

1 回目 情報検索の概要

1. 情報検索が対象とする情報

インターネットから得られる情報（Web 情報, Wikipedia, YouTube など）が、世の中に存在する情報の主体だと思っている人の割合が増加していると思われる。インターネットは手軽に利用でき、生活に役立つ情報を多く含んでいることから、個人にとって非常に有用な情報源である。半面、信頼性に欠ける情報が多くの割合で混在するとともに情報件数が多すぎて必要な情報にたどり着けないという問題点がある。

情報検索は情報収集の有効な 1 つの手段であるが、情報収集の本来の目的とは、情報の収集そのものではなく、収集した情報を活用することにある。従って、情報検索で対象とする情報は、利用することのできる信頼性のある情報であることが必要となる。

情報は人により創造され人により発信されることから、信頼性における情報とは、一般に以下のような条件を満たす情報である。

- ① 情報の作成者が明示してある。著者
- ② 著者の所属と所属する場所が特定できる。所属
- ③ 情報の発信日時が明示されてある。出版年, 雑誌の巻号ページ
- ④ 情報を伝達する出版物・メディアが特定できる。書名, 雑誌名
- ⑤ 情報の発信者と発信者の場所が特定できる。出版社, 放送局, 新聞社
- ⑥ 情報が評価を受けている。学术论文, 特許

上記のような情報の出所を示すものは書誌事項と呼ばれる。そして、これらの条件を満たす情報には、新聞記事情報、雑誌記事情報、図書情報、学術文献情報などがある。インターネット情報はこれらの条件を満たさないことから、判断・意思決定など個人が行動を起こすために使用する情報としては問題が多い。Wikipedia も集合知としての価値を有するが、反面、作成者を明示できないことから責任者が不在となり、レポート作成のために使用する情報としては適当でない情報を含む。Wikipedia の内容を参考にするに問題はないが、レポート作成などの論拠に使用するのは不適當である。少なくとも問題が多いことを前提に使用する必要がある。

情報検索の授業では最初にインターネット情報を扱うが、上記の条件を満たしインターネット情報より信頼性が高いと考えられる以下の情報を主に扱う。

- | | | |
|--------|-----------|-------|
| レポート 1 | 新聞記事情報 | |
| レポート 2 | インターネット情報 | |
| レポート 3 | 新聞記事情報 | 有料 DB |
| レポート 4 | 図書・雑誌記事情報 | 有料 DB |
| レポート 5 | 英語情報 | 有料 DB |

有料DBは有料のデータベースである。信頼性の高いデータベースはメンテナンスに費用がかかることから有料での利用が一般的である。英語情報は、日本語情報より質量ともに圧倒的に勝っていることから、英語の新聞記事情報、英語の雑誌記事情報、英語の図書情報、英語の学術文献情報などを、英語のキーワードで検索する。

2. 情報検索の意味

情報検索を英語で表現すると information search と訳す人が多いのではないと思われるが、情報検索は英語で information retrieval と表現される。retrieval には回収する、回復する、取り戻すという意味が含まれており、information retrieval は、情報を回収するという意味を含んでいる。情報検索により情報の存在場所を捜し出して、さらに情報を利用するために存在場所から、元となる情報、つまり原文を情報検索者の手元まで持ってくることにより（原文の回収で）情報収集が完結する。情報検索は情報収集の前段で情報の存在場所を捜し出す役割を担うことから、情報収集のための有力な1手段といえる。収集する原文は、オリジナル情報、1次情報、原著などと呼ばれる。

information retrieval

情報を捜して原文を持ってくる、(回収する、回復する、取り戻す)

information search

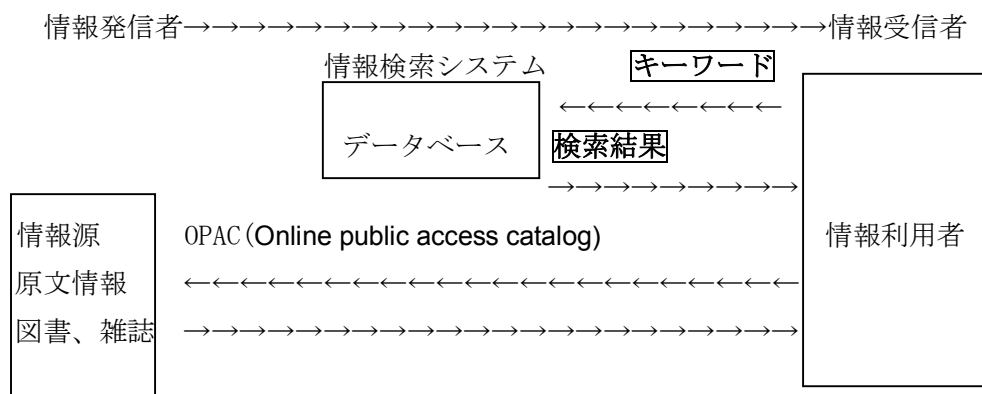
情報を捜して求める、(探る、調べる、たどる)

3. 情報検索の特徴

情報は情報の発信者に主体があり、発信者からメディアを経由して情報の受信者に届くというのが通常概念である。情報検索では、情報発信者とは無関係に、情報の利用者である情報の受信者が、情報発信者の創造した情報の所在を明らかにして、自らの手元に原文を回収する。したがって情報発信者に主体がある通常情報の流れとは逆になる。

一方、インターネットで検索エンジンを使用して情報検索を行う場合には information search という言葉が使用されることが多い。情報を探る、調べる、たどるといった意味を search は持っており、information search の意味は必要な情報がどこにあるか探し出すという意味を含んでいる。インターネットの情報を検索する場合には、検索エンジンも検索結果の原文情報も同じインターネット上に存在する。従って、検索エンジンによる情報検索結果の表示画面からリンクにより原文まで容易にアクセスすることが可能になっている。検索エンジンでは原文情報までのアクセスも含むことから information retrieval ではなく information search でも良いことになる。しかしWebのページには原文あるいは1次情報とはいえない内容のものが多いので、オリジナルの情報を捜したことにならない場合も多いので注意が必要である。

【情報収集と情報検索】



インターネットの Web 情報を対象とする google や yahoo の検索エンジンや図書館の蔵書検索 (OPAC) などが身近で利用するデータベースである。有料のデータベースの場合はプロデューサと呼ばれるデータベース作成機関が、原文情報から一定のフォーマットに従って 2 次情報を作成しデータベースを作成している。データベースは本来、情報検索システムに含まれるプロデューサにより作成された個々のデータベースファイルを意味するが、情報検索システムそのものをデータベースと呼ぶことも多い。また、システムに含まれる全てのデータベースを、単にデータベースと呼ぶこともある。

情報検索とは検索システムを利用して原文情報がどこにあるかを捜すことである。データベースの情報は基本的に要約などの書誌事項がテキスト情報として整理されている。原文情報の一部分にすぎないのでデータベースの検索結果から概要は理解できても、原文の意味する正確な情報は入手できないことがほとんどである。正確な情報を入手するためには、情報検索の後に原文を入手するための作業が不可欠となる。原文情報の入手には図書館を利用したり、文献取り寄せサービスを利用したりする。

4. 原文 (1 次情報)

1 次情報とは、情報の発信者が創造した情報を言う。テキスト情報の場合は**著者**が作成した情報で、図書、学術文献、特許などの著作物が代表的なものである。著作権法では、著作物を、「思想又は感情を創作的に表現したものであって、文芸、学術、美術又は音楽の範囲に属するものをいう」と定義されており、独創性が含まれる。1 次情報は著作物にかぎらず基本的に、情報発信者の独創性、つまり情報発信者が独自に考えた内容が含まれた情報である。オリジナル情報という表現も同じように独創性が含まれるという意味を持っている。情報社会は情報によりさらに新しい価値を創造する社会だと考えることができるが、新しい価値の創造には、創造性を含んだ 1 次情報が必要である。

1 次情報に対して 2 次情報が存在する。2 次情報は、1 次情報を集計、加工したもので、1 次情報を書誌事項で整理したデータベースや抄録誌が代表的なものである。情報検索では、ほとんどの場合、使用するデータベースに整理された 2 次情報である書誌事項を使っ

て情報を検索する。情報検索では2次情報を使って、1次情報を収集することになる。また、2次情報はほとんどがテキスト情報であるので情報検索にはキーワードが使用される。

さらに、2次情報を加工したものは3次情報と呼ばれる。2次情報を集めて傾向等进行分析したもので特定の話題に関する概要を知りたいときに使用される。

他人の情報を単に伝達した伝聞情報は、情報検索で本来対象とする情報ではないが、Web情報には著者が不明で、著者が自分で考えたり確認したりした情報ではなく、伝聞情報に類似する情報が多いので、検索エンジンを使った場合には区別されずに結果が表示されるので、検索結果から伝聞情報を取り除き、著者が自ら創造した1次情報を取り出す必要がある。

5. 書誌事項

キーワード検索の対象項目に相当し、通常のデータベースでは項目を限定した検索が可能である。書誌事項は個別の情報を特定するために必要あるいは効果的な事項と言える。

情報の内容そのものを表示する項目として、表題と抄録(アブストラクト)がある。一部のデータベースでは全文のテキストも検索項目になっているが、キーワードが多くなりすぎることから、詳細な検索以外には不向きである。

書誌事項には内容以外に、付帯情報を示す書誌事項が存在する。著者、所属、表題、抄録、出版年月日、雑誌・図書名、巻・号・ページ、分類などが代表的な付帯情報をあらわす書誌事項である。書誌事項は情報の意味を伝達しないが、情報の所在を明示すると共に、情報の信頼性を保障あるいは推定するものである。

インターネット情報には著者をはじめとして書誌事項がないものが多い。検索エンジンのデータベースでは、HTMLの<title>部分を表題に、ページの最初の文字列を抄録に、URLを所属あるいは著者の代わりに使用しているが、情報の信頼性を担保できる本来の書誌事項ではない。

6. 情報検索に必用な要素

情報検索を実行するために必用な要素は、データベース、情報検索システム、検索課題である。

1) データベース

書誌事項をコンピュータで処理可能にしたテキストデータの集合体である。

2) 情報検索システム

データベースをコンピュータで利用可能にした検索システムである。Web情報を対象とした、Google、Yahooなどの検索エンジン、新聞記事情報、学術文献情報を対象とした商用データベースなどがある。

3) 検索課題

情報収集の目的を認識した上で検索課題を確定させることが必要である。情報の使用目

的が不明確なまま情報収集を行なうと、収集した情報を有効に活用することができなくなる。

7. データベース

データベースは軍事基地の意味で、1950年代にアメリカ軍の資材などのデータを管理する基地の意味で誕生したといわれている。例えば沖縄の米軍基地やグァンタナモ米軍基地と同じ基地の概念であり、軍隊がいる通常の基地のようにデータを集めて管理する部門を一つの基地とみなしたと考えられる。現在では、「データベースはコンピュータに一定の形式で蓄積されたデータ集合体であり、様々な方法で閲覧したり使用できたりするものである (Oxford 現代英英辞典)」と捉えられ、コンピュータの使用が前提となっている。

データベースとして、朝日新聞記事データベース (朝日新聞社)、雑誌記事索引 (国立国会図書館)、日本特許情報データベース (特許庁)、MEDLINE () などがあ
る。() の中は、データベースの作成期間でデータベースのプロデューサーと呼ばれる。

8. 情報検索システム

Google、Yahoo などの検索エンジン、日経テレコン (日経四紙)、朝日DNA (朝日新聞)、日外Web (雑誌記事索引)、NDL-OPAC (国立国会図書館蔵書検索)、ProQuest (主要な世界の専門誌・学会誌 原文情報を含む) などがある。これらのシステムを開発・運営・管理している機関をディストリビュータという。ほとんどが Web 経由で利用可能であり、ネットワークを介した利用が前提となっている。通常1つの情報検索システムで、複数のデータベースが検索可能である。

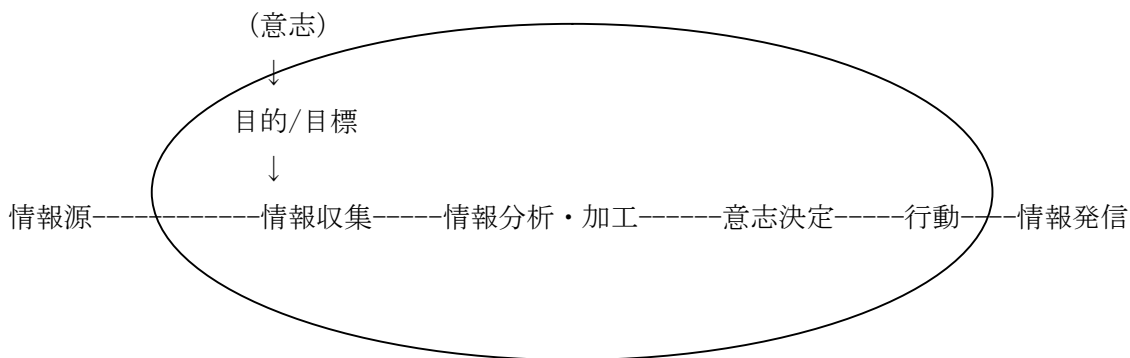
9. 検索課題の設定

情報は本来特定の目的を達成するための判断材料や、判断をするための基礎知識を得るために使用される。情報の使用目的や目標が明確であればあるほど、情報収集を効果的に行なうことができる。また、情報の使用目的や目標が明確であればあるほど、収集した情報をもとに目的を実現するための実行方法を考えたり、目的を達成するためにどのように実行すべきか判断したりすることが容易になり、情報を効果的に利用することができる。

情報収集の手段としては情報検索の他にも、定期刊行物の購読、専門家や友人へのヒアリング、自ら行なう研究や調査 (アンケート調査など) など多くの方法が考えられる。しかし短時間で、目的にあった多量の情報を収集できることから情報検索は情報収集の有力な手段である。特に情報収集の初期の段階で有効である。情報検索は、人が目的を達成するために必要な情報の収集を実行する際の非常に有効な1手段であるといえる。

人が情報を活用する場合の情報の流れは図に示すように、データベースなどを使って情報発信者あるいは情報源から、情報や原文情報を収集し、収集した情報を情報の利用目的に合わせて分析・加工して、目的を達成するための判断材料として利用するといった順に

なる。情報検索やデータベースはこの中の情報収集の段階で利用する有効な手段であるが、あくまでも有効な手段であり、情報の持つ内容を活用するためには情報収集の目的や目標が明確であることが重要である。目的や目標が不明確であると、情報の収集範囲が拡大するにもかかわらず、有効な情報を集めることが困難になる。極端な場合は情報収集の途中で目的や目標が変化し情報収集を再度行なわざるを得ないことになる。



10. 情報検索を利用した情報収集の目的

情報検索は情報収集の1つの手段である。目情報検索の意義は、①多くの必要な情報を収集し、多くの目的解決手段を検討することにより、より正確な判断を行うことができる、②判断をして行動を起こす前に過去の事例や知見などを収集し、スタートレベルを引き上げることにより、目的に早く到達することができる、③関連する情報を入手し、自己の知識に取り入れ、自己の知識レベルを向上させることかできる、④参考となる最新データを得ることにより競争を有利にすることができる、といった効果が期待できることである。情報検索は大量のデータから容易に情報収集を行なえることから、情報収集を効果的に行うための非常に有効な1手段であるといえる。

学問を行う場合に自分で考えることが重要であるが、ゼロから全て自分で考える必要はなく、既にある知識を活用すれば、既に自明になった地点から自分で考えることを開始すればよい。既にある知識を活用し行動のスタートレベルを上げることができれば、早く優れた結果に到達する可能性が高くなる。

