

07. Statystyki wykorzystania Aplikacji Czytelnika i udostępnianych poprzez nią obiektów cyfrowych

Aplikacja Czytelnika systemu dLibra posiada wbudowany mechanizm statystyk dotyczących podstawowych działań użytkowników końcowych, takich jak wyszukiwanie czy wyświetlanie treści obiektów. Dane te gromadzone są wyłącznie w oparciu o dane o żądaniach, jakie otrzymuje Aplikacja Czytelnika. Dzięki temu użytkownicy końcowi nie są w stanie blokować tego monitorowania, co jest możliwe w przypadku wykorzystania do śledzenia dodatkowych zewnętrznych narzędzi, takich jak Google Analytics. Co więcej, w statystykach gromadzonych przez Aplikację Czytelnika zliczane są również bezpośrednie odwołania do plików z treścią, których nie da się w prosty sposób śledzić za pomocą narzędzi takich jak Google Analytics.

Tabela poniżej zawiera zestawienie podstawowych możliwości statystycznych, jakie oferuje wbudowany mechanizm statystyk systemu dLibra, w porównaniu z wykorzystaniem narzędzia Google Analytics.

Dodatkowo zachęcamy co zapoznania się z [podrozdziałem opisującym szczegółowo możliwości analizowania zachowania użytkowników Aplikacji Czytelnika systemu dLibra 6 przy pomocy Google Analytics](#).

	Wbudowane statystyki dLibra 6	Google Analytics
Sposób gromadzenia danych	Po stronie Aplikacji Czytelnika, na podstawie otrzymywanych żądań HTTP.	Po stronie przeglądarki WWW użytkownika, na podstawie kodu śledzącego Java Script.
Kompletność danych	Dane obejmują wszystkie żądania użytkowników, z pominięciem ruchu robotów wyszukiwarek (patrz niżej).	Dane nie obejmują: <ul style="list-style-type: none">• Użytkowników, którzy w swoich przeglądarkach blokują śledzenie ich za pomocą dedykowanych wtyczek.• Użytkowników, którzy odwołują się bezpośrednio do plików publikacji z pominięciem stron WWW biblioteki cyfrowej (np. przejście do pliku PDF bezpośrednio z wyników Google).• Ruchu robotów wyszukiwarek (patrz niżej).
Pomijanie ruchu generowanego przez roboty wyszukiwarek	To jakie programy (przeglądarki) nie są uwzględniane przy zliczaniu statystyk zależy od zawartości pliku konfiguracyjnego ignored_agents.txt . Każdy wpis jest traktowany jako wyrażenie regularne (w pliku znajduje się przykładowy wpis dla GoogleBot'a). Należy umieścić w nim łańcuchy znakowe odpowiadające polu User Agent (nagłówek HTTP) programów które nie powinny być zliczane (przykład dla googlebot'a w pliku). Nazwy user agentów można znaleźć w sieci np. tu http://en.wikipedia.org/wiki/User_agent . Po wprowadzeniu zmian w plikach należy zrestartować Tomcat'a żeby nowe wartości były brane pod uwagę. Wraz z uruchomieniem Tomcat'a do pliku ignored_agents.txt zapisywane są wszystkie ignorowane agenty pobrane z http://dlibra.psn.pl/ignored_agents.txt .	Za pomijanie ruchu robotów odpowiadają wewnętrzne mechanizmy Google Analytics.

Dostępne statystyki	<ul style="list-style-type: none"> Statystyki dotyczące rozwoju zasobu biblioteki cyfrowej <ul style="list-style-type: none"> Łączna liczba dostępnych obiektów w ujęciu czasowym (z dokładnością dzienną) Łączna liczba nowych obiektów w danym miesiącu (z dokładnością miesięczną) Łączna liczba nowych obiektów w danym miesiącu, z podziałem na formaty obiektów (z dokładnością miesięczną) Statystyki dotyczące zachowania użytkowników <ul style="list-style-type: none"> Łączna liczba wygenerowanych stron WWW (z dokładnością miesięczną) Łączna liczba wyszukiwań, z podziałem na proste i zaawansowane (z dokładnością miesięczną) Łączna liczba odwiedzających (sesji) (z dokładnością miesięczną) Łączna liczba wyświetlonych obiektów (z dokładnością miesięczną, kilkakrotne wyświetlenie tego samego obiektu w ramach jednej sesji użytkownika zliczane jest tylko jeden raz) Łączna liczba wyświetlonych obiektów, z podziałem na formaty obiektów (z dokładnością miesięczną, kilkakrotne wyświetlenie tego samego obiektu w ramach jednej sesji użytkownika zliczane jest tylko jeden raz) 	Wyłącznie statystyki dotyczące zachowania użytkowników, bardzo rozbudowane, zależne od konfiguracji Google Analytics (więcej informacji tutaj).
Format danych	Wykresy (PNG), tabele (HTML, CSV).	Interaktywne wykresy i tabele, eksport do formatów: CSV, XLS, PDF.
Sposób dostępu	Domyślnie publicznie dla wszystkich, pod adresem <adres główny biblioteki cyfrowej>/stats (np. http://sbc.org.pl/stats/).	Dostęp tylko dla uprawnionych użytkowników, poprzez serwis Google Analytics .

Dostęp do treści obiektu, a zbieranie statystyk

- W przypadku wielokrotnego dostępu do tych samych treści przez użytkownika w ramach jednej sesji, statystyki wyświetleń/pobrań danego obiektu/pliku są zwiększane tylko raz.
- Statystyki są zliczane niezależnie od tego, czy użytkownik wyświetlił treść obiektu za pomocą przeglądarki osadzonej na stronie Aplikacji Czytelnika, czy też poprzez bezpośrednie odwołanie (link) do pliku treści (np. PDF).
- Jeżeli użytkownik nie ma uprawnień do wyświetlenia treści danego obiektu, to przy próbie dostępu do treści statystyki nie zostaną zwiększone.

Statystyki wyświetlania

- Statystyki wyświetleń treści obiektu zbierane są zarówno na poziomie wydania obiektu, jak i poszczególnych jego plików.
- Dostęp do dowolnego pliku wydania zwiększa statystyki tego wydania.
- Statystyki dla poszczególnych plików zliczane są niezależnie od siebie. Oznacza to, że dostęp do jednego pliku w ramach wydania nie ma wpływu na statystyki pozostałych plików w wydaniu.
- Statystyki wyświetleń treści obiektów w poszczególnych formatach wyliczane są jedynie na podstawie statystyk wyświetleń plików głównych w danym formacie. Przykładowo dla obiektu w formacie GALERIA zawierającego 20 plików raport ze statystykami będzie bazować na wyświetleniach tylko 1 pliku (plik główny).
- W przypadku zmiany treści obiektu (podmiany plików) statystyki dla nowych plików zliczane są od nowa (od zera). Dotychczasowe wyświetlenia dla podmienionych plików, z wyjątkiem statystyki dotychczasowego pliku głównego obiektu/formatu, są usuwane. Statystyki nieaktualnych plików głównych nadal są uwzględniane w raportach.

Statystyki pobierania

- Statystyki pobierania dla wydań są rejestrowane dla wszystkich wydań, ale przechowywane są jedynie na poziomie kolekcji, tzn. liczba pobrań wydań jest zapisywana jako suma pobrań wszystkich wydań w ramach poszczególnych kolekcji.
- Statystyki pobrań dla poszczególnych plików są doliczane do statystyki wyświetleń tych plików.
- W przypadku pobierania wszystkich plików wydania (ZIP), statystyka wyświetleń/pobrań jest zwiększana dla wydania oraz **tylko** dla pliku głównego obiektu.
- W przypadku obiektu wieloformatowego przy pobieraniu wszystkich plików pojedynczego formatu (ZIP), statystyka wyświetleń/pobrań jest zwiększana dla wydania oraz **tylko** dla pliku głównego danego formatu.
- W przypadku obiektu wieloformatowego przy pobieraniu wszystkich plików wszystkich formatów (ZIP), statystyka pobrań jest zwiększana **tylko** dla wydania. Statystyki plików **nie** są zwiększane.