

# 専門用語のクラスター分析による ICT 関連用語の動向把握

長崎大学教育学部 全 炳徳, 白川美織

## 1. はじめに

私たちの暮らす現代社会は、「情報化社会」という呼び名をもつ。特に 1990 年代以降の情報通信技術の急速な進歩により始まったインターネット<sup>1)</sup>は、近代以降の新時代を切り開いていたともいえる。更に、このようなネットによる情報化社会を支える携帯電話の進化も目覚ましいものがある。今や誰でも、いつでも、どこでも情報を手に入れることが実現され高度情報化社会がもうすぐそこまで来ているといえよう。このような情報化は私たちの身の回りの生活の利便性だけでなく、教育現場においても大きな変革をもたらしている<sup>2)</sup>。

文部科学省では教育の情報化を推進し、「世界最先端の IT 国家宣言」のバックアップを進めており、これを教育の情報化に関する基本方針として挙げられている。その中のひとつにある『教育の情報化ビジョン』の概要では、生きる力と共に情報活用能力が、21 世紀を生きる子どもに求められる力として述べられている。このことから「教育の情報化」に対する関心が年々高まっており、教育現場が情報化の重要なキーワードになっていることが分かる。更に、研究分野においても今後ますます「教育の情報化」への関心が高まることが予想される。

さて、近年の教育分野に多用される「ICT」は「情報コミュニケーション技術」をさす用語であるが、主に教育分野で使われていたものの、最近では「IT」の代わりの用語としても利用されるようになってきている。このような ICT 関連用語の時代的な変遷はどのように調べれば良いのだろうか。そこで、本研究では教育研究機関（日本学術振興会及び公益財団法人パナソニック教育財団）のデータベースとして公開されている研究テーマを基に、ICT 関連用語の変遷を追跡してみた。結果から、日本が歩んできた情報化関連用語を整理すると共に、今後注目すべき ICT 関連用語のキーワードを分析した。

## 2. 専門用語の評価手法と定義

ここでは、本研究で説明する「専門用語」についてやや詳しく詳述する。ここで説明する「専門用語」とは、日本学術振興会及び公益財団法人パナソニック教育財団においてそれぞれのホームページにおいて公開しているデータベースにある研究タイトルから、キーワード自動抽出システム「言選 Web」を用いて抽出したキーワードのことを指す。これらの抽出された専門用語には、重要度という値が付けられる。重要度の計算方法は次項の通りである。

$$\text{重要度} = \text{用語の出現頻度} \times \sqrt{\left\{10^{\wedge}(\text{前の語に接続した回数} \times 2)\right\} \times \left\{10^{\wedge}(\text{後の語に接続した回数} \times 2)\right\}}$$

接続回数のカウント方法は「延べ数」，「種類数」，「情報理論的な回数」で行われている．つまり「語の接続」が「多くの語」に分散していれば，カウントが多くなるといえる．以上の計算方法で出された重要度の数値が高いものほど，本研究のデータ分析で重要とされる専門用語だと定義している．

### 3. 「専門用語」の抽出とクラスター分析

#### 3.1 専門用語の抽出 - 言選 web の重要度について -

ここでは，重要度を抽出したキーワード自動抽出システム「言選 Web」について簡潔に述べる．「言選 Web」<sup>3)</sup>の処理は次の手順で行われている．まずはテキストボックスに入力した文章は，MeCab システムを用いて文章を形態素に分解し，用語を組み立てる．更に，語になりえない文字を消去し，残ったものを取り出す．次に，重要性の高い順に並べかえる，という手順である．

重要度の計算に用いている方法は FLR 法であり，用語は単名詞そのものか，複数の単名詞を組み合わせて作られることになる．この理論では，他の単名詞と連結して複合語をなすことが多い単名詞ほど，文書中で重要な概念を示すと考えられている．例として，「漢字文献情報処理」を考えてみると分かりやすい．この語は次のとおり 5 つの単名詞に分割できる．それぞれの単語が他とどれだけ結びつくかを文章中から統計をとり，次のとおりわかったものとする．

表 1 「漢字文献情報処理」の単名詞分解・前後の語との結びつき

| 単語 | 前の語に接続 | 後の語に接続 |
|----|--------|--------|
| 漢字 | 2      | 3      |
| 文献 | 3      | 4      |
| 情報 | 4      | 5      |
| 処理 | 2      | 0      |
| 研究 | 0      | 3      |

用語の重要度はこれらの  $10^{\wedge}$  (単名詞  $\times 2$ ) の数値の平均から求める．平均値は，相乗平均が相和平均よりもより効果的なため，相乗平均を用いている．なお，相乗平均をとる際に接続した回数が 0 回 (表 1 の「処理」「研究」) の単名詞に対応するため，実際には各回数に 1 を加算した値を用いている．こうして得られた単名詞の接続情報に，用語の出現頻度をかけたものが「言選 Web」から出力される重要度である．出現頻度 (Frequency) に左 (Left) と右 (Right) の語の接続情報を組み合わせて使うため，これを FLR 法と呼ぶという．



図2 「言選Web」から抽出した重要度の結果表（エクセル）

1975～2014年度までのデータベースから抽出した研究論文テーマから得られた重要度としてヒットされた単語は、パナソニック教育財団では1857語に及んだ。本研究では「ICT教育の関連研究の動向」を分かりやすく端的に示すために、語の選別を行った。選別方法は次の基準で実施した。(1)重要度の計算上最低値は1になるもの、つまり0と同等の扱いの語はすべて削除する。(2)その語から情報化とICT教育を連想する語のみを選び、ICT関連研究を連想させないと思われる語については除外する。以上2つの選別方法により得られた結果、パナソニック教育財団では248語、日本学術振興会では225語のICT関連用語を得た。

### 3.2 クラスタ分析

本研究では重要度を示す値がデジタル値であることを利用して、ICT関連用語の動向を調べるために、重要度に基づくICT関連用語のクラスタ分析を行った。クラスタ分析とは異なる性質のものが混ざりあっている集団(専門用語)の中から互いに似たものを集めて群落(クラスタ)を作り、対象を分類しようという方法の総称である。このクラスタ分析を用いると客観的な基準に従って科学的に集団の分類ができ、また視覚的にも結果の見える化が可能となる。クラスタ分析の手順は①データの標準化、②類似度測定、③樹形図作成、④濃淡図作成順で行われた。

#### 3.2.1 データの標準化

重要度はFLR法の計算特性上、入力文章が多いほど分割される単名詞数が増える傾向があることから、研究件数が多い年度は重要度が全体的に高くなっている。そこで、ICT関連用語の年代ごとのクラスタ分析の結果に偏りがないようにするため、その年代の研究母数を用いて標準化を行った。標準化を含むクラスタ分析の作業は、R言語のツールを用いており、紙面上詳細の説明を省略する。

#### 3.2.2 専門用語の樹形図・濃淡図

R言語によるクラスタ分析の結果は大きく樹形図と濃淡図に区別できる。

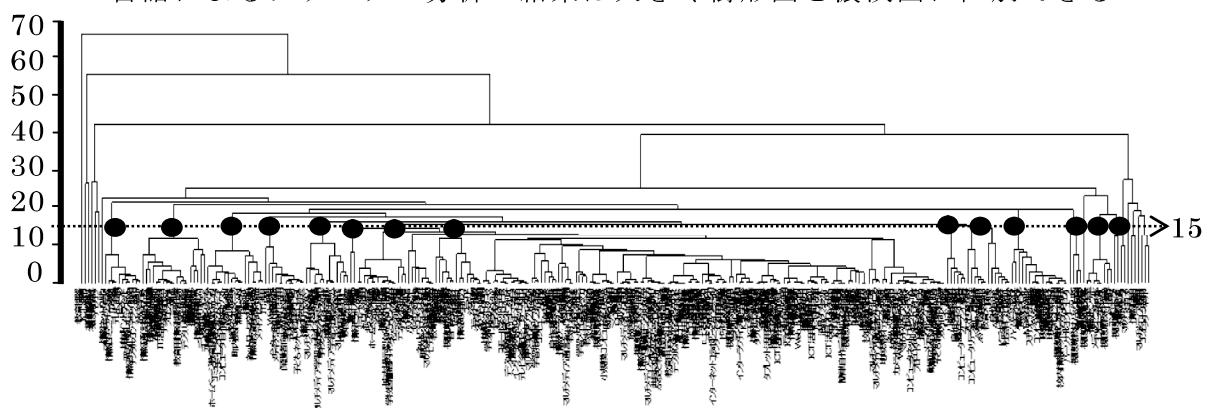


図3 パナソニック教育財団の248語に対する樹形図結果

図3を見ると、クラスター分析結果の高さの値が0から70まで分布しており、そのうち15の値のところを破線として示している。更に、この破線に沿って黒丸の標識を示されてある。それは高さ15を基準にしてみると、少なくとも14種類以上のクラスターが分類できることを示しており、それぞれのクラスターごとに独自の専門用語がぶら下がっている。各黒丸にぶら下がっている用語は専門用語の重要度に沿って、それぞれが独自性を持っており、各自のクラスターとして分類されている。

本研究では図3の全体像から黒丸で示される部分の中で「コンピュータ関連」「電子化関連」「ICT関連」の3種類のみを中心に、ICT関連の専門用語の動向を把握してみることにする。その意味で、以下の結果と評価は主に、上記に提示された3種類の専門用語に注目したものである。

## 4. 結果と評価

### 4.1 日本学術振興会で示された専門用語の動向

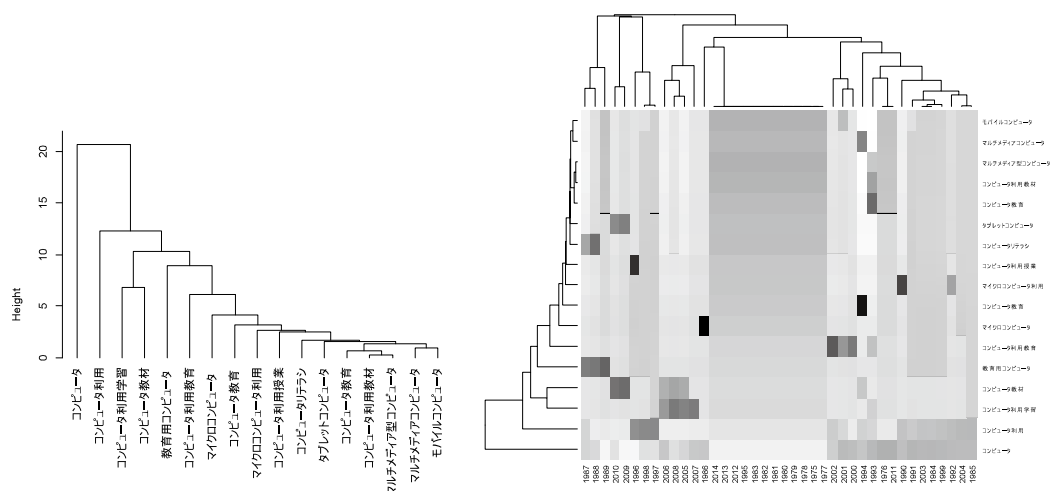


図4 コンピュータ関連のクラスター分析結果と濃淡図

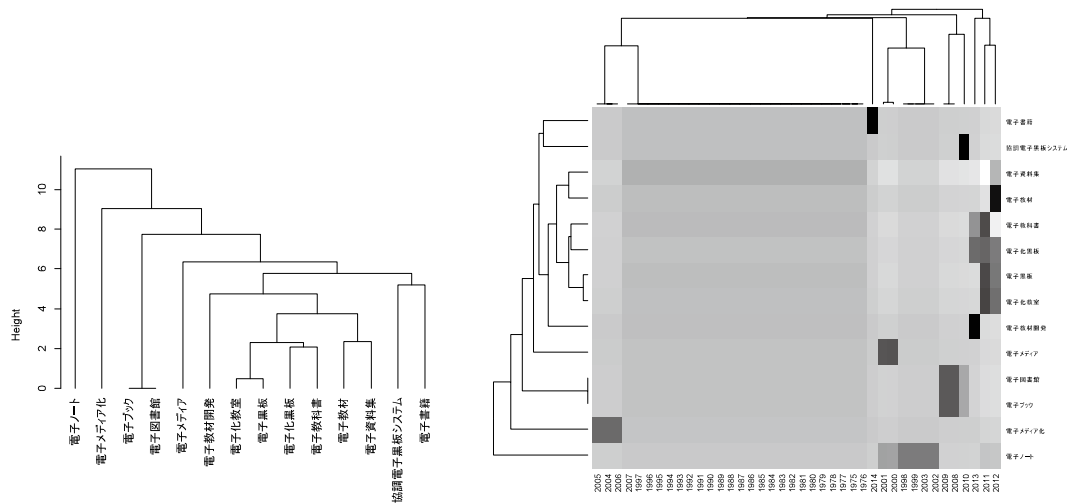


図5 電子化関連のクラスター分析結果と濃淡図

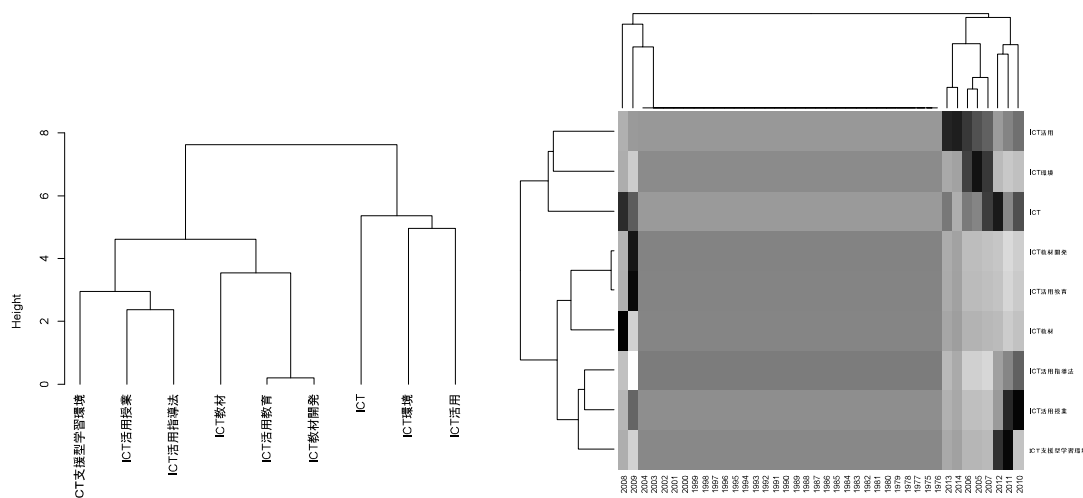


図6 ICT関連のクラスター分析結果と濃淡図

日本学術振興会で示された研究テーマを分析した結果をまとめると、以下のとおりである。この結果は全体的な専門用語の動きを含めて、図4～図6までの結果を踏まえて詳述したものである。

- 1970年代に使われていた教育機器の主流は「OHP」とであるという結果である。これは後に説明するパナソニック教育財団と同様であり、時代の特徴とも言える。
- 1980年代には「CAI：Computer Aided Instruction」が登場している。これが1980年代の後半になると「コンピュータ」という用語と関連付けられ、教育分野の動きが見られるようになる。
- 1990年代では重要度が高い用語が「コンピュータ」や「ハイパーメディア」へと変わっている。インターネットの普及が背景にあると考えられる。メディアリテラシーという概念も重視され始めるのもこの時代の特徴といえる。
- 2000年代に入り「インターネット」、「コンピュータ」に関連する用語が研究テーマの中で大きな割合を占める。また、伸ばす力が「活用能力」から「発信能力」へと移行していることから、コンピュータやインターネットがコミュニケーションツールとして浸透し始めていることが見て取れる。
- 2010年代に入ると「ICT」に関連する用語が種類となっていく。それに加え、2013年には「電子化」の用語も見られるようになる。これはパナソニック教育財団とは異なった結果である。一般企業と公的教育研究機関の特徴が表れている部分でもあるといえる。
- 図4から図6までの濃淡図を見ると、濃い色で示される特徴的な用語の



図9 ICT関連のクラスター分析結果と濃淡図

日本学術振興会の結果を比較しながらパナソニック教育財団の研究テーマから示された用語の特徴をまとめると以下のとおりになる。

- 1970年代，教育機器として「OHP」が主流であった。
- 1980年代に入ると「コンピュータ」や「パソコン」という用語が重要度にヒットされた用語として出現し始める。
- 1990年代では1980年代に続き「コンピュータ」という用語が多く出ている。1999年には「イントラネット」という用語が出ていることから企業間のネットワークが学校現場でも導入され始めた時期であることが考えられる。
- 2000年代は「インターネット」，「ネットワーク」に関連する用語が大半を占めている。同時に「情報活用能力」や「情報コミュニケーション能力」という用語が登場していることからコンピュータが生活や学校現場で普及している中で情報をどう使うか，正しい使い方などのリテラシー面の育成に力を入れ始めている年代であるともいえる。特に，2007年に「ICT」関連用語の重要度が高く現れているのがひとつの特徴といえる。
- 2010年代の中でも2011年に「ICT」関連用語が特に増加している。この年代全体を見ると「ICT」，「マルチメディア」，「コンピュータ」に関連する用語が多い。2013年以降の近年では「タブレット」という用語も出ており，2013～2014年度では「電子化」に関連する「電子黒板」，「電子教室」などの用語が重要度の高いものとして位置づけられていることから，近年研究が進んで行われている分野であり，今後もっとも教育の情報化の中でも重要視するべきものであると考えられる。
- パナソニック教育財団は一般企業の教育研究機関であるため，企業の特徴である「電子化」が日本学術振興会の研究テーマより強く研究内容にも反映されていることが伺える。

これらのことから，「OHP」から始まったICT関連教育は現代の教育現場で「電子化」として進化を遂げ研究が進められていることがよく分かる。日本学術振興会の結果からも今後注目すべきキーワードは「電子化」であることが用意に推測できる。

#### 4.3 日本学術振興会とパナソニック教育財団と結果の比較

上記にも述べたように，日本学術振興会とパナソニック教育財団はそれぞれの所属する機関としての特徴を反映した形で，結果を示している。各機関の研究テーマからの用語の重要度に着目し，クラスター分析の結果や濃淡図から見えるICT関連用語の変遷は特徴的であるといえる。表2にまとめてこれらの用語の自



体的な変遷をまとめている。特に網かけの用語が各教育研究機関の特色である。

表2 クラスタ分析と濃淡図から見える ICT 関連用語の変遷

| 年代   | 日本学術振興会  | パナソニック教育財団  |
|------|--|---|
| 1970 | OHP<br>NHK 教育放送                                      | OHP   |
| 1980 | ビデオ教材<br>CAI<br>コンピュータ<br>パソコン<br>高度情報化社会            | 放送<br>コンピュータ利用<br>コンピュータ学習<br>MSX パソコン                    |
| 1990 | コンピュータ<br>ハイパーメディア<br>情報処理活用能力<br>ネットワーク<br>インターネット  | コンピュータ<br>機器利用<br>イントラネット<br>リテラシー                        |
| 2000 | インターネット<br>コンピュータ<br>マルチメディア<br>情報化<br>情報発信能力<br>ICT | インターネット<br>ネットワーク<br>IT<br>情報活用能力<br>情報コミュニケーション能力<br>ICT |
| 2010 | ICT<br>電子教科書<br>タブレット<br>スマートフォン                     | ICT<br>マルチメディア<br>コンピュータ<br>タブレット<br>電子黒板<br>電子化教室        |

表2の結果から、全体的にみると ICT 関連研究の動向は非常に似ている傾向を示している。「OHP」から始まり、「コンピュータ」、「パソコン」が普及し始めた情報化時代の背景とともに「ネットワーク」、「インターネット」が教育現場にも広がっていることが良く分かる。

1990年代からは「ネットワーク」が日常生活の中へ浸透したことから、「情報活用能力」といったメディアリテラシーに関する用語が出現しはじめた。続いてコンピュータといった情報機器がコミュニケーションツールの一つとなり始めた時期でもある。この時期から「情報発信能力」といった“情報を取捨選択する力”から“自ら発信していく力”へと移り変わっている。このことから、教育現場で身につけるべき力も時代の流れとともに変化していることが分かる。

2000年代後半からは、「ICT」に関連する用語が見られ始めるようになる。「ICT」関連用語は両教育研究機関とも近年にかけて重要度の高い用語として挙げられていることも結果から明らかである。

そして 2013～2014 年代に表れ、共に高い重要度を示しているものが「電子化」に関連する用語である。これは近年始まり重要度が高いことから今後力を入れて取り組まれていく分野であると考えられる。特にパナソニック教育財団では「電子化」に関連する用語が学術振興会の結果よりも多く見られたことから、運営方針を実行している企業としての「専門分野」としていることもあり、より力を入れて研究助成をしているのではないだろうか、と推測できる。パナソニック教育財団の研究はより一層教育現場の電子化に関する研究が多い。一方、用語の数は少ないものの、学術振興会の結果でも同様に「電子化」へ進んでいく傾向がはっきりと見られるようになっている。

以上の2つの機関の分析結果からも分かるように、今後 ICT 関連研究は、学校現場の「電子化」へと進んでいくのではないだろうかと推測できる。実際、現在総務省の ICT 政策の一つとして“フューチャースクール推進事業”が行われている。ICT の利活用を推進し、タブレット PC や電子黒板などを積極的に授業や学校生活の中へ取り入れるように指導している。遠隔地とのテレビ会議や、通常授業の中での発表・まとめ・指導・資料作成などの活動に生かされている部分もある。以上の結果から、現在進んで行われている日本の教育の電子化の背景が分かりやすく一目できるようになったといえる。

## 5. おわりに

1980 年のコンピュータ普及に伴い日本の教育の情報化は急速な発展を遂げてきた。今後はフューチャースクールの推進を背景として、さらに「電子化」の研究が進んでいくことで、タブレット PC・電子黒板・スマートフォンなどで子どもたちが学習する“電子化教室”が現実となっていくであろう。

## 参考文献

1. 村井純, インターネット, 岩波新書 416, pp. 136-169, 1995.
2. <http://jouhouka.mext.go.jp/> 「文部科学省, 教育の情報化」
3. <http://gensen.dl.itc.u-tokyo.ac.jp/> 「言選 Web」