



Digital Tinker Data Cleansing Packet

Description of Data Cleansing Services

Supported Data Formats

The following file formats are supported for raw data **and** output:

- **.csv** Comma Separated Values
- **.doc** Microsoft ® Word document
- **.mdb** Microsoft ® Access database
- **.txt** Plain text file
- **.xls** Microsoft ® Excel spreadsheet

The following file formats are supported for raw data only:

- **.html** Web pages or portions of web pages

The following file formats are supported for output only:

- **.pdf** Adobe® Portable Document Format

Wherever you see file types in the following descriptions they are limited to the supported types.

Conversion

When a file is in one format, such as a spreadsheet (.xls), you may need to convert it to a database table (.mdb).

While the process may be simple, you may want to automate it. I'll write a tool to convert any supported raw data format to any supported output

Conversions are done on a file-by-file basis, and the total amount of raw data may not exceed 10 megabytes.

No substitutions or extractions are performed on the raw data.

If you need those additional services, you must specify that in your submission.

Substitution

Substitution is another way of saying search and replace. You may desire a way to automate tedious and repetitive data management. Another challenge might be a complex substitution that requires rules. Doing it without automation means spending too much time and risking errors during the process.

Simple Search and Replace Automation

I will create a simple automation tool to quickly read your text file, spreadsheet, csv or database and replace *all occurrences of one string of characters with one other string of characters.*

Simple replacements are done on a file-by-file basis, and the total amount of raw data may not exceed 10 megabytes.

Conditional Replacement

When simple substitution isn't enough, I'll create a more flexible tool to read your text file, spreadsheet, csv or database and replace *all occurrences of one or more strings of characters with other strings of characters, based on rules that you specify.* Rules can be as complex as you need them to be and can include conditions like skipping data that is incomplete or marked in some way.

If you are not sure of how to describe the rules, we can work together, with you showing me sample output and then testing the results until they give you the results you need.

Conditional replacements are done on a file-by-file basis, and the total amount of raw data may not exceed 10 megabytes.

Extraction

Extraction involves pulling a subset of data from the raw data file. If you only need a few columns from a spreadsheet, for example, you may want them extracted and placed into another file.

I will perform data extraction from **one** supported raw data file and create a new output file in **the same format.**

Extractions are done on a file-by-file basis, and the total amount of raw data may not exceed 10 megabytes.



Digital Tinker Data Cleansing Packet

Description of Software Services

Formula Analysis

You may request help with your Microsoft ® Excel Spreadsheet formulas and Microsoft ® Access database queries.

I will ask you to provide the spreadsheet or database, along with a description of *expected* output.

I must have at least a duplicate with sample data. If that is not possible, I'll do email consultation if you'll agree to pay for my time.

(The consultation rate is listed below.)

Automation Software

Throughout the descriptions, I've mentioned creating automated software.

If you simply need a one-time data cleansing, then you don't need the software!

However, if you need to have the same type of data cleansed periodically, you may want do one of the following:

- Resubmit the data and wait for me to re-run the software
- Purchase one or more licenses for your company's use in-house
- Purchase volume licenses to resell to clients so that *they* can clean the data that they send you

Depending on the license, you have either a limited-time exclusive use or non-exclusive use of this software.

As a practical matter, not too many people will need to do exactly what you do with your data, so I don't expect that I'll be making too many general-purpose tools.

Finally, creating a user-friendly application is a lot more time-consuming than whipping up an in-house utility.

Therefore, allow plenty of time for development and testing.

Exclusions

Your submission can not be accepted if it:

- involves anything illegal or immoral
- is incomplete or corrupted
- is over 10 megabytes in size (counting all files)
- is in an unsupported file format
- requires output to unsupported format
- involves translation to another language, except where such translation is provided by you in a *supported and usable* format
- violates rules of logic (e.g. Circular reasoning causes an infinite loop)
- is beyond the scope or capabilities of the Digital Tinker Services

Limitations of Liability

By submitting your data, you acknowledge that it is a **copy** and that I can not be held liable for damage to the file which occurs during electronic transmission or while in my possession. If any data is damaged, you agree to retransmit another **copy** or to decline my services without having to pay for incomplete work.



Digital Tinker Data Cleansing Packet

Non-Binding Price List Expires January 31, 2009

Client Class: Small Business (25 employees or less)

For more information, or to request a current price list, email Mitch@morphodesigns.com

Item	Description	Price
Data Conversion		
Convert a file to PDF format	I create a PDF, making minor formatting adjustments	\$ 25.00
Convert raw data into CSV	I create a clean output file in txt or csv format only	\$ 40.00
Completely remove HTML tags from a file	I create a clean output file in txt, doc, xls or mdb	\$ 40.00
Remove non-formatting HTML tags from a file	I create a clean output file in txt, doc, html, xls or mdb	\$ 80.00
:	:	
:	:	
Substitution		
Simple Search and Replace Automation	I create a clean output file in txt, doc, html, xls or mdb	\$ 40.00
Complex Search and Replace with conditional rule	I create a clean output file in txt, doc, html, xls or mdb	\$ 80.00
:	:	
:	:	
Data Extraction		
Extract information from Text or Microsoft ® Word	I create a clean output file in txt, doc, html, xls or mdb	\$ 80.00
Extract information from Microsoft ® Excel	I create a clean output file in txt, doc, html, xls or mdb	\$ 80.00
Extract information from Microsoft ® Access	I create a clean output file in txt, doc, html, xls or mdb	\$ 80.00
Extract information from HTML source file	I create a clean output file in txt, doc, html, xls or mdb	\$ 100.00
Formula Analysis		
Fix formulas in Microsoft ® Excel Spreadsheet(s)	I examine and repair Formulas, VBA code	\$ 40.00
Fix formulas in Microsoft ® Access Database(s)	I examine and repair queries, table joins, VBA code	\$ 40.00
E-mail consultation	Only if you can't provide the spreadsheet or database	\$ 80.00
:	:	
Software to Automate Data Cleanup		
Non-exclusive Single-License Software	You use on one PC, I retain resale rights	\$ 60.00
Non-exclusive 5-License Software	You use or resell 5, I retain resale rights	\$ 250.00
Non-exclusive 25-License Software	You use or resell 25, I retain resale rights	\$ 750.00
Non-exclusive Unlimited License Software	You resell unlimited, I retain resale rights	\$ 1,000.00
1 Year exclusive Single-License Software	You use on one PC, I defer resale rights	\$ 75.00
1 Year exclusive 5-License Software	You use or resell 5, I defer resale rights	\$ 300.00
1 Year exclusive 25-License Software	You use or resell 25, I defer resale rights	\$ 1,000.00
1 Year exclusive Unlimited License Software	You resell unlimited, I defer resale rights	\$ 1,500.00
:	:	
:	:	
:	:	
:	:	
:	:	
:	:	

Client Class: Small Business (25 employees or less)



Digital Tinker Data Cleansing Packet

Submission Form

Paste this form into an email and give as much detail as possible. Read the explanations and delete each one as you answer. Skip questions that don't apply but, **if you don't understand a question, please answer, "I don't understand"**. Don't worry if the lines split up.

BRIEFLY DESCRIBE YOUR DATA CLEANSING NEEDS

ex.: I don't know how to get all the names and phone numbers from this spreadsheet to a neat Word document.

SPECIFY WHICH SERVICES YOU REQUIRE

ex.: Conversion, Substitution, Extraction, Formula Analysis, Software Creation

DOES YOUR DATA FREQUENTLY NEED TO BE FIXED? IF SO, WHY?

ex.: Yes, my client sends a spreadsheet every month

PLEASE LIST ALL DIFFERENT FILE TYPES THAT MAKE UP YOUR RAW DATA TEXT, DOC, HTML, XLS, MDB

ex.: XLS (refer to the supported formats in the descriptions)

ARE YOU ABLE TO SUBMIT COPIES OF THE ACTUAL DATA?

ex.: No.

IF CONFIDENTIALITY IS AN ISSUE, ARE YOU WILLING TO SUBMIT FICTIONAL DATA IN THE SAME FORMATS?

ex.: Sure!

IF NEITHER ACTUAL NOR FICTIONAL DATA CAN BE SUBMITTED, ARE YOU WILLING TO SUBMIT DETAILED INSTRUCTIONS?

ex.: I think so

PLEASE GIVE A FEW EXAMPLES OF THE TYPE OF INSTRUCTIONS YOU WILL GIVE IF YOUR SUBMISSION IS ACCEPTED.

ex.: Spreadsheet has last name in first column, first name in second column and phone number in 8th column. I need first name, last name and phone number from each row, unless there is no phone number.
The output in Word must look like this John Doe, (555) 123-4567.
The phone number is sometimes in the spreadsheet as 555-123-4567.

WHAT IS THE ABSOLUTE DEADLINE FOR YOU TO OBTAIN FIXED DATA?

ex.: By the end of this month!

DO YOU HAVE A FIDUCIARY RESPONSIBILITY TO PROVIDE OR MAINTAIN THIS DATA?

ex: Yes or No

DO YOU REQUIRE AN AUDIT OF CHANGED DATA?

ex.: Yes or No

DO YOU WISH TO HAVE THE SOFTWARE TO PERFORM THE SERVICE OR RESELL?

ex.: Yes or No

THANK YOU FOR SUBMITTING THIS FORM!

Email to mitch@morphodesigns.com