

# **ПРОГРАММНОЕ СРЕДСТВО ДЛЯ ХРАНЕНИЯ И АНАЛИЗА РЕЗУЛЬТАТОВ СОЦИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИИ ASP.NET**

*Лащиновская К.Д.*

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

*Петюкевич Н.С. – ст. преподаватель*

В статье дано краткое описание функциональности программного средства для хранения и анализа результатов социологических исследований. Также приводится обоснование некоторых архитектурных решений, принятых при проектировании программного продукта.

Социологическое исследование представляет собой деятельность, направленную на получение объективных знаний о социальных отношениях, явлениях и процессах. Социологические исследования в зависимости от задач, которые они решают, и методов, которыми они пользуются,

могут быть подразделены на теоретические и эмпирические. Теоретические исследования ставят целью развитие и совершенствование научных представлений об изучаемом предмете, а эмпирические (прикладные) посвящены решению конкретной социальной проблемы. Результаты эмпирических социологических исследований необходимы для более эффективного управления предприятием, отраслью, страной. На сегодняшний день социологический анализ активно применяется для построения рекламных кампаний, оценки мнений сотрудников организаций и учащихся учебных заведений. Одним из наиболее удобных методов сбора информации для социологического исследования является опрос. Анализ данных, полученных в результате проведения опроса, позволяет сформулировать утверждения, характеризующие испытуемую группу лиц.

Для анализа результатов эмпирических социологических исследований часто применяются такие программные средства, как SPSS, Stata и SurveyMonkey. Достоинством SPSS и Stata является очень мощный аналитический инструментарий. Эти программные продукты имеют встроенный командный язык для получения нужных выборок данных и их анализа. Однако при использовании данных программных средств результаты проведения опросов и отчеты хранятся на локальном компьютере, что делает затруднительной работу в команде исследователей. Веб-сервис SurveyMonkey является более современным решением, которое позволяет совместно создавать опросы и анализировать полученные результаты в online-режиме. Однако при проведении опросов в нескольких странах (на нескольких языках) данные, полученные из разных стран, объединяются в одну выборку [1]. Таким образом, отсутствует возможность проводить аналитику в разрезе стран.

В разработанном программном средстве были учтены недостатки аналогов. Данное программное средство представляет собой веб-приложение, которое позволяет работать командам исследователей, а также поддерживает работу с исследованиями, проводимыми в нескольких странах. Приложение ориентировано на мелкие и средние компании, специализирующиеся на проведении эмпирических социологических исследований, а также на организации, которые периодически проводят социологические опросы среди своих сотрудников.

При создании нового исследования существует возможность выбрать пользователей, которые будут иметь доступ к нему. Для каждого пользователя устанавливается уровень доступа к исследованию: наблюдение (только просмотр результатов анализа), работа с данными (загрузка новых данных исследования), аналитика (построение диаграмм и отчетов) и полный доступ к исследованию. Указание уровней доступа позволяет разделить обязанности в команде исследователей.

Для создания исследования необходимо указать опорные вопросы и варианты ответа на них. Доступны следующие виды вопросов: вопросы с единственным выбором, вопросы с множественным выбором, открытые вопросы с текстовым ответом. Если исследование проводится на нескольких языках, для каждого вопроса и вариантов ответа на него нужно указать соответствие для выбранного языка.

Добавление данных в приложение осуществляется путем загрузки файла в формате XLSX, содержащего ответы респондентов на вопросы исследования. Для корректного считывания данные в файле должны быть организованы по единому шаблону. За шаблон файла данных был взят файл, который генерируется сервисом для создания опросов Google Forms. При таком форматировании файла заголовками колонок таблицы являются вопросы исследования на одном из выбранных языков, а каждая строка таблицы содержит ответы одного респондента.

Разработанное приложение предоставляет достаточно широкие аналитические возможности.

Для вопросов с вариантами ответа доступен частотный анализ для каждого варианта ответа. Результат анализа может быть представлен в виде круговой или столбчатой диаграммы, при этом анализ можно проводить в разрезе стран и временных промежутков, в которых проводилось исследование. Также существует возможность строить диаграммы корреляции, для чего выбирается независимый вопрос и вопрос, зависящий от него. Для каждого варианта ответа независимого вопроса строится диаграмма, отражающая распределение по вариантам ответа для зависимого вопроса. Все диаграммы, построенные в программном средстве, могут быть сохранены на компьютер в виде изображения в формате PNG.

Для вопросов с выбором одного варианта ответа существует возможность вычисления коэффициента корреляции. Коэффициент корреляции – это мера оценки статистической взаимосвязи между парой (как правило) или большим количеством переменных. Нужный коэффициент выбирается исходя из типа шкалы переменной (номинальная, порядковая, интервальная, метрическая шкалы). В большинстве случаев каждый вариант ответа на вопрос имеет свой идентификатор, который не зависит от эмпирического значения варианта ответа. К вопросам, имеющим подобные варианты ответа, относятся, например, вопросы по гендерной принадлежности испытуемого, его национальности, семейном положении, роде деятельности. В этом случае переменные, хранящие результаты ответов на вопросы, относятся к номинальной шкале. Для переменных номинальной шкалы может быть получен коэффициент корреляции Крамера [2]. Данный

коэффициент измеряется в пределах от 0 до 1. Чем больше значение коэффициента, тем сильнее связь между переменными.

Для вопросов, предполагающих текстовый ответ, существует возможность создания диаграммы «облако слов». Для построения диаграммы все текстовые ответы анализируются по частоте встречаемости каждого слова. Найденные слова размещаются на диаграмме, причем размер шрифта каждого слова пропорционален частоте встречаемости слова среди всех ответов.

В программном средстве также существует возможность создавать фильтры для выборки данных. Условия фильтра могут быть объединены посредством оператора И либо ИЛИ. Данные, подходящие под фильтр, могут быть загружены на локальный компьютер в формате XLSX.

При проектировании программного средства большое внимание было уделено архитектуре хранилища данных. База данных средства анализа результатов опроса должна поддерживать возможность работы с большими объемами данных, а также должна позволять быстро извлекать ответы респондентов по переданным условиям.

В разработанном приложении для удобства выборки данных таблицы в базе данных построены по архитектуре «звезда»: таблица, содержащая ответы респондентов, является центральной и связана через внешние ключи с таблицами вопросов, вариантов ответов, респондентов и стран. Необходимо отметить, что подобная схема связи таблиц является основой для создания многомерной базы данных (куба) посредством использования службы Microsoft Analysis Services. Многомерная база данных является более совершенным решением для построения аналитических систем, чем реляционная база данных, так как в кубе данные заранее агрегируются по различным измерениям, и существует возможность делать множественные «срезы» данных по нужным параметрам. Однако весь аналитический функционал в разработанном приложении может быть реализован через вызов хранимых процедур и использование куба является избыточным. Поэтому было принято решение отказаться от многомерной базы в пользу обычной реляционной базы.

Так как данное программное средство призвано решить, в том числе, проблему работы над исследованием в команде, результаты проведения социологических опросов должны храниться на компьютере, к которому имеют доступ все исследователи в команде. При этом пользователи должны иметь возможность работать с данными независимо друг от друга. Таким образом, было принято решение о построении приложения по клиент-серверной архитектуре с использованием принципов REST. Серверная часть программного средства была написана при помощи технологии ASP.NET WebApi, которая позволяет создавать RESTful-сервисы. Для работы с файлами в формате XLSX была использована библиотека EPPlus. Преимуществом данной библиотеки является то, что она не имеет сторонних зависимостей и может функционировать без установленного на сервере Microsoft Excel. В качестве базы данных была выбрана MSSQL Server 2019. Для доступа к данным были использованы EntityFramework (с подходом database-first), а также библиотека ADO.NET (использовалась для вызова хранимых процедур). Клиентская часть программного средства была написана на основе JavaScript-фреймворка Vue, характеризующегося высокой скоростью работы и наличием большого количества устанавливаемых компонентов.

**Список использованных источников:**

1. Документация для сервиса SurveyMonkey [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://help.surveymonkey.com>
2. Статистики, рассчитываемые для таблиц сопряженности [Электронный ресурс]. – Электронные данные. – Режим доступа: [https://www.ibm.com/support/knowledgecenter/ru/SSLVMB\\_25.0.0/spss/base/idh\\_xtab\\_statistics.html](https://www.ibm.com/support/knowledgecenter/ru/SSLVMB_25.0.0/spss/base/idh_xtab_statistics.html)