

Recuperación de Información

Implementación de un Sistema de
Recuperación de Información
utilizando Lucene

Alberto Jesús Durán López
Miguel Ángel Cantarero López
Simón López Vico

INDEXACIÓN

Se indexan los siguientes campos:

- Movie Number (LongPoint)
- Release Year (IntPoint)
- Title (TextField)
- Origin/Ethnicity (TextField)
- Director (TextField)
- Cast (TextField)
- Genre (TextField)
- Wiki Page (StringField)
- Plot (TextField)
- ALL, todos los campos concatenados (TextField)

BÚSQUEDA

Tipos de búsquedas realizadas:

- IndexBooleanQuery
- IndexBooleanQueryDicGen
- IndexBooleanQueryMulti (sobre varios campos a la vez)
- IndexYearQuery e IndexRangeYearQuery
- IndexQueryParser
- IndexPhraseQuery

FACETAS

Facetas implementadas (pueden implementarse el resto fácilmente)

- Origin/Ethnicity
- Director (MultiValued)
- Genre (Multivalued)

Todas las búsquedas realizadas pueden mostrar facetas

TERMINAL

```
simon@simon-SATELLITE-L50-B:~/Escritorio/Dropbox/RI/p3/code/IRsystem$ ./IRsystem.sh terminal
Introduzca el campo sobre el que realizar la busqueda:
[0] -> Por defecto, sobre todos los campos
[1] -> Introducir manualmente

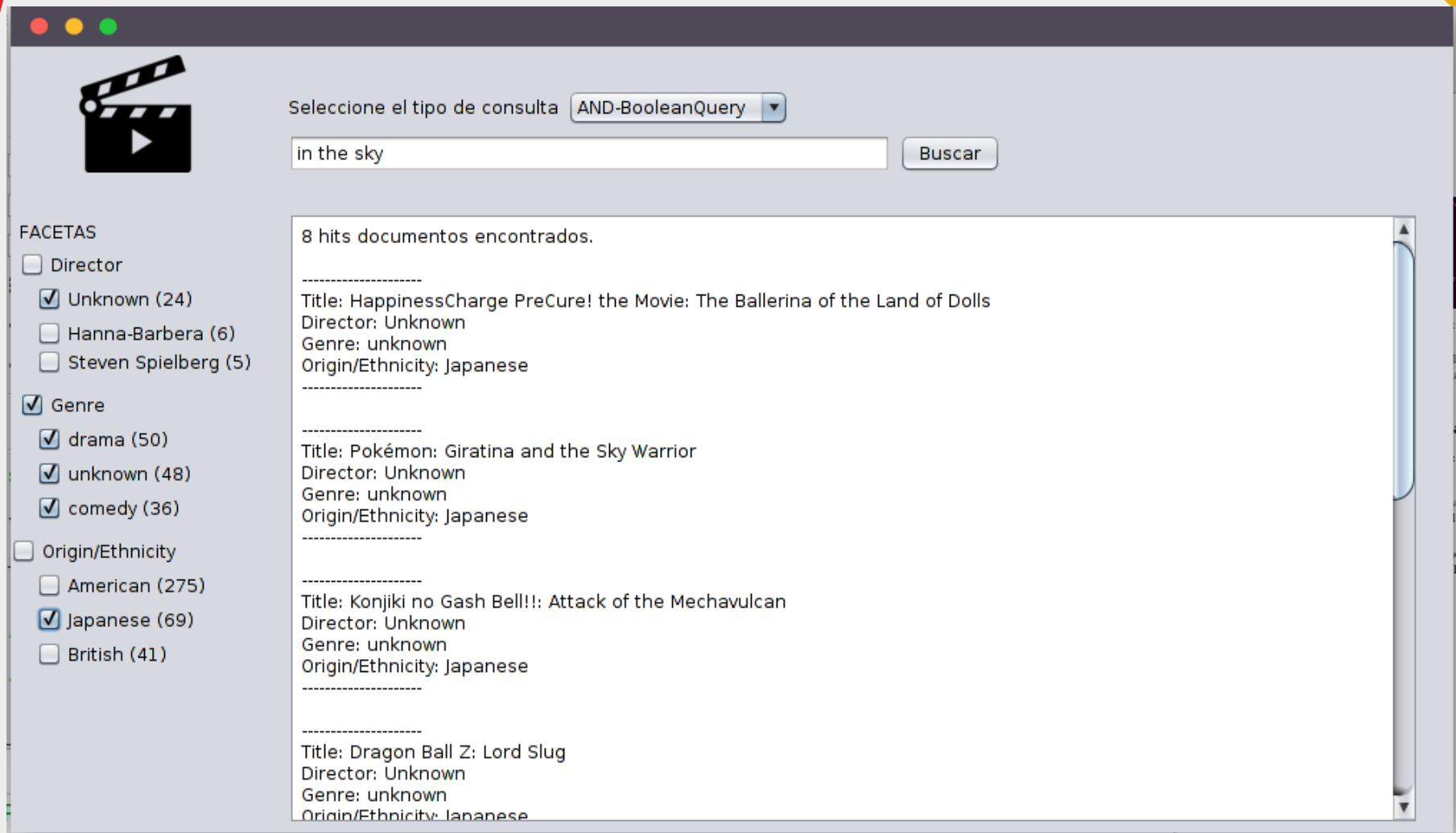
0
Introduzca el tipo de consulta que desea realizar:
[0] -> AND-BooleanQuery
[1] -> PhraseQuery
[2] -> PalabraQuery

0
Introduzca las palabras que deban contener los documentos:
comedy steven spielberg
BooleanQuery - Hay 4 hits documentos que contengan esas palabras
¿Desea mostrar las categorías? (Facetas) si/no
no
¿Desea mostrar los documentos relacionados?
si
-----
Movie Number: 14939
Release Year: 2006
Title: The Groomsmen
Genre: comedy
Director: Edward Burns
Plot: A groom and his four groomsmen wrestle with issues such as fatherhood, homosexuality, homophobia, and
Paulie (Burns), a self-supporting writer, is making [...]
-----

-----
Movie Number: 14568
Release Year: 2004
Title: The Terminal
Genre: comedy
Director: Steven Spielberg
Plot: Viktor Navorski, a traveler from the nation of Krakozhia, arrives at New York's John F. Kennedy International
and finds that his passport is suddenly no longer valid. The United States no longer [...]
-----

-----
Movie Number: 14019
Release Year: 2002
Title: Catch Me If You Can
Genre: comedy, drama
Director: Steven Spielberg
Plot: In 1963, teenaged Frank Abagnale lives in New Rochelle, New York with his father Frank Abagnale Jr. who is on
business loan at Chase Manhattan Bank due to u [...]
-----
```

INTERFAZ GRÁFICA



Selecione el tipo de consulta AND-BooleanQuery

in the sky Buscar

FACETAS

- ☐ Director
 - ☒ Unknown (24)
 - ☐ Hanna-Barbera (6)
 - ☐ Steven Spielberg (5)
- ☒ Genre
 - ☒ drama (50)
 - ☒ unknown (48)
 - ☒ comedy (36)
- ☐ Origin/Ethnicity
 - ☐ American (275)
 - ☒ Japanese (69)
 - ☐ British (41)

8 hits documentos encontrados.

Title: HappinessCharge PreCure! the Movie: The Ballerina of the Land of Dolls
Director: Unknown
Genre: unknown
Origin/Ethnicity: Japanese

Title: Pokémon: Giratina and the Sky Warrior
Director: Unknown
Genre: unknown
Origin/Ethnicity: Japanese

Title: Konjiki no Gash Bell!!!: Attack of the Mechavulcan
Director: Unknown
Genre: unknown
Origin/Ethnicity: Japanese

Title: Dragon Ball Z: Lord Slug
Director: Unknown
Genre: unknown
Origin/Ethnicity: Japanese