

# Методология оценки эффективности вузовской науки

## Введение

Оценка эффективности научных работ вузов в данной работе функционально связывается с целевым и эффективным использованием средств федерального бюджета и внебюджетных источников высшими учебными заведениями и научными организациями Министерства образования Российской Федерации. Вузовская наука представляет собой сложную социально-экономическую систему. Оценку научно-исследовательской работы вуза, по мнению авторов, следует проводить исключительно по результатам проведенных вузом исследований. Ранее был предложен ряд методик оценки эффективности сложных социальных и экономических систем. Известны работы С. А. Айвазяна [1,2,3], О. И. Ларичева [4], Б. Р. Литвака [5]. В данной работе собран и использован положительный опыт многих исследователей в области анализа сложных систем.

Оценка эффективности научных работ основана как на объективных данных о научно-исследовательской работе вузов, так и на экспертных мнениях. Организационная процедура получения экспертных оценок широко используется во всех отраслях народного хозяйства. Для получения экспертных оценок применяются такие известные методы как метод прогнозного графа, метод ПАТТЕРН [6], метод проблемных сетей [7], метод решающих матриц [8] и другие. Последний из упомянутых методов предложен Г. С. Поспеловым в 1966 г. использован при решении проблемы планирования средств, выделяемых на фундаментальные исследования.

При составлении данной методики были учтены особенности, с которыми столкнулись аналитики при проведении экспертиз, в частности, при использовании отчетных данных о научно-исследовательской работе вузов и научно-исследовательских институтов Министерства образования Российской Федерации за 2000-й год.

В соответствии с «Положением об организации научных исследований Министерства образования Российской Федерации» определено требование о достаточности представленных материалов для оценки результатов завершенных НИР. Следует отметить, что для некоторых вузов характерно только перечисление научных конференций и наименований тем научно-исследовательских работ или перечисление сведений о выполненных пунктах технического задания<sup>1</sup>.

В данной работе для нахождения оценки эффективности результатов научных работ поставлены и решены две задачи. Требовалось, во-первых, произвести сравнительный анализ результатов научной деятельности вузов, что сделано посредством выставления экспертных оценок каждой из вузовских работ. Во-вторых, поставлена задача оценить вклад вузовской науки в общероссийскую. Этот вклад оценивался следующим образом. Были построены два сравнимых интегральных показателя — показатели эффективности вузовских и общероссийских научных работ. Искомая оценка является отношением этих двух показателей. В данные интегральные показатели в качестве базовых входят как показатели качества, так и показатели объема или количества научных работ. Качество научных работ в данной методике описывается предложенным ниже интегральным индикатором, а объем работ оценивается экспертами на основании измеряемых данных. По результатам экспертизы была найдена доля вузовских научных работ в общероссийских и, в соответствии с этой оценкой, сделано утверждение о необходимости уточнения бюджета финансирования вузовской науки.

Результаты научных исследований вузов за 2000-й год были классифицированы экспертами по 70-и критическим технологиям (см. таблицу 1).

<sup>1</sup>Как выяснилось, формулировка этой достаточности определяет тривиальные требования заключения комиссии по приемке работ о том, что выполненные работы должны соответствовать требованиям технического задания (см. п. 2. «Положения»); содержать однозначные показатели: «очень высокие», «высокие», «значительные», «незначительные» (см. п. 3.); «полученные результаты готовы к практическому использованию», «частично готовы», «не готовы» (см. п. 4.); «результаты используются в учебном процессе», «могут быть использованы», «не используются». Использовать заключение комиссии, например, «высокие научные достижения» весьма затруднительно в связи с их многозначностью. В этой связи необходимо усовершенствовать методологию составления отчетной документации, предоставляемой вузами Министерству.

В России перечень критических технологий был подготовлен и утвержден правительственной Комиссией по научно-технической политике 21 июля 1996 г. Основным критерием отбора технологий была их необходимость для экономического и социального развития страны, поддержания обороноспособности, научной и технологической инфраструктуры. Критические технологии вошли в семь приоритетных направлений развития науки и техники (по данным ЦИСН Миннауки России)<sup>2</sup>. Также дополнительным приоритетным направлением стали фундаментальные исследования.

Приоритетные направления	КТ
Информационные технологии и электроника	10
Производственные технологии	9
Новые материалы и химические продукты	10
Технологии живых систем	14
Транспорт	5
Топливо и энергетика	16
Экология и природопользование	6
Фундаментальные исследования	—
<i>Всего</i>	70

Таблица 1: Критические технологии России

## 1. Формирование экспертных групп

Для анализа результатов научных исследований по семи приоритетным направлениям развития науки, технологий и техники и по фундаментальным исследованиям было создано восемь рабочих групп экспертов. Каждая группа состояла из 5–6-и человек. Для всех групп определена единая процедура экспертного опроса.

Каждый эксперт, на основании персональных данных, получил квалификационную категорию  $u$ . Всего определены четыре категории<sup>3</sup> —  $u \in \{I, II, III, \delta\}$ .

Персональные данные каждого эксперта есть тройка  $(s, r, h) \in S \times R \times H$ , где  $S, R, H$  — множества альтернативных высказываний, причем  $S = \{s_1, s_2, s_3\}$  определяет высшее и послевузовское образование эксперта,  $R = \{r_1, r_2, r_3\}$  определяет научную подготовку эксперта,  $H = \{h_1, h_2, h_3\}$  задает стаж работы эксперта по данному приоритетному направлению и  $s, r, h \in \{И, Л\}$ . Содержательный смысл высказываний, входящих в множества  $S, R, H$ , определен следующим образом.

Высшее и послевузовское образование:

- $s_1$  — совпадает с профилем приоритетного направления,
- $s_2$  — базовое образование по смежной специальности,
- $s_3$  — базовое образование по иной специальности.

Научная подготовка:

- $r_1$  — академик РАН, член-кор. РАН, академик отраслевой академии;
- $r_2$  — профессор, доктор наук;
- $r_3$  — кандидат наук, с.н.с., доцент.

Опыт работы по профилю:

- $h_1$  — не менее десяти лет,
- $h_2$  — не менее пяти лет,
- $h_3$  — не менее года.

Решающее правило  $u : S \times R \times H \longrightarrow \{I, II, III, \delta\}$ , которое определяет квалификационную категорию эксперта, задано как функция<sup>4</sup>

$$u = \begin{cases} I, & \text{если } h_1 \bar{s}_3 r_1 \vee h_2 r_1 s_1 \vee h_1 r_2 s_1 \equiv II, \\ II, & \text{если } h_3 r_1 s_1 \vee h_2 r_2 s_1 \vee h_1 r_3 s_1 \vee h_2 r_1 s_2 \vee h_1 r_2 s_2 \vee h_1 r_1 s_3 \equiv II, \\ III, & \text{если } h_3 \bar{r}_1 s_1 \vee h_2 r_3 s_1 \vee h_3 r_1 s_2 \vee h_2 r_2 s_2 \vee h_2 r_1 s_3 \vee h_1 r_2 s_3 \equiv II, \\ \delta, & \text{если } h_3 \bar{r}_1 \bar{s}_1 \vee \bar{h}_3 \bar{s}_1 r_3 \vee h_3 r_1 s_3 \vee h_2 r_2 s_3 \equiv II. \end{cases}$$

Эксперты, отнесенные к категории  $\delta$ , в экспертизе участия не принимают.

<sup>2</sup> «Экономика и жизнь», № 33, август 1999, с. 30.

<sup>3</sup> Здесь и далее используется общепринятая математическая нотация.

<sup>4</sup> Как принято в математической логике,  $\vee$  обозначает логическую дизъюнкцию высказываний (соответствующих трехместных конъюнкций), а  $\bar{s}$  — отрицание высказывания  $s$ .

## 2. Проведение экспертизы

Процедура проведения экспертизы основана на методе Делфи [9], который предусматривает ряд мер, обеспечивающих продуктивную работу экспертной комиссии и согласованность экспертной оценки. Процедура проведения экспертизы Делфи заключается в следующем.

Процедура опроса проводится анонимно в несколько этапов. В процессе работ информация о предмете экспертизы пополняется. Эксперт назначает оценки как на основании собственного опыта, приобретенного в процессе работы по данной тематике, так и на основании объективных данных о состоянии научной деятельности вузов и других российских научных организаций в рамках данной критической технологии и в рамках данного приоритетного направления. Кроме того, эксперт в процессе работы корректирует свои оценки, согласуясь с мнением и доводами других экспертов.

Экспертиза проводится в три тура. В первом туре экспертам сообщается цель экспертизы, раздаются анкеты, предоставляется базовая и вспомогательная информация<sup>5</sup>. Цель экспертизы заключается в выставлении оценок качества научно-исследовательских работ вузов и в оценке вклада вузовских научных работ в общероссийские по каждой из критических технологий.

Информация, полученная от эксперта поступает в распоряжение аналитической группы, которая организует опрос, а также обрабатывает промежуточные и окончательные результаты экспертизы. Аналитическая группа на основании полученных данных определяет самые высокие и самые низкие оценки, усредненное мнение экспертов — медиану, разброс экспертных оценок и фиксирует аргументы и особые мнения. Аналитики получают от экспертов следующие оценки: оценку качества научно-исследовательских работ вузов по каждой работе и оценку вклада вузовских научных работ в общероссийские по каждой из критических технологий в рамках одного направления.

На втором и третьем туре экспертам возвращаются усредненное экспертное мнение и аргументы экспертов, давших самые высокие и самые низкие оценки. Обоснования и оценки предъявляются анонимно. Характерной особенностью метода Делфи является последовательно уменьшающийся разброс оценок экспертов, их возрастающая согласованность.

## 3. Данные о состоянии научно-исследовательской деятельности

При проведении экспертизы, в качестве данных, которые подтверждают или опровергают экспертные суждения, экспертам была предоставлена следующая информация.

Базовая информация — отчеты о результатах научно-исследовательской деятельности вузов и научно-исследовательских институтов Министерства образования Российской Федерации. Все отчеты были предварительно классифицированы самими экспертами по критическим технологиям и оценивались по методике, описанной ниже.

Дополнительная информация — описания состояния научно-исследовательских работ по каждому из вузов. Состояние работ описывается следующими показателями.

В качестве результативности научной деятельности вуза предлагалось использовать следующие приведенные показатели:

- 1) количество научных и учебных публикаций,
- 2) лицензионная деятельность,
- 3) количество экспонатов на выставках,
- 4) количество защищенных диссертаций,
- 5) объем привлеченных внебюджетных средств.

Оценку научно-исследовательской работы вуза, по мнению авторов, следует проводить *исключительно по результатам проведенных вузом исследований.*

<sup>5</sup>Экземпляры отчетов и анкет являются отдельными рукописями и в данной работе не приведены.

При этом должны регламентироваться определенные ограничения или условия, которые имели место в различных, ранее разработанных методиках. К их числу относятся следующие:

1. Показатель “доля численности профессорско-преподавательского состава и научных сотрудников вуза принимающих участие в научных исследованиях вуза” не информативен, так как каждый преподаватель и научный сотрудник имеет право принимать участие в научных исследованиях за дополнительные средства. Если сотрудник вуза не заинтересован в получении дополнительных средств, то его никто не имеет права заставить это делать.

2. Показатель “количество НИР, выполняемых в вузе” не информативен, так как каждую НИР платятся бюджетные или внебюджетные средства. Это находится в компетенции только самих потенциальных исполнителей.

3. Показатель “участие в приоритетных исследованиях” не информативен, так как участие в разработке критических технологий федерального уровня приоритетных направлений исследований гражданского назначения определяет заказчик — Министерство образования Российской Федерации. То же в интересах обороны и безопасности страны определяют соответствующие заказчики.

4. Показатели “оптимальность управления научным процессом”, “участие в фондах (отечественных и зарубежных)”, “подготовка научных кадров” являются внутренним делом вузов.

5. Показатель “объем затрат на НИР” не информативен, так как объем затрат на каждую НИР утверждается Министерством образования при проведении конкурса. Практика показала, что фактические затраты на проведение НИР никогда не бывают меньше, чем первоначально утвержденные.

6. Показатель “влияние результатов исследований на учебный процесс” должен обязательно учитываться в методике оценки эффективности результатов исследований в вузе.

Экспертные оценки были выставлены независимыми квалифицированными экспертами. Кроме того, научный уровень результатов исследований подтверждался документально: патентами, премиями, грантами (отечественными или зарубежными), большим количеством изданий (отечественных и зарубежных), где цитируются результаты данных исследований. Сведения о внедрении результатов исследований также были подтверждены документально.

## 4. Оценка качества отдельных НИР

Для оценки качества отдельных научно-исследовательских работ были собраны отчеты о научных работах вузов и НИИ Министерства образования. Каждая работа описывается тройкой  $(a, b, c)$ , где

$A = \{a_1, a_2, a_3\}$  — множество альтернативных оценок научного уровня результатов НИР,

$B = \{b_1, b_2, b_3\}$  — множество альтернативных оценок масштабности исследований,

$C = \{c_1, c_2, c_3\}$  — множество альтернативных оценок качества реализации результатов исследований

и  $a, b, c \in \{И, Л\}$ . Тройка  $(a, b, c)$  содержательно определяется следующим образом.

Научный уровень результатов НИР:

$a_1$  — научные результаты превышают мировой уровень,

$a_2$  — научные результаты на мировом уровне или приближаются к нему,

$a_3$  — научные результаты отстают от мирового уровня.

Масштабность исследований:

$b_1$  — в результате исследований решена ключевая научная проблема, позволяющая перейти к инженерным, практическим разработкам;

$b_2$  — в результате исследований решена вспомогательная научная задача, обеспечивающая решение основной, ключевой проблемы;

$b_3$  — в результате исследований решена частная научная задача в рамках вспомогательных задач.

Качество реализации результатов исследований:

$c_1$  — результаты исследований внедрены в производство, в решение социальных проблем;

$c_2$  — результаты исследований внедрены в ОКР, в массовое российское образование;  
 $c_3$  — результаты исследований использованы в последующих НИР.

Решающее правило  $u$ , отображающее множество  $A \times B \times C$  показателей, описывающих эффективность научной работы отдельного вуза, на множество интегральных индикаторов  $Q$ , строится следующим образом. Каждой тройке  $(a, b, c)$  однозначно соответствует элемент  $q$  множества  $Q = \{1, \dots, 27\}$ . Веса показателей  $a, b, c$  таковы что альтернативы — объекты, получившие экспертные оценки, упорядочены:  $(a_1, b_1, c_1) \succ (a_1, b_1, c_2) \succ \dots \succ (a_1, b_2, c_3) \succ (a_1, b_3, c_1) \dots \succ (a_3, b_3, c_3)$ . Тогда монотонное решающее правило  $u : (a, b, c) \longrightarrow q$  ставит в соответствие каждому элементу  $(a, b, c)$  упорядоченного множества  $A \times B \times C$  элемент  $q$  упорядоченного множества  $Q$ .

## 5. Оценка результатов исследований по отдельным критическим технологиям

После определения интегрального индикатора качества отдельных научных работ был разработан метод для нахождения интегрального индикатора  $\mathcal{I}$  эффективности результатов НИР отдельной критической технологии. Число отчетов о научных работах по каждой критической технологии достигало нескольких сотен. Описанная ниже процедура использована для нахождения интегрального показателя (индикатора) эффективности результатов НИР по отдельным приоритетным направлениям исследований.

В рамках каждой критической технологии рассмотрим работы, каждой из которых поставлен в соответствие индикатор качества  $q$ . Опишем каждое подмножество  $\xi_i$  множества  $\{\xi_1, \dots, \xi_m\}$  всех работ следующим образом. Пусть на произвольном множестве  $Q$  интегральных индикаторов задана конечная решетка  $X = \{x_1, \dots, x_m\}$ . Тогда каждому узлу  $x_j$  решетки можно поставить в соответствие  $y$  работ, которые имеют интегральный индикатор  $q$  такой, что  $x_{j-1} \leq q \leq x_{j+1}$ . Таким образом, каждое подмножество  $\xi_i$  работ описывается вектором  $\mathbf{y}_i = [y_{i1}, \dots, y_{im}]$ . Так как интегральный индикатор не должен зависеть от количества работ, выполненных в рамках критической технологии, нормируем векторы  $\mathbf{y}_i$ , получая  $\mathbf{a}_i = \mathbf{y}_i \mathbf{e} |\xi_i|^{-1}$ , где  $|\xi_i|$  — мощность подмножества  $\xi_i$  и  $\mathbf{e}$  — единичный вектор.

Соответственно всё множество работ по критическим технологиям с номерами  $i = 1, \dots, m$  описывается матрицей  $A = [a_{ij}]_{i,j=1}^{m,n}$ .

Найдем интегральные индикаторы  $\mathcal{I}_1, \dots, \mathcal{I}_m$  с помощью метода главных компонент [2]. Пусть для матрицы  $A$  выполнены условия нормирования и центрирования

$$a_{ij} = (-1)^\vartheta (a_{ij} - \min(a_{.j})) (\max(a_{.j}) - \min(a_{.j}))^{-1} + \vartheta - \check{a}_{.j},$$

где  $\check{a}_{.j} = n^{-1} \sum_{i=1}^n a_{ij}$  и

$$\vartheta = \begin{cases} 1, & \text{если } j \leq j_0, \\ 0, & \text{если } j > j_0. \end{cases}$$

Индекс  $j_0$  узла  $x_{j_0}$  решетки  $X$  выбирается таким образом, что половина элементов множества  $\{\xi_1, \dots, \xi_n\}$  имеет индикатор  $q \leq x_{j_0}$ .

Находим такие коэффициенты  $C = \{c_{ij}\}_{i,j}^{n,n}$ , что линейные комбинации векторов

$$\mathbf{z}_i = \tilde{A} \mathbf{c}_i, \quad i = 1, \dots, n,$$

обладали бы наибольшей дисперсией, то есть

$$\max \left( \frac{D\mathbf{z}_1 + D\mathbf{z}_2 + \dots + D\mathbf{z}_n}{D\check{\mathbf{a}}_1 + D\check{\mathbf{a}}_2 + \dots + D\check{\mathbf{a}}_n} \right)$$

при ограничениях нормировки  $\sum_{i=1}^m c_{ij}^2 = 1, j = 1, \dots, n$  и  $\sum_{i=1}^m c_{ij} c_{ik} = 0, j, k = 1, \dots, n, j \neq k$ . Здесь  $D$  — знак операции вычисления дисперсии соответствующей случайной величины

$$D\mathbf{a}_{.j} = \frac{1}{n} \sum_{i=1}^n (a_{ij} - \check{a}_{.j})^2.$$

где  $\check{a}_{.j}$  — среднее арифметическое значение вектора-столбца.

Процедура нахождения главных компонент [1] заключается в следующем:

1. Задается ковариационная матрица  $\Sigma = \{\sigma_{jk}\}_{j,k=1}^{n,n}$ , элементы которой находятся по формуле

$$\sigma_{jk} = \frac{1}{n-1} \sum_{i=1}^n (a_{ij} - \check{a}_{.j})(a_{ik} - \check{a}_{.k}). \quad (*)$$

2. Определяется наибольшее собственное значение  $\lambda_1$  матрицы  $\Sigma$ , как наибольший по величине корень характеристического уравнения  $|\Sigma - \lambda I| = 0$ , где  $I$  — единичная матрица.
3. Решая уравнение  $(\Sigma - \lambda_1 I)\mathbf{c}_{.1} = 0$ , находим компоненты собственного вектора  $\mathbf{c}_{.1} = (c_{11}, c_{21}, \dots, c_{n1})^T$  матрицы  $\Sigma$ .
4. Для каждого объекта подсчитываем значение первой главной компоненты

$$\mathcal{I}_i = c_{11}(a_{i1} - \check{a}_{.1}) + c_{21}(a_{i2} - \check{a}_{.2}) + \dots + c_{n1}(a_{in} - \check{a}_{.n}).$$

В результате вышеприведенной процедуры мы получаем интегральные индикаторы  $\mathcal{I}_1, \dots, \mathcal{I}_m$  для каждой критической технологии.

Для оценки качества построения интегрального индикатора используется значение величины  $\rho = \lambda_1(\sigma_{11} + \sigma_{22} + \dots + \sigma_{nn})^{-1}$  отношения первого собственного значения  $\lambda_1$  к сумме значений диагональных элементов ковариационной матрицы (\*).

## 6. Согласованность экспертных оценок

При нахождении интегрального индикатора эффективности научных работ вузов для каждой группы экспертов согласованность нескольких экспертных оценок оценивается при помощи коэффициента конкордации. Каждый эксперт выставляет  $i$ -й работе оценку  $r_i$  в ранговой шкале, которая затем стандартизируется так, что сумма рангов для любого эксперта равна сумме чисел натурального ряда  $\sum_{i=1}^n r_i = \frac{1}{2}n(n+1)$ . Обозначим через  $r_i^{(\nu)}$  оценку, выставленную  $i$ -й работе экспертом с номером  $\nu$ .

Коэффициент конкордации оценок, выставленных более чем двумя экспертами имеет вид

$$W = \frac{12}{m^2(n^3 - n)} \sum_{i=1}^n \left( \sum_{\nu=1}^m r_i^{(\nu)} - \frac{m(n+1)}{2} \right)^2.$$

Коэффициент  $W = 1$ , если оценки всех  $m$  экспертов совпадают, и  $0 \leq W \leq 1$ .

Близость оценок  $r_1^{(\nu)}, \dots, r_n^{(\nu)}$  эксперта с номером  $\nu$  к результирующим медианным оценкам  $r_1, \dots, r_n$  находилась с помощью коэффициента ранговой корреляции по Кендаллу

$$\rho = \frac{\sum_{i=1}^n \sum_{j=1}^n (r_j - r_i)(r_j^{(\nu)} - r_i^{(\nu)})}{\sum_{i=1}^n \sum_{j=1}^n (r_j - r_i)^2}$$

и рассматривался как апостериорное качество данного эксперта.

## Заключение

Разработана методика оценки эффективности научно-исследовательской работы вузов и научных организаций Министерства образования России. Целью данной методики является получение сравнительных оценок эффективности научных исследований вузов, а также оценка вклада вузовской научной деятельности в общероссийскую. Методика предполагает, что используются, во-первых, объективные данные о научной работе вузов, и, во-вторых, экспертные оценки.

Объективные данные извлечены из отчетов вузов о своей научно-исследовательской деятельности, ежегодно предоставляемых в Минобразование России. Также при сборе данных использована косвенная информация, такая как количество публикаций научных работ в отраслевых журналах, аналитические обзоры о состоянии науки, ежегодные отчеты о научной деятельности институтов РАН и другие сведения.

Экспертные оценки состояния научной деятельности по приоритетным направлениям исследований выставлены подобранными группами экспертов. При этом для получения достоверных оценок используются специальные технологии сбора экспертных знаний. Предложен метод согласования экспертных оценок при нахождении интегрального индикатора, определена точность этих оценок и согласованность высказываний экспертов.

## Список литературы

- [1] Айвазян С. А. Интегральные индикаторы качества жизни населения: их построение и использование в социально-экономическом управлении и межрегиональных сопоставлениях. М.: ЦЭМИ РАН, 2000.
- [2] Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика /Классификация и снижение размерности. М.: Финансы и статистика, 1989.
- [3] Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики. Учебник для вузов. М.: ЮНИТИ, 1998.
- [4] Ларичев О. И. Объективные модели и субъективные решения. М.: Наука, 1987. 143 с.
- [5] Литвак Б. Г. Экспертная информация: Методы получения и анализа. М.: Радио и связь, 1982.
- [6] Лопухин М. М. ПАТТЕРН — метод планирования и прогнозирования научных работ. М.: Сов. радио, 1971.
- [7] Петровский С. А. Оценка научных исследований и их результатов с точки зрения человеческих потребностей. /В кн. Проблемы и методы комплексного экспертного прогнозирования. М.: ИМЭМО. 1978.
- [8] Поспелов Г. С., Ириков В. А. Программно-целевое планирование и управление. М. Советское радио, 1976.
- [9] Хелмер О. Анализ будущего: метод Дельфи. /В кн. Научно-техническое прогнозирование для промышленности и правительственных учреждений. Пер. с англ. В. П. Козырева. М.: Мир, 1973.
- [10] Основы методологии эффективности научных исследований, выполненных вузами и научными организациями Минобразования. Рукопись. 10 с.
- [11] Методология оценки эффективности вузовской науки. Практическое пособие. Выпуск 3. Часть первая. /Под редакцией профессора, доктора экономических наук Э. Н. Яковлева. — М.: Государственное научное учреждение “Экспертно-аналитический центр Министерства образования Российской Федерации”, 2002. 20 с.