

# FIVE QUESTIONS TO ASK A REDACTION VENDOR

The risk is real. Personal information can get stolen when government and business do not utilize redaction software, or use an inferior product that does not irreversibly black out sensitive information. Quality redaction software can protect all parties concerned and ensure compliance with privacy legislation. Find a good vendor and get started.

COMMONWEALTH OF PENNSYLVANIA  
DEPARTMENT OF TRANSPORTATION  
Bureau of Driver Licensing  
Mail Date: JANUARY 1, 2015

WID # 9000900  
PROCESSING DATE  
DRIVER LICENSE  
DATE OF BIRTH


Dear Mr. [Redacted]  
This is a Special Notice of the Suspension of your  
Driving Privilege authorized by Section 1532B of the Penn  
Vehicle Code. As a result of your 01/01/2015 convi  
violating Section 3736 of the Vehicle Code RECKLESS  
on 01/01/2015:

Your driving privilege is SUSPENDED for a per  
01/01/2015 to 01/01/2015.

Be advised that if you wish to restore your driving  
privilege, follow the instructions in this letter for COMPLY  
THIS SUSPENSION; PAYING THE RESTORATION FEE and P  
PROOF OF INSURANCE. You should follow all instructi  
carefully. Even if you have served all the time  
under this suspension, you must surrender your driv  
ilege until all the requirements are satisfied.

**COMPLYING WITH THIS SUSPENSION**  
You must return all current Pennsylvania driver's l  
learner's permits, temporary driver's licenses  
cards) in your possession on or before 01/01/2015.  
surrender these items before, 01/01/2015, for  
credit; however, you may not drive after these i  
surrendered.





# What experience do you have working on similar redaction projects?

In redaction, as in almost everything in life, experience matters. An experienced vendor with dozens of similar implementations under its belt, knows what to expect, what mistakes to avoid and how to set up your project for success. The labor costs to verify redaction results are often the largest expense related to a project with millions of images that need to be redacted. An experienced redaction vendor understands this and can leverage knowledge from previous projects to set up a smooth and efficient verification workflow, saving you time and money.

Identity theft is in the news and on the radar of many counties and states. New vendors have sprouted up, responding to the increased demand for redaction technology to protect citizens from identity theft. Ask a vendor for a list of customers with similar projects.

## **CONTACT AT LEAST TWO PEOPLE ON THE LIST AND ASK THEM TO RATE THE VENDOR IN THE FOLLOWING CATEGORIES:**

- Expertise in document workflow
- Accuracy achieved
- Fully customizable set of document types and routing logic
- Responsiveness to support needs (How quickly did they respond to your problem?)
- How long did it take for them to resolve your problem?)
- Meeting deadlines

# 2

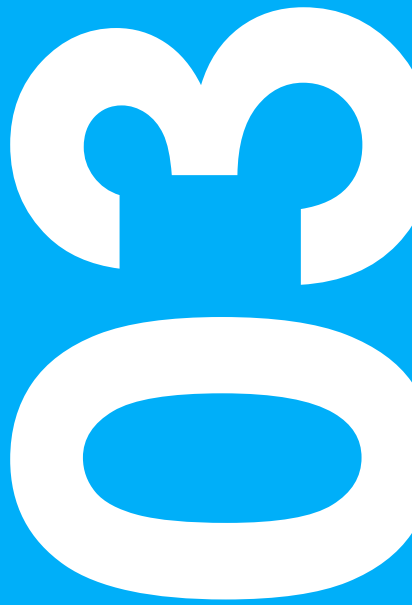
## How do you calculate the accuracy of your software in targeting privacy information—is that with or without manual verification?

You are, in essence, buying accuracy. Other software features and services provided by a vendor do not mean very much if the software does not achieve a high level of accuracy. Vendors have a responsibility to be as transparent as possible when it comes to calculating redaction accuracy rates. Some vendors cloak the calculation of accuracy under a shroud of mystery. The problem with these obscured calculations is that customers may not know exactly what they are getting. Under the worst circumstances, a customer might select a vendor without fully understanding the conditions under which a vendor is promising to deliver 99% accuracy. This customer may have no recourse when the vendor does not deliver the expected results.

Verification is a necessary step for most redaction projects that need to achieve accuracy rates of 98.5% or higher. Vendors have different methods of categorizing documents but most high volume redaction projects make distinctions between low, medium and high level confidence documents. These definitions vary but should be created with your input and with your specific project in mind.

A high confidence document might mean that the redaction software either found zero instances of sensitive information or located sensitive information but is highly confident that the data types meet all the conditions and should be redacted. Medium confidence may mean that the software has located a string pattern of eight-digits that resembles a Social Security Number pattern 123-45-6789 but has not located a clue. Low Confidence may mean that the software has not located a clue word or phrase, but has located a string pattern of nine-digits without hyphens. You quickly get a sense of how these small irregularities require some degree of human intervention to ensure the highest accuracy. Verification workflow becomes a very important part of the project and working with an expert in this area can save you time and money.

An important consideration related to accuracy is the quality of the redacted document. You take pride in delivering clean reports to customers that do not cover more of the record than necessary. Ask your vendor for samples of redacted documents. Make sure the redaction zones extend far enough to cover sensitive text but not beyond the edges of the text. Test drive the vendor's manual redaction tools to make sure they allow users to create precise redaction zones easily and intuitively.



# What can you do now to help us deal with future legislation that may require redaction of additional data types?

Legislative mandates have been issued requiring the redaction of additional data types beyond the scope of the original project. In 2003, the Florida legislature mandated the redaction of Financial Account information including bank, credit and debit card numbers. Similarly, the Nevada legislature issued a revised statute in 2006 to mandate the redaction of Driver's License numbers, Identification Card numbers, and Financial Account information including bank, credit and debit card numbers.

Redaction vendors can help clients successfully navigate these waters in the future with some key technology and deliverables today. One option is to generate reports identifying specific documents containing the additional data fields (credit card number, driver's license number, etc...) that may need to be redacted in the future. This captured information can be used to create a budget for the additional verification process and to isolate suspected images for automatic/manual redaction processing.

The passage of time presents some problems for this approach. It is a fact that documents change and data capture tools and techniques improve rapidly. Using old rules and old capture technology may decrease accuracy and/or increase verification labor costs. To maintain accuracy and keep manual labor costs low, a good solution would be to save the OCR output to avoid incurring the cost of rescanning, and write new custom rules for subsequent mandates as they arise. Ask vendors how they can help you minimize costs and maximize accuracy if additional data types need to be redacted in the future.



# What type of reporting capabilities are available with your software?

To bring a backfile project in on deadline and manage the labor costs associated with day-forward redaction, you need to be able to get a snapshot of your project at any given moment. The reality is that the reporting available with redaction software varies widely.

## **TO ENSURE THAT YOU WILL HAVE THE DATA YOU NEED WHEN YOU NEED IT, HERE IS A CHECK LIST OF REPORTS YOU SHOULD HAVE AT YOUR FINGERTIPS:**

- Accuracy Rate
- Number of files pending processing and verification
- Speed at which files are processing and getting verified
- Verification (productivity) rates of various verification operators
- Percentage of documents with redactions
- Percentage of documents being verified
- Speed at which machines are processing
- Number of documents processed on individual machines last hour, day, week, month



# Can your redaction software be tightly integrated with our document management system?

Vendors who get high marks on this question will undoubtedly be those with years of experience who have seen it all and done it all. The number of technology partners a vendor has is a good indicator of how many different types of integrations they have completed (a high number of partners indicates a high number of successful integrations).

## QUESTIONS TO ASK INCLUDE:

- Has the redaction vendor completed installations with your particular document management software?
- Is the redaction software able to “listen” for newly added files?
- Can the software notify your document management application when files are processed and ready to be archived?
- Can it track the X/Y coordinate locations for the redaction zones?