



UNIVERSIDADE FEDERAL
DE SANTA CATARINA



Persistência Poliglota:

Fundamentos, Desafios & Oportunidades



CENTRO TECNOLÓGICO
Universidade Federal de Santa Catarina

Ronaldo dos Santos Mello

GBD/PPGCC/INE/CTC/UFSC



Programa de
Pós-Graduação
em Ciência da
Computação
U F S C

Agenda



1. Persistência Poliglota: O que é?!
2. SGPD: Arquiteturas de base
3. SGPD: Desafios & Oportunidades de Pesquisa

Agenda

- 1. Persistência Poliglota: O que é?!**
2. SGPD: Arquiteturas de base
3. SGPD: Desafios & Oportunidades de Pesquisa

Persistência Poliglota: O que é?!

● Motivação

- dados hoje são gerenciados em **diferentes contextos/requisitos**
 - dados simples X dados complexos
 - dados E x SE x ÑE
 - consistência X escalabilidade
 - *batch* X tempo real
 - *write-intensive* X *read-intensive* workloads
 - acesso baseado em chave/filtros (OLTP) X consultas analíticas (OLAP) . . .

Polyglot Data Management: State of the Art & Open Challenges

Felix Kiehn
University of Hamburg, Germany
felix.kiehn@uni-hamburg.de

Mareike Schmidt
University of Hamburg, Germany
mareike.schmidt-3@uni-hamburg.de

Daniel Glake
University of Hamburg, Germany
Hamburg University of Applied



Persistência Poliglota: O que é?!

● Motivação

- dados hoje são gerenciados em diferentes contextos/requisitos
 - dados simples X dados complexos
 - dados E x SE x ÑE
 - consistência X escalabilidade
 - *batch* X tempo real
 - *write-intensive* X *read-intensive* workloads
 - acesso baseado em chave/filtros (OLTP) X consultas analíticas (OLAP) . . .
- cada contexto/requisito \Rightarrow 1 ou poucas tecnologias de BD (SGBDs)
 - relacional, OR, XML, chave-valor, grafo, documento, *text search engine*, *array*

Polyglot Data Management: State of the Art & Open Challenges

Felix Kiehn
University of Hamburg, Germany
felix.kiehn@uni-hamburg.de

Mareike Schmidt
University of Hamburg, Germany
mareike.schmidt-3@uni-hamburg.de

Daniel Glake
University of Hamburg, Germany
Hamburg University of Applied



Persistência Poliglota: O que é?!

- **Motivação**

- aplicações geralmente utilizam 1 tecnologia de BD
- mas e se a minha aplicação precisa gerenciar dados em múltiplos contextos/requisitos?



Persistência Poliglota: O que é?!

● Motivação

- mas e se a minha aplicação precisa lidar com dados em múltiplos contextos/requisitos?
- cenário exemplo: **e-commerce**
 - registros de produtos
 - descrições textuais de produtos
 - fotos/vídeos de produtos
 - avaliações de produtos



Persistência Poliglota: O que é?!

● Motivação

- mas e se a minha aplicação precisa lidar com dados em múltiplos contextos/requisitos?
- cenário exemplo: **e-commerce**
 - dados com **naturezas** diferentes:
 - registros de produtos (**estruturado simples**)
 - descrição textual do produto (**não-estruturado**)
 - conteúdo do produto (p.ex: livro) (**estruturado complexo ou semiestruturado**)
 - fotos e vídeos do produto (**não-estruturado**)
 - rede de recomendação cliente-produto ou produto-produto ou cliente-cliente (**estruturado complexo**)
 - dados para *big data analytics* (**DW, DL**)



Persistência Poliglota: O que é?!

● Motivação

- mas e se a minha aplicação precisa lidar com dados em múltiplos contextos/requisitos?
- cenário exemplo: **e-commerce**
 - múltiplos **requisitos de gerência de dados**:
 - ♦ acesso a avaliações de produtos (alta disponibilidade; consistência fraca; *OLTP-read-intensive*)
 - ♦ acesso a produtos (alta disponibilidade; consistência forte; *OLTP-read-intensive*)
 - ♦ vendas (consistência forte; *OLTP-write-intensive*)
 - ♦ análise de perfis de usuários, vendas e recomendação de produtos (*OLAP*)

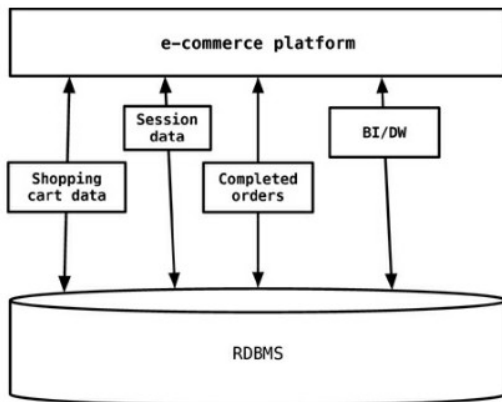


Persistência Poliglota: O que é?!

● Motivação

- mas e se a minha aplicação precisa lidar com dados em múltiplos contextos/requisitos?
- cenário exemplo: *e-commerce*

cenário atual
de gestão
de dados ➡

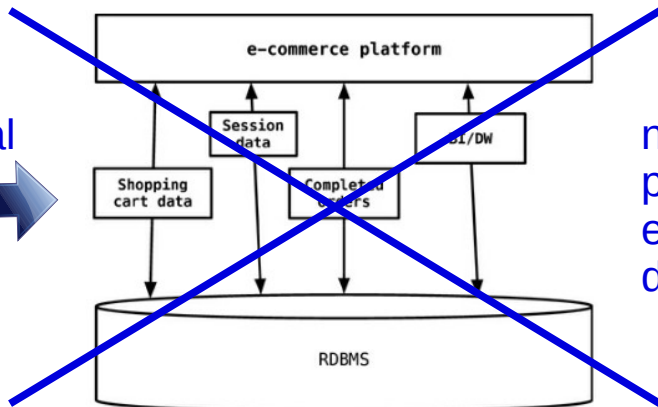


Persistência Poliglota: O que é?!

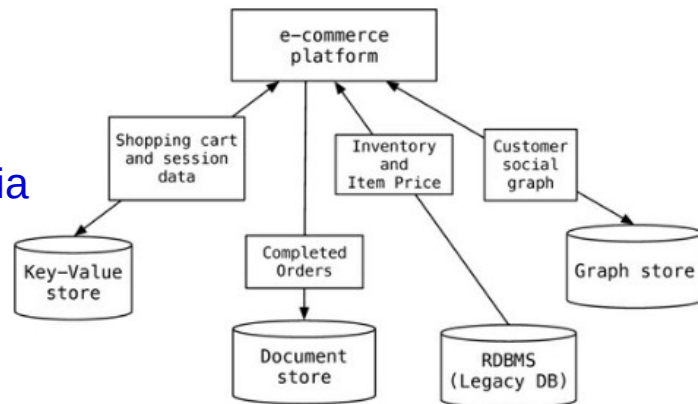
● Motivação

- mas e se a minha aplicação precisa lidar com dados em múltiplos contextos/requisitos?
- cenário exemplo: *e-commerce*

cenário atual
de gestão
de dados



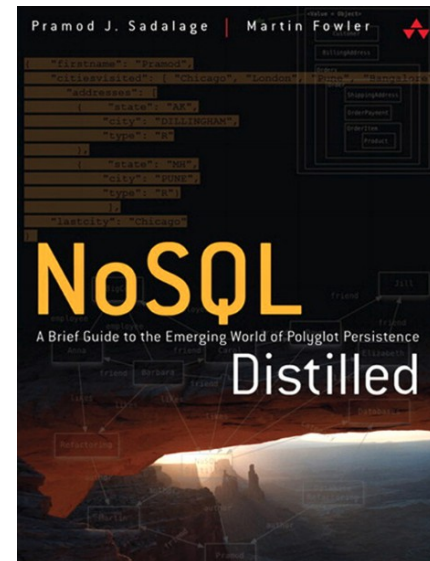
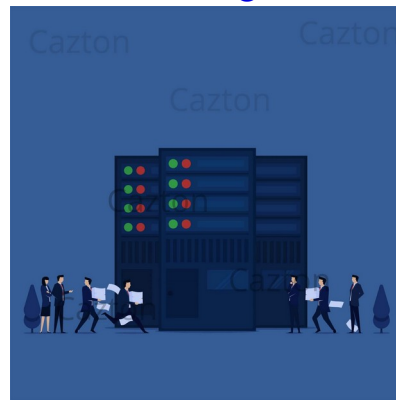
novo cenário
para persistência
e gestão
de dados



Persistência Poliglota: O que é?!

● Persistência Poliglota

- *"Combinar benefícios de diferentes tecnologias de BD evitando seus pontos fracos e deixando transparente a coordenação dessas tecnologias"* (Sadallage & Fowler, 2012)
- materializado como um SGPD (sistema de gerência poliglota de dados)



Persistência Poliglota: O que é?!

- **SGPD – Principais requisitos a considerar**

1. **Poder de modelagem** (modelos de dados suportados & naturezas de dados suportadas)

Persistência Poliglota: O que é?!

- **SGPD – Principais requisitos a considerar**

1. Poder de modelagem (modelos de dados suportados & naturezas de dados suportadas)
2. **Expressividade de linguagens/operações de acesso** (capacidades DDL e DML dos modelos de dados: *schema/schemaless*; restrições de integridade; capacidades de consulta: *key-based*, *set-based*, *navigate paths*, operadores não-relacionais – p.ex: espaço-temporais, *arrays*, grafos – agrupamento/inserção *on demand* de *streams* de dados; 1 ou + interfaces de acesso para esses modelos)

Persistência Poliglota: O que é?!

- **SGPD – Principais requisitos a considerar**

1. Poder de modelagem (modelos de dados suportados & naturezas de dados suportadas)
2. Expressividade de linguagens/operações (capacidades DDL e DML dos modelos de dados: *schema/schemaless*; restrições de integridade; capacidades de consulta: *key-based*, *set-based*, *navigate paths*, operadores não-relacionais – p.ex: espaço-temporais, *arrays*, grafos – agrupamento/inserção *on demand* de *streams* de dados; 1 ou + interfaces de acesso para esses modelos)
3. **Adaptação fácil e transparente a mudanças de contextos/requisitos**
(acionamento de novas tecnologias de BD com novos modelos devido a mudanças nos requisitos de dados da aplicação; migração/replicação de dados entre modelos de BD devido a mudanças no *workload*; replanejamento de acesso devido a mudanças no *workload*; capacidades de adaptação automáticas – *self-managed* SGPD (*machine learning??*))

Agenda

1. Projeto Tradicional de um BD
2. Persistência Poliglota: O que é?!
- 3. SGPD: Arquiteturas de base**
4. SGPD: Desafios & Oportunidades de Pesquisa

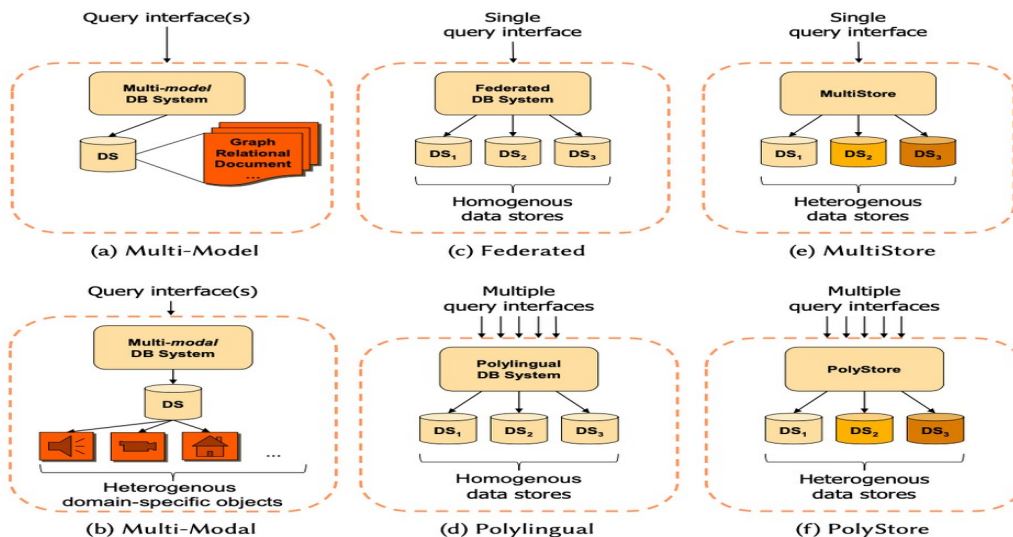
Arquiteturas de Base

- **Diversas abordagens na literatura possuem semelhança com a ideia de um SGPD**
 - heterogeneidade em termos de arquitetura (falta de padronização...)

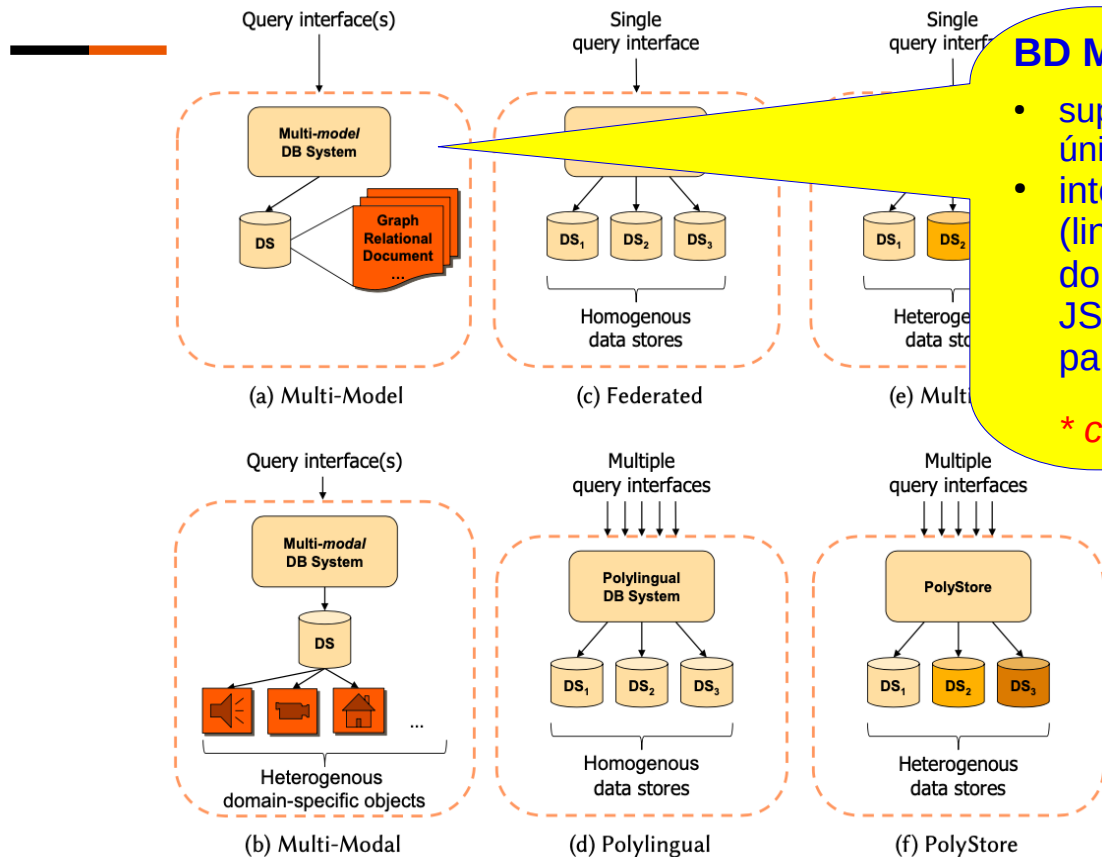


Arquiteturas de Base

- Diversas abordagens na literatura possuem semelhança com a ideia de um SGPD
 - heterogeneidade em termos de arquitetura (falta de padronização...)



Arquiteturas de Base

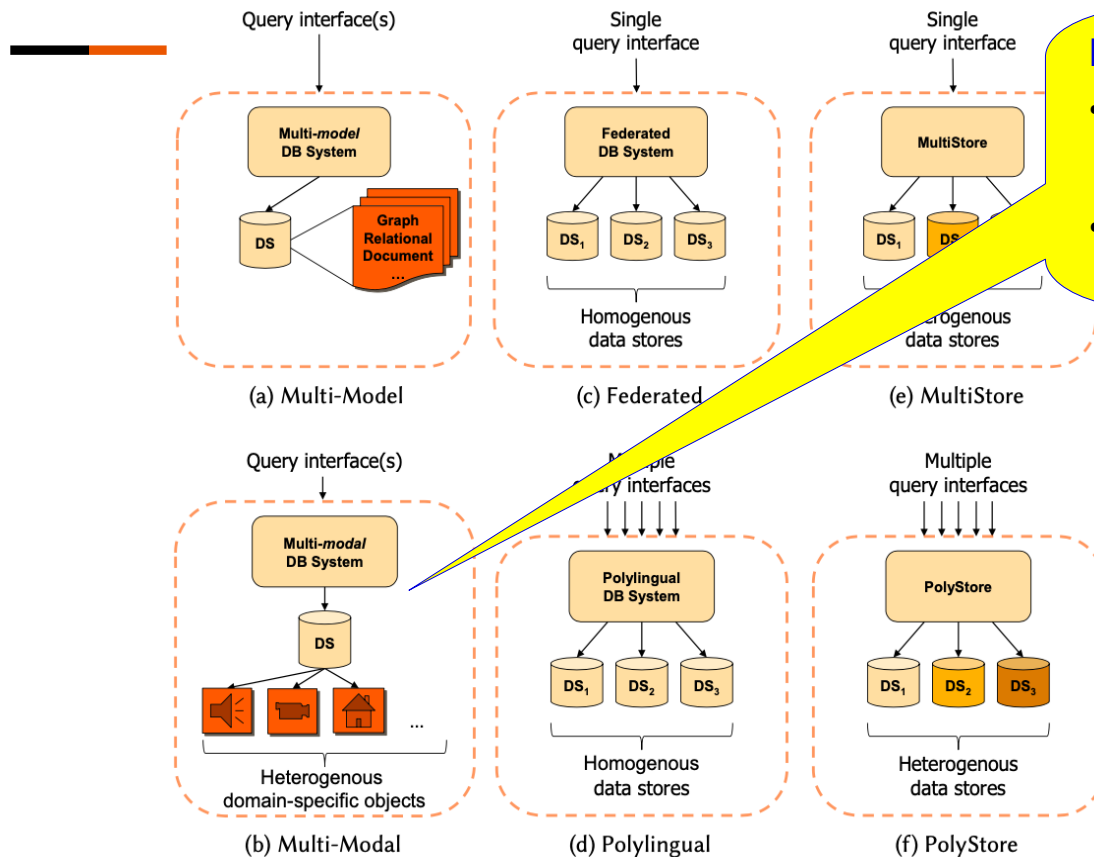


BD Multimodelo

- suporte a vários modelos de BD em um único SGBD
- interface de acesso única (linguagem multimodelo – p.ex: *AQL* do *ArangoDB*: filtro sobre estruturas JSON e grafo)* ou interfaces de acesso para cada modelo de dados

* *caso mais comum*

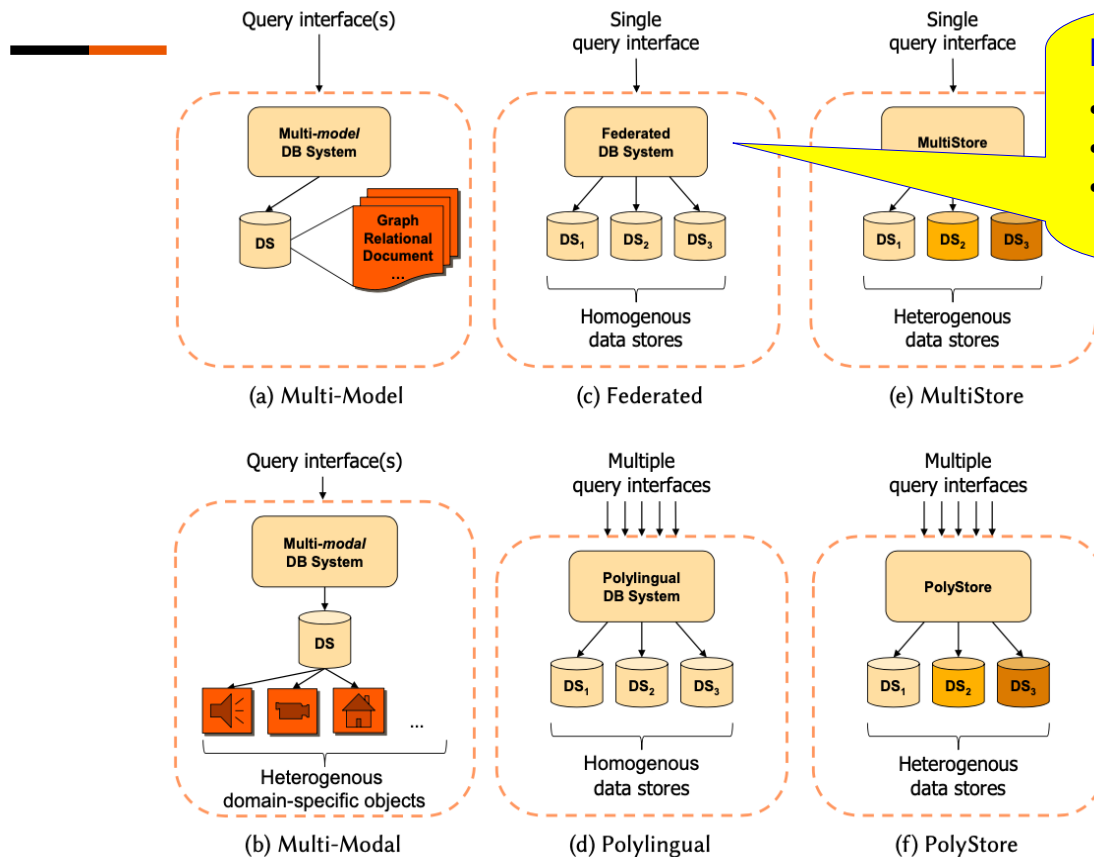
Arquiteturas de Base



BD Multimodal

- organiza dados não por modelo de BD, mas por natureza do dado (texto, imagem, vídeo, ...)
- interfaces de acesso específicas por natureza do dado

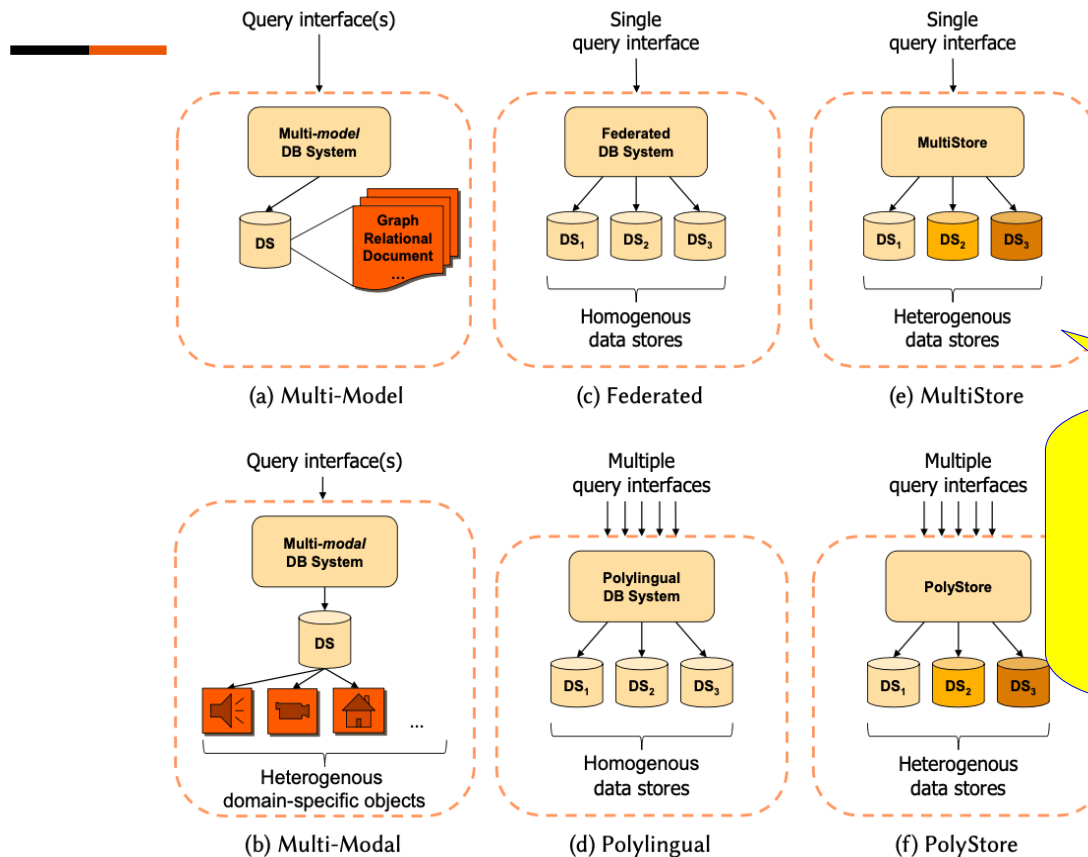
Arquiteturas de Base



BD Federado

- integra fontes de dados
- fontes com o mesmo modelo de BD
- interface de acesso é única e baseada neste modelo de BD

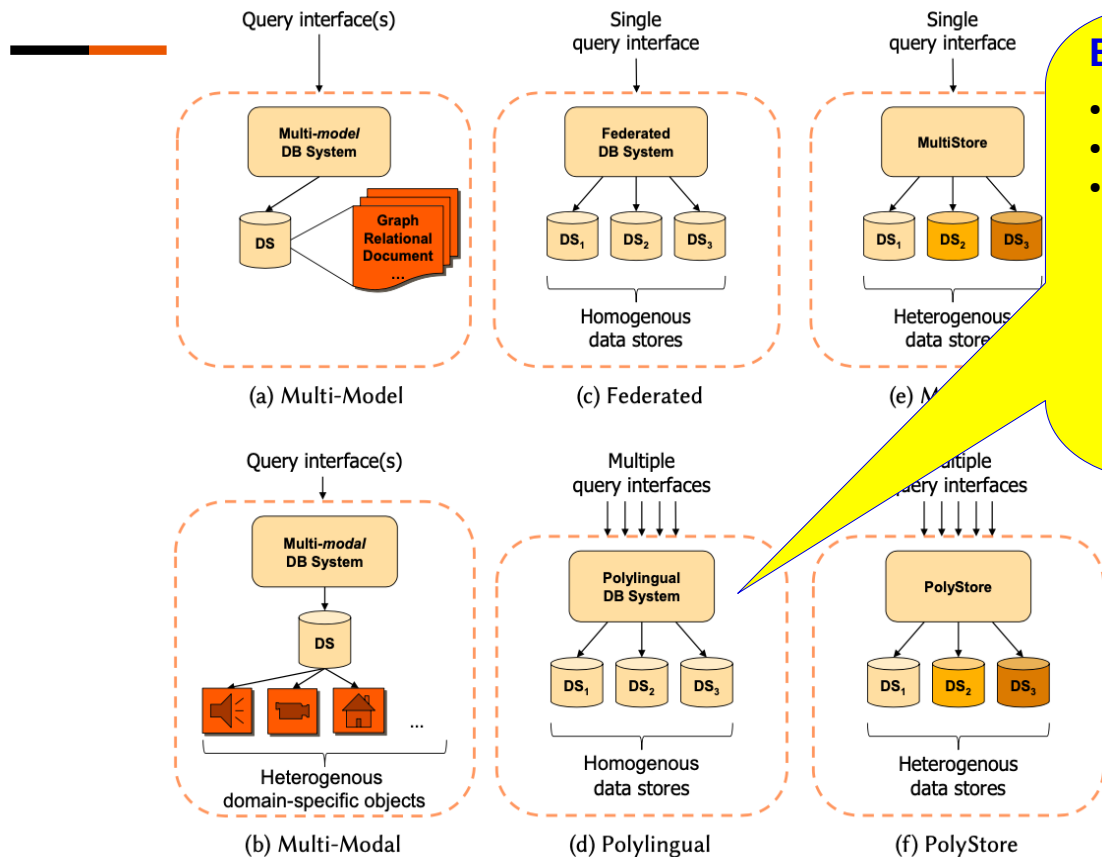
Arquiteturas de Base



BD Multistore

- integra fontes de dados
- fontes com modelos heterogêneos de BDs
- interface de acesso é única e baseada em um modelo canônico

Arquiteturas de Base

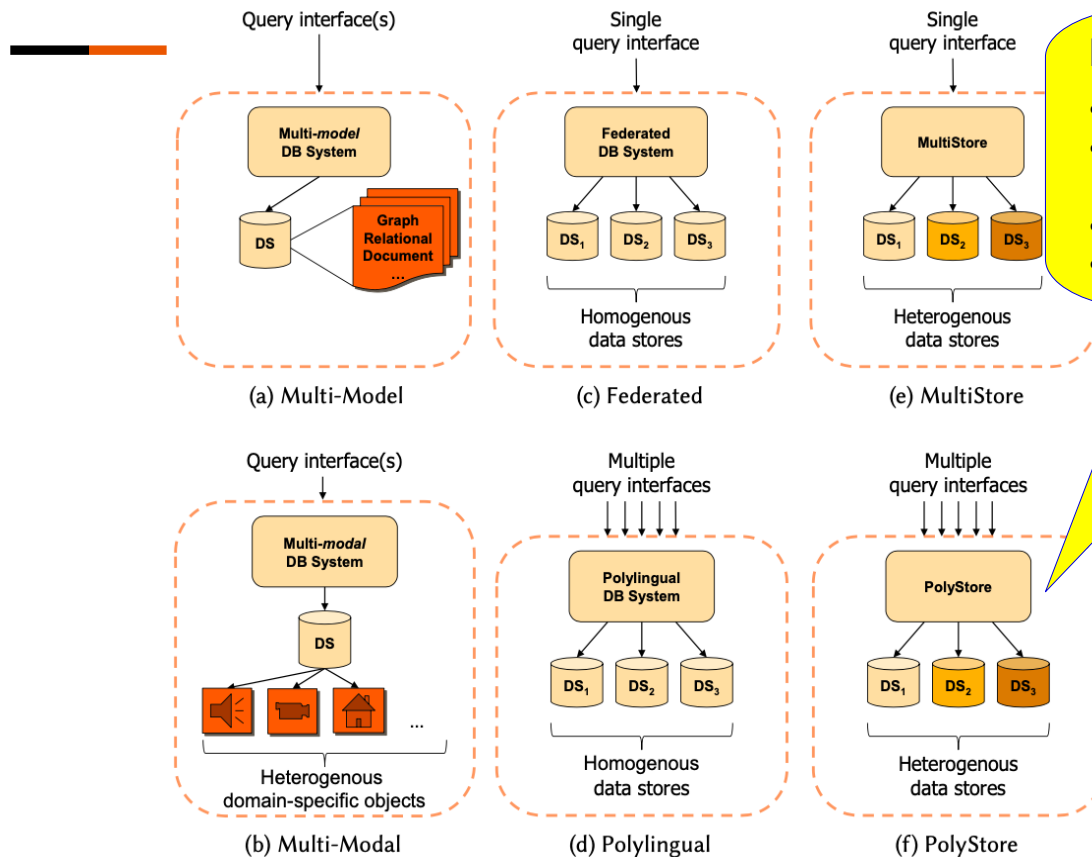


BD Polilíngua

- integra fontes de dados
- fontes com o mesmo modelo de BD
- interfaces de acesso tanto declarativas (p.ex.: SQL-like) quanto procedurais (APIs de acesso) *

* bom para combinar operações analíticas sobre dados recuperados do BD

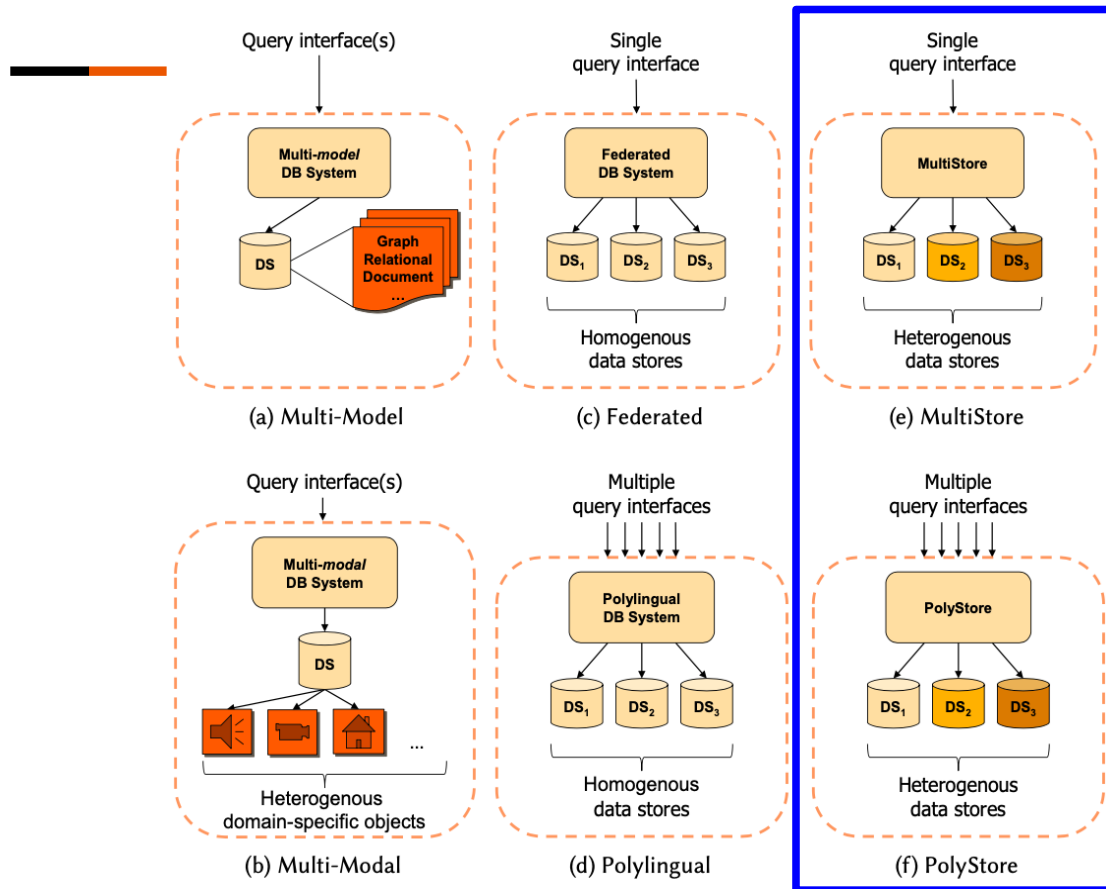
Arquiteturas de Base



BD Polystore

- integra fontes de dados
- fontes com modelos heterogêneos de BDs
- múltiplas interfaces de acesso
- exemplo: *Data Lake*

Arquiteturas de Base



Arquiteturas de Base

- Qual a diferença entre a arquitetura de um **SGPD** e...

Arquiteturas de Base

- Qual a diferença entre a arquitetura um **SGPD** e...
 - a arquitetura de um **Sistema de Integração de Dados (SID)**? Um SID visa acesso integrado a dados *geralmente para fins de read*, enquanto um SGPD pode ser usado *para fins OLTP (read e write) e OLAP (read for analytics)*

Arquiteturas de Base

- Qual a diferença entre a arquitetura de um **SGPD** e...
 - a arquitetura de um **SGBD Distribuído (SGBDD)**? Um SGBDD geralmente possui uma única interface de acesso para os nodos servidores de dados, enquanto um SGPD pode ter múltiplas interfaces de acesso, além do dinamismo para lidar com múltiplos (e considerar novos) modelos de dados e realizar migração/replicação de dados conforme mudam requisitos e *workload*

Arquiteturas de Base

- Qual a diferença entre a arquitetura de um **SGPD** e...
 - a arquitetura de um **Data Lake (DL)**? Um DL provê acesso a fontes de dados heterogêneas através de múltiplas interfaces de acesso, mas **não permite migração/replicação de dados entre/em diferentes fontes** em função da evolução do *workload* e/ou requisitos

Agenda



1. Persistência Poliglota: O que é?!

2. SGPD: Arquiteturas de base

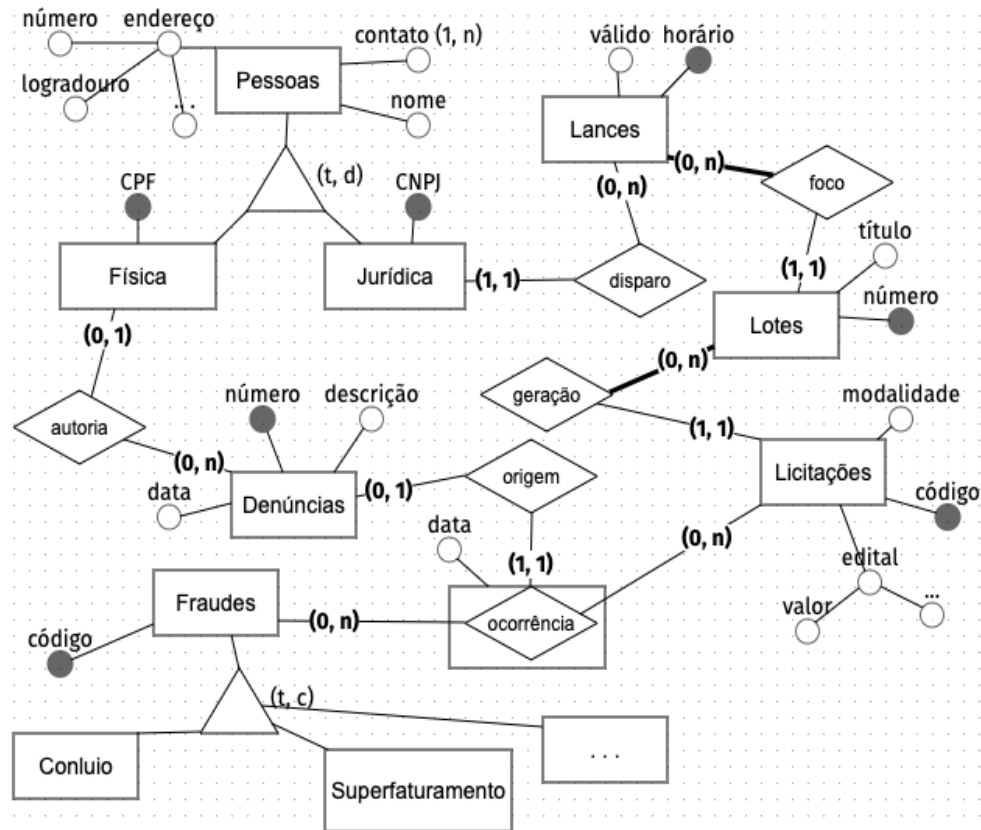
3. SGPD: Desafios & Oportunidades de Pesquisa

Desafios & Oportunidades de Pesquisa

- **Projeto Poliglota de Dados**

Desafios & Oportunidades de Pesquisa

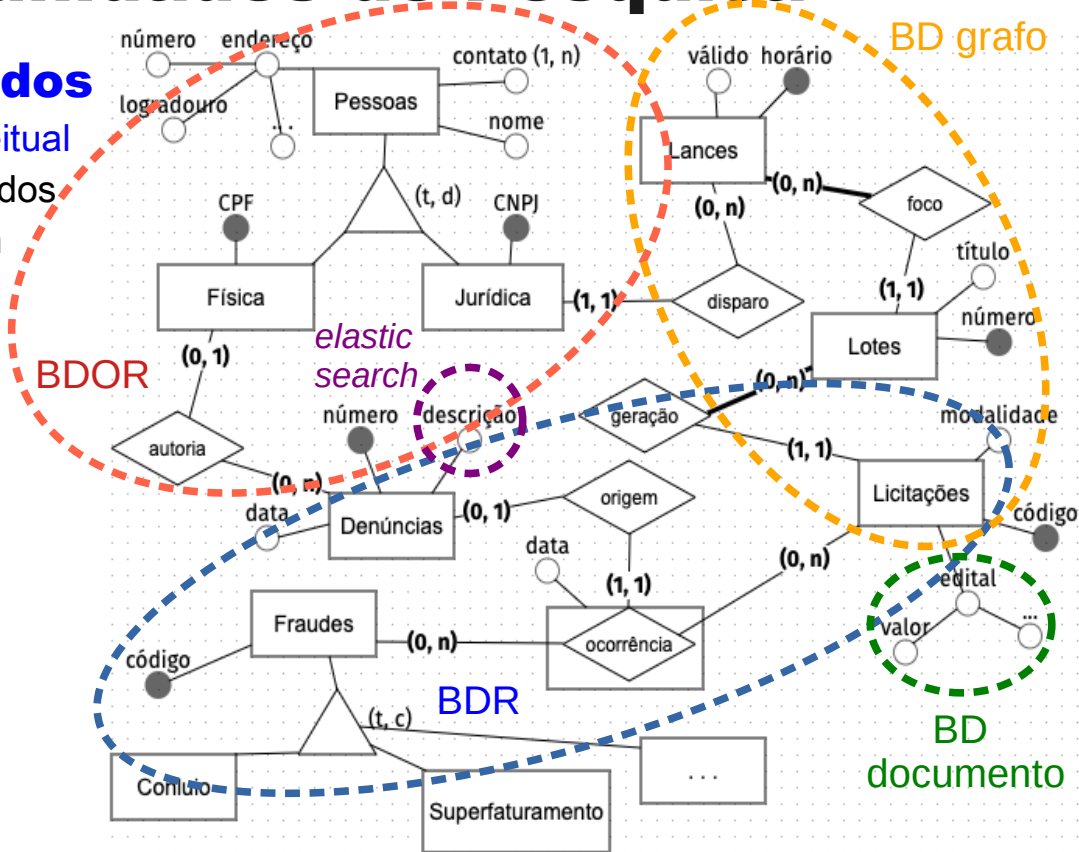
- Projeto Poliglota de Dados**



Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

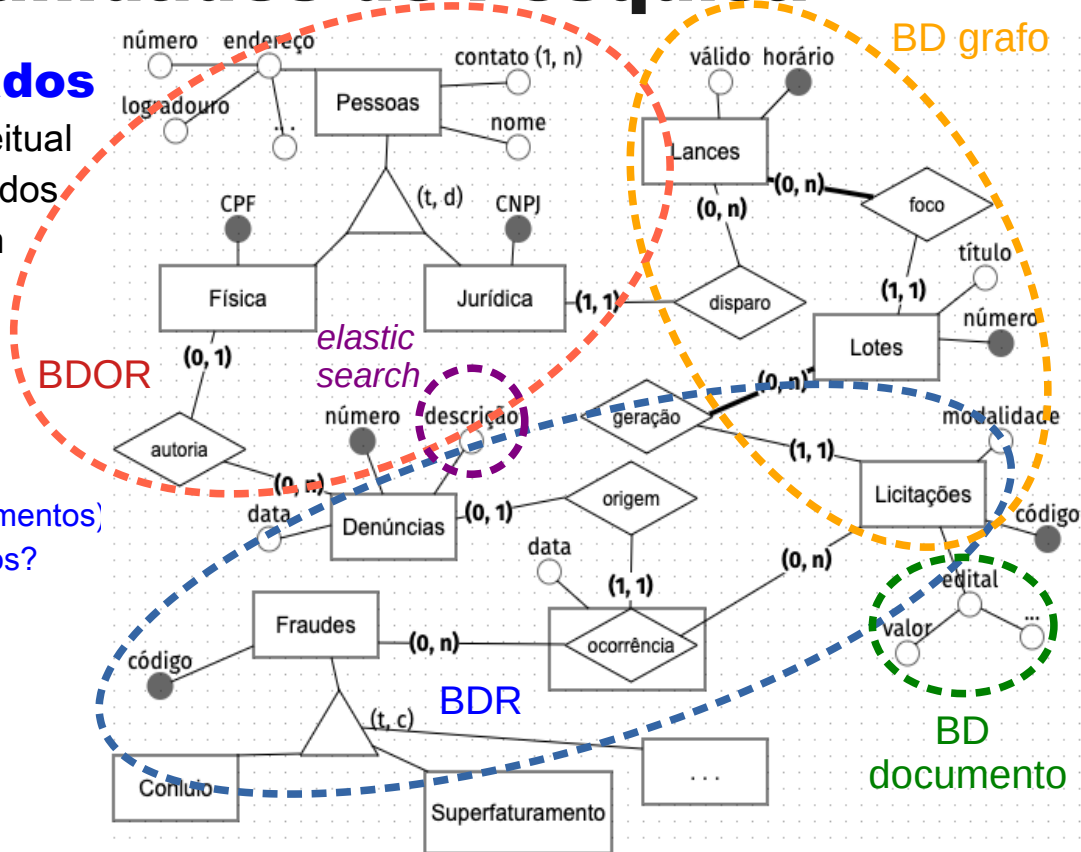
- como “fatiar” uma modelagem conceitual quando o meu objetivo é persistir dados em diversas tecnologias de BD (com diversos modelos de dados)?



Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

- como “fatiar” uma modelagem conceitual quando o meu objetivo é persistir dados em diversas tecnologias de BD (com diversos modelos de dados)?
 - quais níveis de granularidade vamos permitir o fatiamento? (entidade? entidade + relacion.? grupos de atributos? atributos?...)
 - como manter os vínculos (relacionamentos) entre os pedaços? (dados replicados? índices políglotas (intermodelos)?...)



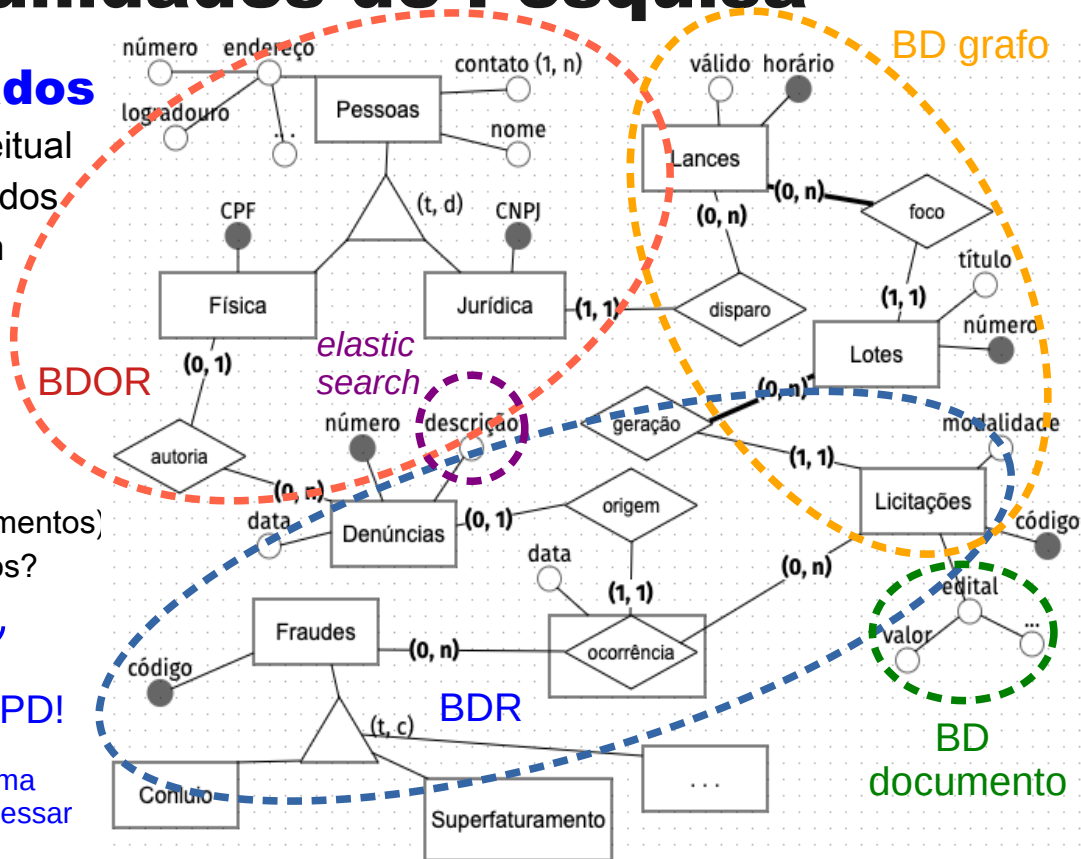
Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

- como “fatiar” uma modelagem conceitual quando o meu objetivo é persistir dados em diversas tecnologias de BD (com diversos modelos de dados)?
 - quais níveis de granularidade vamos permitir o fatiamento? (entidade? entidade + relacion.? grupos de atributos? atributos?...)
 - como manter os vínculos (relacionamentos) entre os pedaços? (dados replicados? índices poliglotas (intermodelos)?...)

valor dos metadados em um SGPD!

(como estão distribuídos os dados? quais são os mecanismos de acesso? como acessar de forma otimizada uma ou mais tecnologias de BD para processar uma certa operação? ...)

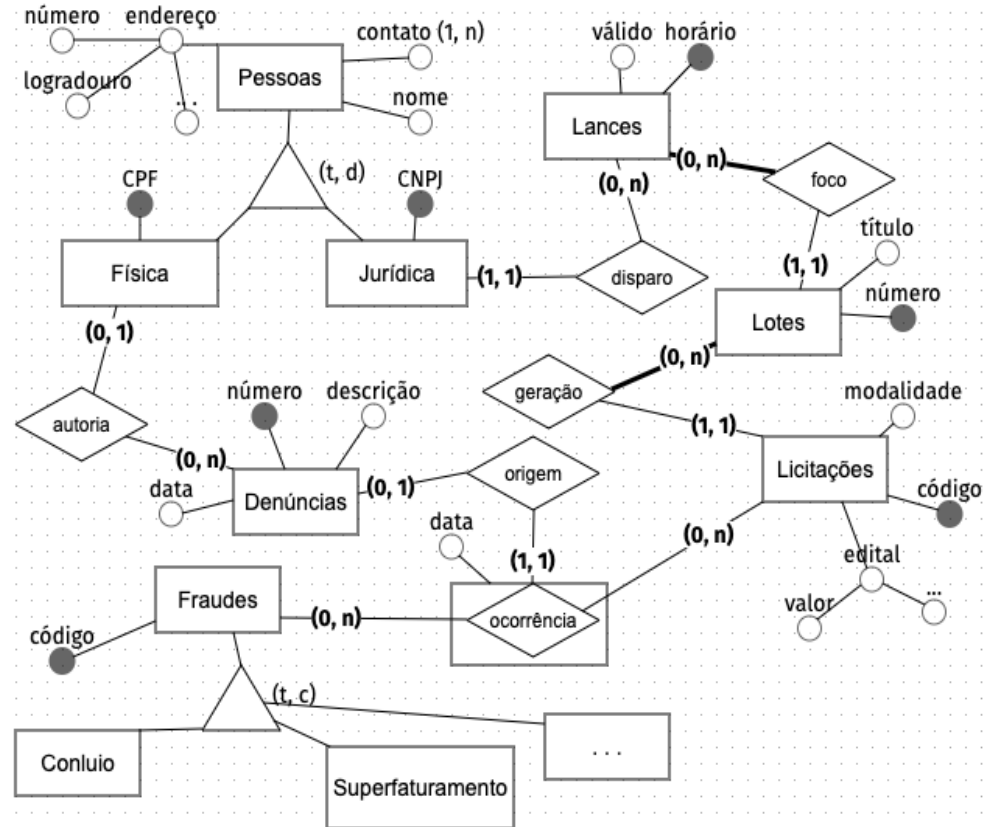


Desafios & Oportunidades de Pesquisa

- **Projeto Poliglota de Dados**
 - esse “fatiamento” deve levar em conta o quê?

Desafios & Oportunidades de Pesquisa

- **Projeto Poliglota de Dados**
 - esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual

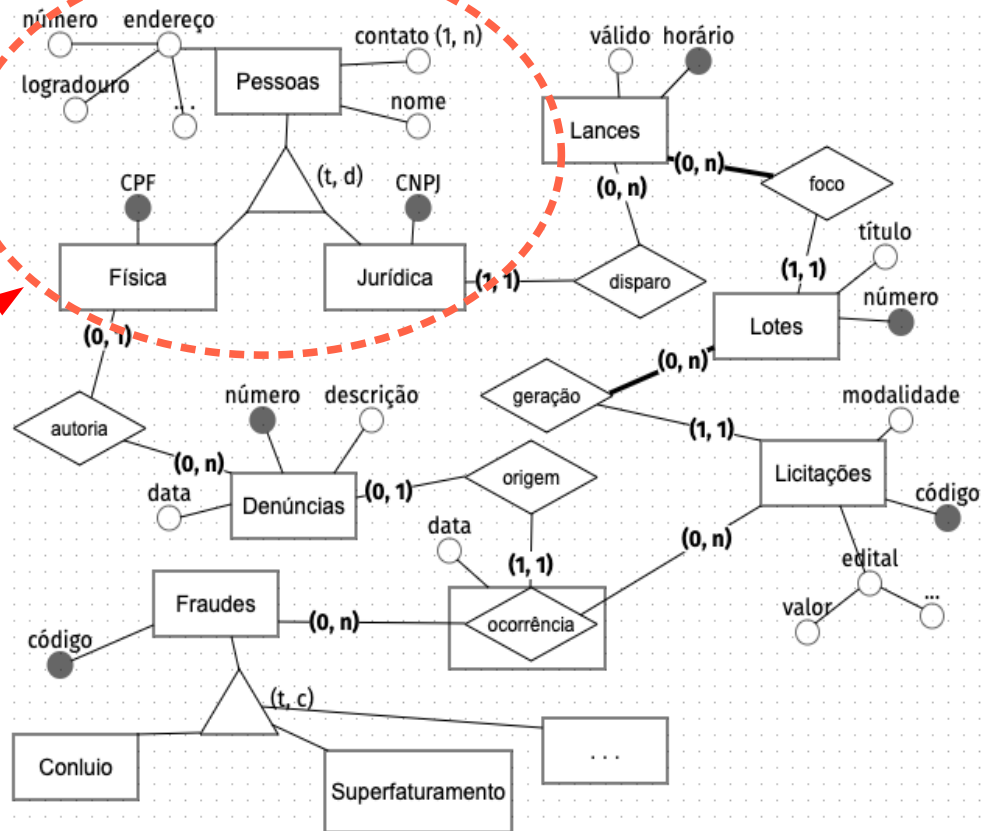


Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual

estruturas complexas



Desafios & Oportunidades de Pesquisa

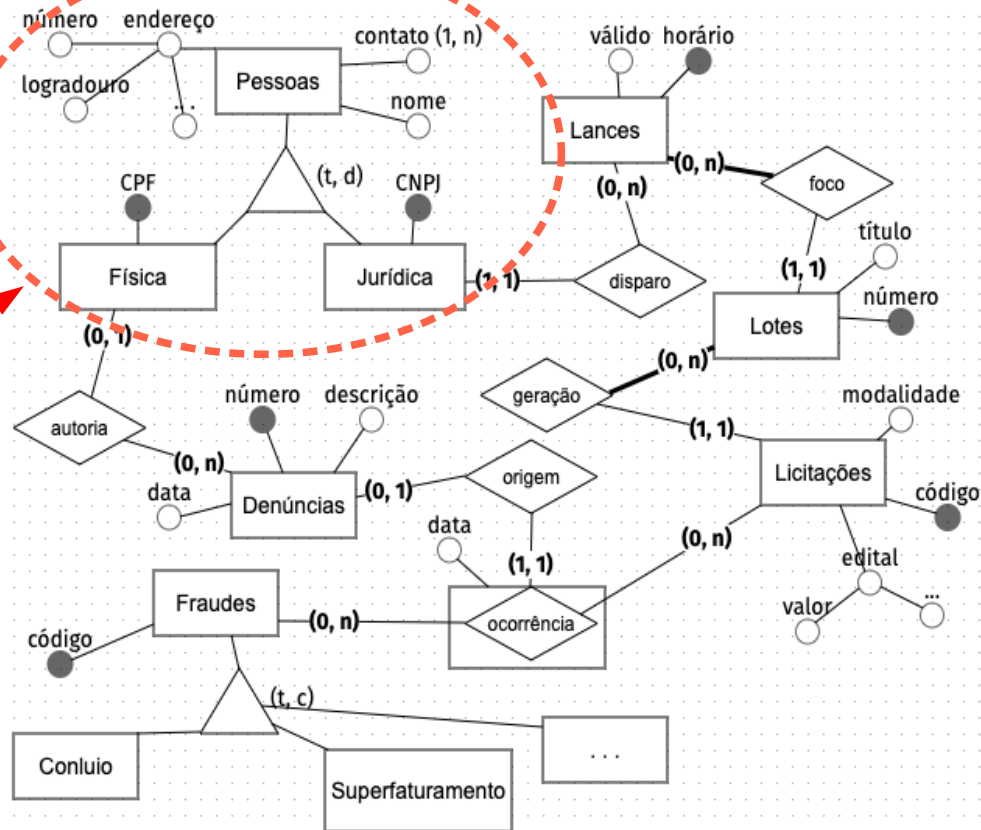
● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual

estruturas complexas



bons candidatos:
BDOR, BD documento,
...



Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

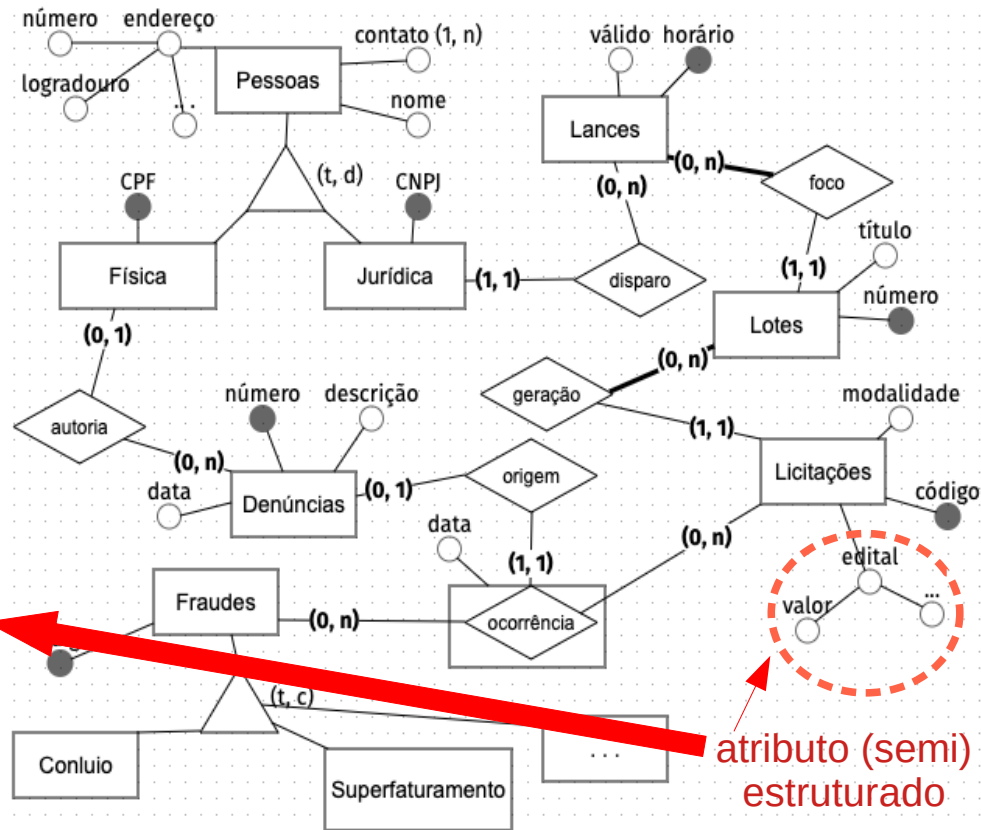
- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual

Desafios & Oportunidades de Pesquisa

● **Projeto Poliglota de Dados**

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual

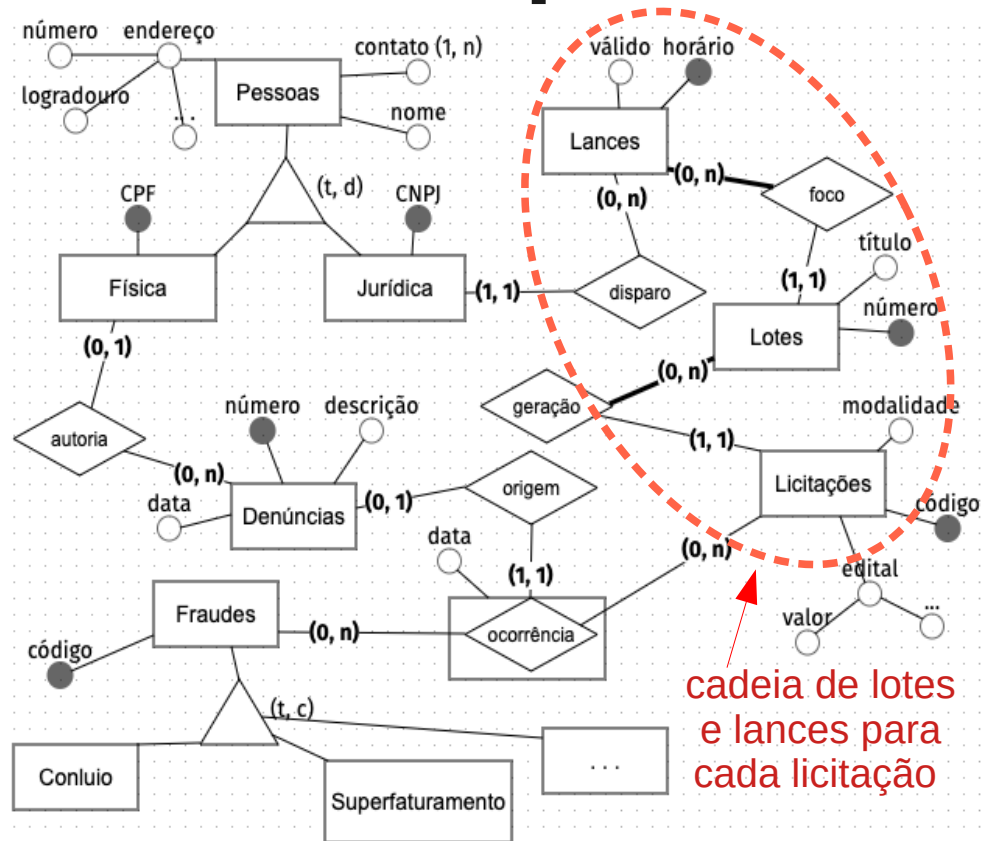
*bons candidatos:
BDR com campo JSON
ou XML, BD XML, BD
documento, ...*



Desafios & Oportunidades de Pesquisa

● **Projeto Poliglota de Dados**

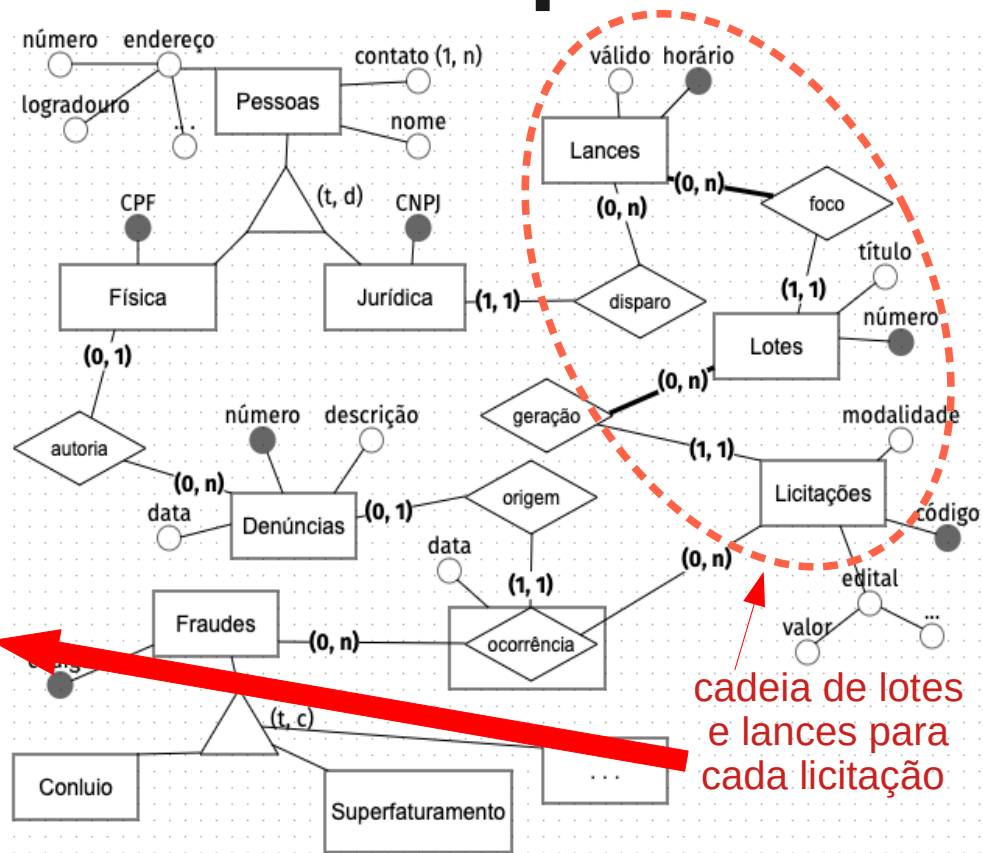
- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual



Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual

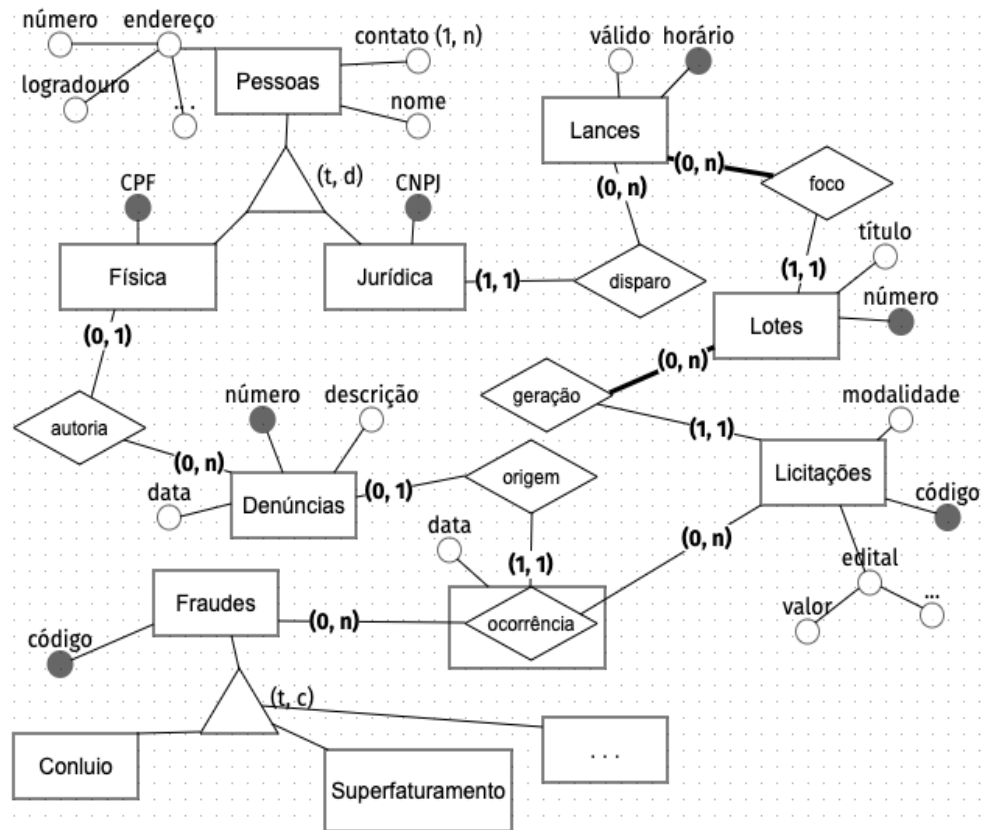


*bons candidatos:
BD grafo, BD para
blockchain, ...*

Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual
 - workload* estimado

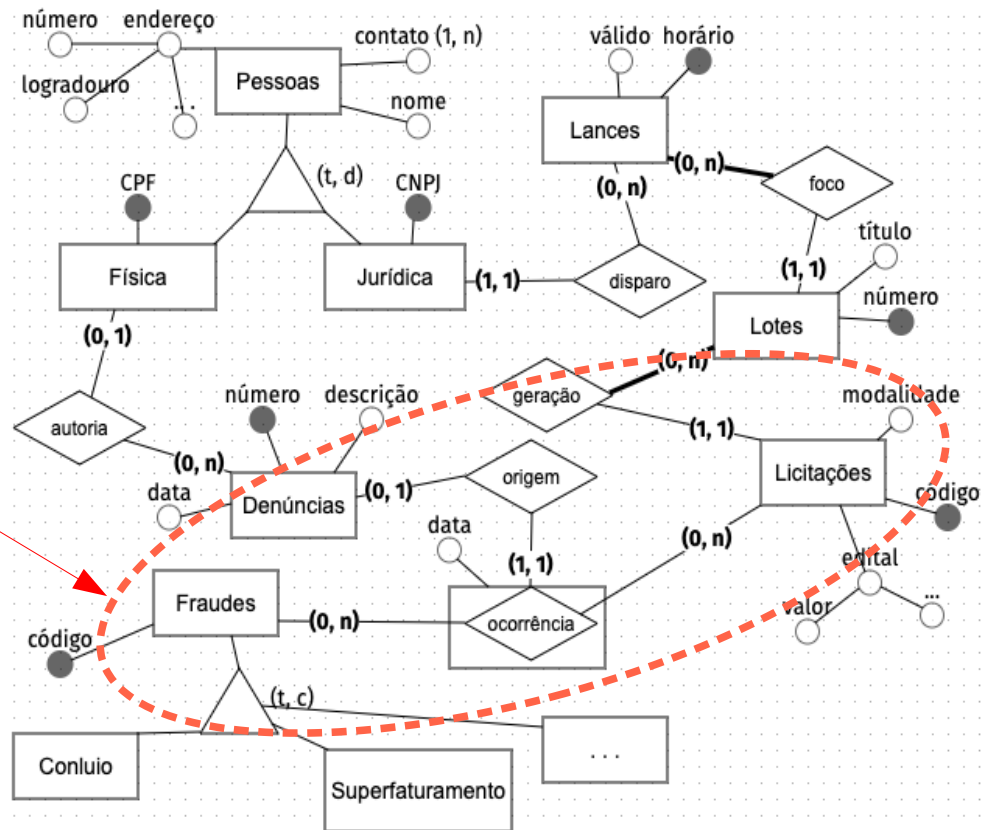


Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual
 - workload* estimado

preciso acessar com frequência
fraudes associadas a licitações para
descobrir correlações/padrões (p.ex.:
certos tipos de fraudes ocorrem muito em
certas modalidades de licitação)



Desafios & Oportunidades de Pesquisa

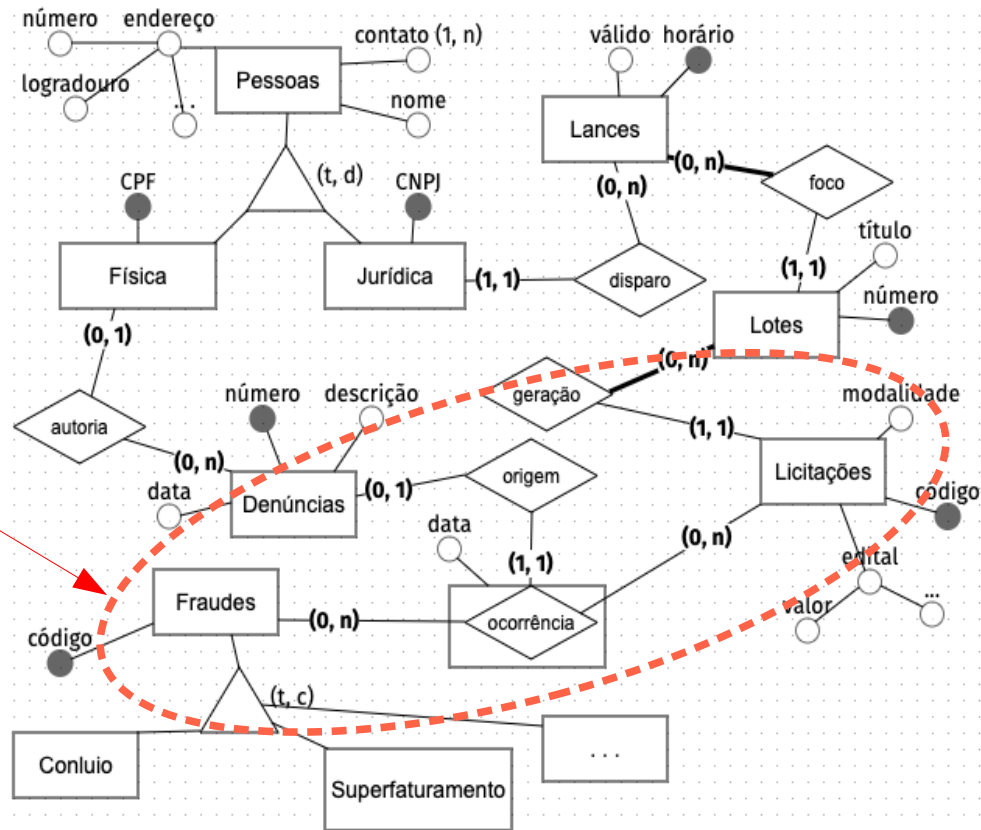
● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual
 - workload* estimado

preciso acessar com frequência
fraudes associadas a licitações para
descobrir correlações/padrões (p.ex.:
certos tipos de fraudes ocorrem muito em
certas modalidades de licitação)



bons candidatos:
BDR (efficient joins), DL (analytics), ...

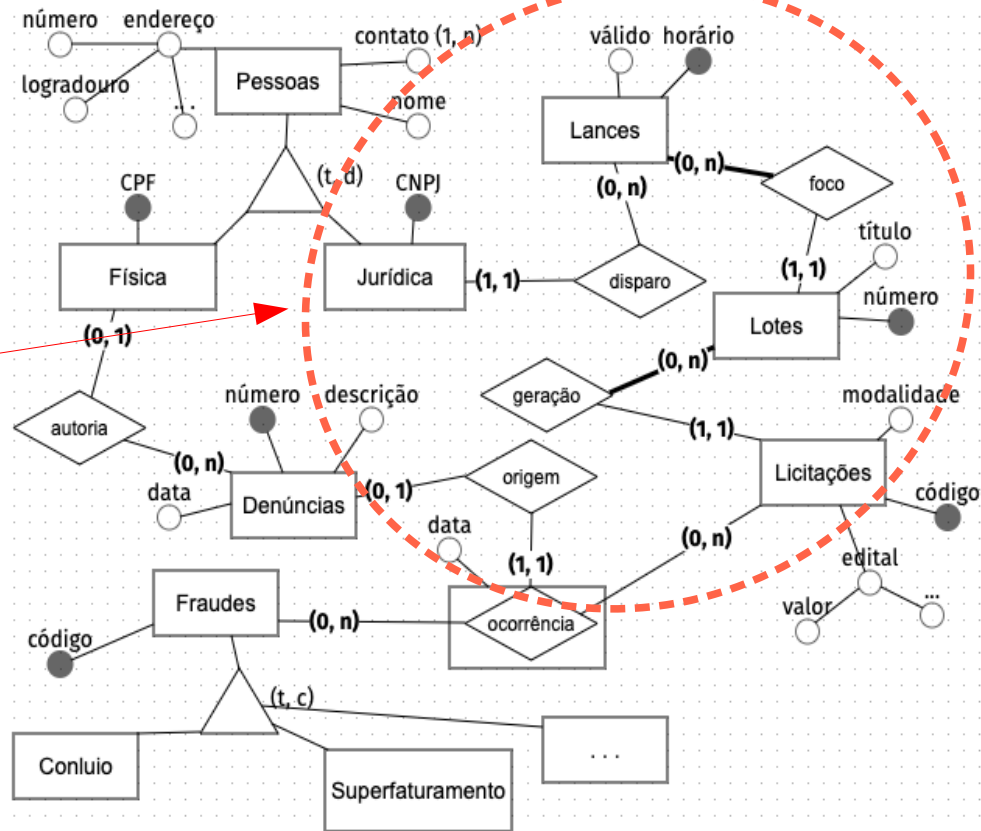


Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual
 - workload* estimado

preciso saber quais empresas têm dado lances em licitações de uma certa modalidade em uma certa região nos últimos anos



Desafios & Oportunidades de Pesquisa

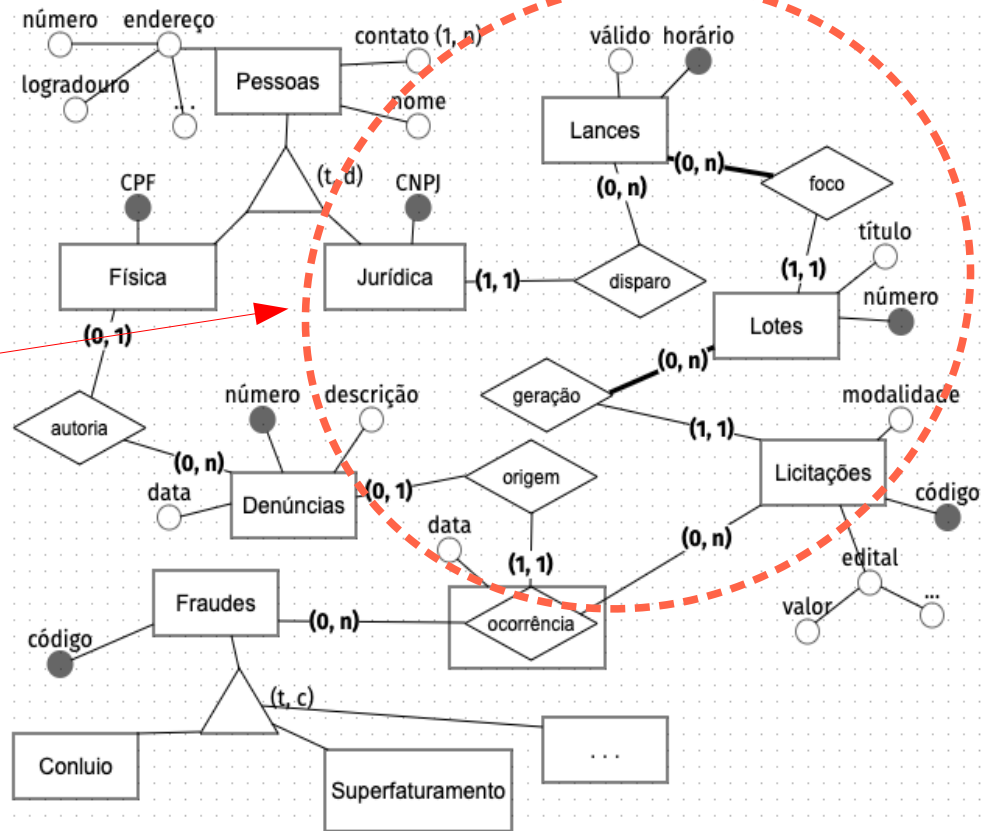
● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual
 - workload* estimado

preciso saber quais empresas têm dado lances em licitações de uma certa modalidade em uma certa região nos últimos anos



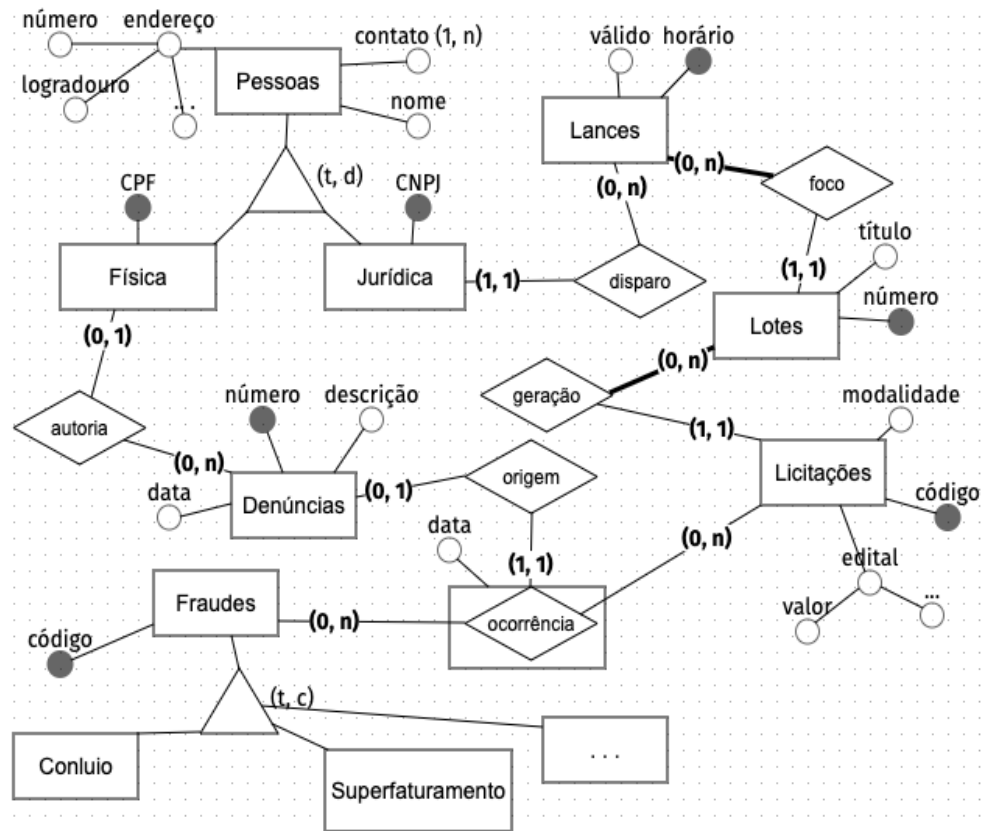
bons candidatos:
BD espaço-temporal, DW, ...



Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual
 - workload* estimado
 - restrições de integridade do domínio (RIs)



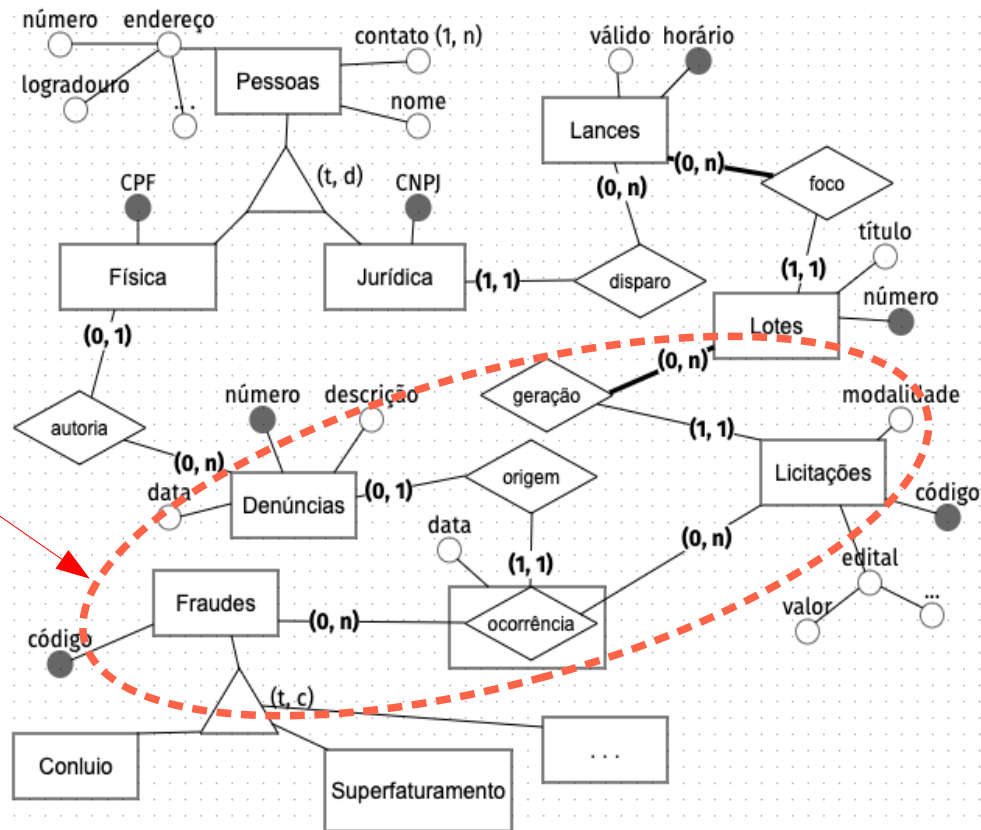
Desafios & Oportunidades de Pesquisa

Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual
 - workload* estimado
 - restrições de integridade do domínio (RIs)

certos tipos de fraudes nunca ocorrem em certas modalidades de licitação

$fraude(tipo = 'X') \Rightarrow \neg(licitação(modalidade = 'Y'))$



Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados

- esse “fatiamento” deve levar em conta o quê?
 - estrutura do esquema conceitual
 - workload* estimado
 - restrições de integridade do domínio (RIs)

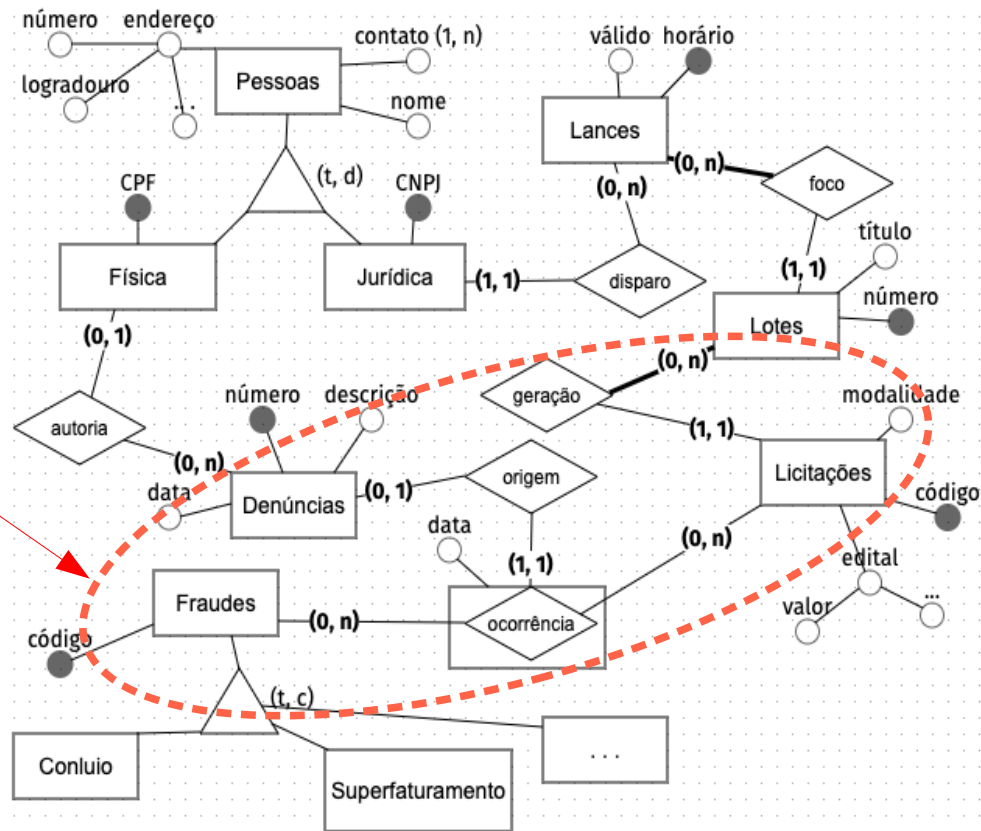
certos tipos de fraudes nunca ocorrem em certas modalidades de licitação

$fraude(tipo = 'X') \Rightarrow \neg(licitação(modalidade = 'Y'))$



bons candidatos:

BDR (consistência forte, DCL), base de regras de associação ou ontologia, ...

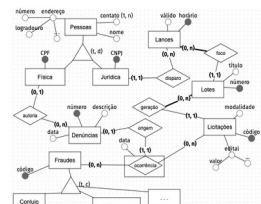


Desafios & Oportunidades de Pesquisa

- **Projeto Poliglota de Dados** ← desafio

Desafios & Oportunidades de Pesquisa

● Projeto Poliglota de Dados



esquema conceitual

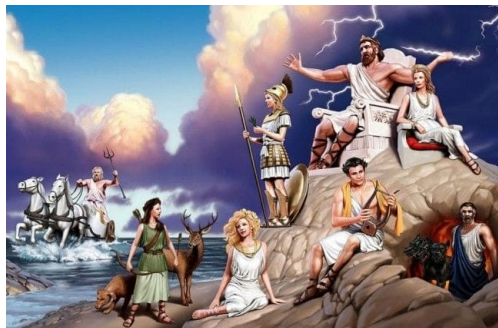


workload estimado

$X \Rightarrow Y$
RIs do domínio

Metodologia para projeto
poliglota de dados

avaliação de
alternativas de
tecnologias de BD



DBAs

desafio

oportunidade de pesquisa



esquema de distribuição
poliglota de dados



esquema de metadados +
relacionamentos inter-tecnologia BD

Desafios & Oportunidades de Pesquisa

- **Interface(s) de acesso para um SGPD**

Desafios & Oportunidades de Pesquisa

- **Interface(s) de acesso para um SGPD**
 - como definir uma linguagem ou conjunto de métodos de acesso poliglota?



Desafios & Oportunidades de Pesquisa

● **Interface(s) de acesso para um SGPD**

- como definir uma linguagem ou conjunto de métodos de acesso poliglota?
 - uma opção: **linguagem universal**
 - ♦ adotada por alguns SGBDs multimodelo



Desafios & Oportunidades de Pesquisa

● Interface(s) de acesso para um SGPD

- como definir uma linguagem ou conjunto de métodos de acesso poliglota?
 - uma opção: **linguagem universal**
 - adotada por alguns SGBDs multimodelo
 - exemplo: **AQL** (*ArangoDB query language*)
 - acesso a um modelo integrado de documentos e grafos, bem como textos longos (*search engine*)

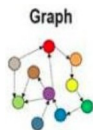
```
INSERT {
  _key: "GilbertoGil",
  "profession": "singer",
  "age": 80,
  "city": {"name": "Fortaleza", "state": "CE"}
} INTO users
```



consulta com filtro sobre
estrutura complexa de
documento

```
FOR u IN users
FOR c IN u.city
FILTER c.state = "CE"
RETURN u.profession
```

```
FOR x IN 1..2
OUTBOUND 'users/GilbertoGil'
RETURN x.profession;
```



consulta envolvendo
navegação em grafo



Desafios & Oportunidades de Pesquisa

● Interface(s) de acesso para um SGPD

- como definir uma linguagem ou conjunto de métodos de acesso poliglota?
 - outra opção: **múltiplas linguagens e tradutores**
 - exemplo: protótipo *Polypheny DB*



```
INSERT INTO Users
(name,profession,age,city,state)
VALUES
("Gilberto Gil","singer", 80,"Fortaleza","CE")
```

```
CREATE (p:Users{name:"Gilberto
Gil", profession:"singer", age:80,
city:"Fortaleza", state:"CE"});
```

x	y	z

consulta de
entrada

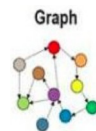
```
SELECT profession
FROM Users
WHERE state = "CE"
```

modelo relacional (SQL)



```
MATCH (u:Users)
WHERE u.state = "CE"
RETURN u.profession
```

modelo de grafo (Cypher)



Desafios & Oportunidades de Pesquisa

● Interface(s) de acesso para um SGPD

Abordagem	Prós	Contras
Linguagem universal	<ul style="list-style-type: none">interface única de acesso	<ul style="list-style-type: none">expressividade/complexidade para atender diversos modelos de dadosmapeamento para vários modelos de dados no caso de uma linguagem abstrata em nível conceitual
Múltiplas linguagens	<ul style="list-style-type: none">maior adaptabilidade à inclusão de novos modelos de dados	<ul style="list-style-type: none">curva de aprendizado de várias linguagensmapeamentos entre linguagens diferentesdificuldade de mapeamento entre linguagens com poder de expressão diferentes



Desafios & Oportunidades de Pesquisa


- **Interface(s) de acesso para um SGPD** ← desafio

Desafios & Oportunidades de Pesquisa

- **Interface(s) de acesso para um SGPD** ← desafio



oportunidade de pesquisa



Projeto de uma linguagem e/ou métodos
de acesso poliglota de BD

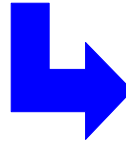
Desafios & Oportunidades de Pesquisa

- **Interface(s) de acesso para um SGPD** ← desafio

oportunidade de pesquisa

Projeto de uma linguagem e/ou métodos
de acesso poliglota de BD

alguns pontos
a avaliar...



Desafios & Oportunidades de Pesquisa

- **Interface(s) de acesso para um SGPD** ← desafio

oportunidade de pesquisa

Projeto de uma linguagem e/ou métodos
de acesso poliglota de BD

alguns pontos
a avaliar...

- potenciais modelos de dados a suportar

Desafios & Oportunidades de Pesquisa

- **Interface(s) de acesso para um SGPD** ← desafio

oportunidade de pesquisa

Projeto de uma linguagem e/ou métodos
de acesso poliglota de BD

alguns pontos
a avaliar...

- potenciais modelos de dados a suportar
- estimativa de *workload* mais estável

Desafios & Oportunidades de Pesquisa

- **Interface(s) de acesso para um SGPD** ← desafio

oportunidade de pesquisa

Projeto de uma linguagem e/ou métodos
de acesso poliglota de BD

alguns pontos
a avaliar...

- potenciais modelos de dados a suportar
- estimativa de *workload* mais estável
- linguagem(ns) de base a estender (padrões mais consolidados (p.ex.: SQL) facilitam *parsing*, otimizações, curva de aprendizado, ...)

Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload*** (Evolução do projeto poliglota de dados)



Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***
 - como gerenciar isso?



Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***
 - como gerenciar isso?
 - manter mapeamentos entre os modelos de dados suportados pelo SGPD?



Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***
 - como gerenciar isso?
 - manter mapeamentos entre os modelos de dados suportados pelo SGPD?
(desafio: definir regras de mapeamento entre modelos de dados – com completude de conceitos e eficiência de representação)



Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***
 - como gerenciar isso?
 - manter mapeamentos entre os modelos de dados suportados pelo SGPD? (desafio: definir regras de mapeamento entre modelos de dados – com completude de conceitos e eficiência de representação)
 - o que fazer no caso de decisão por uma nova tecnologia de BD (novo modelo de dados) para atender melhor esse novo *workload*? realizar engenharia reversa da porção a ser migrada para um modelo conceitual MC, e a partir do MC realizar o projeto lógico/físico na nova tecnologia?



Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***
 - como gerenciar isso?
 - manter mapeamentos entre os modelos de dados suportados pelo SGPD? (desafio: definir regras de mapeamento entre modelos de dados – com completude de conceitos e eficiência de representação)
 - o que fazer no caso de decisão por uma nova tecnologia de BD (novo modelo de dados) para atender melhor esse novo *workload*? realizar engenharia reversa da porção a ser migrada para um modelo conceitual MC, e a partir do MC realizar o projeto lógico/físico na nova tecnologia?
 - outro desafio: como detectar necessidades de migração de dados? Monitoramento de desempenho de transações OLTP e OLAP?



Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***

← desafio

Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***



Mecanismo de migração de dados
em um SGPD

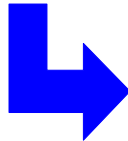
Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***



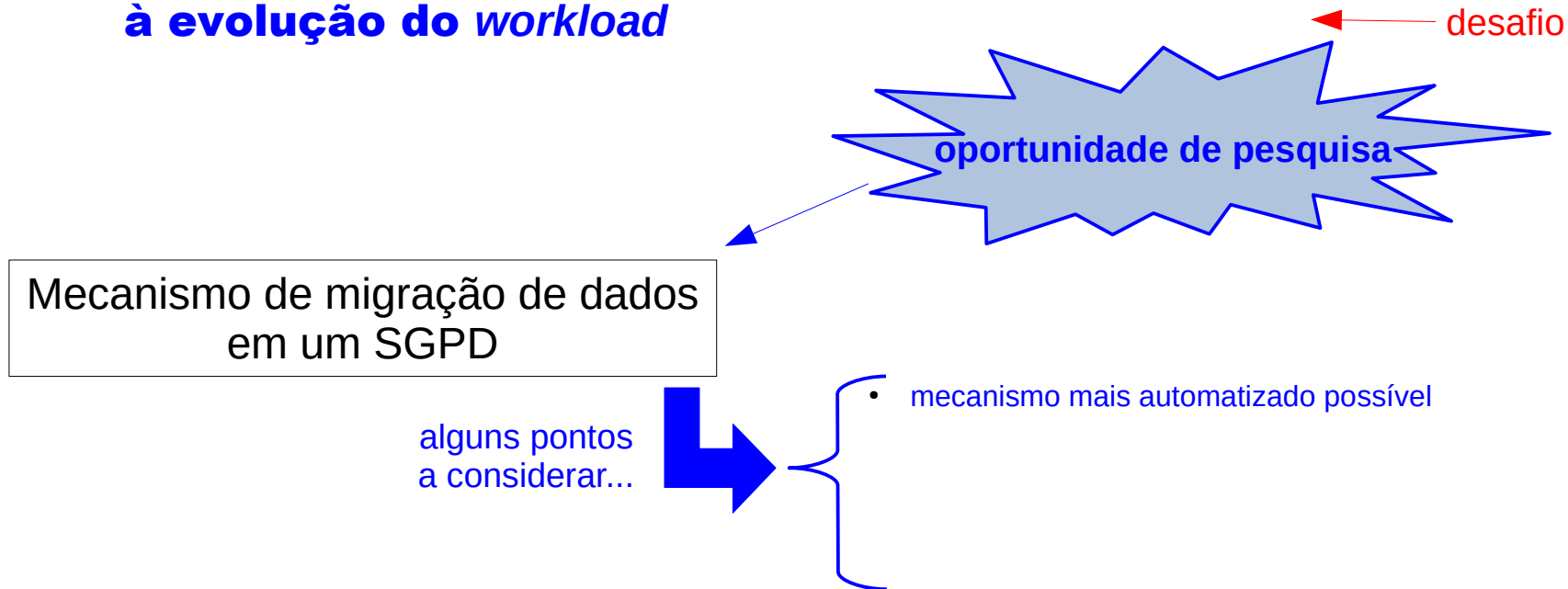
Mecanismo de migração de dados
em um SGPD

alguns pontos
a considerar...



Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***



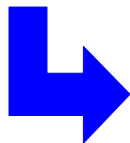
Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***



Mecanismo de migração de dados
em um SGPD

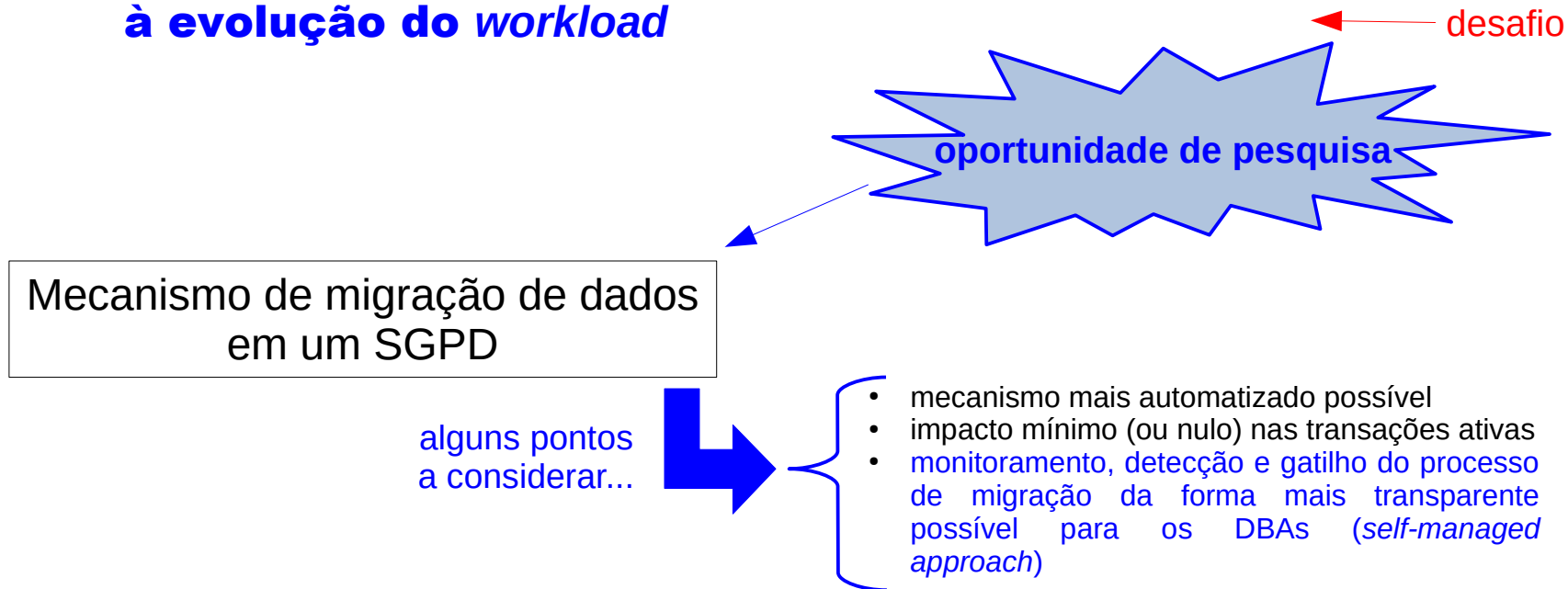
alguns pontos
a considerar...



- mecanismo mais automatizado possível
- impacto mínimo (ou nulo) nas transações ativas

Desafios & Oportunidades de Pesquisa

- **Migração de dados entre tecnologias de BD devido à evolução do *workload***



Referências

- Kiehn, F. et al. [Polyglot Data Management: State of the Art & Open Challenges](#). PVLDB, 15(12): 3750 – 3753 (2022)
- Roy-Hubara, N.; Shoval, P.; Sturm, A. [Selecting Databases for Polyglot Persistence Applications](#). Data Knowledge Engineering (DKE). 137: 101950 (2022)
- Glake, D. et al. [Towards Polyglot Data Stores - Overview and Open Research Questions](#). CoRR abs/2204.05779 (2022)
- Tan, R. et al. [Enabling Query Processing Across Heterogeneous Data Models: A Survey](#). IEEE BigData, p.3211-3220 (2017)



UNIVERSIDADE FEDERAL
DE SANTA CATARINA



Persistência Poliglota:

Fundamentos, Desafios & Oportunidades



CENTRO TECNOLÓGICO
Universidade Federal de Santa Catarina

Ronaldo dos Santos Mello
r.mello@ufsc.br

GBD/PPGCC/INE/CTC/UFSC



Programa de
Pós-Graduação
em Ciência da
Computação

U F S C