

ОБОСНОВАНИЕ ВЫБОРА СУБД ДЛЯ ОПИСАНИЯ МАССИВА ДАННЫХ О СВОЙСТВАХ ПИЩЕВЫХ ПРОДУКТОВ И СЫРЬЯ

О.Н. Мусина, Е.М. Нагорных

Алтайский государственный технический университет им. И.И. Ползунова, Барнаул, Россия

Государство уделяет большое внимание укреплению здоровья населения, в рамках федеральных и национальных проектов формулируются научно обоснованные рекомендации по питанию, разрабатываются программы по популяризации здорового питания.

Современный уровень развития общества требует современного подхода к решению вопросов о разработке продуктов здорового питания и сбалансированных рационов, ориентируясь не только на профессионализм и квалификацию технологов и нутрициологов, но и на возможности информационных технологий, обобщающих багаж накопленных человечеством знаний [1, 2]. Так, благодаря изучению химического состава пищевых продуктов, расширению перечня получаемых данных о содержании в составе пищевых продуктов пищевых веществ, включая минорные биологически активные вещества, в эпоху цифровой нутрициологии стало возможным создание систематизированных баз данных химического состава пищевых продуктов и сырья [3, 4].

На текущий момент самыми известными мировыми базами данных, содержащими информацию о продовольственном сырье и пищевых продуктах, являются: американская база данных Министерства сельского хозяйства США – база данных USDA [5], система баз данных Продовольственной и сельскохозяйственной организации Объединенных Наций FAO [6], база данных EuroFIR AISBL (содержит информацию о составе продуктов 26 европейских стран, Канады, США, Новой Зеландии и Японии [7], канадская база данных питательных веществ [8], английская база данных Quadram Institute [9], российская база данных ФГБУН «ФИЦ питания и биотехнологии» [10].

В рамках выполнения Государственного задания № 075–00316–20–01 нами разработана структура и наполнение базы данных о пищевых продуктах и сырье. В базу вошли сырье и продукты, отсутствующие в традиционно используемых технологами и нутрициологами различных вариантах справочников химического состава пищевых продуктов [11–15]. В процессе создания этой базы данных решали следующие задачи:

- получение собственных данных о химическом составе пищевых продуктов;
- поиск, сбор и компиляция данных из научных изданий с высоким качеством публикаций (рецензируемые научные журналы, диссертации);
- анализ известных систем управления базами данных (СУБД) и выбор СУБД для создаваемой БД.

Проанализировав информацию о зарегистрированных Роспатентах базах данных с помощью поисковой системы Федерального института промышленной собственности [16], установлено, что в этих базах данных собрана информация по определенной группе продуктов, об отдельном регионе, об определенном слое населения, или о группе людей с определенным заболеванием. При этом базы данных преимущественно были разработаны в Microsoft Excel. У такого подхода есть определенные недостатки: Microsoft Excel требует однородности (однотипности) данных в столбцах, что осуществимо только для отдельных групп продуктов и сырья, также Excel при больших объемах данных перестает корректно работать и оперативно выдавать результат. Если файлы будут превышать максимально возможный размер в 2–4 Гб, то они просто перестанут открываться.

Нами при проектировании базы данных использовалась реляционная модель, которая является широко распространенной, подходящей для решения поставленных задач, удобной для реализации на компьютере за счет простоты управления и гибкости структуры.

На рынке программного обеспечения существуют разные инструменты для работы с базами данных. Наиболее популярны – Oracle MySQL, Microsoft SQL Server, Microsoft Access и PostgreSQL.

Oracle MySQL – свободная реляционная система управления базами данных, но в рамках лицензии GNU GPL, по которой автор передает свои наработки в общественную собственность.

Microsoft SQL Server и Microsoft Access – коммерческое программное обеспечение, как, впрочем, и Microsoft Excel.

PostgreSQL – свободная реляционная база данных без ограничений на размер базы данных. Языки программирования, которые поддерживаются данной СУБД, это JavaScript, Go, Java, Net, C++, Python. Однако полнофункциональная интегрированная среда разработки этих языков программирования для написания, отладки, тестирования и развертывание кода на любой платформе – это Microsoft Visual Studio, которая является коммерческим продуктом.

С целью минимизации расходов, в качестве системы управления нашей базой данных была выбрано свободное ПО – бесплатная редакция Oracle Database Express Edition, которая позволяет использовать все свойства реляционной базы данных. Эта редакция не имеет ограничений по функциональности, хорошо масштабируется, при этом дается возможность подключать 3 базы данных, общий размер которых может достигать 12 Гб.

Выбор в качестве СУБД именно Oracle Database Express Edition позволит осуществлять миграцию на основе стандартных пакетов импорта / экспорта данных из других СУБД, а также Microsoft Excel и Microsoft Access, что позволит без лишних усилий перенести информацию, которая хранилась в разных файлах Microsoft Excel и Microsoft Access в единую базу данных, а затем продолжить вводить, хранить, обрабатывать, анализировать данные.

Информацию из Oracle Database Express Edition можно будет получать напрямую, используя язык PL/SQL, или через информационные системы, разработанные с помощью различных языков программирования с поддержкой баз данных Oracle. Мы предполагаем использование среды разработки Oracle Application Express, который реализован как веб-приложение, поддерживаемое большим количеством браузеров на любой платформе.

Логическая модель создаваемой базы данных представлена на рис. 1, а скриншоты ее заполнения – на рис. 2.

Логическая модель базы данных состоит из 5 таблиц.

В связи с тем, что для разных видов пищевых продуктов и сырья существуют разные уровни классификации, в базе данных таблица с условным названием «Сырье» будет представлена древовидной структурой, состоящей из элементов двух типов – узлов и листьев. Рассмотрим данную таблицу подробнее.

Узлами 1 уровня в таблице «Сырье» будут являться следующие элементы: Ингредиенты растительного происхождения; Ингредиенты животного происхождения; Ингредиенты прочие.

Узлами 2 уровня в таблице «Сырье» будут являться следующие элементы: Сырье из мяса и мясопродуктов; Сырье из птиц и яиц; Сырье из продуктов речного и морского промысла; Сырье из продуктов пчеловодства и т. д.

Узлами 3 уровня в таблице «Сырье» будут являться следующие элементы: Мясо и субпродукты; Мясо птицы и продукты переработки птицы; Яичное сырье и т. д.

Листьями в таблице «Сырье» будут являться следующие элементы: мясо овцебыка; мясо верблюда; мясо марала; мясо тюленя и т. д.

Для узла 1 уровня «Ингредиенты прочие» узлами 2 уровня будут витамины, микроэлементы, БАДы, и, пропуская узлы 3 уровня, сразу будут указаны листья.

Поскольку разных показателей существуют разные уровни их классификации, таблица «Показатель» также будет представлена древовидной структурой, состоящей из элементов двух типов узлов и листьев.

Узлами 1 уровня в таблице «Показатель» будут являться следующие элементы: белки; липиды; углеводы; энергетическая ценность и т. д.

Узлами 2 уровня в таблице «Показатель» будут являться следующие элементы: аминокислоты; витамины; насыщенные жирные кислоты; полиненасыщенные жирные кислоты и т. д.

Узлами 3 уровня в таблице «Показатель» будут являться следующие элементы: незаменимые аминокислоты; заменимые аминокислоты; насыщенные жирные кислоты; полиненасыщенные жирные кислоты; минеральные вещества (зольность) и т. д.

Листьями в таблице «Показатель» будут являться следующие элементы: вода; пищевые волокна; крахмал; витамин А кальций; валин; лейцин; лизин и т. д.

Таблица «Единицы измерения» будет заполняться следующими элементами: граммы; проценты; килоджоули; килокалории и т. д.

Таблица «Значение» будет заполняться цифровыми значениями каждого показателя для каждого вида пищевого продукта или сырья, погрешностями и указанием на источник(и) информации для значений.

Таблица «Тип значения» будет заполняться следующими элементами: показатель; погрешность.

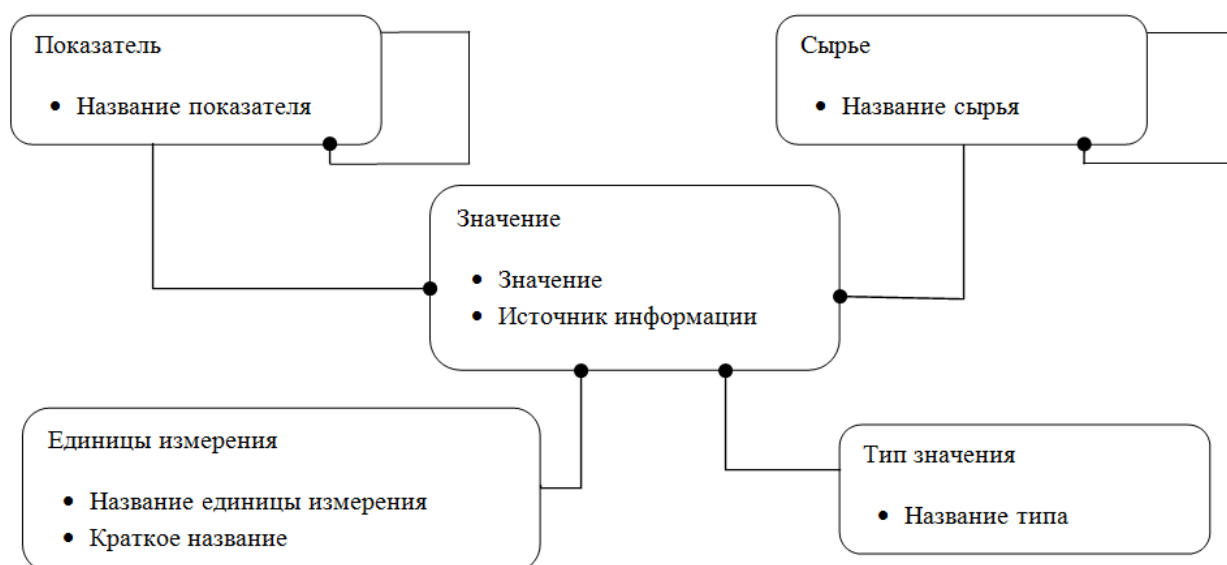


Рисунок 1. Логическая модель базы данных свойств продовольственного сырья

Сырье и продукты растительного происхождения

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Плоды

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Плоды -> Черника обыкновенная

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Плоды -> Черемуха обыкновенная

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Плоды -> Рябина обыкновенная

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Плоды -> Шиповник обыкновенный

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Трава, листья, стебли, побеги

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Трава, листья, стебли, побеги -> Смородина черная (лист)

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Трава, листья, стебли, побеги -> Черника обыкновенная (побеги)

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Цветки и соцветия

Сырье и продукты растительного происхождения -> Лекарственное травяное сырье -> Корни и корневища

Сырье и продукты растительного происхождения -> Плодово-ягодное сырье

Сырье и продукты растительного происхождения -> Орехи

а

-> Общий химический состав
-> Общий химический состав -> Белок
-> Общий химический состав -> Жир
-> Общий химический состав -> НЖК
-> Общий химический состав -> Хол
-> Общий химический состав -> МДС
-> Общий химический состав -> Кр
-> Общий химический состав -> Угл
-> Общий химический состав -> Пищевые волокна (сумма)
-> Общий химический состав -> Пищевые волокна
-> Общий химический состав -> Пищевые волокна -> Клетчатка
-> Общий химический состав -> Пищевые волокна -> Гемичеселлюлозы
-> Общий химический состав -> Пищевые волокна -> Лигнин
-> Общий химический состав -> Пищевые волокна -> Пектин
-> Общий химический состав -> Органические кислоты
-> Общий химический состав -> Зола
-> Общий химический состав -> Энергетическая ценность
-> Общий химический состав -> Вода

б

Рисунок 2. Скриншот заполнения базы данных: а – таблица «Сырье», б – таблица «Показатель»

Данный метод хранения двух основных справочников в древовидной структуре является универсальным и позволит на любом этапе внести новый уровень классификации или несколько новых уровней классификации сырья или показателя, не переделывая уже имеющуюся базу данных.

При разработке базы данных использовались возможности объектно-ориентированной среды Oracle Database Express Edition, позволяющие создавать таблицы, изменять структуры таблиц и редактировать записи. Благодаря такому выбору СУБД, в базе данных возможно без лишних усилий поддержание в актуальном состоянии информации о химическом составе сырья и пищевых продуктов и источнике данных.

Планируется, что эта база данных будет содержать объединенную информацию о химическом составе, аминокислотном составе, включающем в себя незаменимые и заменимые аминокислоты, жирнокислотном составе, углеводном составе, а также минеральном составе, включающем в себя макро- и микроэлементы. Эти показатели в дальнейшем позволят осуществлять расчет пищевой, биологической и энергетической ценности, а также минерального и витаминного состава проектируемых продуктов и рационов. На основе этих данных возможно будет разрабатывать продукцию для отдельных категорий потребителей, составлять индивидуальное меню с учетом особенностей человека или его заболеваний, пересматривать продовольственные корзины с учетом особенностей региона. Применение современных цифровых технологий при проектировании рецептур позволит оперативно произвести выбор сырья и пищевых продуктов с учетом заданных признаков химического состава, выгрузить информацию в Excel, хранить и редактировать данные.

Авторы благодарят за финансовую поддержку Минобрнауки РФ (тема № 075–00316–20–01, FZMMM-2020–0013, мнемокод 0611–2020–013).

Литература

1. Никитина М.А., Сусь Е.Б. Информационная система проектирования пищевых продуктов // Все о мясе. – 2015. – № 1. – С. 36–39
2. Marconi S., Durazzo A., Camilli E., Lisciani S., Gabrielli P., Aguzzi A., Gambelli L., Lucarini M., Marletta L. Food composition databases: considerations about complex food matrices // Foods. – 2018. – Vol. 7. – Issue 1. – P. 2. <https://doi.org/10.3390/foods7010002>
3. Бессонов В.В., Богачук М.Н., Боков Д.О., Макаренко М.А., Малинкин А.Д., Сокурено М.С., Зотов В.А., Шевякова Л.В. Базы данных химического состава пищевых продуктов в эпоху цифровой нутрициологии // Вопросы питания. – 2020. – Т. 89. – № 4. – С. 211–219. <https://doi.org/10.24411/0042-8833-2020-10058>
4. Musina O., Putnik P., Koubaa M., Barba F.J., Greiner R., Roohinejad S., Granato D. Application of modern computer algebra systems in food formulations and development: a case study // Trends in Food Science & Technology. – 2017. – № 64. – С. 48–59. <https://doi.org/10.1016/j.tifs.2017.03.011>
5. База данных The United States Department of Agriculture (USDA). [Электронный ресурс]. URL: <https://fdc.nal.usda.gov/ndb/search> (дата обращения 03.07.2021)
6. Базы данных The Food and Agriculture Organization of the United Nations (International Network of Food Data Systems INFOODS). [Электронный ресурс]. URL: <http://www.fao.org/infoods/infoods/tables-et-bases-de-donnees/russia/en/> (дата обращения 03.07.2021)
7. База данных EuroFIR AISBL. [Электронный ресурс]. URL: <https://www.eurofir.org/> (дата обращения 03.07.2021)
8. База данных The Canadian Nutrient File (CNF). [Электронный ресурс]. URL: <https://www.canada.ca/en/health-canada/services/food-nutrition/healthy-eating/nutrient-data/canadian-nutrient-file-about-us.html> (дата обращения 03.07.2021)
9. База данных Quadram Institute. [Электронный ресурс]. URL: <https://quadram.ac.uk/targets/food-composition/> (дата обращения 03.07.2021)
10. База данных «ФИЦ питания и биотехнологии». [Электронный ресурс]. URL: http://web.ion.ru/food/FD_tree_grid.aspx (дата обращения 03.07.2021)
11. Химический состав российских пищевых продуктов: Справочник / Под ред. член-корр. МАИ, проф. И.М. Скурихина и академика РАМН, проф. В.А. Тутельяна. – М.: ДеЛи принт, 2002. – 236 с.
12. Химический состав пищевых продуктов: Справочные таблицы содержания основных пищевых веществ и энергетической ценности пищевых продуктов. Кн. I: / Под ред. И.М. Скурихина и М.Н. Волгарева. – 2-е изд., перераб. и доп. – М.: Агропромиздат, 1987. – 224 с.
13. Химический состав пищевых продуктов: Справочные таблицы содержания аминокислот, жирных кислот, витаминов, микро- и макроэлементов, органических кислот и углеводов. Кн. II: / Под ред. И.М. Скурихина и М.Н. Волгарева. – 2-е изд., перераб. и доп. – М.: Агропромиздат, 1987. – 360 с.
14. Химический состав пищевых продуктов: Справочные таблицы содержания основных пищевых веществ и энергетической ценности пищевых продуктов / Под ред. А.А. Покровского. – М.: Пищевая промышленность, 1976. – 228 с.
15. Химический состав и энергетическая ценность пищевых продуктов: справочник МакКанса и Уиддоусона / пер. с англ. под ред. Батурина А.К. – СПб.: Профессия, 2006 – 416 с.
16. Федеральный институт промышленной собственности. Программы для ЭВМ, БД и ТИМС. [Электронный ресурс]. URL: <https://www1.fips.ru/iiss/> (дата обращения 03.07.2021)