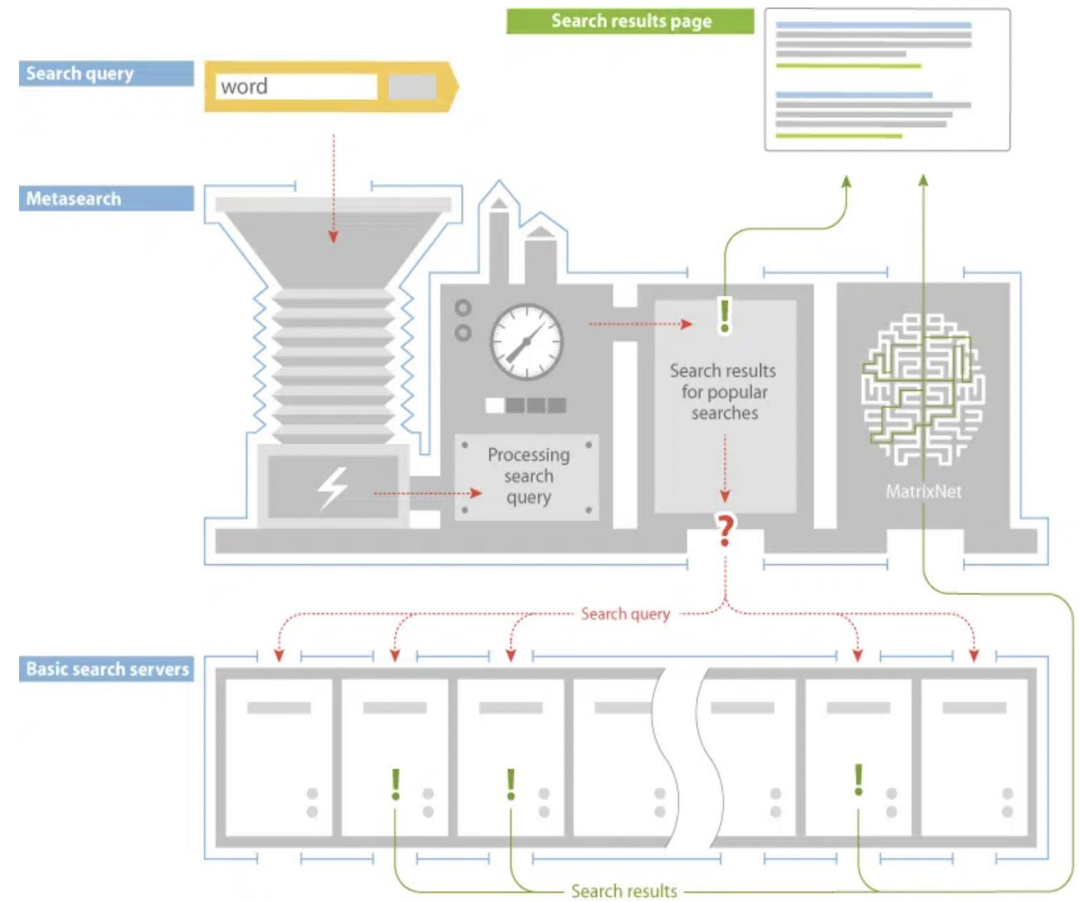


Yandex Signals

Signals

- Signals are factors that Yandex's search algorithm considers while ranking search result URLs
- They indicate the web page's quality/relevance
- Eg. Backlinks (links from other websites pointing to the page), social signals (likes/comments/shares), geolocation/language
- Yandex supposedly has more than 17k signals (17,854 assumed)
- The signals are ranked
- Page ranking factors can be static (eg. number of links leading to a page), dynamic (eg. based on the query words), or specific to search queries (eg. geolocation)

- Yandex has an intermediate layer (Metasearch) that serves cached results for popular queries
- If the results are not found there, the search query is sent to Basic Search (a series of thousands of different machines)
- Each builds a posting list of relevant documents then returns it to MatrixNet, Yandex's neural network application for re-ranking, to build the SERP (Search Engine Results Page)
- Computing the posting lists is the first place to consider the ranking factors
- MatrixNet builds formulas based on tens of thousands of factors, increasing relevance of results significantly
- MatrixNet can also adjust factors (custom ranking formula) for specific query class (eg. music)



Number of slashes in URL

- Yandex uses signal `FI_NUM_SLASHES` as a ranking factor
- Weight: +0.05057609417
- Possible idea:
 - The number of slashes is proportional to the depth of the web page (i.e. how far it is from the root/home page)
 - The further the page is from the home page, the lesser important it should be to the query

Incorporating NUM_SLASHES into Yioop

- As a bonus factor [under Page Options] to be added to Doc Rank scores
- **Declare** `NUM_SLASHES` **in** `configs/Config.php`
- **Add function** `findNumSlashes($key)` **to** `library/IndexDocumentBundle.php`
 - Where `$key` is the document URL
- **Add** `NUM_SLASHES/findNumSlashes($key)` **to** `getDocKeyPositionsScoringInfo()`
[`library/index_bundle_iterators/WordIterator.php`]

References

- <https://www.searchenginejournal.com/yandex-search-ranking-factors/477978/>
- <https://searchengineland.com/yandex-leak-learnings-392393>
- <https://yandex.com/company/technologies/>