Stop! Don't make these mistakes in your document database!

@Lauren_Schaefer


Citytv


Citytv

Learn about and prevent schema design anti-patterns.

Massive arrays

Massive number of collections

4

Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

Massive arrays

Massive number of collections

4

Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

Massive arrays

Massive number of collections

4

Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

Massive arrays

Massive number of collections

4

Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

Massive arrays

Massive number of collections

4

Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

Massive arrays

Massive number of collections

4

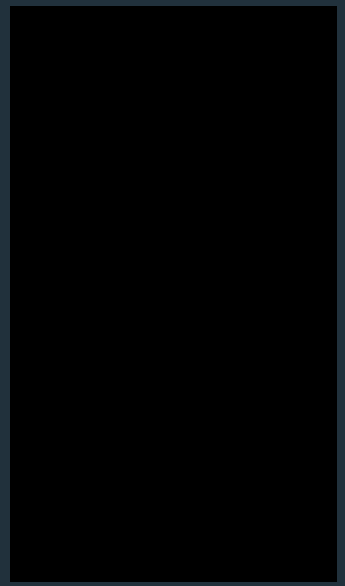
Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

@Lauren_Schaefer

Developer Advocate, MongoDB



@Lauren_Schaefer

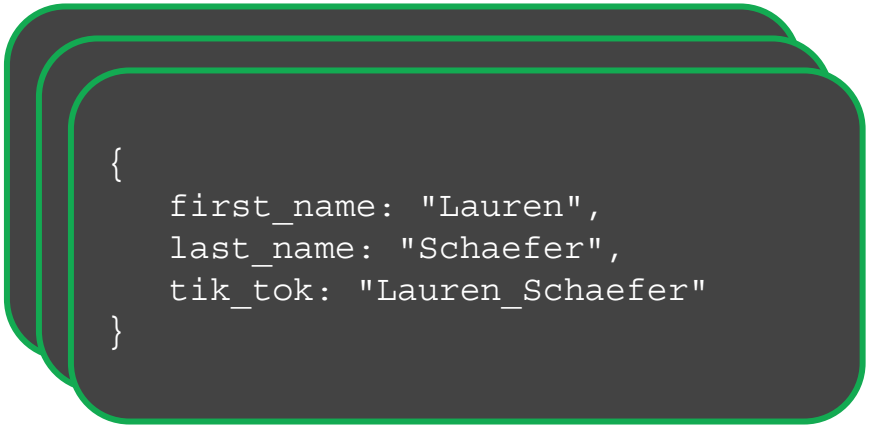
Developer Advocate, MongoDB

Document

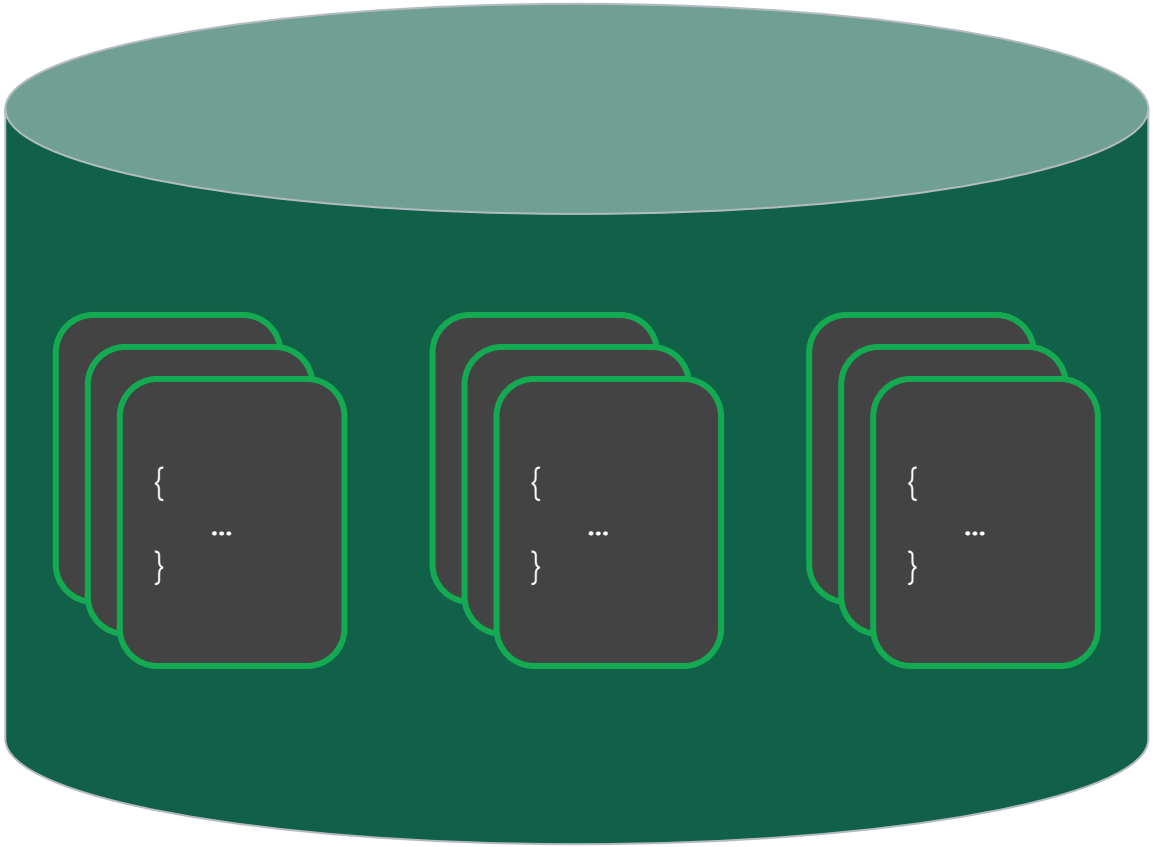
```
{
    first_name: "Lauren",
    last_name: "Schaefer",
    tik_tok: "Lauren_Schaefer"
}
```

Collection

Users



Database



Massive arrays

Massive number of collections

4

Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

Data that is accessed together should be stored together.

Massive Arrays

```
{
    a: "b",
    c: "d",
    e: {
        f: "g",
        h: "i"
    },
    j: ["k", "l", "m"]
}
```

Massive Arrays

```
{
    a: "b",
    c: "d",
    e: {
        f: "g",
        h: "i"
    },
    j: ["k", "l", "m"]
}
```

Massive Arrays

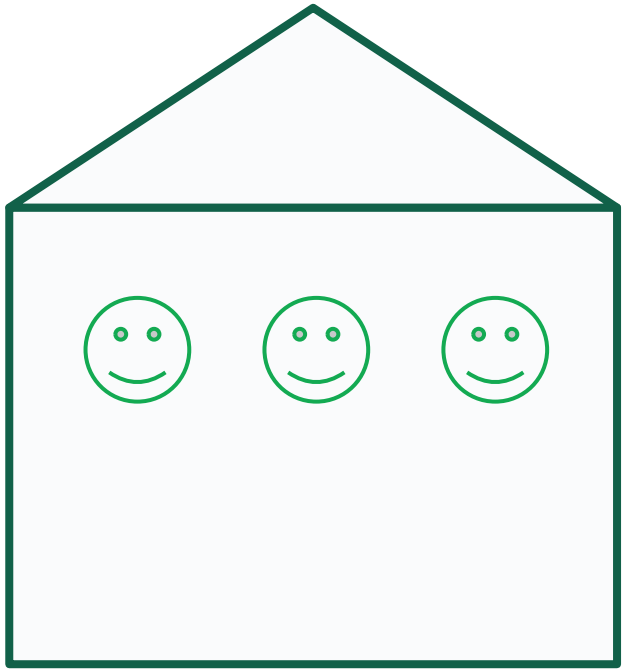
```
{
    a: "b",
    c: "d",
    e: {
        f: "g",
        h: "i"
    },
    j: ["k", "l", "m"]
}
```

The Problem

- May exceed document size limits
- Index performance on arrays decreases as array size increases


I MET
LIL SEBASTIAN

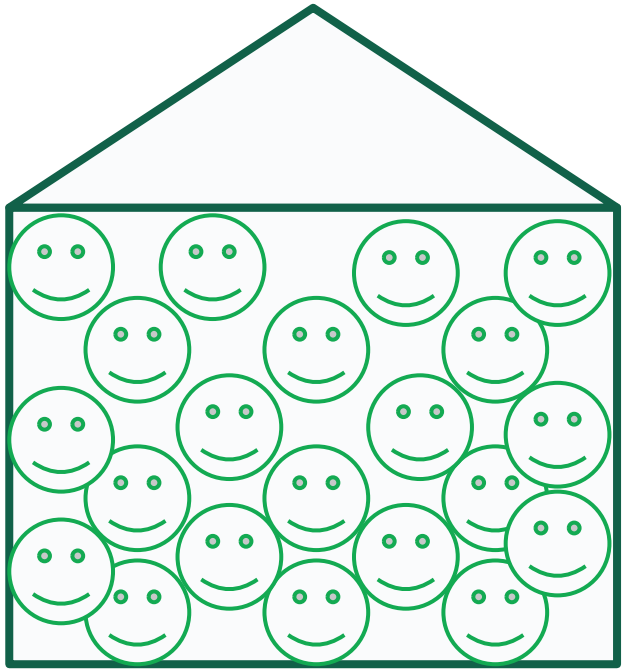
Massive Arrays



Buildings

```
{
   "_id": "city_hall",
   "name": "City Hall",
   "city": "Pawnee",
   "state": "IN",
   "employees": [
      {
         "_id": 123456789,
         "first": "Leslie",
         "last": "Yepp",
         "cell": "8125552344",
         "start-year": "2004"
      },
      {
         "_id": 234567890,
         "first": "Ron",
         "last": "Swandaughter",
         "cell": "8125559347",
         "start-year": "2002"
      }
   ]
}
```

Massive Arrays



Buildings

```
{
    "_id": "city_hall",
    "name": "City Hall",
    "city": "Pawnee",
    "state": "IN",
    "employees": [
        {
            "_id": 123456789,
            "first": "Leslie",
            "last": "Yepp",
            "cell": "8125552344",
            "start-year": "2004"
        },
        {
            "_id": 234567890,
            "first": "Ron",
            "last": "Swandaughter",
            "cell": "8125559347",
            "start-year": "2002"
        },
        {
            "_id": 345678901,
            "first": "Andy",
            "last": "Fryer",
            "cell": "8125552341",
```

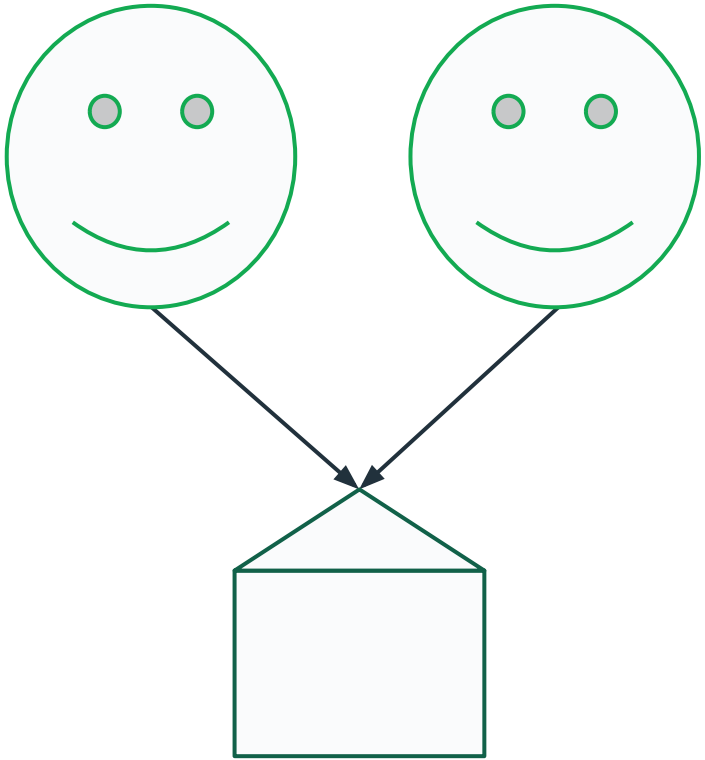
Massive Arrays

Employees

```
{
    "_id": 123456789,
    "first": "Leslie",
    "last": "Yepp",
    "cell": "8125552344",
    "start-year": "2004",
    "building": {
        "_id": "city_hall",
        "name": "City Hall",
        "city": "Pawnee",
        "state": "IN"
    }
}
```

```
{
    "_id": 234567890,
    "first": "Ron",
    "last": "Swandaughter",
    "cell": "8125559347",
    "start-year": "2002",
    "building": {
        "_id": "city_hall",
        "name": "City Hall",
        "city": "Pawnee",
        "state": "IN"
    }
}
```

Massive Arrays



Employees

```
{
    "_id": 123456789,
    "first": "Leslie",
    "last": "Yepp",
    "cell": "8125552344",
    "start-year": "2004",
    "building_id": "city_hall"
}
```

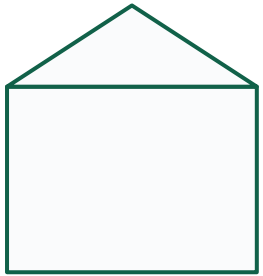
```
{
    "_id": 234567890,
    "first": "Ron",
    "last": "Swandaughter",
    "cell": "8125559347",
    "start-year": "2002",
    "building_id": "city_hall"
}
```

Buildings

```
{
    "_id": "city_hall",
    "name": "City Hall",
    "city": "Pawnee",
    "state": "IN"
}
```

The extended reference pattern

Massive Arrays

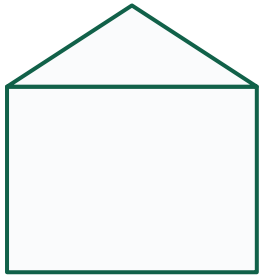


Massive Arrays

Employees

```
{
    "_id": 123456789,
    "first": "Leslie",
    "last": "Yepp",
    "cell": "8125552344",
    "start-year": "2004",
    "building": {
        "name": "City Hall",
        "state": "IN"
    }
}
```

```
{
    "_id": 234567890,
    "first": "Ron",
    "last": "Swandaughter",
    "cell": "8125559347",
    "start-year": "2002",
    "building": {
        "name": "City Hall",
        "state": "IN"
    }
}
```



Buildings

```
{
    "_id": "city_hall",
    "name": "City Hall",
    "city": "Pawnee",
    "state": "IN"
}
```

Summary

Summary

- Do: Store information together that you'll be frequently querying together

Summary

- Do: Store information together that you'll be frequently querying together
- Don't: Store information in massive, unbounded arrays

Massive arrays

Massive number of collections

4

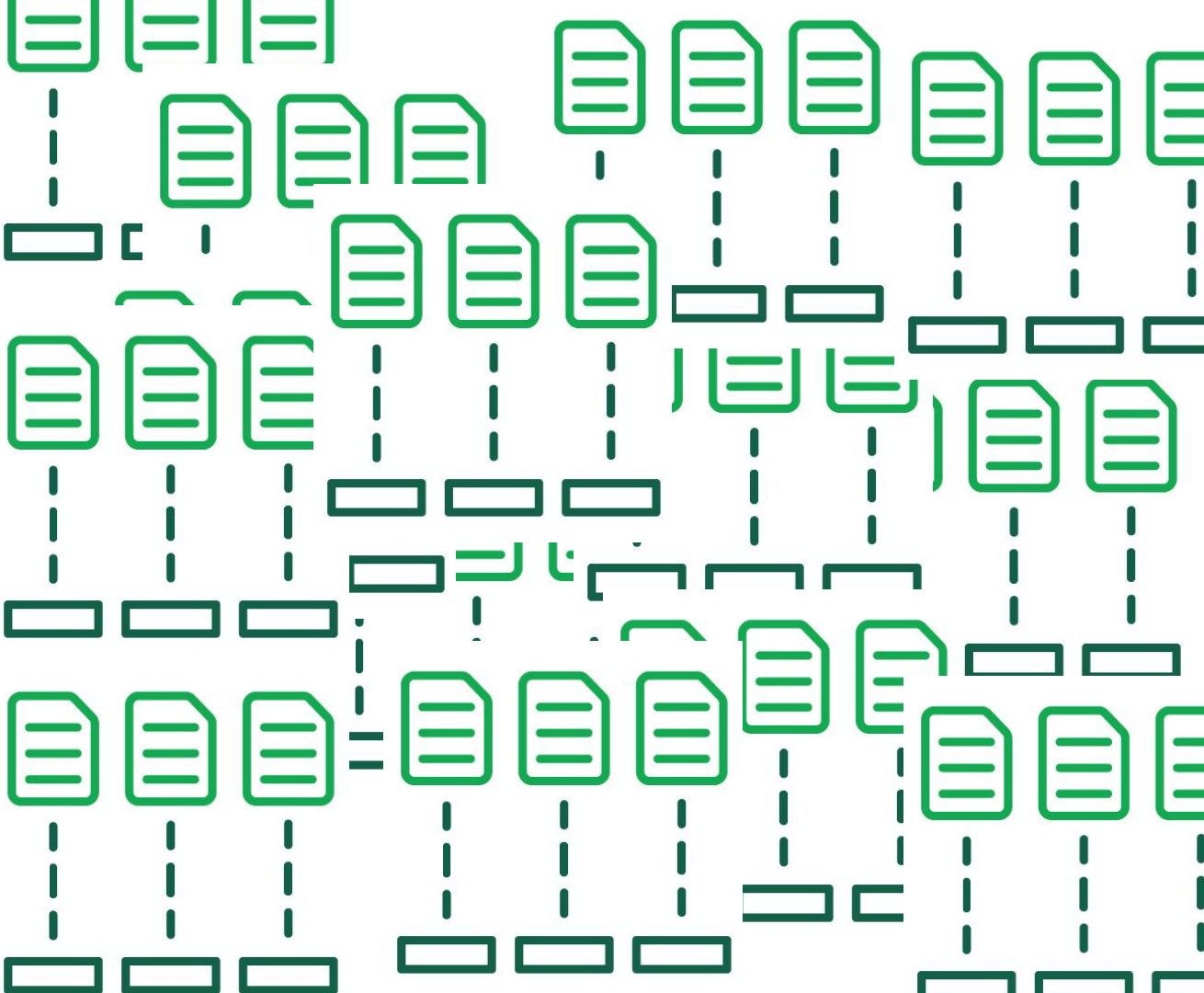
Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

Massive Number of Collections

The Problem



Massive Number of Collections

Massive Number of Collections

Massive Number of Collections

riverstats

DATABASE SIZE: 5.2GB INDEX SIZE: 1.07GB TOTAL COLLECTIONS: 365

riverstats-v2

DATABASE SIZE: 3.07GB INDEX SIZE: 27.45MB TOTAL COLLECTIONS: 1

Massive Number of Collections

Collections To Drop

- Empty collections
- Collections whose size is mostly indexes

Summary

Summary

- Don't: create a massive number of collections

Massive arrays

Massive number of collections

4

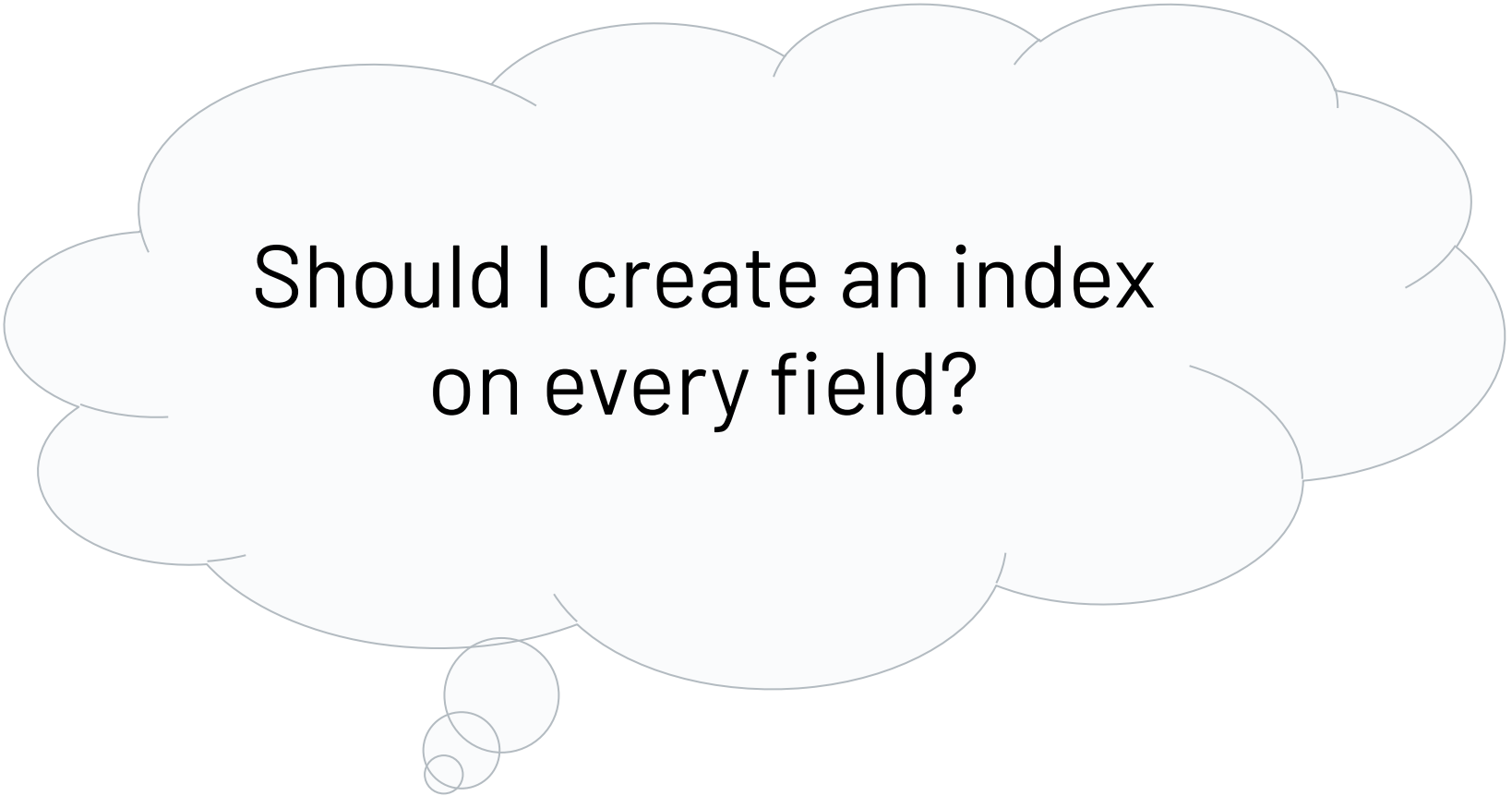
Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

Indexes are good.

Indexes are good.


Should I create an index
on every field?

Unnecessary Indexes

No.

The Problem

- Indexes take up space
- Indexes can negatively impact write performance

Indexes to Drop

- Rarely used indexes
- Redundant indexes

Inspirational Women of the World

Search for your favorite inspirational woman!

Full name:

Search

Last name:

Search

Hobbies:

- [ ] Boating
- [ ] Playing musical instruments
- [ ] Reading
- [ ] Tennis
- [ ] Writing
- [ ] Video games
- [ ] Visiting Li'l Sebastian

Search


I wonder if I'm
gonna make it
awesome...


I wonder if I'm
gonna make it
awesome...

Summary

Summary

- Do: Create indexes that support frequent queries

Summary

- Do: Create indexes that support frequent queries
- Don't: Create unnecessary indexes

Massive arrays

Massive number of collections

4

Schema Design Anti-Patterns

Unnecessary indexes

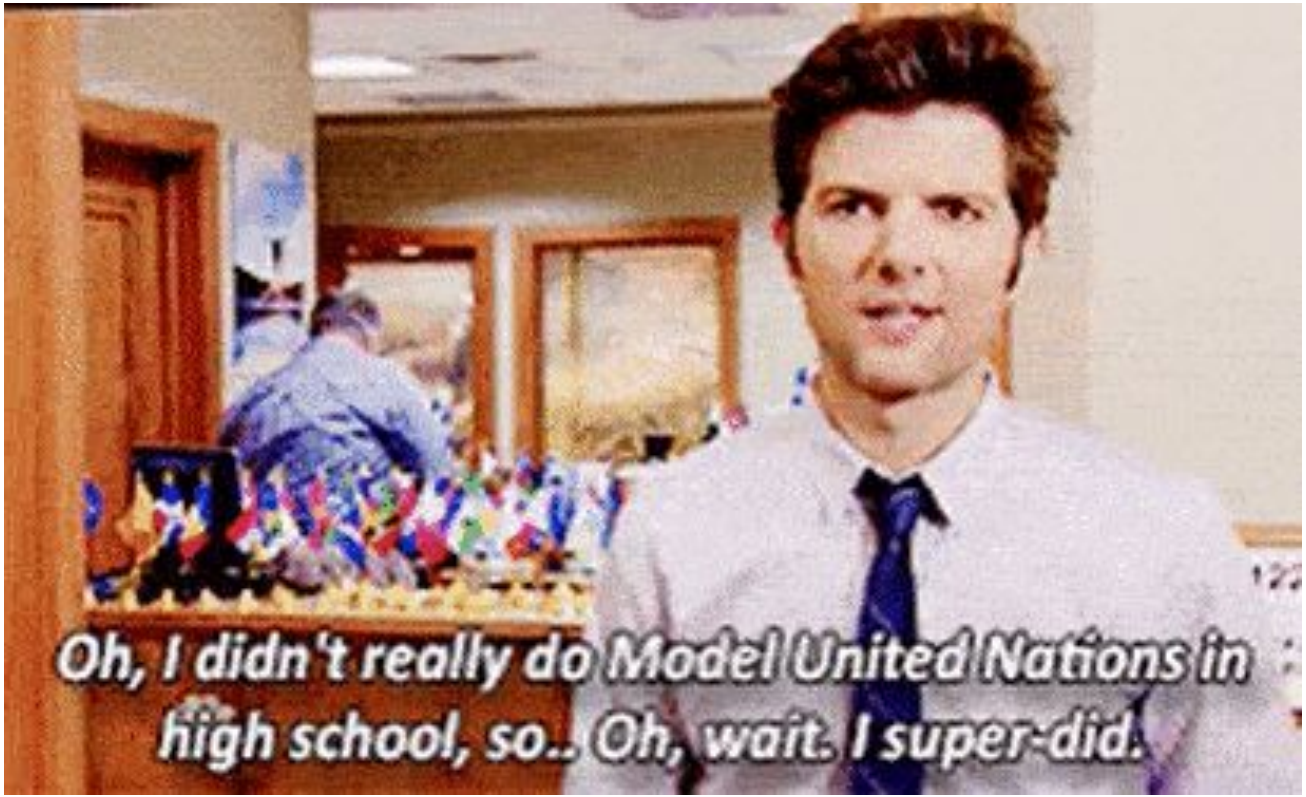
Separating data that is accessed together

The Problem

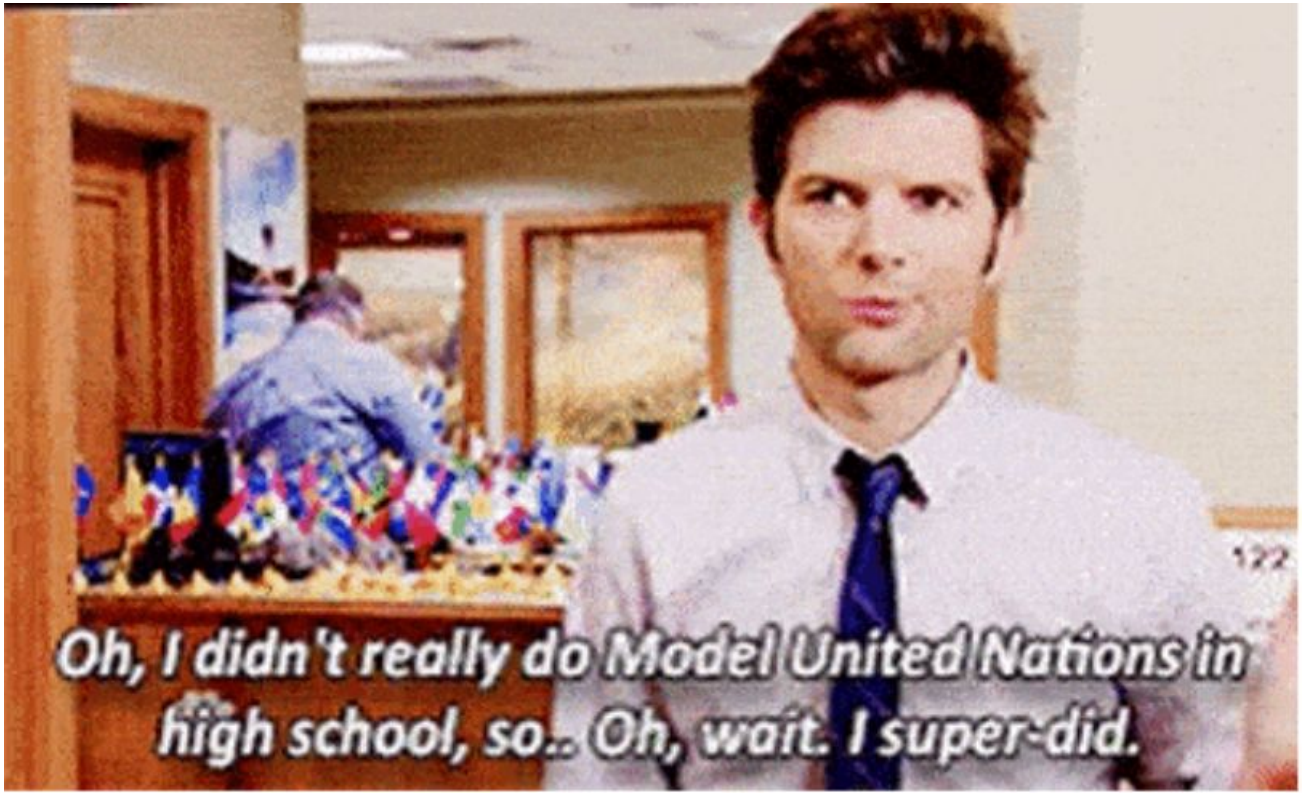
- Not all document databases support joins
- Joins can be slow and resource-intensive

Data that is accessed together should be stored together.

Separating Data That is Accessed Together

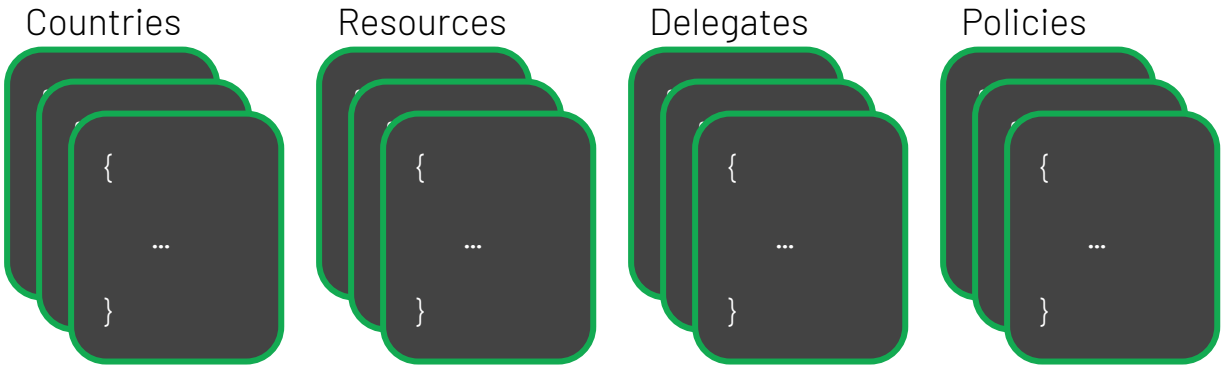


Separating Data That is Accessed Together



Data to store

- Basic stats about each country
- List of resources that each country has available to trade
- List of delegates for each country
- Policy statements for each country



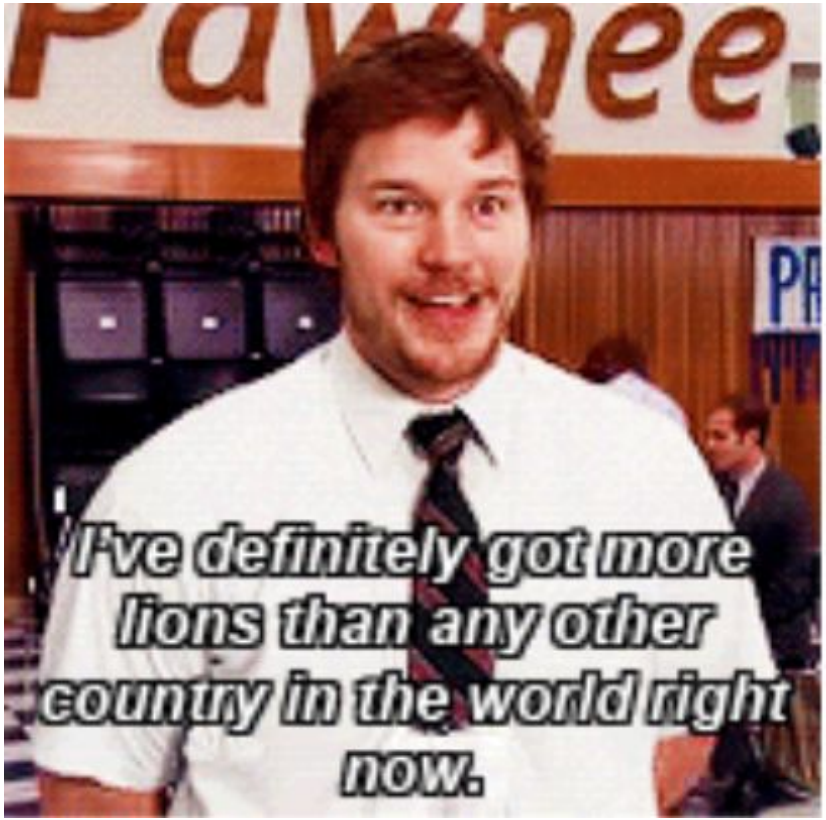
Country Report

- Basic stats
- Resources available to trade
- Delegates
- Names and dates of last five policy documents

Separating Data That is Accessed Together

Separating Data That is Accessed Together

Separating Data That is Accessed Together



Separating Data That is Accessed Together

Countries

```
{
   "_id": "finland",
   "official_name": "Republic of Finland",
   "capital": "Helsinki",
   "languages": [
      "Finnish",
      "Swedish",
      "Sámi"
   ],
   "population": 5528737
}
```

Delegates

```
{
   "_id": ObjectId("5ef0ff710d9314ac117d2036"),
   "country_id": "finland",
   "first_name": "Donna",
   "last_name": "Beagle"
}
```

Resources

```
{
   "_id":  ObjectId("5ef0feeb0d9314ac117d2034"),
   "country_id": "finland",
   "lions": 32563,
   "military_personnel": 0,
   "pulp": 0,
   "paper": 0
}
```

Policies

```
{
   "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
   "date-created":
      ISODate("2011-11-09T04:00:00.000+00:00"),
   "status": "draft",
   "title": "Country Defense Policy",
   "country_id": "finland",
   "policy": "Finland has formally decided to
      use lions in lieu of military for all
      self defense..."
}
```

Country Report

- Basic stats
- Resources available to trade
- Delegates
- Names and dates of last five policy documents

Separating Data That is Accessed Together

Countries

```
{
   "_id": "finland",
   "official_name": "Republic of Finland",
   "capital": "Helsinki",
   "languages": [
      "Finnish",
      "Swedish",
      "Sámi"
   ],
   "population": 5528737
}
```

Delegates

```
{
   "_id": ObjectId("5ef0ff710d9314ac117d2036"),
   "country_id": "finland",
   "first_name": "Donna",
   "last_name": "Beagle"
}
```

Resources

```
{
   "_id":  ObjectId("5ef0feeb0d9314ac117d2034"),
   "country_id": "finland",
   "lions": 32563,
   "military_personnel": 0,
   "pulp": 0,
   "paper": 0
}
```

Policies

```
{
   "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
   "date-created":
      ISODate("2011-11-09T04:00:00.000+00:00"),
   "status": "draft",
   "title": "Country Defense Policy",
   "country_id": "finland",
   "policy": "Finland has formally decided to
      use lions in lieu of military for all
      self defense..."
}
```

Separating Data That is Accessed Together

Countries

```
{
    "_id": "finland",
    "official_name": "Republic of Finland",
    "capital": "Helsinki",
    "languages": [
        "Finnish",
        "Swedish",
        "Sámi"
    ],
    "population": 5528737,
    "resources": {
        "lions": 32563,
        "military_personnel": 0,
        "pulp": 0,
        "paper": 0
    }
}
```

```
{
    "_id": ObjectId("5ef0ff710d9314ac117d2054"),
    "country_id": "finland",
    "lions": 32563,
    "military_personnel": 0,
    "pulp": 0,
    "paper": 0
}
```

Delegates

```
{
    "_id": ObjectId("5ef0ff710d9314ac117d2036"),
    "country_id": "finland",
    "first_name": "Donna",
    "last_name": "Beagle"
}
```

Policies

```
{
    "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
    "date-created":
        ISODate("2011-11-09T04:00:00.000+00:00"),
    "status": "draft",
    "title": "Country Defense Policy",
    "country_id": "finland",
    "policy": "Finland has formally decided to
        use lions in lieu of military for all
        self defense..."
}
```

Separating Data That is Accessed Together

Countries

```
{
    "_id": "finland",
    "official_name": "Republic of Finland",
    "capital": "Helsinki",
    "languages": [
        "Finnish",
        "Swedish",
        "Sámi"
    ],
    "population": 5528737,
    "resources": {
        "lions": 32563,
        "military_personnel": 0,
        "pulp": 0,
        "paper": 0
    }
}
```

Delegates

```
{
    "_id": ObjectId("5ef0ff710d9314ac117d2036"),
    "country_id": "finland",
    "first_name": "Donna",
    "last_name": "Beagle"
}
```

Policies

```
{
    "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
    "date-created":
        ISODate("2011-11-09T04:00:00.000+00:00"),
    "status": "draft",
    "title": "Country Defense Policy",
    "country_id": "finland",
    "policy": "Finland has formally decided to
        use lions in lieu of military for all
        self defense..."
}
```

Separating Data That is Accessed Together

Countries

```
{
    "_id": "finland",
    "official_name": "Republic of Finland",
    "capital": "Helsinki",
    "languages": [
        "Finnish",
        "Swedish",
        "Sámi"
    ],
    "population": 5528737,
    "resources": {
        "lions": 32563,
        "military_personnel": 0,
        "pulp": 0,
        "paper": 0
    }
}
```

Delegates

```
{
    "_id": ObjectId("5ef0ff710d9314ac117d2036"),
    "country_id": "finland",
    "first_name": "Donna",
    "last_name": "Beagle"
}
```

Policies

```
{
    "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
    "date-created":
        ISODate("2011-11-09T04:00:00.000+00:00"),
    "status": "draft",
    "title": "Country Defense Policy",
    "country_id": "finland",
    "policy": "Finland has formally decided to
        use lions in lieu of military for all
        self defense..."
}
```

Separating Data That is Accessed Together

Countries

```
{
    "_id": "finland",
    "official_name": "Republic of Finland",
    "capital": "Helsinki",
    "languages": [
        "Finnish",
        "Swedish",
        "Sámi"
    ],
    "population": 5528737,
    "resources": {
        "lions": 32563,
        "military_personnel": 0,
        "pulp": 0,
        "paper": 0
    },
    "delegates": [
        {
            "first_name": "Andy",
            "last_name": "Fryer"
        },
        {
            "first_name": "Donna",
            "last_name": "Beagle"
        }
    ]
}
```

Delegates

```
{
    "_id": ObjectId("5ef0ff710d9314ac117d2036"),
    "country_id": "finland",
    "first_name": "Donna",
    "last_name": "Beagle"
}
```

Policies

```
{
    "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
    "date-created":
        ISODate("2011-11-09T04:00:00.000+00:00"),
    "status": "draft",
    "title": "Country Defense Policy",
    "country_id": "finland",
    "policy": "Finland has formally decided to
        use lions in lieu of military for all
        self defense..."
}
```

Separating Data That is Accessed Together

Countries

```
{
    "_id": "finland",
    "official_name": "Republic of Finland",
    "capital": "Helsinki",
    "languages": [
       "Finnish",
       "Swedish",
       "Sámi"
    ],
    "population": 5528737,
    "resources": {
       "lions": 32563,
       "military_personnel": 0,
       "pulp": 0,
       "paper": 0
    },
    "delegates": [
       {
          "first_name": "Andy",
          "last_name": "Fryer"
       },
       {
          "first_name": "Donna",
          "last_name": "Beagle"
       }
    ]
}
```

Policies

```
{
    "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
    "date-created":
       ISODate("2011-11-09T04:00:00.000+00:00"),
    "status": "draft",
    "title": "Country Defense Policy",
    "country_id": "finland",
    "policy": "Finland has formally decided to
       use lions in lieu of military for all
       self defense..."
}
```

Separating Data That is Accessed Together

Countries

```
{
    "_id": "finland",
    "official_name": "Republic of Finland",
    "capital": "Helsinki",
    "languages": [
       "Finnish",
       "Swedish",
       "Sámi"
    ],
    "population": 5528737,
    "resources": {
       "lions": 32563,
       "military_personnel": 0,
       "pulp": 0,
       "paper": 0
    },
    "delegates": [
       {
          "first_name": "Andy",
          "last_name": "Fryer"
       },
       {
          "first_name": "Donna",
          "last_name": "Beagle"
       }
    ]
}
```

Policies

```
{
    "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
    "date-created":
       ISODate("2011-11-09T04:00:00.000+00:00"),
    "status": "draft",
    "title": "Country Defense Policy",
    "country_id": "finland",
    "policy": "Finland has formally decided to
       use lions in lieu of military for all
       self defense..."
}
```

Separating Data That is Accessed Together

Countries

```
{
   "_id": "finland",
   "official_name": "Republic of Finland",
   "capital": "Helsinki",
   "languages": [
      "Finnish",
      "Swedish",
      "Sámi"
   ],
   "population": 5528737,
   "resources": {
      "lions": 32563,
      "military_personnel": 0,
      "pulp": 0,
      "paper": 0
   },
   "delegates": [
      ...
   ],
    "recent-policies": [
        {
            "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
            "date-created":
              ISODate("2011-11-09T04:00:00.000+00:00"),
            "title": "Country Defense Policy"
        },
        {
            "_id": ObjectId("5ef357bb3e5f7febbd3ed7fd"),
            "date-created":
              ISODate("2011-11-10T04:00:00.000+00:00"),
            "title": "Humanitarian Food Policy"
        }
    ]
}
```

Policies

```
{
   "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
   "date-created":
      ISODate("2011-11-09T04:00:00.000+00:00"),
   "status": "draft",
   "title": "Country Defense Policy",
   "country_id": "finland",
   "policy": "Finland has formally decided to
      use lions in lieu of military for all
      self defense..."
}
```

Separating Data That is Accessed Together

Countries

```
{
   "_id": "finland",
   "official_name": "Republic of Finland",
   "capital": "Helsinki",
   "languages": [
      "Finnish",
      "Swedish",
      "Sámi"
   ],
   "population": 5528737,
   "resources": {
      "lions": 32563,
      "military_personnel": 0,
      "pulp": 0,
      "paper": 0
   },
   "delegates": [
      ...
   ],
    "recent-policies": [
        {
            "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
            "date-created":
              ISODate("2011-11-09T04:00:00.000+00:00"),
            "title": "Country Defense Policy"
        },
        {
            "_id": ObjectId("5ef357bb3e5f7febbd3ed7fd"),
            "date-created":
              ISODate("2011-11-10T04:00:00.000+00:00"),
            "title": "Humanitarian Food Policy"
        }
    ]
}
```

Policies

```
{
   "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
   "date-created":
      ISODate("2011-11-09T04:00:00.000+00:00"),
   "status": "draft",
   "title": "Country Defense Policy",
   "country_id": "finland",
   "policy": "Finland has formally decided to
      use lions in lieu of military for all
      self defense..."
}
```

Separating Data That is Accessed Together

Countries

```
{
    "_id": "finland",
    "official_name": "Republic of Finland",
    "capital": "Helsinki",
    "languages": [
        ...
    ],
    "population": 5528737,
    "resources": {
        ...
    },
    "delegates": [
        ...
    ],
    "recent-policies": [
        ...
    ],
    "events": [
        {
            "event-id": ObjectId("5ef34faa3e5f7febbd3ed7fc"),
            "event-date":
              ISODate("2011-11-10T05:00:00.000+00:00"),
            "topic": "Global Food Crisis"
        },
        {
            "event-id": ObjectId("5ef35ac93e5f7febbd3ed7fe"),
            "event-date":
              ISODate("2012-02-18T05:00:00.000+00:00"),
            "topic": "Pandemic"
        }
    ]
}
```

Policies

```
{
    "_id": ObjectId("5ef34ec43e5f7febbd3ed7fb"),
    "date-created":
       ISODate("2011-11-09T04:00:00.000+00:00"),
    "status": "draft",
    "title": "Country Defense Policy",
    "country_id": "finland",
    "policy": "Finland has formally decided to
       use lions in lieu of military for all
       self defense..."
}
```

Summary

Summary

- Do: Carefully consider your use case as you design your schema

Summary

- Do: Carefully consider your use case as you design your schema
- Don't: Separate data that is accessed together

Massive arrays

Massive number of collections

4

Schema Design Anti-Patterns

Unnecessary indexes

Separating data that is accessed together

Massive arrays

Massive number of collections

Unnecessary indexes

Separating data that is accessed together

Massive arrays

Massive number of collections

Unnecessary indexes

Separating data that is accessed together

Massive arrays

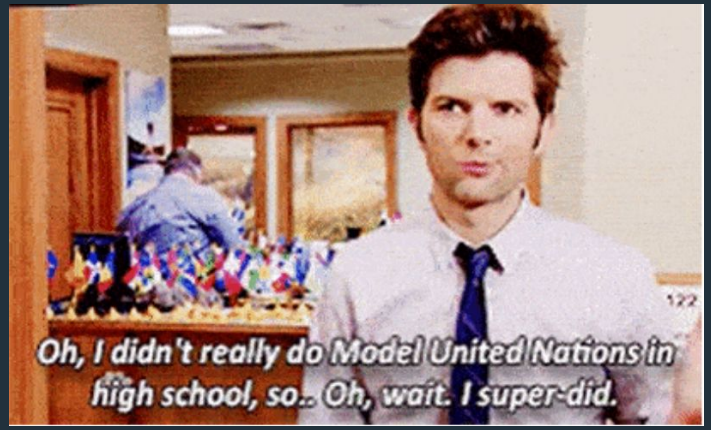
Massive number of collections

Unnecessary indexes

Separating data that is accessed together

Massive arrays

Massive number of collections



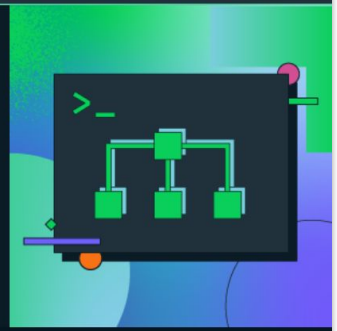
Unnecessary indexes

Separating data that is accessed together

A Summary of Schema Design Anti-Patterns and How to Spot Them

Published: Jul 14, 2020 MONGODB SCHEMA DESIGN

By Lauren Schaefer and Daniel Coupal

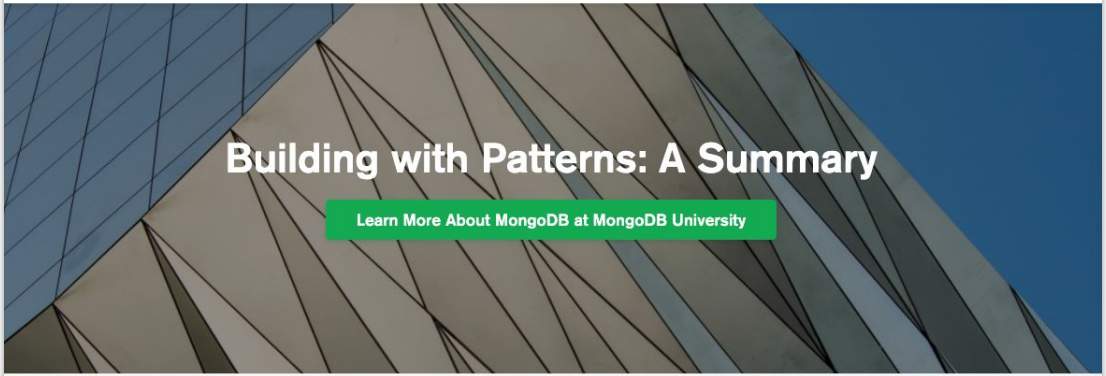


We've reached the final post in this series on MongoDB schema design anti-patterns. You're an expert now, right? We hope so. But don't worry—even if you fall into the trap of accidentally implementing an anti-pattern, MongoDB Atlas can help you identify it.

The Anti-Patterns

Below is a brief description of each of the schema design anti-patterns we've covered in this series.

- Massive arrays: storing massive, unbounded arrays in your documents.
- Massive number of collections: storing a massive number of collections (especially if they are unused or unnecessary) in your database.
- Unnecessary indexes: storing an index that is unnecessary because it is (1) rarely used if at all or (2) redundant because another compound index covers it.
- Bloated documents: storing large amounts of data together in a document when that data is not frequently accessed together.
- Separating data that is accessed together: separating data between different documents and collections that is frequently accessed together.
- Case-insensitive queries without case-insensitive indexes: frequently executing a case-insensitive query without having a case-insensitive index to cover it.




Daniel Coupal and Ken W. Alger
April 26, 2019
#Developer #University

As we wrap up the *Building with Patterns* series, it's a good opportunity to recap the problems the patterns that have been covered solve and highlight some of the benefits and trade-offs each pattern has. The most frequent question that is asked about schema design patterns, is "I'm designing an application to do X, how do I model the data?" As we hope you have discovered over the course of this blog series, there are a lot of things to take into consideration to answer that. However, we've included a *Sample Use Case* chart that we've found helpful to at least provide some initial guidance on data modeling patterns for generic use cases.

Sample Use Cases

The chart below is a guideline for what we've found after years of experience working with our customers of what schema design patterns are used in a variety of applications. This is not a "set in stone" set of rules about which design pattern can be used for a particular type of application. Ensure you look at the ones that are frequently used in your use case. However, don't discard the other ones, they may still apply. How you design your application's data schema is very dependent on your data access patterns.

MongoDB University

All Courses | Certification | Sign In

Free • 5 Chapters • Online

M320: Data Modeling

Learn everything you need to know about data modeling for MongoDB.

Courses > M320

Next Session:

Start: Sep 08, 2020 at 17:00 UTC

End: Nov 10, 2020 at 17:00 UTC

Register

Instructor: Daniel Coupal

Daniel is a Senior Curriculum Engineer on the Education team at MongoDB. Daniel was part of the Technical Support Team at MongoDB before transferring to the Education Department. Prior to MongoDB,

What You'll Learn

After completing this course, you should have a good understanding of how to create data models for MongoDB.

We will go over a few techniques, from a very simple process for simple schemas to more complex ones for large teams and large projects.

Prerequisites:
M001 and software architecture experience with data modeling or MongoDB experience in general.

What You'll Build

You'll build a solid understanding of frequent patterns to apply when modeling and will be able to apply those in your designs.

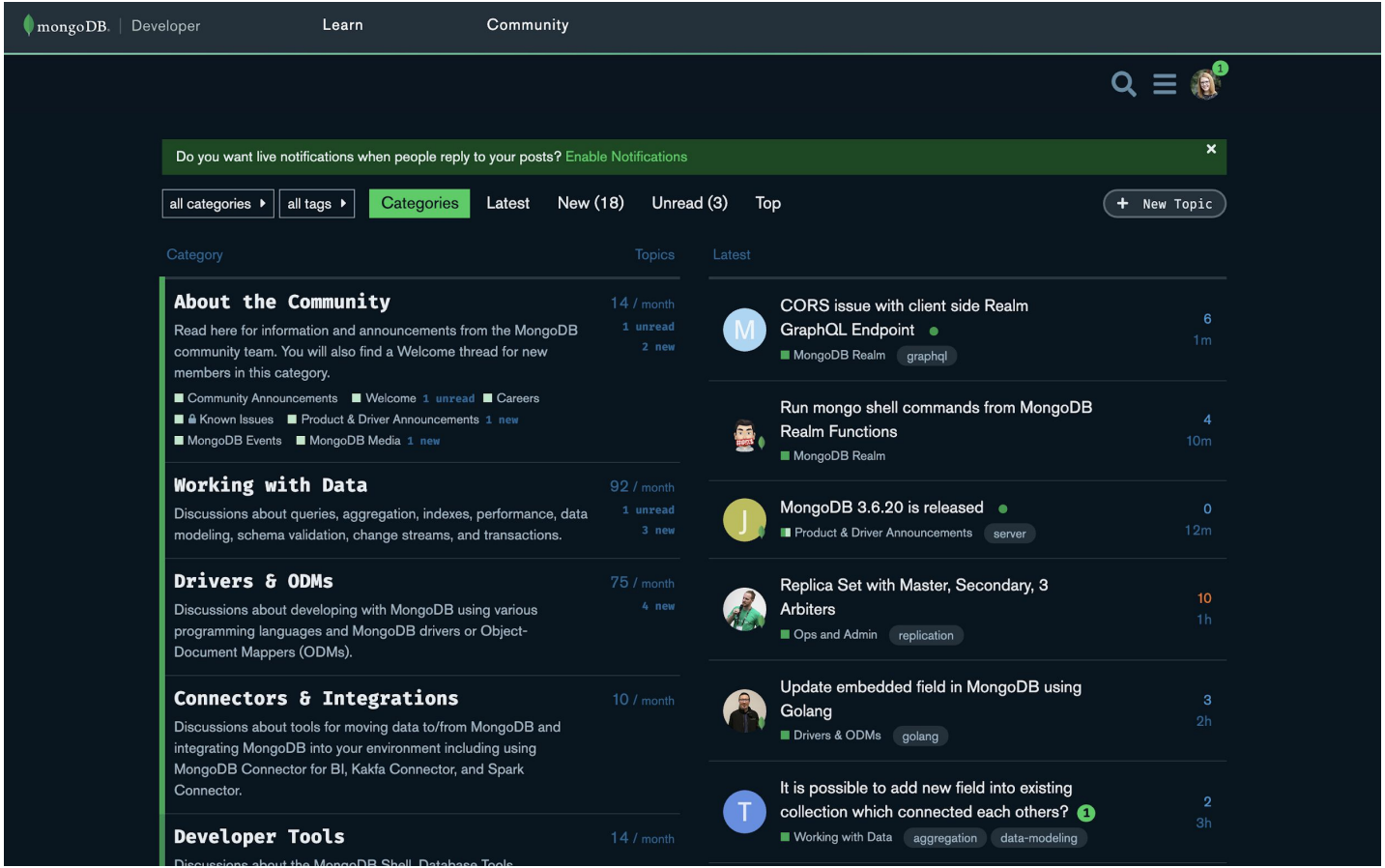
Course Details

Questions to Ask as You Model Your Data

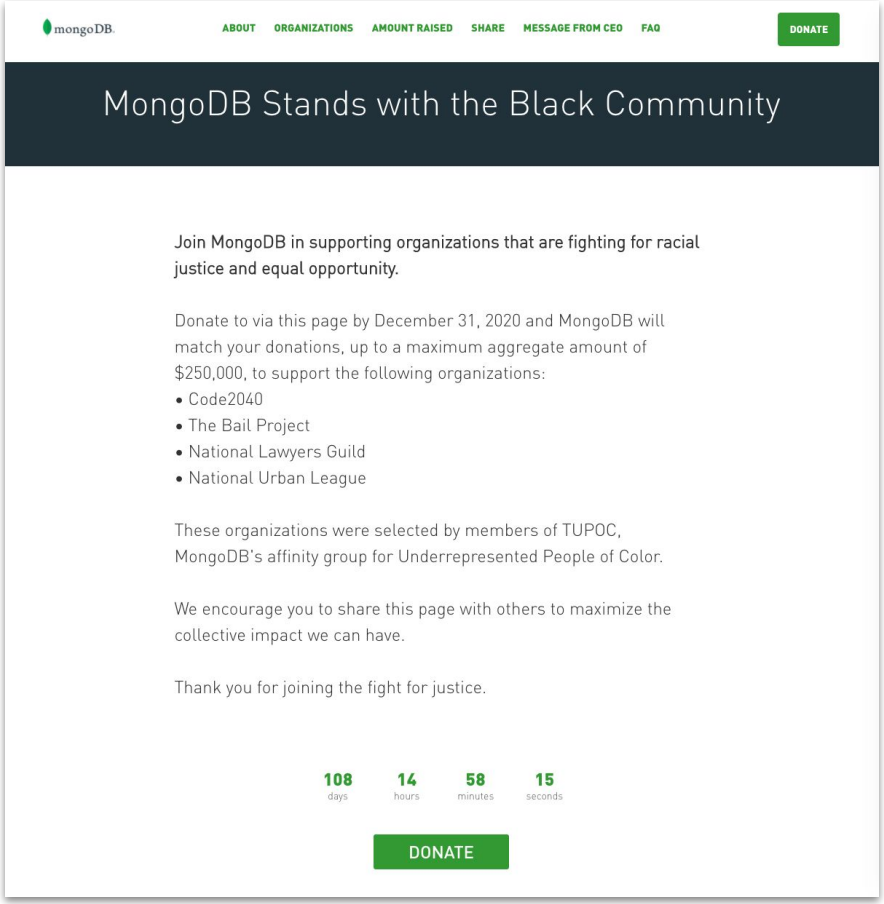
- What data will you need to store?
- What data is likely to be accessed together?
- What queries will be run most frequently?
- What data is likely to grow at a rapid, unbounded pace?

There is no "right" model for your data.


Do it. Fierce. Power.



community.mongodb.com



mongodbforjustice.mongodb.events