

# Exploring highly interconnected humanities data: are faceted browsers always the answer?

Michele Pasin

Information Architect

Nature Publishing Group

*michele.pasin@nature.com*

KING'S  
College  
LONDON

npg nature publishing group

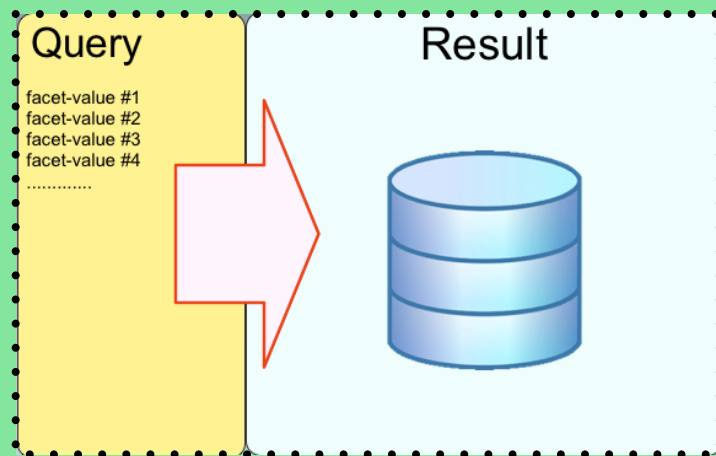
## Outline

1. **Background.** Interaction models in search interfaces: retrieval model vs explorational model
2. **Use-case.** DJFacet, an app targeting multi-result, highly structured DH datasets
3. **Evaluation & Conclusions.** Are faceted browsers always the answer?

# **1 - Background**

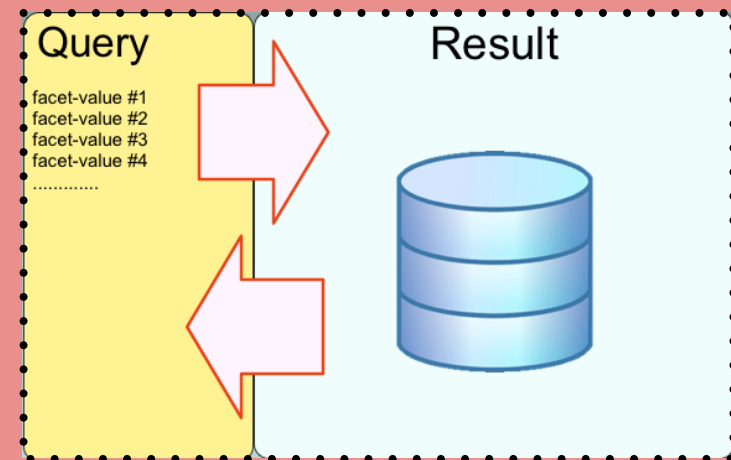
# Two models of interaction

## Retrieval Model



VS

## Exploration model



**DYNAMIC TAXONOMIES AND FACETED SEARCH, Theory, Practice and Experience.** Giovanni Maria Sacco, Yannis Tzitzikas, *Springer-Verlag*, 2009 (Chapter 1, 'The Model').

# Retrieval model: structured search

Query

**Google scholar** **Advanced Scholar Search**

**Find articles**

with **all** of the words  
with the **exact phrase**  
with **at least one** of the words  
**without** the words  
where my words occur

**Author** Return articles written by

**Publication** Return articles published in

**Date** Return articles published between

**Collections**

anywhere in the article

e.g., "PJ Hayes" or McCarthy

e.g., J Biol Chem or Nature

e.g., 1996

## Articles and patents

☒ Search articles in all subject areas (☒ include patents).

☐ Search only articles in the following subject areas:

- ☐ Biology, Life Sciences, and Environmental Science
- ☐ Business, Administration, Finance, and Economics
- ☐ Chemistry and Materials Science
- ☐ Engineering, Computer Science, and Mathematics

**Scholar** Articles and patents anytime include citations Create email alert

### [Implementing faceted classification for software reuse](#)

R Prieto-Diaz - Communications of the ACM, 1991 - portal.acm.org  
LJbrao/Concepts A fundamental problem in software reuse is organizing collections of reusable components for effective search and retrieval. The concept of **faceted** classification introduced in [3, 4] presents a partial solution to this problem. **Faceted** classification offers certain ...  
[Cited by 602](#) - [Related articles](#) - [All 7 versions](#)

### [Multi-faceted regulation of cell differentiation by extracellular matrix](#)

CQ Lin... - The FASEB journal, 1993 - FASEB  
Page 1. Multi-faceted regulation of cell differentiation by ... models could provide powerful tools not only for understanding regulation of normal cell function but also for the studies of tumorigenesis and possibly cancer therapy.-Lin, C. Q, Bissell, M. J. Multi-faceted regulation ...  
[Cited by 419](#) - [Related articles](#) - [BI, Direct](#) - [All 5 versions](#)

### [Faceted metadata for image search and browsing](#)

KP Yee, K Swearingen, K Li... - Proceedings of the SIGCHI ..., 2003 - portal.acm.org  
ABSTRACT There are currently two dominant interface types for searching and browsing large image collections: keyword-based search, and searching by overall similarity to sample images. We present an alternative based on enabling users to navigate along conceptual ...  
[Cited by 583](#) - [Related articles](#) - [BI, Direct](#) - [All 45 versions](#)

### [Case study research: a multi-faceted research approach for IS](#)

ALM Cavaye - Information Systems Journal, 1996 - Wiley Online Library  
... Case study research: a multi-faceted research approach for IS. ALM Cavaye  
online: 28 JUN 2008. ... How to Cite. Cavaye, A. (1996). Case study research approach for IS. Information Systems Journal, 6: 227-242. ...  
[Cited by 340](#) - [Related articles](#) - [All 2 versions](#)

[Clustered versus faceted categories for information exploration](#)

Result

# Retrieval model: guided search

Query

YAHOO! DIRECTORY

☐ Web | ☒ Directory | ☐ Category

## Astronomy Pictures

[Email this page](#)

[Directory](#) > [Science](#) > [Astronomy](#) > [Pictures](#)

### CATEGORIES [\(What's This?\)](#)

- [Astrophotography@](#)
- [Comet Hale-Bopp@](#)
- [Companies@](#)
- [Galaxies@](#)
- [Hubble Space Telescope@](#)

- [Messier Objects@](#)
- [NASA@](#)
- [Planets@](#)
- [Sun@](#)

### SITE LISTINGS [By Popularity](#) | [Alphabetical](#) | [\(What's This?\)](#)

- [Astronomy Picture of the Day \(APOD\)](#)  
Discover the cosmos! Each day there is a different image or photograph of our fascinating universe, along with explanation written by a professional astronomer.  
[antwrp.gsfc.nasa.gov/apod/astropix.html](http://antwrp.gsfc.nasa.gov/apod/astropix.html)
- [NASA Solar System Simulator](#)  
Simulated view of any body in the solar system from any location and time in full-color graphics. A NASA/JPL site.  
[space.jpl.nasa.gov](http://space.jpl.nasa.gov)
- [NASA's Planetary Photojournal](#)  
Images of the planets, the Sun, asteroids, and comets. NASA's Photojournal provides easy access to images from NASA Solar System exploration missions.  
[photojournal.jpl.nasa.gov](http://photojournal.jpl.nasa.gov)
- [Great Images in NASA \(GRIN\)](#) Pick! [\[read review\]](#)  
Collection of the photographs capturing the history of NASA's space program: human exploration, robotic missions, experimental aircrafts, planetary investigations, and deep space studies.  
[grin.hq.nasa.gov](http://grin.hq.nasa.gov)
- [NASA Image eXchange \(NIX\)](#)  
Collection of NASA aircraft, astronaut, spacecraft, space center, and other images.  
[nix.nasa.gov](http://nix.nasa.gov)
- [NASA Images](#)  
See images of the Universe, the Solar System, and the Earth and our planet. NASA Images is a service of Internet Archive, a nonprofit library, to offer free access to these collections.

Result

# Exploration model: faceted search

**amazon.co.uk** Hello Mikele Pasin. We have [recommendations](#) for you. ([Not Mikele?](#))  
Mikele's Amazon.co.uk | [Today's Deals](#) | [Gift Cards](#) | [Gifts & Wish Lists](#)

[Shop All Departments](#) Search  camera

Welcome International Gift Certificates Deals of the Week

**Department**  
**Any Department**  
Electronics & Photo (168,954)  
Books (105,833)  
Computers & Accessories (8,644)  
MP3 Downloads (7,535)  
Kitchen & Home (7,181)  
Toys & Games (2,557)  
Music (1,185)  
[+ See All 30 Departments](#)


**Delivery Option** ([What's this?](#))  
**Any Delivery Option**  
FREE Super Saver Delivery

**Digital Cameras**


**Brand**  
**Any Brand**  
Canon (295)  
Casio (116)  
Fujifilm (282)  
Kodak (302)  
Nikon (163)  
Panasonic (216)  
Samsung (216)  
Sony (269)

**"camera"**  
**Related Searches:** [digital camera](#), [pink camera](#), [video camera](#).  
Showing 1 - 16 of 306,101 Results

**Browse Our Latest Camera Offers**  
• **Deal of the Week:** Kodak EasyShare M530 only **£59.99** | [Toshiba](#)  
• **Offers:** Free case with Flip Camcorders – starts at **£49.99**  
• **Price deals:** Creative Vado 8GB only **£54.98** - [see more](#)

1.  **Sony DSCW320B Cyber-shot Digital Camera**  
Buy new: ~~£114.49~~ **£97.45**  
[31 new](#) from **£97.45** [2 used](#) from **£98.00**  
Get it by **Wednesday, Feb 2** if you order in the next 10 hours  
★★★★☆ (72)  
Eligible for **FREE** Super Saver Delivery.  
**Electronics & Photo:** See all 168,954 items

2.  **Electronics Store**  
Digital Cameras  
Point & Shoot Digital Cameras  
Camera Batteries, Chargers & Power Supplies  
Camera & Camcorder Batteries

3.  **Fujifilm FinePix S1600 Digital Camera**

Result

Query



# Exploration model: faceted search

Result

## EARLY MODERN LONDON THEATRES

Home Introduction **Search the Database** Workspace Learning Zone Project Information REED Team Help

KEYWORD SEARCH **BROWSE** RECORD

Active Filters (1) [Reset Search](#)

**Event & Person** >

Event date > 1607 ✕

Choose more filters in order  
narrow down your query further

TROUPE & VENUE [CLICK TO OPEN](#) ?

EVENT & PERSON [CLICK TO OPEN](#) ?

DOCUMENT: DESCRIPTION [CLICK TO OPEN](#) ?

CITATION [CLICK TO OPEN](#) ?

Query

### Matching \*Events\* Items (16)

Listing items 1 to 16, page 1 of 1

SUMMARY	VENUE(S)	TROUPE(S)	PEOPLE	DATE HAPPENED
Dekker places the player John Bentley among a company of playwrights	Bel Savage; Bell; Bull; Theatre	Queen's Men	Atchelow, Thomas; Bentley, John; Dekker, Thomas; Greene, Robert; Kyd, Thomas; Marlowe, Christopher; Nashe, Thomas; Peele, George; Watson, Thomas	undated
A play about the Sophy of Persia is performed at the Red Bull	Blackfriars (II); Curtain; Red Bull	Children of the Queen's Revels (1604-c1614); Queen Anne's Men	Beaumont, Francis	From: circa 1607
The interior of the Curtain is round	Curtain	...	Day, John; Rowley, William; Wilkins, George	From: 1607
Queen Anne's Men perform 'Sir Thomas	Boar's Head; Curtain; Red	Queen Anne's Men	Dekker, Thomas; Webster, John	From: before 1607



# The Retrieval model explained

## Facet 1

value 1-1  
value 1-2  
value 1-3  
value 1-4  
.....

## Facet 2

value 2-1  
value 2-2  
value 2-3  
value 2-4  
.....

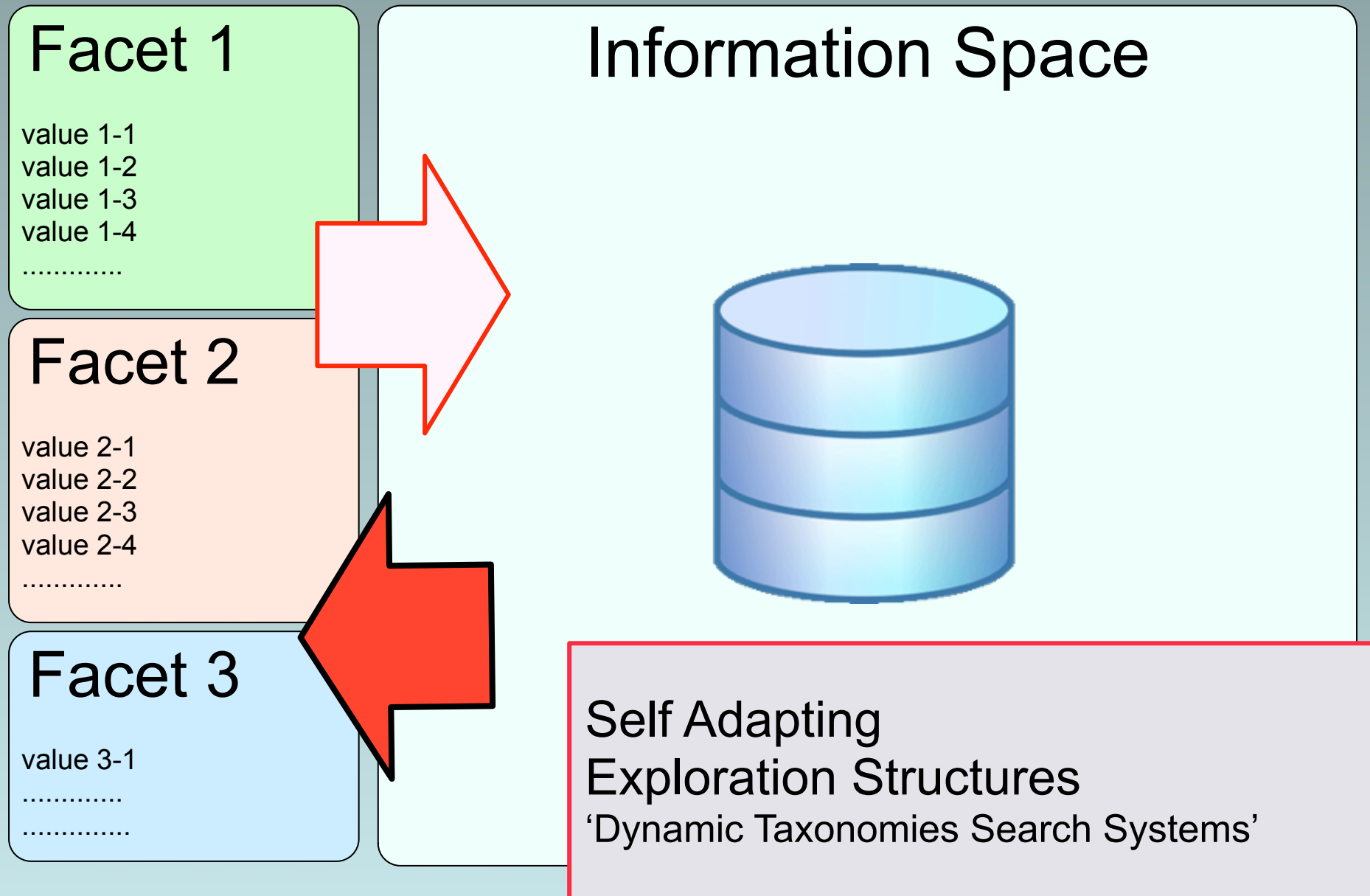
## Facet 3

value 3-1  
.....  
.....

## Information Space



# The Exploration model explained



# Faceted Browsers: key features

- Widely **tested**, many implementations

- Hearst (2002) “*Finding the flow in web site search*”
- /facet (2006); Exhibit (2007); Humboldt (2008); Collex (2007);

- Implement a ‘**schema-less**’ approach

- Nowivskie: allow to “explore lateral relationships” and “possibilities for algorithmic serendipity in research”

- **Easy** to use, **user-centered**

- support both experts and non-experts
- expose domain features

- Highly **scalable** / **convergent**

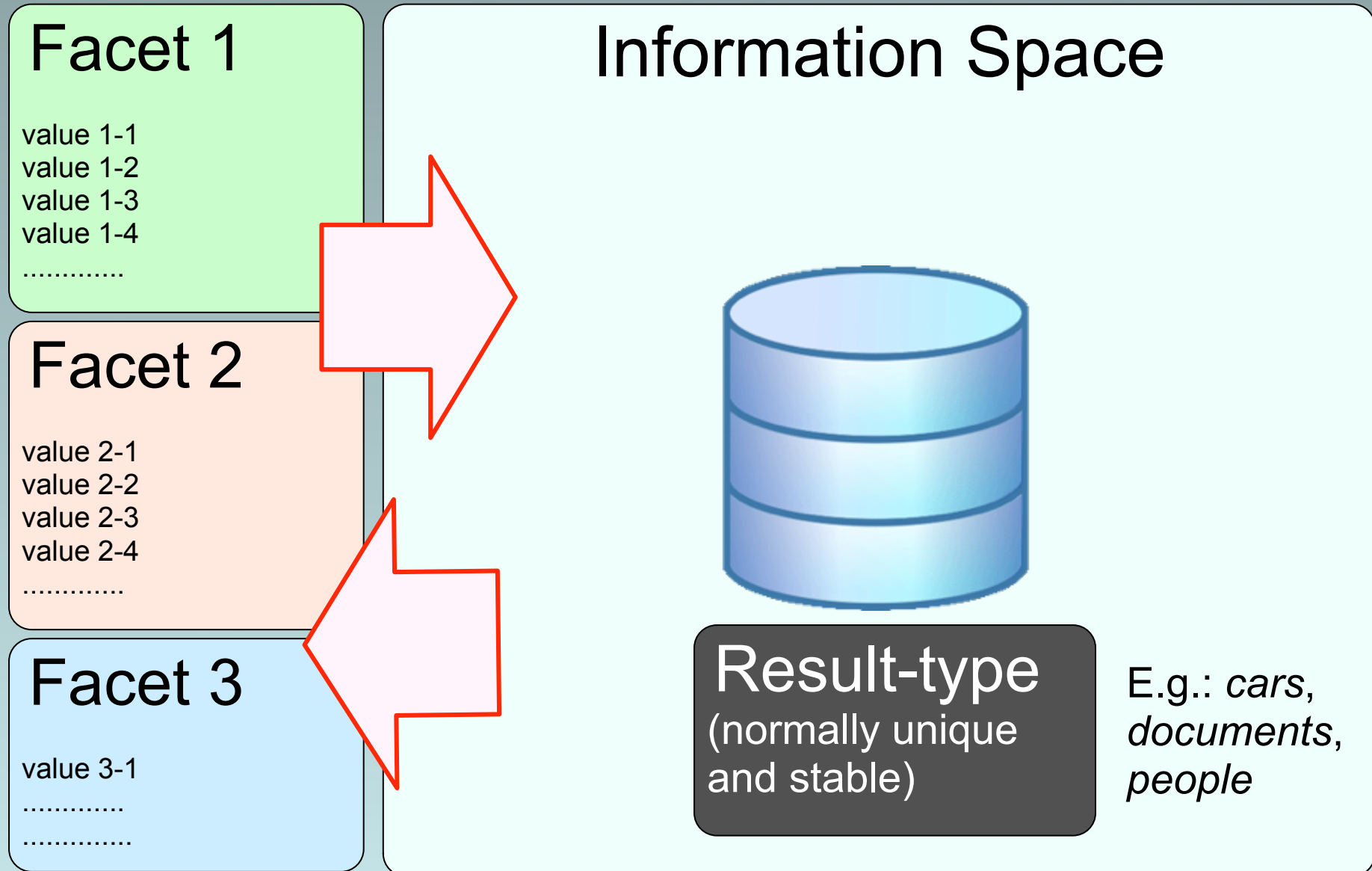
- bottom-up classification / prevent inconclusive searches

- Allow for a ‘**relaxed**’ faceted classification

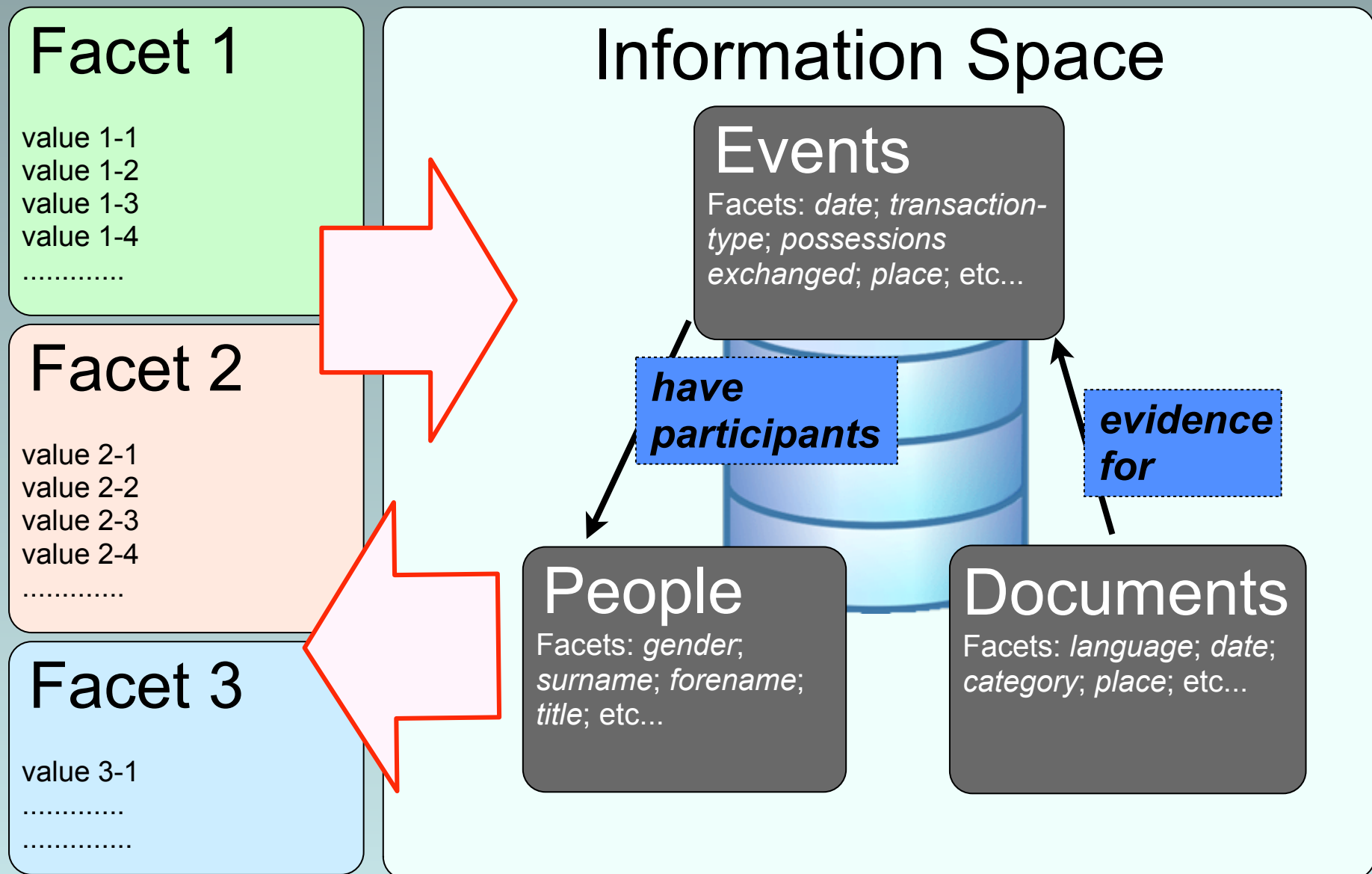
- faceted classifications **vs** dynamic taxonomies search systems
- available structured data are often enough to bootstrap a FB

## **2 - New Directions**

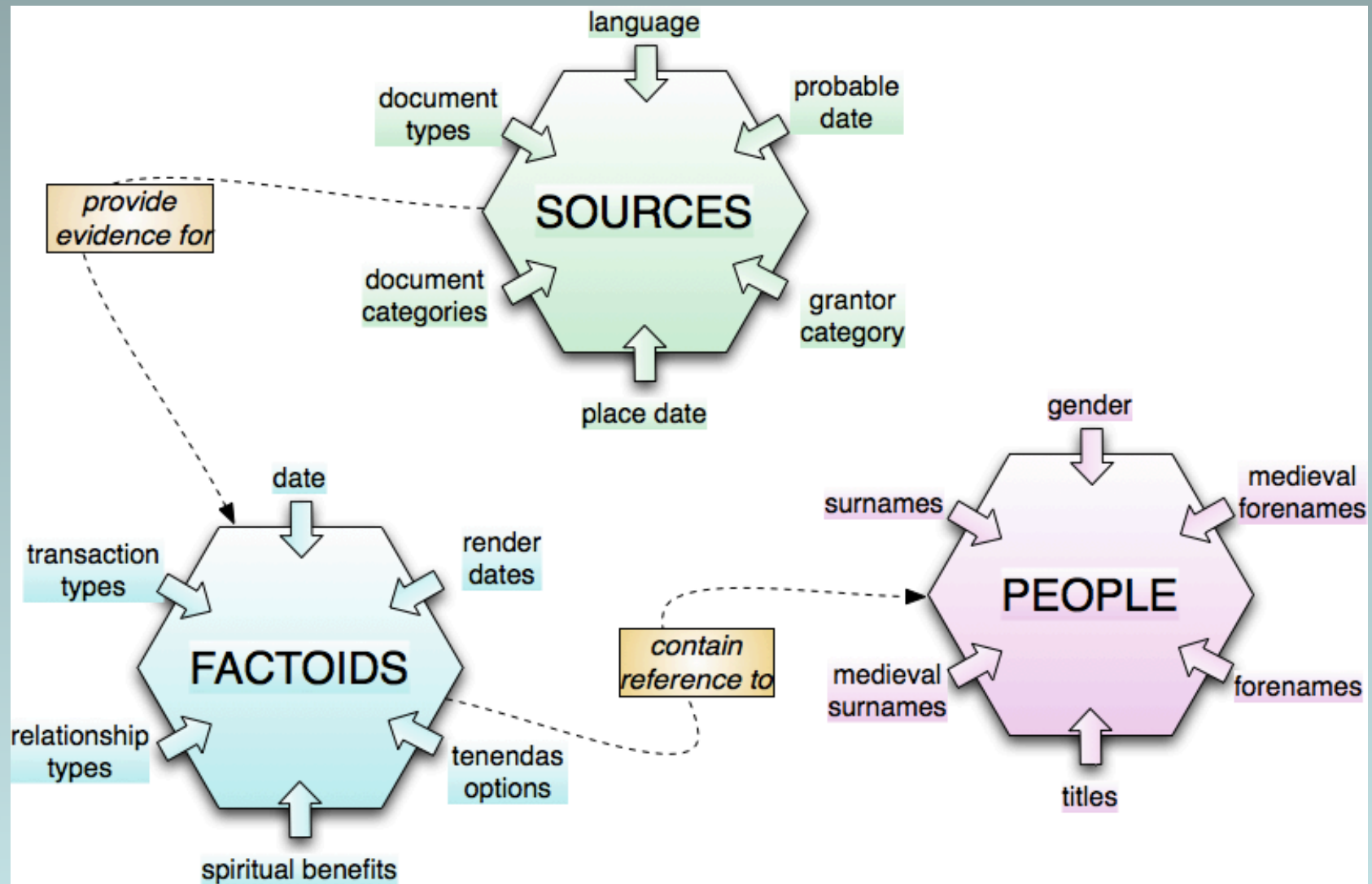
# Extending the model: multiple result types



# Extending the model: multiple result types



# Data Model and facets in POMS





# Django Facet: a Python multi-result FBS

The screenshot shows the Django Facet web application. At the top is the logo 'DJango FACET' with a bar chart icon. Below it is a navigation bar with links: 'Project home', 'Search home', 'Online Docs', and 'Source Code'. The main search area shows 'Searching for: COUNTRIES | Religions' with a current filter of 'Christian (Religion)'. On the left, there are two facet panels. The 'PLACE FACETS' panel has a 'Region name' section with options: Africa (3), Europe (2), Middle East (3), North America (1), and South America (1). It also has a 'Region idbname' section with options: Asia (excluding near east) (1), Latin america and the caribbean (1), Near east (2), Northern africa (2), and Northern america (1), followed by a 'more...' link. The 'RELIGION FACETS' panel has a 'Religion name' section with options: Muslim (7) and Other (6). On the right, a 'Matching' table is partially visible with columns 'ID' and 'N'. The table lists several IDs: 194, 205, 139, 193, 196, 225, 79, 215, 96, and 214. Below the table is a 'Listing items' section.

- Python/**Django** based
- **Easy** to install / integrate
- Back-end **agnostic**
- **Minimal** look and feel
- **REST** architecture
- Supports **pivoting**

And more: caching system, custom facets, hierarchical facets etc..

<http://www.michelepasin.org/software/>

# Django Facet: a Python multi-result FBS

```
facetslist = [  
    {'appearance' : {  
        'label' : 'Person name' ,  
        'uniquename' : 'personname',  
        'model' : Person ,  
        'dbfield' : "name",  
        'displayfield' : "name",  
        'grouping' : ['personinfo'],  
    } ,  
    'behaviour' : [{  
        'resulttype' : 'persons',  
        'querypath' : 'name',  
    } ,  
    {  
        'resulttype' : 'events',  
        'querypath' : 'associatedpeople__name',  
    } ,  
    {  
        'resulttype' : 'documents',  
        'querypath' : 'associatedfactoids__associatedpeople__name',  
    } ,  
    ]},  
]
```

# Case studies: EMLOT <[www.emlot.kcl.ac.uk](http://www.emlot.kcl.ac.uk)>

The screenshot displays the EMLOT website interface. At the top, the title "EARLY MODERN LONDON THEATRES" is prominently displayed. Below this, a navigation bar includes links for Home, Introduction, Search the Database, Workspace, and Learning Zone. A secondary bar offers search methods: KEYWORD SEARCH, FACETED SEARCH (highlighted), and RECORD. The main search area features a "Search for:" input field and a "Showing All \*Events\*" status indicator. On the left, a sidebar contains radio button filters for Transcription Records, Primary Sources, Secondary Sources, People (selected), Events, Troupes, and Venues. A red-bordered box highlights the "Search for:" input field and the "People" filter. On the right, a detailed view of a selected document is shown, with a black-bordered box highlighting its facets. These facets include "DOCUMENT: DESCRIPTION", "CITATION", "EVENT & PERSON", and "TROUPE & VENUE", each with a "CLICK TO OPEN" link and a help icon. The "EVENT & PERSON" facet is expanded, showing fields for Event Type, Event Date, Person Surname, Person Event Role, and Person Date. The "TROUPE & VENUE" facet shows fields for Troupe Name, Troupe Type, Troupe Country of Origin, and Troupe Date.

EARLY MODERN LONDON THEATRES

Home Introduction Search the Database Workspace Learning Zone

KEYWORD SEARCH FACETED SEARCH RECORD

Search for: Showing All \*Events\*

☐ Transcription Records  
☐ Primary Sources  
☐ Secondary Sources  
☒ People  
☐ Events  
☐ Troupes  
☐ Venues

Filter by more filters

DOCUMENT: DESCRIPTION [CLICK TO OPEN](#) ?

CITATION [CLICK TO OPEN](#) ?

EVENT & PERSON [CLICK TO OPEN](#) ?

- + EVENT TYPE ?
- + EVENT DATE
- + PERSON SURNAME ?
- + PERSON EVENT ROLE ?
- + PERSON DATE

TROUPE & VENUE [CLICK TO OPEN](#) ?

- + TROUPE NAME ?
- + TROUPE TYPE ?
- + TROUPE COUNTRY OF ORIGIN ?
- + TROUPE DATE

# Case studies: POMS <www.poms.ac.uk>

PEOPLE OF MEDIEVAL SCOTLAND  
1093-1314

Home About Database Information Schools News Contact Help

SEARCH BROWSE RECORD FAMILY TREES LABS

Search

Facets available:

HISTORY

PEOPLE AND INSTITUTIONS

SOURCES

RELATIONSHIPS

+ RELATIONSHIP TYPES ?

+ ROLES ?

+ SPIRITUAL BENEFITS ?

enter some text to filter

All Relatives (13)  
Ancestors (pro  
anima) (70)  
Ancestors (pro

Friend(s) (pro  
salute) (2)  
Husband (pro  
anima) (1)

FORENAME ?

A. (25)  
Aaron (1)  
Abel (1)  
Abraham (21)  
Absalom (7)


Factoids Sources People and Institutions

Listing items 1 to 50, page 1 of 5

# EMLOT: complex queries made simple

Active Filters (4)  [Reset Search](#)

**Event & Person** >

Event type > **Arrest** 


**Troupe & Venue** >

Troupe type > **Adult players** 

**Citation** >

Repositories > **London Metropolitan Archives,  
London, UK** 

**Document: description** >

Primary source date > **1613** 

 Matching *\*Transcription Records\** Items (2)

Listing items 1 to 2, page 1 of 1

<< First < Previous


SECONDARY SOURCE	AUTHOR(S)	PAGE(S) (TRANSCRIPTION RECORD)	PRIMARY SOURCE
English Professional Theatre, 1530-1660	Berry, Herbert; Ingram, William; Wickham, Glynne William Gladstone	569	LMA, MJ/SR/0519



# EMLOT: complex queries made simple

Active Filters (4)  [Reset Search](#)

Event & Person >

Event type > **Arrest** 


Troupe & Venue >

Troupe type > **Adult players** 

Citation >

Repositories > **London Metropolitan Archives, London, UK** 

Document: description >

Primary source date > **1613** 

## Matching \*People\* Items (3)

Listing items 1 to 3, page 1 of 1

SURNAME	FORENAME	TITLE
Collyns	Nicholas	Esquire
Fulsis	Alexander	...

## Matching \*Events\* Items (1)

Listing items 1 to 1, page 1 of 1

SUMMARY	VENUE(S)	TROUPE(S)	PEOPLE
A purse is snatched at the Red Bull	Red Bull	Queen Anne's Men	Collyns, Nicholas; Fuls Robert

## Matching \*Transcription Record

Listing items 1 to 2, page 1 of 1

SECONDARY SOURCE	AUTHOR(S)	PAGE(S)
English Professional Theatre, 1530-1660	Berry, Herbert; Ingram, William; Wickham, Glynne William Gladstone	569

# EMLOT: complex queries.. simple?

**Facet:**  
**Place of publication:**

**"London"**

**&&**

**Facet:**  
**Venue Name:**

**"Phoenix/  
Cockpit"**

**&&**

**Facet:**  
**Person Role:**

**"Playwright"**

**&&**

**Facet:**  
**Troupe type:**

**"Adult players"**

Tr. records

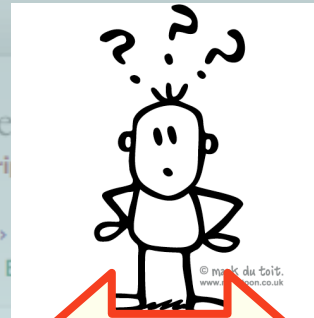
Events

Sources

People

Troupes

Venues





# EMLOT: complex queries.. simple?

**Facet:**  
**Place of publication:**

**"London"**

**&&**

**Facet:**  
**Venue Name:**

**"Phoenix/  
Cockpit"**

**&&**

**Facet:**  
**Person Role:**

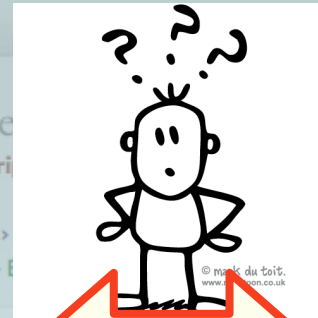
**"Playwright"**

**&&**

**Facet:**  
**Troupe type:**

**"Adult players"**

Events happened at the *Phoenix* that involved *Adult Players* and *Playwrights*, and were reported in Documents published in *London*...



Tr. records

Events

Sources

People

Troupes

Venues

# EMLOT: complex queries.. simple?

**Facet:**  
**Place of publication:**

**"London"**

**&&**

**Facet:**  
**Venue Name:**

**"Phoenix/  
Cockpit"**

**&&**

**Facet:**  
**Person Role:**

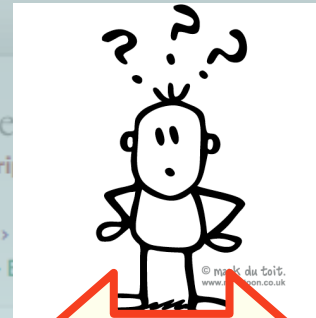
**"Playwright"**

**&&**

**Facet:**  
**Troupe type:**

**"Adult players"**

Events happened at the **Phoenix** that involved **Adult Players** and **Playwrights**, and were reported in Documents published in **London**...



Events documented in sources published in **London**, that are about the **Phoenix** and about people who were simultaneously **Adult Players** and **playwrights**...

Tr. records

Events

Sources

People

Troupes

Venues

## **3 - Evaluation & Conclusions**

# Evaluation

## - Purpose:

- improving the general efficiency of DJFacet
- testing the intuitiveness of the search and navigation facilities;
- testing the comprehension of the specific facets we are using
- testing the comprehension of the 'multi-result' approach

## - Setup:

- 8 people
- face to face sessions of 30-60 minutes
- recorded using screen-casting software
- the performance is analysed and annotated afterwards

## - Tasks:

- incremental difficulty
- level 0: warming up, exploring the interface (facets and result types)
- level 1: queries with 1 facet
- level 2: queries involving 2 facets
- level 3: queries involving 3 or more facets

## Evaluation results

<b>Comprehension of the intended meaning of facets</b>	<ul style="list-style-type: none"><li>- In general, quite positive</li><li>- Document-class and document-type are very ambiguous</li><li>- Some of the terms within the facets are not easy to interpret: eg the 'staging context' event-type.</li></ul>
<b>Generic UI issues</b>	<ul style="list-style-type: none"><li>- Facets' role in a search is more intuitive when they are open</li><li>- Clear separation between controls and results</li><li>- Result-type switches are not obvious, people confuse them with "other" facet controls</li></ul>
<b>Comprehension of the significance of results</b>	<ul style="list-style-type: none"><li>- Pivoting action is not explained properly</li><li>- People with no familiarity with the domain don't get the implicit relations between result-types</li><li>- People with familiarity with the domain perform quite well</li></ul>

## Evaluation results: future work

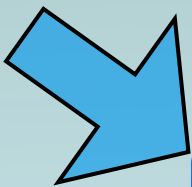
- Cues that help users **understand** the **DB model**:
  - **static** section in the help menu
  - **dynamic** 'query explanation' mechanism
    - via graphical diagram providing a visual representation of the query
    - via a natural language rendering of the query
- Messages that help users notice the '**pivoting**' action:
  - popups or messages before changing result-type
  - make this control less prominent when filters are already selected

## Conclusions

- Are FB useful in the DH?
  - ..definitely YES!
  - a plethora of reasons:
    - **easy to use, scalable, transparent, algorithmic serendipity** etc..
- But.... careful with:
  - **rich domain models** (and we want to reveal this richness)
    - fb may quickly become complicated (eg lots of search facets)
  - **more than one result-type**
    - need to provide ways to tackle ambiguity
  - **domain-specific questions:**
    - other types of *entry points* to the datasets will be more effective (eg interactive visualizations)



... thanks!



[www.michelepasin.org/software/djfacet](http://www.michelepasin.org/software/djfacet)

email me at: *michele.pasin@nature.com*