Краткое изложение факторов ранжирования

<u>Пред.</u> Рейтинг результатов поиска

Следующий

Краткое изложение факторов ранжирования

Таблица 5.1.

Имя	Уровень	Тип	Резюме
max_lcs	запрос	int	максимально возможное значение LCS для текущего запроса
bm25	документ	int	быстрая оценка ВМ25 (1.2, 0) без поддержки синтаксиса
bm25a (k1, b)	документ	int	точное значение ВМ25 () с настраиваемыми константами К1, В и поддержкой синтаксиса
bm25f (k1, b, {поле = вес,})	документ	int	точное значение BM25F () с дополнительными настраиваемыми весами полей
field_mask	документ	int	битовая маска совпадающих полей
query_word_count	документ	int	количество уникальных включающих ключевых слов в запросе
doc_word_count	документ	int	количество уникальных ключевых слов, сопоставленных в документе
LCS	поле	int	Самая длинная общая подпоследовательность между запросом и документом, прописью
user_weight	поле	int	вес пользовательского поля
hit_count	поле	int	общее количество вхождений ключевого слова
word_count	поле	int	количество уникальных совпадающих ключевых слов
tf_idf	поле	плавать	sum (tf * idf) по совпадающим ключевым словам == сумма (idf) по вхождениям
min_hit_pos	поле	int	первая совпавшая позиция вхождения, прописью, на основе 1
min_best_span_pos	поле	int	первая максимальная позиция диапазона LCS, прописью, на основе 1
Exact_hit	поле	bool	ли query == field
min_idf	поле	плавать	min (idf) по совпадающим ключевым словам
max_idf	поле	плавать	max (idf) по совпадающим ключевым словам
sum_idf	поле	плавать	сумма (idf) по совпадающим ключевым словам
Exact_order	поле	bool	все ли ключевые слова запроса были а) сопоставлены и б) в порядке запроса
min_gaps	поле	int	минимальное количество пробелов между совпадающими ключевыми словами в пределах совпадающих интервалов
lccs	поле	int	Самая длинная общая смежная подпоследовательность между запросом и документом, прописью
wlccs	поле	плавать	Взвешенная самая длинная общая смежная подпоследовательность, сумма (idf) по непрерывным интервалам ключевых слов
atc	поле	плавать	Совокупная близость терминов, log (1 + сумма (idf1 * idf2 * pow (distance, -1.75)) по лучшим парам ключевых слов