

SEÇÃO 4

NORMAS E PADRÕES PARA O ARQUIVO DIGITAL

MÓDULO 7

Gestão e preservação de documentos digitais

SEÇÃO 4

Normas e padrões para o arquivo digital

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>.



ARCHIVO NACIONAL
COSTA RICA



UNIVERSIDAD DE
COSTA RICA

Conteúdo

Capítulo 1. Objetivo da seção	4
Capítulo 2. Diferença entre norma e padrão	5
2.1. Norma	5
2.2. Padrão	5
Capítulo 3. Organizaciones de normalización y estandarización	6
3.1. Organismos de normalização	6
3.2. Grupos de padronização	8
Capítulo 4. Do padrão à norma	9
Capítulo 5. Algumas considerações importantes sobre o uso das normas e padrões	10
Capítulo 6. Normas e padrões para o arquivo digital	11
6.1. Primeiro panorama geral	11
6.2. Principales âmbitos normativos que devem ser considerados	12
6.3. As normas e padrões como as bonecas russas (matrioskas)	18
Bibliografia	22

Capítulo 1. Objetivo da seção

Os serviços de informação passaram por uma série de mudanças vertiginosas em sua concepção, finalidade e forma de operar, portanto, as normas e os padrões desempenharão um papel fundamental para a formação do arquivo digital.

Partindo do princípio de que a formação do arquivo digital implica resolver uma série de situações, é fundamental apoiar-se nas normas e padrões específicos que facilitem a criação de soluções. Acima dessas normas e padrões específicos de cada área em particular, também existem normas gerais que se ocupam do arquivo digital como um todo.

O objetivo dessa seção é apresentar uma visão sintética do conjunto de normas e padrões com o objetivo de entender como e quando utilizá-los.

Entretanto, antes de aprofundar-se nessas normas e padrões, é necessário compreender o vocabulário pertinente, assim como ter conhecimento dos principais atores e saber identificar e entender como utilizar as normas e padrões.



GLOSSÁRIO

Padrões: Acordo baseado em um consenso mais limitado do que a norma, geralmente elaborado entre industriais em fóruns ou consórcios (fonte Afnor).

Padrão de fato: um padrão que não foi emitido por nenhum organismo oficial dedicado a isso e que foi imposto pelo uso e aceitação generalizados por uma comunidade. (Barnard, A y Voutssas, J, 2014, p. 108).

Padrão de jure: um padrão emitido por algum organismo dedicado e/ou autorizado para sua emissão. Podem ser nacionais (como NOM, ANSI), multinacionais (como CEN) ou internacionais (como ISO). (Barnard, A y Voutssas, J, 2014, p. 108).

Norma: Documento, elaborado por consenso e aprovado por uma organização reconhecida, que proporciona, para usos comuns e repetidos, normas, diretrizes ou características para as atividades ou seus resultados, garantir um nível ótimo de ordem em um contexto determinado (definição oficial da ISO/CEI).

Normalização Documental: Busca estabelecer, diante de problemas existentes, disposições orientadas a usos frequentes e repetidos em um contexto dado, com o objetivo de alcançar um nível de ordem e qualidade ótimo. (Arquivo Nacional da Costa Rica, 2020, p. 5)

Capítulo 2. Diferença entre norma e padrão

Esses dois termos são utilizados frequentemente como sinônimos, mas na verdade têm significados diferentes.

Essa confusão se deve ao fato de que, no idioma inglês, existe apenas uma palavra para designar esses dois conceitos, e é “STANDARD”.

2.1. Norma

É um documento voluntário que estabelece requisitos, especificações, diretrizes, procedimentos ou características que têm como objetivo garantir que os produtos, os serviços e os sistemas sejam seguros, coerentes e confiáveis. Oferece às organizações um conjunto de ferramentas com o potencial de ajudá-las a ter um melhor desempenho e ser mais competitivas.

As normas são desenvolvidas por meio do consenso das partes interessadas (grupos de especialistas da indústria, consumidores, organizações de pesquisa, governo, organizações setoriais, entre outros).

2.2. Padrão

Os padrões são definidos por grupos que não têm mandatos governamentais oficiais. Ou seja, um padrão é um acordo que, assim como uma norma, possui diretrizes, regras e definições, mas que não foi legitimado por nenhum organismo de padronização.

Estes grupos poderão ser:

- Industrial ou comercial: por exemplo, PostScript ou PDF são padrões de fato que são definidos pela empresa Adobe,
- Colegiados como:
 - W3C (World Wide Web Consortium)²
 - o consórcio Unicode³
 - o consórcio OASIS (Organization for the Advancement of Structured Information Standards)⁴

Um padrão pode ser ou não uma norma.

Em inglês são utilizados os seguintes termos para referir-se aos padrões e as normas, a saber:

“*De jure*”: refere-se a normas que foram aprovadas por um organismo de padronização.

“*De facto*”: refere-se a normas que não foram revisadas ou legitimadas por um organismo de padronização.

1 https://www.inteco.org/page/inteco.standards_development

2 <http://www.w3.org/>

3 <http://unicode.org/>

4 <https://www.oasis-open.org/homepage/>



ATENÇÃO

Para efeitos deste curso, será utilizado o termo 'norma' para designar normas e padrões, a fim de simplificar o texto.

Capítulo 3. Organizações de normalização e padronização

3.1. Organismos de normalização

Os organismos nacionais de normalização estão sob o mandato dos governos.

Estes organismos nacionais contribuirão com sua contribuição e terão direito a voto nos organismos internacionais.



Esquema de princípio do processo de normalização internacional. Elaboração própria a partir de PIAF



COMPLEMENTO

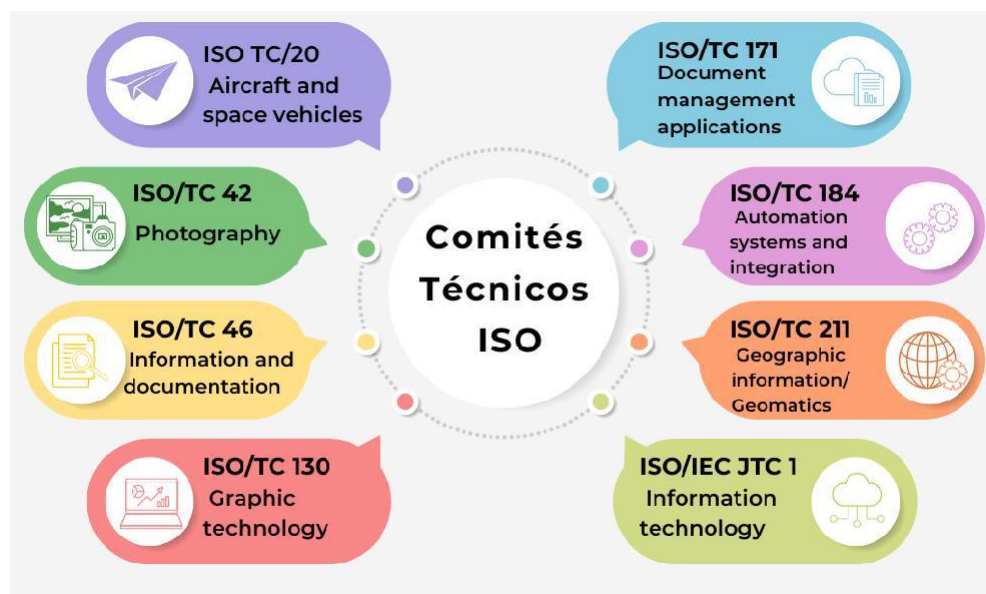
Os organismos nacionais **poderão delegar algumas das suas responsabilidades em associações profissionais.**

- Na França, por exemplo, a BNAE, Escritório Nacional de Aeronáutica e do Espaço, recebeu uma delegação da AFNOR para o desenvolvimento e a aprovação de normas na área da aeronáutica e do espaço.
- Nos Estados Unidos, o Instituto de Engenheiros Elétricos e Eletrônicos (IEEE), com 360.000 membros de 175 países, foi acreditado pela ANSI para elaborar normas em sua área de competência. Além disso, o IEEE produziu o padrão para a representação de números (ANSI/IEEE Std 854-1987 IEEE Standard for Radix-independent Floating-Point Arithmetic), que atualmente é utilizado nos computadores de todo o mundo.

Existem também **organismos de normalização a nível continental.** É o caso do Comitê Europeu de Normalização (CEN).

Esta organização não é adequada para o surgimento de padrões transversais e multifacetados como o arquivo digital, pois cada profissão tende a encontrar suas próprias soluções, mesmo que o problema a ser resolvido seja o mesmo para outras profissões.

A ISO tem mais de 200 comitês técnicos (TC), que estão subdivididos em vários subcomitês. Um número significativo desses comitês está interessado em atividades que produzirão documentos digitais, imagens, planos, descrições de objetos em duas ou três dimensões e que deverão ser conservados a longo prazo.



Alguns dos comitês técnicos da ISO produzem normas que se aplicarão ao arquivo digital. Elaboração própria a partir de PIAF

3.2. Grupos de padronização

Os grupos que produzem padrões têm estruturas muito variadas, a saber:

- Grupos colegiados (Unicode, W3C, OASIS, ECMA...) que reúnem, a nível internacional, colaboradores provenientes do setor industrial, das universidades, instituições de pesquisa, entre outros.
- Grupos “profissionais”, como o caso do Comitê Consultivo sobre Sistemas de Dados Espaciais (CCSDS), que originou a norma OAIS (Open Archival Information System), norma fundamental para o arquivo digital e que será estudada de forma mais detalhada na seção 5.

Outros grupos “profissionais” têm uma vocação de organização muito ampla, de coordenação entre seus membros, e nos quais a normalização constitui apenas uma parte, maior ou menor, de sua atividade.

Exemplo disso é o Conselho Internacional de Arquivos (CIA) que publicou as seguintes normas de descrição arquivística:

- ISAD(G): Norma Geral de Descrição Arquivística
- ISAAR(CPF): Norma Internacional sobre os Registros de Autoridade de Arquivos relativos a Instituições, Pessoas e Famílias.
- ISDF: Norma internacional para a descrição de funções.
- ISDIAH: Norma internacional para descrever instituições que guardam fundos de arquivo.

Outro exemplo é a Sociedade de Arquivistas Americanos (SAA), que mantém o “*Encoded Archival Description*”⁶ (EAD), o qual é um padrão de codificação de instrumentos arquivísticos de pesquisa baseado na linguagem de programação XML.

Da mesma forma, a Biblioteca do Congresso dos Estados Unidos publicou e mantém o padrão METS (Meta Data Encoding Transmission Standard)⁷ para a codificação de dados e metadados.

- Empresas privadas: um exemplo é o caso do Adobe que dá a origem do formato PDF.
- Instituições ou empresas também podem produzir normas internas para homogeneizar certas práticas, desenvolver e melhorar a qualidade, o rendimento, a satisfação dos clientes, entre outros. Esses padrões internos podem cobrir os mais diversos aspectos, desde modelos para a criação de documentos até os procedimentos que devem ser seguidos em caso de algum incidente de produção.

Mas também existem sobreposições:



EXEMPLO

O CCSDS, é um grupo de normalização que também desempenha uma função no subcomitê técnico dentro da ISO: Comitê Técnico 20 (Aeronaves e veículos espaciais), Subcomitê¹³ (sistemas de transferência de informação e de dados espaciais).

6 <http://www.loc.gov/ead>

7 <http://www.loc.gov/standards/mets>

Capítulo 4. Do padrão à norma

É bastante comum que um padrão, quando é reconhecido e utilizado internacionalmente, se torne uma norma.



COMPLEMENTO

HTML, Hypertext Markup Language para as páginas da web

- Padronizado pelo consórcio W3C (versão 4.0 em 1997)
- Normalizado pela ISO em 2000 (ISO 15445:2000)

PNG, Portable network Graphics, para as imagens

- Padronizado pelo W3C em 1999
- Normalizado pela ISO em 2004 (ISO 15948:2004)

PDF - PDF/A

- PDF - formato proprietário (ADOBE)
- PDF/A se tornou a norma ISO 19005-1 em 2005
- PDF 1.7 se tornou a norma ISO 32000-1 em 2008

ODF, Open Document Format, para documentos ofimáticos

- Padronizado por OASIS (maio de 2005)
- Normalizado pela ISO em 2006 (ISO 26300) em maio de 2006. A última versão é de 2015.

OOXML, Office Open XML para documentos ofimáticos

- Padronizado por ECMA em 2006. Atualmente está aceito como o padrão ISO/IEC 22275:2018.
- Normalizado pela ISO em 2008 (ISO 29500). A última versão é de 2015.

O formato de metadados Dublin Core

- Definido em 1995 por um grupo de especialistas e usuários de metadados de Internet,
- Normalizado pela ISO em 2003 (ISO 15836). A última versão é de 2011.

Por último, cabe destacar que, dentro dos grupos de normalização, o W3C reveste uma importância muito particular:

- É um grupo promotor dos padrões abertos e livres de direitos,
- Todos os padrões do W3C estão disponíveis de forma gratuita em <http://www.w3.org/>
- O W3C centra a maior parte de sua atividade na estruturação da informação em formato digital,
- Está desde o início na padronização de HTML, XML (e de todas as linguagens especializadas derivadas (por ex. RDF, MathML, XHTML...), PNG para as imagens rasterizadas, SVG (Scalable vector graphics) para os gráficos vetoriais

Capítulo 5. Algumas considerações importantes sobre o uso das normas e padrões

A norma pode ser um instrumento de poder:

“As normas, ao ter uma capacidade de atuar sobre o comportamento de outros atores e de ajustá-los a um modelo definido, constituem um instrumento de poder. Dentro dos organismos que as elaboram, a escolha das normas (sujeito e conteúdo) constitui, portanto, uma questão de influência e poder. Os membros das organizações de normalização tentam impor sua visão de mundo e reforçar sua própria estratégia de conquista e proteção dos mercados.” “A Igualdade de Armas” Informe ao Primeiro-Ministro por Bernard Carayon, 2006.

O fato de que exista uma norma não significa que se utilize:

Se a norma for utilizada muito pouco ou não for utilizada de forma alguma, não haverá ferramentas de software, ou haverá apenas algumas poucas, que facilitem seu uso, pois sua qualidade não foi testada.



EXEMPLO

Um exemplo particularmente chamativo é o da norma ODA (Open Document Architecture). ODA foi definido como formato de intercâmbio de documentos. É um padrão ECMA publicado em 1985, que foi promovida pela Comissão Europeia e tornou-se a norma ISO 8613 em 1989.

Essa norma está homologada pelo AFNOR como uma norma francesa que está vigente, porém nunca foi utilizada.

Critérios que devem ser considerados na escolha das normas e padrões.

Os principais critérios são:

- A adaptação às necessidades.
- A complexidade e o custo da sua aplicação.
- Não ignorar os aspectos político-industriais.

Também influencia na decisão:

- A possível obsolescência: embora as normas ofereçam maior estabilidade e durabilidade do que as tecnologias patenteadas, isso não significa que não se tornem obsoletas após um certo tempo.
- A dependência das tecnologias que estão a ponto de desaparecer.

**ATENÇÃO****Preço das normas**

Na maioria dos casos, o texto dos padrões é livre e gratuito.

Ao contrário, as normas são de alto custo monetário; seu preço é às vezes um obstáculo para sua difusão e utilização.

É bom saber que, em alguns casos, quando um padrão foi normalizado, o mesmo documento pode ser pago para o organismo de normalização e gratuito para o grupo de normalização que o tenha elaborado.

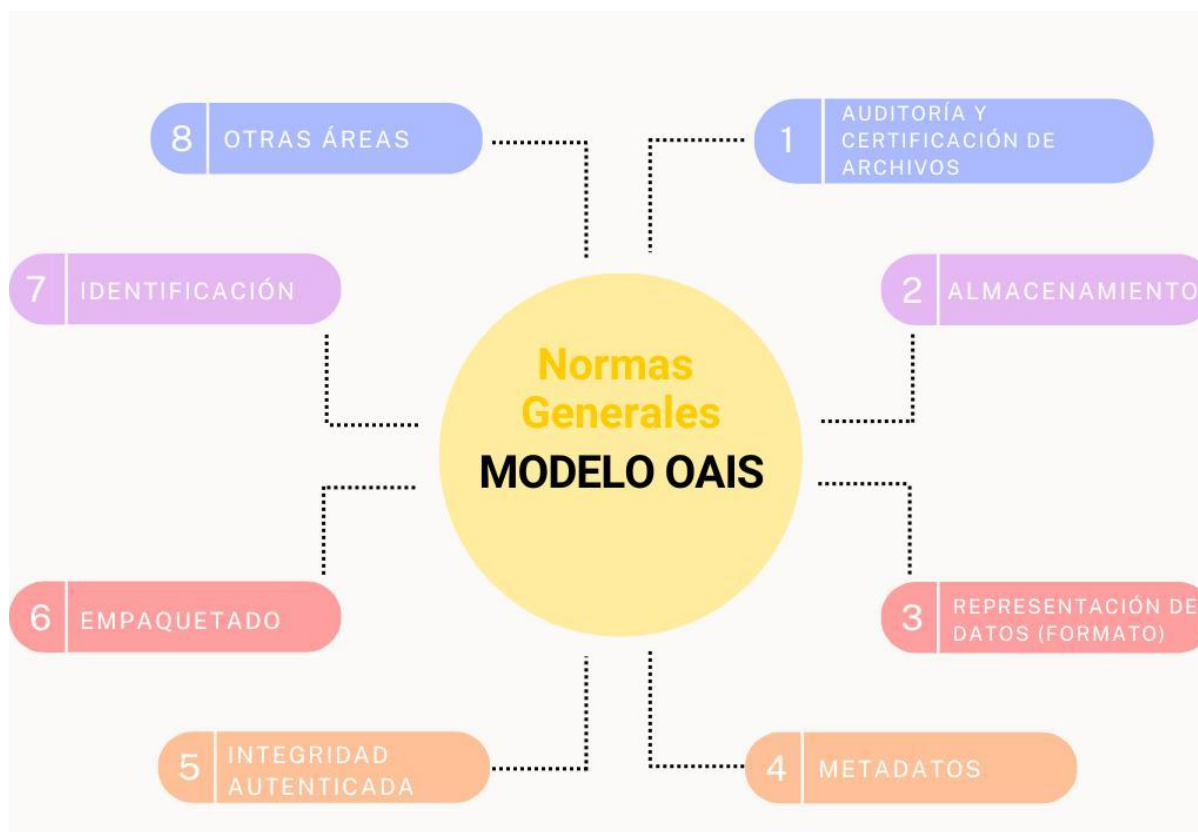
O organismo de normalização se limitará a trocar a capa do documento para colocar o seu nome e o seu logotipo. É o caso, em particular, das normas do CCSDS, entre elas o Modelo de Referência da OAIS, que podem ser consultadas livremente no site do CCSDS.

Capítulo 6. Normas e padrões para o arquivo digital

Diante das tecnologias que estão em constante inovação, surge como principal necessidade a compatibilidade técnica. Nesse sentido, a vinculação entre os atores do arquivo digital, as normas e os padrões proporcionam uma estabilidade na qual podemos nos apoiar. Essas bases consensuadas vão desempenhar um papel essencial na conformação do arquivo digital.

6.1. Primeiro panorama geral

Neste ponto, não se trata de elaborar um inventário complexo em relação às normas e padrões utilizáveis. Propomos uma classificação e identificação das principais áreas normativas. Essas principais áreas estão representadas na seguinte imagem.



Principais âmbitos normativos do arquivo digital. Elaboração própria a partir de PIAF

As normas gerais, especificamente, o modelo de referência OAIS, são as que serão utilizadas para abordar o arquivo digital em sua totalidade. Será estudado em profundidade na seção 5.

As demais esferas serão analisadas com mais detalhe nas seções 6 a 10, no entanto, é conveniente que nesta etapa do curso seja apresentado um primeiro esboço.

6.2. Principales âmbitos normativos que devem ser considerados

Esses âmbitos são numerosos, portanto, é difícil que uma mesma pessoa seja capaz de conhecer profundamente todos esses âmbitos. Entretanto, é indispensável que todo responsável por um arquivo digital tenha uma visão geral de todas essas áreas e possa contar com competências especializadas para cada uma delas quando necessário.

A seguir, são mencionadas uma série de normas que serão aprofundadas nas seguintes seções deste curso.

Serão citadas, neste resumo, uma série de documentos, que mais do que normas ou padrões, são boas práticas ou recomendações.

Armazenamento:

Trata-se da preservação de bits, para isso conta-se com:

- As orientações e recomendações sobre a estratégia de armazenamento e a organização do processo de armazenamento.
- Normas sobre os suportes para registrar a informação.

**EXEMPLO**

Norma ISO 9660:1988/Enm. 2:2020 "Tratamento da informação. Volume e estrutura de arquivos do CD-ROM para o intercâmbio de informações", essa norma será substituída pela norma ISO/IEC PRF 9660 "Volume e estrutura de arquivos do CD-ROM para o intercâmbio de informações" (o rascunho está em processo de aprovação - outubro de 2022)

Norma ISO/IEC 13421:1993 Information technology — Data Interchange on 12,7 mm, 48-track magnetic tape cartridges — DLT 1 format (norma revisada em 2020 e continua vigente).

Normas ISO/IEC 13962:1995 Information technology — Data interchange on 12,7 mm, 112-track magnetic tape cartridges — DLT 2 format (norma revisada em 2020 e continua vigente).

Essas normas sobre os suportes de registro da informação **se complementam de forma útil com as avaliações de alguns desses suportes.**

**EXEMPLO**

A Coalizão de Preservação Digital (DPC, na sigla em inglês) conta com uma série de ferramentas que fornecem orientação detalhada sobre como melhorar ou atualizar as capacidades de preservação digital dentro de uma organização, algumas delas de uso gratuito e outras de acesso exclusivo para seus membros

Além disso, os relatórios de vigilância tecnológica permitem monitorar e fornecer guias de referência detalhadas sobre conteúdo ou tipos de dados específicos, como: dados geoespaciais, registros de construção, audiovisuais, e-mails, etc.

Acesso ao site: <https://www.dpconline.org/>

Formatos de representação da informação:

Independentemente de se tratar de um texto, imagem, imagens 3D, gráficos vetoriais, documentos multimídia, de áudio, de vídeo, de dados técnicos, de dados científicos, **todas as formas de informação digital se referem a formatos.**

Existem formatos gerais e profissionais:

**COMPLEMENTO**

- Nos formatos gerais, encontram-se, por exemplo, o PDF, os principais formatos de escritório ODF e OOXML, a metalinguagem XML, o formato de imagem PNG.

O desenvolvimento do uso da metalinguagem XML constituiu uma evolução importante sobre o acesso, intercâmbio e arquivamento de dados. O PDF, ODF e OOXML atuais baseiam-se em grande medida nesta metalinguagem.

- Entre outros formatos, podemos citar FITS (Flexible Image Transport System) para a astronomia ou até mesmo as diferentes aplicações da linguagem STEP (Standard for the Exchange of Product Model Data ou Padrão para o Intercâmbio de Modelo de Dados de Produto) para a indústria.

As normas relativas aos dados espaciais também ocupam um lugar importante nos formatos comerciais.

Metadados

Em princípio, não se criou algo novo; antes da era digital, já haviam sido inventados e utilizados os metadados. Durante séculos, os catálogos das bibliotecas incluíram metadados específicos para cada livro. Entretanto, com a tecnologia digital surgiram novas necessidades, por exemplo:

- Dispor de metadados técnicos que garantam a capacidade futura de fornecer informações utilizáveis e acessíveis a partir de um conjunto de sequências de bits.
- Encontrar imediatamente os documentos que são solicitados, em meio a milhões de documentos que possivelmente estão armazenados em locais geográficos distantes.

De fato, a expansão da tecnologia digital deu lugar à aparição de numerosos formatos de metadados

- Dublin Core (ISO 15836), que tem um caractere geral.
- Norma ISO 19115 de metadados espaciais, que está orientada a dados específicos e que oferece mais de 400 seções descritivas.
- Norma ISO 23081 de metadados para a gestão de documentos que estabelece os elementos conceituais e de implementação nos processos de gestão de documentos.

A seção 9 deste curso está dedicada aos metadados.

Integridade e autenticidade da informação:

Trata-se de normas que nos permitirão aportar provas da integridade e autenticidade dos documentos. Estas normas abrangem o seguinte:

- A criptografia
- Algoritmos de cálculo de impressões digitais
- Protocolos de comunicação seguros
- Os processos de assinatura digital.

O Instituto Nacional de Padrões e Tecnologia (NIST), chamado entre 1901 e 1988 de Escritório Nacional de Normas, é uma agência da Administração de Tecnologia do Departamento de Comércio dos Estados Unidos, que também propôs desenvolver uma série de padrões, diretrizes, melhores práticas e recursos de cibersegurança para atender às necessidades da indústria americana, das agências federais e do público em geral. Até a data, tem desempenhado um papel crucial na promulgação de padrões de cibersegurança. O NIST emite pautas e padrões, como o «Marco de Cibersegurança» (Cybersecurity Framework), que é amplamente utilizado nos Estados Unidos e internacionalmente como um marco de referência para a gestão de riscos de cibersegurança.

Por sua parte, a ISO (Organização Internacional de Normalização) respalda e normaliza padrões relacionados à segurança da informação, incluindo algoritmos criptográficos e métodos de cálculo de impressão digital, trabalhando em estreita colaboração com organizações como o NIST para garantir a coerência e a interoperabilidade dos padrões em nível mundial.

Este tema será desenvolvido na seção 10, que abrange a integridade, a autenticidade e os aspectos jurídicos, entre outros aspectos.

Empacotamento

Os padrões de empacotamento permitem associar objetos digitais dentro de uma estrutura global, coerente e portátil. Os objetos digitais podem estar integrados por:

- Um ou vários arquivos.
- Metadados que caracterizam ditos objetos e arquivos.
- Vínculos entre os próprios arquivos ou entre os componentes do objeto e os metadados.

Esses padrões também permitem definir mecanismos aplicáveis aos objetos digitais, podendo exigir, por exemplo, que um objeto deva ser aberto utilizando tal aplicativo, ou que os diferentes arquivos do objeto devem ser organizados e tratados de tal forma para permitir a restituição de seu conteúdo.

A título de exemplo e de forma metafórica, quando enviamos um pacote por correio postal. De fato, um pacote postal permite transportar todo tipo de objetos, mas sempre estará acompanhado do endereço do destinatário e do remetente redigidos de forma normalizada; além disso, esse empacotamento deverá cumprir uma série de regras básicas.

Os padrões de empacotamento permitirão identificar, descrever e manipular os componentes elementais digitais de um objeto complexo. Os padrões METS (Metadata Encoding and Transmission Standard) da Biblioteca do

Congresso y XFDU (XML Formatted Data Unit) do CCSDS pertencem a esta área que será desenvolvido na parte 9 sobre metadados.

Identificação

Consiste em identificar um objeto digital de maneira única dentro de um arquivo ou em um âmbito mais amplo, dentro de um conjunto de arquivos. Essa necessidade não é nova e já foi resolvida, por exemplo, com o uso do ISBN (International Serial Bibliografia Number) para as obras publicadas ou do ISSN (International Standard Serial Number) para as revistas.

A questão da identificação de documentos e, em geral, de todo tipo de recursos digitais é abordada de forma diferenciada no âmbito digital. A Coalizão de Preservação Digital (DPC) descreve o persistent identifier (identificador persistente) como um conjunto de números, letras ou símbolos que têm a finalidade de continuar representando ou identificando um dado e quanto à sua condição de persistência, indica que se trata de uma referência duradoura a um recurso digital, composta pelo identificador único e por um serviço que localiza o recurso ao longo do tempo, mesmo quando sua localização muda.

Para o arquivo digital, o uso de identificadores únicos e duradouros considera-se uma boa prática, pois deve-se garantir a disponibilidade e o acesso aos recursos digitais por um longo prazo.

Como parte dos recursos disponíveis conta-se com os Padrões de identificação (URI, Uniform Resource Identifier) e nomes (URN, Uniform Resource Name) que foram definidos pelo consórcio W3C, e que atualmente podem ser interpretados e manipulados pela maioria dos navegadores web.

Atualmente, todo mundo sabe até que ponto as URL não são duráveis através do tempo. Quantas vezes já aconteceu que o endereço de um recurso na Internet não era mais válido?

Existe uma necessidade muito importante de contar com identificadores permanentes. Nos últimos anos, surgiram diferentes serviços de atribuição de identificadores permanentes como ARK «Archival Resource Key» ou os DOI «Digital Object Identifier». Estes serão examinados mais detalhadamente na parte 9 dedicada aos metadados.

A certificação dos arquivos digitais:

O documento eletrônico frequentemente costuma dar um pouco de medo. Sua sustentabilidade apresenta toda uma série de novos problemas. Até que ponto podemos confiar nos organismos encarregados de garantir sua preservação?

Em geral, quando uma empresa ou instituição deseja demonstrar que está fazendo bem o seu trabalho e que é capaz de fornecer produtos e/ou serviços que atendem aos requisitos dos clientes e às normas aplicáveis, busca estabelecer um sistema de controle de qualidade e iniciar um processo de certificação ISO 9001. Essa certificação será emitida por um período de tempo definido por auditores independentes, acreditados por uma organização reconhecida para emitir tais credenciações.

No âmbito do arquivo digital existem várias normas afins à certificação, a saber:

1. Norma UNE-ISO 14641-1:2015 “Arquivo eletrônico. Parte 1: Especificações para o design e funcionamento de um sistema de informação para a preservação de informações digitais.
2. ISO 16363 2012 Auditoria e certificação de repositórios digitais confiáveis.

3. ISO 16919:2014 Requisitos para organismos que fornecem auditoria e certificação de repositórios digitais confiáveis candidatos, que estabelece os requisitos para qualquer organização que certifique os auditores para ISO 16363; y
4. ISO 17021:2015 Requisitos para os organismos que fornecem auditoria e certificação de sistemas de gestão, proporciona um mecanismo para auditar os organismos de acreditação.

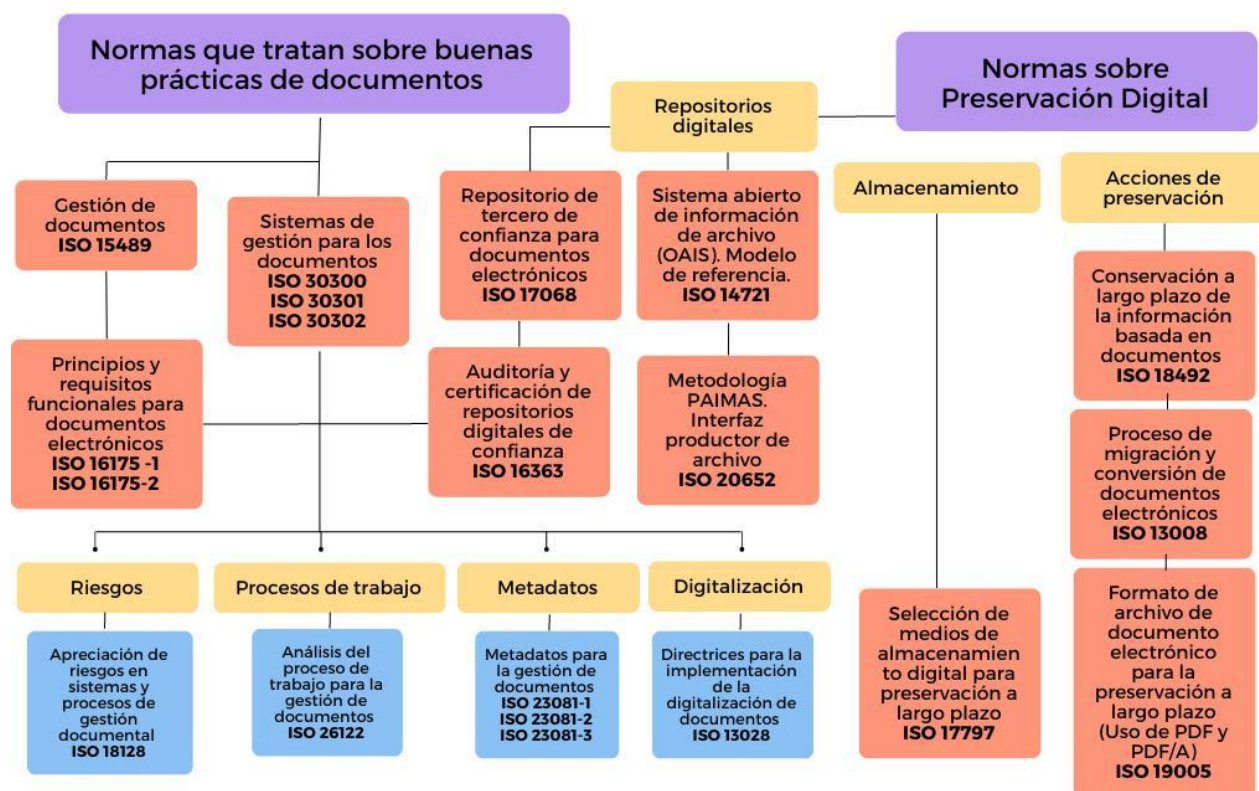
O que se refere à certificação dos arquivos será abordado na seção 5.

Outras esferas normativas: entre as outras áreas normativas que devem ser considerados no arquivo digital, destacam-se:

- Tudo o que se refere aos sistemas de gestão da segurança da informação (incluindo a confidencialidade, a proteção da informação sensível, os planos de continuidade de negócios)
- As questões relativas à externalização de serviços de arquivo,

Adicionalmente, apresenta-se um compêndio das principais normas ISO relacionadas à gestão de documentos e à preservação digital, fundamentais para a implementação de um arquivo digital.

NORMAS ISO



*Resumo de normas ISO de interesse para gestão e preservação de documentos eletrônicos
Adaptação gráfica a partir de conteúdo de curso ministrado pelo Arquivo Geral da Colômbia em 2021.*



COMPLEMENTO

Projeto InterPARES⁸

De 1994 a 1997, foi desenvolvido um projeto conhecido como 'Preservation of the Integrity of Electronic Records' ou 'Projeto UBC', cujo objetivo era “estabelecer as normas para a criação de documentos de arquivo eletrônicos confiáveis e manter sua autenticidade durante sua vida ativa e semiativa”, e que foi desenvolvido por pesquisadores de Arquivística com um enfoque da Diplomática.

A partir de 1998 o grupo de pesquisa foi denominado InterPARES e passou por **quatro etapas**: Na primeira etapa (1998-2002), apesar de a pesquisa ter continuado com um enfoque diplomático, durante esta fase integrou acadêmicos e profissionais de outras disciplinas, além da Arquivística: acadêmicos e profissionais de diversas áreas: Direito, História, Tecnologias de Informação e Comunicações, Engenharia e Química. Na segunda etapa (2003-2006), o grupo se tornou mais colaborativo e transdisciplinar, incluindo cientistas, especialistas do setor bancário e da indústria, artistas e representantes de governos de várias partes do mundo. Na terceira fase (2007-2012), o projeto InterPARES 3 reformulou seu objetivo, metas específicas e marco teórico, com a finalidade de tornar o conhecimento e a metodologia resultantes de suas etapas anteriores sobre a problemática da preservação de arquivos digitais autênticos a longo prazo mais práticos, além de compartilhá-los com organizações médias e pequenas, tanto de instituições públicas quanto privadas, responsáveis por arquivos digitais. Outro aspecto importante é que pela primeira vez incluem-se as necessidades sociais dos usuários dos arquivos digitais, garantindo sua autenticidade e confiabilidade a longo prazo.

6.3. As normas e padrões como as bonecas russas (matrioskas)

É preciso saber que uma série de normas se apresentam como matrioskas, ou seja, que se baseiam em outras normas, que também podem se basear em outras normas.



ATENÇÃO

A aplicação de uma norma de um determinado nível não implica que as normas de nível inferior sejam aplicadas corretamente.

Em muitos casos, a escolha de uma norma não é suficiente, portanto, é conveniente adicionar a essa escolha de norma condições de aplicações especiais obrigatórias



COMPLEMENTO

Vamos analisar o caso da norma ISO/IEC 26300-1:2015 “Open Document Format for Office Applications” a qual tem como predecessora a norma ISO 26300, Open Document Format:

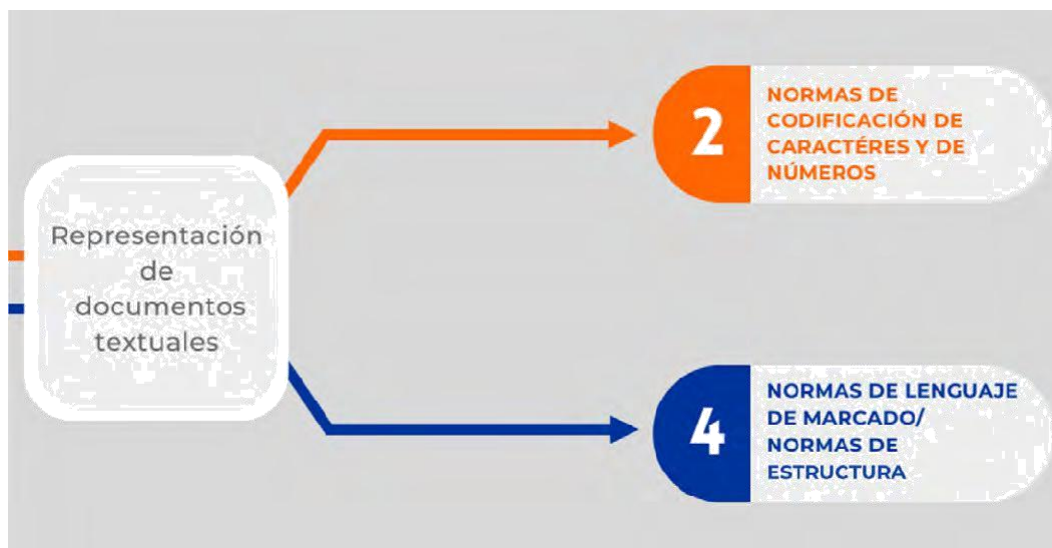
A norma ISO 26300 baseou-se em primeiro lugar em um padrão de estruturação de documento chamado Relax NG, padrão desenvolvido pelo consórcio OASIS e posteriormente normalizado pela ISO ((ISO/IEC 19757-2), Relax NG baseia-se na linguagem XML, padrão do W3C, XML utiliza por defeito a codificação de caracteres UTF-8. UTF-8 é uma forma de codificação diretamente derivada do conjunto de caracteres universal definido pela norma ISO 10646 (também conhecida como Unicode). Porém XML também pode basear-se em outros tipos de codificação, o conjunto de caracteres utilizado é especificado no início do documento XML.

Por sua vez, a norma ISO/IEC 26300-1:2015 “Open Document Format for Office Applications” (em vigor) define um esquema XML projetado para que os documentos possam ser transformados por meio de XSLT e processados com ferramentas baseadas em XML.

Vamos encontrar tais situações nas normas descritivas de formatos de dados, as normas de metadados, identificação, empacotamento e outros.

A aplicação de uma norma de um determinado nível não implica que as normas de nível inferior sejam aplicadas corretamente.

No que diz respeito à representação da informação de forma digital, nossa preocupação deve sempre abranger por completo a cadeia lógica que vai da sequência de bits até a restituição de um documento inteligível.



Normas e padrões no caso dos documentos textuais. Elaboração própria a partir de PIAF

Normas de codificação: Os caracteres numerados entre 0 e 127 são os do código ASCII. Os que têm um código entre 160 e 255 são os do código ISO 8859-1.

O padrão ISO 8859-1, também chamado Latin-1 ou Western Europe, é a primeira parte de um padrão mais completo chamado ISO 8859 (que inclui 16 partes) e que permite codificar todos os caracteres dos idiomas europeus.

Normas de estrutura

A linguagem de marcado mais utilizado na World Wide Web é HTML. Em HTML, as etiquetas pré-definidas são utilizados para especificar dentro de um arquivo de texto elementos como títulos, parágrafos, acrônimos, aspas, etc.

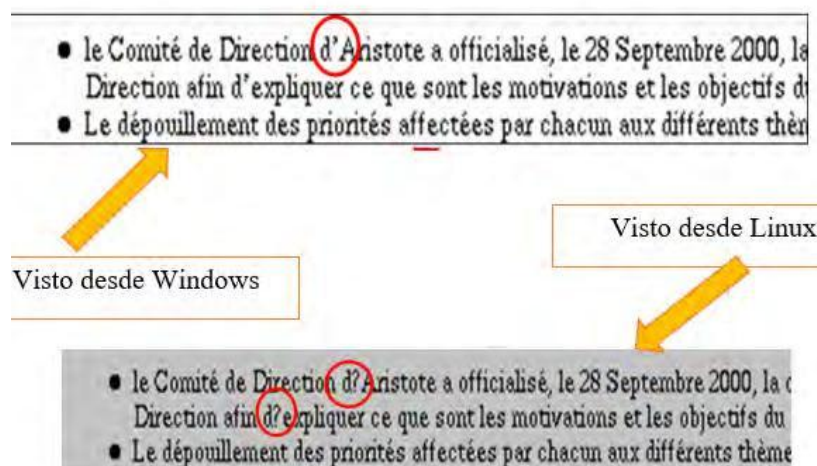
XSD (XML Schema Definition) é uma linguagem utilizada para descrever a estrutura e as restrições dos conteúdos dos documentos XML de uma forma muito precisa, além das normas sintáticas impostas pela própria linguagem. XML.

Representação de documentos textuais TXT, PDF/A, formato aberto, office open XML.

No passado tivemos muitos problemas com esse tema. Um exemplo muito simples é o da gravação, no Windows 2000, com Word 97, de um arquivo em formato HTML, com o objetivo de inserir este arquivo em um site.

Observação: Como ilustrado na imagem a seguir, após a transferência para o site, a visualização desse arquivo aparece corretamente quando examinada a partir do sistema operacional Windows, mas não é correta quando visualizada a partir de outro sistema operacional, como o Linux. Alguns caracteres especiais foram substituídos por sinais de interrogação ou de numeração, como é o caso dos acentos ortográficos.

Exemplo: Seção 3-1 por Seção 3-1



Destaque de anomalias na visualização de certos caracteres pelo navegador. Elaborado por PIAF

Explicação: Examinar o arquivo HTML com um editor de texto. É importante saber que, no início de cada arquivo HTML, um atributo chamado «charset» permite especificar o conjunto de caracteres utilizado no arquivo. Neste caso, podemos ver que o arquivo foi criado não com um conjunto de caracteres normalizado, mas com um conjunto de caracteres proprietário, ou seja, o conjunto de caracteres “Windows-1252”, próprio do sistema operacional Windows. Entretanto, este arquivo continua totalmente conforme com a norma HTML, que não impõe um conjunto de caracteres específico.

```
Fichier MS Word 97 - enregistré au format HTML sous Windows
<HTML>
<HEAD>
<META HTTP-EQUIV="Content-Type"
CONTENT="text/html; charset=Windows-1252">
<TITLE>Association Aristote – Groupe
Pérennisation des Informations
Numériques</TITLE>
```

Demonstração de um conjunto de caracteres específico. Elaborado por PIAF

Exemplo:

Como nos demais casos em que se cria um documento conforme uma gramática pública, é preciso declarar o tipo de documento.

Uma definição de tipo de documento ou DTD é uma descrição da estrutura e sintaxe de um documento XML ou SGML. Sua função básica é a descrição da estrutura de dados, para usar uma estrutura comum e manter a consistência entre todos os documentos que utilizem a mesma DTD.

Neste caso, a declaração seria esta:

```
<!DOCTYPE HTML PUBLIC "ISO/IEC 15445:2000//DTD HTML//FR">
```

Conclusão: Se quisermos arquivar um arquivo em formato HTML a longo prazo, devemos especificar:

- Que o arquivo deve estar estritamente em conformidade com a norma ISO 15445:2000, que define o formato HTML na versão 4.0 (uma versão 5.0 será provavelmente normalizada mais adiante),
- Mas também que o conjunto de caracteres utilizado deve ser um conjunto normalizado pela ISO.

Esse exemplo ajuda a compreender, o que dizíamos antes, **que à escolha de uma norma convém adicionar elementos de aplicação externos obrigatórios.**

Bibliografia

- Arquivo Nacional da Costa Rica. Departamento Servicios Archivísticos Externos (2020). Glosario único de términos, definiciones, conceptos y abreviaturas de las normas técnicas nacionales. https://www.archivonacional.go.cr/web/dsae/glosario_%20unico_terminos.pdf
- Asociación de normalización española. (s.f) “ISO/IEC 26300-1:2015”. Disponible en <https://www.une.org/encuentra-tu-norma/busca-tu-norma/iso?c=066363>
- 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)
- Barnard, A y Voutssas, J (2014). Glosario de Preservación Archivística Digital Versión 4.0. Universidad Nacional Autónoma de México. https://iibi.unam.mx/archivistica/glosario_preservacion_archivistica_digital_v4.0.pdf
- Instituto de Normas Técnicas de Costa Rica. (s.f) ¿Qué es una norma? Disponible en: https://www.inteco.org/page/inteco_standards_development
- Instituto de Normas Técnicas de Costa Rica. (s.f). Preguntas frecuentes. Disponible en: <https://www.inteco.org/page/inteco.faq>
- International Organization for Standardization (s.f) “ISO 9660:1988/Amd 2:2020 Information processing — Volume and file structure of CD-ROM for information interchange — Amendment 2”. Disponible en <https://www.iso.org/standard/78259.html>
- Library of Congress (s.f) “Flexible Image Transport System (FITS), Version 3.0”. Disponible en <https://www.loc.gov/preservation/digital/formats/fdd/fdd000317.shtml>
- Ministerio de Ambiente y Desarrollo Sostenible de Colombia.(s.f) “Definición y entes normalizadores” disponible en: <http://www.ideam.gov.co/web/ecosistemas/normas-estandares#:~:text=En%20un%20C3%A1mbito%20formal%2C%20se,y%20servicios%20cumplen%20con%20su>
- Umaña Alpízar R. (2017). “Preservando documentos digitales auténticos” III Seminario de Legislación Archivística. Disponible en: https://www.uned.ac.cr/actividades/images/sellar/memoria/Raquel_Uma%C3%B1a_-_Costa_Rica.pdf

