E·GOV

Business Gateway
Data Harmonization Approach

September 30, 2005

The Business Gateway Program
Management Office

Table of Contents

I. Executive Summary

II. Previous Harmonization Work
- **The Surface Coal Vertical**
- **The Trucking Vertical**

III. Guiding Principles

IV. Timeline

V. Data Harmonization Approach

VI. Evaluation Criteria
- **Strategic Evaluation Criteria**
- **Technical Evaluation Criteria**

VII. Next Steps

Appendix A: Implementation Methodology

Appendix B: Roles and Responsibilities

I. Executive Summary

The primary purpose of this document is to lay out the evaluation criteria—both strategic and technical—that Business Gateway (BG) will use in choosing the next vertical. Also within this document, Business Gateway defines data harmonization, explains the guiding principles and approach for selecting the next vertical, and provides a timeline for its completion. Finally, in the appendices, BG describes a potential framework for the implementation based upon the Surface Coal harmonization pilot and explains the roles and responsibilities of the various parties involved.

Business Gateway Background

Inherent with any E-Gov Initiative, the Business Gateway has a multi-agency scope. The Governance Board is composed of executive representatives from 22 departments and agencies, and the managing partner is the Small Business Administration (SBA). Consistent with the government-wide Federal Enterprise Architecture (FEA), BG intends to unify work across the agencies to create a more citizen-centered, customer-focused government.

The core objective of Business Gateway is to *make it easier for businesses to do business with the Federal Government*. More specifically, BG (via the portal www.business.gov) is focused on guiding businesses through the maze of government rules and regulations, reducing the burden associated with government compliance by providing simplified access to helpful compliance information/resources.

Consistent with the Small Business Paperwork Relief Act (SBPRA) of 2002, one of the ways that BG is reducing burden on businesses is by employing database technologies that make obsolete the traditional process of submitting dozens of duplicative and overlapping (paper-based) forms to various agencies. This type of burden reduction is called *harmonization*.

"Harmonization" Defined

The core objective of the BG, *making it easier for businesses to do business with the Federal Government,* drives the goal of harmonization.

Harmonization, succinctly defined, is the identification of opportunities to reduce the time it takes for a business to interact with the federal government. The focus of this document is to lay out a strategy for achieving one specific kind of harmonization, the harmonization of *data*.

- Harmonizing data – reducing the amount of data input by businesses and effectively reusing data across federal agencies and state governments.

If the Business Gateway team can implement another vertical that saves time for businesses and streamlines the usage of data across the Federal Government, they will have achieved the objective of *making it easier for business to do business with the Federal Government*.

II. Previous Harmonization Work

While pursuing the next data harmonization opportunity, Business Gateway is keeping in mind the lessons learned from the previous two data harmonization projects the program worked with: the Surface Coal Vertical and the Trucking Vertical.

In describing an industry as a "vertical" industry, BG means that the industry is a highly standardized industry, in which each of the companies must deliver very similar data to the government. Verticals are usually very narrow slices within a sub-sector of an industry, in which the companies have *similar data* that they must report to the government.

A vertical approach to data harmonization can be contrasted with a more "horizontal" approach, in which burden is reduced by finding common data elements across industries. Examples of the types of data that are standard across many different industries are items such as name, address, employer ID number, etc…

Lessons Learned from Previous Data Harmonization Efforts

A. **Single Source Coal Reporting "The Surface Coal Vertical":** BG's first data harmonization pilot is called the Single Source Coal Reporting (SSCR) project, or colloquially, the "Surface Coal Vertical." The project started in October 2002 when SBA became involved (and the OMB 300 Business Case was written) and the pilot solution was released in August 2005. Previous to the SSCR pilot, each surface coal mining company delivered highly redundant data to various federal and state agencies. For example, each coal company would report the tons of coal that they mined in a given month to the state government, to the Department of the Interior Office of Surface Mining, and to the Mine Safety and Health Administration (MSHA) in the Department of Labor. Now that there is a "single source" system for reporting information, a coal company need only enter the tons of coal mined once. Because of the "tightness" of this industry, BG's vertical approach to data harmonization worked extremely well, and as the software is refined during 2006 Surface Coal Mining companies will benefit through greatly reduced reporting times. The ultimate goal is to provide a link to the solution on business.gov (forms catalog).

B. **The Trucking Vertical** (December 2003 – June 2004): The second BG harmonization pilot is called "The Trucking Vertical." Although a solution to this second pilot was never implemented, BG was trying help another highly burdened group of businesspeople—truckers. In order to comply with regulations, trucking companies must fill out myriad forms and then submit them to various state governments and federal agencies. BG surveyed various stake holders in this industry and selected 13 forms to be harmonized. However, after more closely inspecting the forms, it became apparent that the forms did not contain enough common data elements and that therefore the opportunity for harmonization was not as extensive as originally thought.

Because of the Trucking Vertical, the BG project team realized that clearly defining a vertical can be a significant challenge. The team realized that it is crucial to target a very "tight" industry segment in which the businesses in that segment are very similar. Although the Trucking Vertical was focused only on truckers, BG learned that the trucking industry is actually comprised of a number of different types of companies in very different businesses. For example, companies in the trucking industry include small owner/operators of single trucks, large trucking companies, hazardous materials remediation companies, waste haulers, and private fleets (such as Wal-Mart). Because of their divergent roles in the trucking industry, the government requires widely varying information from each.

Additionally, the key players in the Trucking Vertical did not avail themselves of the technologies available to solve the data redundancy problems. They would have profited from including a technology vendor in their meetings so that they would have had a better idea of the automated capabilities of certain technical solutions.

Moving forward

BG intends to apply the framework and lessons learned from the Surface Coal Mining Vertical and the Trucking Vertical methodically to branch out to other highly burdened industries, rather than just creating other verticals ad hoc. In fact, in 2003 BG hired Sytel, Inc. to study and report on other industries (and sub-sectors within those industries) having potential for data harmonization. As several successful verticals are implemented, and a methodology tested, BG hopes that other agencies will be inspired and empowered to start data harmonization efforts of their own. BG's long term objective is to identify common data elements that can be streamlined on a more horizontal basis (across sub-sectors, across industries, compliance issues, etc…). The overall desire is to develop a methodology that will create synergy and momentum within the Federal Government.

In this document BG develops a methodology for selecting an industry sub-sector, in which a data harmonization effort has strong potential for reducing regulatory forms burden.

III. Guiding Principles

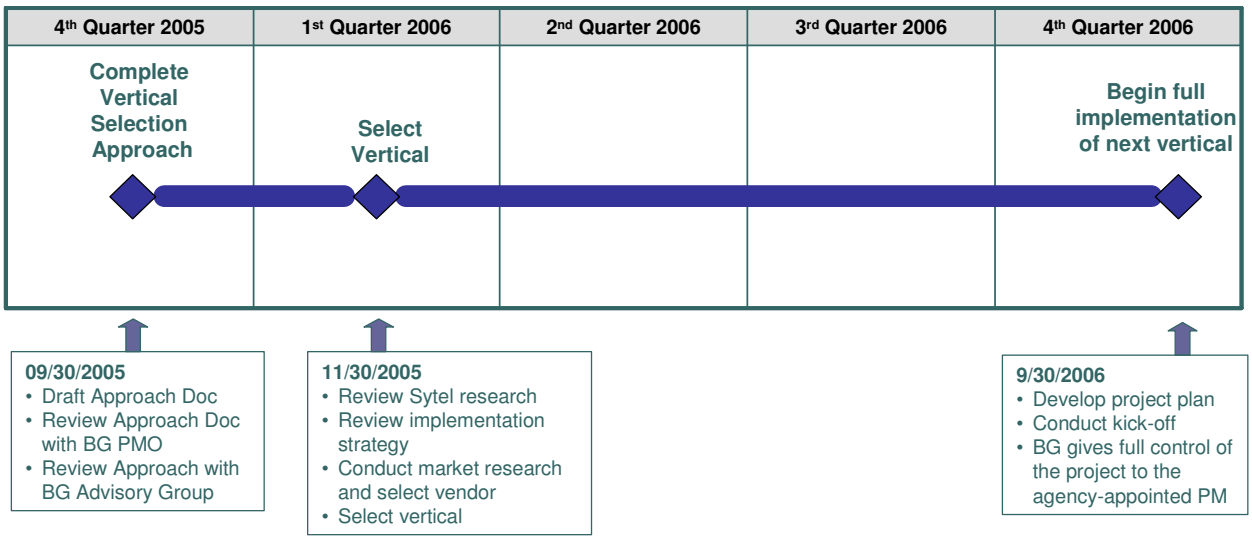
The following principles guide Business Gateway's approach to choosing and implementing the next vertical.

1. **Individual pilot projects should be developed and built to create a framework which acts as an overall harmonization environment**: BG's decision will be informed by and used in conjunction with the Federal Enterprise Architecture. It would be a failure if subsequent pilot projects are non-scalable and do not fit into an overall vision for data harmonization. Therefore, pilot projects should ideally be web-based, open, scalable, and form the beginnings of a standard XML nomenclature.
2. **Build on previous work**: One approach to selecting the next vertical would be to select one of the sub-sectors identified in the "Selection of Industry Sub-sectors for Data

Harmonization Efforts" report put together by Sytel, Inc. on December 11, 2003. In this report Sytel has identified three highly burdened industries (Chemical, Food, and Health), and has identified 11 sub-sectors total within these three industries. Although there are other burdened industries than the ones identified, it might be beneficial for BG to build upon Sytel's research than start researching from scratch.

3. **Maximize the benefit across the Federal Government**: Because the Business Gateway project is funded by 22 federal agencies, the team aims to distribute the benefits among as many different agencies as possible. Further, BG is searching for a vertical that will have a high return on investment.
4. **Use a "vertical" approach**: Business Gateway deems it better to focus on vertical efforts, i.e. specific sub-sectors with lots of data overlap within one particular industry, rather than more horizontal efforts, i.e. redundant data that various companies submit across many industries. Vertical efforts are potentially simpler to organize, have clear champions (owners)—both within federal agencies and within particular industries, and yet still demonstrate that data can be harmonized.
5. **Use a well-defined methodology to drive to the solution, not solely a proprietary solution:** In this report Business Gateway lays out the specific criteria it will use to discern whether a sub-sector is ripe for a harmonization project. These criteria have been created with the best interest of the government and citizens in mind; BG has not *predetermined* a software tool that will meet the project's needs.
6. **Focus on federal forms.** The scope of the project includes federal, state, and local forms, but the primary focus of the project will be on federal forms.

IV. Timeline



V. Data Harmonization Selection Approach

The Business Gateway team has discussed the three approaches listed below as ways of finding an agreeable vertical project. The Business Gateway team will simultaneously pursue all three avenues in order to discover a ripe vertical, which is both heavily burdened and has willing champions to move the project forward.

1. **Call Agencies and talk to their call centers**, asking what kinds of complaints they receive from industry. For example, BG discovered the opportunity to harmonize data in the coal industry (SSCR) because of the large amount of complaints by coal mining companies to DoI. The biggest advantage of this approach is *agency buy-in*: likely the agency would feel the necessity of meeting their constituents' needs and would consequently champion a data harmonization initiative. However, calling scores of bureaus within the 22 agencies would be time consuming and would probably require weeks (or months) of a dedicated worker's time.

2. **Leverage the "Selection of Industry Sub-Sectors for Data Harmonization Efforts" document**, prepared by Sytel, Inc. on December 11, 2003. In this study, Sytel actually interviewed form fillers within industry and asked them whether or not they felt burdened, and which forms were especially duplicative. This paper identifies many sub-sectors ripe for a data harmonization effort and documents which forms are duplicative. Once BG identified a couple of the strongest candidates, they would contact the agency/agencies and industry representatives to see if there was internal impetus for reducing forms burden on industry.

3. **Conduct market research**, discovering what sorts of solutions are offered by different vendors. Vendors potentially already know about heavily burdened industries, which could reduce the amount of research the BG team must perform. Additionally, if one of the vendors finds a good opportunity for a data harmonization project, then potentially we could ask that vendor to do the work.

4. **Leverage the Forms Catalog.** As a part of Business Gateway project, each data element within the Forms Catalog has already been meta tagged, i.e. each field has been labeled with information about it such as "name," "date submitted," etc… Business Gateway plans to leverage this meta tagging system to see what overlaps exist. BG will also analyze the Forms Catalog technical environment, assessing how the Forms Catalog can be used as a framework on which to build vertical opportunities.

5. **Hold a meeting of the key stakeholders** *before deciding on a particular vertical* to ascertain whether or not there is buy-in from industry and the involved federal agencies. Not only should the agencies believe in the project enough to appoint a project manager to run the vertical, but industry leaders should also concur that a data harmonization effort will benefit the companies they represent.

VI. Evaluation Criteria

Based on the different approaches, BG will evaluate sub-sectors based on strategic lessons learned, technical criteria, and the viability of the vendor that can work on that sub-sector. Business Gateway has weighted these criteria by importance (see the below chart) and fully explains them below.

	Individual Weighting (%)	Totals (%)
Strategic Evaluation Criteria		
Data Redundancy	**40**	
Owners/Champions	**25**	
Small Businesses Directly Impacted	**10**	
Value Added to the Government	**5**	80
Technical Evaluation Criteria		
Web-based	**5**	
Open	**5**	
Scalable	**5**	
Model-Driven Architecture (platform independent)	**5**	20

Strategic Evaluation Criteria

1. *Data redundancy*: A measurement of duplicate data takes into account the number of times that data must be submitted—this may be an instance of one form being submitted periodically, where much of the data on that form is the same as the previous time the form was submitted—or this may be an instance of several different forms being simultaneously submitted to a regulatory body, wherein each different form contains common data elements with the others.
2. *Owners/Champions*: For Business Gateway to choose a sub-sector it must have both people inside of the industry that will fight for a data harmonization project ("bottom up support") and government officials that believe in the initiative enough that they will dedicate a project manager to run the project.
3. *Small business especially impacted*: Larger businesses can absorb regulatory costs or pass them onto their customers more easily than smaller companies because in larger companies these costs are dispersed among a larger number of employees. Larger companies have the ability to specialize—delegating much of the forms burden to a few employees who specialize in regulations and are exclusively dedicated to fulfilling this role in the company. SBA is BG's sponsoring agency. Also, larger companies can afford to hire consulting firms to design automated systems to help them fill out forms.
4. *Value added to the government*: The next project should have a large return on investment, a reasonable cost, and should be able to be done in a short time period (about one year).

Technical Evaluation Criteria

1. **Web-based**: Users have access from any geographical location, and can periodically enter in the required information.
2. **Open**: The technology is written in Java, Security Pearl, XML, and other open languages. Also, the tool would need to be able to work with a BEA platform, because this is the platform that the www.business.gov portal will use.
3. **Scalable**: That is, the solution can be easily and quickly integrated with other data storage technologies and systems.
4. **Model-driven architecture**: Information is described and categorized as a certain *kind* of information, and the system does not require the information to have a standard format. The advantage of this capability is that the standardization of data elements becomes less of an issue. Thus, a model-driven integration of data exhibits platform independence.

VII. Next Steps

The ultimate goal for this document was to define a harmonization strategy that can be used as a roadmap that will be easily replicated to implement numerous harmonization opportunities across the government. Within this document BG has explained the evaluation criteria it will use to choose the next vertical. As the Business Gateway team moves forward with selecting the next harmonization opportunity, refinements may be made to implementation methodology appendix, the logical next steps include, but are not limited to, the following:

- Identify the approach to implementing the next vertical (this document)
- Select a vertical
- Conduct sign-off meeting with the key players
- Identify a dedicated Program Manager for the vertical
- Provide funding to selected agency
- Establish an agreed-upon timeline
- Begin developmental work

Appendix A: Implementation Methodology

Although Business Gateway is concerned about prescribing a rigid methodology for the implementation of the next vertical, BG deems it important to outline the high-level steps that the Coal Vertical project team took in their successful implementation. Once the Business Gateway team has chosen a sub-sector on which to focus, they will take certain, specific steps to ensure the successful completion of the project. These steps may or may not resemble those taken by the Coal Vertical team, but those steps are made of record below as a learning tool. When BG drafts a Statement of Work (SOW) document, they will articulate their desire for the software provider to submit a detailed project plan that will detail all necessary phases and tasks for a vertical implementation. During the implementation life-cycle, the Business Gateway team will expect deliverables periodically.

The high-level phases and tasks performed by the Surface Coal Vertical project team are as follows:

Phase 1: Create a Data Reference Model

Tasks:

1.1. Determine data requirements of agencies, including philosophies/approaches to company reporting and use of data, security requirements, and data amendment/change requirements

1.2. Identify areas for standardization, consolidation and harmonization of data elements among agencies, and develop/facilitate standard definitions of data elements.

1.3 Develop logical data model and grouping of data elements (a functional business design)

1.4 Produce complete data reference model

Proposed Deliverables:

A. Data requirements document
B. Consolidated/harmonized data requirements and definitions
C. Logical data model
D. Data reference model

Phase 2: Analyze Design Objectives

Tasks:

2.1. Identify design objectives

2.2. Identify and define design alternatives

2.1.1. Stand Alone System
2.1.2. Shared System
2.1.3. TBD

2.2. Analyze alternatives for cost, feasibility, schedule

2.3. Design selection

Proposed Deliverables:

A. Approved design objectives
B. Alternatives analysis and recommendation

Phase 3: System Analysis, Design, and Development

Tasks:

3.1. Analysis
- 3.1.1. Develop use case scenarios
- 3.1.2. Develop database design
- 3.1.3. Develop business rules

3.2. Design
- 3.2.1. Create prototype
- 3.2.2. Define pages
- 3.2.3. Develop plans for security, development, Q & A, certification, accreditation, training, and implementation

3.3. Development

Proposed Deliverables:

A. Use case scenarios documentation
B. Prototype mock up

Phase 4: Implementation

Tasks:

4.1 Conduct Pilot Test
4.2 Roll-out
4.3 Enhancement

Proposed Deliverables:

A. Pilot Demonstration

In addition to the above mentioned phases and tasks, it will be critical that there is appropriate project oversight that will include such key elements as:

- Development of a Project Charter
- Integrated project plan oversight
- Risk Management
- Requirement Management
- Status Reporting

The implementation of the next vertical will require a clearly developed plan and continuous oversight. However, it is equally as important to coordinate the process with the technical solution. The software solution that is selected should be able to directly support the major objectives associated with data harmonization. The product should be able to meet the following objectives:

- Reduce burden on selected industry by eliminating redundant submission of information
- Reduce burden on industry and agencies by reducing reliance on paper forms
- Produce cross agency ability/methods for information sharing
- Establish data standards to provide consistent definitions for collections and allow more meaningful interpretation of data across agencies

General Operational Requirements

In order to make it easy for both companies within a vertical to submit data and for the participating agencies to receive data, the Business Gateway team requires the following operational capabilities, at a minimum, be available via the software that is selected:

Data collection tools should:

- Be accessible via the Internet
- Have a transparent format
- Have instructions available online
- Allow the user to update information periodically, whenever he/she logs in

Data distribution:

- Should take into account the fact that different agencies have different data interests
- Should take into account the fact that different agencies have varying technical capabilities
- Should happen periodically, according to when the different agencies need the data

By combining a strong process and effective tool, the Business Gateway team firmly believes that the probability of another successful implementation increase

Appendix B: Roles and Responsibilities

BG PMO
Business Gateway's intention is not to run the project, but merely to provide support, up-front financing, and reporting oversight to the project manager. Business Gateway expects the supervisory federal agency to supply a project manager.

Project Manager
Ideally, one person will run the project from inception to completion. The ideal candidate has both *functional* and *technical* expertise: From a functional point of view, this project manager should have close ties on the grass-roots level in the industry and should be able to sell the vision to the federal agency. He or she should be able to identify the key industry leaders and be able to lead them to a common vision for data harmonization.

During implementation, it is important that the acting PM be a technical expert from industry who understands process language and has a background in business analysis. The PM should be familiar with database software and be savvy enough in that respect to work closely with the software application vendor.

Software Application Vendor
Ideally, the software vendor could both provide and integrate the software package that BG selects (as Probaris did on the SSCR project). The software vendor should have a dedicated project manager that works closely with the agency PM.

Participating Agency/Offices/Bureaus
In the Surface Coal Vertical, the Department of the Interior took the lead on the implementation. However, there were several regional offices that provided their process expertise and resource support. It is critical to have similar participation and expertise to any new vertical that is selected.

There will likely be other roles necessary for the implementation that will be determined as the project unfolds.