

研究の性格に応じた研究評価を一ひとつの体験

田畑孝一

図書館情報メディア研究科教授 知的コミュニティ基盤研究センター長

研究開発の評価を「国際的に特に優れている」、「国際的に優れている」、「国内で特に優れている」、「国内で優れている」、「それ以下」の5段階評価（順に5～1点）で行えと言われたら、読者諸氏はどのように思いますか。もはや7、8年前になるが、ある省庁が始めた5年間の企業向けの研究開発プロジェクトで、採択された十数件の応募プロジェクトについてその3年間分の中間評価を行う評価委員として私は関与した。あと2年間研究を継続し、引き続き補助金を出すかどうかを決めるものである。それまでの長きにわたって省庁は民間に補助金を出しっぱなしで、巨額であってもその成果についてまったく評価を行ってきかなかった。誠に奇妙なことであるがそれが当たり前であった。世の中の風向きが変わりはじめ、ついに省庁が重い腰を上げて評価に取り組むことになった。ところが、このプロジェクトの開始時点では評価を行うこ

とは、募集側も応募側もまったく想像していなかった！ そのような状況で急遽行うことになったので、募集側も応募側もびっくり仰天という訳である。募集側は評価をどのようにしたらよいかノウハウをまったく持っていなかった。応募側も評価されたことがなく、また評価されると聞いていなかったのも無防備に、軽い気持ちで研究開発を行ってきたのである。

ともあれ、評価委員の前で各企業の発表会が行われ、渡された評価シートを見て驚いた。冒頭の5段階評価である。これを見て、二つの点で驚いた。一つは、日本は国際的に初めから劣っているという観念で、明治以来の輸入文化の流れがいまでも省庁にあることである。世界的にはだめでも日本国内ではまあがんばっているという観念である。評価基準は国内外を通した一つの基準であるべきにも拘わらず。もう一つは、どの応募プロジェクトもその性格を問わず、

単純に5段階評価を行うということであった。研究開発にはそのアイデアのオリジナリティに優れているものがある一方で、アイデアは既知でも実用化に優れているものもあり、多様な性格の研究開発を一つの尺度で図るべきではない。応募プロジェクトは事実、その性格は多様であった。評価委員としてそれを主張したが、評価方法は変わらなかった。

さて、それからしばらくして、(筑波大学と統合前の)図書館情報大学で2000年度から特別研究の仕組みが導入され、それに伴いその研究成果を評価する必要が生じ、評価方法を提案する立場に立たされた。上に述べた批判的な気持ちを持った以上、何か適切なものを提示しなければならない。それがどのようなものであったかを述べる前に、特別研究の導入に至った当時の状況を述べる。国立大学では、2000年度から教育・研究費を教育研究基盤校費という名目で文部省から配分を受けるように改革された。それまではずっと、教官当積算校費という名目で配分を受けており、それは教官一名当たりで定められた研究費を教官数で積算したものであった。つまり、すべての教官がその研究成果や教育貢献に関係なく、同じ額の配分を受けていた。それゆえ、大学独自の判断で大学全体を活性化するための特別な企画や日常の共通の経費を教官当積

算校費から捻出しなければならず、いわば各教官に常に頭を下げる必要があったのである。これを改善するために、教官当積算校費に学生当積算校費を加えたものが上述の教育研究基盤校費として一括して配分されることになった。これによって大学は時代に合った目的に向かってトップダウンに予算を使用することができるようになった。

当時の図書館情報大学ではその趣旨を生かし、教育研究基盤校費の中から、その15%が教官定額分に、20%が特別研究推進費に配分された。前者はいわば教官の最低限度の研究費を支給するものであり、後者は各教官が応募して研究費を競争的に獲得するものである。特別研究にはその申請額で決まる応募クラス三つが用意された。学内に特別研究推進委員会が設けられ、そこで申請された研究が審査され、採否が決定された。学内にそれとは別に、特別研究評価委員会が設けられ、そこで終了した研究の研究成果を評価することになった。私がおその評価委員会の委員長ということで、前述したように評価方法を提案する立場になってしまった訳である。当時の図書館情報大学の学術分野はいわゆる文系から理系まで多様であった。かといって、それぞれの学術分野への応募の絶対数は少なく、学術分野ごとに区分して研究成果を比較評価

することもできない。そこで、学術分野とは関係なく、研究の性格で区分が可能であるかどうかを探ることにした。

採択された特別研究 84 件を調べて研究の性格を分析した。研究の内容を調べてその性格を表すことばを用意し、それに該当するものを寄せ集めた。その際、ことばが単語であると簡略過ぎてその性格が良く分からなくなるので句で表すようにした。一方、句で分類する際、あまり詳しく過ぎるとそれに該当するものが一課題のみになってしまう。試行錯誤の結果、「小分類」は表 1 に示す「句による」項目となった。全体の半分位の研究課題で「小分類」の項目がほぼ出揃い、残りの半分の研究課題は既に出た項目のどれかに属し、新たな項目は増え

ず、この分析は安定したものとなった。この「小分類」をまとめたところ、表 1 の「分類」の欄に示すように「調査」、「分析・解析」、「理論・方法論」、「実用化」の 4 項目となった。「実証」や「実験」という同じ単語を含んだ句が異なる「分類」に出現していることに注目してほしい。この「分類／小分類」による研究課題の分布は表 1 のようになった。この分布は学術分野や応募クラスによる偏りがなく、学術分野に依存しない研究の性格であることが分かり、今回の目的に都合の良いものになった。

さて、このように性格の異なる研究をどのように評価したら良いか。評価要素として、独創性、知見、信頼性、有用性などが挙げられるが、性格の異なる研究をこれら全部の評価要素で評価を行うのはかえって不都合なこともある。例えば、実用化研究はその元となるアイデアは必ずしもオリジナルでなくてよく、それによって開発したものが信頼性高く動作すれば良い。また、評価要素が多いと評価作業が散漫になるので、なるべく評価要素を絞った方が良い。このことから、ここでは、表 2 のように整理した。調査研究では新たな知見がどの程度の信頼性で得られたか。分析・解析研究ではどのような知見が得られ、それはどの程度有用なものであるか。理論・方法論の研究では独創性はどの程度高く、それはどの

表1 研究の性格
(分類Pにおける小分類O-の表記はO-Pの意。
例えば、実態-は、実態調査)

分類	小分類	研究の件数 (計84件)
調査	収集・分析に基づく-	12
	比較に基づく-	4
	実証的-	2
	実験的-	3
	実態-	1
	フィールドワーカー	3
分析・解析 (説明を含む)	実証による-	3
	実験による-	6
	調査による-	4
理論・方法論 (モデル化、 体系化を含む)	調査に基づく-	6
	解析に基づく-	7
	実験に基づく-	2
	考察に基づく-	8
実用化	データベース等作成	8
	システム開発	10
	試験の開発	5

表2 研究の性格と評価要素

研究の性格	評価要素 (◎を特に重視)			
	独創性	知見	信頼性	有用性
調査		◎	◎	
分析・解析		◎		◎
理論・方法論	◎			◎
実用化			◎	◎

程度有用なものか。実用化研究では実用化したものが信頼でき、かつ有用か。上述の特別研究のうち終了した研究の成果59件を、6人の評価委員がこの方法で評価を行った。最初に研究の性格を確認した後、各評価要素でそれぞれ5段階評価(5～1点)を行った。評価委員会で評価結果をとりまとめたが、評価委員全員が評価作業は大変やりやすかったと報告されている。なお、上述の研究の性格を募集時のカテゴリに設定するのも一方法である。応募の際にどのカテゴリかを自己申告してもらい、カテゴリごとに採否を決定するのである。

以上、研究の評価作業の一実例について述べたが参考になれば幸いである。評価する方も評価される方も、始まったばかりの時代である。経験を積んでノウハウを蓄積していきたいものである。

(たばた こういち/情報システム工学)