

# Автоматизация процесса квантования учебного текста на основе метода составления выдержек

**Александр Рыбанов,**  
Волжский политехнический институт (филиал)  
Волгоградский государственный  
технический университет  
vit@volpi.ru

**В статье анализируются возможности процедуры автоматизированного квантования учебных текстов для повышения качества образовательного процесса. В качестве метода автоматизированного квантования учебного текста предлагается метод составления выдержек с использованием модели линейных весовых коэффициентов.**

*Ключевые слова:* учебный контент, квантирование, принцип системного квантования, смысловая единица кванта, автоматизированное квантование

## Введение

Квантование — это разбиение учебной информации на элементарные фрагменты (учебные единицы, шаги, кадры) различного назначения (информационные, тренирующие, контролирующие, управляющие), направленное на уменьшение сложности усвоения смыслов, содержащихся в каждом фрагменте учебной информации<sup>1,2</sup>. Объём текстовой информации в этих фрагментах должен быть ограничен. IDEF0-модель верхнего уровня процедуры квантования учебного текста представлена на рис. 1.

Процесс квантования — это преобразование:  $T^* = f(T)$ , где  $T = (T_i | i = 1...n)$  — учебная информация, предназначенная для квантования,  $T_i$  — логически законченный фрагмент учебной информации  $T$ ;  $T^* = (T_i^* | i = 1...n)$  — квантованное представление учебной информации, где  $T_i^*$  — квант учебной информации, поставленный в соответствие фрагменту  $T_i$ . Степень квантования текста определяется следующим образом:

$$l = (T^* - T)/T.$$

Принцип системного квантования учебной информации предполагает учёт следующих закономерностей: учебная информация большого объёма

<sup>1</sup> Рыбанов А.А. Оценка качества текстов электронных средств обучения // Школьные технологии. 2011. № 6. С. 172–174.

<sup>2</sup> Рыбанов А. Анализ качества квантования учебного текста // Педагогические измерения. 2014. № 1. С. 3–16.

запоминается с трудом<sup>3</sup>; учебная информация, представленная компактно в определённой системе, лучше воспринимается; выделение

в учебной информации смысловых единиц способствует эффективному запоминанию<sup>4</sup>.

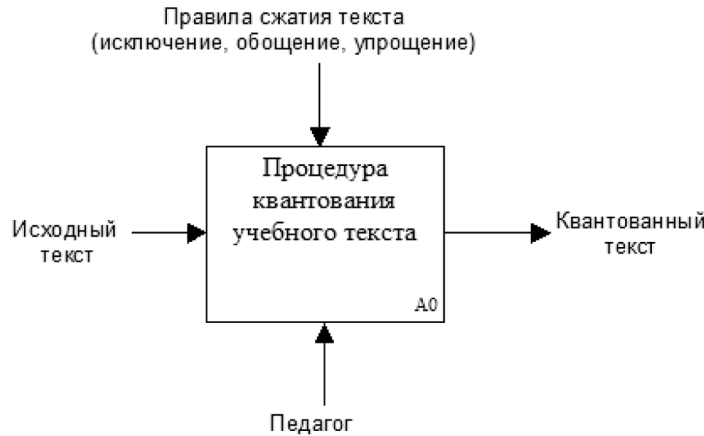


Рис. 1. IDEFO — модель процедуры квантования учебного текста

Учитывая то, что квант  $T_i^*$  учебной информации должен содержать наиболее информативную часть фрагмента  $T_i$ , требования к кванту учебной информации можно формализовать следующим образом: квант  $T_i^*$  учебной информации должен обладать более низкой избыточностью и большей энтропией, чем  $T_i$ ; квант  $T_i^*$  учебной информации по объёму должен быть меньше соответствующего ему фрагмента  $T_i$  учебной информации:  $|T_i^*| \leq T_i$ .

## Процесс построения педагогом кванта учебной информации

Процесс построения педагогом кванта  $T_i^*$  для фрагмента  $T_i$  учебной информации состоит из следующих этапов (рис. 2):

1) подготовительный (чтение и осмысление фрагмента  $T_i$  учебной информации);

2) аналитический (выделение основных смысловых единиц (предложения, слова, словосочетания), построение структуры кванта  $T_i^*$  для фрагмента  $T_i$  учебной информации);

3) непосредственное построение кванта  $T_i^*$  для фрагмента  $T_i$  учебной информации (выделенные ранее единицы располагаются в единый вторичный текст в соответствии со структурой кванта  $T_i^*$ ).

В качестве смысловых единиц кванта  $T_i^*$  для фрагмента  $T_i$  учебной информации могут быть использованы:  $g_1$  — полное (без изменений) ключевое предложение исходного текста  $T_i$ ;  $g_2$  — перефразированное ключевое предложение исходного текста  $T_i$ ;  $g_3$  — предложение из ключевых слов и словосочетаний исходного текста  $T_i$ ;  $g_4$  — предложение, обобщающее несколько предложений исходного текста  $T_i$ .

<sup>3</sup> Рыбанов А.А. Процедура квантования как основа повышения качества учебного материала // Образовательные технологии (г. Москва). 2014. № 3. С. 117–120.

<sup>4</sup> Рыбанов А.А. Квантование учебной информации как средство повышения качества контента в системах дистанционного обучения // Дистанционное и виртуальное обучение. 2014. № 7 (85). С. 4–21.

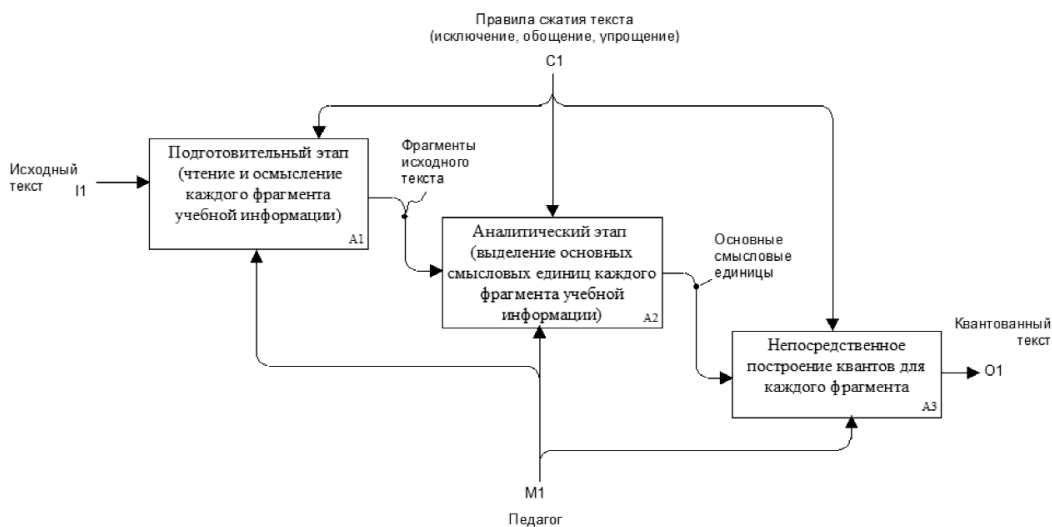


Рис. 2. IDEF0-модель процесса построения педагогом кванта учебной информации

### Автоматизация процесса квантования текста методом составления выдержек

Качественно подготовленные квантованные учебные тексты обеспечивают понимание учебного материала<sup>5</sup> у большинства учащихся, в силу того, что деление материала на части заметно сужает объём непосредственно воспринимаемой информации и смыслов в каждом фрагменте, делая тем самым постигаемым смыслы всего учебного текста. А работа с заданиями в тестовой форме к таким текстам обеспечивает усвоение содержания каждого текста. Поэтому совершенствование технологии квантования учебных<sup>6</sup> текстов является актуальной задачей. Обобщённая архитектура системы автоматизированного квантования текста приведена на рис. 3.

Для автоматизации процесса квантования<sup>7</sup> учебного текста может быть использован метод со-

ставления выдержек, который реализуется в два этапа.

*Этап 1.* Сопоставление исходного текста и фразовых шаблонов, в результате чего выделяются блоки наибольшей лексической и статистической релевантности.

*Этап 2.* Соединение выделенных фрагментов и формирование квантованного текста.

Для реализации первого этапа используют модель линейных весовых коэффициентов. В соответствии с ней каждому блоку  $W$  исходного текста автоматически приписываются весовые коэффициенты:

- 1)  $k_1$ , зависящий от расположения блока  $W$  в исходном тексте;
- 2)  $k_2$ , зависящий от частоты появления блока  $W$  в исходном тексте;
- 3)  $k_3$ , зависящий от частоты использования  $W$  в ключевых предложениях;
- 4)  $k_4$ , отражающий показатели статистической значимости блока  $W$ .

Затем по значениям  $k_1$ ,  $k_2$ ,  $k_3$  и  $k_4$  и коэффициентам настройки про-

<sup>5</sup> Рыбанов А.А. Квантование учебной информации как средство повышения качества учебного материала // Актуальные вопросы профессионального образования. 2014. Т. 12. № 15 (142). С. 28–30.

<sup>6</sup> Rybanov A.A. Educational information quantization for improving content quality in learning management systems // Turkish Online Journal of Distance Education. 2014. Т. 15. № 4. С. 303–321.

<sup>7</sup> Рыбанов А.А. Автоматизированное определение количественных характеристик текста // Современные научные исследования и инновации. 2014. № 2 (34). С. 5.

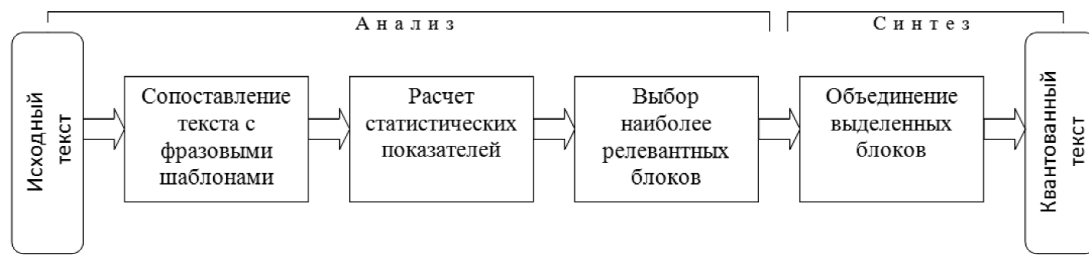


Рис. 3. Обобщённая архитектура системы автоматизированного квантования текста

граммы квантования  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$  и  $\alpha_4$  вычисляется коэффициент важности блока:

$$H(W) = \alpha_1 k_1 + \alpha_2 k_2 + \alpha_3 k_3 + \alpha_4 k_4.$$

По коэффициентам важности выполняется отбор блоков исходного текста в квантованный текст.

Для вычисления каждого весового коэффициента используется своя группа правил:

1) для  $k_1$  они учитывают расположение блока;

2) для  $k_2$  правила учитывают результаты автоматической индексации документа;

3) для  $k_3$  учитывается наличие в блоке таких ключевых фраз и выражений, как «в заключение...», «согласно результатам анализа...», «отличный от...», «малозначащий...» и т.п.;

4) для  $k_4$  правила учитывают вхождение термина в заголовки, ко-

лонтитулы, первый параграф текста, пользовательский профиль запроса и т.п.

Настройка с помощью коэффициентов  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$  и  $\alpha_4$  позволяет управлять *степенью квантования текста*<sup>8</sup>.

## Заключение

Достоинство описанной модели линейных весовых коэффициентов для квантования текстов заключается в простоте её реализации. Недостаток использования данной модели связан с возможностью формирования бессвязных квантованных текстов. Для его устранения вводится этап ручного редактирования результатов автоматизированного квантования текста.

<sup>8</sup> Рыбанов А.А. Повышение качества учебного материала на основе применения процедуры квантования // Школьные технологии. 2015. № 2. С. 169–171.