

02. Rozszerzenia

Niniejszy rozdział zawiera listę rozszerzeń (wtyczek) do systemu dLab, które zostały wykorzystane w domyślnej konfiguracji modelu pracy. Rozszerzenia te wykonują czynności w ramach zadania, które można zautomatyzować.

Rozszerzenia do wykonywania czynności

Rozszerzenia do wykonywania czynności realizują określone czynności. W domyślnej konfiguracji modelu pracy rozszerzenia do wykonywania czynności to: przygotowanie wersji prezentacyjnej, publikacja w dLibrze, oraz archiwizacja.

Przygotowanie wersji prezentacyjnej - PDF

Wtyczka do generowania wersji prezentacyjnej w formacie PDF z plików źródłowych (plików matek). Wtyczka komunikuje się z systemem [ABBYY Recognition Server 3.0](#) za pośrednictwem protokołu usług sieciowych SOAP.

Przygotowanie wersji prezentacyjnej - DjVu

Wtyczka do generowania wersji prezentacyjnej w formacie DjVu z plików źródłowych (plików matek). Wtyczka korzysta z oprogramowania Document Express na zasadzie wywołania pliku wykonywalnego. Może uruchomić program zarówno na lokalnym komputerze, jak i na zdalnej maszynie, używając protokołu SSH. Wtyczkę tę można bardzo łatwo skonfigurować, aby wywołać dowolny inny program.

Przygotowanie wersji prezentacyjnej - kopiowanie plików wejściowych (PDFy wydawnicze)

Ta wtyczka nie przetwarza plików matek w celu stworzenia wersji prezentacyjnej, tylko bezpośrednio kopiuje pliki matki do katalogu z plikami prezentacyjnymi. Wtyczka ta dedykowana jest do sytuacji w której dana pozycja nie musi zostać zeskanowana w celu uzyskania wersji cyfrowej, tylko jest możliwość uzyskania wersji (PDF lub innej) gotowej do publikacji bezpośrednio od wydawnictwa, będącej zarówno postacią źródłową (pliki master/pliki matki) jak i postacią prezentacyjną.

Publikacja w dLibrze

Wtyczka do publikacji w dLibrze dodaje pliki umieszczone w katalogu wersji prezentacyjnej do publikacji planowanej powiązanej z wykonywanym zadaniem. Dodane wydanie jest automatycznie oznaczane jako opublikowane. W przypadku dokumentów PDF oraz DjVu jest też automatycznie tworzona miniaturka wydania z pierwszej strony publikacji. Ponowne uaktywnienie wtyczki w tym samym zadaniu (np. na skutek restartu zadania) skutkuje podmianą plików w już istniejącym wydaniu publikacji.

Archiwizacja

Wtyczka do archiwizacji umieszcza pliki uzyskane w ramach wykonywanego zadania w systemie dArceo, opracowanym na bazie Wielofunkcyjnego Repozytorium Danych Źródłowych, zrealizowanego w ramach etapu A11 projektu [SYNAT](#). Wtyczka archiwizuje oryginalne pliki matki w bibliotece cyfrowej, a także ewentualne wyniki poszczególnych etapów przetwarzania. Pliki z różnych etapów przetwarzania archiwizowane są jako osobne obiekty, powiązane odpowiednimi zależnościami. Ponowne uaktywnienie wtyczki w tym samym zadaniu (np. na skutek restartu zadania) skutkuje aktualizacją (stworzeniem nowych wersji) obiektów już istniejących w archiwum.

Rozszerzenia do automatycznego przetwarzania

Rozszerzenia tego typu nie wykonują konkretnych czynności w ramach zadania, tylko wykonują dodatkowe przetwarzanie przy okazji zmiany stanu określonych czynności. Najczęściej są to krótkie, pomocnicze operacje.

Zmiana nazw plików

Wtyczka do zmiany nazw plików jest uruchamiana w momencie oznaczania czynności Przygotowanie plików matek jako wykonanej. Wtyczka zmienia nazwy wszystkich plików tak, aby składały się z takiego samego przedrostka oraz numeru porządkowego, np. p0000.tiff, p0001.tiff, p0002.tiff, itd. Standaryzacja nazw plików ułatwia ich późniejsze przetwarzanie.

Generowanie miniatur

Wtyczka generuje miniatury ze wszystkich plików matek zapisanych we wspieranych formatach, tzn. tiff, png, jpg lub gif. Miniatury są umieszczane w specjalnym katalogu, który później wykorzystuje aplikacja WWW do wyświetlania zawartości katalogu plików matek. Ma to na celu ułatwienie pracy osobom wykonującym weryfikację i obróbkę graficzną plików matek - mogą one wstępnie zobaczyć zeskanowane pliki i stwierdzić, które z nich chcą pobrać na swój komputer. To rozszerzenie również jest uruchamiane po wykonaniu czynności Przygotowanie plików matek (bezpośrednio po zakończeniu pracy wtyczki zmieniającej nazwy plików), a dodatkowo gdy czynność Obróbka graficzna staje się gotowa do wykonania (na wypadek gdyby pliki matki zostały zmienione).

Zliczanie plików

Ta wtyczka liczy pliki znajdujące się w katalogu powiązany z przetwarzaną czynnością i zapisuje ich liczbę jako wartość parametru tej czynności. Wartości te są później wykorzystywane przy tworzeniu raportów z wykonanych prac. Rozszerzenie można dowolnie konfigurować w zależności od aktualnie przetwarzanego typu czynności - zliczać różne pliki (według ustawionego wzorca nazwy pliku) w katalogu wejściowym lub wyjściowym. Domyślnie wtyczka jest uruchamiana po wykonaniu czynności Przygotowanie plików matek (zlicza wszystkie pliki w katalogu wyjściowym, czyli określa liczbę zeskanowanych stron) oraz po wykonaniu czynności Przygotowanie wersji prezentacyjnej (zlicza wszystkie pliki w katalogu wejściowym, czyli liczbę stron, które zostały poddane procesowi OCR).

Tworzenie kopii zapasowej plików

Rozszerzenie kopiuje katalog przestrzeni roboczej przetwarzanego zadania do określonej lokalizacji - w domyśle na dodatkową macierz dyskową. Proces kopiowania jest uruchamiany po wykonaniu czynności Przygotowanie plików matek i ma to na celu stworzenie kopii bezpieczeństwa plików matek zanim zostaną ostatecznie zarchiwizowane w zewnętrznym repozytorium. Ponieważ kopiowanie może trwać dłuższy czas, wtyczka ta wykonuje kopiowanie w tle i nie blokuje możliwości wykonania kolejnych czynności w zadaniu.

Zmiana uprawnień użytkowników dLibry

Wtyczka zmienia w systemie dLibra uprawnienia użytkowników do obiektu powiązanego z aktualnie przetwarzanym zadaniem. To, w jaki sposób prawa mają być zmienione, podlega konfiguracji. W domyślnym ustawieniu nadawane są prawa odczytu dla użytkowników publicznych (typowych kont założonych przez www oraz użytkowników niezalogowanych). Wtyczka uruchamiana jest podczas akceptacji czynności Publikowanie w dLibrze, dzięki czemu treść publikacji stanie się dostępna dla wszystkich dopiero po tym, jak redaktor zweryfikuje jej poprawność.

Czyszczenie przestrzeni roboczej

Zadaniem tej wtyczki jest usunięcie z przestrzeni roboczej wszystkich plików i katalogów należących do przetwarzanego zadania, aby zwolnić miejsce dla następnych zadań. Operacja ta jest wykonywana przy akceptacji czynności Archiwizacja, czyli gdy wszystkie czynności związane z przetwarzaniem plików są zakończone i zweryfikowane. W przypadku klocków introligatorskich foldery wspólne dla całego klocka są usuwane dopiero po zakończeniu ostatniego zadania w klocku.