

The Clown Shoes Collective presents spacewalr.us

David Balatero

Vjeko Brajkovic

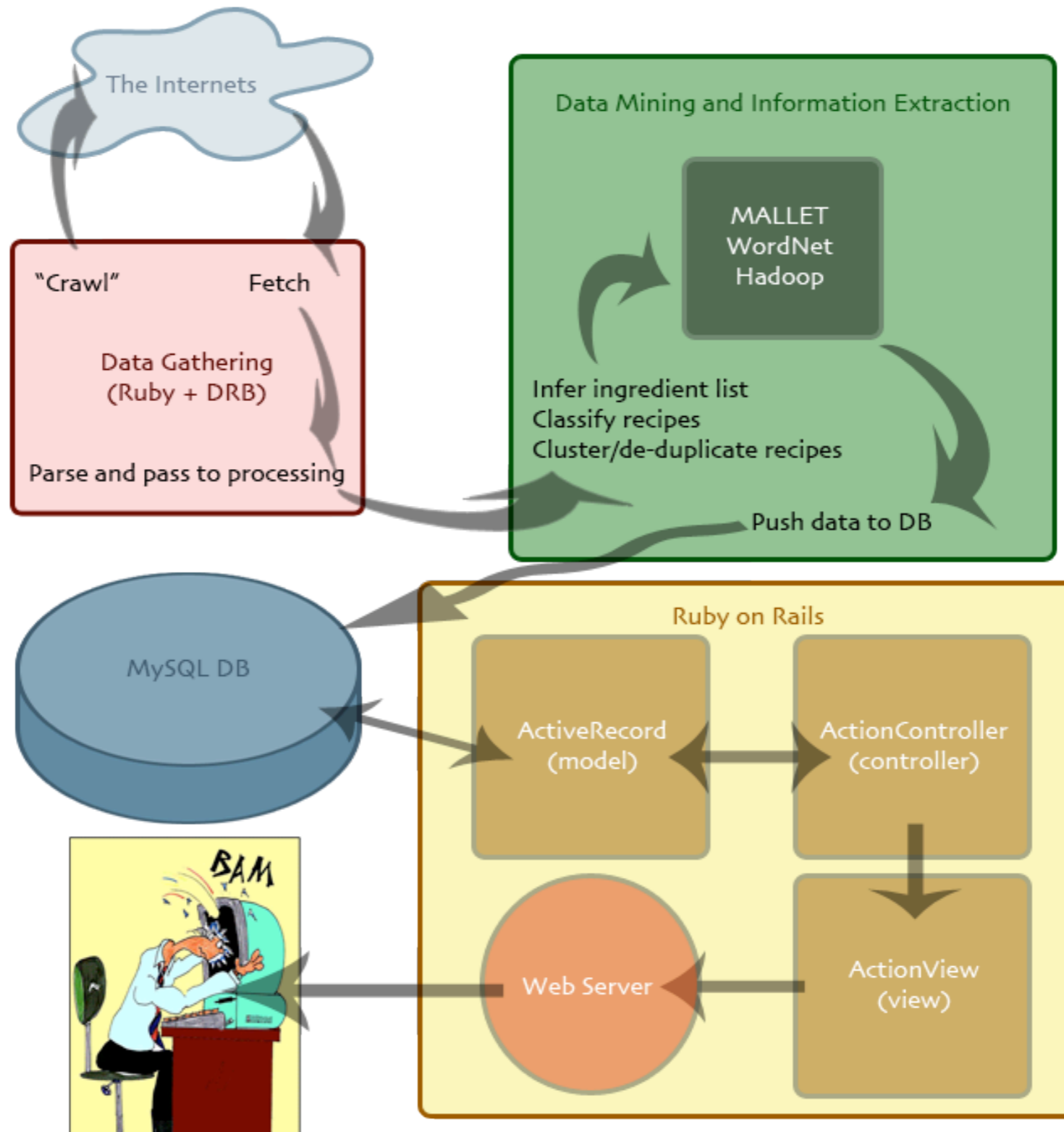
Felix Chu

Daniel Otero

The Goal:

- Help the struggling noob cook
- Provide “deep” recipe search
- Help the indecisive by recommendation
- Give users what today’s recipe websites attempt to provide, but do it well

The Idea



Getting Data

- Identified recipe sites with large collections of “consistent” data
- Designed a scalable fetching/crawling bot using Distributed Ruby
- Created parsers from each data format to a standard format compatible with our typical Hadoop jobs

It's all about the ingredients

- Most of the features we planned to offer relied on meta-knowledge of ingredients
- To illustrate the problem, a few examples:
 - **3 bananas (easy)**
 - **3 cups of bananas (still pretty easy)**
 - **1 1/2 tubs of Philadelphia Lite cream cheese, chilled (tricky)**
- Goal: try to identify a quantity, a unit (or the absence of one), the most naturally identified “ingredient”, and the remaining text.

Getting fancy

- Our framework can easily extract quantity and unit:
- Used WordNet library to find nouns and the neighboring words modifying them

1 1/2 cups of fresh green beans, neatly chopped

quantity: 1.5
unit: Volume.CUP
prefix: "of fresh"
ingredient: "green bean"
suffix: ", neatly chopped"

- Fetched a labeled data set from Epicurious.com
- Trained Maximum Entropy classifiers on the labeled data
- Used Near Duplicates algorithm to find duplicate recipes within categories

What we didn't see coming

- Recommendations are great, but where do we get enough user data?
 - Re-crawled the same sites for their user ratings
- Data was not nearly as consistent as we'd hoped
- Ingredients are pretty hard to extract
 - Trial and error (and hours of Googling) saved the day
- Our evaluation based on random sampling:
 - Ingredient extraction: 95.2% accurate, Recipe classification: 91.3% accurate

Oh yeah, and we wrote a web app

spacewalr.us
cause kung fu is for little girls

[home](#)

[search](#)

[browse categories](#)

[about](#)

[contact](#)

You are logged in as **ohboyotero**. [profile](#) [update profile](#) [favorites](#) [inventory](#) [dislikes](#) [logout](#)



steamed rice

Yield: 4 servings

Timing: 4 minutes prep , 32 minutes cook


Ingredients:

1. 1 cup jasmine rice

Directions:

1. In a heavy pot with a lid, bring 2 c. water to a boil. Add rice and salt and lower heat.
2. Cover the pot and simmer rice for 15 to 20 minutes.
3. After 15 minutes, take off the pot lid and test a few grains of rice - if the center is at all crunchy, simmer a few more minutes.
4. Test rice again. When it's done, fluff it with a fork to let steam escape and help keep grains separate.

Rate it: 
3 ratings

Give me batches. [Go!](#) 

It does everything the other sites do

search for:

include

We're limiting your search to recipes that can be prepared using only your 14 ingredients.

[Manage your inventory](#)

exclude

We're excluding your [Dislikes](#):

- You have no Dislikes. Want to [change that?](#)

Also exclude:

Exclude:

★ minimum

search results

Displaying 3 results, page 1 of 1

Sort by: [rating](#) | [my rating](#) | [title](#)



[Chicken with Mango and Eggs](#)



2 ratings

Yield: 3 servings

Timing: 5 minutes prep , 22 minutes cook



[Pork with Beer](#)



1 ratings

Yield: 5 servings

Timing: 3 minutes prep , 39 minutes cook



[Steamed Rice](#)



3 ratings

Yield: 4 servings

Timing: 4 minutes prep , 32 minutes cook

And then there's recommendation

your profile

This is a happy place. It is full of the familiar. You have at your disposal a new recipe recommendation (if you've rated enough recipes), a list of recipes you've recently rated, and your top favorites.

account features

Your account comes with a few perks. Here are a few:



favorites: You're inevitably going to come across some recipes that you like more than others. To keep track of these, you have a favorites collection. To add a recipe you're viewing to your favorites, just click on the big ol' "Add to favorites" link at the top of the page. You can view, sort, and remove your favorites on the [favorites](#) page.



inventory: To make your life a bit easier, you can keep track of each ingredient you have on hand. Thus, rather than throwing your hands up in the air in an exasperated fashion, furious that you can't find a damn recipe that you can actually cook, you can just tell our search to use your [inventory](#) to restrict the search. That way you'll be able to cook all of your search results.



Dislikes: Similarly, if you never ever want to eat a particular ingredient, you can add it to your [Dislikes](#). Dislikes will always be excluded from all of your searches, so it's a perfect way to handle any allergies or hated foods.

your recommendation

Looking to try something new? Check out this recipe:



[Cheese dog biscuits](#)

★★★★★

Yield: 35 biscuits

Timing: 5 minutes prep , 25 minutes cook

your latest ratings

[Quick and Easy, No Cook Holiday But...](#) was rated a **4** about 2 minutes ago.

[seared tofu with with green beans a...](#) was rated a **5** about 3 minutes ago.

[greek sesame seed candy](#) was rated a **5** about 3 minutes ago.

[Puyallup Fair Scones \(Fisher Scones\)](#) was rated a **5** about about 2 hours ago.

your top-rated favs

[Southwestern Burger With Chipotle Mayonnaise & BBQ Sauce](#)

★★★★★

Yield: 3 servings

Questions?

