

Dariusz MAZUR

Politechnika Śląska, Wydział Organizacji i Zarządzania

## AUTOMATYCZNA AKWIZYCJA ANKIET ZE STRONY WWW DO SYSTEMU INFORMATYCZNEGO

**Streszczenie.** Niniejszy artykuł powstał w wyniku projektu, mającego na celu zaprojektowanie takiego systemu dla działu marketingu, dzięki któremu przedsiębiorstwo umieszczając odpowiednią ankietę na stronie WWW, umożliwiłoby potencjalnemu klientowi złożenie zapytania, które automatycznie wprowadzane byłoby do bazy danych SI przedsiębiorstwa.

## AUTOMATICALLY COLECTING INFORMATION FROM WEB SITES TO MANAGEMENT INFORMATION SYSTEM

**Summary.** The article summarizes the project which purpose was to create IT system for marketing department of an enterprise. This system can enable an enterprise to include a special form on their Web site, which users can use to submit a request and automatically insert it into the enterprise's database.

### 1. Wstęp

Baza danych osobowych jest podstawą działania działów marketingu i reklamy. Jest ona wykorzystywana do prowadzenia reklamy personalnej, badań zachowań klientów, wysyłania personalizowanych materiałów reklamowych i innych działań z zakresu Public Relation. Tego rodzaju bazy danych są podstawą działania systemów klasy CRM (*ang. Customer Relationship Management* - zarządzanie kontaktami z klientami). Baza klientów wraz z ich referencjami stanowi kluczowe narzędzie prowadzenia efektywnego i personalizowanego systemu obsługi klienta, dzięki czemu zwiększa się zadowolenie i lojalność klientów, co

sprzyja wzrostowi sprzedaży. Obniża się również jednostkowe koszty marketingu dzięki kierowaniu działań do ściśle określonych grup.

## **2. Sposoby gromadzenia danych o klientach**

Ważnym aspektem działania bazy klientów jest gromadzenie nowych „namiarów” na potencjalnych klientów. Chodzi o zdobywanie adresów (konwencjonalnych bądź internetowych) w celu nawiązania bezpośredniego kontaktu i przedstawienia szczegółowej oferty, oraz zdobywanie szczegółowych informacji o klientach (w celu personalizacji ofert i wykonania badań statystycznych). Pochodzenie danych o klientach jest bardzo różne. Namiary mogą być pozyskiwane poprzez ręczne wprowadzanie informacji pochodzących z bezpośrednich kontaktów czy ankiet nadsyłanych przez nabywców wybranych produktów. W dobie Internetu jawi się nowa forma pozyskiwania takich namiarów. Mianowicie, na witrynie umieszczona zostaje ankieta w postaci formularza, której wypełnienie może być związane z zamówieniem materiałów informacyjnych, broszur czy prenumeratą wiadomości i nowinek związanych z daną witryną. Treść samej ankiety zależy od przeznaczenia i wyznaczają ją specjaliści od marketingu czy Public Relation. Wspólnym mianownikiem wszystkich ankiet są zbierane w nich dane personalne, pozwalające na kontakt z daną osobą. Pozostałe informacje to dodatkowe dane pozwalające na selekcjonowanie klientów według żądanych kryteriów. Mechanizmy i narzędzia służące tworzeniu takich formularzy na stronach WWW są powszechnie znane. Ważna natomiast jest potencjalna łatwość, z jaką można przekazywać informacje wprowadzone do ankiety przez użytkownika do systemu informatycznego przedsiębiorstwa.

## **3. Model systemu przekazywania danych o klientach z Internetu do systemu informatycznego przedsiębiorstwa**

Aby uzyskać zamierzony efekt, tzn. możliwość automatycznego przekazywania danych o klientach z ankiet do systemu informatycznego przedsiębiorstwa, zaprojektowano następujący model:

- odpowiednią ankietę w postaci formularza umieszcza się w ramach witryny WWW przedsiębiorstwa,
- ankieta po wypełnieniu przez użytkownika generuje plik tekstowy wysyłany jako e-mail pod wskazany adres,



- plik tekstowy jest specjalnie sformatowany: kilka linii nagłówka, następnie poszczególne pola ankiety w formacie:

NAZWA POLA   SEPARATOR   ZAWARTOŚĆ POLA,

- e-mail odbierany jest przez klienta pocztowego, jakim jest Microsoft Outlook, i zgodnie z konfiguracją przekazywany do wskazanego folderu,
- uruchomiona na tym samym komputerze aplikacja wchodząca w skład systemu informatycznego przedsiębiorstwa przegląda oznaczony folder w Microsoft Outlook i każdą nową wiadomość analizuje, odczytane dane wprowadza do bazy danych osobowych, a nadesłaną wiadomość oznacza lub kasuje zgodnie z algorytmem postępowania.

### 3.1. Omówienie wad i zalet w stosunku do innych rozwiązań

Wśród systemów tradycyjnych można wyróżnić wiele rozwiązań znacząco się od siebie różniących: począwszy od prostych skryptów aktualizujących bazę danych po pełną aplikację internetową z wykorzystaniem przeglądarki. Inny niż prezentowany model, który może w tym celu zostać zastosowany, wykorzystuje specjalne narzędzia pozwalające na gromadzenie informacji w bazie danych umieszczonych bezpośrednio na serwerach generujących stronę www z ankietą (on-line). Rozwiązanie to jednak cechuje się kilkoma istotnymi wadami:

- a) baza danych osobowych bezpośrednio umieszczona na serwerach internetowych jest bardziej narażona na niebezpieczeństwo włamania i kradzieży [6]. Pociąga to za sobą konieczność stosowania odpowiednich zabezpieczeń. W proponowanym modelu możliwe jest tylko wysyłanie plików tekstowych, które dopiero po weryfikacji wprowadzane są do bazy danych. Uszkodzenie serwera internetowego czy pocztowego pozostaje bez wpływu na zgromadzone już w bazie dane;
- b) wymagane są dodatkowe narzędzia do realizacji takiej bazy, co pociąga za sobą wzrost kosztów. W wielu wypadkach, z uwagi na różne systemy operacyjne serwerów internetowych, nie jest możliwa implementacja bez ich gruntownej przebudowy. Nie wymagana jest również umiejętność projektowania aplikacji internetowych.

Opisywany wcześniej model systemu przekazywania danych o klientach cechuje się następującymi zaletami:

- a) system informatyczny nie musi być stale połączony z systemem internetowym, dane mogą być pobierane cyklicznie,
- b) dane przed wprowadzenie do bazy danych mogą podlegać szczegółowej analizie, dokonuje się na tym etapie odpowiedniego kodowania i konwertowania informacji zawartej w pliku tekstowym na potrzeby bazy danych,

c) dane mogą pochodzić z wielu różnych ankiet, które aplikacja po właściwej interpretacji wprowadza do wspólnej bazy danych osobowych.

Wymienione cechy w punktach b) i c) można również zaimplementować w systemach online, pociąga to za sobą jednak znaczący wzrost nakładów na tworzenie aplikacji internetowych.

### **3.2. Etapy tworzenia ankiety automatycznie wczytywanej do SI**

Implementacja przedstawianego rozwiązania następuje według przedstawionego schematu:

1. Zaprojektowanie pól do wypełniania, zakodowanie ich w postaci zrozumiałej dla przeglądarki (czyli programów wykorzystywanych do oglądania stron WWW), wprowadzenie mechanizmów wysyłania zawartości wypełnionych pól w postaci e-mail pod wskazany adres.
2. Umieszczenie ankiety na serwerze internetowym, właściwa konfiguracja serwera tak, aby użytkownicy mogli wypełniać ankietę, i po potwierdzeniu zakończenia wypełniania, jej treść w ustalonej formie została wysłana. Przykładowy e-mail wysłany po wypełnieniu ankiety pokazany jest na rysunku 1.
3. Konfiguracja programu Microsoft Outlook tak, aby odbierał wiadomości pochodzące z ankiety i umieszczał je w odpowiednim folderze.
4. Instalacja modułu analizy, który sprawdza, czy nie ma nowych wiadomości z ankietami, a w razie pojawienia się ich dokonuje analizy.
5. Wprowadzenie danych do bazy SI oraz oznaczenie wiadomości jako przetworzonej.

## **4. Konstrukcja modułu analizy ankiet**

Moduł do analizy ankiet składa się z trzech części:

1. Część odpowiedzialna za pobieranie informacji z klienta pocztowego,
2. Część dokonująca interpretacji i kodowania tekstu zawartego w przesyłkach,
3. Część dokonująca zapisów w bazie danych.



```
tytul_formularza  = Ankieta
Imie              = Daniel
Nazwisko          = Plewa
Ulica             = Łukasiewicza 17
Kod               = 33-100
Miejscowosc       = Tarnów
Wojewodztwo       = małopolskie
Telefon           = 014 6220811
E-mail            = dplewa@student.agh.uci.edu.pl
WWW               = http://
Wiek              = 18 - 24 lat
Wykształcenie     = średnie
Zawod             = student
Uprawiane sporty  = turystyka piesza nizinna
Uprawiane sporty  = turystyka piesza górska
Uprawiane sporty  = wspinaczka skałkowa
Alvikę znam z     = reklam w prasie
Alvikę znam z     = wystawy sklepowej
Planuję zakup     = kurtki zewnętrznej
Guzik wysłania    = Wyślij formularz
```

-----  
Form processed at Thu Jan 27 14:32:51 CET 2000

Rys. 1. Przykładowy e-mail z danymi z ankiety

Fig. 1. An example of e-mail

#### 4.1. Pobieranie informacji z klienta pocztowego

Przy tworzeniu modułu wczytywania ankiet zdecydowano się na klienta pocztowego zawartego w programie Microsoft Outlook. Takiego wyboru dokonano, gdyż program ten spełniał następujące założenia: należy zastosować **typowy program pocztowy**, z którego będzie można **automatycznie pobierać** i przetwarzać wybrane wiadomości (korzystając z interfejsu COM).

Interfejs COM został stworzony przez Microsoft w celu umożliwienia bezpośredniej współpracy różnych aplikacji, w tym pochodzących od różnych producentów. Za pośrednictwem tego interfejsu można bezpośrednio ingerować w sposób działania programu Outlook, można pobierać lub zmieniać dane w nim zawarte [8].

#### 4.2. Interpretacja zawartości przesyłki

Druga część modułu analizy ankiet jest odpowiedzialna za interpretację przesłanego tekstu. Można tu wyróżnić kilka etapów:

- weryfikacja, czy wiadomość zawiera informacje pochodzące z ankiety,

- podział tekstu na fragmenty odpowiedzialne za poszczególne pola i identyfikacja tych pól,
- weryfikacja pól zawierających dane personalne i teleadresowe,
- interpretacja i kodowanie pozostałych pól.

Wstępna weryfikacja ma za zadanie odrzucić wszystkie bezużyteczne (z punktu widzenia marketingowego) przesyłki: dotyczy to przesyłek niepełnych lub o nieprawidłowej strukturze. Weryfikacji podlegają adres, kod pocztowy, adres e-mail. Ponieważ nie ma ograniczeń w dostępie do strony z ankietą, to należy się liczyć z możliwością otrzymywania przesyłek-śmieci, które niepotrzebnie byłyby przetwarzane.

#### 4.3. Wprowadzanie informacji do bazy danych

Ostatni fragment nastęrcza najwięcej problemów. Jeżeli bowiem w założeniach projektowych przeznaczenie baza danych osobowych nie jest ściśle określone i ma być możliwość stosowania stosunkowo szerokiego spektrum pytań w ankietach, a dodatkowo ankiety te mogą być modyfikowane, to rodzi się problem takiego zaprojektowania struktury bazy danych, aby istniała możliwość automatycznego wprowadzania informacji, a następnie ich późniejszej analizy. W tym celu wprowadzono następujące ograniczenie: na wszystkie pytania odpowiada się poprzez zaznaczenie (jednej, żadnej lub kilku) odpowiedzi. Przy takim założeniu wystarczy, aby program generujący plik tekstowy zawarł w nim kody pól związanych z wybranymi odpowiedziami i na tej podstawie aplikacja analizująca po odpowiednim przetworzeniu będzie mogła wprowadzić informację do bazy danych. Na podstawie słownika możliwych odpowiedzi stosuje się odpowiednie kodowanie tej informacji i zapis taki wprowadzany jest do bazy danych.

Zanim to jednak się stanie, następuje jeszcze jedna weryfikacja: czy dane osoby nie zostały już kiedyś umieszczone w bazie danych. Dzięki takiej weryfikacji otrzymuje się dodatkowe funkcje spełniane przez program. Odrzucane są dublowane wysyłki tej samej ankiety przez użytkownika, co się zdarza np. przy nieumiejętnym wprowadzaniu danych bądź próbie zakłócenia działania poprzez wielokrotne wysyłanie tej samej wiadomości. Ciekawa właściwość to możliwość automatycznej aktualizacji danych o kliencie uprzednio wprowadzonym, w przypadku gdy przesłane dane są wystarczające do wiarygodnej jego identyfikacji.

### 5. Wnioski

Przedstawiony artykuł powstał na bazie doświadczeń z realizacji zlecenia dla przedsiębiorstwa, którego celem było zautomatyzowanie procesu wprowadzania zamówień



katalogów z ankiety umieszczonej na stronie Internetowej do systemu informatycznego przedsiębiorstwa. Rozwiązanie miało spełniać oczekiwania przy minimalnych nakładach finansowych i bez zmiany architektury komputerowej. Rozwiązanie powstało w wyniku przyjęcia następujących założeń:

- dane będą przetwarzane off-line, aby nie była konieczna stała łączność z Internetem,
- maksymalnie uproszczona będzie struktura programów działających na serwerze,
- większość logiki przetwarzania zostanie wprowadzona do klasycznego systemu informatycznego.

Za podstawę przyjęto wykorzystanie poczty elektronicznej jako metodę wymiany informacji pomiędzy serwerem internetowym a systemem informatycznym przedsiębiorstwa.

Przedstawione rozwiązanie po pierwsze znakomicie spełnia postawione zadania. Po drugie nadaje się do zastosowania w przedsiębiorstwach, które nie stosują dedykowanych aplikacji internetowych, szczególnie gdy serwer internetowy jest dzierżawiony, z ograniczoną ilością usług do wykorzystania. Po trzecie rozwiązanie to można adaptować do innych zastosowań, na przykład do współpracy z aplikacją typu sklep internetowy w zakresie wymiany danych z systemem informatycznym przedsiębiorstwa.

Aby powiększyć skalowalność rozwiązania należy zmienić klienta pocztowego na specjalizowanego, o większej wydajności.

## LITERATURA

1. Zaliwski A.: Korporacyjne bazy wiedzy, Polskie Wydawnictwo Ekonomiczne, Warszawa 2000.
2. Materiały konferencyjne, II Konferencja Zarządzania Relacjami z Klientem, IDG Poland SA, Stryków 2000.
3. Praca zbiorowa pod red. T. Wierzbickiego: Systemy informatyczne zarządzania. PWE. Warszawa 1985.
4. Paszko M.: IBM – twórca technologii CRM, artykuł, Modern Marketing dodatek specjalny: Software w marketingu i sprzedaży, APW Marketer, Warszawa 2000.
5. Tsichritsis D.C., Lochovsky F.H.: Modele danych. Wydawnictwo Naukowo-Techniczne. Warszawa 1990.
6. Ustawa o ochronie danych osobowych.
7. Dokumentacja do Borland Delphi v.5.0.
8. MSDN Library Help.

Recenzent: Dr Jarosław Francik

Wpłynęło do Redakcji 11 kwietnia 2001 r.

**Abstract**

The most important part of CMR (Customer Relationship Management) systems is the client's database. This database combined with client's references is the key tool of effective and personally clients management systems. One of the purposes of this database is collecting information about new and potential clients. This articles describes the solution which uses the Internet to collect this data. The presented system uses questionnaire published on the company's Web site. Visitors can use it to submit a question, but questionnaire also enable them to type in information about their address, among the other, based on the designed form. The next part of this article describes the method of sending this information to the SI system via an electronic mail. For this purpose system makes use of Microsoft Outlook mail client, managed externally via the COM interface. The last part of the article describes problems concerning verification and interpretation of the data being sent, prior to inserting them into the client's database.