

Using ezDL as a search interface for patent searching

Matthias Jordan

Universität Duisburg-Essen

7.6.2011

UNIVERSITÄT
DUISBURG
ESSEN

Existing OS solutions for patent search

- ▶ Patent search is a complex task
 - ▶ Traditional search already complex
 - ▶ Special considerations for patent searching (e.g. completeness, accountability)
 - ▶ Patent applicants sometimes try to hide their documents
- ▶ Existing open source systems are either
 - ▶ search engines with bare-bones interfaces or
 - ▶ non-specific GUI toolkits

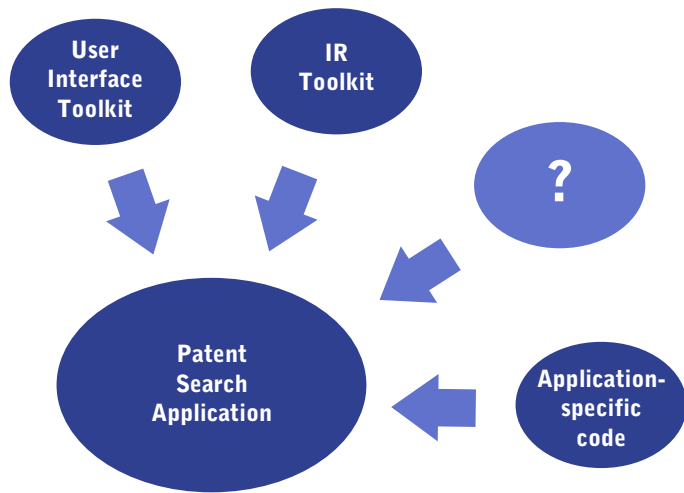
- ▶ Survey among patent professionals (Vanderbauwhede, 2010) showed need for support with
 - ▶ search result assessment
 - ▶ search result navigation
 - ▶ search management, organization and history

Important Patent Search Features

	Lemur	Lucene	MG4J	Terrier
Combining multiple search results	X	X	X	X
Keyword highlighting	X	X	X	X
Search history	X	X	X	X
Organising search queries	X	X	X	X
Timeliness of retrieved documents	X	X	X	X
Saving custom lists of results	X	X	X	X
User defined highlighting	X	X	X	X
Viewing full document image	X	X	X	X
Grouping by patent family	X	X	X	X
Patent family information	X	X	X	X

Based on Vanderbauwhede, 2010

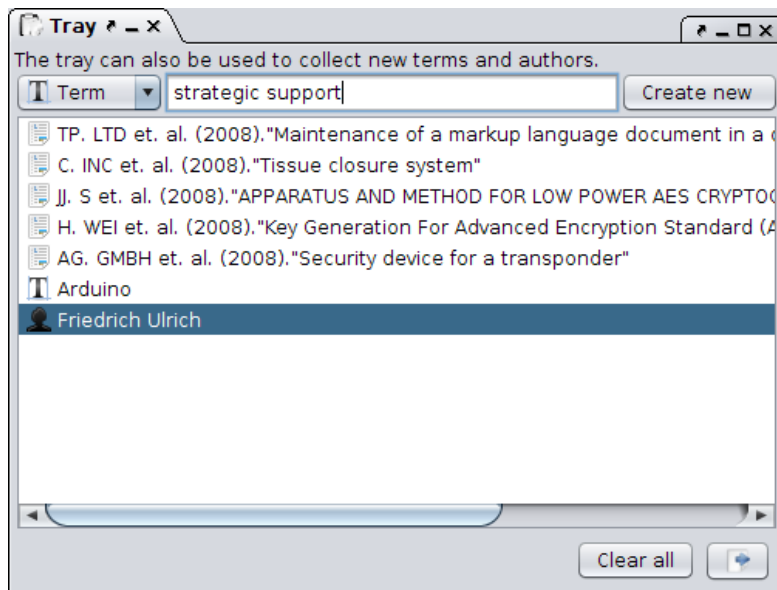
The Great Divide



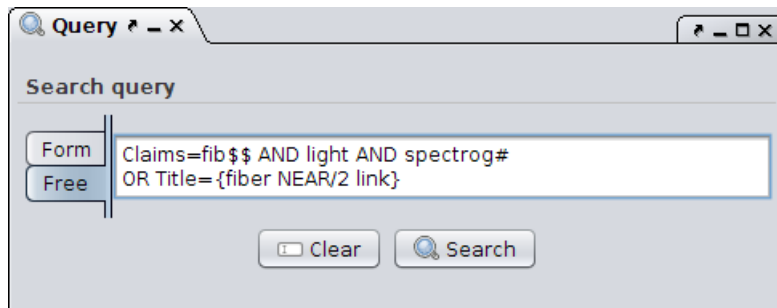
- ▶ DAFFODIL has been continually developed and evaluated since 2000
- ▶ Unified access
 - ▶ to heterogeneous data sources
 - ▶ using a common interface,
 - ▶ a common query syntax and
 - ▶ a combined result view
- ▶ 10 years of experience and evaluation
- ▶ Reimplementation of DAFFODIL system started in 2009
- ▶ *Not* a search engine, but all search engines can be plugged in



What ezDL brings to the table



What ezDL brings to the table



The screenshot shows a web-based search interface for ezDL. The window has a title bar with a magnifying glass icon and the text "Query". Below the title bar, there is a section labeled "Search query". On the left side of this section, there are two buttons: "Form" and "Free". The "Free" button is currently selected, and a text input field is active next to it. The text in the input field is "Claims=fib\$\$ AND light AND spectrog# OR Title={fiber NEAR/2 link}". Below the input field, there are two buttons: "Clear" and "Search". The "Search" button has a magnifying glass icon.

Query

Search query

Form

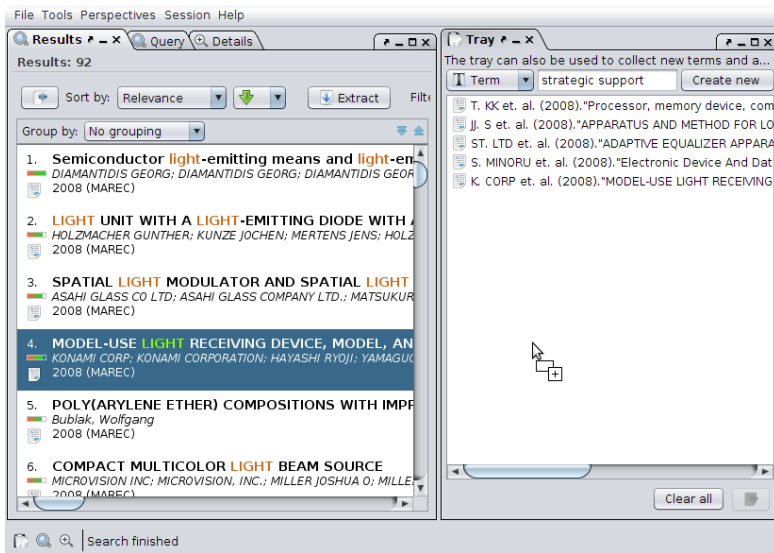
Free

Claims=fib\$\$ AND light AND spectrog#
OR Title={fiber NEAR/2 link}

Clear

Search

What ezDL brings to the table



What ezDL brings to the table

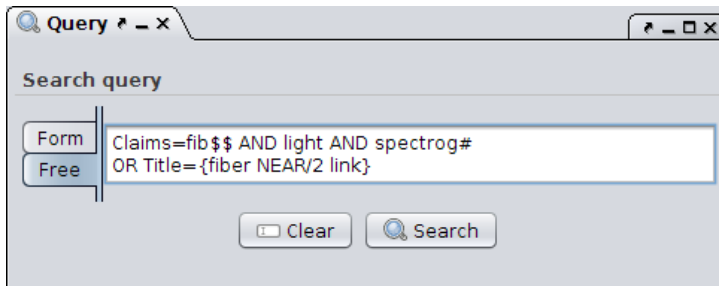


EzDL for patent search

- ▶ Integrated features supporting patent search
- ▶ Showcase for ezDL's extensibility
- ▶ Searches on MAREC collection using SOLR: 10 hours
- ▶ Displaying Patent documents: 12 hours
- ▶ Extracting, grouping: 1 hour

Presented features

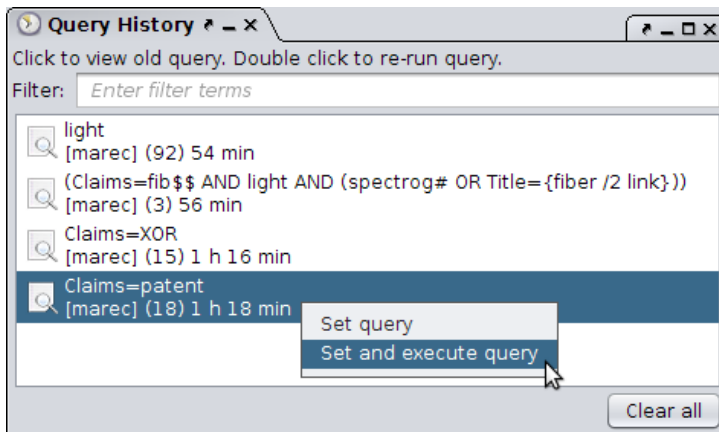
- ▶ Rich query syntax
 - ▶ Boolean queries with fields operators
 - ▶ Phrases and proximity operators, year ranges
 - ▶ Wildcards and truncation operators



The screenshot shows a web browser window titled "Query". Inside the window, there is a section labeled "Search query". Below this label, there are two buttons: "Form" and "Free". To the right of these buttons is a text input field containing the query: "Claims=fib\$\$ AND light AND spectrog# OR Title={fiber NEAR/2 link}". Below the input field, there are two buttons: "Clear" and "Search".

Presented features

- ▶ Query support
 - ▶ Full query history
 - ▶ Filtering
 - ▶ Re-running and combining



The demo

The screenshot displays the ezDL patent search interface with the following components:

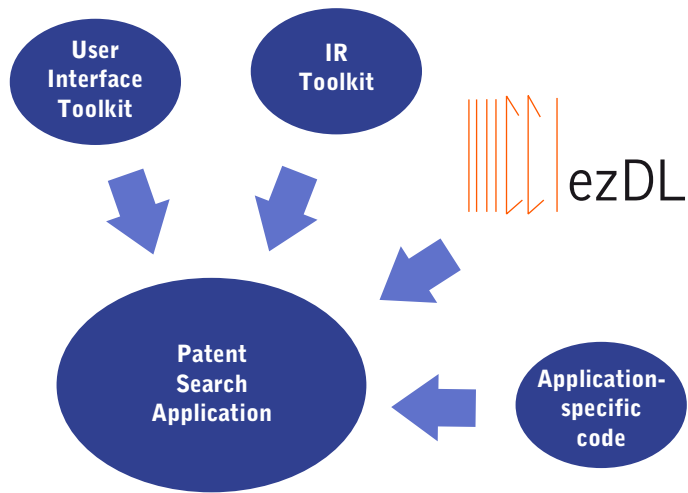
- Query Panel:** Search query: Claims: light. Filter: [Form/Free]. Buttons: Clear, Search.
- Results Panel:** Results: 93. Sort by: Relevance. Filter: Enter filter terms. Group by: By author.
- Query History Panel:** Click to view old query. Double click to re-run query. Filter: Enter filter terms. List of queries: Claims=light [marec] (93) 1 min, light [marec] (92) 1 h 19 min, Claims=lib\$4 AND light AND (spectro\$ OR Title=fr) [marec] (3) 1 h 20 min, Claims=XOR [marec] (15) 1 h 39 min, Claims=patent [marec] (18) 1 h 41 min. Button: Clear all.
- Tray Panel:** The tray can also be used to collect new terms and author... Term: strategic support. Button: Create new.
- Details Panel:** Information recording/reproducing device and information reproducing method. Year: 2008. Inventors: SASAKI, TOSHIO (inventor), SASAKI, TOSHIO (inventor), Gultschew, Gintley, Stockmar & Schwann & Anwaltssozietät (Agent), FUJIFILM CORP (Applicant), FUJIFILM CORPORATION (Applicant). Patent Document: (19) EPO (European Patent Office) (11) EP 1 881 493 A1. (12) (43) Date of publication: 23.01.2008.

Important Patent Search Features Revisited

	Lem	Luc	MG4J	Ter	ezDL
Combining multiple search results	✗	✗	✗	✗	✓
Keyword highlighting	✗	✗	✗	✗	✓
Search history	✗	✗	✗	✗	✓
Organising search queries	✗	✗	✗	✗	✓
Timeliness of retrieved documents	✗	✗	✗	✗	✓
Saving custom lists of results	✗	✗	✗	✗	✓
User defined highlighting	✗	✗	✗	✗	✓
Viewing full document image	✗	✗	✗	✗	✓
Grouping by patent family	✗	✗	✗	✗	✓
Patent family information	✗	✗	✗	✗	✓

Based on Vanderbauwhede, 2010, with extension

The Great Divide – United



Conclusion

- ▶ Patent search is demanding
- ▶ Graphical tools are more helpful than command line tools
- ▶ EzDL is not a patent search application
- ▶ Demo shows features to support patent search
- ▶ More features easily implemented by leveraging existing framework

Architecture

