How to implement the full text search using only what PSQL offers

Jiri Cincura
jiri@cincura.net
http://blog.cincura.net
FT search

- Firebird does not support FT search
- 3rd party/hand made implementations
- Tokenizers, dictionaries, ...
FT search

* 3rd party products
  * i.e. Lucene/Lucene.NET
    * Store in FB or outside FB
* UDFs
* (P)SQL
  * Aka „poor mans FT search“
FT search

* WHERE name LIKE '%term%'
  * Not really FT
  * Cannot use index
* Define the needs
  * Names, products, ...
  * Full words
  * Anywhere-ish – from/to word boundaries
* Tokenize
* Index
* Search
FT search

* Step 1: Split the string on „spaces“
  * Tokenize
  * Regex to save us with [[:ALNUM:]] and collations
* Step 2: Process the pieces
  * LIKE 'term%' is index-searchable (prefix)
  * ❇️ we can do LIKE 'mret%' (suffix)
* Step 3: Build the inverted index
  * (term, ID)
  * Functional indices
* Step 4: Create SPs for searching (optional)