

Rafał Plewa<sup>1)</sup>, Paweł Plewa<sup>2)</sup>

## KONCEPCJA METODYKI AKWIZYCJI DANYCH Z URZĄDZEŃ POMIAROWYCH W LABORATORIUM MOBILNYM

**Streszczenie.** W artykule przedstawiono koncepcję metodyki akwizycji danych z urządzeń pomiarowych zainstalowanych w mobilnym laboratorium badawczym. Propozycja przewiduje ocenę aplikacyjności zastosowania technologii Usług Sieciowych (Web Services) w bezprzewodowej transmisji danych z pomiarów wag przy użyciu wagi Tru-Test 3000. Transmisja danych realizowana jest w oparciu o Usługi Sieciowe zbudowane na platformie Microsoft SharePoint®. Identyfikacja ważonych zwierząt realizowana przy pomocy tagów RFID, zgodnych ze stosowanymi w Unii Europejskiej normami ISO 11784 oraz ISO 11785.

Zebrane w ten sposób dane, poddane procesowi ich oczyszczania, kontroli ich jakości i wstępnej analizy pozwolą na dalsze ich przetwarzanie w celu pozyskania z nich wiedzy w procesie data miningu. Zastosowanie data miningu, pozwoli ocenić wpływ oraz istotność wybranych czynników produkcji wołowiny na te cechy mięsa, które są najbardziej znaczące z punktu widzenia konsumenta.

**Słowa kluczowe:** ekstrakcja danych, eksploracja danych, RFID, Usługi Sieciowe, wagi, akwizycja danych, laboratorium mobilne, produkcja wołowiny, optymalizacja produkcji.

### WSTĘP

Niedawna reforma europejskiej polityki rolnej zaowocowała zmianami w kryteriach, według których naliczane są dopłaty za uzyskanie jakości premium mięsa wołowego. Hodowcy bydła mięsnego, którzy w szczególności byli zależni od tych kwot, w celu utrzymania przychodów, musieli przeewaluować swoje systemy hodowli w celu znalezienia optymalnych metod produkcji [1]. Produkcja wołowiny w systemach zapewnienia jakości ukierunkowana musi być na potrzeby konsumentów, co wynika faktu, iż konsumenci krajów wysoko rozwiniętych są w stanie płacić więcej za produkty cechujące się wysoką, powtarzalną jakością. Również na rynku polskim obserwuje się coraz większe zainteresowanie mięsem o wysokiej jakości, które jest produkowane w systemach gwarantujących powtarzalną jakość. Produkcja produktów żywnościowych o wysokiej i powtarzalnej jakości musi odbywać się w oparciu o rzetelną wiedzę naukową, w tym o nowoczesne narzędzia statyczne po-

<sup>1)</sup> Zakład Techniki w Żywnieniu, Katedra Żywności Funkcjonalnej i Towaroznawstwa, Wydział Nauk o Żywnieniu Człowieka i Konsumpcji, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie, e-mail: rafal\_plewa@sggw.pl

<sup>2)</sup> Katedra Podstaw Inżynierii, Wydział Inżynierii Produkcji, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie, e-mail: pawel\_plewa@sggw.pl

zwalające na eksplorację danych uwzględniających informację zwrotną od konsumenta [10, 8, 3, 9].

Za konieczne uznano zatem stworzenie modelu matematycznego, pozwalającego na identyfikację optymalnych systemów produkcji wołowiny. Pierwszym krajem, w którym podjęto tego typu działania była Irlandia, gdzie funkcją celu tego typu modelowania była maksymalizacja zysków z produkcji, a parametrami wejściowymi były różne sposoby karmienia zwierząt. Model uwzględniał złożone zależności pomiędzy kosztami karmienia, utrzymania zwierząt, ceną wołowiny, wymaganiami odnośnie podaży, wielkością gospodarstwa rolnego oraz polityką ochrony środowiska. Wyniki badań pokazały, że wahania cen wołowiny skutkują dopasowywaniem się systemów produkcji do nowych warunków i, w rezultacie, prowadzą do powstania dużych różnic między przychodami poszczególnych gospodarstw [4].

Stworzenie optymalnego modelu możliwe jest przy wykorzystaniu inteligentnej eksploracji danych (np. przy zastosowaniu narzędzi statystycznych – data miningu), która ma na celu wydobycie zależności z dużej ilości informacji. W terminologii firmy SAS data mining oznacza procesy selekcji, eksploracji i modelowania, wykonane na dużej ilości danych, prowadzących do odkrycia dotychczas nieznanymi wzorców [13]. Data mining jest zarazem procesem wydobycia korelacji, wzorców, trendów na podstawie dużych wolumenów danych przechowywanych w repozytoriach, przy wykorzystaniu technologii rozpoznawania wzorców oraz technik statystycznych i matematycznych, jak również jest technologią pozyskiwania wiedzy poprzez stosowanie modeli sieci neuronowych i algorytmów genetycznych, drzew decyzyjnych, technik statystycznych, modeli fraktalnych, czy algorytmów segmentacji [2].

W artykule poruszone zostaną zagadnienia zintegrowania i ujednoczenia mechanizmów komunikacji w specjalistycznych urządzeniach pomiarowych, które dotychczas nie były używane w zintegrowanym systemie, w którym konieczne jest zebranie danych z wielu niezależnych źródeł na potrzeby późniejszych analiz statystycznych.

## ZASTOSOWANIE I ELEMENTY LABORATORIÓW MOBILNYCH

W skład laboratorium mobilnego wchodzi:

- elektroniczna waga tensometryczna ze zintegrowanym czytnikiem RFID (ang. *Radio Frequency Identification* – radiowy system automatycznej identyfikacji)
- analizator pasz AgriNIR,
- ultrasonograf ze zintegrowanym czytnikiem RFID,
- komputer przenośny o podwyższonej wytrzymałości,
- zestaw narzędzi do pomiarów zoometrycznych.

Elektroniczna waga tensometryczna TruTest 3000 posiada zintegrowany czytnikiem tagów RFID, połączony z anteną montowaną na klatce wagowej dla bezdo-

tykowej identyfikacji zwierzęcia. Waga posiada również konsolę z ekranem LCD pokazującą aktualne pomiary i identyfikatory zwierząt, dającą możliwość odczytu statystyk w oparciu o poprzednie pomiary oraz konsolę alfanumeryczną, pozwalającą na ręczne wprowadzenie danych. Z klatką wagową zintegrowany jest również poskrom zwierząt umożliwiający dokonanie pomiarów zoometrycznych oraz zakolczykowanie zwierząt kolczykami z tagami RFID.

Analizator pasz AgriNIR daje możliwość przeprowadzenia badania wartości takich jak: białko, popiół, skrobia, wilgotność, ADF (ang. *Acid Detergent Fiber* – włókno kwaśno-detergentowe), NDF (ang. *Neutral Detergent Fiber* – włókno obojętno-detergentowe)

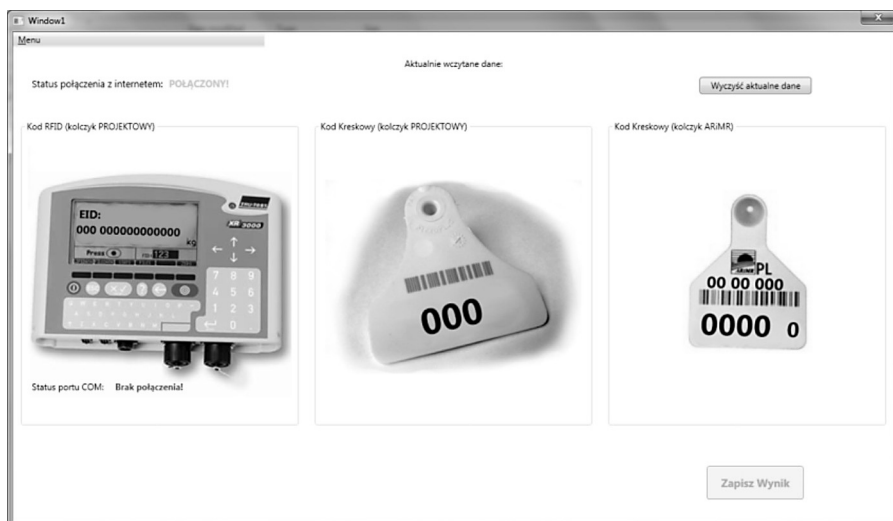
Ultrasonograf ze zintegrowanym czytnikiem RFID, daje możliwość zapisania zdjęć mięśni (np. mięśnia najdłuższego grzbietu *m. longissimus dorsi*) zwierząt w pamięci urządzenia w celu ich późniejszej analizy. Parametry zdjęcia, jak i numer identyfikacyjny zwierzęcia pozyskany ze zintegrowanego czytnika RFID są zapisywane w postaci metadanych. W skład zestawu do badań USG poza samym urządzeniem z wysokiej rozdzielczości wyświetlaczem LCD, zasilaniem, okablowaniem zestawem głowic i przystawek wchodzi wszelkie inne niezbędne do wykonania badania elementy.

W skład narzędzi do pomiarów zoometrycznych, wchodzi elementy, w których odczyt wartości pomiarów odbywa się w sposób wizualny i nie mają one połączenia elektronicznego z resztą systemu. Dane pozyskane przy pomocy tych narzędzi wprowadzane są do systemu przez pracowników technicznych poprzez odpowiednie formularze.

W skład zestawu wchodzi również komputer przenośny o podwyższonej wytrzymałości na działanie wody, kurzu i wstrząsów, wyposażony w moduł bezprzewodowej transmisji danych w oparciu o technologie GPRS/EDGE/HSDPA.

## PROCES BADANIA I IDENTYFIKACJI ZWIERZĄT W LABORATORIUM MOBILNYM

Każda sztuka bydła w Polsce musi posiadać tak zwany kolczyk IRZ (kolczyk w systemie Identyfikacji i Rejestracji Zwierząt) nadany przez Agencję Restrukturyzacji i Modernizacji Rolnictwa (ARiMR), co jest podyktowane przepisami obowiązującymi w Unii Europejskiej. Kolczyk IRZ posiada wydrukowany w postaci numerycznej oraz kodu kreskowego numer identyfikacyjny według określonego formatu, zwanego dalej numerem IRZ. Numer IRZ jest obowiązującym numerem identyfikacyjnym zwierzęcia w ramach całego procesu produkcji wołowiny, od urodzenia zwierzęcia, do przygotowania odpowiednich porcji mięsa używanego do konsumpcji. W celu przyspieszenia pomiarów zwierzęta kolczykuje się kolczykami zawierającymi tagi RFID, co pozwala na bezkontaktową identyfikację zwierzęcia (rys. 1).



Rys. 1. Interfejs stworzonej aplikacji do parowania kolczyków (opracowanie własne)

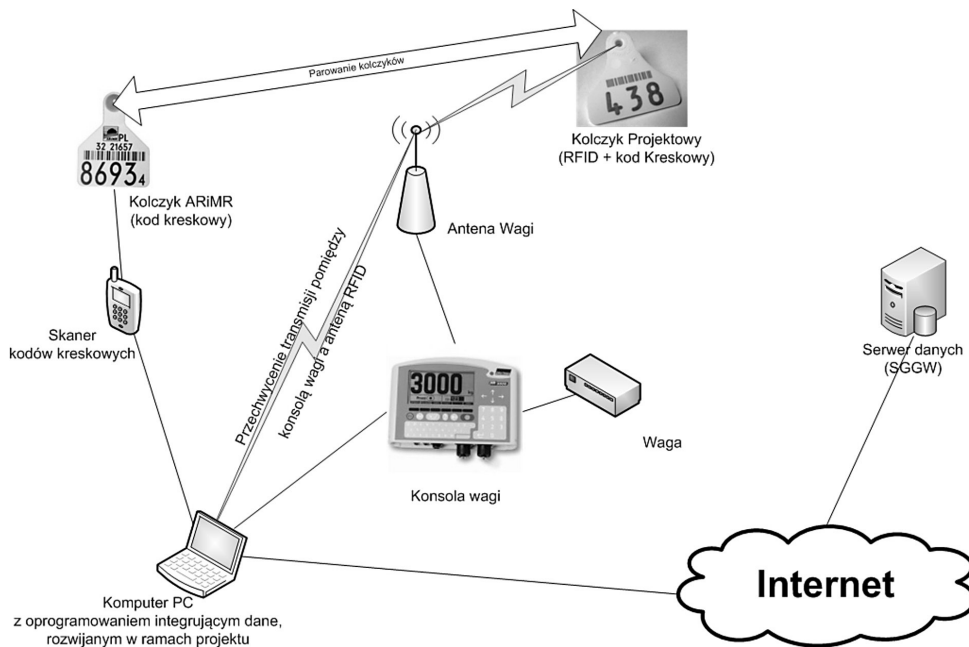
Przy pierwszym badaniu zwierząt w danym gospodarstwie, przeprowadzony jest proces kolczykowania zwierząt kolczykami z tagami RFID.

Aby kolczyki RFID były użyteczne, konieczne jest przeprowadzenie procesu, który skojarzy kolczyk zawierający tag RFID z kolczykiem IRZ, przy pomocy specjalnie stworzonego oprogramowania, transferującego w czasie rzeczywistym (przy pomocy transmisji GPRS/EDGE), skojarzone ze sobą pary kolczyków do centralnego serwera.

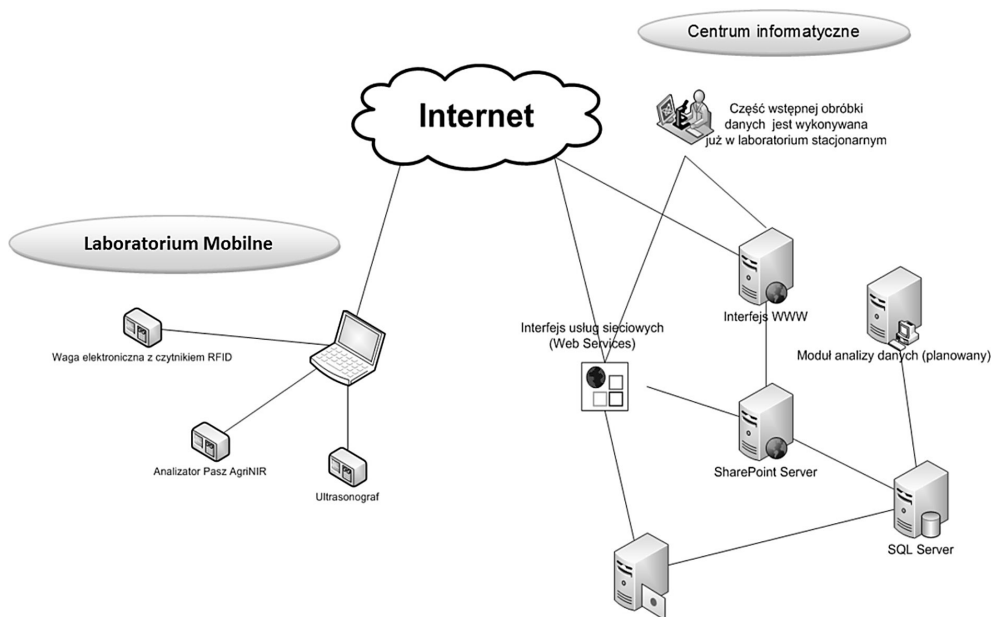
Podczas badania, w laboratorium mobilnym, zwierzęta są przepędzane przez wagę opisaną w poprzedniej części, która pozwala na automatyczne przeprowadzenie pomiaru wagi i skojarzenia wyniku pomiaru z numerem identyfikacyjnym (RFID) zwierzęcia. Po zakończeniu pomiaru w całym gospodarstwie i podłączeniu konsoli wagi do komputera, dane opatrzone stemplem czasowym i powiązane z gospodarstwem pomiary, są transferowane do centralnego serwera gdzie znajdują się już skojarzone pary kolczyków, co pozwala na przeglądanie danych według numerów IRZ zwierząt (rys. 2).

Wyniki badań pasz, przy pomocy specjalnej aplikacji, również są automatycznie przesyłane do centralnego serwera. Dane z pomiaru parametrów pasz skojarzone są gospodarstwem zarejestrowanym wcześniej w systemie, tak samo jak to jest w przypadku wyników pomiarów zwierząt, przez co jest możliwe późniejsze sprawdzenie parametrów paszy, którą zwierzę było karmione.

Wynikiem badania ultrasonografem są zdjęcia mięśni zwierząt (mięśnia najdłuższego grzbietu – *m. longissimus dorsi*). Również te dane są opatrzone stemplem czasowym i numerem tagu RFID, dzięki nowatorskiemu rozwiązaniu integracji czytnika RFID z ultrasonografem (rys. 3). Ultrasonograf posiadający ekran LCD



Rys. 2. Przykładowy element systemu – moduł do pomiaru wag (opracowanie własne)



Rys. 3. Uproszczony schemat połączeń w laboratorium mobilnym (opracowanie własne)

wysokiej rozdzielczości i zaopatrzone w manipulatory daje możliwość natychmiastowego zbadania długości i pola powierzchni przekroju mięśni, jednak ze względu na trudne zwykle warunki podczas badania, zdjęcia transferowane są do centralnego serwera bez metadanych zawierających numeryczne wyniki pomiaru. Mierzenie długości i pola przekroju mięśnia możliwe jest poprzez aplikację działającą w oparciu o zdjęcia pobierane automatycznie z centralnego z serwera.

## **MOŻLIWOŚCI WYKORZYSTANIA TECHNIK DATA MININGU DO OPTIMALIZACJI PROCESU PRODUKCJI MIĘSA WOŁOWEGO**

Natura badanych danych, pochodzących zarówno z urządzeń pomiarowych laboratoriów mobilnych wzbogacona o dane z innych źródeł, takich jak ankiety konsumentów, dane z laboratoriów stacjonarnych dotyczące badań sensorycznych czy komputerowej analizy obrazu, wymaga operacji na bardzo dużych zbiorach danych, o różnorodnej strukturze. Uniemożliwia to wykorzystanie konwencjonalnych metod analizy danych, które nie działają wystarczająco efektywnie w przypadku dużych baz danych o niejednorodnym formacie.

## **OPIS PRZEBIEGU PROCESU ODKRYWANIA WIEDZY Z DANYCH**

### **1. Zrozumienie dziedziny problemu**

Złożoność natury danych, a także problemów stawianych przy okazji ich analizy, nie pozwala na natychmiastowe sformułowanie pytań, na które odpowiedź przybliża do osiągnięcia głównego celu – optymalizacji procesu produkcji wołowiny. Analizując dane z produkcji mięsa wołowego należy ustalić, jakie informacje są dostępne oraz jakich badacz oczekuje uzyskać po procesie analizy. Często ostateczny cel tego typu analizy krystalizuje się dopiero podczas pracy z danymi. Jego określenie dokonuje się zatem w trakcie interaktywnego procesu, w którym istnieje stała możliwość uwzględniania wiedzy oraz preferencji użytkownika.

### **2. Budowa roboczego zbioru danych**

Dane gromadzone w procesie badań w laboratoriach mobilnych są możliwie pełne – czyli w przypadku niezawodnego działania sprzętu pomiarowego, wszystkie zaprojektowane komórki zawierają dane. Dziedzina problemu może jednak wymagać włączenia do analizy informacji o samych docelowych konsumentach produktów badanych w laboratoriach mobilnych. Dane takie nie zawsze są w pełni dostępne. Dobór źródeł danych zależy od pytań, na które należy uzyskać odpowiedź, jednakże trzeba uwzględnić fakt, że różnorodność zasobów informacji nie zawsze pozwala na natychmiastową specyfikację odpowiednich jej źródeł [11].

### 3. Oczyszczenie i przygotowanie danych

Wszelka dalsza analiza musi być poprzedzona selekcją tych danych, które wydają się w ustalonym rozumieniu adekwatne. Wybrane źródła muszą być zapisane w postaci w miarę zunifikowanego formatu, umożliwiającego pracę algorytmom przeszukującym dane pod kątem wykrywania interesujących zależności czy regularności.

### 4. Przekształcanie i redukcja danych

Nawet na etapie selekcji i przygotowania danych do dalszej analizy nie można ostatecznie wyszczególnić tych składowych opisu, które okażą się najbardziej przydatne w rozwiązaniu postawionego problemu. Wszelkie zbierane informacje o każdym badanym zwierzęciu czy też produktach pochodnych można opisać za pomocą bardzo wielu różnych cech, z których nie wszystkie okażą się być przydatne, jako zapis wiedzy stosowalnej w praktyce [11].

### 5. Eksploracja danych – data mining

Po przeprowadzeniu wyżej opisanych czynności, otrzymuje się dane zapisane w miarę możliwości, w formie tablicy, której wiersze odpowiadają obiektom (np. zwierzętom, porcejom mięsa), zaś kolumny – opisującym je cechom. Eksploracja danych to algorytmiczne wykrywanie zależności pomiędzy wartościami tych cech, które może być realizowane przy użyciu bardzo różnorodnych technik, opartych na statystyce, sztucznej inteligencji, czy też odwołujących się do metod uczenia maszynowego.

Przykładowo, zakładając, że projektowanym zadaniem jest stworzenie klasyfikatora, który – otrzymawszy na wejściu informacje o tempie i dynamice wzrostu danego mięśnia zwierzęcia – wskaże na wyjściu proporcje rodzajów paszy, przy pomocy których najprawdopodobniej zwierzę było karmione. Jest to problem zaprojektowania tak zwanej „czarnej skrzynki”, której jedynym celem jest poprawne rozpoznawanie i klasyfikacja nowo badanych obiektów. Tak rozumiany klasyfikator można uzyskać stosując regresję liniową, sieci neuronowe, czy klastering [7].

Przykładem innego celu może być zrozumiały opis profilów zwierząt pod kątem np. wartości rzeźnej czy udziału poszczególnych udziałów tłuszczu w elemencie kulinarnym. W takim przypadku, wynikiem musi być opis w postaci drzewa decyzyjnego, bądź zbioru reguł asocjacyjnych. Trzeba jednak pamiętać o weryfikacji otrzymanych wyników pod kątem poprawności, ponieważ niektóre wyniki mogą zadowalająco sprawdzać się na danych, ale są raczej nieprzydatne dla osiągnięcia celu – optymalizacji produkcji wołowiny. Konieczne jest więc zatem zastosowanie podejść bardziej wyrafinowanych, korzystających z głębszej informacji o naturze danych [12].

### 6. Interpretacja znalezionych zależności

Na tym etapie, po konsultacjach z ekspertami w dziedzinie produkcji mięsa wołowego, decyduje się o przydatności otrzymanych wyników do realizacji zakłada-

nych na samym początku celów. Ewentualne sugestie, czy krytyczne uwagi mogą prowadzić do konieczności powtórzenia nawet całego opisanego procesu odkrywania wiedzy z danych, po uwzględnieniu odpowiednich modyfikacji. Taka kolejna iteracja jest normalnym procesem, ponieważ złożoność rozpatrywanych zadań analitycznych, a także różnorodność danych, przekracza możliwości ludzkiej percepcji.

## 7. Istota procesu eksploracji danych

Eksploracja danych (data mining) rozumiana jest jako proces wykrywania w danych „regularności”, interesujących z punktu widzenia osiągnięcia celu. To właśnie one niosą w sobie informacje o tym, które cechy są tak naprawdę istotne. Otrzymana w ich formie wiedza jest czymś nowym i stanowi klucz do oceny jakości procesu pozyskiwania wiedzy z danych i może być źródłem sugestii dotyczących modyfikacji etapów poprzednich [14].

Problemem jest fakt, że takich regularności jest w danych praktycznie „nieskończenie wiele”, zaś dla osiągnięcia celu – optymalizacji procesu produkcji wołowiny interesujące będą tylko niektóre z nich i to w różnym stopniu. Dlatego też odpowiednie narzędzia oraz metody analizy wydają się być niezbędne a celu przeprowadzania procesu optymalizacji produkcji wołowiny w oparciu o pożądane przez konsumentów cechy wołowiny.

## WNIOSKI

Stosowanie wielu wyspecjalizowanych urządzeń pomiarowych, których pierwotnym przeznaczeniem nie była integracja czy współpraca z innymi urządzeniami oraz fizyczne rozbudowanie istniejących urządzeń pomiarowych o niestosowane wcześniej w Polsce rozwiązania – takie jak zintegrowanie z układami elektronicznymi i firmware urządzenia czytnika tagów RFID – wymaga zbudowania odpowiedniej platformy programowo-sprzętowej. Platforma taka pozwala zebranie danych z urządzeń, znajdujących się w zdalnej lokalizacji, wstępne ich oczyszczenie i zregulowanie, co stanowi punkt wyjścia do dalszej analizy. Należy podkreślić, że wszelkie zbierane dane są logicznie połączone ze sobą, co jest zapewnione dzięki zaprojektowanemu systemowi identyfikacji badanych elementów od urodzenia – cielaka do wyprodukowania gotowego do konsumpcji elementu kulinarnego.

W takim przejściowym repozytorium, zawierających duże woluminy danych, które mimo pochodzenia z różnych urządzeń pomiarowych, są logicznie powiązane, zawarta jest już wiedza, która dzięki zastosowaniu technik data miningu może zostać odpowiednio eksplorowana. Zastosowanie tych technik i wykorzystanie tak pozyskanej wiedzy, pozwoli na wyprodukowanie mięsa wołowego o zaprojektowanym, zadany przez producenta i spodziewanym przez konsumenta poziomie jakości, w tym o oczekiwanym zestawie cech produktu (sensorycznych, fizykochemicznych, czy wartości odżywczej).



Techniki data miningu z powodzeniem stosowane są już w takich dziedzinach jak medycyna, ekonomia, marketing, bankowość, technika, technologie informatyczne, jak i inżynieria rolnicza. W omówionym artykule, w tak dużym i odpowiednio zdefiniowanym repozytorium różnorodnych danych, data mining jest idealną techniką dającą podstawy do optymalizacji procesu produkcji wołowiny w Polsce.

## BIBLIOGRAFIA

1. Berentsen P.B.M., Giesen G.W.J., Renkema J.A.: Introduction of seasonal and spatial specification to grass production and grassland use in a dairy farm model. *Grass Forage Sci.* 55, 2000: 125–137.
2. Berry M. J. A., Linoff G.: *Data mining techniques: for marketing, sales and customer support.* Willey & Sons 1997.
3. Cros M.J., Duru M., Garcia F., Martin-Clouaire R.: Simulating management strategies: the rotational grazing example. *Agric. Syst.* 80, 2004: 23–42.
4. Crosson P., O’Kiely P., O’Mara F.P., Wallace M.: The development of a mathematical model to investigate Irish beef production systems *Agricultural Systems.* 89 (2-3), 2006: 349-370
5. Dunham M.H.: *Data Mining. Introductory and Advanced Topics.* Prentice Hall 2003.
6. Fayyad U.M., Piatetskiy-Shapiro G., Smith P., Ramasasmy U.: *Advances in Knowledge Discovery and Data Mining*”. AAAI Press / MIT Press 1996.
7. Han J., Kamber M.: *Data Minig: Concepts & Techniques,* Morgan Kaufmann 2000.
8. McGilloway D.A., O’Riordan E.G.: The potential for grassland based ruminant production systems beyond 2000. *Irish Grassland Animal Production Association J.* 33, 1999: 3–16.
9. Nielsen B.K., Kristensen A.R., Thamsbourg S.M.: Optimal decisions in organic steer production – a model including winter feed level, grazing strategy and slaughtering policy. *Livestock Production Sci.* 88, 2004: 239–250.
10. O’Riordan E.G., O’Kiely P.: Potential of beef production systems based on grass. *Irish Grassland Animal Production Association J.* 30, 1996: 185–217.
11. Pyle D.: *Data Preparation for Data Mining.* Morgan Kaufmann, Harcourt Intl. 1999.
12. Sheikholeslami G., Chatterjee S., Zhang A.: WaveCluster: A Multi-resolution clustering approach for very large spatial databases“, *Proceedings of the XXIV VLDB Conference,* NY 1998.
13. STATISTICA Data Miner. StatSoft Inc. 2002.
14. Weiss S.M, Indurkha N.: *Predictive data mining. A practical guide,* Morgan Kaufman Publishers 1998.

Badania zostały wykonane w ramach Projektu „Optymalizacja produkcji wołowiny w Polsce zgodnie ze strategią *od widelca do zagrody*” współfinansowanego przez Europejski Fundusz Rozwoju Regionalnego w ramach Programu Operacyjnego Innowacyjna Gospodarka (nr umowy UDAPOIG.01.03.01-00-204/09-03).

## THE CONCEPT OF DATA ACQUISITION METHODS FOR MOBILE LABORATORY MEASURING DEVICES

### Summary

This article presents the concepts of methods used for data acquisition in mobile laboratory equipment. The proposal includes feasibility study and evaluation of Web Services technology used for wireless data transmission in electronic animal weigh scale Tru-Test 3000. Data transmission is accomplished through the use of Web Services built on Microsoft SharePoint® platform. The identification of examined animals is accomplished by use of RFID tags, that conforms the European Union's standards ISO 11784 and ISO 11785.

The acquired data, processed for data cleaning, quality control and preliminary analysis will enable further processing in purpose of knowledge extraction in data mining process. The use of data mining will enable the evaluation of influence and significance of selected beef production factors on the traits that are most important for consumers.

**Keywords:** Data mining, RFID, Web Services, weigh scales, data acquisition, mobile laboratories, beef production, production optimization.