

Finding relevancy

Improving your application search bar!

Alexander Reelsen alex@elastic.co | @spinscale

Today's goal

Learn about search, relevancy and the user perspective!

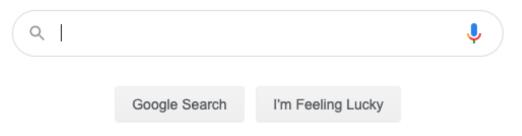




The ultimate search experience







Google offered in: Deutsch

Germany

Advertising Business How Search works Privacy Terms Settings





Advanced Search

Find pages with		7	To do this in the search box.
all these words:	foo	≜	Type the important words: tri-colour rat terrier
this exact word or phrase:			Put exact words in quotes: "rat terrier"
any of these words:			Type OR between all the words you want: miniature OR standard
none of these words:			Put a minus sign just before words that you don't want: -rodent, -"Jack Russell"
numbers ranging from:	to		Put two full stops between the numbers and add a unit of measurement: 1035 kg, £300£500, 20102011
Then narrow your results by			
language:	any language	•	Find pages in the language that you select.
region:	any region	*	Find pages published in a particular region.
last update:	anytime	-	Find pages updated within the time that you specify.
site or domain:			Search one site (like wikipedia.org) or limit your results to a domain like .edu, .org or .gov
terms appearing:	anywhere in the page	•	Search for terms in the whole page, page title or web address, or links to the page you're looking for.
SafeSearch:	Show explicit results	•	Tell SafeSearch whether to filter sexually explicit content.
file type:	any format	*	Find pages in the format that you prefer.
usage rights:	all	*	Find pages that you are free to use yourself.

Advanced Search

The ultimate search experience: Query intent

- stock crash 2008
- tomorrow's temperature
- barack obama age
- twitter
- cheap nike hoodie red xl



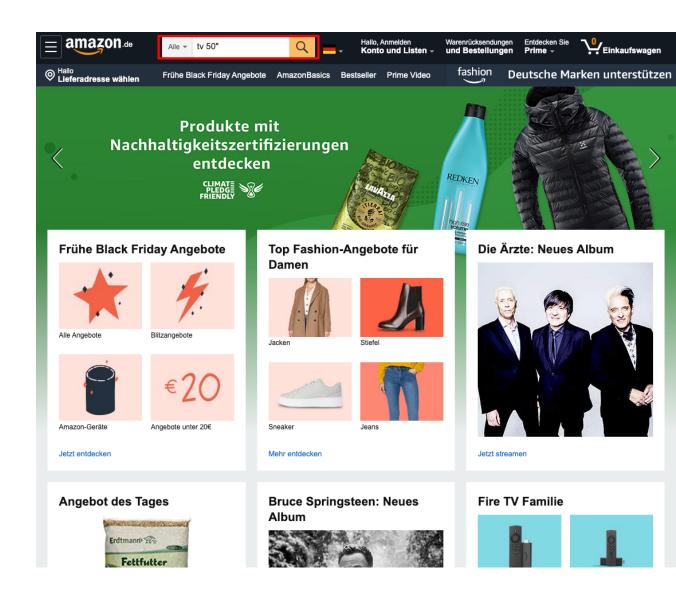
The ultimate search experience: Query intent

- Extremely generic
- Rarely needed for most web applications
- Many search use-cases are specialized
 - E-Commerce
 - News
 - Social network



E-Commerce

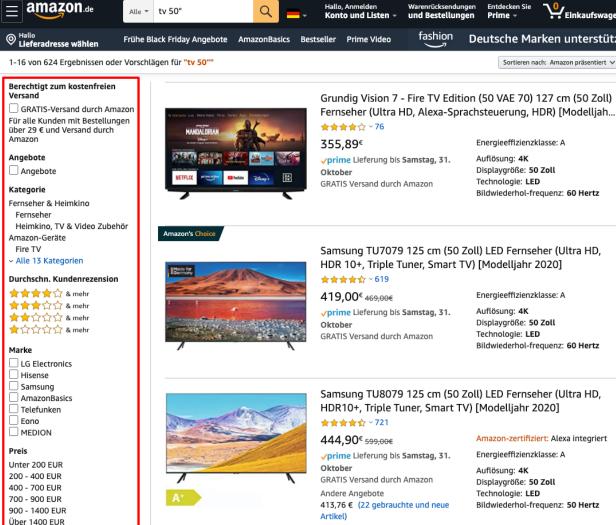
Start with a full text search





E-Commerce

- Start with a full text search
- Drill down via facetted navigation



Einkaufswagen

Deutsche Marken unterstützen

Energieeffizienzklasse: A

Displaygröße: 50 Zoll

Energieeffizienzklasse: A

Displaygröße: 50 Zoll

Bildwiederhol-frequenz: 60 Hertz

Amazon-zertifiziert: Alexa integriert

Bildwiederhol-frequenz: 50 Hertz

Energieeffizienzklasse: A

Displaygröße: 50 Zoll

Technologie: LED

Auflösung: 4K

Technologie: LED

Auflösung: 4K

Bildwiederhol-frequenz: 60 Hertz

Technologie: LED

Auflösung: 4K

Sortieren nach: Amazon präsentiert V

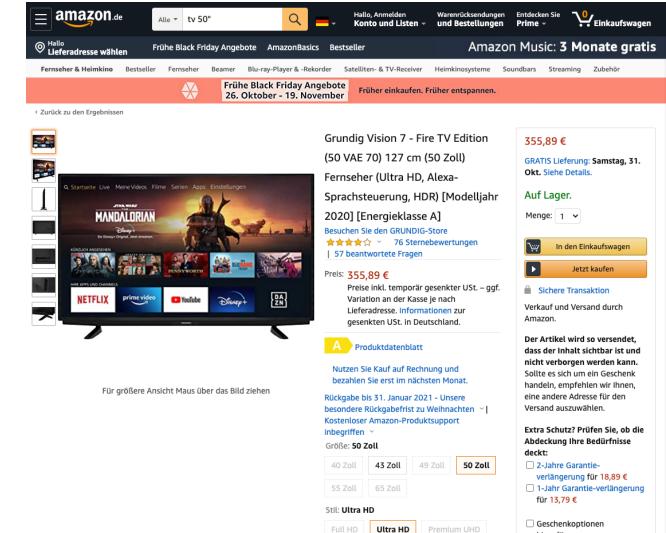
und Bestellungen

fashion



E-Commerce

- Start with a full text search
- Drill down via facetted navigation
- Select product



hinzufügen



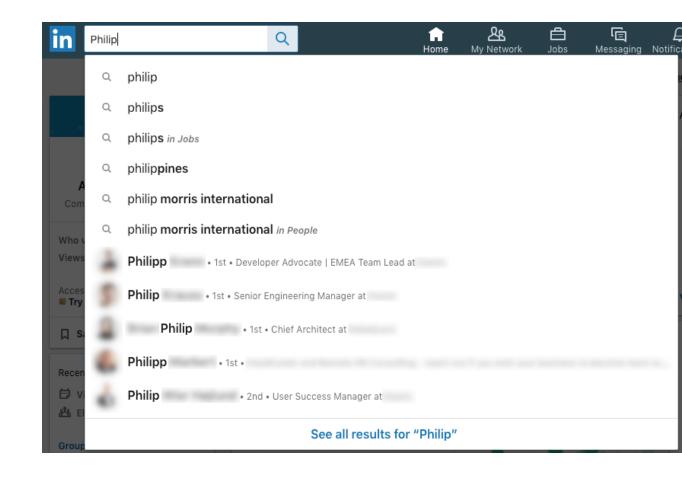
The ultimate search experience: Commercial Intent

- Scoring is never neutral, influencing factors (E-Commerce)
 - Popularity
 - Profit
 - Availability
 - Previous searches
 - Previous purchases
 - Cross/Upselling products



LinkedIn Search

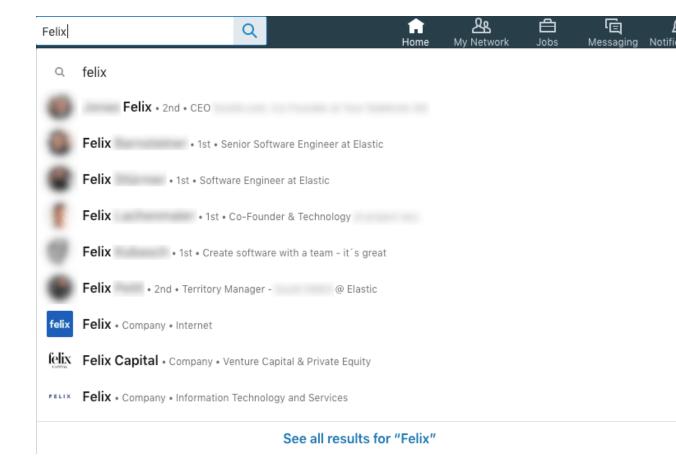
- Search for first name
- Search for last name
- Prefix search
- 1st degree > 2nd degree
- 2nd degree > 3rd degree





LinkedIn Search

 Rare surname more important than first name

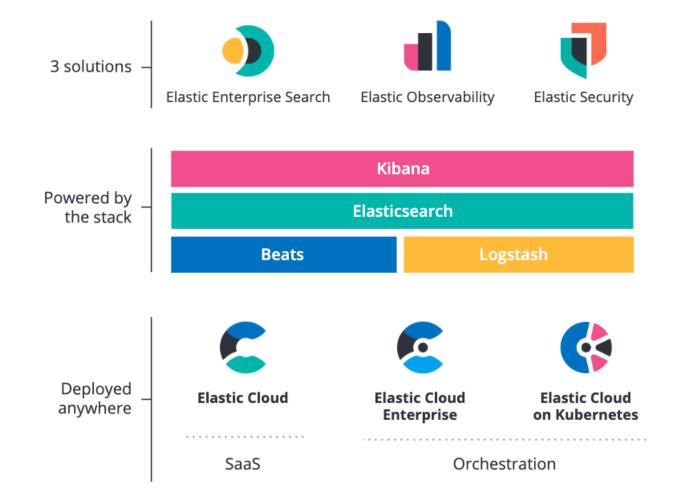




Let's build such a person search using Elasticsearch!

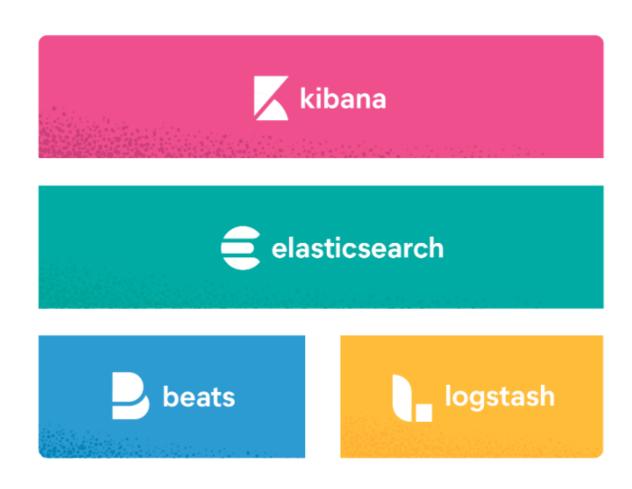


Product Overview





Elastic Stack





Elasticsearch in one minute

- Search Engine (FTS, Analytics, Geo), near real-time
- Distributed, scalable, highly available, resilient
- Interface: HTTP & JSON



Sample app: Social network

• Entity: Person

• Fields:

Matching: First & last name

Scoring: Employer & direct contacts

Meta: Title & image



Sample app: Social network

```
"name" : {
    "last" : "Reelsen",
    "first" : "Alexander"
},
    "employer" : "Elastic",
    "title" : "Community Advocate",
    "image" : "1",
    "contacts" : [ "philippk", "philippl" ]
}
```



Demo



Run node application

```
git clone https://github.com/spinscale/social-network-search-sample
cd social-network-search-sample
npm install
# create static build
npm run build
npm run start
# or for development mode w/ livereload
npm run build
```



Install Elasticsearch

```
# https://www.elastic.co/downloads/elasticsearch
wget https://artifacts.elastic.co/downloads/elasticsearch/elasticsearch-7.9.3-darwin-x86_64.tar.gz
# wget https://artifacts.elastic.co/downloads/elasticsearch/elasticsearch-7.9.3-linux-x86_64.tar.gz
# wget https://artifacts.elastic.co/downloads/elasticsearch/elasticsearch-7.9.3-windows-x86_64.zip

tar zxf elasticsearch-7.9.3-darwin-x86_64.tar.gz
cd elasticsearch-7.9.3
./bin/elasticsearch
```



Install Kibana

```
# https://www.elastic.co/downloads/kibana
wget https://artifacts.elastic.co/downloads/kibana/kibana-7.9.3-darwin-x86_64.tar.gz
# wget https://artifacts.elastic.co/downloads/kibana/kibana-7.9.3-linux-x86_64.tar.gz
# wget https://artifacts.elastic.co/downloads/kibana/kibana-7.9.3-windows-x86_64.zip

tar zxf kibana-7.9.3-darwin-x86_64.tar.gz
cd kibana-7.9.3
./bin/kibana
```

Point your browser to http://localhost:5601/



Profile



Alexander Reelsen Community Advocate at Elastic 3 Contacts



Contacts



Philip Kredible Team Lead at Elastic



Philipp Krenn Community Advocate at Elastic



Philipp Laughable Senior Software Engineer at FancyWorks



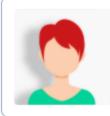
Search

Score first/last name, score own contacts higher



Show query

philip



Philip Kredible

Team Lead at Elastic

Score: 2.7641907



Admin

THIS IS NOT SECURE, DO NOT RUN IN PRODUCTION, BUT ONLY FOR DEMO PURPOSES. NEVER EXPOSE ELASTICSEARCH DIRECTLY SO IT IS REACHABLE FROM YOUR BROWSER

First, you need to ensure, that Elasticsearch is set up properly by allowing CORS. This requires adding the following settings to your configuration

```
http.cors.enabled: true
http.cors.allow-origin: /https?:\/\/localhost(:[0-9]+)?/
```

Delete & Reindex Data



Summary

- Coming up with a simple query is easy
- Figuring out user intents is hard
- Coming up with a proper query will take time and tries
- Analytics & metrics are vital



Next steps

- rank_feature: Boost score based on numeric field value
- vector : Similarity search based on a feature vector



? Can you quantify search experience ?



? How to measure search improvements ?



? Continuous feedback loop ?



Thanks for listening

Q & A

Alexander Reelsen
Community Advocate
alex@elastic.co | @spinscale





Resources

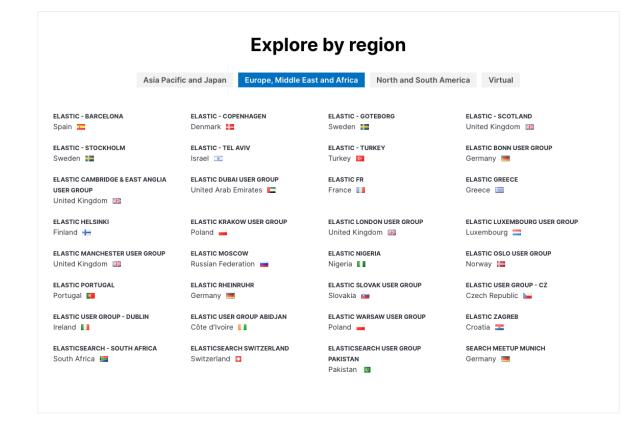
- GitHub Sample Repo
- rank_feature field type documentation
- vector field type documentation
- search_as_you_type field type documentation



Community & Meetups

https://community.elastic.co







Discuss Forum

https://discuss.elastic.co







Category	Topics	Latest	
Announcements Release announcements, end of life notifications and other bits about Elastic products that we think will be useful to	385 5 unread	■ ▼ Notes on Using These Forums ■ Meta Elastic	2 Apr 2017
everyone. Community Ecosystem		Couldn't push logs to elasticsearch using filebeat	1 3m
Any questions regarding Beats, forwarders and shippers for various types of data.	61 / week 1 unread 15 new	SarSeries configuration • Kibana	0 6m
■ Filebeat 1 unread ■ Packetbeat 1 new ■ Metricbeat 3 new 7 new ■ Winlogbeat 2 new ■ Heartbeat 1 new ■ Auditbeat ■ Functionbeat ■ Journalbeat ■ Beats Developers ■ Community Beats 1 new ■ Topbeat ■ Central Management ■ Elasticsearch	178 / week	Dec 15th, 2019: [EN] Elasticsearch Snapshot Lifecycle Management (SLM) with Minio.io S3 Advent-staging	0 7m
Any questions related to Elasticsearch, including specific features, language clients and plugins. Rally 1 unread	831 unread 36 new	Invalid IP network, skipping {:network=>"10.13.7.0/10.13.7.24" • Logstash	0 10m
Logstash Everything related to your favorite centralized logging platform, including plugins and recipes.	95 / week 29 unread 24 new	FScrawler stuck at 2.6gb index size • Elasticsearch	2 11m
Kibana All things about visualizing data in Elasticsearch & Logstash, including how to use Kibana and extending the platform.	113 / week 42 unread 19 new	Elastic APM Java agent - sanitize_fields_names on application/json* data APM java	1 21m
APM Everything related to APM – whether it is the APM Server, the Kibana dashboards, or the agents.	12 / week 5 new	Metricbeat Failed to connect EOF Metricbeat Mix free and paid licenses Elasticsearch license	5 22m 0 23m
Logs Everything related to the Logs app – setup with Filebeat, Filebeat modules, and using the Kibana Logs app.	55	Filebeat CPU utilization metrics are not normalized by default Beats stack-monitoring	2 23m
Metrics	1 / week	How do i aggregate these documets Logstash	6 26m
Everything related to metrics - Metricbeat, integrations and modules, Kibana dashboards and the Metrics app.		Metricbeat error	1 28m

Thanks for listening

Q & A

Alexander Reelsen
Community Advocate
alex@elastic.co | @spinscale



