

問 1 を選択

まず、文書の主題をどのように表現するのかという問題がある。そして、文献データベース作成時に、文書の主題索引を作成する時に指定する主題表現(分類記号やキーワード)と、文献データベースの利用者が文献を検索する時に指定する主題表現を一致させることが可能なのかという問題がある。

主題索引の作成は、索引付与者(コンピュータによる自動抽出を含む)が文献(タイトル、抄録、本文)から、主題索引となる主題表現を抽出する。検索で使用する主題表現の定義が明確に定義される意味を持ち、かつ、文献がこれらの主題表現で代替できる主題を持つ場合に限り、適切に主題表現を抽出することが可能である。ただし、この前提がすべての文献に当てはまるわけではない(自然科学領域の文献は、専門領域内での語彙の意味が共有されており、主題表現の共通化はある程度可能と考えられものの他の領域の文献の主題表現の共通化は困難と考えられる)。

文献データベースの検索で使用する主題表現(分類記号やキーワード)の検索言語として、自然語にするか統制語(シーソラス)にするかという問題がある。自然語の場合、同義語や同型異義語、複合語の扱い等に難がある。統制語の場合、索引付与者による誤りが生じる可能性や新しい語に対応できない等の問題がある。