

Disciplina:

BANCOS DE DADOS NoSQL

Professor: Augusto Zadra





1.3 Tipos de Bancos NOSQL e principais bancos do mercado

LINGUAGENS DE MARCAÇÃO

- Antes de iniciarmos com os tipos de bancos de dados NoSQL faz-se necessário fazermos uma introdução a algumas questões que são de fundamentos.
- Iniciaremos falando sobre a estruturação de arquivos, e a maior parte de nós já ouviu dizer algo sobre XML ou JSON.
- Mas não é hora de perguntar de é de comer ou de passar no cabelo.
- É hora de conhecermos do que se trata.

LINGUAGENS DE MARCAÇÃO

- **LINGUAGENS DE MARCAÇÃO** são sistemas usados para definir **padrões e formatos** de exibição dentro de um documento.
- Funcionam para definir como um determinado conteúdo vai ser visualizado na tela ou como os dados serão distribuídos.
- Essa codificação interna é feita pelo uso de marcadores ou *tags*.
-  **TAGS:** são palavras-chave para relacionar informações semelhantes.
-  **METADADOS:** são dados usados para classificar e organizar arquivos, páginas e outros conteúdos.

LINGUAGENS DE MARCAÇÃO

- Arquivos estruturados de dados são arquivos que seguem padrões para entendimento dos dados que estão ali armazenados.
- Qualquer que seja o arquivo estruturado, existem diversas maneiras de representá-la por formatos como CSV (*Comma/Character Separeted Values*) – **Esse quase todo mundo já utilizou pelo menos uma vez, não é?** – ou por formatos ODS, XLSX entre outros.
- Todos os citados acima, respeitam diferentes graus a estrutura linha coluna que vem dos bancos de dados relacionais.

LINGUAGENS DE MARCAÇÃO – EXEMPLO CSV

Descrição da coluna

Para entendimento do
significado

Ano,Cidade,País,Ouro,Prata,Bronze

1896,Atenas,Austria,2,1,2

1896,Atenas,França,5,4,2

1896,Atenas,Hungria,2,1,3

1896,Atenas,Grã Bretanha,2,3,2


Dados da linha
separados por vírgula.

Apontamento da coluna alinhado pelo separador,
nesta caso a “,”.

LINGUAGENS DE MARCAÇÃO

- Porém há casos de arquivos estruturados que não seguem esta lógica da linha vs. coluna que é o caso dos tipos JSON e XML.
- ***JavaScript Object Notation – JSON (RFC 4627)***: formato de padrão aberto que utiliza texto legível a humanos para transmitir objetos de dados consistindo de pares atributo-valor.
- ***eXtensible Markup Language – XML (W3C)***: formato flexível e simples, que possibilita a criação de *tags* personalizadas concentrando na estrutura da informação e não em sua aparência.

LINGUAGENS DE MARCAÇÃO

- **JSON** e **XML** são os dois formatos mais utilizados pelas linguagens de programação principalmente no que se refere a comunicação assíncrona entre aplicações como é o caso das requisições **assíncronas**.
-  ***Asynchronous JavaScript and XML - AJAX***: conjunto de técnicas de desenvolvimento voltado para a web que permite que aplicações trabalhem de modo assíncrono, processando qualquer requisição ao servidor em segundo plano.

LINGUAGENS DE MARCAÇÃO – XML vs JSON

Característica	XML	JSON
Natureza auto-descritiva: fácil entendimento do significado	X	X
Representar informação complexa: objetos compostos, relações de hierarquia, atributos multivalorados	X	X
Transportar informações em aplicações AJAX	X	X
Independentes de linguagem	X	X
Codificação em vários formatos	X	-
Comentários no arquivo	X	-
Leveza do arquivo e maior agilidade no tráfego em rede	-	X
Executar requisições AJAX em domínio cruzado: hospedadas fora do servidor	-	X

HASHING

- Conceituação importante para entendimento sobre os bancos de dados NoSQL é a utilização de **HASHING**, também conhecido como **tabela de dispersão** ou **método de cálculo de endereço**.
- **HASHING**: constituem uma abordagem comum para o problema de armazenamento de dados e sua forma de representação mais simples é um vetor de registros para organizá-los.

Tipos de bancos de dados NoSQL

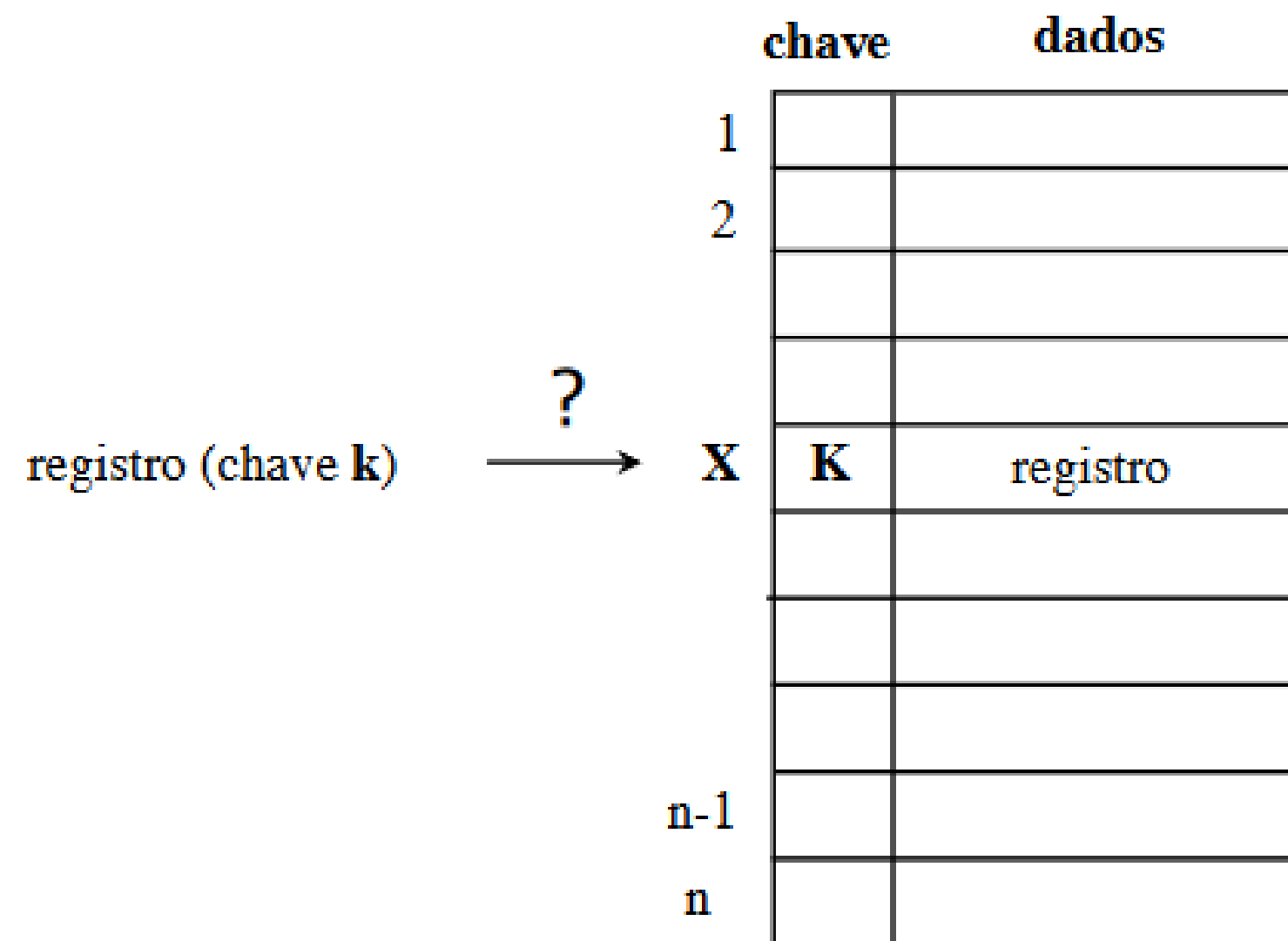
HASHING

- **Hash** é uma generalização de um array simples comum, sendo uma estrutura de dados do tipo dicionário.
- Dicionários são estruturas especializadas em prover as operações de **inserir, pesquisar e remover**.
- A técnica utiliza uma **função**, aplicada sobre parte da informação (chave), para retornar o índice onde a informação deve ou deveria estar armazenada.
- A este cálculo dá-se o nome de **Função de Hashing**.

Tipos de bancos de dados NoSQL

HASHING

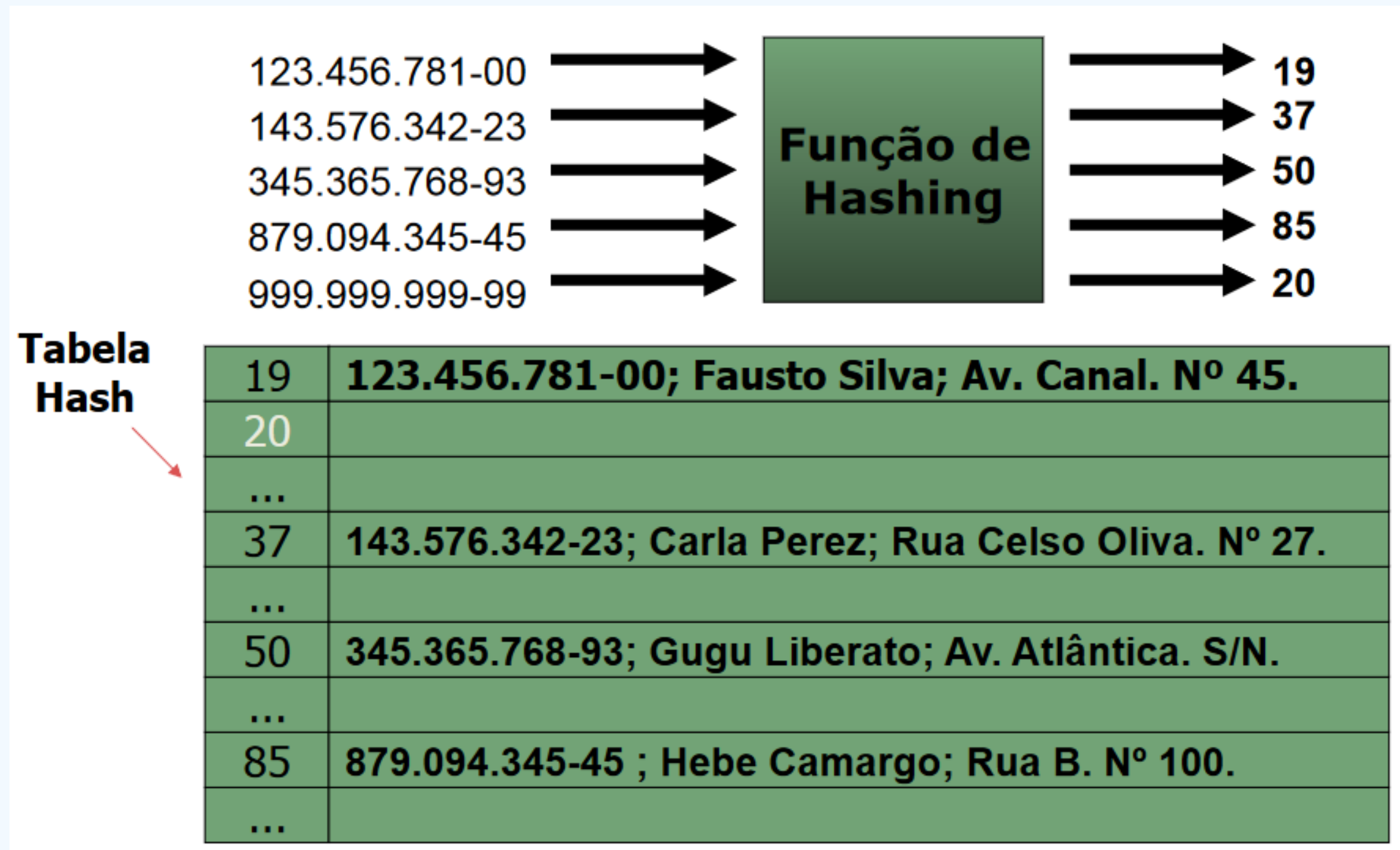
- Através do valor que será armazenado utiliza-se a função para calcular qual a posição que o registro será inserido.



Fonte: BOERES, 2021.

Tipos de bancos de dados NoSQL

HASHING



Fonte: BOERES, 2021.

TABELAS HASH

- Em resumo, ao invés de organizarmos a tabela segundo o valor relativo de cada chave em relação às demais, leva-se em o seu valor absoluto, interpretado como um valor numérico.
- Através de uma função conveniente, a chave é transformada em um endereço de uma tabela tornando a busca mais rápida.
- Em caso de volumes de dados muito altos a probabilidade do resultado da *função hash* se repetir é **alto**, desta forma há técnicas de tratamento de colisões que devem ser levadas em consideração para a manutenção da eficiência.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL

- Os aplicativos NoSQL usam uma variedade de tipos de armazenamento de dados que variam do armazenamento de valor-chave simples que associa uma chave única a um valor, aos armazenamentos de gráficos usados para associar relacionamentos.
- Oferta-se também armazenamento de documentos usados para dados variáveis e, cada tipo NoSQL de armazenamento de dados tem atributos exclusivos.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL

- Existem 5 modelos de armazenamento para os bancos de dados NoSQL:
 - ✓ Orientado a Coluna (tabular)
 - ✓ Armazém de Documentos
 - ✓ Chave/Valor
 - ✓ Grafo
 - ✓ Motor de Busca
- A escolha do banco de dados NoSQL depende do tipo de dados que você precisa armazenar, seu tamanho e complexidade.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Chave/Valor

- São os mais simples de todos os bancos de dados NoSQL e armazenam apenas pares de chaves com valores associados, e permitem obter os valores quando uma consulta é realizada a uma chave.
- A chave e os valores podem ser qualquer coisa, desde objetos simples até objetos compostos complexos.
- A estrutura de dados básica é um dicionário ou mapa e pode armazenar um valor, como um inteiro, uma string, uma estrutura JSON ou uma matriz, junto com uma chave usada para fazer referência a esse valor.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Chave/Valor

Vantagens

- São altamente particionáveis e permitem escalabilidade horizontal que outros tipos de bancos de dados não conseguem alcançar.
- Utiliza uma tabela hash ou uma tabela com varias chaves no qual cada chave referencia um valor.
- Facilidade de inserção e recuperação, pelo uso da chave de indexação para mapear um valor.
- O acesso pela chave-primária, geralmente, fornece um ótimo desempenho e escalabilidade.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Chave/Valor

Desvantagens

- Não são empregados em aplicações mais complexas devido a sua simplicidade.
- A busca imprecisa por campos, sem usar o índice, perde a performance conforme a quantidade de dados aumenta.
- O modelo de banco de dados não foi projetado para fornecer consistência quando várias transações são executadas.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Chave/Valor

Modelo



Fonte: BOERES, 2021.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Chave/Valor

Utilização

- Os dados da sessão podem incluir informações de perfil do usuário, mensagens, dados e temas personalizados, recomendações, promoções direcionadas e descontos.
- Cada sessão de usuário tem um identificador exclusivo. Os dados de sessão nunca são consultados por nada além de uma chave primária, então um armazenamento de chave-valor rápido é mais adequado para dados de sessão.
- Proporcionam menor sobrecarga por página do que bancos de dados relacionais.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Armazém de Documentos

- A principal característica é a organização de dados livre de esquemas e contém uma coleção de pares de valores-chave armazenados em documentos.
- Os registros não precisam ter uma estrutura uniforme e registros distintos podem ter coleções de atributos diferentes.
- Tipos de dados nas colunas individuais podem ser diferentes para cada registro sendo é possível armazenar mais de um valor em uma mesma coluna.
- Foram inspirados pelo BD Lotus Notes para armazenar documentos.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Armazém de Documentos

- Na prática os atributos podem ser multivalorados e podem conter dados atômicos, documentos aninhados ou ainda listas.
- Bancos de dados orientados a documento utilizam um formato de armazenamento que suporta que uma estrutura seja aninhada em outra (ou seja, um documento embutido em outro), com um **XML** ou **JSON**.
- Os bancos de dados de documentos são os mais flexíveis para armazenar uma grande coleção de documentos distintos e não relacionados.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Armazém de Documentos

- Oferecem suporte a índices secundários e vários tipos de documentos ou listas aninhados.
- Em vez de colunas com nomes e tipos de dados, um documento contém uma descrição do tipo de dados e o valor dessa descrição.
- Para adicionar outros tipos de dados a um banco de dados de documentos, não há necessidade de modificar todo o esquema do banco de dados como fazemos com um banco de dados relacional.
- fornece um mecanismo de consulta para pesquisar coleções de documentos com atributos específicos.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Armazém de Documentos

MODELO

DOCUMENTO XML

```
< artist >
  < artistname > Iron Maiden </artistname >
  < albums >
    < album >
      < albumname > The Book of Souls </albumname >
      < datereleased > 2015 </datereleased >
      < genre > Hard Rock </genre >
    </album >
    < album >
      < albumname > Killers </albumname >
      < datereleased > 1981 </datereleased >
      < genre > Hard Rock </genre >
    </album >
    < album >
      < albumname > Powerslave </albumname >
      < datereleased > 1984 </datereleased >
      < genre > Hard Rock </genre >
    </album >
    < album >
      < albumname > Somewhere in Time </albumname >
      < datereleased > 1986 </datereleased >
      < genre > Hard Rock </genre >
    </album >
  </albums >
</artist >
```

Fonte: Autor, 2021.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Armazém de Documentos

MODELO

DOCUMENTO JSON

```
{
  '_id' : 1,
  'artistName' : { 'Iron Maiden' },
  'albums' : [
    {
      'albumname' : 'The Book of Souls',
      'datereleased' : 2015,
      'genre' : 'Hard Rock'
    }, {
      'albumname' : 'Killers',
      'datereleased' : 1981,
      'genre' : 'Hard Rock'
    }, {
      'albumname' : 'Powerslave',
      'datereleased' : 1984,
      'genre' : 'Hard Rock'
    }, {
      'albumname' : 'Somewhere in Time',
      'datereleased' : 1986,
      'genre' : 'Hard Rock'
    }
  ]
}
```

Fonte: Autor, 2021.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Armazém de Documentos

Vantagens

- Modelagem de dados flexível que atende aos aplicativos emergentes e em evolução contínua como IoT, nuvem, analytics e outros.
- Alta performance de escrita pois priorizam a disponibilidade de gravação em relação à consistência estrita de dados.
- Desempenho rápido de consulta com poderosos mecanismos e recursos de indexação que fornecem consultas rápidas e eficientes.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Armazém de Documentos

Desvantagens

- Embora seja bom no armazenamento de documentos, não foi projetado para transações de fluxo empresarial que necessitem de durabilidade.
- Não são performáticos na execução de consultas complexas ou transações múltiplas.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Armazém de Documentos

Utilização

- Quaisquer dados que tenham uma estrutura de contêiner natural:
 - ✓ Documentos de escritório,
 - ✓ Pedidos de vendas,
 - ✓ Faturas,
 - ✓ Descrições de produtos,
 - ✓ Formulários e páginas da web,
 - ✓ Publicação,
 - ✓ Troca e pesquisa de documentos.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Orientado a Coluna

- Neste tipo de sistema de gerenciamento, os dados são armazenados em registros com a habilidade de manter um número elevado de colunas dinâmicas.
- Os nomes das colunas e das chaves dos registros não são fixos, e um registro pode ter literalmente bilhões de colunas.
- Por conta disso, um banco orientado a coluna é similar a um sistema chave/valor de duas dimensões.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Orientado a Coluna

MODELO

- Veja a evolução do armazenamento:

ID	name	address	zip code	phone	city	country
1	Benny Smith	23 Workhaven Lane	52683	14033335568	Lethbridge	Canada
2	Keith Page	1411 Lillydale Drive	18529	16172235589	Woodridge	Australia
3	John Doe	1936 Paper Blvd.	92512	14082384788	Santa Clara	USA

Organização tradicional: armazenados sequencialmente.

```
1,Benny Smith,23 Workhaven Lane,52683,14033335568,Lethbridge,Canada;  
2,Keith Page,1411 Lillydale Drive,18529,16172235589,Woodridge,Australia;  
3,John Doe,1936 Paper Blvd.,92512,14082384788,Santa Clara,USA;
```

Tabela bidimensional: orientado a linhas.

```
1,2,3;  
Benny Smith,Keith Page,John Doe;  
23 Workhaven Lane,1411 Lillydale Drive,1936 Paper lvd.;  
52683,18529,92512;  
14033335578,16172235589,14082384788;  
Lethbridge,Woodridge,Santa Clara;  
Canada,Australia,USA;
```

Tabela bidimensional: orientado a colunas.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Orientado a Coluna –

Vantagens

- As consultas podem ser mais rápidas;
- Permite que você ignore todos os dados não relevantes muito rapidamente;
- Como o tipo de dados para cada coluna é semelhante, você obtém melhor compactação ao executar algoritmos de compactação em cada coluna.
- Pense em um exemplo de cálculo de média destes dados.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Orientado a Coluna –

Desvantagens

- São ineficientes para casos onde precisa-se recuperar vários campos em uma linha.
- São ineficientes para consultas de valores específicos.
- São ineficiente para altos volumes de inserções e atualizações.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Orientado a Coluna –

Utilização

- Bancos de dados desta categoria são mais interessantes para:
 - ✓ Ferramentas de Analytics ;
 - ✓ Ferramentas de Mapas;
 - ✓ Armazenamento de dados de preferência de usuários;

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Grafo

- É um sistema que contém uma sequência de nós e relacionamentos que, quando combinados, criam um gráfico.
- O modelo orientado a grafos possui três campos de dados:
 - ✓ Nós: são os vértices do grafo;
 - ✓ Relacionamentos: são as arestas;
 - ✓ Propriedades (ou atributos) dos nós e relacionamentos;

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Grafo

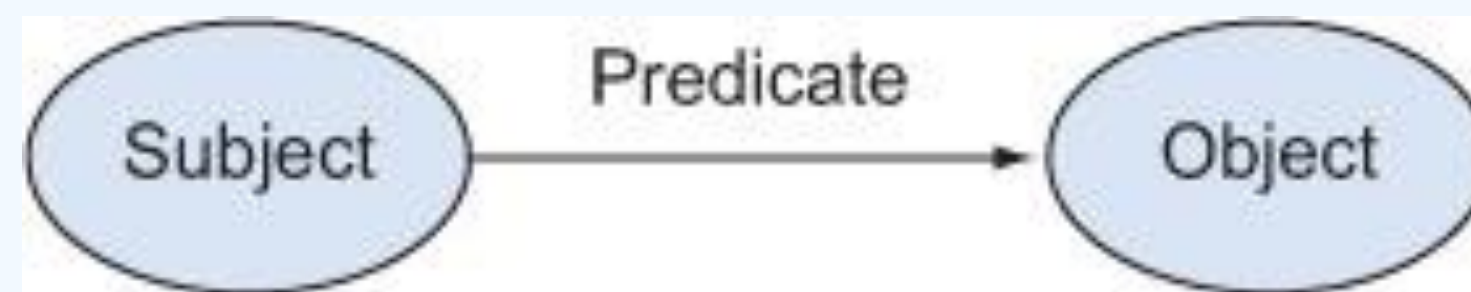
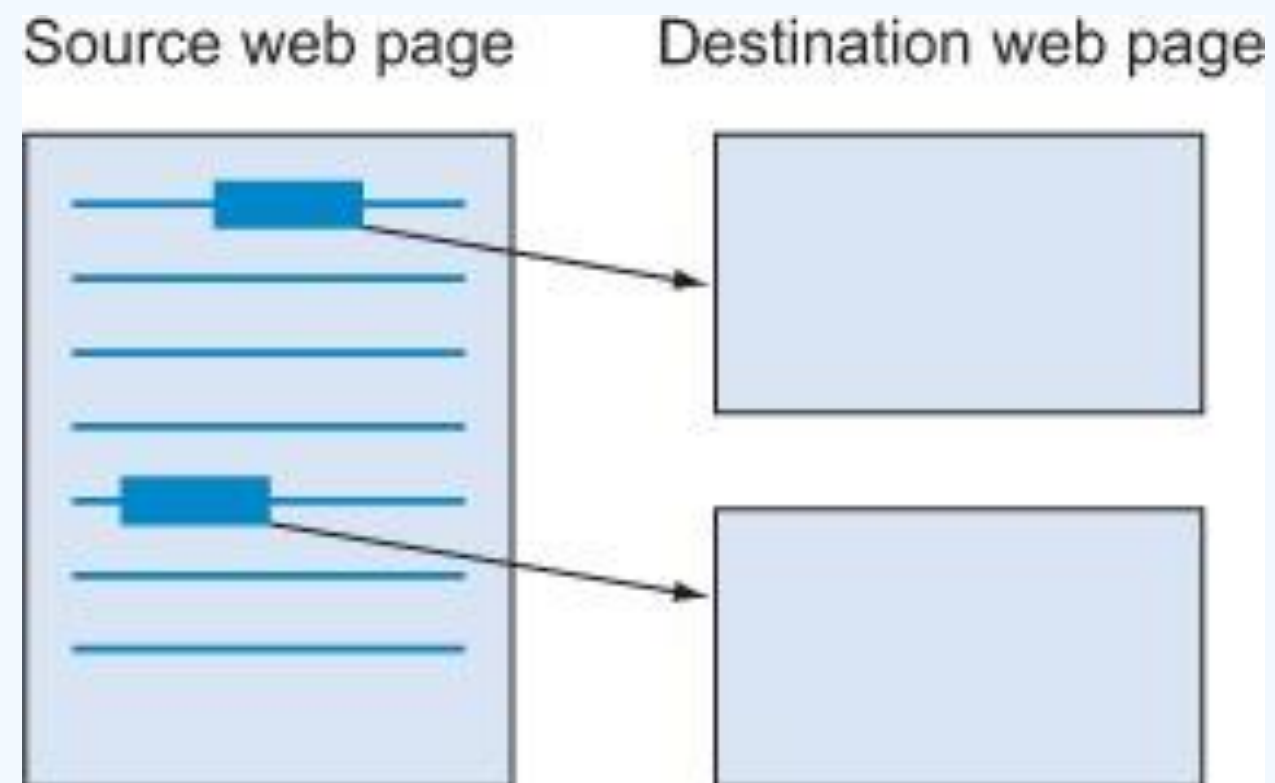
- São ideais quando você tem muitos itens relacionados entre si de maneiras complexas e esses relacionamentos têm propriedades.
- Permitem que você faça consultas simples que mostram os nós vizinhos mais próximos, bem como consultas que examinam profundamente as redes e encontram padrões rapidamente.
- Informam que existe um relacionamento entre os nós e podem fornecer relatórios detalhados sobre cada um de seus relacionamentos.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Grafo

MODELO



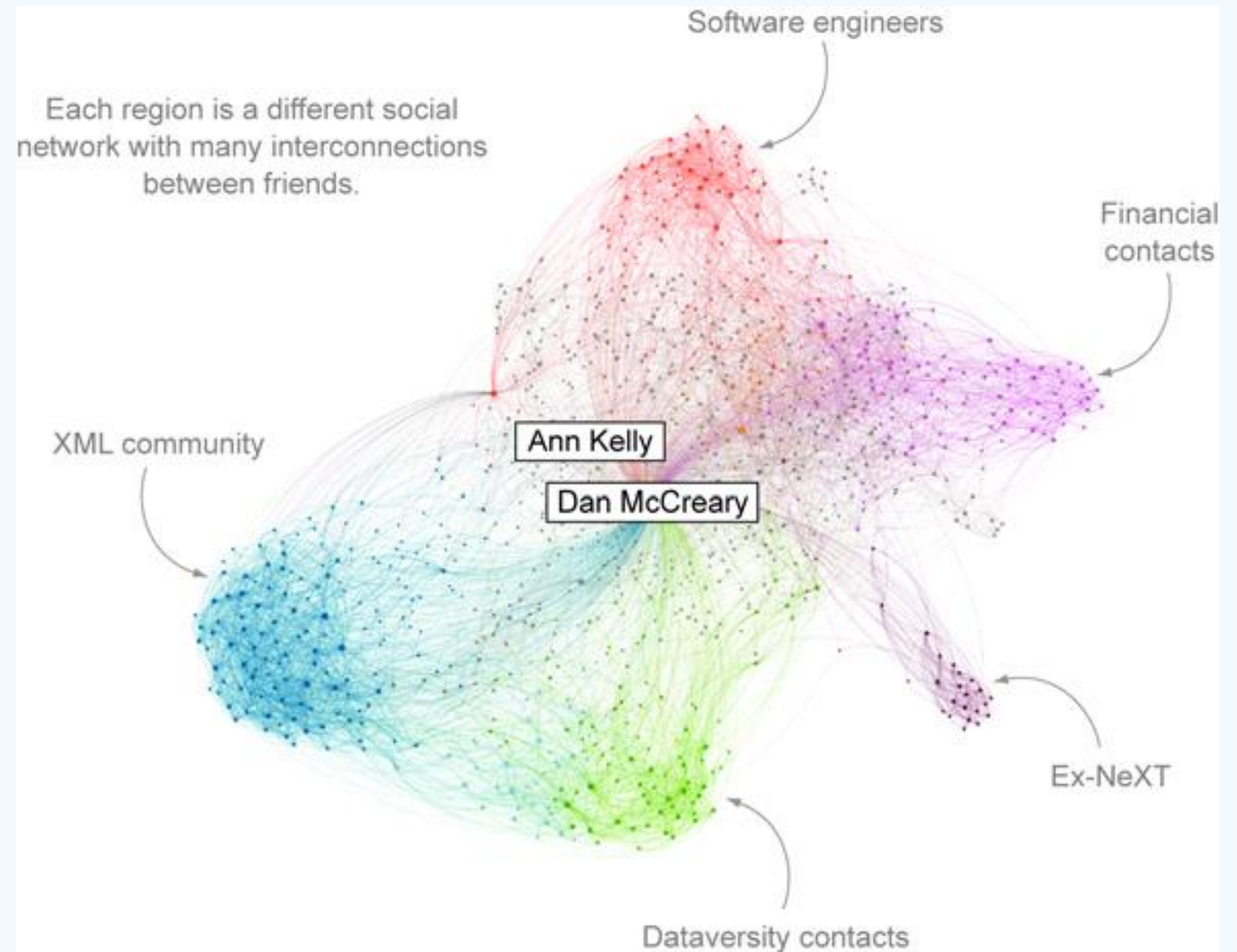
Fonte: MCCREARY, 2013.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Grafo

MODELO



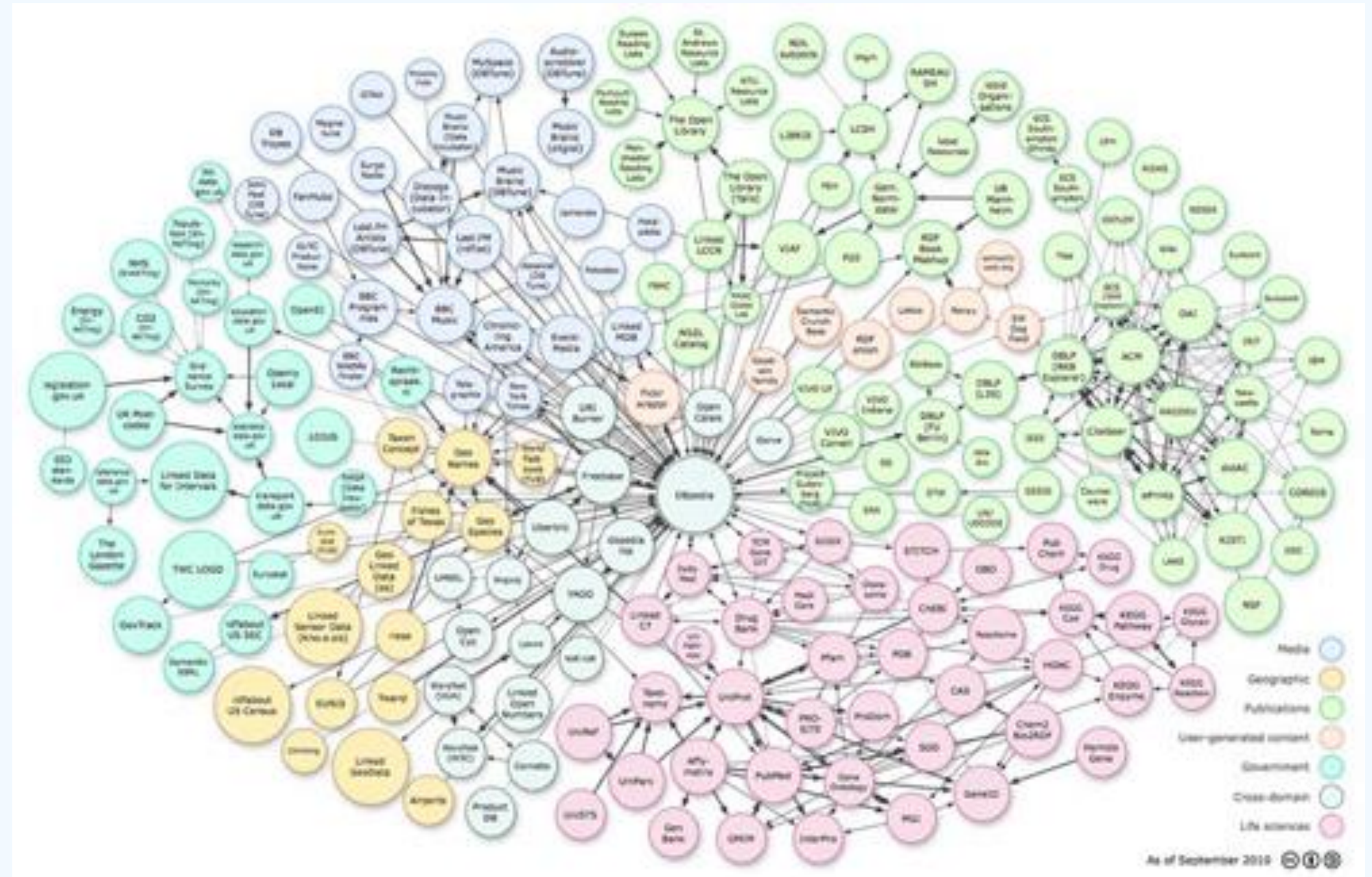
Fonte: MCCREARY, 2013.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Grafo

MODELO



Fonte: MCCREARY, 2013.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Grafo

Vantagens

- Representação rica de dados e suas relações com ações reais dos negócios com ampliação de detalhes.
- Habilita o conjunto de dados para descoberta de padrões e informações “ocultas”.
- Flexíveis, modificáveis e sustentáveis.
- Agilidade pois há alinhamento com o modelo de programação ágil implementado na maioria das empresas.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Grafo

Desvantagens

- Difíceis de dimensionar em vários servidores devido à estreita conexão de cada nó em um gráfico.
- Os dados podem ser replicados em vários servidores para aprimorar o desempenho de leitura e consulta, mas as gravações em vários servidores e as consultas em grafos que abrangem vários nós são complexas de implementar.

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Grafo

Utilização

- Detecção de fraude.
- Master Data Management (MDM).
- Rede e operações de TI.
- Gestão de Identidade e Acesso a serviços (IAM).

Tipos de bancos de dados NoSQL

Classificações de bancos NoSQL:

Motor de Busca

- É um banco de dados NoSQL cuja função é permitir a pesquisa de conteúdo de dados.
- Possui características bem peculiares, como o suporte à busca com o emprego de expressões complexas, busca de texto completo, ranqueamento e agrupamento de resultados, busca geoespacial e pesquisa distribuída, além de serem altamente otimizados para responder à alta demanda de pesquisas que podem receber.

*Now, we'll advance to
more concepts and rating NoSQL DB!*

REFERÊNCIAS BIBLIOGRÁFICAS

W3C. AJAX Introduction. Disponível em: < https://www.w3schools.com/js/js_ajax_intro.asp>. Acesso em 23 de nov. de 2021.

SOUZA, Ivan de. Afinal, o que é JSON e para que ele serve? Descubra agora! Disponível em: < <https://rockcontent.com/br/blog/json/#4> >. Acesso em 23 de nov. de 2021.

REIS, Fábio dos. O que são Bancos de Dados NoSQL. Disponível em: < <http://www.bosontreinamentos.com.br/nosql/bancos-de-dados-o-que-e-nosql/> >. Acesso em 23 de nov. de 2021.

BOERES, Cristina. ESTRUTURA DE DADOS E ALGORITMOS Hashing (Tabela de Dispersão). Disponível em: < http://www2.ic.uff.br/~boeres/slides_ed/ed_TabelaHash.pdf >. Acesso em 23 de nov. de 2021.

BARANAUSKAS, Augusto. Tabelas Hash, 2011. Disponível em: < https://dcm.ffclrp.usp.br/~augusto/teaching/icii/_vti_cnf/Hash-Tables-Apresentacao.pdf > Acesso em 23 de nov. de 2021.

MCCREARY Dan, KELLY Anna. Making Sense of NoSQL. Manning Publications. Shelter Island, NY. 2013.