



**INFORME TECNICO PREVIO DE EVALUACION DE
SOFTWARE N° 003 -2007- 2A2000**

1. NOMBRE DEL AREA:

OFICINA DE INVESTIGACIÓN TECNOLÓGICA

2. RESPONSABLES DE LA EVALUACION:

OMAR SEGUNDO GUTIÉRREZ COAGUILA

3. CARGOS :

ANALISTA DE TECNOLOGIA

4. FECHA

31 DE MAYO DE 2007

5. JUSTIFICACIÓN

La Gerencia de Desarrollo de Sistemas Aduaneros a través del Memorando Electrónico 00093-2006-2A6000 solicitó la evaluación de herramientas de búsqueda con la finalidad de usar dichas herramientas con los sistemas informáticos que desarrollan.

6. ALTERNATIVAS

De acuerdo a lo solicitado en el Memorando Electrónico 00093-2006-2A6000. Se identifican como alternativas:

- Apache Lucene
- Oracle Intermedia Text

7. ANALISIS COMPARATIVO TECNICO

Teniendo en consideración el siguiente contexto.

En la Administración Pública, tenemos,

- Por medio de la Resolución Ministerial 274-2006-PCM del 25.JUL.2006 se aprueba la Estrategia Nacional de Gobierno Electrónico, elaborada por la Oficina Nacional de Gobierno Electrónico e Informática; donde se especifica la necesidad de una definición de estándares abiertos que permitan la interoperatividad de los sistemas dentro del sector público y con el sector privado; y que consiste en "Tecnologías y componentes de arquitectura abierta y modular que permitan la interoperatividad de aplicaciones de Gobierno Electrónico (protocolos, interfases, mensajes, estructura de datos, encriptación)".



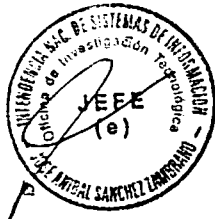
En la SUNAT, tenemos,

- La directiva “Desarrollo de aplicaciones utilizando tecnología web en un modelo de ‘n’ capas, aprovechando la reutilización, portabilidad y escalabilidad de los componentes de software” de la categoría “Procesos Internos” perteneciente a las directivas de la INSI del Plan Maestro, versión 2005.
- La directiva “Selección de herramientas sobre la base de estándares abiertos, buscando las mejores soluciones que se integren con la plataforma tecnológica evitando así la dependencia de proveedores” de la categoría “Tecnología” perteneciente a las directivas de la INSI del Plan Maestro, versión 2005.

Para realizar el análisis comparativo de las herramientas se han definido los factores técnicos de evaluación, los cuales representan a los criterios mínimos que la herramienta debe cumplir para poder funcionar adecuadamente en nuestra Arquitectura Informática.

La herramienta que no supere el estimado mínimo deberá ser descartada. La siguiente matriz muestra la evaluación:

	Puntaje Mínimo	Puntaje Máximo	Apache Lucéne	Oracle Intermedia Text
TOTAL	74	100	100	77
ATRIBUTOS INTERNOS	40	54	54	46
Permite búsquedas por aproximación PUNTAJE Si cumple máximo puntaje, si no cumple cero		7	7	7
Permite búsquedas con eliminación de palabras comúnmente usadas que sean previamente definidas PUNTAJE Si cumple máximo puntaje, si no cumple cero		7	7	7
Permite búsquedas usando operadores lógicos PUNTAJE Si cumple máximo puntaje, si no cumple cero		7	7	7
Permite búsquedas por múltiples campos PUNTAJE Si cumple máximo puntaje, si no cumple cero		8	8	8
Permite búsquedas por aproximación fonética PUNTAJE Si cumple máximo puntaje, si no cumple cero		7	7	7
Métodos de actualización de información una vez indexado. Manual: Ejecución a discreción de una persona Automático: Ejecución automática del procedimiento, actualiza toda la información desde la última ejecución. Se debe especificar la frecuencia o momento de actualización Incremental: Procedimiento automático que detecta los cambios y los indexa. PUNTAJE Permite procedimiento manual: 10% del puntaje Permite procedimiento automático: 30% del puntaje Permite procedimiento incremental: 60% del puntaje	6	10	10	10
Es posible realizar modificaciones a los fuentes del software para realizar adecuaciones a las necesidades propias. PUNTAJE Si cumple máximo puntaje, si no cumple cero.		8	8	0



ATRIBUTOS EXTERNOS:	19	26	26	16
El software debe poder licenciarse independientemente PUNTAJE Si cumple máximo puntaje, si no cumple cero.		10	10	0
El software debe poseer soporte para el buen funcionamiento PUNTAJE De tipo comercial y público (listas de interés y otros) : 100% del puntaje Sólo de tipo comercial : 70% del puntaje Sólo de tipo público (listas de interés y otros) : 30% del puntaje		10	10	10
Debe poder ejecutarse en plataformas de sistema operativo Windows 2000 Advanced Server o alguna versión de Red Hat Linux o UNIX PUNTAJE Se ejecuta en una de las opciones descritas: un tercio del puntaje Se ejecuta en dos de las opciones descritas: dos tercios del puntaje Se ejecuta en todas las opciones descritas: 100% del puntaje		6	6	6
ATRIBUTOS DE USO	15	20	20	15
Porcentaje de crecimiento de espacio en disco al momento de indexar, respecto de los datos originales PUNTAJE [100%,150%>: 100% del puntaje [150%,200%>: 50% del puntaje 200% o más: Cero		10	10	5
Posee API's y/o interfaces que permitan la realización de consultas al motor de búsqueda desde aplicativos desarrollados por el usuario. PUNTAJE Si cumple máximo puntaje, si no cumple cero.		10	10	10

8. ANALISIS COMPARATIVO DE COSTO - BENEFICIO

Los costos asociados al producto incluyen:

- **Licenciamiento:** La estimación de costos de licenciamiento de Oracle Text se realiza sobre un equipo monoprocesador; se supone que a más procesadores el costo aumentará. Para usar el Oracle Text es necesario contar con una licencia de Oracle, por lo que se realizarán los cálculos con una licencia monoprocesador, que se encuentra valorizada en USD \$ 5,944.05. El licenciamiento del Apache Lucene es cero.
- **Hardware:** Se cuenta con ambientes para alojar el motor de búsqueda que se seleccione, por lo que no es necesario realizar inversiones adicionales.
- **Soporte al buen funcionamiento y mantenimiento:** Este costo da derecho a la actualización permanente de versiones; así como al soporte para buen funcionamiento. En el caso del Apache Lucene sólo contempla el soporte para buen funcionamiento ya que la actualización de versiones es gratuita. Los costos para Apache Lucene a corto, mediano y largo plazo son USD \$ 3,427.20, USD \$ 10,281.60 y USD \$ 17,136.00 y los de Oracle Text son USD \$ 1,307.69, USD \$ 3,923.07 y USD \$ 6,538.46. Los costos de licenciamiento y capacitación ocurren al momento de la implementación; sin embargo el costo de soporte al buen funcionamiento y mantenimiento es un costo que se desembolsa mensualmente durante un plazo definido, por lo que se debe determinar el Valor Presente Neto para poder sumarlos con los otros costos. Para determinar el Valor Presente Neto se usará la fórmula "VA" de Microsoft Excel y la tasa mensual se encontrará en función de la Tasa Social de Descuento¹.



¹ La Tasa Social de Descuento mide el costo efectivo para la sociedad de utilizar capital en una inversión y es utilizada para descontar los beneficios y los costos futuros de un proyecto de inversión gubernamental y convertirlos en valores presentes. Actualmente es del 11% y se encuentra definida en <http://www.mef.gob.pe/DGPM/docs/indicadores/informe-tsd.pdf>. La tasa mensual equivalente es 0.8734594%

- Capacitación e Implementación: Ambas cotizaciones contemplan costos de capacitación e implementación, que se detallan en el cuadro adjunto.
- Debido a que se solicita un módulo que posibilite la funcionalidad de búsqueda, el beneficio es contar con esta funcionalidad para uso en el desarrollo de los sistemas de información, que no puede expresarse cuantitativamente. Por este motivo, se realizará un análisis de costos a corto, mediano y largo plazo; el beneficio puede considerarse como una constante.

El siguiente cuadro muestra el análisis de costos:

	Corto (1 año)		Mediano (3 años)		Largo (5 años)	
	Apache Lucene	Oracle Text	Apache Lucene	Oracle Text	Apache Lucene	Oracle Text
Análisis de Costos	3.240.30	10.399.38	8.789.39	12.516.70	13.293.15	14.235.17
Costo de licenciamiento	0.00	5.944.05	0.00	5.944.05	0.00	5.944.05
Costo de soporte y mantenimiento	3.240.30	1.236.38	8.789.39	3.353.70	13.293.15	5.072.17
Costo de capacitación e implementación	0.00	3.218.95	0.00	3.218.95	0.00	3.218.95

De lo expuesto se identifica:

- Ambas alternativas cumplen con superar el puntaje técnico mínimo. La alternativa de Apache Lucene obtuvo 100 puntos y la de Oracle, 77.
- La alternativa de Apache Lucene tiene menor costo de implementación en los tres escenarios.
- Apache Lucene proporciona mayor flexibilidad en crecimiento de instalaciones; ya que ante un aumento de demanda, al contar con licenciamiento libre es posible realizar las implementaciones adicionales necesarias; con la otra alternativa debe realizarse un proceso de adquisición que adiciona tiempo a la implementación.

9. CONCLUSIONES

El Apache Lucene es el producto que obtiene el mejor puntaje técnico y es el que tiene menores costos de implementación. Por este motivo se sugiere que se incluya al Apache Lucene como una opción de motor de búsqueda a ser usada por los sistemas de información.

10. FIRMAS

