
ПРОГРАМНА СИСТЕМА „АНКЕТИ“

МАРГАРИТА В. ВАСИЛЕВА, АНДРЕЙ И. АНДРЕЕВ, КОРНЕЛИЯ Т. ЦОНЕВА

SOFTWARE SYSTEM "INQUIRY"

MARGARITA V. VASILEVA, ANDREY I. ANDREEV, CORNELIA T. TSONEVA

ABSTRACT: *This paper examines the purpose and functioning of a software system for recognition and processing of inquiries, developed with the participation of the authors. The system belongs to a class of software products for recognition and interpretation of manually entered information into known areas of formalized documents (Document Capturing). It is designed for automatic data capturing of six different in their structure and purpose types of questionnaires for the needs of the University of Shumen. The captured information is being processed to obtain reports on the results of the inquiry.*

KEYWORDS: *Document Capturing, processing inquiries, forms processing, handwritten character recognition, handwritten character capturing handwritten text recognition,, image recognition*

I. ПРЕДНАЗНАЧЕНИЕ

Софтуерната система “АНКЕТИ” е система за разпознаване и обработка на анкетни карти за контрол на качеството на обучение в Шуменския университет.

Тя принадлежи към класа програмни продукти за разпознаване на формуляри [3] (на Запад се използва терминът Document Capturing — "прихващане на документ"), предназначени за разпознаване на определен формат на документа, съдържащ печатна и ръчно въведена информация. Задачата за разпознаване се състои в намиране и интерпретация на информацията в някои предварително известни области. Разпознатата информация се обработва с цел получаване на справки за резултатите от анкетирането по различни критерии.

Основно предимство на програмната система е замяна на неефективния ръчен труд с автоматизирана обработка на анкетите, значително съкращаване на времето за обработка, намаляване на възможността за допускане на грешки, повишаване на точността и достоверността на обработваните данни.

II. ТЕХНОЛОГИЯ НА РАБОТА

Технологията на работа включва следните действия:

1. Разработка и тестване на анкетната форма

Комисията по качество разработва и утвърждава анкетната форма, която се тества преди използване. В момента са разработени шест различни вида анкетни карти, отговарящи на определени изисквания [1]:

- Анкетна карта за центъра за кариерно развитие;
- Анкетна карта за качеството на обучението по дисциплина;
- Анкетна карта за мнението на преподавателя за учебния процес;
- Анкетна карта за работодатели;
- Анкетна карта за новоприети студенти;
- Анкетна карта за завършващи студенти.

Тестването дава възможност да се изключат проблемните области на дизайна на формата и да се елиминират възможни грешки, например:

- много близко разположени полета за отговори на анкетираните (квадратчета), което дава възможност за маркиране и на съседното квадратче;
- тънки контурни линии на квадратчетата, които се накъсват след сканиране и не могат да се разпознаят;
- несанкционирано променяне на местоположението на квадратчетата в копията на анкетните карти, заради веригата от състояния – текстов файл (Word) → принтер (хартия, формат А4) → ксерокс (хартия А4) → скенер (графични файлове) и др.

След одобряване на анкетната форма, тя се отпечатва на лазерен принтер и се размножава на ксерокс.

2. Анкетиране на целевата група

При анкетирането на целевата група трябва:

- да се извърши предварителен инструктаж на анкетираните, като им се обърне внимание върху важността на коректното маркиране на избрания отговор за следващото разпознаване;
- кодовете на специалностите да се въвеждат съгласно инструкцията;
- отговорите да се маркират с химикал, а не с флумастер, маркер или мастило, които оставят следа на обратната страна на листа и водят до грешно разпознаване.

За програмата няма значение дали анкетирания отбелязва отговорите в квадратчетата с „X“, „V“ или с друг символ. Важното е в квадратчето да има достатъчно на брой точки от химикала или молива. Използваният символ за маркиране на верния отговор не трябва да излиза извън нужното квадратче и да навлиза в съседно квадратче, което анкетирания не е предвиждал да маркира. Понякога анкетирания (особено ако използва „X“), задрасква „с размах“ верния отговор и част от маркировката попада в съседно квадратче. Ако анкетата се обработва ръчно или с компютърна програма чрез въвеждане на данните от клавиатурата, за оператора е ясно кое квадратче е маркирано, но при автоматизираното въвеждане може неправилно да бъдат разпознати като маркирани две съседни квадратчета. Такава ситуация не може да бъде обработена програмно поради факта, че част от въпросите допускат повече от един отговори.

Верните отговори се маркират задължително в квадратчетата, разположени в дясната страна на страницата, а не чрез ограждане на поредния въпрос в ляво:

Грешно

Вярно

o

Доволен/а ли сте, че сте завършили специалност в ШУ? (един отговор)

- а) да, напълно
- б) да, отчасти
- в) не
- г) не мога да преценя

X

3. Сканиране

Сканирането се извършва на скенер Epson във формат bmp с резолюция 200 dpi.

За предпочитане е използването на черно-бял (black&white) режим, а не цветен (color) или автоматичен (auto). При цветен режим графичния файл е 20-25 пъти (!) по-

голям, а от гледна точка на програмата не съдържа абсолютно никаква информация в повече. В този режим опцията Image Adjust Threshold трябва да има стойност 120-160.

След избор на анкета, с която ще се работи (този избор е задължителен и както бе казано по-горе, е първото действие след стартиране на програмата), се избира командата «скенер» от главното меню.

4. Разпознаване

В резултат на разпознаването [2] програмата създава файл с разпознатите анкети. Поради некоректно маркиране на отговорите на въпросите от анкетата е възможно част от анкетите да не се разпознаят правилно. Например, анкетираният е избрал грешен отговор, който е зачеркнал и след това е маркирал верния (според него) отговор. В такива ситуации е невъзможно системата да разпознае еднозначно отговорите на анкетирания, но това може да направи операторът. Проблемът е разрешим чрез използване на модула за ръчно въвеждане, включен с системата.

5. Обработка

В резултат на обработката се пресмятат абсолютните и относителни честоти на отговорите на въпросите на анкетата.

Системата извежда резултатите във вид на справки, определени от предварително зададените критерии – за университета, по факултети, по специалности, по професионални направления.

6. Експорт в други системи

За улеснение на потребителите справките могат да се експортират в Word и в Excel.

III. РАБОТА С ПРОГРАМАТА

1. Общи изисквания

Съгласно заданието, програмата трябва да обработва шест различни по тематика анкети. Този брой може да е един, шест, десет и пр., стига отделните анкети да са в следната рамка :

- Всяка анкета е съставена от произволен брой въпроси;
- Всеки въпрос от дадена анкета има определен брой (повече от един) възможни отговори.
- Анкетирания може да посочи един или няколко възможни отговори (максималния допустим брой отговори е посочен в конкретния въпрос).

След като се стартира програмата и на екрана се появи основната форма, първото, което трябва да направи потребителя е да посочи анкетата, с която ще работи в момента.

2. Режими на работа

Работата с програмата е структурирана в три режима :

1. Четене и графичен анализ на файлове, които изобразяват страниците на анкетните карти на хартиен носител (A4), т.е. файловете, получени от сканиране. Получаване на работен текстов файл.
2. Редактиране (евентуално) на текстовия работен файл и на друга информация.
3. Обработка на работния файл за получаване на резултати. Извеждане на резултатите – на екран, принтер, Word/Excel.

4. Инструкция за работа в различните режими

Следват инструкции за ползване на програмата в посочените режими.

- **Четене на сканираните файлове.**

Предполага се, че предварително анкетните карти са обработвани на скенер и са получени графичните файлове – образи на страниците от анкетните карти на хартия. За правилната работа на програмата е необходимо да се посочи името на директорията (папката) в която са разположени тези графични файлове. Посочването става в директен режим. Диалогът е същия, както избирането на папка в известния File Explorer, който е инсталиран на всеки компютър, работещ в операционна среда Windows. Избирането става с еднократно натискане на левия бутон на мишката (Mouse Left Button – МЛБ).

Когато посочите нужната папка, в дясната част на появилата се форма се извежда списъка със сканираните файлове в тази директория (това е само за информация, че сте избрали не някоя директория с .exe, .txt, .doc и прочие по тип файлове). Обработката започва след като посочите бутона «Избор» с МЛБ.

Това е най-трудоемката операция в работата на програмата. Съответно тя е и най-продължителната откъм машинно време. За всяка анкетна карта трябва да се обработят няколко графични файла – колкото са страниците в картите на този вид анкета.

Времето за работа зависи от броя на анкетните карти. При апробация на софтуерната система бяха обработени анкетни карти на завършващите студенти от две проведени анкети – през лятото и през есента. При първото анкетиране участваха 133 абсолвенти, при второто – 113. Времето за обработка бе около 1.5 – 2 минути, т.е. около три анкетни карти за 2 секунди. За сведение, времето за получаване на таблиците с резултатите – преброяване и проценти на няколко таблици (примерно общо за университета и по факултети е по-малко от 1 секунда !). Времето за отпечатване на резултатите не зависи от програмата, а от принтера.

Времето за експортиране на резултатите към Word или Excel е няколко секунди. Вероятно по-голямата част от това време е не самия експорт, а установяване на връзка с Word (примерно) и стартиране на продукта.

- ***Работа с данните от клавиатурата.***

По горе посочихме за какво е нужен такъв режим, въпреки модула за обработка на сканираните страници от анкетни карти. Редактиране на работен файл се извършва след като се отвори за работа точка от менюто Файлове/Зареждане. На екрана се появява съдържанието на конкретен запис от файла. Навигацията между записите е интуитивно ясна. Става чрез бутоните за първи, последен, предишен и следващ запис след текущия. Има също бутон за избиране на запис от файла по номер. В дясно има още два бутона за реално променяне на информацията – запис на промените и изтриване на текущия запис. Редактирането е съвсем леко и бързо. Единственото писане от клавиатурата е само при редактиране кода на специалността. Останалите промени касаят промяна на отговорите на въпросите. За всеки отговор има само два варианта „да“ или „не“, т.е. маркирано или немаркирано. Това се прави с натискане на ляв бутон на мишката. При еднократно натискане на левия бутон на мишката в квадратчето за отговор, ако отговора не е маркиран – се маркира и обратно – ако е бил маркиран, се размаркира.

В менюто „Файлове“ има още една манипулация за работните файлове – конкатениране, или сливане на файлове от един вид анкета. Според Възложителя, може да се наложи данните от две или повече проведени анкети от един тип да се обработват обединено. Примерно, анкети провеждани през м. януари, юли и ноември да се разгледат и общо, като изследване на една и съща по тип анкета. Избирането на файлове за сливане става отново в диалогов режим. Действията съвпадат с избиране на файлове в File Explorer – избирането на първия файл – с МЛБ върху името му, избирането на следващите файлове

– с комбинацията CTRL+МЛБ, ако някой файл вече е маркиран - с МЛБ върху името му той алтернативно се размаркира.

Накрая, в този режим можете да се заяви създаването на нов файл от тип – файл за тази анкета. Точка от менюто «Файлове», подточка «Нов». Това ще се налага рядко, примерно скенера не работи, а се налага да се направи обработката.

- Обработка на работния файл за получаване на резултати

Служи за статистическа обработка на данните от разпознатите анкети за получаване на резултати. Обработката се стартира след избор на командата „Статистики“ от главното меню.

Програмата калкулира брой и проценти за отговорите на отделните въпроси. След като разполагаме с информацията (евентуално с натрупване) възложителят може да постави като задача всякакви други статистически изчисления, диаграми и пр.

В момента броя и процентите мога да бъдат получени общо за всички записи, по направления, по факултети, или в рамките на посочен факултет – по специалности. Допълнително са запрограмирани два филтъра – по образователно – квалификационна степен (бакалаври, магистри) и по форма на обучение (редовно, задочно). Например, може да се заяви обработка за целия университет, включваща само бакалаври или само магистри. По същият начин само за анкетираните от редовно обучение, или от задочно обучение. Когато се извърши дадено изчисление (за по-малко от секунда) на екрана се появяват резултатите в табличен вид. Тези резултати могат да се отпечатат на принтер или да се експортират към продуктите на Microsoft Office (Word, Excel). Това става чрез избиране на съответните бутони (с МЛБ). Бутоните не са надписани, но са интуитивно ясни. В тях са поставени малки икони, които посочват това, което ще се направи с табличните резултати.

БЛАГОДАРНОСТИ

Статията е финансирана по научноизследователски проект № РД-08-85/04.02.2016 «Мониторинг на дейностите за оценка и управление на качеството на образованието в ШУ в контекста на новите образователни политики и динамиката на пазара на образователни услуги“ на Шуменския университет "Епископ Константин Преславски" с ръководител доц. д-р Стефка Янчева Калева.

ЛИТЕРАТУРА

1. Лечева Д., Харизанов К., Йорданова П. Уеб-базиран софтуер за обработка на анкети. Научни трудове на Русенския университет - 2013, том 52, серия 6.1.
2. Местецкий Л. Математические методы распознавания образов Курс лекций МГУ, ВМиК, кафедра «Математические методы прогнозирования» ©, <http://www.ccas.ru/frc/papers/mestetskii04course.pdf>
3. Обработка анкет и бланков. <http://idr.in.ua/info/obrabotka-anket.html>