

データ収集作業のガイドライン

－効率的・効果的な評価作業のためのデータ収集の課題と対応－

大学評価コンソーシアム

1. はじめに

評価作業、即ち、現状把握を行う場合、適切なデータが必要である。データには状況が記述された文書などのテキストデータと、数量的なデータがある。状況の記述は、最悪、現場に直接出向いて聞き込みを行えば、評価担当者自らでも収集可能であるが、数量的なデータ、特に経年的なデータについては、適切に管理していたものを、適切に編集する必要がある。

しかしながら、評価担当者だけでなく、数量データを取り扱う大学関係者の多くが、データを集めようと思っても集まらない、集まっても使えない、精度があまり高くない、などの思いを持っているようで、それを改善するためにどうしたらよいのか、全国の評価担当者で議論してみることが必要ではないか、と考えた。

米国の IR では、多様なデータを収集し分析するところが多い。これは、学内外に提出しなければならぬ数量データが日本に比べ多いので、データを一元的に管理したり、複数の拠点で管理されているデータにアクセスできる権限を持っていたりすることとも関連する。もっとも、米国の IR のデータマネジメントは、必要に迫られて発達した、とも言えなくもない。また、米国においては、国家レベルで運用しているデータベースや、州レベルで運用しているデータベースに各大学がデータを提供し、それを（他大学分を含め）利用するということが一般に行われているため、必然的にデータの提出やデータベースの利活用について、そのノウハウ、方法論が発達していったと思われる。

しかし、我が国の場合、学校基本調査、大学情報データベースなど国レベルで運用しているデータベースはあるものの、提出したデータについて、大学では利活用できない。そのため、いかに楽に効率的にデータを提出するか、ということに最適化したデータの取扱いが志向される場合が多い。また、学内に研究者情報データベース、人事データベース、財務データベースがあっても、それぞれの方針、それぞれのルール、それぞれのフォーマットで運用されており、学内外でデータニーズ（問い合わせ）が発生する度に、それを不幸にも（？）受け付けてしまった担当者が、学内を右往左往しながらデータをかき集める、ということを繰り返していることが多い。

このような状況に対して、データを生データレベルで一元化し、データ担当者が求められたフォーマットに合わせて出力するようなデータセンターの構築なども検討されているが、あまり顕著な成果が聞こえてこないことも確かである。もっとも、データの管理を一元化してしまうと、現場で問題意識をあまり持たない方は、ますますデータを見なくなってしまうのでは、という懸念もあるので、この辺の考え方については、少しつつこんだ議論が必要だろう。

そのような中で平成 23 年 9 月 16 日、九州大学西新プラザに於いて大学評価担当者集会 2011 の第一分科会「評価と IR：データ収集編」において、教育情報の公表や内部質保証にむけたデータ整備を事例に、データ収集、蓄積、活用の課題とその解決策について全国の評価担当者で議論

を行った。ここでは、そこで得られた知見をガイドラインの形にして公表することで、各評価担当者、データ管理・分析担当者の業務向上の一助になれば、と考えている。

また、このガイドラインは、これで完成というわけではなく、みなさんのご意見や、大学を取り巻く環境の変化に応じて柔軟に改訂、更新していきたいと考えています。ゆくゆくは理論的なとりまとめや、他の部分のガイドラインも含め、評価担当者向けハンドブックのような形にまとめて行ければ、とも考えております。ご意見やご要望などありましたら、お気軽に大学評価コンソーシアム事務局までお知らせください。

2. ガイドライン

データの収集、蓄積については、大学レベルで解決できる問題と、大学間や大学と政府機関等との間で調整しなければ解決できない問題がある。ここでは、大学レベルで解決できる問題をガイドラインとして示し、解決できない問題は、今後の課題としてまとめた。

その1：データ収集の目的とデータ定義を明確にしよう。[評価業務で求められるデータの類型]

データをなぜ集めるのか、ということがはっきりすれば、自ずと定義も基準日も見えてくると思われる。しかし、問題は、その「なぜ」の部分である。評価担当者が集めるデータ（もしくは動機）は大きく分けて2つある。

A：学外からフォーマットが示され必ず提出することが求められているデータ

B：業務上必要なデータ

B1：自分で持っていないデータ（認証評価や法人評価などの評価作業で必要だが評価セクションでは保有していないので各部局に依頼して収集するデータ）、

B2：教員総覧など評価セクションにおいてデータベースで管理しているデータ

B3：評価セクションで所有しているデータ（さすがにこれは大丈夫だと思われるので、ここでは言及しない。）

A：学外からフォーマットが示され必ず提出することが求められているデータを扱う際の留意点

学校基本調査や大学情報データベース、大学ランキングや新聞社等からの調査がこれに該当する。これらのデータについては、定義が明示されているものと、明示されていないものがある（ランキングの精度を落とすことになりかねない？）。定義が明示されていようがまいが、評価セクションで学生や教員に関するデータを網羅した基幹データベースを運用している場合には、適当にデータを編集すれば作業は完了である。そのような大学はあまりないので、多くの大学では、定義が明示されていれば「定義書」を添付して、そういったものが無ければ評価担当者自身が定義を考えて各部局にデータの提供を依頼するわけだが、毎回、なぜか、ここから大変な思いをすることになる。クロスチェックをすると計算が合わない、計算しなくとも見た目で辻褃が合わないデータが手元に集まってくる。問い合わせると別の数字が出てきて、前回と違うのはなぜかと問い合わせると、さらに違った数値が出てくることも少なくない。これを繰り返す中で段々と收拾がつかなくなっていくが、締め切りが迫り、気ばかり焦るが、Excelのセルは埋まらない。

これは（当たり前だが）、現場の担当者が評価担当者に嫌がらせをしようとして、わざとおかしな数値を出してくるわけではない。（一人一人に確認したわけではないが、たぶんそうだろうと信じたい。）これは、評価担当者側のデータを集める手順か、部局担当者側のデータを整理格納する手順に根本的な問題がある、ということを示唆している。

米国の IR でデータ管理が発達した理由の一つは、それが必然だった、ということは既に述べた。エンロールメント・マネジメントなどをしっかりやらなければ、本当に困る状況に陥るのだ（例えば、IPEDS や適格認定からのデータ提出を怠ると、その大学へ進むは学生の奨学金授与の資格を剥奪されかねない）。では、なぜ、我が国ではデータ管理が不十分なのだろうか。それは、たぶん、正確なデータを確保、収集しなくても困らないからである。誰も必然性がないことを真剣には取り組まない。（ただし、上述の括弧内と矛盾するようだが、実際に米国においても、現場の部署（学部・学科など）ではデータの管理が甘いところもあるようである。その場合、IR がデータ管理をするという覚悟で統括し、逆に学部・学科に正確なデータを提供しているという姿勢になっている）。

では、どうしたらよいのか。将来的なところは今後の課題に記すが、とりあえずは、なぜこのデータが必要なのか、現場の担当者にもしっかり理解してもらうことと、定義を明示することであろう。次のステップは、データを活用する機会を増やすことであろう。使われないデータなら適当にやっても大丈夫だろう、と普通は思う。経年変化や他学部との比較、可能であれば他大学との比較を機会があれば現場の方々に示し、様々な場面で活用してもらうことである。業務の中の様々な決定にデータを用いる習慣や意識を、いままで以上にお持ちいただくことが肝要である。自分たちがデータを利用するのならば、当然、気を配るであろう。

B 1 : 各部局が保有するデータを入手する際の留意点

これらのデータは、主として学外に提出する評価書の根拠として用いられるものである。我が国の評価現場の多くでは、何らかの観点に従って現状把握を行う場合、データ分析を行って客観的に状況を導き出すよりは、直感的な現状把握があって、それをデータで確認する、という手法が、その是非はともかく、どちらかといえば支配的である。即ち、何か主張したいことがあって、そのためにデータを探してくる、ということになる。従って、評価セクションで現状把握を行う場合、部局から寄せられた状況の記述を読み、根拠資料の過不足を確認し、足りなければ具体的に「足りないデータ」を求めればよい。このような場合には、評価担当者の中で欲しいデータのイメージはできあがっているのだから、それをいかに具体的に部局に要請・要求できるか、ということがポイントになる。

B 2 : データベースに格納されているデータを入手する際の留意点

人事管理システムなど厳密に運用されているデータベースは、それこそ入力ミスが大きな問題になることが多いためか、あまりトラブルになっているのは聞かない。しかし、教員が各種業績を自ら入力することを基本とするようなデータベースでは、課題がない大学のほうが少ないだろう。

課題として、多くの大学がまず挙げるのは、入力率である。入力率ないし更新の頻度は、デー

データの一次発生源である教員に依存するという脆弱な制度であるが故に仕方の無い部分がある。教員評価の根拠資料にデータベースを指定して入力を義務づけようが、入力しない教員にペナルティを課そうが、何をしようとも、絶対に 100%の入力はない。100 人の教員がいればデータの管理母体が 100 箇所増えるわけであり、その 100 箇所のデータの管理母体が一糸乱れず遺漏無く動くことは不可能である。さらに怠惰の問題に加え、入力ミスに関するチェックは極めて困難である。

入力率はいくつかの方法で改善することが可能である。まず、大学本部で持っているデータは、本部の事務職員、技術職員で手分けして通常業務の一環として入力する。科研費などの外部資金の受入、兼業・兼職、授業実施などは大学本部が持っていないわけがない。研究業績については、「買う」という手がある。次に、データベースに入れたデータはとにかく使い倒す、ということである。教員に対して、何らかの調査を行うにも、データベースに入っているデータは、既に入力された状態で渡す。とにかくデータベースに入れておけば自分の時間の節約になる、ということを教員に徹底し、本部は「二度手間はかけさせない」という約束は必ず守ることである。

しかし、これでも結局、入力率（データの捕捉率）は 100%にはならないので、何かの分析をしようと思っても、精度の問題が常につきまとう。従って、このデータで何をしたいのか、というデータの目的を考えて、どのような精度確保を行うのか、という話になる。絶対確実なデータが欲しいなら、教員には入れさせないくらいの覚悟が必要である。またデータのチェックを不定期で監査的に行うという手もなくはないが、どこまでのコストをかけるか、という話である。

その 2：執行部の理解を得て、協力してもらおう。

データ収集に対して、学長を始めとする大学執行部がどの程度、その重要性を理解してくれているか、ということも収集効率には関係する。いくら決裁を取って学長名でデータの提供依頼を行っても、現場の担当者が昨日や今日、大学職員になったばかりの人間でなければ、その指示書も評価担当者が起案し、「学長の名において」やっていることがすぐにわかるので、いささかパンチに欠ける。また、現場の教員なら学長名の依頼文を見ても何も感じない方もいるだろう。

重要なのは、経営層がデータ収集の重要性を理解し、さまざまな場面でデータの収集に協力してくれるよう全学の構成員に呼びかける、ということである。教授会の開始前に時間をもらってひとことを話すのだってよい。大学執行部は本気だ、という思いは意外と伝わる（米国でも案外同じらしいです）。ただし、この場合、まず大学執行部に「データ収集の重要性を理解」してもらったところから始めなくてはならない。その際に重要なのは、データを集めるとどういうメリットがあるのか、ということを大学執行部に具体的に示すことである（事例は、今後追加します。好事例があれば、お知らせください）。データの経年変化や学部間比較、他大学とのベンチマーク結果などを示し、自分たちがしなければならぬ意思決定に対して、データを収集することによって有効なヒントが得られるのだ、ということを示すのである。そのような形で、大学執行部や学部執行部からデータに対するニーズが生まれてくれば、その効率的、効果的な収集、精度向上のためには、たいていのことをやってくれるはずである。例えば、学部ごとにローカルルールがあって、データがうまく集まらないなら、調整する方向で動いてくれるだろうし、少なくとも評価担当者の苦勞は知ってくれるので、あまり無茶なことは言わなくなってくれるはずである。

その3：各部署との連携を強化して協力してもらおう。

どこにどんなデータがあるのか、誰に聞けばそのデータがもらえるのか。経験の長い評価担当者には自身の頭の中でのデータベースが構築されていることが多い。これは、データを長年もっていることから自然と知見として蓄積されているということもあるが、データのやりとりだけでなく、人と人とのコミュニケーションを行っているという面も無視できないだろう。いきなりメール1本でデータを要求するようなことはせずに人間関係のなかで仕事を進めている。

結局、連携体制が構築できるかどうかは、日頃からのつきあいが重要である。困ったときだけ頼んできて、あとは知らんぷり、というような姿勢では、なかなか連携は強化できないだろう。

3. 今後の課題

全国的なデータベース（大学ポートレート）が導入されれば、自ずとデータの入力とデータの利活用について関心が集まるだろう。データの定義の問題も全国基準が示されれば、解決の方向に進むのかもしれないし、ベンチマークなど新たな大学経営支援の方法が発達することも考えられる。

大学情報データベースが、あまり大学関係者の中で注目されなかった理由としては、データを提供する一方で、自分たちが利活用できなかったことが挙げられる。利活用ができなかった故に、利用者である評価担当者が大学情報データベースに意見を言ったり、大学情報データベース側から頻繁に利用者に意見を聞いたり、依頼をしたり、という「共に育てる」ということができなかったことも大きいだろう。今後、複数の大学にまたがる共通プラットフォームとして大学ポートレートを育てていくなれば、大学ポートレートと利用者との橋渡しが不可欠である。今後、そのような橋渡しは多くの機関によって行われるだろうが、大学評価コンソーシアムとしても、大学のデータ収集が現状認識（評価）の一つとして有効に機能するように努力したいと考えている。

4. 謝辞

このガイドラインは、大学評価担当者集会 2011 第一分科会に参加されたみなさまのアイデアやご意見をまとめたものです。当日の活発なご議論やそのあとのまとめ作業によってこのようなガイドラインをまとめることができました。みなさま、おつかれさまでした。今後とも、よろしくお願いします。

5. ガイドライン作成（第一分科会運営）の主たるスタッフ

浅野 昭人（立命館大学）、浅野 茂（神戸大学）、大野 賢一（鳥取大学）、小湊 卓夫（九州大学）、佐藤 仁（福岡大学）、○ 寫田 敏行（茨城大学）、関 隆宏（新潟大学）、難波 輝吉（名城大学）、本田 寛輔（ニューヨーク州立大学エンパイア・ステイト・カレッジ）

○：文書作成チーム

6. 更新履歴

平成 25 年 1 月 11 日版：初版

平成 25 年 2 月 12 日版：改訂（スタッフによるコメントにもとづき加筆・修正を実施。以下、未対応項目：文中に読点が多いこと、口語体が多いことや、部局「さん」や「・・・していただく」など、学内者に敬語が使われているのは不自然である。国立と私立などでの用語の違い（例えば、学部か部局、後者は事務部署も入る、入らない？）などの調整ができていないために分かりにくい箇所がある。）