

From Zero to CPAN

Adding A New Product Search to Magazines.com

Cory G Watson: **gphat**

- *Magazines.com*: Group Director, Development
- *Cold Hard Code*: Creative Director
- Contributor and Color Commentator

Objectives

- Research
- Refactor
- Reuse
- Return

Magazines.com

Virtual Middle Man

What We Do

- We focus on the Internet.
- Publishers focus on publishing.
- We take your address & money.
- We give address to publisher, we keep some of the money.

Search

Quite the Rabbit Hole

Research

What I Inherited:

Flat file + customized grep + dedicated server

What I Made:

Custom SQL + “Index” table

What I Refined To:

Custom SQL + “Index” table + Serialization + Cache

New Problem

50 - 100 x increase in products.

Shit.

So much for my awesome SQL.

Now What?

CPAN

**Lucene, Plucene,
Kinosearch, Xapian, etc.**

Tried, but never a good fit.

Search

- Magazine name is very important
- Categories and synonyms
- Keywords & descriptions

Problems

- W
- “cosmo” “gentleman’s quarterly”
- kids magazines

Solr

<http://lucene.apache.org/solr/>

Solr Features

- Fast
- Custom schemas
- Full text
- Highlighting
- Faceting
- Complex queries & functions.
- Filtering
- Caching
- Replication

Solr Downsides

- Java
- Lots of schizophrenic docs

CPAN! (bricas++)

WebService::Solr

Refactor / Reuse

What can we leverage?

- `WebService::Solr`
- Relationship with bricas
- Enlightened / Modern Perl
- Moose, Roles

What to Refactor?

- Third search switch... abstraction?
- Caching
- Updating?

Data::SearchEngine

- Moose Role
- search method
- Results & Item object
- Serializable

Data::SearchEngine::Solr

- Thin wrapper around `WebService::Solr`
- Added roles to `Data::SearchEngine`:
Faceted, Updateable

Return

REFINE YOUR RESULTS

Category

Any Category

- Entertainment & TV (10)
- Fashion & Style (2)
- Women's (9)
- Ethnic (4)
- Enrichment (17)
- News & Politics (4)
- Business & Finance (14)
- Cooking, Food & Beverage (3)
- Home & Gardening (6)
- Lifestyle (40)
- Computers & Electronics (2)

+ See More Categories

Price

Any Price

- Under \$10 (5)
- \$10 - \$20 (54)
- \$20 - \$30 (25)
- \$30 - \$50 (12)
- Over \$50 (13)
- Free Business Magazines (3)

Rating

Any Rating

- 4 Stars & up (32)
- 3 Stars & up (37)
- 2 Stars & up (38)
- 1 Star & up (39)

You Searched For "people".

Showing 1 - 12 of 112 results.

Items per page: 12 | 24 | 48

Sort by



People

★★★★★

26 issues: \$56.94

You Save 45%

[Details](#)



StyleWatch, by the Publishers of People

★★★★★

10 issues: \$19.95

You Save 50%

[Details](#)



People En Espanol

★★★★★

11 issues: \$19.97

You Save 39%

[Details](#)



Native Peoples

★★★★★

6 issues: \$19.95

You Save 32%

[Details](#)



Time

★★★★★

28 issues: \$16

You Save 88%

[Details](#)



In Touch Weekly

★★★★★

6 issues: \$9.97

You Save 44%

[Details](#)



BusinessWeek

★★★★★

10 issues: \$9.97

You Save 80%

[Details](#)



Better Homes & Gardens

★★★★★

12 issues: \$14.97

You Save 64%

[Details](#)



Results

- Code on CPAN
- Less internal code to maintain
- 10x to 30x performance increase
- Lots of new features
- A few new boxes to maintain

Why This Is Great

- Used existing code
- I learned a lot
- Made something new and interesting
- Company saved time
- Company contributed back
- Author's module improved

Summary

- Solr is *Awesome*
- Moose is *Awesome*
- Roles are *Awesome*

Questions?

Colophon

- Titling: Tungsten 84pt, Hoefler & Frere-Jones
- Body: Hoefler Text 42pt, Hoefler & Frere-Jones
- Presentation: Keynote '09, Apple, Inc.