

DISA@LNU - Information and Software Quality

"Smarter tools"

Morgan Ericsson¹

¹ email: morgan.ericsson@lnu.se, twitter: [@morganericsson](https://twitter.com/morganericsson), GitHub: [morganericsson](https://github.com/morganericsson)

```

from spacy.en import English
import sqlite3

parser = English()
conn = sqlite3.connect('RC_elec.db')
conn2 = sqlite3.connect('tmp.db')

curs = conn.execute('SELECT post_id,sub_id,user_id,
                    link_id,body FROM posts')

for row in curs:
    res = parser(row[4])
    for i,token in enumerate(res):
        if token.is_stop:
            continue
        if token.is_oov:
            continue
        if token.is_punct or token.is_space:
            continue

    conn2.execute('INSERT INTO titlewords
                  VALUES (?, ?, ?, ?, ?, ?, ?, ?)',
                  (row[0],row[1],row[2],row[3],token.lemma_,
                  token.pos_, token.tag_, i))

conn2.commit()

```

```

<h1 class = "standard-page-heading">Detta &#228;r
Linn&#233;universitetet</h1>
<p class = "standard-page-lead
lead">Linn&#233;universitetet &#228;r en kreativ och
internationell kunskapsmilj&#246; som odlar nyfikenhet,
nyt&#228;nkande, nytta och n&#228;rhet.</p>
</header>
<div class="main-content-body">
<div class = "standard-page-main-body
clearfix"><p>Linn&#233;universitetet &#228;r ett av Sveriges nyaste
l&#228;ros&#228;ten och finns i Kalmar och V&#228;xj&#228;. H&#228;r l&#228;ser ungef&#228;r
31.000 studenter vid 150 utbildningsprogram och 2.000
frist&#228;ende kurser. Forskningen h&#228;ller h&#228;g kvalitet och har
f&#228;tt genomslag nationellt som internationellt. Den sp&#228;nner
&#228;ver humaniora, samh&#228;llsvetenskap, naturvetenskap och
teknik, och inneh&#228;ller en rad v&#228;letablerade
forskningsomr&#228;den. Allt ifr&#228;n arbetsmarknadspolitik,
v&#228;lf&#228;rdsfr&#228;gor och entrepren&#228;rskap till biovetenskap,
akvatisk ekologi och tr&#228; och energiteknik.</p>

```

```
from spacy.en import English
import sqlite3
```

```
parser = English()
conn = sqlite3.connect('RC_elec.db')
conn2 = sqlite3.connect('tmp.db')

curs = conn.cursor()
curs.execute('SELECT post_id,sub_id,user_id,link_id,body FROM posts')

for row in curs:
    parser.parse_doc(row[4])
    tokens = parser.doc.get_tokens()
    for token in tokens:
        if token.is_stop:
            continue
        if token.is_punct or token.is_space:
            continue

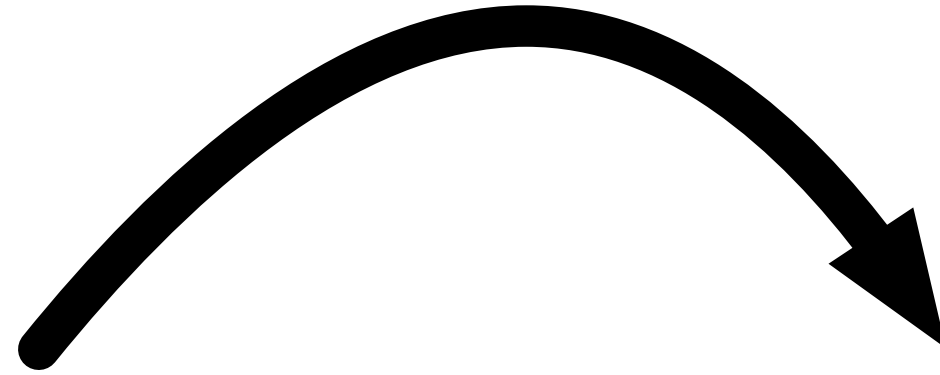
        conn2.execute('INSERT INTO titlewords
            VALUES (?, ?, ?, ?, ?, ?, ?, ?, ?)',
            (row[0], row[1], row[2], row[3], token.lemma_,
            token.pos_, token.tag_, i))

conn2.commit()
```

Quality?

```
<h1 class = "standard-page-heading">Detta &#228;r
Linn&#233;universitetet</h1>
<p class = "standard-page-lead
lead">Linn&#233;universitetet &#228;r en kreativ och
ern&#228;ll kunskapsmilj&#246; som odlar
&#228;ande, a och n&#228;rhet.</p>
ead
y c
y c
"st
l-page
-b
arf
>Linn
versite
av Sveriges nyaste
s&#228;
n fi
Kalmar
. H&#228;r l&#228;r
gef&#228;r
000
nter
utbi
rogram oc
)
st&#228;nde
kurser.
&#228;ningen
&#228;r h&#228;g kvalite
och har
f&#228;tt genomslag nationellt
nationellt. Den sp&#228;nner
&#228;ver humaniora, samh&#228;llsve
naturvetenskap och
teknik, och inneh&#228;ller en rad valetablerade
forskningsomr&#228;den. Allt ifr&#228;n arbetsmarknadspolitik,
v&#228;lf&#228;rdsfr&#228;gor och entrepren&#228;rskap till biovetenskap,
akvatisk ekologi och tr&#228; och energiteknik.</p>
```

Metrics, ...

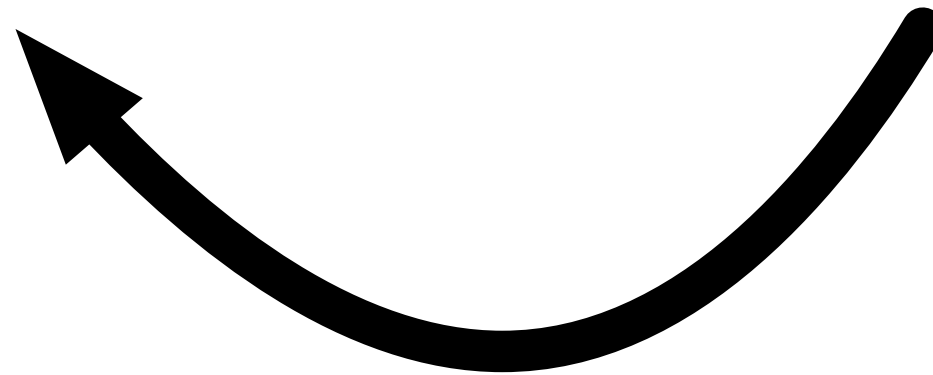


Software

Information

quality

quality



Analytics, ...

Story time!

#	entrytype	author/editor	title	year	journal/booktitle	bibtexkey	ranking
1	Article	Garrido et al.	CORDIC II: A New Improved CORDIC Algorithm	2016	IEEE Trans. on Circuits and S...	GarridoKallstromKummEtAl2...	
2	Inproceedings	Geiger et al.	On the Evolution of BPMN 2.0 Support and Implementation	2016	10th International IEEE Symp...	GeigerHarrerLenhardEtAl2016	★★★★☆
3	Inproceedings	Geiger et al.	Process Engine Benchmarking with Betsy -- Current Status and Future Directions	2016	Proceedings of the 8th Centr...	GeigerHarrerLenhard2016	
4	Inproceedings	Nikol et al.	Service-oriented Multi-tenancy (SO-MT): Enabling Multi-tenancy for Existing Servic...	2016	10th International IEEE Symp...	NikolTraegerHarrerEtAl2016	
5	Phdthesis	Kopp	Partnerübergreifende Geschäftsprozesse und ihre Realisierung in BPEL	2016		Kopp2016	
6	Article	Wettinger et al.	Streamlining DevOps automation for Cloud applications using TOSCA as standa...	2016	Future Generation Comp. Syst.	WettingerBreitenbuecherKop...	★★★★☆
7	Inproceedings	Geiger et al.	BPMN Conformance in Open Source Engines	2015	Proceedings of the 9th IEEE I...	GeigerHarrerLenhardEtAl2015	
8	Inproceedings	Harrer et al.	Improving the Static Analysis Conformance of BPEL Engines with BPELLint	2015	Proceedings of the 9th IEEE I...	HarrerGeigerPreisingerEtAl2...	★★★★★
9	Inproceedings	Kopp et al.	A Domain-Specific Modeling Tool to Model Management Plans for Composite Ap...	2015	Proceedings of the 7th Centr...	KoppBinzBreitenbuecherEtAl...	
10	Inproceedings	Breitenbücher et al.	A Modelling Concept to Integrate Declarative and Imperative Cloud Application Pr...	2015	CLOSER 2015 - Proceedings...	BreitenbuecherBinzKoppEtAl...	★★★★☆
11	Inproceedings	Breitenbücher et al.	A situation-aware workflow modelling extension	2015	Proceedings of the 17th Inter...	BreitenbuecherHirmerKepes...	
12	Inproceedings	Wagner et al.	Choreography-based Consolidation of Interacting Processes Having Activity-bas...	2015	CLOSER 2015 - Proceedings...	WagnerKoppLeymann2015	
13	Inproceedings	Thomsen et al.	Darstellung des Konzeptes - DMA Decentralised Market Agent - zur Bewältigung ...	2015	45. Jahrestagung der Gesell...	ThomsenHartmannKlumpEt...	
14	Inproceedings	Gustafsson and Johansson	Decimation filters for high-speed delta-sigma modulators with passband constr...	2015	2015 IEEE International Sym...	GustafssonJohansson2015	
15	Inproceedings	Johansson and Gustafsson	Filter-bank based all-digital channelizers and aggregators for multi-standard vid...	2015	2015 IEEE International Conf...	JohanssonGustafsson2015	★★★★☆
16	Inproceedings	Alam and Gustafsson	Generalized division-free architecture and compact memory structure for resamp...	2015	European Conference on Cir...	AlamGustafsson2015	
17	Inproceedings	Breitenbücher et al.	Hybrid TOSCA Provisioning Plans: Integrating Declarative and Imperative Cloud ...	2015	Cloud Computing and Servic...	BreitenbuecherBinzKoppEtAl...	
18	Inproceedings	Ingemarsson and Gustafsson	On fixed-point implementation of symmetric matrix inversion	2015	European Conference on Cir...	IngemarssonGustafsson2015	
19	Inproceedings	Johansson and Gustafsson	On frequency-domain implementation of digital FIR filters	2015	2015 IEEE International Conf...	JohanssonGustafsson2015a	
20	Article	Eshuis et al.	Service Outsourcing with Process Views	2015	IEEE Trans. Services Comput...	EshuisNortaKoppEtAl2015	★★☆☆☆
21	Inproceedings	Kopp et al.	Towards a Cloud-based Platform Architecture for a Decentralized Market Agent	2015	45. Jahrestagung der Gesell...	KoppFalkenthalHartmannEtAl...	
22	Inproceedings	Harrer et al.	Automated and Isolated Tests for Complex Middleware Products: The Case of B...	2014	Proceedings of the 7th IEEE I...	HarrerRoeckWirtz2014	★★☆☆☆
23	Inproceedings	Harrer et al.	BPEL Conformance in Open Source Engines: The Case of Static Analysis	2014	Proceedings of the 7th IEEE I...	HarrerPreisingerWirtz2014	
24	Inproceedings	Röck et al.	Performance Benchmarking of BPEL Engines: A Comparison Framework, Statu...	2014	Proceedings of the 26th Inter...	RoeckHarrerWirtz2014	
25	Incollection	Harrer	Process Engine Selection Support	2014	On the Move to Meaningful Int...	Harrer2014	
26	Techreport	Preißinger and Harrer	Static Analysis Rules of the BPEL Specification: Tagging, Formalization and Tests	2014		PreisingerHarrer2014	
27	Techreport	Röck and Harrer	Testing BPEL Engine Performance: A Survey	2014		RoeckHarrer2014	
28	LNUG project proposal.		TOSCA: Portable Automated Deployment and Management of Cloud Applications	2014	Advanced Web Services	BinzBreitenbuecherKoppEtAl...	6
29	Inproceedings	Harrer et al.	Towards a Robustness Evaluation Framework for BPEL Engines	2014	Proceedings of the 7th IEEE I...	HarrerNizamicWirtzEtAl2014	
30	Inproceedings	Kelb and Wirtz	Towards Application Portability in Platform as a Service	2014	Proceedings of the 8th IEEE I...	KelbWirtz2014	

```
package net.sf.jabref.model;

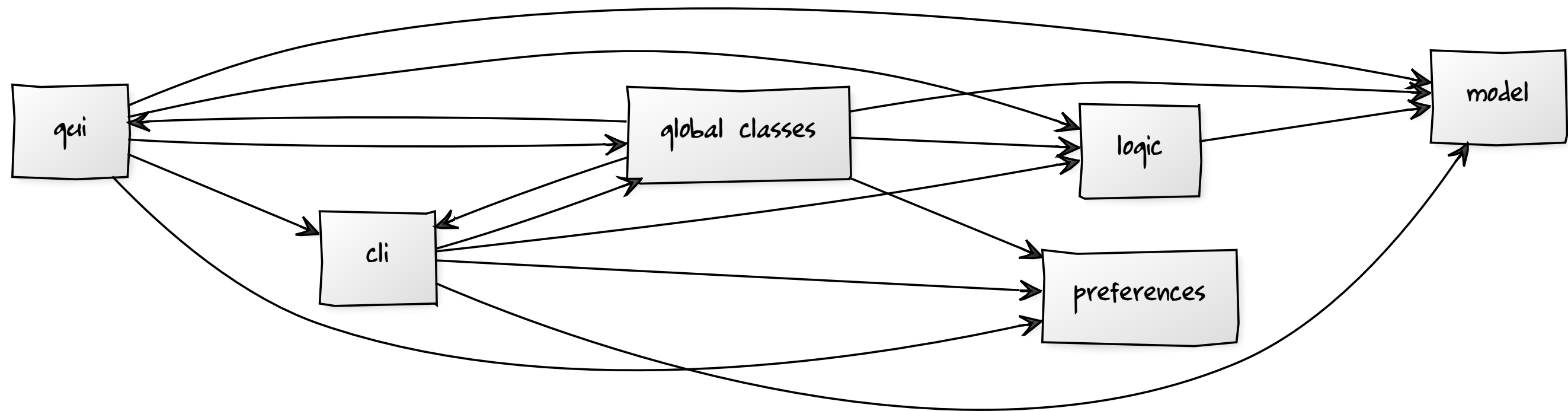
import java.util.Arrays;
import java.util.Collection;
import java.util.HashSet;
import java.util.List;
import java.util.Map;
import java.util.Optional;
import java.util.Set;
import java.util.TreeMap;

import net.sf.jabref.model.database.BibDatabaseMode;
import net.sf.jabref.model.entry.BibLatexEntryTypes;
import net.sf.jabref.model.entry.BibtexEntryTypes;
import net.sf.jabref.model.entry.CustomEntryType;
import net.sf.jabref.model.entry.EntryType;
import net.sf.jabref.model.entry.IEEETranEntryTypes;

public class EntryTypes {

    /**
     * This class is used to specify entry types for either BIBTEX and BIBLATEX.
     */
    private static class InternalEntryTypes {
```

Language	files	blank	comment	code
Java	1,158	24,949	18,512	113,031
XML	100	23	23	3,987
XSD	5	125	631	3,194
TeX	2	121	9	976
Groovy	5	95	37	901
Markdown	8	247	0	808
Python	2	108	110	269
Perl	2	81	92	201
Bourne Again Shell	1	20	21	128
Bourne Shell	4	18	24	104
YAML	3	16	23	85
DOS Batch	1	23	2	59
XSLT	1	5	1	55
Ruby	3	4	3	23
SUM:	1,295	25,835	19,488	123,821



```
package net.sf.jabref.model;

import java.util.Arrays;
import java.util.Collection;
import java.util.HashSet;
import java.util.List;
import java.util.Map;
import java.util.Optional;
import java.util.Set;
import java.util.TreeMap;

import net.sf.jabref.model.database.BibDatabaseMode;
import net.sf.jabref.model.entry.BibLatexEntryTypes;
import net.sf.jabref.model.entry.BibtexEntryTypes;
import net.sf.jabref.model.entry.CustomEntryType;
import net.sf.jabref.model.entry.EntryType;
import net.sf.jabref.model.entry.IEEETranEntryTypes;

public class EntryTypes {

    /**
     * This class is used to specify entry types for either BIBTEX and BIBLATEX.
     */
    private static class InternalEntryTypes {
```

```

public static Iterable<Object[]> data() {
    return Arrays.asList(
        new Object[][] {
            {PACKAGE_NET_SF_JABREF_LOGIC, PACKAGE_JAVA_AWT},
            {PACKAGE_NET_SF_JABREF_LOGIC, PACKAGE_JAVAX_SWING},
            {PACKAGE_NET_SF_JABREF_LOGIC, PACKAGE_NET_SF_JABREF_GUI},
            {PACKAGE_NET_SF_JABREF_LOGIC, CLASS_NET_SF_JABREF_GLOBALS},

            {PACKAGE_NET_SF_JABREF_MODEL, PACKAGE_JAVA_AWT},
            {PACKAGE_NET_SF_JABREF_MODEL, PACKAGE_JAVAX_SWING},
            {PACKAGE_NET_SF_JABREF_MODEL, PACKAGE_NET_SF_JABREF_GUI},
            {PACKAGE_NET_SF_JABREF_MODEL, PACKAGE_NET_SF_JABREF_LOGIC},
            {PACKAGE_NET_SF_JABREF_MODEL, CLASS_NET_SF_JABREF_GLOBALS}
        }
    );
}

```

Good enough?

Good enough?

What does this mean?

Good enough?

- Does it find all violations? No! Only checks imports...
 - but it does enforce a structure
- What if it did? Diagram is not perfect... What is logic, what is model?
 - but there is a design

Good enough?

How can we improve?

Many tools claim to do better?

- But require a significant time/effort, sometimes even changes to the development process / toolchain.
- Suffer from bit rot (many tested tools would not parse code that follows the current Java standard)

"Smarter tools"

"Smarter tools"

- All violations are not equal, we should focus on the important ones.
- We should move towards a *perfect* design.
- Tools should evolve with minimal effort

Another example

Cohesion within vs without

External dependencies



ADDNODE GROUP



Collaborators and Users

