



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Repositório Digital da Memória Científica do INPE

Gerald Jean Francis Banon

<banon@dpi.inpe.br>

URL do documento original:

<<http://urlib.net/3ERPFQRT3W/3A32NUE>>

Fóruns Permanentes Ciência e Tecnologia
Repositórios Confiáveis de Documentos Arquivísticos Digitais

Unicamp, 11 de agosto de 2011



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

MISSÃO DO INPE

Produzir ciência e tecnologia nas áreas espacial e do ambiente terrestre e oferecer produtos e serviços singulares em benefício do Brasil.

Conteúdo

Recomendações/Normas

Marcos históricos

Perspectivas

Recomendações/Normas

(1/2)

- **CCSDS 650.0-B-1; 2002**
Reference Model for an Open Archival Information System (OAIS).
- **ABNT NBR 15472; 2007**
Modelo de referência para um sistema aberto de arquivamento de informação (SAAI).
- **CCSDS 651.0-M-1; 2004**
Producer-Archive Interface Methodology Abstract Standard.
- **ABNT NBR 15862; 2010**
Interface Produtor-Arquivo – Padrão de metodologia abstrata.

Fórum aberto

CCSDS = *Consultative Committee for Space Data Systems*
(Comitê Consultivo para Sistemas Espaciais de Dados)

Recomendações/Normas

(2/2)

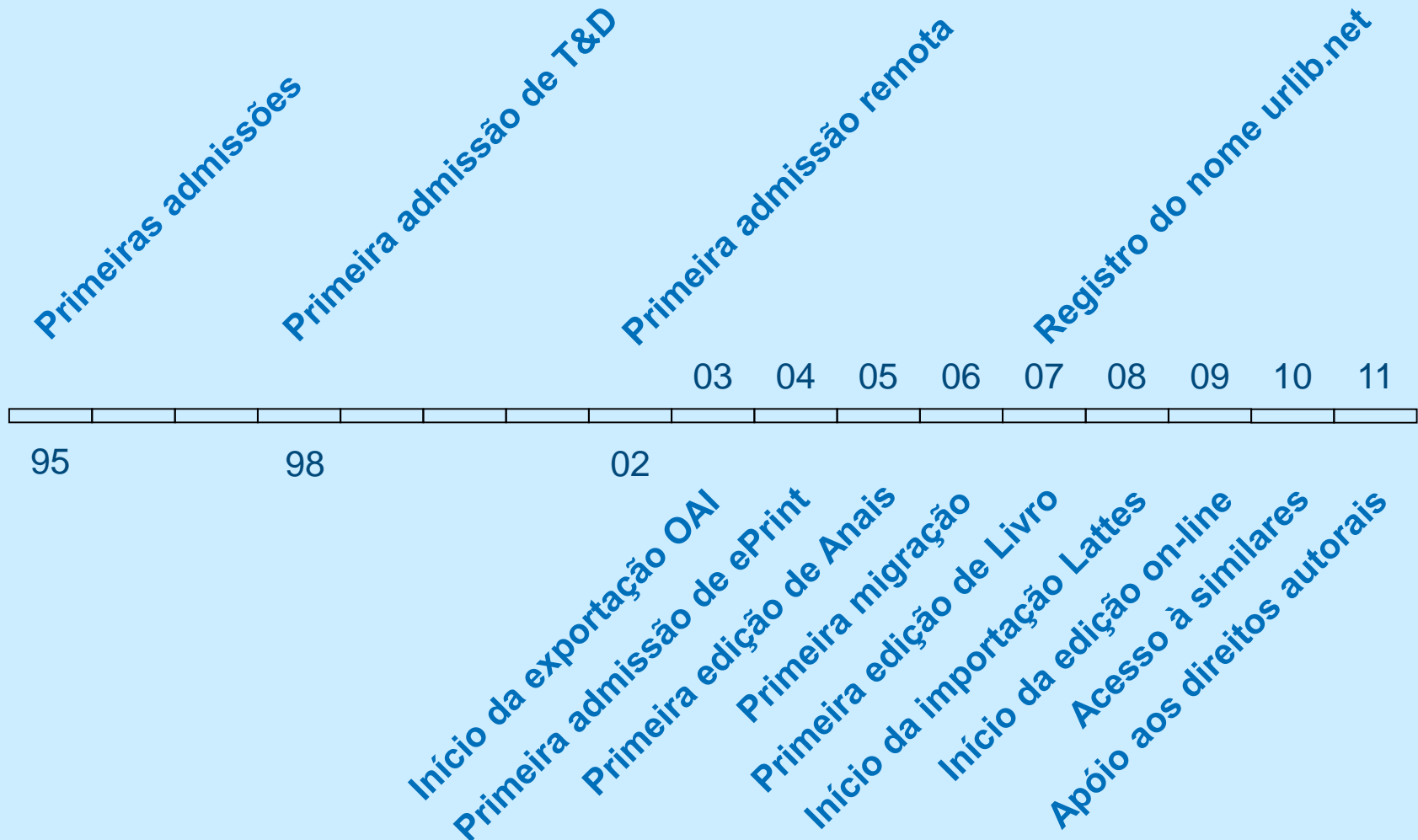
Recomendações em revisão:

- **CCSDS 650.0-P-1.1; 2009**
Reference Model for an Open Archival Information System (OAIS).
- **CCSDS 652.0-R-1; 2009**
*Audit and Certification of **Trustworthy Digital Repositories**.*
- **CCSDS 652.1-R-1; 2010**
Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories.

~ 90 quesitos

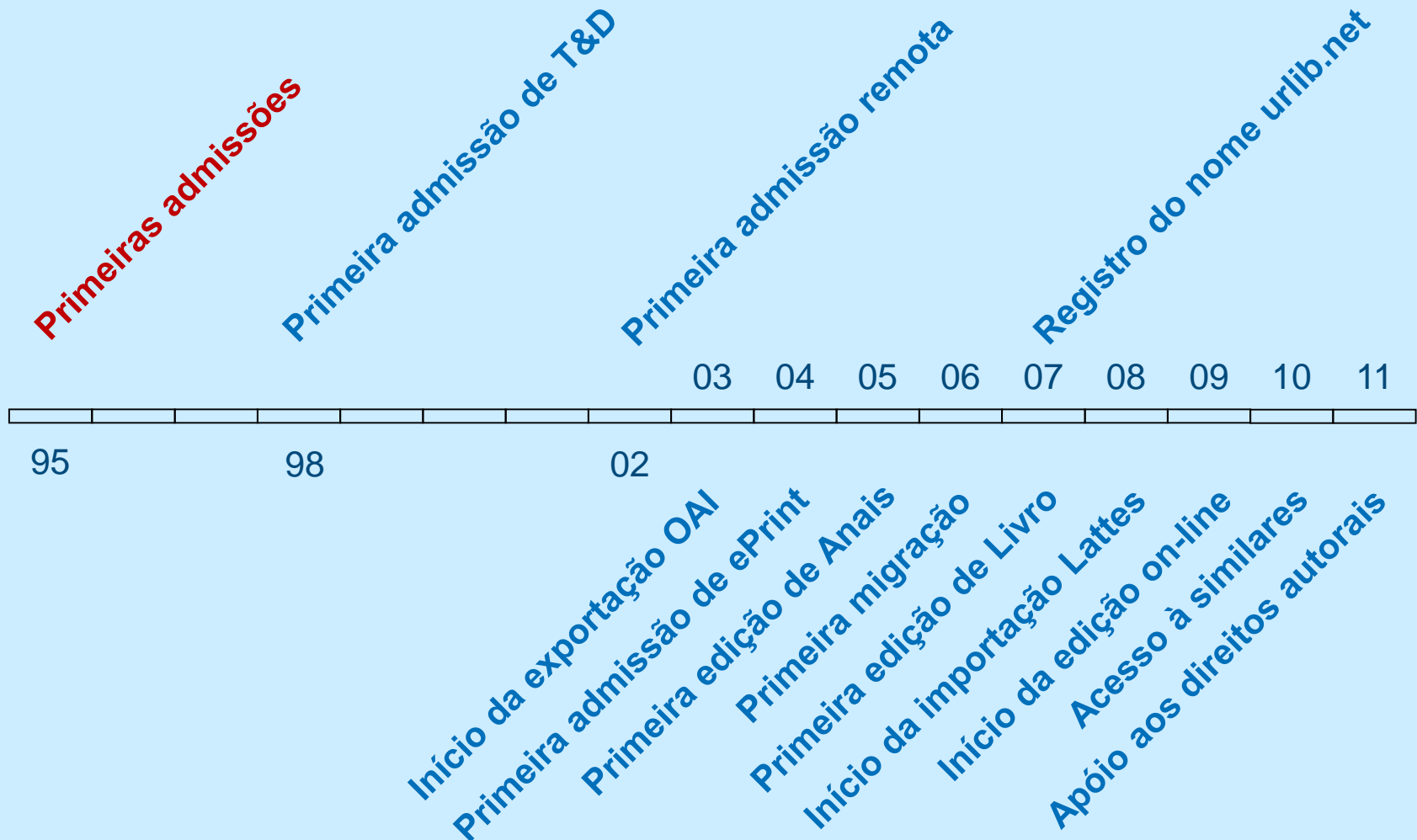
Marcos históricos

(1/1)



Marcos históricos

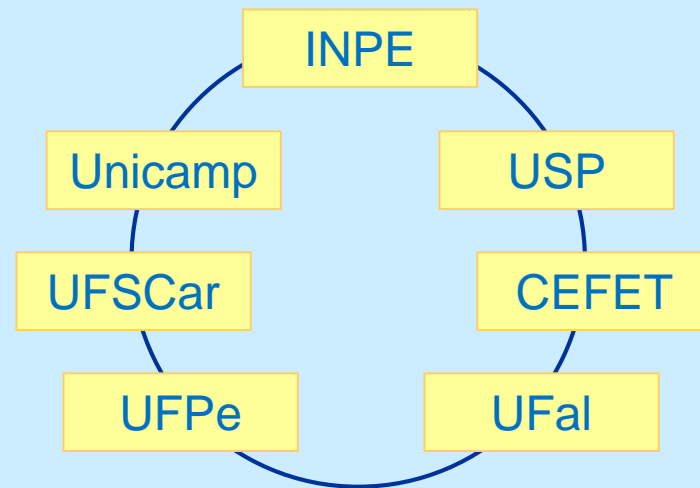
(1/13)



Primeiras admissões (1995)

(1/5)

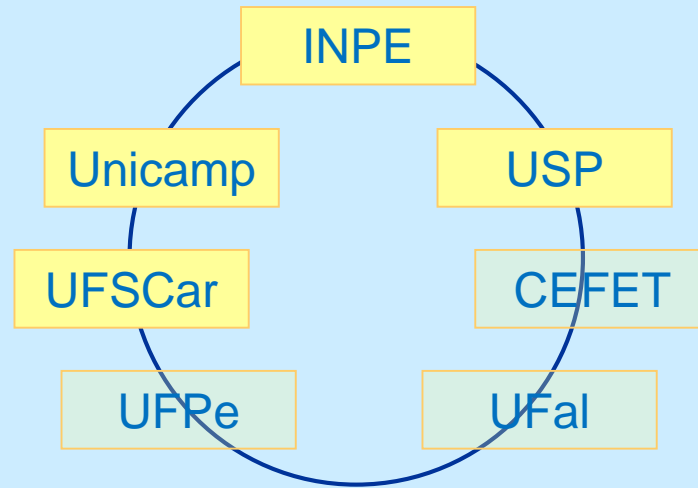
Motivação: criar uma **Federação de Repositórios** com nomes únicos de **PAI** para cada **PAI** da federação, para depositar e preservar as contribuições da 7 instituições envolvidas no projeto AnIMoMat do ProTeM-CC/CNPq.



PAI = Pacote de Armazenamento de Informação

Primeiras admissões

(2/5)



Plataforma para hospedar uma **Federação de Repositórios**

As primeiras **admissões** ocorreram em 1995 na plataforma *URLib*, após a instalação da primeira versão do *URLibService* em 4 instituições.

URLib = Repositório Uniforme para uma Biblioteca

Software para **Repositório** membro de uma **Federação**

Primeiras admissões

(3/5)

IBI = *Internet Based Identifier*
IBI = Identificador com Base na Internet

Exemplos de IBI para nome de PAI

`dpi.inpe.br/banon/1995/09.01.10.50`

`dpi.inpe.br/banon/1995/10.20.17.45`

`ime.usp.br/jb/1996/03.27.18.32`

`dca.fee.unicamp.br/lotufo/1996/04.18.16.42`

`dc.ufscar.br/jander/1996/04.29.10.48`

Identifica um computador

prefixo

sufixo

Identifica um item no computador

Primeiras admissões

(4/5)

Federação
é
Repositório

O *URLibService* atribui à cada **PAI** da **Federação** dois IBIs.

Exemplo:

foram atribuídos ao **PAI** contendo esta apresentação os IBIs:

```
urlib.net/www/2011/07.06.23.08
```

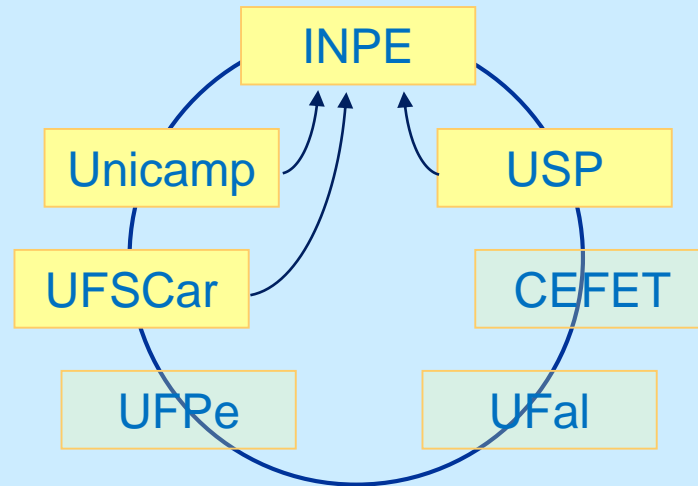
```
3ERPFQRT3W/3A32NUE
```

O Repositório deve identificar cada PAI dentro do Repositório.

(CCSDS 652.0 - 4.2.4.1)

Primeiras admissões

(5/5)



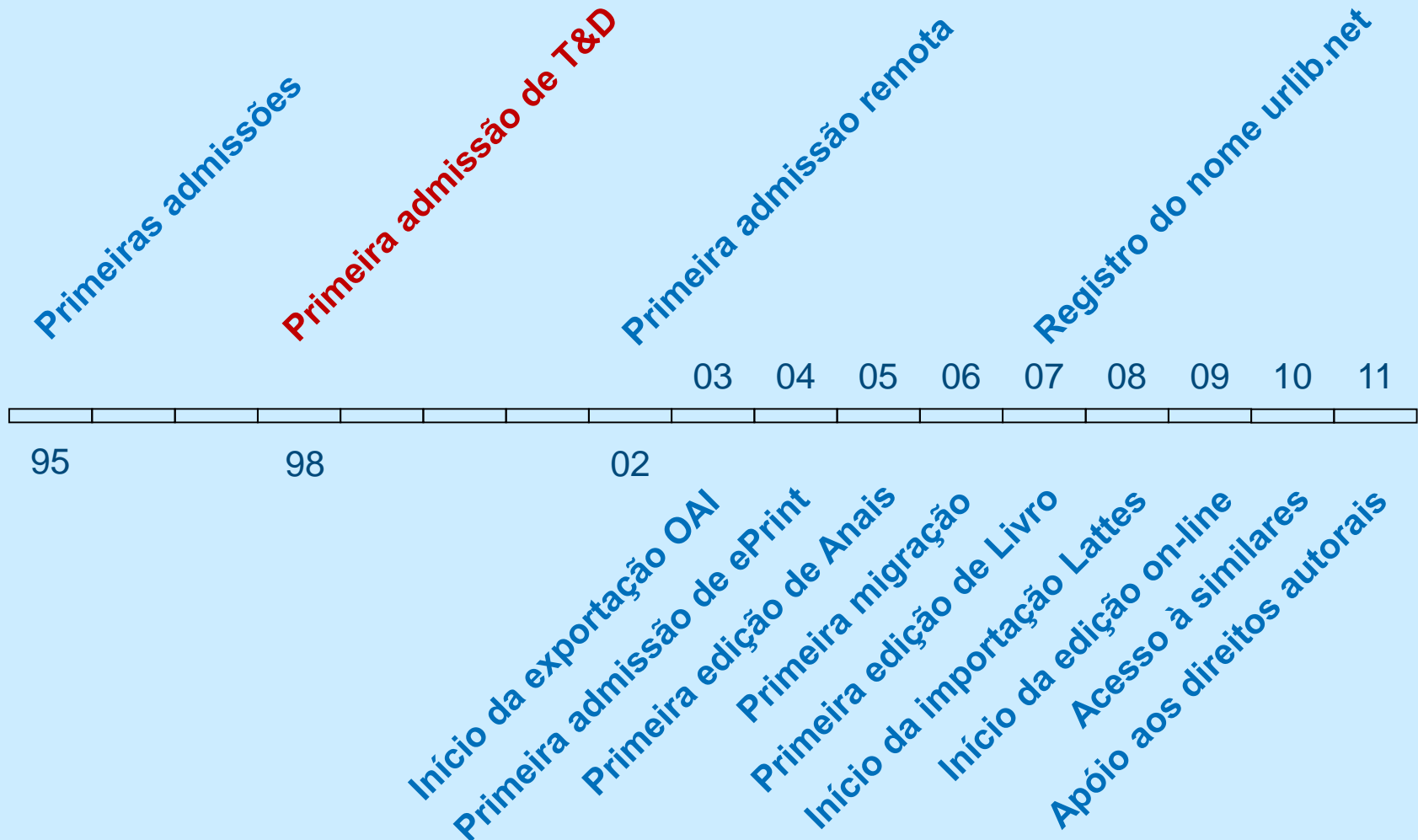
Em 1998, no término do projeto AnIMoMat, os **PAIs** foram preservados por meio de **Migrações Digitais** para o repositório do INPE e os demais repositórios foram encerrados.

O repositório do INPE se tornou o:

Repositório Digital da Memória Científica do INPE.

Marcos históricos

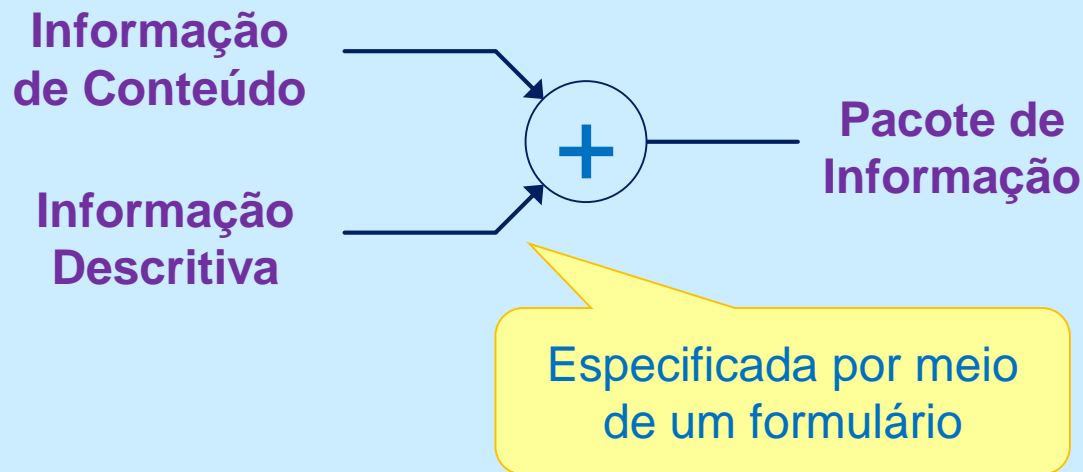
(2/13)



Primeira admissão de T&D (1998)

(1/2)

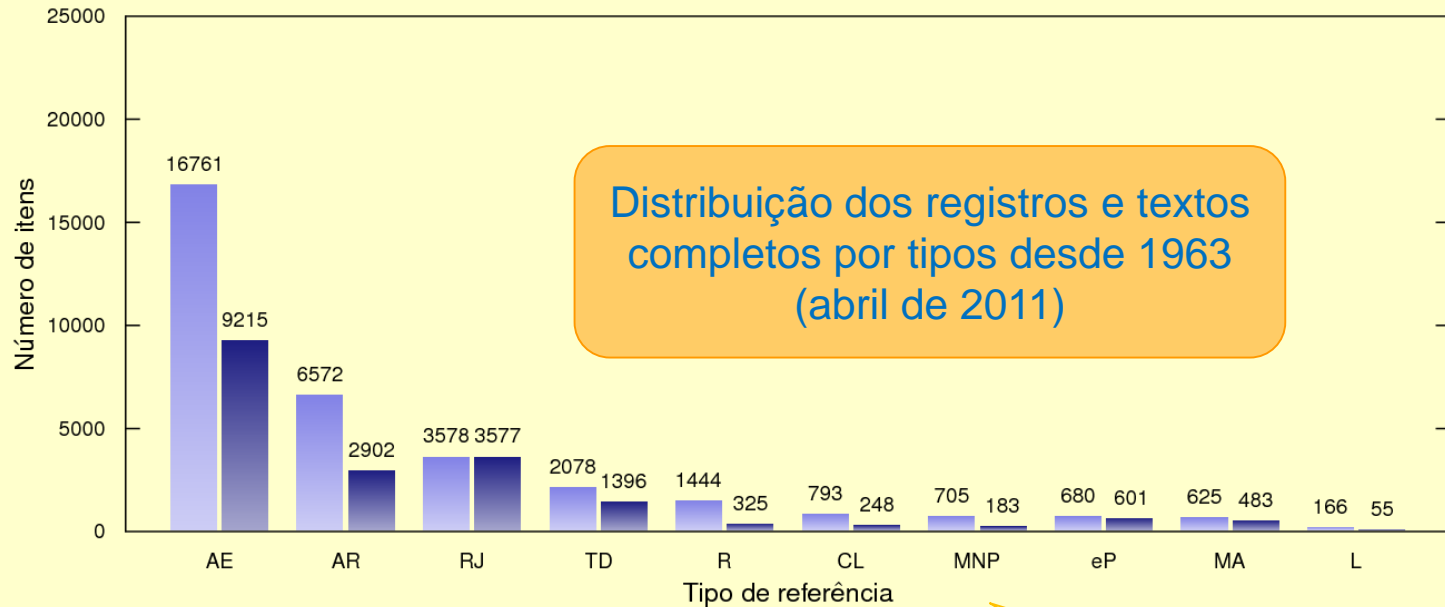
Motivação: tornar o acervo existente da produção científica da Instituição um acervo digital indexado e acessível *on-line*.



O Repositório deve especificar claramente a informação que precisa ser associada à Informação de Conteúdo específica no momento do seu depósito. (CCSDS 652.0 - 4.1.2)

Primeira admissão de T&D (1998)

(2/2)



Abrangência:

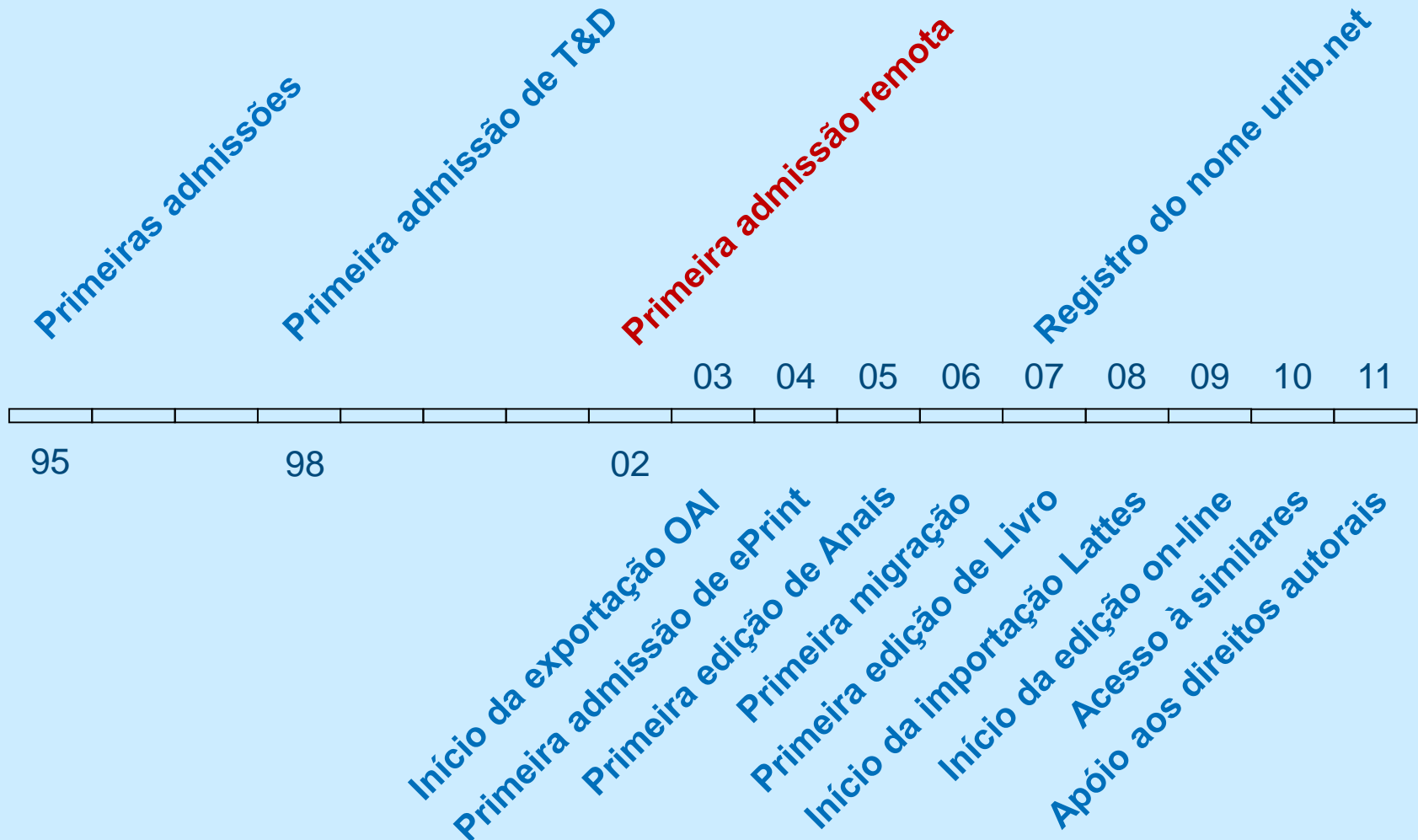
- AE - Artigo em evento
- AR - Artigo em revista
- RJ - Recorte de jornal
- TD - Tese e dissertação
- R - Relatório
- CL - Capítulo de Livro
- MNP - MAN NTC PUD
- eP - ePrint
- MA - Material Audiovisual
- L - Livro

Registro - Total: 33402
 Texto completo - Total: 18985

Em 2005 foram importados 9500 registros do Micro ISIS

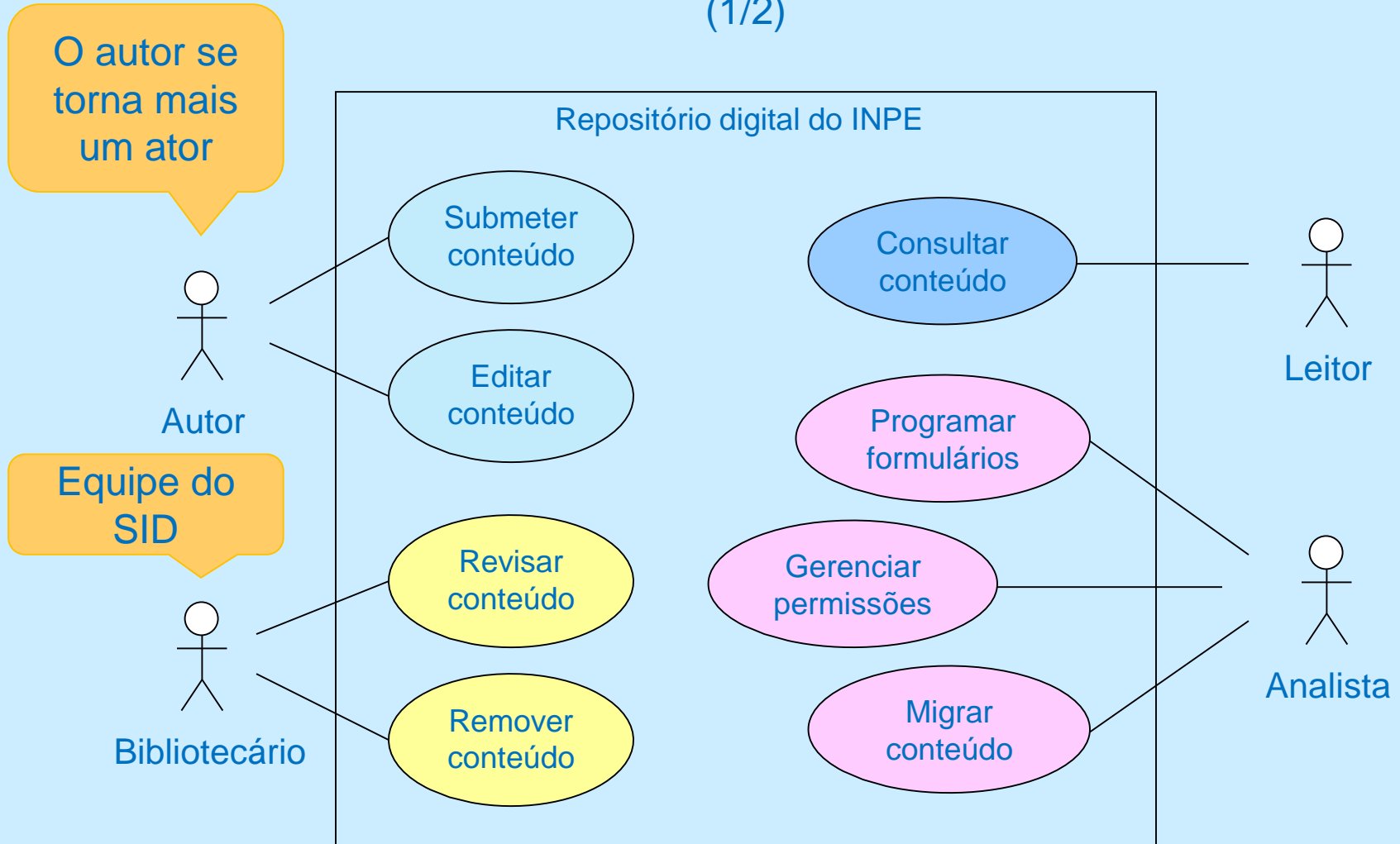
Marcos históricos

(3/13)



Primeira admissão remota (2002)

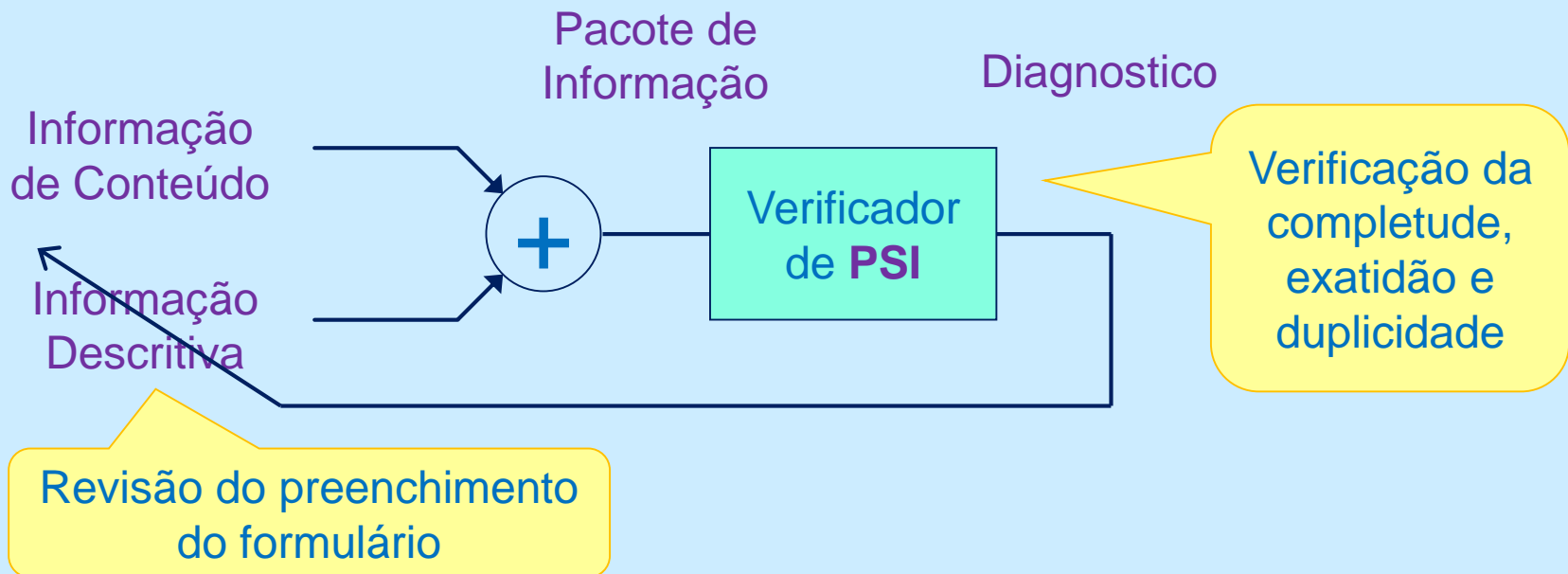
(1/2)



Primeira admissão remota

(2/2)

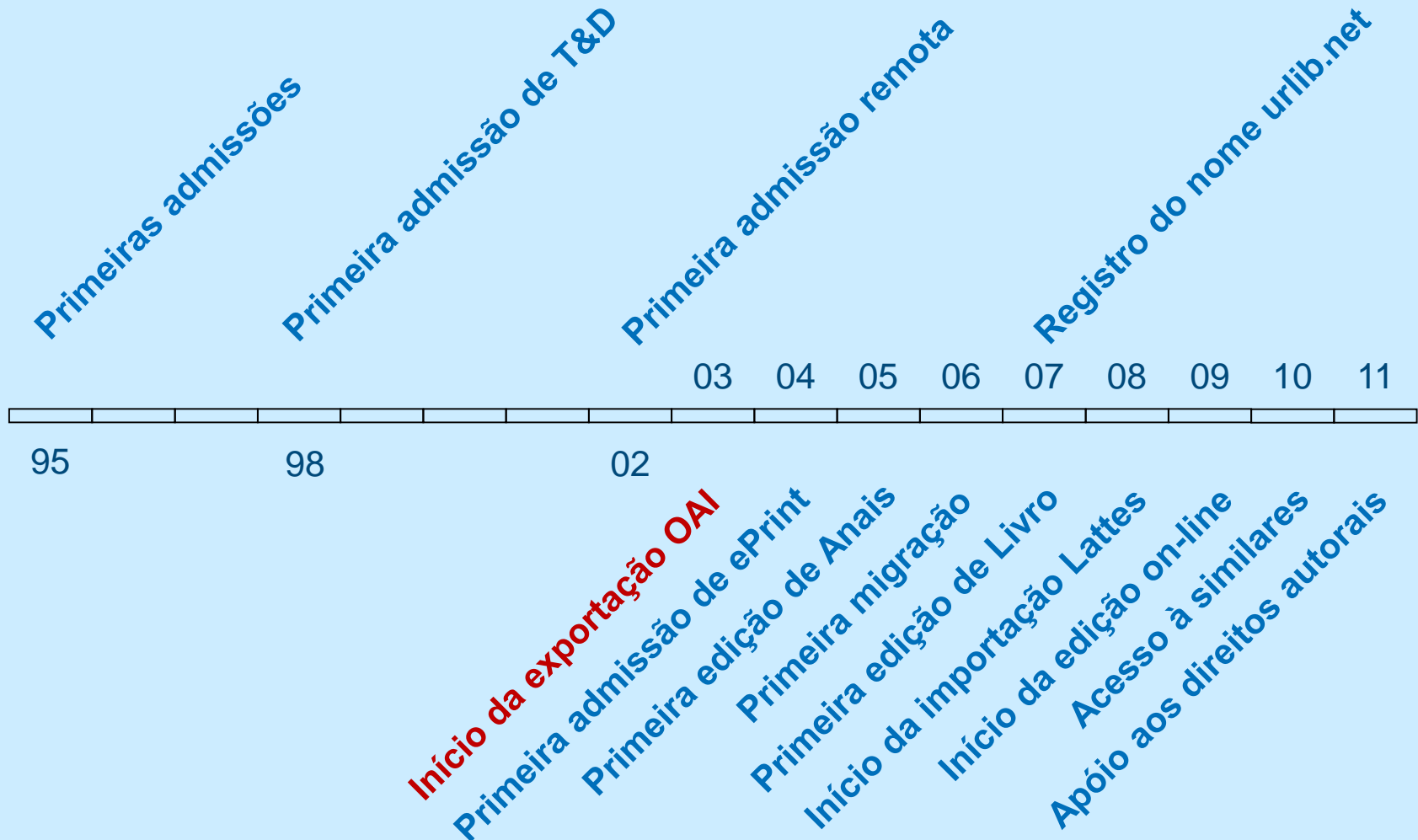
PSI = Pacote de Submissão de Informação



*O Repositório deve ter um processo de admissão que verifica a completude e exatidão de cada **PSI**. (CCSDS 652.0 - 4.1.5)*

Marcos históricos

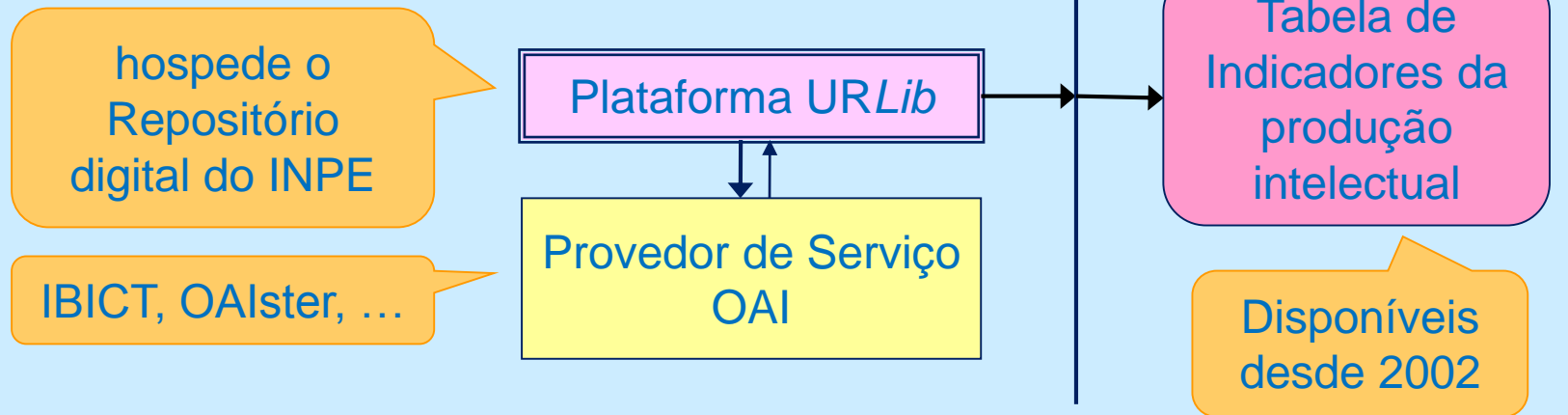
(4/13)



Início da exportação OAI (2003)

(1/1)

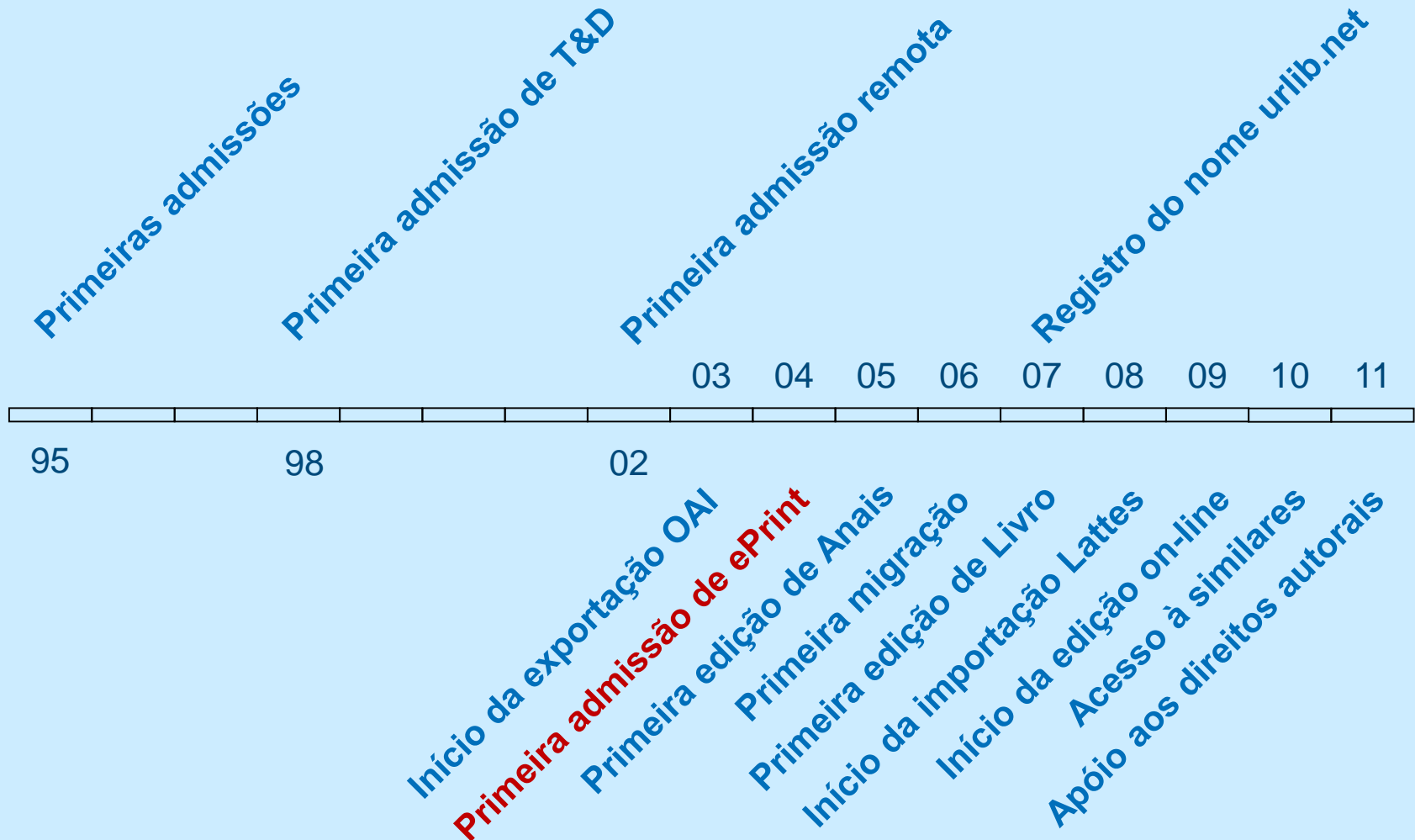
Motivação: tornar o repositório digital um provedor de metadados via o protocolo OAI_PMH como forma de atender à Portaria no 13/2006 da CAPES



O Repositório deve ter identificado e estabelecido as obrigações que ele precisa cumprir. (CCSDS 652.0 – 3.2.1.1)

Marcos históricos

(5/13)



Primeira admissão de ePrint (2004)

(1/3)

Motivação: oferecer um serviço para registrar, divulgar e preservar trabalhos em formato de artigo, não, ou ainda não, publicados por terceiro.

Um documento é o original se for editado pelo dono do **Repositório** no momento da sua admissão

		Original arquivado no Repositório	
		Sim	Não
cessão direitos patrimoniais	Com	X	X
	Sem	X	X

ePrint do INPE

Primeira admissão de ePrint (2004)

(2/3)

Os ePrints do INPE são exibidos como originais

IBI | **é original** | ePrint | Metadados | Como citar? | Acessar | Baixar | 302 KiB | OK | urlib.net | Recarregar | Ocultar

INPE ePrint: dpi.inpe.br/banon-pc2@1905/2005/12.07.19.19 v5 2011-08-02

IBI data

Biblioteca Digital da Memória Técnico-Científica do INPE

A primeira versão desse documento foi elaborada por Gerald Jean Francis Banon em dezembro de 2005 para ser incluído no Anexo A intitulado: Exemplos de arquivos existentes, do Projeto 08:001.06-059 da ABNT intitulado: Sistemas espaciais de dados e informações – Modelo de referência para sistema aberto de arquivamento de informação (SAAI).

A segunda versão inclui as melhorias de português sugeridas por Katia P Thomaz.

Cabeçalho inserido na submissão

Primeira admissão de ePrint (2004)

(3/3)

Especificidade de um ePrint do INPE:

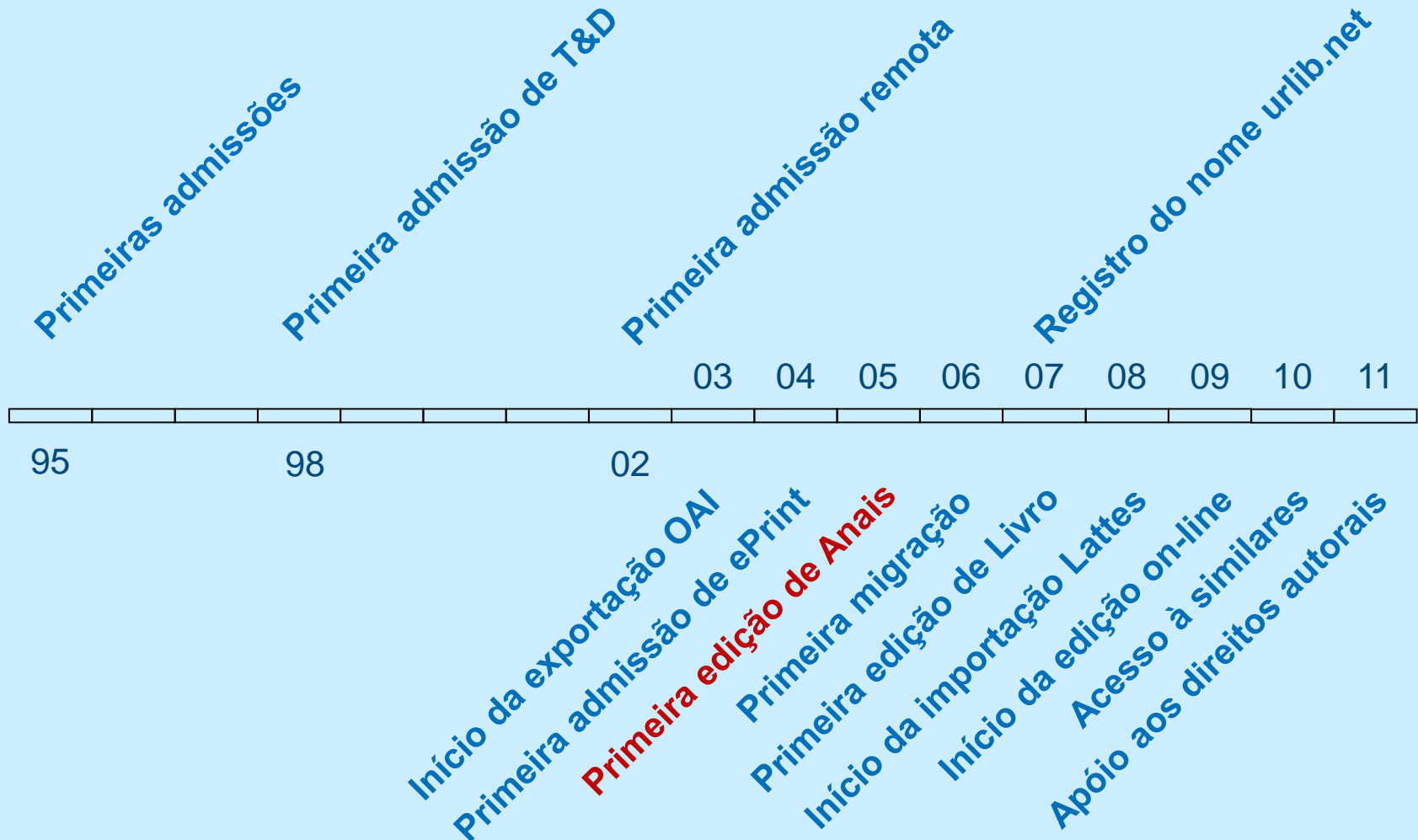
1. não há geralmente revisão pelos pares;
2. recebe um cabeçalho com IBI e data de submissão;
3. não há cessão de direitos patrimoniais;
4. seu acesso pode ser livre ou restrito (\neq arXiv).

O INPE não tem exclusividade

O Repositório deve ter e manter acordos ou contratos de depósito adequados para os materiais digitais que ele gerencia, preserva e/ou para os quais fornece acesso. (CCSDS 652.0 – 3.5.1)

Marcos históricos

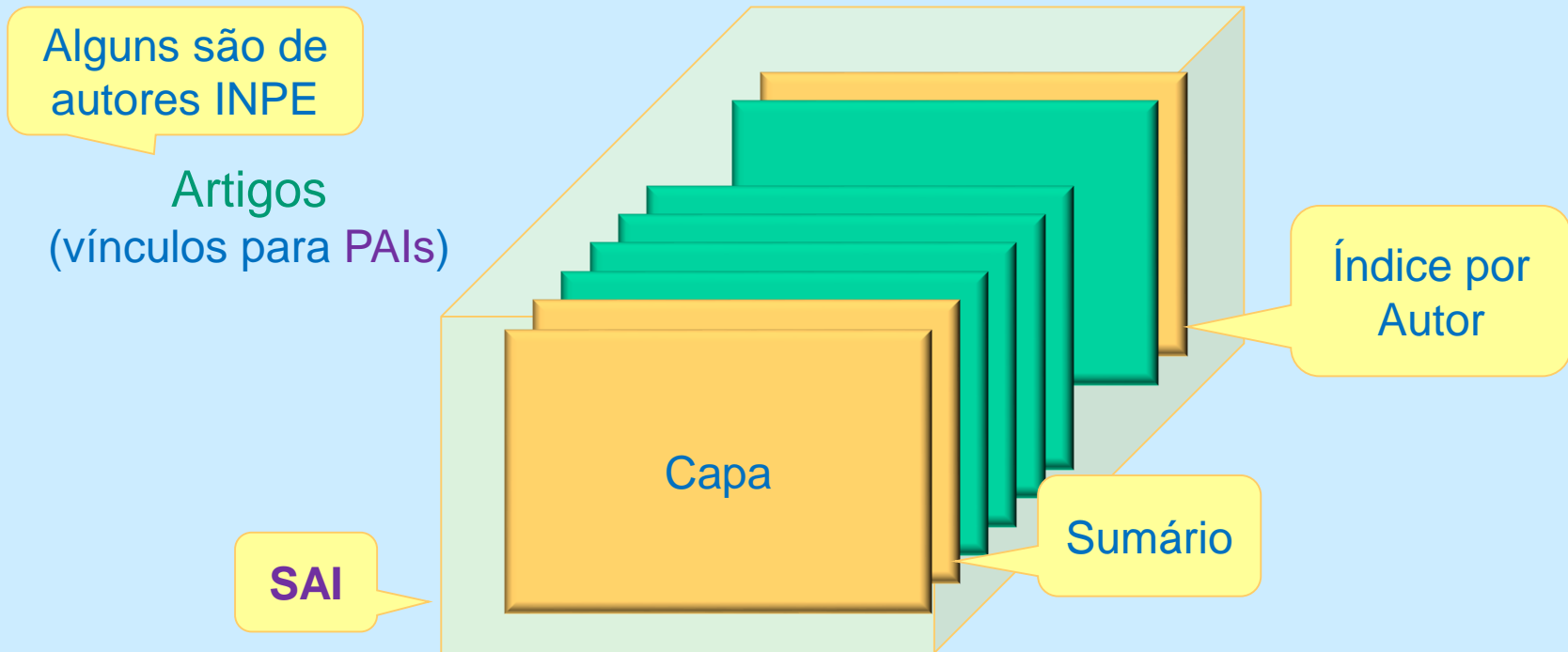
(6/13)



Primeira edição de anais (2005)

(1/3)

Motivação: preservar trabalhos aceitos em eventos, oferecer um acesso ágil e sem restrição.



SAI = serie de arquivamento de informação

Primeira edição de anais (2005)

(2/3)

The image shows two overlapping screenshots of a digital archive interface. The top screenshot displays a 'Sumário' (Table of Contents) for the 'VIII Simpósio Brasileiro de Sensoriamento Remoto - SBSR' held in Salvador, BA, from April 14-19, 1996. It lists navigation options: 'Ficha Catalográfica', 'Prefácio', 'Comissão', and 'Índice por Autor'. The bottom screenshot shows the 'Índice por Autor' (Index by Author) page, which includes an alphabetical index from A to Z and a search input field with the letter 'A' entered. A yellow callout bubble on the right points to the top screenshot, and another on the left points to the bottom screenshot.

Exemplo de Anais digitalizadas

Gerados automaticamente

Primeira edição de anais (2005)

(3/3)

Uma das formas
é o INPE se
tornar editor de
anais

Política de acesso (própria à Instituição):

“Cabem à administração pública, na forma da lei, a gestão da documentação governamental e as providências para franquear sua consulta a quantos dela necessitem.”

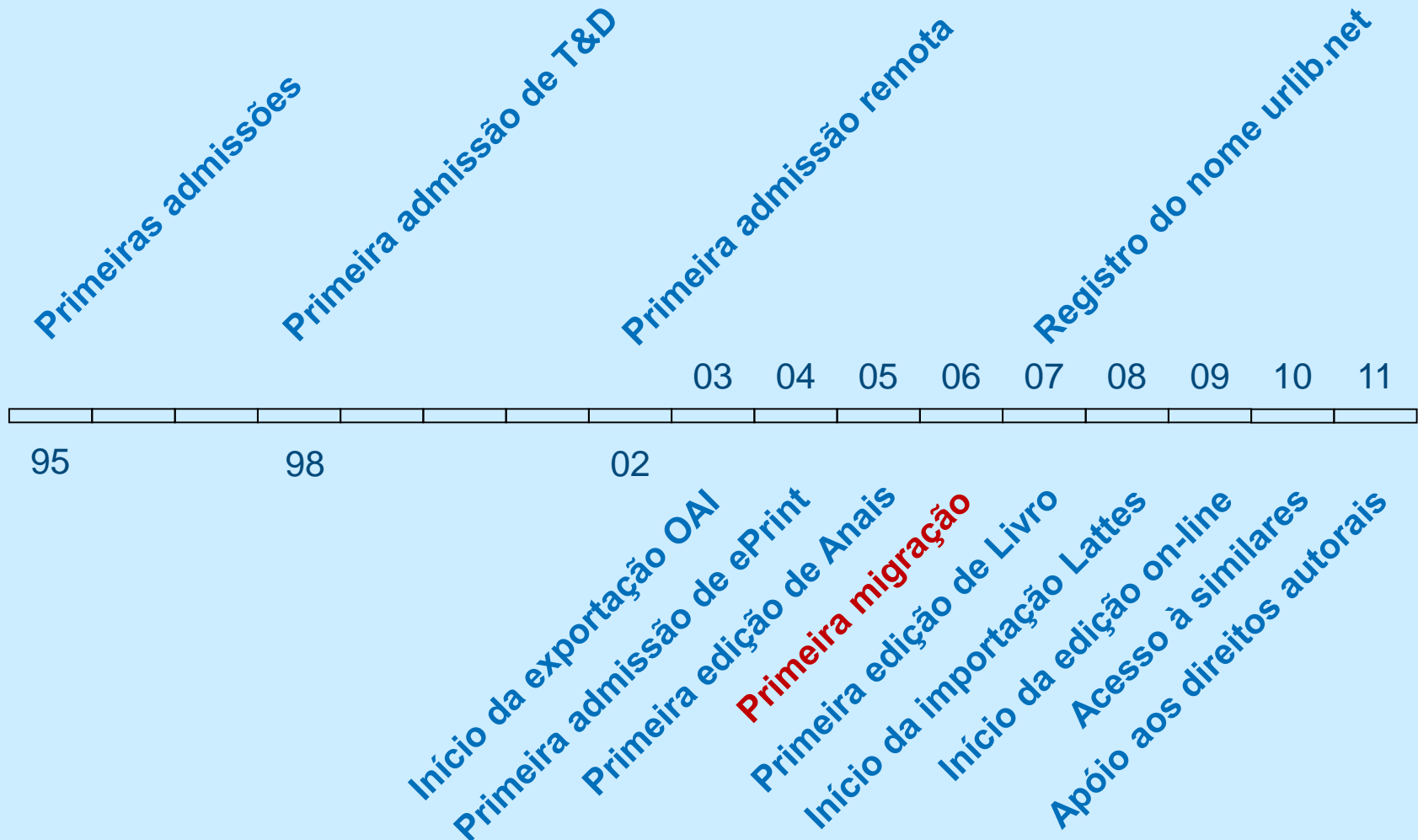
(Par. 2º do Art 216º da Constituição da República Federativa do Brasil de 1988)

O Repositório deve cumprir as Políticas de Acesso.

(CCSDS 652.0 – 4.6.1)

Marcos históricos

(7/13)



Primeira migração (2006)

(1/4)

Motivação: preservar todo o conteúdo da informação da **Federação** em presença de mudanças do suporte computacional.

Estratégia: remodelar a **Federação**:

1. criando novos repositórios para hospedar novos conteúdos;
2. unindo repositórios por **Migração Digital** para simplificar a manutenção.

O Repositório deve ter estratégias de preservação documentadas relevantes para suas posses. (CCSDS 652.0 – 4.3.1)

Primeira migração (2006)

(2/4)

Servidora SUN
Entreprise 250
(Apoio FAPESP 1998)

Nome
n. registros
repositório

2005 (3 mq.)

CPT 1294	LCP 119	SRE 461
SBSR 1067	prelo 362	exp. 226
outros2 1395	outros 387	Rep. espelho
WorCap 180	ePrint 270	MAO 17

2011 (12 mq.)

m05 6303	m09 3871	m12 3596
m15 2256	m16 3004	m17 3930
m18 4206	m19 2147	ismm 349
plutao 2309	martes 6027	bib 139

Rep. espelho

Mais de um Repositório pode ser hospedado num mesmo computador

Parte da
instalação
no SID
08/2011

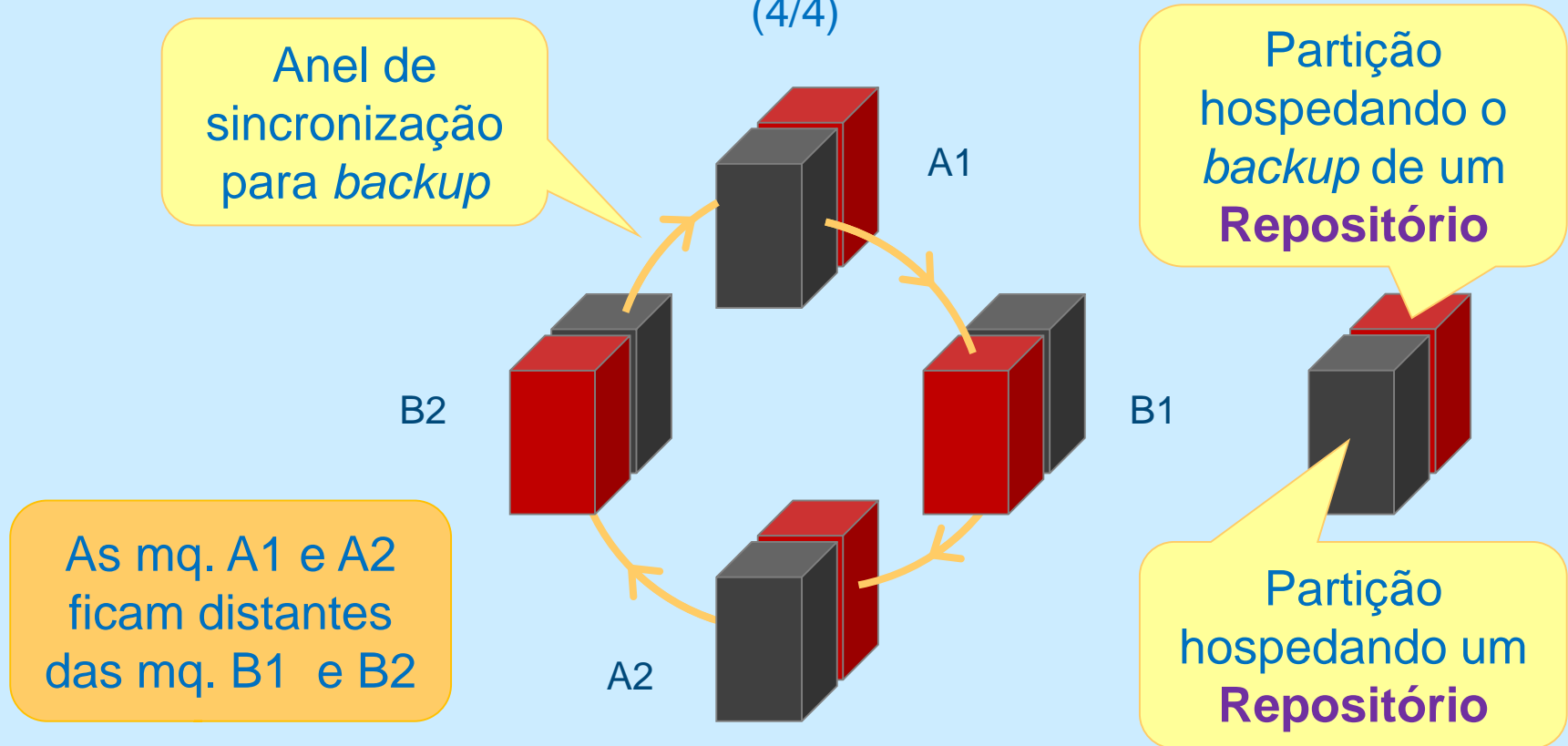
Primeira migração (2006)

(3/4)



Primeira migração (2006)

(4/4)

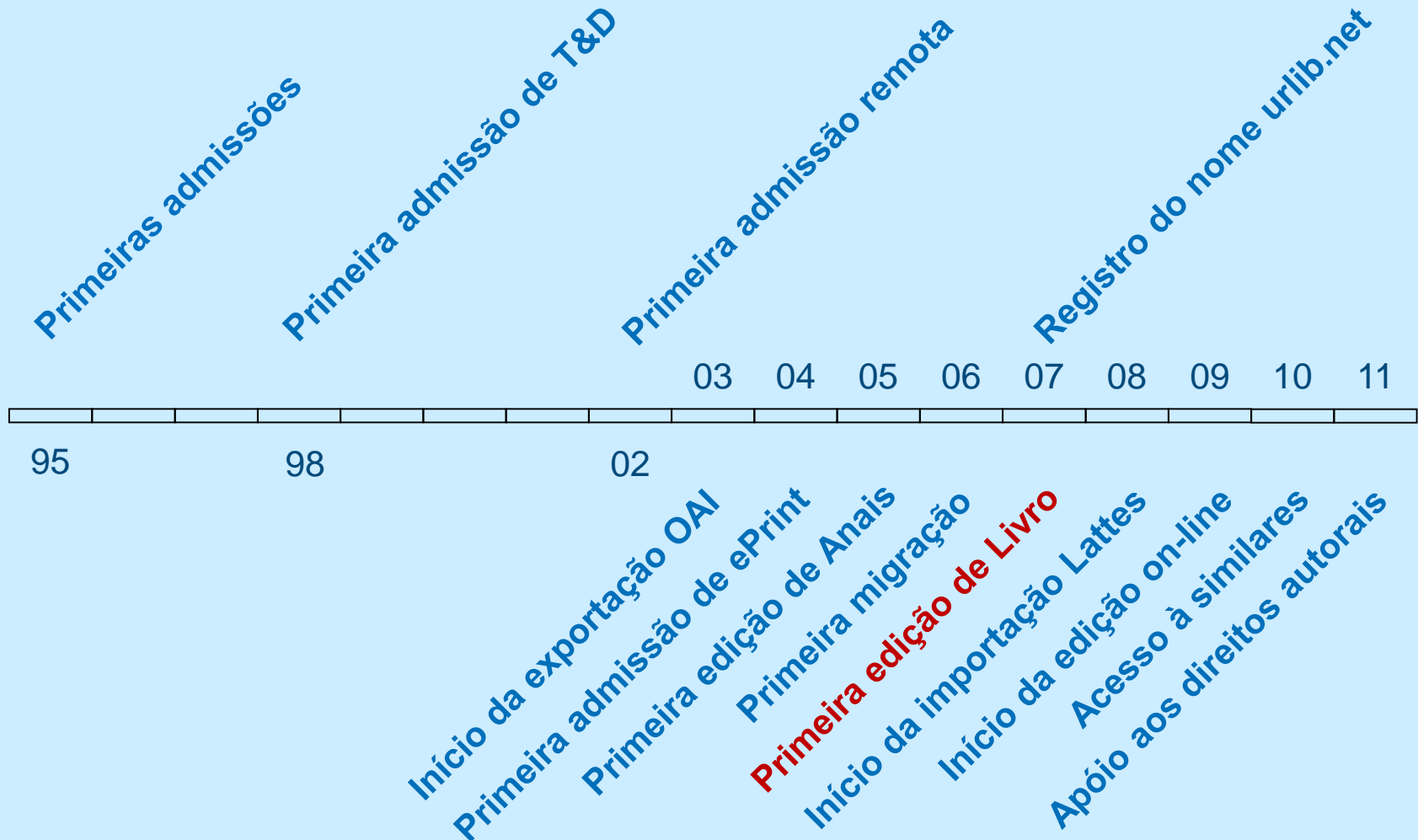


*O **Repositório** deve ter o hardware e suporte de software adequados para a funcionalidade de backup suficiente para preservar o conteúdo do **Repositório** e controlar suas funções.*

(CCSDS 652.0 – 5.1.1.2)

Marcos históricos

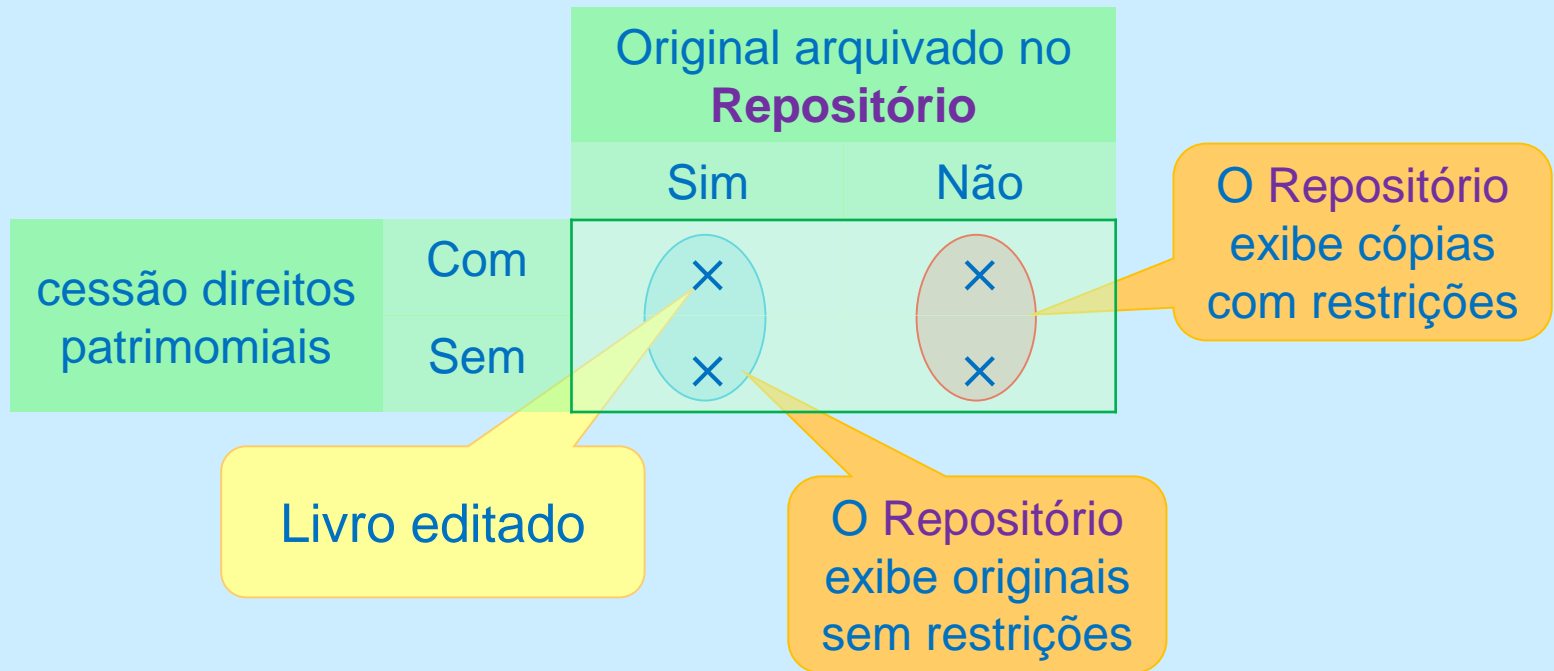
(8/13)



Primeira edição de livro (2007)

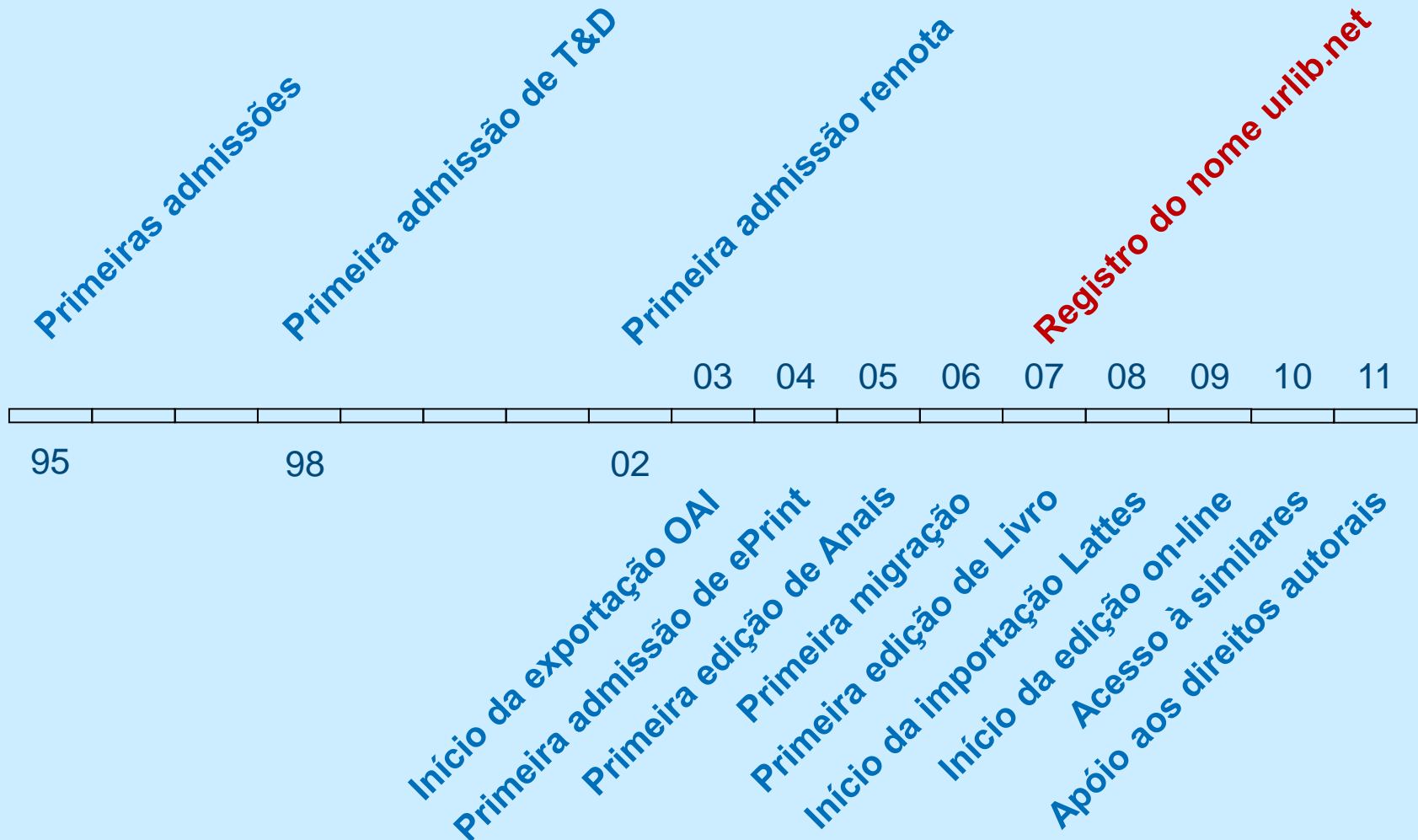
(1/1)

Motivação: oferecer um acesso sem restrição.



Marcos históricos

(9/13)



Registro do nome urlib.net (2007)

(1/2)

Motivação: oferecer um acesso a cada **PAI** via uma URL persistente.

Exemplo: acesso à uma tese por meio de seu IBI:

`http://urlib.net/8JMKD3MGP7W/3973S4S`



resolvedor



IBI

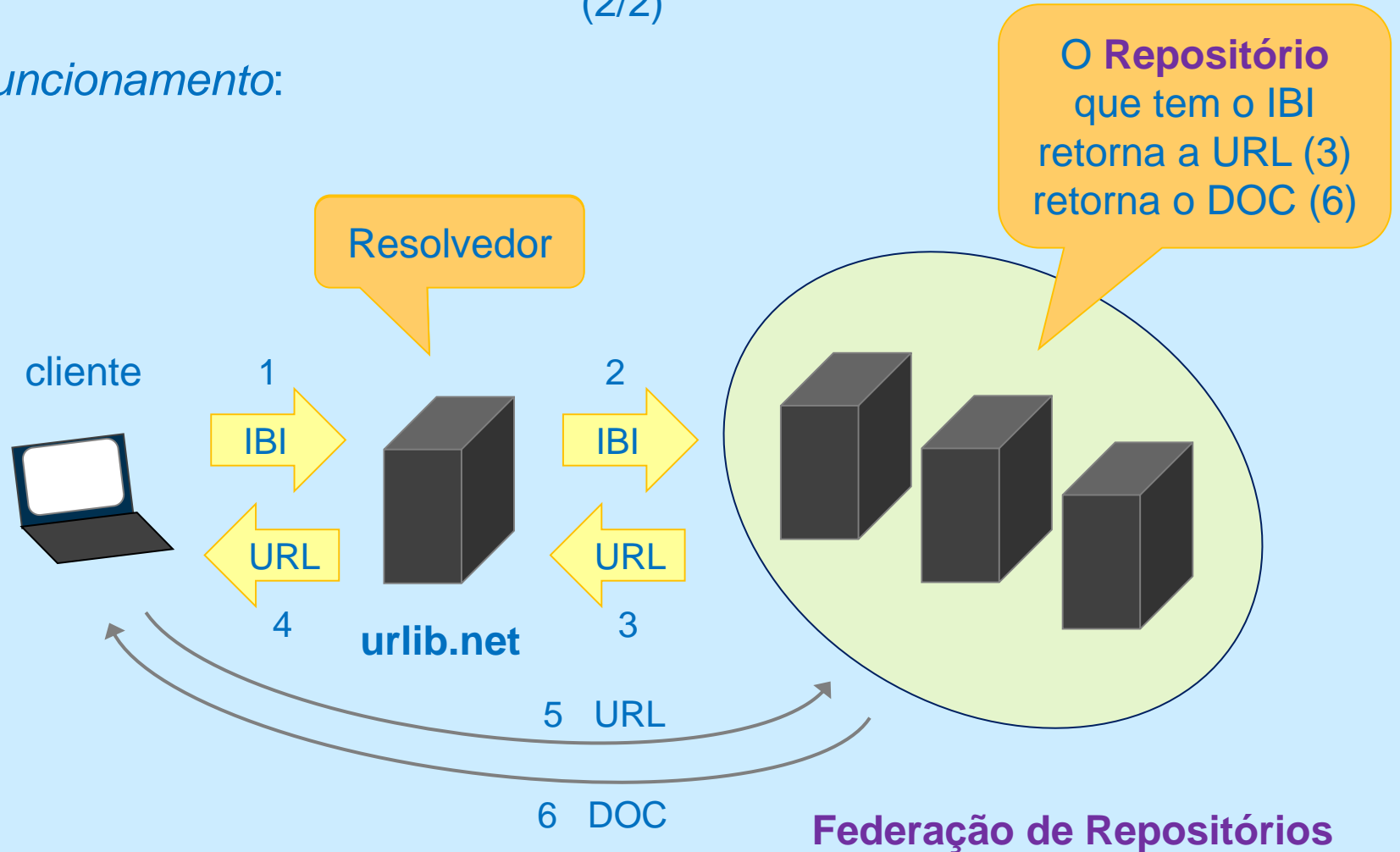
O Repositório deve ter um sistema confiável de serviços de vinculação/resolução a fim de encontrar o objeto identificado com unicidade, independentemente da sua localização física.

(CCSDS 652.0 – 4.2.4.2)

Registro do nome urlib.net (2007)

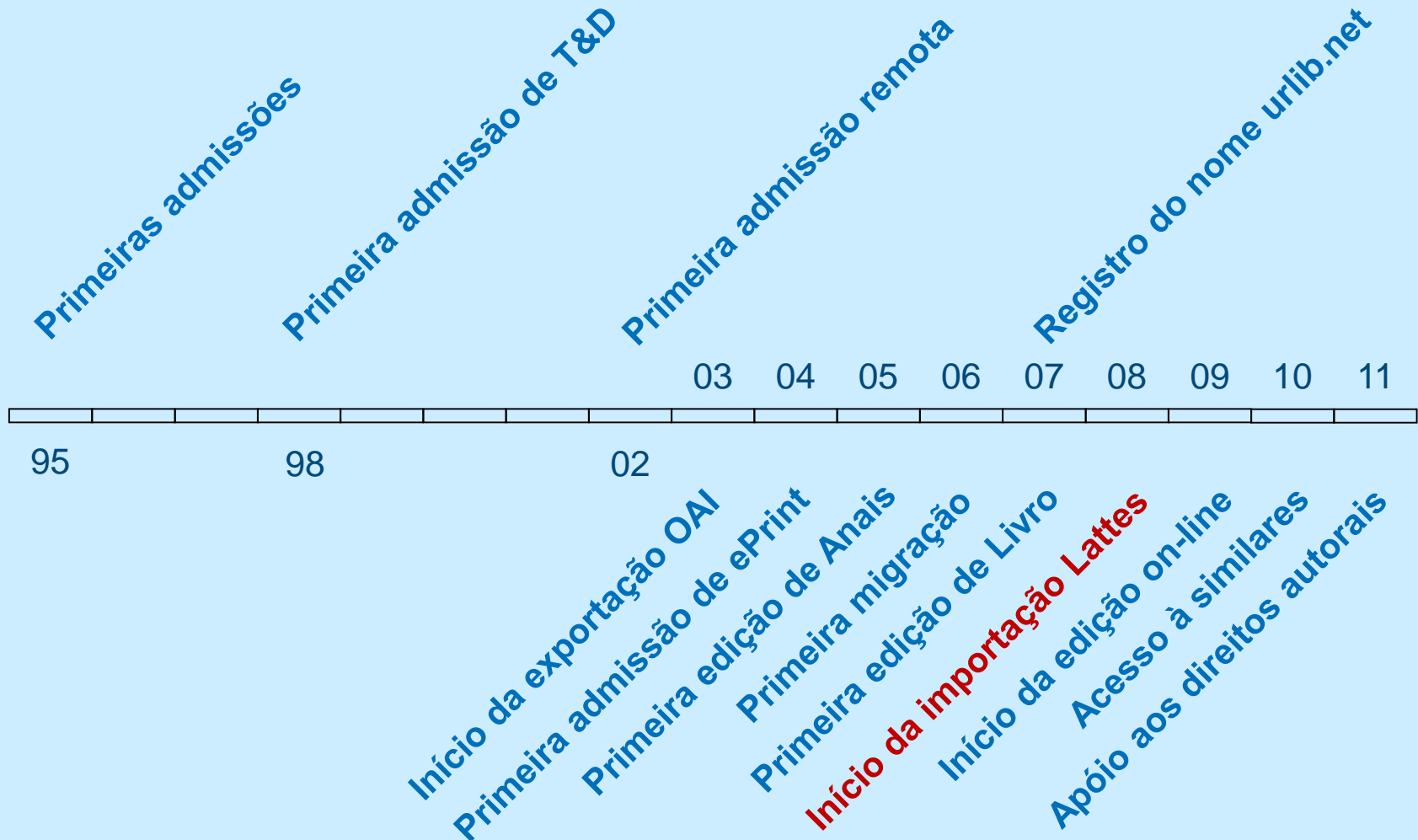
(2/2)

Funcionamento:



Marcos históricos

(10/13)



Início da importação Lattes (2008)

(1/2)

Motivação: oferecer um serviço de apóio à captura da produção científica do Instituto.

Especificidade:

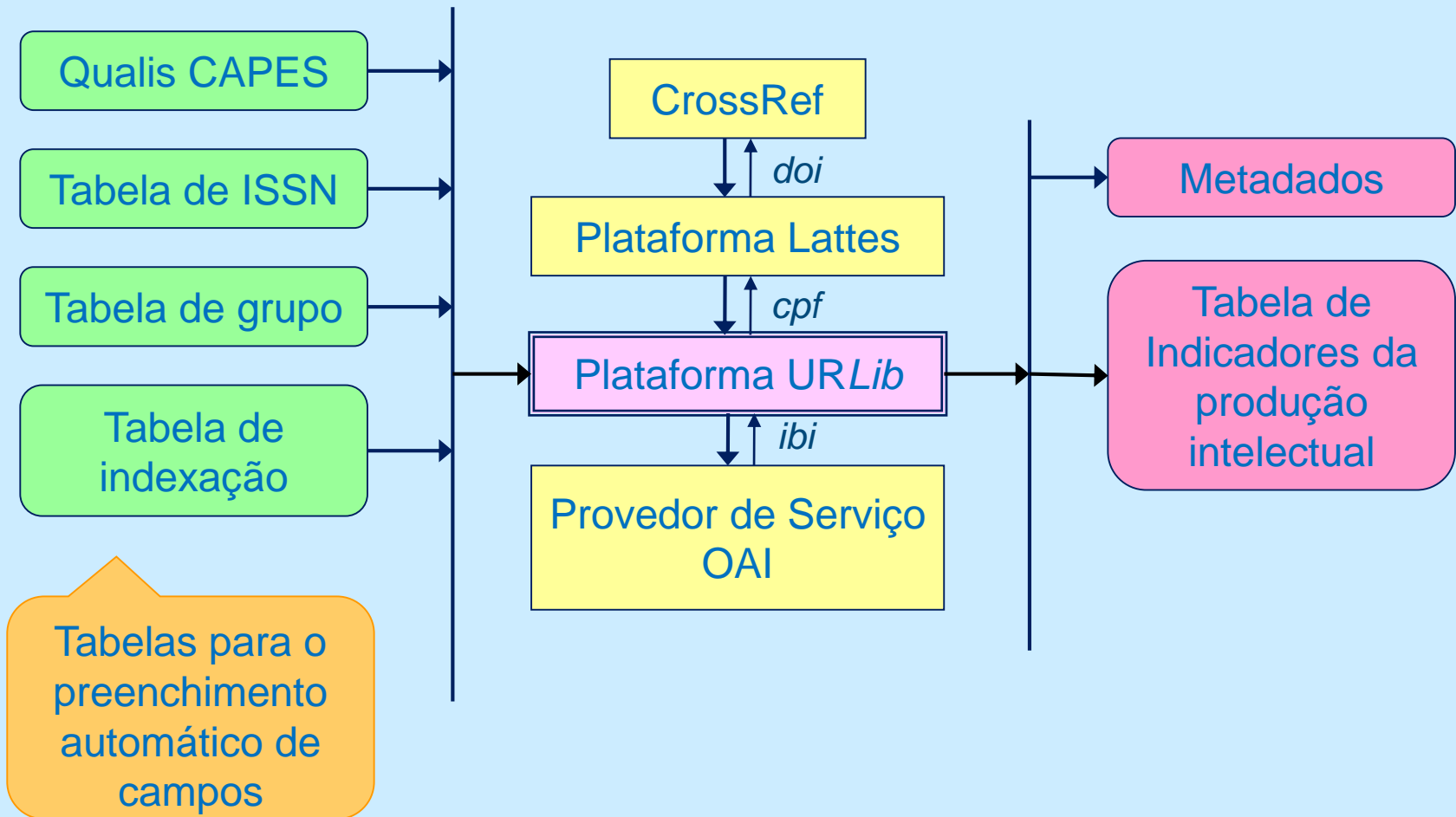
junho e
dezembro

1. importações realizadas duas vezes ao ano;
2. preenchimento automática de alguns campos;
3. dados curriculares obtidos com base no CPF.

O Repositório deve ter mecanismos para verificar adequadamente a identidade do Produtor de qualquer material.
(CCSDS 652.0 – 4.1.4)

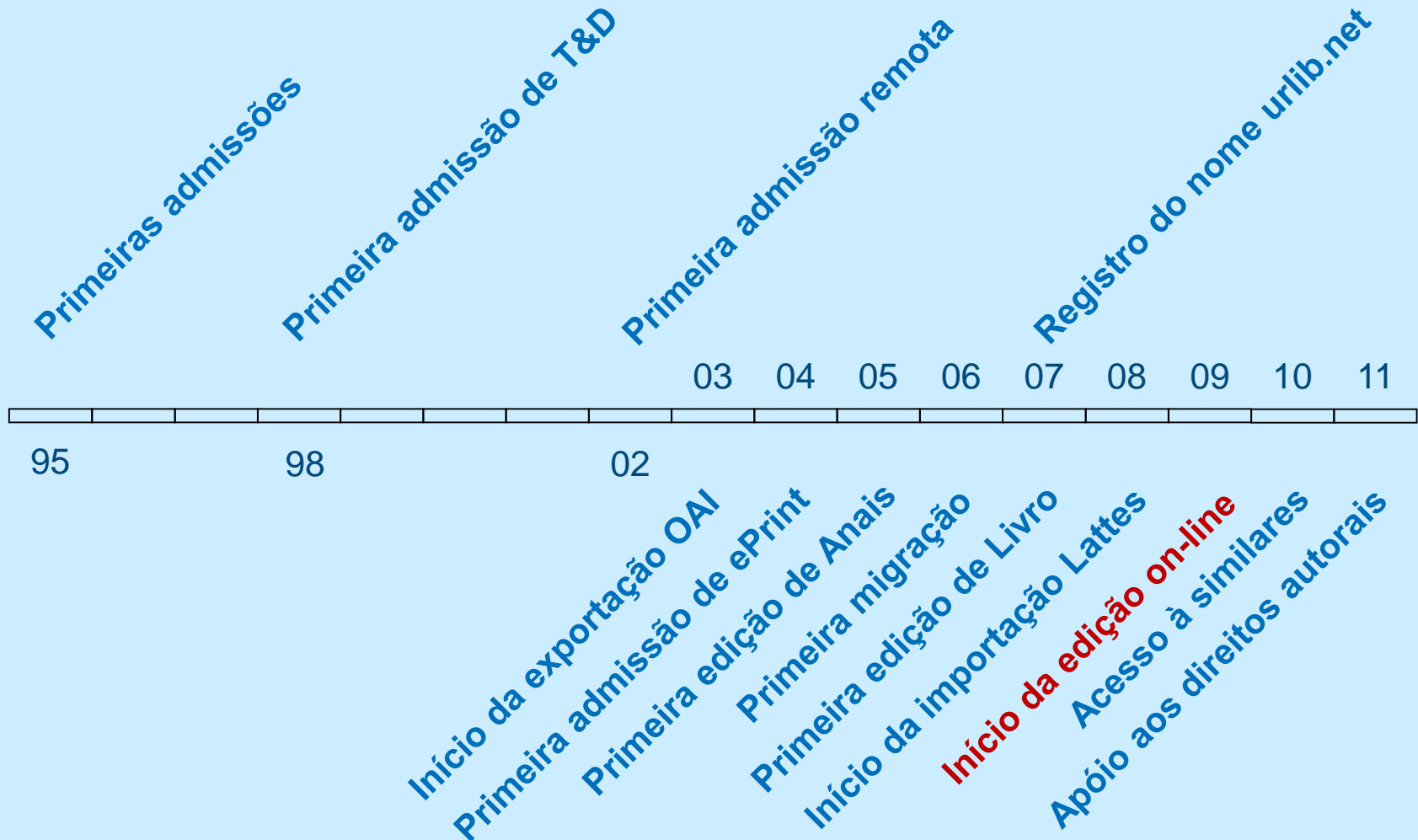
Início da importação Lattes (2008)

(2/2)



Marcos históricos

(11/13)



Início da edição *on-line* (2009)

(1/2)

Motivação: oferecer um serviço orientado para a preservação por longo prazo.

Especificidade:

1. edição com base no LaTeX;
2. compilação remota no próprio **Repositório**;
3. preservação dos vínculos, dados e programas.



Geração
do PDF

*O **Repositório** deve ter uma definição de cada **PAI** que é adequada para a preservação por longo prazo, permitindo a identificação e análise de todos os componentes necessários a esse **PAI**.*

(CCSDS 652.0 – 4.2.1.2)

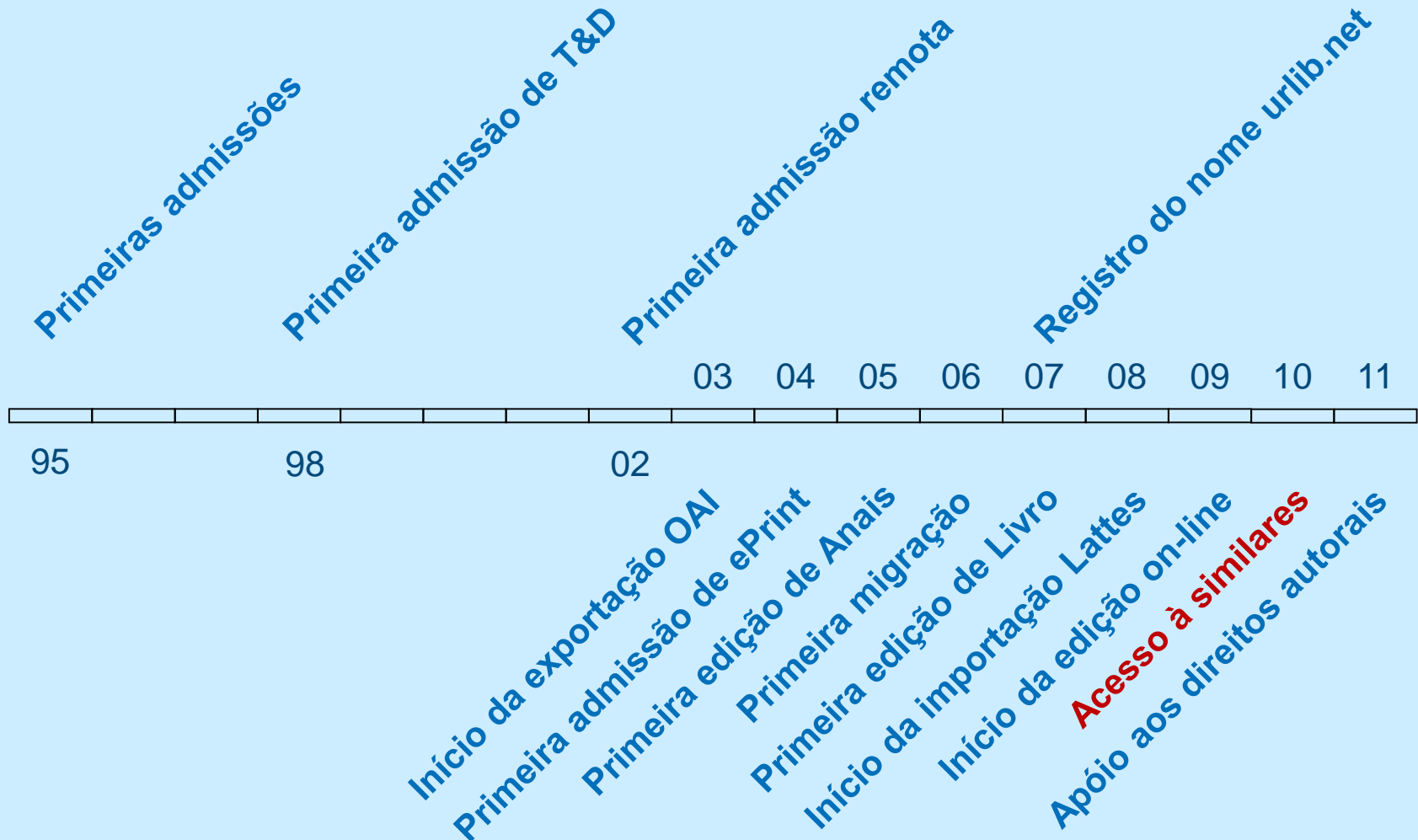
Início da edição *on-line* (2009)

(2/2)

The screenshot shows a web browser window with a document viewer. The browser's address bar shows a URL from 'banon-pc3.dpi.inpe.br'. The document viewer has a menu bar with 'Arquivo', 'Editar', 'Exibir', 'Histórico', 'Favoritos', 'Ferramentas', and 'Ajuda'. The document content includes a page number '174' and a histogram. The histogram's y-axis is labeled 'Número de animais' (0 to 10) and the x-axis is labeled 'VHS' (9.0 to 15.0). The histogram bars represent the following data points: (9.0-9.5, 4), (9.5-10.0, 8), (10.0-10.5, 9), (10.5-11.0, 5), (11.0-11.5, 1). Three callout boxes highlight specific features: '1. Botão para baixar o TEX' points to the 'Baixar' button; '2. Botão para anexar o TEX' points to the 'Selecionar arquivo...' button; '3. Botão para geração do PDF' points to the 'Executar' button. A yellow callout box on the left states: 'O gráfico é redesenhado na geração do PDF com base nos dados'.

Marcos históricos

(12/13)



Acesso à similares (2010)

(1/3)

Motivação: oferecer um serviço para a detectar duplicados no mesmo idioma.

Especificidade:

1. detecção de similares com base nas palavras mais importantes do título;
2. Ordenação com base num grau de correlação entre títulos.

Duplicações
ocorrem na
importação
Lattes

Acesso à similares (2010)

(2/3)

Registro importado do Lattes de forma errônea pela falta de um coautor no CV

o da Pesquisa

A busca foi <related:dpi.inpe.br/plutao/2010/12.02.14.06.44-0:en:title:2:observations
convective production sprite2006:observations prolific transient luminous event production above
mesoscale convective system argentina during sprite2006 campaign brazil:>. 2 referências encontradas buscando em 12 dentro 12 sites.

SãoSabbasTPCASTPSSFFC:2010:ObPrTr J8LNKAN8RW/38MQUDJ Artigo em Revista Científica (1.00)

São Sabbas, F., et al. :2010:

Observations of prolific transient luminous event production above a mesoscale convective system in Argentina during the Sprite2006 Campaign in Brazil

referência (BibTeX | Refer | BibINPE | XML)

<dpi.inpe.br/plutao/2010/12.02.14.06.43> atualizar | similares

<plutao.dpi.inpe.br>

SãoSabbasTPBCASTPSSFFC:2010:ObPrTr 8JMKD3MGP7W/3822C5L Artigo em Revista Científica acesso restrito (1.00)

São Sabbas, F. T., et al. :2010:

Observations of Prolific Transient Luminous Event Production Above a Mesoscale Convective System in Argentina during the Sprite2006 Campaign in Brazil

referência (BibTeX | Refer | BibINPE | XML | xrefer | oai_dc | Capa)

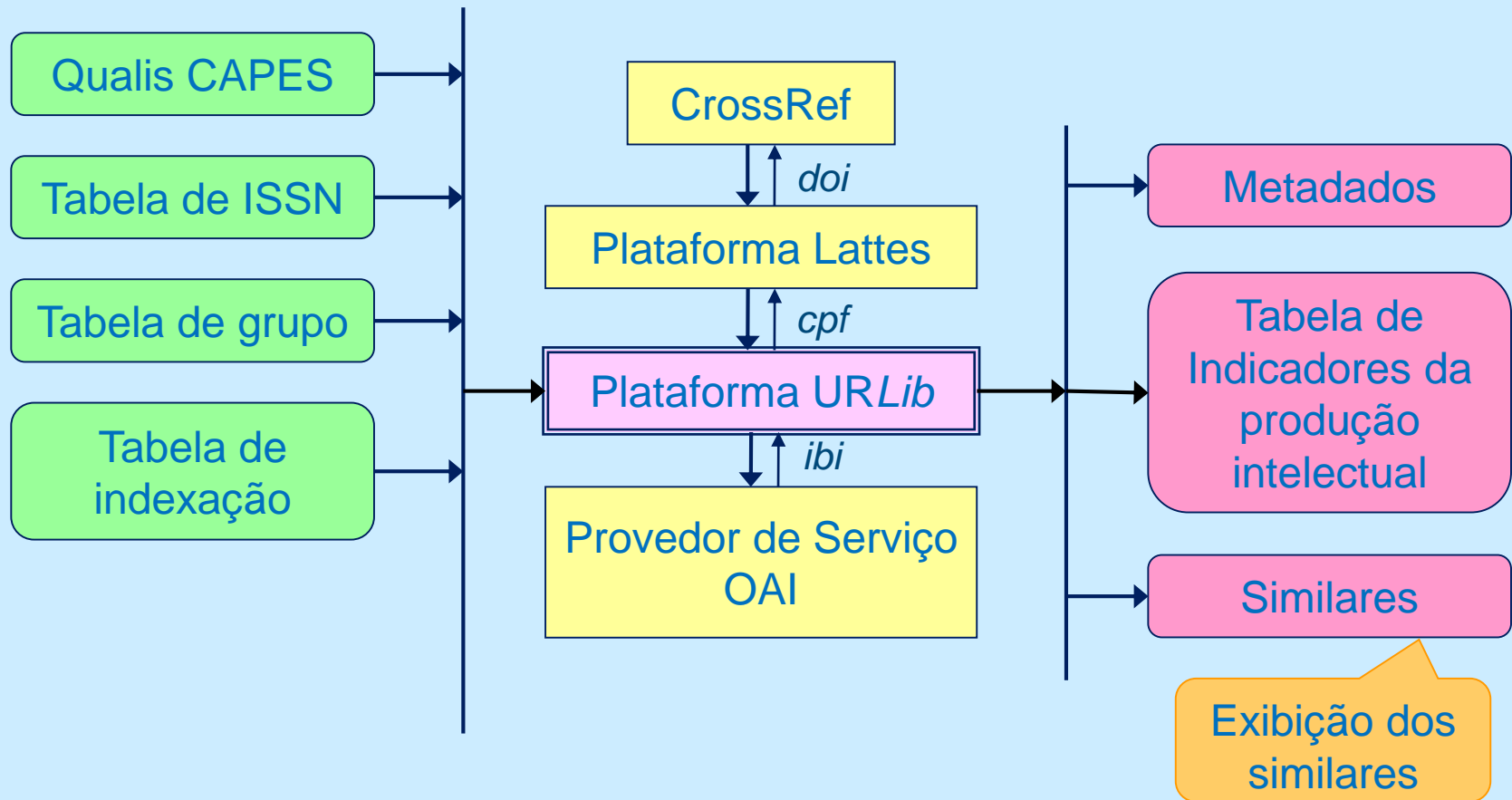
<sid.inpe.br/mtc-m19@80/2010/08.02.13.42> (98 KiB) estatísticas | acesso restrito | atualizar | similares

<mtc-m19.sid.inpe.br>

Registro anterior a importação

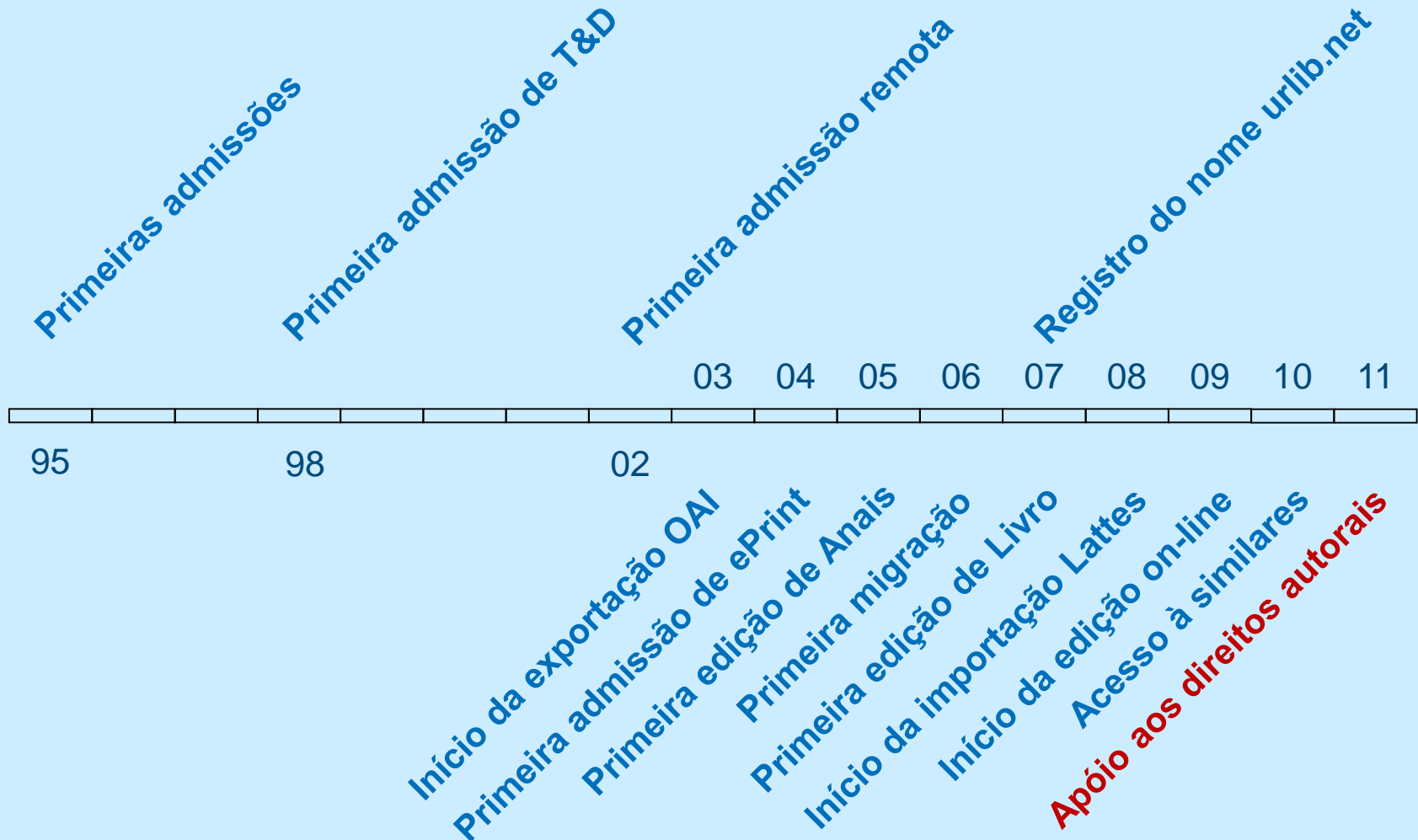
Acesso à similares (2010)

(3/3)



Marcos históricos

(13/13)



Apóio aos direitos autorais (2011)

(1/2)

Motivação: oferecer um serviço de apoio ao cumprimento das políticas de auto-arquivamento das revistas.

Política de acesso (própria à Terceiros):

Exemplos de
possíveis
valores

allowpublisher allowfinaldraft
denypublisher allowfinaldraft
allowpublisher denyfinaldraft
denypublisher denyfinaldraft

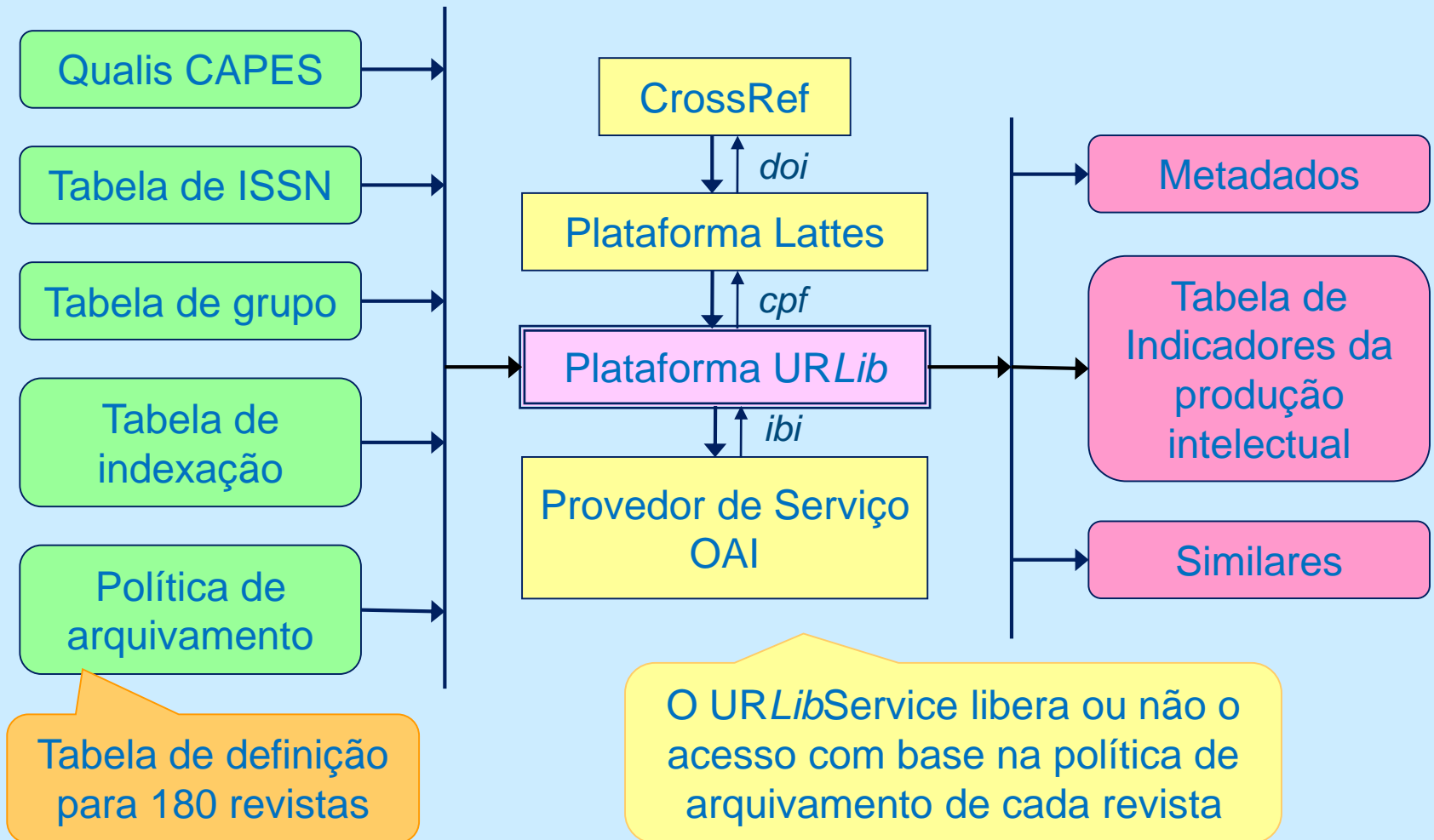
denypublisher6 allowfinaldraft
allowpublisher denyfinaldraft12

O Repositório deve cumprir as Políticas de Acesso.

(CCSDS 652.0 – 4.6.1)

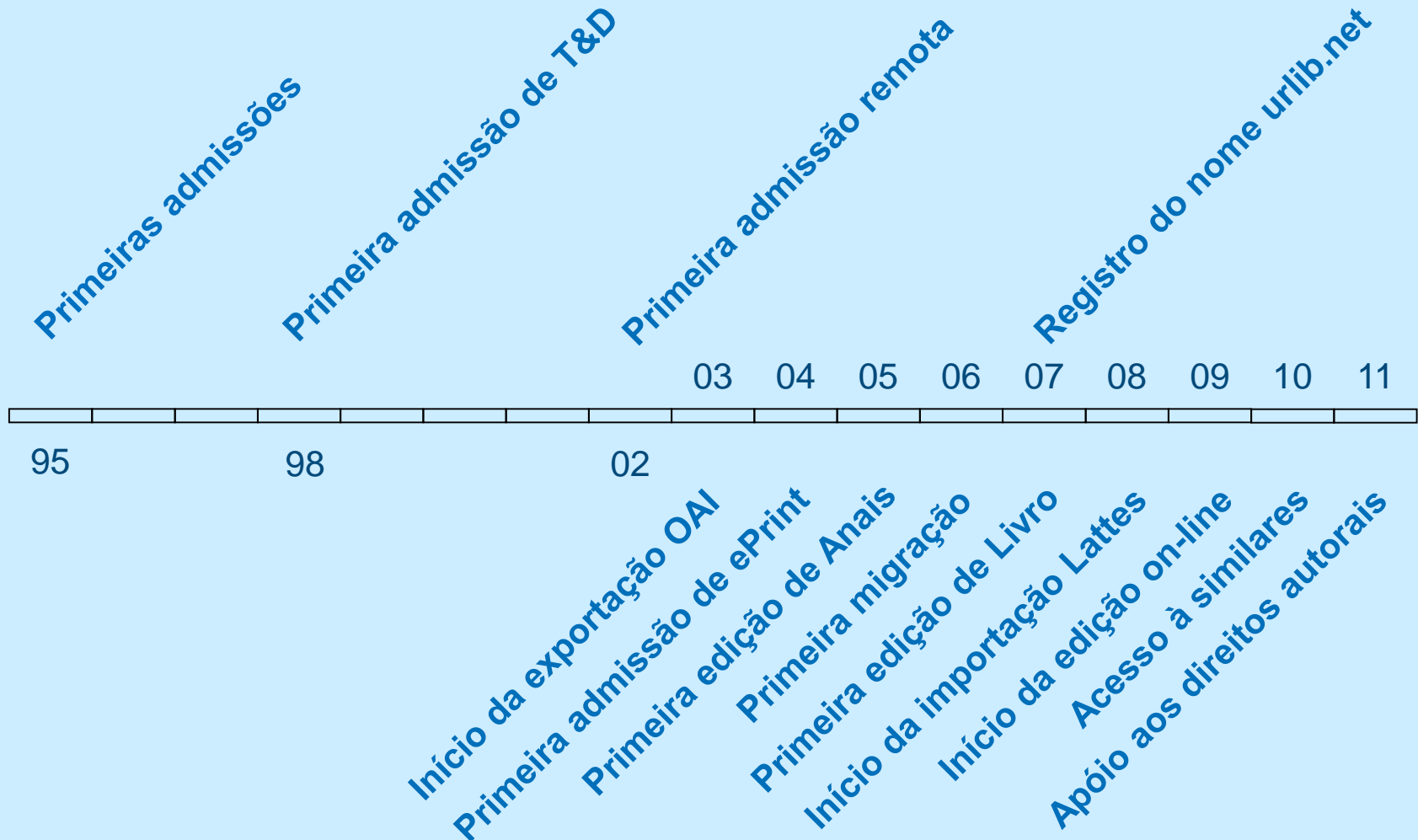
Apóio aos direitos autorais (2011)

(2/2)



Marcos históricos

(1/1)




Perspectivas

(1/1)

Identificar parceiros para tornar a **Federação de Repositórios** um repositório interinstitucional.

Motivos:

1. unir recursos entre instituições;
2. evitar duplicidade no cadastramento da produção científica das instituições.



Minimiza
riscos

O Repositório deve ter implementado controles para tratar adequadamente cada um dos riscos de segurança definidos.

(CCSDS 652.0 – 5.2.2)

Vínculos úteis

(1/1)

Auditoria interna sobre o repositório digital da Memória Científica do INPE

<http://urlib.net/J8LNKAN8PW/36CT2G2>

Biblioteca Digital da Memória Técnico-Científica do INPE

<http://urlib.net/J8LNKAN8PWT2T/QNQFSH>

Repositório digital do INPE

<http://urlib.net/sid.inpe.br/bibdigital@80/2006/04.07.15.50.13>

<http://bibdigital.sid.inpe.br>

Agradecimentos

(1/1)

Queremos aqui registrar a inestimável contribuição da equipe do Serviço de Informação e Documentação (SID/INPE), e em particular, da Marciana Leite Ribeiro, na montagem, desde o primeiro dia, do Repositório Digital da Memória Científica do INPE, assim como a participação da Lise Christine Banon no desenvolvimento do módulo de importação dos CV Lattes.

Obrigado!