

WARSZAWA 2016	Biblioteka Narodowa	 BIBLIOTEKA NARODOWA
	Repozytorium cyfrowe BN – moduł statystyk	

Repozytorium Cyfrowe BN

Moduł Statystyk

WARSZAWA 2016	Biblioteka Narodowa	 BIBLIOTEKA NARODOWA
	Repozytorium cyfrowe BN – moduł statystyk	

System umożliwia analizowanie pracy workflow i gromadzonych zbiorów dzięki różnym statystykom jakie można wygenerować w zadanym okresie czasu. Dostępne są następujące statystyki:

1. Statystyki workflow

Statystyki tranzycji workflow

Data wygenerowania raportu Fri Jun 24 08:41:06 CEST 2016

Użytkownik	Planowane	Skanowane	Kontrola	Usunięte	Publikacje suma	Publikacje bez ograniczeń	Publikacje z ograniczeniami	Digitalizacje bez publikacji	Publikacje w przyszłości
admin	76	20	8	2	21	21	0	0	0
m.dmitruk	289	72	0	0	8	8	0	0	0
k.karpinski	21	1	1	0	0	0	0	0	0
a.leszynska	58	2	0	0	1	1	0	0	0
k.wilczak	1882	2	0	0	1	1	0	0	0
skanarzysla1	0	0	0	0	0	0	0	0	0
provider-test1	0	0	0	0	0	0	0	0	0
redaktor1	0	0	0	0	0	0	0	0	0
Suma	2326	97	9	2	31	31	0	0	0

Statystyki te pozwalają na przeprowadzenie analizy ile obiektów przeszło przez poszczególne etapy workflow. Analiza pozwala na zbadanie tych ilości zarówno w odniesieniu do pojedynczego użytkownika jak i sumarycznie. Statystyki dotyczą okresu czasu wskazanego przez użytkownika. Ujęte są następujące etapy workflow (oraz ich odmiany):

- planowanie – statystyki ilości obiektów które przeszły etap planowania, czyli ile obiektów zostało utworzonych
- skanowanie – statystyki ilości obiektów które przeszły przez etap skanowania
- kontrola – statystyki ilości obiektów które zostały skontrolowanych
- usunięte – statystyki ilości obiektów które zostały w danym okresie usunięte
- publikacje suma (publikowanie, jeżeli użyty jest tryb uproszczony tych statystyk) – wartości w tej kolumnie wskazują, ile obiektów w danym okresie zostało opublikowanych, czyli w rzeczywistości zakończyło ścieżkę workflow
- publikacje bez ograniczeń – wskazuje ile obiektów zostało opublikowanych z uprawnieniami pozwalającymi przeglądać pozycję bez ograniczeń, czyli przez dowolnego nawet niezalogowanego użytkownika. Kwalifikują się w tym przypadku wszystkie obiekty którym nadano dla grupy „anonymous” uprawnienie „Ładowanie cyfrowych obiektów” z celem „*”
- publikacje z ograniczeniami – wskazuje ile obiektów zostało opublikowanych z uprawnieniami ograniczającymi przeglądanie pozycji, np. możliwość przeglądania

WARSZAWA 2016	Biblioteka Narodowa	 BIBLIOTEKA NARODOWA
	Repozytorium cyfrowe BN – moduł statystyk	

tylko w czytelni. Kwalifikują się w tym przypadku wszystkie obiekty którym nie nadano dla grupy „anonymous” uprawnienia „Ładowanie cyfrowych obiektów” z celem „*”

- digitalizacja bez publikacji – statystyka wskazuje liczbę obiektów dla których na etapie publikowania wybrano opcję „nie publikuj”. Są to obiekty które w rzeczywistości nie zakończyły ścieżki workflow i czekają na opublikowanie.
- publikacje w przyszłości – statystyka wskazuje liczbę obiektów które zostały opublikowane w późniejszym terminie, tzn. dla których najpierw wybrano opcję „nie publikuj” a dopiero w przyszłości dokonano ich opublikowania.

Poprzez opublikowanie rozumie się przejście przez etap workflow „publikowanie obiektu”, niezależnie od tego jakie uprawnienia zostaną ustawione, czyli końcowo – niezależnie od tego w którym miejscu i dla kogo dany obiekt będzie możliwy do wyświetlenia.

2. Statystyki skanowania

Statystyki te mają na celu przede wszystkim pracę skanerzystów. W poszczególnych kolumnach są zawarte wartości dotyczące ilości wrzucanych do systemu plików. Zaznaczyć należy iż nie jest ważne czy użytkownik umieścił w systemie plik graficzny (skan) czy dowolnie inny rodzaj pliku (np. PDF). Brane jest natomiast pod uwagę czy na formacie wysyłania plików do systemu użytkownik oznaczył że są to pliki skanów/pliki struktury paginacyjnej – ponieważ tylko wówczas liczą się do tych statystyk. Statystyki pozwalają zarówno na analizę ilości umieszczonych plików przez wybranych użytkowników jak i analizę sumaryczną w wyznaczonym przez twórcę statystykę zakresie.

Poszczególne kolumny odzwierciedlają następujące wartości:

- całkowita ilość skanów – liczba plików umieszczonych w systemie, brane są pod uwagę tylko pliki oznaczone jako pliki skanów (struktury paginacyjnej)
- ... w obiektach – ilu obiektów dotyczą wartości z kolumny „całkowita ilość skanów”
- bez usuniętych skanów – liczba plików umieszczonych w systemie z pominięciem plików które zostały później usunięte (ważne jest które pliki zostały usunięte a nie kiedy, w związku z tym istnieje możliwość, że statystyki wykonane za ten sam okres w różnych odstępach czasu mogą się różnić, taka sytuacja będzie miała miejsce jeżeli w tym czasie zostaną usunięte pliki wrzucone w okresie którego dotyczą wygenerowane statystyki). Statystyka ta najbardziej odzwierciedla rzeczywistą wydajność pracy.
- ... w obiektach – ilu obiektów dotyczą wartości z kolumny „bez usuniętych skanów”
- w trybie poprawy – liczba plików umieszczonych w systemie w trybie poprawy, czyli po negatywnym wyniku etapu kontroli

WARSZAWA 2016	Biblioteka Narodowa	 BIBLIOTEKA NARODOWA
	Repozytorium cyfrowe BN – moduł statystyk	

- ... w obiektach – ilu obiektów dotyczą wartości z kolumny „w trybie poprawy”
- usunięte – liczba plików które zostały usunięte, liczy się okres kiedy usunięty plik wrzucano a nie kiedy był usuwany, tym samym statystyki te podobnie jak „bez usuniętych skanów” wygenerowane za ten sam okres czasu w różnym odstępie czasu mogą się od siebie różnić
- ... w obiektach – ilu obiektów dotyczą wartości z kolumny „usunięte”
- usunięte w przedziale – liczba plików które zostały usunięte, liczy się okres kiedy zostały usunięte a nie kiedy usuwane pliki były umieszczane w systemie
- ... w obiektach – ilu obiektów dotyczą wartości z kolumny „usunięte w przedziale”

Statystyki skanowania

Data wygenerowania raportu Fri Jun 24 08:41:06 CEST 2016

Użytkownik	Całkowita ilość skanów	..w obiektach	Bez usuniętych	..w obiektach	W trybie poprawy	..w obiektach	Usunięte	..w obiektach	Usunięte w przedziale	..w obiektach
admin	335	28	335	28	5	3	0	0	0	0
m.dmitruk	810	30	810	30	0	0	0	0	0	0
k.karpinski	165	19	165	19	3	2	0	0	0	0
a.leszyska	21	3	21	3	0	0	0	0	0	0
k.wilczak	267	37	267	37	0	0	0	0	0	0
skanerysta1	17	13	17	13	0	0	0	0	0	0
provider-test1	1	2	1	2	0	0	0	0	0	0
redaktor1	0	0	0	0	0	0	0	0	0	0
Suma	1616		1616		8		0		0	

3. Statystyki obiektów

Statystyki te przedstawiają analizę typów plików umieszczanych w systemie w zadanym okresie czasu. W wierszach są wylistowane poszczególne typy plików (wartości mimetype) jakie były przesyłane, wraz z liczbą sumaryczną na końcu. Natomiast w poszczególnych kolumnach są zaprezentowane następujące wartości:

- pliki główne – liczba plików danego typu umieszczona w systemie jako pliki struktury paginacyjnej
- ... w obiektach – ilu obiektów dotyczy wartość z kolumny „pliki główne”
- pliki zbiorcze – liczba plików danego typu umieszczona w systemie jako pliki zbiorcze
- ... w obiektach – ilu obiektów dotyczy wartość z kolumny „pliki zbiorcze”

WARSZAWA 2016	Biblioteka Narodowa	 BIBLIOTEKA NARODOWA
	Repozytorium cyfrowe BN – moduł statystyk	

Statystyki zbiorcze obiektów w systemie

Data wygenerowania raportu Fri Jun 24 08:41:07 CEST 2016

Typ pliku	Pliki główne	..w obiektach	Pliki zbiorcze	..w obiektach
image/jpeg	2582	57	1	1
image/gif	1	1	0	0
video/mp4	1	1	1	1
audio/x-wav	1	1	0	0
video/3gpp	1	1	0	0
video/quicktime	1	1	0	0
audio/mpeg	1	1	0	0
video/x-flv	1	1	0	0
application/octet-stream	1	1	0	0
image/tiff	4464	159	52	5
image/png	7	7	0	0
application/pdf	4	2	10	4
application/zip	0	0	2	1
Suma	7065		66	

4. Statystyki użytkownik-skaner

Statystyki prezentują korelację użytkownik-skaner, co oznacza iż zaprezentowana jest liczba umieszczonych w systemie plików przez użytkownika z informacją na jakich skanerach te pliki były wytworzone. W pierwszej kolumnie wylistowani są użytkownicy a w kolejnej lista skanerów na których pliki były przez tego użytkownika skanowane.

Poszczególne kolumny z wartościami liczbowymi prezentują następujące statystyki:

- pliki główne – liczba plików umieszczonych w systemie strukturze paginacyjnej
- pliki zbiorcze – liczba plików umieszczonych w systemie jako pliki zbiorcze
- łącznie – łączna ilość umieszczonych w systemie plików
- w publikacjach – ilu obiektów dotyczy łączna ilość umieszczonych w systemie plików

Jednym z wylistowanych skanerów będzie „Born digital” – są to pliki umieszczone w systemie w ramach obiektów które były w systemie umieszczane ścieżką born digital, czyli nie były skanowane tylko wydawca wydał je w formie elektronicznej.

WARSZAWA 2016	Biblioteka Narodowa	 BIBLIOTEKA NARODOWA
	Repozytorium cyfrowe BN – moduł statystyk	

Statystyki relacji użytkownik–skaner

Data wygenerowania raportu Thu Jun 30 12:15:06 CEST 2016

Użytkownik	Skaner	Pliki główne	Pliki zbiorcze	Łącznie	W publikacjach
admin	Haselblat	0	0	0	0
	Optiplex	28	2	30	9
	Digital Born	307	7	314	20
k.karpinski	Haselblat	0	0	0	0
	Optiplex	128	3	131	14
	Digital Born	37	0	37	5
Suma		500	12	512	

5. Statystyki użytkownik-magazyn

Statystyki prezentują korelację użytkownik-magazyn, co oznacza iż zaprezentowana jest liczba umieszczonych w systemie plików przez użytkownika z informacją z jakich magazynów pochodzą obiekty z których pliki były skanowane. W pierwszej kolumnie wylistowani są użytkownicy a w kolejnej lista magazynów z których pochodziły obiekty z których wrzucane pliki były wytworzone.

Poszczególne kolumny z wartościami liczbowymi prezentują następujące statystyki:

- pliki główne – liczba plików umieszczonych w systemie strukturze paginacyjnej
- pliki zbiorcze – liczba plików umieszczonych w systemie jako pliki zbiorcze
- łącznie – łączna ilość umieszczonych w systemie plików
- w publikacjach – ilu obiektów dotyczy łączna ilość umieszczonych w systemie plików

Statystyki relacji użytkownik–magazyn

Data wygenerowania raportu Thu Jun 30 12:15:07 CEST 2016

Użytkownik	Magazyn	Pliki główne	Pliki zbiorcze	Łącznie	W publikacjach
admin	Magazyn czasopism	0	0	0	0
	Magazyn Książek	33	2	35	10
	Digital Born	302	7	309	19
k.karpinski	Magazyn czasopism	0	0	0	0
	Magazyn Książek	128	3	131	14
	Digital Born	37	0	37	5
Suma		500	12	512	

6. Statystyki skaner-użytkownik

WARSZAWA 2016	Biblioteka Narodowa	 BIBLIOTEKA NARODOWA
	Repozytorium cyfrowe BN – moduł statystyk	

Statystyki prezentują korelację skaner-użytkownik, co oznacza iż zaprezentowana jest liczba umieszczonych w systemie plików wykonanych na konkretnym skanerze z informacją przez jakiego użytkownika te pliki były umieszczane. W pierwszej kolumnie wylistowane są poszczególne skanery na których wrzucane pliki były wytwarzane a w kolejnej lista użytkowników którzy z danego skanera pliki umieszczali.

Poszczególne kolumny z wartościami liczbowymi prezentują następujące statystyki:

- pliki główne – liczba plików umieszczonych w systemie strukturze paginacyjnej
- pliki zbiorcze – liczba plików umieszczonych w systemie jako pliki zbiorcze
- łącznie – łączna ilość umieszczonych w systemie plików
- w publikacjach – ilu obiektów dotyczy łączna ilość umieszczonych w systemie plików

Jednym z wylistowanych skanerów będzie „Born digital” – są to pliki umieszczone w systemie w ramach obiektów które były w systemie umieszczane ścieżką born digital, czyli nie były skanowane tylko wydawca wydał je w formie elektronicznej.

Statystyki relacji skaner-użytkownik

Data wygenerowania raportu Thu Jun 30 12:21:28 CEST 2016

Skaner	Użytkownik	Pliki główne	Pliki zbiorcze	łącznie	W publikacjach
Haselblat					
	admin	0	0	0	0
	k.karpinski	0	0	0	0
Optiplex					
	admin	28	2	30	9
	k.karpinski	128	3	131	14
Digital Born					
	admin	307	7	314	20
	k.karpinski	37	0	37	5
Suma		500	12	512	

7. Statystyki skaner-magazyn

Statystyki prezentują korelację skaner-magazyn, co oznacza iż zaprezentowana jest liczba umieszczonych w systemie plików wykonanych na konkretnym skanerze z informacją z jakich magazynów pochodzą obiekty z których pliki były skanowane. W pierwszej kolumnie wylistowane są poszczególne skanery na których wrzucane pliki były wytwarzane a w kolejnej lista magazynów z których pochodziły obiekty z których wrzucane pliki były wytworzone.

Poszczególne kolumny z wartościami liczbowymi prezentują następujące statystyki:

- pliki główne – liczba plików umieszczonych w systemie strukturze paginacyjnej
- pliki zbiorcze – liczba plików umieszczonych w systemie jako pliki zbiorcze

WARSZAWA 2016	Biblioteka Narodowa	 BIBLIOTEKA NARODOWA
	Repozytorium cyfrowe BN – moduł statystyk	

- łącznie – łączna ilość umieszczonych w systemie plików
- w publikacjach – ilu obiektów dotyczy łączna ilość umieszczonych w systemie plików

Statystyki relacji skaner–magazyn

Data wygenerowania raportu Fri Jun 24 08:51:59 CEST 2016

Skaner	Magazyn	Pliki główne	Pliki zbiorcze	łącznie	W publikacjach
jeszcze nowszy skaner					
	Magazyn czasopism	0	0	0	0
	Magazyn Książek	0	0	0	0
nowy skaner					
	Magazyn czasopism	0	0	0	0
	Magazyn Książek	1526	17	1543	58
Suma		1526	17	1543	

8. Statystyki magazyn-użytkownik

Statystyki prezentują korelację magazyn-użytkownik, co oznacza iż zaprezentowana jest liczba umieszczonych w systemie plików wytworzonych z obiektów pochodzących z konkretnego magazynu z informacją przez jakiego użytkownika te pliki były umieszczane. W pierwszej kolumnie wylistowane są poszczególne magazyny z których pochodzą obiekty z których umieszczane w systemie pliki były wytwarzane a w kolejnej lista użytkowników którzy pliki powiązane z danym magazynem umieszczali.

Poszczególne kolumny z wartościami liczbowymi prezentują następujące statystyki:

- pliki główne – liczba plików umieszczonych w systemie strukturze paginacyjnej
- pliki zbiorcze – liczba plików umieszczonych w systemie jako pliki zbiorcze
- łącznie – łączna ilość umieszczonych w systemie plików
- w publikacjach – ilu obiektów dotyczy łączna ilość umieszczonych w systemie plików

WARSZAWA 2016	Biblioteka Narodowa	 BIBLIOTEKA NARODOWA
	Repozytorium cyfrowe BN – moduł statystyk	

Statystyki relacji magazyn-użytkownik

Data wygenerowania raportu Thu Jun 30 12:21:28 CEST 2016

Magazyn	Użytkownik	Pliki główne	Pliki zbiorcze	Łącznie	W publikacjach
Magazyn czasopism	admin	0	0	0	0
	k.karpinski	0	0	0	0
Magazyn Książek	admin	33	2	35	10
	k.karpinski	128	3	131	14
Digital Born	admin	302	7	309	19
	k.karpinski	37	0	37	5
Suma		500	12	512	

9. Statystyki magazyn-skaner

Statystyki prezentują korelację magazyn-skaner, co oznacza iż zaprezentowana jest liczba umieszczonych w systemie plików wytworzonych z obiektów pochodzących z konkretnego magazynu z informacją na jakich skanerach te pliki były wytworzone. W pierwszej kolumnie wylistowane są poszczególne magazyny z których pochodzą obiekty z których umieszczane w systemie pliki były wytwarzane a w kolejnej lista skanerów na jakich pliki te były skanowane.

Poszczególne kolumny z wartościami liczbowymi prezentują następujące statystyki:

- pliki główne – liczba plików umieszczonych w systemie strukturze paginacyjnej
- pliki zbiorcze – liczba plików umieszczonych w systemie jako pliki zbiorcze
- łącznie – łączna ilość umieszczonych w systemie plików
- w publikacjach – ilu obiektów dotyczy łączna ilość umieszczonych w systemie plików

Statystyki relacji magazyn-skaner

Data wygenerowania raportu Thu Jun 30 12:21:29 CEST 2016

Magazyn	Skaner	Pliki główne	Pliki zbiorcze	Łącznie	W publikacjach
Magazyn czasopism	Haselblat	0	0	0	0
	Optiplex	0	0	0	0
Magazyn Książek	Haselblat	0	0	0	0
	Optiplex	1526	17	1543	58
Suma		1526	17	1543	