

Wybrane aspekty przetwarzania zasobu PODGiK do postaci cyfrowej w powiecie wołomińskim

KROK DO OPERATU Z INTERNETU

Wielką korzyścią wynikającą z przekształcenia zasobu geodezyjnego i kartograficznego do postaci cyfrowej jest automatyzacja docierania do operatów. Nie tylko eliminowana jest konieczność przeszukiwania wielu regałów i półek, ale powstaje również możliwość udostępniania dokumentów przez internet.

PIOTR GŁĘBICKI
WALDEMAR IZDEBSKI

Przejdźcie od zasobu prowadzonego w postaci klasycznej do zasobu numerycznego (przechowywanego w bazach danych) jest obecnie jednym z podstawowych zadań powiatowych ośrodków dokumentacji geodezyjnej i kartograficznej (PODGiK-ów). W obowiązujących przepisach geodezyjnych [1, 2] trudno znaleźć motywację do podjęcia takiego wyzwania, więc w większości są to działania lokalne, na miarę skromnych możliwości technicznych, finansowych i kadrowych.

Przez przekształcenie zasobu geodezyjnego i kartograficznego do postaci numerycznej powszechnie rozumie się działania związane z zasobami mapowymi (mapa zasadnicza). Każdy, kto widział ośrodek dokumentacji, wie jednak, że oprócz map są tam dziesiątki tysięcy zgromadzonych przez lata papierowych operatów. Pozostawienie ich w formie tradycyjnej nie daje możliwości pełnego rozwoju PODGiK, dlatego po zinformowaniu zasobów mapowych, naturalnym krokiem jest podjęcie działań w kierunku informatyzacji operatów.

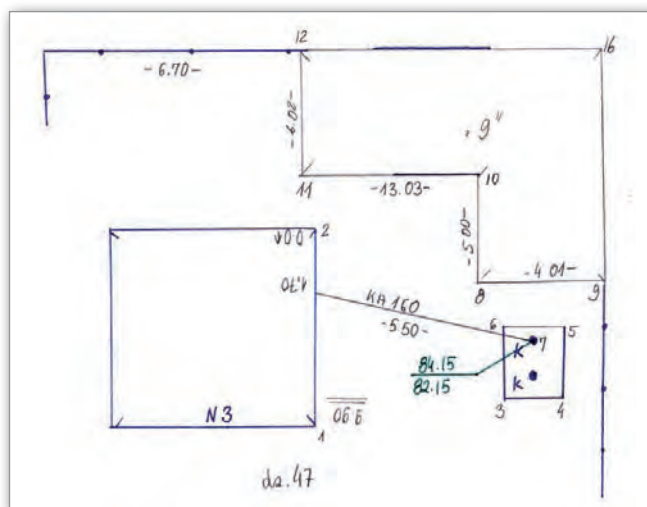
W powiecie wołomińskim skanowanie operatów rozpoczęto w styczniu 2005 roku, kiedy już od kilku lat wszystkie zasoby mapowe funkcjonowały w postaci numerycznej. Podjęcie działań spowodowane było wzrostem zainteresowania klientów PODGiK skanowanymi mate-

riałami. I nie tyle chodzi tu o samą informację, co raczej o czas i sposób, w jakim będzie ona dostarczona. Tradycyjne wyszukiwanie dokumentów, kopiowanie etc. nie może się równać z szybkością i łatwością przetwarzania zbiorów komputerowych. Obecnie zakres i możliwości wykorzystywania danych numerycznych znacznie się poszerzyły i być może w całkiem nieodległej przyszłości rzeczywistość wizyta wirtualna złożona o dowolnej porze i z dowolnego miejsca w kraju.

Oczywiście kilka czy kilkanaście zeskanowanych operatów nie ma większego znaczenia, ale kilka tysięcy rocznie i determinacja w działaniu, powoli zaczynają przynosić zamierzone efekty. Do dzisiaj metodą tą zarchiwizowano ponad 26 tys. kompletnych operatów z lat 2004-07 i blisko 30 tys. stron dokumentów, głównie z lat 1999-2003. Podsumowując niespełna trzyletni okres działalności w tym zakresie, zarchiwizowane materiały można ocenić na 10-15% objętości zasobu geodezyjno-kartograficznego w Wołominie.

• CO SKANOWAĆ

Bardzo ważną kwestią, którą niewątpliwie należy rozważyć przed przystą-



Rys. 1. Przykładowy szkic polowy zeskanowany w rozdzielczości 200 dpi, 256 kolorów

pieniem do archiwizacji, jest określenie merytorycznego zakresu pozyskiwanych danych. Przekazywana do ośrodków dokumentacji składa się z materiałów o dość zróżnicowanej przydatności w procesie dalszego przetwarzania. W sposób naturalny nasuwa się tu skojarzenie z metodyką przewidzianą w stosownym rozporządzeniu [3]. W takim przypadku do skanowania można by podchodzić jak do brakowania zasobu, stosując odpowiednie kategorie archiwalne – takie były zresztą pierwotne założenia archiwizacji w Wołominie. Rozwiązanie to ma jednak swoje bardzo istotne wady, o których należy tu wspomnieć:

- „Brakowanie dokumentacji” wymaga wykwalifikowanych pracowników o dużym doświadczeniu zawodowym z zakresu geodezji przy jednoczesnej znajomości nowoczesnego sprzętu elektronicznego.

- Skrupulatna analiza materiałów wymaga czasu (czynności związane ze skanowaniem przeciętnego operatu 10-20 stron zajmują ok. 10 min), a i tak ist-



Rys. 2. Zestaw do fotografowania dokumentacji zastosowany w Powiatowym Ośrodku Dokumentacji Geodezyjnej i Kartograficznej w Wołominie

nieje wysokie ryzyko, że pewne istotne strony zostaną pominięte.

● Czynniki ludzkie – wątpliwość odbiorcy informacji, czy aby na pewno wszystko, co ważne w operacie, zostało zeskanowane i odszukanie oryginalnej dokumentacji w celu rozstrzygnięcia ewentualnego sporu zaprzecza istocie numerycznego zasobu.

Mając powyższe na względzie, obrano zupełnie odmienną strategię polegającą na skanowaniu całych operatów. Niewątpliwie unikamy w ten sposób najcięższego zarzutu, że tak zarchiwizowana dokumentacja jest niekompletna, a w razie potrzeby zawsze można „oczyścić” ją ze zbędnej treści.

Na początku były obawy, że pliki tworzone z całych operatów będą niepotrzebnie obciążały system komputerowy. W 2005 r. dostępne na rynku maksymalne pojemności twardych dysków oscyływały w granicach 100-160 GB, a objętość zasobu szacowaliśmy na przynajmniej 10-krotnie więcej. Potwierdziło się jednak przyjęte wówczas założenie, że po-

stęp w dziedzinie informatyki zdecydowanie przewyższy tempo gromadzenia danych. Zeskanowane dotąd dokumenty zajmują nieco ponad 140 GB.

● JAK SKANOWAĆ

W dużej mierze sposób skanowania zależy od jakości i rodzaju dokumentów. Na przykładzie zasobu wołomińskiego, a jest on typowym przedstawicielem tej grupy archiwów, całą posiadaną dokumentację podzielić można na trzy główne grupy:

1. dokumentacja komputerowa, powstająca w wyniku zastosowania technik komputerowych, tj. drukowana lub kreślona na ploterze,
2. dokumentacja tworzona ręcznie, przeważnie pożąłką, często na niesformatowanym papierze z dużą liczbą odręcznych zapisów,
3. dokumentacja oprawiona introligatorsko (szyta lub klejona).

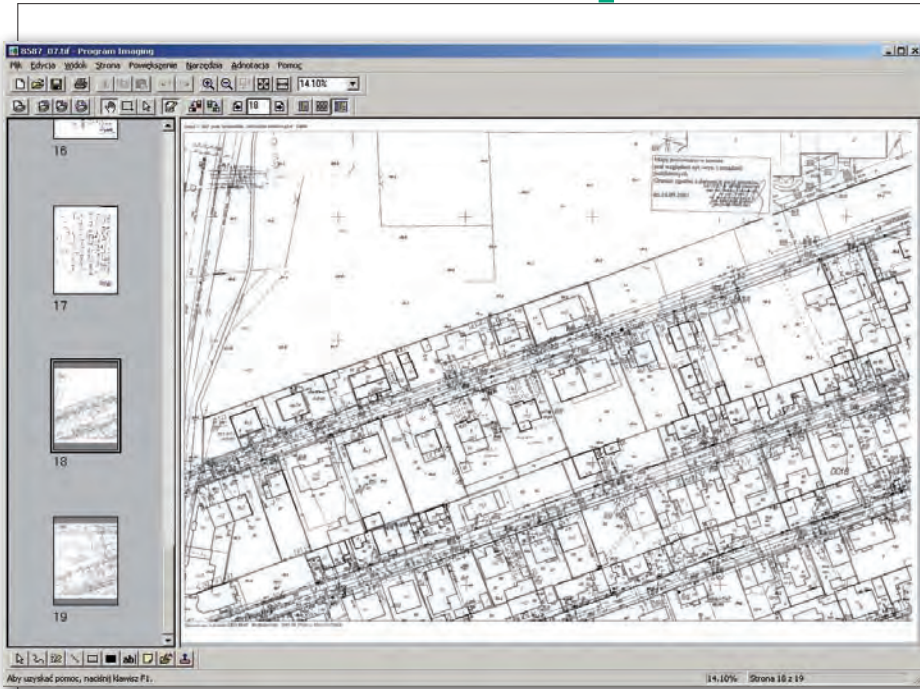
Do pierwszej grupy dokumentacji należą operaty komputerowane od połowy lat 90. ubiegłego wieku – rysunek czytelny, papier formatu A4, idealnie, jeśli nie

jest nadmiernie pozszywany. Taka dokumentacja doskonale nadaje się do skanowania z automatycznego podajnika, a zacięcia papieru zdarzają się sporadycznie. W tym miejscu warto powiedzieć parę słów na temat skanerów, bo to od ich pracy zależą efekty powziętego zamierzenia. Absolutnie należy wykluczyć stosowanie płaskich skanerów do zdjęć wykonanych aparatem fotograficznym! Zastosowana w nich technologia ma za zadanie dorównać rozdzielczości zdjęć, co powoduje, że skanowanie pojedynczej kartki może trwać stosunkowo długo. Utworzenie zestronnicowanego (połączonego z pojedynczych obrazów w jeden spójny dokument) operatu zarchiwizowanego przy użyciu takich skanerów jest mało praktyczne.

Z naszych doświadczeń polecić możemy niedrogie, biurowe skanery do dokumentów, z automatycznym podajnikiem papieru na około 50 kartek (rzadko operaty liczą więcej) i dwiema głowicami skanującymi umożliwiającymi skanowanie obydwu stron dokumentu przy jednym przebiegu papieru. Wydajność na poziomie 50 obrazów (25 kartek) na minutę jest do tych celów w zupełności wystarczająca. Dostępne ustawienia rozdzielczości i parametrów skanowanego obrazu są zoptymalizowane pod kątem dokumentów. Z doświadczeń wynika, że najlepszym kompromisem między jakością uzyskanego obrazu a objętością jest skanowanie w rozdzielczości 200 dpi przy 256 kolorach (rys. 1). Formatka A4 przy takich ustawieniach zajmuje przeciętnie około 200-300 kilobajtów.

Druga grupa dokumentów jest trudniejsza do przetwarzania i wymaga większego zaangażowania od operatora. Konieczne jest tu przejście wszystkich skanowanych stron, ocena ich czytelności oraz wybranie odpowiednich technik skanowania dla poszczególnych stron dokumentacji. Spore utrudnienia w automatyzacji pracy sprawia podniszczony, niesformatowany papier.

Ostatnia grupa to dokumenty oprawione introligatorsko (szyte lub klejone). W ich przypadku użycie skanera jest mocno ograniczone, a czasem wręcz niemożliwe. Do tej grupy dokumentacji wykorzystujemy metody stosowane w bibliotekach do archiwizacji starodruków. Na rys. 2 przedstawiony został zestaw stosowany przez nas do fotografowania dokumentacji składający się m.in. ze statywu, stolika oraz zestawu lamp umożliwiających równomierne oświetlenie fotografowanego przedmiotu. Cyfrowy aparat fotograficzny posiada odchylany



Rys. 3. Przykład operatu udostępnionego w GEO-MAP z wykorzystaniem programu Imaging

obrotowy wyświetlacz oraz wyzwalacz z pilota.

Poczynione w tej dziedzinie obserwacje wskazują, że najlepiej sprawdzają się aparaty (oczywiście cyfrowe) posiadające następujące cechy:

- matryca ok. 3,5 megapiksela,
- zoom optyczny maksymalnie 4-krotny,
- obrotowy wyświetlacz i wyzwalacz z pilota (rys. 2) – doskonałe opcje przy ro-

bieniu zdjęć do pamięci aparatu (np. bezpośrednio w pomieszczeniu archiwum),

- pełne sterowanie aparatem przez komputer.

● ZABEZPIECZENIE DANYCH

Nawet przy najlepszej polityce bezpieczeństwa nie można wykluczyć uszkodzenia zgromadzonych zbiorów komputerowych spowodowanych niewłaściwą obsługą czy też awarią sprzętu. Chcąc

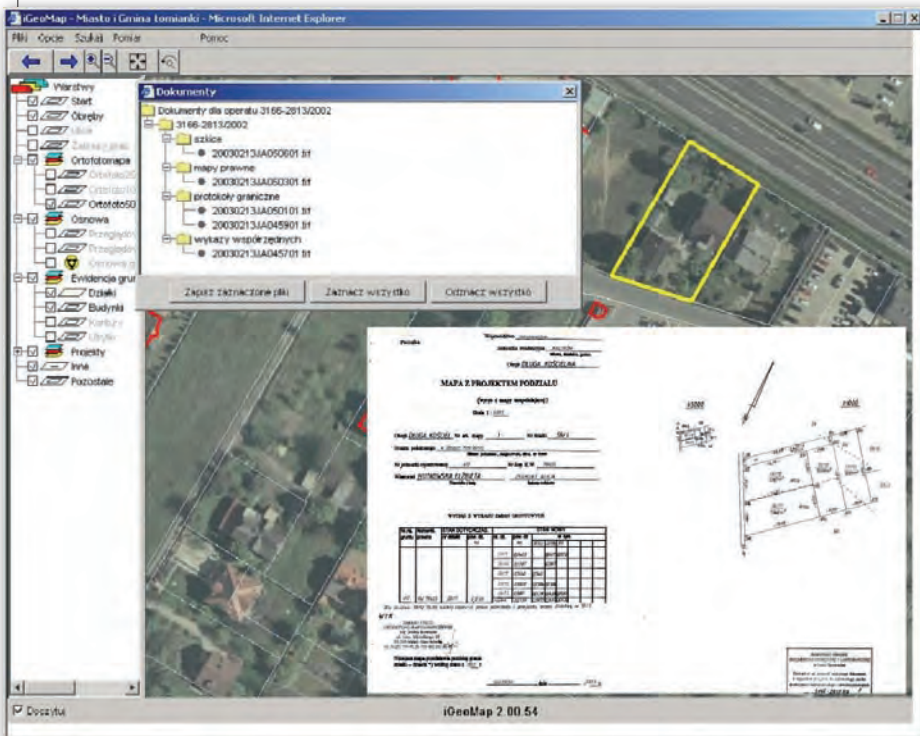
uniknąć utraty posiadanych zbiorów, przechowujemy je w czterech różnych lokalizacjach. Podstawowy zbiór gromadzony jest na głównym serwerze POD-GiK, a replika znajduje się na dysku komputerowym umieszczonym poza pomieszczeniem serwerowni. Dodatkowo bieżące skany gromadzone są na jednym ze stanowisk skanujących do momentu, kiedy uzbiera się zestaw danych o wielkości zbliżonej do pojemności płyty DVD. Wówczas wykonywana jest archiwizacja na płycie, a jej obraz przechowywany jest dodatkowo na dysku komputerowym. Na serwerze operaty gromadzone są w strukturze katalogów odpowiadających rocznikom, a ich zawartość stanowią zestronnicowane pliki poszczególnych operatorów, w kolejności narastającej liczby dziennika. Natomiast na płytach archiwizowane są operaty w kolejności przyjmowania do zasobu (skanowania).

● UDOSTĘPNIANIE DANYCH

Jak już wielokrotnie podnoszono, zeskanowanie zasobu nie jest celem samym w sobie, lecz przekłada się na konkretne korzyści w funkcjonowaniu PODGiK. Należy do nich zaliczyć automatyzację w docieraniu do operatorów bez konieczności przeszukiwania wielu regałów i półek. Zeskanowane operaty podłączane są do zakresów prac geodezyjnych i w każdej chwili mogą być wyświetlone w systemie GEO-MAP, który jest wykorzystywany do prowadzenia zasobu.

Technologia udostępniania danych w internecie oparta na oprogramowaniu iGeoMap daje podobne możliwości. Dodatkowo skutkuje komfortem docierania do operatorów bez wychodzenia z domu. Wymagane jest jedynie internetowe zgłoszenie pracy geodezyjnej.

PIOTR GŁĘBICKI
PODGiK w Wołominie,
WALDEMAR IZDEBSKI
GEO-SYSTEM Sp. z o.o.



Rys. 4. Przykład udostępniania zeskanowanych operatorów w serwisie internetowym

Literatura:

- [1] Ustawa Prawo geodezyjne i kartograficzne, DzU z 1989 r. nr 30, poz. 163;
- [2] Instrukcja Techniczna O-4 Zasady prowadzenia państwowego zasobu geodezyjnego i kartograficznego, zarządzenie prezesa GUGiK z 10 lipca 1987 r.;
- [3] Rozporządzenie ministra rozwoju regionalnego i budownictwa w sprawie klasyfikowania, kwalifikowania i porządkowania materiałów wyłączanych z państwowego zasobu geodezyjnego i kartograficznego, DzU z 2001 r. nr 74, poz. 796;
- [4] Izdebski W., 2007: Rola internetu w nowoczesnym ośrodku dokumentacji geodezyjnej i kartograficznej. Materiały Konferencji Narzędzia Informatyczne Geodezji i Kartografii. Nowe regulacje geoinformacyjne. Warszawa, 30 sierpnia 2007 r.