

■ 検索効率 (retrieval performance)

再現率と精度を総称して、検索がどのくらいうまくいったかを示す。検索性能とも訳され、検索有効性と同義。適合度順に検索結果が出力されるシステムでは、(全適合文献の)平均精度、11点再現率精度、上位ランクでの精度、R精度などの尺度を使用し、それぞれについて、テストコレクション中の全検索課題に対する値の平均値を用いる。

■ 検索有効性 (search effectiveness)

検索効率と同義。

■ 再現率 (recall)

検索対象文書集合中の検索要求に適合する文書の総数に対する、検索された適合文書数の比率。検索漏れの少なさを示す尺度。

■ 精度 (precision)

検索された文書中で、検索要求に適合する文書数の比率。検索ノイズの少なさを示す尺度。適合率とも訳される場合があるが、情報検索ではかつて「適合率 (relevance ratio)」という別の尺度があったので、混乱をさけるため本特集では「精度」に統一している。

■ 平均精度 (average precision over all relevant documents)

適合度順出力をするシステムにおいて、各検索課題ごとに、全適合文書の精度の平均を求めたもの。実際は、全検索課題の平均精度の平均 (mean average precision over all relevant documents) を算出する。11点平均精度は補間して算出するのに対して、「補間なしの平均精度 (non-interpolated average precision)」ともいう。

■ R精度 (R-precision)

適合度順出力をするシステムにおいて、当該検索課題の正解文書総数 (R) と同じ数だけ文書を出力した時点での精度。比較的上位の検索有効性を示す。

■ 11点再現率精度グラフ (11pt recall-precision graph),
11点平均精度 (11pt average precision)

11点再現率精度は、適合度順出力をするシステムにおいて、各検索課題ごとに、再現率が0, 0.1, 0.2, ..., 0.9, 1.0の11点における精度を算出した値。全検索課題の平均を算出し、再現率をx軸、精度をy軸としたグラフにする。システムの検索有効性の全体バランスを示す。再現率が0のときの精度は、直感的には適合度順出力で第1位の文書の精度と考えてさしつかえない。各検索課題ごとに、この11点での精度の平均を求めたものを11点平均精度というが、最近では、平均精度では、補間なしのものを用いる場合が多い。

■ ベースライン (baseline)

情報検索システムの検索性能評価は、同じテストコレクションを用いた異なるシステム間の相対評価として行う。基準となるシステムをベースラインとし、それに評価すべきアルゴリズム、パラメータの改良、要素技術などを組み込んだ提案手法システムと検索性能を比較し、ベースラインより何%検索性能が向上したかを示す。ベースラインシステムの検索性能が低すぎる場合、提案手法における検索性能改善の一般化可能性について疑問が呈されることが多い。

■ レlevance (relevance)

(検索された) 文書の検索要求への適合性。情報検索システムの評価において最も基本となる概念。システムに投入された

検索質問との適合性、利用者の内部にある情報要求との適合性、あるいは、トピックとして客観的に捉えられるものから、心理的、状況に依存したもの、システムとのやりとりの中でシフトするものなど多様な観点がある。

■ TREC (Text REtrieval Conference)

米商務省国立標準・技術院 (National Institute of Standardization and Technology; NIST) が1992年から、毎年、主催している情報検索システムの評価会議。この評価会議を通じて、大規模なテストコレクションが構築され、その多くは研究目的使用で公開されている。

■ トラック (track)

TRECの分科会。特定の検索や検索に関連する技術、情報検索システムの特定の側面などを取り上げて評価を行う。IREXやNTCIRの「タスク」と同じ。TRECでは、いままで、対話型、自然言語処理、フィルタリング、超大規模 (100GB)、言語横断検索、高精度検索、Q&A、Web文書などのトラックが企画され、関連技術について集中的に研究を促進する契機となってきた。

■ 語分割 (segmentation)

文書や検索質問のテキストから、検索システムが使用する索引語および検索語を切り出す操作。単語やフレーズに切り分ける方式と、1文字ずつずらしながら一定文字数の文字列に分割するNグラム方式などがある。

■ バイグラム (bi-gram)

Nグラム語分割の一方式で、テキストを、1文字ずつずらしながら、2文字の文字列に分割する方式。biは2の意味。たとえば、「検索システムの評価」は、検索、索シ、シス、ステ、テム、ムの、の評、評価に切り分けられる。文字種別、出現頻度によってNを変化するなどのバリエーションがある。

■ 言語横断検索 (CLIR: cross-lingual information retrieval)

検索質問とは異なる言語で書かれた文書も一括して検索する技術。translingual, cross language も cross-lingual と同義。

■ 多言語情報検索 (multilingual information retrieval)

二以上の言語の文書を含む文書集合に対する検索の総称。検索質問とは異なる言語で書かれた文書も検索できる言語横断検索機能を含む場合もあるが、あらかじめ文書データベースを翻訳し、検索質問を同じの言語の文書だけを検索対象とする「単言語検索の集合」を意味する場合もある。

■ 要約 (summary, summarization)

大意を保持したまま、テキストの長さを短くする処理、あるいはその結果のテキスト。

■ 報知的要約 (informative summary)

原文の代わりとして用いることを意図して作成された要約。

■ 指示的要約 (indicative summary)

原文の適合性を判断するなど、原文を参照する前の段階で、原文を参照すべきかどうかの判断材料として用いられることを意図して作成された要約。

■ 抜粋 (extract)

重要文抽出など、元のテキストから抜き出すことによって作成した要約。

