

Г.Д. Павленок✉, Р.Д. Лукьянов, К.К. Павленок

Институт археологии и этнографии СО РАН
Новосибирск, Россия

E-mail: lukianovagalina@yandex.ru

IT-Fact – новый программный продукт для создания унифицированных баз данных в археологии

Атрибутивный анализ – востребованный и удобный инструмент при обработке массового археологического материала в тех случаях, когда необходимо фиксировать зависимость одного индивидуального признака артефакта от другого, или, что чаще встречается – от определенного комплекса признаков. Для работы с большими базами данных необходимо тщательно следить за унифицированным характером информации с этапа первичного ввода данных. Наиболее эффективное решение этой задачи предполагает использование специализированной компьютерной программы, ограничивающей тип введенной информации (текстовая или цифровая), количество возможных вариантов значений признаков и пр. До недавнего времени с этими задачами справлялись продукты серии Entrer, которые начали разрабатываться еще в середине 1990-х гг., однако устаревшая платформа программы и непредвиденные ошибки, появляющиеся при ее установке на более современные версии оперативных систем, приводят к значительным трудностям работы в программе при настройке или редакции списка признаков и их значений. В связи с этим была начата разработка программного обеспечения, способного работать в актуальных версиях современных операционных систем и учитывающего тенденцию к использованию мобильных устройств в качестве рабочих инструментов. В результате проведенной работы был создан и апробирован на коллекциях каменного века обшей численностью более 2000 экз. программный продукт IT-Fact, предназначенный для эффективного атрибутивного анализа технологически значимых признаков артефактов. При разработке программного продукта в качестве базовых компонентов были использованы язык программирования Python 3 и кроссплатформенная среда для разработки пользовательского интерфейса Qt 5. На настоящий момент IT-Fact представляет собой унифицированный графический интерфейс для ввода, и в перспективе – обработки большого объема данных (заранее определенный исследователем набор качественных и количественных описательных признаков артефактов с перечнем возможных значений).

Ключевые слова: археологическая коллекция, атрибутивный анализ, программное обеспечение, язык программирования.

Galina D. Pavlenok✉, Roman D. Lukianov, Konstantin K. Pavlenok

Institute of Archaeology and Ethnography SB RAS,
Novosibirsk, Russia

E-mail: lukianovagalina@yandex.ru

IT-Fact – the New Software for Specialized Archaeological Databases

Attributive analysis is a popular and convenient tool for processing large amount of archaeological material evidence when it is needed to establish dependence of one individual feature of the artifact on another, or more often, on a certain set of features. In working with large databases, scholars should carefully monitor the unified nature of information since the initial stage of data entry. The most effective solution to this problem is using specialized software which limits the type of information entered (textual or digital), number of possible options for feature values, etc. Until recently, these tasks were solved by the Entrer software series developed in the mid 1990s. However, the outdated software platform and unforeseen errors appearing after its installation on modern operating systems lead to significant difficulties in setting up or editing the list of features and their values. This is why new IT-Fact software for current operating systems was developed with the purpose of conducting effective attributive analysis of technologically significant artifact features and tested on over 2000 lithic objects of the Stone Age. The Python3 programming language and Qt5 cross-platform graphic interface for development of user environment were basic components in the design of the IT-Fact software. Currently, IT-Fact is a unified graphical user interface for entering and, in the future, processing a large amount of data (a set of qualitative and quantitative descriptive features of artifacts predetermined by the researcher with the list of possible values).

Keywords: archaeological collection, attributive analysis, software, programming language.

Современный уровень развития археологической науки требует детальной и тщательной обработки всех предметов из археологических коллекций, в т.ч. индивидуального описания находок из массовых категорий, таких как нестандартизированные отходы производства изделий из каменного сырья, мелкие фрагменты керамических изделий и пр. («атрибутивный подход» к описанию археологических коллекций) [Павленок К.К., Белоусова, Рыбин, 2011]. Для облегчения и ускорения ввода данных в единую базу для последующей их статистической обработки, а также с целью минимизации ошибок при индивидуальном описании археологических предметов, европейскими археологами Ш. МакФерроном и Г.Л. Дибблом был создан программный продукт Entrer, версии 1–4 (An OldStoneAge.Com Production) [McPherron, Holdaway, 1996]. Эта программа представляет собой графический интерфейс для системы управления базами данных Microsoft Access, позволяющий выстроить строгий неизменяемый алгоритм описания конкретной категории артефактов на основе индивидуальных описательных признаков. Данный программный продукт очень быстро приобрел популярность в среде профессионалов-археологов и стал одним из наиболее удобных способов обработки археологических коллекций в рамках атрибутивного подхода.

Однако разнообразие современных технических приспособлений (электронные планшеты, тахеометры, штангенциркули и другие измерительные приборы с возможностью автоматического переноса данных в компьютер), которые используются в современной археологии, нередко приводит к ситуации, когда исследователь стоит перед необходимостью работать с различными (зачастую конфликтующими) форматами данных при изучении материалов одного археологического комплекса. Это обуславливает необходимость разработки специализированного программного обеспечения, способного унифицировать процесс обработки данных, полученных при использовании различного программного обеспечения и технических средств.

Дорабатывать продукты серии Entrer нецелесообразно (даже в условиях открытого программного кода), поскольку они являются оболочкой к единственному типу баз данных (Microsoft Access), которая поддерживается только одной операционной системой (Microsoft Windows). Это накладывает значительные ограничения на использование продукта, а также делает невозможным применение системных функций иных операционных систем. Кроме того, отсутствие поддержки и регулярных обновлений этого продукта привело к ситуации, когда сочетание его с более новыми версиями

операционных систем приводит к ошибкам и значительным трудностям работы в программе при настройке или редакции списка признаков и их значений.

Разработка программного продукта, который по функционалу может заменить продукты серии Entrer, а также соответствует характеристикам и возможностям современных технических средств и программного обеспечения, является актуальной задачей в рамках общего тренда внедрения цифровых технологий в археологию.

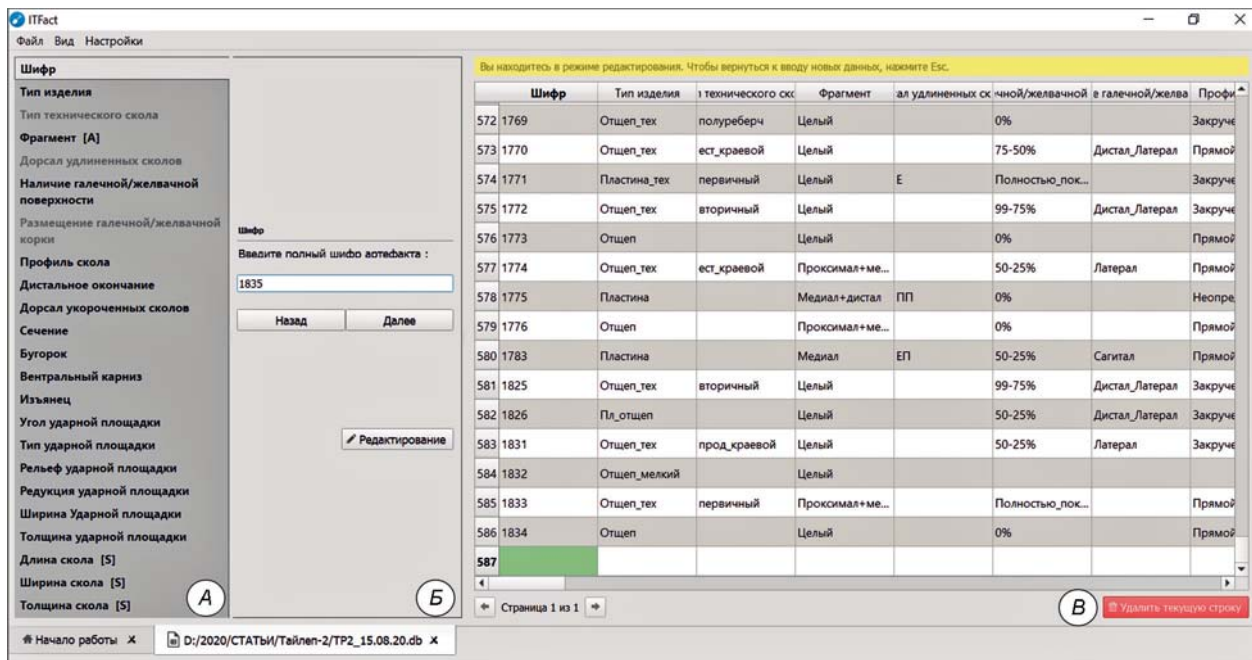
Новый программный продукт, который сейчас разрабатывается в ИАЭТ СО РАН, представляет собой унифицированный графический интерфейс для ввода и обработки большого объема данных (заранее определенный исследователем набор качественных и количественных описательных признаков артефактов с перечнем возможных значений). В качестве базовых компонентов были использованы высокоуровневый язык программирования Python 3 и кроссплатформенная среда для разработки пользовательского интерфейса Qt 5.

Разработка велась в соответствии с предлагаемой в официальной документации фреймворка Qt 5 архитектурой модель/представление. Такой подход позволяет лучше структурировать кодовую базу и сосредоточить код, отвечающий за хранение, получение и отображение данных в различных программных сущностях. Вместо единого слабоструктурированного файла разработчик имеет дело с серией отдельных файлов с кодом, каждый из которых отвечает за выполнение конкретной задачи. Это значительно облегчает дальнейшую разработку и поддержку продукта.

Приоритетными задачами, которые решались при разработке программного продукта, являлись:

- ускорение процесса ввода индивидуальных данных археологического предмета;
- минимизация случайных ошибок при вводе информации;
- обеспечение совместимости с конфигурационными файлами продуктов серии Entrer;
- обеспечение возможности экспорта итоговых данных в сторонние программные продукты;
- обеспечение совместимости с актуальными версиями операционной системы Microsoft Windows;
- обеспечение совместимости с различными системами управления базами данных (поддержка импорта/экспорта информации, в т.ч. полученной со стороннего высокотехнологичного оборудования, в популярные форматы баз данных).

Проблемы удобства ввода данных и возможных ошибок ввода для различных типов признаков решены строгими ограничениями возможных



Интерфейс программы IT-Fact.

А – поле признаков; Б – поле значений признаков; В – поле базы данных.

значений в поле ввода. Также предусмотрено поведение системы при обнаружении проблем, связанных с неконсистентностью данных. Подобные проблемы, например, могут возникнуть при удалении из описания какого-либо признака или значения, что может вызвать несоответствие конфигурации (схемы описания) и итоговой базы данных. В этом случае система после получения одобрения от пользователя обеспечит удаление этого признака и связанных с ним данных из всех элементов программы.

Экспорт данных в текущей версии реализован путем выгрузки итоговых данных в файл с расширением .xlsx. Выбор обусловлен тем, что формат таблицы Microsoft Excel не требует никакой дополнительной обработки данных и поддерживается всеми современными операционными системами. Также была реализована возможность импорта файлов конфигурации (файлы .cfg) и баз данных (файлы .mdb), созданных в Entrer-4, что позволяет в полном объеме использовать ранее зафиксированную информацию и наработки.

Используемый фреймворк Qt 5 благодаря тесной интеграции с операционной системой, на которой запущено приложение, позволяет, помимо прочего, обеспечить поддержку специфических устройств ввода. Это дает возможность, например, использовать для быстрого ввода метрических данных описываемого предмета цифровые штангенциркули с поддержкой подключения к устройству через порт USB.

Кроссплатформенность программы достигается за счет возможностей языка Python 3, имеющего свободно распространяемые реализации интерпретатора под разные операционные системы. Для операционной системы Microsoft Windows в текущей версии реализована система сборки программы в наиболее распространенный исполняемый файл .exe, а также удобный графический инсталлятор. Для этого используются также свободно распространяемые продукты NSIS и PyInstaller. В перспективе планируется переход к использованию продукта Pyqtdeploy, позволяющего унифицировать и автоматизировать процесс сборки приложения для всех заявленных платформ (Microsoft Windows, Android, UNIX-подобные системы и др.).

В результате был получен программный продукт IT-Fact, к достоинствам которого (в сравнении с продуктами серии Entrer) можно отнести следующие характеристики.

Простота настройки списка признаков и их значений при подготовке системы описания (конфигурации). Конфигурация позволяет создать неограниченное количество описательных признаков (см. рисунок, А), задать значения каждого признака в формате заранее определенного списка (для унифицированной информации), либо произвольного текста или числового значения для уникальных значений признаков, таких, например, как метрические данные, примечания и пр. (см. рисунок, Б). При этом IT-Fact позволяет легко редактировать значения конкретного признака, наведя курсор на

нужную ячейку в уже сформированной базе данных (см. *рисунок, В*). За счет того, что в новом программном продукте файл конфигурации и собственно базы данных является единым целым, отсутствует необходимость каждый раз перезагружать программу при редактировании списка признаков, чтобы они сочетались с аналогичным списком базы данных. Подобные проблемы постоянно возникали в *Entrer-4*, где конфигурация и база данных были разбиты на два файла – *.cfg* и *.mdb* и редакция одного из них автоматически вызывала ошибку чтения второго. В *IT-Fact* синхронизация списка признаков и базы данных происходит автоматически.

Еще одним важным достижением стала возможность добавления нескольких условий зависимости одного признака от других, тогда как в *Entrer-4* действительно рабочим было лишь одно условие на один признак. Эту функцию можно проследить на примере описания каменных орудий: например, в *IT-Fact* возможно признак «угол наклона фасеток ретуши второго участка» поставить в зависимость от двух условий – от признаков «наличие вторичной обработки» и «количество участков вторичной обработки». Поскольку условия-признаки иерархичны, при отсутствии положительного значения первого признака («наличие вторичной обработки») либо конкретного значения второго признака («количество элементов вторичной обработки» – 1), третий признак («угол наклона фасеток ретуши второго участка») не будет запрашиваться программой, что ускорит процесс обработки коллекции.

При подготовке программы *IT-Fact* специально была разработана функция, позволяющая импортировать конфигурации (файл *.cfg*) и базы данных (файл *.mdb*), созданные в *Entrer-4*, что позволяет в полном объеме использовать ранее зафиксированную информацию и наработки.

Выбор языка и базового фреймворка, а также современный подход к проектированию архитектуры приложения *IT-Fact* позволяют утверждать, что данное программное обеспечение имеет также и перспективы улучшения. Так, в дальнейшем возможно обеспечение работы как с локальными, так и удаленными базами данных (в т.ч. несколькими пользователями в синхронном режиме), а также полная поддержка других современных операционных систем: UNIX-подобные системы,

Android и др. («кроссплатформенность»). Кроме того, в рамках разработки модуля для удаленной работы с базами данных планируется реализация внешнего программного интерфейса (API), который предоставит сторонним разработчикам возможность расширения базового функционала программы посредством добавления собственных программных модулей. В отдаленной перспективе возможно добавление функции голосового ввода значений признаков, что позволит существенно увеличить скорость обработки коллекций. Расширение возможностей программного продукта *IT-Fact* за счет добавления этих функций позволит создать гибкий и универсальный инструмент обработки коллекций, превосходящий по эффективности существующие в настоящий момент аналоги.

Благодарности

Работа выполнена при поддержке гранта РФФИ (проект № 18-39-20003).

Список литературы

Павленок К.К., Белоусова Н.Е., Рыбин Е.П. Атрибутивный подход к реконструкции «операционных цепочек» расщепления камня // *Вестн. Новосиб. гос. ун-та. Сер.: История, филология.* – 2011. – Т. 10, вып. 3. – С. 35–46.

McPherron S., Holdaway S.J. *Entrer Trois* // *A multimedia companion to the Middle Paleolithic site of Combe-Capelle Bas (France).* – Philadelphia: The Univ. Museum, 1996. – CD-ROM.

References

McPherron S., Holdaway S.J. *Entrer Trois*. In *A multimedia companion to the Middle Paleolithic site of Combe-Capelle Bas (France)*. Philadelphia: The Univ. Museum, 1996. CD-ROM.

Pavlenok K.K., Belousova N.E., Rybin E.P. Atributivnyi podkhod k rekonstruktsii “operatsionnykh tsepochek” rasschepleniia kamnia. *Novosibirsk State University Bulletin. Series: History and Philology*, 2011, vol. 10, iss. 3, pp. 35–46. (In Russ.).

Павленок Г.Д. <https://orcid.org/0000-0003-3727-776X>

Лукьянов Р.Д. <https://orcid.org/0000-0003-0840-8054>

Павленок К.К. <https://orcid.org/0000-0003-0205-2077>