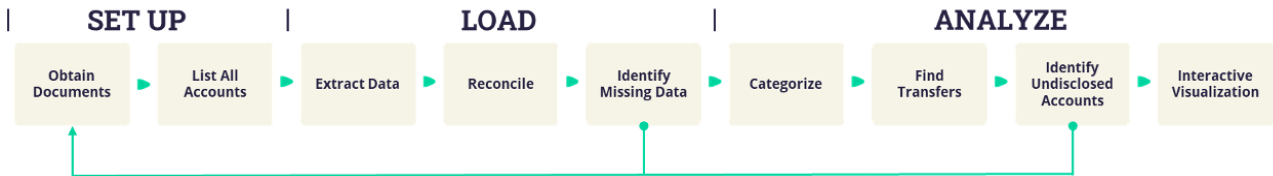# Dispute and Investigation Best Practice Guide

Tod McDonald, CPA, CIRA

Founder, Valid8 Financial

SUMMARY

Speed and accuracy are the most valuable assets that an accounting professional can bring to any financial dispute or investigation. Based on over 20 years of CPA experience and dozens of high stakes, high profile investigations and disputes, this guide attempts to establish a framework to help accounting professionals avoid pitfalls, follow best practices, and deliver the fastest, most accurate analysis for financial disputes and investigations.

The guide is structured to follow the procedures shown in the graphic below.

SET UP		LOAD		ANALYZE	
Obtain Documents	List All Accounts	Extract Data	Reconcile	Identify Missing Data	Categorize

Find Transfers	Identify Undisclosed Accounts	Interactive Visualization

VALID|8

CONTENTS

VALID|8

GETTING STARTED

The first step for any engagement is obtaining the documents. This is often done via subpoena. Documents can arrive in a variety of formats with some being easier to work with than others. This section walks through common pitfalls and best practice techniques to deal with the variety of document types and formats.

OBTAIN STATEMENT AND DOCUMENT FILES

The format type of documents received determines the types of software tools required for the engagement. This first section addresses dealing with the various formats. The next section "Load Data" will cover how to deal with the different types of documents like brokerage statements, credit card statements, check images, and deposit slip images.

Hard Copies

In contentious matters, professionals can be forced to "dumpster dive" for documents. Worst case scenario, professionals search through garages, attics, boxes, and filing cabinets to find old, incomplete, marked-up documents. Regardless of quality, the documents need to be digitized. Any modern printer/copier/scanner can be used to digitize the hard copies into PDF files.

• **Pitfall:** Running stacks of hard copy documents through the scanner without reviewing and organizing first. The result will be a single file with thousands to tens of thousands of pages which creates time-intensive complexity during data extraction, reconciliation, and QA procedures.

• **Best Practice:** At a minimum, organize the documents by account. If time permits, further organize by individual statement period. Taking the time to create one file for each account or one file per statement period will dramatically simplify the ability to do quality assurance on the data extracted from the files.

Scanned Images

PDF documents may already be available, or the financial institution may provide the files to you. Often this will be a single file with multiple accounts spanning many months or years.

• **Pitfall**: Attempting data extraction from a scanned image file without first performing a thorough review to understand how many different accounts, financial institutions, account types, and statement periods exist.

• **Best Practice:** Use a PDF editor/image editor to split the document into separate files that contain information for each single account. Be sure to pay attention to pages that contain data from two different accounts to be sure data extraction is not duplicated.

Native PDF Files

Native PDF files are the easiest to work with. Many tools exist to extract data and error rates are extremely low.   Also, native files will likely be organized by individual account period which helps avoid the pitfalls associated with hard copies and scanned images.

Other Formats

On occasion, financial institutions may provide transaction lists via XLSX, CSV, or TXT/JSON among others. Excel handles dozens of file types and is very useful.

• **Pitfall**: Digitized data is often incomplete and requires integration with other types of data sources. In addition, because data has already been digitized, there may be issues with the chain of custody.

• **Best Practice:** Use a software database solution that allows CSV (or other text file format) uploads so you can integrate with other data sources.

VALID|8

DOCUMENT ACCOUNT DETAILS

Taking the time to organize and review all documents, statements, images, and transaction lists provides a good perspective on what type of accounts and other details are needed to help organize future work. The final step to starting an engagement is to list out all accounts that require analysis. Recommended data fields for each account are listed below:

- Legal Entity
- Financial Institution
- Currency
- Account Number
- Account Name
- Account Type
- Description

LOAD AND EXTRACT DATA

Do not start to load or extract data into a database until all documents are reviewed, organized into files, and accounts are listed.   The next step is to establish an accurate and complete foundation or starting point, typically with bank and brokerage statements. The objective for these procedures is to ensure data is 100% accurate before analysis begins.

Extract Bank Statements

Many tools exist to help professionals extract transaction data from PDF or other image type statement files using specialized OCR (optical character recognition) technology. All tools work well with native PDF files. Pitfalls can occur with non-native PDF file types, potentially resulting in more work than simply manually transcribing data. Generally, for small projects, these pitfalls are minor annoyances. As you get into engagements that require analysis of several to dozens of accounts, multiple legal entities, and a variety of document types and data sources, the work can become insurmountable without the right software tools.

   • **Pitfall:** Using OCR solutions that require users to define a template per statement format. Templates are more prone to error and require more set up work. Statement formats for institutions change over time and may be different from branch to branch, each requiring a different template.

   • **Best Practice:** Modern solutions that use machine learning to recognize text strings extracted by OCR technology avoid the need to develop templates for each statement format. Also, they simplify large single files that contain dozens of accounts and hundreds of periods by identifying the individual account periods and additional fields like balances and period dates.

   • **Pitfall:** Manual integration of data from different data sources and document types. Most OCR solutions are "pass through". The user loads a statement file and gets a CSV or EX. Then the user is responsible for clean-up and integration into some other master data set. Clean up and integration can require a significant amount of time and effort.

   • **Best Practice:** Master, editable data base, not pass through.

Extract Brokerage Statements

In many cases, analysis includes brokerage accounts which can make data extraction especially complicated. Most specialized OCR tools for statements do not support brokerage accounts as they focus on banking and credit card statements. The most advanced solutions will support brokerage statements by extracting only cash in and cash out transactions. Without this functionality, professionals are often left to manual transcription.

Reconcile Account Periods

Extracting transaction data from documents is not an exact science. No matter what solution you use, errors can occur so professionals must prove accuracy. The only way to do this is to compare the extracted transactions with account period dates and balances. As you can imagine, trying to do this in Excel for a project with hundreds, or even thousands of account periods is daunting.

• **Pitfall:** Extracting transaction data without account period dates and balances.

• **Best Practice:** Use software solutions that identify individual account periods in PDF/image files and extract period balances and dates in addition to transactions. The software solution can also check to make sure that the beginning balance plus inflows minus outflows equals ending balance for each account period.

Find and Fix Incorrect or Duplicate Data

No matter how data gets extracted from documents, errors are likely. Professionals can burn too much time chasing, updating, and verifying corrections.

> • **Pitfall:** Adding new data to your master database or file before it has been reconciled for accuracy and checked for duplicates. Disputes and investigations are an ever-expanding universe of data that is acquired at different times. Adding incorrect or duplicate data into a master copy creates complexity, errors, and extra cleanup work.

> • **Best Practice:** Fix all extraction errors and check for duplicate statement files and account periods before integrating new data into an existing master file or database.

Load or Integrate Data with Master Copy

After extracted data is confirmed as 100% accurate with no duplicate transactions, load the data into the master database or integrate with the master file.

Identify Missing Data (and repeat)

Examine all data and all accounts to determine the timeline of data that exists. Look for continuity between all account periods. The prior account period ending date should match with current account period beginning date. Check for prior ending and current beginning balances. Identify any inconsistencies or gaps indicating missing data.

> • **Pitfall:** Maintaining an inventory in Excel, updating inventory each time the data set changes, and reporting back to your team on data set status.

> • **Best Practice:** Use software to map out each account period in a visual timeline that is updated in real time as data is added. All users can visually identify gaps and check data set status in real time.

VALID|8

At this point, the transaction database is a clean foundation or starting point. Keep in mind, more work may be required before analysis can start, especially if wires, checks, or deposit slips were used to initiate banking transactions. In these cases, the transactions on the bank statements will not contain the right information for proper analysis.

Extract Check and Deposit Slip Images

If required, subpoena wire transaction details, and/or check and deposit slip images from each financial institution. Wire details are likely to be provided via a CSV or text file. If not already included on the statement, check and deposit slips will likely come via a multi-page TIFF or PDF file. There are two key capabilities required to integrate these important details with ease:

1. The ability to extract transaction information from handwritten check and deposit slip images
2. The ability to integrate details from the extracted check or deposit slip images and wire transaction details with the appropriate bank statement transactions

Often, this is the most difficult and the most crucial step as wire details and checks contain information about where the money went and where it came from.

• **Pitfall:** Recognizing late in an engagement that check, deposit slip, and wire details are required for analysis. These details can be the most difficult and time-consuming aspect of any engagement resulting in missed deadlines and additional costs.

• **Best Practice:** Use software solutions that can extract data not only from statements but from handwritten check and deposit slip images. These software solutions will match the check items and wire details back to the banking transactions so they can be appended with the additional information.

VALID|8

ANALYSIS

Data prep is the hard work. Analysis is where professional firms distinguish themselves. The key is getting the data prep done as fast as possible so that more time can be spent working on the analysis. Every case is different and has different requirements. This section lists a few of common types of analysis done for disputes and investigations.

Threshold Analysis

If the client is price sensitive or does not know if it's worth doing a deep dive investigation, a threshold analysis can be useful. Output is typically a table that breaks out transactions into cohorts based on amounts and counts. See the example below.

| Year | All | | >$100 | | >$150 | | >$250 | | >$500 | | >$1000 | |
	Sum	Txn Count	Sum	Txn Count	Sum	Txn Count	Sum	Txn Count	Sum	Txn Count	Sum	Txn Count
2016	$ 1,931,865	1,074	$ 1,917,947	759	$ 1,909,116	689	$ 1,890,267	594	$ 1,824,910	411	$ 1,712,809	267
Percent of All:			99%	71%	99%	64%	98%	55%	94%	38%	89%	25%
2017	$ 1,784,307	1,372	$ 1,751,778	656	$ 1,738,183	549	$ 1,717,413	444	$ 1,670,235	313	$ 1,592,186	201
Percent of All:			98%	48%	97%	40%	96%	32%	94%	23%	89%	15%
2018	$ 1,451,025	1,428	$ 1,415,552	687	$ 1,399,960	562	$ 1,380,131	463	$ 1,330,452	319	$ 1,253,899	205
Percent of All:			98%	48%	96%	39%	95%	32%	92%	22%	86%	14%
2019	$ 1,598,334	1,441	$ 1,563,843	725	$ 1,544,045	562	$ 1,522,135	453	$ 1,479,418	330	$ 1,393,479	212
Percent of All:			98%	50%	97%	39%	95%	31%	93%	23%	87%	15%

Clients can help the professional pare down the amount of work to a reasonable size by only looking at transactions of value beyond a floor.

Even $ Check Analysis

More useful than Benford's law is looking at all even dollar transactions (example $5,000, $7,000, $250, $600, etc.). Proper business transactions are rarely even dollar amounts. Sales tax and multiple line items create natural complexity. Like threshold analysis, this is typically done via a table showing counts of even dollar transactions.

Benford's Law

Benford's law is a curve that represents the natural distribution of digits as they appear in a data set. For example, because 1 always comes before 2, the number 1 will appear more often than 2 in a naturally occurring number set. If numbers have been tampered with, there will be a deviation from the Benford's law curve.

Categorize and Group Transactions

Categories and groups of transactions allow simple summary tables and easy interpretation. Several factors can make this difficult. Each engagement is different and requires different categories. The amount of transactions can be vast. Reading and categorizing each one can take a long time.

• **Pitfall:** Updating categories in a spreadsheet. In many cases, documents are processed or extracted at different times during an engagement. Spending time defining and updating transaction categories can create complexity with trying to integrate new data from new documents.

• **Pitfall:** Relying on a proprietary database user interface. Some databases allow the creation and editing of custom fields and categories. This approach typically allows easy integration of additional data but can be less efficient when updating categories for many transactions, particularly if there are hundreds of different categories to apply.

• **Best Practice:** A hybrid approach where a master database is integrated with spreadsheet software provides the best of both worlds. Users can leverage the efficiency of data input via a spreadsheet while maintaining a single, version-controlled master database. Look for solutions that have tight integration with spreadsheet products.

VALID|8

Identify Transfers

Using spreadsheets to find transfers between accounts and legal entities is complicated at best. Considering date ranges and same day, same dollar amount scenarios, it can be almost impossible to get a comprehensive view of all transfers. Software solutions that identify transfers and allow users to confirm or reject matched pairs is more effective and efficient than using a spreadsheet. A comprehensive solution should include the following capabilities to identify transfers accurately and efficiently between accounts and entities:

• Manual match
• Reject match
• Confirm match
• Auto-confirm match based on confidence level
• Reset all matches
• Do not match rejected pairs
• Match given date range

VALID|8

Below is an example of a transfer match engine user interface.

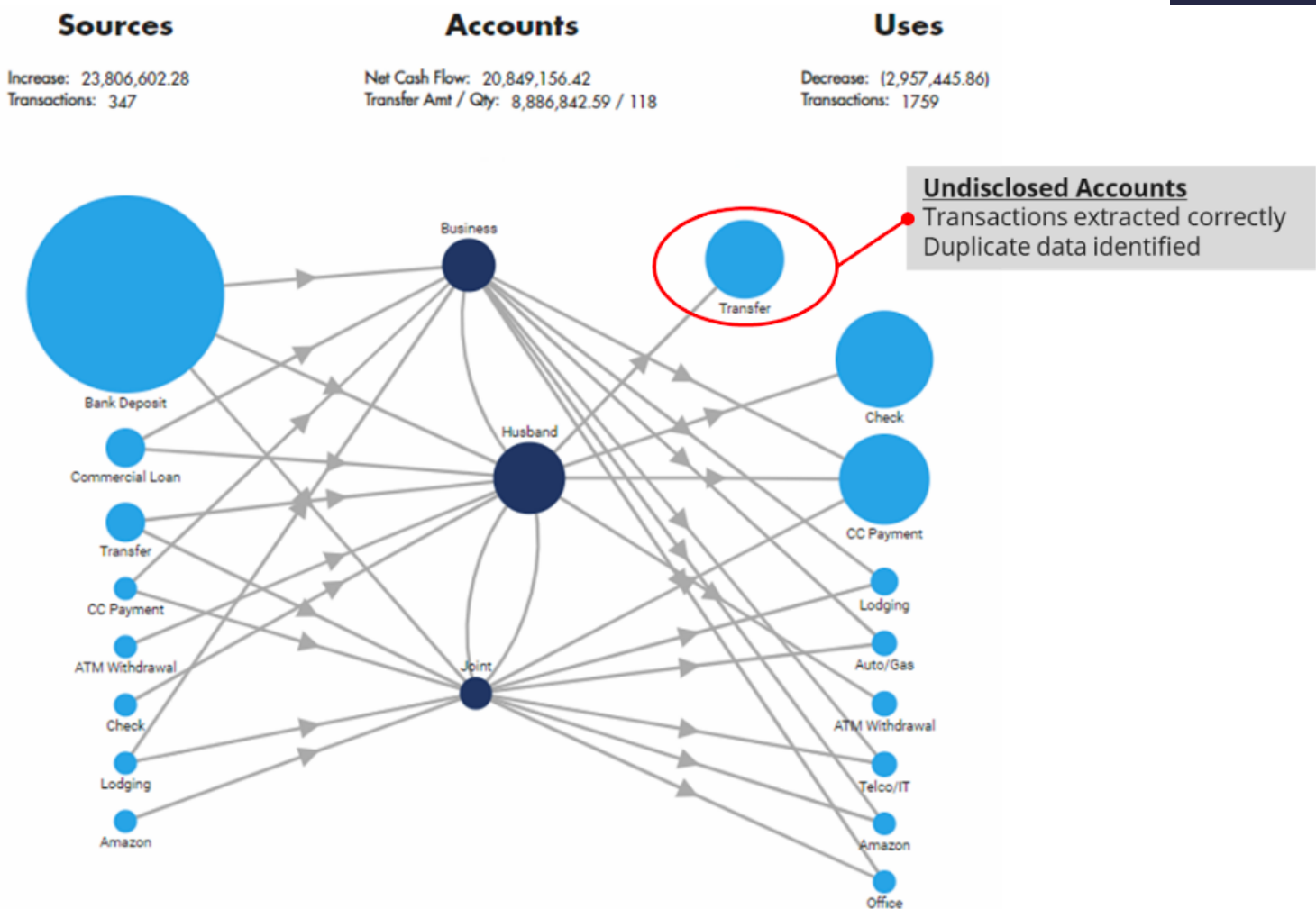| | Found 30 | Confirmed 120 | Rejected 1 | Approve All | Find New Matches | Reset All Matches | Update |

Approve Page | Reject Page

ID	Amount	Date	Description	Account	Entity	Match
1	$25.00	07/19/2012	CARD/DEPOSIT RELATIONSHIP	40039010013511...	Joint	✓ ⊘
	($25.00)	07/13/2012	32133	38971719	Husband	
2	($1,500.00)	10/11/2012	32236 813006292155277 3S	38971719	Husband	✓ ⊘
	$1,500.00	10/05/2012	MARRIOTT UNIVERSITY PK SL SALT LAKE CTYUT	40039010013511...	Joint	
3	($2.00)	03/13/2013	Adp Payroll Fees Des:Adp Fees ID:13X26 3833951 902571008...	38971719	Husband	✓ ⊘
	$2.00	03/11/2013	Manual transfer test - Dep before W/D	153595420008	Business	
4	($129.30)	03/29/2013	813009992739547 32814 813005442671084	38971719	Husband	✓ ⊘
	$129.30	04/01/2013	PAYMENTS AND OTHER CREDITS	40039010013511...	Joint	
5	$2,086.61	04/01/2013	WA BANKING CENTER PAYMENT	4003 9010 0229 0...	Subsidary	✓ ⊘
	($2,086.61)	03/29/2013	813005992014568 32817 813005442671083	38971719	Husband	
6	$2.00	04/12/2013	Manual transfer test - W/D before Dep	153595420008	Business	✓ ⊘
	($2.00)	04/10/2013	Adp Payroll Fees Des:Adp Fees ID:13Xz6 4991152 902599005...	38971719	Husband	

• **Pitfall:** Using manual techniques or spreadsheet functions like VLOOKUP in spreadsheets to identify transfers between accounts and entities.

• **Best Practice:** Leverage software solutions that can identify matches and allow users to provide manual dispositions on each. Having the ability to confirm or reject a match between transactions provides professionals with the flexibility to get everything exactly right.

VALID|8

Identify Undisclosed Accounts (and repeat)

Undisclosed accounts are easily identified if all transfers have been properly categorized and each transfer between accounts has been identified (meaning, you have associated the outflow from one account to a corresponding inflow from another account). Performing a simple search on all transfers not included in the set of matched inflow or outflow pairs indicates a transfer out of the data set and something that may need to be investigated.

• **Pitfall:** Integrating various spreadsheets and using VLOOKUPS can be complicated and cumbersome. Results can be difficult to understand especially if used as a court exhibit.

• **Best Practice:** Data visualizations simplify and accelerate the process. See the screen shot example in the Interactive Data Visualization section where transfers out of the data set are clearly identified.

Sources

Increase: 23,806,602.28
Transactions: 347

Accounts

Net Cash Flow: 20,849,156.42
Transfer Amt / Qty: 8,886,842.59 / 118

Uses

Decrease: (2,957,445.86)
Transactions: 1759

Undisclosed Accounts
Transactions extracted correctly
Duplicate data identified

REPORTING

Critical to the success in any engagement is the ability for professionals to create a compelling case narrative. The ability to report on the results and articulate what happened is the key factor for establishing a narrative. Reporting usually falls into one of two categories:

1) Preliminary assessment where the investigation is still underway
2) Final exhibits to support opinions/conclusions

Filtering, sorting, and organizing the data in real time is crucial for preliminary assessments. Interactive data visualizations are especially helpful in this scenario and can clarify what questions to ask next.

Final exhibits often take a lot more time and thoughtfulness to develop. Whether using a spreadsheet, a proprietary software platform, or data visualization techniques, simplicity is the key to effectiveness.

Tod McDonald, CPA, CIRA
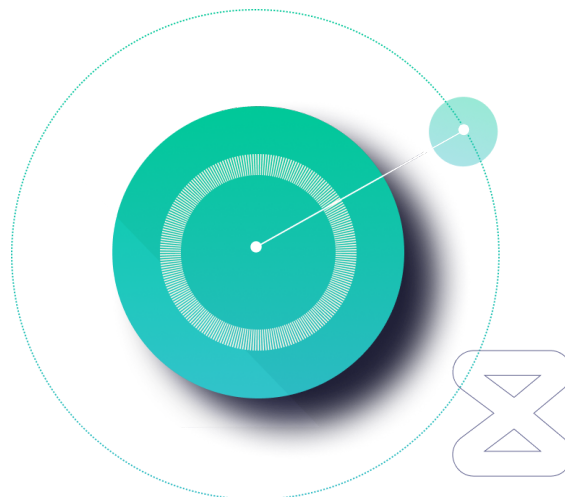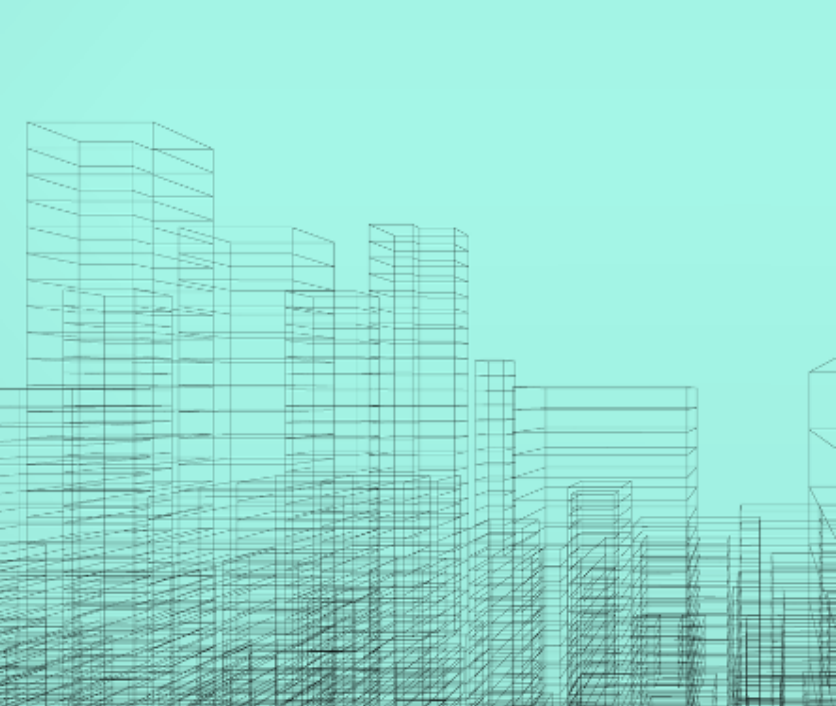
https://valid8financial.com/
t.mcdonald@valid8financial.com
(206) 920-1144

VALID|8