BOARD OF SUPERVISORS
AGENDA LETTER

Clerk of the Board of Supervisors
105 E. Anapamu Street, Suite 407
Santa Barbara, CA 93101
(805) 568-2240

Department Name:	Clerk-Recorder Assessor
Department No.:	062
For Agenda Of:	10/28/2008
Placement:	Administrative
Estimated Tme:	5 min/
Continued Item:	No
If Yes, date from:	
Vote Required:	Majority Vote

TO: Board of Supervisors

FROM: Department Director(s): Joseph E. Holland, Clerk-Recorder Assessor
568-2558
Contact Info: Melinda Greene, Project Manager 568-2252

SUBJECT: **Waiver of Competition to Purchase OCR Software for the Recorder Information Imaging Management System (RIIMS)**

County Counsel Concurrence	**Auditor-Controller Concurrence**
As to form: Yes	As to form: Yes

Other Concurrence: Purchasing
As to form: Yes

Recommended Actions:

That the Board of Supervisors:
1. Authorize the Purchasing Agent to procure OCR Redaction and Indexing tools for implementation within the existing Recorder Information Imaging Management System (RIIMS) without a formal bid process and authorize the proprietor of RIIMS, DFM Associates, as the sole source vendor in the amount of $180,000 (up to $32,000 for ongoing maintenance).

Summary Text:

The Clerk-Recorder Assessor would like to expand existing services with DFM Associates to implement Optical Character Recognition (OCR) automation technology for two distinct products offered: redaction of Social Security Numbers (SSN) from the official records inventory and automated data entry for indexing.

One of the core functions for the Recorder is to maintain a permanent archive of documents related to recordable transactions, such as deeds and maps relating to real property sales, known as official records. Based on testing, approximately 7% of these documents have social security information. On April 22, 2008, the Board implemented the provisions of California Government Code 27361 and created a new county program to redact social security information from the Recorder's official records

retroactive to 1980, and prospectively starting in January 1, 2009. The Board established a $1.00 recording fee with potential to generate $800,000 over the life of the project. In this board letter, the Board was also alerted that the Recorder would need to implement OCR technology as part of the program and provided notice of intent that "This software will be procured through the existing blanket purchase order with DFM Associates, and implemented within the existing RIIMS system." Due to the cost of the OCR tools exceeding the materiality threshold of the existing purchase order, and since the services are unique to the basic support of RIIMS, the Recorder is requesting authority for the purchasing agent to procure the OCR services without a formal bid process.

There are certain economies of scale and cost savings realized if the Automated Indexing software is implemented concurrently with the redaction project since it relies on similar OCR technology and analyst services. This technology will reduce data entry and help the Recorder cope with peak volumes at month end and unanticipated high volume days.

Background:

The Recorder has a long standing relationship with DFM who designs and supports the Recorders' Information Imaging System (RIIMS) in operation since 1989. DFM has partnered with a premiere developer of OCR software, Extract Systems, and has customized the tools to work with RIIMS. DFM's experience with the Recorder's data and ability to train and provide support services to our staff is unparalleled due to our long standing relationship. Although the Recorder has tested a manual solution for redaction for implementation by the January 1, 2009 legislative deadline, our goal is to implement the OCR software for a higher confidence level of 98-99% effectiveness. While it may be possible to find an alternative partner willing to contract with DFM, it is not practical for DFM to build customized interfaces with several vendors within the timeframe allotted. Also, the Recorder prefers to maintain the information technology support services through DFM instead of dealing with multiple vendors' help desks.

REDACTION
Identity theft continues to occur with alarming frequency, with the social security number as the single most useful tool. To help protect social security numbers (SSNs), the California State Legislator updated Government Code Section 27301 to require local government agencies to truncate (redact the first five digits of) SSNs that appear in their records that may be disclosed to the public unless full disclosure of the SSN is required by state or federal law. Among other things, this legislation requires the recorder of each county to establish a social security number truncation program to create an electronic public record version of each official record. The public version is the same as the official, except that the first five digits of the SSN will be redacted. The new law specifically states that the public version has the same legal force and effect as the official record. It also requires that the Recorder provide the official record only in response to a subpoena or court order.

The potential redaction population from 1980 through September 2008 is approximately 2.9 million documents (8.7 million pages). Of these documents an estimated 7%, 192,000 documents, have social security numbers which will need to be redacted.

In order to sample the redaction population, documents were divided into two categories based on document types. These categories were documents with a high probability of having a social security number and those that had a remote possibility. The samples were further stratified into one year

increments under the assumption that the inclusion of social security numbers on documents would change over time. In a random sampling, social security numbers appeared in the "remote" population 4% of the time. Since social security numbers were not exclusive to the "high probability" population of documents, a 100% review of all 1980-2008 documents is needed to ensure a high level of completeness.

Based on the projected revenues over the life of the program, the Recorder anticipates that the revenues will be sufficient to fully redact the backlog of images, using a combination of OCR technology and staff validation.

AUTO-INDEXING

The Recorder uses the RIIMS to manage the official records archives available to the public. The process for storing these documents includes imaging or scanning the document and keying pertinent data such as grantor or grantee and other indexed fields in order to recall the image at a later date. This data entry process is called indexing. If a field is not keyed properly, it could result in the inability to find this record at a later date. Recorders have a low tolerance for error and implement a technique called blind key verification in order to mitigate this risk. The process involves a first person identifying the indexed fields from a document and keying them into the RIIMS database, then a different person, unaware of the indexer's work, keys the same indexed fields for comparison. If there are differences with the two sets of indexed data, the verifier must review and select which set of data is correct. All indexing must be completed by the end of each business day, which complicates matters on a high volume day.

The concept of Auto-Indexing attempts to automate some portion of the manual data entry by identifying these fields from an OCR scanned document. The data is still manually verified and reconciled with the Auto-indexed data so the risk of error does not rise in our official records index if the technology does not perform optimally.

There are several motivations to implement Auto-Indexing including cost, convenience of managing peak volume and our hesitancy to fill vacant position in light of the budget challenges facing the County. The return on investment would be less than four years resulting in an ongoing savings of $20,000-$30,000 assuming that the volume of recording will grow at a reasonable pace. One of the challenges that Recorders are faced with is the erratic volume in work load day to day, particularly on the last day of the month. On a large volume day where the work cannot be completed by the scheduled clerk hours, the Recorder may need to authorize overtime, or have the work performed by a more expensive manager who delays other more strategic projects for the division.

The Recorder will model implementation of Auto-Indexing through RIIMS after the Santa Cruz Recorder's Office, which has been in production for almost two years.

Performance Measure:

Auto-Indexing:
Staff will compare the actual rate of OCR indexing to manually keyed indexed data for the target OCR population.
Track productivity rate of minutes of staff time/100 documents indexed.

Redaction:
Staff will compare the rate of discovery for manual redaction compared to the OCR redaction.

Fiscal and Facilities Impacts:

Fiscal Analysis:

Funding Sources	Current FY Cost:	Annualized On-going Cost:	Total One-Time Project Cost
General Fund	$ 7,000.00	$ 7,000.00	
State			
Federal			
Fees	$ 127,000.00	$ 25,000.00	$ 113,000.00
Other:	$ 70,000.00		$ 70,000.00
Total	$ 204,000.00	$ 32,000.00	$ 183,000.00

Narrative:
Redaction expenditures include maintenance fees to DFM at $2,000 per month, a per image fee of $0.013, and hardware for $3,000.

One-Time fees for DFM include scanning the estimated 8.5 million retroactive inventory (1980-current year) at $110,000 plus $3,000 department expenses for the acquisition of hardware for a total of $113,000.

Ongoing, the maintenance services for the program includes $24,000 in maintenance fees, and $1,000 for prospective OCR scanning for a total of $25,000 annually.

The Recorder is using a bridge loan established in the April 22, 2008 Board Letter, funded by the Recorder's Automation Designation (Line Item 9744), to provide the capital to fund the program for the first few years. The current year fiscal cost is $2,000 for prospective scanning, $12,000 for half a year of maintenance, and $113,000 in one-time costs for a projected $127,000 in redaction expenditures.

Auto-Indexing
Estimated auto-indexing expenditures for FY 08-09 include an initial set up fee of $70,000 and $7,000 in annual maintenance fees to DFM. The funding source for the one time set up fee is the Recorder's Automation Designation (Line Item 9744) while the ongoing maintenance will be absorbed by the Recorder's operating budget.

Budget: The Recorder will defer existing budgeted projects that use the same funding source (Automation Designation), in order to fund both projects.

Staffing Impacts:

Legal Positions: FTEs:

Redaction
Although partially automated, an estimated 2250 clerk hours are required for validation to complete the redaction workload retroactive to 1980. The ongoing workload is estimated at 100-130 clerk hours per year. The work will be performed using existing staff.

In addition, the project will require project management and systems analyst skills since the project will involves implementation of software and hardware.

Auto-Indexing
The staff time would be reduced by 40% ($31,000/year) for the task of indexing since this is an ongoing function that would be partially automated. This savings in staff time would be redirected to the division manager and supervisor to implement projects and perform quality assurance testing, and reduce the potential for overtime. Implementation of the project would also require project management and systems analyst skills since the project involves implementation of software and hardware.

Special Instructions:

Send copy of minute order to Rose Rodarte and Melinda Greene in the Clerk-Recorder-Assessor

Attachments:

Authored by: Melinda Greene

cc: