Description of document:	Federal Election Commission (FEC) Web Search Study Project, A Study Report For Enterprise Search Solution, 2011
Requested date:	13-March-2019
Release date:	04-November-2019
Posted date:	23-March-2020
Source of document:	FOIA Request Federal Election Commission Attn: FOIA Requester Service Center 1050 First Street, NE Washington, DC 20463 Fax: 202-219-1043 Email: FOIA@fec.gov

From: Steve Hajjar <shajjar@fec.gov>
Cc: Press Office <press@fec.gov>
Sent: Mon, Nov 4, 2019 3:34 pm
Subject: Freedom of Information Act Request to the Federal Election Commission FOIA No. 2019-048
VIA ELECTRONIC MAIL

Re: Your Freedom of Information Act Request to the Federal Election Commission FOIA No. 2019-048

This email is in response to the request you filed for information under the Freedom of Information Act (FOIA) dated and received by the Federal Election Commission's (FEC) FOIA Requester Service Center on March 13, 2019. Specifically, you requested:

> "a copy of the report from the Static/Dynamic Web Search Study performed under contract FE11G096, parent award GS35F5565H, by contractor TFS Group Inc., supplied to FEC in 2011/2012."

We have searched our records and have located responsive documents, which we are releasing in part. Attached to this letter is a twenty-four (24) page PDF containing responsive records the Agency located that are not exempt from disclosure. We have withheld two (2) pages of responsive records in their entirety under FOIA Exemptions 5. We have also released ten (10) pages with redactions under Exemption 5. Please note that our response to your request does not include documents or publications publicly available on our website or compilations of publicly available news articles.

Exemption 5 protects from disclosure inter- or intra-agency memoranda or letters that would not be available by law to a party other than an agency in litigation with the agency, including documents covered by the attorney work-product, deliberative process, and attorney-client privileges. See 5 U.S.C. § 552(b)(5).

Accordingly, your FOIA request has been granted in part.

To the extent that the records you requested concern an ongoing FEC enforcement matter, we can neither confirm nor deny that any such records exist. To protect the interests of those involved in the Agency's enforcement process, the Federal Election Campaign Act (FECA) requires that any Commission action on an enforcement matter be kept confidential until the case is resolved. See 52 U.S.C. § 30109(a)(12). Therefore, records concerning an open enforcement matter are exempt from disclosure under FOIA Exemption 3(A) pursuant to section 30109(a)(12)(A) of the FECA. FOIA Exemption 3(A) protects from disclosure information that is specifically exempted from disclosure by statute. 5 U.S.C. § 552(b)(3)(A).

You may contact our FOIA Public Liaison, Christine McClarin, at (202) 694-1485, for any further assistance and to discuss any aspect of your request. Additionally, you may contact the Office of Government Information Services (OGIS) at the National Archives and Records Administration to inquire about the FOIA mediation services they offer. The contact information for OGIS is as follows:

Office of Government Information Services
National Archives and Records Administration
8601 Adelphi Road-OGIS
College Park, MD 20740-6001
e-mail at ogis@nara.gov
telephone at 202-741-5770
toll free at 1-877-684-6448; or
facsimile at 202-741-5769.

You may appeal any adverse FOIA determination. Any such appeal must be filed in writing and should follow the guidelines set forth in 11 C.F.R. § 4.8. If you have any questions, please contact the FOIA Requester Service Center at FOIA@fec.gov, or (202) 694-1650.

Sincerely,
Patricia Washington

Federal Election Commission

Web Search Study Project

A Study Report For

Enterprise Search Solution

Evaluation | Assessment | Recommendation

Conducted by

TFSG

8605 Westwood Center Dr., Suite 207, Vienna VA 22182 | 703.992.6723 | tfsginc.net

Table of Content

1 Introduction

As an organization grows so does the information it collects, stores, catalogs or categorizes for later use or benefit. Along with this growth, the need to access the collected information also begins to grow proportionally. Meanwhile, the influx of newly added information also ramps up at a phenomenal pace. All of the above factors related to the information usage eventually come to a head where the ease with which collected information could be recalled begins to decline while the other factors continue to increase. This phenomenon commonly referred to as "Information Overload" then begins to sap an organization's productivity due to the lack of relevant information available at the time of need. Thus, a well-designed and efficient enterprise search and retrieval system is paramount for the success of an organization and its implementation must be considered an integral part of an organization's business infrastructure.

While a web search applies to documents and information available on the public internet and desktop search applies to the documents and information contained on a single computer, enterprise search applies to the documents, data and information contained within the enterprise itself. Thus, the enterprise search realm may traverse multiple servers, networks, systems and repositories in the logical and physical sense to amass ready collections that can be swiftly and easily served to the inquiring users on demand.

Not only does the Federal Election Commission (FEC) own, manage and maintain the federal election campaign finance data, but also oversees a whole set of other administrative, business, and compliance related data as well.

The campaign finance data also known as the disclosure data is mostly open to the general public and actually is mandated to be disclosed within 48 hours of its initial collection. However, all other types of information or data, be it administrative, business, compliance and or disclosure related, may or may not be open to the general public access. Data and information collected and preserved as a result of processing the disclosure data or exercising the core business functions of FEC mission is also needed for access and usage, however, this type of data remains under strict control and security by FEC. As such, this type of data could only be accessed or used by internal FEC personnel or by those within FEC who are authorized to view it.

1.1 Background and Issues in Content Provisioning

Currently, FEC maintains all content in loosely bound yet fragmented repository structures consisting of directory folders, various databases and a data warehouse. Provisioning secure access to its own user community and general public access to the disclosure data then becomes a significant challenge that FEC has consistently faced. It has come to a point where access to this fragmented content by internal users and the general public for ready search and retrieval of pertinent information has become challenging if not near impossible.

This challenge prompted FEC to deploy custom developed search and retrieval solutions for each specific content repository. Even though it has solved the content access issue, it has, by default compartmentalized each repository paired with its significant users thus preventing cross-repository access or search capabilities. This limitation not only prevents users from accessing content in other repositories, but also eliminates the

possibility for users to even find out about any related content that may exist in other repositories from the one they normally access or search. The end result has been the formation of information silos that aggregate redundant or duplicate information within the organization which does not serve FEC stakeholders well in the process.

1.1.1 Usage and Access Profile of Content Search Users

FEC mainly provides content to the following two types of user groups:

- **Internal FEC User Community** – This user group is further divided across content access and privileges as determined by FEC according to their specific roles within the organization and functions they perform. In addition to the disclosure data, this group may also access enterprise data related to audit, compliance, enforcement, review, management, reporting, internal work flow, performance and case tracking information. Aside from the disclosure data available to the general public, all other data may or may not be classified, yet remains off-limits to the public unless and until it is approved by FEC or FOIA requests.
- **General Public** – This user group can only access the disclosure data this is made available through the 48-hour information disclosure mandate and is unclassified. This group cannot access any classified or non-disclosure related information.
- **Watchdog and Special Groups** – These types of user groups also only gain access to the disclosure information, however, they can access this information through direct downloads of large data sets or special requests.

1.2 Panacea for Content Compartmentalization Syndrome

FEC's current leadership at the Information Technology Division (ITD) has recognized the content provisioning and redundancy issues highlighted within their enterprise. ITD's leadership is actively moving to introduce a solution that can leverage an efficient use strategy of the organizational content allowing both internal and external users access to all content repositories through a singular search portal. Not only will such a solution enable accessibility of all content fragments to the users, but also provide access to related content in other repositories that remain hidden from user view in the current environment.

1.2.1 The Web Search Study

The existing content search deployed through the fec.gov website is the primary focus that must be revamped and/or re-structured to enable the organizational goals mentioned above. In keeping with its traditional approach of eliciting technical solutions from the industry, FEC has contracted TFSG to produce:

1. A thorough study of the FEC Web Site, databases and Google Mini implementation in light of the existing search environment
2. Conduct market level research to investigate, review and select industry standard enterprise search tools or solutions
3. Develop assessment and evaluation models to produce an impartial comparison of all selected tools that are to be presented to FEC for review

4. Recommendations on new search functionality including architecture, design and infrastructure
5. Recommendations on an industry standard and commercially developed enterprise search solution that is efficiently deployable and exerts a minimum resource commitment on FEC's part

Upon the conclusion of this study, FEC will review and select a particular solution that will be procured for implementation in the Phase-II of this contract. TFSG is also expected to assist FEC in implementing, tuning and optimizing the solution after initial implementation for enterprise wide production readiness. Also, as part of Phase-II, TFSG will arrange for a training program such that the implemented solution could be easily managed and maintained by FEC in-house.

From this study, FEC aims to gain valuable foresight regarding the enterprise search solutions landscape, as well as a gauge of each solution's capabilities and how it could enable FEC in achieving its mission. Besides offering a glimpse of industry strength solutions in the enterprise search arena, this study will also empower FEC with a well-informed and balanced view to arrive at the best decision for its enterprise mission.

1.3 FEC Content and the Enterprise Search

The following outlines the type of content that is currently searchable and the areas FEC would like to add to make their enterprise search more comprehensive and useful:

1.3.1 Existing Content Search Scope

Currently, the existing searchable content remains within the scope of the fec.gov website as further elaborated in section 3.1.1 FEC.GOV Website Content Overview.

1.3.2 Future Requirements for Enterprise Search

The above mentioned issues have given rise to the need for an advanced and comprehensive search environment that will include dynamic content in addition to the static content currently available for search.

1.3.2.1 Dynamic Content Search

As listed in the statement of work, the FEC dynamic content systems comprise:

- Compliance Map – including Contributions, Candidate Search, Campaign Finance Summary
- View or Download Electronic Filings
- View Images of All Finance Reports
- Disclosure Blog
- Downloadable Data Files
- Independent Expenditure Notices
- Audit Reports
- Enforcement Query System
- Commission Calendar
- Advisory Opinions Search

It was discovered that some of the content in the dynamic area is also of Non-Database nature. The following are examples of some of this content based on the information received from FEC upon discovery request:

Table 1.3.2.1 Non-Database Dynamic Content

Item	Content Type	Dynamic Content	Remarks
1	**Audit Reports**	Non-Database Content, XML Driven	Predominantly static
2	**Enforcement Query System**	No Information Available	This content resides outside of FEC Control (with NIC)
3	**Advisory Opinions**	No Information Available	
4	**Matters Under Review**	No Information Available	
5	**Administrative Fines**	No Information Available	
6	**Commission Meetings**	Non-Database Content	Static PDF, SHTML files (1,500)
7	**Public Hearings**	Non-Database Content	
8	**Calendar**	Oracle Database	Very small database
9	**Audio Recordings**	Non-Database Content	Audio files MP3 (800)

The above list points to the need for considerable content and data cleansing and re-structuring pertaining to each of the various content types.

1.3.2.2 Unified Enterprise Search – Static & Dynamic Content

Considering the shape of the above listed environment, the content provisioning issues and the need for searching dynamic content points to a hypothetical search solution where a singular search portal may be leveraged to extract multiple static content repositories, file systems, and dynamic content residing in an enterprise database, a data warehouse, and in any other form.

Furthermore, considering the FEC's unique data and data usage environment across internal and external users, the search solution must provide an added utility of answering basic and fundamental business questions in the form of drill down, targeted or narrowing functions, and topic or category based clustering and grouping. These capabilities should then be seamlessly integrated into the overall search layout such that any user while searching the FEC data may also discover additional related information or data in an area the user may not have envisioned. This particular functionality will greatly improve the data usage, introduce efficiency in the FEC business workflows and enable users to fully leverage the power of FEC's data for their specific needs. It is important to note that although these needs may be of different attitude and nature, nonetheless, empowering the users would certainly be a significant accomplishment by FEC.

1.3.2.3 Identification of Minimum Technical Requirements

Although minimum technical requirements have be laid out in the Statement of Work, it is important to explicitly define them at the beginning of this study such that they become a measure for the study as well as the resulting assessment and evaluation of the tools reviewed.

The technical requirements for the enterprise search initiative have been categorized in the following basic areas:

- Search Architecture

- User Interface
- Standard Built-In Features
- Administrative Functionality

The following table defines and outlines these requirements in detail:

Table 1.3.2.3 Minimum Technical Requirement for Enterprise Search Tool/Solution

Item	Requirement Category	Requirement Detail	Remarks
1	**Architecture**	(b) (5)	
2		(b) (5)	(b) (5)
3		(b) (5)	
4		(b) (5)	
5		(b) (5)	
6		(b) (5)	(b) (5)
7	**User Interface**	(b) (5)	(b) (5)
8		(b) (5)	
9		(b) (5)	
10		(b) (5)	(b) (5)
11		(b) (5)	(b) (5)
12		(b) (5)	(b) (5)
13		(b) (5)	
14	**Built-in Standard**	(b) (5)	
15	**Search Features**	(b) (5) te	
16		(b) (5)	
17		(b) (5)	
18		(b) (5)	
19		(b) (5)	
20	**Admin**	(b) (5)	
21	**Functionality**	(b) (5)	

1.4 Organization of this Study Report

The purpose of this study is to not only document the existing search environment and recommend an enterprise search tool to FEC, but also to craft a rationale for selecting an optimum solution. The resulting study aims to empower FEC in weighing the proposed solution in the light of:

- Its capabilities and feature set
- FEC's resource limitations
- Need for a maintenance free search environment, and

- Organizational fiscal constraints

We strongly feel that by using this study, FEC will be in a position to exercise a well-balanced and informed decision in choosing the best search solution.

This study includes various sections that touch upon the need for such a study to be conducted, the existing environment, the methodology adopted, and the actual study. The following provides a snapshot of these sections and their flow such that the purpose, the methods and the results are clear and easily discernible:

- **Section 1. Introduction**: Provides a background and overview of the FEC enterprise content, its variations and specific need for different groups of users to access it. In addition, this section describes the type of content that is currently searchable, also defines the future search needs, as well as the actual minimum requirement. This section also provides a road map and flow of the overall study report.
- **Section 2. Enterprise Search – An Overview**: Provides a preamble to actual study and defines what should constitute enterprise search within the context of FEC and according to general industry guidelines.
- **Section 3. Study: TFSG Approach to Researching, Evaluating and Recommending Enterprise Search Tools**: Besides containing the actual study, this section also provides TFSG's research methodology and the evaluation and assessment mechanisms developed and used to select industry leading tools. Furthermore, this section establishes the foundation to create a fair and impartial assessment comparison that offers clear and concise information for FEC to make informed decisions.

  This section also offers highlights from the proof-of-concept demonstrations conducted by the selected finalists and provides TFSG's recommendations based on the FEC requirements for future search functionalities and the requisite architecture involved.

In its entirety, this study report documents the existing environment, search related issues encountered due to its inherent limitation, future content dissemination and access expectations, and the methodology adopted to conduct the study and chronicle its results. Once submitted, this study report should provide the proof-of-concepts in the context of FEC data using the industry leading enterprise search tools so that FEC can intelligently select and adopt the most suitable solution while adhering to its fiscal responsibilities and resource limitations.

2 Enterprise Search - An Overview

The concept of provisioning a concise and defined set of information contained within an organization's data repositories, directory structures, databases or data warehouse, and the organizational intranet entails the enterprise search system and all of the above listed areas constitute the enterprise search realm. In addition, enterprise search accommodates a multitude of roles and access privileges across the content user community such that inherent security features filter out search results based on the user's access profile.

The implementation of a comprehensive enterprise search realm, however, involves a defined set of activities, processes and functions that must be carried out to assure that all stored, changed, updated and continuously added information within the enterprise remains available at all times to serve the users' search demands.

2.1 Components of an Enterprise Search Implementation

In an enterprise search systems, content goes through various activities and processes from a source repository to search results. These functions and activities could be loosely termed as components of an enterprise search system. In addition, the various content repositories to be searched and the activities from ingest of content all the way to the output of search results are also considered components of the enterprise search system.

The following briefly outlines these components which must be carried out diligently to properly conform to the needs of the enterprise information search and retrieval requirements:

- **Enterprise Search Realm** – includes file systems, repositories, networks and systems that must be searched to gain the collected information
- **Enterprise Search Product** – includes the hardware, software or combination thereof that holds and processes all of the enterprise search activities and functions
- **Enterprise Content Ingestion** – Content Input
- **Content Processing and Analysis** – Meta Tagging
- **Content Indexing** – Building master content catalog
- **Query Management**

2.1.1 Enterprise Search Realm

The following constitutes an enterprise search realm and includes not only the content, but also the functions and the search product being implemented to provide search results:

- Enterprise Content including file systems, document repositories/directory structure, databases, data warehouse, backup systems, and data archives
- Search Infrastructure hardware including search appliance and/or host servers running the search software, network appliances enabling the search and all systems that support the search infrastructure
- Search Software and/or Application that creates the search environment either through the appliance or by running over a host server system
- Search Support Personnel responsible for operating, managing and administering, optimization, and tuning of enterprise search systems thus customizing search environment, look and feel, layout, and results customization
- Search Authority – Individual(s) responsible for approving requests for access to privileged content. The search authority also manages and maintains a record of users' access profiles that may or may not coincide with their overall organizational security and information access profile.

2.1.2 Enterprise Search Product

Any commercially available industry strength tool or solution in software, hardware or a hybrid form can be considered an enterprise search product. These tools or solutions can effectively and efficiently manage and control the various activities and functions that take place during the implementation and operation of an enterprise search system. In addition, these tools also provide interfaces for the search users as well as the enterprise search system administrators and managers. Although considered a part of an enterprise search system, the implementation of any such tool within an enterprise can itself be consider the enterprise search system.

2.1.3 Enterprise Content Ingestion (Input)

The content collection process deals with enabling the enterprise search system to crawl a specific or set of content areas located either on a file system as documents or within particular databases. Once the search system has cataloged the position of this content, it builds its own record of content and relative locations such that upon search it can efficiently access the particular piece of information on demand.

The crawling function that is part of the content ingestion into an enterprise search system is either accomplished via parsing and traversing a file system or utilizing a connector to a database repository. Most commercially available enterprise search solutions internally maintain connector interfaces for common industry strength databases. Most of the modern enterprise search solutions come with built-in with a scheduler function that an organization can customize to create a schedule that allows the enterprise search solution to crawl and gather content while the content usability is at its lowest, most preferably during the early morning hours. This way of collecting and ingesting content into the enterprise search system allows for a smooth operation and enables the search system to serve the most current and up-to-date content.

2.1.4 Processing and Analysis – Meta Tagging

Certain type of content may require additional data descriptors attached so that the search system could first understand its type, and then be able to catalog and connect it to any other content that may be related to it. Images, document scans, PDF images, multimedia documents (audio and/or video) that do not carry any descriptor information, fall into this category.

If an enterprise maintains content repositories with these document types or information, it becomes necessary to process all such content through a system of tagging specific content with descriptor information also known as Meta tags. Commercially available enterprise search solutions may sometimes identify this type of content and auto tag them for later processing or even batch tag such content based on specific pre-established synonyms or references. However, it is desirable to conduct this operation manually such that the most attainable resolution possible could be achieved by selectively assigning highly descriptive and targeted Meta data.

In addition, the processing function also parses each document whenever and wherever possible to extract fundamental unit of text (in case of documents, English language words) also referred to as tokens. These tokens are then added to the master catalog also known as the index to form the based matching unit.

Whenever a search is performed against any tokenized content, the search system matches the search keywords presented by the user with the available tokens in its index to retrieve the locations and thus the actual content in the search results.

2.1.5 Indexing

As previously mentioned, the process of amassing a master system of records such that any given document could be searched and retrieved upon demand is generally known as indexing. The final result of the indexing activity is a massive index store that serves as the pointer catalog of locations for all of the enterprise content such that once accessed and demanded, the index points to the actual data, document or information which is quickly and efficiently retrieved without a loss of time or performance compromise. It is this facility that makes an enterprise search such a pivotal and integral function for any enterprise. The index may contain the dictionary of all unique words in the corpus as well as information about ranking and term frequency.

2.1.6 Query Management – Issuance, Parsing, Matching, Results Building

By utilizing a web page, the user issues a query to the system. The query consists of any terms the user enters as well as navigational actions such as faceting and paging information. The processed query is then compared to the stored index, and the search system returns results (or "hits") referencing source documents that match. Some systems are able to present the document as it was indexed besides relevance and other meta data information.

It must be noted that all of the above components of an enterprise search system although listed separately commonly do exist within commercially available enterprise search solutions. These solutions make it easier to manage and perform all of these functions through a user friendly and easy to navigate interfaces thus reducing the search management efforts and introducing efficiency in provisioning enterprise search applications.

2.2 Content Relevance in Enterprise Search

Content relevance plays an important role in directing the users' attention towards the most relevant and notably the most important results returned from a particular search. Where relevancy plays a key role in the World Wide Web search engines in leveraging propagated link structures across the web, enterprise search solutions are devoid of such collective and associative properties within the enterprise content that may or may not contain markers for such associations. In this case, the enterprise search solutions are relying on content contribution factors as well as available hyperlinking features within the enterprise to calculate and offer content relevancy. New and improved linking algorithms exploiting hyperlink structures promote documents and attribute authorities which could then be utilized as query-independent relevance factors. The marked difference of enterprise search functions and their rich user interfaces, as well as information clustering and faceting has proven much more useful within the enterprise environment in directing users' attention to specific content.

3 The Study

3.1 TFSG's Approach to Research, Evaluation, Assessment and Recommendations

This study has been divided into the following four areas in conformance with the requirements of the Phase-I deliverables and to provide a sufficient framework for FEC to choose a suitable solution from the selected set of tools evaluated:

- Discovery
- Research
- Evaluation & Assessment
- Recommendation

The following section presents TFSG's findings and observations broken down across these study areas:

3.2 Discovery – Fec.Gov Website

As part of the Phase-I deliverable, TFSG conducted a thorough inventory of the fec.gov website. The searchable content available on this website is primarily of static nature include HTML documents, PDF files and text documents. TFSG performed a crawl of the fec.gov to extract and create a map of the information being maintained within this website.

3.2.1 FEC.GOV Website Content Overview

TFSG took great care in inventorying the FEC website (www.fec.gov) during late night hours to assure this activity did not affect the website functionality. The following diagram is a sample snapshot of most of the content that was discovered as a result in a site map format showing only the top level folders.

(b) (5)

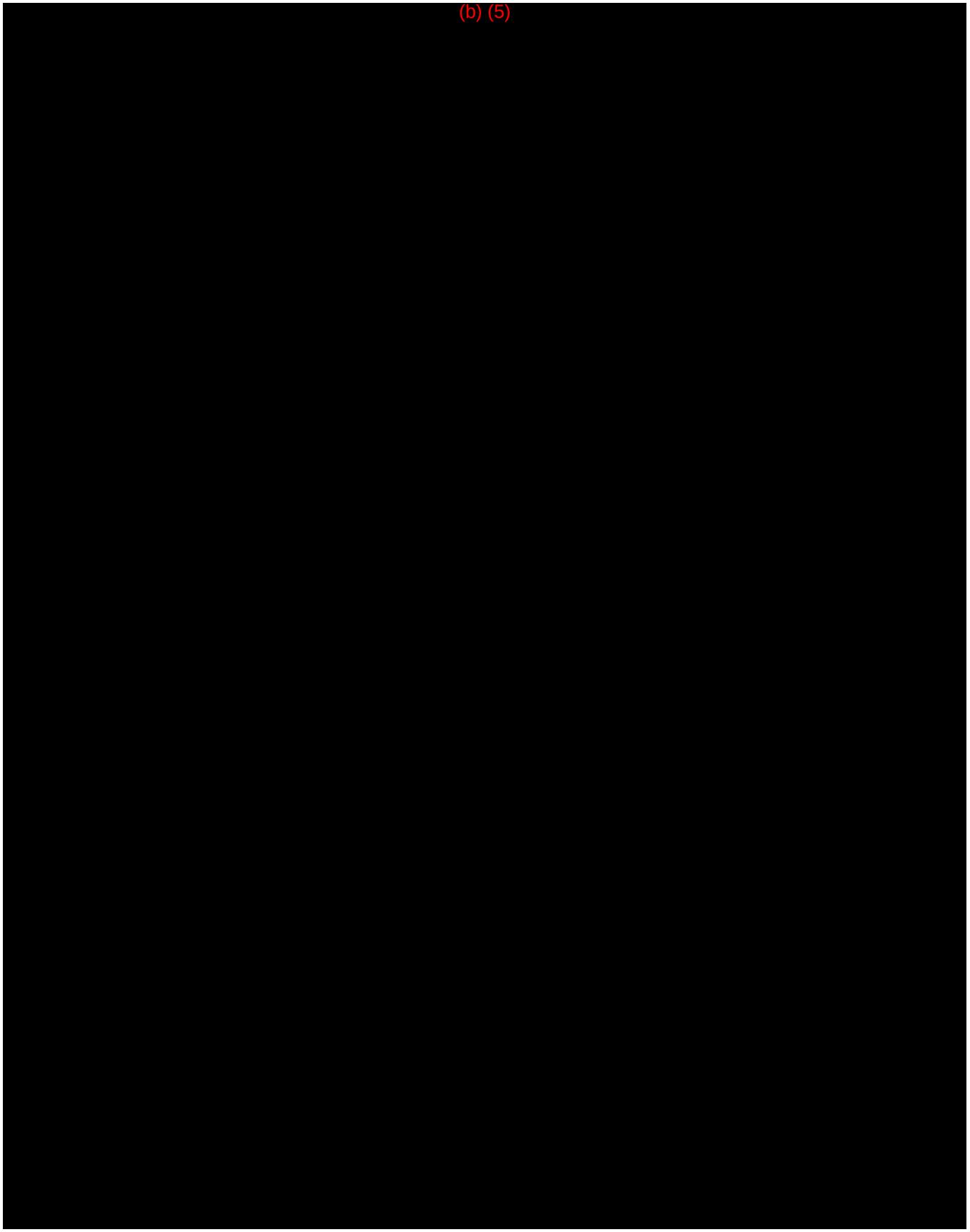

Table 3.1.1: FEC.GOV Website Content Breakdown											
ITEM	CONTAINER NAME	LEVEL	SUB FOLDERS	CONTENT TYPE (Counts)							
12	DISCLOSUREHS	II		2		15	6				
13	DISCLOSUREP	II		3		15	7				
14	EEO	II		2				1			
15	ELECFIL	II		6							
16	EM	II	1	4				2			
17	FECIG	II	1	42				52			
18	FINANCE	II	1	7							
19	GENERAL	II	2	3				3			1
20	IMAGES	II	1								26
21	INFO	II	5	28				4			
22	LAW	II	11	21				26		1	
23	LINKS_FILES	II		1							26
24	MEMBERS	II	5	2				5			
25	MENUS	II									1
26	MUR	II		1							
27	PAGES	II	11	46				36		2	
28	PDF	II	16	2				70			
29	PRESS	II	13	218	3			9			
30	PUBREC	II	20	16				13			
31	RSS	II		1							
32	SCRIPTS	II				1					
33	STYLESHEETS	II					1				
34	SUNSHINE	II	5	3				44			
35	SUPPORT	II		1							
36	VOTREGIS	II		1							
	CUMMULATIVE			**502**	**13**	**31**	**14**	**311**	**5**	**3**	**55**

In addition to the above listed content, certain fec.gov website content could not be crawled due to permission issues where the content is only viewable through an established link on a specific page. For example, a crawl on the following folder returned zero documents, whereas close to 4,500 PDFs in the form of Matters under Review (MUR) documents reside there:

http://www.fec.gov/disclosure_data/MUR

Also, only few sub folders belonging to the years 2004, 2006 and 2008 were discovered while crawling the folder:

http://www.fec.gov/audits

Whereas, through the prepared link on page (http://www.fec.gov/auditsearch/auditsearch.do) many more audit year reports (939 mixed HTML and PDF documents) were available.

> *NOTE: The FEC.GOV website was inventoried using public access and as such the results attained do not reflect the exact number of documents that are currently residing within the fec.gov structure due to limited direct access privileges to all FEC content for the general*

public. However, the inventoried sample does highlight the fact that currently only HTML and PDF documents are searchable.

In addition to the above listed content, it was discovered that a few links on the www.fec.gov website actually pointed to outside domains, in one case referencing a Code of Federal Regulations (CFR) volume with a link to the Government Publishing Office (GPO) website (www.gpo.gov), however, the Enforcement Query System (EQS) and Advisory Opinions (AO) requests pointed to:

- (b) (5)
- (b) (5)

While inventorying (b) (5) link, 1 sub folder containing (b) (5) documents was found, whereas (b) (5) link returned 1 sub folder containing (b) (5) documents.

Finally, a search link inventoried (search.fec.gov) pointed to an html page containing links that presented:

- A html page listing links for static content again listed in the form of links
- Reference to the Enforcement Query System page which again pointed to the outside domain listed above.

NOTE: (b) (5)

3.3 Research on Commercial Enterprise Search Solutions

In the past and as recently as few years ago, TFSG has been extensively involved in architecting, designing, developing, deploying and leveraging full text searches, proximity searches, faceted and federated search mechanism to help its clients and customer base use search in myriad of ways within their missions and enterprise.

It is with this forethought we decided to bring to bear the expertise we have gathered in this field and began investigating, researching, and selecting the industry leaders in the enterprise search arena. Although TFSG has implemented solutions based on search leaders of the time including ASKSAM Free Form Database Search, Dt-Search Engine, and Zy-Index to name a few, TFSG decided to conduct an impartial and unbiased comparison of the industry leader in the enterprise search space. To this end, the following approach was adopted to initiate the research.

- Development of tool selection criteria (further refined as the study progressed)
- Investigation and review of industry journals and trade articles
- Broad search conducted online using a number of common web search engines
- Targeted research using the focus groups including Gartner, Forrester Research Reports, etc.

3.3.1 Initial Tool Selection Criteria

TFSG developed and maintained a simple, yet focused selection criteria for the initial selection of the industry standard tools being used in the enterprise search space. The selection criteria dictated that:

- All selected tools must be as maintenance free as possible to accommodate one of FEC's primary constraints in this project – resource limitation
- Ease-of-Use in terms of user as well as administration and management perspectives
- Proven and demonstrated track record, market penetration and industry references

In developing the research, TFSG maintained an hour-glass approach, where the initial set of enterprise search tools selected was kept broad including as many products as possible that met the initial selection criteria. This resulted in a wide selection, however, with shallow information for each candidate within the initial selection group.

3.3.1.1 *The Initial Set of Enterprise Search Tools Selected for the Study*

The initial set of tools selected and our observations gleaned from available information and data were captured in the following grid:

Table 3.2.1.1-1 Initial Enterprise Search Tools Comparison

(b) (5)

Up to this point, all selection and research was conducted using a passive mode of gleaning from the available material in print, online, or through references. Going forward, TFSG adopted an active stance of approaching, calling, and/or emailing the vendors while at the same time conducting a targeted search for online articles, reports, or white papers written on any of the candidates on selection. The hour-glass approach enabled TFSG to

follow through with further research and investigation into the initial selection set causing additional candidates to be removed from further consideration using the following trimming criteria:

- Lack of response from vendors initially approached via phone calls, emails or request for additional information submitted through vendors' websites
- Low references or unfavorable remarks discovered in various industry articles, online industry reviews stemming from Return-On-Investment (ROI) issues, Total Cost of Ownership (TCO), Ease-of-Use, or high maintenance
- Lack of industry references or brand awareness pointing to immaturity of the tool
- Acquisition or take-over of vendor by another company causing doubts about service, quality and response

The following table depicts the various factors that were considered in trimming the above selection to a more manageable and useful comparison group:

Table 3.2.1.1-2 Initial Enterprise Search Tools Comparison

(b) (4)

3.3.1.2 Narrowing the Tools Selection

The above comparison resulted in a reduced set of finalists to be tested and evaluated against the FEC minimum technical requirements defined earlier:

- Coveo – **Coveo Enterprise Search**
- Google – **Google Search Appliance**
- ISYS – **ISYS Enterprise Search**
- Thunderstone – **Texis Parametric Search**
- Vivisimo – **Velocity 8 Search Engine**

3.4 Evaluation and Assessment Mechanism

Upon arriving at a reduced set of (5) tools, TFSG began conducting evaluations of each product through direct interface with the vendor representatives and wherever possible by using the trial or development editions of the solutions for testing and review purposes.

3.4.1 Conformance to FEC Technical Requirements

The TFSG project team held regular conference calls and discussions with each of the vendor representatives to gain valuable details and for the first time measure each tool against FEC's minimum technical requirements. As a result of the detailed interviews conducted with representatives from the 5 finalist, TFSG was able to develop a requirements conformance profile of each tool as depicted in the following table:

Table 3.2.1.1-3 Conformance to FEC Minimum Technical Requirements

Vendor	Coveo	Google	ISYS	Thunderstone	Vivisimo
		(b) (5)			

Table 3.2.1.1-3 Conformance to FEC Minimum Technical Requirements					
Vendor	Coveo	Google	ISYS	Thunderstone	Vivisimo
(b) (5)					

3.5 Tools Evaluation

In reviewing the technical requirements, it was observed that almost all of the tools in the final group conformed to the requirements as depicted in Table 3.2.1.1-3 making each one eligible for selection as the tool of choice. It was then determined that actual proof-of-concepts using the real FEC data must be developed to review and properly assess each tool for:

- Search performance
- Overall feature set
- Ease-of-use
- Administration and Manageability
- Footprint
- Scope, and
- Cost factors

In order to develop the proof-of-concepts, TFSG designated a test system environment which was loaded with static as well as dynamic data types in the form of folders containing HTML and PDF documents as well as a sample (b) (5) database containing a few tables and views. For testing sake, the test database contained only (b) (5) records.

3.5.1.1 Initial Proof-of-Concepts Developed in Conjunction with the Vendors

TFSG engaged the vendor technical teams extensively and provided remote access to its test system for crawling purposes, as well as instructions on what features the proof-of-concepts should relay. All vendors were given enough time (a period of 1 week) to access and crawl the TFSG test system and then create a search proof-of-concept for initial demonstrations and reviews. Vendors were informed that once the proof-of-concepts were stable, client demonstrations will be scheduled where each vendor will demonstrate their proof-of-concept directly to FEC. During the proof-of-concepts development, the TFSG team worked with all vendor technical teams to assure that:

- All vendors used the same sample data to develop the proof-of-concepts
- All proof-of-concepts exhibited single search box implementation and returned results from static as well as dynamic content
- All proof-of-concepts met the FEC minimum technical requirements
- In addition to meeting the requirements, additional capabilities that each tool could demonstrate
- All proof-of-concepts demonstrated not only the user aspect, but also focused on the administration modules and exhibited the administrative features in detail
- [redacted] (b) (5)

After a week's period, TFSG began reviewing each vendor's proof-of-concept which was demonstrated by the vendors using web conferencing technologies and other online presentation methods. TFSG provided various levels of feedback to each vendor to assure the proof-of-concepts were stable and could be demonstrated to FEC.

3.5.1.2 Schedule for FEC Demonstrations

Once the proof-of-concepts were stabilized, TFSG arranged and published the following demonstration schedule for on-site as well as online vendor demos:

Table 3.2.1.1-3 Schedule of Vendor Demos for FEC

[redacted] (b) (5)

By using the aggressive schedule above, TFSG was able to complete all vendor demonstrations in the first week of December and began compiling the assessment and study report thereafter.

As part of the vendor demonstration, TFSG also provisioned the vendor links for FEC where the presented demos were stored. This allowed FEC personnel to further review and in some case also get a hands-on feel of the admin modules of each specific tool.

3.6 Tools Assessment Criteria and Results

Meanwhile TFSG continued to build and compile the tools assessment. The following assessment criteria was used as a framework to determine the viability and acceptability of any particular feature or tool capability based on the presented demonstrations and the overall tool's review and study:

- Overall Feature Set and Capabilities
- Performance and scope
- Operation
- Support
- Total Cost of Ownership
- Required Enterprise Resource Commitment

The following table depicts TFSG's assessment results for each tool using the above assessment criteria:

Table 3.3.3 Tools Assessment

Vendor	Coveo	Google	ISYS	Thunderstone	Vivisimo
		(b) (5)			

Table 3.3.3 Tools Assessment

Vendor	Coveo	Google	ISYS	Thunderstone	Vivisimo
(b) (5)					

3.7 Final Recommendation

In concluding the final recommendation, TFSG has considered the following forethoughts and realities as they apply to the FEC environment:

- (b) (5)
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

3.7.1 Conclusion

TFSG views itself in a partnering role with FEC in serving FEC's mission and business functions. As such, the TFSG team exerts its effort to investigate, develop, identify and recommend what is best for FEC not only currently, but also in the future, while recognizing FEC's limitations, constraints, acceptability standards, requirements and fiscal pressures. While offering this recommendation, TFSG is fully aware that it is FEC's prerogative to select a tool as a result of this study irrespective of the recommendations contained in it. TFSG also recognizes that it is further tasked to implement and support whichever tool FEC select for this purpose, and TFSG would like to assure its firm commitment to comply with the latter task in time.