

About @null (ノ◻ノ)



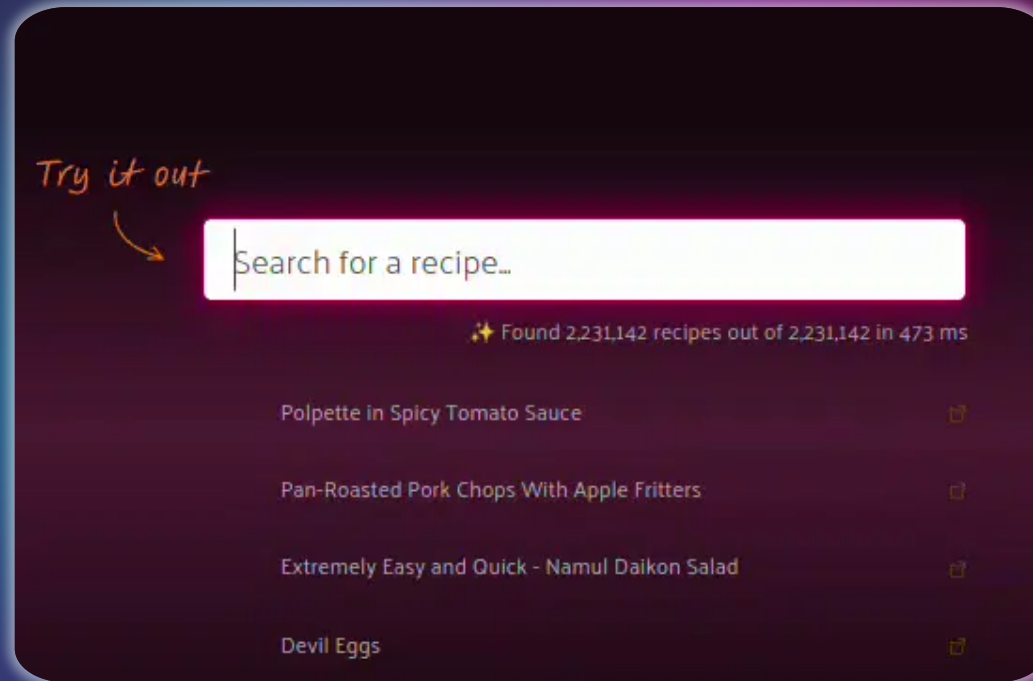
- Senior Developer, Catalyst IT, Wellington NZ
- Web development since 1996
- 20+ years PHP & search
- Loves corgis

TL;DR 🧠

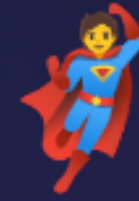
Typesense is a search engine for your website that does clever things much faster than anything else you've seen

TL;DR: DEMO 🤯

typesense.org



What you can do



🔥 Instantaneous sub-ms response times

- ✓ Faceting
- ✓ Sorting
- ✓ Curation
- ✓ Synonyms
- ✓ Scoped search
- ✓ Geospatial
- ✓ Decentralised management dashboard
- ✓ AI / Machine learning with GPU support
- ✓ Automatic schema detection
- ✓ Analytics
- ✓ Document text extraction
- ✓ Multi-language support
- ✓ Federated search

How does it work?

- Search as you type
- Peerable and distributed for redundancy
- Performance. Cost. Ease-of use

About Typesense

- Created in 2016 by Jason Bosco and Kishore Nallan
- Alternatives were hard to learn or expensive (or both)
- Written in C++

Typesense ❤️ FLOSS

- Free to use, GPLv3
- SaaS hosting available
- "CDN for extremely fast search"

Performance

Amount	RAM	Time
2.2M Addresses	600MB	~5 minutes
2.2M Recipes	900MB	3.6 minutes
28M Books	14GB	76 minutes
250 concurrent search queries per second		



Curated search results 🪜

- Promote one product over another
- Change order of specific search results
- EXAMPLE: Featured pages during a sale

Curated search results 🪜

- Exclude unsupported products
- Demote pages in specific search results
- EXAMPLE: Products out of stock

Synonyms

- "Silverstripe", "Drupal", and "Laravel" \Leftrightarrow "CMS".
- "smart phone" \Rightarrow "android", "iphone".
- Curations always precede synonyms

Built-in analytics 🤖

- Popular queries
- "No-hit" queries
- Track how often a particular result is clicked on
- Integrates with Matomo and Google



Is that it?

Typesense Dashboard

- Decentralised and offline, or locally installed
- Connects to any installation on Earth with CORS
- Search collections, update schema, server metrics



Federated search

- Concurrent searches in a single HTTP request
- Site-wide search with different result sets

Scoped search

- Safely offer authenticated privileged search
- Role-based access
- 1 user, 1 key, 1 collection
- 1 group, 1 key, 1 collection

Geospatial searches

- Radius, or within a polygon, or outside of them
- GIS applications, such as council plans, Homes.co.nz, etc

Language support

- Full unicode support, including te reo Māori
- non-Latin alphabets, (Cyrillic, Chinese, etc)
- Snowball Stemmer
- Stop-words for different languages



Text extraction with Tika

- Works perfectly with [silverstripe/textextraction](#)
- Extract text from documents
- Containerized, no dependency on Solr

AI and machine-learning

- Adds context to results.
- Index-time calculations.
- LLM reads and summarises.

AI and machine-learning

- Image search.
- Audio / video transcriptions.
- Huge win for accessibility



...My webserver needs a GPU? 🤔

- No! Typesense can use your CPU instead
- If you have a GPU available, it can use it
- Compatible with OpenAI, Google, and others
- Turn it off when you're done with it

Key takeaways

- Easy to deploy
- Save on development time
- Performance and user experience
- Cost savings

Time for demos!

- FindXKCD
- Paint colours to CSS variables
- NZ Address Lookup
- Audiobook search
- Federated search with text extraction
- Silverstripe Typesense module

FindXKCD



elliotsawyer - sawyer.nz - catalyst.net.nz

Paint colours to CSS variables



NZ Address Lookup



Audiobook search

- OpenAI Whisper produces a transcript
- Parsed line-by-line



Federated search

- Replaced CMS Search (one of them)
- Automatically searches every collection
- Multi-threaded search, near-immediate response

Silverstripe Typesense module

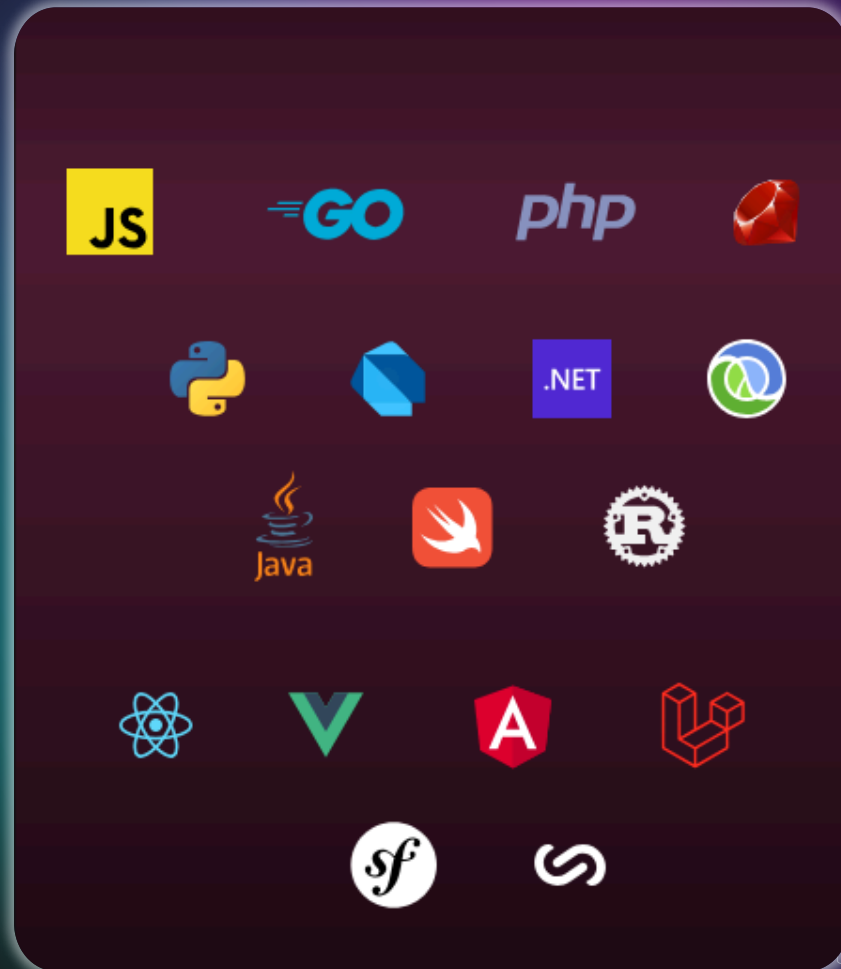
- Configure with YAML or inside of the CMS
- Syncs remote or local, auto-update on save
- Synonyms and translations
- Under development, contributions welcome!



AWESOME

How do you get started?

Batteries included!



Getting started



Very simple JSON based schema:

```
let schema = {
  'name': 'companies',
  'fields': [
    {
      'name': 'company_name',
      'type': 'string',
      'facet': false
    },
    {
      'name': 'num_employees',
      'type': 'int32',
      'facet': false
    },
    {
      'name': 'country',
      'type': 'string',
      'facet': true
    }
  ],
  'default_sorting_field': 'num_employees'
}
```

Automatic schema detection ✨

...or just let Typesense do it for you.

```
{  
  "name": "companies",  
  "fields": [  
    { "name": ".*", "type": "auto" }  
  ]  
}
```

Field names are regex-based, and so are facets; can provide as much or little detail as you want

Automatic schema detection ✨

Disorganised or poorly structured data set

Typesense will still create a working collection for you.

More info about Typesense

- typesense.org/
- cloud.typesense.org/
- github.com/typesense/typesense

Credits!

Rosemary	Mum and Dad	Lindsay D
Simon E	David M	Catalyst IT



Thank you!

- Slack: @null
- github.com/elliott-sawyer
- codeberg.org/0x
- linkedin.com/in/elliott-sawyer/
- hello@cashware.nz