
Dockerizing Indri

Claudia Hauff

A white die with black pips is shown on a wooden surface. The die is positioned in the upper right quadrant of the image. The background is a blurred wooden surface. A semi-transparent white rectangular box is overlaid on the lower half of the image, containing the text.

Installing Indri on
Ubuntu without issues

Robust

Robust 04

.z .1z .2z



GOV2



CORE18
JSON

```
<query>
<number>301-LM</number>
<text>#combine( international organized crime )</text>
</query>
```

```
<query>
<number>301-SD</number>
<text>#weight( 0.9 #combine(international organized crime)
0.05 #combine(#1(international organized) #1(organized crime) )
0.05 #combine(#uw8(international organized) #uw8(organized crime)) )</text>
</query>
```

```
<query>
<number>301-DESC</number>
<text>#combine( identify organizations that participate in international criminal activity the activity and
if possible collaborating organizations and the countries involved )</text>
</query>
```

```
<query>
<number>301-BM25</number>
<text>international organized crime</text>
</query>
```

```
<query>
<number>301-LM</number>
<text>#combine( international organized crime )</text>
</query>
```

Indri doesn't know TREC
topic file formats.

```
<query>
<number>301-SD</number>
<text>#weight( 0.9 #combine(international organized crime)
0.05 #combine(#1(international organized) #1(organized crime) )
0.05 #combine(#uw8(international organized) #uw8(organized crime)) )</text>
</query>
```

Indri doesn't like special
chars in its topic files.

```
<query>
<number>301-DESC</number>
<text>#combine( identify organizations that participate in international criminal activity the activity and
if possible collaborating organizations and the countries involved )</text>
</query>
```

```
<query>
<number>301-BM25</number>
<text>international organized crime</text>
</query>
```

Not every retrieval model
ingests the same topic
format in Indri.



LM + PRF + SD

>>

LM + PRF

LM + SD

>>

LM



Indri fails silently