

xdoc2txt - конвертор PDF, Word, Excel, RTF
в текст из различных документов, таких
как Ichitaro

ВНИМАНИЕ! *Использован машинный перевод с японского ибо в ином виде описание найти не удалось. Посему и стилистические огрехи, ну а в основном (кроме лицензионной части) выполнено техническое редактирование текста.*

xdoc2txt

Обзор

xdoc2txt это универсальный конвертер для извлечения текстовых элементов из различных форматов документов таких как PDF, Word, Excel, Ichitaro работающий из командной строки Windows.

- xdoc2txt анализирует структуру различных документов самостоятельно не используя такие программы как WORD или Acrobat.
- Он быстро работает что делает его идеальным для фильтров движков полнотекстового поиска.
- Поддерживаются следующие форматы документов:

.rtf	Форматированный текст (Microsoft RTF)
.docx	Microsoft Word 2007/2010/2013 (OOXML)
.xlsx	Microsoft Excel 2007/2010/2013 (OOXML)
.pptx	Microsoft PowerPoint 2007/2010/2013 (OOXML)
.doc	Microsoft Word ver5.0 / 95/97/2000 / XP / 2003
.xls	Microsoft Excel ver5.0 / 95/97/2000 / XP / 2003
.ppt	Microsoft PowerPoint 97/2000 / XP / 2003
.sxw / .sxc / .sxi / .sxd	OpenOffice.org
.odt / .ods / .odp / .odg	Open Document v1.2
.jaw / JTW	Ichitaro v5

.jbw / juw	Ichitaro v6
.jfw / Jvw	Ichitaro v7
.jtd / JTT	Ichitaro ver8/9/10/11/12
.oas / OA2 / OA3	OASYS / Win
.bun	New Pine / Pine 5 / Pine 6
.wj2 / wj3 / WK3 / WK4 / 123	Lotus 123
.wri	Windows3.1 Write
.pdf	Adobe PDF
.mht	Веб-архив
.html	W3C HTML
EML	Формат экспорта OutlookExpress

Начиная с v2.0 поддерживаются Ifilter так что если формат документа не поддерживается xdoc2txt, вы можете извлечь текст если есть соответствующий Ifilter.(функция доступна только в EXE версии)

Для EXE / DLL версии COM-компонента возможности по извлечению текста аналогичны.

Операционная система:

xdoc2txt работает в следующими ОС.

вер	Операционная система
ver1.x (MBCS)	Windows 95/98 / ME / NT 4.0 / 2000 / XP / Vista / Windows 7 (32bit / 64bit) / Windows 8 (32bit / 64bit) / Server 2003 / Windows Server 2008 R2 (64-разрядная) / Windows Server 2012 (64-разрядная) / Windows Server 2012 R2 (64-разрядная)

ver2.x (Unicode)	2000 / XP / Vista / Windows 7 (32bit / 64bit) / Windows 8 (32bit / 64bit) / Server 2003 / Windows Server 2008 R2 (64-разрядная) / Windows Server 2012 (64-разрядная) / Windows Server 2012 R2 (64-разрядная)
---------------------	--

- Для v2.0, требуется установка общих библиотек времени исполнения (для Vista и новее обычно не нужно)

Xdoc2txt 32bit (x86) версия требуется 32/64 битная редакция ОС и

[Microsoft Visual C++ 2010 Redistributable Package \(x86\)](#)

Для работы xdoc2txt 64bit (x64) версии требуется 64 битная редакция ОС и

[Microsoft Visual C++ 2010 Redistributable Package \(x64\)](#)

Если ОС выводит сообщение о невозможности запуска xdoc2txt из-за ошибок в конфигурации приложения, то скачайте и установите *Microsoft Visual C++ 2010 Redistributable Package (runtime)* с сайта Microsoft.

Авторское право и условия использования

- xdoc2txt быть свободно использован для некоммерческих и личных целей, в том числе в

- некоммерческих общественных организациях
- Для любого коммерческого использования xdoc2txt необходимо приобрести лицензию для чего обращайтесь к автору.

[xdoc2txt коммерческой лицензии](#) (24 мая, издание 2007 г.)

- Xdoc2txt, что входит в Hyper Estraier, если вы хотите распространять с Hyper Estraier, он будет исключен из коммерческой лицензии.
- Если вы хотите, чтобы перераспределение xdoc2txt, спасибо вам всегда принимать разрешение распределения автору. В случае свободного программного обеспечения, это не позволит принцип распределения.

Если вы хотите, чтобы перераспределить xdoc2txt, пожалуйста распространяться без изменения весь файл, включенный в пакет. Также, пожалуйста, указать, где пользователь известен руководства, такие, как то, что вы с помощью xdoc2txt.

Следует отметить, что если программное обеспечение спецификации, которые могут включать xdoc2txt в качестве внешнего фильтра (если вы используете отдельный xdoc2txt, что вы скачали), свяжитесь с автором абсолютно ненужным.

- xdoc2txt и авторских прав и сопроводительная документация являются собственностью их hishida.
- xdoc2txt предоставляется как есть без гарантии. любые убытки, возникающие из использования или неиспользования xdoc2txt (в отношении упущенной выгоды, прерывания бизнеса, в том числе других денежных убытков и потери деловой информации), автор не несет никакой ответственности.
- Почта и включение журнала Software, чтобы позволить перепечатка в Интернете. Если он будет опубликован, пожалуйста, свяжитесь с нами для журнала публикации до автора, потому что это прекрасно в докладе сообщению.

Параметры командой строки

```
xdoc2txt.exe [опции ..] <имя файла ...>

-h : вывести справку

-s :
    выходная кодировка ShiftJIS (по умолчанию)

-e :
    выходная кодировкаEUC

-j : JIS
-u : UTF-16 LE) (рекомендуется использовать с Ifilter)
-8 : UTF-8

-f : выводить в файл
```

```
-p : показать свойства документа
-r=(0|1|2) : стиль скобок (0: как есть, 1:
, 2: aozoga bunko)
-o=0 : прочие опции;
-o=0 : не выводить в номера страниц PDF
-o=1 : удалить разрывы строк PDF
-g=# :
диапазон масштаба PDF в процентах (по умолчанию 92%)
-v : вывести номер версии
-x : вывести только видимые ячейки (для EXCEL2007)
-i : по возможности использовать IFilter (совместимо с UTF-8 , UTF-
```

Следующие опции были удалены в версии 2.0 как устаревшие

и игнорировать ограничения доступа в документах PDF (н

-с кэш PDF (по умолчанию выключена)

*<имя файла ...> имя выходного файла (если содержит пробелы, то обяза
Допустимо использовать шаблоны подстановки "*" и "?".*

Как использовать

- Следующий пример, пишет текст, включенный в sample.doc документа MS-Word на стандартный вывод.

```
xdoc2txt sample.doc
```

При перенаправлении потока вывода он может быть сохранены в файл.

```
xdoc2txt sample.doc> sample.txt
```

Если использовать опцию -f, то можно сохранить вывод в файл с автоматическим расширением .txt.

```
xdoc2txt -f sample.doc sample.xls
```

Шаблоны *? можно использовать в пакетном задании для указания набора входных файлов.

```
xdoc2txt -f *.xls
```

Для документов Office и Ichitaro (ver8 и новее), можно отобразить свойства документа задав опцию -p. Будут отображены только установленные элементы.

```
xdoc2txt -p руководство .doc
```

Результат выполнения:

```
<Title> KWIC Finder руководство </Title>
<Autor> hishida </Autor>
<Template> Normal.dot </Template>
<LastAuthor> hishida </LastAuthor>
<RevisionNumber> 1 </RevisionNumber>
<AppName> Microsoft Word 9.0 </AppName>
<Lastprinted> 2004/03/23 19:39:00 </Lastprinted>
<Created> 2004/03/23 19:35:00 </Created>
<LastSaved> 2004/03/23 19:44:00 </LastSaved>
```



```
<PageCount> 1 </PageCount>  
<WordCount> 21 </WordCount>  
<CharCount> 121 </CharCount>
```

- Защищённые паролем документы Word / Excel / PowerPoint / Ichitaro не обрабатываются.

Перетаскивание мышью

Создайте ярлык на рабочем столе со следующими свойствами:

1. Правая кнопка мыши → [отправить (N)] на xdoc2txt.exe в Проводнике Windows → [Рабочий стол (создать ярлык)]
2. Правая кнопка мыши на иконке, которая была создана на рабочем столе → [Свойства (R)]
3. В конце Target (T)], и добавить -f.
Пример: "C:\Program Files\kwic\xdoc2txt.exe" -f
4. Если вы перетащите файлы на этот значок, то выходные файлы с расширением .txt будут сохранены в той же директории что и входные.

Ссылка на статью:

[Http://Www.Forest.Impress.Co.Jp/article/2003/11/19/xdoc](http://www.Forest.Impress.Co.Jp/article/2003/11/19/xdoc)

.

О Ifilter

- В Ver2.0 или новее поддерживает Ifilter. Если указать опцию -i, Ifilter могут быть использованы для преобразования.
- Мы проверили следующий Ifilter.
 - Ichitaro для IFilter 32-битной ОС
 - DocuWorks Content-фильтр
 - Microsoft Office Update filter
 - Adobe Reader 9.5 PDF Ifilter
 - ※ Adobe Reader 10 Ifilter вышел после выхода Adobe PDF IFilter v6.0, Adobe PDF IFilter 9 для 64-битных ОС не доступен.
- Ifilter работает только с EXE сборкой и с другими (DLL/COM) не применим.

Скачать

Ver2.x (Unicode)

Новое! 2017/07/06

- [xdoc2txt 2.16.1 \(xd2tx2161_x64.zip\) - x64 \(64 бит\) версия](#)
- [xdoc2txt 2.16.1 \(xd2tx2161.zip\) x86 \(32 бит\) версия](#)

Ver1.x (MBCS)

- [xdoc2txt1.52 \(d2txt152.zip\)](#)

Фильтры

Имя	Лицензия	Применение	
GoogleXdoc (подключение xdoc2txt к GoogleDeskTop как плагина)	бесплатно	Полнотекстовый поиск	Soft far
Namazu для Win32	бесплатно	Полнотекстовый поиск	Образец фильтрации до (Mr. by a.hanai) Система полнотекстовой
Hyper Estraier	бесплатно	Полнотекстовый поиск	http://hyperestraier.sourceforge.net/
Meadow2	бесплатно	Редактор	Meadow memo Wiki
MiGrep	бесплатно	Поиск	M & I page of
VxEditor	бесплатно	Редактор	X-Labo WebPage
smoory	бесплатно	Просмотр вертикального текста	http://www.vector.co.jp/soft/win32/other/s000101.html
Transwise	бесплатно	Поддержка перевода	http://www6.ocn.ne.jp/~vtranswise/
EBView	бесплатно	Словарь текстовый поиск	http://ebview.sourceforge.net/
Search Cross	Продукт	Полнотекстовый поиск	Village Center Co., Ltd.
KOA Direct Server	бесплатно (с ограничениями)	Система обмена данными	KOA Project
HNXgrep	бесплатно	Поиск с помощью Grep	http://www.vector.co.jp/soft/win32/other/s000101.html

* Известные автору программы использующие xdoc2txt в качестве фильтра.

История

Ver2.x (Unicode)

- 2.16.1 2017/07/06
 - Добавлена 64 битная (x64) версия. * Требуется Microsoft Visual C ++ 2010 Redistributable Package (x64).
- 2.16.1 2016/06/28
 - Удалена зависимость от VC90.CRT в манифесте
- 2.16 2016/04/26
 - Исправлена ошибка отображения в списке When кодов форматирования в XLSX
- 2.15 2016/04/07
 - Исправлена ошибка показа в тексте невидимых управляющих символов DOCX
- 2.14 2015/11/19
 - Устранён крэш на некоторых файлах PDF
- 2.13 2015/8/25
 - Исправлена ошибка потери символов в прозрачном Scan Snap-тексте
 - Исправлена ошибка чтения некоторых PDF зашифрованных AES-128 (требуется cryptlib.dll)
- 2.12 2015/7/18
 - Уточнение параметров английского текста
 - Переход на C++ 2010 с Visual C++ 2008

- | | | |
|------|------------|---|
| 2.11 | 2015/5/29 | <ul style="list-style-type: none"> • Excel2007 улучшение производительности Формат извлечения текста из (.xlsx) |
| 2.10 | 2015/4/15 | <ul style="list-style-type: none"> • ошибка фиксирует ком версии файла после PDF в ExtractText() не был выпущен • Добавить функцию ExtractTextEx() для COM и Dll версии. Опция использовать командной строки. (-г - о -g X действителен только) |
| 2.09 | 2015/4/09 | <ul style="list-style-type: none"> • Дело модификации, которые не могут быть отображены лигатуры в формате PDF • Исправлена проблема, что аварийное завершение часть .mht • Добавить программу образец FileFind помощью Dll |
| 2.08 | 2015/3/11 | <ul style="list-style-type: none"> • Перейти Если расширение содержание в формате .pdf не в формате PDF • Улучшенная утечка памяти в момент извлечения PDF • Явное Cdecl попробовать программу для вызова Dll |
| 2.07 | 2014/10/28 | <ul style="list-style-type: none"> • Исправлена проблема, что аварийное завершение в определенной XLSX |
| 2.06 | 2014/10/09 | <ul style="list-style-type: none"> • Исправлена ошибка, которая не может извлечь характер изменения только цвет символов часть строки в ячейке в XLSX |
| 2.05 | 2014/08/31 | <ul style="list-style-type: none"> • Насколько это возможно, избегайте аварийному завершению в поврежденной PDF |

- 2.04 2014/07/29
 - Фиксированная часть японской электронной почте искажены в .eml
 - Команда версия 256 предельный размер входного файла (-z=, который может быть настроен)

- 2.03 2014/07/16
 - Исправлена проблема, когда некоторые .xlsx к краху

- 2.02 2014/06/14
 - [Рекомендуется, если вы хотите, чтобы постоянно использовать] образец добавил динамически загружать и выпустить в LoadLibrary() и FreeLibrary() версии Dll

- 2.02 2014/05/04
 - Исправлена проблема, что аварийное завершение часть PDF

- 2.01 2014/02/16
 - Половина ширины Кана соответствие СКП код

- 2.00 2013/01/23
 - Официальный версия

- 2.00β4 2012/12/28
 - исправления

- 2.00β3 2012/12/24
 - Исправлена новая линия стандартный вывод стал для \r \r \n проблему
 - Исправлена проблема с аварийного завершения Если вы используете опцию -f в письменной форме не каталога
 - Фиксированная часть рубина документа Word, не отображаются
 - Установленный проблема, которая не может быть извлечена часть характера PDF для вывода в Word 2007

- 2.00β2 2012/12/19
 - Ошибка фиксирует быть прекращено в какой-то ненормально .odt
 - Исправления к аварийному завершению в документ длины 0

- 2.00β1 2012/12/01
 - Добавлен ресурс версия для xdoc2txt.exe (чтобы отличить его от серии 1.x)

- 2.00β0 2012/11/26
 - Исправлена проблема, чтобы заморозить часть PDF

- 2.00α3 2012/11/17
 - Office 2007 / 2010 относительно фиксированной лицо не интерпретирована в документе

- 2.00α2 2012/11/15
 - VC++ образец того, делает аргумент заказа из DLL версии же, как COM версии.

- 2.00α1 2012/11/14
 - Добавить версию COM-компонент (xd2txcom.dll)

- 2.00α0 2012/11/13
 - Мы делаем внутреннюю снижение Unicode. Изменить компилятор из VC++ 6.0 для VC++ 2008.
 - -и к опциям (utf-16), - добавлено -8 (UTF-8)
 - соответствующая Ifilter (опция -i). даже расширение xdoc2txt не соответствуют родной, она может быть выполнена извлечение текста, если есть соответствующий IFilter.
 - / Соответствует закодированной в формате PDF в LZWDcode (патент Unisys истек).
 - Она объединяет cryptlib.dll, соответствующий стандарту с зашифрованным PDF с паролем бесплатно.

- Это обеспечивает DLL версию (xd2txlib.dll). Он также поставляется с образцом звонить из C # и VB.Net

Ver1.x (MBCS версия)

**Разработка MBCS версии 1.x закончена.
Пожалуйста, используйте v2.x.**

- | | | |
|------|------------|---|
| 1.52 | 2015/11/19 | <ul style="list-style-type: none">• Устранён краш на некоторых файлах PDF |
| 1.51 | 2015/08/25 | <ul style="list-style-type: none">• Исправлена ошибка потери символов в прозрачном Scan Snap-тексте• Исправлена ошибка чтения некоторых PDF зашифрованных AES-128 (требуется cryptlib.dll) |
| 1,50 | 2014/10/28 | <ul style="list-style-type: none">• Исправлена проблема, что аварийное завершение в определенной XLSX |
| 1.49 | 2014/10/09 | <ul style="list-style-type: none">• Исправлена ошибка, которая не может извлечь характер изменения только цвет символов часть строки в ячейке в XLSX |
| 1.48 | 2014/05/04 | <ul style="list-style-type: none">• Исправлена проблема, что аварийное завершение часть PDF |
| 1,47 | 2013/11/30 | <ul style="list-style-type: none">• Исправления и в аварийном завершении включены длинный ссылку в документе Word, |

- 1.46 2012/12/24 • Установленная проблема, которая не может быть извлечена часть характера PDF для вывода в Word2007
- 1.45 2012/11/26 • Исправлена проблема, чтобы заморозить часть PDF
- 1.44 2012/11/17 • Поддержка искажен некоторых персонажей в DOCX
- 1.43 2012/10/17 • "Office Open XML, форматы файлов", чтобы исправить проблему, если будет прекращена ненормально в .xlsx которые не соответствуют
- 1.42 2012/05/17 • Исправлена проблема, которая может быть ненормальное завершение в рамках PDF
- 1.41 2011/07/30 • Исправления, что последний символ отсутствует в EXCEL в текстовом поле (ошибка, которая была смешанной с 1,37)
- 1.40 17, 2011 • Исправлена проблема, когда есть случай, когда отображение текста в некоторых из PDF без вести.
- 1.39 2011/04/28 • Исправлена проблема, когда дополнительные символы в части PDF отображается документ (/ , которая была использована в язык, указанный Lang)
- 1.38 2010/12/21 • модификация (часть условий в случае PDF1.5 или поздно / XRef используется) в аномальных проблем терминации часть PDF

- 1.37 2010/05/16
 - Чтобы справиться с проблемой PDF, которое было выходом части PDF писателя (Брава! Desktop) искажен
 - Исправлена случай, когда пыль поступает в EXCEL из текстового поля

- 1.36 2010/01/09
 - Исправлена проблема, когда кодирование СКП PDF искажен

- 1.35 2009/08/28
 - В соответствии с документом Office 2007 пустой

- 1.34 2009/06/22
 - Office 2007 документ, датированный пароль Я хочу, чтобы отображаться с "зашифрованного файла".
 - Zlib.dll от этой версии не требуется (изменение статической ссылке)

- 1.33 2009/06/07
 - Исправлена ошибка, когда аварийно с поврежденной PDF
 - Исправлена проблема, искаженные в документе OpenOffice с паролем (Я хочу, чтобы отобразить "зашифрованный файл").
 - Исправлена случай, когда сбой в крайне большого документа Excel2007 число и номер строки сиденья.
 - -х добавлена опция. Показывать только самые клетки, которые существуют в Excel 2007

- 1.32 2008/12/01
 - Исправлена повреждены, что есть случай, когда бесконечный цикл в PDF
 - Из Acrobat7.0 позже, читая соответствие PDF, который был зашифрован с 128-битным AES (необходимо ввести cryptlib.dll)

- Исправлен искажённый в поврежденном файле

- 1.31 Ноябрь 2008 Office2007.
- Excel2007 Исправлена количество мест завершилась аварийно, если он превышает около 100 в
- 1.30 R2 2008/08/18
- Добавить ActiveX версии xdoc2txt.osx. Условия распространения такие же, как исполняемый версии.
- 1.30 2008/05/22
- Опция -p: добавить "название компании", "объявление", "имя администратора" на дисплее собственности документа офиса
- 1.29 2008/05/18
- Исправлена проблема, когда есть, чтобы быть прекращено аномально в PDF файле, который был создан в формате PDF создания программного обеспечения, кроме Acrobat
 - Исправлена проблема, что размер аварийно с 0 байт PDF
- 1,28 2008/03/18
- 1.27 2008/01/24
- Более именем пути составляет 256 байт входного файла и "ошибка в имени файла" и, кажется, решить эту проблему, которая не может быть обработана
- 1.26a 2007/10/21
- Новая строка исправить 0x0D, что есть Переполнение буфера в случае, HTML не 0x0A
- 1.26 2007/05/11
- Исправления, которые не видят часть колонны в Microsoft Office Excel 2007
 - Ошибка фиксирует два обзоры PDF (утечка дисплей, аномальные меры прекращения)

- 1.25 2007/04/13
 - Соответствующий дисплей имени поля ввода в Microsoft Word

- 1.24 2007/02/18
 - Microsoft Office Word 2007 / Excel 2007 / PowerPoint 2007, OpenOffice.org, и соответствует открытый документ
 - Ошибка исправлена аварийному завершению Если вы используете несколько большим количеством цифр, 1E+275 в EXCEL.

- 1.23 2006/08/29
 - Он поддерживает проблем, которые не могут быть извлечение текста из PDF, созданный с AntenaHouse PDF Driver2.0 (PDF1.5 соответствующих последующих перекрестных ссылок потоков)
 - По PDF, исправить последнюю проблему, что есть случай, который не может быть извлечение текста в файл

- 1.22 2006/05/28
 - Исправления по PDF кодирования " "" искажен

- 2006/05/10
 - Условия и условия изменения коммерческого использования

- 1.21 2006/05/08
 - Исправлена ошибка, что, когда вы ищите пароль с документа Ichitaro Ver6 из памяти

- 1.20 2006/02/17
 - В соответствии с картирования Unicode лигатур (FF, Интернет и т.д.) в формате PDF

- 1.19 2006/02/08
 - Профилактическая модификация Переполнение буфера в формате PDF

- 1.18 2006/02/04
 - PowerPoint95 переписка
 - Исправления в том, что содержимое строки не появляется все в EXCEL

- 1.17 2005/09/19
 - Добавить персонажа между параметрами настройки -g из PDF

- 1.16 2005/05/02
 - Добавить тип в /H /V от CMAP в поддержке PDF (кодирование JIS)
 - Если огромные цифры в формате PDF сжимаются в /FlateDecode, Исправлена ошибка, не выделить память
 - Дополнительный дополнительный в формате PDF -o = 0 в PDF - - Вы не хотите, чтобы отобразить номер страницы формы
 - o = 1, чтобы удалить разрывы строк в формате PDF (если это новая линия для каждого символа в вертикальном письменной форме)
 - Параметры вывода на рублином HTML
 - не -r = 0
 - r = 1 ()
 - r = 2 "" Формат Aozora Bunko

- 1.15 2005/04/23
 - Исправлена ошибка, которая не может быть частью текста в формате PDF, который был создан в Acrobat4
 - Исправлена ошибка, в формате PDF не может быть текст носить печать даже один раз в Acrobat

- 1.14 2005/01/31
 - Исправления к аварийному завершению в формате PDF созданы с Justsystem PDF Creator
 - ошибка фиксирует, что изображение только дело прекратить ненормально в PDF без текста в

- 1.13 2004/05/30
 - И регулировки расчет между PDF ди
 - WK4 ошибка фиксирует, что есть случай, чтобы быть в аварийно (123)

- 1.12 2004/05/05
 - Возможность игнорировать установку привилегий PDF документа доступа (-n)
 - Исправления, которые не видят половину ширины CID в формате PDF
 - Когда вы в стандартный вывод, исправления, что отображается дополнительная перевозка

- 1.11 2004/04/04
 - Шифрование без пароля была поддержка PDF (до 128-битное шифрование). Необходимо скачать отдельную cryptlib.dll
 - Корреспондент в формате PDF, который вы создали в "easyPDF 3.1" "Jaws PDF Creator"
 - Для решения проблем, вкладка характер удаленные в Ichitaro V7 позже
 - Добавить параметр -r. Просмотр свойств документа офиса

- 1.10 2004/03/13
 - Корреспондент в формате PDF, который был создан в OpenOffice.org.1.1

- 1.09 2004/02/25
 - Корреспондент в формате PDF, что было выходом ActiveReports 2.0J
 - Исправления к аварийному завершению в частности PDF

- 1.08 2004/01/28
 - Если появились строки результата формулы EXCEL двойного байта " $\pm \times \div$ ", исправления должны быть искажен полуширина Кана
 - Удаление дополнительной перевозки

- 1.07 2004/01/26
 - Word, Excel, есть случай, когда "± × ÷" двухбайтных символов перевранные в половину ширины Кана в PowerPoint исправлений
- 2004/01/18
 - Это будет предусматривать "Об авторском праве и условиями использования."
- 1.06 2003/11/09
 - Исправления, которые первая линия RTF, что вы сохранили в WordPad не отображаются
 - Если расширение не включены в документ Word, в документе OLE из .doc, аномальной расторжении исправить ошибку
- 1.05 2003/07/15
- 1.04 2003/03/26
 - Уточнение вычисления интервалов в японских PDF.
- 1.03 2002/11/23
 - Поддержки кодирования PDF Unicode
- 1.02 2002/10/18
 - соответствие MHT / HTML
- 1.01 2002/9/9
 - добавлена опция -c
- 1.00 2002/7/8
 - Отделение извлечения указанной в качестве фильтра части текста из KWIC Finder.