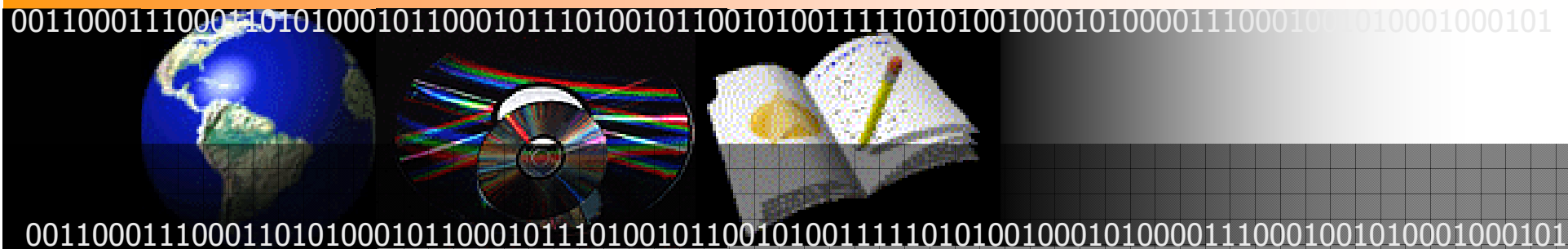


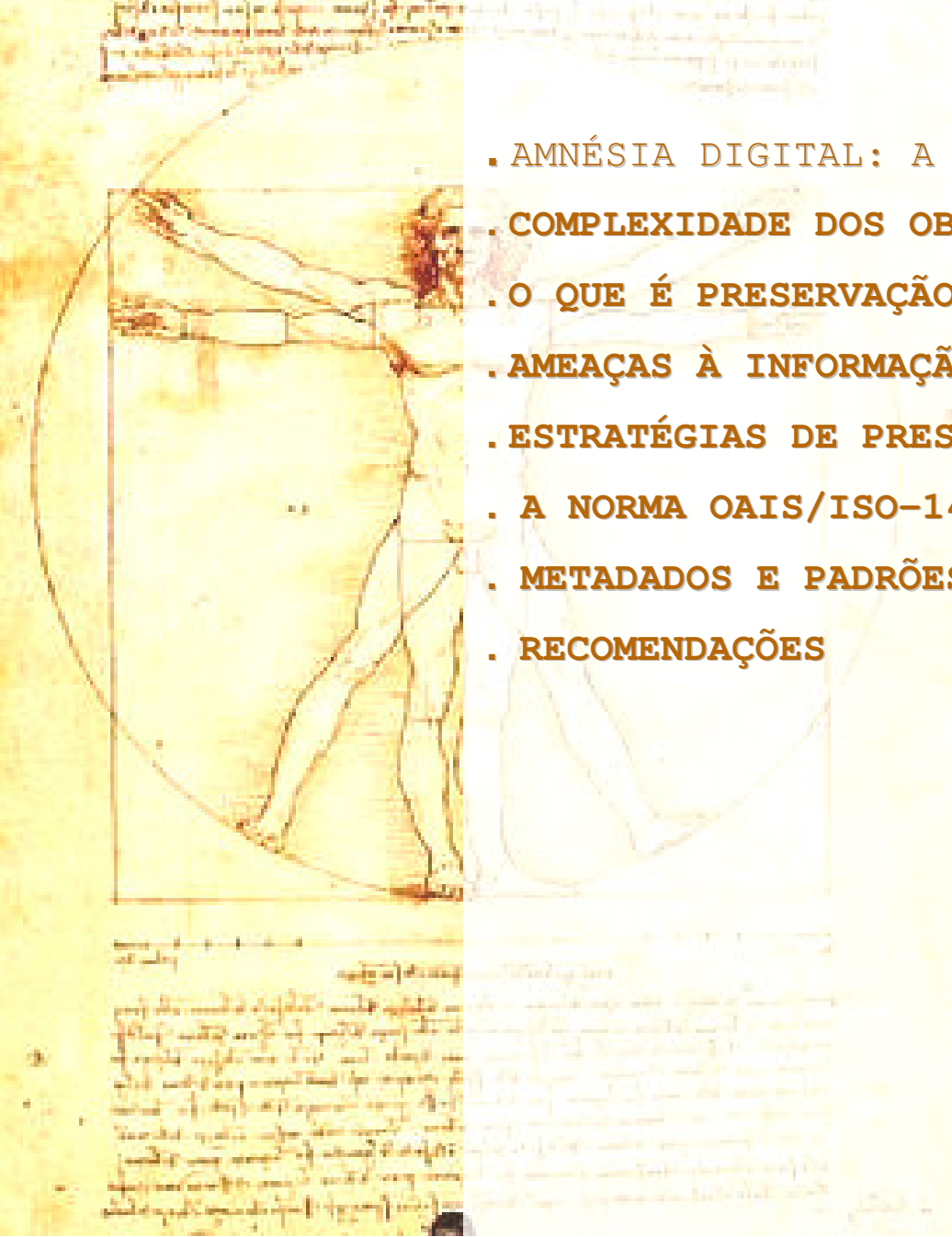
Preservação digital: uma brevíssima introdução



Luis Fernando Sayão, D.Sc.
Comissão Nacional de Energia Nuclear
Centro de Informações Nucleares
lsayao@cnen.gov.br

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

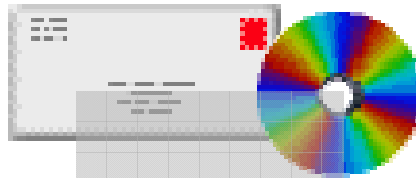
Agenda

- 
- . AMNÉSIA DIGITAL: A ERA DO ESQUECIMENTO
 - . COMPLEXIDADE DOS OBJETOS DIGITAIS E DA PRESERVAÇÃO
 - . O QUE É PRESERVAÇÃO DIGITAL
 - . AMEAÇAS À INFORMAÇÃO DIGITAL
 - . ESTRATÉGIAS DE PRESERVAÇÃO DIGITAL
 - . A NORMA OAIS/ISO-14721:2003
 - . METADADOS E PADRÕES PARA A PRESERVAÇÃO DIGITAL
 - . RECOMENDAÇÕES

0011000111100011010100010110001011101001011001010011111010100100010100001110001001010001000101

::: A era do esquecimento

:: O caso dos netos de Jeff Rothenberg



- >> ano: 2045
- >> carta escrita em 1995 e um CD-ROM
- >> problema: como ler e interpretar o conteúdo do CD
- >> a carta pode ser lida sem nenhuma ferramenta: é a única coisa imediatamente inteligível

00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

!!! O foco da preservação digital não está nas mídias!!!

A preservação digital tem pouca relação com a longevidade de discos óticos, fitas magnéticas e outros meios de armazenagem
(bem guardado o meu CD de música pode durar 100 anos, mas daqui a 10 anos haverá aparato para tocá-lo?)

O objetivo da preservação digital não é preservar os objetos digitais independente da sua **USABILIDADE**

“O SANTO GRAL DA PRESERVAÇÃO DIGITAL É O ACESSO CONTÍNUO AO CONTEÚDO INTELECTUAL DOS DOCUMENTOS DIGITAIS”

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: O foco da preservação digital é outro

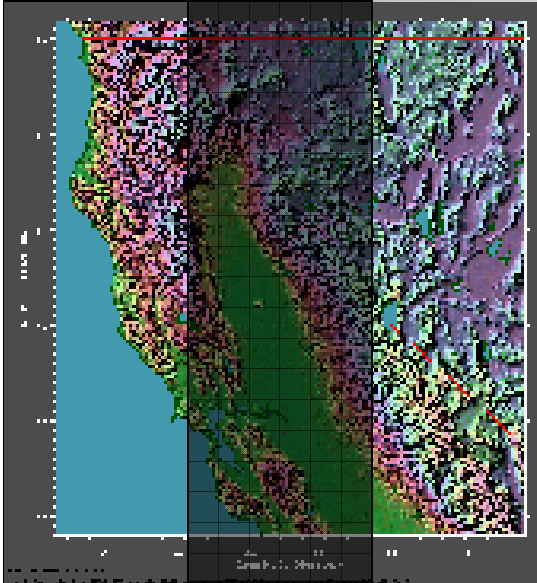
No mundo analógico - do papel e do microfilme a **PRESERVAÇÃO** e **ACESSO** são atividades relacionadas, porém distintas – e muitas vezes antagônicas ...

“ No mundo digital a **PRESERVAÇÃO** e o **ACESSO** são indissociáveis – a preservação digital se confunde com o próprio acesso ”



00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

:: Os problemas já estão em curso...



:: acidente nuclear no Canadá

:: fotos de satélite da Amazônia

**:: dados da exploração de Marte
pela Viking em 1976**

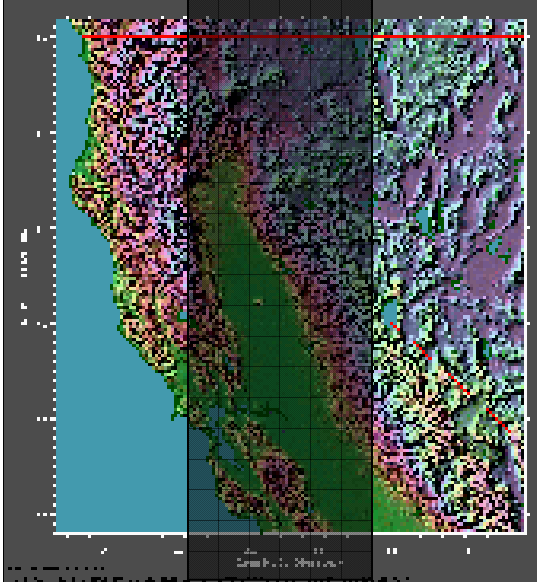
**:: caso clássico:
censo americano de 1960 ...**

:: o primeiro e-mail enviado em 1964



00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

:: Os problemas já estão em curso...



:: acidente nuclear no Canadá

:: fotos de satélite da Amazônia

:: dados da exploração de Marte
pela Viking em 1976

:: caso clássico:
censo americano de 1960 ...

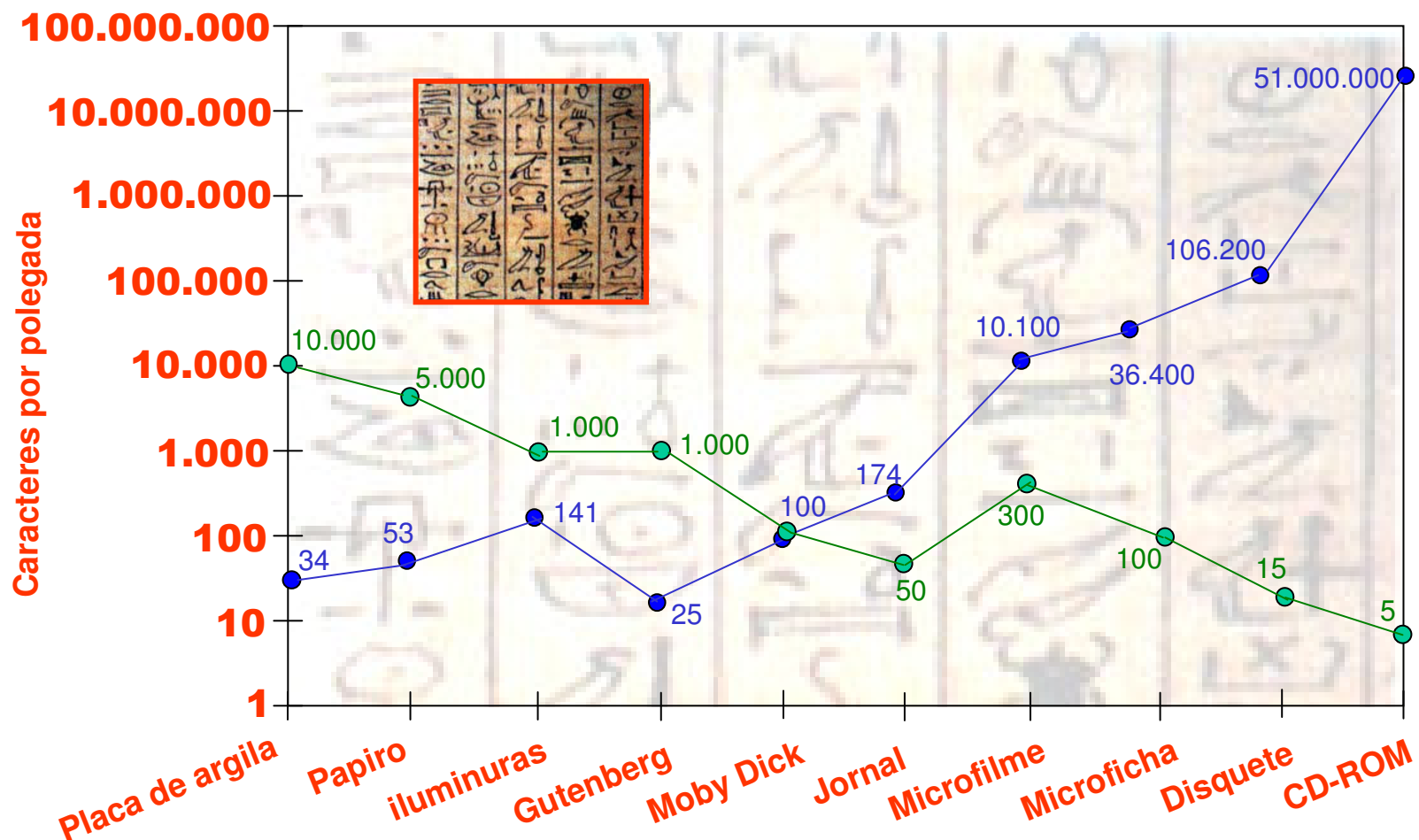
:: o primeiro e-mail enviado em 1964



**O hardware e/ou o software usados
para acessar os registros não
estão mais disponíveis e a tecnologia
atual não é compatível**

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101 ::: Densidade de informação X longevidade das mídias

A informação no formato digital é mais frágil do que os fragmentos de papiro encontrados nas tumbas dos faraós



Baseado em Paul Conway – Preservação no universo digital

00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

:: A vulnerabilidade das mídias digitais...



:: Degradação dos bits devido a condições hostis de armazenamento

efeitos das condições ambientais – efeitos físicos, químicos, magnéticos – sobre as mídias.

Exemplo: flutuação de temperatura e umidade.

:: Obsolescência tecnológica

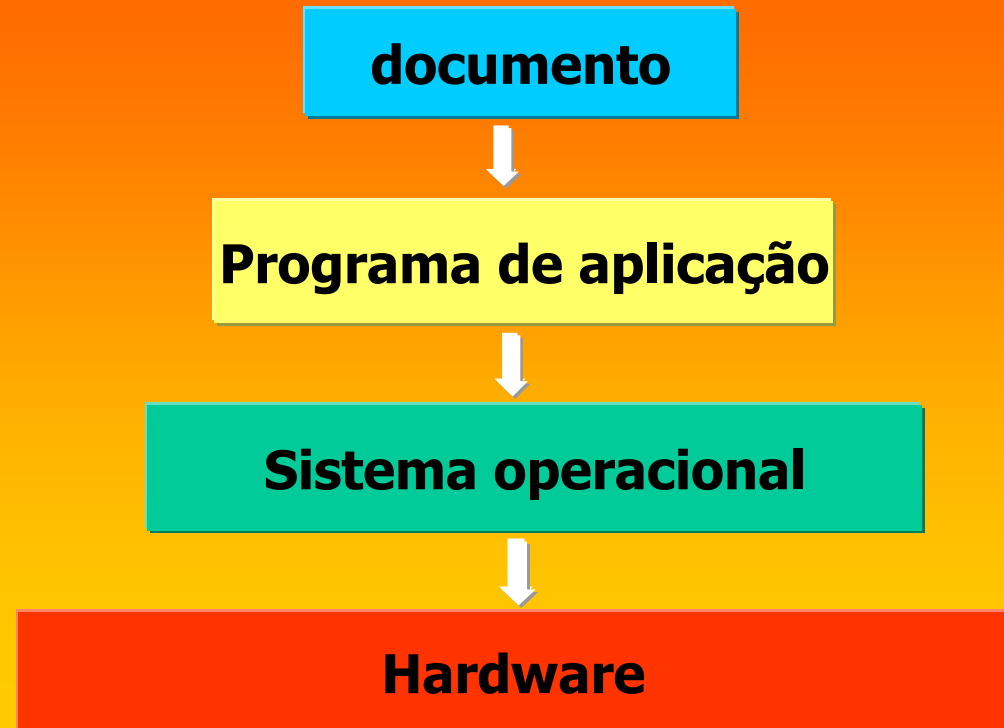
- . obsolescência da mídias
- . obsolescência de formatos
- . dependência a softwares incompatíveis ou descontinuados
- . dependência a ambientes de softwares obsoletos – S.O. I/O drivers, etc.
- . dependência a hardwares obsoletos

... causada pelas suas inúmeras dependências

[gerências e tecnológicas]



001001100010010
011011000010010
110011000110010
001001011000111
011110001001011
000011001110110
111001101100011
001001100010010
011011000010010
110011000110010
001001011000111
011110001001011
001001100010010
011011000010010



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Quais são as ameaças às informações digitais?

. **OBSOLESCÊNCIA DE MÍDIAS e de HARDWARE**

.. no sentido de não serem mais capazes de comunicação com outros componentes do sistema ou serem substituídos

. **OBSOLESCÊNCIA DE SOFTWARE e FORMATOS**

.. Embora os bits permaneçam legíveis, a informação não pode ser decodificada em uma forma legível

. **FALHAS DAS MÍDIAS**

.. degradação com tempo, efeitos físicos, químicos, magnéticos
.. perda catastrófica e irreversível de massa de dados: falha no disco

. **FALHAS DO HARDWARE**

.. falhas temporárias recuperáveis por queda de energia
.. falha catastrófica e irreversível: queima da fonte

. **FALHAS NO SOFTWARE**

.. bugs que podem por em risco os dados armazenados

. **ERROS NO SISTEMAS DE COMUNICAÇÃO**

.. falhas transferência de dados

. **FALHAS NO SERVIÇO DE REDE EXTERNO**

.. Exemplo: resolvers de domínio e de PURL

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Quais são as ameaças às informações digitais?

. ERROS DO OPERADOR

. AGRESSÕES EXTERNAS

.. vírus e worms

. AGRESSÕES INTERNAS

.. sabotagem de pessoas com acesso autorizado

. DESASTRES NATURAIS

.. Inundações, incêndios, terremotos: afetam o ambiente computacional

. PROBLEMAS ECONÔMICOS

.. Informação digital é mais vulnerável a instabilidades econômicas do que a informação em papel

. COLAPSO DA ORGANIZACIONAL

.. Acabar, falir, mudar de missão

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

... mas temos que considerar as ameaças conjunturais



- :: Rápido crescimento do número de objetos digitais e a proliferação de padrões e formatos
- :: Complexidade crescente dos objetos digitais
- :: Falta de planejamento para a preservação digital
- :: Falta de perspectiva sobre as exigências de acesso de longo prazo na criação de produtos digitais
- :: Ausência de padrões largamente aceitos que assegurem acesso através do tempo
- :: Direitos associados à propriedade intelectual e copyright que interferem na preservação digital
- :: Falta de expertise técnica em formação de coleções digitais e preservação
- :: Ênfase na criação/aquisição de material digital numa época de recursos escassos, ao invés de foco no desenvolvimento da preservação e acesso do estoques existentes
- :: Custo, falta de modelos de custos para a preservação

:: O risco de uma amnésia digital é real...



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

... porque os documentos se tornam digitais rapidamente

GIORDA
NO BRVNO
Nolano.

*Del infinito vniuerso
et Mondi.*

*Ad Illustrissimam Signor de
Mauuicicra.*



*In Stampato in Venetia.
Anno, M. D. LXXXIII.*

A urgência da sociedade moderna em transformar ou produzir toda a informação de que precisa – texto, imagens, vídeos, som – para formatos digitais é justificada

.. pela enorme economia de espaço físico

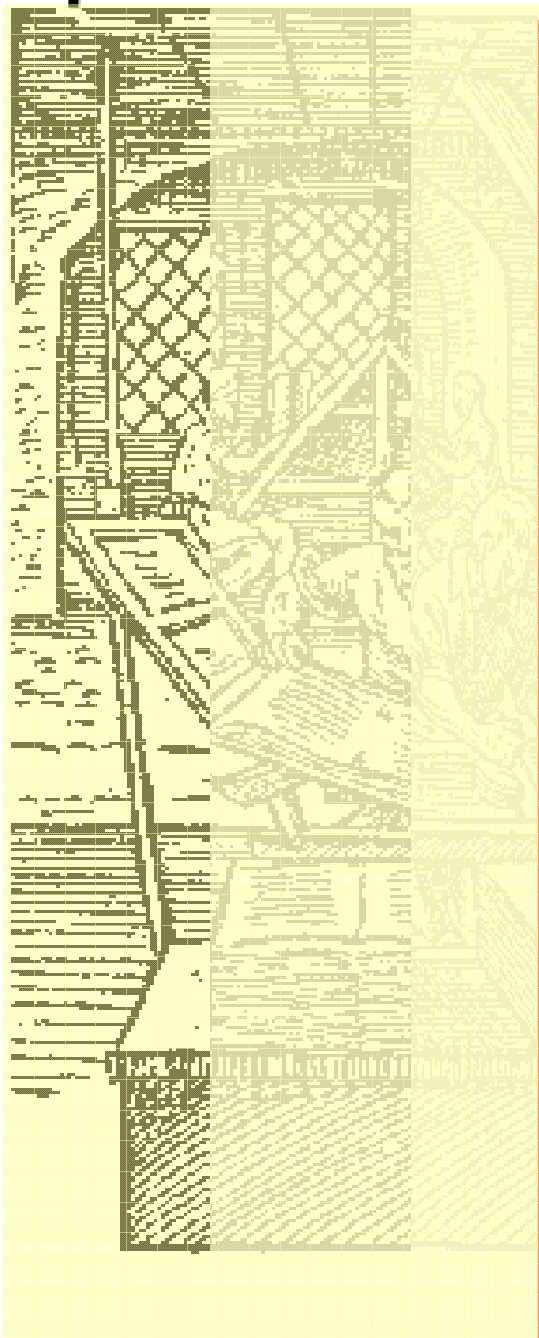
.. pelo extraordinário ganho de produtividade e eficiência proporcionado pela otimização dos fluxos de trabalho

.. pela facilidade de acesso aos estoques de informação públicos e privados

.. pela facilidade de edição, publicação e distribuição de dados e informações digitais.

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

...pela sua extraordinária versatilidade e funcionalidade



- . **Facilmente copiados**
- . **Facilidade de criação, acesso e distribuição**
- . **Facilidade de processamento – pesquisa, reformatação, edição**
- . **Facilidade de integração e reutilização – a partir de muitas fontes**
- . **Apresentam novas características que são intrinsecamente digitais, por exemplo, hipermídia distribuída, interatividade, geração dinâmica**
- . **Economiza dinheiro e papel**

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Documentos e informações digitais são críticos para o nosso futuro



001001100010010
011011000010010
110011000110010
001001011000111
011110001001011
000011001110110
111001101100011
001001100010010
011011000010010
110011000110010
001001011000111
011110001001011
001001100010010
011011000010010

:: Dependência social da informação digital

O governo, a administração pública e privada, os negócios, a pesquisa científica e a inovação tecnológica, a expressão artística e cultural depende cada vez mais do registro digitais e documentos não disponíveis de outra forma

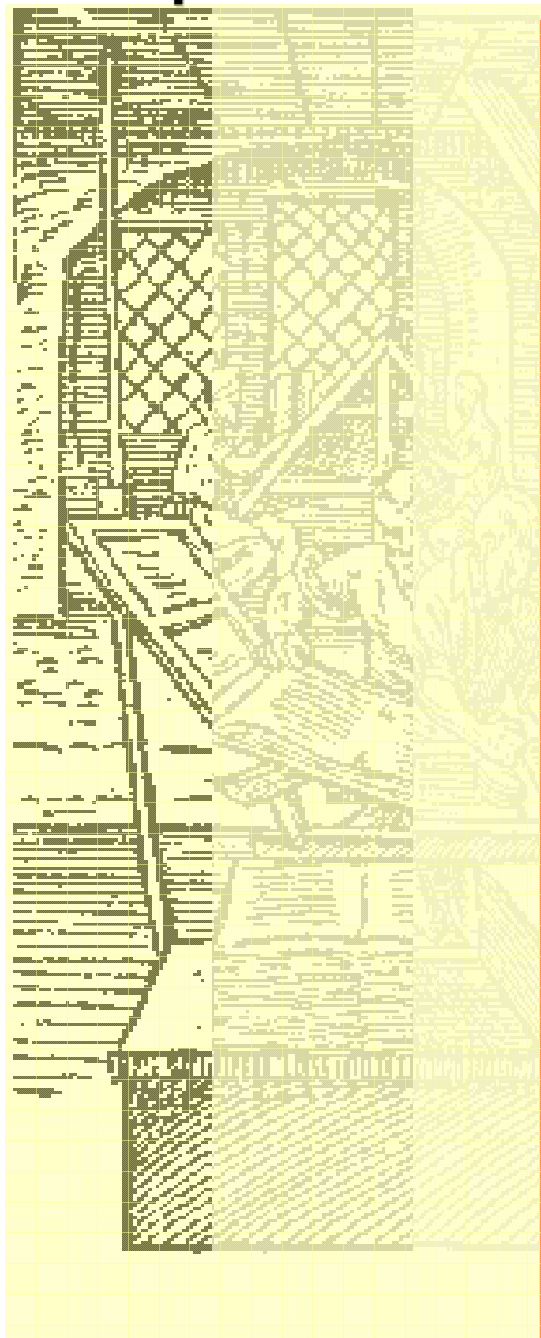
:: Apesar da Fragilidade intrínseca do armazenamento digital

Incapacidade dos sistemas de informação em assegurar a preservação de longo prazo num contexto de rápido avanço tecnológico

Constatação de que é a informação digital é difícil de preservar, mas também é difícil de controlar

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

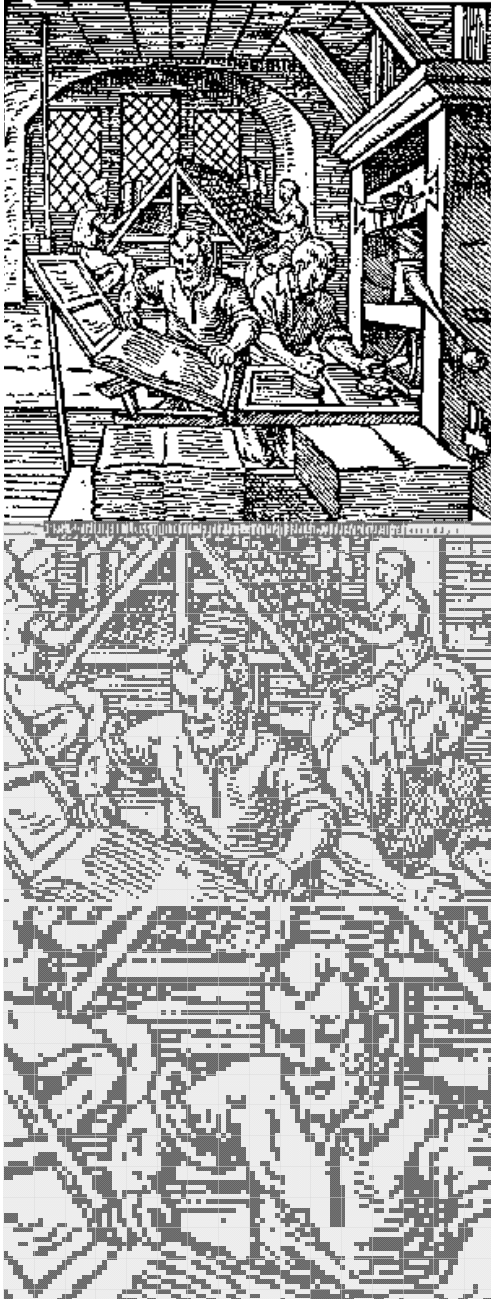
...e a apesar da nossa incapacidade preservar e controlar a informação digital



“A nossa compulsão em produzir informações digitais é infinitamente maior do que a nossa capacidade de garantir acesso a ela, e não há em contrapartida estratégias consolidadas para garantir o acesso de longo prazo as informações de valor contínuo”

:: Considerando a vulnerabilidade do ambiente digital

[é necessário...]



Garantir que a herança cultural, histórica, científica, técnica, econômica, administrativa da humanidade, expressa cada vez mais por manifestações distintas de documentos digitais, possa ser:

Continuamente acessada, em termos de conteúdo e funcionalidades

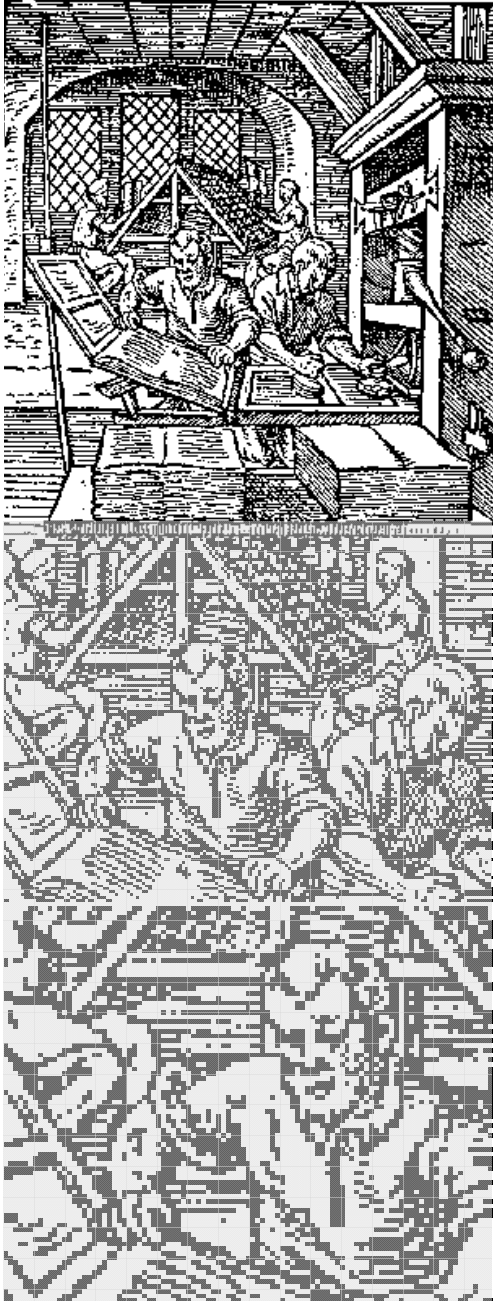
Através dos recursos tecnológicos disponíveis na época

Tenha garantida sua:

- ... integridade
- ... autenticidade
- ... confiabilidade
- ... sem que jamais ocorra perda do registro, da prova, do testemunho ou da memória

00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

... o acesso à informação digital de valor contínuo



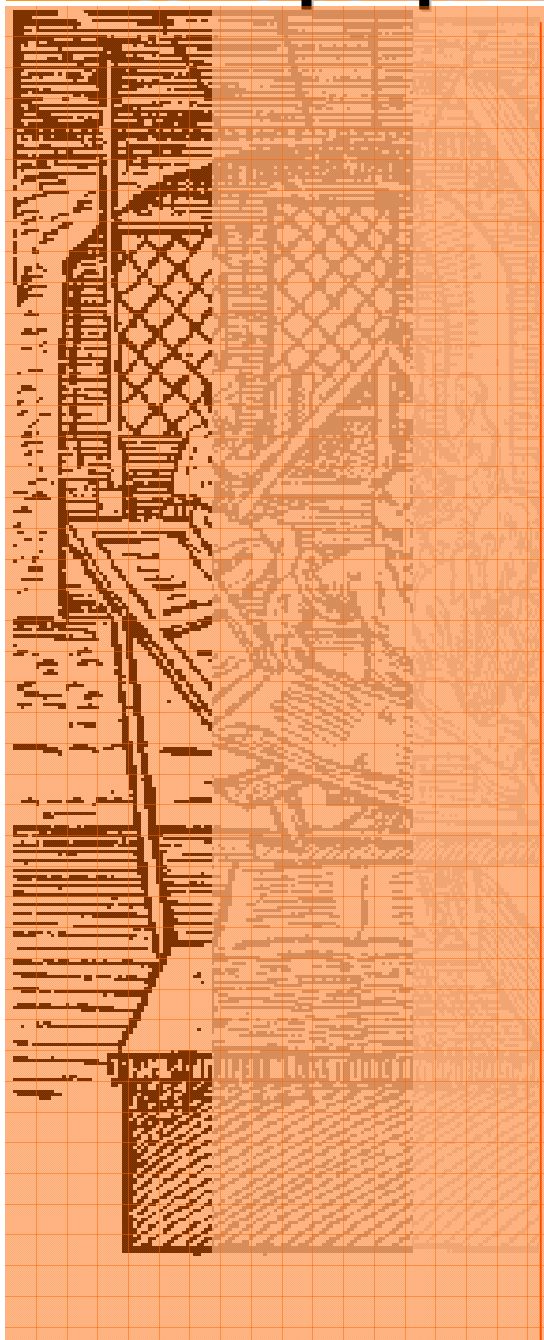
Talvez não haja maior desafio para os profissionais de informação do que garantir meios de acesso às informações digitais para as futuras gerações

A questão essencial que se coloca para a sociedade da informação pode bem ser como salvar a nossa memória digital

“ Gestão no presente para garantia da acessibilidade no futuro ”

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

...mesmo porque tudo mudou na forma de registrar informação



REGISTROS INVISÍVEIS

Pela primeira vez em 3.500 anos de atividade de gestão de documentos, produzimos registros que não existem para o olho humano. Completamente diferentes das placas de argila da Babilônia, dos papiros egípcios, dos pergaminhos romanos, do papel moderno e mesmo do microfilme

GESTÃO

Pela primeira vez, profissionais das mais diversas áreas, sem capacitação para a gestão de estoques informacionais, são responsáveis por criá-los e armazená-los

DESMATERIALIZAÇÃO dos SUPORTES

E o mais importante, pela primeira vez não estamos, não estamos produzindo, gerenciando e guardando artefatos físicos, mas tentando entender e preservar padrões virtuais que dão a informação digital seu conteúdo, estrutura, contexto e assim o seu significado, que são completamente controlados por software

:: Resumindo!!!

PROBLEMA PRIMORDIAL DE PRESERVAR DOCUMENTOS DIGITAIS TEM ORIGEM NA NATUREZA “DIGITAL” DESSES DOCUMENTOS

DEPENDÊNCIA AO AMBIENTE TECNOLÓGICO

Ao contrário dos documentos convencionais tais como livros, periódicos, manuscritos ou microfilmes, os objetos digitais só podem ser acessados por intermédio de uma combinação de hardware e software

CICLO RÁPIDO DA INOVAÇÃO TECNOLÓGICO

No ambiente competitivo da T.I., o ciclo de inovação tecnológica é muito rápido, significando que hardware e software se tornam obsoletos em ciclos de menos de três anos

ACOMPANHAR AS MUDANÇAS TECNOLÓGICAS

Assegurar acesso contínuo aos estoques de informações digitais, portanto, requer acompanhar as mudanças tecnológicas e deslocar a base tecnológica dos objetos – formatos, dispositivos de armazenamento, sistemas operacionais, etc.- para a base corrente.



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Mas como preservar?

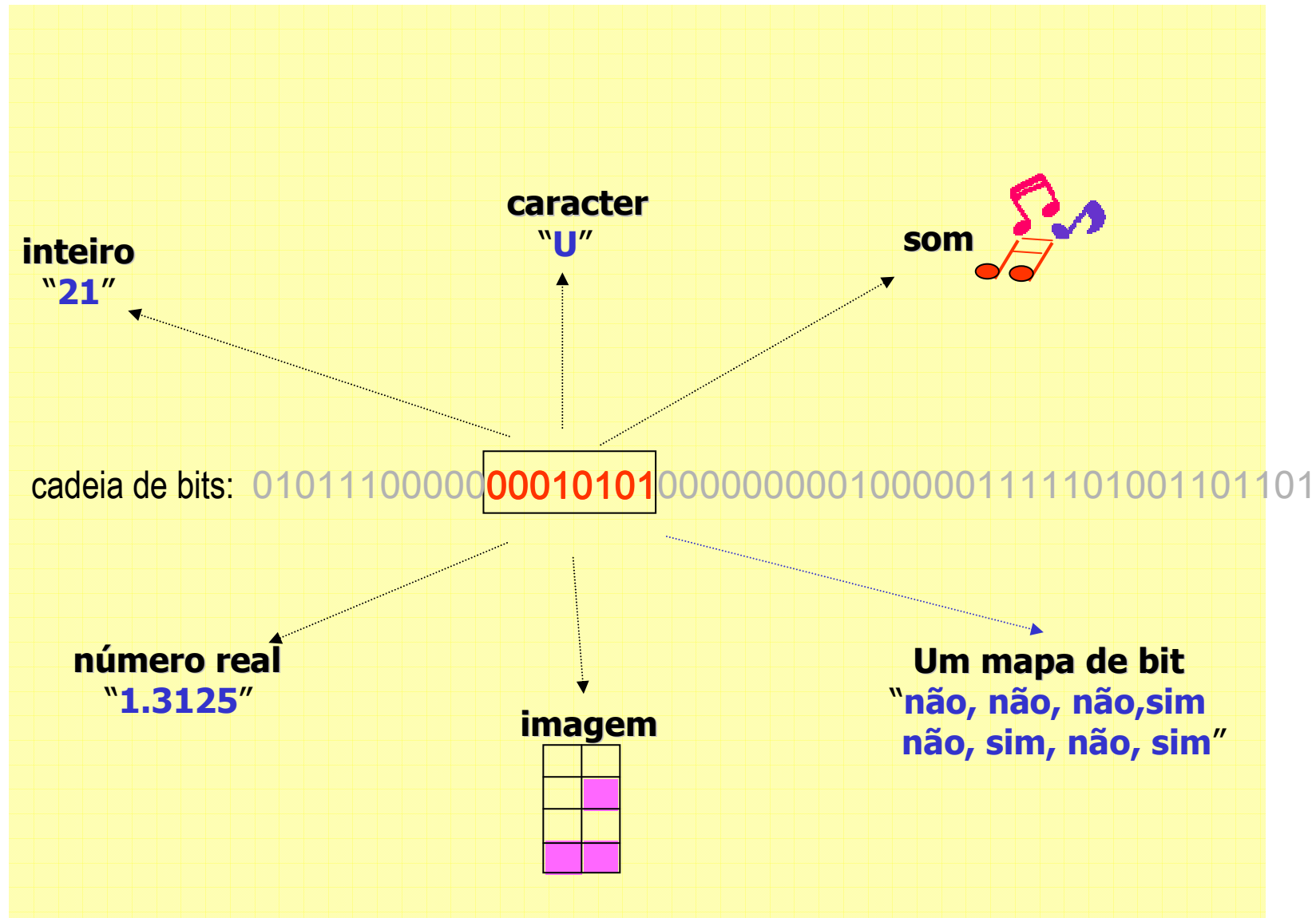


A natureza digital de uma informação é dada pela forma como ela é armazenada: seqüências de 0's e 1's, isto é, cadeias de bits (bitstream)

Podemos, portanto, concluir que a forma mais básica de preservar a integridade de um objeto informacional é preservar a configuração de bits que o define como um objeto único

:: Um byte numa cadeia de bits pode representar qualquer coisa
[e deve ter o seu significado interpretado]

[e deve ter o seu significado interpretado]



copiado de Jeff Rothenberg
Scientific American – Jan. 1995

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

... e precisa ser interpretada pelo software que a criou

Data	Operação	Valor	Saldo	Movimento
05/04/03	depósito	500,00	500,00	Conta Corrente
26/04/03	cheque#314	100,00	400,00	
27/04/03	depósito	60,00	460,00	
03/11/03	cheque#315	100,00	360,00	

Para compreender uma cadeia de bits

é necessário conhecer o formato usado para criá-la

- data: 6 dígitos (ddmmaa)
- no cheque: 4 dígitos
- depósito: 0000
- valores: 11 dígitos

0504030000000000050000000000050000
260403031400000010000000000040000
270403000000000006000000000046000
031103031500000010000000000036000

050403000000000005000000000005000026040303140000001000000000004000270403000000000006000000000046000031103031500000010000000000036000

As entradas concatenadas produzem uma cadeia de dígitos decimais

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

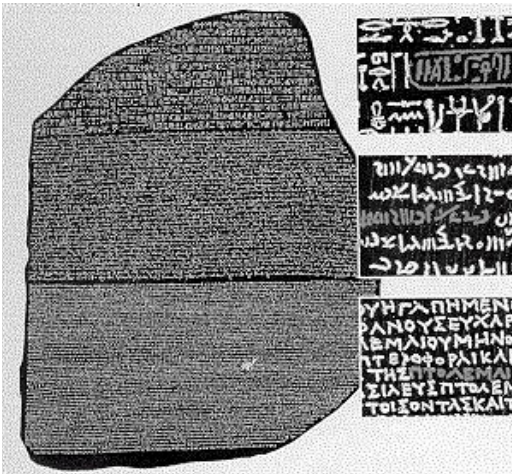
:: Salvar os bits é necessário mas não suficiente

“ Recuperada uma cadeia de bits ela deve ter o seu significado interpretado, pois a maioria dos arquivos contém informações cujo significado só pode ser explicitado pelo software que as criou. ”

:: essa interpretação é fortemente dependente de formatos, códigos e estruturas – que não podem ser explicitamente representadas na cadeia de bits

:: portanto, para se **compreender um documento** é necessário conhecer o **significado** do seu conteúdo na linguagem para o qual ele foi construído

:: no caso de um documento digital essa linguagem é um **programa de computador**



:: **salvar a cadeia de caracteres sem salvar o seu interpretador é como salvar o hieróglifos sem salvar a Pedra de Rosetta**

00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101
:: Todos os documentos digitais são dependentes de software
que por sua vez, são dependentes de hardware

“ Documentos digitais só podem ser vistos através do processamento de um programa de computador ”

- :: eles estão armazenados de forma codificada
- :: sua estrutura e conteúdo só podem ser entendidas pelo programa que o criou
- :: **sem este programa não podem ser acessados, lidos ou mesmo impressos**
- :: têm que ser interpretados para serem compreendidos pelos humanos
- :: **eles são essencialmente programas**
- :: **sem um software o documento digital é inútil e inexistente**



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Por que é tão difícil preservar documentos digitais?

“ A preservação digital envolve não somente a retenção do objeto informacional em si, mas também do seu significado. É necessário, portanto, que as técnicas de preservação sejam capazes de compreender e recriar a forma original ou a função do objeto de forma que seja assegurada sua autenticidade e acessibilidade

:: não podemos simplesmente salvar documentos digitais como documentos físicos

”



00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

:: Por que é tão difícil preservar documentos digitais?

“Além do mais, as estratégias de preservação devem ter uma amplitude tal que incorporem vários outros aspectos, incluindo custo-benefício, restrições legais, gestão e requisitos de acesso para o usuário-final, etc.”
:: preservação digital não é somente um problema tecnológico



:: Mas o que devemos preservar?



O problema primordial da preservação é o conteúdo, ou seja, a substância intelectual contida nos objetos informacionais.

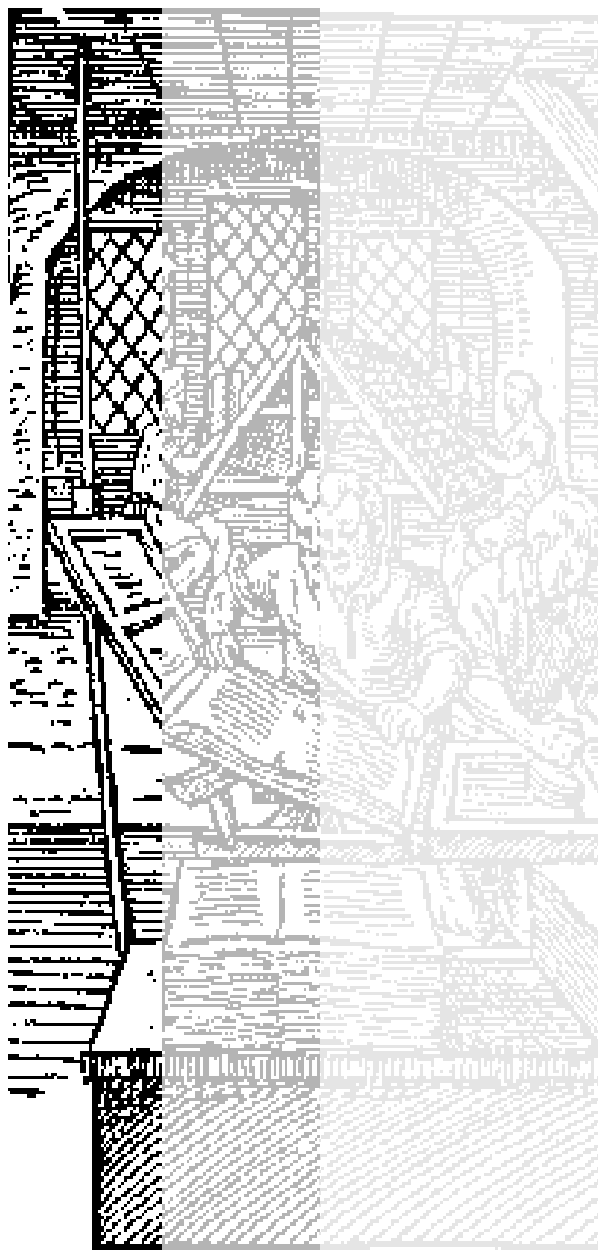
Quando um livro é preservado no seu formato original, todos os aspectos do livro são preservados, seu layout, seu formato, seu conteúdo. Todos os elementos estão inextricavelmente unidos.

... e o processamento ocorre na cabeça do leitor

Quando um documento é preservado, suas assinaturas, carimbos, autenticação, também são preservados

00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

:: Mas o que devemos preservar?



Mas no mundo digital a noção de conteúdo é complexa e estratificada

Objetos digitais são facilmente decompostos em elementos

No mundo digital, a primeira tarefa é identificar os aspectos múltiplos que devem ser preservados

A preservação digital deve garantir que o maior número possível desses aspectos devam persistir no tempo.

no mundo digital a noção de conteúdo é complexa e estratificada

Fixar os limites do objeto

um livro é um livro, mas qual é o limite de um objeto digital?

Preservar a cadeias de bits

é a presença física, formada por séries de “0”s e “1”s

Preservar o conteúdo

capacidade de acessar o conteúdo no seu nível mais baixo
tal como um texto ASCII

Preservar a apresentação

forma, layout, fontes, tamanho, margens, colunas, cores, paginação...
em muitos casos está especificado separadamente – folha de estilo

Preservar as funcionalidades

componentes multimídia, conteúdo dinâmico, interoperabilidade,
busca

Preservar a autenticidade

verificar se o documento é o que diz ser, se as transformações preservaram
a sua forma original; se sofreu mudanças não autorizadas

Localizar e referenciar ao longo do tempo

nomear e referenciar um objeto digital de forma a distingui-lo de outras versões, cópias e
e edições

Preservar a proveniência

origem e cadeia de custódia confirmam a autenticidade e a integridade

Preservar o contexto

objetos digitais são também definidos por suas dependências a hardware,
software, modo de distribuição e links com outros objetos



no mundo digital a noção de conteúdo é complexa e estratificada

Fixar os limites do objeto

um livro é um livro, mas qual é o limite de um objeto digital?

Preservar a cadeias de bits

é a presença física, formada por séries de “0”s e “1”s

Preservar o conteúdo

capacidade de acessar o conteúdo no seu nível mais baixo
tal como um texto ASCII

Preservar a apresentação

forma, layout, fontes, tamanho, margens, colunas, cores, paginação...
em muitos casos está especificado separadamente – folha de estilo

Preservar as funcionalidades

componentes multimídia, conteúdo dinâmico, interoperabilidade,
busca

Preservar a autenticidade

verificar se o documento é o que diz ser, se as transformações preservaram
a sua forma original; se sofreu mudanças não autorizadas

Localizar e referenciar ao longo do tempo

nomear e referenciar um objeto digital de forma a distingui-lo de outras versões, cópias e
e edições

Preservar a proveniência

origem e cadeia de custódia confirmam a autenticidade e a integridade

Preservar o contexto

objetos digitais são também definidos por suas dependências a hardware,
software, modo de distribuição e links com outros objetos

METADADOS

011011000010010

110011000110010

001001011000111

011110001001011

001001100010010

011011000010010

1100110001010010

001001100010010

011110001001011

000011001110110

1110011001100011

0010011000110010

011011000010010

110011000110010

001001100010010

011110001001011

001001100010010

011011000010010

001001100010010

011011000010010

00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

::: Mas afinal o que é Preservação digital???

É a capacidade de manter documentos e arquivos digitais disponíveis por períodos de tempo que **transcendam os avanços tecnológicos** sem alterações ou perda de **legibilidade**

[The Association for Information and Image management]

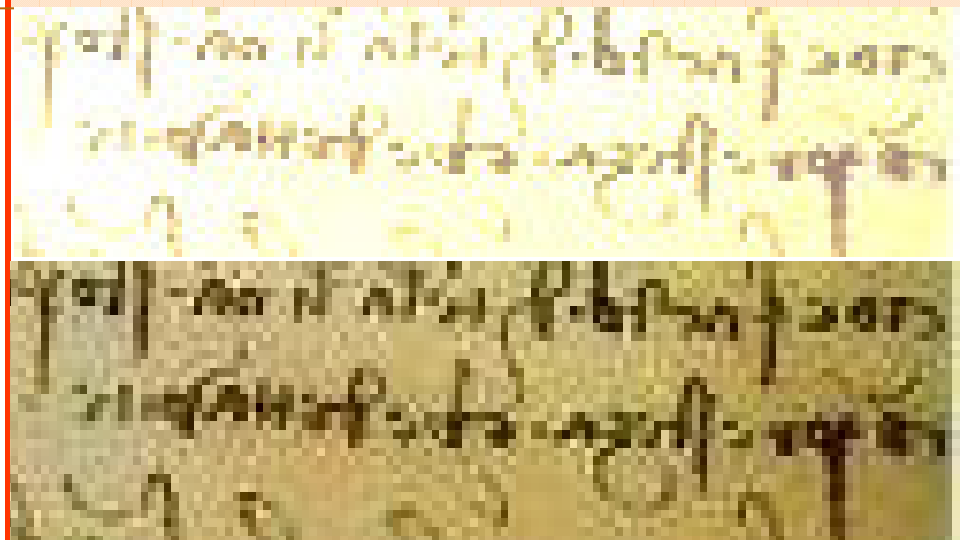
É o conjunto de **atividades gerenciais** necessárias para assegurar o **acesso contínuo** e a preservação de materiais **digitais**

[RLG/OCLC Report]

Planejamento, alocação de recursos e aplicação de métodos de preservação e tecnologias necessárias para que a **informação digital de valor contínuo permaneça acessível por longo prazo.**

[Hedstrom]

Capacidade de manter a **integridade e acessibilidade** da informação digital por longo prazo. Esta preservação de integridade e acessibilidade não se limita apenas a **proteger a informação digital contra o acesso não autorizado**, mas também contra o uso inadequado resultante da **má interpretação ou má representação da informação** por parte dos sistemas computacionais.
:: Waters e Garret (1996)



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Afinal o que é preservação digital?

:: Preservação Física

Foco nas mídias e na sua renovação

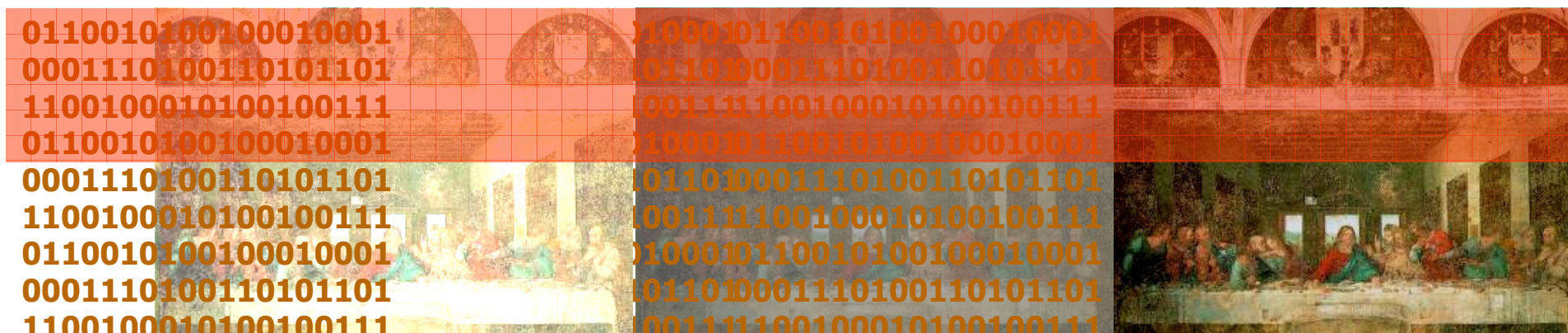
:: Preservação Lógica

Foco nos formatos e dependência de hardware e software que mantenham legíveis e interpretáveis a cadeia de bits

:: Preservação intelectual

Foco no conteúdo intelectual e na sua autenticidade e integridade

[é necessário ainda preservar do aparato – na forma de metadados – necessários para localizar, recuperar e interpretar a informação digital; assim como o monitoramento e a instrumentalização da comunidade alvo, para que ela possa compreender plenamente a informação no momento do seu acesso]



00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

Por que a preservação digital é um problema tão complexo?



Porque envolve muitas variáveis...

- :: conceituais, modelagem, terminológicos
- :: representação, metadados, recuperação
- :: normas, padrões, protocolos
- :: jurídicas, legais
- :: políticas, estratégicas
- :: comerciais, modelo de custo, de negócio
- :: gestão, planejamento, administração
- :: pesquisa, educação, treinamento
- :: metodologias, operação
- :: pesquisa, educação, treinamento
- :: cooperação, intercâmbio, interoperabilidade
- :: segurança, integridade
- :: usuário
- :: tecnologia de informação

:: COMPROMISSOS DE LONGO PRAZO

001001100010010
011011000010010
110011000110010
001001011000111
011110001001011
000011001110110
111001101100011
001001100010010
011011000010010
110011000110010
001001011000111
011110001001011
001001100010010
011011000010010

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Preservação digital é também um paradoxo...

Tradicionalmente preservar significa manter imutável e intacto, entretanto no ambiente digital preservar significa, na maioria dos casos, mudar, recriar, renovar: mudar formatos, renovar mídias, hardware e software

:: por um lado, queremos manter a informação intacta, como ela foi criada

:: por outro, queremos acessá-la dinamicamente e com as mais avançadas ferramentas



00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

:: Não há uma estratégia única de preservação

“Até o presente momento – e provavelmente num futuro indefinido – não teremos disponível uma estratégia única. O que temos são soluções específicas para casos específicos”

a comunidade interessada foi capaz de desenvolver métodos efetivos para a preservação dos conteúdos informacionais dos materiais digitais cujos formatos e estruturas são bem conhecidos e mais simples, mas falha ou obtém resultados pífios para materiais mais complexos e dinâmicos, ou novas formas



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: A solução mais adotada é não fazer nada...

Ausência de solução

:: não fazer nada: migração por necessidade dos documentos ativos e deixar os outros desaparecerem

:: arqueologia digital: deixar para o futuro; "criptografia" para decifrar as cadeias de bits, pois já haverá capacidade de processamento barata no futuro

0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111

100010110010100100010001
011010001110100110101101
000111100100010100100111
100010110010100100010001
011010001110100110101101
000111100100010100100111
0100010110010100100010001
011010001110100110101101
000111100100010100100111



00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

... ou retornar ao conforto do mundo analógico

Soluções parciais

- :: salvar imagens de páginas de documentos: em formas digitais (TIFF, PDF, etc.) e/ou analógicas - microfilme, impressão, hd-rosetta, etc. (low-tech)
- :: extrair e salvar o "conteúdo essencial" do documento
- :: reformatar documentos segundo formatos padronizados ou "canônicos"
- :: salvar os códigos-fonte do softwares que processam para prática de engenharia reversa no futuro



00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

... ou retornar ao conforto do mundo analógico

Endereço  <http://www.norsam.com/hdrosetta.htm>

HD-Rosetta
Data Preservation

MEMS and
Micromaching

Jewelry Applications

Failure Analysis

Our Partners

Investor Info

Contact Us

HD-Rosetta Archival Preservation Services

Norsam's High Density Rosetta (HD-Rosetta) provides analog storage of information and images on nickel plates that last for thousands of years. HD-Rosetta is patented worldwide by Norsam under exclusive license from Los Alamos National Laboratory.



HD-ROSETTA IS:

Compact

On a plate no larger than 2" square and 1/4" thick, the High Density Rosetta has immense storage capabilities: Based on a page size of 8.5x11", you may store about 196,000 pages with electron microscope retrieval between 5,000 - 18,000 pages with optical microscope retrieval.

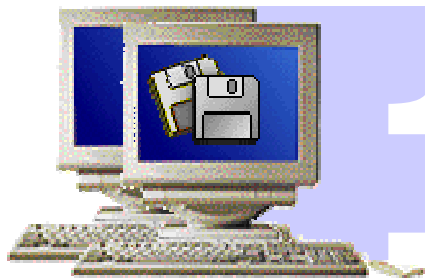
Technologically Secure

Unlike digital data, which require medium-specific software and operating systems to retrieve information, the Rosetta needs only a simple magnifying lens to view the analog data. It will never face obsolescence.

Magnetically Stable

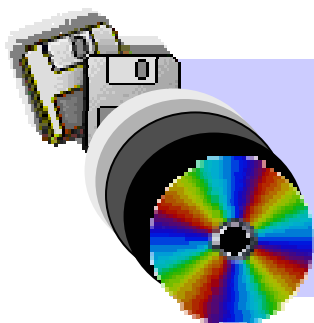
Since the Rosetta's information is etched into nickel, it is never affected by electromagnetic

:: Principais abordagens de preservação digital



Preservar o ambiente tecnológico

- :: preservação tecnológica**
- :: emulação tecnológica**



Contornar a obsolescência tecnológica dos formatos dos arquivos

- :: migração**
- :: encapsulamento**

:: Principais abordagens na preservação digital



Preservação Tecnológica

[museu completo de equipamentos]

- :: manutenção de museus equipamentos
- :: preservar o programa de aplicação original, sistema operacional, e plataforma de hardware...

:: Principais abordagens na preservação digital



Preservação Tecnológica

- + preserva o comportamento e aparência do objeto digital**
- + pode ser uma boa estratégia a curto prazo e para situações críticas**

- espaço, manutenção e custo**
- documentação**
- suporte, assistência técnica**
- deterioração das mídias e do equipamento**
- acesso restrito – museus de computadores**
- perda do aspecto distribuído da informação digital**

:: Principais abordagens na preservação digital

Emulação da tecnologia

o que é...

Emulador no contexto geral

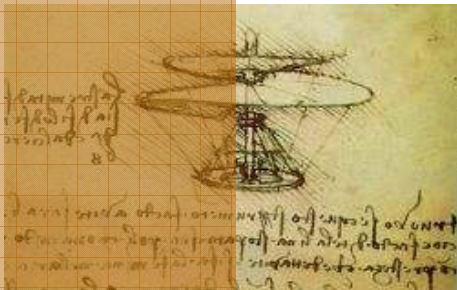
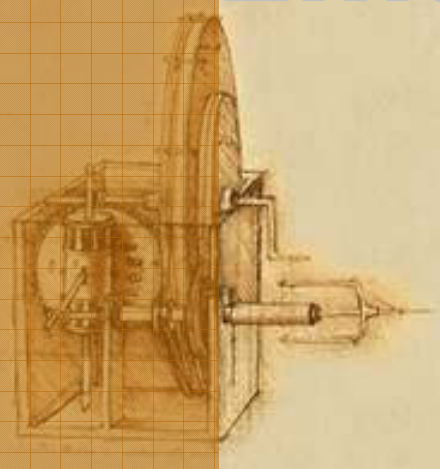
é alguma coisa que desempenha a função e o comportamento de outra coisa

Emulador no nosso contexto

é um programa que faz um computador agir como se fosse um outro – e diferente - computador [tornando-o capaz de rodar programas que rodam no outro computador]

um emulador de um computador obsoleto pode ser rodado num computador do futuro

um emulador é uma máquina virtual [é um programa (software) que cria virtualmente uma máquina (hardware)]

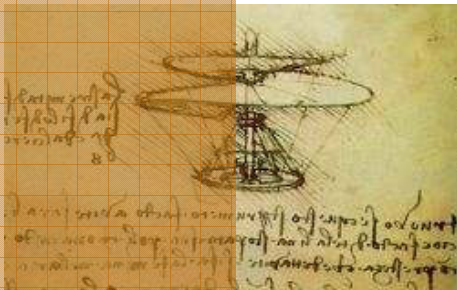
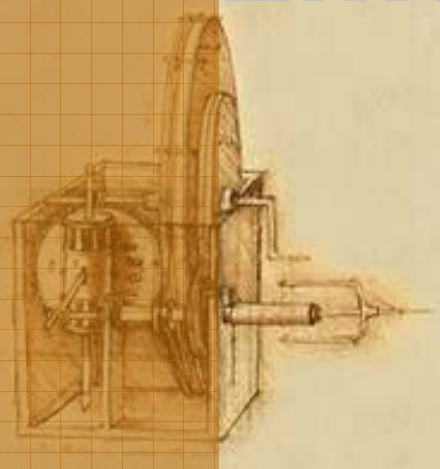


0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111

:: Principais abordagens na preservação digital

Emulação da tecnologia

características



0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111

Tem muitos pontos em comum com a estratégia de preservação tecnológica

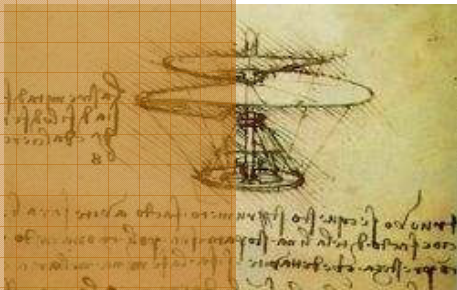
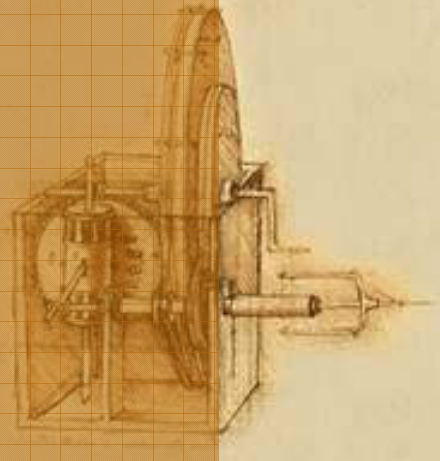
Envolve preservar o programa de aplicação original

Não envolve preservar antigos computadores nem sistemas operacionais originais

Quando a presença do objeto digital original é de grande importância e não pode ser convertido para formatos independentes ou sofrer processo de migração

:: Principais abordagens na preservação digital

Emulação da tecnologia domínio



0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111

Programas de aplicação

Recursos digitais complexos que contenham arquivos executáveis

Recursos sobre os quais faltam conhecimentos suficientes

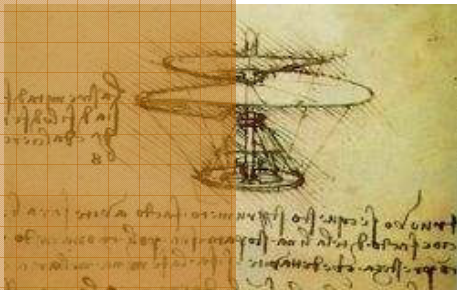
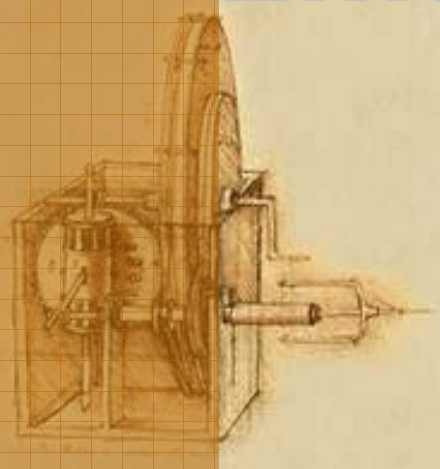
Recursos cujo valor é desconhecido e cujo uso no futuro é improvável

Recursos cuja aparência e o comportamento são importantes

:: Principais abordagens na preservação digital

Emulação da tecnologia

pontos positivos



0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111

Preserva potencialmente o recurso propriamente dito na sua versão original

Pode ser uma boa estratégia a longo prazo para acesso a recursos complexos como arquivos executáveis

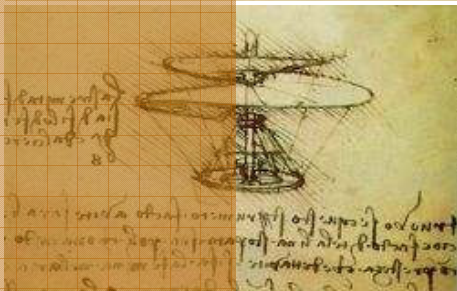
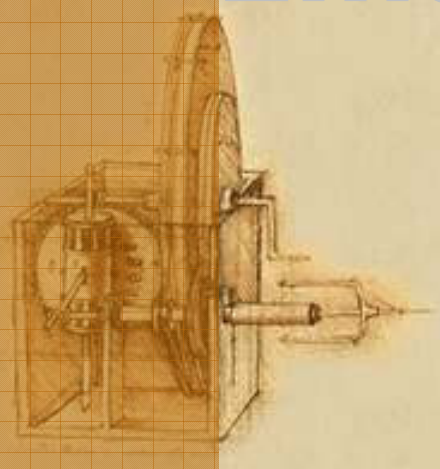
Evita a necessidade de compreender os formatos originais de documentos individuais

[exceto saber que conjunto de software é necessário para abrir o documento]

:: Principais abordagens na preservação digital

Emulação da tecnologia

pontos positivos



0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111

Requer (teoricamente) esforço zero por documento, inicialmente e ao longo do tempo
[exceto para fazer "rejuvenescimento" da cadeia de caracteres para novas mídias]

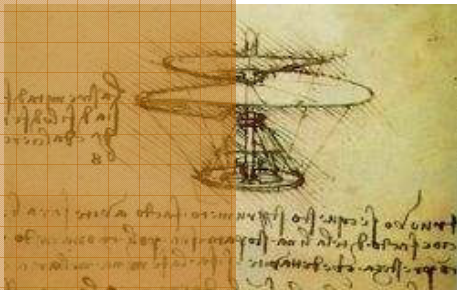
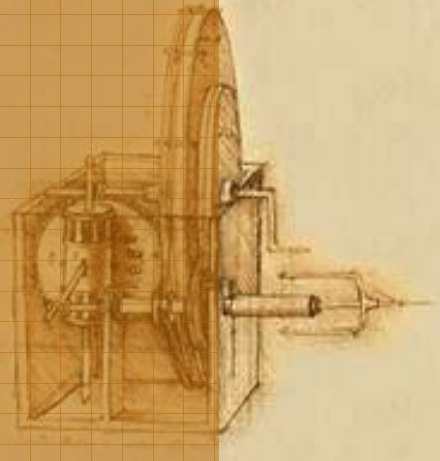
Dispensa a necessidade de conversão para novos formatos, a menos que se queira acessar no futuro o documento em tais formatos

Mantém a comportamento/aparência do recurso

:: Principais abordagens na preservação digital

Emulação da tecnologia

alguns problemas



0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111
0110010100100010001
0001110100110101101
1100100010100100111

Complexidade de se criar softwares emuladores

[requer especificação detalhada de hardware e software]

Os usuários precisam saber como rodar os softwares antigos

[decifrar helps, documentação, tutoriais etc. como antigos manuscritos]

A cadeia de caracteres tem que permanecer intacta

[não ser corrompida por compressão, transformações, cópias imperfeitas, etc.]

Pode ser necessário emular mais que somente o processador

[mas também placas aceleradoras, de som, gráficas, periféricos, interfaces de I/O, etc.]

Requer um emulador e um ambiente de emulação para cada plataforma

[cada máquina antiga numa máquina nova ou recursivamente]

[compromisso com fornecedores ou autônomos na criação de emuladores]

:: Principais abordagens na preservação digital



Migração

o que é

Transferência periódica do recurso digital de:

Uma mídia que está se tornando obsoleta ou fisicamente deteriorada, ou ainda menos estável para para um suporte mais novo

De um formato ultrapassado para um formato mais atual ou padronizado

De uma plataforma computacional – hardware e software - em vias de descontinuidade para outra mais mais moderna

O objetivo da migração é preservar a integridade dos objetos digitais e e manter a capacidade de eles serem recuperados, exibidos e usados face as constantes mudanças tecnológicas

:: Principais abordagens na preservação digital



Migração

estratégias

**Transferência do recurso digital para mídias
mais estáveis – digitais ou analógicas**

**Uso de softwares retrospectivamente
compatíveis**

**Conversão do recurso para um número
limitado
de formatos padronizados independentes de
hardware e de software**

OAIS:

:: rejuvenescimento (refreshment)

:: replicação (replication)

:: re-empacotamento (repackaging)

:: transformação (transformation)

:: Principais abordagens na preservação digital



Migração

características

Foco no conteúdo intelectual e em assegurar a acessibilidade usando a tecnologia corrente

Considerada a mais promissora por muitos autores e a único que tem funcionado operacionalmente até o momento

Para recursos acessados com frequência e sob gerência: dados científicos e bases de dados

Para recursos cujos formatos são bem conhecidos

:: Principais abordagens na preservação digital



Migração

problemas

A migração não preserva originais

[potencialmente garante o acesso, mas não ao original, é a cópia da cópia da cópia...]

Recursos complexos como multimídia podem sofrer perdas severas de funcionalidade

Pode ser inviável diante de mudanças de paradigmas

[modelo hierárquico para relacional, a conversão não é naturalmente viável]

Deve ser aplicada repetidamente para cada documento individualmente

[independente do nível de acesso; procedimentos específicos para cada tipo de documento, aplicação, formato, etc.]

:: Principais abordagens na preservação digital



Migração

problemas

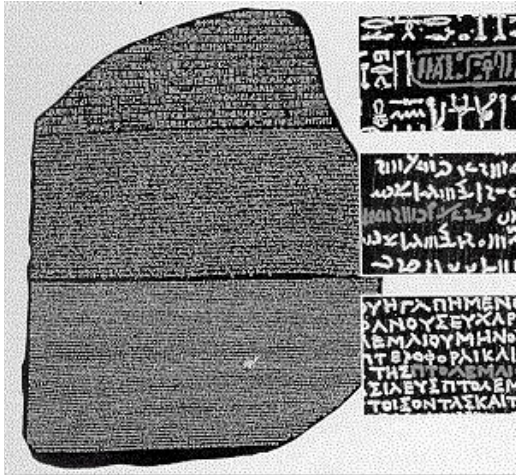
**Custos ao longo do tempo podem exceder
o custo de preservar a tecnologia e de
detalhar a especificação técnica para
uma futura emulação**

Grau de risco bastante considerável

[em relação a viabilidade, os requisitos, as perdas, repetição das
conversões, mudanças rápidas de padrões]

Degradação da informação

:: Principais abordagens na preservação digital



Encapsulamento

Todos os detalhes de como interpretar o objeto digital fazem parte da informação encapsulada

- :: Uso de estruturas físicas ou lógicas chamadas “containers” ou “empacotadores”**
- :: Uso de metadados**
- :: Uso de formatos mais documentados**
- :: UPF – Universal Preservation Format**

00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

!!! Os padrões são muito importantes!!!

Não importando a estratégia de preservação digital adotada, os padrões – principalmente os abertos – permitem que os documentos digitais sejam representados em formatos mais duradouros e estáveis, reduzindo o ciclo de obsolescência.

- :: MAIS FÁCIL
- :: MAIS BARATO
- :: MENOS FREQUENTE

“ IDEALMENTE A PADRONIZAÇÃO DEVE PRECEDER A PRÓPRIA CRIAÇÃO DO OBJETO DIGITAL ”

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

... mas têm limitações



- :: Os padrões e protocolos, mesmo abertos, se tornam obsoletos
 - . mudanças causadas pelo desenvolvimento
 - . mudanças de paradigmas
 - . mercado competitivo
 - . necessidade de conversão

- :: Existem setores nos quais padrões técnicos inexistem ou são conflitantes

::: aplicação dos padrões

[illegible]

A partir de de janeiro de 2007 os órgãos do governo dos Estados Unidos só poderão produzir e aceitar documentos eletrônicos em formatos não-proprietários ou de padrões abertos

PDF
TIFF
JPEG

ODF – Open Document Format for Office Applications, baseado no XML

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Metadados

Independente da estratégia adotada, a preservação de longo prazo envolve a criação e manutenção de metadados

(Metadados para preservação e uma área ativa de pesquisa e desenvolvimento na comunidade internacional de arquivos e bibliotecas digitais)



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Metadado: o que é

A NISO – National Information Organization – entende metadados como:

“ Informação estruturada que descreve, explica, localiza, ou ainda possibilita que um recurso informacional seja fácil de recuperar, usar ou gerenciar. O termo metadados freqüentemente designa dados sobre dados, ou informação sobre informação. ”

Metadados: funções

1

a descoberta de recursos - que permite que recursos sejam identificados, localizados, selecionados por critérios de relevância e distinguidos por diferenças e similaridades;

2

a organização de recursos;

3

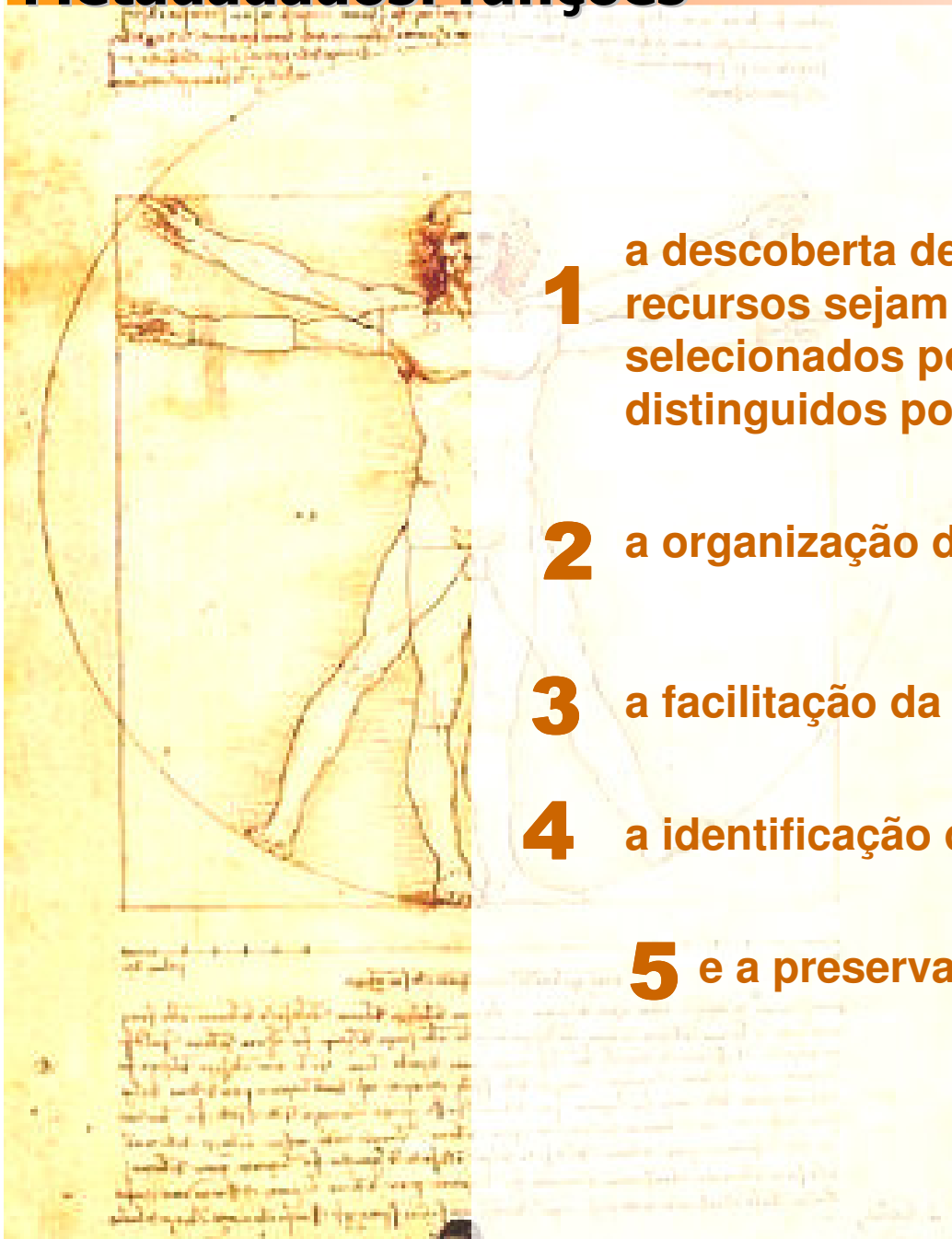
a facilitação da interoperabilidade;

4

a identificação digital;

5

e a preservação digital.



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Metadados: tipos

Descritivos

facilita a descoberta, identificação e seleção; facilita a identificação do assunto e o conteúdo do objeto

Administrativos

ajuda ao administrador da coleção na gestão de arquivos, pode incluir metadados sobre gestão de direitos; metadados de preservação e metadados técnicos que descrevem características físicas do objeto

Estruturais

descreve a estrutura interna de objetos complexos; documenta ao relacionamento entre objetos: artigos, itens, volumes, páginas e capítulos de livros

No espaço da Internet, metadados devem ser legíveis tanto por pessoas como também por programas computacionais.

Programas podem ajudar a organizar a Internet e tornar mais fácil a atividade de encontrar a informação relevante

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Metadados

The SciELO Brazilian Scientific Journal Gateway and O...

Home Feeds Imprimir

The SciELO Brazilian Scientific Journal Gateway and Open Archives A Repor

Arquivo Editar Formatar Exibir Ajuda

Search | Back Issues | Author Index | Title Index | Contents

ARTICLES

D-Lib Magazine
March 2003

Volume 9 Number 3

ISSN 1082-9873

Os metadados podem estar embutidos num objeto digital – inscritos na sua codificação, como é comum nos documentos HTML; ou podem estar armazenados separadamente, estruturados em base de dados e vinculados ao objeto que eles descrevem, facilitando a busca e a recuperação.

The SciELO Brazilian Scientific Journal Gateway and Open Archives

A Report on the Development of the SciELO-Open Archives Data Provider Server

[Carlos Henrique Marcondes](#)

Information Science Department, Federal Fluminense University, Brazil

marcon@vm.uff.br

[Luis Fernando Savão](#)

Nuclear Information Center, Nuclear Energy National Commission, Brazil

lsayao@cnen.gov.br

Abstract

SciELO, the Scientific Electronic Library Online, uses a methodology developed by BIREME/PAHO/WHO [1] that enables implementation of web digital libraries of scientific journal collections of full text articles. Various SciELO gateways are now providing access to academic journals from Brazil and other countries in Latin America and the Caribbean. SciELO plays an important role in the worldwide dissemination of the technical and scientific literature published in developing countries, thus increasing the visibility of this literature that otherwise would be accessible only within the borders of those developing countries.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
<!-- saved from url=(0059)http://www.dlib.org/dlib/march03/marcondes/03marcondes.meta.xml" rel=metadata><LINK title="Default"
href="The SciELO Brazilian Scientific Journal Gateway and Open Archives"
type=text/css rel=stylesheet>
<STYLE type=text/css></STYLE>
```

```
<META NAME = "dc.creator" CONTENT = "MARCONDES, Carlos Henrique"
<META NAME = "dc.creator" CONTENT = "Savão, Luis Fernando"
<META NAME = "dc.title" CONTENT = "The SciELO Brazilian Scientific Journal Gateway and Open Archives"
<META NAME = "dc.subject" CONTENT = "Digital Libraries"
<META NAME = "dc.subject" CONTENT = "Brazil"
<META NAME = "dc.subject" CONTENT = "SciELO"
<META NAME = "dc.subject" CONTENT = "Open Archives Protocol"
<META NAME = "dc.subject" CONTENT = "Interoperability"
<META NAME = "dc.publisher" CONTENT = "DLib Magazine, ISSN 1082-9873"
<META NAME = "dc.date" CONTENT = "2003"
<META NAME = "dc.language" CONTENT = "en"
<META NAME = "dc.identifier" CONTENT = "DOI: 10.1045/march2003marcondes"
<META NAME = "dc.type" CONTENT = "Journal Article"
<META NAME = "dc.format" CONTENT = "HTML"
<META NAME = "dc.description" CONTENT = "SciELO, the Scientific Electronic Library Online, reports on the results of a pilot-project in which a general
```

```
</HEAD>
<BODY bgcolor=#ffffff>
<DIV class=center>
<TABLE cellSpacing=0 cellPadding=0 width=700 align=center border=1>
<TR>
<TD class=center>
<P><FONT style="FONT-WEIGHT: bold"><A class=menu href="http://www.dlib.org/Archivtext/AT-dlib2query.html" target=_top>Search</A> &nbsp; <A name=Top|</A> &nbsp; <A href="http://www.dlib.org/back.html" target=_top>Back</A> &nbsp; <A class=menu href="http://www.dlib.org/author-index.html" target=_top>Author Index</A> &nbsp; <A class=menu href="http://www.dlib.org/title-index.html" target=_top>Title Index</A> &nbsp; <A class=menu href="http://www.dlib.org/dlib/march03/03contents.html" target=_top>Contents</A> </FONT></P></TD></TR></TBODY>
<P align=center><IMG height=16 alt=Articles<br>The SciELO Brazilian Scientific Journal Gateway and Open Archives<br>width=500> </P></DIV><!-- Begin Article Header -->
<TABLE cellSpacing=0 cellPadding=0 width=100% border=0>
<COLGROUP>
<COL width="6%">
```

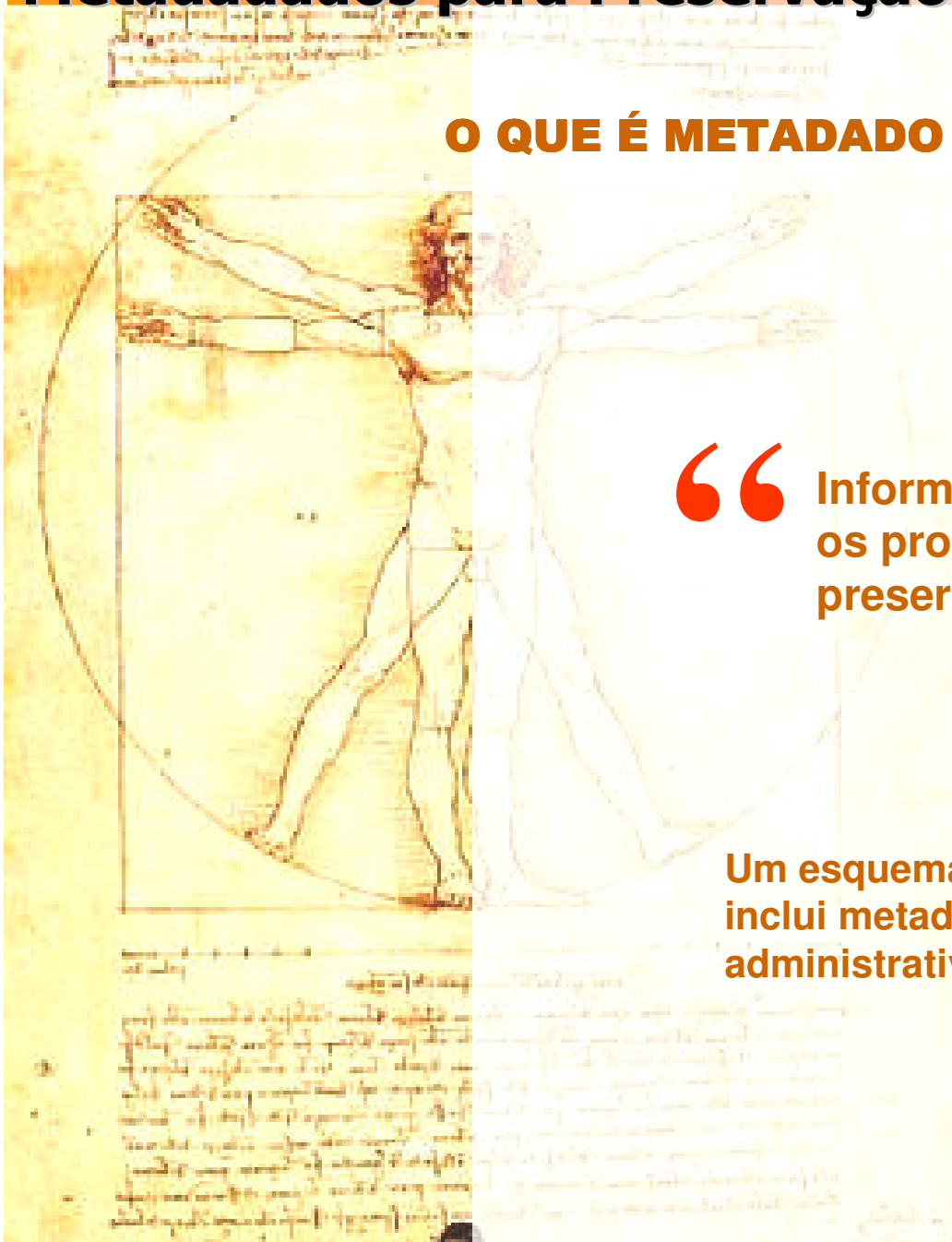
001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Metadados para Preservação

O QUE É METADADO PARA PRESERVAÇÃO DIGITAL

“ Informação que apóia e documenta os processos associados com a preservação digital de longo prazo ”

Um esquema de metadados de preservação digital inclui metadados descritivos, estruturais e administrativos



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Metadados para Preservação

PODE SER USADO PARA

Armazenar informações técnicas e administrativas sobre decisões e ações de preservação

Documentar as ações de preservação tomadas, como políticas de migração ou de emulação

Registrar os efeitos das estratégias de preservação

Assegurar a autenticidade dos recursos digitais ao longo do tempo

Registrar informações sobre gestão de coleções e de direitos

Dois objetivos funcionais:

Preservar a Cadeia de BITS

assistir os gerentes de preservação com conhecimento suficiente para tomar as decisões corretas no sentido de preservar a cadeia de bits de um objeto por longo prazo;

Garantir a Interpretação

assegurar que o conteúdo de um objeto arquivado possa ser apresentado e interpretado, mesmo diante de mudanças tecnológicas

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Metadados para Preservação

HÁ MUITA DISCUSSÃO SOBRE QUE TIPO DE INFORMAÇÃO É NECESSÁRIA PARA APOIAR OS PROCESSOS DE PRESERVAÇÃO DIGITAL...O CONSENSO SE CONCENTRA EM 5 ÁREAS RELEVANTES

Proveniência

Autenticidade

Atividades de Preservação

Ambiente Técnico

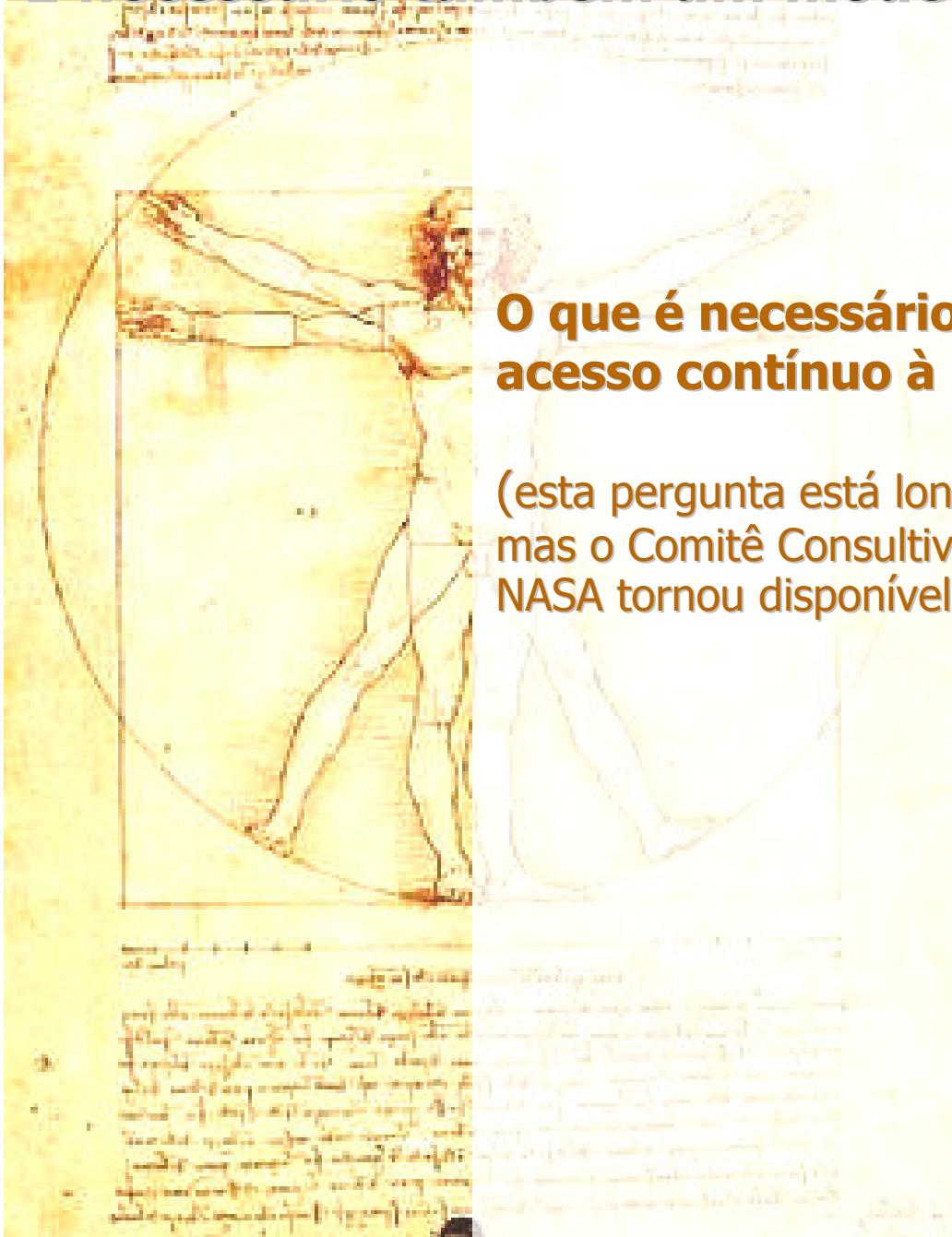
Gestão de Direitos

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

É necessário também um modelo de gestão

O que é necessário para preservar e manter o acesso contínuo à informação digital?

(esta pergunta está longe de ser satisfatoriamente respondida, mas o Comitê Consultivo para Sistemas de Dados Espaciais da NASA tornou disponível um patamar comum para discussões)



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

O que é o Modelo de Referência OAIS?

O Modelo Referencial OAIS – Open Archival Information System – é uma estrutura conceitual que disciplina e orienta um sistema de arquivo dedicado a preservar e manter o acesso a informação digital por longo prazo



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Como é definido um OAIS segundo a Norma ISO



Um OAIS consiste numa organização de pessoas e sistemas que aceitaram a responsabilidade de preservar informação e torná-la disponível para uma **Comunidade Alvo**

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Quais são os objetivos de um OAIS?

O objetivo do modelo de referência é aumentar o grau de consciência e compreensão dos conceitos relevantes para o arquivamento de objetos digitais, especialmente entre instituições não arquivísticas.

Ampliar a consciência e a compreensão dos conceitos arquivísticos relevantes necessários para a preservação e o acesso de longo prazo

Aprover conceitos necessários para organizações não arquivísticas participarem efetivamente no processo de preservação

Permitir comparações: arquitetura, operação, estratégias e técnicas de preservação

Estabelecer fundamentos que possam ser expandidos para informações que não estão no formato digital

Ampliar o consenso em torno dos elementos e processos voltados para a preservação de longo prazo e acesso a informação digital, promover a expansão do mercado e o apoio dos fornecedores

Orientar a identificação e produção de padrões relacionados ao OAIS

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Ambiente OAIS

O Modelo de Referência OAIS é uma **infra-estrutura conceitual** para um arquivo digital. O Modelo estabelece **terminologia e conceitos** relevantes para o arquivamento digital; identifica os **componentes chave e processos** inerentes a maioria das atividades de arquivo, e propõe um **modelo de informação** para objeto digitais e seus metadados associados.

Não especifica implementação

É **neutro** em relação aos objetos digitais e questões tecnológicas envolvidas

PRODUTORES

Pessoas ou sistemas clientes que fornecem a informação a ser preservada

ADMINISTRAÇÃO

Entidade responsável pelo estabelecimento de políticas mais gerais do arquivo.

CONSUMIDORES

Pessoas ou sistemas clientes que interagem com os serviços do OAIS para recuperar e adquirir informações preservadas de interesse. Uma classe especial de Consumidores é a Comunidade Alvo (Designated Community).

O Ambiente OAIS



00110001110001101010001011000101110100101100101001111010100100010100001110001001010001000101

Modelo de Informação do OAIS

<INTERPETAÇÃO DO OBJETO DE DADOS>

:: A interpretação do Objeto de Dados como informação significativa pela Comunidade Alvo é obtida por meio da combinação da **Base de Conhecimento** dessa comunidade e a **Informação de Representação** associado ao **Objeto de Dados**

:: Podemos concluir que de modo geral que “Dado interpretado usando sua Informação de Reapresentação produz Informação (**Objeto de Informação**)”



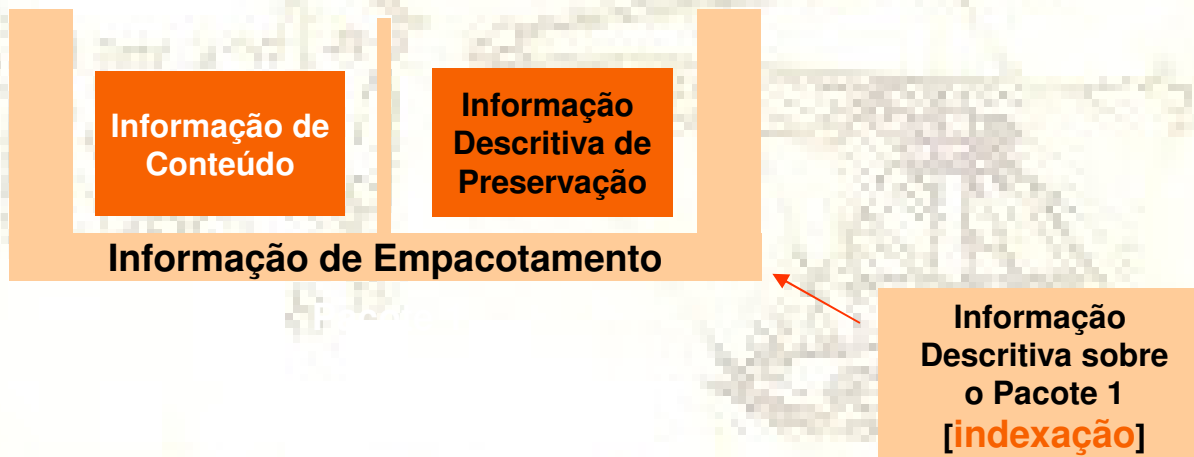
:: Para um Objeto de Informação ser preservado com sucesso é crítico para o OAIS claramente identificar e compreender o **Objeto de Dados** e a sua Informação de Representação

:: Isto é um desafio importantíssimo para a preservação digital !!!



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

Modelo de Informação do OAIS



Informação de Conteúdo

[consiste de dados de interesse primário – i.e., a informação que o arquivo tem a missão de preservar – associado com a informação de representação]

Informação Descritiva de Preservação

[contém informações necessárias para gerenciar a preservação da informação de conteúdo]

- Informação de Referência
- Informação de Proveniência
- Informação de Contexto
- Informação de Estabilidade

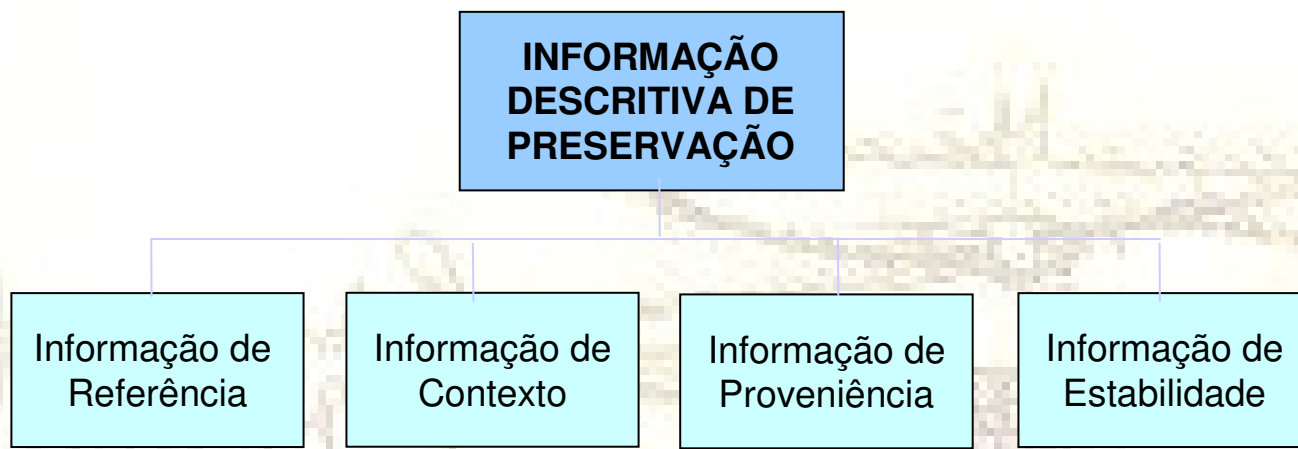
Informação de Empacotamento

[liga o objeto digital e seus metadados com uma unidade identificável ou pacote, i.e., um pacote de informação de arquivamento]

Informação descritiva

[facilita o acesso à informação de conteúdo através de ferramenta de busca e recuperação]

Modelo de Informação do OAIS



Em resumo, a Informação Descritiva de Preservação registra a identidade, os relacionamentos, a história e a Integridade do Objeto de Dados de Conteúdo arquivado

Referência *Identidade*

[descreve ao sistema de identificação e os mecanismos para assinalar identificadores usados para identificar a Informação de Conteúdo sem ambigüidades, interna e externamente ao arquivo]

Contexto *relacionamentos*

[documenta o relacionamento da Informação de Conteúdo com o seu ambiente, incluindo a razão de criação e a relação com outros objetos de Informação de Conteúdo]

Proveniência *história*

[Documenta a história da Informação de Conteúdo, incluindo sua origem, mudanças no tempo no objeto ou no seu conteúdo, e sua cadeia de custódia]

Estabilidade *integridade*

[Provê verificação da Integridade de Dados ou chaves para Validação/ Verificação usadas para assegurar que um objeto de Informação de Conteúdo particular não foi alterado de forma não documentada]

Modelo de Informação do OAIS

Estrutura Completa da Informação Descritiva de Informação



:: Política mínima de preservação digital

Reconhecido o caráter efêmero da informação digital é necessário:

1 criar informação para durar

:: Criar informação digital com **qualidade**

:: A capacidade de preservar a informação digital depende fortemente na maneira como ela foi criada originalmente, como ela esta organizada, controlada e quão bem **documentada** ela está

:: Muitas decisões importantes que sustentam a preservação de longo prazo foram tomadas quando o sistema foi planejada e **antes da criação** do primeiro registro

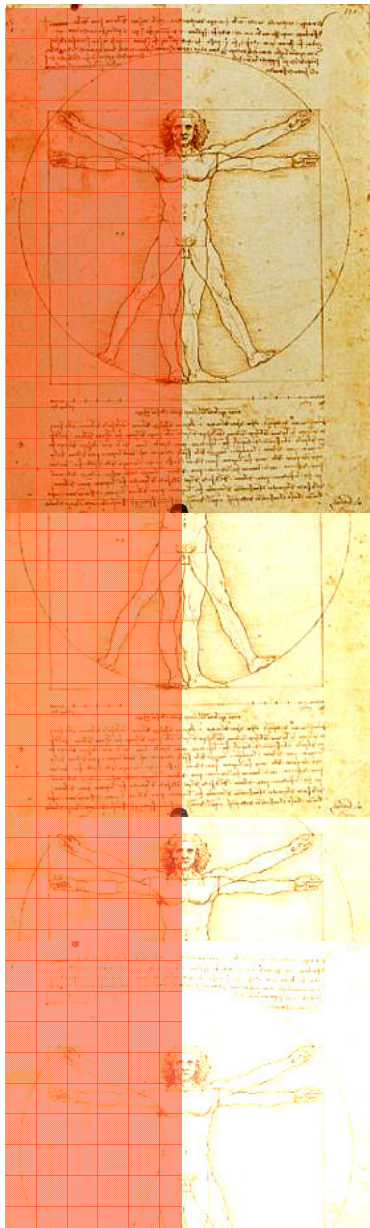
:: Profissionais de informação – bibliotecários e arquivistas - devem estar sempre na gerencia dos projetos de serviços e produtos de informação

:: Considerações sobre valor de longo prazo da informação devem ser levado em conta nos projetos de serviços e produtos de informação

:: Identificar as responsabilidades das instituições arquivísticas

:: Adotar diretrizes para a aquisição de material digital, definir **formatos aceitáveis**

:: **Descrever** de forma completa os objetos digitais



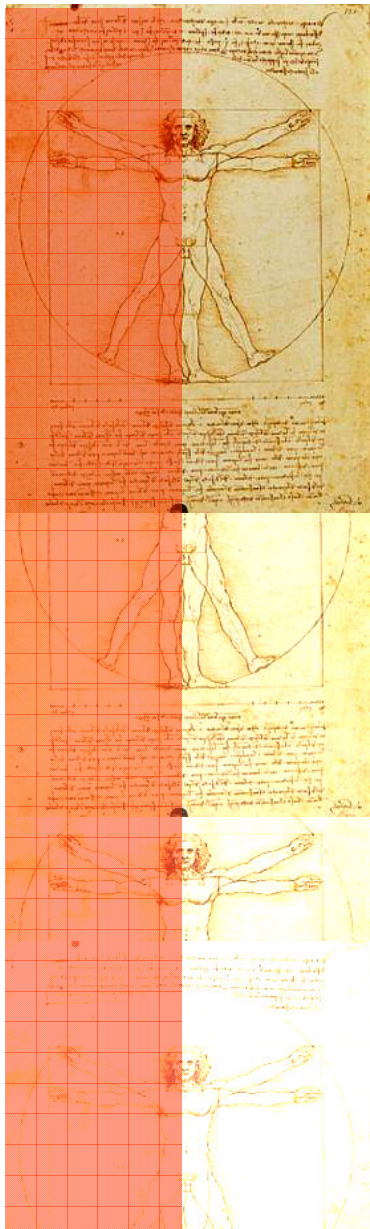
001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Política mínima de preservação digital

Reconhecido o caráter efêmero da informação digital é necessário:

2 padrões & protocolos

- ::: Um dos desafios importantes é criar **diretrizes, políticas e normas** que o governo, empresas, C & T & I possam usar para assegurar a longevidade dos registros digitais
- ::: Promover o uso de **padrões e protocolos abertos**, estáveis e de uso amplo
- ::: Adotar padrões na criação, armazenamento e transmissão de documentos digitais
- ::: Monitorar o surgimento de **novos padrões** e migrar quando necessário
- ::: Arquivos e bibliotecas digitais devem manter-se **à frente do desenvolvimento de padrões** e garantir que a sua própria estrutura estejam em conformação com os padrões largamente adotados



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Política mínima de preservação digital

Reconhecido o caráter efêmero da informação digital é necessário:

3 gestão documental

- ::: Aplicar procedimentos e estratégias de **gestão documental** quando da criação, tratamento, transmissão e preservação de documentos em formatos digitais
- ::: Desenvolver **planejamento de longo prazo**
- ::: Definir critérios para **seleção do patrimônio digital** , pois nem tudo poderá ser salvo
- ::: Desenvolver modelos de **custo** e de minimização de **riscos** das conversões
- ::: Uso e o estabelecimento de **melhores práticas**; conhecer as experiências, manuais de procedimentos, especificações, requisitos, etc; desenvolver manuais
- ::: Para ajudar os **profissionais de informação do futuro**, precisamos de métodos para organizar a informação digital hoje



:: Política mínima de preservação digital

Reconhecido o caráter efêmero da informação digital é necessário:

4 tecnologia

::: Aplicar as **técnicas apropriadas** de preservação digital respeitando as especificidades de cada problema, pois não há até agora solução genérica ou solução que funcione em grande escala

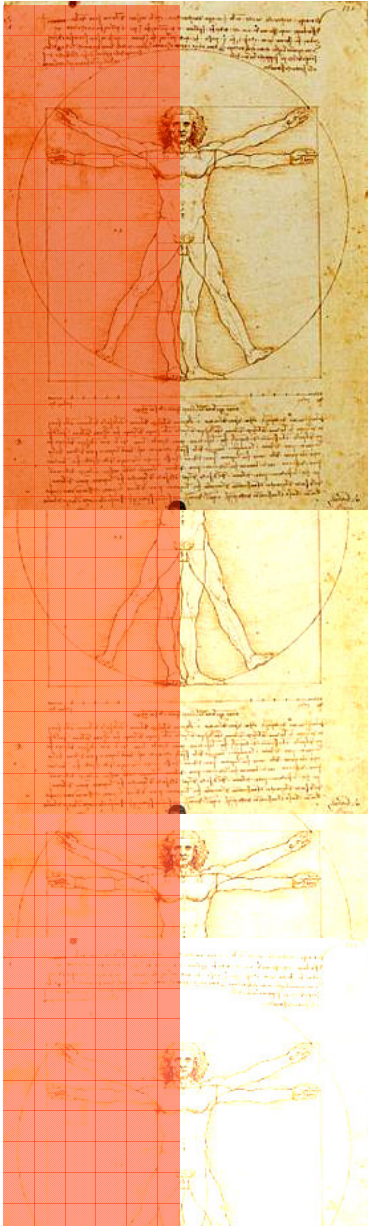
::: Cada instituição deve desenvolver **seu próprio plano de migração** baseado:

- .. na natureza do material digital adquirido/gerenciado
- .. nas necessidades, expectativas e habilidades do usuário
- .. na capacidade técnica interna

::: Desenvolver **soluções abertas** em cooperação com a indústria de TI

::: Monitorar as inovações tecnológicas para a área e **validar as novas tecnologias** quando apropriado

::: Manutenção e migração são muito **mais complicadas** quando a organização não usa padrões e métodos de gestão



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Política mínima de preservação digital

Reconhecido o caráter efêmero da informação digital é necessário:

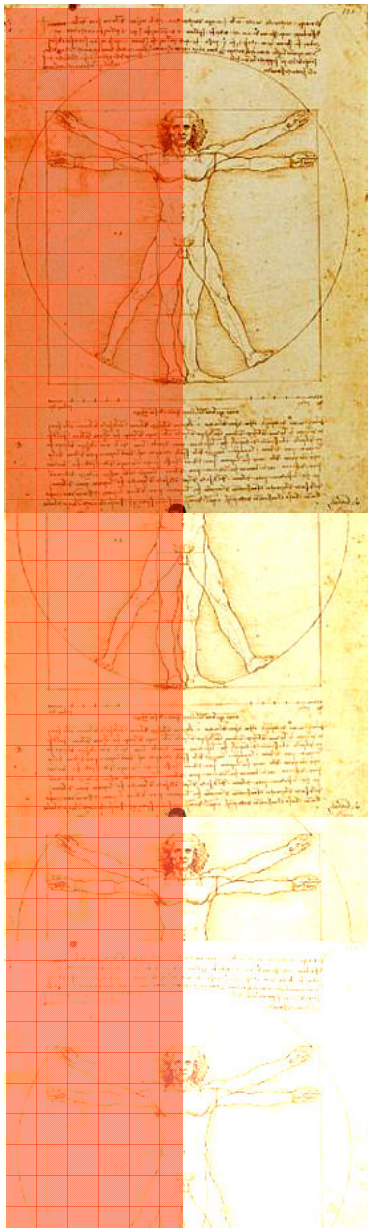
5 segurança digital

Apoiar o uso de instrumentos oriundos da tecnologia da informação e de ampla aceitação que garantam a integridade, a confiabilidade, a privacidade e a autenticidade dos documentos digitais, e que os proteja contra acidentes e intervenções não autorizadas

::: criptografia, assinatura digital, certificação digital;...

6 agenda de pesquisa

Desenvolver uma **agenda nacional de pesquisa** para a preservação e longevidade digital alinhada com as principais iniciativas internacionais



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Política mínima de preservação digital

Reconhecido o caráter efêmero da informação digital é necessário:

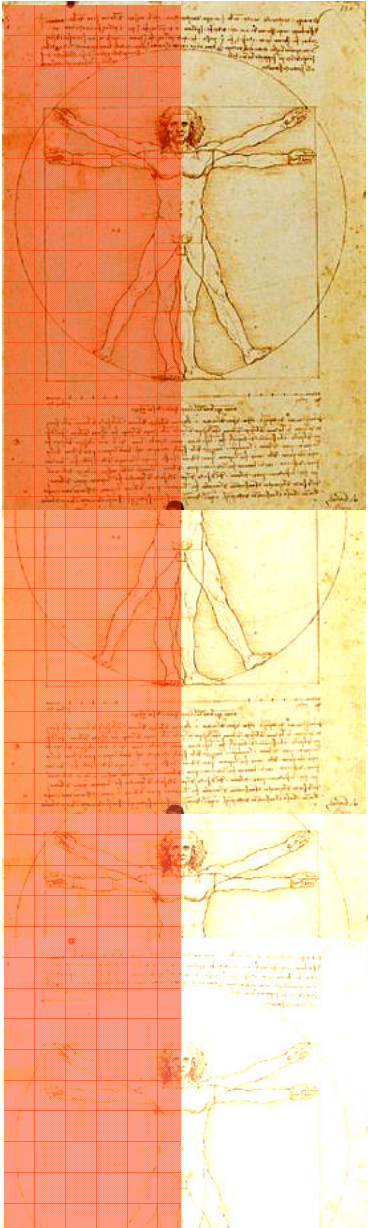
7 legislação

Disponer de um **corpo de leis** que garanta a proteção do patrimônio digital e o seu reconhecimento como valor de prova

::: depósito legal, direitos autorais...

8 arquivos e bibliotecas nacionais

Fortalecer e instrumentalizar – em termos de equipamentos, metodologias e pessoal especializado - os arquivos e bibliotecas para que possam desempenhar um papel ativo e de liderança na gestão da preservação digital dos documentos, principalmente os sob suas custódias



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Política mínima de preservação digital

Reconhecido o caráter efêmero da informação digital é necessário:

9 políticas públicas

Definir estratégias e **políticas públicas** para a preservação do patrimônio digital digital

::: alocação de recursos

::: diretrizes e ordenamento

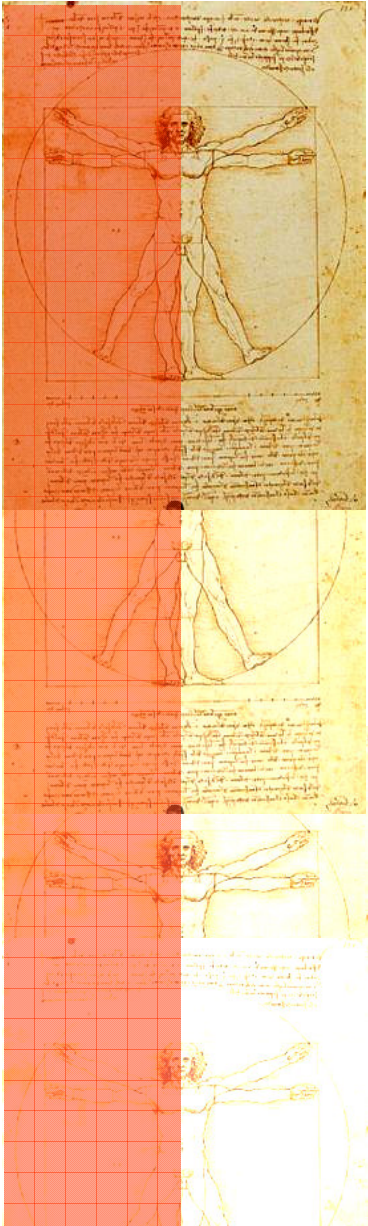
::: conscientização da sociedade

10 recursos humanos

::: **educação e treinamento** para a preservação digital

::: inclusão nos currículos das escolas de arquivologia, biblioteconomia e informática

::: pós-graduação



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: Política mínima de preservação digital

Reconhecido o caráter efêmero da informação digital é necessário:

11 alianças e cooperação

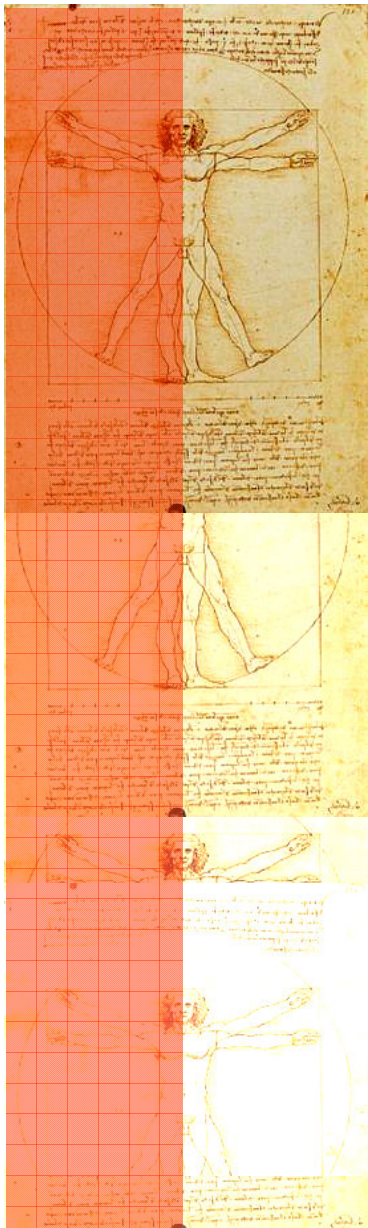
Estabelecer pactos de **cooperação** entre governo, editores, indústria de TI, bibliotecas, arquivos, museus, universidade, instituto de pesquisa...

12 metadados

Incentivar o uso de estruturas padronizadas de **metadados** orientadas para a gestão da preservação digital e para a acessibilidade dos documentos digitais

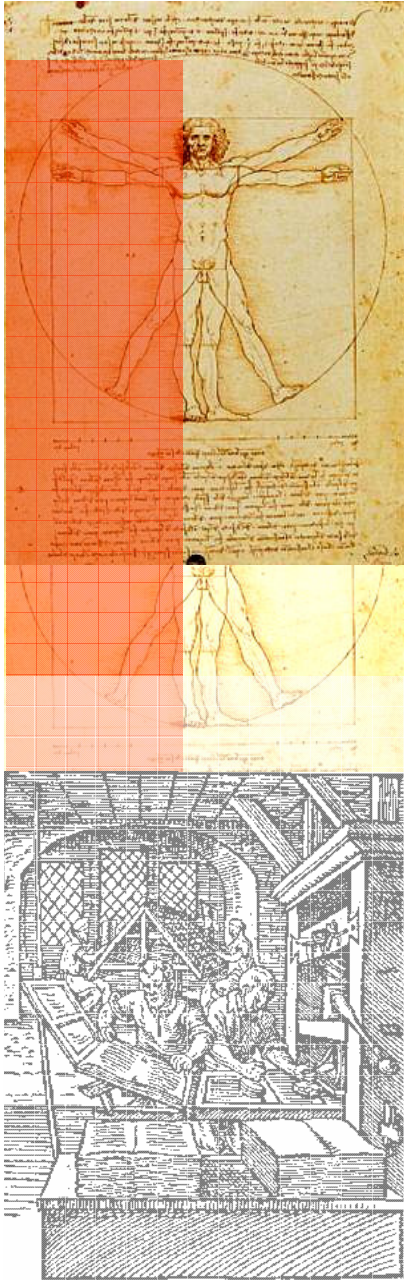
13 acesso

Por fim o mais importante: garantir acesso aos estoques de informações **digitais** de valor contínuo às suas respectivas comunidades alvo. Não menos importante, garantir que essa informação – enquanto válida – permaneça compreensível para esta comunidade



001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: À guisa de conclusão



A INFORMAÇÃO DIGITAL DURA PARA
SEMPRE – OU CINCO ANOS, O QUE
VIER PRIMEIRO

Jeff Rothemberg

OBRIGADO!

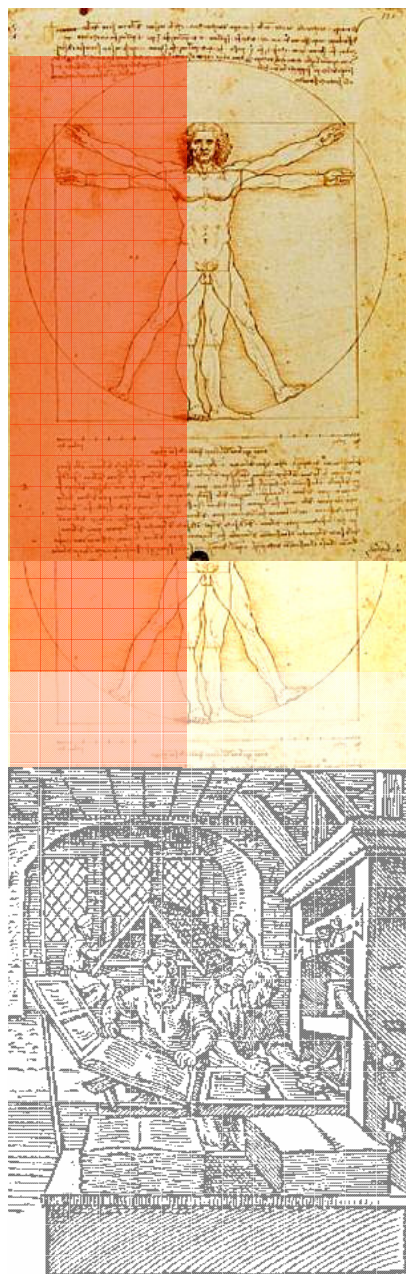
Luis Sayão

lsayao@cnen.gov.br

001100011100011010100010110001011101001011001010011111010100100010100001110001001010001000101

:: À guisa de conclusão

<http://www.arquivonacional.gov.br/>



BRASIL.GOV

Conselho Nacional de Arquivos

Conarq

Câmara Técnica de Documentos Eletrônicos

O QUE É

INTEGRANTES

PRODUÇÃO DA CTDE

LEGISLAÇÃO

BIBLIOGRAFIA

EVENTOS

SÍTIOS DE INTERESSE

PERGUNTAS MAIS FREQUENTES

ATUALIZADO EM
23/03/04

5 6 8 9 0

←
Outras Câmaras Técnicas

Conarq

Arquivo Nacional

PRODUÇÃO DA CTDE

ATAS

GESTÃO

PRESERVAÇÃO

PRESERVAÇÃO DE DOCUMENTOS ELETRÔNICOS

CARTA PARA PRESERVAÇÃO DO PATRIMÔNIO ARQUIVÍSTICO DIGITAL

O Conselho Nacional de Arquivos apresenta a Carta de Preservação do Patrimônio Arquivístico Digital. Este documento tem o objetivo de conscientizar e ampliar a discussão sobre a instabilidade do legado digital, que se encontra em perigo de perda e de falta de confiabilidade. Manifesta ainda a necessidade de estabelecer políticas, estratégias e ações, que garantam a preservação de longo prazo e o acesso contínuo aos documentos arquivísticos digitais.

Considerando que, a Carta de Preservação do Patrimônio Digital da UNESCO alerta sobre a possibilidade de desaparecimento do legado digital e sugere que os estados membros adotem um conjunto de medidas para salvaguardar esse patrimônio, o CONARQ está propondo um documento que atenda às especificidades do documento arquivístico digital, fonte de prova e de informação, que necessita de proteção especial.

O CONARQ conclama a comunidade arquivística e a todos os interessados no tema a adotarem os princípios e as propostas apontadas no documento.

- Carta de Preservação do Patrimônio Arquivístico Digital [Download: Formato PDF \(tamanho: 94 kb\)](#)
[Download: Formato RTF \(tamanho: 98 kb\)](#)
- Documentos de referência:

Carta para la Preservación del Patrimonio Digital - UNESCO, outubro 2003

[Download: ESPANHOL Formato PDF \(tamanho: 18,3 kb\)](#)

[Download: INGLÊS Formato RTF \(tamanho: 29,9kb\)](#)

:: Projetos internacionais

Austrália

- VERS – Victorian Electronic Record Strategy – padrões para gestão e preservação de registros eletrônicos
- PANDORA – Preserving and Acessing Networked Documentary Resources of Australia
Projeto da National Library of Australia, cujo principal objetivo é capturar, arquivar e proporcionar preservação de longo termo de publicações eletrônicas importantes; preocupa-se com migração HTML.

CEDARS

- liderado pelo Consortium of University Research Libraries – UK e Irlanda
- Projeto baseado no Modelo **OAIS**;
- guidelines para desenvolvimento de políticas de gestão de coleções digitais e preservação
- projetos-pilotos: migração, emulação, e “refreshing” – comparações

CAMiLEON

- JISC - Joint Information Systems Comittee (UK) e NSC - National Science Foundadion (USA)
- emulação comparado com outras metodologias

NEDLIB

- liderado pela National Library of the Netherlands
- tem como objetivo desenvolver uma arquitetura comum e ferramentas básicas para a construção de depósitos para publicações eletrônicas
- segue o modelo de referência do **OAIS**
- experiência com emuladores

:: Projetos internacionais

Kulturarw Heritage

- projeto da Royal Library - Suécia
- testar métodos de coletar, arquivar e proporcionar acesso a documentos eletrônicos suecos
- crawlers e robots são usados para coletar automaticamente páginas suecas

Library of Congress

NARA – National Archives and Records Administration

- desenvolver um arquivo persistente
- preservar todo o contexto
- uso de supercomputadores e computação distribuída

InterPARES – International Research on Permanent Authentic Records Electronic Systems

- projeto multinacional de pesquisa – desenvolver conhecimento teórico e metodologias necessárias para a preservação permanente de registros autênticos

PRISM – Preservation, Realiability, Interoperability, Security, Metadata

- Cornell University
- investigar e desenvolver políticas e mecanismos necessários para integridade de informações em bibliotecas digitais