



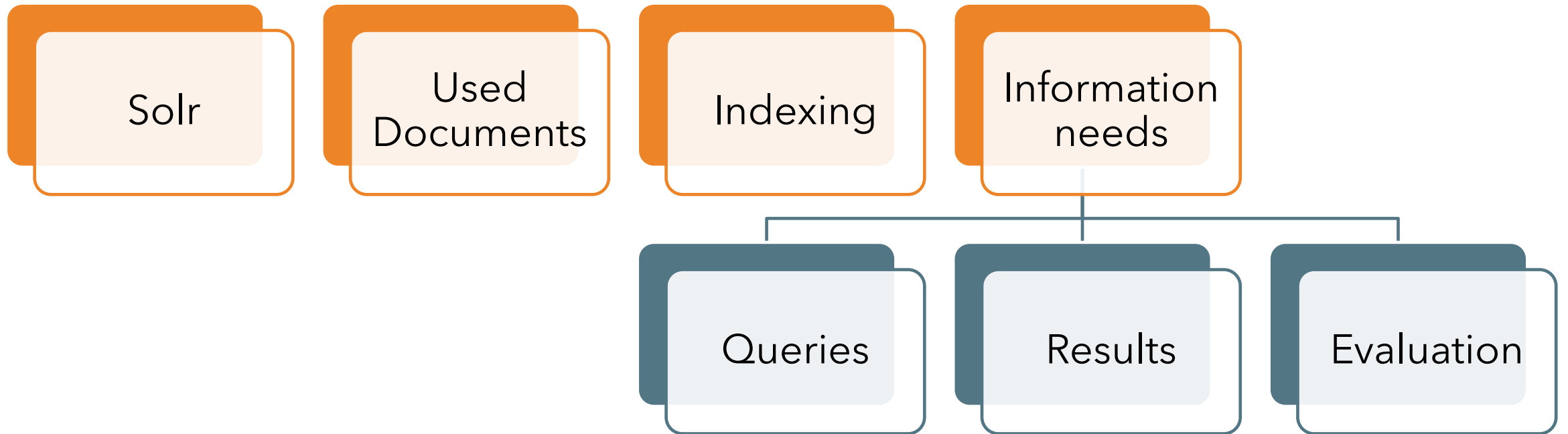
PITCHFORK MUSIC REVIEWS

MILESTONE #2

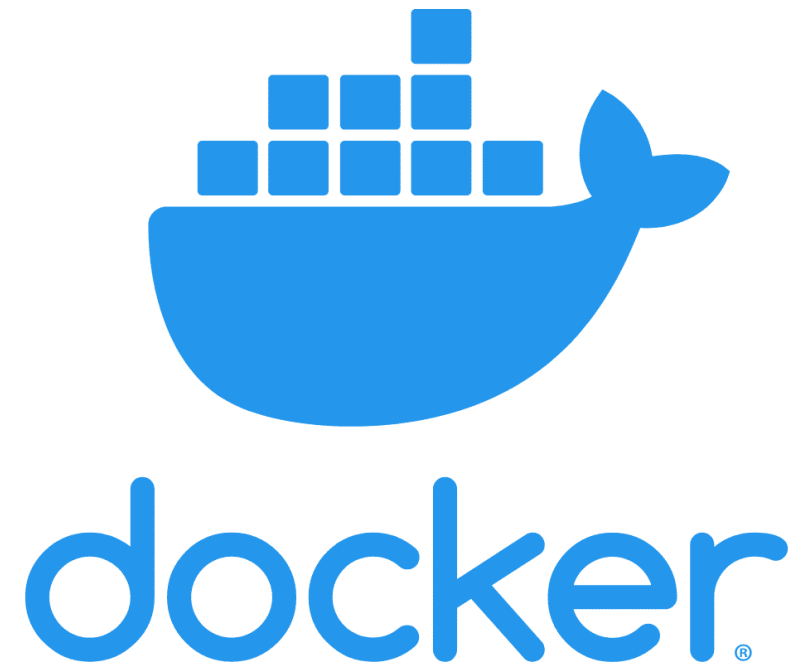
INFORMATION PROCESSING AND RETRIEVAL

FABIO HUANG - UP201806829,
LISA SONCK - UP202202272,
TEMACO - UP202200606

CONTENT



SOLR



USED DOCUMENTS

```
{
  "reviewid": [13088],
  "title": ["sketches of spain: legacy edition"],
  "artist": ["miles davis"],
  "rating_score": [8.0],
  "author": ["mark richardson"],
  "review_publication_date": [1244160000000],
  "genre": ["jazz"],
  "song_release_year": "2009",
  "record_label": ["legacy"],
  "id": "bd92093a-9f2b-4405-9a75-80ba845e12a4",
  "review_content": ["Is this even jazz? Sketches of Spain was p"],
  "_version_": 1749576559401893888,
  "score": 8.093981},
{
```

INDEXING

Type	Field	Indexed
standard_text	artist	TRUE
	author	TRUE
	genre	TRUE
	song_release_year	TRUE
	record_label	TRUE
pdoubles	rating_score	TRUE
plongs	review_publication_date	TRUE
free_text	review_content	TRUE
	title	TRUE
string	reviewid	FALSE

INDEXING

Field Type	Tokenizer	Filter
titles	solr.ClassicTokenizerFactory	solr.ASCIIFoldingFilterFactory
free_text		solr.LowerCaseFilterFactory
		solr.ASCIIFoldingFilterFactory
		solr.LowerCaseFilterFactory
		solr.EnglishMinimalStemFilterFactory
		solr.EnglishPossessiveFilterFactory
		solr.HyphenatedWordsFilterFactory
		solr.KStemFilterFactory

INFORMATION NEEDS



Find albums that are perceived as boring.



Find songs and albums that are ideal for studying.

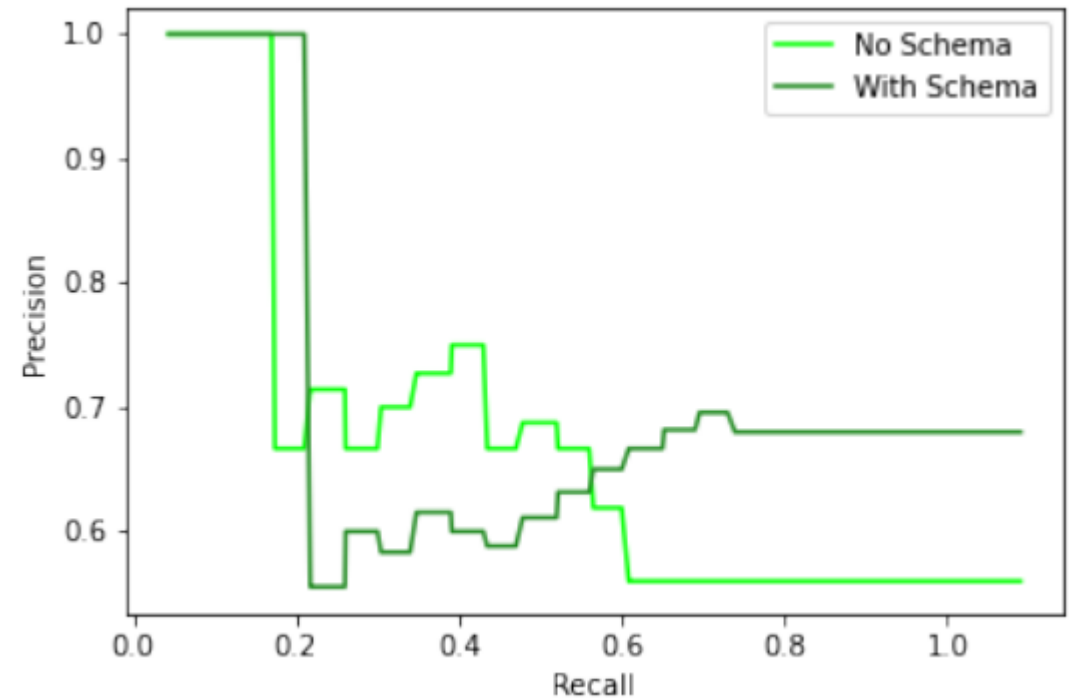
QUERY 1: BORING ALBUMS

```
review_content:boring^5  
-title:boring  
-review_content:"doesn't boring"~5  
-review_content:"not boring"~5  
-review_content:"never boring"~5  
review_content:"is boring"~5  
-review_content:" "(" boring "(" " " ~3^2
```


QUERY 1: BORING ALBUMS - RESULTS

Metric	Values With Schema	Values Without Schema
Average Precision	0.719355	0.740282
Precision at 5 (P@05)	1.0	0.8
Precision at 10 (P@10)	0.6	0.7
Precision at 15 (P@15)	0.6	0.666667
Precision at 20 (P@20)	0.65	0.65
Precision at 25 (P@25)	0.68	0.56

Table 1. Results with and without an applied schema.



QUERY 2: SONGS AND ALBUMS FOR STUDYING

q:

- genre:jazz^2
- review_content: calm
- review_content: quiet
- review_content: ambient
- review_content: relaxing
- review_content: "classical music"
- !review_content: "violent"

fq:

- rating_score: [8.0 TO 10.0]

QUERY 2: SONGS AND ALBUMS FOR STUDYING

Metric	Values With Schema	Values Without Schema
Average Precision	0.475046	0.48338
Precision at 5 (P@05)	0.6	0.8
Precision at 10 (P@10)	0.6	0.6
Precision at 15 (P@15)	0.466667	0.466667
Precision at 20 (P@20)	0.35	0.35
Precision at 25 (P@25)	0.28	0.28

Table 2. Results with and without an applied schema.

System	AUC value
Query 2 Without Schema	0.4964
Query 2 With Schema	0.5309

Table 3. AUC for Query 2.

