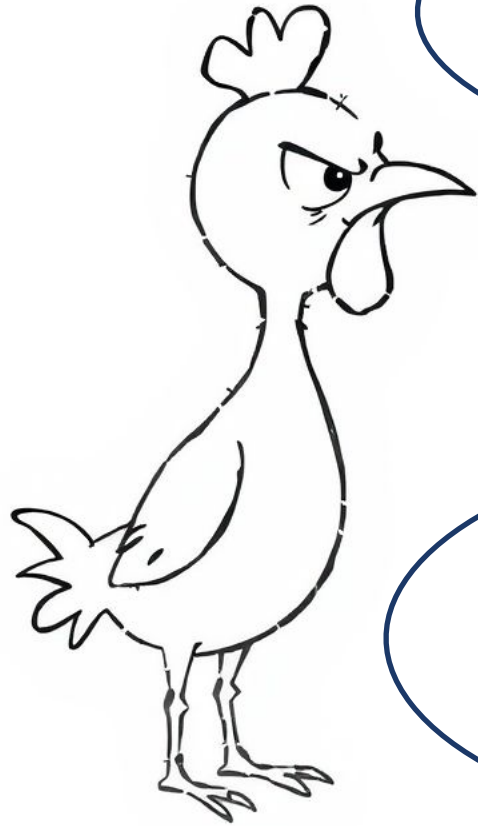


User Behavior Insights (UBI)

Aruna Govindaraju (she/her)
OpenSearch Specialist @AWS

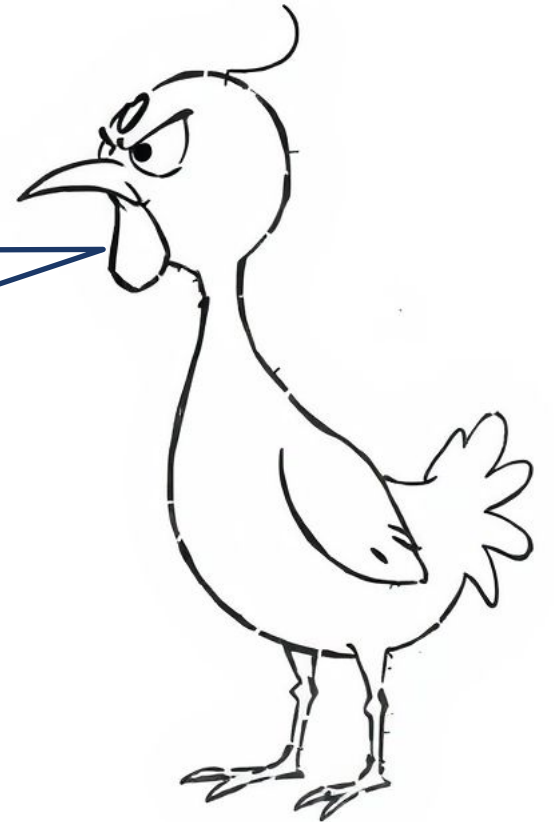
Supercharging Search in OpenSearch:
Harnessing User Behavior insights for Continuous Improvement

Search is Hard !



WHY CAN'T THE SEARCH
ENGINE JUST SHOW ME
WHAT I MEANT?

BECAUSE YOUR
QUERY WAS: '*SHOE
THING I SAW ONCE
WITH BLUE MAYBE*'





wantdo Women's Hooded Winter
Coat Waterproof Warm Long Puffer..

And yet it never seems to work right

All ▾

jacket for warm weather



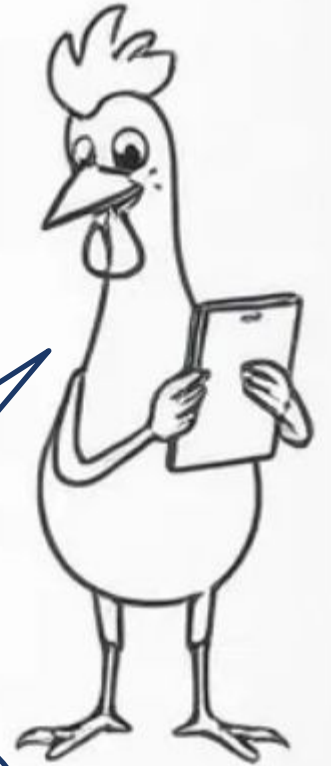
YOU ARE
KIDDING!



Matching words $\neq$ Understanding intent!



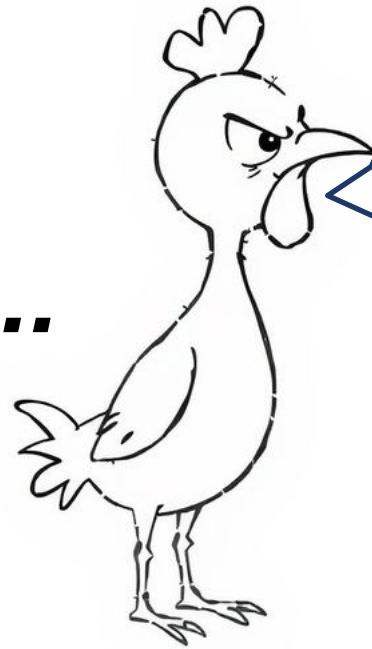
WHY ARE YOU SHOWING
ME 200 RESULTS FOR
'BLACK SNEAKERS'?



THEY ALL CONTAIN THE
WORD 'BLACK'... SEEMS
RELEVANT TO ME

How to fix this?

***Great Search...
or just***

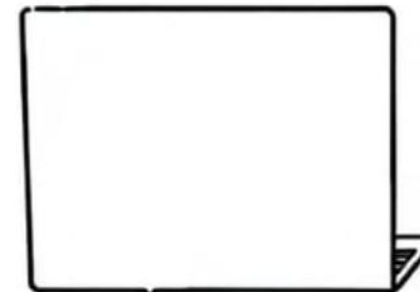


WHAT DID YOU DO ?
NOW EVERY SEARCH
RETURNS "SOCKS",
EVEN FOR "LAPTOP" !

Great Guesses?

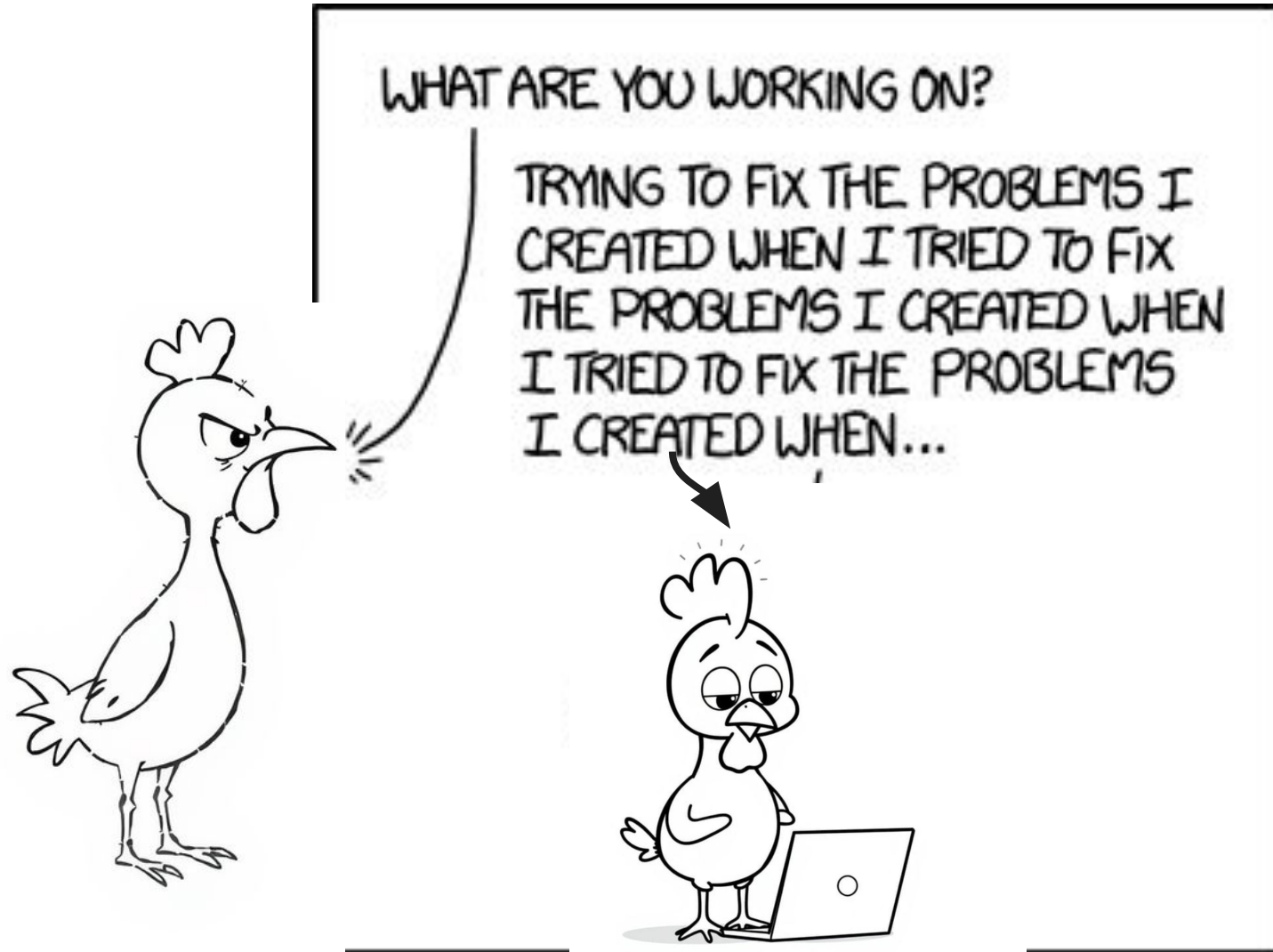


HMMM.. I JUST CHANGED
THE SYNONYM LIST AND
BOOSTED RELEVANCE BY
5%....



Tuning search

*without understanding
user-behavior
is just creative chaos.*



What search needs?

Standardize
schema

Capture
queries,
events

Analyze &
build
insights

A/B Testing

Trace &
Correlate

Understand
Application
Data

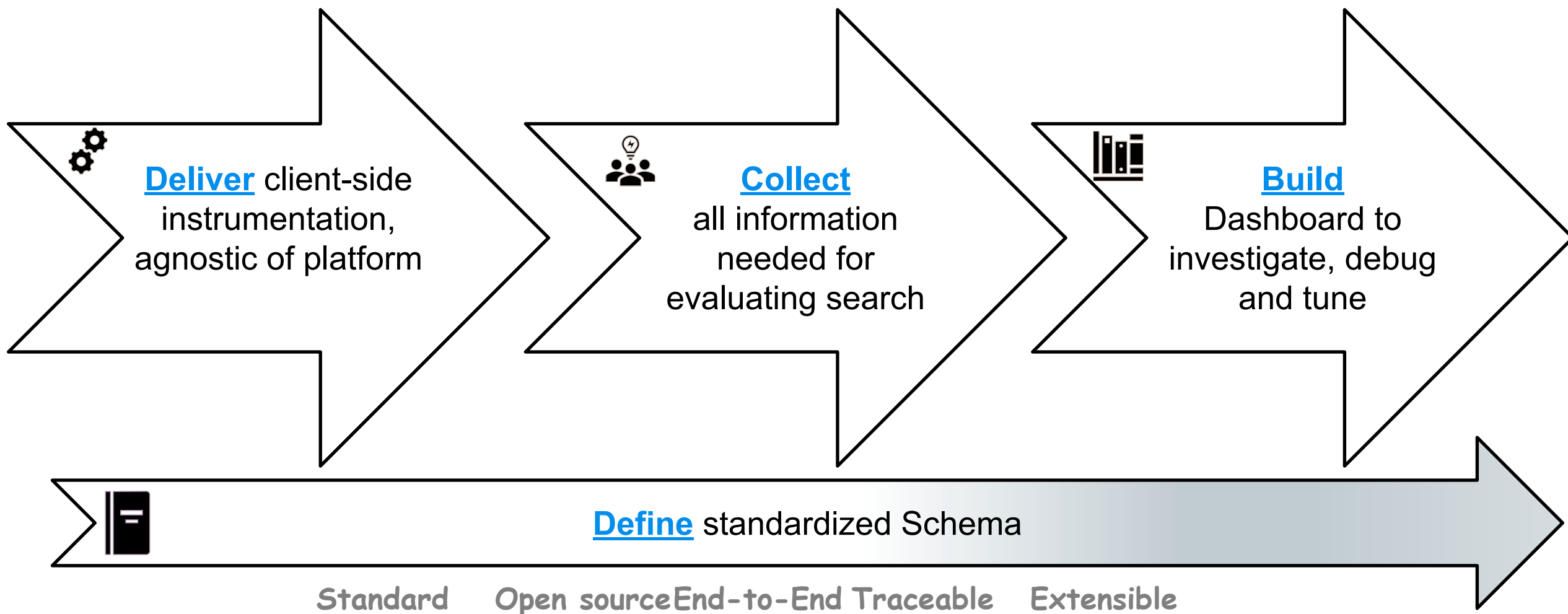
Tuning

Logging

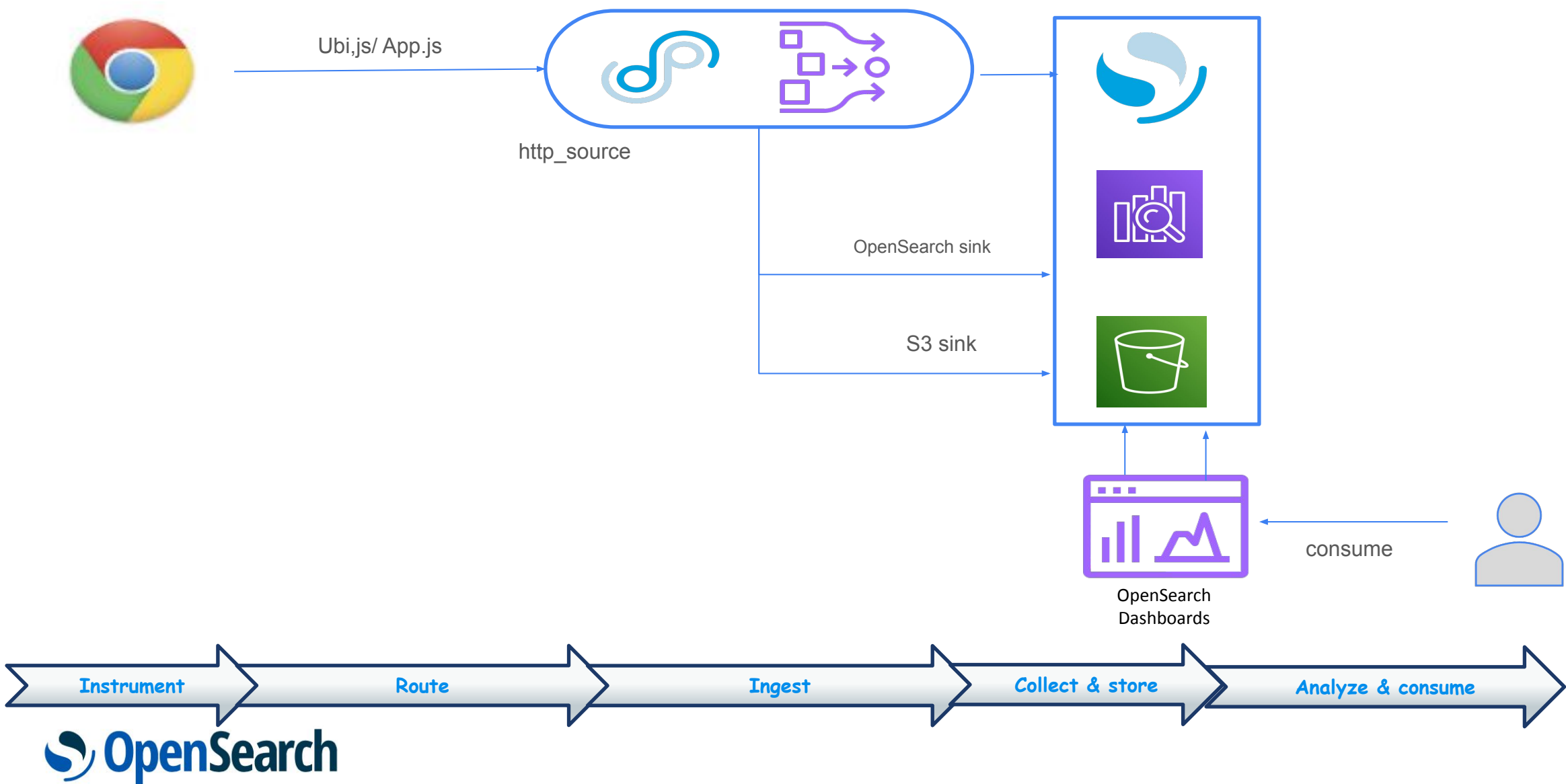
Orchestrate
the right
signal

**Quality
evaluation**

User Behavior insights



UBI Implementation pattern – with OpenSearch



Setting UBI plugin

Installing UBI plugin for OpenSearch

Step 1: Install the plugin

```
bin/opensearch-plugin install file:/opensearch-ubi-2.18.0.0.zip
```

Step 2: Initialize the plugin

Creates **ubi_queries** and **ubi_events**

```
curl -s -X POST http://opensearch:9200/_plugins/ubi/initialize
```

Step 3: Instrument queries

Instrument queries to OpenSearch to include **ubi** section in the **ext** block

For events, send the action on the client directly to OpenSearch

```
curl -s -X POST http://localhost:9200/ubi_events/_doc/ -H "Content-Type: application/json" -d' {}
```

Step 4: Install dashboards

Run the shell script:

```
./install_dashboards.sh localhost:9200 localhost:5601 , installs the dashboard templates
```



UBI Query Index

```
{
  "dynamic": false,
  "properties": {
    "timestamp": { "type": "date", "format": "strict_date_time" },
    "query_id": { "type": "keyword", "ignore_above": 100 },
    "query": { "type": "text" },
    "query_response_id": { "type": "keyword", "ignore_above": 100 },
    "query_response_hit_ids": { "type": "keyword" },
    "user_query": { "type": "keyword" },
    "query_attributes": { "type": "flat_object" },
    "object_id_field": { "type": "keyword", "ignore_above": 100 },
    "client_id": { "type": "keyword", "ignore_above": 100 },
    "application": { "type": "keyword", "ignore_above": 100 }
  }
}
```

Schema



```
{
  "_index": "ubi_queries",
  "_id": "K4cAP5YBuO9ZoB79iubT",
  "_score": 1,
  "_source": {
    "query_response_id": "514755ea-e1b0-482f-8368-55b5f90",
    "query_id": "a82b33dc-829e-4a6c-9e75-97bf7ee1bb3a",
    "application": "Chorus",
    "query": {
      "size": 0,
      "query": {
        "bool": {
          "must": [
            {
              "multi_match": {
                "query": "s.Brand^1.0",
                "attrs.Color^1.0",
                "bullets^1.0",
                "category^1.0",
                "description",
                "type": "best_fields",
                "operator": "OR",
                "slop": 0,
                "prefix_length": 0,
                "max_expansion": 1,
                "query": "NONE",
                "auto_generate_synonyms_phrase_query": true,
                "fuzziness": 1,
                "adjust_pure_negative": true,
                "boost": 1,
                "adjust_pure_negative": true,
                "source": {
                  "includes": ["*"],
                  "excludes": [],
                  "aggregations": {
                    "category_filter": {
                      "filter": {
                        "size": 20,
                        "min_doc_count": 1,
                        "shard_min_doc_count": 0,
                        "show_terms": true,
                        "order": [
                          { "_count": "desc" },
                          { "_key": "asc" }
                        ]
                      },
                      "ext": {
                        "query_id": "a82b33dc-829e-4a6c-9e75-97bf7ee1bb3a",
                        "user_query": "design",
                        "client_id": "CLIENT-0a5851d1-2c27-48e4-a79d-87eda8fe",
                        "application": "Chorus",
                        "object_id_field": "asin",
                        "query_attributes": {}
                      }
                    }
                  }
                }
              }
            }
          ]
        }
      }
    },
    "user_query": "design",
    "query_response_hit_ids": [],
    "query_attributes": {},
    "client_id": "CLIENT-0a5851d1-2c27-48e4-a79d-87eda8fe",
    "timestamp": "2025-04-16T14:28:42.323+0000"
  }
},
```

Sample document

UBI Events Index

Schema

```
{
  "properties": {
    "application": { "type": "keyword", "ignore_above": 256 },
    "action_name": { "type": "keyword", "ignore_above":
100 },
    "client_id": { "type": "keyword", "ignore_above": 100 },
    "query_id": { "type": "keyword", "ignore_above": 100 },
    "message": { "type": "keyword", "ignore_above": 1024 },
    "message_type": { "type": "keyword", "ignore_above": 100
},
    "user_query": { "type": "keyword" },
    "timestamp": {
      "type": "date",
      "format": "strict_date_time",
      "ignore_malformed": true,
      "doc_values": true
    },
    "event_attributes": { "type": "object", "ignore_above": 100
  }
}
```



Sample document

```
{
  "_index": "ubi_events",
  "_id": "a9d8bd92-e8be-478f-af7e-8c29a34c5309",
  "_score": 1,
  "_source": {
    "application": "Chorus",
    "action_name": "add_to_cart",
    "query_id": "a82b33dc-829e-4a6c-9e75-97bf7ee1bb3a",
    "session_id": "SESSION-027fa977-422d-4083-a135-9ee89830f5ef",
    "client_id": "CLIENT-0a5851d1-2c27-48e4-a79d-87eda8fe643f",
    "user_id": "",
    "timestamp": "2025-04-16T14:28:42.417Z",
    "message_type": "CONVERSION",
    "message": "Fashion Angels Fashion Design Sketch Kit - Compact Portfolio Sketchbook
Kids Ages 6+ and Up, Comes with Stencils and Stickers (undefined)",
    "event_attributes": {
      "object": {
        "object_id": "B07BGNP86Q",
        "object_id_field": "asin",
        "description": "Fashion Angels Fashion Design Sketch Kit - Compact Portfolio Sketchb
for Kids Ages 6+ and Up, Comes with Stencils and Stickers"
      },
      "position": {
        "ordinal": 1
      }
    },
    "user_query": "design"
  }
},
```

Samples

```
curl -s http://localhost:9200/your-index/_search -H
"Content-Type: application/json" -d'
{
  "ext": {
    "ubi": {
      "query_id": "12300d16cb-b6f1-4012-93ebcc49cac90426",
      "user_query": "Toner",
      "client_id": "c4af7ca2-d6d2-4326-a41f-e616ebdd3d7b",
      "object_id_field": "product_name",
      "query_attributes": {
        "experiment_id": "12345"
      }
    }
  },
  "query": {
    "match": {
      "name": "toner"
    }
  }
}'
```

```
curl -s -X POST http://localhost:9200/ubi_events/_doc/ -H
"Content-Type: application/json" -d'
{
  "action_name": "page_exit",
  "user_id": "1821196507152684",
  "query_id": "00112233-4455-6677-8899-aabbccddeeff",
  "session_id": "c3d22be7-6bdc-4250-91e1-fc8a92a9b1f9",
  "page_id": "/docs/latest/",
  "timestamp": "2024-05-16T12:34:56.789Z",
  "message_type": "INFO",
  "message": "On page /docs/latest/ for 3.35 seconds",
  "event_attributes": {
    "position": {},
    "object": {
      "idleTimeoutMs": 5000,
      "currentIdleTimeMs": 250,
      "checkIdleStateRateMs": 250,
      "isUserCurrentlyOnPage": true,
      "isUserCurrentlyIdle": false,
      "currentPageName": "http://localhost:4000/docs/latest/",
      "timeElapsedCallbacks": [],
      "userLeftCallbacks": [],
      "userReturnCallbacks": [],
      "visibilityChangeEventName": "visibilitychange",
      "hiddenPropName": "hidden"
    }
  }
}'
```

Demo

Pick your Algo

Keyword ▾

Filter by Brands

- ☐ Amazon Basics 77
- ☐ Funko 54
- ☐ LEGO 50
- ☐ Melissa & Doug 34
- ☐ Corsair 21
- ☐ DEWALT 20
- ☐ MyGift 20

Filter by Product Types

- ☐ Home & Kitchen 4399
- ☐ Kitchen & Dining 1672
- ☐ Tools & Home Improvement 1663
- ☐ Toys & Games 1632
- ☐ Electronics 1463
- ☐ Clothing, Shoes & Jewelry 1310
- ☐ Accessories 1116

 Search for products, brands or ASIN

10000 results found in 55ms



Amazon Basics Kid's Sheet Se...
16.72 \$ | Amazon Basics
Add to  | rank: 0



Autocastle Electric Battery He...
26.99 \$ | Autocastle
Add to  | rank: 1



110 Pcs Outer Space Party Su...
9.99 \$ | Gemwi
Add to  | rank: 2



Plantar Fasciitis Posterior Nig...
31.99 \$ | MARS WELLNESS
Add to  | rank: 3



Saint Mossi 9-Lights Modern ...
149.99 \$ | SM Saint Mossi
Add to  | rank: 4



AVIRGO Luggage Tag Flexible ...
14.89 \$ | Hermes
Add to  | rank: 5



Lush Decor, Blue Cynthia Jaco...
22.99 \$ | Lush Decor
Add to  | rank: 6



USB C to 3.5mm Headphone A...
7.99 \$ | Vienon
Add to  | rank: 7



Salt Rookie Crank Set 175mm ...
69.99 \$ |
Add to  | rank: 8



LBRO2M Bed Sheet Set Queen...
32.39 \$ | LBRO2M
Add to  | rank: 9

Demo Instrumentation & Collection

Visualizing User Behavior - Sample Dashboards

Insights – Dashboards



Leveraging UBI data

UBI Enables Use Cases

Ecommerce

Personalization

Relevancy Improvements

Churn Prediction

Product Inventory Gaps

User Segmentation

Chunk Selection in RAG

Knowledge Base Content Gaps

Migration to OpenSearch

Analytical Use Cases

Measure Search Quality over Time

Behavioral Analytics

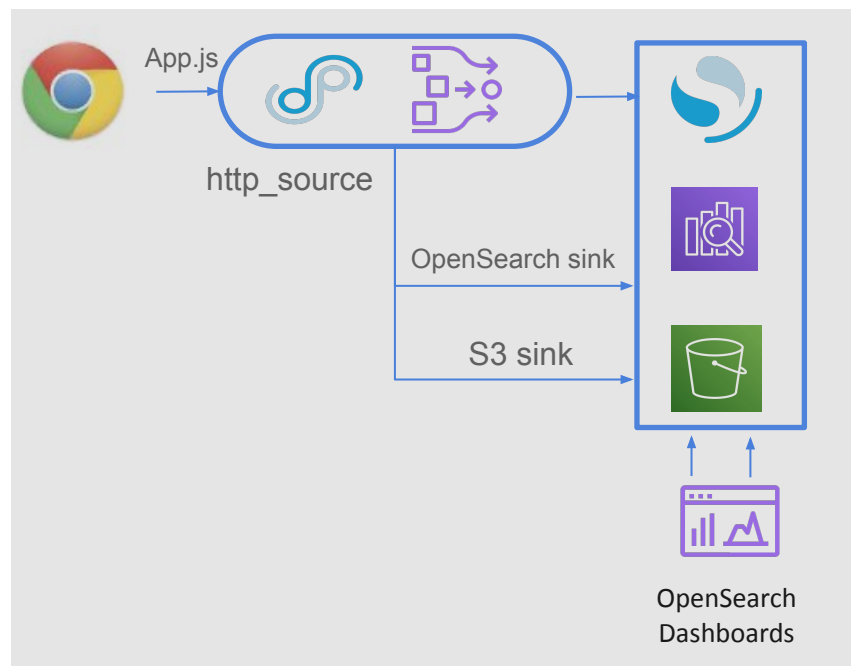
Increased Attractiveness for OS
Observability

Fraud Detection

Anomaly Detection

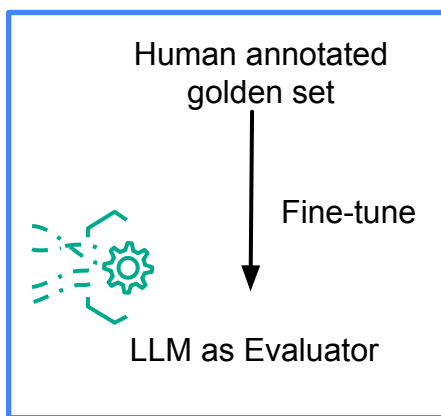
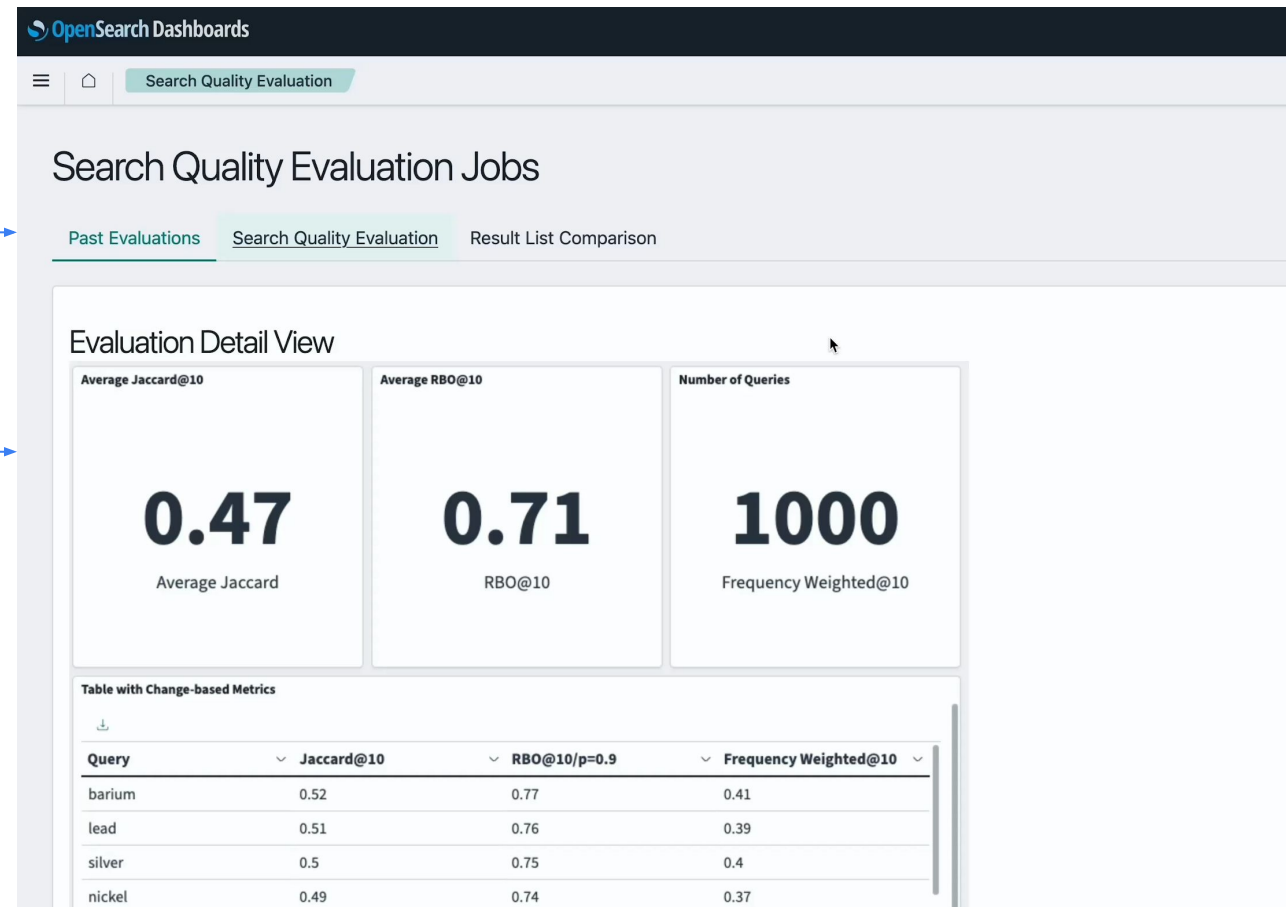
Search Relevancy workbench & Search Quality

Reference architecture 1



Collection of top queries

Implicit judgements

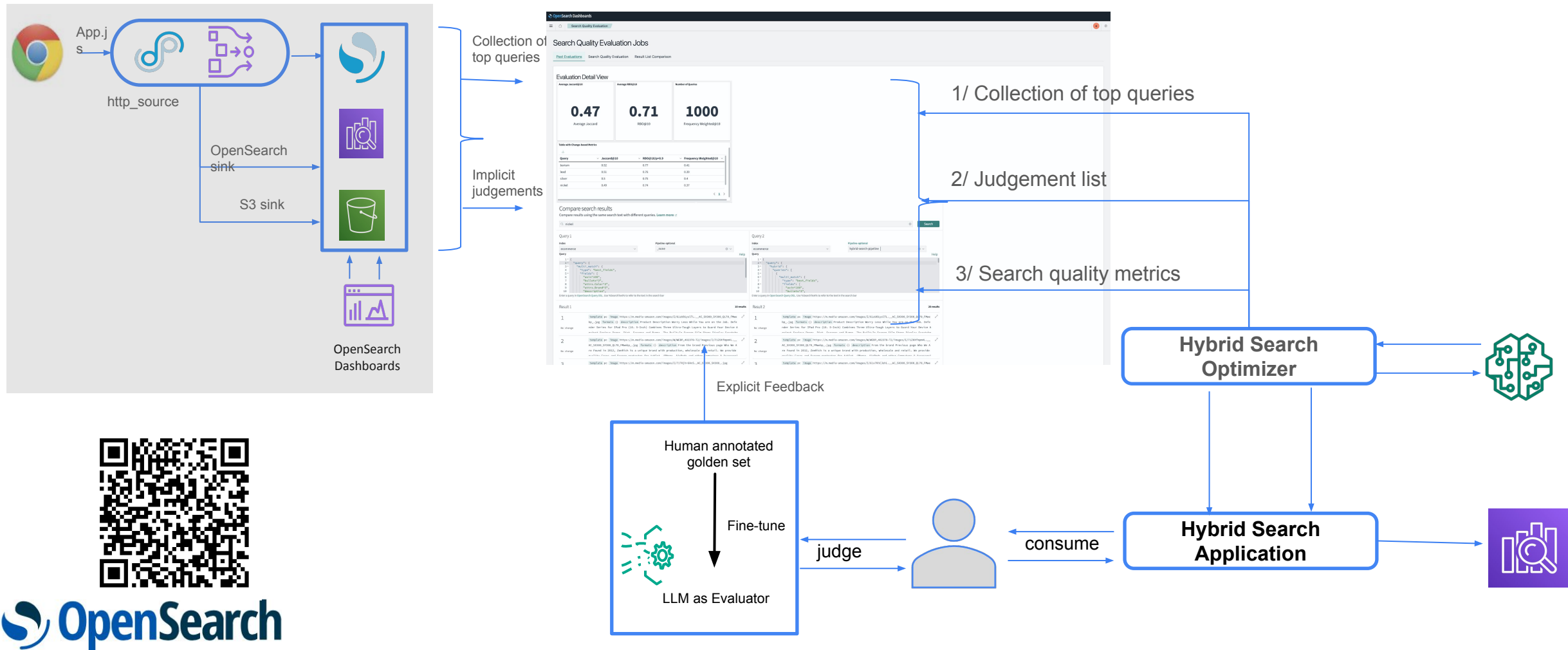


Explicit Feedback



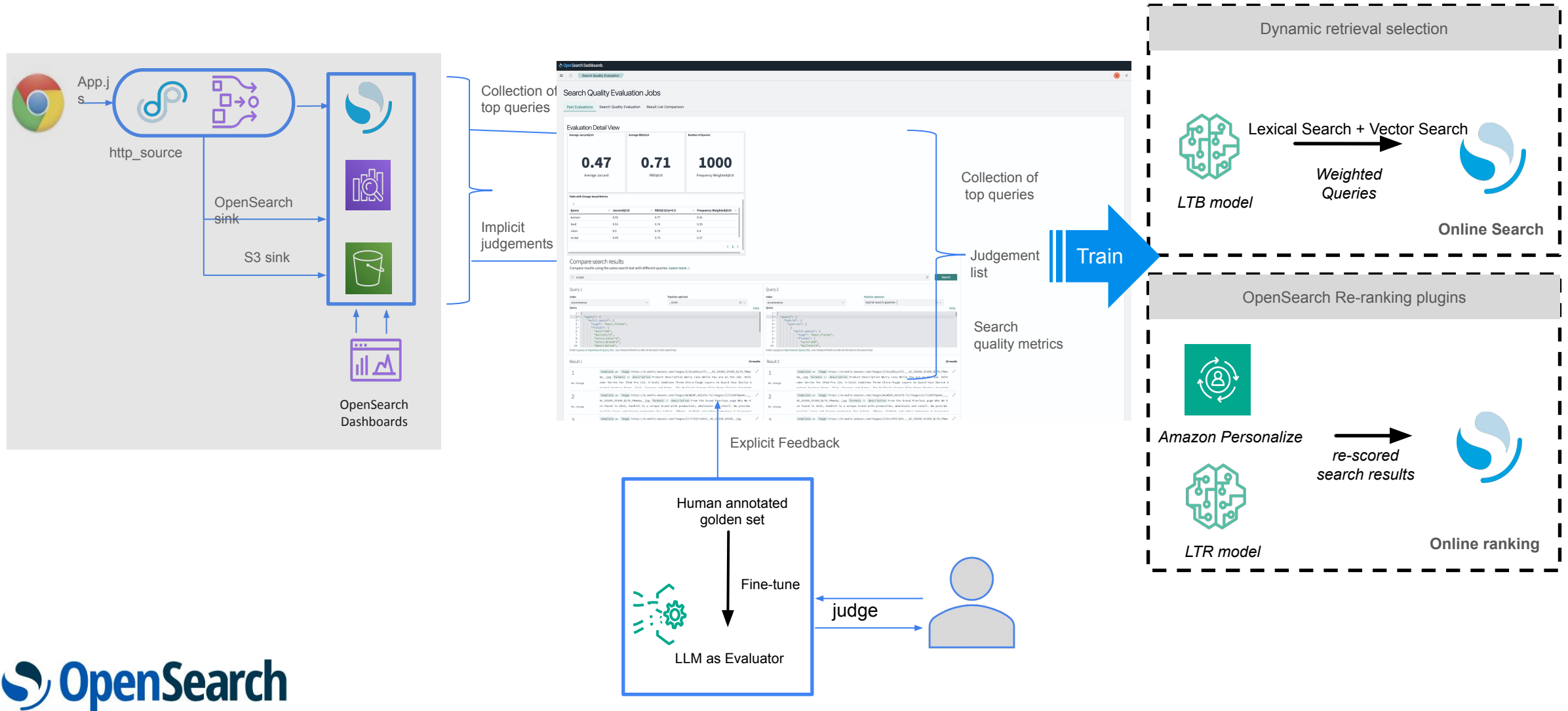
Optimizing Hybrid Search

Reference architecture 2

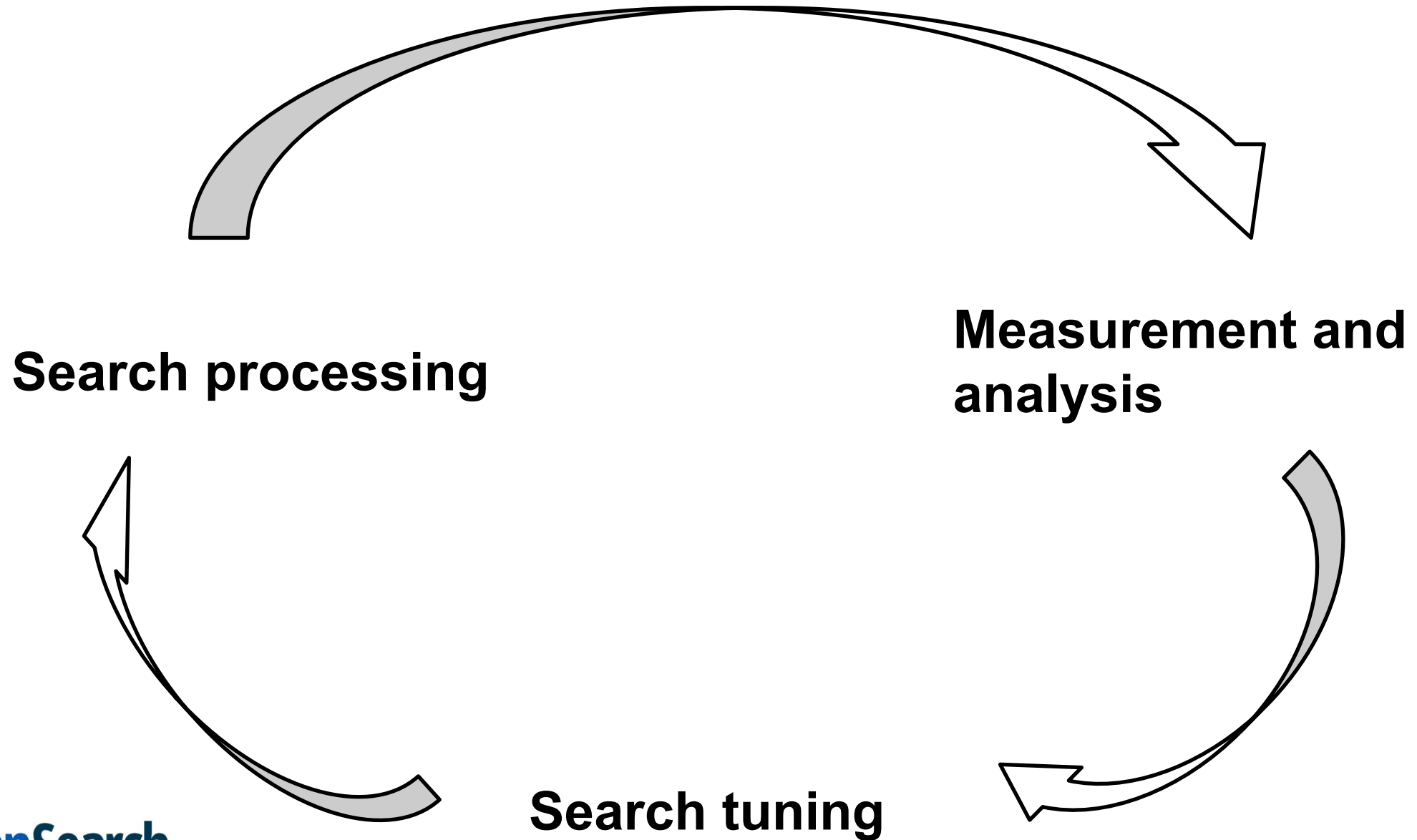


Integration with LTB & Re-ranking

Reference architecture 3



The virtuous circle of search improvement



The virtuous circle of search improvement

Search processing

- Pipeline processor
 - Modular query processing
 - Query rewriter
 - Modular result processing
 - Relevance integrations
 - Personalization integrations
- ✓ • Integrated semantic search
- Combining semantic and lexical (BM25)
- Sparse neural

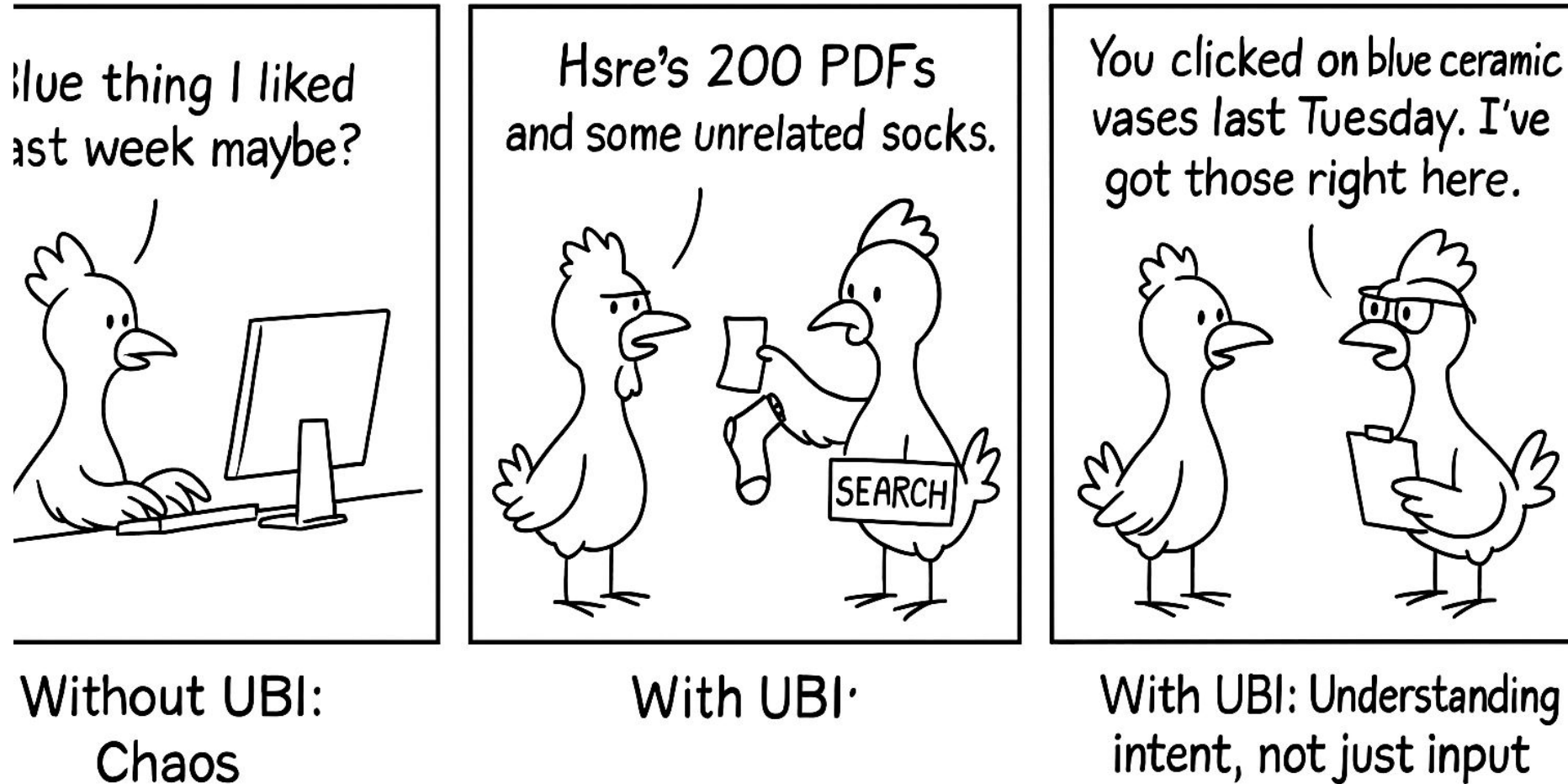
Measurement and analysis

- ✓ • End-to-end search behavior (UBI)
- ✓ • Online evaluation tools
- ✓ • Results eyeballing
- ✓ • Creation of judgement sets
- ✓ • Offline evaluation tools
- ✓ • Search Metrics – DCG, NDCG
- A/B testing

Search tuning

- Search configurations
- Regression evaluation
- Manual tuning, Bayesian optimization
- Semantic model fine-tuning
- ✓ • Learning to Rank

User Behavior Insights – Collecting and Analyzing Intent



Thank you!

Aruna Govindaraju

