DATA AND MIGRATION

WHAT IS A DOCUMENT DATABASE?

JSON documents

Hierarchical
key-value pairs

WHAT IS A DOCUMENT DATABASE?

JSON documents

Hierarchical
key-value pairs

Two API choices

1) SQL API
(aka core API)

2) MongoDB API

RELATIONAL VS. DOCUMENT

Relational Database	Document Database

RELATIONAL VS. DOCUMENT

Relational Database	Document Database
Rows	Documents

Table

Row 1	Data
Row 2	Data
Row 3	Data
...	:

Container

Document 1
```
{
  data
}
```

Document 2
```
{
  data
}
```

Document 3
```
{
  data
}
```
...

RELATIONAL VS. DOCUMENT

Relational Database	Document Database
Rows	Documents
Columns	Properties

Col1	Col2	Col3	Col4
Data	Data	Data	Data
Data	Data	Data	Data
Data	Data	Data	Data

Document 1
```
{
  "prop1": data,
  "prop2": data,
  "prop3": data,
  "prop4": data
}
```

Document 2
```
{
  "prop1": data,
  "prop2": data,
  "prop3": data,
  "prop4": data
}
```

Document 3
```
{
  "prop1": data,
  "prop2": data,
  "prop3": data,
  "prop4": data
}
```

RELATIONAL VS.DOCUMENT

Relational Database	Document Database
Rows	Documents
Columns	Properties
Strongly typed schemas	No defined schema

ID	Name	IsActive	Dob
1	Dan Smith	True	8/30/1964
2	Sue Jones	False	2/18/2002
3	Adam Stark	True	7/13/1987

Document 1

```
{
  "id": 1,
  "name": "Dan Smith",
  "isActive": true,
  "dob": "1964-30-08"
}
```

RELATIONAL VS. DOCUMENT

Relational Database	Document Database
Rows	Documents
Columns	Properties
Strongly typed schemas	No defined schema

ID	Name	IsActive	Dob
1	Dan Smith	True	8/30/1964
2	Sue Jones	False	2/18/2002
3	Adam Stark	True	7/13/1987

Document 1
```
{
  "id": 1,
  "name": "Dan Smith",
  "isActive": true,
  "dob": "1964-30-08"
}
```

Document 2
```
{
  "id": 2,
  "fullName": "Sue Jones",
  "dob": "2002-02-18"
}
```

Document 3
```
{
  "id": 3,
  "fullName":
  {
    "first": "Adam",
    "last": "Stark"
  },
  "isActive": true,
  "dob": "2015-04-19"
}
```

RELATIONAL VS. DOCUMENT

Relational Database	Document Database
Rows	Documents
Columns	Properties
Strongly typed schemas	No defined schema

ID	Name	IsActive	Dob
1	Dan Smith	True	8/30/1964
2	Sue Jones	False	2/18/2002
3	Adam Stark	True	7/13/1987

Document 1
```
{
  "id": 1,
  "name": "Dan Smith",
  "isActive": true,
  "dob": "1964-30-08",
  "type": "user"
}
```

Document 2
```
{
  "id": 2,
  "fullName": "Sue Jones",
  "dob": "2002-02-18",
  "type": "user"
}
```

Document 3
```
{
  "id": 3,
  "fullName":
  {
    "first": "Adam",
    "last": "Stark"
  },
  "isActive": true,
  "dob": "2015-04-19",
  "type": "user"
}
```

RELATIONAL VS. DOCUMENT

Relational Database	Document Database
Rows	Documents
Columns	Properties
Strongly typed schemas	No defined schema

ID	Name	IsActive	Dob
1	Dan Smith	True	8/30/1964
2	Sue Jones	False	2/18/2002
3	Adam Stark	True	7/13/1987

Document 1
```
{
  "id": 1,
  "name": "Dan Smith",
  "isActive": true,
  "dob": "1964-30-08",
  "type": "user",
  "version": 1
}
```

Document 2
```
{
  "id": 2,
  "fullName": "Sue Jones",
  "dob": "2002-02-18",
  "type": "user",
  "version": 2
}
```

Document 3
```
{
  "id": 3,
  "fullName":
  {
    "first": "Adam",
    "last": "Stark"
  },
  "isActive": true,
  "dob": "2015-04-19",
  "type": "user",
  "version": 3
}
```

DENORMALIZING THE MODEL

User Table

UserID	Name	Dob
1	Dan Smith	8/30/1964

Holdings Table

StockID	UserID	Qty	Symbol
1	1	100	MSFT
2	1	75	WMT

Document

```
{
  "id": 1,
  "name": "Dan Smith",
  "dob": "1964-30-08",
  "holdings": [
    {"qty": 100, "symbol": "MSFT"},
    {"qty": 75, "symbol": "WMT"}
  ]
}
```

DENORMALIZING THE MODEL

Document
```
{
  "postid": "1",
  "title": "My blog post",
  "body": "Post content...",
  "comments": [
      "comment #1",
      "comment #2",
      "comment #3",
      "comment #4",
           :
      "comment #1598873",
           :
```

Document
```
{
  "postid": "1",
  "title": "My blog post",
  "body": "Post content..."
}
```

Document
Document
Document
```
{
  "postid": "1",
  "comment": "comment #3"
}
```

Document
```
{
  "postid": "1",
  "title": "My blog post",
  "body": "Post content...",
  "comments": [
      "comment #1",
      "comment #2",
           :
      "comment #100"
  ]
}
```

Document
Document
Document
```
{
  "postid": "1",
  "comments": [
      "comment #301",
      "comment #302",
           :
      "comment #400"
  ]
}
```

DENORMALIZING THE MODEL

Document

```
{
  "id": "1",
  "prop1": "Typically  read",
  "prop2": "Typically  read",
  "prop3": "Typically  read",
  "prop4": "Typically  read",
  "prop5": "Typically  read",
  "prop6": "Typically  read",
  "prop7": "Typically  read",
  "prop8": "Typically  read",
  "prop9": "Typically  read",
  "prop10": "Typically read",
     :
  "prop200": "Typically read",
  "prop201": "Typically updated",
  "prop202": "Typically updated",
  "prop203": "Typically updated",
  "prop204": "Typically updated",
  "prop205": "Typically updated"
}
```

Document

```
{
  "id": "1-1",
  "prop1": "Typically  read",
  "prop2": "Typically  read",
  "prop3": "Typically  read",
  "prop4": "Typically  read",
  "prop5": "Typically  read",
  "prop6": "Typically  read",
  "prop7": "Typically  read",
  "prop8": "Typically  read",
  "prop9": "Typically  read",
  "prop10": "Typically read",
     :
  "prop200": "Typically read"
}
```

Document

```
{
  "id": "1-2",
  "prop201": "Typically updated",
  "prop202": "Typically updated",
  "prop203": "Typically updated",
  "prop204": "Typically updated",
  "prop205": "Typically updated"
}
```
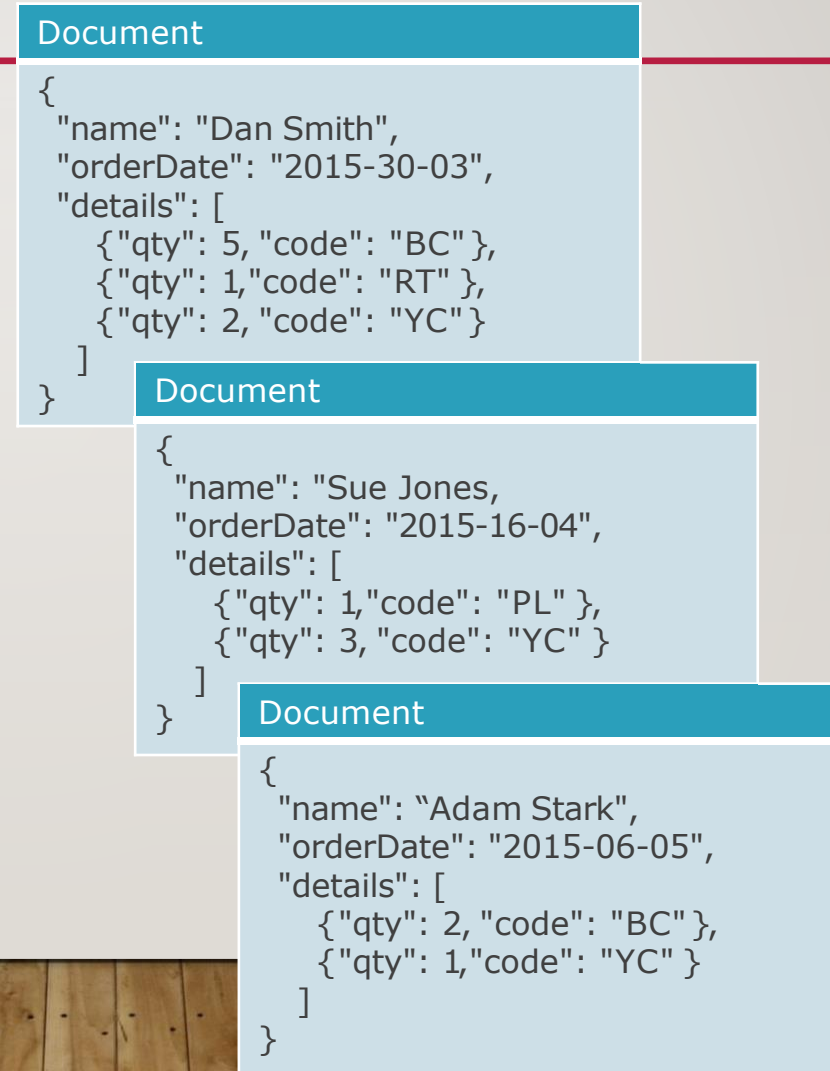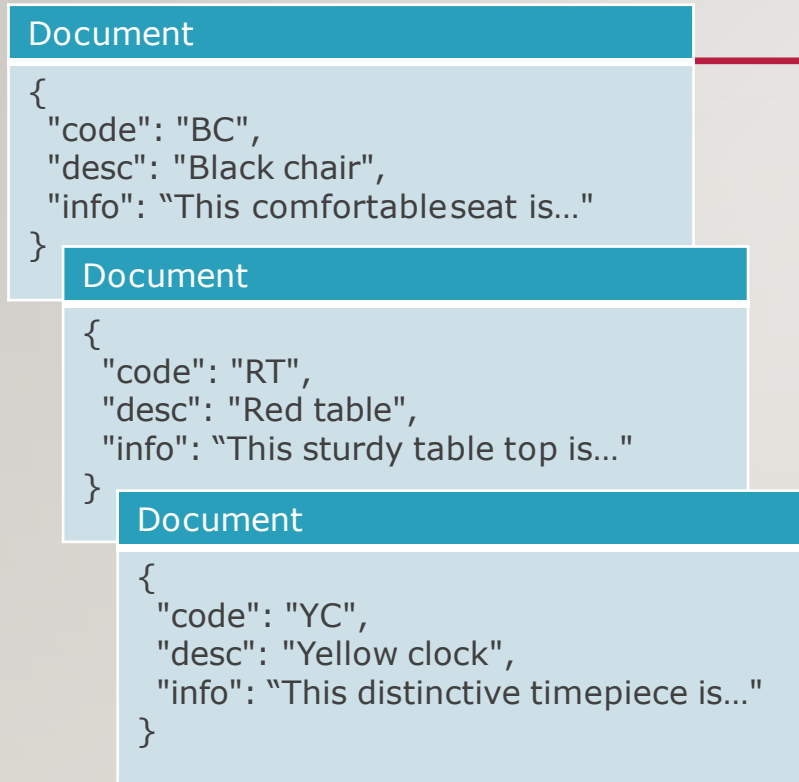
DENORMALIZING THE MODEL

Document

```
{
  "code": "BC",
  "desc": "Black chair",
  "info": "This comfortable seat is…"
}
```

Document

```
{
  "code": "RT",
  "desc": "Red table",
  "info": "This sturdy table top is…"
}
```

Document

```
{
  "code": "YC",
  "desc": "Yellow clock",
  "info": "This distinctive timepiece is…"
}
```

Document

```
{
  "name": "Dan Smith",
  "orderDate": "2015-30-03",
  "details": [
      {"qty": 5, "code": "BC"},
      {"qty": 1,"code": "RT" },
      {"qty": 2, "code": "YC"}
  ]
}
```

Document

```
{
  "name": "Sue Jones,
  "orderDate": "2015-16-04",
  "details": [
      {"qty": 1,"code": "PL" },
      {"qty": 3, "code": "YC" }
  ]
}
```

Document

```
{
  "name": "Adam Stark",
  "orderDate": "2015-06-05",
  "details": [
      {"qty": 2, "code": "BC"},
      {"qty": 1,"code": "YC" }
  ]
}
```

DENORMALIZING THE MODEL

Document

```
{
  "code": "BC",
  "desc": "Black chair",
  "info": "This comfortable seat is..."
}
```

Document

```
{
  "code": "RT",
  "desc": "Red table",
  "info": "This sturdy table top is..."
}
```

Document

```
{
  "code": "YC",
  "desc": "Yellow clock",
  "info": "This distinctive timepiece is..."
}
```

Document

```
{
  "name": "Dan Smith",
  "orderDate": "2015-30-03",
  "details": [
    {"qty": 5, "code": "BC", "desc": "Black chair" },
    {"qty": 1, "code": "RT", "desc": "Red table" },
    {"qty": 2, "code": "YC", "desc": "Yellow clock" }
  ]
}
```

Document

```
{
  "name": "Sue Jones,
  "orderDate": "2015-16-04",
  "details": [
    {"qty": 1, "code": "PL", "desc": "Purple lamp" },
    {"qty": 3, "code": "YC", "desc": "Yellow clock" }
  ]
}
```

Document

```
{
  "name": "Adam Stark",
  "orderDate": "2015-06-05",
  "details": [
    {"qty": 2, "code": "BC", "desc": "Black chair" },
    {"qty": 1, "code": "YC", "desc": "Yellow clock" }
  ]
}
```

Change Feed

DATA MIGRATIONTOOL

Open source project

Microsoft Download Center
http://www.microsoft.com/en-us/download/details.aspx?id=46436

Documentation and How-To
https://docs.microsoft.com/en-us/azure/cosmos-db/import-data

Source code on GitHub
https://github.com/azure/azure-documentdb-datamigrationtool

Import from
- SQL Server
- JSON
- CSV
- MongoDB
- Azure Table Storage
- and more…

Demo

Importing data from SQL Server

Source Information

Specify source information

Import from:

SQL

Connection String

data source=.;initial catalog=AdventureWorks2017;integrated security=true;   [Verify]

◉ Enter Query   ○ Select Query File

```
SELECT
 Name    AS [name],
 AddressType   AS [address.addressType],
 AddressLine1  AS [address.addressLine1],
 City    AS [address.location.city],
 StateProvinceName AS [address.location.stateProvinceName],
 PostalCode   AS [address.postalCode],
 CountryRegionName AS [address.countryRegionName]
 FROM
 Sales.vStoreWithAddresses
 WHERE
 AddressType = 'Main Office'
```

Nesting Separator ⑦

.

Navigation (left panel):
- Welcome
- Source Information
- Target Information ⚠
- Advanced
- Summary
- Results

[Previous] [Next]

Results

Import results

Elapsed time: 0:00:07.3

Transferred: 701

Failed: 0

Failure Information

Export ▼

Record	Error

New Import   Cancel

SPECIAL DOCUMENT PROPERTIES

```json
{
    "id": "Sample",
    "familyName": "Jones",
    "address": {
        "addressLine": "789 Harbor Boulevard",
        "city": "Chicago",
        "state": "IL",
        "zipCode": "60603"
    },
    "_rid": "IwdPAIDPIS0BAAAAAAAAAA==",
    "_self": "dbs/IwdPAA==/colls/IwdPAIDPIS0=/docs/IwdPAIDPIS0BAAAAAAAAAA==/",
    "_etag": "\"09008ef9-0000-0700-0000-5d220a430000\"",
    "_attachments": "attachments/",
    "_ts": 1562511939
}
```

SPECIAL DOCUMENT PROPERTIES

Property	Value
id	User-defined unique ID
user-definable	Partition key

```json
{
    "id": "Sample",
    "familyName": "Jones",
    "address": {
        "addressLine": "789 Harbor Boulevard",
        "city": "Chicago",
        "state": "IL",
        "zipCode": "60603"
    },
    "_rid": "IwdPAIDPIS0BAAAAAAAAAA==",
    "_self": "dbs/IwdPAA==/colls/IwdPAIDPIS0=/docs/IwdPAIDPIS0BAAAAAAAAAA==/",
    "_etag": "\"09008ef9-0000-0700-0000-5d220a430000\"",
    "_attachments": "attachments/",
    "_ts": 1562511939
}
```

SPECIAL DOCUMENT PROPERTIES

Property	Value
id	User-defined unique ID
user-definable	Partition key
_rid	Resource ID

```json
{
    "id": "Sample",
    "familyName": "Jones",
    "address": {
        "addressLine": "789 Harbor Boulevard",
        "city": "Chicago",
        "state": "IL",
        "zipCode": "60603"
    },
    "_rid": "IwdPAIDPIS0BAAAAAAAAA==",
    "_self": "dbs/IwdPAA==/colls/IwdPAIDPIS0=/docs/IwdPAIDPIS0BAAAAAAAAA==/",
    "_etag": "\"09008ef9-0000-0700-0000-5d220a430000\"",
    "_attachments": "attachments/",
    "_ts": 1562511939
}
```

SPECIAL DOCUMENT PROPERTIES

Property	Value
id	User-defined unique ID
user-definable	Partition key
_rid	Resource ID
_self	URI path to the resource

```json
{
    "id": "Sample",
    "familyName": "Jones",
    "address": {
        "addressLine": "789 Harbor Boulevard",
        "city": "Chicago",
        "state": "IL",
        "zipCode": "60603"
    },
    "_rid": "IwdPAIDPIS0BAAAAAAAAA==",
    "_self": "dbs/IwdPAA==/colls/IwdPAIDPIS0=/docs/IwdPAIDPIS0BAAAAAAAAA==/",
    "_etag": "\"09008ef9-0000-0700-0000-5d220a430000\"",
    "_attachments": "attachments/",
    "_ts": 1562511939
}
```

SPECIAL DOCUMENT PROPERTIES

Property	Value
id	User-defined unique ID
user-definable	Partition key
_rid	Resource ID
_self	URI path to the resource
_etag	GUID (optimistic concurrency)

```
{
    "id": "Sample",
    "familyName": "Jones",
    "address": {
        "addressLine": "789 Harbor Boulev
        "city": "Chicago",
        "state": "IL",
        "zipCode": "60603"
    },
    "_rid": "IwdPAIDPIS0BAAAAAAAAA==",
    "_self": "dbs/IwdPAA==/colls/IwdPAIDPIS0=/docs/IwdPAIDPIS0BAAAAAAAAA==/",
    "_etag": "\"09008ef9-0000-0700-0000-5d220a430000\"",
    "_attachments": "attachments/",
    "_ts": 1562511939
}
```

SPECIAL DOCUMENT PROPERTIES

Property	Value	
id	User-defined unique ID	
user-definable	Partition key	
_rid	Resource ID	
_self	URI path to the	resource
_etag	GUID (optimist	ic concurrency)
_attachments	URI suffix to the	attachments

```
{
    "id": "Sample",
    "familyName": "Jones",
    "address": {
        "addressLine": "789 Harbor Bouleva
        "city": "Chicago",
        "state": "IL",
        "zipCode": "60603"
    },
    "_rid": "IwdPAIDPIS0BAAAAAAAAA==",
    "_self": "dbs/IwdPAA==/colls/IwdPAIDPIS0=/docs/IwdPAIDPIS0BAAAAAAAAA==/",
    "_etag": "\"09008ef9-0000-0700-0000-5d220a430000\"",
    "_attachments": "attachments/",
    "_ts": 1562511939
}
```

SPECIAL DOCUMENT PROPERTIES

Property	Value
id	User-defined unique ID
user-definable	Partition key
_rid	Resource ID
_self	URI path to the resource
_etag	GUID (optimistic concurrency)
_attachments	URI suffix to the attachments
_ts	Last updated timestamp (epoch)

```json
{
    "id": "Sample",
    "familyName": "Jones",
    "address": {
        "addressLine": "789 Harbor Bouleva
        "city": "Chicago",
        "state": "IL",
        "zipCode": "60603"
    },
    "_rid": "IwdPAIDPIS0BAAAAAAAAAA==",
    "_self": "dbs/IwdPAA==/colls/IwdPAIDPIS0=/docs/IwdPAIDPIS0BAAAAAAAAAA==/",
    "_etag": "\"09008ef9-0000-0700-0000-5d220a430000\"",
    "_attachments": "attachments/",
    "_ts": 1562511939
}
```

SPECIAL DOCUMENT PROPERTIES

Property	Value
id	User-defined unique ID
user-definable	Partition key
_rid	Resource ID
_self	URI path to the resource
_etag	GUID (optimistic concurrency)
_attachments	URI suffix to the attachments
_ts	Last updated timestamp (epoch)
ttl	Time to Live (expiration)

```
{
    "id": "Sample",
    "familyName": "Jones",
    "address": {
        "addressLine": "789 Harbor Bouleva
        "city": "Chicago",
        "state": "IL",
        "zipCode": "60603"
    },
    "_rid": "IwdPAIDPIS0BAAAAAAAAAA==",
    "_self": "dbs/IwdPAA==/colls/IwdPAIDPIS0=/docs/IwdPAIDPIS0BAAAAAAAAAA==/",
    "_etag": "\"09008ef9-0000-0700-0000-5d220a430000\"",
    "_attachments": "attachments/",
    "_ts": 1562511939
}
```

https://cosmos.azure.com/?key=xPM6KkLVk7zOpHhhp3yt%2fSCgyquyRfUAonfzNhlEdMoYfiFO0...

Microsoft Azure | Cosmos DB > cosmos-west-us

Sign In

Save   Discard

SQL API

- Families
  - Families
    - Items
    - **Scale & Settings**
    - Stored Procedures
    - User Defined Functions
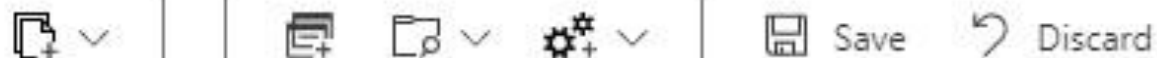   - Triggers
    - Conflicts

Scale & Settin... ✕

Settings

Time to Live

| Off | On (no default) | On |

0   0   0

Microsoft Azure | Cosmos DB > cosmos-west-us

Sign In

SQL API

Scale & Settin... ✕

⚠️ Changing the TTL or Indexing Policy impacts query results while the index transformation occurs. When a change is made and the indexing mode is set to consistent or lazy, queries return eventual results until the operation completes. For more information see, TTL and index interaction.

- Families
  - Families
    - Items
    - **Scale & Settings**
    - ▸ Stored Procedures
    - ▸ User Defined Functions
    - ▸ Triggers
    - Conflicts

Save   Discard

▾ Settings

Time to Live

| Off | On (no default) | On |

🔄 0   ✖️ 0   ℹ️ 0

https://cosmos.azure.com/?key=xPM6KkLVk7zOpHhhp3yt%2fSCgyquyRfUAonfzNhlEdMoYfiFO0...

Microsoft Azure | Cosmos DB > cosmos-west-us

Sign In

SQL API

Scale & Settin... ✕

- 🪐 Families
  - 📄 Families
    - Items
    - **Scale & Settings**
    - ▶ Stored Procedures
    - ▶ User Defined Functions
    - ▶ Triggers
    - Conflicts

⚠️ Changing the TTL or Indexing Policy impacts query results while the index transformation occurs. When a change is made and the indexing mode is set to consistent or lazy, queries return eventual results until the operation completes. For more information see, TTL and index interaction.

▼ Settings

Time to Live

| Off | On (no default) | **On** |

86400   second(s)

💾 Save   ↶ Discard

🔄 0   ⊗ 0   ⓘ 0

SPECIAL DOCUMENT PROPERTIES

Property	Value	
id	User-defined unique ID	
user-definable	Partition key	
_rid	Resource ID	
_self	URI path to the	resource
_etag	GUID (optimist	ic concurrency)
_attachments	URI suffix to the	attachments
_ts	Last updated t	imestamp (epoch)
ttl	Time to Live (e	xpiration)

```
{
    "id": "Sample",
    "familyName": "Jones",
    "address": {
        "addressLine": "789 Harbor Bouleva
        "city": "Chicago",
        "state": "IL",
        "zipCode": "60603"
    },
    "_rid": "IwdPAIDPIS0BAAAAAAAAAA==",
    "_self": "dbs/IwdPAA==/colls/IwdPAIDPIS0=/docs/IwdPAIDPIS0BAAAAAAAAAA==/",
    "_etag": "\"09008ef9-0000-0700-0000-5d220a430000\"",
    "_attachments": "attachments/",
    "_ts": 1562511939
}
```

Summary

DOCUMENT DATABASE

RELATIONAL VS. DOCUMENT

DENORMALIZED DATA MODEL

DATA MIGRATION TOOL

Special document properties