



Zasady i mierniki oceny jakości danych

Ocena jakości danych CORINE Land Cover

Elżbieta Bielecka

WAT, Wydział Inżynierii Lądowej i Geodezji
Instytut Geodezji i Kartografii

III Konferencja naukowo-techniczna

WYKORZYSTANIE WSPÓŁCZESNYCH ZOBRAZOWAŃ SATELITARNYCH, LOTNICZYCH
I NAZIEMNYCH DLA POTRZEB OBRONNOŚCI KRAJU I GOSPODARKI NARODOWEJ



JAKOŚĆ



- **Definicje**

- **Wg Platona (*gr. poiotes*)** - "pewien stopień doskonałości,, - kategoria filozoficzna
- Zgodność z celem
- Zgodność ze specyfikacją (zero braków)
- Stopień doskonałości wyrobu lub usługi
- Zespół cech i charakterystyk wyrobu lub usługi, które noszą w sobie zdolność zaspokojenia określonej potrzeby



JAKOŚĆ DANYCH



- **Jakość danych przestrzennych**

- **Ogół cech produktu** ukierunkowany na jego zdolność do zaspokojenia wyrażonych i ukrytych potrzeb (ISO 19101)
- W modelu jakości danych jakość jest **ogólnym wskaźnikiem** stosowanym do ustalenia przydatności zbioru danych do użycia w określonym celu
- **Dane są wysokiej jakości** jeżeli nadają się do użycia zgodnie z przeznaczeniem w zakresie działania, podejmowania decyzji i planowania
- **Dane nadają się do użycia** zgodnie z przeznaczeniem, jeżeli nie zawierają defektów i posiadają pożądane cechy (Redman T.C. "Data Quality. The Field Guide". 2001)



JAKOŚĆ DANYCH



- **Elementy jakości danych przestrzennych (ISO 19113)**
 - **Kompletność** - obecność i nieobecność obiektów, ich atrybutów i związków (nadmiar, niedomiar danych)
 - **Spójność logiczna** - stopień zgodności z logicznymi regułami struktury danych, atrybutów i związków (pojęciowa, dziedziny topologiczna, formatu)
 - **Dokładność** - dokładność położenia obiektów
 - **Dokładność czasowa** - dokładność czasowych (związanych z czasem) atrybutów i związków obiektów
 - **Dokładność tematyczna** - dokładność atrybutów ilościowych i poprawność atrybutów nieilościowych oraz klasyfikacji obiektów i ich związków



CORINE Land Cover



- Baza zawierająca informacje o pokryciu terenu i użytkowaniu ziemi w Europie
- Cyklicznie aktualizowana
1990, 2000, 2006, 2009/2010
- Jednolite założenia metodyczne
 - Nomenklatura
 - Dokładność
 - Materiały źródłowe
 - Technologia



ZAŁOŻENIA METODYCZNE



Legenda CORINE Land Cover

Poziom europejski

1. Tereny antropogeniczne
2. Tereny rolne
3. Lasy i ekosystemy seminaturalne
4. Tereny podmokłe
5. Wody

Poziom regionalny

15 klas

Poziom krajowy

44 formy (w Polsce - 31)



ZAŁOŻENIA METODYCZNE



- **Materiał źródłowy – obrazy satelitarne**
 - Landsat 5 TM (1985-1993)
 - Landsat 7 ETM+ (1999-2001)
 - SPOT 4, 5; IRS (2005-2007)
 - Bezchmurne
- **Materiały dodatkowe**
 - Mapy topograficzne i tematyczne
 - Zdjęcia lotnicze
 - Opracowania monograficzne



ZAŁOŻENIA METODYCZNE



- Interpretacja wizualna wspomagana komputerowo (klasy mieszane, zdjęcia z różnych satelitów, lat i okresów wegetacyjnych)
- Wiarygodność interpretacji 85%
- Elementy powierzchniowe
 - Najmniejsza powierzchnia kartowanych form 25 ha
 - Najmniejsza powierzchnia zmiany pokrycia terenu 5 ha
 - Najmniejsza szerokość 100 m
- Precyzja digitalizacji form pokrycia terenu – 50 m



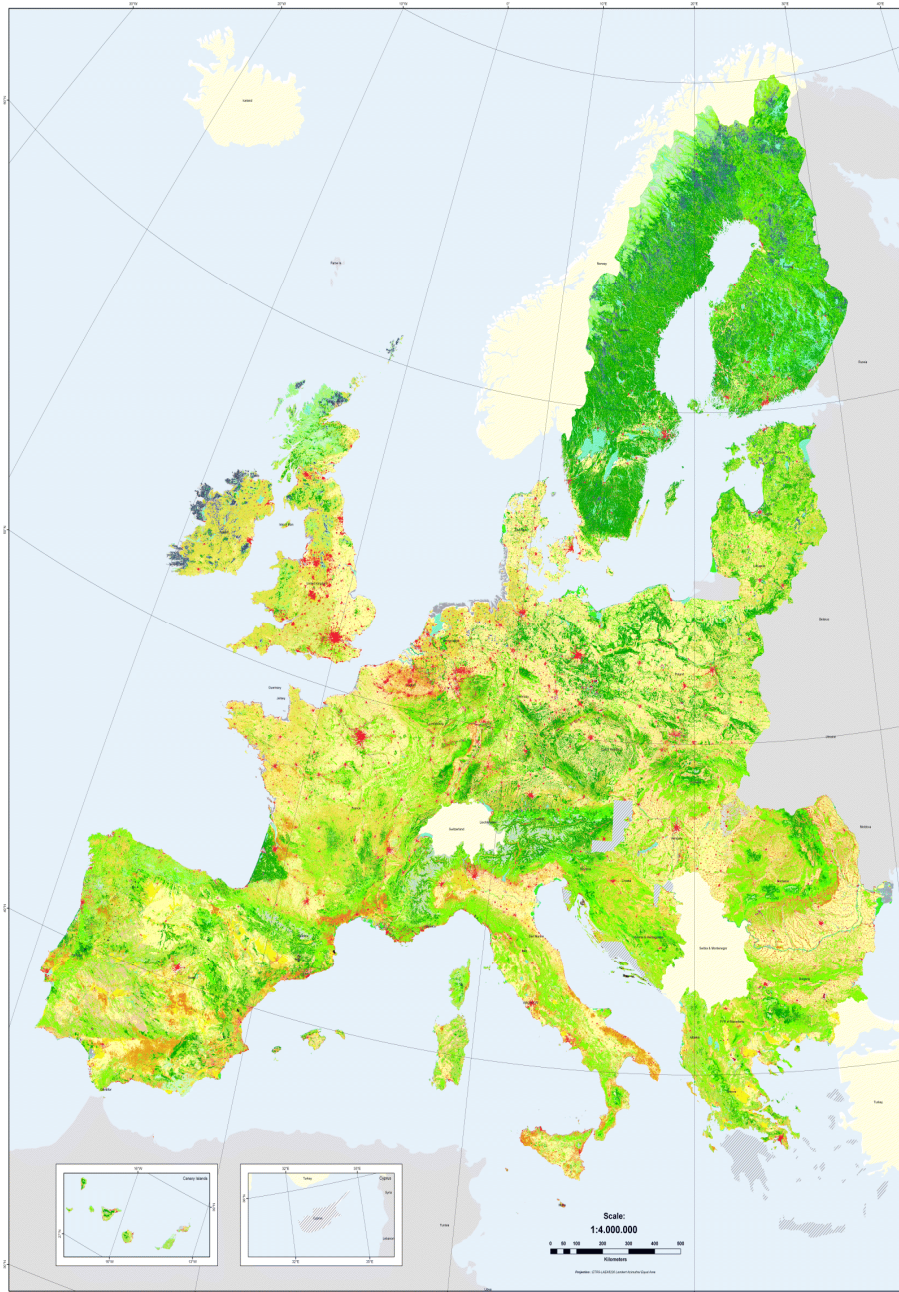
Założenia metodyczno-realizacyjne



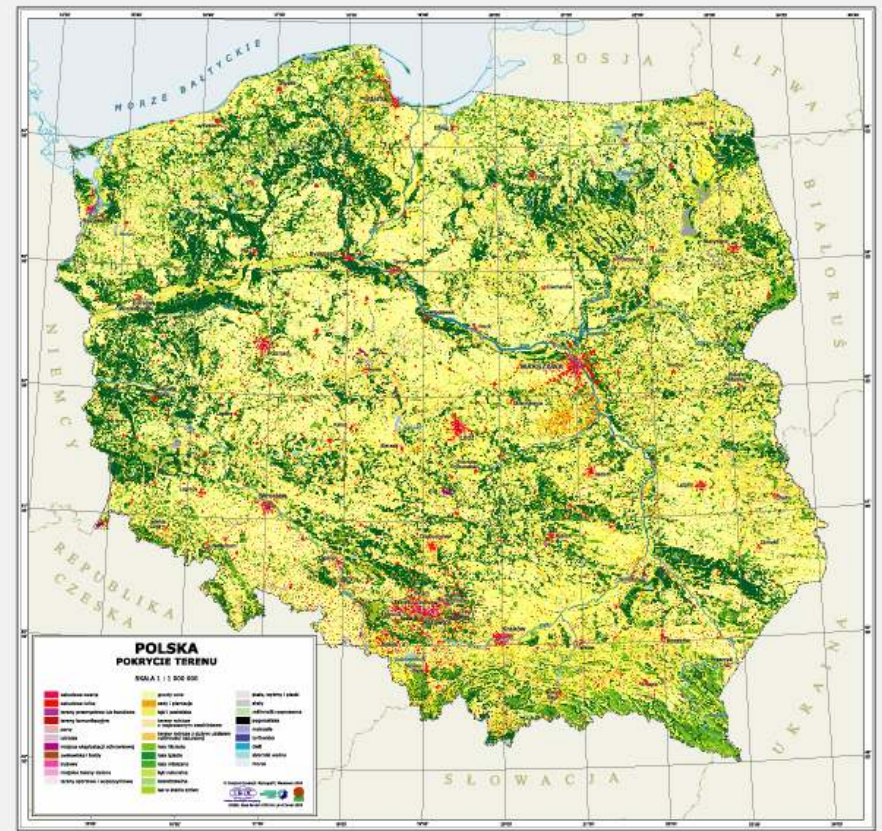
- Baza danych wektorowa, topologiczna - ARC/INFO export file
- Wykorzystanie oprogramowania typu GIS umożliwiającego:
 - digitalizację na ekranie
 - jednoczesne wyświetlanie w dwóch sąsiadujących oknach zdjęć satelitarnych z lat 90. i roku 2000
 - Wyświetlanie pojedynczych kanałów i kompozycji barwnych
 - wzmocnienie treści
- Metadane wg standardu EN-ISO 19 115



Corine land cover 2000



Wyniki realizacji projektu



OCENA JAKOŚCI



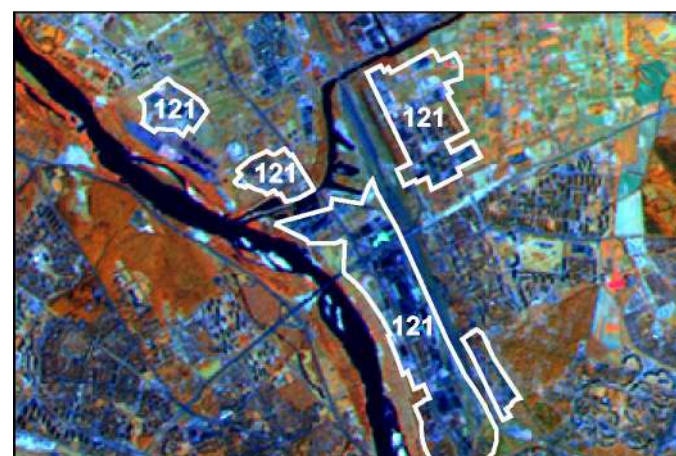
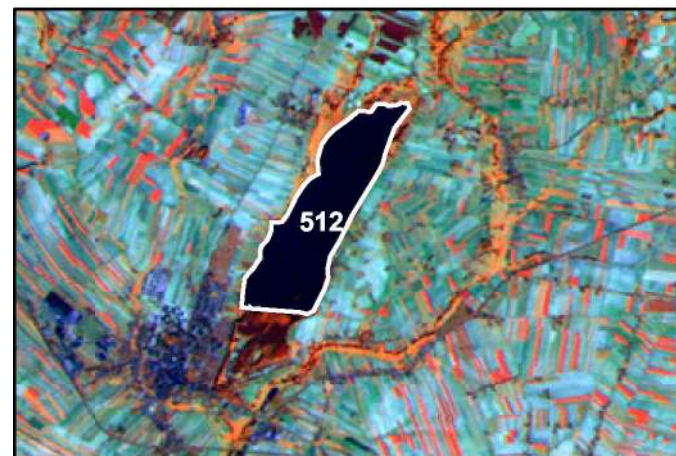
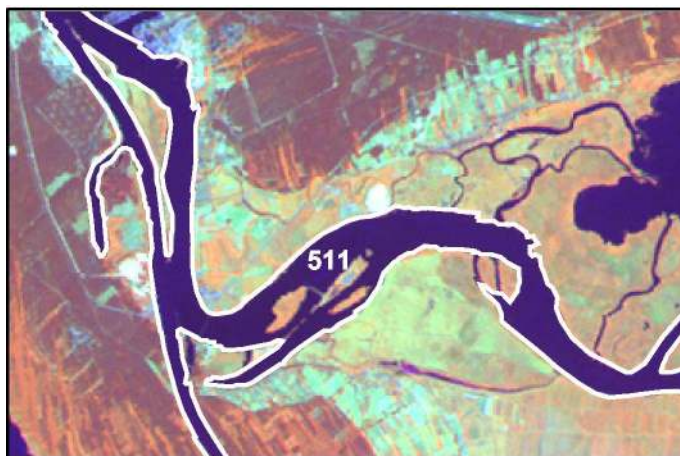
- **Kompletność bazy danych**
 - Europa (bez Rosji, Ukrainy i Białorusi), Turcja
- **Spójność logiczna**
 - Dziedziny (100%)
 - Pojęciowa (90%)
 - Topologiczna (100%)
 - Formatu (100%)
- **Dokładność tematyczna zróżnicowana**
 - 95% - 70%
- **Dokładność położenia 100m**
- **Szczegółowość (25 ha, 5 ha)**
- **Aktualność**
 - 1985-1993; 1999-2001; 2005-2007



DOKŁADNOŚĆ TEMATYCZNA



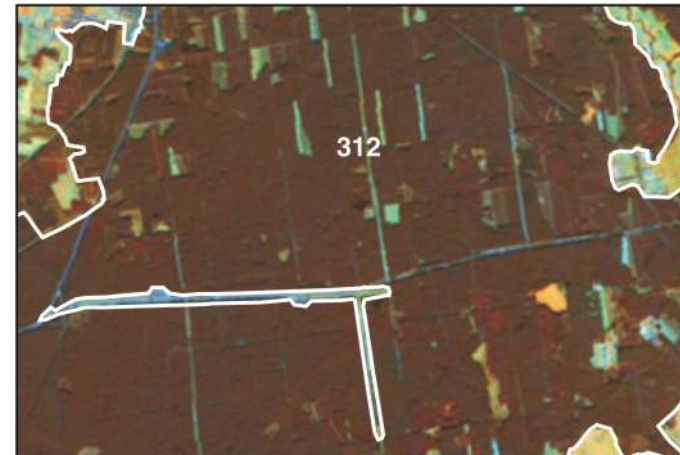
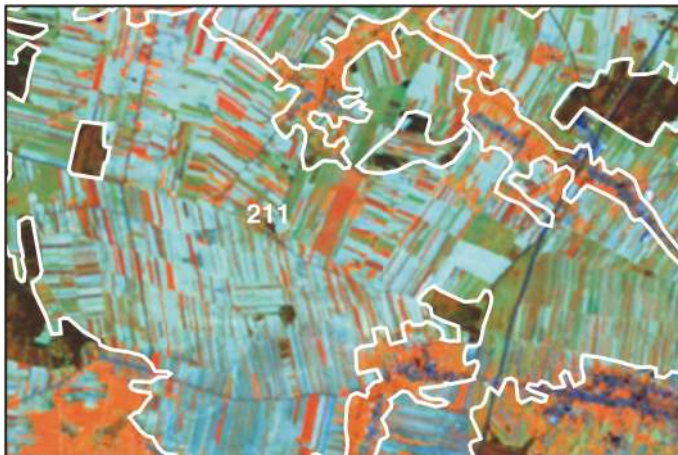
Dokładność bardzo wysoka >90%



DOKŁADNOŚĆ TEMATYCZNA



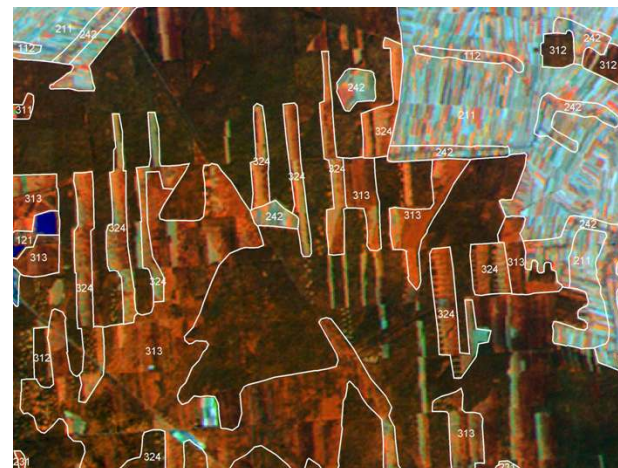
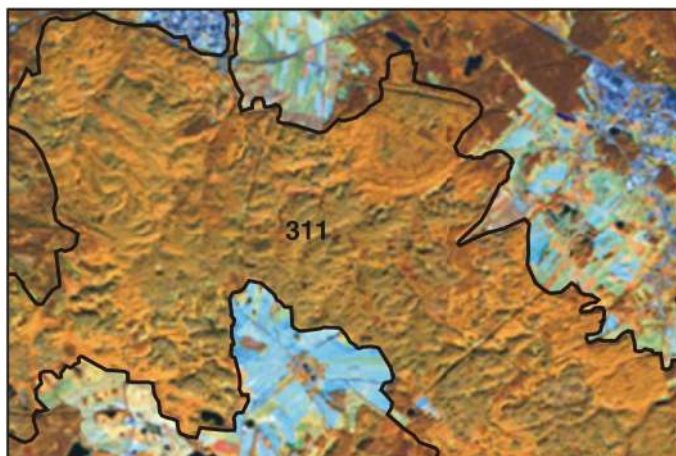
Dokładność bardzo wysoka 90-95%



DOKŁADNOŚĆ TEMATYCZNA



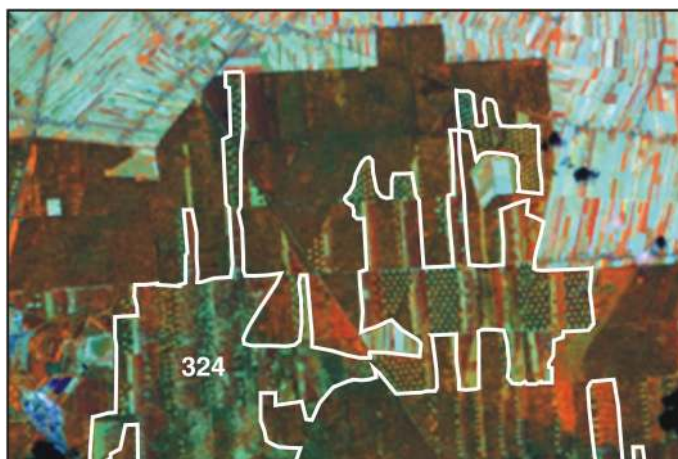
Dokładność dobra 85-90%



DOKŁADNOŚĆ TEMATYCZNA



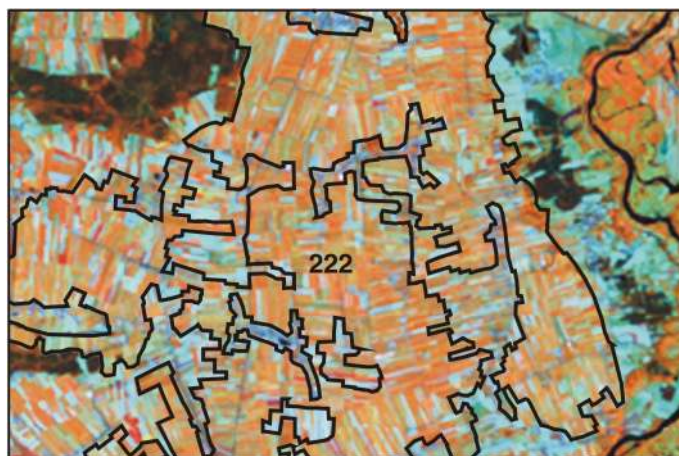
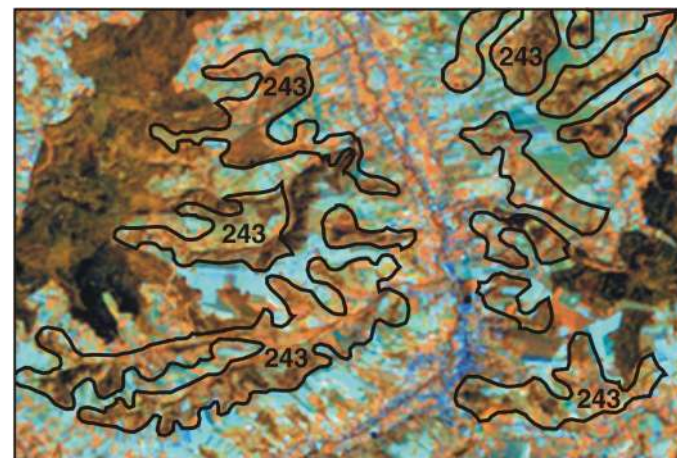
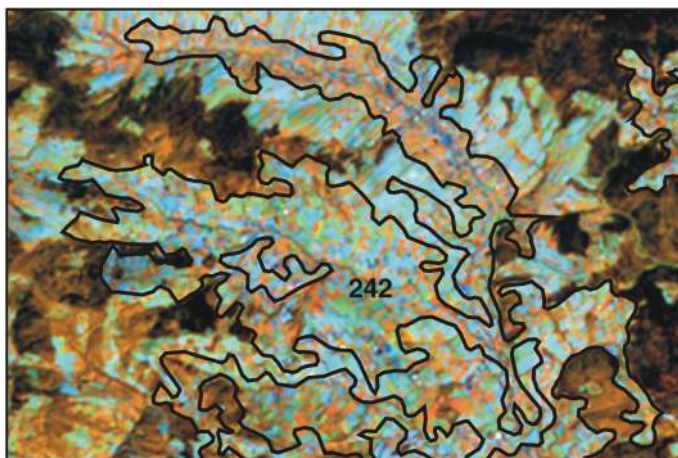
Dokładność średnia 80-85%



DOKŁADNOŚĆ TEMATYCZNA



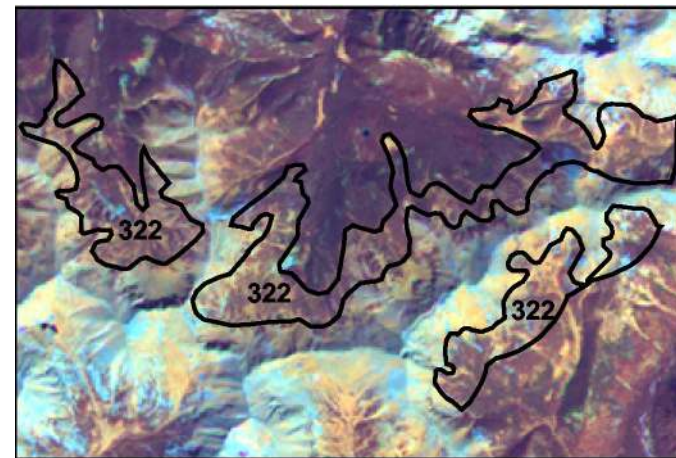
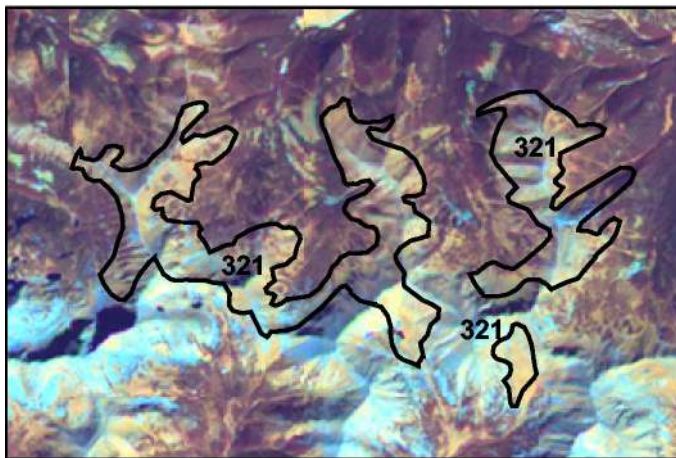
Dokładność niska 70-80%



DOKŁADNOŚĆ TEMATYCZNA



Dokładność niska 70-80%



Dokładność bardzo niska <70%



BŁĘDY W DANYCH



Źródła defektów w danych

Spowodowane **czynnikami ludzkim**

- błędy użytkowników
- wady aplikacji
- wielokrotne migracje i integracje danych

Powstałe wskutek **czynników zewnętrznych**

- awarie sprzętowe
- samoistne starzenie się danych



PODSUMOWANIE



- Jakość danych jest pojęciem **subiektywnym**
- Nie można zastosować jednego kryterium oceny do wszystkich możliwych sytuacji
- W celu zbadania jakości danych potrzebna jest:
 - analiza i poznanie charakteru danych
 - zapoznanie się z przeznaczeniem danych i procesów, jakim są one poddawane

