

# Remove stop words during query transformation

## Context and Problem Statement

When querying for a title of a paper, the title might contain stop words such as “a”, “for”, “and”. Some data providers return 0 results when querying for a stop word. When transforming a query to the Lucene syntax, the default Boolean operator `and` is used. When using IEEE, this often leads to zero search results.

## Decision Drivers

- Consistent to the Google search engine
- Allow reproducible searches
- Avoid WTFs on the user’s side

## Considered Options

- Remove stop words from the query
- Automatically enclose in quotes if no Boolean operator is contained

## Decision Outcome

Chosen option: “Remove stop words from the query”, because comes out best.

## Pros and Cons of the Options

### Remove stop words from the query

- Good, because good search results if no Boolean operators are used
- Bad, because when using complex queries and stop words are used alone, they are silently removed

### Automatically enclose in quotes if no Boolean operator is contained

- Good, because good search results if no Boolean operators are used
- Bad, because silently leads to different results
- Bad, because inconsistent to Google behavior

