Schema Design

schema, n. – a representation of a plan or theory in the form of an outline or model.

Data Modeling

- How do we represent real world relationships and properties in our program?
  - ...in a way that makes writing the program easy
  - ...while remaining flexible for future changes
  - ...oh, it also has to be fast (enough).

Designing a Schema

- Analysis
  - What does my program need to output?
  - What data will I need to produce that output?
- Conceptual Design
  - Conceptual entities and their relationships
- Logical Design
  - What does the structure of the database look like?
- Physical Design
  - How do we get there? CREATE TABLE statements, etc.

Designing for Different Data Stores

- **Relational:** Tables, columns, keys, relationships, constraints
- **Hierarchical:** Keys, value types, references, nesting
- **Javascript code:** Type signatures

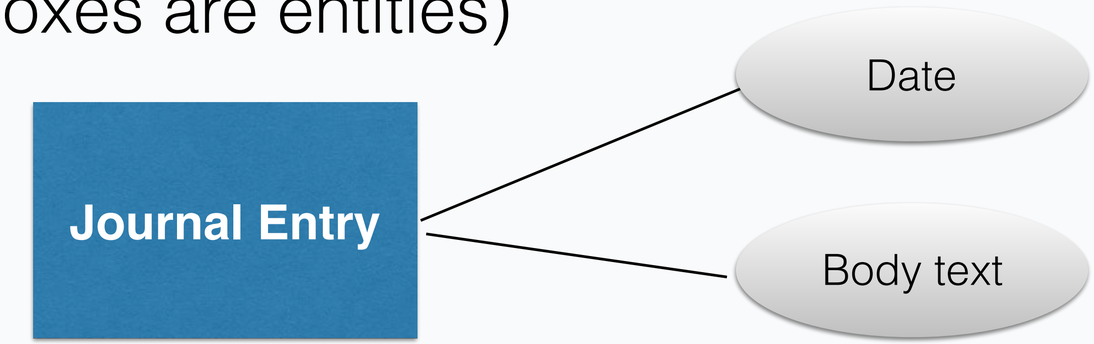
Example: A Journal Analysis

- I want a program to keep my journal in.
- I want to be able to enter the text of each journal entry.
- I want to be able to see journal entries chronologically.

Example: A Journal Conceptual Design

(circles are properties)

(boxes are entities)



Example: A Journal Relational Design

entries	
id	int, primary key
date_created	date
text	text

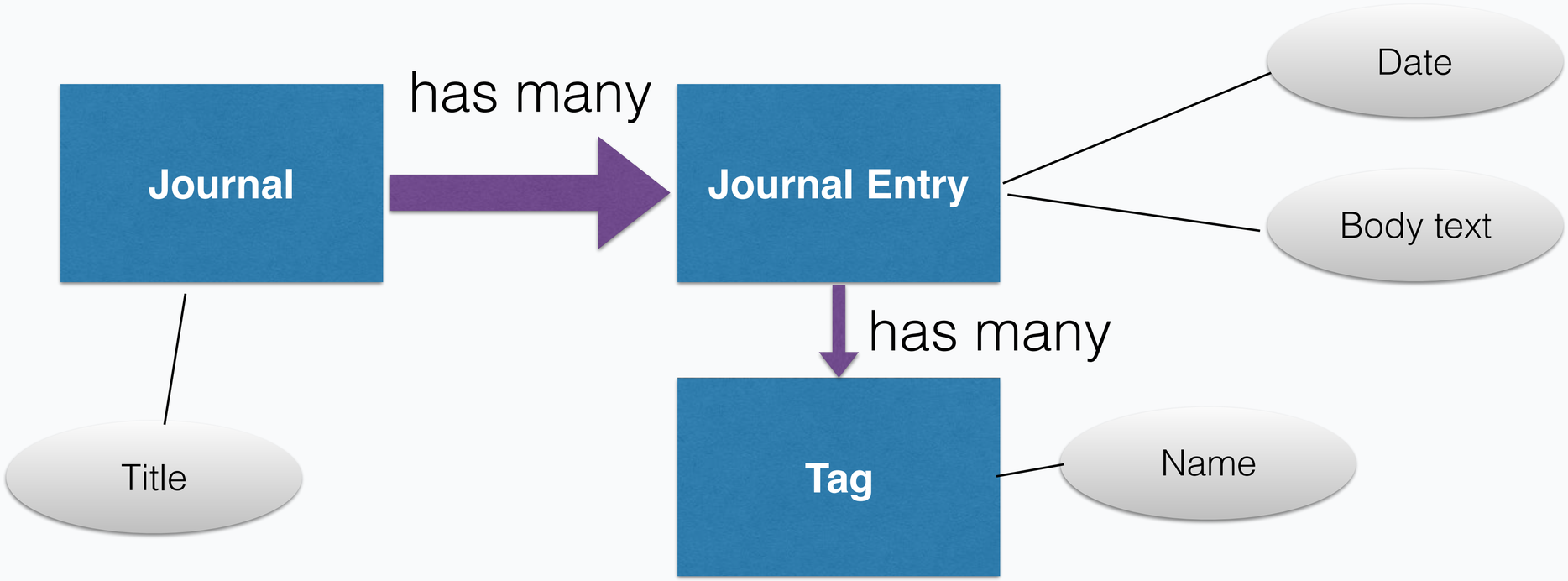
Example: A Journal Hierarchical Design

```
/
  /entries
    /$key
      date_created: Time
      text: String
```

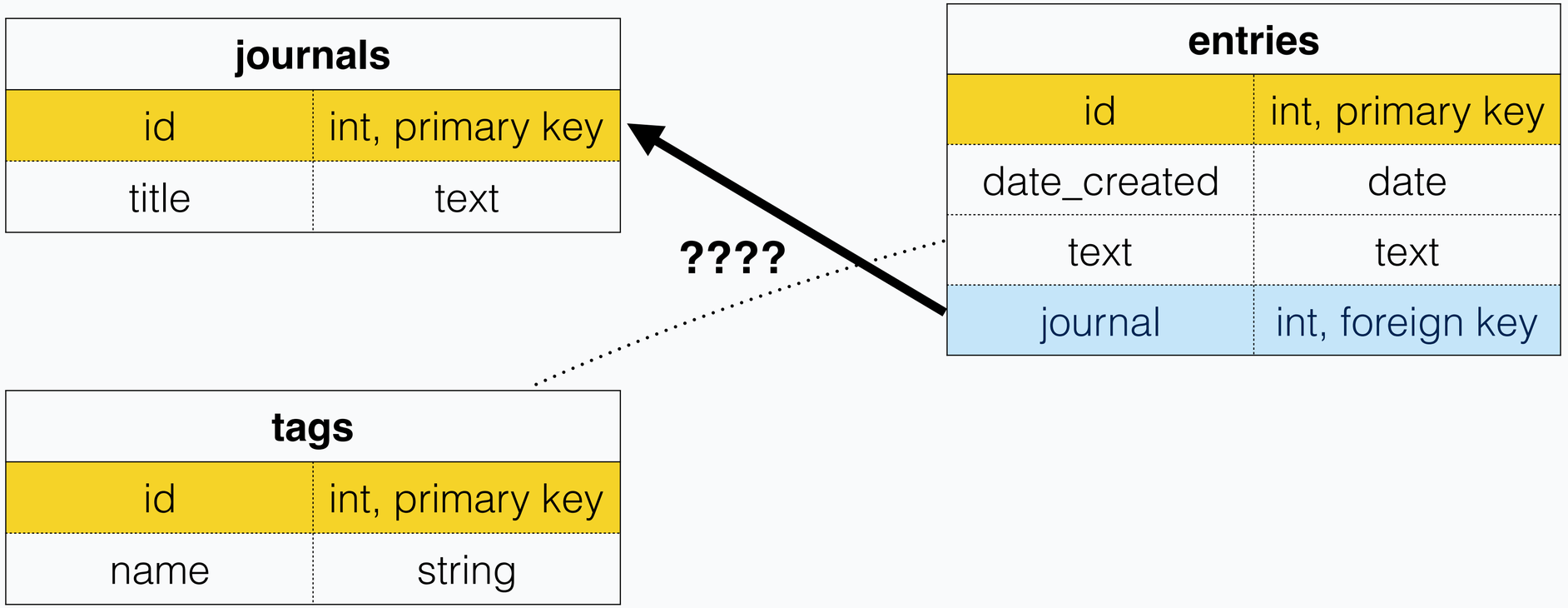
All done!

- Oh wait, I forgot a couple of things
  - I want to be able to have multiple journals
  - I want to be able to #tag entries and find all entries with a particular #tag
- Analysis?

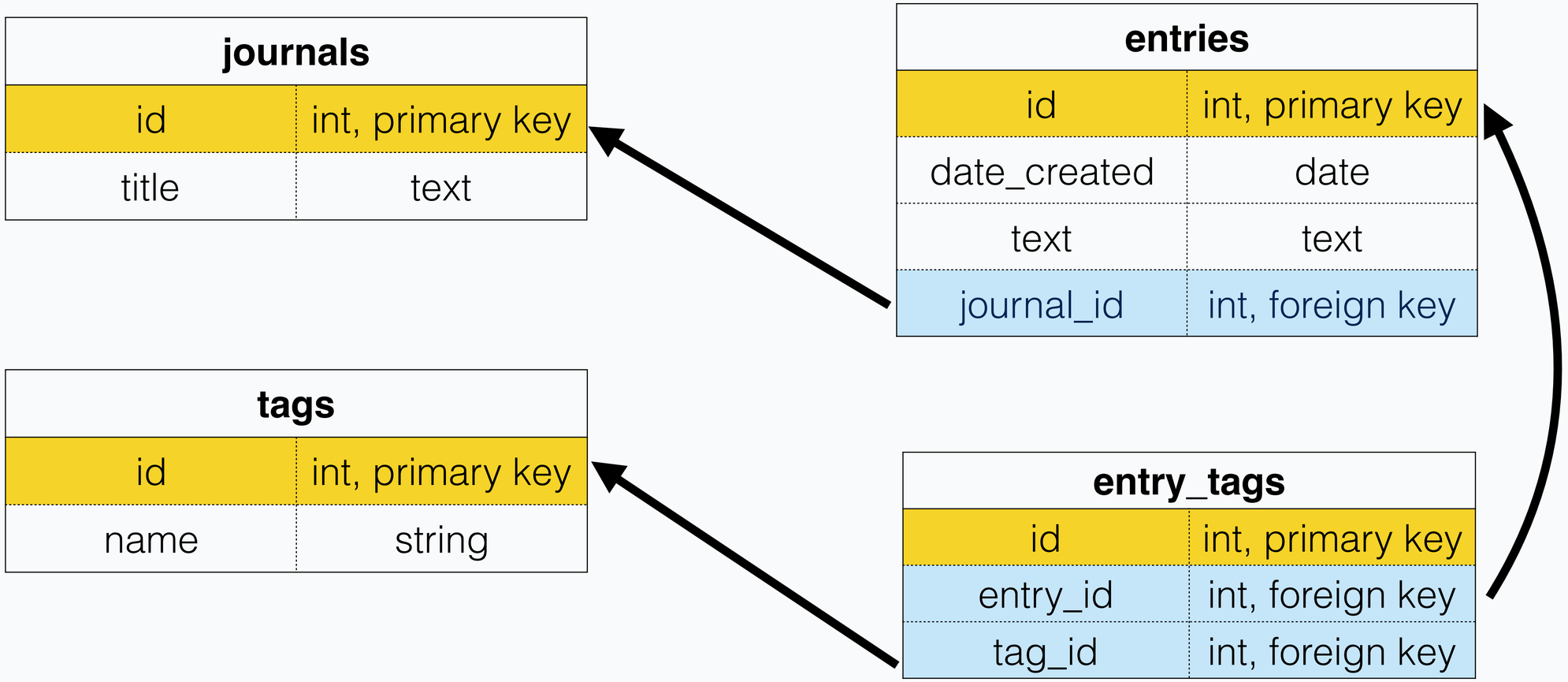
Example: A Journal
Conceptual Design, Take 2



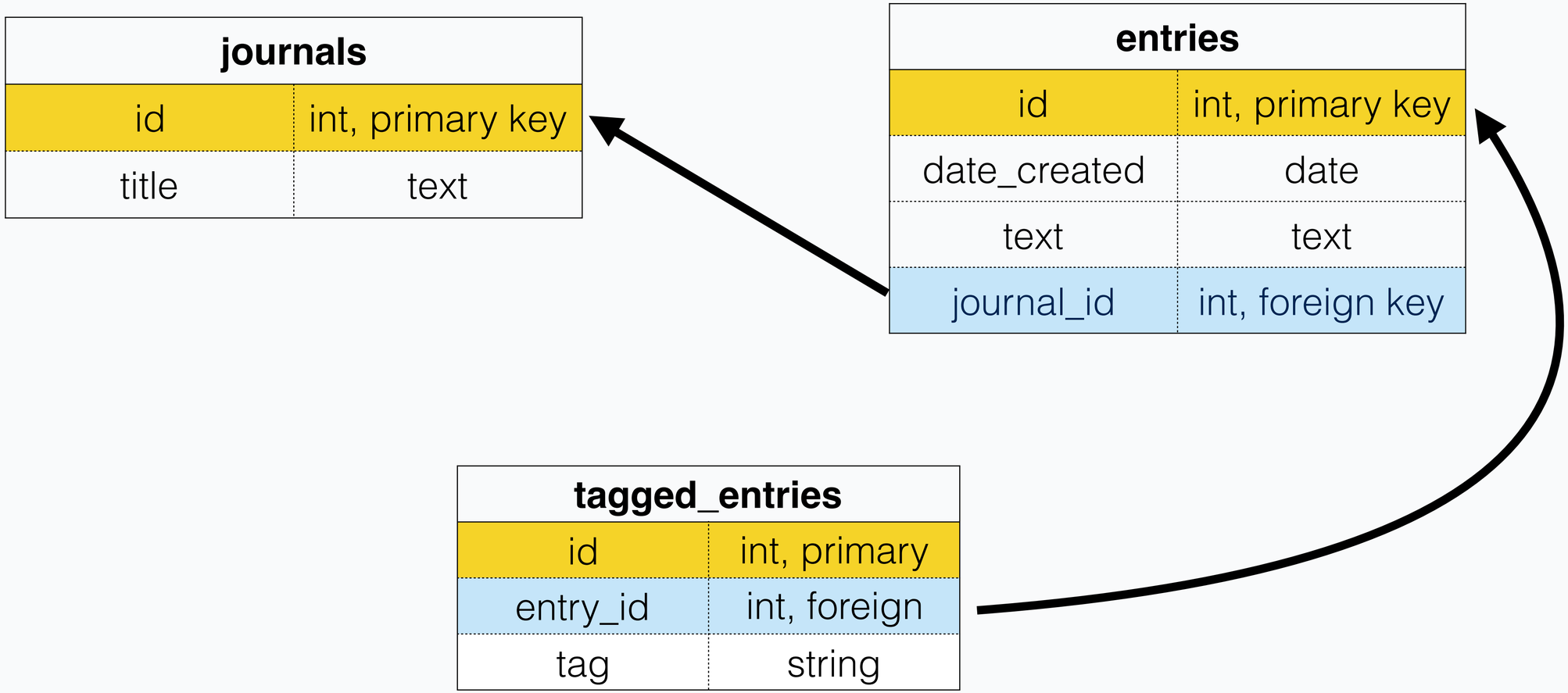
Example: A Journal
Logical Design, Take 2?



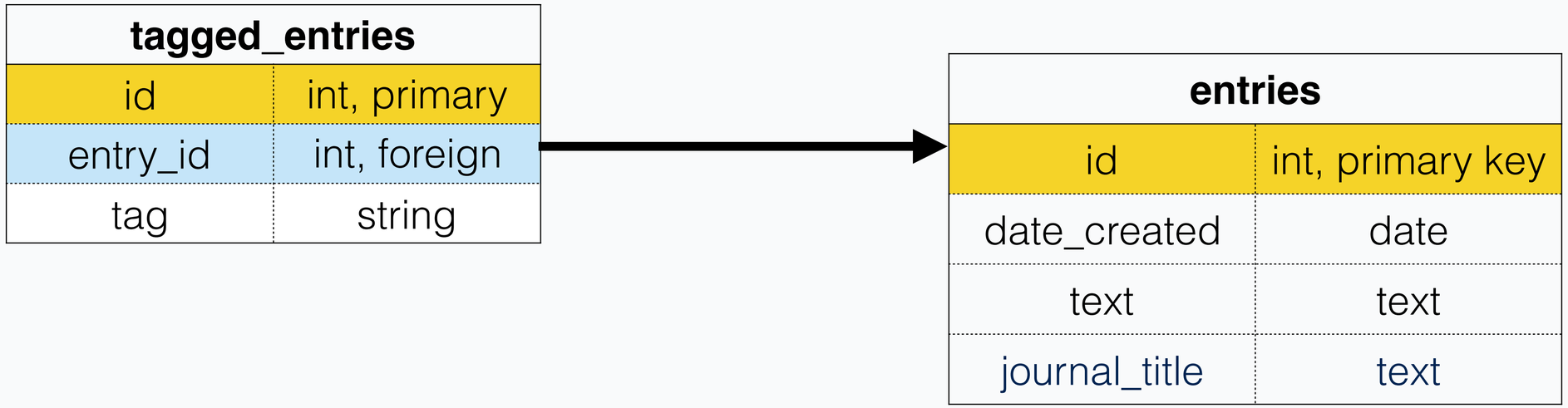
Example: A Journal
Logical Design, Take 2



Example: A Journal
…But why not?



Example: A Journal
…Or even just this?



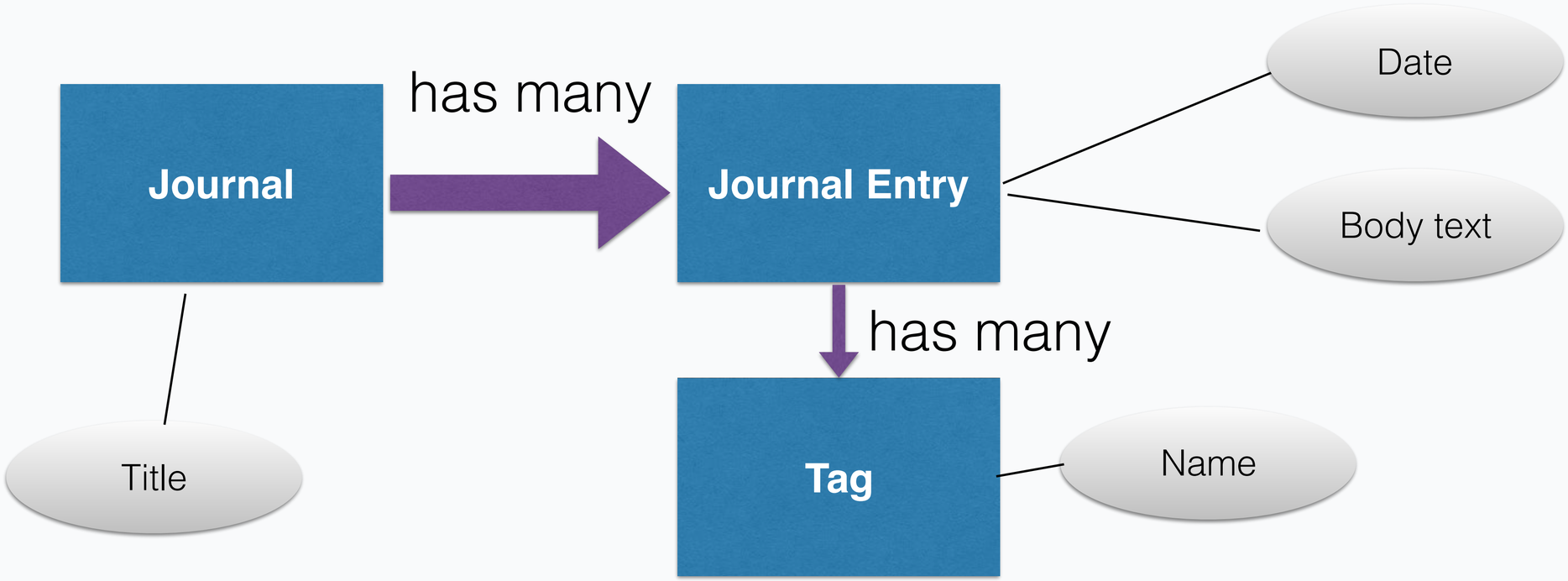
Normalization

- "The process of organizing the columns and tables of a relational database to minimize data redundancy." ~ *wikipedia, source of all truth.*
- What happens when I want to change the name of "happy times" to "sadness"?
- How many comparisons does the database have to do to find all the entries for a given journal?

select * from entries;

id	date_created	text	journal_title
0	2016-04-01	I am happy	happy times
1	2016-04-02	I am very happy	happy times
2	2016-04-03	Despair fills me	happy times
3	2016-04-03	Sadness is my life	an anatomy of pain

Example: A Journal
Conceptual Design, Take 2



Example: A Journal
Hierarchical Design (Deeply Nested)

a read here

```
/
  /journals
    /$key
      title: String
      /entries
```

must fetch the text of all entries

```
        /$key
          date_created: Time
          text: String
          /tags
            mermaid: true
            happy: true
```

Example: A Journal Hierarchical Design (Flat)

```
/
  /journals
    /$key
      title: String
      /entries
        $key: Boolean
  /entries
    /$key
      date_created: Time
      text: String
```

Design one!

- Twitter
- Gmail
- Facebook
- Instagram
- Wordpress
- Wikipedia
- AirBnB
- Google (search)