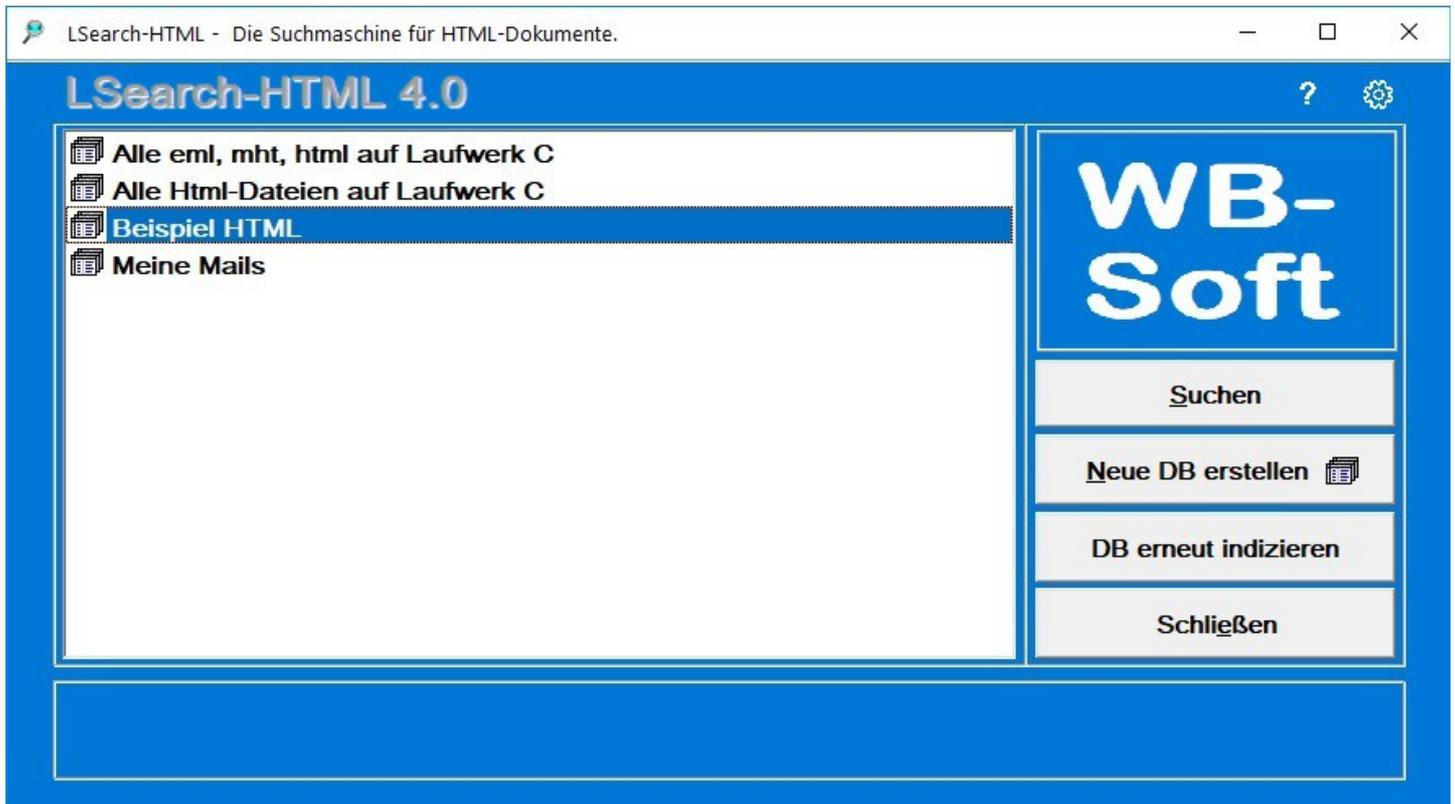


# LSearch-HTML

## Textsuche wie im Internet, aber lokal auf der Festplatte Ihres Computers.

LSEARCH-HTML ist eine Suchmaschine, die Ihnen das schnelle Durchsuchen von HTML, MHT, MHTML und EML-Dokumenten ermöglicht. Sie geben einen oder mehrere Suchbegriffe ein und bekommen eine Trefferliste von jenen Dokumenten angezeigt, in denen die Suchbegriffe zu finden sind. Die Trefferliste erscheint in Form einer HTML Seite in Ihrem Standard-Browser. Klicken Sie in der Trefferliste einen Verweis an, so bekommen Sie das HTML-Dokument im Browser angezeigt und können dort mit der Suchfunktion das Gesuchte markieren lassen. Damit Sie das Suchwort dort einfach über die Zwischenablage in das Eingabefeld einfügen können, kopiert LSEARCH-HTML es vorher in die Zwischenablage. Unter Einstellungen können Sie diese Funktion aber deaktivieren.



LSEARCH-HTML lässt sich überall dort einsetzen wo lokale Suchfunktionalität benötigt wird (z.B. auf einer Festplatte, Wechseldatenträger, Speicherkarte oder Speicherstick die mit dem Dateisystem NTFS formatiert wurden). Es wird eine Datenbank erstellt, welche die Suchbegriffe enthält. Die Suche erfolgt durch die Indizierung sehr schnell.

### **Was sind HTML-Dokumente:**

HTML (HyperText Markup Language) ist eine plattformunabhängige Dokumentbeschreibungssprache, in der Web-Seiten geschrieben sind. Dateien dieses Typs erhalten die Endung .htm oder .html

### **Was sind MHT-Dokumente:**

MIME Encapsulation of Aggregate HTML Documents (MHTML) ist ein Standard zum Speichern einer kompletten HTML-Seite in einer einzigen Archivdatei. Dateien dieses Typs erhalten die Endung .mht

### **Was sind MHTML-Dokumente:**

Das selbe wie MHT. Dateien dieses Typs erhalten die Endung .mhtml

### **Was sind EML-Dokumente:**

EML-Dateien sind gespeicherte E-Mail-Nachrichten. Dateien dieses Typs erhalten die Endung .eml

## **Das Indizieren der HTML-Dateien:**

Bevor Sie mit der Suche beginnen können, muss zuerst eine Datenbank angelegt werden (die HTML-Dateien werden indiziert). Auch eine Liste der Pfade und Dateinamen eines Laufwerks wird erstellt. Sie nennt sich „Haupt\_Dateinamen\_Liste\_Laufwerk\_C.txt“ für die Festplatte C: und wird nur bei Bedarf erstellt. Zum Beispiel wenn sie schon etwas älter ist und auf dem Laufwerk HTML-Dateien hinzugekommen oder gelöscht wurden. Dazu wird das Programm „ErstelleDateiliste.exe“ gestartet. Es braucht Administratorrechte.



### **So erstellen Sie Ihre erste Datenbank (Schnellerklärung):**

1. Betätigen Sie den Schalter „Neue DB erstellen“.
2. Geben Sie einen Namen für die Datenbank ein.
3. Die Endungen \*.htm und \*.html sind schon eingetragen. Durch betätigen des Schalters „Andere Endung auswählen“ können Sie diese ändern. Der Schalter „X“ rechts neben der Endungen-Eingabezeile löscht deren Inhalt.
4. Laufwerk bzw. den Pfad auswählen in dem nach den HTML-Dateien gesucht werden soll.
5. Betätigen Sie den "OK" Schalter. Externes Programm „ErstelleDateienliste.exe“ wird gestartet.
6. Das blinkende Symbol in der Taskleiste anklicken und die Frage mit "Ja" bestätigen. Die Hauptdateienliste wird erstellt. Danach werden die Dateien indiziert.

**Wenn Sie „Windows Defender“ als Antivirens Scanner nutzen, dann wird vor dem Indizieren Lsearch vom Defender überprüft. Das kann mehrere Minuten in Anspruch nehmen und Lsearch kann während dieser Überprüfung nicht arbeiten. Lsearch erkennt den „Windows Defender“ und zeigt an, dass Lsearch überprüft wird. Andere Antivirens Scanner erkennt Lsearch nicht. Also nicht wundern, wenn es am Anfang einer Indizierung minutenlang nicht weitergeht.**

### **Vorhandene Datenbank erneut indizieren (Schnellerklärung):**

Wenn sich Dateien verändert, gelöscht oder neue hinzugekommen sind, sollten Sie die Datenbank erneuern, indem die DB neu erstellt wird.

1. Betätigen Sie den Schalter „DB erneut indizieren“.
2. Der Namen für die Datenbank ist schon eingetragen.
3. Die Endungen sind schon eingetragen.
4. Das Laufwerk bzw. der Pfad ist schon eingetragen.
5. Betätigen Sie den "OK" Schalter.
6. Sie werden gefragt, ob die Hauptdateienliste erstellt werden soll.

### **So erstellen Sie Ihre erste Datenbank (mit Erklärung):**

Betätigen Sie den Schalter „Neue DB erstellen“, es erscheint obiges Fenster. Dort geben Sie einen Namen für die Datenbank ein. Danach noch das Laufwerk bzw. den Pfad in dem nach den HTML-Dateien gesucht werden soll auswählen. Als Voreinstellung sind die Endungen \*.htm und \*.html eingetragen. Wenn Sie \*.eml-Dateien oder/und \*.mht-Dateien indizieren möchten klicken Sie auf den Schalter „Andere Endung auswählen“. Es erscheint eine Liste mit den Endungen. Dort wählen Sie die entsprechenden Endungen aus. Sie werden nacheinander in die Endungen-Eingabezeile eingetragen. Mit dem Schalter „X“ rechts neben der Endungen-Eingabezeile können Sie den Inhalt dieser Zeile löschen.

Nun betätigen Sie den OK Schalter. Nach dem Sie OK betätigt haben, wird eine Hauptdateienliste erstellt. Dafür wird das Programm „ErstelleDateienliste.exe“ mit Administratorrechten gestartet. In der Taskleiste von Windows blinkt nun das Symbol „ErstelleDateienliste.exe“. Klicken Sie es an, erscheint ein Fenster in dem Sie vom Betriebssystem gefragt werden „Möchten Sie zulassen....“. Sie sollten die Frage der Benutzerkontensteuerung mit "Ja" bestätigen sonst kann die Datenbank nicht indiziert (erstellt) werden.

Warten Sie mit der Bestätigung nicht zu lange, denn Sie haben nur 60 Sekunden Zeit dies zu tun, sonst wird die aktuelle Indizierung abgebrochen. Sie müssen dann den Vorgang „Neue DB erstellen“ noch mal von vorne beginnen. Falls Sie innerhalb der 60 Sekunden „Nein“ wählen, müssen Sie warten bis von Lsearch-HTML die Meldung kommt, dass die Vorgabezeit überschritten ist, bevor Sie weitermachen können.

Wenn Sie mit JA bestätigt haben, erstellt das Programm die Hauptdateienliste im Ordner ...\\Datenbank\\. Dies geht relativ schnell. Für 300.000 Dateien etwa 25 Sekunden (ist abhängig von der Ausstattung Ihres PC).

Danach geht es los mit dem Indizieren. Auch dies geht recht flott voran.

Zum Beispiel für 2.907 HTML-Dateien auf Laufwerk C mit einer Gesamtgröße von 24 MB in 31 Sekunden (ohne Erstellung der Hauptdateiliste).

Die Datenbank wird in dem Datenbank-Ordner, indem sich LSEARCH-HTML befindet angelegt.

Die Datenbank trägt den Namen den Sie angegeben haben und die Endung „.dat“. Auch eine Log-Datei wird erstellt. Darin wird angegeben, wenn eine Datei nicht indiziert werden konnte und wie lange die Indizierung gedauert hat. Auch die Anzahl und Gesamtgröße der Dateien steht darin.

Wenn Sie eine Datenbank zum Beispiel „Alle Html-Dateien auf Laufwerk C“ erstellen möchten und einen oder mehrere Ordner vom Indizieren ausnehmen wollen, dann tragen Sie diese Ordner in die Datei „AusnahmeIndizierung.txt“ mit einem Editor ein. In dieser stehen schon ein Ausnahmeordner:  
C:\\Windows.old

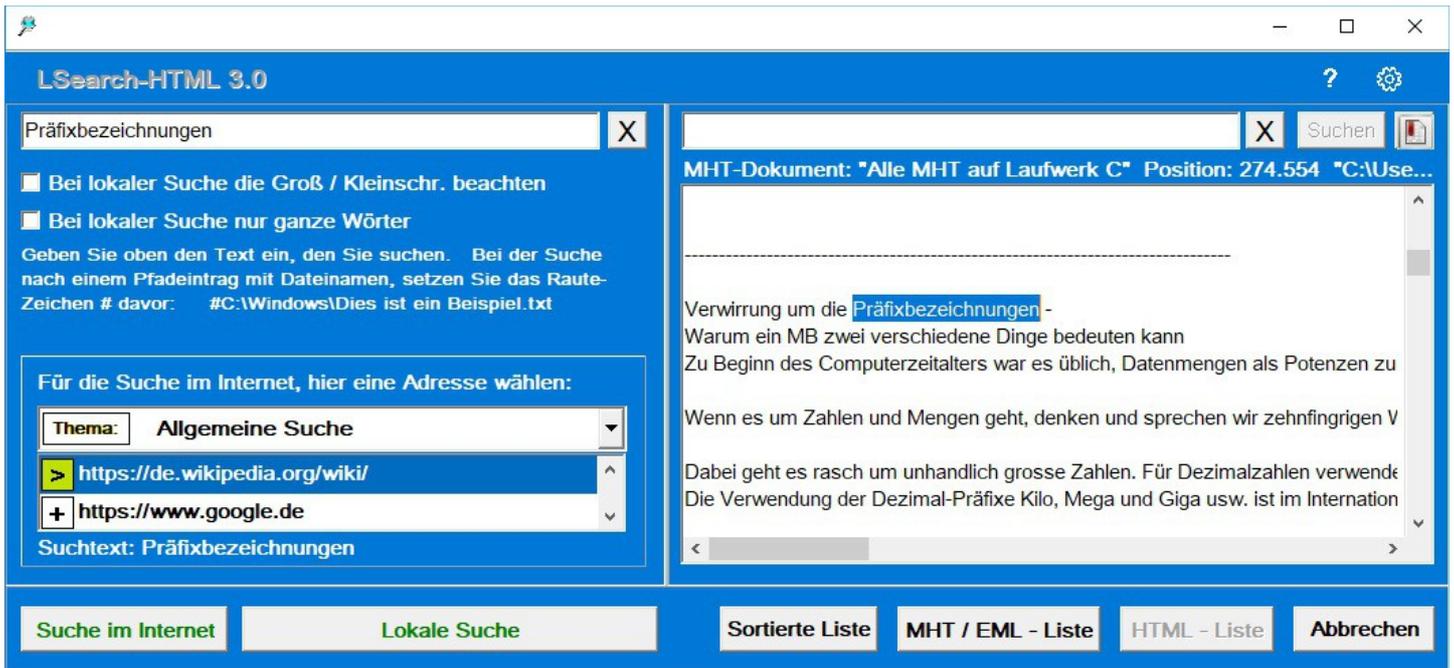
Diese Einträge können Sie ändern oder auch löschen und eigene einfügen.

Wenn MHT oder EML-Dateien indiziert werden wird das externe Programm „MHTtextOut.exe“ gestartet. Dieses benutzt den „Internet-Explorer“ zum Indizieren. Viele MHT-Dateien laden online Inhalte nach. Da dies lange dauern kann bis der jeweilige Server diese sendet, ist das online Nachladen abgeschaltet. Das Indizieren würde sonst zu lange brauchen.

Damit EML-Dateien (Mails) in „MHTtextOut.exe“ bzw. „Internet-Explorer“ angezeigt werden können, wird die Datei temporär in den Ordner ...\\Datenbank\\Temp\\ kopiert und deren Endung .eml in .mht umbenannt.

## Die Suche:

Betätigen Sie den Schalter „Suchen“, dann erscheint diese Oberfläche:



Mit den Schalter „MHT / EML - Liste“ und „Html - Liste“ laden Sie die jeweilig Liste in den Viewer.

Der Schalter „Sortierte Liste“ lädt, wenn vorhanden die sortierte Liste des vorher geladenen Dokuments. Bei den MHT / EML und Html -Listen handelt es sich um die originalen Textdokumente die nacheinander in ein Gesamtdokument eingefügt wurden.

Doppelklicken Sie auf einen Eintrag wird die ganze Zeile markiert. Oder markieren Sie einen Textbereich mit der linken Maustaste. Betätigen Sie dann die rechte Maustaste. Es erscheint dann ein kleines Menü an der Mausposition, in dem Sie durch Klicken auf den Eintrag „**Text in die Suchleiste links übertragen**“ den Text übertragen lassen können.

Betätigen Sie den Schalter „**Lokale Suche**“, dann wird mit der Suche begonnen und das Ergebnis in Ihrem Standard-Browser oder dem von Ihnen eingestellten Browser angezeigt.

Wird ein Suchbegriff gefunden, dann wird er in die Suchbegriffsliste eingefügt. Wenn Sie oben links in das Sucheingabefeld klicken und die Suchbegriffsliste der aktuellen Datenbank Suchbegriffe enthält wird diese unterhalb des Eingabefeld angezeigt. In dieser Liste können Sie für eine erneute Suche einen Begriff anklicken und dieser wird dann in das Eingabefeld eingefügt.

Die Suchbegriffslisten sind im Ordner „Datenbank\Suchwoerterlisten\gespeichert“.

Sie können auch im Internet nach einem Begriff suchen lassen. Zum Beispiel wenn Sie bei einem Fremdwort die Bedeutung nicht kennen oder zu einem Thema Informationen benötigen. Dazu klicken Sie auf „**Suche im Internet**“. Dies geht nur wenn im linken Suchtexteingabefeld Text eingegeben bzw. eingefügt wurde.

In dem rechten Suchtexteingabefeld können Sie Text eingeben, den Sie in dem Textviewer suchen möchten. Dazu klicken Sie dann **rechts** auf den Schalter „**Suchen**“.

Neben dem Schalter „**Suchen**“ ist der Schalter zum speichern eines Lesezeichens:

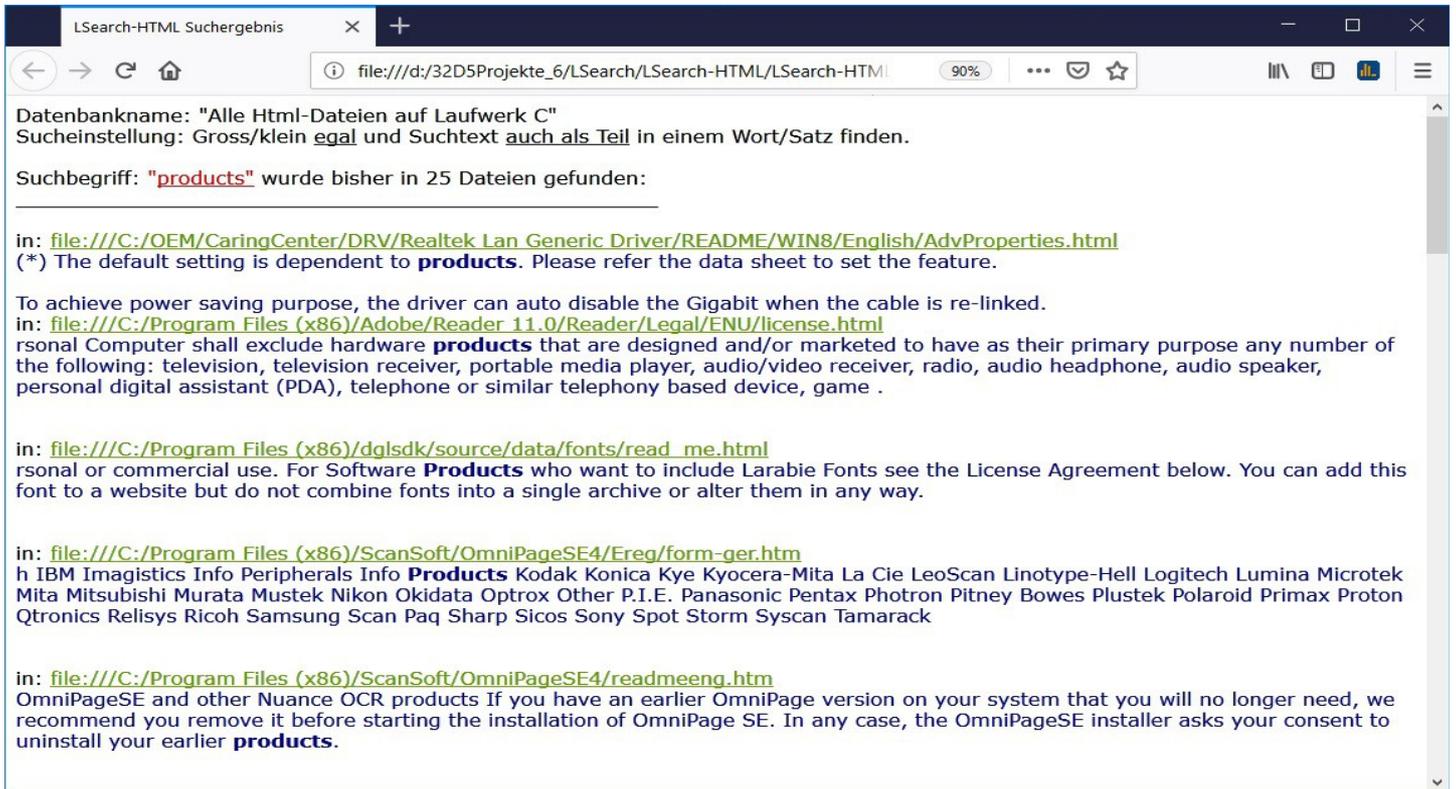


Wenn Sie auf diesen Schalter klicken, speichern Sie die aktuelle Leseposition. Wenn Sie später die Textliste erneut im Textviewer öffnen, zeigt sich der Text automatisch an dieser Stelle.

Ist eine „Sortierte Liste“ geladen ist das Lesezeichen Symbol deaktiviert:



## Die Suchergebnisse werden im Browser angezeigt:



LSearch-HTML Suchergebnis

file:///d:/32D5Projekte\_6/LSearch/LSearch-HTML/LSearch-HTML/

Datenbankname: "Alle Html-Dateien auf Laufwerk C"  
Sucheinstellung: Gross/klein egal und Suchtext auch als Teil in einem Wort/Satz finden.

Suchbegriff: "**products**" wurde bisher in 25 Dateien gefunden:

in: <file:///C:/OEM/CaringCenter/DRV/Realtek Lan Generic Driver/README/WIN8/English/AdvProperties.html>  
(\* ) The default setting is dependent to **products**. Please refer the data sheet to set the feature.

To achieve power saving purpose, the driver can auto disable the Gigabit when the cable is re-linked.

in: [file:///C:/Program Files \(x86\)/Adobe/Reader 11.0/Reader/Legal/ENU/license.html](file:///C:/Program Files (x86)/Adobe/Reader 11.0/Reader/Legal/ENU/license.html)  
Personal Computer shall exclude hardware **products** that are designed and/or marketed to have as their primary purpose any number of the following: television, television receiver, portable media player, audio/video receiver, radio, audio headphone, audio speaker, personal digital assistant (PDA), telephone or similar telephony based device, game .

in: [file:///C:/Program Files \(x86\)/dgl/sdk/source/data/fonts/read\\_me.html](file:///C:/Program Files (x86)/dgl/sdk/source/data/fonts/read_me.html)  
Personal or commercial use. For Software **Products** who want to include Larabie Fonts see the License Agreement below. You can add this font to a website but do not combine fonts into a single archive or alter them in any way.

in: [file:///C:/Program Files \(x86\)/ScanSoft/OmniPageSE4/Ereg/form-ger.htm](file:///C:/Program Files (x86)/ScanSoft/OmniPageSE4/Ereg/form-ger.htm)  
IBM Imagistics Info Peripherals Info **Products** Kodak Konica Kye Kyocera-Mita La Cie LeoScan Linotype-Hell Logitech Lumina Microtek Mita Mitsubishi Murata Mustek Nikon Okidata Optrox Other P.I.E. Panasonic Pentax Photron Pitney Bowes Plustek Polaroid Primax Proton Qtronics Relisys Ricoh Samsung Scan Paq Sharp Sicos Sony Spot Storm Syscan Tamarack

in: [file:///C:/Program Files \(x86\)/ScanSoft/OmniPageSE4/readmeeng.htm](file:///C:/Program Files (x86)/ScanSoft/OmniPageSE4/readmeeng.htm)  
OmniPageSE and other Nuance OCR products If you have an earlier OmniPage version on your system that you will no longer need, we recommend you remove it before starting the installation of OmniPage SE. In any case, the OmniPageSE installer asks your consent to uninstall your earlier **products**.

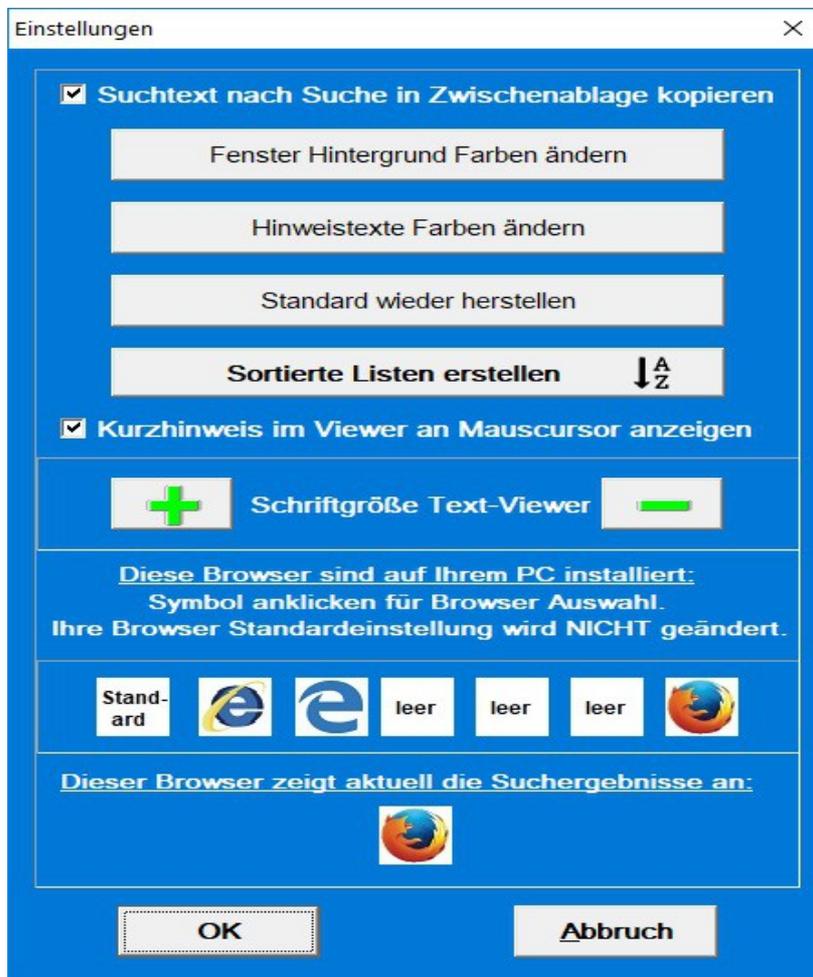
## Einstellungen ändern:

Klicken Sie oben rechts auf das Symbol:



Es erscheint dann die Oberfläche für das Ändern der Standard-Einstellungen.

Dort können Sie diverse Einstellungen vornehmen. Es ist alles selbsterklärend.



Einstellungen

Suchtext nach Suche in Zwischenablage kopieren

Fenster Hintergrund Farben ändern

Hinweistexte Farben ändern

Standard wieder herstellen

Sortierte Listen erstellen ↓ A Z

Kurzhinweis im Viewer an Mauscursor anzeigen

+ Schriftgröße Text-Viewer -

Diese Browser sind auf Ihrem PC installiert:  
Symbol anklicken für Browser Auswahl.  
Ihre Browser Standardeinstellung wird NICHT geändert.

Stand-ard e e leer leer leer

Dieser Browser zeigt aktuell die Suchergebnisse an:

OK Abbruch

Haben Sie mehrere Browser auf Ihrem PC installiert können Sie für die Anzeige der Suchergebnisse Ihren Lieblings-Browser einstellen.

Mit dem Schalter „Sortierte Liste erstellen“ wird eine sortierte Liste von der zuvor beim Indizieren erstellte Dokumentenliste „MHT /EML oder Html -Listen“ angefertigt.

### **Erstellung sortierter Listen:**

Um Listen mit Sätzen (Phrasen) und Wörter zu erstellen, klicken Sie in den Einstellungen auf den Schalter „Sortierte Listen erstellen“.

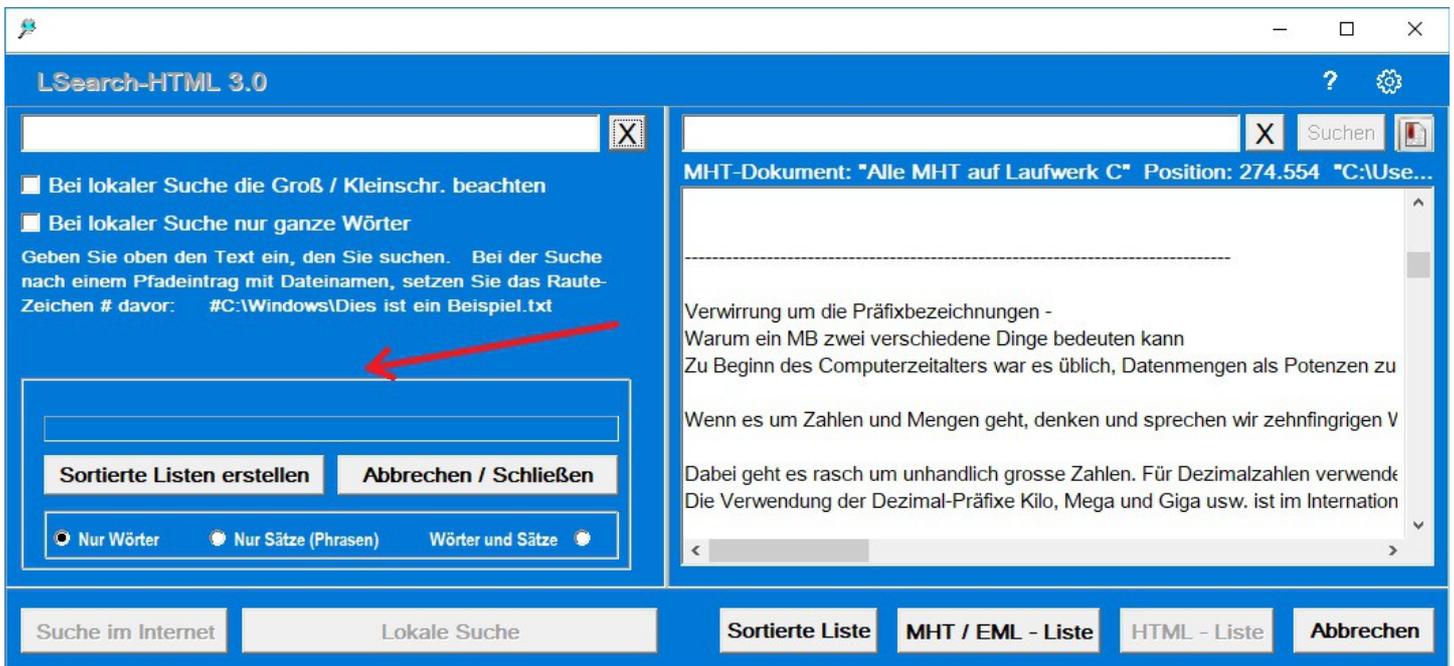
Es wird eine sortierte Liste von der zuvor beim Indizieren erstellte „ MHT /EML oder Html -Listen“ angefertigt.

Es gibt dabei keine doppelten Einträge in den sortierten Liste.

Am Ende der Listen stehen die Pfade der Originaldateien. Davor steht das Rautezeichen #. Dieses Zeichen dient als Kennzeichnung. Wenn Sie einen dieser Eintragungen in die Suchleiste links übertragen, werden Sie gefragt ob Sie die Originaldatei mit dem zugehörigen Programm öffnen möchten, denn entgegen anderem Text können Sie diese Pfadeintragungen nicht suchen lassen.

Für die spätere Suche nach Text in den Dateien sind diese Listen ganz hilfreich wenn Sie den Suchbegriff nicht genau kennen. Scrollen Sie die Liste bis der gesuchte Begriff erscheint. Markieren Sie diesen mit der linken Maustaste. Betätigen Sie dann die rechte Maustaste. Es erscheint ein kleines Menü an der Mausposition, in dem Sie durch Klicken auf den Eintrag **„Text in die Suchleiste links übertragen“** den Text in das linke Sucheingabefeld übertragen lassen können.

Haben Sie in den Einstellungen auf den Schalter „Sortierte Listen erstellen“ geklickt und die Bestätigung mit JA beantwortet, erscheint die Suchen-Oberfläche etwas verändert. Wählen Sie aus ob in der Liste „Nur Wörter“, „Nur Sätze“ oder „Wörter und Sätze“ gespeichert werden sollen. Danach klicken Sie auf den Schalter „Sortierte Listen erstellen“.



### **Löschen einer Datenbank:**

Klicken Sie mit der rechten Maustaste auf den Eintrag des Datenbanknamen. Es erscheint dann der Menüeintrag „Datenbank löschen“. Klicken Sie diesen an, dann werden Sie gefragt ob Sie die Datenbank wirklich löschen möchten.

## **Sie möchten Ihre Firmen- Preislisten oder Informationen zusammen mit den erstellten Suchdatenbanken auf einem Speicherstick an Ihre Kunden weitergeben?**

**Wenn Ja, dann habe ich eine Lösung für Sie! LSEARCH-HTML-LIGHT. Es ist Freeware und enthält nur die Suchfunktion. LSEARCH-HTML-LIGHT kann kostenlos mit Ihren Dokumenten verteilt werden.**

### **Vorgehensweise zum Erstellen eines Speicher-Sticks oder Karte mit der Vollversion LSEARCH-HTML:**

Kopieren Sie LSEARCH-HTML-LIGHT mit der Hilfedatei, der „Adressen\_Webseiten.txt“ (wird für die Suche im Internet gebraucht) und der „Lizenz.txt“ auf den Stick und starten Sie es. LSEARCH-HTML-LIGHT erstellt nun auf dem Stick den Ordner "\\Datenbank". Da noch keine Daten darin sind, wird auch keine in LSEARCH-HTML-LIGHT angezeigt. Beenden Sie nun das Programm.

Erstellen Sie ein Verzeichnis mit Ihren HTML-Dokumenten und speichern dieses dann auf den Stick. Angenommen der Laufwerksbuchstabe des Sticks in Ihrem PC lautet H. Auf dem Stick ist nun das Verzeichnis "H:\Unsere Produkte" und darin Ihre HTML-Dateien.

Sie können aber auch das Verzeichnis auf Ihrer Festplatte indizieren und dann dieses auf den Stick kopieren.

Nun starten Sie von Festplatte die Vollversion LSEARCH-HTML und betätigen den Schalter "Neue DB erstellen". Geben Sie dann einen Namen für die Datenbank ein (zum Beispiel: "Unsere Produkte") und wählen Sie das Verzeichnis "H:\Unsere Produkte" aus.

Betätigen Sie dann den Schalter OK.

Nach dem Sie OK betätigt haben, wird eine Hauptdateienliste erstellt. Dafür wird das Programm „ErstelleDateienliste.exe“ mit Administratorrechten gestartet. In der Taskleiste von Windows blinkt nun das Symbol „ErstelleDateienliste.exe“. Klicken Sie es an, erscheint ein Fenster in dem Sie vom Betriebssystem gefragt werden „Möchten Sie zulassen.....“.

Wenn Sie mit JA bestätigt haben, erstellt das Programm die Hauptdateienliste im Ordner ...\\Datenbank\\ auf der Festplatte. Dies geht relativ schnell. Für 300.000 Dateien etwa 25 Sekunden (ist abhängig von der Ausstattung Ihres PC).

Danach geht es los mit dem Indizieren. Auch dies geht recht flott voran.

An dem angegebenen Ort werden die HTML-Dateien nun gesucht und indiziert. Die Datenbank wird auf der Festplatte erstellt. Auf der Festplatte am gleichen Ort wo sich LSEARCH-HTML befindet ist das Verzeichnis "\\Datenbank". In diesem Ordner befinden sich dann die Dateien:

„NameDatenbank.dat“, „NameDatenbank\_UTF8Dokument.txt“ und wenn MHT-Dateien indiziert wurden „NameDatenbank\_MHTtext.txt“.

Es sind auch die Dateien:

„NameDatenbank\_Utf8\_TitelPositionen.txt“ und eventuell „NameDatenbank\_MHT\_TitelPositionen.txt“ vorhanden.

Lassen Sie nun noch die „Sortierte Listen“ von der Datenbank erstellen. Danach sind die Dateien:

„NameDatenbank\_Phrasen\_Woerter\_Utf8Liste.txt“ und eventuell „NameDatenbank\_Phrasen\_Woerter\_MHTListe.txt“ vorhanden.

Diese Dateien kopieren Sie nun auf den Stick in den Ordner "\\Datenbank" der von LSEARCH-HTML-LIGHT (siehe oben) erstellt wurde. Einfach alle Dateien die mit dem Namen, den Sie angegebenen haben in den Ordner "\\Datenbank" kopieren.

Die Hauptdateienliste brauchen Sie nicht auf den Stick zu kopieren.

Wenn Sie nun LSEARCH-HTML-LIGHT auf dem Stick starten wird die Datenbank "Unsere Produkte" angezeigt. Sie können nun mit dem Schalter "Suchen" Text suchen lassen.

### **Hinweis:**

Da sich der Laufwerksbuchstabe des Sticks an einem anderen PC wahrscheinlich ändert, passt LSEARCH-HTML-LIGHT ihn in diesem Fall an. Falls Sie das Verzeichnis auf der Festplatte indiziert haben müssen Sie zum Testen des Sticks dieses Verzeichnis umbenennen, sonst wird bei der Suche auf dem Stick diese Dokumente angezeigt.

Möchten Sie in der Vollversion LSEARCH-HTML oder in der Freeware LSEARCH-HTML-LIGHT Ihr Firmenlogo bzw. Werbebanner eingebracht haben?

Dann erfragen Sie bei [baecker@wb-soft.de](mailto:baecker@wb-soft.de) den Preis dafür.

Ich weise daraufhin, dass die in dieser Hilfedatei verwendeten Soft- und Hardwarebezeichnungen und Markennamen der jeweiligen Firmen im allgemeinen warenzeichen-, marken- oder patentrechtlichem Schutz unterliegen.

(c) 2022 by Wolfram Bäcker, 66904 Börsborn.

Internet: [www.wb-soft.de](http://www.wb-soft.de)

---