

### *Литература*

1. Афанасьев В.В. Применение методов математической статистики в научных исследованиях / В.В. Афанасьев // Ярославский педагогический вестник. – №4. – 2006. – С. 5-12.
2. StatSoft, Inc. (2001). Электронный учебник по статистике. – М.: StatSoft. – Режим доступа: <http://www.statsoft.ru/home/textbook/default.htm> (дата: 1 авг. 2013г.)
3. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Учебное пособие / А.Д. Наследов. – СПб.: Речь, 2004. – 392 с.
4. Антонов А.В. Системный анализ. Учеб. Для вузов / А.В. Антонов. – М.: Высш. Шк. 2004 – 454 с.:ил.

### *References*

1. Afanas'ev V.V. Primenenie metodov matematicheskoy statistiki v nauchnyh issledovaniyah / V.V. Afanas'ev // Yaroslavskij pedagogicheskij vestnik. – №4. – 2006. – S. 5-12.
2. StatSoft, Inc. (2001). Ehlektronnyj uchebnyj po statistike. – M.: StatSoft. – Rezhim dostupa: <http://www.statsoft.ru/home/textbook/default.htm> (data: 1 avg. 2013g.)
3. Nasledov A.D. Matematicheskie metody psihologicheskogo issledovaniya. Analiz I interpretaciya dannyh. Uchebnoe posobie / A.D. Nasledov. – SPb: Rech', 2004. – 392 s.
4. Antonov A.V. Sistemnyj analiz. Ucheb. Dlya vuzov / A.V. Antonov. – M.: Vyssh. Shk. 2004 – 454 s.:il.

УДК 004.67

doi: 10.18101/978–5–9793–0814–2–258–261

## **Разработка информационной системы сбора социологических данных\***

\* Исследование выполнено при финансовой поддержке РГНФ. «Общественное сознание современной молодежи Монголии», проект № 15-23-03001

### **© Дармаев Тумэн Гомбоцыренович**

кандидат физико-математических наук, доцент, директор Научно-образовательного и инновационного центра системных исследований и автоматизации Бурятского государственного университета, г. Улан-Удэ

E-mail: [dtg@bsu.ru](mailto:dtg@bsu.ru);

### **© Хабитуев Байр Викторovich**

инженер Научно-образовательного и инновационного центра системных исследований и автоматизации Бурятского государственного университета, г. Улан-Удэ

E-mail: [bairinc0@gmail.com](mailto:bairinc0@gmail.com);

### **© Нямаахүүгийн Батжаргал**

профессор Института иностранного языка Монгольского университета науки и технологий, Улан-Батор, Монголия

E-mail: [intel\\_bat@yahoo.com](mailto:intel_bat@yahoo.com)

Сбор данных посредством опросников – популярный метод получения статистических данных для социологии. Спецификой такого рода исследований является большое количество респондентов. Проведение масштабного тестирования означает большое количество проблем связанных как с непосредственным проведением тестирования, так и с последующим сбором, систематизацией и первоначальной обработкой данных. Внедрение современных информационных технологий (систем сбора данных опроса) в данный процесс позволит автоматизировать большую часть рутинной работы. В работе рассматривается разработка системы сбора и анализа данных масштабных социологических опросов.

**Ключевые слова:** система сбора, базы данных, анализ данных, системы тестирования, социологические опросы.

## **Development of the information system for sociological data collection**

### **Tumen G. Darmaev**

PhD in Physics and Mathematics, A/Professor, Ulan-Ude Buryat State University, Scientific and Educational Centre of System Research and Automatization

### **Bair V. Khabituev**

Engineer, Ulan-Ude, Buryat State University, Scientific and Educational Centre of System Research and Automatization

### **Nyamaahuugiin Batgargal**

Professor, Mongolia, Ulaanbaatar, Mongolian University of Science and Technology, Institute of Foreign Language

Data collection through testing is very popular method of statistics data collection for sociological research. Such type of research often means that large bulk of respondent data needs to be collected. Global experiments involve some technical problems such as: testing, data collecting, systematization and data pre-processing. System of data collecting can automate routine in the process of data collection. The paper is dedicated to approaches for system for social data collecting development.

**Keywords:** collecting system, database, data analysis, testing system, head count.

**Введение.** В рамках проекта РГНФ №15-23-03001 «Общественное сознание современной молодежи Монголии» авторами была поставлена задача – определить в какой мере у молодёжи Монголии присутствуют знания и уважение к родной культуре, родным традициям, родному языку, уважение и терпимость к людям иной национальности и культур.

Для решения поставленной задачи необходимо провести социологическое исследование. При проведении подобного исследования одной из главных задач является обеспечение массовости. В ходе первоначального анализа задачи было вычислено необходимое число опрашиваемых, оно составило более 800 человек – с учётом разных социальных слоёв, районов проживания и т. д.

#### **Требования к системе**

Важнейшей задачей при реализации проекта является сбор данных анкетирования. Проведение анкетирования среди такого числа респондентов связано с серьёзными проблемами в организации сбора и первичной обработки данных. В связи, с чем авторами было принято решение – разработать систему сбора данных анкетирования.

В рамках проекта коллективом авторов была разработана анкета. В ходе анализа анкеты было выявлено, что:

- 1) некоторые вопросы подразумевают один ответ, а некоторые несколько вариантов ответа;
- 2) некоторые вопросы подразумевают наличие открытого варианта ответа;
- 3) некоторые вопросы представляют собой набор вопросов. Например:

Вопрос: *Какую оценку Вы бы дали данным государственным институтам Вашей страны?*

Вариант ответа	Положительная	Скорее, положительная	Скорее, отрицательная	Отрицательная
Армия				
Органы безопасности				
милиция, суд, прокуратура				
Парламент				
Правительство				
Президент				
Церковь				
Средства массовой информации				
Политические партии				

Фактически данный вопрос представляет из себя оптимизированный (для пользователей) набор вопросов одного типа:

Вопрос: *Дайте оценку государственному институту вашей страны Название института?*

Варианты ответов:

*Положительная*

*Скорее положительная*

*Отрицательная*

*Скорее отрицательная*

Не менее важным вопросом при разработке системы являлся вопрос обеспечения безопасности и валидности сбора данных. В качестве вариантов были рассмотрены:

1) Открытая система анкетирования, в этом случае тесты доступны в открытом доступе – заполнять их может любой желающий. Явным плюсом данного варианта является полное отсутствие дополнительных этапов для респондентов – нужно просто зайти и пройти тест, однако этому варианту присущи все вышеперечисленные минусы.

2) Стандартная авторизация/регистрация с использованием электронной почты/SMS – уведомлений. Данное решение позволяет практически полностью отсеять разного рода роботов, но также «заставляет» респондентов регистрироваться в системе (что им, по сути, не нужно, так как их задача – один раз пройти тест), то есть создаёт дополнительные сложности.

3) Использование «агентов». Агент – специально проинструктированный человек, которому объяснены все вопросы в анкете, а также он обучен работе с системой. Идея состоит в том, что тестирование также проводится онлайн, но доступ к анкетам осуществляется через уникальные промокоды, генерируемые агентом в системе (рис. 1).

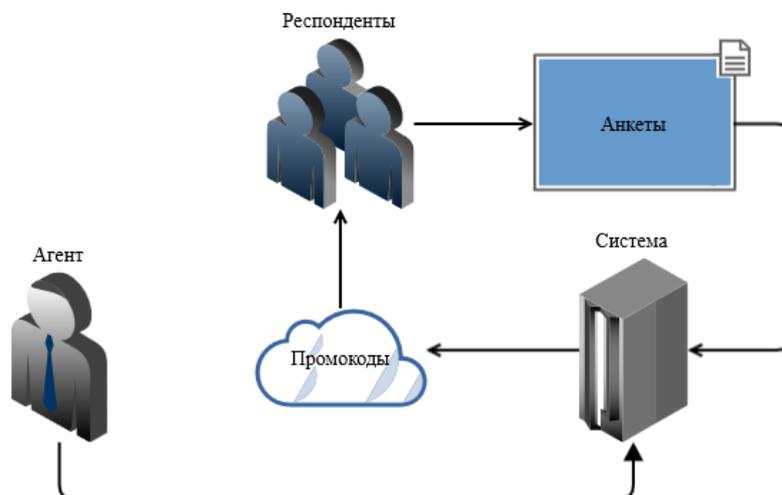


Рисунок 1. Авторизация респондентов через агентов

Примерная схема работы такова:

- 1) Агент запрашивает в своём личном кабинете промокоды;
- 2) Система генерирует запрашиваемое число промокодов и привязывает их к Агенту;
- 3) Агент производит инструктаж респондентов по работе с системой и по вопросам;
- 4) Агент раздаёт каждому респонденту промокод;
- 5) Респонденты авторизуются в системе, активируя промокод, и заполняют анкету.

Такая система авторизации имеет ряд плюсов:

- 1) Минимальная работа с системой авторизации респондентов – нужно только ввести промокод выданный Агентом;
- 2) Личная заинтересованность Агентов вкупе с подробными инструкциями Агентов (как по работе с системой, так и по вопросам анкеты) позволит с большой долей вероятности говорить о валидности заполнения анкет;
- 3) Возможность контроля работы Агентов – результаты анкет привязаны к Агентам через промокоды, в связи с этим существует возможность оценки работы Агентов и исключения ряда результатов в различных ситуациях (например, при соответствующем отчёте Агента);
- 4) Система активируемых промокодов позволяет исключить проблему с роботами, так как даже в случае если активация произведена не респондентом, нежелательные данные можно вычислить (ввиду привязанности кодов к Агентам).

Собранные данные будет необходимо обработать. Можно выделить два уровня обработки:

- «Базовый» уровень подразумевает наличие встроенной подсистемы анализа, для осуществления автоматического/полуавтоматического анализа данных анкетирования (например, по ключам), а также базовый статистический анализ (в том числе средствами СУБД). Для реализации данного вида обработки будет разработана подсистема базового анализа данных, встроенная в систему, которая позволит производить анализ «на лету» – сразу после внесения данных в систему.

- «Продвинутый уровень» – использование стороннего программного обеспечения для реализации затратных (по вычислительным ресурсам и времени) алгоритмов и методов анализа. Для использования подобного программного обеспечения требуется создать подсистему выгрузки данных (с выбором необходимых данных в таблицах). Анализ программных комплексов говорит о том, что данная подсистема должна выгружать в формат CSV, но также должна иметь возможность быстрого добавления выгрузки в любой нужный формат.

Исходя из перечисленных требований, разрабатываемая система должна включать в себя следующие подсистемы [1]:

Подсистема	Назначение, требования к реализации
Авторизация и регистрация	Реализация схемы Агент – респонденты, с упрощённой авторизацией респондентов. Закрепление респондентов за агентами, с возможностью отслеживания работы агентов
Тестирование	Создание, неограниченного числа анкет с разными типами вопросов. Сохранение результатов анкет.
Анализ данных	«Базовый» анализ данных анкетирования.
Выгрузка	Выгрузка данных в различные форматы с возможностью выбора необходимых данных. Масштабируемость – добавление новых форматов выгрузки без модификации базовых классов системы.

Общая архитектура системы представлена на рисунке 2.

### Реализация системы

В соответствии с итерационной моделью разработки программного обеспечения [2] был реализован прототип системы сбора данных социологических опросов. Прототип реализован на языке программирования PHP. При реализации системы использованы шаблоны проектирования Command, Data Mapper, Singleton. Для хранения данных используется база данных, в качестве СУБД была выбрана СУБД MySQL.

На данный момент прототип полностью реализован и размещён в сети Интернет по адресу rfhmong.net.

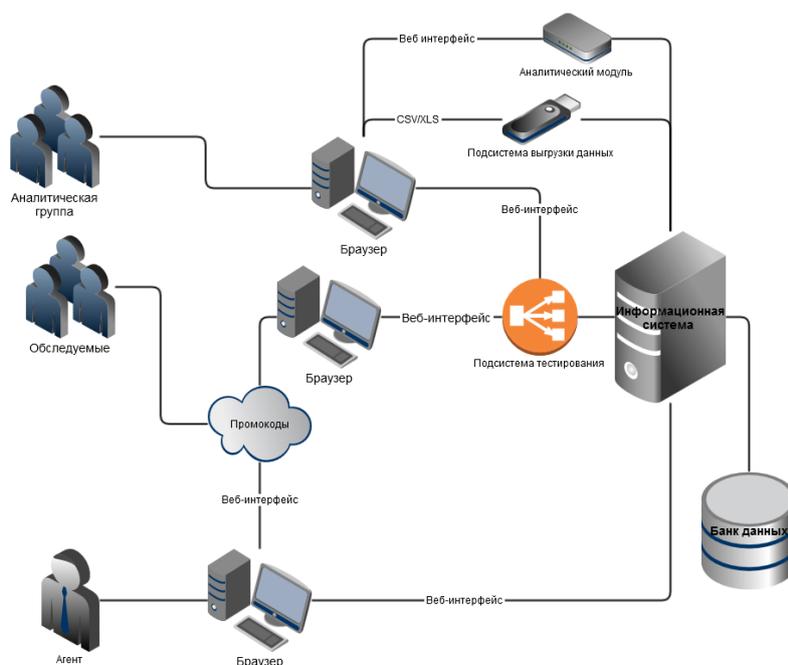


Рисунок 2. Архитектура системы

### Литература

- Хабитуев Б.В. Информационная система сбора социологических данных/ Очирбат Б., Дармаев Т.Г., Хабитуев Б.В.// Актуальные вопросы вещественного и функционального анализа: материалы семинара молодых учёных с международным участием в рамках Международной конференции «Дифференциальные уравнения и математическое моделирование»(г.Улан-Удэ – оз. Байкал, 20-27 июня 2015 г. ) – Улан-Удэ: Изд-во Бурятского госуниверситета, 2015. – 68-73 с.
- Ларман К. Применение UML и шаблонов проектирования, 2-е изд. - М.: Вильямс, 2013. – 736 с.

### References

- Habituev B.V. The information system for social data collectiong/ Ochirbat B., Darmaev T.G., Habituev B.V.//Current issues of real and functional analysis: materials of the seminar of young scientists with international participation within the International Conference “Differential Equations and Mathematical Modeling” (Ulan-Ude – Lake Baikal, June 20-27, 2015) – Ulan-Ude:Buryat State University Publishing Departament, 2015. – 68-73 p.
- Larman C. Applying UML and Patterns: An Introduction to Object-Oriented Analysis and Design and Iterative Development (3<sup>rd</sup> Edition) 3<sup>rd</sup> Edition. – М.: Publishing House Williams, 2013. – 736 p.