

Legimi S.A.

Legimi S. A.

Technologia znaku wodnego Legimi Watermark

Technical specification.

22 lutego 2018

Spis treści

1	Wprowadzenie	3
1.1	Cel dokumentu	3
1.2	Zakres produktu	3
1.3	Omówienie dokumentu	3
2	Opis produktu	3
2.1	Kontekst funkcjonowania	4
2.2	Sposób działania	4
2.3	Trwałość znaku wodnego	4
2.4	Poprawność i zgodność z oprogramowaniem czytników	4
2.5	Licencje	5
2.6	Instalacja	6
2.7	Sposób wykorzystania	6
2.8	Błędy	9
2.9	Ograniczenia	9
A	Dodatki	10
A.1	Przykłady argumentów uruchomienia	10
A.2	Przykład kompilacji środowiska Mono w systemie Red Hat	10

1 Wprowadzenie

1.1 Cel dokumentu

Niniejszy dokument prezentuje funkcje oraz sposób użycia konsolowego narzędzia do zabezpieczania publikacji elektronicznych (w formatach EPUB, PDF oraz MOBI) za pomocą znaku wodnego.

1.2 Zakres produktu

Produkt *Legimi Watermark*, opisany w tym dokumencie, adresuje problem ochrony własności intelektualnej twórcy poprzez nanoszenie na publikacje elektroniczne znaku wodnego (*watermarku*). W odróżnieniu od DRM, który opiera się na zaszyfrowaniu publikacji w taki sposób, że jej odczytanie jest możliwe wyłącznie za pomocą określonego oprogramowania i na ściśle określonym zbiorze urządzeń, watermark nie nakłada na kupującego tego rodzaju ograniczeń. Zamiast szyfrowania treści, wprowadzane są do niej dodatkowe informacje, które pozwalają zidentyfikować właściciela książki.

Legimi Watermark pozwala na automatyczne naniesienie znaku wodnego na określony plik, a także na odczytanie znaku wodnego umieszczonego wcześniej w pliku. W zależności od formatu pliku, wykorzystywane są różne kanały kodowania informacji o właścicielu pliku. Znak wodny tworzony jest zarówno w postaci widocznej (krótki tekst z informacją o zabezpieczeniu pliku) jak i niewidocznej dla kupującego.

1.3 Omówienie dokumentu

Dalsza część niniejszego dokumentu szczegółowo omawia funkcje, sposób uruchomienia oraz interpretacji wyników programu. Określone zostają również ograniczenia technologii. Dokumentacja dotyczy narzędzia *Legimi Watermark* w wersji 3.0.0 lub nowszej.

2 Opis produktu

Legimi Watermark jest narzędziem konsolowym przeznaczonym do wdrożenia na serwerach umożliwiających użytkownikom pobieranie zakupionych publikacji. Działa on w dwóch trybach:

- tryb kodowania: na podstawie pliku w formacie EPUB, MOBI lub PDF oraz informacji identyfikującej kupującego, generuje on drugi plik, w formacie identycznym z formatem pliku wejściowego. Plik wynikowy zawiera zakodowane informacje o kupującym.
- tryb dekodowania: z pliku w formacie EPUB, MOBI lub PDF odczytywany jest zapisany tam wcześniej znak wodny (o ile został on wcześniej zapisany w plik; w przeciwnym wypadku program stwierdza brak znaku wodnego).

Legimi Watermark wspiera następujące formaty:

- EPUB w wersji 2
- EPUB w wersji 3
- MOBI w wersji 6
- hybrydowe pliki MOBI, zawierające dane w wersji 6 oraz w wersji 8 (takie właśnie pliki są standardowo generowane przez *Kindlegen* 2.9)
- PDF w wersji 1.7 lub niższej (ISO 32000-1)

2.1 Kontekst funkcjonowania

Technologia znaku wodnego powinna być wykorzystywana w momencie pobierania zakupionego pliku przez użytkownika. Zakładając, że oryginalny plik publikacji elektronicznej oraz dane identyfikujące kupca są takie same, plik ze znakiem wodnym również będzie niezmienny. Stąd możliwe jest cache'owanie plików wynikowych. Cache plików nie jest jednak częścią *Legimi Watermark*.

2.2 Sposób działania

Działanie *Legimi Watermark* zależne jest od formatu pliku wejściowego. Dla plików EPUB oraz MOBI, informacja o znaku wodnym kodowana jest w kilku niezależnych kanałach. Dzięki temu usunięcie znaku wodnego z zabezpieczonego pliku staje się trudniejsze. Sposób kodowania znaku wodnego nie wpływa na wygląd książki - znak wodny jest niewidoczny w przeglądarkach plików EPUB i MOBI. Poza tym, w celu wywarcia dodatkowego efektu psychologicznego, dodawana jest jawna informacja o zabezpieczeniu pliku znakiem, umieszczana na początku tekstu. W przypadku plików PDF, umieszczana jest widoczna informacja, w postaci tekstu dodawanego u dołu każdej ze stron pliku, modyfikowane są metadane pliku, a także wstawiany jest niewidoczny tekst.

2.3 Trwałość znaku wodnego

Istotną cechą technologii nanoszenia znaku wodnego na publikacje elektroniczne jest trwałość tego zabezpieczenia, tj. odporność na próby usunięcia znaku, w szczególności za pomocą ogólnie dostępnych narzędzi. *Legimi Watermark* dla plików EPUB oraz MOBI został pod tym kątem przetestowany za pomocą narzędzia *Calibre* (<http://calibre-ebook.com>) w wersji 3.17 oraz *Kindlegen* (<https://www.amazon.com/gp/feature.html?docId=1000765211>) w wersji 2.9. W obu przypadkach testowym środowiskiem była stacja robocza z 64 bitowym systemem operacyjnym Windows 8.1.

Calibre jest jednym z najczęściej wykorzystywanych programów do zarządzania elektroniczną biblioteką użytkownika. Swoją popularność zawdzięcza wsparciu dla ogromnej liczby formatów plików, a także możliwością automatycznego zdejmowania zabezpieczeń stosowanych przez *Adobe Digital Editions* (po zainstalowaniu odpowiedniego pluginu). Testy wykazały, że znak wodny *Legimi* jest odporny na konwersję w *Calibre* między formatami EPUB i MOBI. Oznacza to, że znak wodny naniesiony na plik EPUB można nadal odczytać po przekonwertowaniu zabezpieczonego pliku w *Calibre* do formatu MOBI (lub EPUB). To samo dotyczy zabezpieczonych plików MOBI. Jednocześnie znak wodny w zabezpieczonych plikach pozostaje niewidoczny po konwersji.

Kindlegen używany jest głównie przez wydawców - pozwala na on konwersję pliku EPUB do formatu MOBI (nie ma możliwości konwersji w drugą stronę). Znak wodny utworzony w programie *Legimi Watermark* jest również odporny na konwersję wykonaną tym, firmowanym przez Amazona, programem.

W chwili powstawania tego dokumentu, nie było znane żadne ogólnie dostępne narzędzie pozwalające na automatyczne wykrycie lub usunięcie znaku wodnego *Legimi* z zabezpieczonej tą technologią publikacji elektronicznej. Jak każde inne zabezpieczenie, *Legimi Watermark* może zostać usunięty poprzez "ręczną" edycję pliku. Proces taki wymaga jednak specjalistycznej wiedzy.

2.4 Poprawność i zgodność z oprogramowaniem czytników

Program *Legimi Watermark* generuje poprawne, tj. zgodne ze specyfikacją formatu, pliki wynikowe, pod warunkiem, że plik oryginalny jest również poprawny. Dla plików EPUB oraz PDF

publicznie dostępna jest specyfikacja formatu, można więc weryfikować poprawność plików za pomocą ogólnie dostępnych walidatorów.

W przypadku plików MOBI, nie jest znana oficjalna specyfikacja - jest to format zamknięty, będący w posiadaniu firmy Amazon. Walidację można przeprowadzić pośrednio, wykorzystując funkcję przesłania pliku MOBI do chmury Amazona poprzez załączenie go do wiadomości email (<https://www.amazon.com/gp/sendtokindle>). Przesłany plik jest wówczas weryfikowany; w przypadku błędów, otrzymywany jest zwrotny email (nie zawiera on szczegółów walidacji). Jeśli plik dotrze na urządzenie (czytnik Kindle lub smartfon albo tablet z zainstalowaną aplikacją Kindle), należy go uznać za poprawny.

Pliki zabezpieczone *Legimi Watermark* są wyświetlane prawidłowo we wszystkich popularnych programach do odczytu plików EPUB, MOBI lub PDF oraz w dedykowanych urządzeniach typu e-czytnik. Poniżej znajduje się lista aplikacji, na których zabezpieczone pliki zostały przetestowane pod tym kątem. W tych aplikacjach niewidoczny znak wodny (dla formatów EPUB i MOBI) pozostaje niezauważalny dla użytkownika.

Czytniki formatu EPUB:

- Adobe Digital Editions 3.0
- Calibre 3.17
- Moon+ Reader 3.5.4
- Google Play Books 4.0.47.184185130
- iBooks dla iOS 11.2.5

Czytniki formatu MOBI:

- Kindle for PC 1.21.0
- Kindle for Android 7.0.0.49
- Kindle for iOS 6.4
- czytnik Kindle 4
- czytnik Kindle Paperwhite 3
- czytnik Kindle 8

Właściwości znaku wodnego Legimi są prawdopodobnie zachowane we wszystkich popularnych czytnikach, jednak ze względu na ograniczone zasoby, nie ma możliwości sprawdzenia wszystkich konfiguracji oprogramowania i sprzętu.

2.5 Licencje

Legimi umożliwia zakup technologii zabezpieczania plików znakiem wodnym w postaci 3 rodzajów licencji:

- płatnej jednorazowo
- w modelu prowizyjnym - płatność za każdą transakcję, w wyniku której zostaje udostępniony zabezpieczony plik (*pay per file*)
- modelu prowizyjnym Software as a Service (płatność również *pay per file*)

Od strony technicznej, wdrożenie dwóch pierwszych licencji jest bardzo podobne: należy uruchomić program *Legimi Watermark* z odpowiednimi parametrami. W przypadku licencji prowizyjnej konieczne jest podanie dodatkowych argumentów, na podstawie których zostanie później obliczona należność za okres rozliczeniowy. Korzystanie z narzędzia w modelu prowizyjnym wymaga również umożliwienia łączności z usługą rozliczeniową Legimi.

Trzecia opcja oznacza korzystanie z API, które Legimi udostępnia poprzez HTTP (REST API). Należy przygotować skrypt lub program, który będzie wysyłał żądanie zabezpieczenia/sprawdzenia pliku, wraz z plikiem, i odbierał odpowiedź. Szczegółowa dokumentacja techniczna REST API dostępna jako osobny dokument.

2.6 Instalacja

Legimi Watermark nie wymaga żadnej specjalnej konfiguracji - wystarczy rozpakować archiwum do wybranego katalogu. Do uruchomienia programu wymagane środowiska .NET w wersji co najmniej 3.5 lub środowiska Mono w wersji co najmniej 2.8. W środowisku Windows .NET jest zainstalowany domyślnie. Na komputerach działających pod kontrolą innych systemów operacyjnych, konieczna jest instalacja Mono. W tym celu należy pobrać odpowiedni pakiet binarny ze strony <http://www.mono-project.com/download/stable/#download-lin>. Dostępne są dystrybucje binarne m.in. dla Debiana, Ubuntu, Mac OSX oraz Fedory. W przypadku pozostałych systemów operacyjnych, możliwa jest instalacja mono poprzez skompilowanie źródeł. Przykład kompilacji dla dystrybucji Red Hat można znaleźć w dodatkach.

Jeśli *Legimi Watermark* został zakupiony w wersji *pay per file* lub w licencji ograniczonej czasowo, program wymaga łączności z serwerem Legimi w celu sprawdzania aktualności licencji. W tym celu należy umożliwić komunikację poprzez TCP na porcie 80 do hosta www.legimi.pl. Dodatkowo program zapisuje informację o licencji w aktualnym katalogu (aby nie było konieczne odpytywanie serwera przy każdym uruchomieniu), zatem trzeba nadać odpowiednie uprawnienia dla zapisu w katalogu programu.

2.7 Sposób wykorzystania

Program *Legimi Watermark* uruchamiany jest w jako program konsolowy. Wymagane argumenty (w nawiasach podano ich krótkie formy, o ile istnieją):

- `--mode (-m)`: ustawia tryb działania programu. Możliwe wartości to *encode (e)* lub *decode (d)*.
 - *encode* ustawia program w tryb kodowania, tj. umieszczania znaku wodnego w publikacji elektronicznej. W tym trybie wynik działania programu stanowi plik zapisany we wskazanej lokalizacji
 - *decode* ustawia program w tryb dekodowania, tj. odczytania treści znaku wodnego z pliku publikacji elektronicznej. Jeśli uda się znaleźć znak wodny, zostanie on wypisany na standardowe wyjście.
- `--type (-t)`: podaje format pliku wejściowego. Możliwe wartości to *epub*, *mobi* lub *pdf*.
- `--input (-i)`: ścieżka do pliku wejściowego
- `--password (-p)`: opcjonalne hasło używane do zaszyfrowania (i odszyfrowania w trybie dekodowania) znaku wodnego. Zastosowanie hasła uniemożliwia odczytanie informacji ze znaku wodnego osobom trzecim, nawet jeśli użyją one *Legimi Watermark* (będzie możliwe

jedynie stwierdzenie obecności znaku wodnego w plikach EPUB i MOBI). Hasło powinno składać się z maksymalnie 32 znaków ASCII. Dla plików PDF ten argument wymagany jest kiedy zabezpieczony dokument ma mieć ograniczone kopiowanie lub drukowanie - wówczas w przypadku dekodowania znaku wodnego należy podać to samo hasło, co przy tworzeniu zabezpieczonego pliku. W innych przypadkach (tj. kiedy na plik PDF nie nakłada się ograniczeń kopiowania lub drukowania), hasło jest ignorowane podczas zabezpieczania plików PDF.

Dodatkowe argumenty wymagane w trybie kodowania (są one ignorowane w trybie dekodowania):

- `--output (-o)`: ścieżka, pod którą zostanie zapisany plik wynikowy
- `--message (-msg)`: treść znaku wodnego, który zostanie zakodowany w pliku wejściowym. Znak wodny powinien składać się z maksymalnie 48 znaków ASCII. Wyjątkiem są pliki PDF, dla których długość znaku wodnego nie jest ograniczona.
- `--no-transparent`: jedną z technik nakładania znaku wodnego dla plików EPUB oraz MOBI jest dodawanie przezroczystego tekstu na końcach rozdziałów (tuż przed tagiem `mbp:pagebreak` w przypadku pliku MOBI). Niektóre programy do przeglądania tego typu plików (FBReader, Kindle for PC) niepoprawnie obsługują przezroczystość i w efekcie tekst, który powinien być ukryty, jest widoczny. Użycie tej opcji pozwala wyłączyć tę technikę wstawiania znaku wodnego. W ten sposób informacja o zabezpieczeniu pozostaje zawsze niewidoczna, jednak samo zabezpieczenie zostaje osłabione (brak jednej z metod kodowania informacji). Opcja nie ma efektu dla plików PDF.

W zależności od podanego formatu pliku wejściowego, w trybie kodowania dostępne są również dodatkowe argumenty dotyczące tylko tego formatu. Dla plików EPUB:

- `--epub-watermark-page`: ścieżka do pliku XHTML, który zostanie dodany do pliku EPUB jako strona z informacją o zabezpieczeniu pliku znakiem wodnym. Strona z informacją wstawiana jest zaraz po okładce, jako druga strona pliku EPUB. Domyślnie wykorzystywany jest plik `default-info.xhtml`, który jest częścią aplikacji Legimi Watermark. Można zastosować dowolny plik XHTML, jeśli będzie on zawierał referencje do zewnętrznych plików (style lub obrazki), pliki te zostaną automatycznie dodane do `epub`.
- `--epub3-watermark-page`: działa analogicznie jak powyższa opcja, z tym, że podany plik XHTML zostanie użyty jeśli wejściowy plik EPUB będzie w formacie EPUB 3.
- `--epub-no-watermark-page`: podanie tej opcji spowoduje brak strony z informacją o zabezpieczeniu znakiem wodnym w wynikowym pliku EPUB.
- `--epub-encoding-heuristic`: domyślnie, kiedy pliki HTML, wchodzące w skład archiwum EPUB, nie mają zadeklarowanego kodowania (w tagu `meta` pliku HTML), niektóre metody kodowania znaku wodnego nie zostaną wykorzystane. Podanie tej opcji spowoduje użycie heurystyki do wykrycia kodowania pliku.
- `--epub-encoding-force`: służy do podania nazwy kodowania, która zostanie użyta zawsze kiedy plik HTML wchodzący w skład pliku EPUB nie ma zadeklarowanego kodowania. Opcja ta jest alternatywą dla `--epub-encoding-heuristic`. W przypadku podania obu tych opcji, heurystyka nie zostanie użyta.

Dla plików PDF:

- `--pdf-font`: służy do podania rodzaju czcionki, który zostanie wykorzystany do umieszczenia informacji o znaku wodnym w pliku PDF. Dopuszczalne wartości: *courier*, *courier-bold*, *helvetica*, *helvetica-bold*, *times*, *times-bold*, *times-italic*, *times-bold-italic*. Domyślnie wykorzystywana jest czcionka *helvetica*.
- `--pdf-font-size`: rozmiar czcionki wykorzystanej do umieszczenia informacji o znaku wodnym w pliku PDF. Domyślnie 16.
- `--pdf-font-color-r`: czerwona składowa koloru czcionki, w systemie RGB. Domyślnie 0 (domyślny kolor to czcionki to czarny).
- `--pdf-font-color-g`: zielona składowa koloru czcionki, w systemie RGB. Domyślnie 0 (domyślny kolor to czcionki to czarny).
- `--pdf-font-color-b`: niebieska składowa koloru czcionki, w systemie RGB. Domyślnie 0 (domyślny kolor to czcionki to czarny).
- `--pdf-no-copy`: wyłącza możliwość kopiowania tekstu z zabezpieczonego pliku. Użycie tej opcji wymaga podania hasła (parament `--password`). Domyślnie kopiowanie jest dozwolone.
- `--pdf-no-print`: wyłącza możliwość drukowania zabezpieczonego pliku. Użycie tej opcji wymaga podania hasła (parament `--password`). Domyślnie drukowanie jest dozwolone.
- `--pdf-text-position`: ustawia odległość tekstu dodawanego do każdej strony od dolnej krawędzi strony, w pikselach. Dozwolone są wartości ułamkowe (z kropką jako separatorem dziesiętnym) i nieujemne.

Dla plików MOBI:

- `--mobi-watermark-message`: służy do podania tekstu informującego o zabezpieczeniu pliku znakiem wodnym. Tekst ten zostanie umieszczony na początku pliku wynikowego. Tekst powinien być poprawnym tagiem HTML. Możliwe jest zastowanie w nim stylu CSS, przy czym należy pamiętać, że MOBI nie wspiera w pełni zarówno standardu HTML jak i CSS. Domyślnie używany jest tekst: `<p>Plik jest zabezpieczony znakiem wodnym</p>`. Tekst podany w ten sposób jest widoczny na czytnikach korzystających z wersji 6 plików MOBI.
- `--mobi-no-watermark-message`: podanie tej opcji spowoduje, że tekst z informacją o zabezpieczeniu pliku znakiem wodnym nie zostanie dodany do pliku MOBI.
- `--mobi-kf8-watermark-message`: służy do podania tekstu informującego o zabezpieczeniu pliku znakiem wodnym. Tekst ten zostanie umieszczony na początku pliku wynikowego. Tekst powinien być poprawnym tagiem HTML. Zostanie on wyświetlony w czytnikach używających wersji 8 pliku MOBI (Kindle Fire 8)
- `--mobi-leave-srcs`: wiele plików MOBI zawiera rekord SRCS, w którym znajduje się plik oryginalny EPUB, na podstawie którego MOBI zostało utworzone. Jest on domyślnie tworzony podczas konwersji MOBI za pomocą programu *Kindlegen*. *Legimi Watermark* usuwa ten rekord, co pozwala istotnie zmniejszyć rozmiar wynikowego pliku oraz zapobiega wyciekowi niezabezpieczonego pliku EPUB z rekordu SRCS. Jeśli z jakichś względów rekord SRCS powinien być zachowany, należy użyć tej opcji.

Przykładowe komendy uruchomienia programu są podane w załączniku tego dokumentu.

2.7.1 Wdrożenie w modelu prowizyjnym

W przypadku uruchamiania programu zakupionego w ramach licencji prowizyjnej, konieczne jest podanie dodatkowych argumentów podczas zapisywania informacji w zabezpieczanym pliku dowolnego formatu:

- *-transaction*: identyfikator transakcji, dla której generowany jest zabezpieczony plik
- *-price*: cena brutto zabezpieczanej książki. Jako separatora dziesiętnego należy użyć kropki.

Powyższe informacje są od razu przesyłane na serwer Legimi. W celu zminimalizowania czasu pracy programu, przesyłanie odbywa się równolegle (w osobnym wątku) do procesu nakładania zabezpieczenia. Brak możliwości komunikacji z serwerem Legimi spowoduje zakończenie programu błędem.

2.8 Błędy

Program informuje o błędach wypisując komunikaty na standardowe wyjście błędu (*stderr*), jeden błąd w pojedynczej linii wyjścia. Rodzaj błędu można rozpoznać po początku komunikatu:

- *WARN* - oznacza ostrzeżenie, które prawdopodobnie nie przeszkodzi w dalszym wykonaniu programu. Może to być podanie nadmiarowych argumentów przy uruchomieniu narzędzia (np. podanie pliku wyjściowego w trybie dekodowania) lub wykrycie drobnych błędów w strukturze plików wejściowych.
- *ERROR* - oznacza wystąpienie błędu, który powoduje przerwanie wykonania programu. Może być spowodowany brakiem wymaganych argumentów lub nieprawidłowym formatem pliku wejściowego.
- *FATAL* - oznacza wystąpienie krytycznego błędu, który powoduje przerwanie działania programu. Wystąpienie tego rodzaju błędu należy zgłosić Legimi.

Informację o ewentualnym błędzie można również uzyskać sprawdzając kod wyjścia programu:

- 0: operacja zakończona poprawnie
- 1: niepoprawne argumenty
- 2: błąd w trakcie działania programu
- 3: błąd krytyczny

2.9 Ograniczenia

- małe pliki EPUB i MOBI: z uwagi na konieczność zakodowania znaku wodnego w sposób niewidoczny dla użytkownika końcowego, informacja o nim musi być odpowiednio rozproszona w pliku. Stąd dla pliku z krótkim tekstem i bez obrazków, niemożliwe może się okazać umieszczenie w nim niewidocznego znaku wodnego.
- zgodność z czytnikami: mimo dołożenia wszelkich starań i sprawdzenia zabezpieczonych plików na szerokiej gamie urządzeń i programów, nie można zagwarantować poprawnego wyświetlania plików z watermarkiem w każdych okolicznościach. Wpływ na to mogą mieć m.in. błędne czcionki, błędy w oprogramowaniu czytników lub niestandardowe ustawienia użytkownika.

A Dodatki

A.1 Przykłady argumentów uruchomienia

Zakodowanie znaku wodnego w pliku EPUB:

```
watermark.exe -t epub -m encode -msg "znak wodny" -i in.epub -o out.epub
```

Dekodowanie znaku wodnego z pliku EPUB:

```
watermark.exe -t epub -m decode -i in.epub
```

Zakodowanie znaku wodnego w pliku EPUB z wykorzystaniem hasła:

```
watermark.exe -t epub -m encode -msg "znak wodny" -i in.epub -o out.epub -p haslo
```

Dekodowanie znaku wodnego z pliku EPUB z wykorzystaniem hasła:

```
watermark.exe -t epub -m decode -i in.epub -p haslo
```

Zakodowanie znaku wodnego w pliku MOBI z użyciem innego niż domyślny tekstu o zabezpieczeniu pliku znakiem wodnym:

```
watermark.exe -t mobi -m encode -msg "znak wodny" -i in.mobi -o out.mobi --mobi-watermark-message "<p>tekst</p>"
```

A.2 Przykład kompilacji środowiska Mono w systemie Red Hat

```
yum install gcc libtool bison pkg-config libglib2.0-dev gettext make bzip2 g++  
  
wget http://origin-download.mono-project.com/sources/mono/mono-2.10.2.tar.bz2  
tar xvjf mono-2.10.2.tar.bz2  
cd mono-2.10.2  
./configure --prefix=/opt/mono-2.10  
  
# this might take 30-60 minutes  
make  
make install
```