



OSTIS-2012

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:514

ОЦЕНКА КАЧЕСТВА ИНТЕЛЛЕКТУАЛЬНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

Бахтизин В.В., Неборский С.Н.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

bww@bsuir.by

sergei.neborski@gmail.com

Рассматривается проблема оценки качества интеллектуальных информационных систем в контексте современных международных стандартов в области качества, приводятся примеры мер для оценки качества.

Ключевые слова: качество интеллектуальных информационных систем, оценка качества интеллектуальных информационных систем.

ВВЕДЕНИЕ

Проблема оценки качества информационных систем (ИС) является актуальной и важной, т. к. именно результаты оценки качества дают понимание того, что пользователь может ожидать от рассматриваемой ИС. Оценка качества системы должна производиться на всех этапах жизненного цикла ИС, ведь только так разработчик может корректировать как техническую составляющую процесса разработки (принимать архитектурные решения), так и организационную (например, усилить команду разработчиков специалистами, способными решить конкретную проблему в конкретном модуле, выявленную в результате оценки).

1. Проблема оценки качества интеллектуальных информационных систем

Оценка качества программного обеспечения (ПО) не является новой. Первые модели начали появляться еще в 80-х годах XX века (модели Боэма, Дроми, FURPS), и сегодня существует ряд международных стандартов, посвященных именно качеству ПО (ISO/IEC 9126, ISO/IEC 15504, ISO/IEC 15939, ISO/IEC 25000). Однако данные стандарты не учитывают особенностей интеллектуальных информационных систем (ИИС), а сама проблема оценки качества ИИС не является проработанной и хорошо изученной. Данная работа стремится соотнести текущие наработки в области оценки качества ИИС с актуальными действующими стандартами качества ПО.

Оценивая качество ИИС, в настоящее время

оценивают преимущественно функциональность. Очевидно, оценка функциональности имеет свои особенности, и классические подходы к оценке (в частности, использование метрик оценки качества программных средств (ПС) по стандарту 9126-2 [ISO/IEC TR 9126-2, 2003]) не всегда применимы. Существующие подходы к оценке качества ИИС можно условно разделить на два типа [Лахутин, 1970]:

- 1) качество интеллектуальной системы посредством формальной процедуры ставится в зависимость от меры ошибок системы, определяемой тем или иным способом;
- 2) качество интеллектуальной системы оценивается экспертами.

В силу проблем, связанных с вовлечением эксперта в процесс оценивания (“эффект персоны”, когда психологический настрой человека, испытывающего систему, ориентирован прежде всего не на знакомство с ней, а на подтверждение собственных сложившихся взглядов и идей или на подтверждение своей компетенции, и “эффект персонализации”, когда в процессе знакомства эксперта с системой происходит отождествление им потенциальных возможностей системы с возможностями человека [Григорьев, 2004]), процедура оценки должна быть максимально объективна и не зависеть от участвующих в ней лиц.

В литературе можно встретить отождествление качества ИИС и качества реализации ИИС, однако их необходимо разделить. Качество реализации ИИС зависит от характеристик качества используемых моделей знаний и обрабатываемых концептуальных информационных ресурсов

[Shalfeeva, 2010].

Существуют структурные подходы для оценивания качества различных классов информационных ресурсов интеллектуальных программных систем. Эти подходы позволяют обнаруживать дефекты и недостатки представления информации, несогласованности между определениями и фактами. Однако данные подходы в большей степени перекликается с качеством процесса разработки, но не непосредственно качеством ИИС.

2. Стандарты качества программного обеспечения

Актуальными стандартами в области качества ПО являются стандарты ISO/IEC 25000 SQuaRE (Software Quality Requirements and Evaluation – требования к качеству и оценка программного продукта). В Республике Беларусь введен в действие стандарт СТБ ISO/IEC 25000-2009 «Разработка программного обеспечения: Требования к качеству и оценка программного продукта (SQuaRE): Руководство по SQuaRE» [СТБ ISO/IEC 25000, 2009]. Отдельные группы стандартов SQuaRE приведены на рисунке 1.

Требования к качеству 2503n	Модель качества 2501n	Оценка качества 2504n
	Управление качеством 2500n	
	Измерение качества 2502n	

Рисунок 1 – Группы стандартов SQuaRE 25000

Данные стандарты регламентирует иерархическую модель качества ПС. На верхнем уровне находятся характеристики, каждая из которых разделяется на подхарактеристики. Подхарактеристики определяются мерами. Мера – это атрибут объекта и метод его измерения. Концептуально, такая модель качества справедлива и для ИИС.

Однако если стандарты ISO/IEC 25000 жестко определяют характеристики и подхарактеристики качества, то для оценки конкретных ИИС определенные стандартами характеристики не всегда подходят в полной мере, и необходимо определять актуальные характеристики и ранжировать их по степени важности.

Стандарт ISO/IEC 25000 SQuaRE [СТБ ISO/IEC 25000, 2009] выделяет две модели качества ПО:

- качество в использовании;
- качество продукта.

Модель качества в использовании (рисунок 2) определяет 5 характеристик качества, связанных с

взаимодействием пользователя с системой: продуктивность, эффективность, способность к удовлетворению, отсутствие риска, покрытие контекста.

Модель качества продукта декомпозирует свойства ПО по 8 характеристикам:

- функциональная пригодность,
- производительность,
- совместимость,
- удобство использования,
- надежность,
- защищенность,
- сопровождаемость,
- мобильность.



Рисунок 2 – Модель качества в использовании

Так как одним из характерных признаков ИИС является способность к самообучению, можно утверждать, что ИИС является постоянно развивающейся системой. Значит, для ИИС больший интерес имеет модель качества продукта (рисунок 3), нежели модель качества в использовании.



Рисунок 3 – Модель качества продукта

3. Метод оценки качества интеллектуальных информационных систем

Для оценки качества ИИС предлагается использовать метод оценки качества по ISO/IEC 25020 [ISO/IEC 25020, 2006], который по сути приходит на смену стандарту ISO/IEC 15939:2002 [ISO/IEC 15939, 2002].

Схематично модель, положенная в основу данного метода, показана на рисунке 4: меры определяют подхарактеристики качества, на основе подхарактеристик получается оценка отдельных характеристик, и уже на основе характеристик определяется интегральная оценка качества.

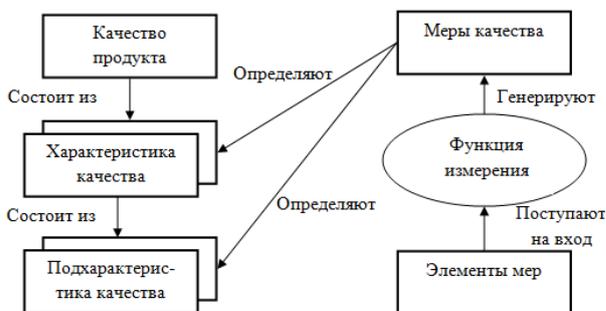


Рисунок 4 – Модель оценки качества

Интегральная оценка качества P^Q рассчитывается по формуле (1):

$$P^Q = \frac{\sum_{k=1}^K (P_k^X \cdot V_k^X)}{\sum_{k=1}^K V_k^X}, \quad (1)$$

где P_k^X - оценка k -й характеристики качества,

V_k^X - весовой коэффициент k -й характеристики качества, причем $\sum_{k=1}^K V_k^X = 1$.

Для определения степени важности отдельных подхарактеристик и характеристик используются весовые коэффициенты V_k , определяемые методами экспертных оценок (например, могут быть использованы методы коллективной работы экспертной группы или метод Дельфи).

4. Выбор мер качества

И все же, основной проблемой оценки качества ИИС можно считать правильный выбор мер качества, т. к. меры позволяют дать объективную количественную оценку конкретных свойств системы. Наиболее проработанной здесь является оценка качества поисковых машин. Можно привести методику оценки, суть которой состоит в получении интегральной оценки экспертом на основе значений вычисленных мер качества [Масевич, 1996]. Основная идея состоит в сборе нескольких параметров поисковой машины, описывающих поведение пользователей и свойства самой системы:

- насколько быстро выполняются запросы;
- сколько занимает поисковый индекс;
- насколько хорошо удовлетворяются информационные потребности пользователя;
- сколько информации удалось получить по запрашиваемой теме;

- насколько удобно пользоваться системой.

Соответственно, можно привести примеры следующих мер для оценки поведения пользователей при поиске в интернете [Rambler, 2011]:

- вероятность возврата в поисковую выдачу после перехода на найденную страницу;
- вероятность просмотра результата менее чем 20 секунд и возвращения в поисковую выдачу;
- средняя позиция ссылок на страницы, на которые осуществляется переход пользователя;
- вероятность полного отсутствия переходов на страницы поисковой выдачи.

Приведенные меры согласуются с требованиями к мерам стандарта ISO/IEC 25020 [ISO/IEC 25020, 2006], они могут быть представлены в относительных единицах и нормированы с тем, чтобы значения изменялись на отрезке $[0, 1]$, причем чем ближе к 1, тем выше качество ИИС. Однако поисковые машины – лишь один класс ИИС, и для других классов характеристики и меры качества будут значительно отличаться.

ЗАКЛЮЧЕНИЕ

Обобщая вышесказанное можно отметить, что проблема оценки качества ИИС остается нерешенной в силу многогранности ИИС и узкой специфики их отдельных классов. В то же время, модели и методы оценки качества ПО, определяемые новыми международными стандартами ISO/IEC 25xxx, могут и должны быть адаптированы с целью их применения к ИИС.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- [ISO/IEC TR 9126-2, 2003] ISO/IEC TR 9126-2:2003: Программная инженерия – Качество продукта – Часть 2: Внешние метрики. – Предст. 2003-07-01. – Женева: ISO/IEC-IEEE, 2003. – 86 с.
- [Лахутин, 1970] Лахутин Д.Г. О проблеме оценки поисковых систем. Ч. 1. – Д.Г. Лахутин, С.О. Чернявский // НТИ. Сер. 2. №1. 1970. С. 24-34.
- [Григорьев, 2004] Григорьев В.А. Оценка качества интеллектуальных систем, оперирующих неформальными понятиями / В.А. Григорьев // Наука и образование, – 2004 г. – №10
- [Shalfeeva, 2010] Shalfeeva E.A. Monitoring of conceptual informational resources for intelligent software systems / E.A. Shalfeeva // Proc. First Russia and Pacific Conference on Computer Technology and Applications, – 2010. – С.140-144
- [СТБ ISO/IEC 25000, 2009] СТБ ISO/IEC 25000-2009: Разработка программного обеспечения. Требования к качеству и оценка программного продукта (SQuaRE). Руководство по SQuaRE. – Введ. 2010-01-01. – Минск : Госстандарт Респ. Беларусь, 2009. – 44 с.
- [ISO/IEC 25020, 2006] ISO/IEC 25020: Software engineering – Software product quality requirements and evaluation (SQuaRE) – Measurement reference model and guide. – Предст. 2006-08-08. – Канада: ISO/IEC-IEEE, 2006. – 24 с.

[ISO/IEC 15939, 2007] ISO/IEC 15939: Software engineering – Software measurement process – Предст. 2007-07-23. – Женева: ISO/IEC-IEEE, 2007. – 38 с.

[Масевич, 1996] Масевич А.Ц. Качество информационного продукта автоматизированных библиотечных систем / А.Ц. Масевич, В.П. Захаров // Материалы 3-й Международной конференции “Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества. Крым – 96” (Форос, 1—9 июня 1996г.) / Ялта, 1996. Т. 1. С. 178—189.

[Rambler, 2011] Rambler pFound - метрика качества поиска [Электронный ресурс] / Рамблер Интернет Холдинг, 2011. – Режим доступа: http://rcdl2011.vsu.ru/doc/full_text/Report.5.6.pdf – Дата доступа: 17.12.2011.

INTELLIGENT SYSTEMS QUALITY EVALUATION

Bakhtizin V.V., Niaborski S.N.

*Belarusian State University of Informatics and
Radioelectronics, Minsk, Republic of Belarus*

bww@bsuir.by

sergei.niborski@gmail.com

This paper analyses the problem of intelligent systems quality evaluation considering current stage of development of international standards devoted to software quality. It gives examples of quality measure for intelligent systems and shows how to apply quality evaluation method defined in ISO/IEC 25020 to intelligent systems.

INTRODUCTION

The problem of quality evaluation of intelligent systems is essentially important because results of quality estimation allow to understand what a user can expect from the system. Quality evaluation must be conducted on all stages of system development, thereby allowing a developer to adjust technical and administrative approaches to the project. This paper aims to provide a link between international quality standards for software evaluation and intelligent systems and shows how to apply those standards to measure intelligent systems quality.

MAIN PART

When trying to evaluate intelligent system quality, in most cases staff deal with functionality. The functionality of intelligent system is observed from two perspectives:

- 1) intelligent system quality is defined as a function of error quantity the number of which is captured in a given way;
- 2) intelligent system quality is estimated by an expert.

System quality, however, is not only functionality – it includes much more characteristics each one of which is further decoupled into sub-characteristics. Software quality evaluation techniques can be applied to intelligent systems estimation as well.

In particular, international standard ISO/IEC 25000-2009 SQuaRE can be used to define product quality model by choosing 8 characteristics: functional suitability, performance efficiency, compatibility, usability, reliability, security, maintainability, and portability. These 8 after certain modifications can be applied to intelligent systems as well.

To evaluate intelligent system quality an evaluation method defined in ISO/IEC 25020 can be used. The standards ISO/IEC 25020 is coming into replacement of ISO/IEC 15939:2002 with some improvements and adaptation to ISO/IEC 25000 line. The model which lies underneath of the evaluation method is the following: measures define sub-characteristics, having estimate on a given sub-characteristics, quality characteristic is defined. From a given estimates of quality characteristics an integrated quality estimate is coming.

The method of intelligent system quality evaluation implies that given measures, sub-characteristics and characteristics participate in the estimate with specific weights. The weights are defined by experts considering experience, statistical results of previous developments, and other factors.

Major problem of quality evaluation is not characteristic definition – but definition of quality measures. For intelligent systems most developed is the area of search machines. Examples of measures for Internet searches are:

- probability of return to search results;
- probability of result analysis for 20 seconds and returning back to result set;
- average position of link to target pages in result set;
- probability of no clicks on result set links.

The given measures are compatible with requirements and recommendation of ISO/IEC 25020 standard to quality measures. They are defined in relative numbers and normalized in a way that values are changing on [0, 1] interval, closer to 1 corresponds to better intelligent system quality.

Search machines, however is only one class of intelligent systems and for other systems characteristics and measures to evaluate quality will differ significantly.

CONCLUSION

The problem of intelligent system quality evaluation remains unresolved because of different applications of intelligent systems and their variety. At the same time, international standards ISO/IEC 25000 in software quality can be applied to intelligent systems as well, providing guidance on precious estimation of intelligent systems quality.