

**Find
clarity in
a world of
noise**

**Courting MongoDB,
committing to
PostgreSQL: Our
data love story!!**



About me: Valeria Alexandra Haro Valle



- From Ecuador
- Senior Data Engineer at Datamaran
- Expertise in managing databases, ETL, pipelines, Data Warehouse, combined with DevOps expertise, automate operations and supporting tech team
- Enjoys dancing and handcrafts



Stages of Love



Falling in Love



Beginning of the Relationship



Disillusionment



True Love

Datamaran requirements



Easy to use



Scalability &
high availability



Flexible schema



Licensing and
cost-effectiveness

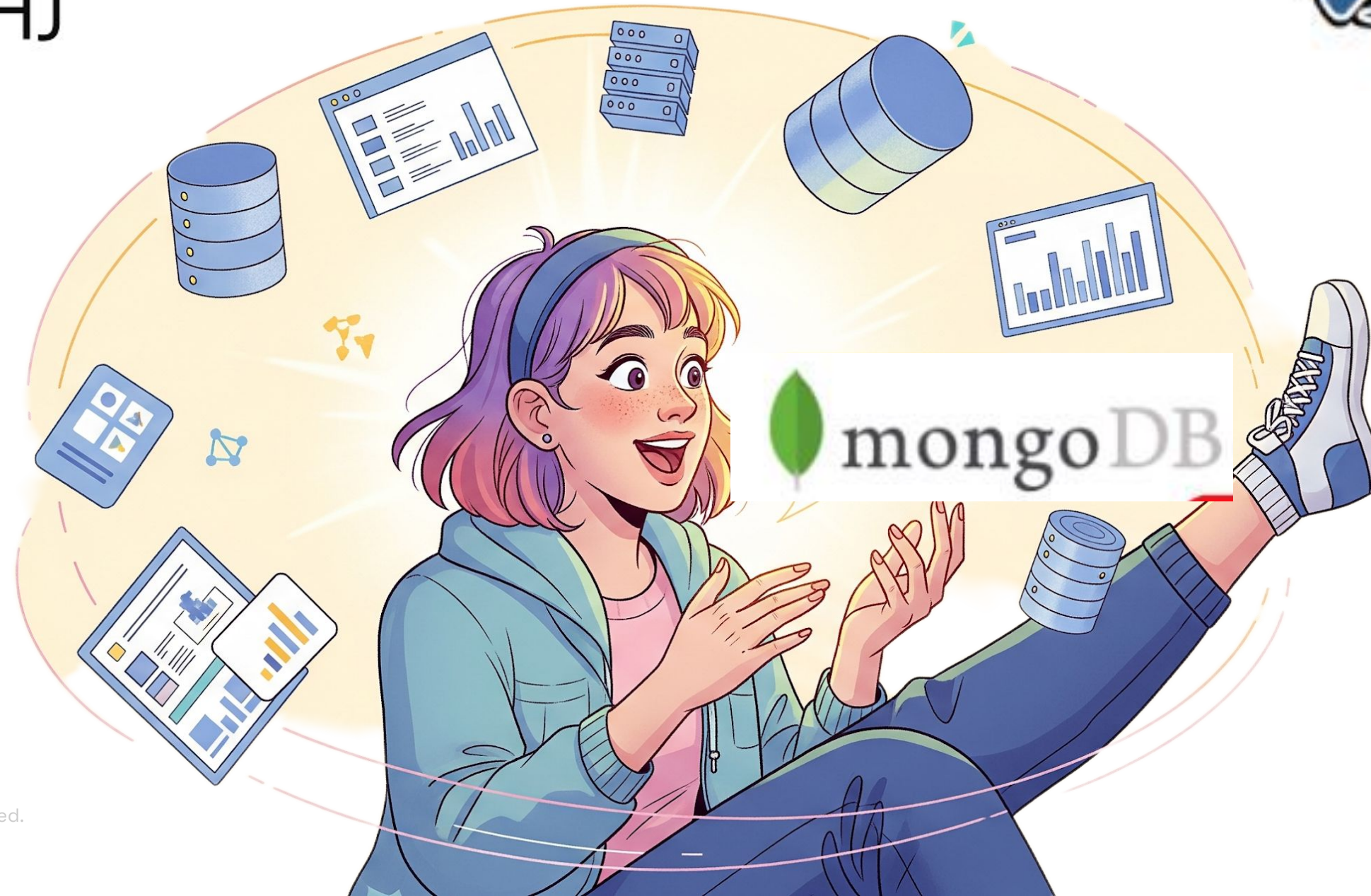


Faster
prototyping

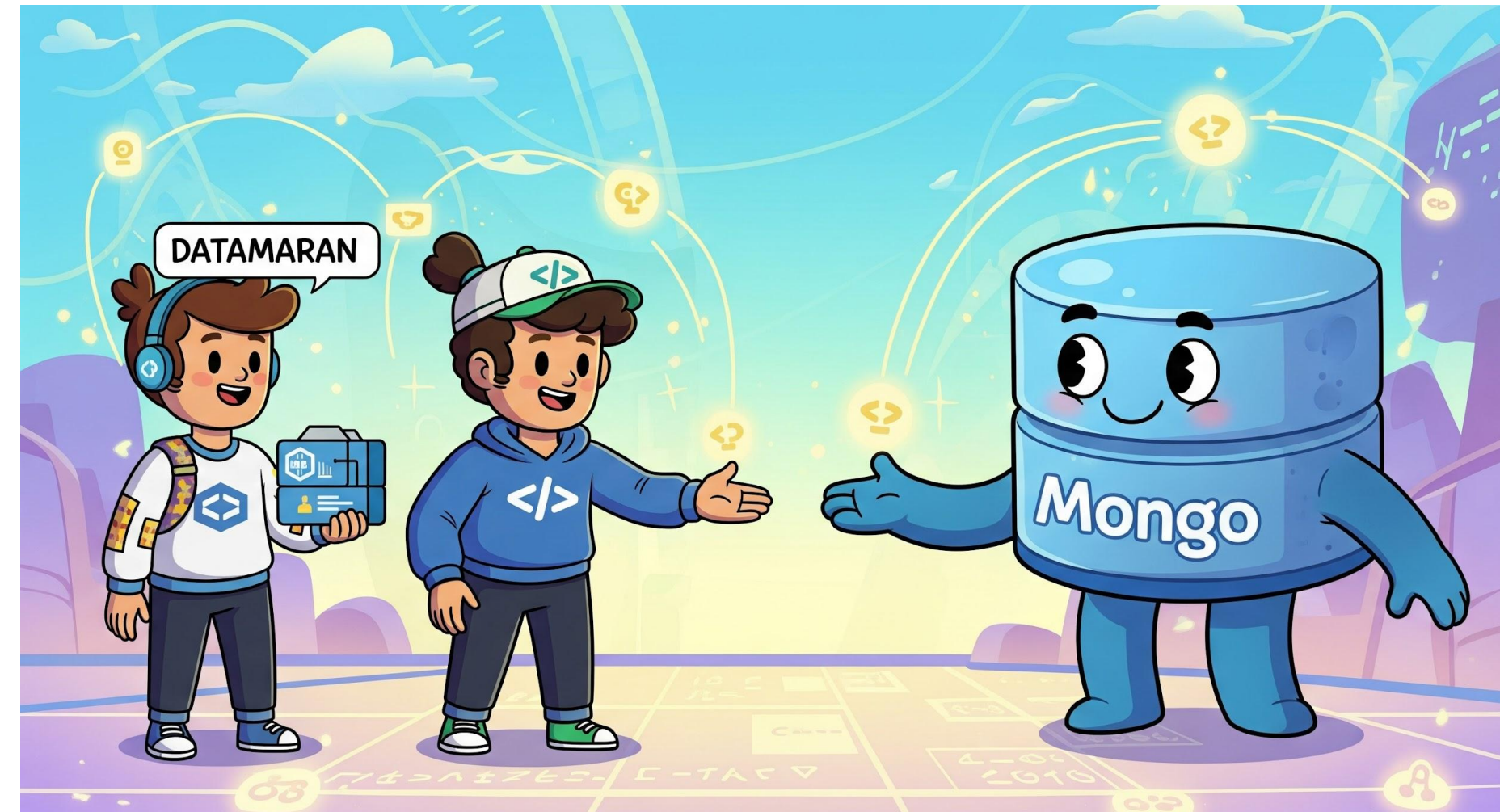


Cloud-managed

Choosing a Database



Beginning of relationship



Key feature we appreciate

Flexible schema

- Organization

```
{
  "_id": ObjectId("..."),
  "name": "APPLE INC",
  "CIK": 320193,
  "country": {
    "name": "United States",
    "cca3": "USA"
  },
  "industry": "Software and Services"
}
```

- NGO

```
{
  "_id": ObjectId("..."),
  "name": "Climate Action Network Europe",
  "country": {
    "name": "Belgium",
    "cca3": "BEL"
  }
}
```

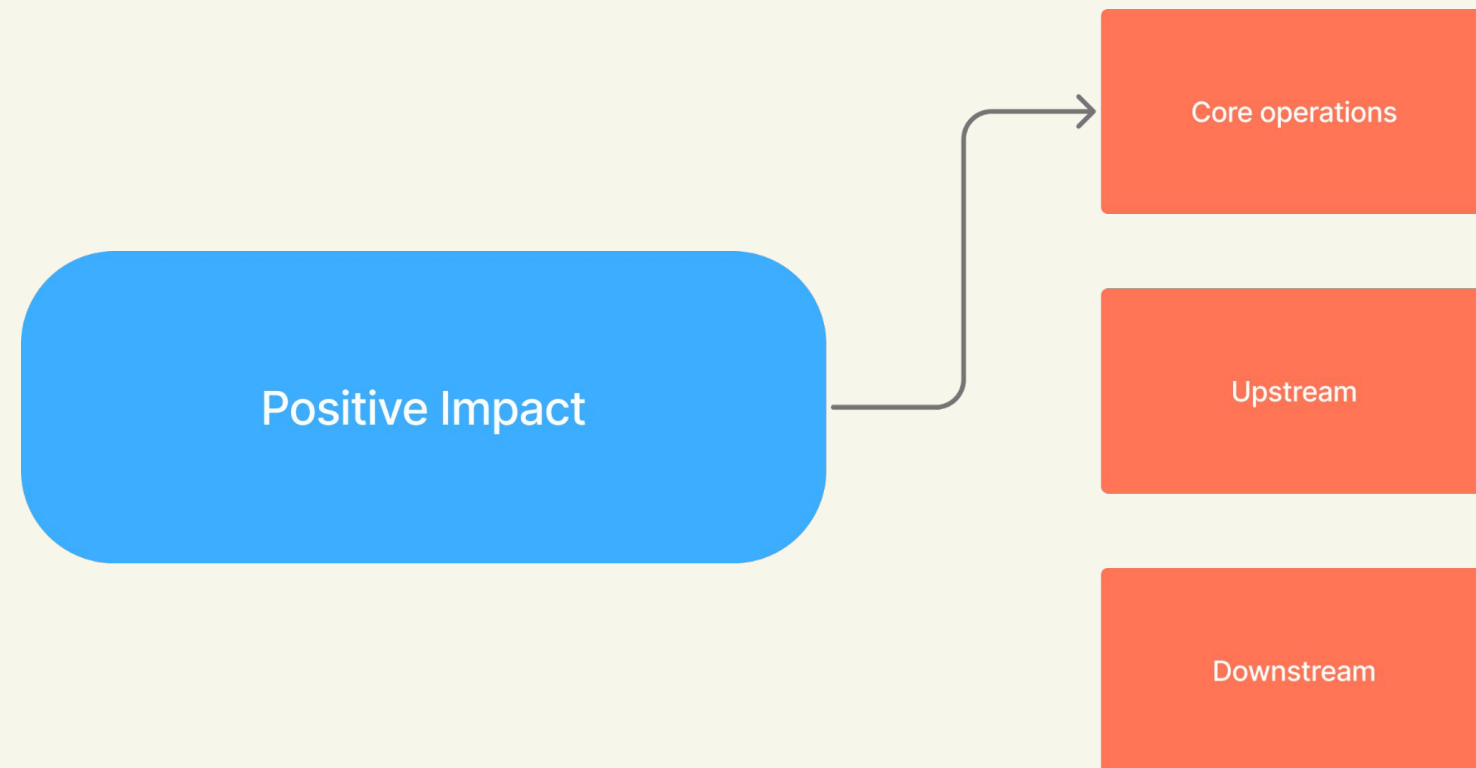
Quickly develop prototypes

Collections map naturally to your web app's data structures

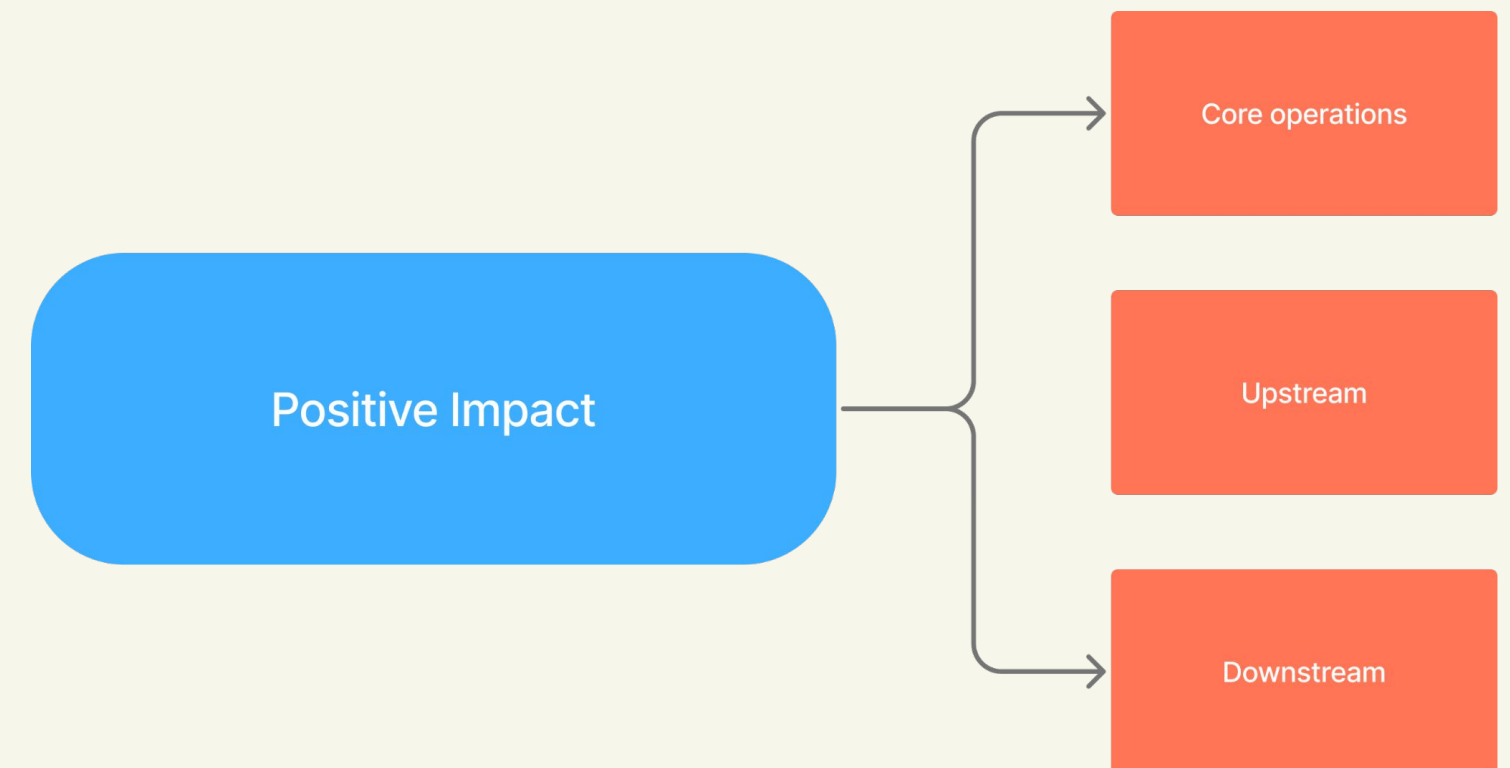
```
_id: ObjectId('660d5877698bab5d4cfb1676')
error: null
▸ inputs: Object
▸ issue_mapping: Object
▼ matrix_data: Array (91)
  ▾ 0: Object
    issue_id: 1
    issue_name: "Non-hazardous waste management"
    norm_score_x: 0.484
    norm_score_y: 0.749
    rank_x: 68
    rank_y: 19
    dm_recommended: false
  ▸ 1: Object
  ▸ 2: Object
  ▸ 3: Object
  ▸ 4: Object
  ▸ 5: Object
  ▸ 6: Object
  ▸ 7: Object
  ▸ 8: Object
```

Key Features We Appreciate

- Embedded documents allow storing related data together
- Freedom to easily update the data model as needed



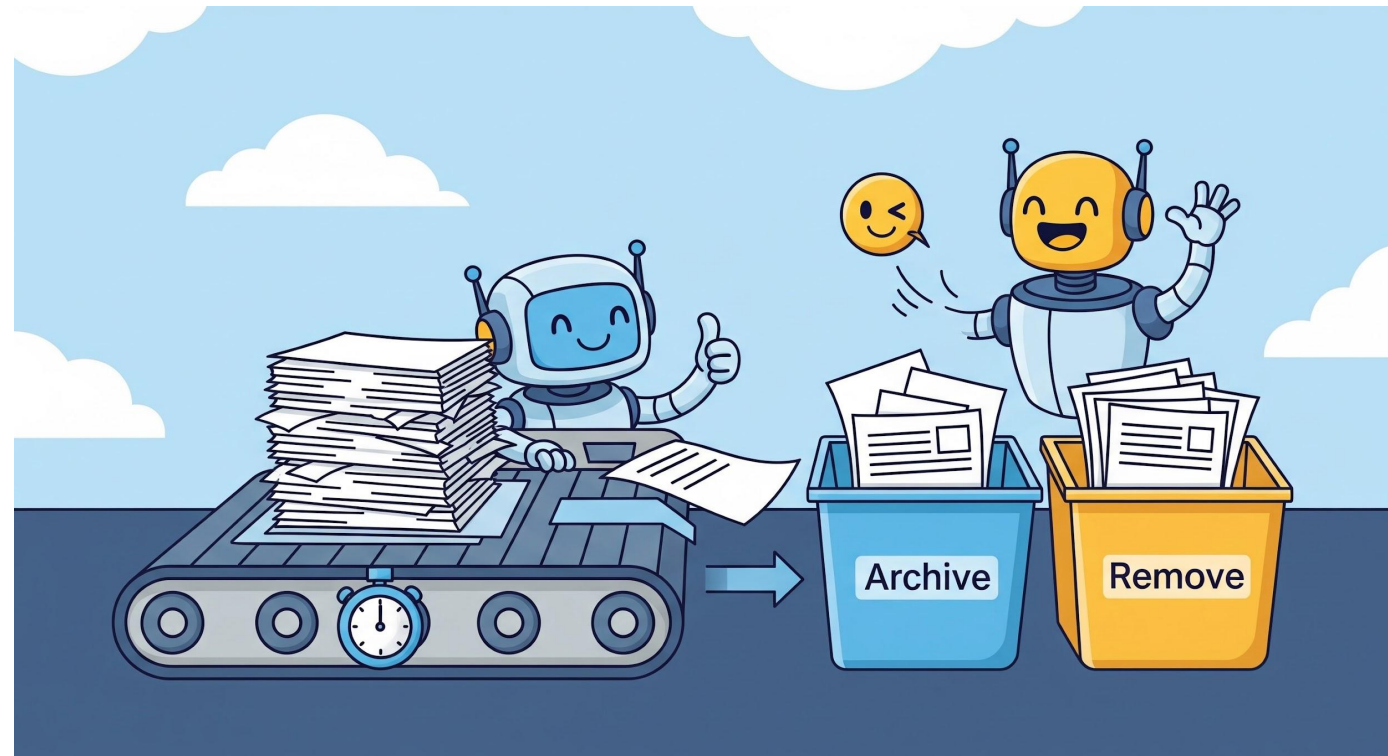
One - to - one



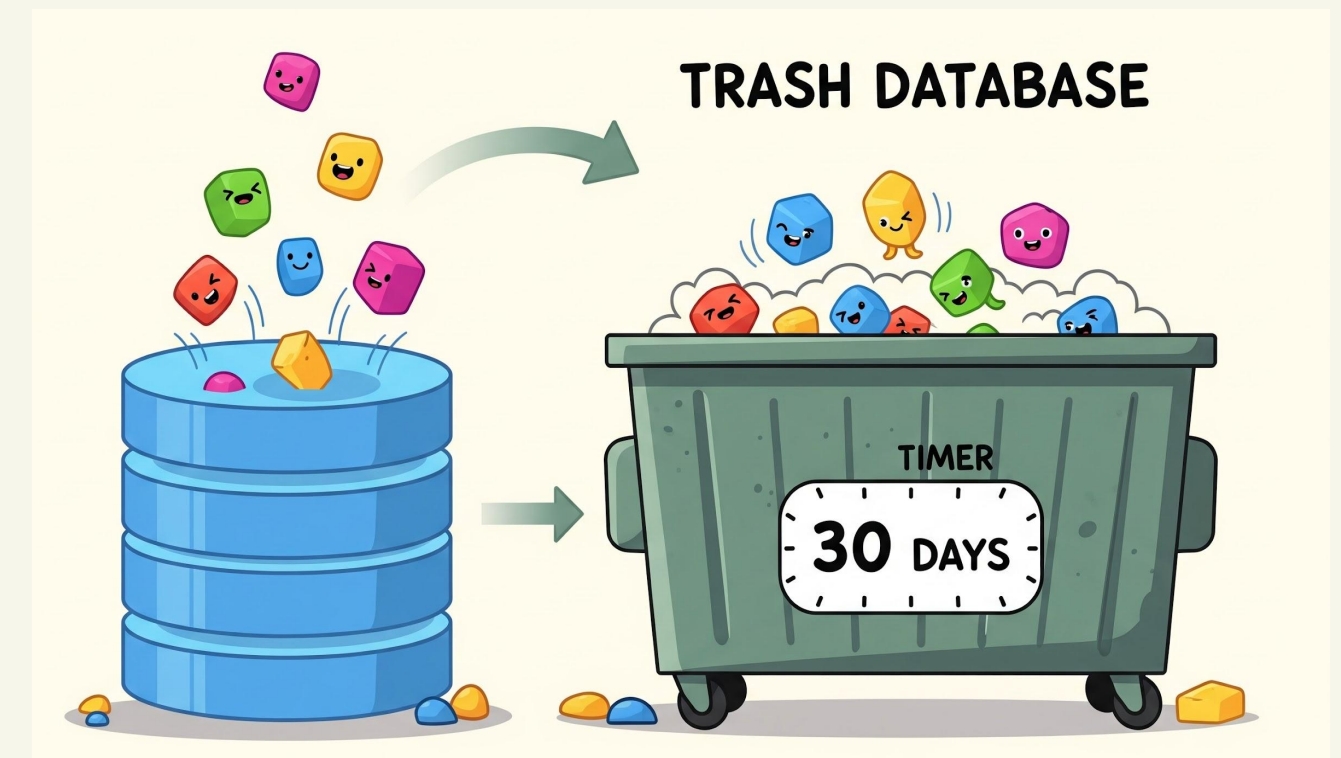
One - to - many

Key Features We Appreciate

- TTL index
 - Automatically remove documents from a collection after a certain amount of time

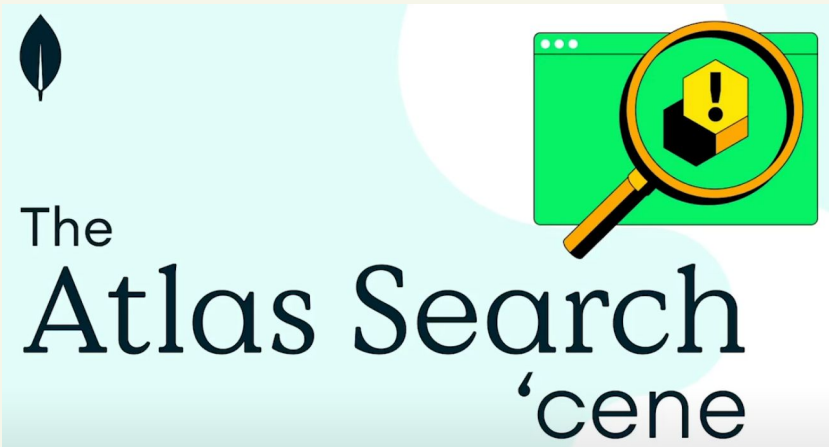


- Recovery process (soft-delete)



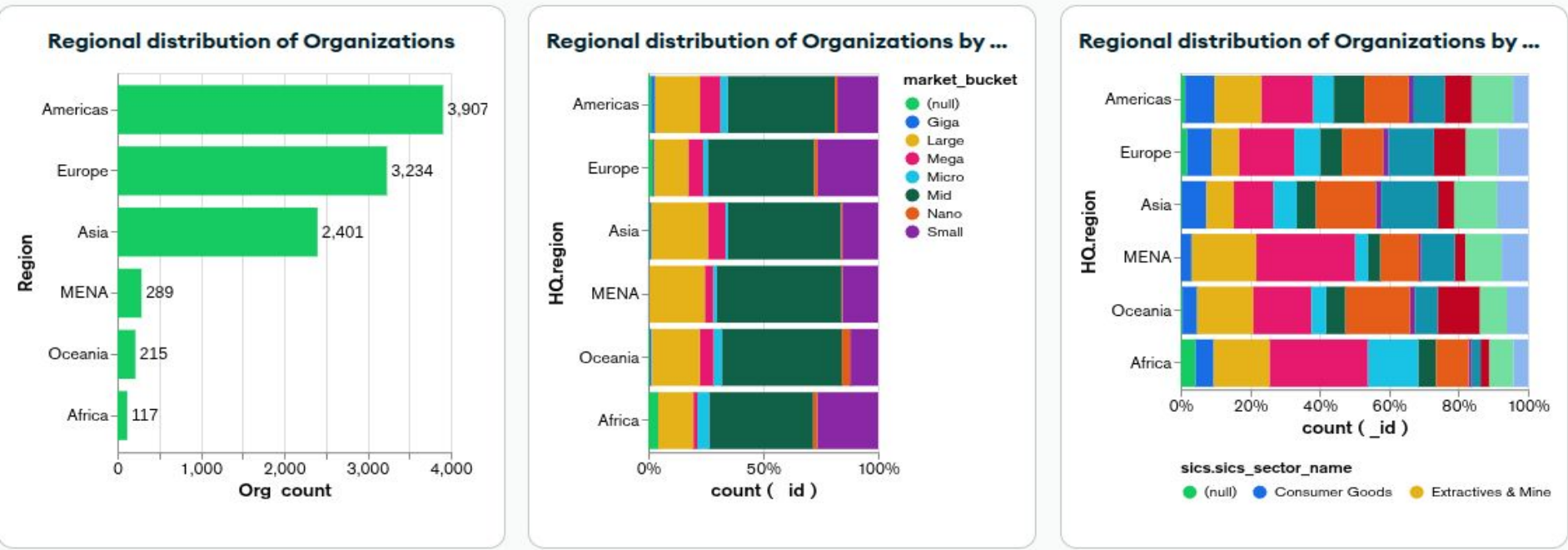
Key Features We Appreciate - Mongo Atlas

- MongoDB is up-to-date with market needs



- Atlas charts

← Industry_Dashboard



Stage 3: Disillusionment



Data anomalies



- Values with different data types

1 "1"

- Missing values(required not force)

country field required

- Non standard values

org_type: "corp"

org_type: "Corporation"

- Non standards fields names

org_name

name

organization_name



Data Relationship

- Embedding

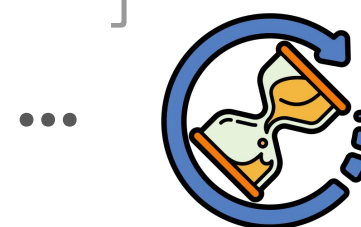
- Lack of proper handling during updates or deletes

Ej: Facebook, Inc is updated to Meta Platforms, Inc

```
{
  "_id": ObjectId("..."),
  "name": "Facebook, Inc",
  "name": "Meta Platforms, Inc"
  "CIK": 1326801,
  "country": {
    "name": "United States",
    "cca3": "USA"
  },
  "industry": "Technology"
}
```



```
{
  "_id": ObjectId("..."),
  "analysis name": "2004 analysis"
  "orgs": [
    {"Name": "APPLE INC", "cca3": "USA"},
    {"Name": "Facebook, Inc ", "cca3": "USA"}
  ],
}
```



```
{
  "_id": ObjectId("..."),
  "Analysis name": "2025 analysis"
  "orgs": [
    {"Name": "APPLE INC", "cca3": "USA"},
    {"Name": "Meta Platforms, Inc", "cca3": "USA"}
  ],
}
```

Data model out of control

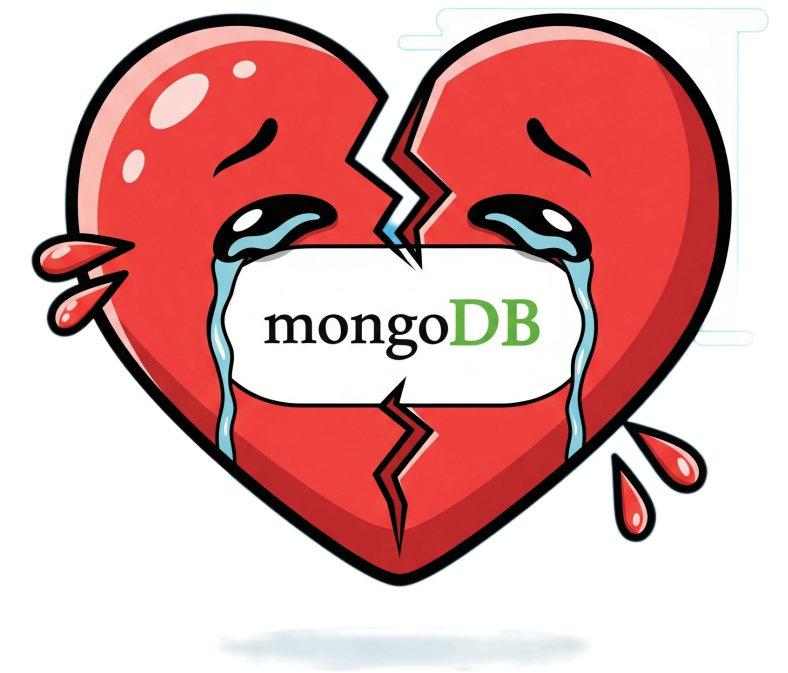
- Nested objects and arrays out of control



```
▼ stakeholders : Array (13)
  ▼ 0: Object
    stakeholder_name : "News Negative Sentiment (- Impact)"
    stakeholder_type : "newsflow_topics"
    axis : "y"
    weight : 12.5
  ▼ inputs : Object
    sources : null
  ▼ industries : Array (1)
    ▼ 0: Object
      id : "TC-HW"
      industry_cd : "TC-HW"
      sector_cd : "TC"
      sector_name : "Technology & Communications"
      industry_name : "Hardware"
    countries : null
  ▼ peers : Array (3)
    ▼ 0: Object
      id : 41
      country_name : "United States"
      org_name : "AMAZON.COM INC"
    ▶ 1: Object
    ▶ 2: Object
    sentiment : "negative"
  ▼ filter_topics : Array (1)
    ▼ 0: Object
      id : "X822"
      topic_name : "Negative impacts"
      topic_cd : "X822"
```

MongoDB Deployment Challenges (Atlas)

- No support for local deployments for key features like:
 - Atlas Search
 - Triggers
 - Vector Search
- Dependency on shared environments, often leading to:
 - Conflicts between teams
 - Harder isolation for testing
- Complex data sync required from production to shared developer environment
- Max 100 users per project



Datamaran requirements



Easy to use



Test environments



Local environments



Scalability &
high availability

~~Flexible schema~~
Data consistency



Encryption



Compatible with
vector

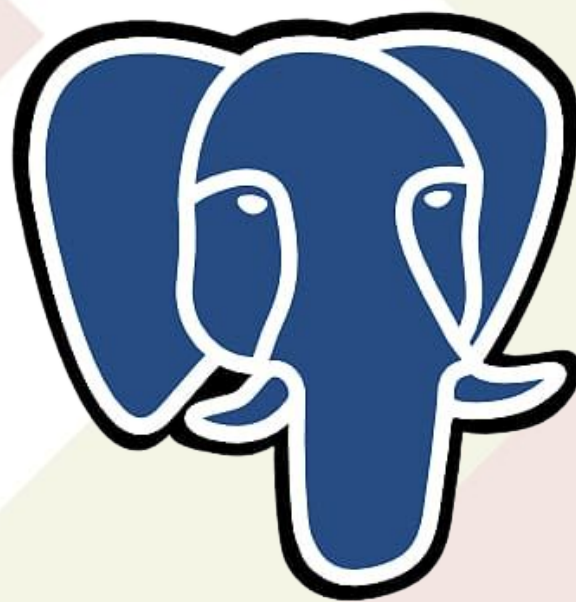


Licensing and
cost-effectiveness



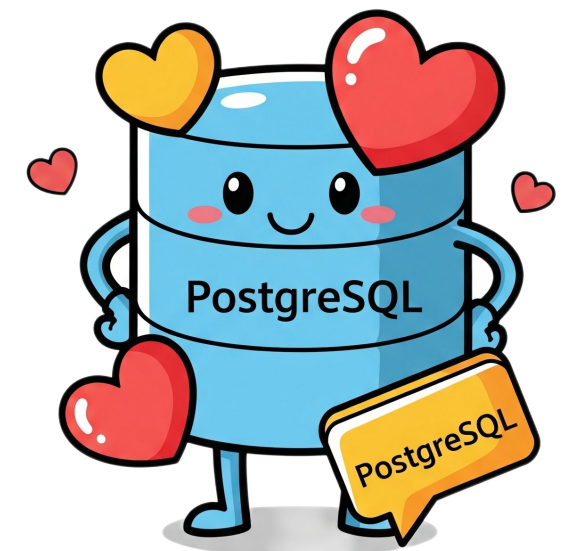
Cloud-managed

POSTGRESQL: The Next Chapter



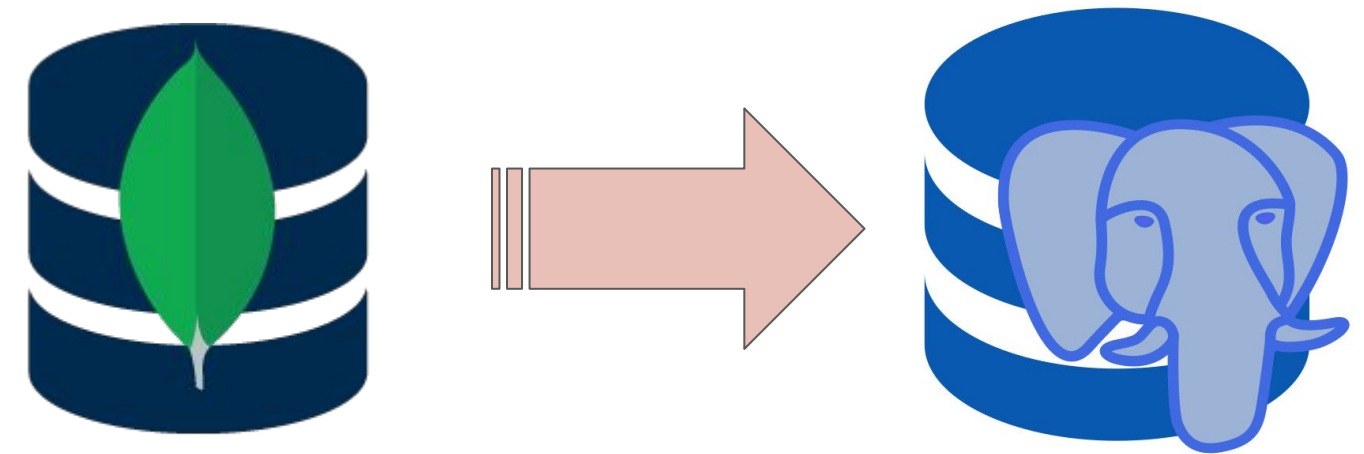
PostgreSQL

- Database templates -> Local development & testing
 - Local environments easy to deploy with docker
 - Integration with CI pipelines for automated testing
- Pre-Production with replication
- Powerful extensions included
 - pgcrypto
 - pgvector
 - pg_cron



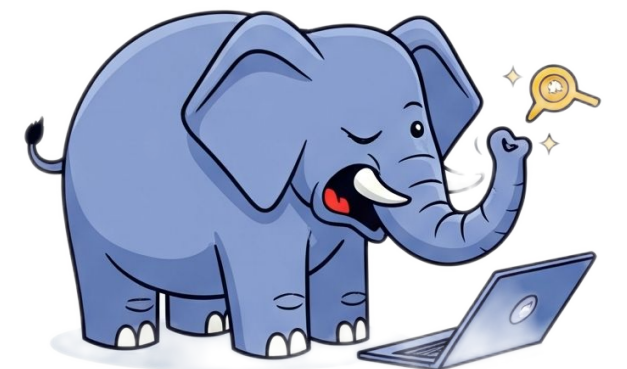
Challenges moving from MongoDB to PostgreSQL

- Applications refactoring required
 - Lack of data layer
- Applications are utilizing both databases
- No migration tools
 - Adhoc python scripts
- Deprecated soft-delete
- Adapting developer practices for SQL environments



Issues encountered in PostgreSQL

- Permission management per database
- Connection handling in application side
- Query performance on large tables
- Schemas changes with replication
 - Risk of breaking replication with certain schema updates(WAL)
- Lack of enforced constraints
 - Inconsistent use of FK and delete constraint



Lessons Learned

1. Flexible schema != schemaless
2. Not following best practices can break even the most robust systems



Lessons Learned

3. Use the proper database according your use case

- MongoDB (Application focus on domain) for prototyping, small applications, micro-services, IoT
- PostgreSQL (Centralize data) for prioritizing data integrity, and supporting system-agnostic use cases, better for solid products



London, HQ



Valencia



New York City



Leeuwarden



info@datamaran.com