



Problemy digitalizacji na podstawie doświadczeń z realizacji projektu SUW

mgr Katarzyna Skotnicka
mgr Helena Gałek

Początki

Digitalizacja zbiorów do Biblioteki Cyfrowej PK.

Digitalizacja zbiorów w ramach projektu *„Zintegrowany System Wymiany Wiedzy i Udostępniania Akademickich Publikacji z Zakresu Nauk Technicznych”*.

Zbiory objęte digitalizacją:

- zbiory specjalne
- rozprawy doktorskie

Selekcja zbiorów specjalnych

- zbiory, do których wygasły prawa autorskie
(70 lat od śmierci autora)
- zbiory, które nie są wypożyczane na zewnątrz
(możliwość skorzystania w czytelni, tylko za zgodą kierownika Oddziału Zbiorów Zwartych lub Dyrekcji)
- zbiory, które wymagają szczególnego traktowania
(wartość i stan zachowania)
- zbiory o formacie większym niż A3

Selekcja rozpraw doktorskich

Rozprawy doktorskie nieobjęte ZARZĄDZENIEM nr 12
Rektora PK, z dnia 9 marca 2006 r.

(w sprawie gromadzenia, opracowania i udostępniania przez BG PK
rozpraw doktorskich obronionych na Politechnice Krakowskiej).

§ 1 pkt. 1. Doktoranci składający przed obroną ukończone rozprawy doktorskie **mają obowiązek przekazania** jednego egzemplarza wersji drukowanej i **jednego egzemplarza wersji elektronicznej rozprawy doktorskiej do Biblioteki Głównej PK** do Oddziału Obsługi Zbiorów Zwartych i Specjalnych.

§ 1 pkt. 4. Egzemplarz wersji elektronicznej będzie **publicznie dostępny** w zasobach Biblioteki Cyfrowej PK (BCPK) po obronie.

Proces digitalizacji

- skanowanie zbiorów w Bibliotece w Oddziale Informatyzacji - Repozytorium (max. format A3)

- skanowanie zbiorów przez profesjonalne firmy zewnętrzne (zbiory wielkoformatowe)



Biblioteka Politechniki Krakowskiej



Skanowanie zbiorów wielkoformatowych

- przekazanie zbiorów firmie zewnętrznej
(transport, ochrona)
- skanowanie zbiorów
- dostarczenie obiektów cyfrowych do Biblioteki PK
(format TIFF)
- obróbka graficzna - pracownicy Biblioteki PK
(kadrowanie, prostowanie, sklejanie, retusz)
- zapis do formatu prezencyjnego PDF

Skanowanie rozpraw doktorskich

- skanowanie zbiorów w Oddziale Informatyzacji -
Repozytorium
(skanery małegoformatowe)
- zapis obiektów cyfrowych
(format TIFF)
- obróbka graficzna
(kadrowanie, prostowanie, sklejanie, retusz)
- zapis do formatu prezencyjnego PDF

Problemy obiektów wielkoformatowych

- nieprofesjonalne podejście firm zewnętrznych:
 - pominięte strony
 - zaburzona kolejność stron
 - niewyraźne skany
 - fragmentaryczne skanowanie stron
- problemy z obróbką plików macierzystych:
 - sprzęt komputerowy o niskiej wydajności
 - brak profesjonalnego oprogramowania do obróbki graficznej

Protokół

Protokół zdawczy usługi digitalizacji dla Biblioteki Politechniki Krakowskiej

Lp	Sygnatura	Liczba stron	odbior prac doktorskich: 14.12.2011 UWAGI V	odbior prac doktorskich: 21.03.2012 (nasze prace doktorskie) 04.04.2012 (skany prac doktorskich) UWAGI VI
1	1	76	do przyjęcia	
2	16	116	Strona 4a zeskanowana kilka razy i nigdy dobrze, należy zeskanować rozłożoną; str. 64 jest dwa razy, str. 80, 94, 108 niewyraźne, część stron nie jest obcięta	Strona 4a zeskanowana kilka razy i nigdy dobrze, należy zeskanować rozłożoną; str. 64 jest dwa razy, str. 80, 94, 108 niewyraźne, część stron źle docięta, strony zeskanowane w częściach, krzywo zeskanowane - ponowne skanowanie
3	19 zał. 1,2	271	Strony nie są docięte, widać ciemne tło; str. 93 cień palca w prawym, dolnym rogu; plik 230 widoczny palec, plik 246 niewyraźny, są dwie części, które powinny być w dwóch podfolderach, podcinane fragmenty tekstu, wykresów od strony grzbietu	Strony źle docięte, widać ciemne tło; str. 93 cień palca w prawym, dolnym rogu; plik 230 widoczny palec, zawinięta strona, plik 246 niewyraźny, są dwie części, które powinny być w dwóch podfolderach, podcinane fragmenty tekstu, wykresów od strony grzbietu; Oba woluminy powinny być zamieszczone w dwóch osobnych podfolderach, a są umieszczone w jednym - ponowne skanowanie, cz. 19/1 nie zeskanowana
4	20	253	Strona tytułowa 2 razy, plik 143 i 144 odwrotna kolejność; plik 031 krzywo i niewyraźnie, strony nie są docięte	plik 143 i 144 odwrotna kolejność; plik 031 krzywo i niewyraźnie, strony źle docięte, widać że praca została wyjęta z oprawy mimo to krzywo zeskanowana, np. 52, 53 - część stron należy ponownie zeskanować, większość ponownie/właściwie podcinać
5	22	79	strony nie są docięte	strony źle docięte, praca wyjęta z oprawy, a mimo to krzywo zeskanowana, różne tła - ponowne skanowanie
6	25	173	strony nie są docięte	strony źle docięte, ucięty tekst np. str. 11, źle dobrane parametry skanowania, krzywe strony, odbicia światła/lampy - ponowne skanowanie

Problemy wewnętrzne

- problemy techniczne:
 - skanery
(problemy z oprawą zbiorów, prześwietlanie skanów)
 - sprzęt komputerowy
(przetwarzanie wsadowe plików)
 - brak profesjonalnego oprogramowania do obróbki graficznej
 - brak systemu zarządzania plikami
 - brak szkoleń dla pracowników Biblioteki PK


Problemy wewnętrzne

- nietypowe zbiory:
 - niestandardowe formaty, rozkładane karty
 - oprawy utrudniające digitalizację
 - niewyraźny druk
 - zabrudzone karty
 - kalki, przebitki
 - tekst schowany w oprawie

niestandardowe formaty, rozkładane karty



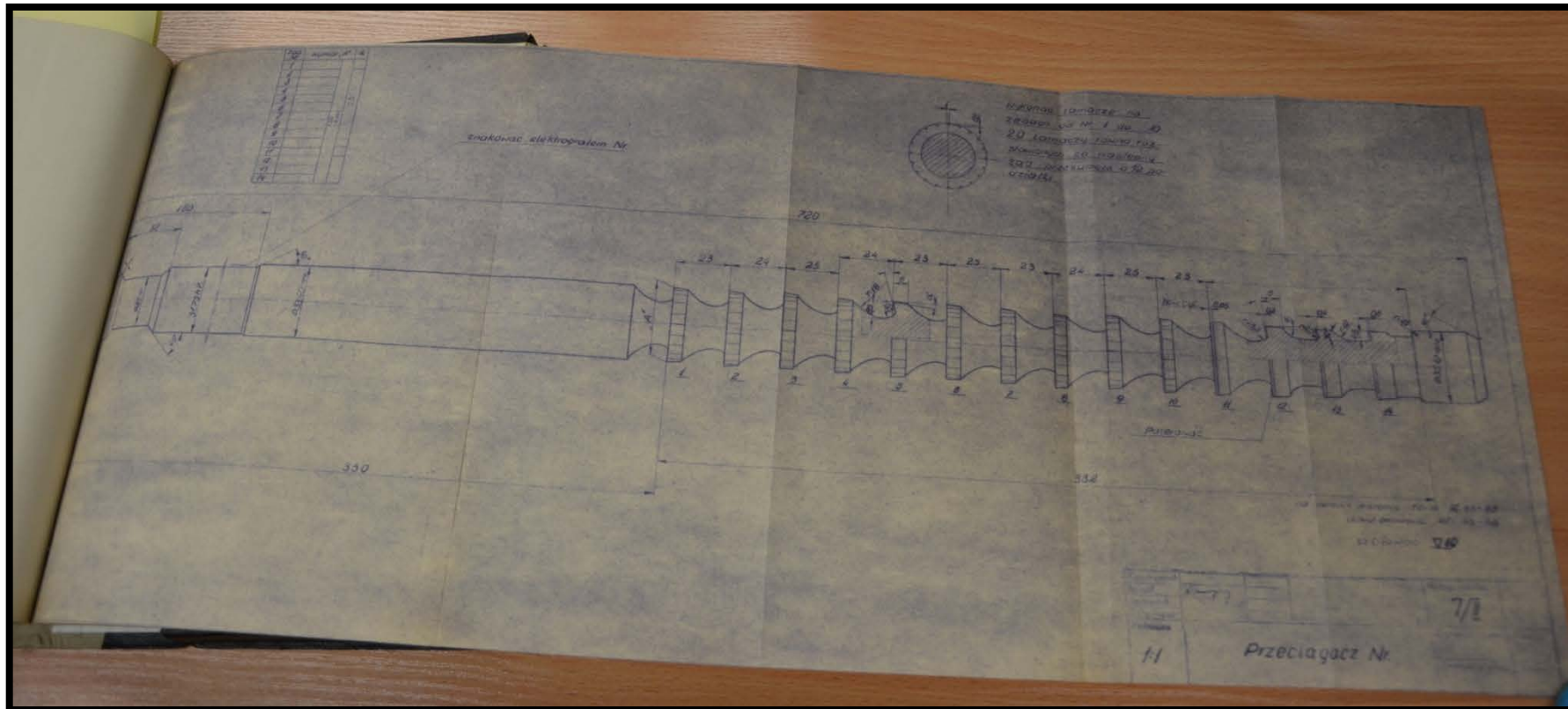
niestandardowe formaty, rozkładane karty



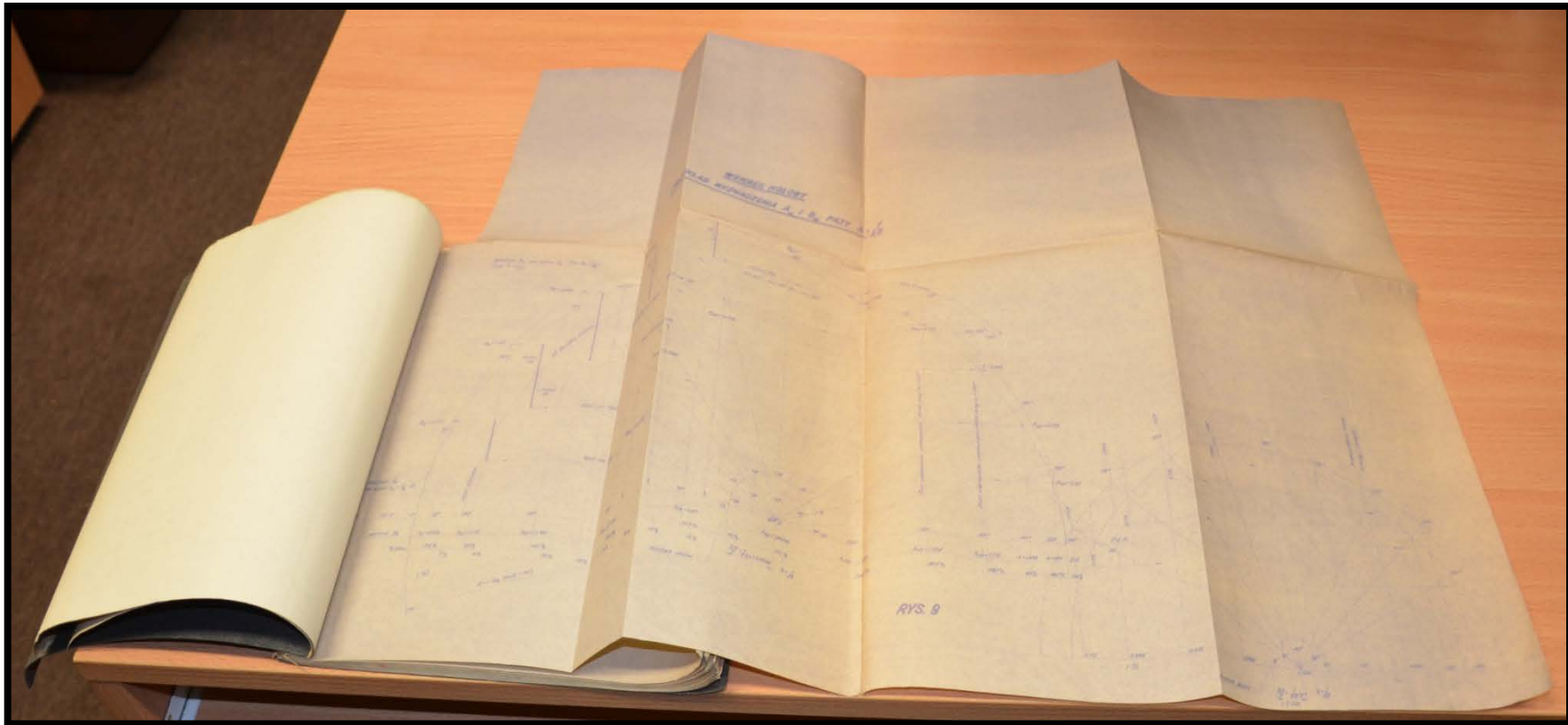
SZEROKOŚĆ ROZWARCIA RYS UKOŚNYCH a_f [mm] BELKA DK-2b

SIEĆ 20	①				②				③				④				⑤				⑥				⑦				⑧				HARTOŚĆ ŚREDNIA				a_{max}		
	od zewnątrz	$\frac{1}{4} H$	$\frac{1}{2} H$	$\frac{3}{4} H$	max	od zewnątrz	$\frac{1}{4} H$	$\frac{1}{2} H$	$\frac{3}{4} H$	max	od zewnątrz	$\frac{1}{4} H$	$\frac{1}{2} H$	$\frac{3}{4} H$	max	od zewnątrz	$\frac{1}{4} H$	$\frac{1}{2} H$	$\frac{3}{4} H$	max	od zewnątrz	$\frac{1}{4} H$	$\frac{1}{2} H$	$\frac{3}{4} H$	max	od zewnątrz	$\frac{1}{4} H$	$\frac{1}{2} H$	$\frac{3}{4} H$	max	od zewnątrz	$\frac{1}{4} H$	$\frac{1}{2} H$	$\frac{3}{4} H$	max				
4000					0.02	0.02				0.02	0.02	0.01			0.02	0.03																			0.02	0.03		0.02	1.50
5000	0.05	0.05	0.03		0.05	0.04	0.03			0.04	0.04	0.02			0.04	0.05																		0.04	0.03	0.03	0.04	1.25	
6000	0.05	0.07	0.05		0.07	0.04	0.03	0.05		0.05	0.05	0.02	0.02		0.05	0.05	0.03				0.05	0.04	0.01										0.04	0.03	0.03	0.05	1.40		
7000	0.05	0.10	0.07		0.10	0.05	0.05	0.06		0.06	0.05	0.10	0.10		0.10	0.10	0.10	0.01			0.10	0.05	0.08	0.02			0.01	0.01	0.01			0.01	0.01	0.03	0.03	0.03	1.40		
8000	0.05	0.10	0.10		0.10	0.05	0.05	0.10	0.01	0.10	0.05	0.15	0.15		0.15	0.10	0.10	0.01			0.15	0.05	0.10	0.03			0.01	0.01	0.03	0.03		0.03	0.05	0.05	0.05	1.50			
9000	0.10	0.10	0.10	0.02	0.15	0.05	0.05	0.10	0.02	0.15	0.10	0.20	0.25	0.01	0.25	0.10	0.10	0.03			0.10	0.05	0.25	0.05	0.02	0.25	0.05	0.10			0.10	0.02	0.02	0.03	0.03	1.25			
10000	0.10	0.15	0.20	0.05	0.25	0.05	0.05	0.10	0.05	0.25	0.20	0.40	0.45	0.05	0.50	0.10	0.10	0.03			0.20	0.10	0.20	0.10	0.03	0.25	0.05	0.10			0.10	0.05	0.20	0.03	0.20	1.75			
4000	0.05	0.01			0.05	0.02				0.02	0.02				0.02	0.01					0.01													0.02	0.01		0.02	2.50	
5000	0.05	0.04	0.04		0.05	0.03	0.02			0.03	0.03				0.03	0.02					0.02													0.03	0.03	0.01	0.03	1.67	
6000	0.05	0.04	0.02		0.05	0.02	0.04			0.04	0.05	0.03			0.05	0.03					0.03	0.02											0.04	0.03	0.02	0.04	1.25		
7000	0.07	0.05	0.05		0.07	0.04	0.05	0.05		0.05	0.05	0.05	0.05		0.05	0.05	0.03				0.05	0.05	0.03										0.04	0.04	0.05	0.05	1.40		
8000	0.05	0.05	0.07		0.10	0.05	0.05	0.10		0.10	0.05	0.10	0.10		0.10	0.05	0.03				0.05	0.03											0.05	0.05	0.05	0.05	1.25		
9000	0.07	0.10	0.10		0.10	0.05	0.05	0.10	0.02	0.15	0.05	0.20	0.20		0.25	0.05	0.02				0.05	0.03											0.05	0.05	0.05	0.05	1.25		
10000	0.10	0.10	0.12	0.03	0.12	0.05	0.05	0.10	0.07	0.20	0.05	0.30	0.40		0.30	0.05	0.02				0.05													0.20	0.25	0.50	0.50	1.75	

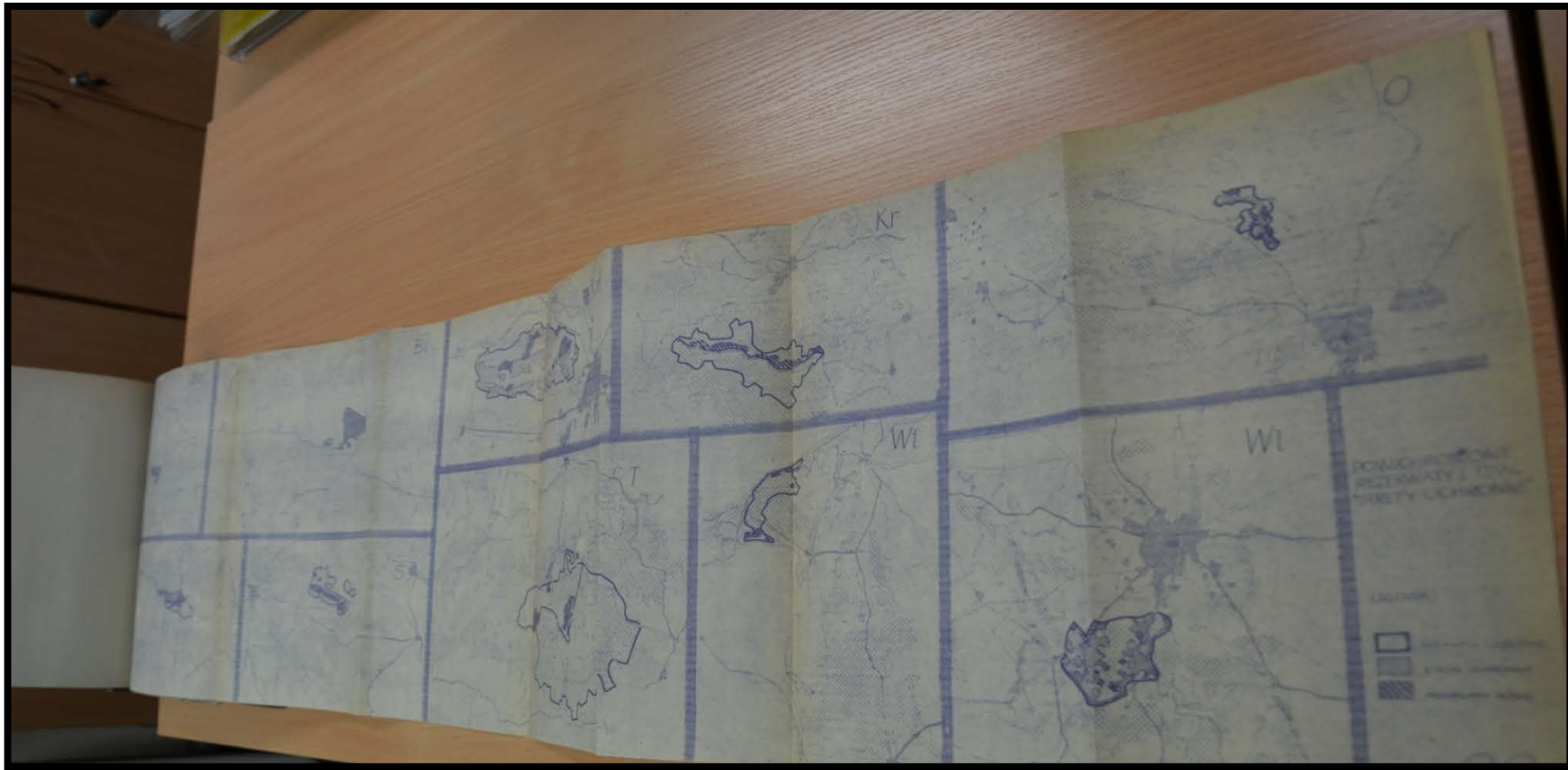
niestandardowe formaty, rozkładane karty



niestandardowe formaty, rozkładane karty



niestandardowe formaty, rozkładane karty



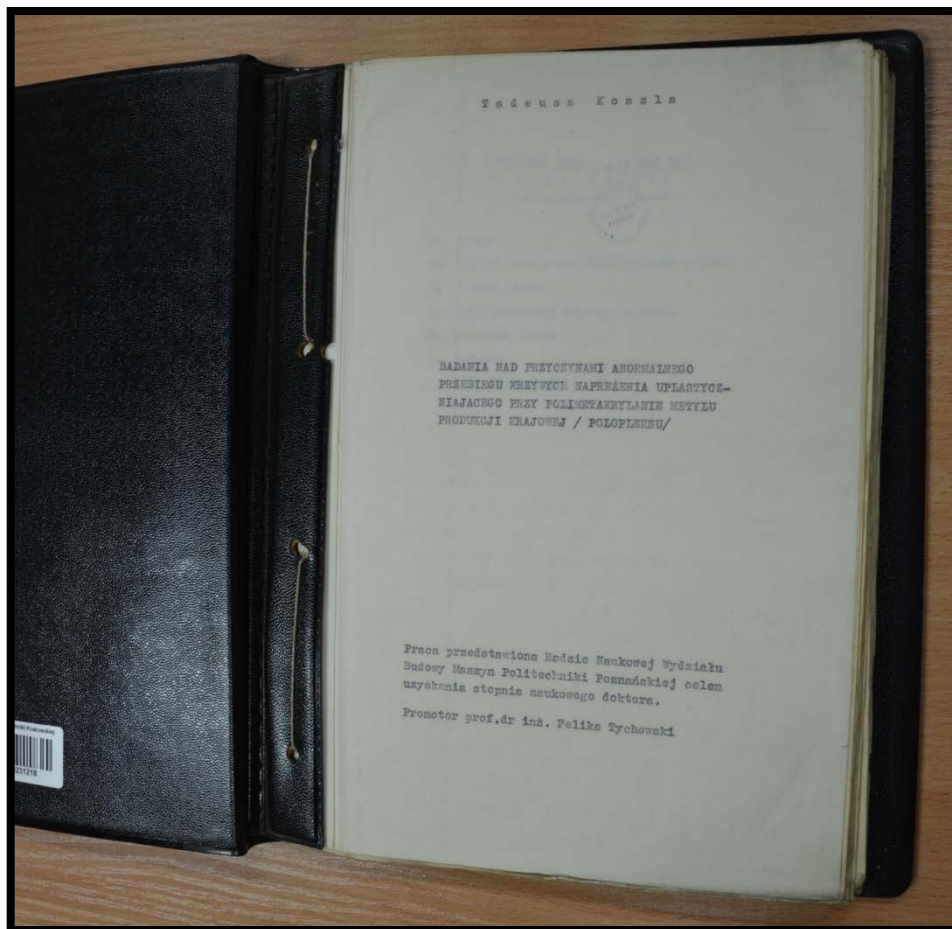
oprawy utrudniające digitalizację



oprawy utrudniające digitalizację



oprawy utrudniające digitalizację



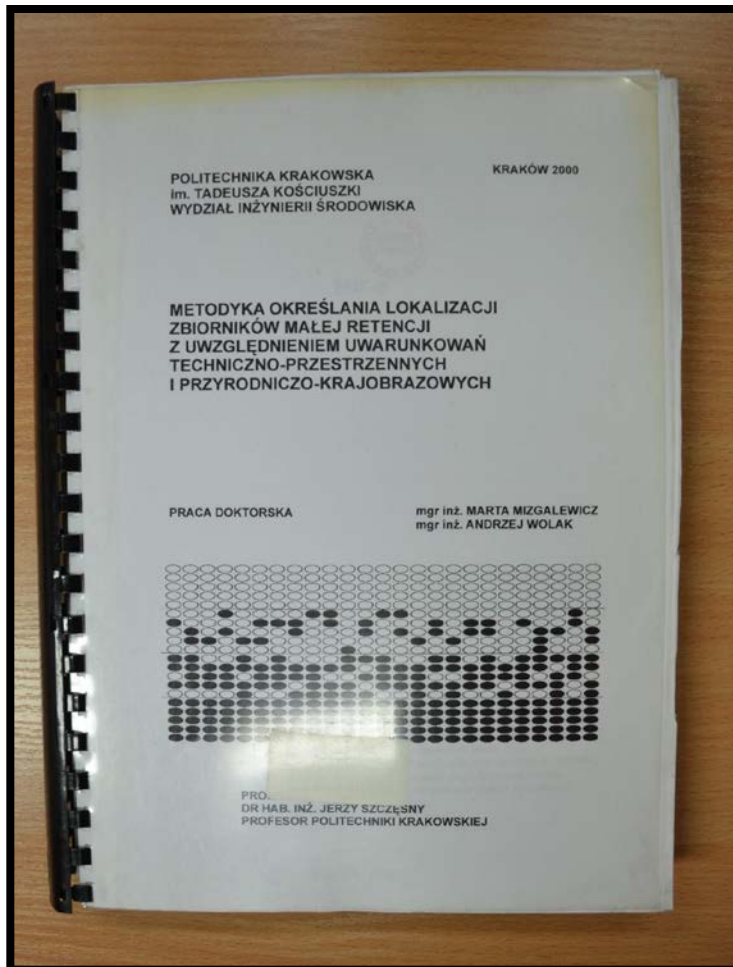
oprawy utrudniające digitalizację



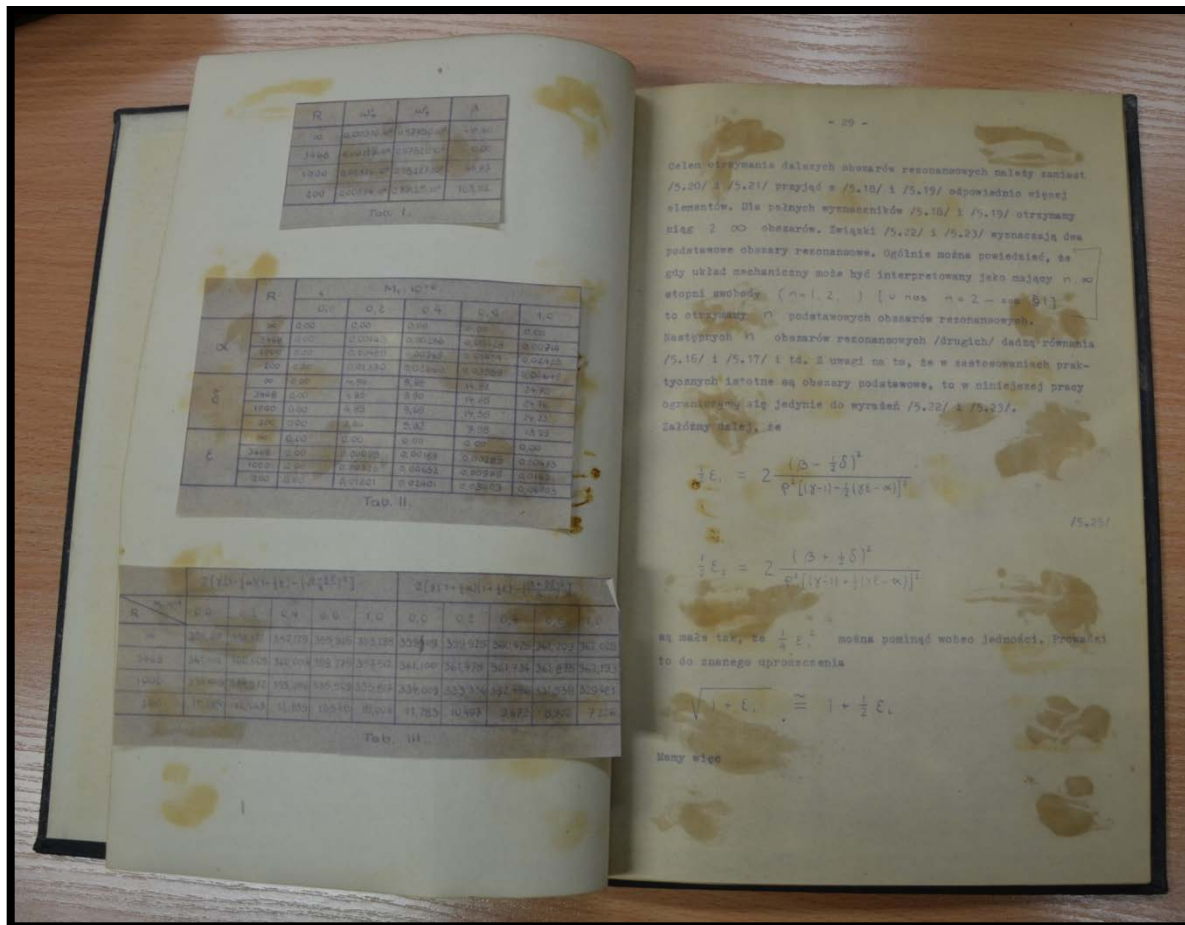
oprawy utrudniające digitalizację



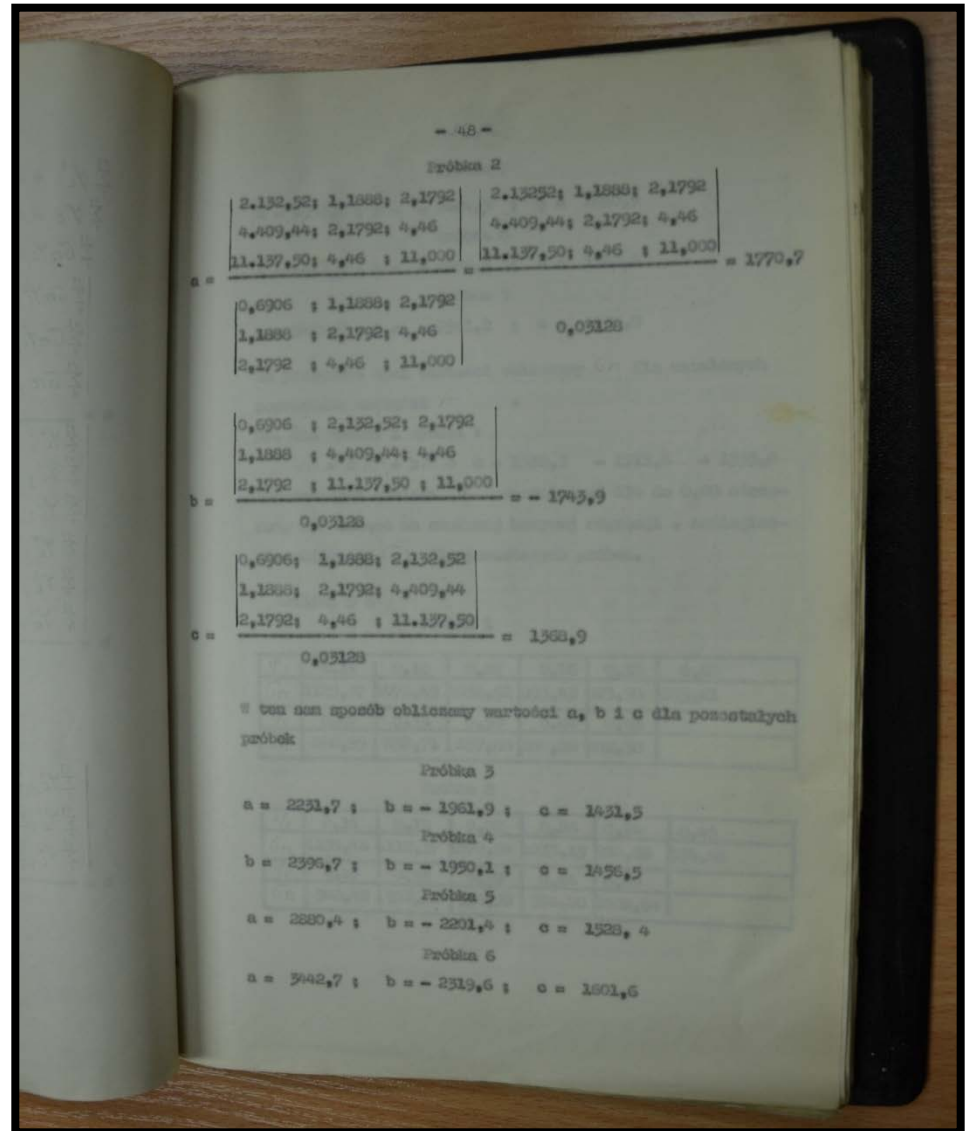
oprawy utrudniające digitalizację



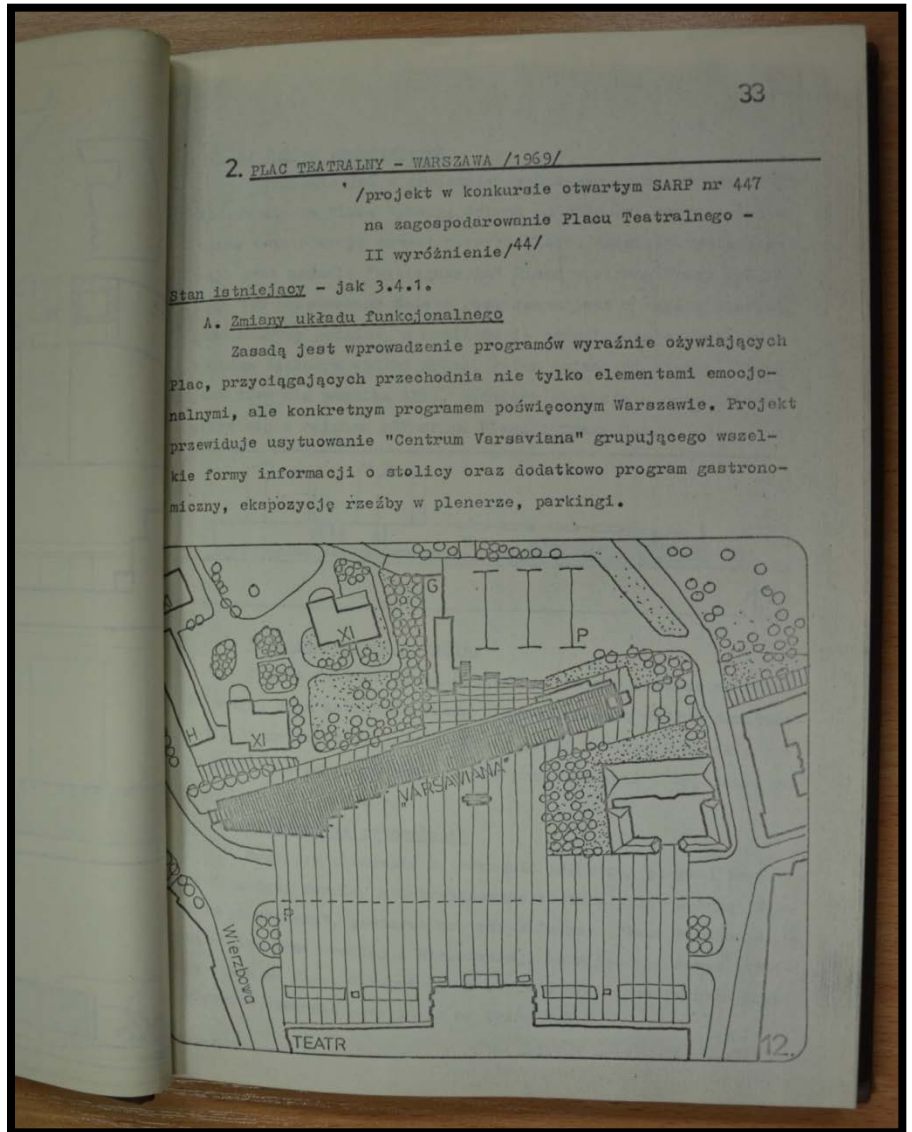
zabrudzone karty, niewyraźny druk



kalki, przebitki



tekst schowany w oprawie



Wnioski

- zwiększenie nakładów finansowych na sprzęt:
 - profesjonalne skanery
 - wydajny sprzęt komputerowy
 - profesjonalne oprogramowanie graficzne
- szkolenia pracowników
- system zarządzania plikami
- długoterminowa archiwizacja plików

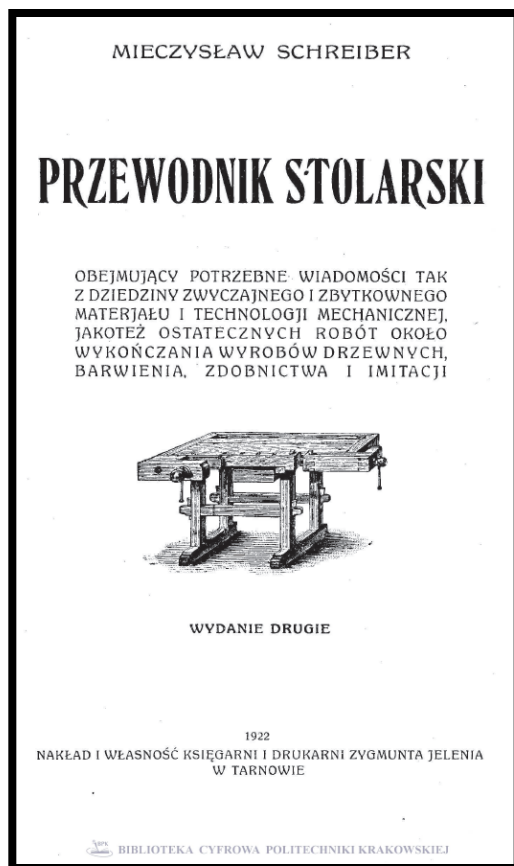
Profity BPK

- zabezpieczenie zagrożonych zniszczeniem lub dostępnych w ograniczonym zakresie zbiorów
- udostępnienie wersji cyfrowych nieograniczonej grupie odbiorców
(umieszczenie w Repozytorium, przeszukiwanie przez Google Scholar, nieograniczona, jednorazowa ilość czytających użytkowników)
- wsparcie w poszukiwaniu literatury z zakresu nauk technicznych

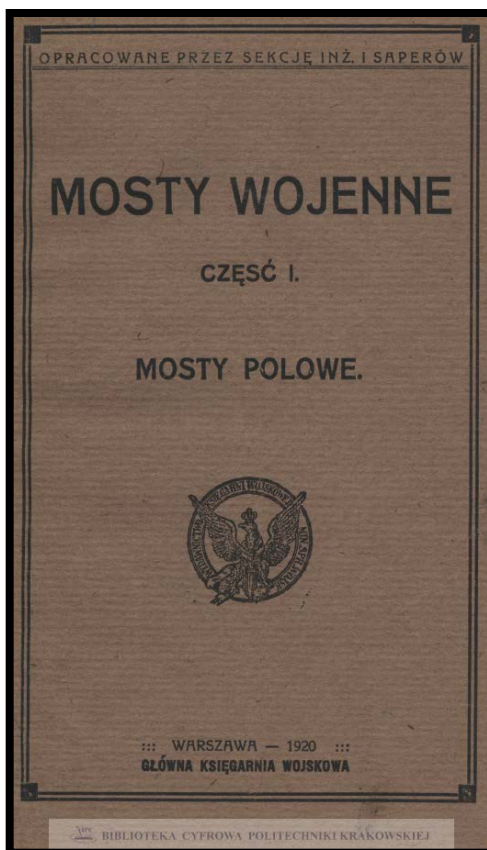
Statystyka zbiorów specjalnych

język publikacji	ilość pobrań				
	polski	niemiecki	łacina	inne	
2045-2852	3	0	0	0	3
1023-1997	14	0	1	0	15
501-995	23	11	0	2	36
155-496	12	87	0	34	133
3-49	8	2	0	1	11
	60	100	1	37	198

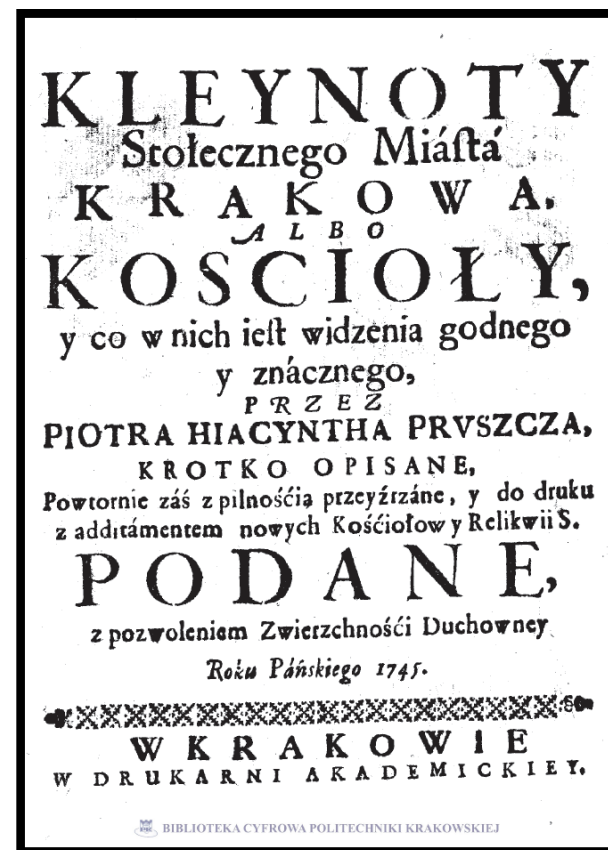
Najczęściej pobierane zbiory specjalne



2 852



2 151



2 045

Statystyka rozpraw doktorskich

rodzaj dostępu \ rodzaj licencji	Licencje Creative Commons	Licencja PK	Licencja PK. Brak możliwości edycji i druku	Licencja w ramach tzw. dozwolonego użytku (art. 28 u. o pr. a.)	
Zasób dostępny dla wszystkich	9	117	77	0	203
Zasób dostępny z komputerów BPK	0	22	46	1054	1122
Zasób dostępny dla załogowanych użytkowników oraz z komputerów w domenie PK	0	23	212	0	235
	9	162	335	1054	1560

Najczęściej pobierane rozprawy doktorskie

Warunki i możliwości usuwania fosforanów i fosforu ogólnego ze ścieków przemysłowych / Lucyna Przywara

(2006, Licencja PK, Zasób dostępny dla wszystkich), **Liczba pobrań: 47 302**

Komputerowa analiza obrazu RTG zatok czołowych jako podstawa identyfikacji osób / Dariusz Karpisz

(2008, Licencja PK, Zasób dostępny dla wszystkich), **Liczba pobrań: 24 786**

Sprawność i toksyczność spalin silnika z zapłonem iskrowym zasilanego ubogą mieszanką propanu-butanu z powietrzem i wyposażonego w reaktor magazynujący tlenki azotu / Michał Mareczek

(2007, Licencja PK, Zasób dostępny dla wszystkich) **Liczba pobrań: 24 763**



Problemy digitalizacji na podstawie doświadczeń z realizacji projektu SUW

Dziękuję za uwagę

mgr Katarzyna Skotnicka
mgr Helena Gałek