

Supported file formats include but are not limited to:

Adobe Framemaker MIF (*.mif)

Adobe Photoshop images (metadata only) (*.psd)

Ami Pro (*.sam)

Ansi Text (*.txt)

Apple iWork KeyNote 2009 (*.key)

Apple iWork Numbers 2009 (*.numbers)

Apple iWork Pages 2009 (*.pages)

ASCII Text

ASF media files (metadata only) (*.asf)

CSV (Comma-separated values) (*.csv)

DBF (*.dbf)

EBCDIC

EML (emails saved by Outlook Express) (*.eml)

Enhanced Metafile Format (*.emf)

Eudora MBX message files (*.mbx)

Flash (*.swf)

GZIP (*.gz)

HTML (*.htm, *.html)

iCalendar (*.ics)

Ichitaro (versions 5 and later) (*.jtd, *.jbw)

JPEG (*.jpg)

Lotus 1-2-3 (*.123, *.wk?)

MBOX email archives such as Thunderbird, including attachments (*.mbx)

MHT archives (HTML archives saved by Internet Explorer) (*.mht)

MIME messages, including attachments

MSG (emails saved by Outlook), including attachments (*.msg)

Microsoft Access 95, 97, 2000, 2003, 2007, 2010, 2013, and 2016 MDB (*.mdb, *.accdb)

Microsoft Document Imaging (*.mdi)

Microsoft Excel for Mac 2.2, 3, 4, 5, 98, 2001, X, 2004, 2008, 2011

Microsoft Excel for Windows 2, 3, 4, 5

Microsoft Excel 95, 97, 2000, XP, 2003, 2007, 2010, 2013, 2016 (*.xls)

Microsoft Excel 2003 XML (*.xml)

Microsoft Excel Office Open XML 2007, 2010, 2013, and 2016 (*.xlsx)

Microsoft OneNote 2007, 2010, and 2013 (*.one) (dtSearch 7.72 and later)

Microsoft Outlook 97, 2000, 2003, 2007, 2010, and 2013 data files, including attachments (*.PST, *.OST)

Microsoft Outlook/Exchange Messages, Notes, Contacts, Appointments, and Tasks

Microsoft Outlook Express 5 and 6 (*.dbx) message stores

Microsoft PowerPoint 3, 4, 95, 97, 98, 2000, 2001, 2002, 2003, 2004, 2007, 2008, 2010, 2011, 2013, 2016 (*.ppt)

Microsoft PowerPoint Office Open XML 2007, 2010, 2013, and 2016 (*.pptx)

Microsoft Rich Text Format (*.rtf)

Microsoft Searchable Tiff (*.tiff)

Microsoft Word for DOS 1, 2, 3, 4, 5, 6 (*.doc)

Microsoft Word for Mac 1, 3, 4, 5, 6, 98, 2001, X, 2004, 2008, 2011

Microsoft Word for Windows 1, 2, 6 (*.doc)

Microsoft Word 95, 97, 98, 2000, 2002, 2003, 2007, 2010, 2013, 2016 (*.doc)

Microsoft Word 2003 XML (*.xml)

Microsoft Word Office Open XML 2007, 2010, 2013, 2016 (*.docx)

Microsoft Works WP (*.wks)

MP3 (metadata only) (*.mp3)

Multimate Advantage II (*.dox)

Multimate version 4 (*.doc)

OpenOffice/LibreOffice versions 1, 2, 3, 4, and 5 documents, spreadsheets, and presentations (*.sxc, *.sxd, *.sxi, *.sxw, *.sxg, *.stc, *.sti, *.stw, *.stm, *.odt, *.ott, *.odg, *.otg, *.odp, *.otp, *.ods, *.ots, *.odf) (includes OASIS Open Document Format for Office Applications)

PDF files (*.pdf)

PDF Portfolio files (*.pdf), including embedded non-PDF documents.

Quattro Pro (*.wb1, *.wb2, *.wb3, *.qpw)

QuickTime (*.mov, *.m4a, *.m4v)

RAR (*.rar)

TAR (*.tar)

TIFF (metadata only) (*.tif)

TNEF (winmail.dat)

Treepad HJT files (*.hjt)

Unicode (UCS16, Mac or Windows byte order, or UTF-8)

Visio XML files (*.vdx)

Windows Metafile Format (*.wmf)

WMA media files (metadata only) (*.wma)

WMV video files (metadata only) (*.wmv)

WordPerfect 4.2 (*.wpd, *.wpf)

WordPerfect (5.0 and later) (*.wpd, *.wpf)

WordStar version 1, 2, 3 (*.ws)

WordStar versions 4, 5, 6 (*.ws)

WordStar 2000

Write (*.wri)

XBase (including FoxPro, dBase, and other XBase-compatible formats) (*.dbf)

XML (*.xml)

XML Paper Specification (*.xps)

XSL

XyWrite

ZIP (*.zip) (PKZIP 2.0-compatible)

Automatically-detected fields

The dtSearch Engine automatically detects fields in the following file formats:

File format	Fields
Email files (Outlook Express, Eudora, MBOX, EML)	To, CC, BCC, From, Sent Via, Sender, Recipient, Subject, Date, Attachments
Outlook items and .MSG files	To, CC, BCC, From, Sent Via, Sender, Recipient, Subject, Date, Sent Date, Delivered Date, Attachments, contact fields (StreetAddress, CompanyName, etc.)
Microsoft Word, Excel, PowerPoint	Document summary information fields
OpenOffice/Open Document Format	Document properties fields
HTML	META tags; <TITLE> is indexed as HtmlTitle field; <H1>, <H2>, <H3> are indexed as HtmlH1, HtmlH2, HtmlH3, etc.
XML	All fields
DBF	All fields
CSV	All fields (CSV, or comma-separated values, files must have a .csv extension, a list of field names in the first line, and must use tab, comma, or semicolon delimiters)
PDF	Document Properties
WordPerfect	Document summary information fields
MP3	All metadata fields
JPG, TIFF	EXIF and IPTC metadata fields; XMP (Vista) metadata supported in version 7.40
ASF, WMA, WMV	All metadata fields

Embedded Object and Attachment Extraction

Embedded objects and attachments are indexed as part of the document that contains them. For example, a spreadsheet object embedded in a PowerPoint presentation would be treated as part of the PowerPoint presentation.

Known Non-searchable file types

exe	jpg	jpeg	bmp	gif	dll	386
3g2	3gp	3gp2	3gpp	ac3	acb	acw
adts	ahu	aif	aifc	aip	air	alv
amc	ani	asf	au	ava	avi	aw
awf	bau	cda	cdc	cer	chm	clr
cov	cpe	cpl	crt	ctl	ctx	cue
cur	der	desklink	dess	deu	dib	dif
drv	dsx	duck	dun	dv	dvr-ms	dwi
dwp	elm	emf	eps	fh10	fh11	fh8
fh9	fhx	fla	lp	flv	fon	fra
ft10	ft11	ft8	ft9	gcsx	gi	glox
gqsx	icc	icm	ico	iii	ipa	ipg
ipsw	itdb	ite	itl	itlp	itls	itms
itpc	its	jfif	jpe	lex	m1v	m2v
m3u	m4r	mac	mid	midi	mml	mmm
mmw	mod	mov	mp2	mp2v	mp3	mp4
mpa	mpe	mpeg	mpf	mpg	mpv2	mqv
msi	msp	msrcincident	msstyles	mswmm	nra	nsd
oca	ocx	odc	odccubefile	otf	p10	p12
p7b	p7c	p7m	p7r	p7s	p951	pag
pal	pct	pfx	pgx	pic	pict	pko
png	pnt	pntg	psd	pt	pxr	qpa
qt	qti	qtif	qtl	qtp	qtr	qts
qtx	rar	rat	raw	rmf	rmi	rnk

Roxio	sd2	skin	skpkmz	snd	spa	spl
Swf	swt	theme	tif	tiff	tlb	torrent
ttc	ttf	vxd	wav	wbmp	wdp	webpnp
wm	wma	wmd	wmdb	wmf	wms	wmv
wmx	wmz	yimg	yps	7z		

Notes

Lotus Notes NSF Searching is available through Email options. Harvester will keyword search NSF files as loose files, but will not extract messages unless Notes Searching is selected under Email options