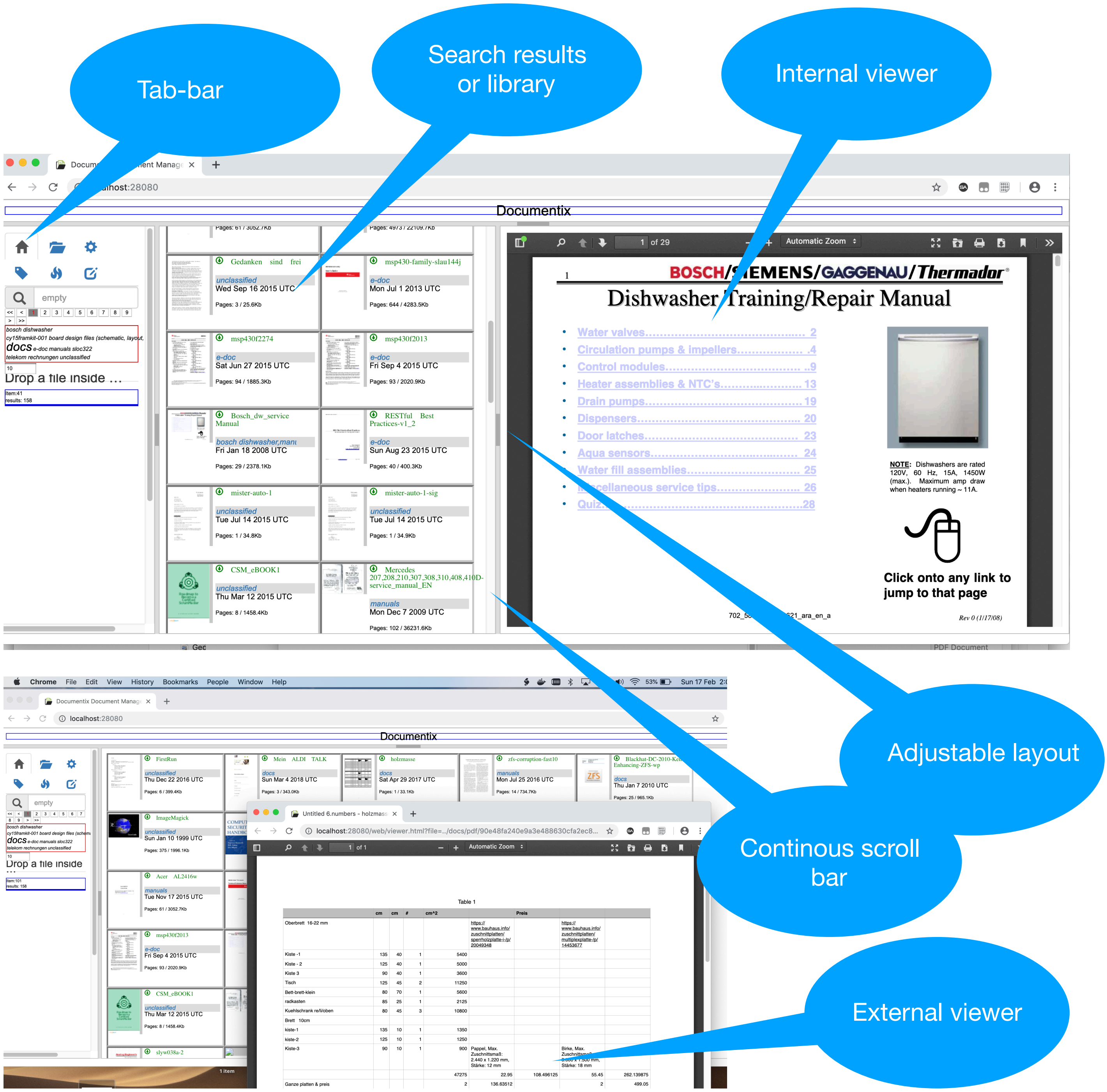
Documentix

the Document manager

Overview

Documentix quickly searches your collection using a full-text search engine
Documentix lets you classify documents
Documentix automatically classify new documents based on previous training
Documentix automatically OCR’s new documents to enable searching
Documentix manages multiple types of documents (pdf/word/ppt/jpeg..)
Documentix presents all archived documents through a web-interface
Documentix allows easy sharing of documents

User interface



Docum…ent Manage ×

…lhost:28080

Documentix

empty

<< < 1 2 3 4 5 6 7 8 9 > >>

bosch dishwasher
cy15framkit-001 board design files (schematic, layout,
docs *e-doc manuals sloc322*
telekom rechnungen unclassified

10

Drop a file inside …

Item:41
results: 158

Pages: 61 / 3052.7Kb

Pages: 4973 / 22109.7Kb

Gedanken sind frei
unclassified
Wed Sep 16 2015 UTC
Pages: 3 / 25.6Kb

msp430-family-slau144j
e-doc
Mon Jul 1 2013 UTC
Pages: 644 / 4283.5Kb

msp430f2274
e-doc
Sat Jun 27 2015 UTC
Pages: 94 / 1885.3Kb

msp430f2013
e-doc
Fri Sep 4 2015 UTC
Pages: 93 / 2020.9Kb

Bosch_dw_service Manual
bosch dishwasher,manu
Fri Jan 18 2008 UTC
Pages: 29 / 2378.1Kb

RESTful Best Practices-v1_2
e-doc
Sun Aug 23 2015 UTC
Pages: 40 / 400.3Kb

mister-auto-1
unclassified
Tue Jul 14 2015 UTC
Pages: 1 / 34.8Kb

mister-auto-1-sig
unclassified
Tue Jul 14 2015 UTC
Pages: 1 / 34.9Kb

CSM_eBOOK1
unclassified
Thu Mar 12 2015 UTC
Pages: 8 / 1458.4Kb

Mercedes 207,208,210,307,308,310,408,410D-service_manual_EN
manuals
Mon Dec 7 2009 UTC
Pages: 102 / 36231.6Kb

1 of 29 Automatic Zoom

1 BOSCH/SIEMENS/GAGGENAU/Thermador®

Dishwasher Training/Repair Manual

- Water valves.. 2
- Circulation pumps & impellers.................. .4
- Control modules....................................... ..9
- Heater assemblies & NTC’s....................... 13
- Drain pumps.. 19
- Dispensers.. 20
- Door latches... 23
- Aqua sensors.. 24
- Water fill assemblies............................... 25
- Miscellaneous service tips........................ 26
- Quiz...28

NOTE: Dishwashers are rated 120V, 60 Hz, 15A, 1450W (max.). Maximum amp draw when heaters running ~ 11A.

Click onto any link to jump to that page

702_5…621_ara_en_a

Rev 0 (1/17/08)

PDF Document

Chrome File Edit View History Bookmarks People Window Help 53% Sun 17 Feb 2:

Documentix Document Manage ×

localhost:28080

Documentix

FirstRun
unclassified
Thu Dec 22 2016 UTC
Pages: 6 / 399.4Kb

Mein ALDI TALK
docs
Sun Mar 4 2018 UTC
Pages: 3 / 343.0Kb

holzmasse
docs
Sat Apr 29 2017 UTC
Pages: 1 / 33.1Kb

zfs-corruption-fast10
manuals
Mon Jul 25 2016 UTC
Pages: 14 / 734.7Kb

Blackhat-DC-2010-Ke… Enhancing-ZFS-wp
docs
Thu Jan 7 2010 UTC
Pages: 25 / 965.1Kb

ImageMagick
unclassified
Sun Jan 10 1999 UTC
Pages: 375 / 1996.1Kb

Acer AL2416w
manuals
Tue Nov 17 2015 UTC
Pages: 61 / 3052.7Kb

msp430f2013
e-doc
Fri Sep 4 2015 UTC
Pages: 93 / 2020.9Kb

CSM_eBOOK1
unclassified
Thu Mar 12 2015 UTC
Pages: 8 / 1458.4Kb

slyw038a-2

Untitled 6.numbers - holzmass ×

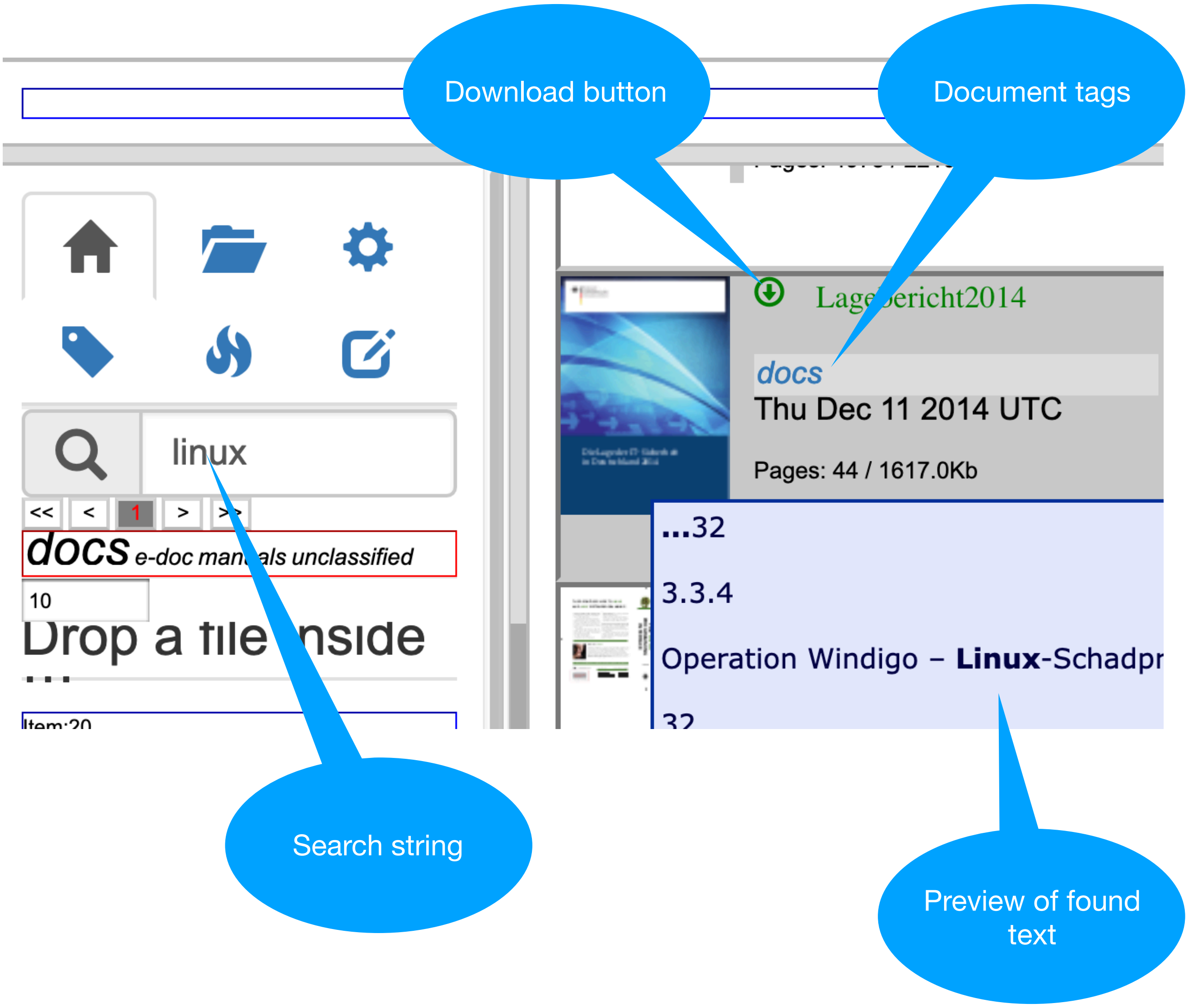
localhost:28080/web/viewer.html?file=../docs/pdf/90e48fa240e9a3e488630cfa2ec8...

1 of 1 Automatic Zoom

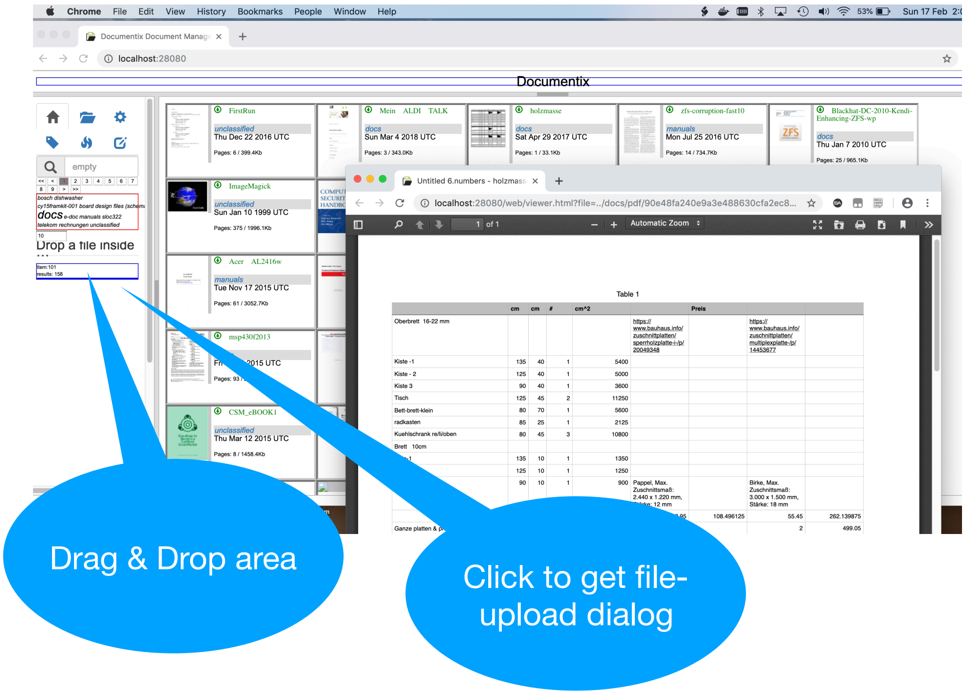
Table 1

	cm	cm	#	cm^2		Preis		
Oberbrett 16-22 mm					https://www.bauhaus.info/zuschnittplatten/sperrholzplatte-i-/p/20049348		https://www.bauhaus.info/zuschnittplatten/multiplexplatte-/p/14453677	
Kiste -1	135	40	1	5400				
Kiste - 2	125	40	1	5000				
Kiste 3	90	40	1	3600				
Tisch	125	45	2	11250				
Bett-brett-klein	80	70	1	5600				
radkasten	85	25	1	2125				
Kuehlschrank re/li/oben	80	45	3	10800				
Brett 10cm								
kiste-1	135	10	1	1350				
kiste-2	125	10	1	1250				
Kiste-3	90	10	1	900	Pappel, Max. Zuschnittsmaß: 2.440 x 1.220 mm, Stärke: 12 mm		Birke, Max. Zuschnittsmaß: 3.000 x 1.500 mm, Stärke: 18 mm	
				47275	22.95	108.496125	55.45	262.139875
Ganze platten & preis				2	136.63512		2	499.05

Search function



Upload new documents



Chrome File Edit View History Bookmarks People Window Help — 53% Sun 17 Feb

Documentix Document Manage

localhost:28080

Documentix

empty

bosch dishwasher cy15framkit-001 board design files (schema docs e-doc manuals sloc322 telekom rechnungen unclassified

10

Drop a file inside

Item:101
results: 158

FirstRun — unclassified — Thu Dec 22 2016 UTC — Pages: 6 / 399.4Kb

Mein ALDI TALK — docs — Sun Mar 4 2018 UTC — Pages: 3 / 343.0Kb

holzmasse — docs — Sat Apr 29 2017 UTC — Pages: 1 / 33.1Kb

zfs-corruption-fast10 — manuals — Mon Jul 25 2016 UTC — Pages: 14 / 734.7Kb

Blackhat-DC-2010-Kendi-Enhancing-ZFS-wp — docs — Thu Jan 7 2010 UTC — Pages: 25 / 965.1Kb

ImageMagick — unclassified — Sun Jan 10 1999 UTC — Pages: 375 / 1996.1Kb

Acer AL2416w — manuals — Tue Nov 17 2015 UTC — Pages: 61 / 3052.7Kb

msp430f2013 — Fri 2015 UTC — Pages: 93 /

CSM_eBOOK1 — unclassified — Thu Mar 12 2015 UTC — Pages: 8 / 1458.4Kb

Untitled 6.numbers - holzmass

localhost:28080/web/viewer.html?file=../docs/pdf/90e48fa240e9a3e488630cfa2ec8...

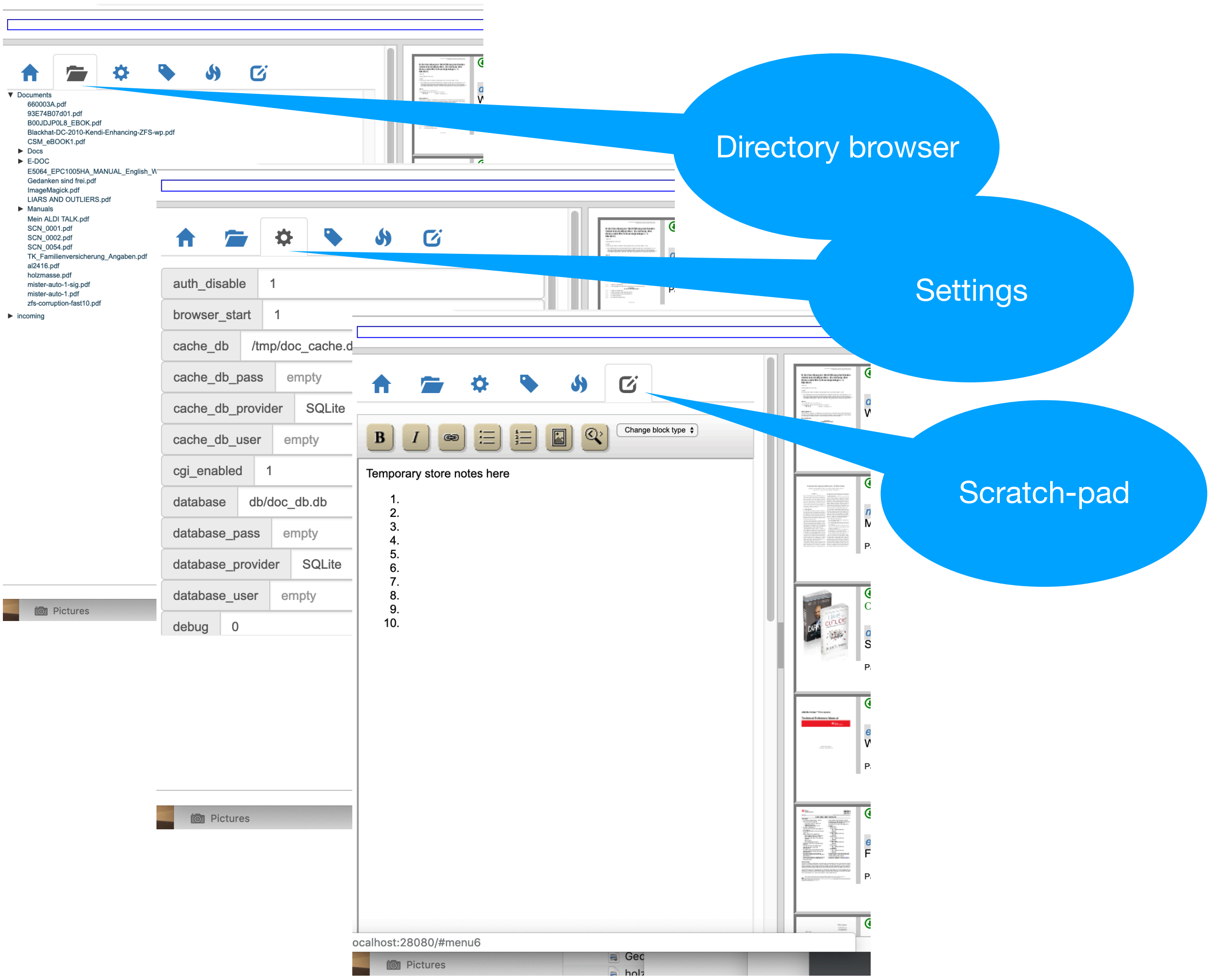
1 of 1 Automatic Zoom

Table 1

	cm	cm	#	cm^2		Preis		
Oberbrett 16-22 mm					https://www.bauhaus.info/zuschnittplatten/sperrholzplatte-i-/p/20049348		https://www.bauhaus.info/zuschnittplatten/multiplexplatte-/p/14453677	
Kiste -1	135	40	1	5400				
Kiste - 2	125	40	1	5000				
Kiste 3	90	40	1	3600				
Tisch	125	45	2	11250				
Bett-brett-klein	80	70	1	5600				
radkasten	85	25	1	2125				
Kuehlschrank re/li/oben	80	45	3	10800				
Brett 10cm								
-1	135	10	1	1350				
	125	10	1	1250				
	90	10	1	900	Pappel, Max. Zuschnittsmaß: 2.440 x 1.220 mm, Stärke: 12 mm		Birke, Max. Zuschnittsmaß: 3.000 x 1.500 mm, Stärke: 18 mm	
					.95	108.496125	55.45	262.139875
Ganze platten & pr							2	499.05

Drag & Drop area

Click to get file-upload dialog



▼ Documents
- 660003A.pdf
- 93E74B07d01.pdf
- B00JDJP0L8_EBOK.pdf
- Blackhat-DC-2010-Kendi-Enhancing-ZFS-wp.pdf
- CSM_eBOOK1.pdf
- ▶ Docs
- ▶ E-DOC
- E5064_EPC1005HA_MANUAL_English_W
- Gedanken sind frei.pdf
- ImageMagick.pdf
- LIARS AND OUTLIERS.pdf
- ▶ Manuals
- Mein ALDI TALK.pdf
- SCN_0001.pdf
- SCN_0002.pdf
- SCN_0054.pdf
- TK_Familienversicherung_Angaben.pdf
- al2416.pdf
- holzmasse.pdf
- mister-auto-1-sig.pdf
- mister-auto-1.pdf
- zfs-corruption-fast10.pdf

▶ incoming

auth_disable	1
browser_start	1
cache_db	/tmp/doc_cache.d
cache_db_pass	empty
cache_db_provider	SQLite
cache_db_user	empty
cgi_enabled	1
database	db/doc_db.db
database_pass	empty
database_provider	SQLite
database_user	empty
debug	0

Change block type

Temporary store notes here

1.
2.
3.
4.
5.
6.
7.
8.
9.
10.

ocalhost:28080/#menu6

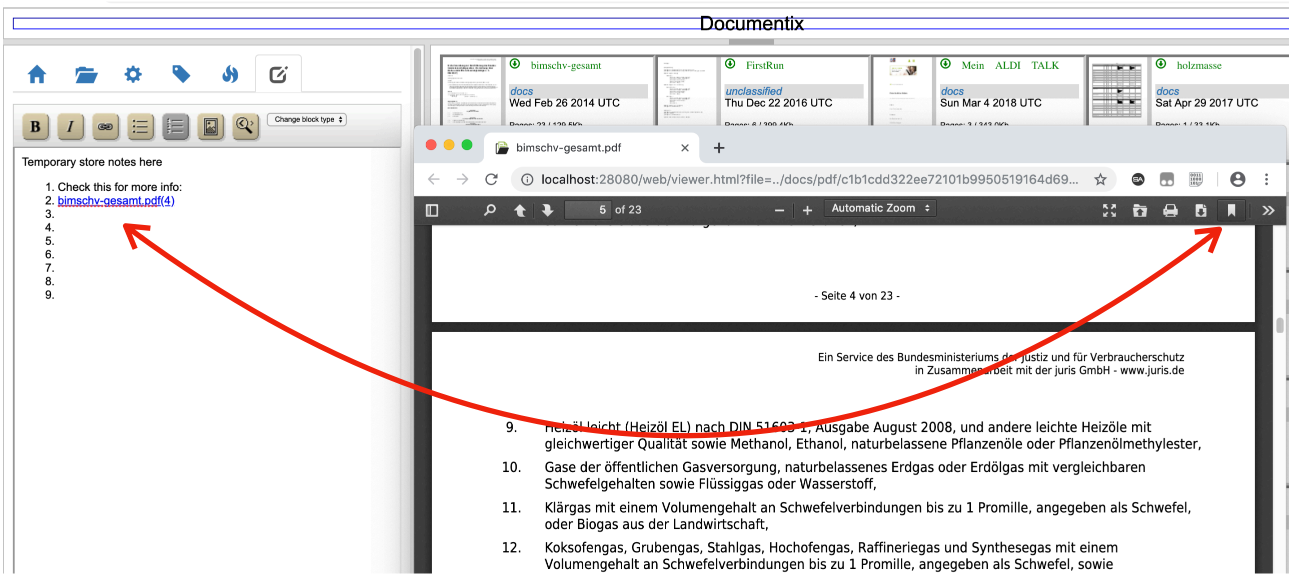
Pictures

Directory browser

Settings

Scratch-pad

Drag & Drop Bookmark-icon to capture document & location in scratch-pad



Temporary store notes here

1. Check this for more info:
2. bimschv-gesamt.pdf(4)
3.
4.
5.
6.
7.
8.
9.

- Seite 4 von 23 -

Ein Service des Bundesministeriums der Justiz und für Verbraucherschutz
in Zusammenarbeit mit der juris GmbH - www.juris.de

9. Heizöl leicht (Heizöl EL) nach DIN 51603-1, Ausgabe August 2008, und andere leichte Heizöle mit gleichwertiger Qualität sowie Methanol, Ethanol, naturbelassene Pflanzenöle oder Pflanzenölmethylester,
10. Gase der öffentlichen Gasversorgung, naturbelassenes Erdgas oder Erdölgas mit vergleichbaren Schwefelgehalten sowie Flüssiggas oder Wasserstoff,
11. Klärgas mit einem Volumengehalt an Schwefelverbindungen bis zu 1 Promille, angegeben als Schwefel, oder Biogas aus der Landwirtschaft,
12. Koksofengas, Grubengas, Stahlgas, Hochofengas, Raffineriegas und Synthesegas mit einem Volumengehalt an Schwefelverbindungen bis zu 1 Promille, angegeben als Schwefel, sowie

Internals

Documentix uses SQLite and the full-text-search engine as a database backend.
All documents are stored using the MD5 hash as an index. (Don't tell me it's broken - it does not matter here)

Documents send to the processing engine are converted to PDF'files (if not already). If the PDF does not contain more than a certain number of words, tesseract as an OCR engine is started on the document and a searchable PDF is created.

An interface to a Canon-mx870 exists but is not supported in the docker image.

It should be easy to create a curl command-line to upload new scans like:

```
> curl  -F file=@o-page-14.pdf -H x-file-name:MyPage.pdf http://localhost:28080/upload
```

Which will upload the file “o-page-14.pdf” and name it on the server “MyPage.pdf”

Retrieving files is also possible:

If you know the MD5 has of a file.

Lets assume the MD5=“009914b6b6a4cb4a37f2fc5bdbc561ca” is.
And the server is reachable under: http://localhost:28080

Then:

http://localhost:28080/docs/raw/009914b6b6a4cb4a37f2fc5bdbc561ca/MyPage
Will return the original document as it was uploaded

http://localhost:28080/docs/pdf/009914b6b6a4cb4a37f2fc5bdbc561ca/MyPage
Will get the file as a PDF (converted if necessary including OCR)

http://localhost:28080/web/viewer.html?file=../docs/pdf/009914b6b6a4cb4a37f2fc5bdbc561ca/MyPage
Will show the file as a PDF in the html viewer.

http://localhost:28080/docs/ico/009914b6b6a4cb4a37f2fc5bdbc561ca/MyPage
Will return the original document as it was uploaded