Freebasing for Fun and Enhancement

Sean Hannan

The Sheridan Libraries

Johns Hopkins University

Freebase

- "Freebase is a vast, free, open online database of structured knowledge, powered and maintained by Metaweb Technologies“
- Data is stored as nodes and relationships

Freebase APIs

- It’s a REST interface, so use whatever you want
- Accessible through http://api.freebase.com/api/service/mqlread?query={"query" : MQLQUERYSTRING }
- (/api/service/mqlwrite also exists to write to freebase…and there’s a bunch of other ones too)

Freebase APIs

- Query language is MQL
- Returns plaintext JSON
- Errors are even neatly serialized as JSON
- JSON JSON JSON

MQL

- JSON based
- Looks like…

```
[
 {
   "date_of_birth" : null,
   "name" : "Venetian Snares",
   "type" : "/people/person"
 }
]
```

Let’s make something

- How about an embeddable snippet of code that links Academy Awards data to entries in our catalog?
- Pain in the ass to make by hand
- Even more of a pain in the ass to remember to update it every year
- So, let’s grab the titles of best picture winners.

Clone the App

http://code4lib.mrdys.user.dev.freebaseapps.com/

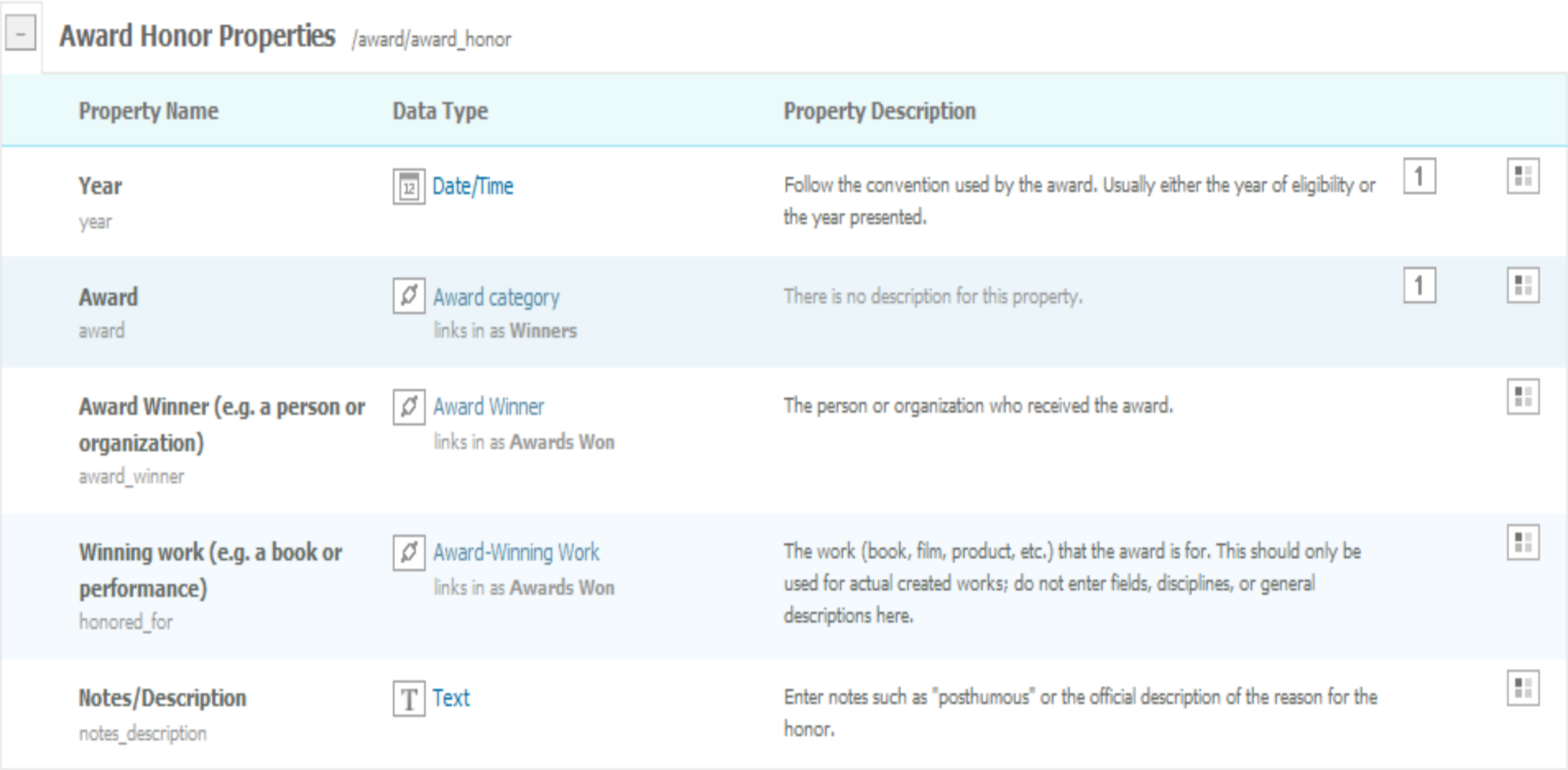
In SQL…

- You’d want something like “SELECT title, year FROM Awards WHERE award = “Best Picture”;
- In MQL this looks like…

MQL

```
[
    {
      "award" : "Academy Award for Best Picture",
      "honored_for" : null,
      "id" : null,
      "sort" : "-year",
      "type" : "/award/award_honor",
      "year" : null
    }
  ]
```

- How do we know what “fields”? Take a look at the /award/award_honor schema:
- http://www.freebase.com/type/schema/award/award_honor



Award Honor Properties /award/award_honor

Property Name	Data Type	Property Description	
Year year	Date/Time	Follow the convention used by the award. Usually either the year of eligibility or the year presented.	1
Award award	Award category links in as **Winners**	There is no description for this property.	1
Award Winner (e.g. a person or organization) award_winner	Award Winner links in as **Awards Won**	The person or organization who received the award.	
Winning work (e.g. a book or performance) honored_for	Award-Winning Work links in as **Awards Won**	The work (book, film, product, etc.) that the award is for. This should only be used for actual created works; do not enter fields, disciplines, or general descriptions here.	
Notes/Description notes_description	Text	Enter notes such as "posthumous" or the official description of the reason for the honor.	

Results

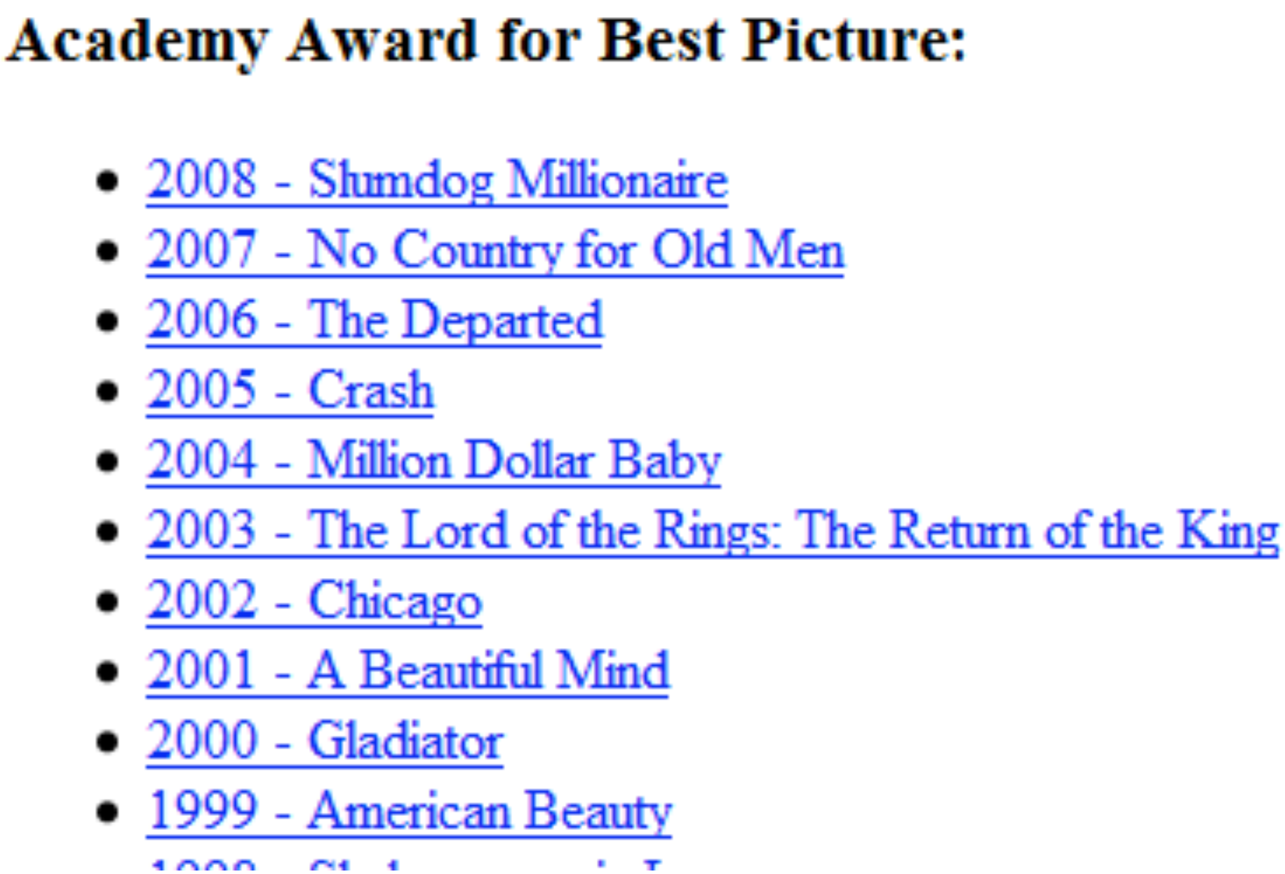
```
{
  "award" : "Academy Award for Best Picture",
  "honored_for" : "Slumdog Millionaire",
  "id" : "/guid/
  9202a8c04000641f800000000af060fb",
  "type" : "/award/award_honor",
  "year" : "2008"
 }, …
```

Templating and Linking w/ Acre

```
<acre:script>
      var query = acre.require("query").query;
      var awards = acre.freebase.MqlRead(query).result;
    </acre:script>

    <h3>${awards[0].award}:</h3>
    <ul>
      <li acre:for="award in awards">
        <a href="https://catalog.library.jhu.edu/ipac20/ipac.jsp?
 menu=search&npp=30&ipp=20&spp=20&aspect=power&profile=ge
 neral&ri=&limitbox_=&index=.VT&term=$
 {escape(award.honored_for)}">$award.year -
 $award.honored_for</a>
      </li>
    </ul>
```

Gets you…



Academy Award for Best Picture:

- 2008 - Slumdog Millionaire
- 2007 - No Country for Old Men
- 2006 - The Departed
- 2005 - Crash
- 2004 - Million Dollar Baby
- 2003 - The Lord of the Rings: The Return of the King
- 2002 - Chicago
- 2001 - A Beautiful Mind
- 2000 - Gladiator
- 1999 - American Beauty

Let's get more information…

```
[
   {
    "award" : "Academy Award for Best Picture",
    "honored_for" : {
     "*" : null,
     "type" : "/film/film"},
    "id" : null,
    "sort" : "-year",
    "type" : "/award/award_honor",
    "year" : null
   }
  ]
```

Subqueries!

- Just open up the JSON block for that ‘field’ and do more selects/filtering to get what you need for that datatype
- I don’t know what’s in there, so let’s wildcard it

Wildcards!

- “*” : null will get you *everything* (for that context)
- Great for debugging, but be careful
- Everything looks like…

```
…
 "prequel" : [],
 "produced_by" : [
       "Christian Colson"
 ],
 "production_companies" : [
       "Fox Searchlight Pictures",
       "Warner Bros."
  ],
  "rating" : [],
…
```

Templating again…

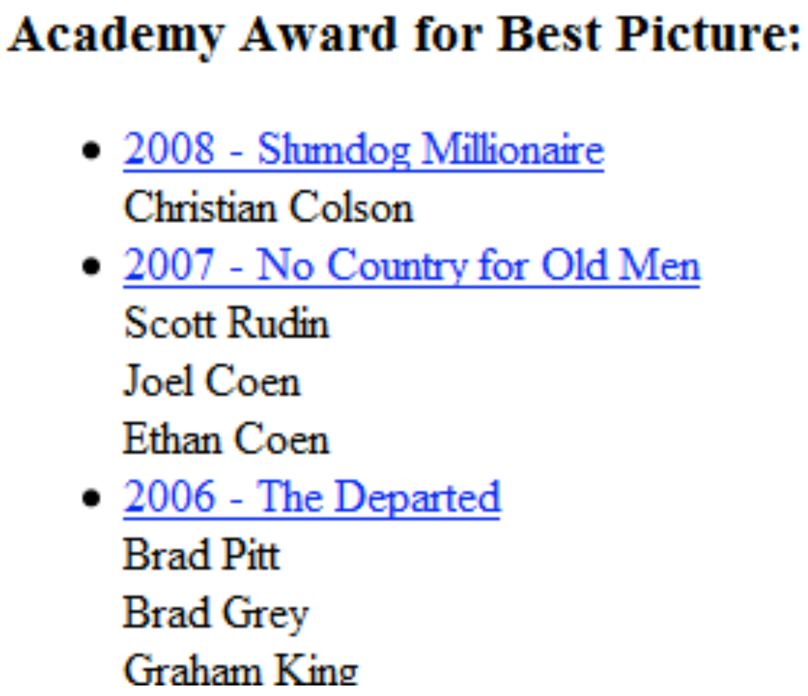
```
<acre:script>
        var query = acre.require("query2").query;
        var awards = acre.freebase.MqlRead(query).result;
      </acre:script>

      <h3>${awards[0].award}:</h3>
      <ul>
        <li acre:for="award in awards">
          <a href="https://catalog.library.jhu.edu/ipac20/ipac.jsp?
  menu=search&npp=30&ipp=20&spp=20&aspect=power&profile=general&ri=&limi
  tbox_=&index=.VT&term=${escape(award.honored_for.name)}">$award.year -
  $award.honored_for.name</a>
        <div acre:for="producers in award.honored_for.produced_by">
          $producers
        </div>

        </li>
      </ul>
```

Attribution is fun.

- Producers are awarded in the case of Best Picture, so give them some love:



Academy Award for Best Picture:

- 2008 - Slumdog Millionaire
  Christian Colson
- 2007 - No Country for Old Men
  Scott Rudin
  Joel Coen
  Ethan Coen
- 2006 - The Departed
  Brad Pitt
  Brad Grey
  Graham King

What about the rest of the categories?

```
[
    {
      "presenting_organization" : "Academy of Motion Picture Arts and
    Sciences",
      "type" : "/award/award_category",
      "name" : null,
      "winners" : [{
        "type" : "/award/award_honor",
        "id" : null,
        "year" : null,
        "/type/reflect/any_master" : [{
          "name" : null,
          "type" : "/film/film",
        }],
        "sort" : "-year"
      }]
    }
  ]
```

Reflection!

- Because the subquery gets information out of the /award/award_honor type, you don’t have all of the fun information associated with the /film/film type.
- “/type/reflect/any_master” will give you all of the objects that link to this particular point in the graph (/award/award_honor/winner in this case)
- Sort of like SQL JOINs (sort of)

Results!

```
…
"name" : "Academy Award for Best Short Film - Live Action - 2 Reels",
  "presenting_organization" : "Academy of Motion Picture Arts and Sciences",
  "type" : "/award/award_category",
  "winners" : [
   {
    "/type/reflect/any_master" : [
     {
      "name" : "The Bespoke Overcoat",
      "type" : "/film/film"
     }
    ],
    "id" : "/guid/9202a8c04000641f80000000091c205b",
    "type" : "/award/award_honor",
    "year" : "1956"
   },
…
```

One more pass!

```
[
      {
       "presenting_organization" : "Academy of Motion Picture Arts and Sciences",
       "type" : "/award/award_category",
       "name" : {
        "value" : null,
        "value|=" : ["Academy Award for Best Director", "Academy Award for Best Actor", "Academy Award for Best
     Picture", "Academy Award for Best Actress", "Academy Award for Writing Original Screenplay"],
        "lang" : "/lang/en"
       },
       "winners" : [{
        "type" : "/award/award_honor",
        "id" : null,
        "year" : null,
        "awardee:/type/reflect/any_master" : [{
         "name" : null,
         "type" : "/award/award_winner",
         "optional" : "optional"
        }],
        "title:/type/reflect/any_master" : [{
         "name" : null,
         "type" : "/film/film"
        }],
        "sort" : "-year"
       }],
       "sort" : "name.value"
      }
     ]
```

New things!

- "value|=" : ["Academy Award for Best Director", "Academy Award for Best Actor", "Academy Award for Best Picture", "Academy Award for Best Actress", "Academy Award for Writing Original Screenplay"],
- “value” can be any of these possible strings.

More new things!

-
  ```
  "awardee:/type/reflect/any_master" : [{
          "name" : null,
          "type" : "/award/award_winner",
          "optional" : "optional"
        }],
  ```
- You can reflect on specific properties!
- “optional” means that it’s cool if it doesn’t return anything for this select

Last of the new things!

- "sort" : "-year"

  ```
  }],
  "sort" : "name.value“
  ```

- Sorting within subqueries, and then sorting within the main query
- Sort properties can be further addressed, based on what is returned.

All Gussied up

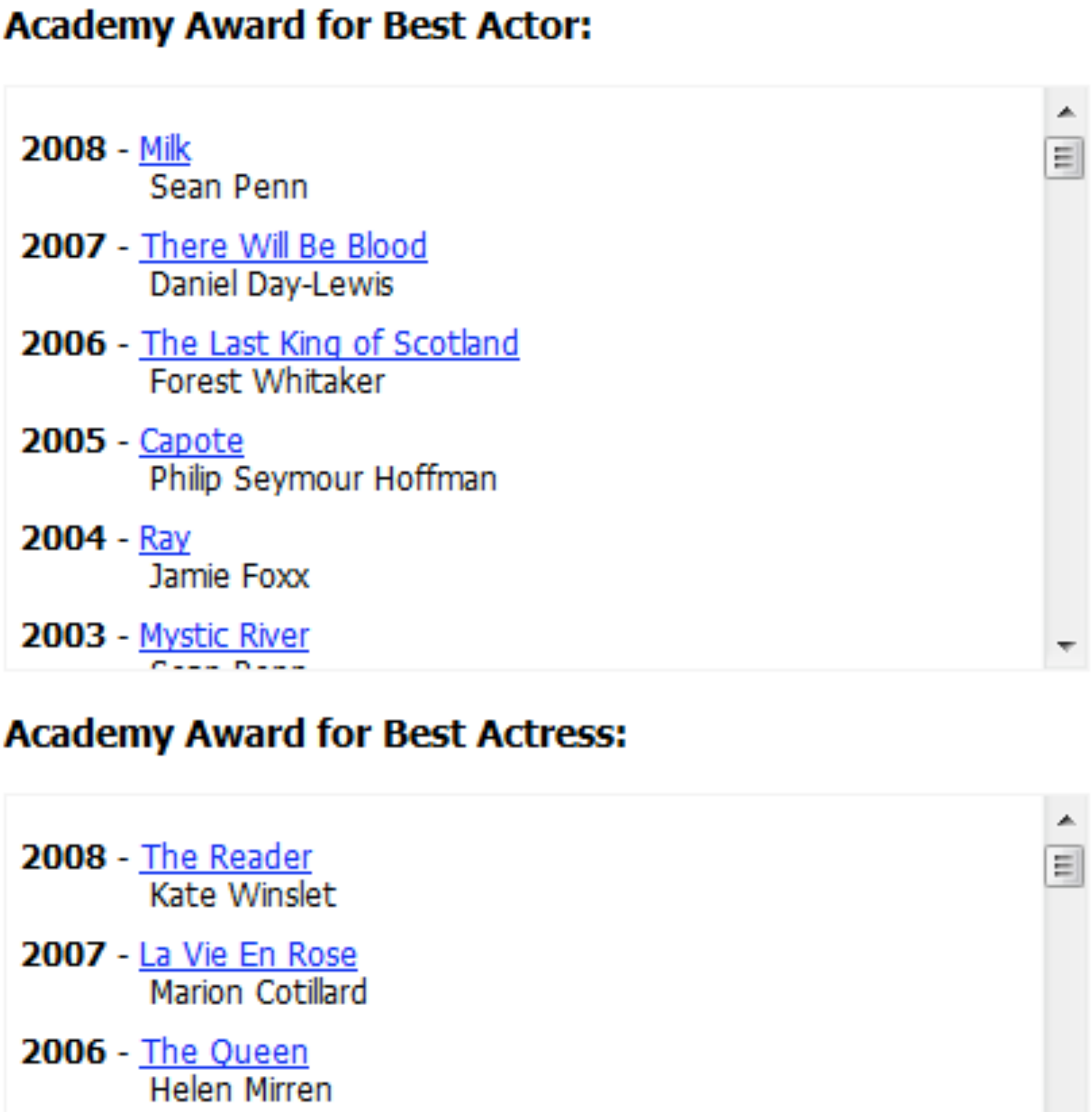
```
<html>

        <head>
          <link rel="stylesheet" type="text/css" href="css"/>

          <script type="text/javascript" src="http://jqueryjs.googlecode.com/files/jquery-1.2.6.min.js"></script>
          <script type="text/javascript">
            $(document).ready(function()
              {
              //hide the all of the element with class msg_body
              $(".scroller").hide();
            //toggle the componenet with class msg_body
            $("h3").click(function()
              {
              $(this).next(".scroller").slideToggle(200);
            });
            });
          </script>
        </head>
        <body>

          <acre:script>
            var query = acre.require("query4").query;
            var awards = acre.freebase.MqlRead(query).result;
          </acre:script>


          <div acre:for="category in awards">
          <h3>$category.name.value:</h3>
          <div class="scroller">
          <ul>
            <li acre:for="winner in category.winners">
              <b>$winner.year</b> - <span acre:for="title in winner['title:/type/reflect/any_master']"><a href="https://catalog.library.jhu.edu/ipac20/ipac.jsp?
       menu=search&npp=30&ipp=20&spp=20&aspect=power&profile=general&ri=&limitbox_=&index=.VT&term=${escape(title.name)}">${title.name}</a> </span>
              <div class="awardee" acre:for="awardee in winner['awardee:/type/reflect/any_master']">
                ${awardee.name}
              </div>
            </li><br />
          </ul>
          </div>
          </div>
        </body>
       </html>
```



Academy Award for Best Actor:

2008 - Milk
Sean Penn

2007 - There Will Be Blood
Daniel Day-Lewis

2006 - The Last King of Scotland
Forest Whitaker

2005 - Capote
Philip Seymour Hoffman

2004 - Ray
Jamie Foxx

2003 - Mystic River
Sean Penn

Academy Award for Best Actress:

2008 - The Reader
Kate Winslet

2007 - La Vie En Rose
Marion Cotillard

2006 - The Queen
Helen Mirren