

Экспертная оценка программных продуктов для аннотирования документов

*Бурмистров Александр Сергеевич
Волжский политехнический институт
студент*

*Свиридова Ольга Викторовна
Волжский политехнический институт
Кандидат технических наук, доцент кафедры «Информатика и технология программирования»*

Аннотация

В статье рассмотрены понятия и представлены результаты сравнительного анализа по методу Саати программных продуктов для аннотирования документов.

Ключевые слова: аннотирование, метод Саати.

Expert evaluation of software products for annotation of documents

*Burmistrov Aleksandr Sergeevich
Volzhsky Polytechnical Institute
student*

*Sviridova Olga Victorovna.
Volzhsky Polytechnical Institute
Ph.D., Associate Professor, Department «Computer technology and programming»*

Abstract

The article deals with concepts and presents the results of a comparative analysis using the Saati method of software products for annotation of the document.

Keywords: annotation of the document, Saati method.

В данной статье проведен анализ данных программных продуктов (ОРФО 5.0, Либретто, МедиаЛингва Аннотатор SDK 1.0, Oracle Context, Intelligent Text Miner), позволяющих автоматически составлять аннотацию, и определены весовые коэффициенты выбранных критериев качества продуктов методом Саати.

В результате анализа потребностей пользователей по отношению к программным средствам, создающим аннотацию, были выбраны перечисленные ниже показатели в качестве критериев для сравнительного анализа программ аналогов:

1. A1 - Объём вводимых данных
2. A2 - Количество типов расширения файлов
3. A3 - Количество возможных диапазонов для конвертирования
4. A4 - Качество логического смысла выводимой аннотации
5. A5 - Удобство интерфейса

Используя аналитическую процедуру иерархического метода Саати, определим вес каждого критерия качества. Матрица парных сравнений заполняется по правилам, которые представлены в таблице 1.

Таблица 1. Значения коэффициентов матрицы парных сравнений

X_{ij}	Значение
1	i -ый и j -ый критерий практически равноценны
3	i -ый критерий на немного предпочтительнее j -го
5	i -ый критерий предпочтительнее j -го
7	i -ый критерий имеет значительное предпочтение в отличие от j -го
9	i -ый критерий явно предпочтительнее j -го

Полученная матрица парных сравнений изображена в таблице 2, в которой представлены веса критериев и среднее геометрические значения.

Таблица 2. Матрица парных сравнений, средние геометрические и веса критериев

	A1	A2	A3	A4	A5	Среднее геометрическое	Веса критериев
A1	1	3	5	7	9	3,94	0,50
A2	1/3	1	5	5	7	2,26	0,29
A3	1/5	1/5	1	3	5	0,90	0,11
A4	1/7	1/5	1/3	1	7	0,58	0,07
A5	1/9	1/7	1/5	1/7	1	0,21	0,03
Сумма						7,89	1,00

На рисунке 1, представлена диаграмма весовых коэффициентов для критериев A1, A2, A3, A4, A5.

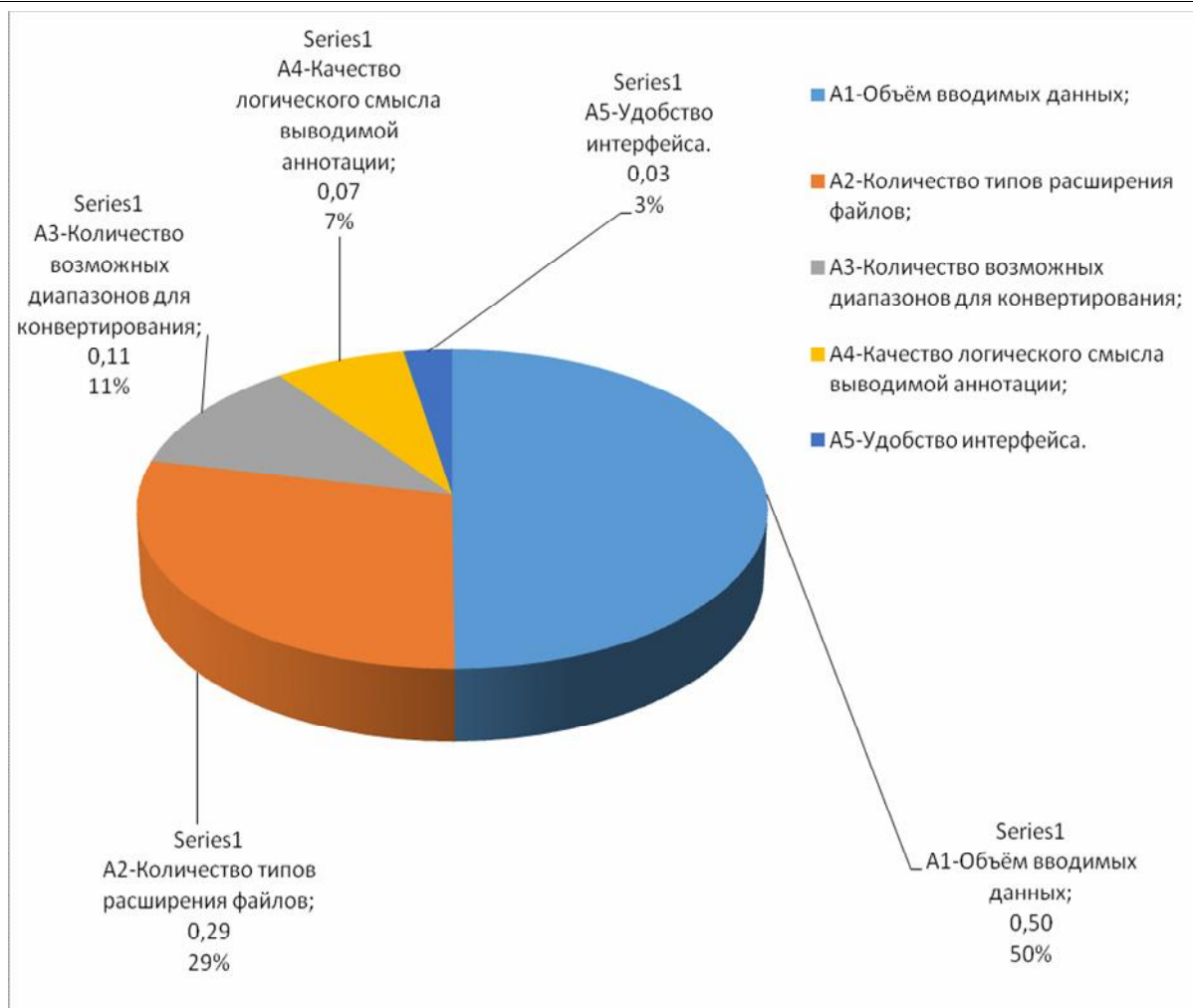


Рисунок 1. Весовые коэффициенты критериев качества

Проверим непротиворечивость матрицы парных сравнений. Рассчитаем суммы столбцов в матрице парных сравнений:

$$R_1=1,79; R_2= 4,54; R_3= 11,53; R_4= 16,14; R_5=29.$$

Следующий шаг расчёт вспомогательной величины L . Получим сумму ряда, произведения сумм столбцов матрицы и весовых коэффициентов: $L = 5,49$. По формуле вычисляем индекс согласованности $ИС = (L-N)/(N-1) = 0,12$.

Величина случайной согласованности для размерности матрицы парных сравнений: $СлС = 1,12$. Вычисленное отношение согласованности по формуле $ОС=ИС/СлС = 0,11$. не превышает 0.2, поэтому уточнение матрицы парных сравнений не требуется.

Полученные коэффициенты используем для определения интегрального показателя качества следующих программных продуктов, направленных на аннотирование документов:

1. ОРФО 5.0;
2. Либретто;
3. МедиаЛингва Аннотатор SDK 1.0;
4. Oracle Context;
5. Intelligent Text Miner.

Распишем категориальную шкалу от 0 до 7 (где число 0 означает, что качество не удовлетворительно, а число 7 указывает на максимально достижимый уровень качества) для функциональных возможностей выше перечисленных программ.

Полученные значения весовых коэффициентов (a_i), соответствуют функциональным возможностям программ аналогов:

1. Объём вводимых данных: $a_1 = 0,50$;
2. Количество типов расширения файлов: $a_2 = 0,29$;
3. Количество возможных диапазонов для конвертирования: $a_3 = 0,11$;
4. Качество логического смысла выводимой аннотации: $a_4 = 0,07$;
5. Удобство интерфейса: $a_5 = 0,03$;

где $\sum a_i = 1$.

Определим по выбранной шкале количественные значения функциональных возможностей X_{ij} (таблица 3) и для выбранных программ вычислим интегральные показатели качества.

Таблица 3. Интегральные показатели качества

Критерии	Весовые коэффициенты	Программные продукты					Базовые значения
		ОРФО 5.0	Либретто	МедиаЛингва Аннотатор SDK 1.0	Oracle Context	Intelligent Text Miner	
A1-Объём вводимых данных;	0,50	7	7	7	7	6	6,8
A2-Количество типов расширения файлов;	0,29	6	6	5	7	5	5,8
A3-Количество возможных диапазонов для конвертирования;	0,11	4	3	3	6	4	4
A4-Качество логического смысла выводимой аннотации;	0,07	6	5	6	5	6	5,6
A5-Удобство интерфейса.	0,03	6	5	5	5	6	5,4
Интегральные показатель качества Q		6,27	6,05	5,84	6,68	5,49	6,07

где $Q_j = \sum a_i * X_{ij}$ – интегральный показатель качества для j -го программного продукта.

Отообразим на лепестковой диаграмме интегральный показатель качества каждого выбранного программного продукта (рисунок 2) .

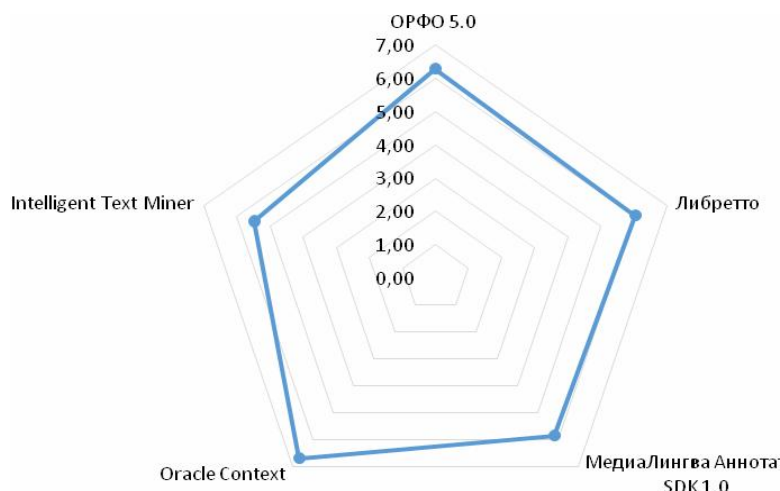


Рисунок 2. Лепестковая диаграмма интегральных показателей качества программ

Значения характеристик функциональных возможностей (критериев) представлены в виде лепестковой диаграммы на рисунке 3.

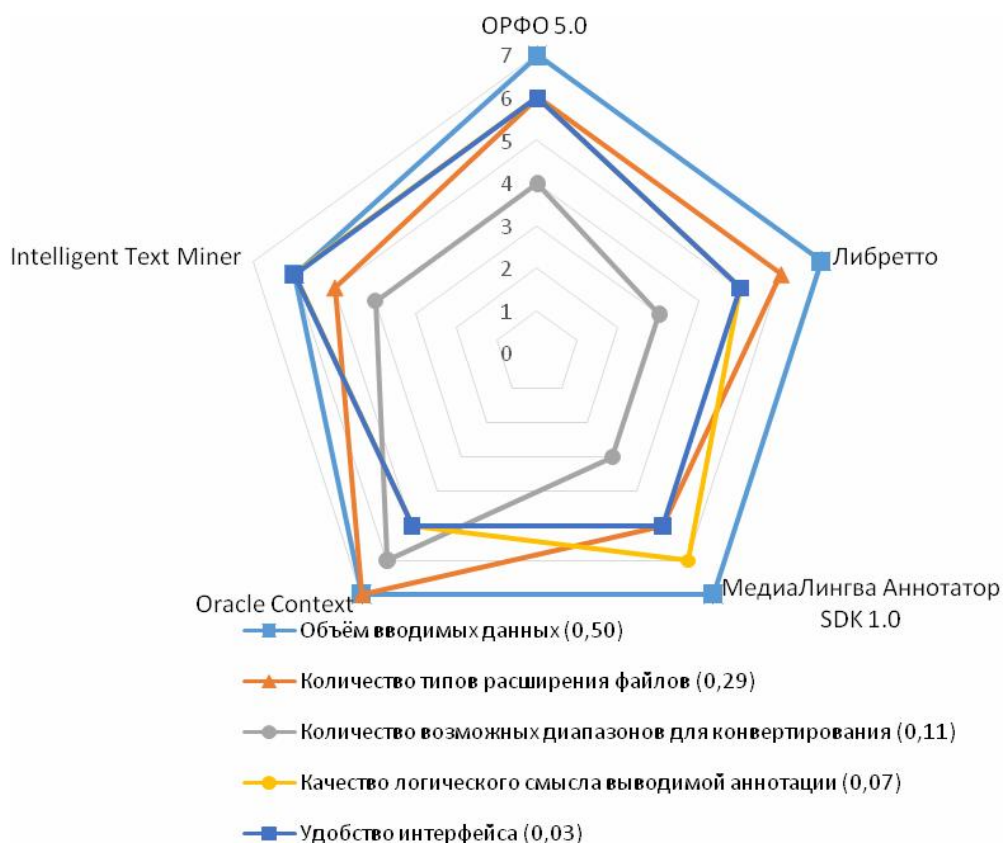


Рисунок 3. Лепестковая диаграмма значений функциональных характеристик

Сравнительный анализ программных продуктов показал, что из всех программных аналогов только Oracle Context и ОРФО 5.0 имеют значение интегрального показателя качества, превосходящее базовое, а у остальных программных аналогов показатель оказался ниже базового. Так же анализ показал, что для Oracle Context и ОРФО 5.0 надо улучшать показатели по таким критериям как: «Удобство интерфейса» и «Количество возможных диапазонов для конвертирования».

Предлагаемая методика экспертной оценки программных продуктов позволила количественно оценить их качество с точки зрения уровня реализуемых функций и выявила функции, которые необходимо улучшить в некоторых рассматриваемых программных продуктах.

Библиографический список

1. Кондрацкий Д.Е., Рыбанов А.А. Исследование методов и алгоритмов автоматизированной системы оценки альтернативных вариантов методом Т.Саати // NovaInfo.Ru. 2016. Т. 3. № 46. С. 107-116.
2. Рыбанов А. Определение весовых коэффициентов сложности тем учебного курса на основе алгоритма Саати // Педагогические измерения. 2014. № 4. С. 21-28.
3. Рыбанов А.А., Макушкина Л.А. Технология определения весовых коэффициентов сложности тем дистанционного курса на основе алгоритма Саати // Открытое и дистанционное образование. 2016. №1 (61). С. 69-79.
4. Рыбанов А.А., Усмонов М.С.О., Попов Ф.А., Ануфриева Н.Ю., Бубарева О.А. Информационные системы и технологии. Центр научной мысли (г. Таганрог); Научный редактор И.А. Рудакова; Редакционная коллегия: Рудакова И.А., Гребенщиков Г.Ф., Акутина С.П., Краснолуцкий В.П.. Москва, 2013. Том Часть 4 Информационные системы и технологии
5. Яцко В.А. Алгоритмы и программы автоматической обработки текста // Вестник ИГЛУ №1 (17), 2012. URL:<http://cyberleninka.ru/article/n/algoritmy-i-programmy-avtomaticheskoy-obrabotki>