

Searching ORM

Sam Farmer

Me

- Senior Applications Architect at FirstComp
- ColdFusion, JavaScript, yada, yada...
- Hobbies:
 - Soccer. Play and Arsenal fan.
 - Code. oldjam.com
- Live in Washington, DC with my wife and two daughters

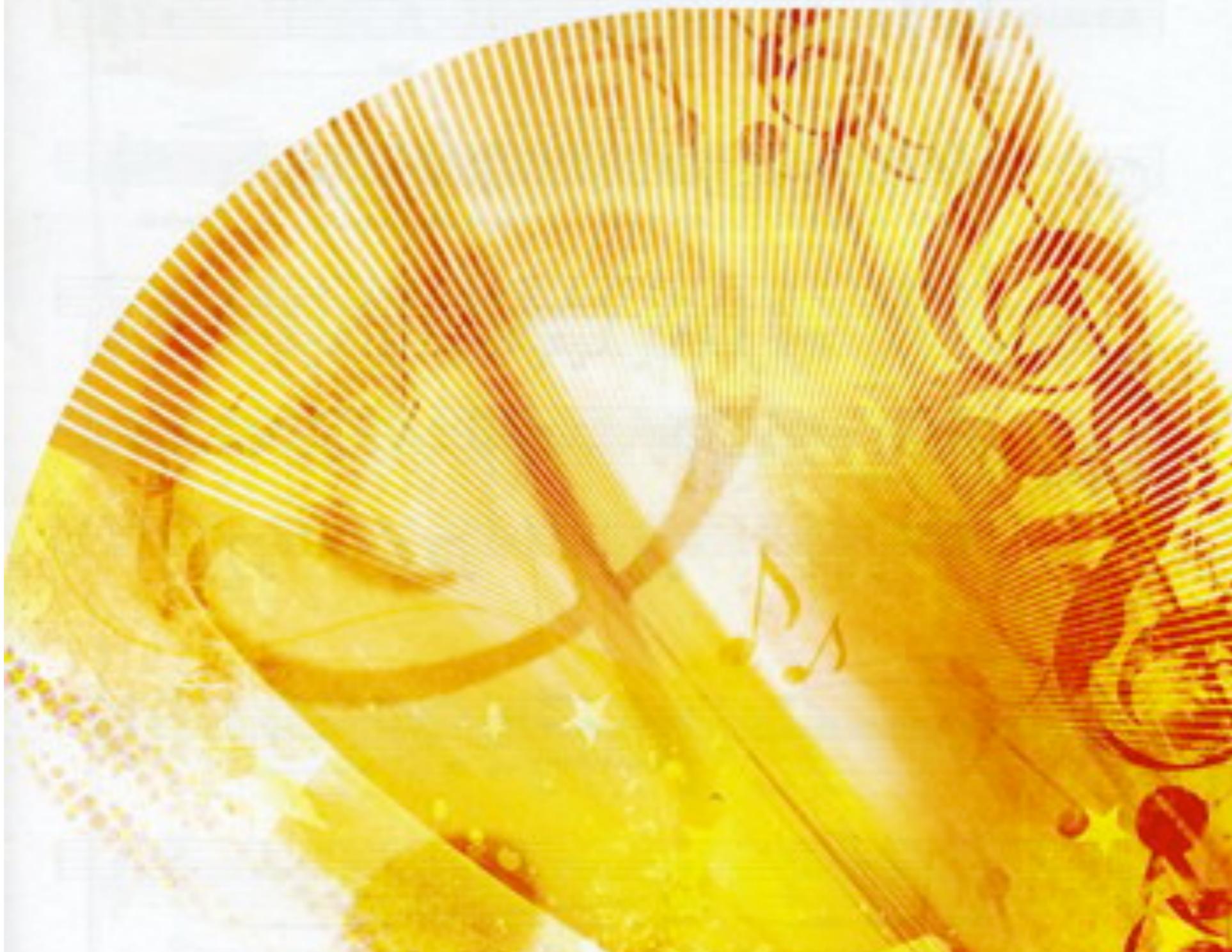
Agenda

- Why search
- Pre-CF10 Enterprise
- With CF10 Enterprise

SSA/PIANO
NOV941006

I STILL HAVEN'T FOUND WHAT I'M LOOKING FOR

WORDS & MUSIC BY U2



Other Sessions

- A/B Testing with Squabble (Mark Mandel)
- Creating ColdFusion Builder Extensions (Simon Free)
- Authentication made Easy using Twitter, Facebook, Google, and more (Billy Cravens)
- Node.js and You - Server Side Javascript (Simon Bateman)

Who likes Objects?

Who likes Objects?

Who likes cf.Objective()?

Who likes ORM?

Using or planning to use
CF10?

Retrieving Objects

Based on criteria

```
entityLoad( "user", 1 );  
  
ormExecuteQuery( "FROM user  
WHERE state = 'DC' or state='NE' " );
```

Why Searching

Get what you are looking for

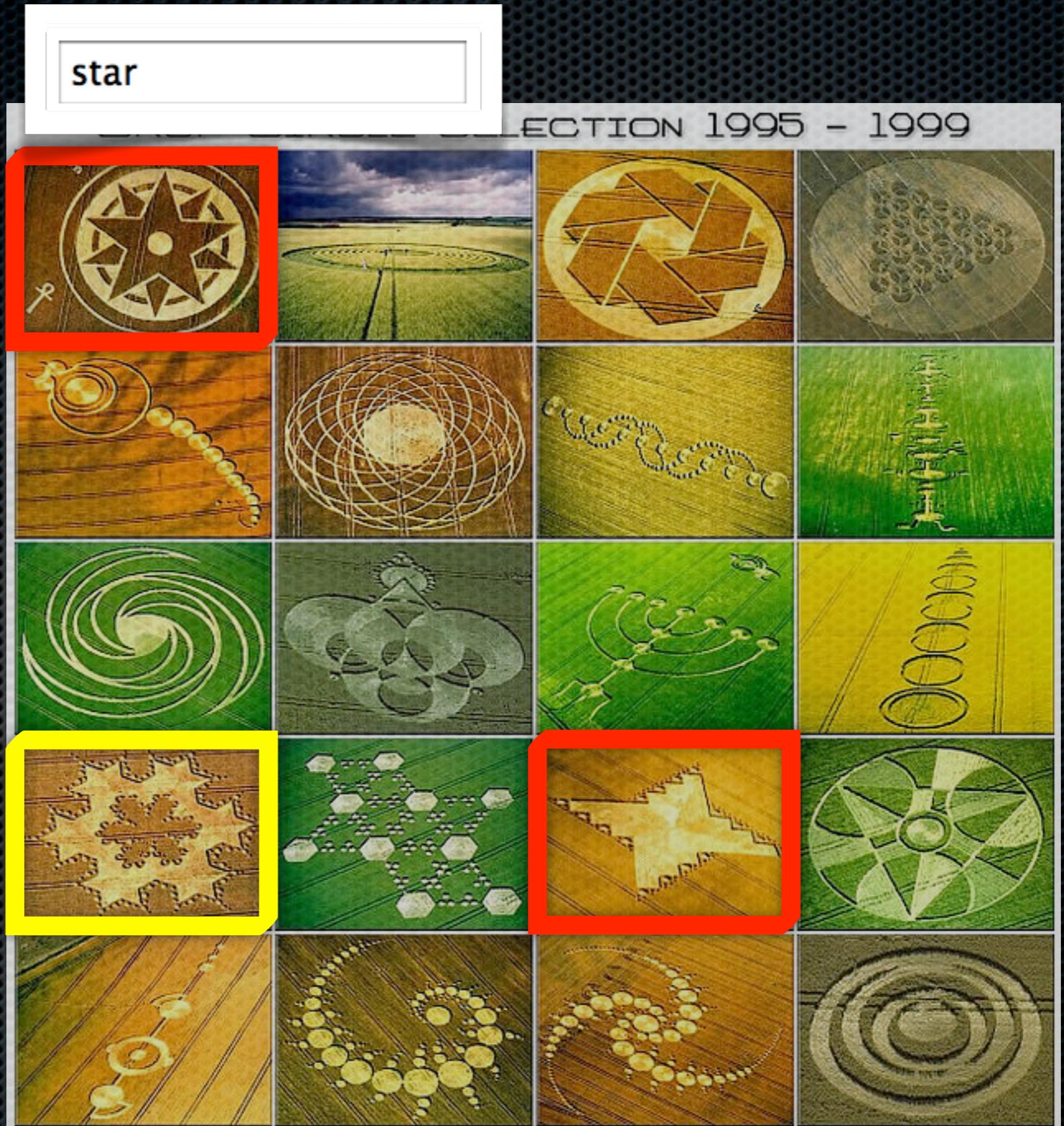
1998



CROP CIRCLE SELECTION 1995 - 1999



That star looking shape



What we used to have

Product:	<input type="text"/>	Product Area:	<input type="text"/>
Version:	<input type="text"/>	Platform:	<input type="text"/>
Found in Build:	<input type="text"/>	Fixed in Build:	<input type="text"/>
State:	<input type="text"/>	Status:	<input type="text"/>
OS Language:	<input type="text"/>	App Language:	<input type="text"/>
Severity:	<input type="text"/>	Reported By:	<input type="text"/>
Date Created:	<input type="text"/> = <input type="button"/>	<input type="text"/>  (mm/dd/yyyy)	
Title:	<input type="text"/>		
Description:	<input type="text"/>		
	<input type="button" value="Search"/>		

Now we have

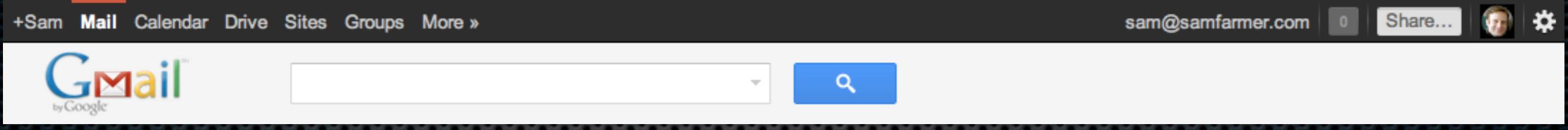
facebook



Search



Sam Farmer | Home | ▾



Search

Lets look at the options

Artist and Art



Looking for



Michael



Mychael
by Night



Face of
Mikael

Like
Everything has its place

LIKE

- Don't snicker
- Premature optimization

With HQL

- Similar to SQL

```
<cfquery dbtype="hql" name="getArt">
FROM ART
WHERE name like
<cfqueryparam value="Mi%">
</cfquery>
```

Found: 1

- Michael by Aiden Donolan

array																	
1	component presentations.ormsearch.non.ART																
	PROPERTIES																
	<table border="1"><tr><td>id</td><td>2</td></tr><tr><td>name</td><td>Michael</td></tr><tr><td>description</td><td>Pastels/Charcoal</td></tr><tr><td>price</td><td>13900</td></tr><tr><td>largeimage</td><td>aiden02.jpg</td></tr><tr><td>mediaid</td><td>1</td></tr><tr><td>issold</td><td>NO</td></tr><tr><td>artist</td><td></td></tr></table>	id	2	name	Michael	description	Pastels/Charcoal	price	13900	largeimage	aiden02.jpg	mediaid	1	issold	NO	artist	
id	2																
name	Michael																
description	Pastels/Charcoal																
price	13900																
largeimage	aiden02.jpg																
mediaid	1																
issold	NO																
artist																	

With Hibernate Criteria API

- More "object-ey"

```

order = createObject( 'java',
"org.hibernate.criterion.Order" );
Restrictions = createObject( 'java',
"org.hibernate.criterion.Restrictions" );

artSearch =
ormGetSession().createCriteria( "art" );
artSearch.add( Restrictions.like( "name", "Mi
%" ) );
artSearch.addOrder( order.asc( "name" ) );

getArt = artSearch.list();

writeDump( getArt );

```

Found: 1

- Michael by Aiden Donolan

array																	
1	component presentations.ormsearch.non.ART																
PROPERTIES																	
	<table border="1"> <tr> <td>id</td><td>2</td></tr> <tr> <td>name</td><td>Michael</td></tr> <tr> <td>description</td><td>Pastels/Charcoal</td></tr> <tr> <td>price</td><td>13900</td></tr> <tr> <td>largeimage</td><td>aiden02.jpg</td></tr> <tr> <td>mediaid</td><td>1</td></tr> <tr> <td>issold</td><td>NO</td></tr> <tr> <td>artist</td><td></td></tr> </table>	id	2	name	Michael	description	Pastels/Charcoal	price	13900	largeimage	aiden02.jpg	mediaid	1	issold	NO	artist	
id	2																
name	Michael																
description	Pastels/Charcoal																
price	13900																
largeimage	aiden02.jpg																
mediaid	1																
issold	NO																
artist																	

(If you like
look at
ColdBox)

coldbox.org



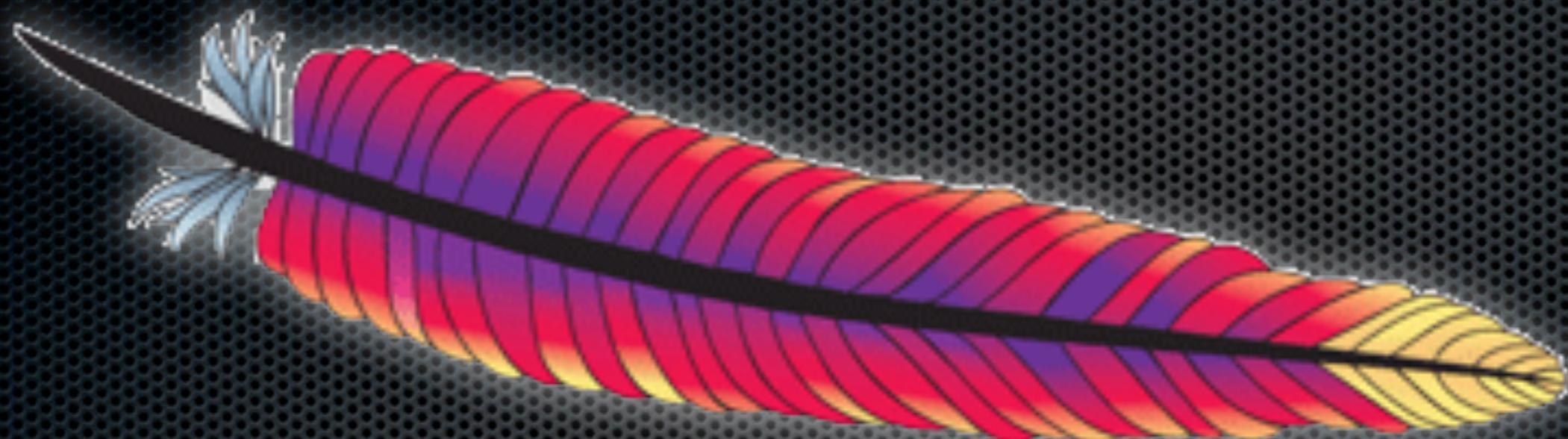
ColdBox

Search Engines

Lets look at the options

Database (Full Text Search) Options

- I couldn't make work
- Smarter people may have
- Or, double query to
 - get Ids
 - get entities



Apache Lucene

A "flagship sub-project"

"Our core algorithms along with the Solr search server power applications the world over, ranging from mobile devices to sites like Twitter, Apple and Wikipedia."

And soon your apps! :)

Lucene & Solr

- Both from Apache
- Solr is a servlet
- Lucene is an index
- Solr needs Lucene but Lucene does not need Solr

What can Lucene do?

- Fuzzy
- Words near each other
- Boost
- Range

SolrEntity

- Nathan Mische proof of concept
 - ▶ <http://cfml.us/z3>

How it works

- Base cfc
- Metadata defines configuration
- Event Handlers
 - postUpdate(), postInsert(), postDelete()
 - update Solr

Entry.cfc

```
component persistent="true" table="entry"
extends="SolrEntity" solrcollection="blog" {

    property name="entryid" fieldtype="id"
    generator="native" solrfield="key";
    property name="author" fieldtype="many-to-one"
    fkcolumn="authorid" cfc="Author";
    property name="title" solrfield="title";
    property name="content" solrfield="body";

}
```

ColdFusion 10 ORM Search

Putting the pieces together

```
searchArt = ormSearch("name:Mi*", "art");
```

```
writeDump( var=searchArt, top=3, showudfs=false );
```

Found: 1

- Michael by Aiden Donolan

struct	
DATA	array
	1 struct
context	struct
	name array
	1 Michael
entity	component presentations.ormsearch.search.ART
	PROPERTIES
	id 2
	name Michael
	description Pastels/Charcoal
	price 13900
	largeimage aiden02.jpg
	mediaid 1
	issold NO
	artist component presentations.ormsearch.search.artist
	score 1.0
MaxTotalCount	1
dumpCount	1
dumpCountText	1
	score 1.0

```
searchArt = ormSearch("name:Michael~", "art" );
```



Found: 3

- Michael by Aiden Donolan
- Mychael by Night by Elicia Kim
- Face of Mikael by Jeff Baclawski

struct															
DATA	array														
1	<table border="1"><tr><td>struct</td><td></td></tr><tr><td>context</td><td><table border="1"><tr><td>struct</td><td></td></tr><tr><td>name</td><td>array</td></tr><tr><td>1</td><td>Michael</td></tr></table></td></tr><tr><td>entity</td><td></td></tr><tr><td>score</td><td>3.4162211</td></tr></table>	struct		context	<table border="1"><tr><td>struct</td><td></td></tr><tr><td>name</td><td>array</td></tr><tr><td>1</td><td>Michael</td></tr></table>	struct		name	array	1	Michael	entity		score	3.4162211
struct															
context	<table border="1"><tr><td>struct</td><td></td></tr><tr><td>name</td><td>array</td></tr><tr><td>1</td><td>Michael</td></tr></table>	struct		name	array	1	Michael								
struct															
name	array														
1	Michael														
entity															
score	3.4162211														
2	<table border="1"><tr><td>struct</td><td></td></tr><tr><td>context</td><td><table border="1"><tr><td>struct</td><td></td></tr><tr><td>name</td><td>array</td></tr><tr><td>1</td><td>Mychael by Night</td></tr></table></td></tr><tr><td>entity</td><td></td></tr><tr><td>score</td><td>1.2200791</td></tr></table>	struct		context	<table border="1"><tr><td>struct</td><td></td></tr><tr><td>name</td><td>array</td></tr><tr><td>1</td><td>Mychael by Night</td></tr></table>	struct		name	array	1	Mychael by Night	entity		score	1.2200791
struct															
context	<table border="1"><tr><td>struct</td><td></td></tr><tr><td>name</td><td>array</td></tr><tr><td>1</td><td>Mychael by Night</td></tr></table>	struct		name	array	1	Mychael by Night								
struct															
name	array														
1	Mychael by Night														
entity															
score	1.2200791														
3	<table border="1"><tr><td>struct</td><td></td></tr><tr><td>context</td><td><table border="1"><tr><td>struct</td><td></td></tr><tr><td>name</td><td>array</td></tr><tr><td>1</td><td>Face of Mikael</td></tr></table></td></tr><tr><td>entity</td><td></td></tr><tr><td>score</td><td>0.56937003</td></tr></table>	struct		context	<table border="1"><tr><td>struct</td><td></td></tr><tr><td>name</td><td>array</td></tr><tr><td>1</td><td>Face of Mikael</td></tr></table>	struct		name	array	1	Face of Mikael	entity		score	0.56937003
struct															
context	<table border="1"><tr><td>struct</td><td></td></tr><tr><td>name</td><td>array</td></tr><tr><td>1</td><td>Face of Mikael</td></tr></table>	struct		name	array	1	Face of Mikael								
struct															
name	array														
1	Face of Mikael														
entity															
score	0.56937003														
MaxTotalCount															
3															

Lucene Query Syntax

- This is the "magic"
- <http://cfml.us/Ax>

Wildcard Searches

? = one character

Arsen?l (finds Arsenal)

* = many characters

Can't be first

Arse*

Ar*al

Fuzzy Searches ~

~ will find similar spellings

Arsenal~ (finds Arsenal)

~0-1 indicates *how* similar

0=least similar, 0.5=default, 1=most similar

Arsenal~0.1

Arsenal~1

Boost ^

^ boost makes one term more important/relevant

Arsenal^6 Chelsea (makes Arsenal more relevant)

^0.1-99 indicates *how much more relevant*

1 = default

Normally whole numbers

Proximity Searches ~9

Find two words close to each other

"one two"~NUM where NUM is how close

"Arsenal Champions"~6 (finds Arsenal)

How it works

ColdFusion / JVM

Hey Lucene can you look this up?

Sure, here are matching Ids

Thanks! Hibernate entities please

Sure, hey database can I get data?

Here you go

Database

Most def.
SELECT blah
FROM table
WHERE Id
= 999

Let's see that...

Is it slow?

- Two fast operations
 - Lucene search
 - Look up by Ids
- Load Testing

Skip the database

- `ormSearchOffline()`
- Returns only fields specified by `indexStore`

```
searchArt = ormSearchOffline("name:Michael~",
"ART", ["id","name"] );
```

```
writeDump( var=searchArt, top=3, showudfs=false );
```



Setup

Application.cfc

```
component {  
    this.name = "ormSearchCriteria";  
    this.datasource = "cfartgallery";  
    this.ormEnabled = true;  
    this.ormSettings = {  
        searchEnabled = true | false,  
        search = { [autoIndex = true],  
                   [indexDir = filepath],  
                   [language = "english"] }  
    };  
}
```

Application.cfc

USED IN SAMPLES

```
component {  
    this.name = "ormSearchCriteria";  
    this.datasource = "cfartgallery";  
    this.ormEnabled = true;  
    this.ormSettings = {  
        searchEnabled = true,  
        search = { autoIndex = true },  
        logsql = true  
    };  
    this.invokeImplicitAccessor = true;  
}
```

Entity - Attributes

```
component persistent = true  
    indexable = true | false  
    indexLanguage = "english"  
    autoIndex = true | false {
```

Entity - Properties

```
component persistent = true
    indexable = true {
property name="name"
    indexable = true | false
indexTokenized = true | false
indexFieldName = ""
indexBoost = 0-1
indexStore = true | false | compressed
indexLanguage = "english"
```

Art.cfc

USED IN SAMPLES

```
component persistent="true" indexable="true" {  
    property name="id" column="ARTID" fieldtype="id";  
    property name="name" column="ARTNAME" ormtype="string"  
        indexable="true" indexStore="true";  
    property name="description" ormtype="text"  
        indexable="true";  
    property name="price" ormtype="double"  
        indexable="true";  
    property name="largeimage" ormtype="string";  
    property name="mediaid" ormtype="int";  
    property name="issold" ormtype="boolean";  
  
    property name="artist" fieldtype="many-to-one"  
        cfc="artist";  
}
```

Indexes

Creation and Update

- Automatic on all entity operations
- Build
- Lucene is fast

Maintenance

- Changes to properties indexed
- Functions provided act on:
 - whole index
 - entities
 - entity
- `ormIndex()` and `ormIndexPurge()`

Pitfalls?

What are the downsides?

Pitfalls

- Enterprise only
- Single server (at the moment)

What can we do now?

Add intelligent search to our apps

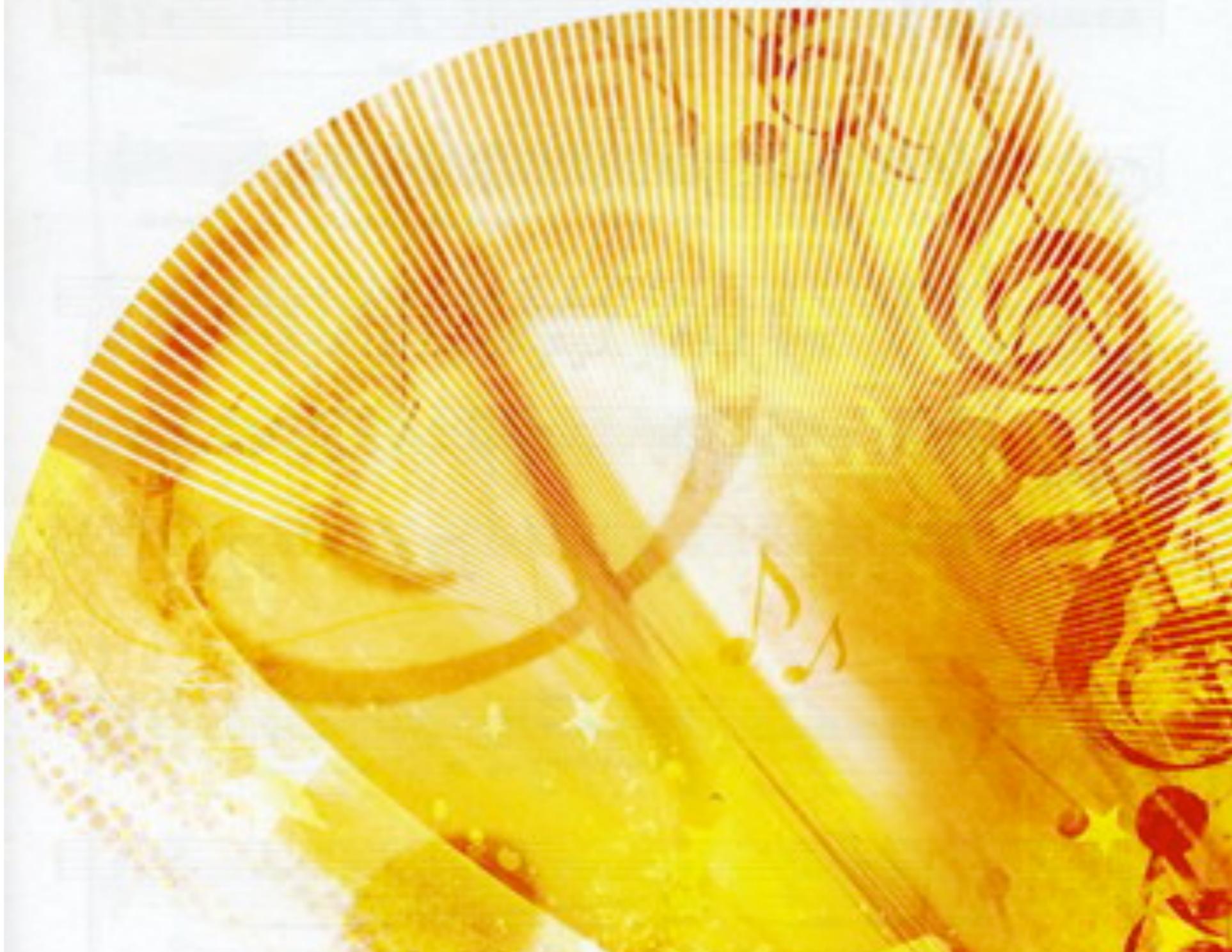
Guess vs Guide

- Make your own search algorithm
- Use Boost, Fuzzy and others
- Refine

SSA/PIANO
NOV941006

I [REDACTED] FOUND
WHAT I'M LOOKING FOR

WORDS & MUSIC BY U2



Thank you!

Q&A

@sam_farmer
sam@samfarmer.com
www.samfarmer.com