Redaction 101

Extract Systems, LLC
Updated: April 2008

Extract Systems
www.extractsystems.com

Identity Theft is the fastest growing crime in the U.S. the Federal Trade Commission reports. Over 800,000 fraud and identity theft complaints were received by the FTC in 2007.

32% of all FTC complaints last year were Identity Theft related; the largest single category of reported fraud. With the average cost of identity theft at more than $500 per incident, the collective fraud loss in 2007 was in excess of $1.2 Billion.

Security experts suggest the figure is actually much higher.

A Growing Legislative Concern

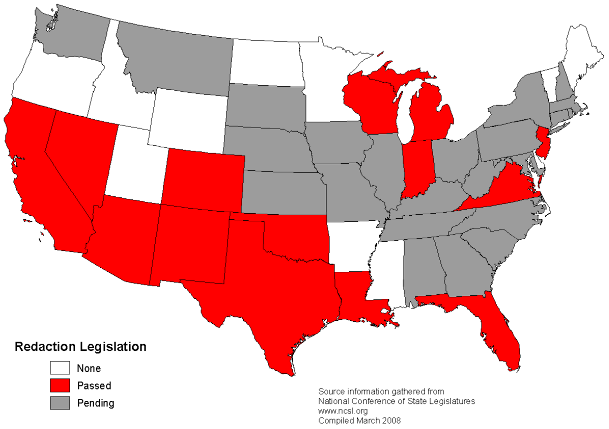
Across the United States, strict legislation is being passed requiring county governments to redact sensitive information, such as social security numbers and credit card numbers from the official and public record. In some states, forward-thinking local government officials have independently determined that it is their responsibility to protect their constituents from identity theft.

What is Redaction?

Redaction is the process of removing social security numbers or other sensitive information from images stored as Official Records by County Recorders, County Clerks and Clerks of the Court that may be accessed electronically via the Internet.

The traditional technique of redacting confidential material from a paper document before its public release involves crossing out portions of text with a wide black pen, followed by photocopying the result. This manual processing of thousands or millions of document pages is a time-consuming process that can strain staff resources. The question is rapidly becoming not whether to redact, but how to accomplish it in the most cost-effective and timely manner.

REDACTION LEGISLATION



Source information gathered from
National Conference of State Legislatures
www.ncsl.org
Compiled March 2008

Technology

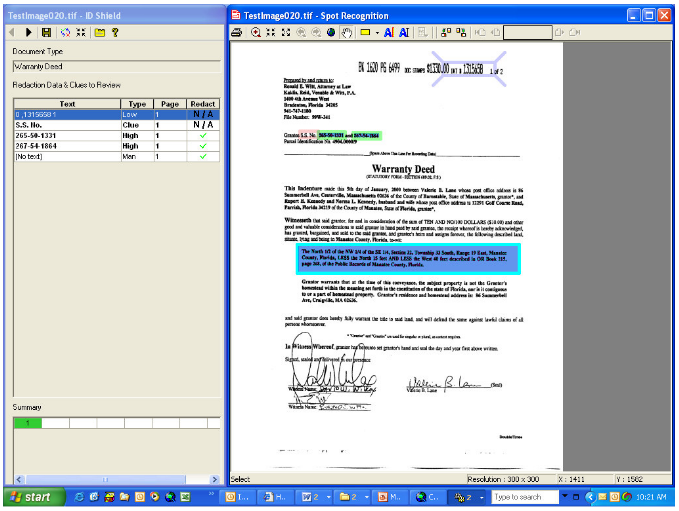
Significant technology advances during the past five years offer an automated option to the traditionally manual document redaction process. Instead of using a black marker on paper documents, many counties are now using OCR (Optical Character Recognition) to process their scanned images of the paper documents. OCR converts scanned images into a digital format which then allows rules-based search engines to locate sensitive information in the OCR results. The search engine uses a combination of words, phrases, patterns of text, proximity and location to identify potentially sensitive information. For example, the engine may find the clue word “SSN:” within a document followed by a pattern of numbers such as xxx-xx-xxxx. The combination of the clue word with the format of text provides high confidence that the potentially sensitive information needs to be redacted. More than 20 document management software providers have partnered with Extract Systems to provide automated redaction technology to their customers.

Verification Flexibility

After locating the sensitive information on an image, ID Shield can be configured to either automatically redact the data and post the newly protected image or send the information to an end user (verifier) to manually confirm the redaction. In the case of automated redaction, ID Shield creates a copy of the image and then "burns-in" the redaction zone. The advantage of this solution is that the original is maintained should a certified copy of the document be requested at the counter, while the redacted image is published for viewing via the Internet.

In a semi-automated verification workflow, ID Shield allows a step for an end user to verify the automatically located redaction zones prior to "burn-in". Both the image and sensitive information are presented to the verifier to "accept or reject" potential redaction zones. Each piece of sensitive information is presented to the verifier using different colored highlights to signify higher or lower confidence levels determined by the software. These confidence levels are based on user-defined criteria. If a verifier locates sensitive information on the image that was not identified automatically, a built-in redaction marker and "rubberband" tool allows him to quickly add the proper redaction zone at his discretion.

I.D. Shield



TestImage020.tif - ID Shield

Document Type: Warranty Deed

Redaction Data & Clues to Review

Text	Type	Page	Redact
0 ,1315658 1	Low	1	N / A
S.S. No.	Clue	1	N / A
265-50-1331	High	1	✓
267-54-1864	High	1	✓
[No text]	Man	1	✓

Summary: 1

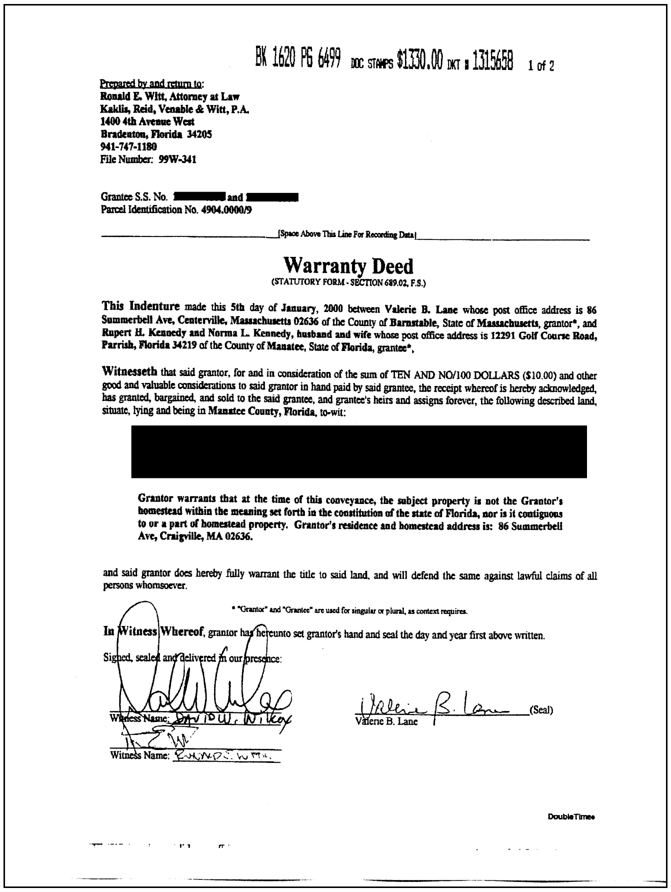
TestImage020.tif - Spot Recognition

Select | Resolution : 300 x 300 | X : 1411 | Y : 1582

Fig. 1 - ID Shield Verification Module displays sensitive data to the end user using different colors to indicate redaction zone confidence levels

Masking vs. Redacting

Because a true redaction process can be complex, some technologies utilize a process whereby a masking layer is placed over the sensitive information on a document.



BK 1620 PG 6499 DOC STAMPS $1330.00 DKT # 1315658 1 of 2

Prepared by and return to:
Ronald E. Witt, Attorney at Law
Kaklis, Reid, Venable & Witt, P.A.
1400 4th Avenue West
Bradenton, Florida 34205
941-747-1180
File Number: 99W-341

Grantee S.S. No. [redacted] and [redacted]
Parcel Identification No. 4904.0000/9

[Space Above This Line For Recording Data]

Warranty Deed

(STATUTORY FORM - SECTION 689.02, F.S.)

This Indenture made this **5th** day of **January, 2000** between **Valerie B. Lane** whose post office address is **86 Summerbell Ave, Centerville, Massachusetts 02636** of the County of **Barnstable**, State of **Massachusetts**, grantor*, and **Rupert H. Kennedy and Norma L. Kennedy, husband and wife** whose post office address is **12291 Golf Course Road, Parrish, Florida 34219** of the County of **Manatee**, State of **Florida**, grantee*,

Witnesseth that said grantor, for and in consideration of the sum of TEN AND NO/100 DOLLARS ($10.00) and other good and valuable considerations to said grantor in hand paid by said grantee, the receipt whereof is hereby acknowledged, has granted, bargained, and sold to the said grantee, and grantee's heirs and assigns forever, the following described land, situate, lying and being in **Manatee County, Florida**, to-wit:

[redacted]

Grantor warrants that at the time of this conveyance, the subject property is not the Grantor's homestead within the meaning set forth in the constitution of the state of Florida, nor is it contiguous to or a part of homestead property. Grantor's residence and homestead address is: 86 Summerbell Ave, Craigville, MA 02636.

and said grantor does hereby fully warrant the title to said land, and will defend the same against lawful claims of all persons whomsoever.

* "Grantor" and "Grantee" are used for singular or plural, as context requires.

In Witness Whereof, grantor has hereunto set grantor's hand and seal the day and year first above written.

Signed, sealed and delivered in our presence:

Witness Name: DAVID W. Wilcox

Valerie B. Lane (Seal)
Valerie B. Lane

Witness Name:

DoubleTime

Fig. 2 - Redacted Image

While this process may look like the document has been redacted, trained identity thieves have a bevy of software solutions available to strip the masked layer from the image and gain access to the sensitive information "underneath." ID Shield does not use "masking" technology. "Burned-in" redaction zones become part of the image. ID Shield's black redaction zones mimic the traditional image redaction process and cannot be stripped out to gain access to the sensitive information.

Accuracy

Manual redaction processes have produced accuracy rates ranging from 85-95%. The accuracy of ID Shield is based upon a number of factors, including image quality and verification process. ID

Shield has demonstrated the ability to redact images at a rate in excess of 95%.

Configuration

ID Shield is capable of running in a multi-processor, multi-threaded environment, meaning that multiple instances of the software can run at the same time on the same server increasing speed and throughput. In addition, ID Shield is installed on standard Windows-based server and workstation hardware.

Case Study:
Historical Backfile Redaction Service

Florida Association of Court Clerks and Comptrollers (FACC) Florida was the first state to require redaction of the official record. Extract partnered with FACC to set up a redaction service for 16 of their association counties. The backfile redaction of seven million images was completed in six months with no verification.

For the Florida counties it is a simple process: images are sent to ID Shield for automated redaction processing. Once the rules-based processing is complete, the new redacted images are accepted and are copied back into the image repository. Throughout the process care is taken to insure the security of the images and to fine-tune the redaction so that the highest quality is maintained.

Florida's State Statute's required that county recorders publish their land records online by the end of 2006, and that all publicly available images be redacted by January 1, 2011. Russell Curtis, the Director of Technical Services for the FACC recently recognized the effectiveness of ID Shield: "Officials across Florida are realizing that redaction is not an insurmountable problem. With ID Shield, the benefits far outweigh the small cost."

Identity Theft Facts

- In 2008, nearly 8 million Americans will be victims of Identity Theft
- Average amount of fraud per victim: $1,685
- Median consumer out-of-pocket expense: $349
- Average time to resolve the situation: 40 hours
- Nearly 10% of ID Theft originates with government records

Conclusion

With the alarming rise in identity theft across the U.S., redaction has become necessary to protect your constituents. While the redaction of thousands or millions of images may seem like a daunting task, it doesn't have to be. Today's technology allows you to continue offering access to public records, while protecting the residents of your jurisdiction.

ID Shield has been used to redact hundreds of millions of images for state and local governments throughout the United States. Whether or not state or federal mandates require the redaction of sensitive information on your images, proactive redaction is sound public policy. We encourage you to contact us today to learn how ID Shield's technology can be used to protect the sensitive data on your images.

Data Collected from:
1. Federal Trade Commission, Consumer Fraud and Identity Theft Data, January – December 2007, released February 2008.
2. National Conference of State Legislatures, www.ncsl.org
3. Javelin Strategy & Research, 2008 Identity Fraud Survey Report, released February 2008.