elastic | The Search Analytics Company

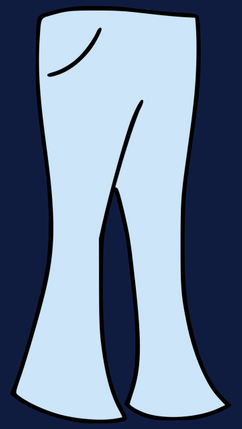
Retro Relevance:

Lessons Learned Balancing Keyword and Semantic Search

Kathleen DeRusso

April 23, 2024

What's this talk about, anyway?



- The text search toolbox — Reminders about tried and true concepts that are still.... *relevant* today. 🥁
- Semantic search — It's new and shiny, but it's not without its challenges
- Hybrid search — The best of both worlds
- Simplifying search through tools — Let's put that toolbox to good use

elastic

Who am I?

Kathleen DeRusso | Principal Software Engineer

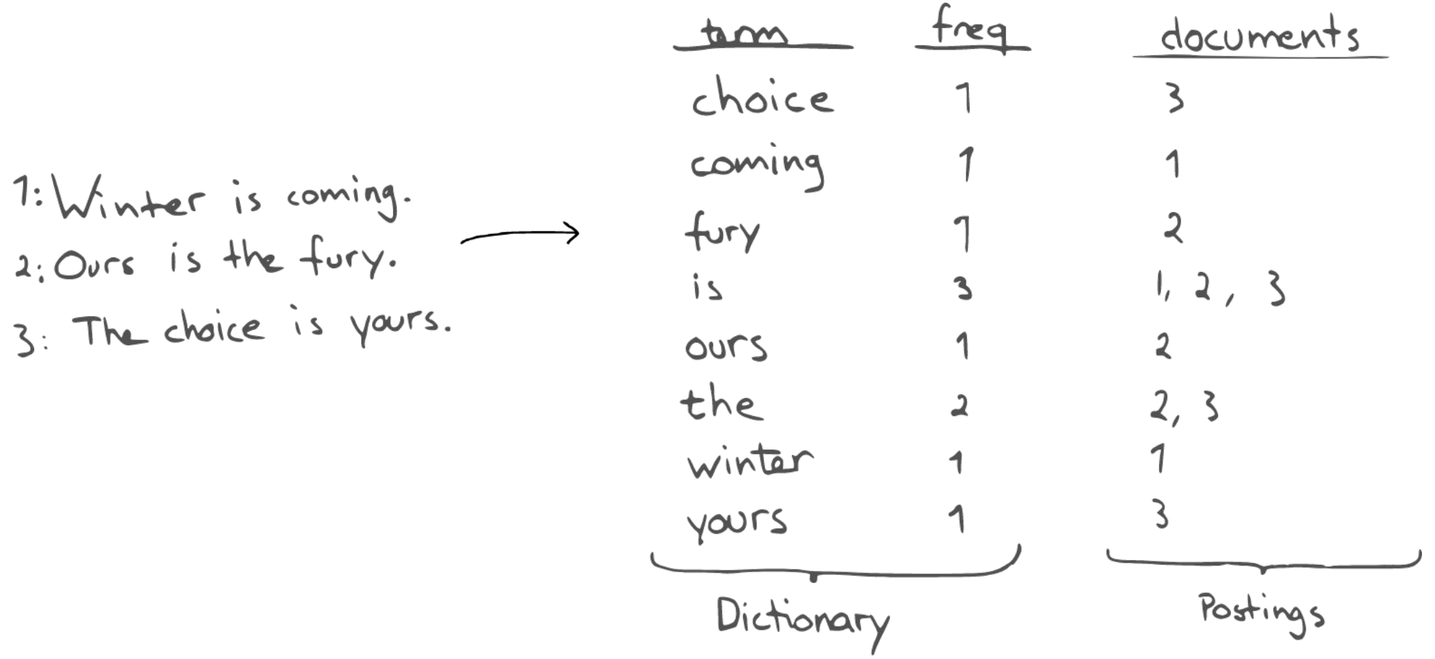
- Software engineer and search practitioner
- B2B search application experience
  - Mostly lexical search
  - Complexity added through business rules
- SaaS experience
  - Store and find data from incremental backups
- Elasticsearch
  - Examples use Elasticsearch, but general concepts translate

elastic

Let’s talk about text search

Text → Tokens → Term Frequencies → Postings Lists

1: Winter is coming.
2: Ours is the fury.
3: The choice is yours.



elastic

TF/IDF - 1972

The Godfather

VCRs

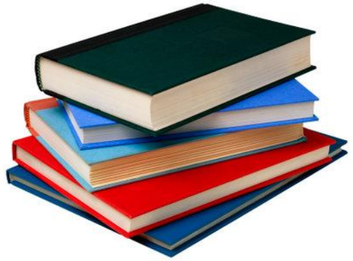
Pong

elastic

BM25 - 1994

Forrest Gump

Beanie Babies were an investment!



Amazon was launched

...as a bookstore!

elastic

Lexical Search Toolbox

Analyzers & Tokenizers

Whitespace
Punctuation
Stemming

Filters, Field Weights & Boosts

“Title” > “Description” or “Meta”

Reranking

Learn to Rank

Synonyms

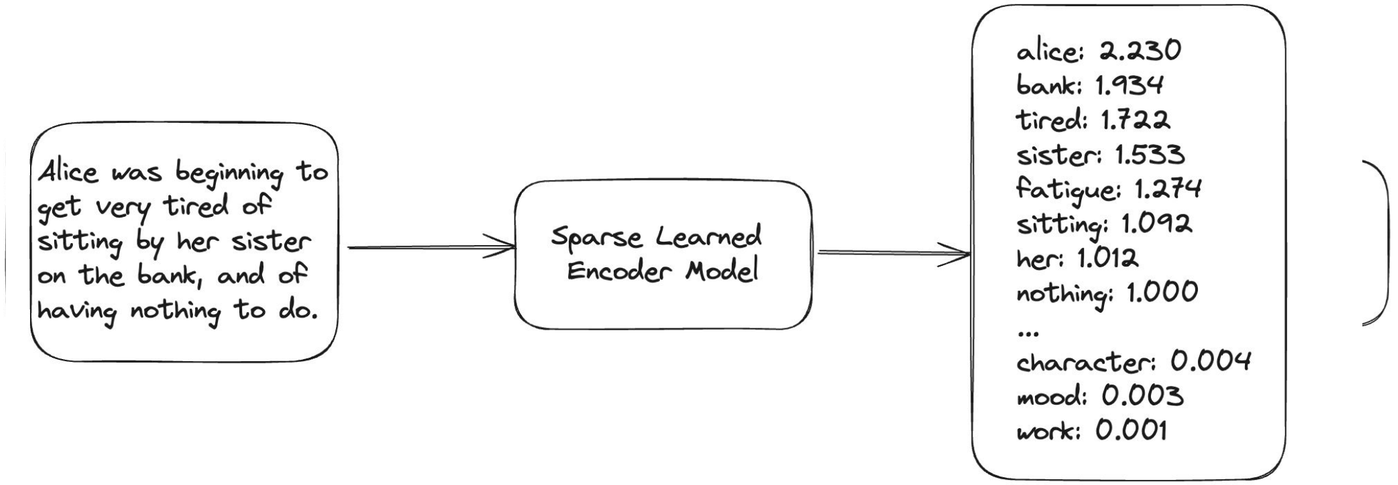
“Soda” vs. “Pop” vs. “Cola” vs. “Coke”
“Sub” vs. “Hoagie” vs. “Grinder”
“Runners” vs. “Sneakers” vs. “Trainers”

Business Rules

Diversification of results, campaigns to promote specific content over relevant matches, etc.

elastic

Semantic Search



elastic

Semantic search isn’t perfect

Cost

Money, time

- RAM, compute, latency
- Convenience vs. time: Inference can be slow
- Developer time

Features

Expected functionality

- Highlighting
- Spelling correction
- Typo tolerance: HODL

Queries

Niche queries

- Exact matches: ACA979GL
- Domain-specific jargon
- Popularity, conversions, campaigns
- Query understanding: half a kilo vs 500g

elastic

Hybrid Search to the Rescue!

Best of both worlds,
providing the best recall

"Modern farmhouse with lots of land and an inground pool in the 12866 zip code"

"Comfortable Skechers with memory foam insoles in purple"

"Fantasy movie with Sean Astin with elves and monsters"

"Remote software engineer jobs using Elasticsearch and cloud native technologies"

"Fantasy series with dragons where the Hand of the King dies"

What does this look like?

```
{
  "query": {
    "bool": {
      "should": [
        {
          "multi_match": {
            ...
          }
        },
        {
          "knn": {
            ...
          }
        }
        {
          "text_expansion": {
            ...
          }
        },
        {
          "rule_query": {
            ...
          }
        }
      ]
    }
  }
}
```

elastic

What does this look like?

```
GET my-index/_search
{
  "sub_searches": [
    {
      "query": {
        "match": {
          "my_text_field": "gluten free pizza recipe"
        }
      }
    },
    {
      "query": {
        "text_expansion": {
          "my_tokens": {
            "model_id": ".elser_model_2",
            "model_text": "gluten free pizza recipe"
          }
        }
      }
    }
  ],
  "rank": {
    "rrf": {}
  }
}
```

elastic

What does this look like?

```
GET example-index/_search
{
    "retriever": {
        "rrf": {
            "retrievers": [
                {
                    "standard": {
                        "query": {
                            "term": {
                                "text": "purple sneakers"
                            }
                        }
                    }
                },
                {
                    "knn": {
                        "field": "vector",
                        "query_vector": [1.25, 2, 3.5],
                        "k": 50,
                        "num_candidates": 100
                    }
                }
            ],
            "window_size": 50,
            "rank_constant": 20
        }
    }
}
```

elastic

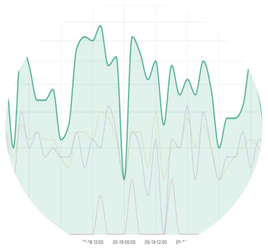
Bringing it all together

Combining multiple result sets is **HARD**

- Linear Combination
  - Technology we all know and love
  - Requires tuning
- RRF
  - Easier to use with minimal/no tuning
  - Makes pagination harder
  - No support for pinned documents
  - Less control over result sets

elastic

Tuning Results



Business Metrics and User Engagement

Click through rate, conversion rate, time on site, bounce rate...

Precision and Recall

MAP, NDCG, F1, ...
Judgment Lists

Don’t tune for a pet query

Judgment lists and relevance metrics are your friend!

elastic

Query rules

Contextual rules based on input in query

- Remember that pet query?
  - If you already know the best result, pin it!
- Why not go further?
  - Head queries
  - No, it’s not cheating!
- Managed by non-search practitioners

```
PUT /_query_rules/my-ruleset
{
  "rules": [
    {
      "rule_id": "rule1",
      "type": "pinned",
      "criteria": [
        {
          "type": "fuzzy",
          "metadata": "query_string",
          "values": [ "puggles", "pugs" ]
        },
        {
          "type": "exact",
          "metadata": "user_country",
          "values": [ "us" ]
        }
      ],
      "actions": {
        "ids": [
          "id1",
          "id2"
        ]
      }
    }
  ]
}
```

elastic

Query rules

Contextual rules based on input in query

- Remember that pet query?
  - If you already know the best result, pin it!
- Why not go further?
  - Head queries
  - No, it’s not cheating!
- Managed by non-search practitioners

```
GET /my-index/_search
{
  "query": {
    "rule_query": {
      "organic": {
        "query_string": {
          "query": "puggles"
        }
      },
      "match_criteria": {
        "query_string": "puggles",
        "user_country": "us"
      },
      "ruleset_id": "my-ruleset"
    }
  }
}
```

elastic

Synonyms

Bridge the gap from domain-specific and models

- Domain specific terminology
- Slang
- Aliases
- Translations
  - Coke vs. Coca-Cola
- API and analyzer in Elasticsearch

```
PUT _synonyms/my-synonyms-set
{
  "synonyms_set": [
    {
      "id": "pc",
      "synonyms": "pc => personal computer"
    },
    {
      "id": "computer",
      "synonyms": "computer, pc, laptop, desktop"
    }
  ]
}
```

elastic

Synonyms

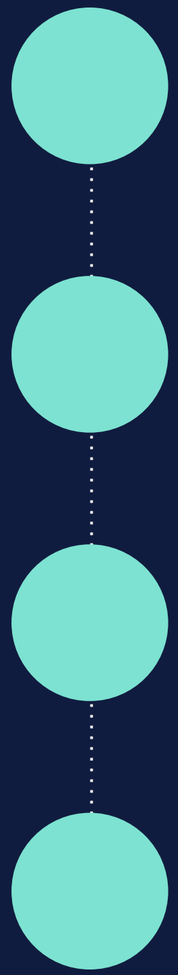
Bridge the gap from domain-specific and models

- Domain specific terminology
- Slang
- Aliases
- Translations
  - Coke vs. Coca-Cola
- API and analyzer in Elasticsearch

```
PUT /my-index
{
  "settings": {
    "index": {
      "analysis": {
        "analyzer": {
          "synonym_analyzer": {
            "tokenizer": "whitespace",
            "filter": ["my_synonyms"]
          }
        },
        "filter": {
          "my_synonyms": {
            "type": "synonym",
            "synonyms_set": "my-synonyms-set",
            "updateable": true
          }
        }
      }
    }
  }
}
```

elastic

Wrapping up...



Hybrid search

BM25 search and associated tools help bridge the gap for semantic search

Tools

Tools like query rules and synonyms make your life easier

Evolving landscape

We’re continuously solving problems and making search easier

Ultimate goal:

Deliver the best search results … with as low effort and cost (money, time) as possible

elastic

Thank you!

... and, we’re hiring!