

PDF No Save・HTMLモード コマンドライン対応機能 マニュアル

2014年3月11日更新

株式会社プランセス

1. 基本書式

pdf_html.exe ①設定ファイルのパス ②PDFファイルの格納フォルダーのパス ③出力フォルダーのパス ④(第4引数)

①第一引数 設定ファイル(拡張子pnh)のフルパス **必須**

設定ファイルは、詳細設定ウィザードの最終ステップで保存ボタンが表示されますので、そこで保存してください。

パスにスペース(空白文字)が含まれる場合は、全体を「”」で囲まないといけません、これはPDF No Saveとは関係のないことであり、コマンドライン操作ではどんなアプリでもそのようにしないといけません。

日本語のパスもOKですが、必ずShift_JISでエンコードされている必要があります(URLエンコードはしないでください)。結果的に、Shift_JISに含まれない漢字や韓国語のパスでは動作しません。

ex. E:¥homepage¥desktop¥download¥hoge¥sample20140210_taiken.pnh

②第二引数 PDFファイルの格納フォルダーのフルパス **必須**

ドライブ名は必ず大文字で指定してください。

その他、フォルダー名などは全て大文字・小文字を意識して指定してください。

パスにスペース(空白文字)が含まれる場合は、全体を「”」で囲まないといけません、これはPDF No Saveとは関係のないことであり、コマンドライン操作ではどんなアプリでもそのようにしないといけません。

日本語のパスもOKですが、必ずShift_JISでエンコードされている必要があります(URLエンコードはしないでください)。結果的に、Shift_JISに含まれない漢字や韓国語のパスでは動作しません。

必ず存在するパスでなければなりません。自動生成はされません。あらかじめ作成しておいてください。存在しなければエラーになります。

ex. E:¥homepage¥desktop¥download¥株式会社A社¥pdf20140214¥kdjcajdyHTYHBGFDSNBV7HR1QpMhCyt89

③第三引数 出力フォルダーのフルパス **必須**

ドライブ名は必ず大文字で指定してください。

その他、フォルダー名などは全て大文字・小文字を意識して指定してください。

パスにスペース(空白文字)が含まれる場合は、全体を「`”`」で囲まないといけません、これはPDF No Saveとは関係のないことであり、コマンドライン操作ではどんなアプリでもそのようにしないといけません。

日本語のパスもOKですが、必ずShift_JISでエンコードされている必要があります(URLエンコードはしないでください)。結果的に、Shift_JISに含まれない漢字や韓国語のパスでは動作しません。

必ず存在するパスでなければなりません。自動生成はされません。あらかじめ作成しておいてください。存在しなければエラーになります。

ex. E:¥homepage¥desktop¥download¥株式会社A社¥html20140214

※PDF No Save HTMLモードの基本動作原理は、第二引数(PDFファイルのパス)と第三引数(出力フォルダー内のHTMLページ)の関係から、PDFファイルの相対パスが決まり、PDFファイルがブラウザ上に表示される仕組みです。第二引数で指定されたPDF格納ディレクトリーと第三引数で指定された出力フォルダーの相対パス関係がそのままWEBサーバ上に展開されるように、FTPソフトなどでサーバに転送していただく手順となります。

ですから、第二引数と第三引数が違うドライブ(EドライブとFドライブなど)に存在する場合はエラーにしています。また、PDF格納ディレクトリーは、キャッシュ対策を使っている場合以外は、DocumentRootより上のディレクトリーにおくことはできませんので、出力フォルダーから見たPDFファイルの相対パスが「../../../../」のように「../」が4つ以上ある場合は、誤った配置になっている可能性が高いので、エラーにしています。

また、この相対パスの中に日本語のフォルダー名(ディレクトリー名)が出てきたり、PDFファイルのファイル名が日本語だとトラブルの元ですので、避けてください。上記のex.の例で言えば、「株式会社A社」というフォルダーがDocumentRootに該当するフォルダーであるとすれば、「株式会社A社」というフォルダー名は相対パスの中に出てこないで(「../pdf20140214/~」のようになります。)、何の問題もありません。よくわからない場合は、「日本語のパスは全て避けてください」。また、PDF格納フォルダーを、出力フォルダーのサブフォルダー(の中のサブフォルダー)に指定してください。そうすれば出力フォルダーのみを丸ごとアップロードすればPDFファイルも一緒にアップロードされますので、アップロー度忘れを防いで良いです。

④第四引数 さまざまなパラメータ (**省略可能**)

ログのパスを指定したり、エラーログのパスを指定することができます。詳細は次項を参照してください。

2. 第四引数(省略可能)について

※(重要)パラメータは「&」ではなく、「_」(アンダーバー2つ)で結んでください。第4引数に関しては、日本語を含む場合は必ずURLエンコード(Shift_JISで)してください。第1から第3引数に関しては、日本語を含んでも、URLエンコードしないでください。

ex.

```
batch_limit=200_error_log_path=e%3A%5Chomepage%5Cdesktop%5Ctest%5Cerror%93%FA%96%7B%8C%EA20140214a.txt_log_path=e%3A%5Chomepage%5Cdesktop%5Ctest%5C%8F%88%97%9D%83%8D%83O20140214a.txt_preview_url=https%3A%2F%2Ftest64.hogehoge.jp%3A8081%2Fhtml20140210%2F_note=%83e%83X%83g20140214%81i%91%CC%8C%B1%94%C5%81j
```

上記のパラメータは、

batch_limit: 200

error_log_path(エラーログのパス): e:¥homepage¥desktop¥test¥error日本語20140214a.txt

log_path(処理ログのパス): e:¥homepage¥desktop¥test¥処理ログ20140214a.txt

preview_url(ローカルサーバのプレビュー用URL): https://test64.hogehoge.jp:8081/html20140210/

note(識別用メモ): テスト20140214(体験版)

をエンコードしたものになります。

パラメータ	説明
batch_limit	一度にバッチ処理するPDFファイルの数の上限。デフォルト(未指定の場合)では200。最大値500。いくらを指定しても、結局は指定されたフォルダー内のPDFファイルの処理が終わるまで処理は続けられますから、たいした意味はないですが、batch_limitがたとえば20に指定されると、20ファイルの処理が終わるごとに5秒間の休止を入れます。PCの性能がイマイチであることなどが原因で、仮に、謎のエラーが起こる場合、連続処理が続いているがゆえにかもかもしれませんので、一度にバッチ処理するPDFファイルの上限を下げてみてください。
batch_sub	第二引数で指定されたPDF格納フォルダーのサブフォルダーを含むかどうか。含む場合はbatch_sub=1としてください。デフォルトでは、サブフォルダーは含みません。

パラメータ	説明
error_log_path	デフォルト(未指定の場合)の場合、エラーログは、PDF No Save・HTMLモード・コマンドライン対応版(exe)がインストールされているフォルダーのサブフォルダー「log」の中にあります「log/error_log日付.txt」に記録されますが、これを変えることができます。案件ごとに変えたい場合などは、指定してください。パスはフルパスで指定してください。普通は拡張子は.txtになるでしょうが、それにこだわる必然性はないです。
log_path	デフォルト(未指定の場合)の場合、処理ログは、PDF No Save・HTMLモード・コマンドライン対応版(exe)がインストールされているフォルダーのサブフォルダー「log」の中にあります「log/report日付.txt」に記録されますが、これを変えることができます。案件ごとに変えたい場合などは、指定してください。パスはフルパスで指定してください。普通は拡張子は.txtになるでしょうが、それにこだわる必然性はないです。
note	識別用メモです。日付や設定についての一口メモ、プロセスIDなど、何でもご記入ください。使用できる文字は、Shift_JISに含まれる文字です。日常使うたいの文字はOKです。
same_name	<p>同名のHTMLファイルやJSファイルが既に存在している場合の処理。値は下記の1から4になります。未設定の場合は、1として処理されます。</p> <ol style="list-style-type: none"> 1. 上書きする(デフォルト) 2. プログラムで別のファイル名※に自動で行います。 <ul style="list-style-type: none"> ※ 元のファイル名に、時間を数字にしたものを追加します。 ※ 出力フォルダーに前回行ったバッチ処理のファイルが残っている場合、大量に別名のHTMLファイルやJSファイルができることになり、また、そのHTMLファイルにリンクする別のHTMLファイルのソースも手動で書き換えなければならないこととなりますので、御注意下さい。 3. エラーメッセージを表示し、その場でバッチ処理を中止する。 4. スキップする <p>処理したいPDFファイルをこまめに追加されるシステムでご利用の場合などは、余計な時間をかけることなく、重宝すると思いません。</p> <ul style="list-style-type: none"> ※ 最初に処理済のPDFファイルをリストから削除する処理を行いますので、ボタンクリック後、実際の処理開始までに少し時間がかかる場合があります。 ※ PDF No Saveをバージョンアップさせた場合は選ばないでください。)

パラメータ	説明
batch_safari	<p>Safari 5.1.2以上6.0.5以下を利用している場合、Safari内蔵のPDFビューアでの閲覧を許可する場合で、Mac Safari内蔵のPDFビューアで正常に表示できないタイプのPDFファイルの場合、どうしますか？値は下記の1から4になります。未設定の場合は、1として処理されます。</p> <ol style="list-style-type: none"> 1. Mac Safari内蔵のPDFビューアの場合は、Adobe Reader 10.1.3以上にアップグレードするように促すページ(safari_snow_sh.html)に転送します。(デフォルト) 2. Mac Safari内蔵のPDFビューアの場合は、PDFファイルが90度もしくは180度回転している状態で表示されているのは理解しているが、そのまま表示する。 3. Mac Safari用に補正して、別ファイル名で保存したPDFファイル(ファイル名に「_safari_sh」が付きます。)とオリジナルのPDFファイルの両方を使用します。 ※ Mac Safari5.1.2以上6.0.5以下(正確には、Adobe Reader10.1.3以上がインストールされていない環境)でアクセスした場合には、補正されたPDFファイル(ファイル名に「_safari_sh」が付いているもの)を表示し、IEやFirefoxなど他のブラウザには、オリジナルのものが表示されることになります。
copy_no_error_page	<p>(上級者用設定)エラー用共通ページをコピーしない設定の場合はcopy_no_error_page=1を指定してください。あらかじめ、別アプリ(有償)のコピーツールでコピーしておくか、後から別アプリのコピーツールでコピーするかどちらかしないと、100%動作しないことになりますので御注意ください。 指定がない場合は、エラー用共通ページはコピーされます。</p>
preview_url	<p>ローカルサーバを構築されている場合、処理済のHTMLファイルをすぐにプレビューしたい場合、</p>

パラメータ	説明
html_tag_path New!!	<p>デフォルト(未指定の場合)の場合、PDFページ・リンク用のHTMLタグ(リストタイプのサンプル)は、PDF No Save・HTMLモード・コマンドライン対応版(exe)がインストールされているフォルダーのサブフォルダー「log」の中にあります「log/tag日付.txt」に記録されますが、これを変えることができます。案件ごとに変えたい場合などは、指定してください。パスはフルパスで指定してください。普通は拡張子は.txtになるでしょうが、それにこだわる必然性はないです。</p> <p>また、出力が不要な場合は、「html_tag_path=none」を指定してください。処理時間が少しだけ短縮されたり、負荷がいくぶん軽減されるはずですので、未処理のファイルなどが出ている場合は、一度お試しください。</p>
wait_msec New!!	<p>一つのPDFファイルを処理後、2つ目のPDFファイルを処理するまでの休止時間(ミリセカンド)を指定できます。1秒は1000ミリセカンドです。デフォルト(未指定の場合)は200ミリセカンドです。数値は100以上1000以下の数字を指定できます。半角数字で指定してください。</p> <p>弊社テストでは、この数値が110や120の場合、2000件のデータ処理で1件、未処理のPDFファイルが発生しました。この数値が100未満の場合、2000件のうち5件近く未処理が発生しましたので、100未満の数値は指定できないようにしてあります。200(ミリセカンド)の場合は、未処理のPDFファイルはゼロになりましたので、未指定の場合(デフォルト)は200にしてありますが、もし、それでも未処理のエラーが発生するようなら、300など数値を大きめにしてください。未処理のPDFファイルが発生する場合の、トラブルシューティング的なパラメータです。必要なければ、使わなくていいパラメータです。</p>

3. エラーコードとエラーメッセージ一覧

エラーコード	エラーメッセージ	備考
cmd001	コマンドライン版の引数の数が間違っています。書式をご確認ください。	引数の数は3つもしくは4つになります。1つか2つか指定されていない場合、このエラーメッセージがalert形式で表示されます。
cmd002	何らかの理由で設定ファイルの読み込みに失敗しました。設定ファイルが存在しないか、壊れている可能性があります。	指定された設定ファイル(拡張子pnh)が存在しないか、壊れている可能性があります。引数を見直してください。パスが正しい場合は、設定ファイルを作成しなおしてください。
cmd003	保護対象のPDFファイルの入ったフォルダーのドライブ名は大文字で指定してください。	e:¥homepage¥desktop¥hoge¥pdfなどを指定した場合。Eドライブであれば、「E:¥～」と指定してください。
cmd004	出力フォルダーのドライブ名は大文字で指定してください。	e:¥homepage¥desktop¥hoge¥outputなどを指定した場合。Eドライブであれば、「E:¥～」と指定してください。
cmd005	一度にバッチ処理するファイル数の上限は半角数字で指定してください。	第4引数の「_batch_limit=」の右辺の数字が全角であったり、数字以外の文字が混じっています。
cmd006	一度にバッチ処理するファイル数の上限は500です。	第4引数の「_batch_limit=」の右辺の数字が501以上である場合。「_batch_limit=500」なら500ずつ処理されるという意味であり、ファイル数が501以上あっても、最終的には全ファイルが処理されますので、ご心配不要です。
cmd007	PDFファイルのフォルダーパス:「●●●」の出力フォルダー(△△△)から見た相対パスは、「../../../~/」のように上位階層を表す「..」が4つ以上になります。	オリジナルのPDFファイルが格納されている入力フォルダーと出力フォルダーとが離れすぎています。相対パスで「../../../../」のようなパス関係の場合、DocumentRootの位置にも拠りますが、PDFファイルが表示されないディレクトリーに設置されている可能性も大きいいため、エラーになります。ディレクトリー構造を見直してください。

エラーコード	エラーメッセージ	備考
cmd008	指定されたログファイルの格納フォルダー: 「●●●」は存在しません。	ログファイルのテキストファイルは自動生成されますが、フォルダーまでは作成しません。フォルダーはあらかじめ存在しているフォルダーを指定してください。
cmd009	指定されたエラーログファイルの格納フォルダー: 「●●●」は存在しません。	このエラーメッセージは、デフォルトのエラーログ (HTMLモードの格納ディレクトリーの下にlogというフォルダーの中にあるerror_log日付.txt) に記録されます。
cmd010	プレビュー用のURLはスラッシュで終わるようにしてください。	第4引数の「_preview_url=」の右辺のURLの末尾がスラッシュ(/)で終わっていません。出力フォルダーに対応するWEBサーバ上のディレクトリーのURLですから、スラッシュで終わるように指定してください。
cmd011	プレビュー用のURLはhttp://もしくはhttps://で始まるようにしてください。	第4引数の「_preview_url=」の右辺のURLが不正である場合。
cmd012	オリジナルのPDFファイルの格納フォルダーのドライブ名と出力フォルダーのドライブ名が異なります。	オリジナルのPDFファイルがEドライブにあるのに、出力フォルダーがJドライブに指定されている場合など。
cmd013	指定された出力フォルダー: 「●●●」 は存在しません。	引数に指定された出力フォルダーが存在しない場合。
cmd014	指定されたPDFフォルダー: 「●●●」 は存在しません。	引数に指定されたPDFフォルダーが存在しない場合。
cmd015	指定されたPDFフォルダー: 「●●●」 には、処理可能なPDFファイルが一つも存在しません。	引数に指定されたPDFフォルダーにPDFファイルが存在しない場合。

エラーコード	エラーメッセージ	備考
cmd016	<p>指定されたPDFファイル: 「●●●」 は存在しません。バッチ処理自体は行われていますが、恐らくPDFファイルは表示されません。</p>	<p>日本語を使っている、Shift_JISに含まれない文字(「搾菜.pdf」など)や特殊な記号などが使われている場合。通常、ありません。該当PDFファイルのファイル名を変更してください。 その他の可能性としては、バッチ処理中に、PDFファイルを削除したり、PDFファイルのファイル名をリネームした場合は考えられます。</p>
cmd017	<p>PDFファイルのパスの取得に失敗しましたので、リロードします。</p>	<p>負荷がかかっているときなどに、なりやすいです。Debug情報。このエラーメッセージは、error_log_pathを指定している場合でも、多くの場合、PDF No Saveのインストールディレクトリの下のlogフォルダーに記録されます。</p>
cmd018	<p>保護対象のPDFファイル(●●●)は、iPhone/iPad/iPod touchや、Mac版Safari+Previewでの表示の最適化に必要な情報を得ることができませんでした。これらの環境をサポートする設定にされた場合は、該当環境でのテスト・微調整を念入りに行ってください。 可能性としては、PDFファイルが壊れているか、オープンパスワードが設定されているPDFファイルである可能性があります。</p>	<p>ページ数やページサイズの取得に失敗しています。</p>
cmd019	<p>●●●.pdfの処理に失敗している可能性があります。 error_code=cmd019 HTMLファイルのファイル名は○○○.htmlなのに、JSファイルのファイル名が▲▲▲.jsであり、ファイル名が異なります。特に注意してご確認ください。 ファイル名に日本語などの2バイト文字や、1バイトの文字でも一部特殊な文字が含まれている場合は確実にこのメッセージが記録されますが、その場合は多くの場合、心配されなくても大丈夫です。</p>	<p>Debug情報。未処理が疑われるケースです。出力フォルダーに「○○○.html」が存在するかどうかご確認ください。</p>

cmd019a New!!	(【御注意】未処理の可能性があるので、このファイルについては再チャレンジします。error_code=cmd019a)	このログは、エラーログではなく、処理ログに記録されます。エラーログに載せるまでもないが、注意が必要なため、掲載しています。このエラーが多発する場合(100件に1件以上)、wait_msecパラメータの値が小さすぎる可能性があります。
エラーコード	エラーメッセージ	備考
cmd019b New!!	●●●.pdfの処理に失敗している可能性があったので再チャレンジしましたら、3回目の再チャレンジで成功しました。	このエラーが多発する場合(250件に1件以上)、wait_msecパラメータの値が小さすぎる可能性があります。 未処理が疑われるようなケースでは結局、cmd019a→cmd019b→cmd019の順に問題が重大であることを示しています。cmd019aやcmd019bはともに警告的な意味合い(黄信号)であり、失敗(赤信号)ではありません。
cmd020	下記のファイルは、日本語を含むファイルパスですが、お客様の設定では、このファイルパスは処理できません。これらのファイル进行处理するためには、出力フォルダーの設定画面で、日本語を含むファイルパスを許可するように設定変更していただく必要があります。	HTMLモードをGUIで起動していただき、出力フォルダーの設定で日本語のパスを許可する設定にされると、このエラーは表示されなくなります。
cmd021	一つのPDFファイルを処理後、2つ目のPDFファイルを処理するまでの休止時間(ミリセカンド)は半角数字で指定してください。	第4引数の「_wait_msec=」の右辺の数値が不正である場合。
cmd022	一つのPDFファイルを処理後、2つ目のPDFファイルを処理するまでの休止時間(ミリセカンド)は100以上1000以下の数値で指定してください。	第4引数の「_wait_msec=」の右辺の数値が不正である場合。
cmd023	定されたPDFページ・リンク用のHTMLタグ(リスト)の格納フォルダー: 「●●●」は存在しません。	第4引数の「html_tag_path=」の右辺の数値が防いである場合。HTMLタグリストは自動生成されますが、フォルダーまでは作成しません。フォルダーはあらかじめ存在しているフォルダーを指定してください。

cmd024	●●個PDFファイルのうち、●●個全部のPDFファイルが処理済みであったため、バッチ処理を中止します。	第4引数で「same_name=4」を指定している場合で、全てのPDFファイルが処理済であるケースに記録されます。
エラーコード	エラーメッセージ	備考
cmd025 New!!	バッチ処理リストの作成に失敗しました。多重起動されていますか？	バッチ処理リストのファイル名がかぶってしまっている場合です。多重起動されている場合のみ起こりうるエラーです。また、確率的にも、理論的には1億分の1の可能性もないです。(かといって、多重起動での動作を保証するものではありません。)
cmd026batch New!!	予期しないエラーが発生しました。バッチ処理リストファイル: tmp/●●●.txtが存在しません。	バッチ処理リストのファイルが何らかの理由で存在しません。PDF No Saveのexeファイルのインストールフォルダーの下のサブフォルダー tmpフォルダーにバッチ処理リストファイルは本来なら存在するはずで
cmd026start New!!	予期しないエラーが発生しました。バッチ処理リストファイル: tmp/●●●.txtが存在しません。	バッチ処理リストのファイルが何らかの理由で存在しません。PDF No Saveのexeファイルのインストールフォルダーの下のサブフォルダー tmpフォルダーにバッチ処理リストファイルは本来なら存在するはずで
cmd026save New!!	予期しないエラーが発生しました。バッチ処理リストファイル: tmp/●●●.txtが存在しません。	バッチ処理リストのファイルが何らかの理由で存在しません。PDF No Saveのexeファイルのインストールフォルダーの下のサブフォルダー tmpフォルダーにバッチ処理リストファイルは本来なら存在するはずで

4. その他の御注意

- ・多重起動に関しましては、動作保証できません。
- ・処理中、ウィンドウを非表示にすることはできません。あらかじめサイズを指定していただければ、極力小さなウィンドウにすることは可能です。
- ・(重要)1ファイルの処理に約3秒かかります。1つ目のPDFファイルの処理の際、共通ページのコピーもなされるので、1つめのPDFファイルのみ処理に時間が余計にかかります。
- ・弊社では約2,200のファイルでの一括処理のテストを行いました。この数について動作保証するものではありません。
- ・(既知の問題)バッチ処理中に「バッチ処理を中断する」ためのリンクが出現しますが、効きが悪いです。止めたい場合は、ウインド右上の×印をクリックして、ウインドウごと閉じてください。
- ・(仕様)same_name=4を指定している場合でも、日本語のファイル名の場合、毎回HTMLファイル名が変わるため、処理済かどうか判断できないため、結果的に日本語のPDFファイルの場合、毎回スキップされませんが、バグではありません。
- ・もし、処理中に謎のスクリプトエラーやその他PDF No Saveが出していると思われるエラーメッセージが表示される場合は、(可能であれば、その画面をキャプチャーしていただいた上で)バッチ処理中、PC上で他の作業をできるだけストップするなどしてみてください。負荷がかかると想定外のエラーが起こりやすいです。
- ・余計なトラブルを避けるため、PDF No Save・HTMLモードのインストールディレクトリーのパスは、「日本語のパス」「C:\Program Files以下」「空白を含むパス」は避けてください。
- ・第4引数は一応任意ですが、処理ログのパス及びエラーログのパスの設定は強く推奨します。また、処理のたびに必ず確認するようにしてください。
- ・仮に複数の端末でバッチ処理をする場合でも、必ず、その端末で一度はGUIで処理を行ってください。設定ファイルだけ、新しい端末に移動しても動作しません。また、詳細設定ウィザードのステップ2などに注意書きを読まないで次に進めないところがたくさんありますが、それらを全部先にその端末で読んでおいて頂かないと、バッチ処理が途中で止まります。
- ・バッチ処理リストは、一時ファイルとして、PDF No Saveのexeのインストールフォルダーの下の「tmp」フォルダーの下に数字14桁のファイル名もしくは、数字とアンダーバーからなるテキストファイルです。バッチ処理後、バッチ処理リストファイルは通常削除されますが、異常終了した場合や強制的に途中終了させた場合、一時ファイルが残ります。もし、セキュリティ的に問題がある(情報漏洩の)場合は、定期的に手動でファイルを削除してください。タイムスタンプを見ていただき、12時間以上経過しているファイルは削除して問題ありません。バッチ処理中でなければ、tmpフォルダー内のファイルは全部削除してもらってもいいです。