

# 小豆島町学校ICTレポート

## インターネット検索その①～検索エンジンとは

### ○Webページの規模

インターネット調査会社のNetcraft社によると、2021年7月現在で、世界には約12億のWebサイトが存在します。それぞれのWebサイトの下には数十から数万、場合によっては数千万以上というWebページが存在します。検索エンジン最大手のGoogleは約3兆のWebページの情報を持っているともいわれています。このほぼ無限ともいえるWebページから目的に合ったWebページを探すサービスを検索エンジンと呼びます。

※なお、インターネット上で使用されている言語で、日本語の占める割合は約2.4%（W3Techsによる2020年3月の調査）。

### ○検索エンジンの仕組み

検索エンジンはあらかじめ情報の索引（インデックス）をもっています。インデックスは、ロボット方式により自動的に生成されます。ロボット方式とは、クローラー（またはスパイダーやボット）とよばれるプログラムを使います。クローラーは、最初にくつつかのWebページを読み込み、それらの中にあるリンク情報からリンク先のWebページの情報を読みに行きます。これを繰り返して、各Webページに含まれている単語からインデックスをつくり、これをデータベースに保存します。

例えば「カブトムシ 捕まえ方」を検索したとしましょう。検索エンジンは、入力された言葉（キーワード）に似た言葉を探します（類義語検索）。「カブトムシ」なら「昆虫」「クワガタムシ」など、「捕まえ方」なら「捕獲」「採取」などです。次にデータベースからキーワードやキーワードの類義語が含まれているか、インデックスを探し、絞り込みを行います。見つかったインデックスに、キーワードが何回でくるか、キーワードがタイトルに使われているか、キーワードの類義語が含まれているか、他のWebページからリンクされている件数（ページランク）などによって採点し、総合得点から順位を決定します。検索エンジンはこれらの処理を、キーワードの入力から1秒以内で実行します。

検索順位は、キーワードからユーザーの要求を推測して計算の方法が変わることもあります。例えば、「高校野球の結果」というキーワードの場合、更新された日時が新しいものを上位に表示します。これは、多くの人が同じキーワードを検索し始めると、急上昇キーワードに登録されて、ページの更新日時が現日時により近いものを高い得点として計算するためです。また検索結果のパーソナル化も行われています。例えば、検索した人の位置情報から地理的に近いと考えられる情報が上位に選ばれます。



### 小豆島町の目標

全児童生徒が1日1時間はiPadを使った授業を受ける