

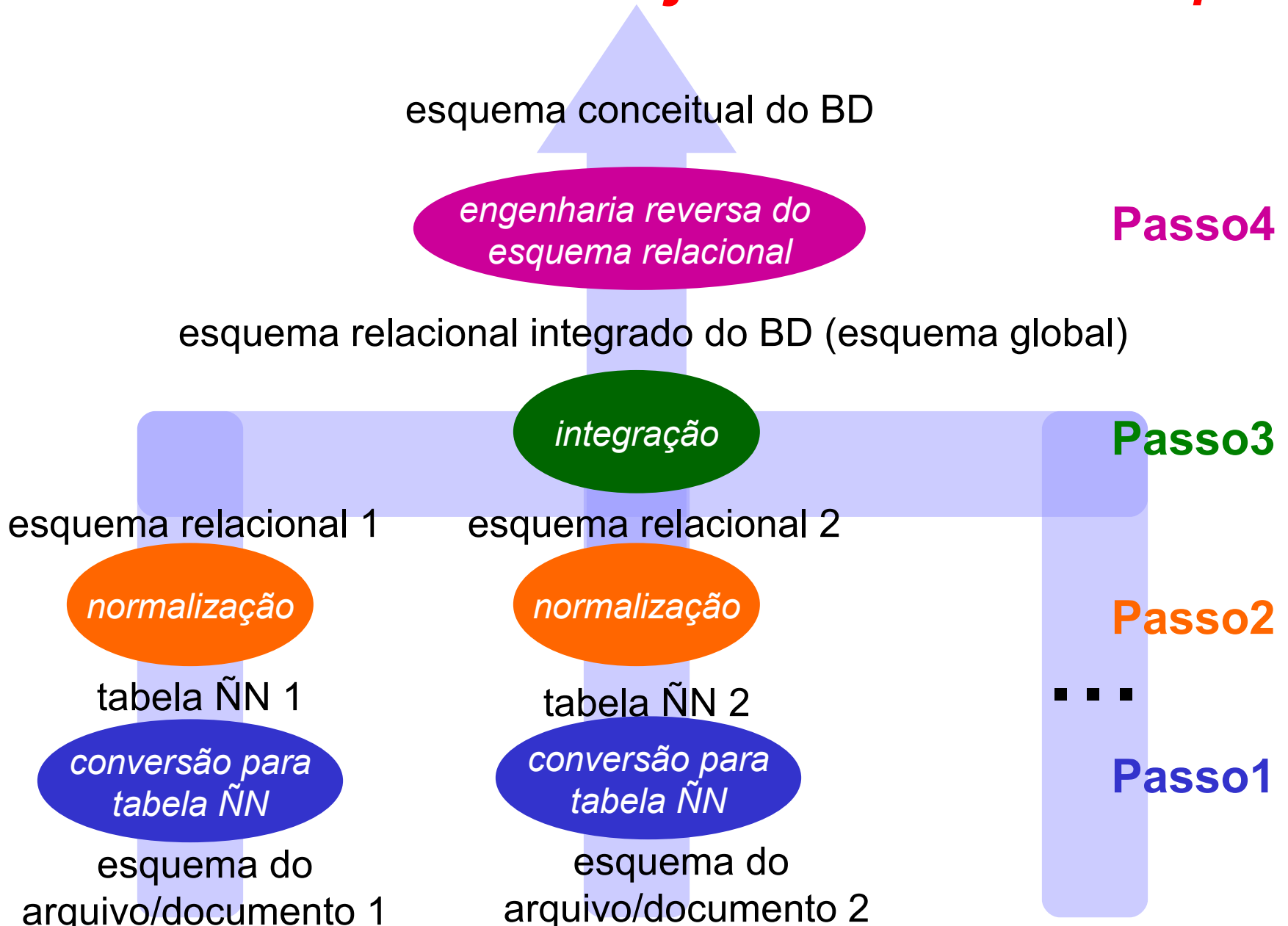
# Projeto *Bottom-Up* de BD

- Uma organização que deseja definir um BD pode já conter descrições de dados
  - sistemas legados
    - gerenciamento de arquivos de dados
  - ambientes não-informatizados
    - dados manuscritos
      - fichários, relatórios, documentos, ...
  - ...

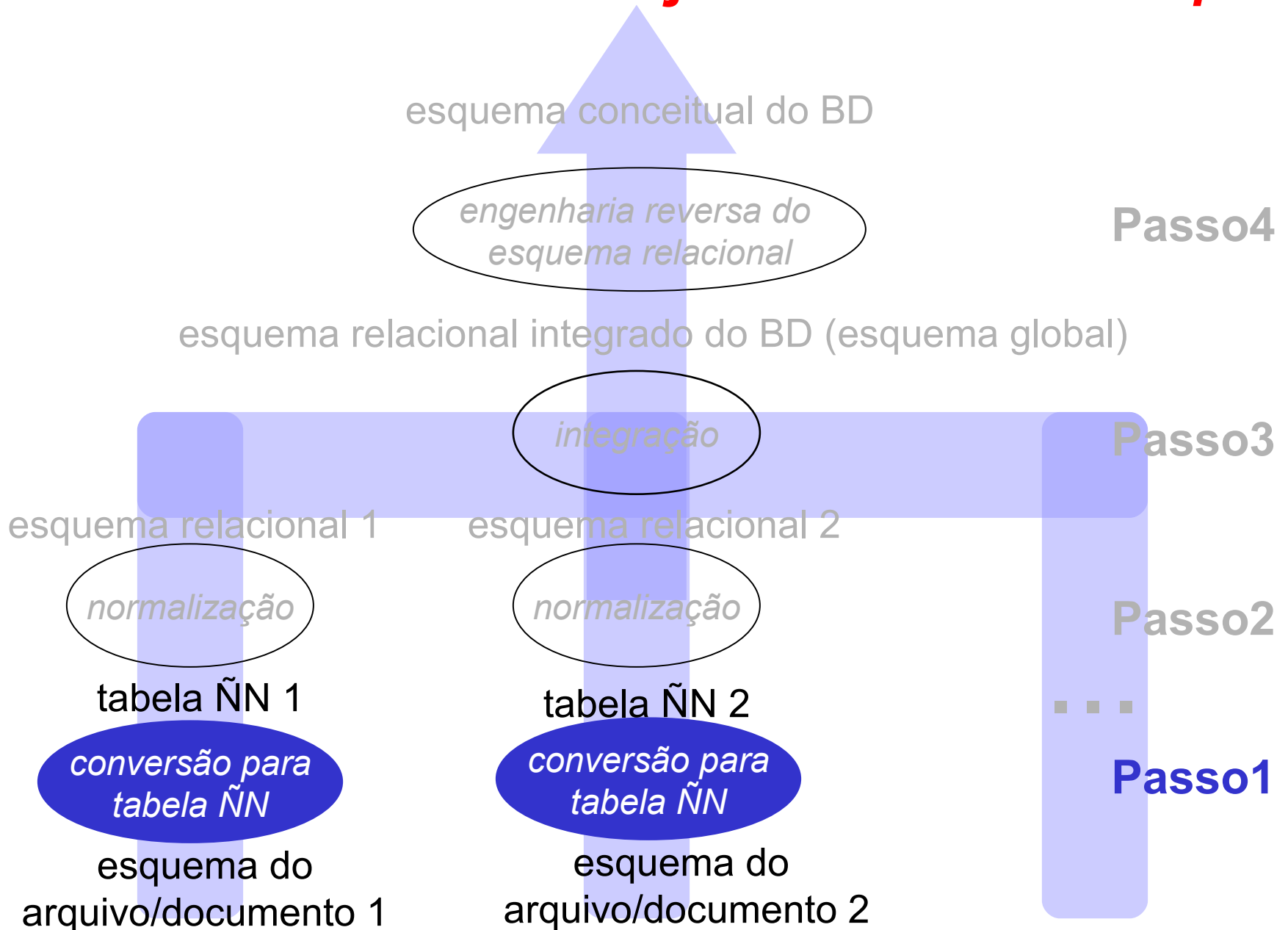
# Projeto *Bottom-Up* de BD

- Objetivo
  - construção dos esquemas conceitual e lógico do BD a partir de fontes de dados da organização
- Projeto *Bottom-Up* é também chamado de processo de Engenharia Reversa de BD
- Foco
  - engenharia reversa de BD relacional

# Processo de Projeto *Bottom-Up*



# Processo de Projeto *Bottom-Up*



# Passo1 – Conversão para Tabela ÑN

- Objetivo
  - obtenção de uma descrição da fonte de dados independente do seu formato
    - relatório, arquivo, ...
  - padronização da representação das fontes para facilitar o processo de projeto
- Forma de representação: Tabela ÑN
  - pode conter tabelas aninhadas e/ou atributos multivalorados
    - representam dados que se repetem em um fato sendo analisado

# Exemplo de Tabela ÑN: Relatório de Alocações de Projetos

CódProj	Tipo	Descr	Emp					
			CodEmp	Nome	Cat	Sal	DataIni	TempAl
LSC001	Novo Desenv.	Sistema de Estoque	2146	João	A1	4	1/11/91	24
			3145	Sílvio	A2	4	2/10/91	24
			6126	José	B1	9	3/10/92	18
			1214	Carlos	A2	4	4/10/92	18
			8191	Mário	A1	4	1/11/92	12
PAG02	Manutenção	Sistema de RH	8191	Mário	A1	4	1/05/93	12
			4112	João	A2	4	4/01/91	24
			6126	José	B1	9	1/11/92	12

ocorrência de tabela aninhada em uma linha de projeto

# Representação na Forma de Tabela ÑN

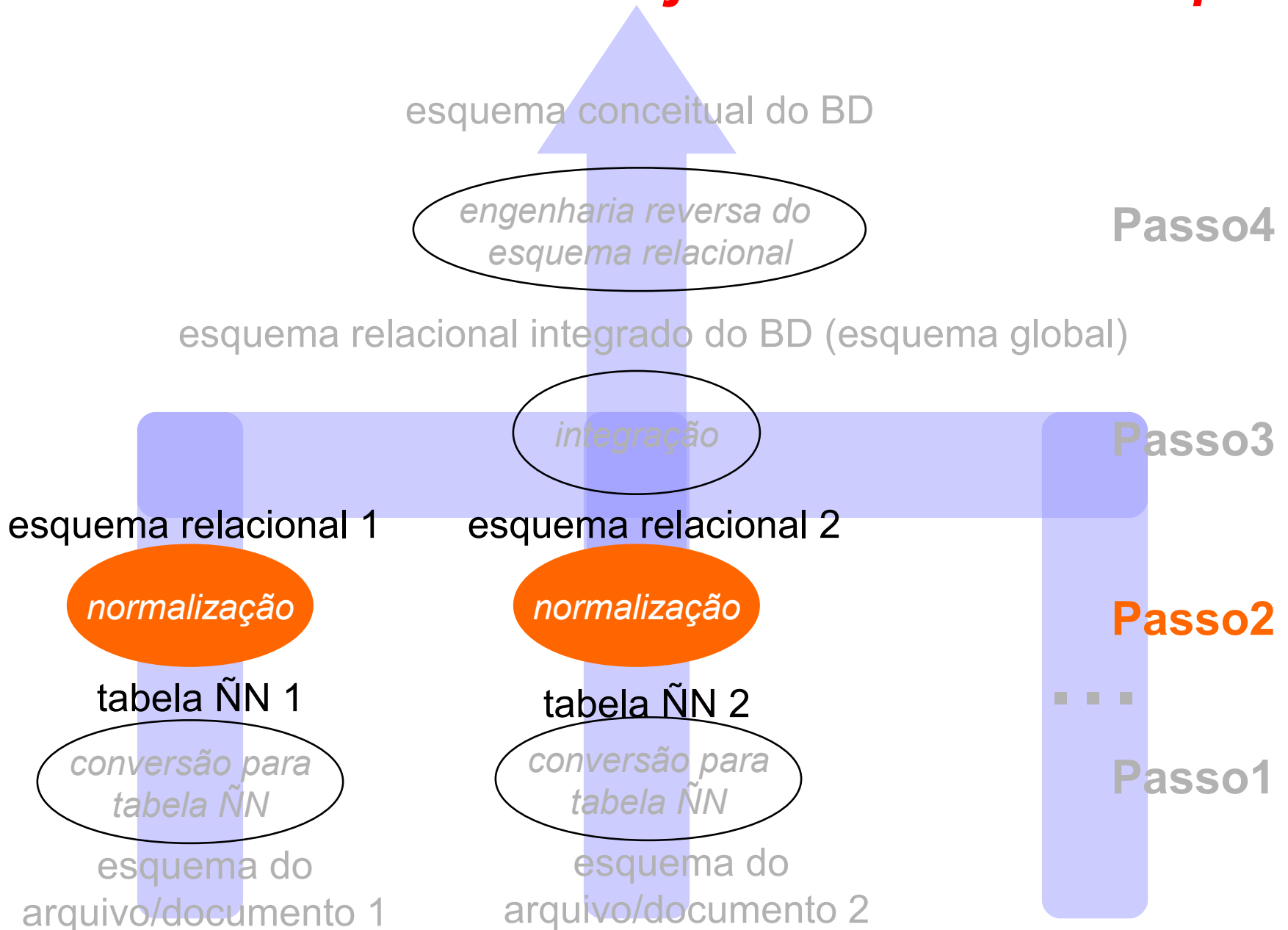
Projetos (codProj, tipo, descr,  
(codEmp, nome, categ, sal, dataIni, tempoAloc))



indicam-se as chaves primárias (CPs) de cada tabela

Observação: a determinação da chave primária de uma tabela aninhada leva em conta a análise dos dados de cada ocorrência de aninhamento e não do conjunto de ocorrências como um todo

# Processo de Projeto *Bottom-Up*





# Processo de Normalização

- Objetivos

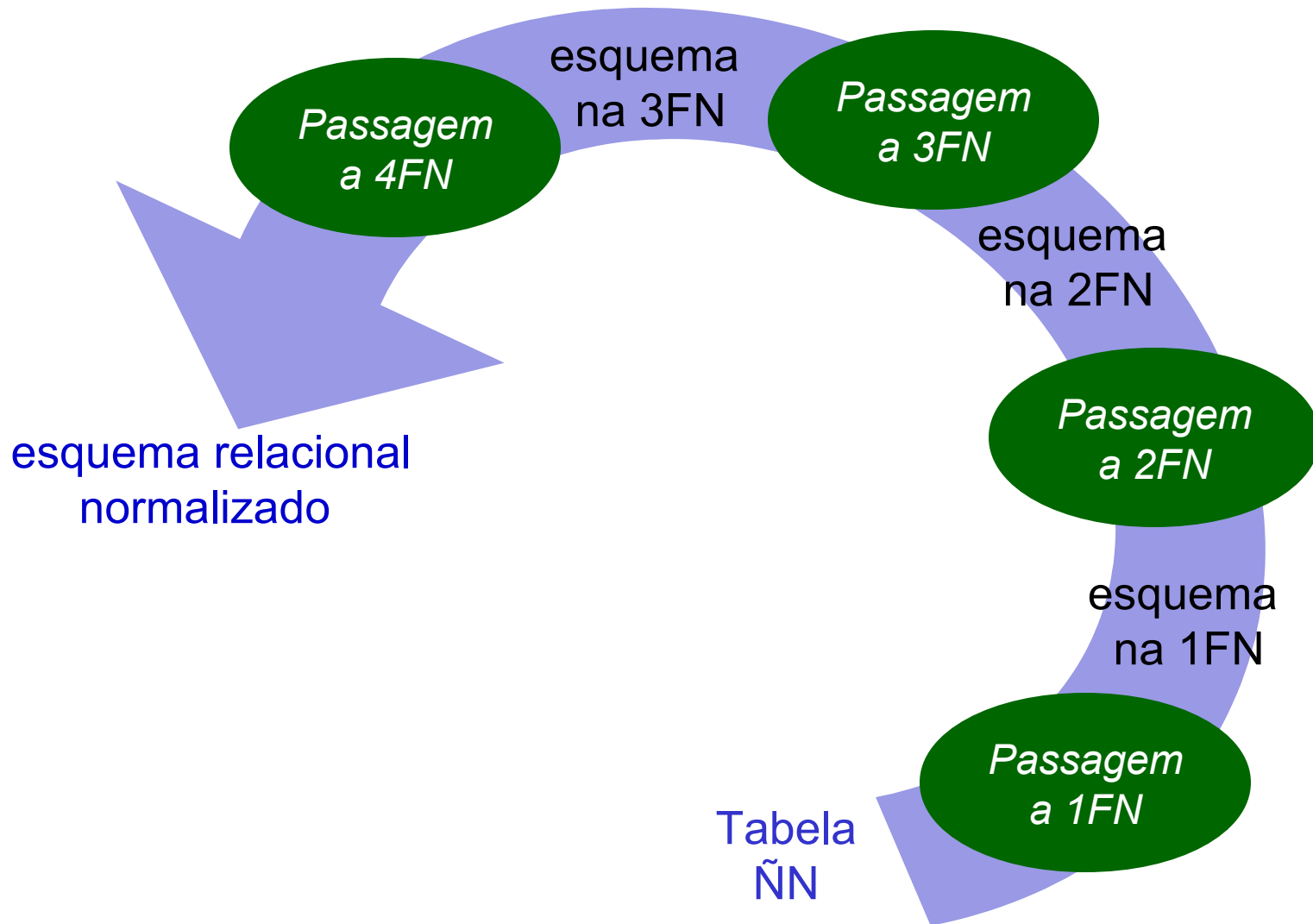
- eliminar redundâncias em uma tabela e evitar controles de atualização

- geração de novas tabelas não-redundantes

- agrupar dados relevantes em uma mesma tabela de modo a constituir uma futura entidade ou relacionamento do esquema conceitual

- análise de *dependências funcionais*

# Processo de Normalização



# 1ª Forma Normal (1FN)

- Objetivo: remoção de aninhamentos
- Exemplos

a) ÑN: Projetos (codProj, tipo, descr,  
(codEmp, nome, categ, sal, dataIni, tempoAloc))



1FN: Projetos (codProj, tipo, descr)  
Alocações (codProj, codEmp, nome, categ, sal,  
dataIni, tempoAloc)

b) ÑN: Departamentos (codD, nome, (CPF, nome, salário))



1FN: Departamentos (codD, nome)  
Empregados (codD, CPF, nome, salário)

# 2ª Forma Normal (2FN)

- Objetivo: remoção de dependências funcionais parciais da chave primária com atributos não-chave
- Exemplo

1FN: Alocações (codProj, codEmp, nome, categ, sal, dataIni, tempoAloc)

**DFs:**  $\text{codEmp} \rightarrow \text{nome, categ, sal}$  (DF Parcial!)  
 $(\text{codProj}, \text{codEmp}) \rightarrow \text{dataIni, TtempoAloc}$



2FN: Alocações (codProj, codEmp, dataIni, tempoAloc)  
Empregados (codEmp, nome, categ, sal)

# 3ª Forma Normal (3FN)

- Objetivo: remoção de dependências funcionais entre atributos não-chave
- Exemplo

2FN: Empregados (codEmp, nome, categ, sal)

**DFs:**  $categ \rightarrow sal$




3FN: Empregados (codEmp, nome, **categ**)  
CategoriasSalariais(categ, **Sal**)

# 4ª Forma Normal (4FN)

- Objetivo: remoção de dependências funcionais multivaloradas
  - aplica-se a tabelas em que todos os atributos fazem parte da chave primária e existem pelo menos 3 atributos
- Exemplo

3FN: DistribuiçãoLivrosAutor(nroEstante, ISBN, codAutor)

**DF Multivalorada:** ISBN  $\twoheadrightarrow$  codAutor



4FN: DistribuiçãoLivros(nroEstante, ISBN)  
Autoria(ISBN, codAutor)

# Normalização – Questões

- Inclusão de chave primária
  - Para tabelas que não possuam atributos que garantam identificação única de suas tuplas

ÑN: Projetos (codProj, tipo, descr,  
(nome, categ, sal, dataIni, tempoAloc))



ÑN: Projetos (codProj, tipo, descr,  
(codEmp, nome, categ, sal, dataIni, tempoAloc))

# Normalização – Questões

- Exclusão de atributos irrelevantes
  - não necessitam ser mantidos no BD

ÑN: Projetos (codProj, tipo, descr, nroEmps, dataRel,  
(codEmp, nome, categ, sal, dataIni, tempoAloc))



ÑN: Projetos (codProj, tipo, descr,  
(codEmp, nome, categ, sal, dataIni, tempoAloc))



# Normalização – Questões

- Inclusão de atributos relevantes implícitos

*a ordem dos dados na  
tabela aninhada determina  
a classificação do candidato*

ÑN: Aprovação (codCurso, nome,  
(codCand, nome, endereço))



ÑN: Aprovação (codCurso, nome,  
(codCand, nome, endereço, ordemClass))

# Normalização – Questões

- Validação de relacionamentos incorretos

## Relatório de Pedidos

ÑN: Pedidos (nroPed, dataPed, codCli, nomeCli,  
(nroTel), (codPeça, descrPeça, qtdePedida))



1FN: Pedidos (nroPed, dataPed, codCli, nomeCli)  
Telefones (nroPed, nroTel)  
Peças (nroPed, codPeça, nomePeça, qtde))



...

Validação: Telefones (codCli, nroTel)

# Normalização – Questões

- Validação de chaves primárias em tabelas associativas
  - evitar chaves primárias incorretas

## Relatório de Escalas de Passeios de Guias Turísticos

ÑN: Guias (codG, nomeG, celular,  
(codPas, nomeLocal, (data, hora, nroPessoas)))



1FN: Guias (codG, nomeG, celular)  
GuiasPasseios (codG, codPas, nomeLocal)  
Escalas (codG, codPas, data, hora, nroPessoas)

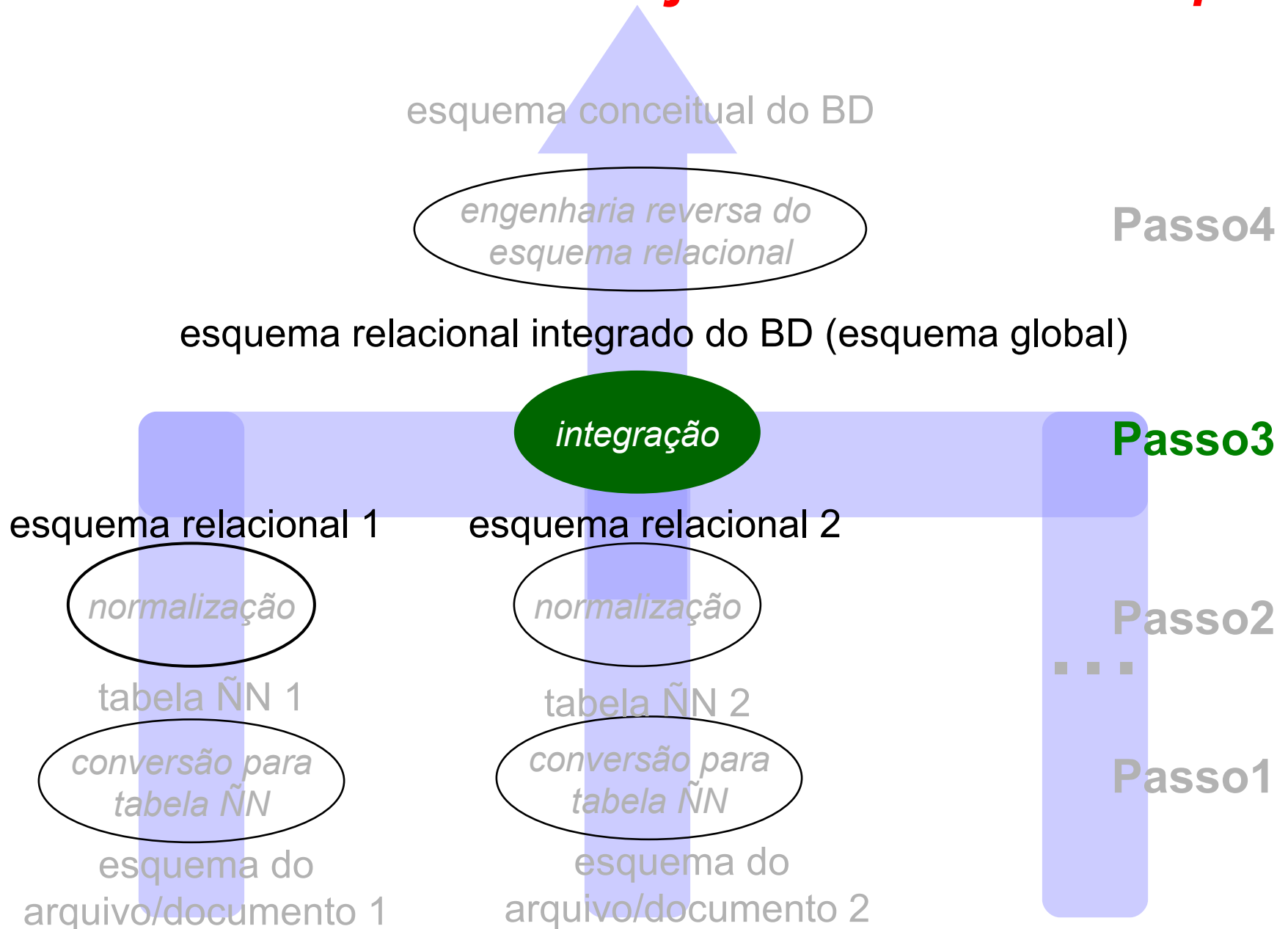
problema com esta chave primária: um guia pode conduzir mais de um passeio na mesma data/hora!



...

Validação: Escalas (codG, codPas, data, hora, nroPessoas)

# Processo de Projeto *Bottom-Up*



# Integração de Esquemas Relacionais

- Integração de tabelas resultantes da normalização
- Objetivos
  - integrar uma mesma entidade ou relacionamento que aparece em tabelas diferentes
  - eliminar atributos redundantes advindos em tabelas diferentes

# Integração de Esquemas Relacionais

- Etapas
  1. integração de tabelas com mesma CP
  2. integração de tabelas com chave contida
  3. verificação da 3FN
- Etapas 1 e 2 se aplicam a tabelas semanticamente equivalentes
  - entidades ou relacionamentos que representam o mesmo fato do mundo fato

# Integração de Tabelas com Mesma CP

- Tabelas que possuem CPs com os mesmos domínios e conteúdos

Tabela1: **Projeto** (CodProj, **Tipo**, **Descrição**)

Tabela2: **Proj** (Código, **Descr**, **DataIni**, **CodDepto**)



**EsquemaGlobal:**

**Projeto** (CodProj, **Tipo**, **Descr**, **DataIni**, **CodDepto**)

# Integração de Tabelas com CP contida

- Todos os atributos de uma tabela  $X$  são CP e esta CP está contida na CP de uma tabela  $Y$ 
  - tabela  $X$  pode ser removida

Tabela1: **DiscCursadas** (CodA, CodD)

Tabela2: **Histórico** (Aluno, Disc, Sem, Nota)



**EsquemaGlobal: Histórico**(Aluno, Disc, Sem, Nota)

**Observação:** estas 2 tarefas de integração podem também ser aplicadas durante a normalização de uma fonte de dados



# Integração de Tabelas com CP contida

- Estar atento a casos em que nomes são sinônimos mas os conteúdos são diferentes!

Tabela1: Matrícula (CodA, CodD)

Tabela2: Histórico (CodA, CodD, Sem, Nota)



*matrícula e histórico são  
relações ≠ entre alunos e  
disciplinas*

**EsquemaGlobal:**

**Matrícula (CodA, CodD)**

**Histórico (CodA, CodD, Sem, Nota)**

# Verificação da 3FN

- A integração de dados em uma tabela pode provocar dependências entre atributos não-chave

Tabela1: Depto (CodD, Nome, CodGerente)

Tabela2: Departamento (Código, Local, NomeGer)



EsquemaGlobal:

Departamento(Código, Nome, Local, CodGerente, NomeGerente)

*volta à 2FN!*

# Verificação da 3FN

**EsquemaGlobal:**

**Departamento(Código, Nome, Local, **CodGerente**,  
**NomeGerente**)**



*aplica-se novamente a 3FN!*

**EsquemaGlobal:**

**Departamento(Código, Nome, Local, **CodGerente**)**  
**Gerente(Codígo, Nome)**

**Observação:** a geração de uma nova tabela no esquema global exige que se verifiquem os demais casos de integração para ela!

# Exercício – Fonte Dados 1

Arquivo XML –  
Quartos.xml

```
<quartos>
  <quarto numero="101">
    <tipo ID="1">
      <descricao>duplo</descricao>
      <preco>110.00</preco>
    </tipo>
    ...
    <acessorio ID="15">frigobar</acessorio>
    <acessorio ID="22">TV</acessorio>
    <acessorio ID="35">banda larga</acessorio>
    ...
  </quarto>
  <quarto numero="102">
    <tipo ID="2">
      <descricao>simples</descricao>
      <preco>90.00</preco>
    </tipo>
    <acessorio ID="02">ar condicionado</acessorio>
    <acessorio ID="15">frigobar</acessorio>
    <acessorio ID="22">TV</acessorio>
    ...
  </quarto>
  ...
</quartos>
```

# Exercício – Fonte Dados 2

Data do Relatório: 17/06/07

## Relatório de Consumo de Hóspedes

Quarto: 101

Andar: 1

Hóspedes:

Nome: João Silva CPF: 1111100000 Entrada: 15/06/07 Saída Prevista: 17/06/07

Nome: Mário Sá CPF: 1111111110 Entrada: 15/06/07 Saída Prevista: 17/06/07

Consumo:

Data: 15/06/07	Item Cardápio: 01 – Almoço	Valor: R\$ 15,00
Data: 15/06/07	Item Cardápio: 22 – Refrigerante	Valor: R\$ 1,50
Data: 16/06/07	Item Cardápio: 06 – Caviar	Valor: R\$ 18,50
	Total:	<u>R\$ 35,00</u>

Quarto: 102

Andar: 1

Hóspedes:

Nome: Maria Souza CPF: 1010101010 Entrada: 14/06/07 Saída Prevista: 18/06/07

Consumo:

Data: 15/06/07	Item Cardápio: 22 – Refrigerante	Valor: R\$ 1,50
Data: 15/06/07	Item Cardápio: 22 – Refrigerante	Valor: R\$ 1,50
Data: 16/06/07	Item Cardápio: 01 – Almoço	Valor: R\$ 15,00
Data: 17/06/07	Item Cardápio: 22 – Refrigerante	Valor: R\$ 1,50
	Total:	<u>R\$ 19,50</u>

Quarto: . . .

---

**Observação:** supor que o BD não irá manter histórico de hospedagens.