

# EDCO4B

# ESTRUTURAS DE DADOS 2

Aula 12 B - Índices Secundários

Prof. Rafael G. Mantovani

# Licença

Este trabalho está licenciado com uma Licença CC BY-NC-ND 4.0:



maiores informações:

[https://creativecommons.org/licenses/by-nc-nd/4.0/deed.pt\\_BR](https://creativecommons.org/licenses/by-nc-nd/4.0/deed.pt_BR)

# Roteiro



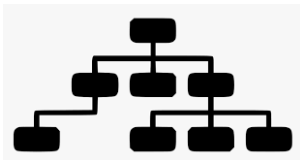
- 1** Introdução
- 2** Índices Secundários
- 3** Operações
- 4** Revisão
- 5** Referências

# Roteiro

- 1 Introdução**
- 2 Índices Secundários**
- 3 Operações**
- 4 Revisão**
- 5 Referências**

# Introdução

- O que já fizemos:

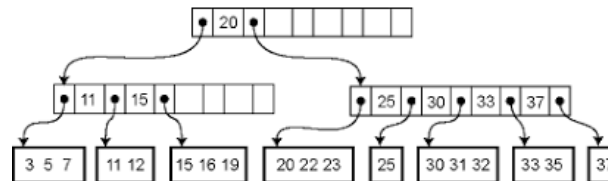


Estruturas de Arquivos



1. arquivos de índices

2. árvores B

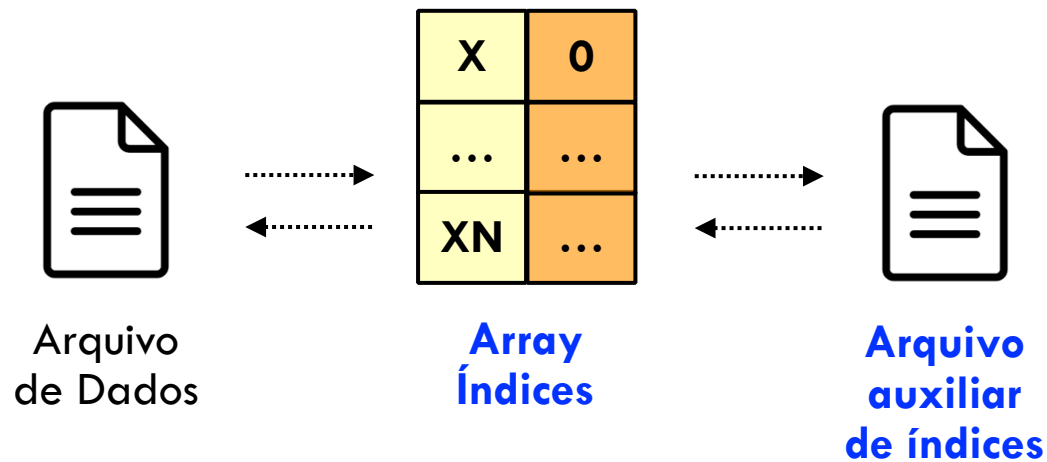


3. Hashing

Timeline

# Introdução

## □ Estrutura:



# Introdução

ANG3795	167
COL31809	353
COL38358	211
DG139201	396
DG18807	256
FF245	442
LON2312	32
MER75016	300
RCA2626	77
WAR23699	132

**Arquivo de índice  
(disco)**

32	LON   2312   Romeo and Juliet   Prokofiev ...
77	RCA   2626   Quartet in C sharp Minor ...
132	WAR   23699   Touchstone   Corea ...
167	ANG   3795   Symphony No. 9   Beethoven ...
211	COL   38358   Nebraska   Springsteen ...
256	DG   18807   Symphony No. 9   Beethoven ...
300	MER   75016   Co1 d'or Suite   Rimsky ...
353	COL   31809   Symphony No. 9   Dvorak ...
396	DG   139201   Violin Concerto   Beethoven ...
442	FF   245   Good News   Sweet Honey in The ...

**Arquivo de dados (disco)**

# Introdução

ANG3795	167
COL31809	353
COL38358	211
DG139201	396
DG18807	256
FF245	442
LON2312	32
MER75016	300
RCA2626	77
WAR23699	132

**Arquivo de índice  
(disco)**

32	LON   2312   Romeo and Juliet   Prokofiev ...
77	RCA   2626   Quartet in C sharp Minor ...
132	WAR   23699   Touchstone   Corea ...
167	ANG   3795   Symphony No. 9   Beethoven ...
211	COL   38358   Nebraska   Springsteen ...
256	DG   18807   Symphony No. 9   Beethoven ...
300	MER   75016   Co1 d'or Suite   Rimsky ...
353	COL   31809   Symphony No. 9   Dvorak ...
396	DG   139201   Violin Concerto   Beethoven ...
442	FF   245   Good News   Sweet Honey in The ...

**Arquivo de dados (disco)**



# Introdução

ANG3795	167
COL31809	353
COL38358	211
DG139201	396
DG18807	256
FF245	442
LON2312	32
MER75016	300
RCA2626	77
WAR23699	132

Arquivo de índice  
(disco)

32	LON   2312   Romeo and Juliet   Prokofiev ...
77	RCA   2626   Quartet in C sharp Minor ...
132	WAR   23699   Touchstone   Corea ...
...	...
396	DG   139201   Violin Concerto   Beethoven ...
442	FF   245   Good News   Sweet Honey in The ...

Arquivo de dados (disco)

Difícilmente vai se procurar por algo usando uma chave primária (**LON2312**). Porque não procurar as músicas pelo título, compositor ou artista?

# Introdução

ANG3795	167
COL31809	353
COL38358	211
DG139201	396
DG18807	256
FF245	442
LON2312	32
MER75016	300
RCA2626	77
WAR23699	132

Arquivo de índice  
(disco)

32	LON   2312   Romeo and Juliet   Prokofiev ...
77	RCA   2626   Quartet in C sharp Minor ...
132	WAR   23699   Touchstone   Corea ...

Difícilmente vai se procurar por algo usando uma chave primária (**LON2312**). Porque não procurar as músicas pelo título, compositor ou artista?

**Chaves secundárias !!!**

396	DG   139201   Violin Concerto   Beethoven ...
442	FF   245   Good News   Sweet Honey in The ...

Arquivo de dados (disco)

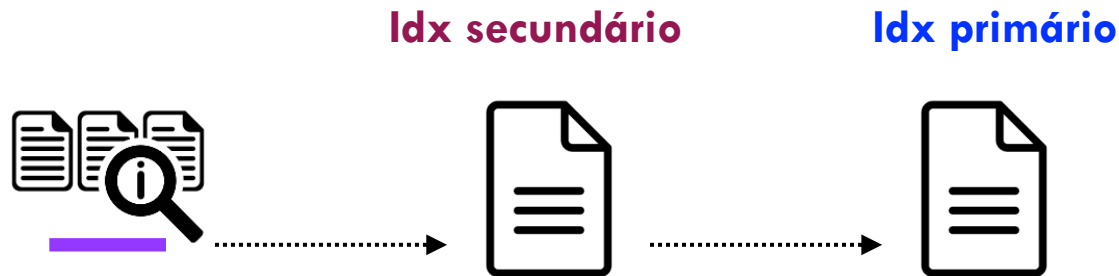
# Roteiro



- 1 Introdução
- 2 Índices Secundários
- 3 Operações
- 4 Revisão
- 5 Referências

# Índice Secundário

- **Ideia:** índice(s) auxiliar(es) para dar mais flexibilidade na busca e recuperação de registros
- Pesquisa



- Adiciona mais um nível no processo de busca
- Busca por meio de chaves secundárias

# Índice Secundário



# Índice Secundário

BEETHOVEN	ANG3795
BEETHOVEN	DG139201
BEETHOVEN	DG18807
BEETHOVEN	RCA2626
COREA	WAR23699
DVORAK	COL31809
PROKOFIEV	LON2312
RIMSKY-KORSAKOV	MER75016
SPRIGNSTEEN	COL38358
SWEET HONEY IN THE R	FF245

# Índice Secundário

BEETHOVEN	ANG3795
BEETHOVEN	DG139201
BEETHOVEN	DG18807
BEETHOVEN	RCA2626
COREA	WAR23699
DVORAK	COL31809
PROKOFIEV	LON2312
RIMSKY-KORSAKOV	MER75016
SPRIGNSTEEN	COL38358
SWEET HONEY IN THE R	FF245

**Idx Secundário: organizado por compositor**

# Índice Secundário

## Chave Secundária

BEETHOVEN	ANG3795
BEETHOVEN	DG139201
BEETHOVEN	DG18807
BEETHOVEN	RCA2626
COREA	WAR23699
DVORAK	COL31809
PROKOFIEV	LON2312
RIMSKY-KORSAKOV	MER75016
SPRIGNSTEEN	COL38358
SWEET HONEY IN THE R	FF245

**Idx Secundário: organizado por compositor**



# Índice Secundário

**Chave  
Secundária**

**Chave  
Primária**

BEETHOVEN	ANG3795
BEETHOVEN	DG139201
BEETHOVEN	DG18807
BEETHOVEN	RCA2626
COREA	WAR23699
DVORAK	COL31809
PROKOFIEV	LON2312
RIMSKY-KORSAKOV	MER75016
SPRIGNSTEEN	COL38358
SWEET HONEY IN THE R	FF245

**Idx Secundário: organizado por compositor**

# Índice Secundário

**Chave  
Secundária**

**Chave  
Primária**

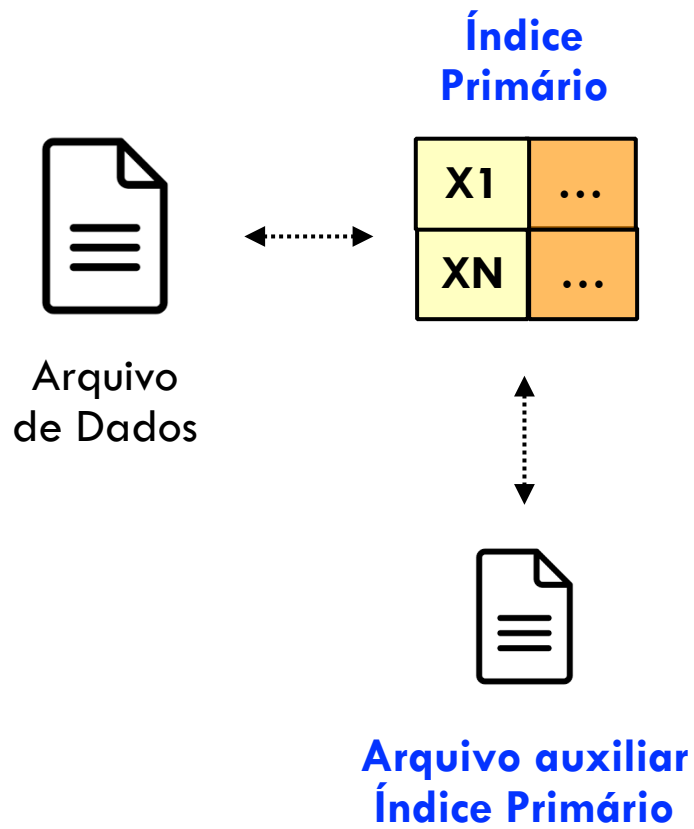
BEETHOVEN	ANG3795
BEETHOVEN	DG139201
BEETHOVEN	DG18807
BEETHOVEN	RCA2626

chave **secundária** não precisa  
ser única

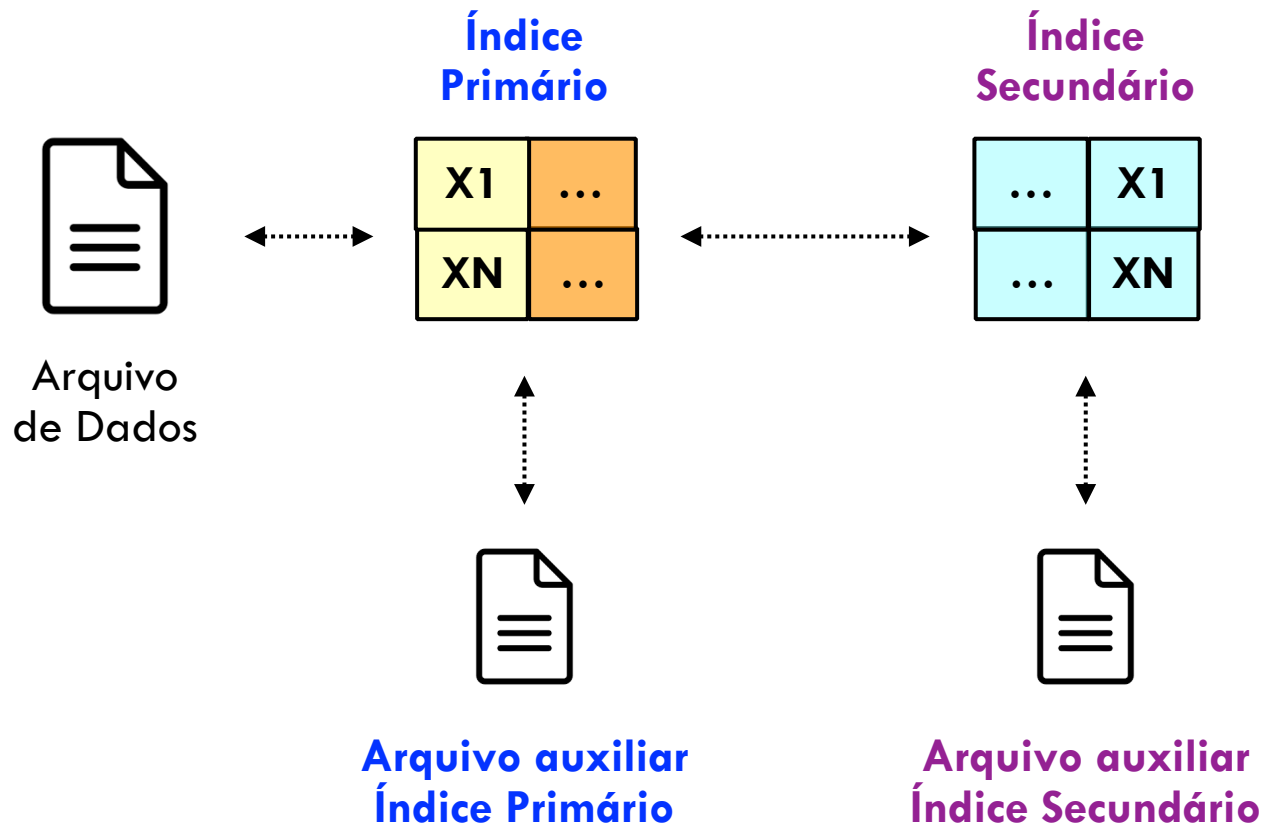
RIMSKY-KORSAKOV	MER75016
SPRIGNSTEEN	COL38358
SWEET HONEY IN THE R	FF245

**Idx Secundário: organizado por compositor**

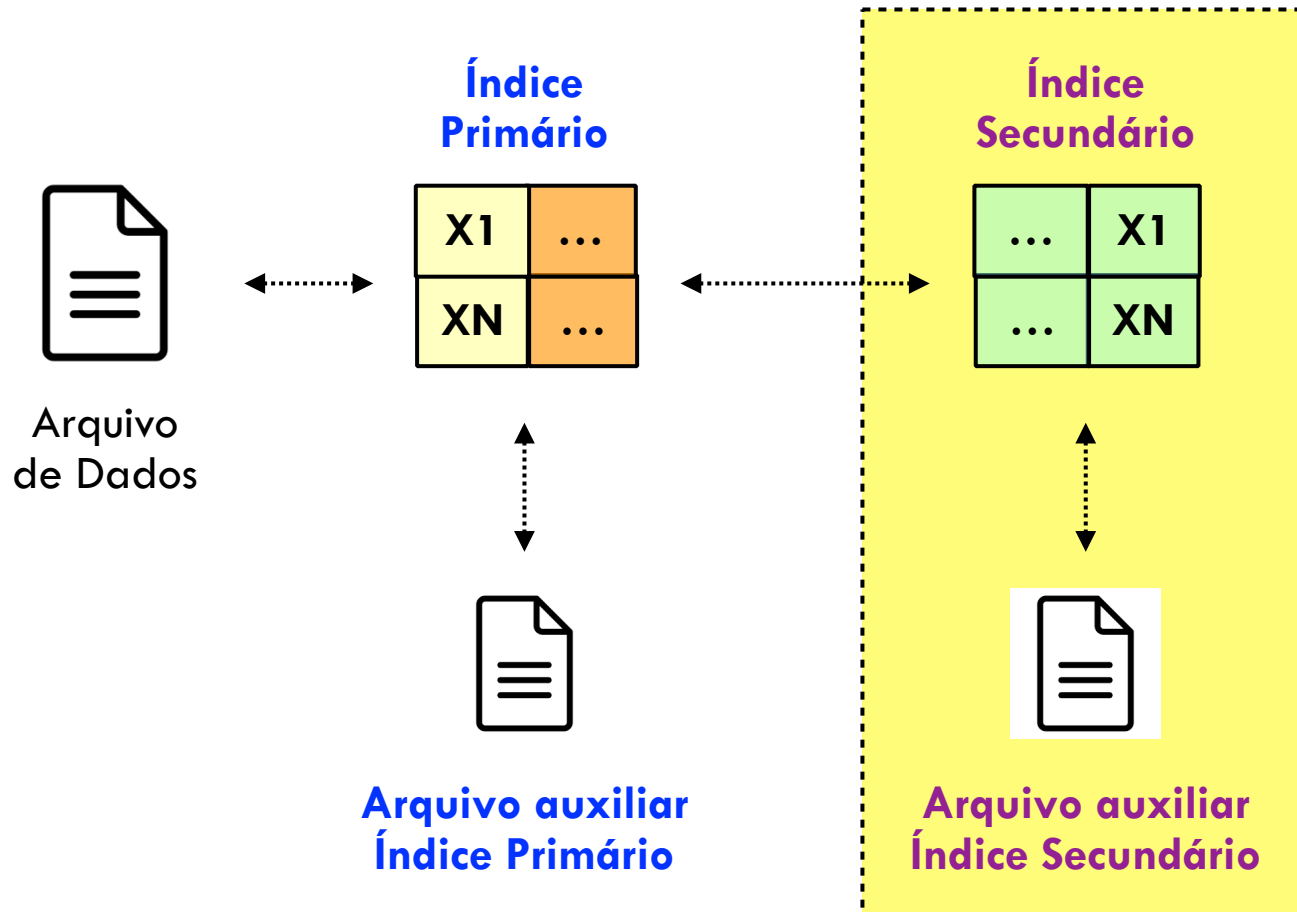
# Índice Secundário



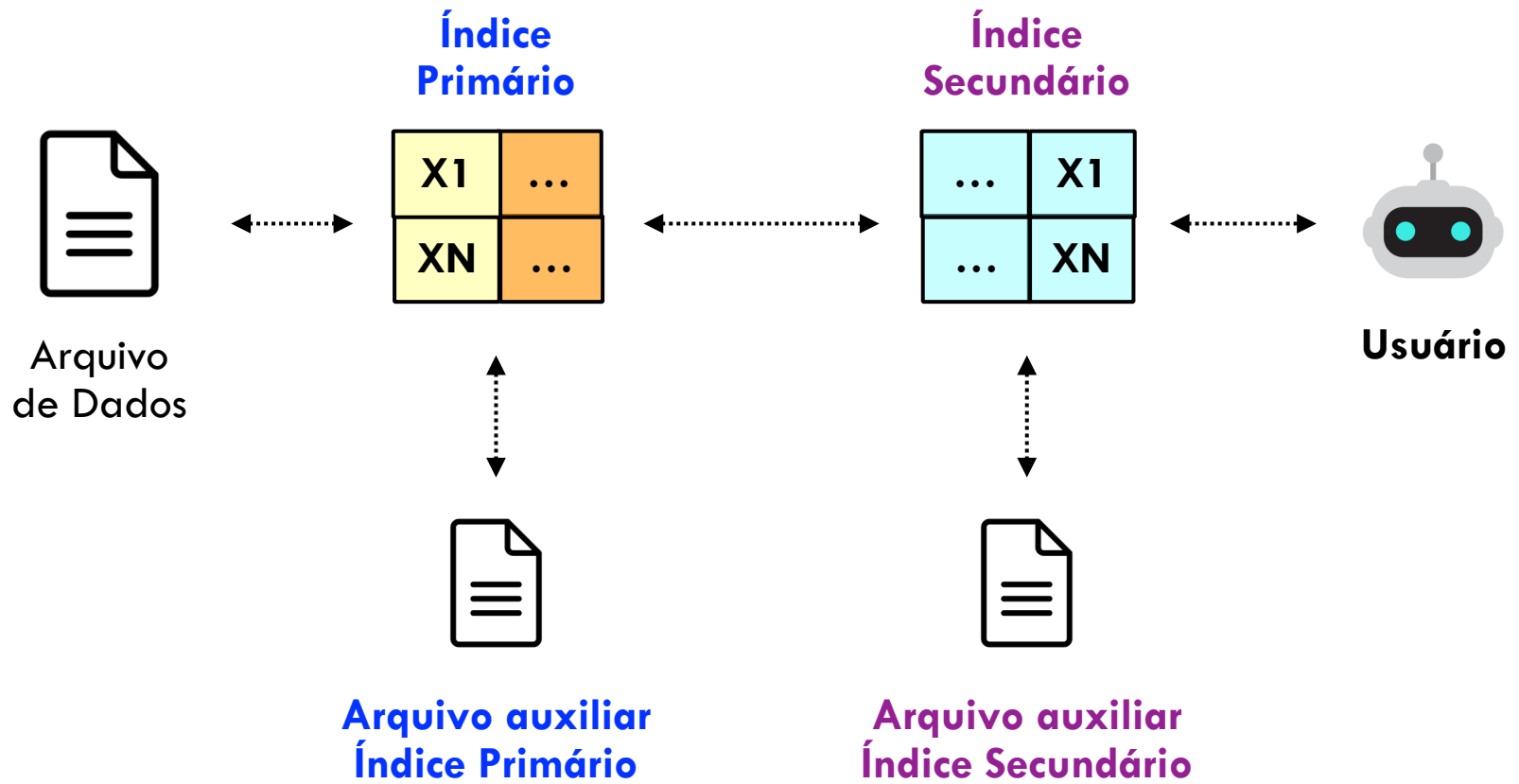
# Índice Secundário



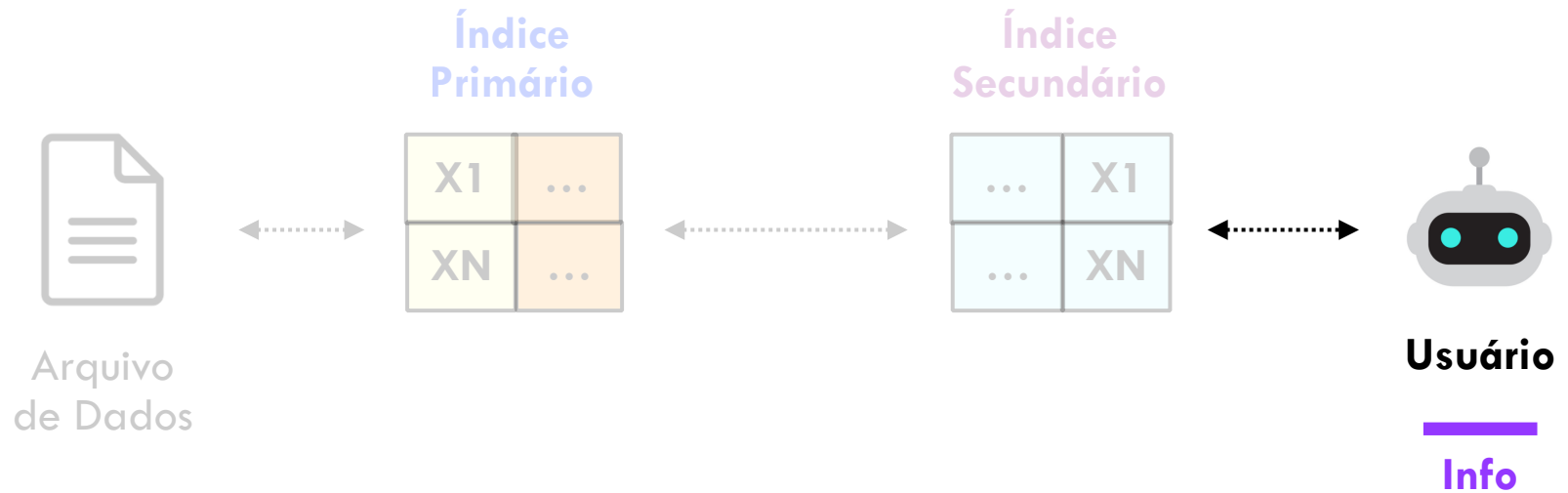
# Índice Secundário



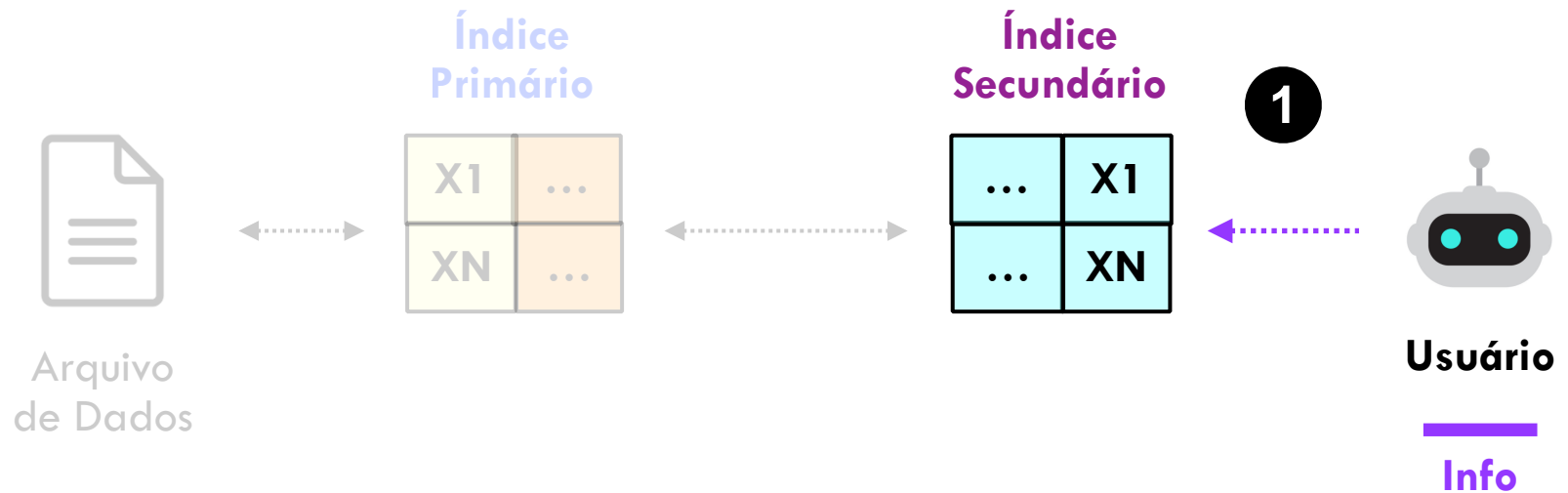
# Índice Secundário



# Índice Secundário

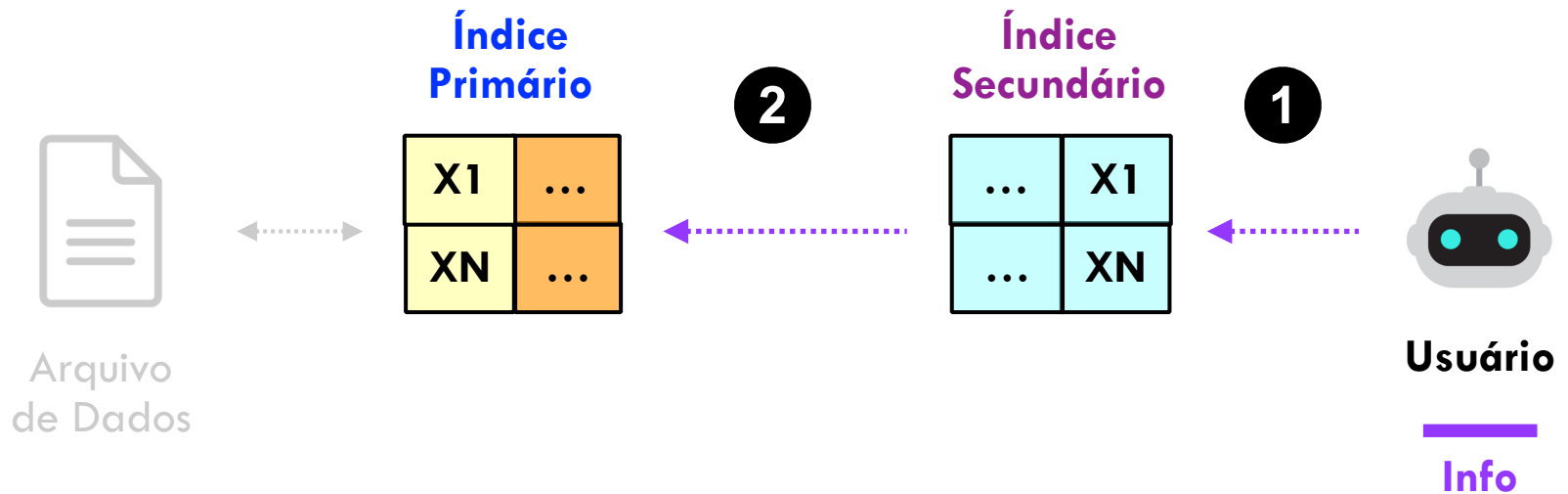


# Índice Secundário

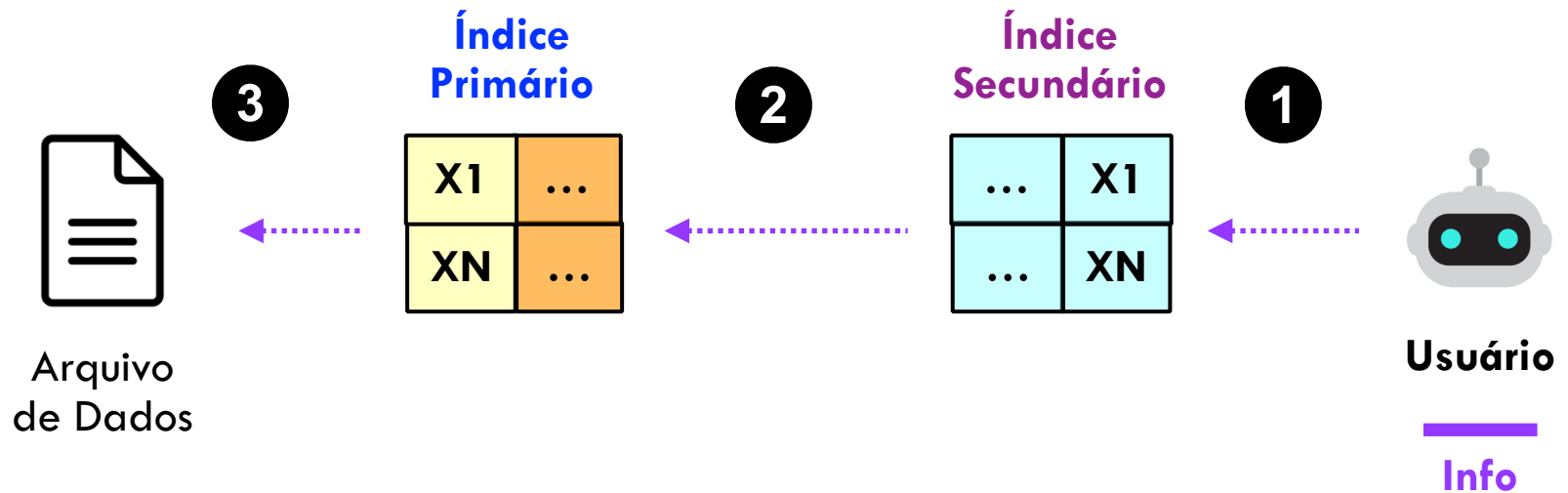




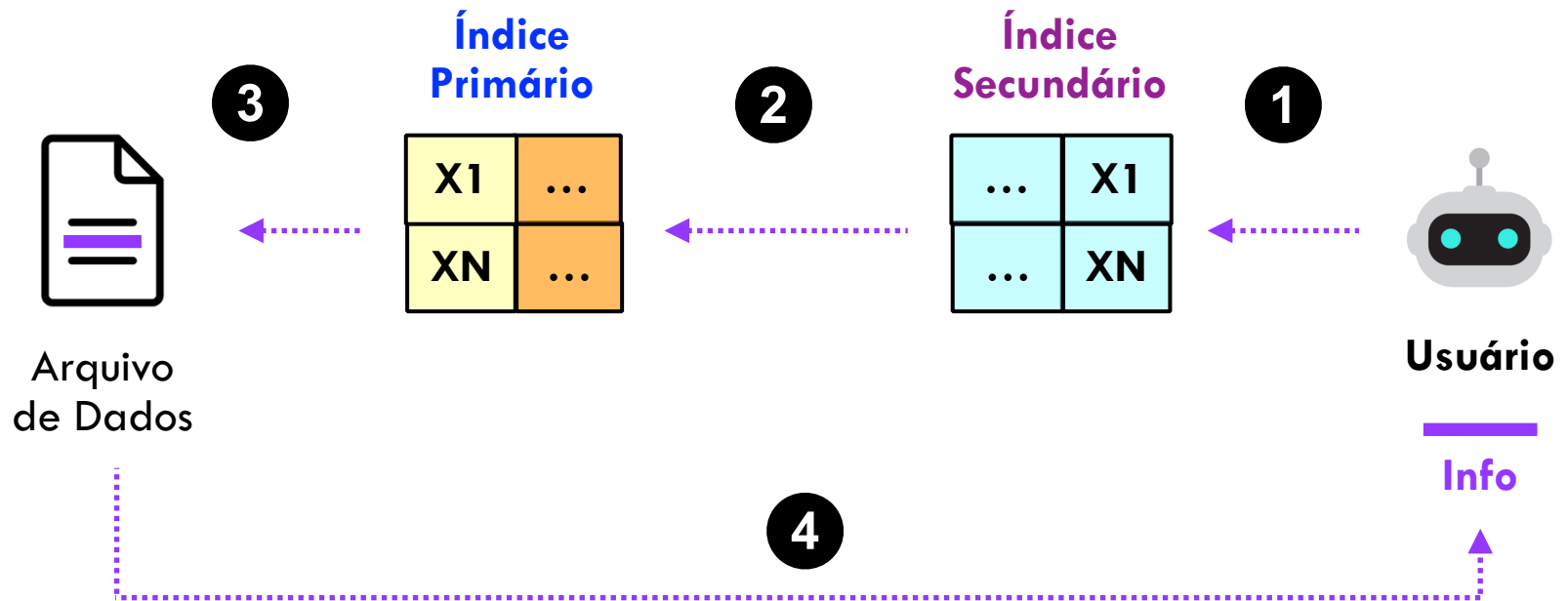
# Índice Secundário



# Índice Secundário

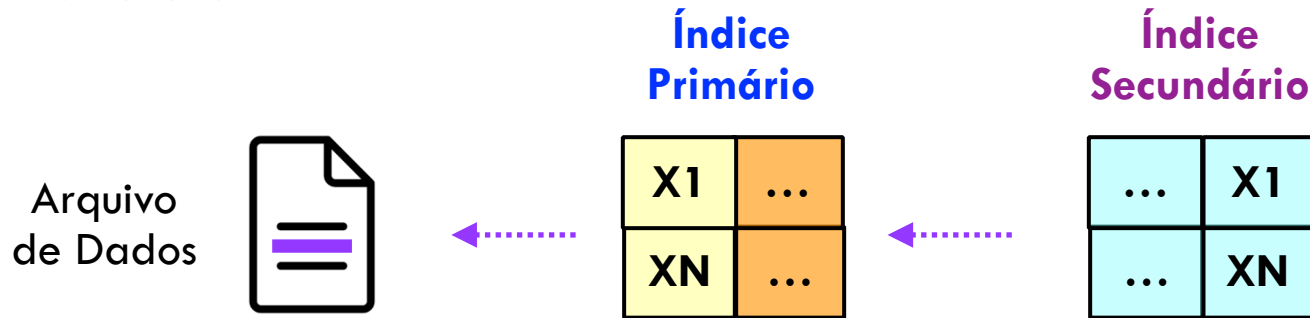


# Índice Secundário



# Índice Secundário

## □ Estrutura:

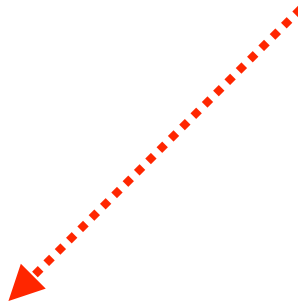


# Índice Secundário

## □ Estrutura:

Índice  
Primário

X1	...
XN	...

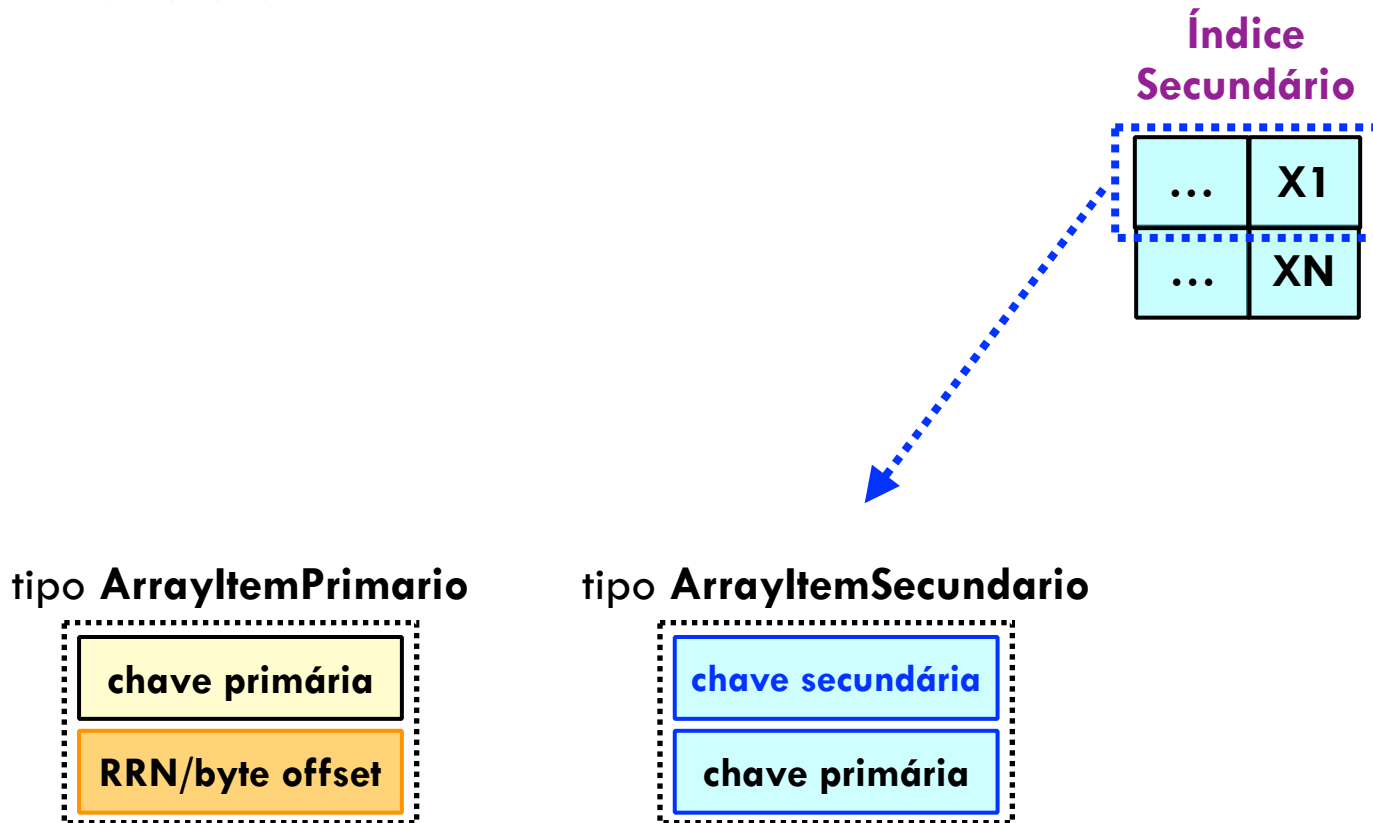


tipo **ArrayItemPrimario**

chave primária
RRN/byte offset

# Índice Secundário

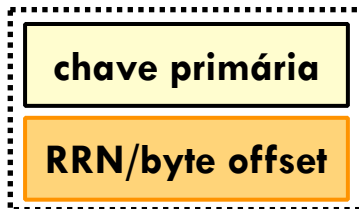
## □ Estrutura:



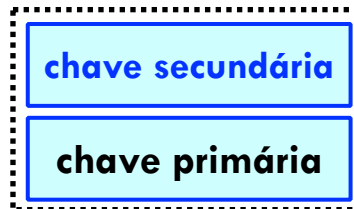
# Índice Secundário

## □ Estrutura:

tipo **ArrayItemPrimario**



tipo **ArrayItemSecundario**

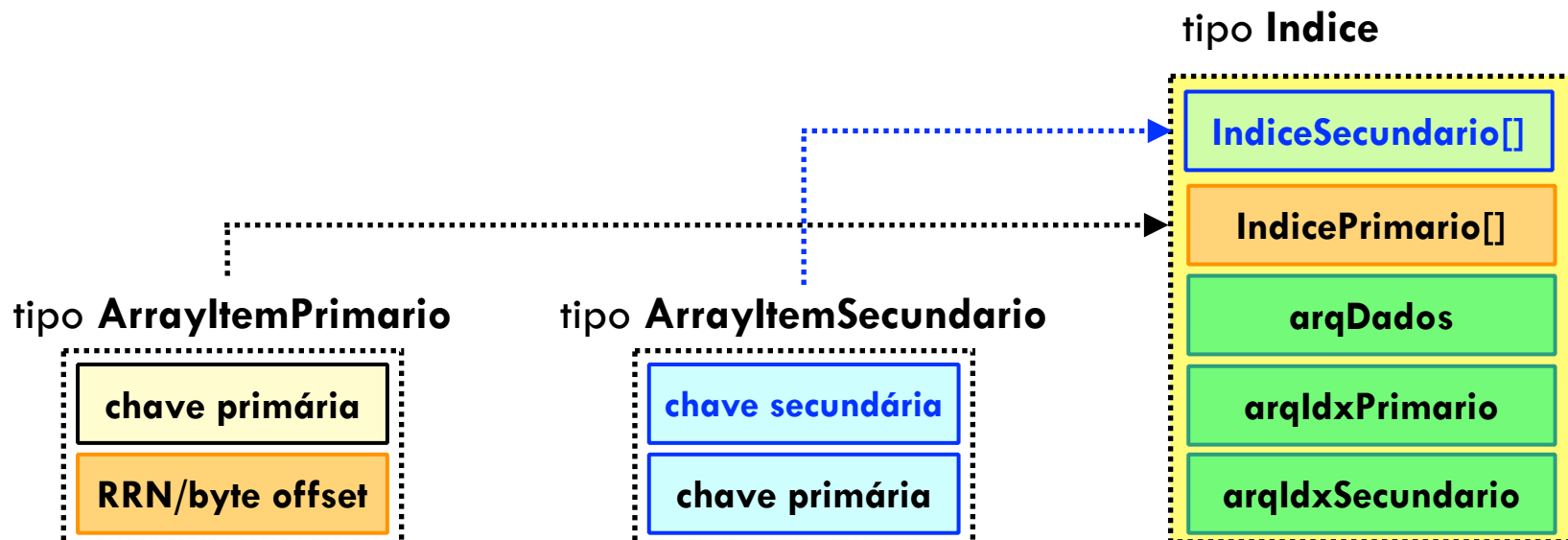


tipo **Indice**



# Índice Secundário

## □ Estrutura:





# Índice Secundário

## □ Estrutura:

índice secundário ordenado (RAM)

índice primário ordenado (RAM)

arquivo com os registros (disco)

arquivo auxiliar com índice primário (disco)

arquivo auxiliar com índice secundário (disco)

tipo **Índice**

**ÍndiceSecundario[]**

**ÍndicePrimario[]**

**arqDados**

**arqIdxPrimario**

**arqIdxSecundario**

# Índice Secundário

## □ Estrutura:

Três tipos necessários: item para índice primário (**ArrayItemPrimario**), item para índice secundário (**ArrayItemSecundario**) e o índice (**Indice**). Índice contém três arquivos: um para os dados (registros), e outros dois auxiliares para gravar os índices. Além disso, mantém uma cópia do arrays de índices atualizada na memória (RAM) enquanto o programa executar.

índice secundário ordenado (RAM)  
índice primário ordenado (RAM)  
arquivo com os registros (disco)  
arquivo auxiliar com índice primário (disco)  
arquivo auxiliar com índice secundário (disco)

### tipo **Indice**



# Exercícios



- 1) Defina os tipos de dados necessários para codificar uma estrutura de índices que use índice secundário.

# Roteiro



- 1 Introdução
- 2 Índices Secundários
- 3 Operações
- 4 Revisão
- 5 Referências

# Operações

**Criar (S)**

**Criar (S, arq)**

**Carregar (S)**

**Inserir (S, reg)**

**Remover (S, k)**

**Pesquisar (S, k)**

**AtualizarIdx (S)**

**AtualizarReg (S, reg)**

**Destruir (S)**

# Operações

**Criar (S)**

Cria a estrutura vazia

**Criar (S, arq)**

Cria a estrutura de um arquivo com dados já existentes

**Carregar (S)**

Carrega os **arquivos** de **índices** na memória antes de usá-los

**Inserir (S, reg)**

Insere um registro no arquivo (e **Índices**)

**Remover (S, k)**

Remove um registro do arquivo (e **Índices**)

**Pesquisar (S, k)**

Procura um registro no arquivo (e **Índices**)

**AtualizarIdx (S)**

Atualiza os **Índices** da memória para refletir as mudanças no arquivo

**AtualizarReg (S, reg)**

Atualiza um registro (e **Índices**)

**Destruir (S)**

Destrói a estrutura, grava os **Índices** nos arquivos auxiliares

**Operações de modificação**

# Operações

**lerRegistroComRRN (arq, RRN)**

ler registro no arquivo baseado no RRN

**lerRegistroByteOffset (arq, byteOffset)**

ler registro no arquivo baseado no byte offset

**numeroRegistros (arq)**

ler número de registros do cabeçalho do arquivo (header)

**ordenaArrayIdx (arrayIdx[])**

Ordena o Índice na memória

**gravalIdxArqAux(arrayIdx[], arqIdx)**

Salva o Índice da memória para o arquivo auxiliar (disco)

...

**Operações de auxiliares**

# Operações

**Criar (S)**

**Criar (S, arq)**

**Carregar (S)**

**Inserir (S, reg)**

**Remover (S, k)**

**Pesquisar (S, k)**

**AtualizarIdx (S)**

**AtualizarReg (S, reg)**

**Destruir (S)**

**Operações de modificação**

**Operações de auxiliares**

**lerRegistroComRRN (arq, RRN)**

**lerRegistroByteOffset (arq, byteOffset)**

**numeroRegistros (arq)**

**ordenaArrayIdx (arrayIdx[])**

**gravalIdxArqAux(arrayIdx[], arqlIdx)**

...



# Pesquisa

**Pesquisar ( $S$ ,  $k$ )**

: baseada na chave de busca ( $k$ )

# Pesquisa

**Pesquisar ( $S$ ,  $k$ )**

: baseada na chave de busca ( $k$ )

*/\*\* Ideia principal do algoritmo \*/*

# Criar

**Criar (S)**

: cria Índice juntamente do arquivo de dados  
(só o cabeçalho é gerado)

**Criar (S, arq)**

: cria Índice baseado em um arquivo já existente

# Criar

**Criar (S)**

: cria Índice juntamente do arquivo de dados  
(só o cabeçalho é gerado)

**Criar (S, arq)**

: cria Índice baseado em um arquivo já existente

*/\*\* Ideia principal do algoritmo \*\*/*

# Inserção

**Inserir (S, reg)**

: inserir um registro no arquivo de dados. Modifica também os **Índices**

# Inserção

**Inserir (S, reg)**

: inserir um registro no arquivo de dados. Modifica também os **Índices**



# Inserção

**Inserir (S, reg)**

: inserir um registro no arquivo de dados. Modifica também os **Índices**



**Inserção de um novo  
registro no arquivo  
de dados**

# Inserção

**Inserir (S, reg)**

: inserir um registro no arquivo de dados. Modifica também os **Índices**



1:1



X	0
..	..



Inserção de um novo  
registro no arquivo  
de dados

Inserção de um novo  
registro no Índice Primário  
e arquivo de idx primário



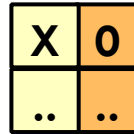
# Inserção

**Inserir (S, reg)**

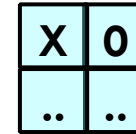
: inserir um registro no arquivo de dados. Modifica também os **Índices**



1:1



1:1



Inserção de um novo  
registro no arquivo  
de dados

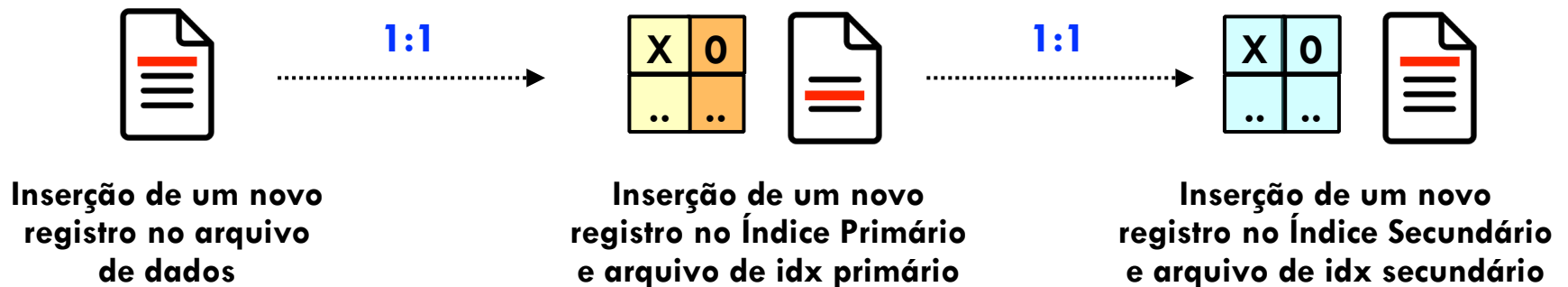
Inserção de um novo  
registro no Índice Primário  
e arquivo de idx primário

Inserção de um novo  
registro no Índice Secundário  
e arquivo de idx secundário

# Inserção

**Inserir (S, reg)**

: inserir um registro no arquivo de dados. Modifica também os **Índices**



Manter Índice Primário e Secundário **ordenados** de acordo com **forma canônica** das chaves

# Inserção

**Inserir (S, reg)**

: inserir um registro no arquivo de dados. Modifica também os **Índices**

```
/** Ideia principal do algoritmo */
```

# Remoção

**Remover ( $S, k$ )**

: remove registros dos **Índices** devido às remoções no arquivo de dados

# Remoção

**Remover ( $S, k$ )** : remove registros dos **Índices** devido às remoções no arquivo de dados

*/\*\* Ideia principal do algoritmo \*/*

# Atualização/Update

**AtualizarReg (S, reg)**

: modifica registros nos **Índices** devido às modificações no arquivo de dados

# Atualização/Update

**AtualizarReg (S, reg)**

: modifica registros nos **Índices** devido às modificações no arquivo de dados

**A**

update muda a **chave secundária**

# Atualização/Update

**AtualizarReg (S, reg)**

: modifica registros nos **Índices** devido às modificações no arquivo de dados

*/\*\* Ideia principal do algoritmo - update chave secundária \*\*/*



# Atualização/Update

**AtualizarReg (S, reg)**

: modifica registros nos **Índices** devido às modificações no arquivo de dados

**A**

update muda a chave secundária

**B**

update muda a **chave primária**

# Atualização/Update

**AtualizarReg (S, reg)**

: modifica registros nos **Índices** devido às modificações no arquivo de dados

*/\*\* Ideia principal do algoritmo - update chave primária \*\*/*

# Atualização/Update

**AtualizarReg (S, reg)**

: modifica registros nos **Índices** devido às modificações no arquivo de dados

**A**

update muda a chave secundária

**B**

update muda a chave primária

**C**

update muda **outros campos**

# Atualização/Update

**AtualizarReg (S, reg)**

: modifica registros nos **Índices** devido às modificações no arquivo de dados

*/\*\* Ideia principal do algoritmo - update muda outros campos \*/*

# Exercícios

Para os próximos exercícios, acesse o seguinte arquivo e adicione um registro com a informação de uma música de seu gosto: **Arquivo Músicas**

Cada música deve possuir os seguintes campos:

**Musica**

<b>Ano</b>
<b>Duração</b>
<b>Título</b>
<b>Artista</b>
<b>Gênero</b>
<b>Idioma</b>

# Exercícios

Para os próximos exercícios, acesse o seguinte arquivo e adicione um registro com a informação de uma música de seu gosto: **Arquivo Músicas**

Cada música deve possuir os seguintes campos:

**Musica**

<b>Ano</b>	Ano de lançamento da música
<b>Duração</b>	Duração em min:seg
<b>Título</b>	Título da música
<b>Artista</b>	Artista/Cantor(a) que gravou a música
<b>Gênero</b>	Gênero da música
<b>Idioma</b>	Idioma da gravação

# Exercícios

A chave primária será composta pela combinação dos seguintes campos:

Ano + Duração

**Exemplo:**

Ano	Duração	Título	Artista	Gênero	Idioma
-----	---------	--------	---------	--------	--------

1977	4:14	Hell Ain't a Bad Place to Be	ACDC	rock	inglês
------	------	------------------------------	------	------	--------

Chave canônica = 1977414

# Exercícios

02) Crie o **Índice Primário** do arquivo de registros de músicas, usando o RRN de cada entrada como a referência de leitura da informação. Ao final do processo, lembre-se de manter o Índice Primário ordenado de acordo com as chaves primárias.



# Exercícios

03) Crie um **Índice Secundário** do arquivo de registros de músicas, usando o campo **ARTISTA** como chave secundária. Associe cada chave secundária com suas correspondente chave primária. Lembre-se de manter o Índice Secundário ordenado ao final do processo.

# Exercícios

04) Crie um **Índice Secundário** do arquivo de registros de músicas, usando o campo **GENERO** como chave secundária. Associe cada chave secundária com suas correspondente chave primária. Lembre-se de manter o Índice Secundário ordenado ao final do processo.

# Exercícios

05) Insira o seguinte registro no arquivo de dados e faça as alterações necessárias em todos os índices já criados:

2013 | 3:54 | I Sold My Bed, But Not My Stereo | Capital Cities | electro indie | inglês

# Exercícios

06) Remova o seguinte registro no arquivo de dados e faça as alterações necessárias em todos os índices já criados:

1977   4:14   Hell Ain't a Bad Place to Be   ACDC   rock   inglês
---

# Roteiro



- 1 Introdução
- 2 Índices Secundários
- 3 Operações
- 4 Revisão
- 5 Referências

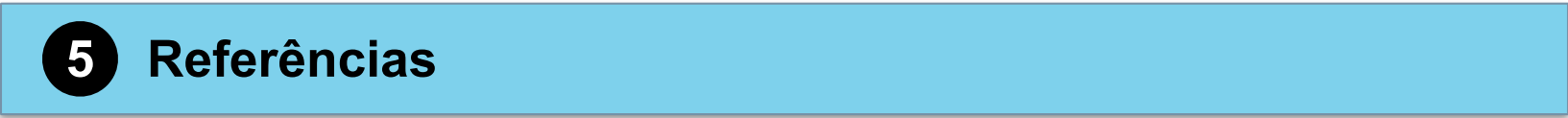
# Revisão



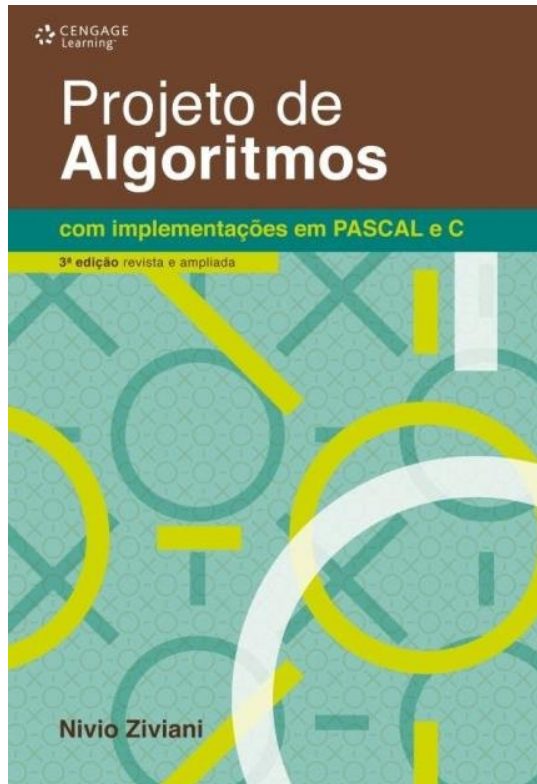
- **Índices Secundários**

# Roteiro

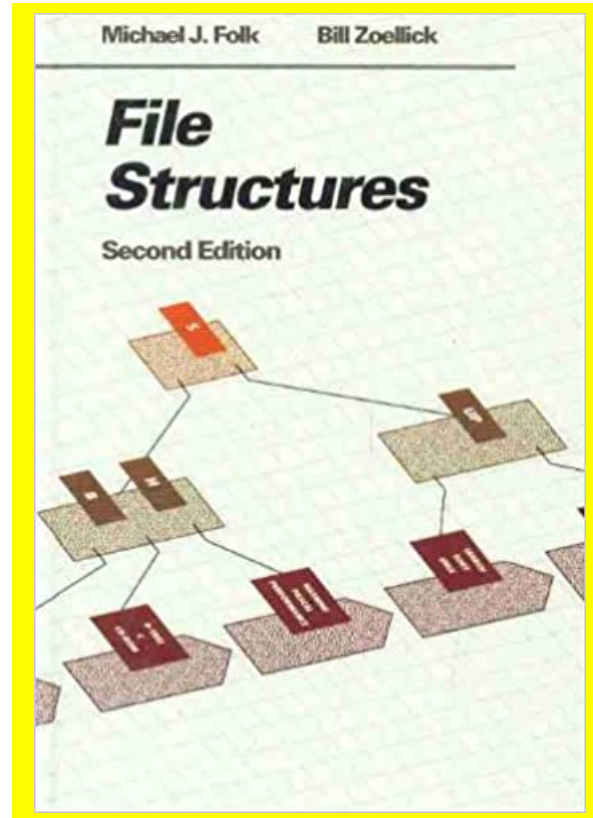


- 1 Introdução
  - 2 Índices Secundários
  - 3 Operações
  - 4 Revisão
  - 5 Referências
- 

# Referências sugeridas



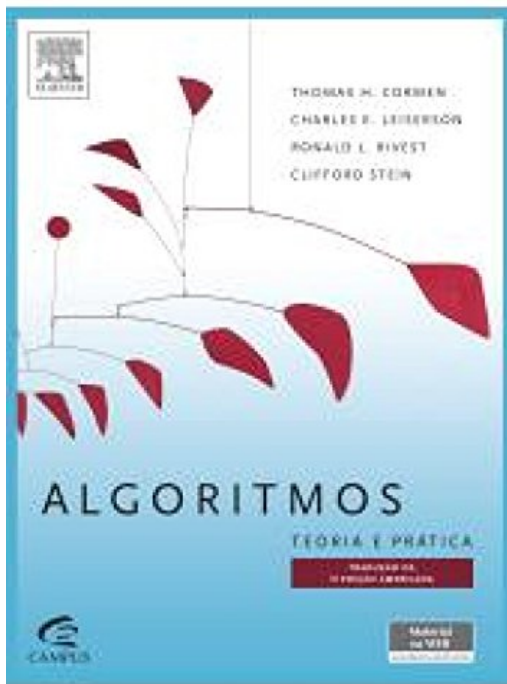
[Ziviani, 2010]



[Folk & Zoellick, 1992]



# Referências sugeridas



[Cormen et al, 2018]



[Drozdek, 2017]

# Perguntas?

Prof. Rafael G. **Mantovani**

[rafaelmantovani@utfpr.edu.br](mailto:rafaelmantovani@utfpr.edu.br)