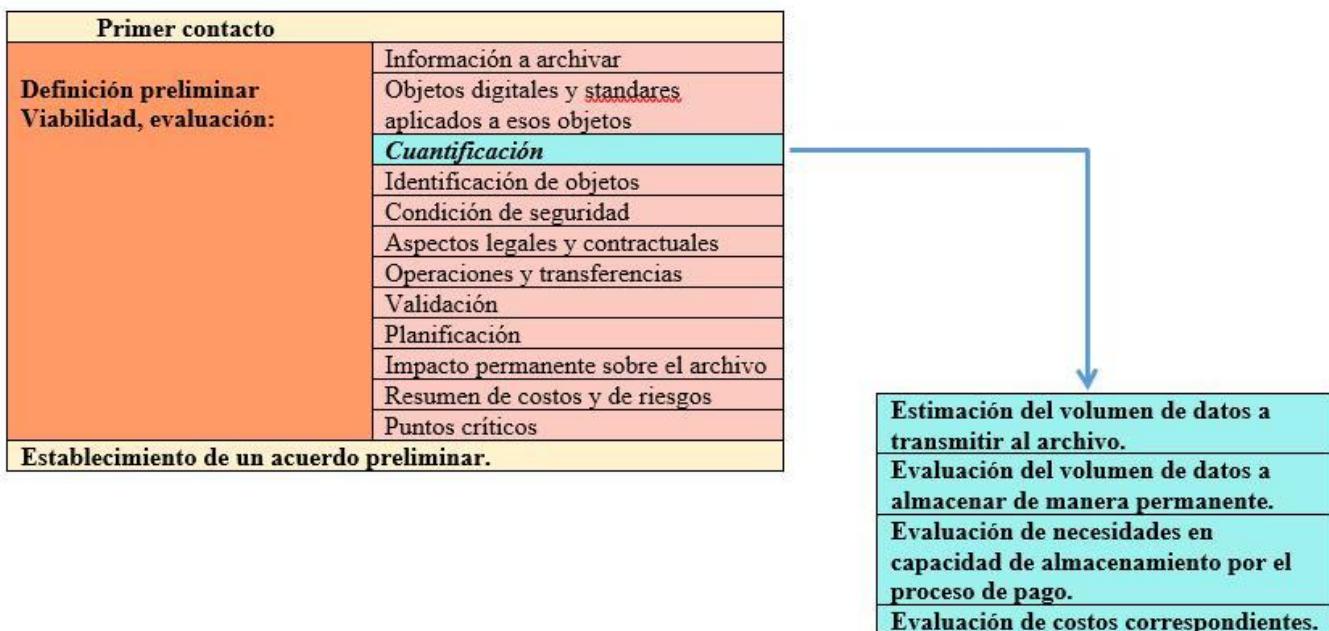
Aqui abordaremos a **norma PAIMAS (Producer archive interfaz Methodology abstract standard, ISO 20652)**, que se deriva diretamente do modelo OAIS. A norma PAIMAS pretende definir para o Arquivo e o Produtor uma metodologia de trabalho comum e estruturada em quatro fases.



Âmbito metodológico do PAIMAS. Elaboração própria a partir de PIAF

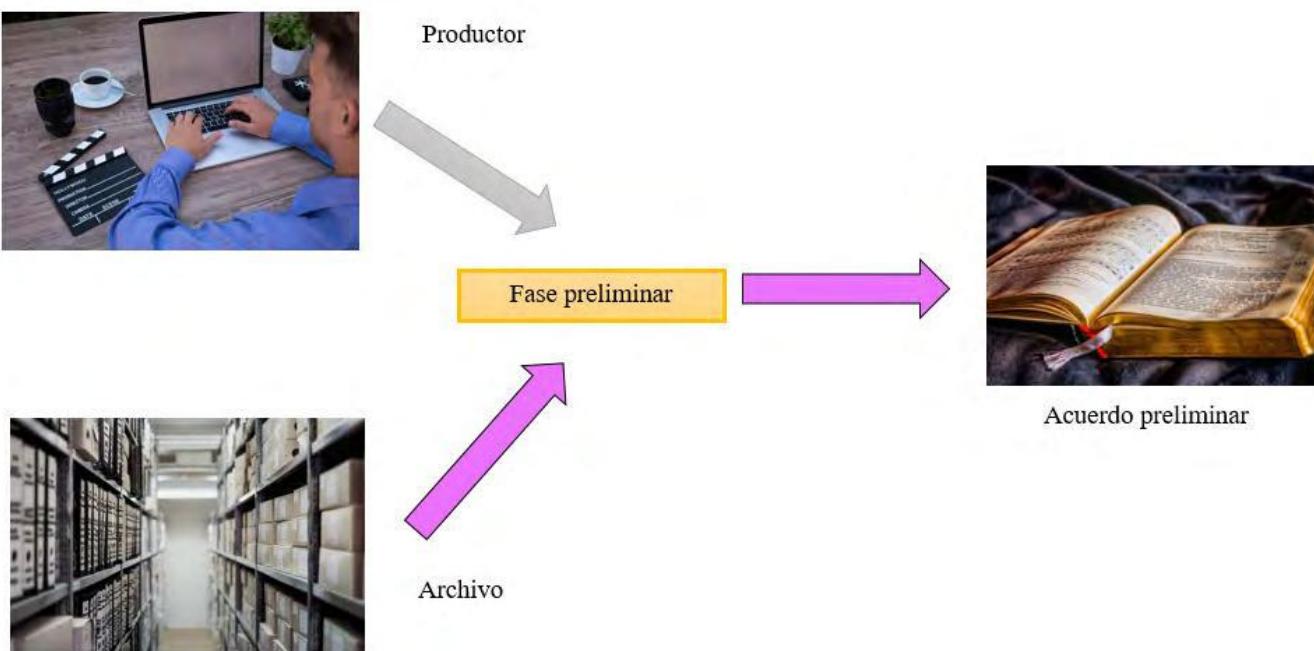
Cada fase é caracterizada principalmente por uma tabela de tarefas a serem realizadas; cada tarefa é, por sua vez, dividida em uma série de ações que devem ser executadas por um ou outro ou ambos os parceiros no processo de transferência. PAIMAS se apresenta como um guia genérica que pode ser especializada de acordo com o contexto da aplicação e as características específicas dos parceiros.

As quatro fases do PAIMAS:



PAIMAS, as tarefas da fase preliminar e a segregação da tarefa quantificação em ação. Elaboração própria a partir de PIAF

Mais concretamente, esta fase inclui uma lista de tarefas a serem realizadas e, para cada tarefa, uma lista de ações.



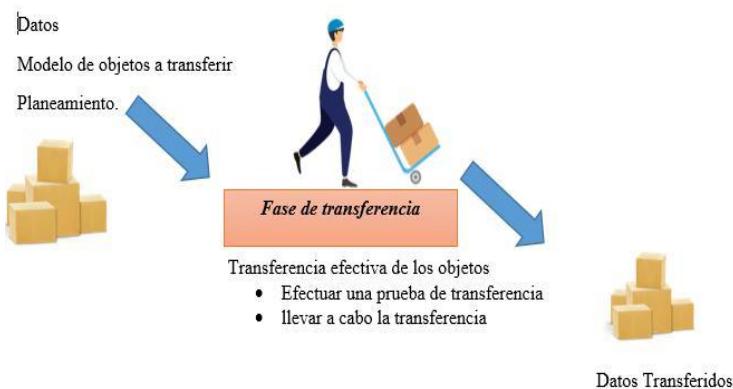
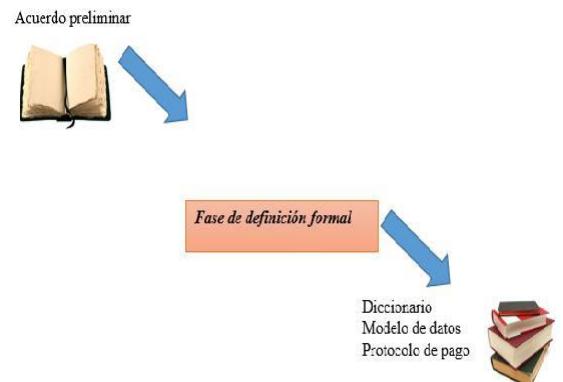
PAIMAS, fase preliminar. Elaborado por PIAF

Cada ação será objeto de uma descrição precisa. Além disso, será indicado se a ação deve ser realizada pelo Arquivo, pelo Produtor ou por ambos. O mesmo aplica-se às outras três fases:

Legenda

A **fase de definição formal** tem como objetivo negociar o protocolo de transferência. Este protocolo deverá incluir uma descrição completa, precisa e formal dos objetos digitais, o produtor e um cronograma.

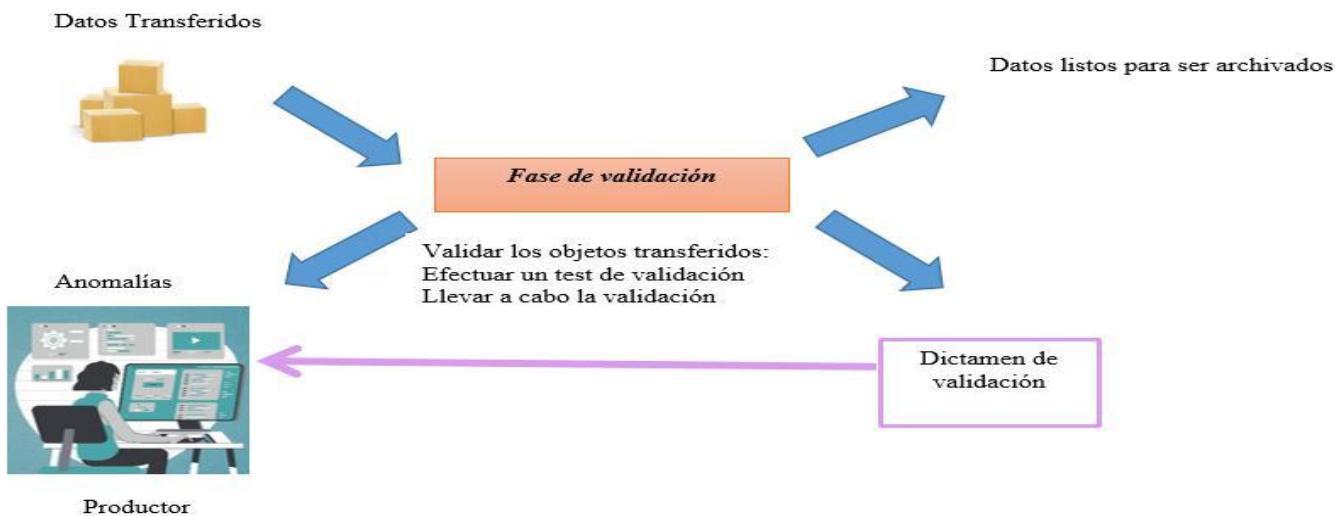
Essa descrição deve ser objeto de um modelo formal, ou seja, um modelo construído com um formalismo que permita a um indivíduo ou a uma pessoa analisar e compreender sua coerência e que também permita à máquina, ao computador, fazer um uso automatizado dela durante o processo de transferência. Esse formalismo pode ser baseado em um dicionário que descreve *PAIMAS*, a fase de definição formal. Elaboração própria a partir de PIAF



Legenda

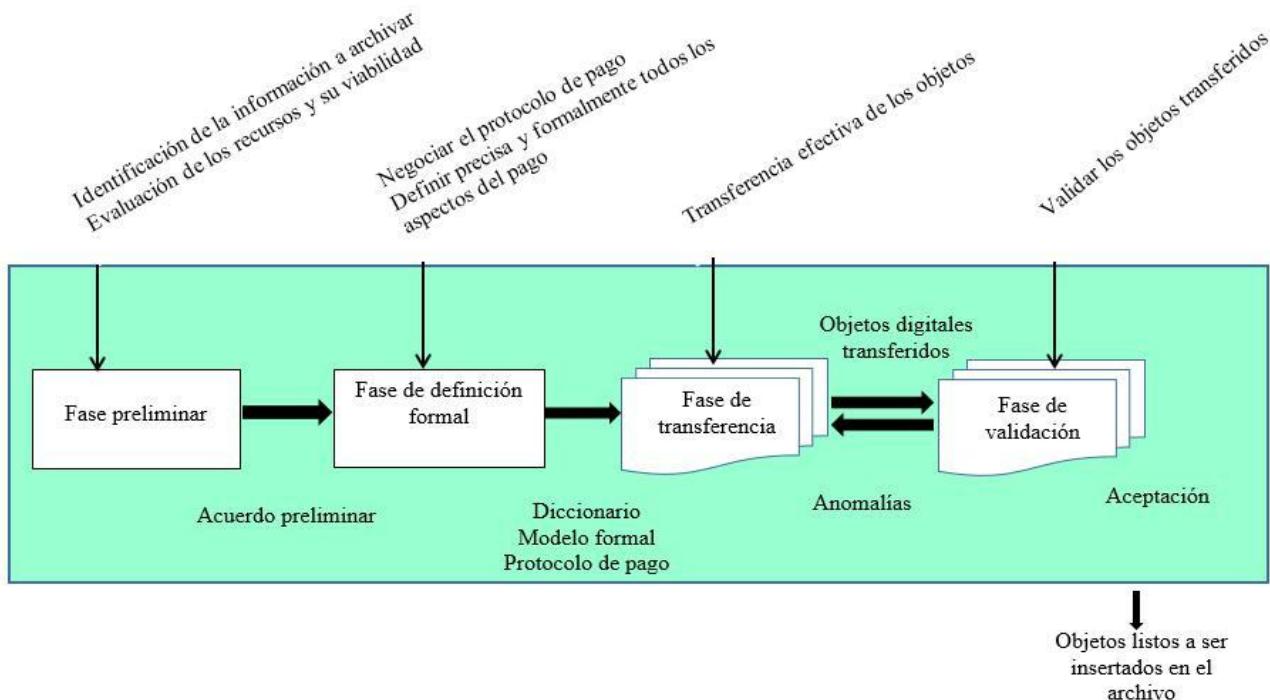
A **fase de transferência** corresponde à transferência efetiva de pacotes de informação de transferência. Inclui as provas de transferência e colocada em funcionamento.

PAIMAS; fase de transferência. Elaboração própria a partir de PIAF



PAIMAS; a fase de validação. Elaboração própria a partir de PIAF

A **fase de validação** inclui a validação de pacotes recebidos pelo Arquivo e todas as ações que esta validação pode implicar. Diferentes tipos de validação de pacotes são considerados: validação contínua de cada pacote, validação global de um conjunto de pacotes que constituem uma lógica, validação sistemática, validação em profundidade por amostragem. Uma visão sintética do conjunto das fases pode apresentar-se da seguinte maneira:



PAIMAS, todo o processo de transferência. Elaboração própria a partir de PIAF



COMPLEMENTO

Todas as tarefas e ações definidas na norma não são relevantes em todos os contextos, ou seja, o PAIMAS previu um mecanismo que permita especializar e adaptar a norma para um domínio ou uma instituição específica. Nessa especialização, poderão ser eliminadas as tarefas não relevantes, adicionadas outras tarefas, levando em consideração outras normas, ferramentas e vocabulário específico do campo em questão.

Capítulo 7. Norma UNE-ISO 14641:2015 (Norma Afnor NF Z 42-013)

O título desta norma é o seguinte: Especificações para o design e funcionamento de um sistema de informação para a preservação de informação digital.



Norma UNE ISO 14641-1:2015

Esta norma oferece um conjunto de especificações técnicas e políticas organizacionais para implementar a captura, o armazenamento e o acesso a documentos eletrônicos. Foca nas **características do sistema informático** no qual o Arquivo Digital vai apoiar suas atividades. Situa-se no nível das especificações técnicas enquanto que o modelo OAIS situa-se no nível conceitual. Pretende:

- Otimizar a preservação de documentos eletrônicos a longo prazo, seu arquivo e integridade.
- Oferecer facilidades de busca de informação.
- Assegurar a acessibilidade e o uso de documentos eletrônicos.

A norma foca especificamente na integridade dos documentos arquivados e na capacidade do arquivo de comprovar essa integridade. Um ponto chave abordado e detalhado na Norma 14641 é a necessidade de dispor de uma descrição detalhada de todos os componentes do sistema de arquivo e uma rastreabilidade completa de todas as alterações dentro do Arquivo. Deve ser possível demonstrar a continuidade do registro de todos os eventos realizados.

Cada documento arquivado estará vinculado a um perfil de arquivo definido como um conjunto de regras aplicáveis aos documentos que compartilham os mesmos critérios de confidencialidade, duração de conservação, destruição e direitos de acesso.

Uma diminuição dos níveis de exigência permite adaptar os procedimentos ao nível de risco que o Arquivo aceita assumir. Uso diferenciado da marca de tempo, o cálculo de impressões digitais, criptografia e assinatura digital serão utilizados quando necessário.



DEFINIÇÃO: MARCA / ESTAMPA DE TEMPO

Objeto que vincula uma representação de dados a um momento específico (expresso em UTC, tempo universal coordenado), além de evidenciar que os dados existiam nesse momento.

Ato de atribuir uma hora precisa a uma transação, com base em uma fonte externa de tempo indiscutível e assinando os elementos de datação assim fornecidos (carimbos de tempo).

Em conclusão, pode-se dizer que a norma ISO 14641 se baseia nas experiências existentes e em outras normas gerais (ISO 14721, ISO 15489), aporta, no plano da aplicação, um grande número de princípios, normas e soluções técnicas que proporcionam ao sistema de arquivo todas as garantias de segurança e integridade que possam ser exigidas para os documentos digitais.



ATENÇÃO

Naturalmente, quanto maior for o nível de segurança requerido, maiores serão os custos de desenvolvimento e funcionamento do Arquivo. Portanto, é essencial saber situar-se no nível adequado em relação com essas exigências.

Capítulo 8. Certificação dos Arquivos Digitais

Dar a um arquivo digital um selo de qualidade e confiança, ou seja, garantir que se possui garantias efetivas suficientes quanto à capacidade desse Arquivo para preservar documentos digitais a longo prazo, garantir sua autenticidade, manter sua integridade e, ao mesmo tempo, mantê-los acessíveis à comunidade de usuários, são os objetivos da certificação.

8.1. O que é a certificação dos Arquivos Digitais?

A seguir, abordaremos uma questão de grande importância, mas ainda não estabilizada quanto às possíveis soluções.

Em que podemos nos basear para ter confiança em um Arquivo Digital?

Nada do que acontece no nosso computador é visível, exceto o resultado, que coincidirá ou não com as nossas expectativas.

Quem nunca viu seu PC travado sem nenhum motivo aparente? E nesses casos é muito difícil, senão impossível, conhecer a causa exata do que aconteceu. Quem nunca perdeu tempo porque seu software de processamento de texto havia congelado antes mesmo de salvar o que não queria perder? Quantos particulares, mas também, quantas empresas não perderam informações devido a um vírus ou à pirataria?

Agora, deseja-se poder contar com essa tecnologia para conservar informações valiosas, vitais, a fim de manter nosso conhecimento, preservar nossa memória? É evidente que é preciso encontrar uma maneira de criar essa confiança. É necessário que um Arquivo possa demonstrar que realmente faz tudo o que é necessário para conservar a informação da qual é responsável, que possui as competências para isso, que tem os meios necessários e que o faz corretamente. **Esta demonstração é o que se chama certificação dos Arquivos.**



COMPLEMENTO

Este processo de certificação dos Arquivos pode ser definido com o mesmo espírito dos processos de certificação ISO 9000 (sistema de gestão da qualidade). De fato, o processo de certificação ISO 9000 permite mostrar que a organização faz bem o que se supõe que deve fazer, que o faz de maneira satisfatória para os clientes e que tem ou iniciou um processo contínuo de melhoria da qualidade.

Foi adotada uma abordagem semelhante no âmbito específico da segurança dos sistemas de informação. Mais uma vez, surge a questão da confiança que se pode ter em um sistema de informação quando ele gerencia transações bancárias, contém segredos de fabricação, tem informações de defesa nacional. Essa abordagem deu origem à definição dos requisitos aplicáveis aos sistemas de gestão da segurança da informação (ISO/IEC 27001:2005) e ao estabelecimento de um código de boas práticas para a gestão da segurança da informação (ISO/IEC 27002:2005).

O objetivo é **voltar à prática da avaliação da aplicação e das organizações**.

A reflexão sobre esse tema para o Arquivo Digital começou em 2003 – após a publicação do Modelo OAIS – por um grupo de trabalho internacional criado por iniciativa da National Archives and Records Administration (NARA) e Research Library Group (RLG). Depois, continuou sob os auspícios de OCLC (Online Computer Library Center).

O objetivo era elaborar requisitos de certificação dos arquivos com vistas à normalização internacional. A ideia principal que foi mantida e retomada em todos os trabalhos posteriores consistiu em estabelecer uma lista de critérios que o Arquivo deve cumprir e classificar esses critérios por domínio. Estes critérios devem poder dar garantias de confiança, confiabilidade e sustentabilidade do Arquivo.

Este trabalho culminou com a publicação, em janeiro de 2007, do documento Trustworthy Repositories Audit & Certification: Criteria and Checklist conhecido como TRAC.

Por outro lado, uma colaboração entre o Digital Curation Center (DCC) na Grã-Bretanha e o programa europeu Digital Preservation Europe (DPE), levou ao desenvolvimento de uma metodologia de auditoria dos arquivos digitais baseada na avaliação de risco. Seu título é Digital Repositories Audit Method Based on Risks Assessment (DRAMBORA).

Planning Tool for Trusted Electronic Repositories (PLATTER) é uma ferramenta elaborada pela Digital Preservation Europe (DPE) que proporciona uma base para que um repositório digital planeje o desenvolvimento de seus objetivos e metas de desempenho ao longo de sua vida útil, além de contribuir para que o repositório estabeleça um estado de confiança entre as partes interessadas. PLATTER está projetado para complementar as iniciativas existentes de auditoria e certificação, como o kit de ferramentas DPE e DRAMBORA.

A norma ISO 16363:2017 sobre auditorias e certificação de repositórios digitais confiáveis fornece uma estrutura para a análise da qualidade, consistência e integridade das informações armazenadas nos repositórios, além disso, a norma tem o compromisso de garantir o acesso e a preservação, a longo prazo, das informações contidas, essa norma pode ser usada como base para a certificação.

Por fim, há a norma ISO 17068:2020 que, embora não seja especificamente voltada para certificação, define os requisitos que um repositório confiável de terceiros deve atender para oferecer suporte ao serviço de custódia da informação de forma íntegra, autêntica e confiável.

3 - <http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91>

4 - <http://www.dcc.ac.uk/>

5 - <http://www.digitalpreservationeurope.eu/>

6 - <http://www.digitalpreservationeurope.eu/>

Um processo de certificação não se resume ao exame da conformidade do Arquivo com um conjunto de critérios.

Esse processo levanta outras questões que devem ser perfeitamente esclarecidas.

- Quem pode realizar uma auditoria de certificação?
- Quais são as competências necessárias?
- Quais são as garantias de neutralidade e confidencialidade?
- Quem pode emitir essa certificação?
- Com que acreditação ou legitimidade?

Todas essas perguntas também devem receber respostas padronizadas e indiscutíveis.

8.2. Projetos sobre certificação e avaliação

Com o objetivo de tomar decisões sobre as necessidades ao iniciar ações em projetos de arquivos digitais ou mesmo reavaliar os requisitos quando já estão em execução, existem ferramentas de avaliação a nível internacional e até normas ISO para medir o nível de maturidade de uma organização especificamente em preservação digital. Dessa forma, as organizações podem identificar sua situação atual e estabelecer uma projeção de onde desejam estar.

Modelo de Avaliação Rápida (Rapid Assessment Model- RAM)¹

Ferramenta proposta pela DPC: Coalizão de Preservação Digital, organização sem fins lucrativos sediada no Reino Unido, que está disponível em espanhol desde junho de 2021, tanto o guia quanto o modelo de aplicação.

Seu objetivo consiste em monitorar o progresso à medida que se desenvolve e melhora a capacidade e a infraestrutura de preservação e com isso estabelecer metas de maturidade futura. Ou seja, fornece uma descrição geral de alto nível, mas também completa dos problemas que a preservação digital deve enfrentar.

Para a aplicação do modelo apresentam-se 11 seções divididas em capacidades organizacionais e técnicas ou de serviço:

Capacidades organizativas:

Viabilidade organizacional

A viabilidade organizacional refere-se a questões relacionadas à governança, à estrutura organizacional, à alocação de pessoal e aos recursos das atividades de preservação digital. Convencer as partes interessadas pode ser um dos desafios mais importantes que devemos enfrentar, mas é necessário para obter apoio, alocação de recursos e, não apenas isso, um enfoque sustentável de financiamento que permita a continuidade dos planos de preservação digital.

Política e estratégia

Corresponde à políticas, estratégias e procedimentos que regem o funcionamento e a gestão do arquivo digital, já se dispõe de uma política de preservação digital adotada oficialmente para estabelecer um programa de preservação digital sustentável e coerente. Uma política estabelecerá suas principais propósitos e objetivos e atuará como um ponto

¹ Disponível em diferentes idiomas: <https://www.dpconline.org/digipres/implement-digipres/dpc-ram>

de referência para a tomada de decisões futuras, assegurando a coerência e o foco. O desenvolvimento de sua política, assim como as estratégias e procedimentos subsequentes de implementação, pode ser auxiliado pelo uso de técnicas como a gestão de riscos para avaliar suas prioridades, além de fazer referência aos padrões estabelecidos que descrevem as boas práticas.

Base legal

As boas práticas na preservação digital exigem a consideração de uma série de questões legais. Isso pode incluir a gestão de direitos contratuais, de licenças e outros direitos legais, bem como as responsabilidades relacionadas à aquisição, preservação e acesso a conteúdo digital (por exemplo, licenças, direitos autorais, termos e condições de uso, regulamentação de proteção de dados).

Capacidade de TI

Embora a tecnologia não seja o único fator importante na preservação digital, ela continua sendo uma peça fundamental do quebra-cabeça, e deve-se dar a devida atenção às capacidades de tecnologia da informação da sua organização para apoiar as atividades de preservação digital. Em primeiro lugar, você deve considerar se tem acesso a computadores com a potência de processamento necessária para realizar as tarefas de preservação digital.

Será necessário estabelecer requisitos para as necessidades da sua infraestrutura de TI, seja armazenamento, ferramentas, hardware ou até mesmo um sistema completo de repositório.

Melhora contínua

Processos para a avaliação das capacidades atuais de preservação digital, definição de objetivos e acompanhamento dos avanços.

Portanto, é útil contar com processos para a avaliação das capacidades atuais de preservação digital, o estabelecimento de objetivos e o acompanhamento do progresso. Um tipo de ferramenta que pode ser utilizado para isso são os modelos de maturidade, e existem vários para a preservação digital, portanto, você deve conseguir encontrar um que funcione para você.

Comunidade

Colaboração e contribuição à comunidade da preservação digital em geral.

Considerando que a preservação digital é uma disciplina relativamente jovem, os profissionais têm descoberto repetidamente que o aprendizado e o apoio entre pares, por meio da troca de experiências e recursos, são produtivos e úteis.

Capacidades técnicas/ operacionais:

Aquisição, transferência e ingestão

Processos para a aquisição ou transferência de conteúdo e ingestão ao arquivo digital.

Em primeiro lugar, estão os processos que abrangem a aquisição e/ou transferência de conteúdo e sua incorporação a um arquivo digital. Isso abrange tudo, desde a negociação com os depositantes e/ou o contato com os criadores de conteúdo digital, passando pela transferência do conteúdo digital, a ingestão, a avaliação e o processamento para sua preservação. Como parte da aquisição, você deverá considerar que orientações deve fornecer aos depositantes/criadores de conteúdo, quais acordos serão necessários e quais metadados e documentações precisarão acompanhar o material.

o conteúdo digital para permitir que você processe e comprehenda o conteúdo digital.

Para a transferência, você deverá decidir quais métodos utilizará para receber os dados. Eles serão entregues em um meio físico? Ou serão transferidos por meio de um protocolo como FTP (Protocolo de Transferência de Arquivos)? Ou talvez você utilize um serviço baseado na nuvem, como o Dropbox? Você também deverá considerar uma abordagem de verificação de vírus para garantir que o conteúdo digital recebido não infecte o sistema da sua organização com nada indesejado. Manter a autenticidade e verificar que não tenham ocorrido alterações ou erros no processo de transferência também será importante, e isso pode ser gerenciado por meio de um processo chamado verificação de integridade. Se sua organização possui um sistema de repositório, haverá processos incorporados para administrar a ingestão, mas também é possível criar seu próprio fluxo de trabalho utilizando ferramentas menores que estão disponíveis gratuitamente ou a baixo custo.

Preservação do fluxo de bits

Isso abrange os processos para garantir as boas práticas no armazenamento e na integridade do conteúdo digital a ser preservado. Suas necessidades de armazenamento dependerão das políticas e práticas existentes, dos recursos financeiros disponíveis e das decisões de preservação que você tomou sobre questões como os formatos e a quantidade de cópias que devem ser mantidas. Deverá considerar que tipos de armazenamento utilizar: servidor? Fita? Nuvem? E é sempre recomendável não depender de um único tipo de armazenamento nem de um único fornecedor/fabricante.

Além do armazenamento, você deve considerar a segurança da informação e quem (dentro da sua organização) terá permissão para acessar o conteúdo digital de suas coleções. Você deve tomar medidas para garantir que seus arquivos digitais não possam ser alterados ou excluídos accidentalmente ou de forma maliciosa, por isso existem vantagens em manter um controle rigoroso de acesso. Os controles de integridade gerenciados e regulares permitirão que você administre qualquer erro que possa ocorrer no seu conteúdo digital, além de fornecer um método para garantir a autenticidade. O curso de iniciante a saber fazer foca principalmente na preservação do fluxo de bits.

Preservação de conteúdo

A preservação do conteúdo vai além de manter apenas o fluxo de bits do conteúdo digital e abrange os processos para preservar o significado ou a funcionalidade, garantindo sua acessibilidade e usabilidade contínuas ao longo do tempo. Os aspectos que podem ser considerados ao planejar a preservação a nível de conteúdo incluem os formatos de arquivo do conteúdo digital, o software e hardware necessários para acessá-lo, além da compreensão da importância da informação contida, do impacto de como ela é exibida e funciona na interpretação do usuário do conteúdo.

Dois conceitos-chave de preservação digital com os quais você deve se familiarizar ao preservar conteúdo são Technology Watch y Preservation Planning. A vigilância tecnológica refere-se ao processo de monitoramento das mudanças na tecnologia para identificar problemas que desencadeiem a necessidade de ações de preservação. Um exemplo pode ser uma nova versão de software que já não oferece compatibilidade com a funcionalidade das versões anteriores.

O planejamento da preservação ajuda você a decidir como irá lidar com quaisquer problemas identificados como resultado da vigilância tecnológica e a considerar o que fará para mitigar os riscos. Existem muitos enfoques diferentes para as preocupações tecnológicas na preservação de conteúdo, mas os dois principais que você ouvirá são a migração e a emulação. A migração faz com que o conteúdo digital seja transferido através de diferentes

versões ao longo do tempo, talvez por meio de novas versões de um formato de arquivo ou convertido para um formato mais estável e acessível. A emulação é o uso de software para reciar o entorno original no qual o conteúdo digital foi criado.

Um tema importante para a preservação do conteúdo é o asseguramento da qualidade. Isso deve ser incorporado em muitos aspectos da preservação digital, mas é particularmente importante para a preservação do conteúdo onde podem ser feitas alterações no conteúdo digital, nos metadados e na documentação que o acompanha. Todos os planos de conservação devem ser bem testados antes de sua implementação e devem incluir controles de qualidade quando forem ativados.

Gestão de metadados

Os metadados já foram mencionados várias vezes neste módulo, e os processos para criar e manter metadados suficientes para apoiar a preservação, a gestão e o uso do conteúdo digital preservado são outra seção importante do DPC RAM. A preservação digital requer a criação e manutenção de uma ampla gama de metadados que podem ser agrupados nos seguintes cabeçalhos:

Administrativo

Técnico

Descriutivo

Preservação

Estrutural

Com todos esses problemas a serem abordados, é fácil se perder em esquemas de metadados complicados, mas sempre é melhor tentar desenvolver uma abordagem de metadados que seja apropriada para seu conteúdo, o mais simples possível, porém tão completa quanto necessário. Não existe um padrão de metadados que abranja tudo para a preservação digital, mas deve-se buscar adotar elementos de diferentes padrões para atender às suas necessidades dentro dos cabeçalhos mencionados anteriormente. O padrão PREMIS cobre elementos de metadados administrativos, técnicos e de conservação, o METS fornece orientação sobre metadados estruturais e, para a descrição, é possível alinhar-se com os padrões que você já utiliza (por exemplo, ISAD (G), MARC ou Dublin Core). Você também pode considerar adotar uma abordagem de identificadores persistentes para o conteúdo digital que está preservando.

Ao começar, existem dois métodos principais para capturar alguns dos metadados necessários que você pode precisar manter. O primeiro é um manifesto de arquivo verificável, que captura informações básicas a nível de arquivo, permitindo que você saiba o que tem e onde está. A captura dessas informações pode ser automatizada utilizando ferramentas de caracterização. Em um nível mais amplo, também é uma boa prática manter um Registro de Ativos Digitais, que captura informações sobre temas como procedência, riscos principais, valor, direitos e retenção. A criação de um manifesto de arquivo verificável e de um registro de ativos digitais é abordada no curso “Documentação de coleções digitais”.

Busca e acesso

A seção final do DPC RAM aborda os processos necessários para permitir a descoberta do conteúdo digital e fornecer acesso aos usuários. Não existe uma forma "correta" de fornecer acesso, e os métodos que você escolher dependerão dos recursos disponíveis, dos tipos de conteúdo digital nas coleções, das licenças que você possui para reutilização e das necessidades dos seus usuários. É aqui que deve ser utilizada a análise do usuário para garantir que estamos preservando o conteúdo digital correto em um formato que permita o acesso e o uso. Também será necessário pensar em capturar e manter os metadados e a documentação necessários para permitir que os usuários compreendam e utilizem o conteúdo.

O acesso pode ser fornecido de muitas formas, desde copiar o material para mídias e repassá-lo aos usuários, até disponibilizar o acesso em uma estação de trabalho dedicada em sua organização, ou ainda por meio de um portal web de acesso aberto que ofereça funcionalidades para buscar, recuperar e utilizar conteúdo digital. Compreender as necessidades dos usuários é importante na hora de decidir como proporcionar acesso e é uma boa prática fazer com que os usuários testem sistemas potenciais e forneçam comentários. A capacidade de descoberta também é importante e afetará as decisões que você tomar sobre os processos de catalogação incluídos em 'Metadados'. Por fim, é importante considerar cuidadosamente como serão geridos os direitos de propriedade intelectual e a proteção de dados ao fornecer acesso, assegurando que isso esteja em consonância com as questões tratadas na "Base jurídica".

NDSA Levels:

Níveis de Preservação Digital propostos pela Aliança Nacional de Gestão Digital (National Digital Stewardship Alliance), organização hospedada na Federação de Bibliotecas Digitais dos Estados Unidos, é um recurso para ajudar os profissionais da preservação digital a construir ou avaliar seu programa de preservação digital. Criada originalmente em 2013, a versão 2.0 foi lançada em 2019 junto com documentação e recursos de apoio adicionais.

Conta-se com uma matriz para o processo avaliativo da versão de 2019 traduzida para o espanhol:

NDSA%		Niveles de Preservación Digital			
Área Funcional	Nivel				
	Nivel 1 - (Conocer su contenido)	Nivel 2 - (Proteger su contenido)	Nivel 3 - (Controlar su contenido)	Nivel 4 - (Mantener su contenido)	
Almacenamiento	Tener dos copias completas en ubicaciones separadas Documentar todos los medios de almacenamiento donde este almacenado el contenido Poner el contenido en soportes de almacenamiento estables	Tener tres copias completas con al menos una copia en una ubicación geográfica distinta Documentar el almacenamiento y medios de almacenamiento, indicando los recursos y las dependencias que estos requieren para funcionar	Tener al menos una copia en una ubicación geográfica con amenaza de desastre diferente a las otras copias Tener al menos una copia en un medio de almacenamiento de diferente tipo Rastrear la obsolescencia del almacenamiento y los medios	Tener al menos tres copias en ubicaciones geográficas distintas, cada una con una amenaza de desastre diferente Maximizar la diversificación del almacenamiento para evitar puntos únicos de falla Tener un plan y realizar acciones para abordar la obsolescencia del hardware, software y medios de almacenamiento	
Integridad	Verificar que la información de integridad se ha proporcionado con el contenido Generar información de integridad si esta no ha sido proporcionada con el contenido Se verifica virus en todo el contenido; se aísla el contenido en cuarentena según sea necesario	Verificar la información de integridad al mover o copiar contenido Usar bloqueadores de escritura cuando se trabaja con medios originales Hacer una copia de seguridad de la información de integridad y almacenar una copia en una ubicación separada del contenido	Verificar la información de integridad del contenido en intervalos fijos Documentar los procesos y resultados de verificación de información de integridad Realizar una auditoría de la información de integridad bajo demanda	Verificar la información de integridad en respuesta a eventos o actividades específicas Reemplazar o reparar el contenido dañado según sea necesario	

Control	Se determinan los agentes humanos y de software que deben estar autorizados para leer, escribir, mover y eliminar contenido	Documentar a los agentes humanos y de software autorizados para leer, escribir, mover y eliminar contenido y aplicar estos	Mantener los registros (logs) y se identifican a los agentes humanos y de software que realizaron acciones sobre el contenido.	Se realizan revisiones periódicas de acciones / registros (logs) de acceso
Metadatos	Crear un inventario de contenido, documentando también la ubicación de almacenamiento actual de estos Hacer una copia de respaldo del inventario y se almacena al menos una copia por separado	Almacenar suficientes metadatos para saber cuál es el contenido (esto podría incluir alguna combinación de aspectos administrativos, técnicos, descriptivos, de preservación y estructurales)	Determinar qué estándares de metadatos aplicar Encuentra y completa los vacíos en sus metadatos para cumplir con esos estándares	Registrar las acciones de preservación asociadas con el contenido y cuándo ocurren esas acciones Implementa los estándares de metadatos elegidos
Contenido	Documentar los formatos de archivo y otras características de contenido esenciales, incluido cómo y cuándo fueron identificados	Verificar los formatos de archivo y otras características de contenido esenciales Establecer relaciones con los creadores de contenido para fomentar la elección sostenible de archivos	Monitorear la obsolescencia y los cambios en las tecnologías de las que depende el contenido	Realizar migraciones, normalizaciones, emulación y actividades similares que garanticen el acceso al contenido

Levels of Digital Preservation v2.0 - 26 Septiembre 2019.

Copyright © 2019 by National Digital Stewardship Alliance. This work is licensed under a Creative Commons Attribution 3.0 Unported License. | DOI 10.17605/OSF.IO/QGZ98

Leija, David; Térmens, Miquel. (2019). Tradução de Níveis de Preservação Digital NDSA 2019

CoreTrustSeal:

Conforme estabelece o site desta organização: trata-se de uma entidade jurídica de direito holandês, internacional, comunitária, não governamental e sem fins lucrativos, que promove infraestruturas de dados sustentáveis e confiáveis. Que oferece a qualquer repositório de dados interessado uma certificação de nível básico baseada nos requisitos essenciais de repositórios de dados confiáveis.

A certificação CoreTrustSeal é concebida como o primeiro passo em um quadro global para a certificação de repositórios, que inclui a certificação de nível estendido (nestor-Seal DIN 31644) e a certificação de nível formal (ISO 16363). Em última instância, o CoreTrustSeal também se empenhará em fornecer uma certificação de nível básico para outras entidades de pesquisa, como serviços de dados e software.

8.3. Outras abordagens previstas

A hipótese do CCSDS em matéria de certificação refere-se a um Arquivo que suporta o conjunto de responsabilidades definidas no modelo OAIS. Podem existir contextos mais simples ou limitados, por exemplo:

- uma parte das atividades de arquivo é terceirizada e o terceiro arquivista deseja ser certificado no perímetro de responsabilidade que lhe corresponde;
- Os documentos a arquivar são documentos recorrentes e as principais questões a serem resolvidas se limitam à preservação dos bits e ao valor probatório dos documentos.

Capítulo 9. Conclusões

O ambiente normativo se enriqueceu consideravelmente ao longo destes anos e constitui uma contribuição insubstituível para o arquivo digital; no entanto, o quadro normativo continua difícil de aplicar na prática devido à diversidade de pontos de vista e perímetros cobertos, mas também devido às recuperações parciais existentes entre determinadas normas.

Essa situação complexa tem como origem essencial a multiplicidade dos setores de atividade que precisam enfrentar necessidades de arquivo digital. Cada um desses setores se dirige naturalmente aos organismos de normalização e, dentro desses organismos, aos comitês técnicos dos quais normalmente dependem, enquanto que uma abordagem transversal provavelmente beneficiaria a todos.

Bibliografia

Asociación Española de Normalización y Certificación UNE-ISO 16363. (2017). Sistemas de transferencia de información y datos espaciales. Auditoría y certificación de repositorios digitales de confianza. Recuperado de Bases de Datos del SIBDI: AENORMÁS.

Asociación Española de Normalización y Certificación UNE-ISO 14721. (2015). Sistemas de transferencia de datos e información espaciales Sistema abierto de información de archivo (OAIS) Modelo de referencia. Recuperado de Bases de Datos del SIBDI: AENORMÁS.

Asociación Española de Normalización y Certificación UNE-ISO 14641. (2012). Especificaciones para el diseño y funcionamiento de un sistema de información para la preservación de información digital. Parte 1. Recuperado de Bases de Datos del SIBDI: AENORMÁS.

Asociación Española de Normalización y Certificación UNE-ISO/TR 18492 IN. (2008). Conservación a largo plazo de la información basada en documentos. Recuperado de Bases de Datos del SIBDI: AENORMÁS.

BANAT-BERGER F., HUC C., DUPLOUY L., *L'Archivage numérique à long terme, les débuts de la maturité?* (Primera obra de síntesis sobre el archivo digital en lengua francesa) Paris, La Documentation française, 2009

BANAT-BERGER F., HUC C., Module 7 - Gestion et archivage des documents numériques. Portail International Archivistique Francophone. 2011. <https://www.piaf-archives.org/se-former/module-7-gestion-et-archivage-des-documents-numeriques> (Se identifica en el texto como PIAF)

Coalición de Preservación Digital (Digital Preservation Coalition). Modelo de evaluación rápida. Disponible en <https://www.dpconline.org/digipres/implement-digipres/dpc-ram>

Core Trust Seal. Disponible en: <https://www.coretrustseal.org/>

Leija, David; Términos, Miquel. (2019). Traducción de Niveles de Preservación Digital NDSA 2019: Traducción al Español de Versión 2.0. APREDIG - Asociación Iberoamericana de Preservación Digital. V2.0 en Español: <https://osf.io/egik8> | <http://www.apredig.org/npndsa2019/>



IBERARCHIVOS
IBERARQUIVOS



ARCHIVO NACIONAL
COSTA RICA



UNIVERSIDAD DE
COSTA RICA