Center for
Technology in Government

Opening Gateways:

A Practical Guide for Designing Information Access Programs

Opening Gateways:
A Practical Guide for Designing Information Access Programs

Theresa A. Pardo
Sharon S. Dawes
Anthony M. Cresswell

Third Edition
Alan Kowlowitz,ed

April 2012

This material is based upon work supported in part by the National Historical Publications and Records Commission under Grant No. 98027.

Notes on the Third Edition

by Alan Kowlowitz, Editor

This *Guide* was originally published under the title of *Opening Gateways: A Practical Guide for Designing Electronic Records Access Programs* in 2000 and revised in 2002. Since it was issued, technological advances have given us a much broader array of tools and approaches to providing access to information. A critical change is the ubiquitous use of the Internet and Internet-based tools as vehicles for delivering, analyzing, and manipulating information as well as for two-way communication between users and information providers. More recently, social networking tools have also facilitated multi-directional and group communications that is immediate and interactive. Users can now immediately provide feedback on the usefulness of information and suggest ways to improve information access and open data programs. These advances have created a broader and in some ways more sophisticated community of potential users and stakeholders whose expectations of ease of access and immediacy of information have grown exponentially. These changes, combined with a social and political environment that demands public sector entities be more open and transparent in their operations, have put increased pressures on government to provide access to more and better information through readily accessible means such as the Internet.

Technological advances have also provided information access program planners with a greater variety of tools that have expanded available program models. In 2000, it was assumed that most entities seeking to establish an access program would take physical custody of the data provided by participating organizations through formal or informal data flows and provide any value added tools themselves. What was then called "portal technology," which allows linked access to information from multiple sources, was in its infancy. Today the technological aspects of providing access to linked data residing with many different entities is much less challenging. What hasn't changed since 2002 are the planning, organizational, and management considerations necessary to develop a successful access program that meets the needs and expectations of both users and participating organizations. In fact, the advances in technology, increased users' expectations, and increased pressure on public sector entities to make more and more useful and relevant information available has made the type of planning processes facilitated by this Guide more relevant than ever.

Given the changing technological and social environment, the revisions to the *Guide* have focused on updating many of the examples provided and language used as well as including an expanded discussion of program models available due to technological advances. Specifically, the revision includes the following:

- A new introduction that recognizes the pressures on government entities to make information readily available to an expanding public.
- Updated language that expands the relevance of the Guide by broadening it beyond electronic records to government information in general, adding references to new technologies, and downplaying the custodial orientation of the original.
- Added examples to highlight successful initiatives to open access to government information and the diversity of program models available.
- Added references to the CTG's *Government Portfolio Public Value Assessment Tool* (PVAT).
- The *Assessment Tool* has been renamed the *Profile of Characteristics*.
- New explanatory material and part in the *Program Design Tool* that directly addresses custodial and non-custodial program models.
- Additional prompts in the *Preliminary Program Description* section, *Profile of Characteristics*, and *Diagnostic Tool* to get program planners to think about program models.

Despite these changes, the basic material and flow of the *Guide* have remained intact and many sections that have stood the test of time have not been revised. It is hoped that this edition of the *Guide* will be consulted by a new generation of government officials wrestling with the issues of providing useful, relevant, and high quality information to a demanding public and be revisited by officials who have used it in the past.

SPECIAL THANKS

CTG would like to extend special thanks and acknowledgement to Alan Kowlowitz, who took on the task of going through the 2002 edition of *Opening Gateways* and updating it for the new challenges faced today by governments working to increase access to information.

Alan was a member of the original *Expert Panel* who shared in the creation of this *Guide*, and was an indespensible partner sharing his expertise from his positions at both the New York State Archives and the NYS Office for Technology.

Retired from state service, Alan has now brought his 32 years of experience to CTG as a Government Fellow. Alan is applying his expertise and deep knowledge of NYS government and its critical challenges to identifying key themes across past projects, taking the lead on repackaging past reports, and researching related topics to help address emerging issues in digital government.

Theresa Pardo
Director
Center for Technology in Government

ACKNOWLEDGEMENTS*

Many people and organizations shared in the creation of this *Guide*. Our Advisory Committee and Expert Panel members offered both conceptual insight and practical advice.

The dedicated staff of state and local agencies in New York lived through and shared with us the lessons of three electronic access projects that are collectively represented in our hypothetical case example. They are the NYS Geographic Information System Clearinghouse, the Homeless Information Management System, and the Kid's Well-Being Indicator Clearinghouse.

The leaders and managers of existing electronic access programs around the world gave their time and expertise in lengthy interviews about their experiences.

At the New York State Archives, Alan Kowlowitz and Geof Huth were indispensable partners.

Finally, members of the staff of the Center for Technology in Government, Donna Canestraro, Meghan Cook, Linda Keane, Kristine Kelly, Carol Murray, Richard Sloma, Fiona Thompson, and Jing Zhang worked for more than two years on the underlying research. Sally Goodall designed the final Web and printed documents.

ADVISORY COMMITTEE

Tora Bikson, Rand Corporation
Lisa Brzezicki, MasterCard International
Philip B. Eppard, Information Science and Policy, University at Albany
Richard Farrell, University at Albany
Paul Fisk, NYS Division of the Budget
David Goodall, NYS Department of Motor Vehicles
Steven Gorrie, State Farm Insurance Companies
Eleanor Gossen, University at Albany
Peter Hirtle, Cornell Institute for Digital Collections
Robert Hord, Xerox Corporation
Norman J. Jacknis, Westchester County
Lloyd Lininger, School of Public Health, University at Albany
Terrence Maxwell, NYS Forum for Information Resource Management
Charles McClure, School of Information Studies, Syracuse University
George Musa, NYS Psychiatric Institute
Will Pelgrin, NYS Office for Technology
Lisa Radcliffe, HELP USA
Pamela Rinehart, US Patent & Trademark Office
Kenneth Thibodeau, National Archives & Records Administration

EXPERT PANEL

Tora Bikson, Rand Corporation
Phillip Coombs, Washington State Archives
Philip B. Eppard, Information Science and Policy, University at Albany
Robert Horton, Minnesota Historical Society
Geoffrey Huth, New York State Education Department
Alan Kowlowitz, NYS Office for Technology
Ekkehard Mochmann, Central Archive for Empirical Social Research
Andy Potter, National Archives & Records Administration

The individuals acknowledged in this section provided their expertise and advice for the original **Opening Gateways Project. Many, if not most of them, have moved on and no longer occupy the positions they held in 2002.*

Table of Contents

INTRODUCTION

The idea that easy access to government information, facilitated by information technology, can make government more open, accessible, and transparent has been an enduring theme for decades. This idea recently received renewed emphasis through the combination of government reform efforts and the emergence of advanced technology tools for information access. Federal and many state governments see the use of information access and dissemination programs as an essential strategy to increase openness and transparency in government. However, the design of such programs is no simple matter and is fraught with significant management, technology, and policy challenges. Unfortunately, governments' rush to make information available may actually frustrate open government goals of openness and transparency by providing information that is not useable or relevant to the using public and various stakeholders.

It is for this reason that the Center for Technology in Government has chosen to update and reissue *Opening Gateways: A Practical Guide for Designing Information Access Programs.* This *Guide* provides several practical tools to help governments create information access programs that are effective, manageable, and affordable. Each tool includes both a description and an example of the tool in action. The *Guide* is designed to be used from the point of view of access providers. It can be used to develop new access programs or to revise existing ones. The assessment, diagnostic, program design, and cost estimation tools presented address issues that remain and will continue to become more relevant to any access program regardless of technology advances or how it is labeled or marketed.

We use a single hypothetical case example throughout the *Guide* to illustrate the use of the tools presented. The case is a state government initiative to create a Web-based repository of information pertaining to the status of children. We call it the **Children's Health and Well-being Data Repository** (CHWDR) and assign responsibility for it to the fictional State Commission on Human Services (Commission), a small policy and educational agency that provides research and information on social issues and trends to government officials and the public. Because the Commission will be the access provider for the Repository, all of the examples are prepared from its point of view.

The Commission works with federal, state, and nonprofit organizations to carry out this mission. These independent players will have an important role in the CHWDR. This is the Commission's first effort to make data and information more widely available in digital form. Currently, it sponsors or develops research and statistical reports. In the children's area, the main product is an annual compilation of statistics drawn from nearly twenty organizations, which is issued in the form of a fixed format publication, sometimes referred to as "the book" or "data book." The Commission realizes that this publication is not meeting the needs of its stakeholders or fulfilling its mission to provide research and information

CASE EXAMPLE OVERVIEW

State Commission on Human Services (Commission) is a small policy and educational agency that provides research and information on social issues and trends to government officials and the public.

Children's Health and Well-being Data Repository (CHWDR) is the proposed Web-based repository for information pertaining to the status of children.

The book or data book is the Commission's current annual, fixed-format publication that contains a compilation of statistics drawn from nearly twenty organizations.

Inter-County Planning Group is comprised of the twenty organizations who contribute the data and meet infrequently when convened by the Commission.

Data Committee is an informal internal group that coordinates the collection of data and compilation into the current data book.

on social issues and trends to government officials and the public. It is also feeling pressure from child advocacy groups to provide more detailed information in a form that can be easily analyzed. The Commission sees the creation of an expanding online repository of information as a way to address these issues. To create the CHWDR the Commission will have to adopt new policies and practices and learn more about data, meta data, technology, customer service, and inter-organizational management. It presently convenes an Inter-County Planning Group that meets infrequently and an informal Data Committee to coordinate its work. If this first effort is successful, it may be expanded beyond children's issues to encompass other topics such as aging, education, or community development.

GENERAL DESCRIPTION OF THE PROCESS

The process begins with a *Preliminary Program Description* that sets the stage for the analysis that follows. The *Profile of Characteristics* consists of several introductory questions and 15 program dimensions. Each dimension is a continuum of characteristics that range from less to more problematic or resource-intensive. Eight dimensions address information users, suppliers, content, or use. Seven additional dimensions consider aspects of the access program and its organizational context. The P*rofile of Characteristics* provides a rough big picture of the considerations that an access provider must take into account in designing a new program or improving an existing one. This profile of characteristics tells an important story about the challenges an access provider will face.

Within the big picture and the story that underlies it, groups of dimensions tend to interact. The *Diagnostic Tool* helps planners understand how the situations they face on some dimensions affect or are affected by others. These interactions suggest alternative policies, management mechanisms, and technologies that will help them build and operate a successful program.

The *Program Design Tool* takes the foregoing characteristics, interactions, and alternatives into account and helps planners specify the main features of their access program. These include the general program model (custodial, distributed, mixed) to be implemented, key services to be offered, guiding policies, legal requirements, staff skills required, technologies to be employed, work flows, and other elements that define the access program and shape its operation. This basic design can be presented in modest, moderate, and elaborate terms.

The results of the program design effort lead to the identification and estimation of the various costs of implementing and operating the access program. The *Cost Estimation Tool* assists in this phase. It identifies the cost categories associated with the design, its implementation, and ongoing operation. This tool also allows for the program to be specified at modest, moderate, and elaborate levels. A comparison of the costs and expected performance at each level helps planners choose the right level of investment for achieving their goals.

GETTING READY WITH A PRELIMINARY PROGRAM DESCRIPTION

Planners creating a new program or revising an existing one should answer these questions as completely as possible as the first step:

- What is the purpose of the access program?
- What are the main information types to be provided?
- Who are the expected users and how will it be determined what information these various types of users value?
- What uses will likely be made of the information?
- Who are the information suppliers and how do they collect their information?
- What laws, regulations, or court decisions govern the use of the information?

- How long is the information useful?
- How will technology be used?
- What staff skills are needed?
- What does past experience tell us?
- What type of program models are available and what are the implications of these models?

The answers to these questions are the foundation for the program design effort. They describe the initial conditions for your program. As the design proceeds and specific choices are made, the answers may change to reflect new information.

The example on the next page shows how these questions would be answered by the Commission for the Children's Health and Well-being Data Repository (CHWDR).

Organizations contemplating open government initiatives may also want to use the *Open Government Portfolio Public Value Assessment Tool* (PVAT) to determine the public value of those initiatives. Public value is a broad approach to viewing how IT investments can produce results of value to citizens or society as a whole. This concept of value includes more than the usual financial or economic metrics common in ROI analysis. It is a new and expanded way of understanding the results of government IT expenditures.

The PVAT provides government leaders with a structured way to assess an open government initiative's public value and allows them to review expected public value across their portfolio of open government initiatives. The information generated from using this tool can support decisions about the mix of initiatives in a portfolio and how to adjust the mix to enhance the agency's public value. It can also be used to develop powerful arguments to support open government initiatives, particularly those involving open access to information, with resource allocators, stakeholders, and the public at large. Government officials can read about the PVAT and request access to it at www.ctg.albany.edu/publications/online/pvat/.

EXAMPLE

Preliminary Program Description for the Children's Health and Well-being Data Repository

What is the purpose of the access program?

The CHWDR project addresses the need for more rapid and timely access to data about the status of children in our state. The goal of CHWDR is to make existing program data available in a more timely and accessible way to state, local, nonprofit, and individual users. There is a particular desire to expand both the kinds of data available and its usefulness for smaller jurisdictions and communities. The project is also part of the state's open government initiative to increase transparency on the outcomes of intervention programs focused on children.

What are the main information types to be provided?

Statistical information about a variety of topics such as the number of low birth weight babies, children living in poverty, youngsters without medical insurance, high school dropouts, and teen pregnancies that occur in the state every year. Data will also be available on the success rates of programs to prevent teen pregnancy prevention, improve child nutrition, and prevent dropout prevention, as well as a new health insurance program aimed at low income families. Most of the information is available at the county level. Some is available for smaller geographic areas.

Who are the expected users and how will you know what information these various types of users value?

State and local government program managers, program evaluators, nonprofit service providers, academic and nonprofit researchers, members of the public. Focus groups and a Web-based survey will be used to determine what information known user groups value. Social networking tools and techniques such as crowdsourcing will also be used to solicit and analyze input from a broad range of potential users and possibly unidentified user groups.

EXAMPLE *Continued...*

What uses will likely be made of the information?

Policy development, planning, budgeting, evaluation, grants, advocacy, research, and accountability.

Who are the information suppliers and how do they collect their information?

The data in CHWDR comes from a variety of sources including ten state agencies, three federal agencies and half a dozen professional research institutes. They collect the data in the course of operating their own programs. These programs include public education, family support and preservation, delinquency and drug prevention, mental health and chemical dependency treatment, juvenile justice, early childhood education and child care, services for the disabled, employment and training, health maintenance and wellness, and recreation and social development. These organizations periodically arrange this program information in database or spreadsheet format and submit it to the Commission for use in an annual fixed format publication on trends in the status of children.

What laws, regulations, or court decisions govern use of the information?

Information supplied to the Commission is statistical and publicly available without legal restrictions. The underlying sources are subject to a variety of restrictions mostly related to personal privacy.

How long is the information useful?

The data in CHWDR is most useful for the most recent three years; 3-10 years of data is useful for trend analysis.

How will technology be used?

CHWDR will be a publicly-accessible Web-based information repository and hopes to offer data analysis tools to allow searching, combining, comparisons, table generation, and some limited GIS applications for users to manipulate data and generate reports. CHWDR will also use various social media resources to inform potential users about the data and will explore using this technology to provide access to some information.

What staff skills are needed?

Leadership and negotiation, group facilitation, project management, data management, Web design, Web development, Web analytics, and New Media skills.

What does past experience tell us?

It is very difficult to get information from all contributing organizations on the same schedule. After initial submission, data gaps or errors may be corrected at the home organization, but these corrections are not always reported to the Commission. A lack of authority over contributing agencies results in a need to invest heavily in cooperative relationships. User-oriented descriptions and other documentation of the information sources are currently inadequate or nonexistent.

What program models are available and what are their implications for the effort?

Planners have considered a custodial model where they receive and take custody of the information from information producers, a distributed model where they link to information sources maintained by multiple information producers, or something in-between that combines elements of both models. They have weighed the implications for the relationship with information producers and the technologies needed to support the access program. They have tentatively decided to pursue a custodial approach to their information access program. However, this decision will be re-evaluated as the Commission works through the various evaluation tools.

Profile of Charateristics

The *Profile of Characteristics* uses two sets of dimensions. The first set pertains to the users, uses, suppliers, and information content. The second set addresses the access program structure and organizational context. Each dimension can be thought of as a continuum ranging from low cost/low risk to high cost/high risk characteristics.

The users-uses-suppliers-content dimensions are:

- Characteristics of users
- Predictability of uses
- Sensitivity of content
- Frame of reference needed to interpret and use content
- Status of meta data
- Uniformity of information sources
- Degree of integration among information sources
- Usefulness of content over time

The access program dimensions include:

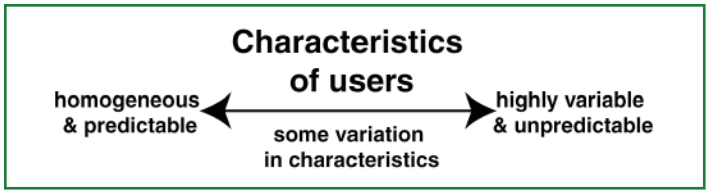
- Structure of relationships with information suppliers
- Structure of relationships with information users
- Involvement of access provider in original data collection
- Extent of data analysis or other manipulation conducted by the access provider
- Nature of data flows
- Suitability of existing technology
- Relationship of the access program to overall organizational mission

In this phase, each dimension is considered independently. The collection of independent ratings produces a *Profile of Characteristics* that must be taken into account in program design. (Later tools consider the relationships among the dimensions.) Each dimension is briefly described on the pages that follow.

DIMENSIONS RELATED TO USERS, USES, SUPPLIERS, AND CONTENT

Characteristics of users

The first dimension deals with the degree to which user characteristics are consistent and predictable. User characteristics include ability to understand the data content, its limitations, and the conditions under which it was

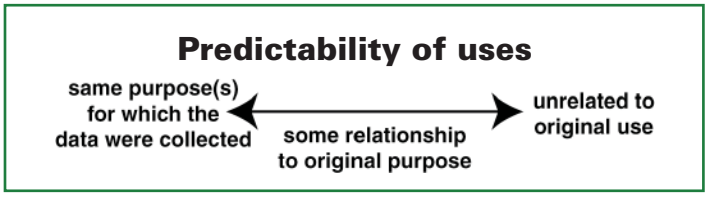


collected; data-handling and analytical skills; technological capabilities and tools; and interests in the data content and what it can be used for. Registered or licensed users who provide information about themselves would fall on the left side of the continuum. Random members of the general public, whose characteristics vary most and are least predictable, serve as the anchor at the opposite extreme. The middle range accounts for users whose characteristics are more or less known or readily predicted. For example, even if they do not register, users of a repository devoted to organic chemistry are likely to be chemists, medical professionals, science teachers, science or medical students, or people with similar interests or knowledge. A repository of popular music might attract users with a much greater variety of characteristics. Program planners can increase what they know about users' characteristics by investing more time and resources in identifying and analyzing users through various traditional (e.g., focus groups, surveys) and newer Web 2.0 techniques (e.g., crowdsourcing).

Predictability of Uses

Information may be used in ways that are very close to or far removed from the uses for which they were created. The degree to which use can be predicted is therefore a key dimension to consider. At one end of the continuum are the uses for which the data were originally collected or

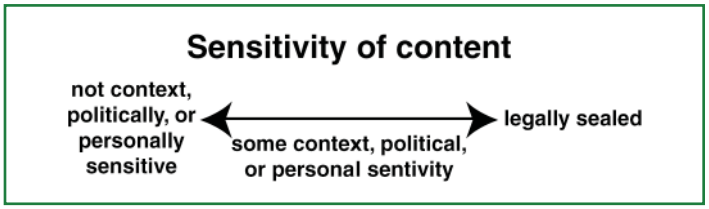
the record was created. At the other end are uses that have no relationship to the original purpose for data collection. Various degrees of relationship to the original purpose lie at points in between. For example, real property records are created to document the history of ownership of land



parcels. They are also used as the basis for local property taxes, a related use. The same records could be used to identify high-income neighborhoods for a marketing campaign–a use that is not at all like the original. As with *Characteristics of users*, *Predictability of uses* could be increased if program planners invest time and resources in identifying potential users and what information user groups value using the type of techniques described above.

Sensitivity of Content

The subject matter or content of data will have characteristics that allow it to be placed along a continuum that has, on the one extreme, factual content that is not controversial or subject to much interpretation and, on the other, content that is so sensitive that laws constrain its use. Between these two extremes lies information of varying sensitivity that must be handled by a range of appropriate policies or management tools. Low sensitivity content might be a daily record of air temperature and wind velocity. Sensitive content includes traditional records such as divorce decrees and adoption records (which are generally sealed) as well as data where unauthorized access or disclosure could severely impact the organization, its critical functions, its employees, its customers, third parties, or citizens. This includes information that can be used to identify a person including Social Security Number, driver's license number or non-driver identification card number and financial account identifier(s).

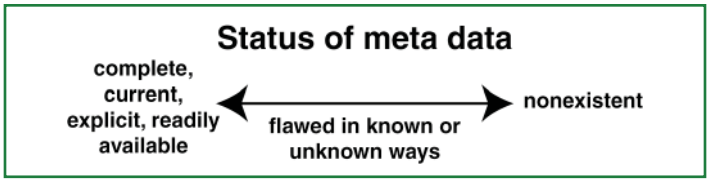


Frame of reference needed to interpret & use content

Information content can vary widely in its need for an expert frame of reference. At one end of the continuum lies information that is readily understandable by a lay person. At the other is information that cannot be used reliably without the knowledge and experience of a subject matter expert. In between is information of increasing nuance or complexity that requires increasing amounts of contextual knowledge in order to be used effectively. A lay person generally has the background knowledge to make good use of a library catalog, news stories, or straightforward numerical data in tables or graphs. It takes more background and training to interpret research reports, complex statistical presentations or inferences, or highly technical information pertaining to various professions or scientific disciplines.

Status of meta data

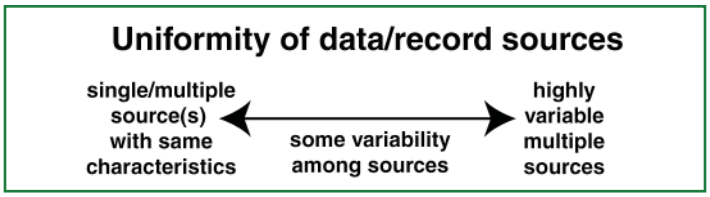
The method and quality of description constitutes the meta data dimension. Meta data, or information about information, can be characterized by its completeness, accuracy, explicitness, currency, and availability to users. At one end of the continuum is meta data which has been made explicit, is current and complete, accurate, and readily



available to potential users of the data. At the other end of the continuum, meta data is nonexistent. At various points along the continuum, meta data is flawed or incomplete, with different levels of understanding about its shortcomings. For example, meta data may describe when the information was collected, how and by whom, but it may not provide definitions of key terms or explain how definitions changed from one year to the next.

Uniformity of data/record sources

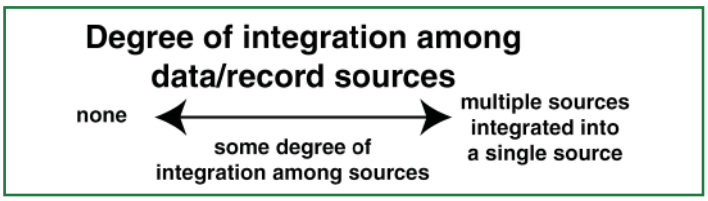
This dimension refers to the uniformity of important aspects of the source data being made available for use. These characteristics include physical format, original purpose, method of data collection, and the meaning of various data elements. At one end of the continuum are data that come from a single source or from multiple sources



that are exactly alike in these characteristics. At the other end lie information that come from multiple sources with disparate characteristics. Income tax withholding instructions are collected by every employer in the country, but in a very uniform way, using a standard form, the W-4. By comparison, case records about social services to individual clients vary widely from one service agency to another.

Degree of integration among data/record sources

If the repository is composed of information from multiple sources, the degree of integration to be achieved among these sources is an important consideration. At one end of this continuum are repositories that maintain multiple sources as separate entities. This program lies at the high



end of the continuum. For example, the National Spatial Data Infrastructure collects or points to a virtual collection of many separate spatial data sets. At the other end are repositories in which information from multiple sources is integrated into a comprehensive single secondary source, such as a data warehouse or a Web-based open government portal. The US federal government and many states are creating open government websites that provide access to a wide array of government produced data from multiple sources. The State of Georgia, for example, created Open Georgia, a gateway for information and key documents about how Georgia spends tax dollars and other revenues to provide services. The portal includes state financial information from various state agencies and is updated annually.

Usefulness of content over time

This dimension recognizes that the content of information sources can vary widely in pertinence and value over time. Some information has only current or short-lived usefulness; other information may be of enduring social, legal, or historical import, and worthy of indefinite preservation. These are the two anchors of this time dimension. At points in between is information whose value to users diminishes over a medium to long term time interval. For instance,

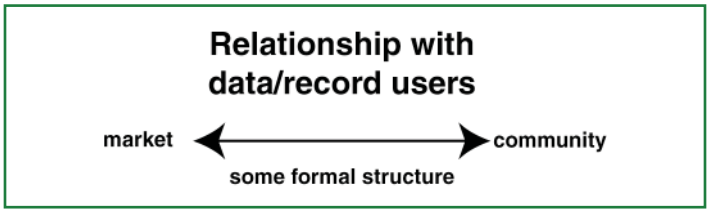
correctional institutions maintain various records essential to accomplish their work, including logs and prisoner case files. A log might document the rounds a prison guard makes on a daily basis, and these have value in the short term, proving that the guards fulfilled their responsibilities or providing a way to estimate when an unwitnessed activity (an escape, a fight, a suicide) took place. Shortly after the production of these records, their value decreases until it disappears. On the other hand, prisoner case files might have a good deal of continuing value. These document each prisoner, including dates of incarceration and release, age, ethnicity, offense,

behavior in prison, medical condition, etc. The prison uses these records to track prisoners, and these records have value to the prison during the time of incarceration and for a short period after incarceration. However, these records also often have permanent value as a way to document prison conditions, ethnicity, and other historical information over time.

DIMENSIONS RELATED TO ORGANIZATIONAL STRUCTURE AND CONTEXT OF THE ACCESS PROGRAM

Relationship with data/record users

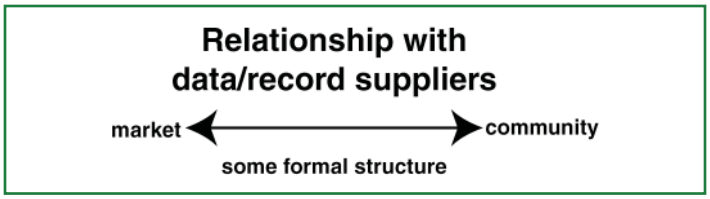
The structure of relationships between the access provider and information users could be described as varying from a simple market kind of mechanism, to a more formal rules-based arrangement, to a more community-like relationship. Market relationships are based on low transaction costs, mutual exchange, short-term involvement, and little or no



need for shared identity or values. Libraries operate under this model and so do most government websites and portals. Bureaucratic, legal, or contractual relationships are based on formal agreements or policies, and are characterized by longer-term involvement and higher costs to establish and maintain. Community relationships are based on long-term familiarity and trust, with shared identity, values, and mutual interests. The establishment and maintenance of these relationships is more costly than the others. More than one kind of relationship can exist with different user groups.

Relationship with data/record suppliers

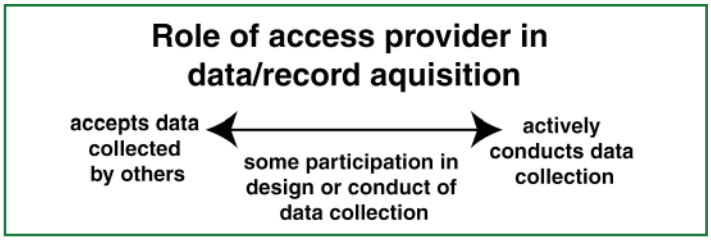
The same relationship structures that apply to users also apply to suppliers of information. They range from simple market transactions between access providers and



information suppliers (such as information brokers who buy state vehicle registration files in order to serve the information needs of vehicle manufacturers) to formal arrangements (such as the information required of businesses by the US Securities and Exchange Commission) to complex community structures (like the CHWDR example we use in this guide). As with users, access providers can have one type of relationship with one set of data suppliers, and a different one with others. The relationship between access provider and information supplier will be one determining factor in selecting a custodial, distributed, or mixed program model. A distributed model may not be practical where formal relationships or community structures do not exist or where entities are not willing or able to change existing relationships.

Role of access provider in data/record acquisition

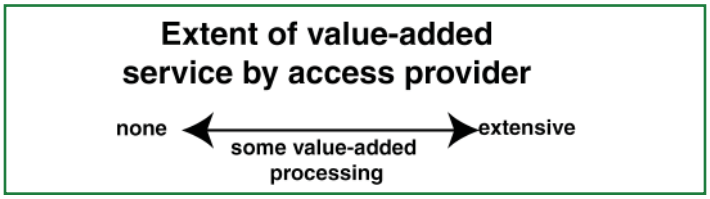
As an access provider, do you play a role in the original data collection? At the low end, the information creation work of the suppliers is independent from the work of the access provider. Examples are data libraries that do no original data collection, but accept or provide links to the information resources that are created or collected by others. The Inter-University Consortium for Political and Social Research (ICPSR), for example, specifies standards for acceptance of data sets, but is otherwise largely unconnected with the work of the data suppliers. The NYS Geographic Information System Cooperative operates in a similar manner accepting or providing links to data created by other entities. Typically,



government archives also fall at this end of the spectrum. At the high end, the access provider plays a significant role in data collection or creation. The Central Archive, an institute of the Cologne Association for Social Research at the University of Cologne, collaborates with the research community to design data collection methods and meta data requirements for new data sets. Some access providers, such as the National Center for Educational Statistics (NCES) and the US Census Bureau, are the main or only data collection agent for their repositories. These programs would also be at the high end of this dimension.

Extent of value-added service by access provider

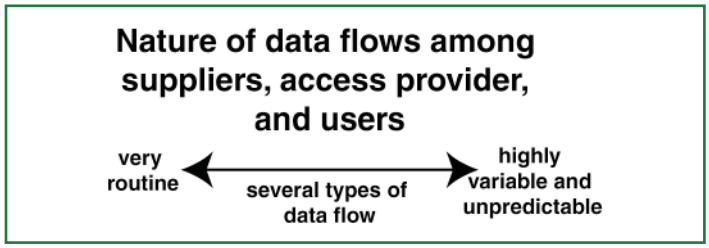
This dimension captures the extent to which the access provider performs work that changes or adds value to the information supplied by others. At the high end of this dimension, the access provider transforms and analyzes the data in substantial ways (e.g., aggregation, constructing indicators, statistical analyses, data mining). As a result, what is available to users is significantly different from or



enhanced beyond the original sources. The changes in the data may involve both format and content as well as ways of presentation. Analytical products such as reports or summaries may be made available to users. The US Census Bureau, for example, provides not only basic census data, but many different subsets, analytical reports, and analysis tools for users. At the low end of the dimension the provider does not transform or analyze the data, providing only access to the original data provided. In the middle would fall programs that conduct modest value-added activities such as providing a search capability or categorizing information according to theme or source. The New York GIS Clearinghouse is an example. It does not process the contributions of information suppliers, but provides several ways to search through the data, including by source, by theme, by coordinates, and so on.

Nature of data flows among suppliers, access provider, and users

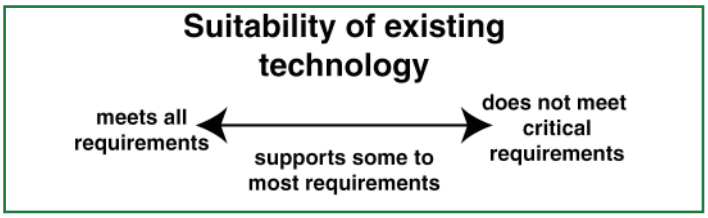
This dimension addresses the way in which information flows from the suppliers or producers to the access program provider to the end users. This information flow can be a virtual flow based on links or a physical flow of data and information from one custodian to another. At the high end, the flow of data into a repository or through links from a provider's portal and the demand for access to that data would be unpredictable and highly variable in timing,



volume, and other characteristics. A general government archive would be an example of a mid-range variability of the dimension. For a highly routine case, both the nature and flow of the data into the repository would be consistent and predictable, as would be the demand for access by users. The Federal Deposit Insurance Corporation (FDIC) is an example of a highly routine program, with regular, rigid requirements for submission of data by banks and routine reporting and access provisions for a large volume of users.

Suitability of existing technology

Electronic access programs necessarily rest on technology foundations. This dimension assesses the degree to which the access provider's existing technology can support the desired access program. On the left side fall programs which already have sufficient infrastructure and technical support to operate a program with the desired features. On the right side are situations in which existing technology does not meet even the most basic critical requirements.



For example, an organization which does not have a Web service capable of supporting applications may only be able to offer downloads of fixed format data unless there are also new resources invested in Web development. The interoperability of technology between data suppliers and access providers needs to be considered here as well as the compatibility of technology available to the expected users. Available technology and technical expertise will also be a factor in determining what program model is most suitable for an access program. Security is an important factor here as well. Security technologies must be explored and understood to ensure that the appropriate technologies have been employed to provide the desired level of security.

Relationship of access program to organizational mission

This characteristic describes the degree to which the access program is central to the core business or mission of the provider organization. If the organization's primary purpose or mission is to provide access to information, it will have a low (or non-problematic) rating on this dimension. Archives and data libraries would be examples of such organizations. Presumably, so would agencies tasked with leading an open government or similar initiative. For the high end of the dimension, the access program would be considered a minor, unimportant, or even unrelated part of the overall organizational mission. This program will have to compete with other, higher priority activities for attention and resources. Certain advocacy organizations would fall in the middle of this dimension. They may collect or acquire information and perform analysis primarily to advance the policy agenda of the organization, and only secondarily to provide data to other users.

HOW TO USE THE PROFILE OF CHARACTERISTICS

For each dimension, place a mark at the point on the continuum that approximates your situation. Occasionally, a single dimension warrants more than one mark. For example, the NYS Geographic Information System Cooperative has two clearly different groups of users. The first group comprises members of a GIS Information Sharing Cooperative who have signed formal agreements to participate in certain ways. This is a well-known, highly predictable group of users. The second group is made up of members of any organization or of the general public who are not members of the co-op. This group is much more variable and less predictable. Since there is a wide difference between these two main groups of users it is important to account for them both. In general, though, you should strive first for a single mark that captures the dominating characteristic for each dimension.

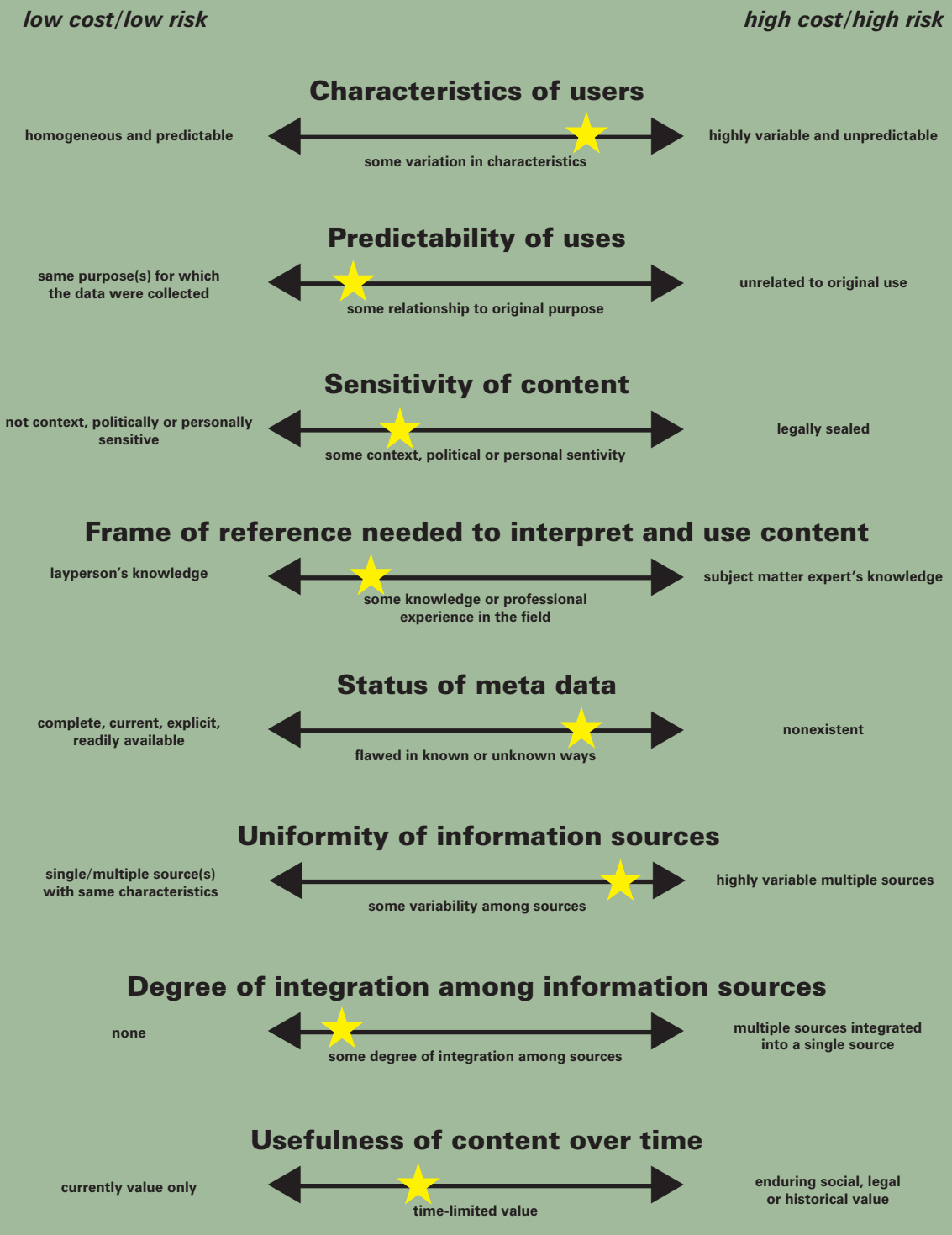
The tool can be used by a single planner and then reviewed by others, or it can be completed by a planning group led by a facilitator. In either case, it is important to explain (and agree on) the reasons for each mark. Since most people tend to underestimate the difficulties they will face in a project, err toward the conservative (i.e., right-hand) side whenever you are in doubt.

A quick glance through the completed profile will give you a rough idea of the challenge that lays before you. If most or all of the marks line up on the left, your job will face fewer difficulties than if the marks are scattered, or worse, align mostly on the problematic, resource-intensive right side.

Our case example is shown on the following pages. It consists of a profile and a narrative description of the situation for each dimension.

Profile of Characteristics applied to the Children's Health and Well-being Data Repository

EXAMPLE



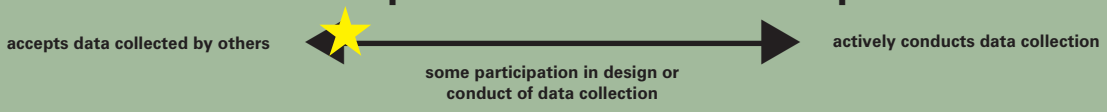
EXAMPLE *Continued...*

low cost/low risk *high cost/high risk*

Relationship with information users

Relationship with information suppliers

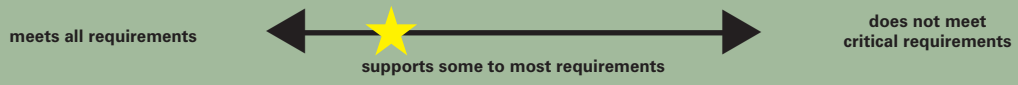
Role of access provider in information acquisition



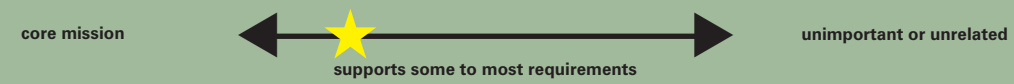
Extent of value-added service by access provider

Nature of data flows among suppliers, access provider and users

Suitability of existing technology



Relationship of access program to organizational mission



Narrative assessment of current CHWDR situation

EXAMPLE

Characteristics of users

CHWDR is intended to support the information needs of the segment of the health and human services community that focuses on programs and services for children. As such, the users are likely to be at least somewhat knowledgeable about the issues and programs relevant to that population. They will vary widely in their analytical and technological capabilities since the users will range from experts in large government agencies, universities, and philanthropic institutions to part-time or volunteer workers in small local governments or nonprofit service organizations. Some general public use is expected. Although this is not likely to be a sizable group it may represent very motivated stakeholders interested in child welfare and the effectiveness of programs aimed at children. The CHWDR will need to analyze and understand these various potential users and determine what they value in and want from the information that will be provided.

Predictability of uses

Interests in the use of CHWDR data will vary. Some will be interested in broad program or policy evaluation, others in winning grants for their localities, still others in comparing their communities' data to others. Some will use the information to inform advocacy efforts to criticize, change, or enhance programs.

Sensitivity of content

The data to be incorporated in CHWDR is largely aggregate statistical information that does not reveal the identity of individuals or report small numbers that can be inferred to apply to a particular person or place. However, it will provide users the ability to integrate data from different sources. Consequently, user-generated analysis might reveal something that is politically or personally sensitive even though the individual source files do not. The fact that data is organized at a county level may raise some political sensitivity when jurisdictions begin to compare themselves to one another. Future plans for CHWDR include providing data at a community or neighborhood level. This level of detail may raise personal sensitivity in small communities where neighborhood and individual identification might be possible.

Frame of reference needed to interpret & use content

A lay person interested in the topics contained in CHWDR would be able to make effective use of most of the information provided. However, there is a fair amount of specialized terminology with which a lay user may be unfamiliar. In addition, users need to be familiar with simple statistics and be able to interpret tabular and graphical presentations of data.

Status of meta data

Traditionally, meta data has consisted of footnotes to tables provided by the various contributing organizations. The importance of meta data is becoming well-recognized and more information on the data collection methods, data descriptions, and caveats to prevent misinterpretation or miscorrelation is being developed. This includes such items as population source, agency source, data source, date compiled, and any specific information that might assist in the analysis of the data

Uniformity of data/record sources

CHWDR includes 50 data sets provided by 19 organizations. Each data set is created by the organization for its own purposes and then provided to CHWDR for secondary use. The data may have statewide coverage only, or may have county or smaller geographic breakdowns. Time periods may vary from one data set to another. Definitions and data categories may vary for the same data from one time period to another. All data sets are provided in electronic form, usually in database format, although some are in spreadsheet format.

Degree of integration among data/record sources

While the data sets provided by CHWDR will not be combined, the planned data analysis tools will allow users to combine information from different sources in order to complete their own integrated analyses.

EXAMPLE *Continued...*

Usefulness of content over time

The data sets available through CHWDR have current value plus time-limited value for purposes of trend analysis. It is likely that users will be most concerned with the last three years of data for their current needs. Program evaluations are likely to need trend data for up to ten years. CHWDR data duplicates or summarizes information held by the supplier agencies that may have longer term or enduring value. Therefore, the supplier agencies, not CHWDR, are responsible for managing access and long-term preservation.

Structure of relationships with information users

Although the information suppliers for CHWDR will also be data users, the Commission expects that the vast majority of users will be local governments, nonprofit service organizations, advocacy groups, academic researchers, and the general public with which it has no official relationship. However, the Commission is actively involved in communities of practice, which include some of the organizational users.

Structure of relationships with information suppliers

The Commission staff must maintain ongoing relationships with all 19 suppliers in order to acquire and maintain information for CHWDR. The Commission has no formal authority over these organizations so these relationships require constant attention. It needs to offer encouragement and incentives to agencies so that they provide their information in usable form, with good user-oriented descriptions, and in a timely way. The suppliers must also agree to maintain the integrity of the information as changes are made to their source data. A community relationship exists with information suppliers. However, it is a relatively formalized and routinized relationship. Therefore, the changes required to implement a distributed program structure with all 19 suppliers may not be possible in the short-term.

Involvement of access provider in original data collection

The Commission does not participate in, coordinate, or influence the original data collection of any of the information sources provided by CHWDR.

Extent of value-added service by the access provider

The Commission will not process or transform the information contributed by suppliers in any substantial way, but it will need to conduct data quality checks, organize and index the information, and provide user analysis tools. The greatest value-added service, at least initially, will be expanded and readily available meta data.

Nature of data flows

Information flows into the Commission for CHWDR from 19 well-known sources. It is a relatively routine and well-understood process. Competing priorities facing information providers can lead to delays. Outflows from CHWDR will be user driven and episodic, but limited to a few routine Web-based mechanisms such as downloads, report generation, and printing.

Suitability of existing technology

The Commission currently has neither the technology infrastructure nor the technical staff skills to build or operate CHWDR. The Repository is being prototyped by a consultant but will need to be transferred to the Commission when it becomes operational. Financial, human, and technical resources must be found to do this. The 19 information suppliers do not conform to any technical standards as far as system platforms and vary as to levels of technical proficiency. It would be difficult for many of them to participate in a technically complex distributed program model.

Relationship of access program to overall organizational mission

CHWDR fits well with the Commission's overall mission to advise and educate policy makers on important social trends. It carries out this mission chiefly by sponsoring and conducting research and issuing paper reports and statistical summaries. CHWDR is a natural extension of that effort, although it applies to only one of many human service areas of concern.

Diagnostic Tool

The *Profile of Characteristics* asked you to treat each dimension independently. In reality, of course, they interact. The *Diagnostic Tool* takes these interactions into account and helps you identify ways to set priorities, make trade-offs, or create options that deal with them in a realistic way.

When the *Profile of Characteristics* indicates a problematic situation with one dimension, it is often possible to adjust others to compensate. The discussion below shows how different situations can be addressed by adopting policies or practices, setting limits, or establishing certain requirements. These brief explanations of the interdependence among dimensions are not exhaustive, but they do illustrate key relationships and possible actions.

Users The more homogeneous and predictable the user population, the more focused the implementation of the program can be. Issues related to the sensitivity of the data may be more easily addressed when the user population is known and can be asked to agree to behave in certain ways. Your ability to predict the nature of use is also likely to be greater. Meta data can be developed to meet the known user community's needs rather than incurring the cost of developing meta data that is broad and detailed enough to serve an unidentifiable general population.

Uses When data collected for one purpose is used for a different purpose, there is potential for misuse or misunderstanding. The intended use, the nature and skill of the user, and the status of meta data are therefore highly interdependent dimensions. Users must be made aware of the limitations of the data as well as its description. The more removed secondary users and uses are from the original purpose for data collection, the more they will need ready access to complete, accurate, and timely meta data and perhaps some expert advice about data use.

Sensitivity of content Politically, personally, and context sensitive content will have a strong impact on design and implementation. More sensitive content will require more stringent governance and access policies, regular review of their effectiveness, and well-trained staff to handle the data appropriately. Sensitive content will also require the use of technical safeguards that ensure security and prevent improper access.

Needed frame of reference When an expert frame of reference is needed to interpret and use content, planners need to consider the capabilities of expected users and enhance the quality and usability of meta data and user support accordingly. They could also consider repackaging the content to make it more suitable for less expert users or provide some ready-made analysis for the most common uses or questions.

Meta data Meta data is a critically important dimension and the one over which you often have the most control. The quality, completeness, and user-friendliness of meta data can be adjusted in many ways to account and compensate for variations in source data, the needed frame of reference for responsible use, or the unpredictability of users and the uses they have in mind. The more these characteristics tend toward the high or problematic end, the greater the importance of good meta data.

Uniformity and integration of sources These related dimensions have important implications for the design and operation of an access program. Consider the following factors. Content from multiple sources or in multiple formats increases the overhead associated with managing relationships, handling the information, and making it available for use. The larger the number of the sources, the larger the number of relationships that will need to be maintained with suppliers. If a distributed or partially distributed program model is used, these relationships could become quite complex. The greater the variation among the sources, the more work will be needed to describe and maintain them. If the content from different sources will be integrated, the level of effort will rise as uniformity among sources diminishes.

The skill required to integrate multiple disparate data sources into a new information resource can be enormous. Integration demands a fine-tuned understanding of the content and clear expectations about intended users and uses. Planners must also consider the frame of reference required to understand and interpret the integrated information and its sources and design user support services and meta data that are appropriate to the users and the information.

Usefulness over time Meta data is critical to the long term, effective use of information with enduring social, legal, or historical value. High quality and complete meta data that addresses context and frame of reference will help ensure that the content remains understandable to future users. A design for information of enduring value must also emphasize standards and make technology choices to ensure migration and preservation long into the future.

User and supplier relations If relationships are market-like, planners need to pay less attention to administrative activities, which are much more important in formal arrangements that need rules or contracts to guide them. Community relationships demand considerably more staff and leadership attention because they rely on long-term shared activities that build trust for joint efforts. This is especially the case if a distributed program model is used. These more labor-intensive relationships will require larger commitment of resources and must lie closer to the heart of the organization's mission than a program that will operate on a simple transaction basis.

Access provider involvement in data collection If the access provider plays no role in original data collection, it will be important to require information suppliers to include good meta data with their information sources. If the access provider participates in the original data collection strategy or work, the access program can benefit from staff who have a much deeper understanding of the information resources they are making available to users. However, the cost of the program is likely to rise to accommodate this additional role and the complex relationships with suppliers that it implies.

Data transformation and other value-added services Value-added services can compensate for inexperienced users, highly variable data sources, and the need for an expert frame of reference. By providing indicators, normalized data, analytical reports and summaries, and user-oriented tools, instructions, and support services, an access provider makes complex or voluminous data more consumable for more users. These services, however, add greatly to the cost of the access program and demand a broader range of staff skills and technical tools than in a program that does not provide these services.

HOW TO USE THE DIAGNOSTIC TOOL

This tool lists all 15 dimensions from the assessment profile broken into two tables. One table includes the dimensions related to use, user, content and suppliers and the other includes the dimensions related to organizational structure and context.

To use the tool, complete the following steps:

1. Put a check mark next to the dimensions over which you have the least influence. For example, some elements of your program may be required by law. Perhaps a long-standing agreement made at the leadership level of your organization limits your discretion to serve some users but not others. Certain data may be absolutely necessary to your program goals, and so on. Since these cannot be changed, or can be influenced very little, they set the key constraints on your program.

2. Put a check mark next to those dimensions over which you have some influence. These are the factors that can be adjusted to compensate or account for the key constraints.

3. Refer to the assessment profile of your situation to see how these items were characterized. Use that characterization to describe the source and nature of the constraint or flexibility.

4. Diagnose the situation using your descriptions of the constraints or flexibility of the dimensions. Document your diagnosis of the interdependencies and begin to identify options or alternatives and set priorities for program design. The description may provide the foundation and logic for necessary trade-offs among program elements. Refer to the discussion above to find ways to adjust some elements in order to deal with others and consider the additional interactions among dimensions that you may have discovered in the assessment.

In this diagnostic process, look particularly for patterns that can lead to serious problems (e.g. information that needs an expert's frame of reference, but users who are not experts; a requirement to integrate information from many sources, but no meta data about the sources). Here are some helpful things to remember:

- Meta data is often your most powerful compensating factor. Good quality meta data can help you deal well with several other dimensions including inexperienced users, data that comes from many different sources, and data sources that must be integrated.

- Policies can be helpful in defining or restricting access, limiting the kinds of data that will accepted or linked to, or requiring actions on the part of data suppliers to make their information more usable.

- Pay close attention to the place that your program plays in the larger mission of your organization (the more closely the goals and intensity of your program match the overall mission of your organization, the better). A mismatch here means you must fit your program to the likely level of support you will get from your leadership–or you must develop a convincing business case that demonstrates why your program should get more attention.

- Value-added services can sometimes make complex data more useable and understandable for a wide audience of users, but they do cost time and money to produce and maintain.

- Compelling compliance for information providers to provide timely, well documented, and high quality data is a resource intensive process.

The case example that follows illustrates the use of the *Diagnostic Tool*.

THIS PAGE LEFT INTENTIONALLY BLANK

Diagnostic Tool Applied to the Children's Health and Well-being Data Repository (CHWDR)

EXAMPLE

The Diagnostic Tool – Use, User, Content, and Supplier Dimensions				
	Dimension	*Nature of Dimension*		*Source and Nature of Constraint or Flexibility*
		Key Constraint	*Adjustable*	
User, Use, Content, and Supplier Dimensions	Characteristics of users	☒		Political leaders have committed to a repository available for wide public use.
	Predictability of uses	☒		Can't predict all of the potential uses that the public may make of the data.
	Sensitivity of content		☒	There is no expectation or commitment to provide data at the detail level, therefore, no individual data will be available.
	Frame of reference needed to interpret and use content		☒	We can adjust the amount of information provided to the user about the data based on the level necessary to interpret and use the content.
	Status of meta data		☒	Meta data is not consistent in availability and quality.
	Uniformity of data sources	☒		The information is not uniform in content, time dimension, geographic dispersion, or format.
	Degree of integration among data sources		☒	Separate sources are grouped into indicator categories but not integrated.
	Usefulness of content over time		☒	We can ensure that the repository contains information that has value to users and is that those resources remain accessible and usable over time.
Diagnosis of Dimension Interdependencies	The very broad set of characteristics of users makes it unlikely that the uses to which the data will be put can be predicted and prepared for. Therefore, the development of the site requires taking into account this potential broad range of uses and ensuring that users are provided with enough information to assist them in use questions. The main concern here is the lack of adequate meta data available to use as the foundation of the information provided to users. An important interdependence is the lack of uniformity among data and sources. Many agencies provide data to the Repository. Each has its own priorities and procedures for data collection. The resources necessary to work with each to acquire the necessary meta data and to influence future data collection efforts will be considerable. The data is linked conceptually to different indicator areas. The data sets themselves are not physically or virtually integrated. Therefore, we have flexibility to expand and modify the data sets provided under each indicator area. The effort associated with growth, in terms of the technical work required to add additional data sets, is more limited and predictable. However, the issues of meta data still exist for each data set.			
	Priorities	1. Acquiring sufficient meta data, or at least more meta data, to support the development of the necessary supporting materials for information users. 2. Developing an interface and set of functional capabilities that satisfy both casual or intermittent users as well as research-oriented, high-frequency users.		
	Trade-offs	1. Commitment to providing access to large numbers of data sets versus a commitment to providing sufficient information to inform users. Due to the unpredictability of uses and users, more time will need to be invested in developing comprehensive support information to provide users with the necessary frame of reference to use the data. This will be directly affected by the available meta data. 2. Information to provide frame of reference versus additional data manipulation capability. We must decide what our commitment is to providing information to provide a frame of reference relative to each potential type of manipulation technique provided by CHWDR. This will directly affect the amount of time the team can spend developing enhanced data manipulation capability given the level of effort required to ensure sufficient information for frame of reference.		
	New options			
	Other			

The Diagnostic Tool – Organizational Structure and Context Dimensions

	Dimension	Nature of Dimension		Source and Nature of Constraint or Flexibility
		Key Constraint	Adjustable	
User, Use, Content, and Supplier Dimensions	Structure of relationships with data suppliers	☒		We have no formal authority over the information suppliers.
	Structure of relationships with data users	☒		We have a good relationship with many users of the Touchstones' data. These relationships are based in Commission activities such as the *Inter-County Planning Group* and the *Data Committee*. New relationships could be developed with new users who emerge as a result of the increased accessibility of the data.
	Involvement of access provider in original data collection		☒	The Commission receives the data from the agencies who do the original data collection for other agency-specific purposes. The data is shared with the Data Committee for compilation and delivery in the Touchstone's context.
	Extent of value-added service by access provider		☒	There is no formal limit to the services that can be provided.
	Nature of data flows		☒	We have no authority to compel suppliers to act in any way.
	Suitability of existing technology for the envisioned access program	☒		We don't have the necessary technologies at Commission to house and maintain the Repository.
	Relationship of access program to overall mission		☒	Our program goals are consistent with the overall mission of the organization.
	Usefulness of content over time		☒	We can ensure that the repository contains information that has value to users and is that those resources remain accessible and usable over time.

Diagnosis of Dimension Interdependencies

The nature of the relationships with the data suppliers constrains the nature of the data flows. We can't control or compete well with the other priorities of the suppliers so we will continue to have difficulty compiling complete and timely data. We have no formal authority so we must work through the Inter-County Planning Group to encourage agencies to support the Repository (in terms of timely data sharing, improving meta data, and potentially even considering the needs of the Repository users when designing data collection protocols) as a high priority task. Doing a good job at identifying and communicating the benefits that users are deriving from the site will help make the case here. We must invest in a approach to do this.

Giving stakeholders the opportunity to influence the future is critical to the success of the Repository. It is also resource intensive. We will need to develop mechanisms to support this kind of consensual decision making and testing of stakeholder satisfaction. We have an existing network of relationships with potential users and providers. The Advisory Committee formed to identify user needs early in the project could be restructured to support the consensual decision making desired. Other options include the Data Committee or the Inter-County Planning Group.

A process for learning about user and provider needs and preferences must inform decisions about the extent of value-added services to add to CHWDR. This process can provide the forum for establishing priorities for the Repository and how it is operated and expanded. It can also provide the forum for collecting benefit data and for communicating the value of the site.

CHWDR will demand a substantial amount of new resources, especially for technology. The Commission does not have the hardware or the necessary human resources to house and manage the site. Therefore, significant time and energy must be committed to establishing partnerships with other member agencies. Contracting is also an option, but one that requires resources in terms of dollars to pay for these contracted services.

In reconsidering a program model, it is clear that the Commission and information providers are not organizationally or technically ready to embark on a distributed approach at this time. However, the possibilities of moving in that direction in the future will be left open.

Priorities	1. Marketing the site as a new and critical information resource for those involved in the health and well-being of children and families in the state. 2. Developing a process for consensus-based decision making regarding the operation and expansion of the Repository.
Trade-offs	Time for consensus-based decision making process versus rapid decision making.
New options	Partnering with other agencies to share the cost of the necessary infrastructure. Possibly forming a small agency coalition to identify ways to share the cost of the human resources necessary to provide such a program.
Other	

Program Design Tool

The *Program Design Tool* helps planners to specify the main features of their program design. This tool helps you identify design options that take the foregoing characteristics, interactions, constraints, and flexibility into account. For example, in resource-poor situations or where a number of key constraints exist, you may want to invest slowly and carefully in just the key aspects of your desired program. In environments where privacy and confidentiality are paramount, you may want to begin the program with a modest design and expand it incrementally while building knowledge and confidence in security techniques and technologies.

Program designs can be developed at different levels of aspiration. Making the different levels explicit allows you to compare the costs relative to the benefits at each level. For example, if a moderate level of technical infrastructure will sufficiently meet your security requirements, then it makes no sense to push for an elaborate infrastructure to ensure security. On the other hand, if only an elaborate level of meta data will serve the needs of your users, you must find the resources to pay for it.

Any number of approaches to using this tool may make sense in your environment. A single individual can complete the tool and share it with a larger design team for refinement, or a formal facilitated design session could be used involving all participants at the same time. Regardless of the logistics, the most effective use of the Program Design Tool will result from a process of reviewing and refining the responses.

PROGRAM MODEL

A critical dimension for designing an access program is whether it will follow a custodial, distributed, or mixed model. Programs can range from custodial programs, where the access provider has actual custody of information to highly distributed, programs, where the access provider does not take custody of information but links or provides access to information maintained by information producers or collectors. Programs can also fall anywhere in between these poles. The program model pursued will have implications for the relationship between information producers and access providers, their roles and obligations, and the technologies selected to support the access program. The choice of program model may or may not have implications for users, as making access to information seamless would usually be a goal regardless of the program model. Some of the implications of the potential models are described below.

- **Custodial Model:** The access provider receives and takes custody of the information. It is responsible for maintaining and preserving access to the information during the period it is being made available to users. The information producer literally delivers the information to the access provider based on prearranged methods or agreements. The access provider would likely select technologies appropriate to its role as information custodian such as simple download or complex data warehouse. In taking custody, it would also be taking responsibility for ongoing information management and preservation, as well as access. State archives, data libraries, and often repositories take this custodial (and often centralized) approach. For example, the Inter-university Consortium for Political and Social Research (ICPSR) has taken this approach; some open government sites such as Data.gov and Open Georgia have taken this approach as well.
- **Distributed Model:** The access provider links to information sources maintained by the information producer or collector. The access provider serves as the central access point using portal or similar technology to provide a seamless experience to users. Information providers assume the responsibility of maintaining information and preserving access for the agreed upon period of time. They would likely have to

agree to various standards and levels of services as well as providing value-added tools. Examples of this model include the United Kingdom's open government site Data.gov.uk, which provides links to government information provided by various data producing ministries; as does the Ministry of Environment in British Columbia, which provides access to climate change data through linking to various sources provided by information producing entities.

- **Combined Model:** The information provider takes custody of some data and provides links to others. This model can combine elements of both the centralized and decentralized models to varying degrees. The New York State Geographic Information System (GIS) Clearinghouse has taken this approach. Such mixed models offer a degree of flexibility in terms of relationships with data producers or collectors and the use of technology.

LEVELS OF ASPIRATION

Modest Program features and functionality at a modest level are the minimum investments worth making. A modest set of services would probably include relatively inexpensive features and limited information sources. Modest objectives would be consistent with a somewhat skeptical or inexperienced management team that is willing to "test the water" but not "take the plunge."

Moderate At the moderate level, a plan for information access services might include additional features and a wider range of information sources. Delivery might be targeted at several distinctly different types of customers. Overall, this level would offer mid-range functionality with some economy-of-scale advantages built in by expanding beyond the modest level. The moderate level of aspiration, however, should not be viewed as an excessive commitment of resources to the initiative.

Elaborate Realistically, what is the most you could hope for? The highest level of aspiration could be a set of objectives that offers a wide range of services, and/or technically sophisticated design goals, consistent with the resources potentially available to the program. Of course, one organization's modest level of aspiration might be the elaborate version of service objectives for another agency. In detailing the elaborate level, be imaginative enough to be called an optimist but not a dreamer.

HOW TO USE THE PROGRAM DESIGN TOOL

The *Program Design Tool* has three parts. The first part allows planners to draw together information collected previously and revisit the issue of which program model is most appropriate for their access program. The second specifies the features and functionality of three levels of design complexity. The third identifies the benefits likely to be achieved by implementing each level. (The *Cost Estimation Tool,* which comes next, completes the picture by allowing you to compare these results on the basis of their costs.)

Part 1 – Specifying the program model

This first step requires you to bring together the relevant information collected using the Profile of Characteristics and Diagnostic Tool to answer the question: *Which program model seems best suited to the goals of the program and present capabilities?* To some extent this requires you to conduct an informal capability analysis. The key capabilities have largely been elicited through addressing issues in the Organizational Structure and Context Dimensions used in the Profile of Characteristics and Diagnostic Tool. In other words, this part of the Program Design Tool does not require you to collect new information but to review and reconsider existing information in the context of selecting an appropriate program model for your program. This part of the tool also

requires you to consider dimensions from the point of view of other organizations that may be involved in an access program, particularly information suppliers.

Part 2 – Specifying features and functionality

The second step requires you to be as explicit as possible about the features and functionality of the program you plan to deliver, given the results of Part 1. Building directly on the results of the assessment and diagnostic tools, you will answer 13 questions to produce modest, moderate, and elaborate designs. As with the assessment and diagnostic tools, the questions fall into two categories:

Users, uses, suppliers, and content:

- Who are your customers?
- What will customers be able to do?
- What information sources will be included and what are their characteristics?
- How extensively will the sources be integrated?
- What meta data will be provided?
- What security and confidentiality measures must be implemented?

Program structure and organizational context:

- How will customers get access to these services?
- How will information flow from data providers through the access program to users?
- How will relationships with data providers be managed?
- What involvement will you have in original data collection?
- What value-added services will you provide to users?
- What technologies are needed?
- What activities will be outsourced?

It usually works best to answer all the questions for the modest design first. This gives you a complete baseline picture. You can then expand from that baseline by answering the questions for the moderate and elaborate versions. As with the other tools, one person can do a first draft followed by refinement from the entire team or a facilitated group effort can be used to both create and improve the design.

For each question be as direct and explicit as possible. For example, when answering *Who are your customers?,* try to include information on their important characteristics. Don't stop with something as terse as "local officials" when you can add useful detail: "county, city, and town officials responsible for the operation of service programs to children and families." Use numbers whenever you can to help you make later comparisons across the modest, moderate, and elaborate levels. For example, try to estimate the size of each customer group, the frequency of data updates, or the number of customer service engagements to be handled.

Part 3 – Identifying likely benefits

The appropriateness of a particular program design rests on its ability to achieve desired performance benefits. When you have settled on the best design, you may want to consider the large body of literature and advice on performance measurement. You can make a solid start now, however, with this highly simplified approach, which entails three general benefit categories: cheaper, faster, and better.

Cheaper refers to all the ways that your services may save resources such as time or money. Remember that an initiative may not produce savings immediately but only over the long term, sometimes by avoiding increased or perhaps new costs in the future. *Faster* refers to shortening processes, and response and waiting times. Providing information and services more quickly also can be considered as an increase in efficiency, even though no cost savings may accrue to the organization. *Better* refers to all the other ways in which performance may be improved beyond increasing the efficiencies of cost and speed. These

improvements may be viewed as more "qualitative," though they can also be measured.

You should not limit your thinking about performance improvements to your organization alone. The new services may make processes and outcomes cheaper, faster, or better for users and suppliers as well. Keep all important stakeholders and constituencies in mind. The table at right shows a short list of variables to illustrate cheaper, faster, and better performance, although it is far from an exhaustive set. Don't let this list constrain your own creativity.

To complete *Part 2*, list the cheaper, faster, and better benefits that you can expect to achieve by implementing each level of program design. Again, be direct and explicit. Consider how you would measure these benefits or otherwise know they were achieved. Try not to rely on benefits whose existence cannot be verified in some practical way. This is another place to use whatever numbers you have to show the progression of benefit from modest to moderate to elaborate. Organizations who do this are often surprised that the incremental benefits of moving from one level to the next are much larger or much smaller than they expected.

An example from the CHWDR project illustrates the process of identifying program features and functionality at modest, moderate, and elaborate levels of aspiration. Keep in mind that a final design may include a mixture of modest, moderate, and elaborate aspirations for different questions. For example, a modestly defined set of customers may have access to an elaborate set of services.

The PVAT described in *Getting Ready with a Preliminary Program Description* can supplement the results of the analysis described in this section by providing a public value view of the benefits of an information access program, particularly one designed to open government. Visit the CTG website to find out more about the PVAT and to request access to it: www.ctg.albany.edu/publications/online/pvat/

Illustrative Benefits	
Cheaper	Reduce or avoid time spent on staff-supported information retreival.
	Reduce or avoid staff time spent on customer support.
	Reduce current telephone, mailing, printing, travel, data acquisition, data distribution cost, or other direct costs.
	Reduce or redirect human resource costs through automation of manual tasks related to customer service and data acquisition.
	Generate revenue that offsets costs.
Faster	Reduce response/waiting time for customers due to faster processing of inquiries and requests.
	Streamline internal processes for information acquisition, processing, and quality control.
Better	More information available to users in electronic form.
	More information available to users through self-directed electronic access.
	Increase use of services (more people use services; same people use more services).
	Consolidate services: one-stop shopping, fewer steps in a process.
	Add convenience through central location, more accessible locations.
	Enhnace quality: more useful, relevant or practical information or service.
	Create innovative new serices, new ways of using information.
	Allow more frequent communication (with same people).
	Allow wider communication (to more people).
	Generate larger number of inquiries, requests, processing, transactions from new or existing customers.
	Develop human resources through enhanced professional abilities and improved work skills.
	Achieve additional visibility, positive media coverage, and public relations advantages.

Program Design Tool applied to the Children's Health and Well-being Data Repository

EXAMPLE

Program Model Tool, Part 1			
Implications for:	**Custodial**	**Mixed**	**Distributed**
Structure of relationships with data suppliers	We have no formal authority over information suppliers. Therefore, we can not compel timely delivery of information.	Faces the some issues as with both Custodial and Distributed Model.	We cannot compel cooperation in a distributed model but might be able to build on existing cooperative relationships given time.
Structure of relationships with data users	The good relationships with users is based on the existing information product, which involves us taking custody of information from suppliers.	Unsure how this model will affect our relationship with data users.	Unsure on how this will affect our relationship with data users.
Involvement of access provider in original data collection	We do not collect the data. This puts us at a disadvantage in interpreting and presenting it.	Mixed Model might allow flexibility in interpretation and presentation.	Data providers collect data and may be in a better position to present and interpret it.
Extent of value-added service by access provider	The Commission presently provides some value added services and would like to provide more to user communities.	Faces the some issues as with both Custodial and Distributed Model.	The commitment and ability of data providers to provide value added services is unclear. Some clearly do not have the resources or ability to do so.
Nature of data flows	Despite our lack of authority to compel compliance, providers do deliver the required data to the Commission through a fairly formal, routine process for the hardcopy report. This data flow will support a Repository based on a Custodial Model.	The custodial side of a mixed model could be supported by the present data flow. However, it would take time to develop the distributed aspects of the program.	There is presently no commitment on the part of data suppliers to a link-based virtual data flow. This issue can be raised but it would take a considerable amount of time to reach consensus.
Suitability of existing technology for the envisioned access program	The Commission does not presently have the technology or technical expertise to support a clearing house. New resources would be necessary to establish a custodial program based solely on a Commission-operated or contracted infrastructure.	A Mixed Model would face all the issues encountered by both a Custodial and Distributed Models.	Many of the information providers, especially the non-government entities, are expertise and technical infrastructure poor and could not host publically available data on their websites. An infrastructure of technical and data standards does not exist and would have to be developed to support a Distributed Model.

EXAMPLE

Continued...

Program Model Tool, Part 1 (cont.)			
Implications for:	**Custodial**	**Mixed**	**Distributed**
Relationship of access program to overall mission	Part of the Commission's mission is providing information to the public on programs for children. A custodial model centered at the Commission would be in line with this mission.	A Mixed Model might be practical to address some of the issues raised for the Distributed Model.	Most information providers are focused on providing services or oversight. The provision of useful data to the public is secondary or tertiary to the mission of most entities. Some entities might be hostile to providing such information through a distributed access program. Most would not have the resources to participate in such a program immediately.
Analysis	Each of the potential program models presents challenges for developing the Repository. The model most feasible for the short-term would be a custodial one. The present data flows and relationship with data suppliers would support such a model with very limited modifications or disruption. In addition, the provision of access to data to the public is squarely part of the Commission's mission. Both a Distributed and Mixed Model would take much longer to develop and present more challenges. Both would require a changed relationship between the Commission and the data suppliers or at least some of the suppliers in a Mixed Model. This relationship would take time to develop and it is unclear if all the supplying entities would accept it. The Commission presently does not possess the technical infrastructure and expertise to develop a custodial program. However, the Distributed Model presents similar technical challenges as well as the development of a standard infrastructure. A distributed or more likely mixed approach, given the varying capabilities of the 19 information suppliers, does present some advantages in terms of value in interpreting data, presenting data, and, perhaps value-added services. Therefore, the Commission will continue to work with its suppliers to explore such a model for the future.		

EXAMPLE *Continued...*

Program Design Tool, Part 2 – Features and Functionality			
Features and functionalities	**Modest**	**Moderate**	**Elaborate**
Users, uses, suppliers and content			
Who are your customers?	State agencies, local governments, for profit and non profit service providers, the research community, the public	Same	Same
What will customers be able to do?	CHWR's website will provide a PDF version of the book.	Ad hoc query, links to other sites, display trends in multiple formats – tables and graphs.	Add: more advanced query capability, link current issues and hot topics, clickable county maps, download selected data in standard format.
What information sources will be included and what are their characteristics?	Data and meta data currently included in the book.	Add more available data and meta data not previously included in the book	Add: newly identified relevant data sets and associated meta data.
How extensively will the sources be integrated?	Sources will be integrated in the same way as in the book.	Data will be grouped according to indicator. Query capability will allow for comparison across data sets and across indicator areas.	Add: data sources will be integrated, to the extent possible, at the database or element level allowing more extensive manipulation capability.
What meta data will be provided?	The meta data that is currently provided in book.	Additional meta data will be provided due to no space limitations. Agencies will be encouraged to fill gaps in meta data.	Add: broad, systematic effort to improve meta data for all data sets and meta data development procedures for new data sources.
What security and confidentiality measures must be implemented?	None beyond the basic system securities necessary to ensure sound system management.	Same	Same

EXAMPLE

Continued...

Program Design Tool, Part 2 – Features and Functionality (cont.)			
Features and functionalities	**Modest**	**Moderate**	**Elaborate**
Program structure and organizational context			
How will customers get access to these services?	Over the Web	Same	Same
How will information flow from the data providers through the access program to users?	The Commission loads PDF version of published reports.	Agencies provide physical databases to the Commission to load into local database and forward for inclusion in staging database. Manual review of staging database prior to release.	The Commission develops a Web-based application that links to data on participating agencies' websites. It provides analytical tools for users and agencies and present data in a user friendly form consistent with agreed upon standards.
How will relationships with data providers be managed?	Relationships will stay the same	Need to encourage development of more comprehensive meta data. New roles for users and suppliers in governance.	Relationships may continue to change based on need to encourage the development of different or more extensive data sets.
What involvement will you have in original data collection?	None	Work with selected data suppliers to communicate need for comprehensive meta data.	Work systematically with suppliers about changes in data collection procedures to produce better data and meta data.
What value-added services will you provide to users?	None	Provide related links and resources plus basic user and provider support services.	Identify customized links and resources based on query, clickable maps, and user support for enhanced data manipulation.
What technologies are needed?	Microcomputer, Internet access, Web server, HTML authoring tools	Add: multiuser database, forms technology, email	Add: GIS
What activities will be outsourced?	None	Possibly hosting of site and technical development	Add: technical maintenance and enhancement of the Repository.

EXAMPLE *Continued...*

Program Design Tool, Part Part 3 - Benefits			
	Modest	**Moderate**	**Elaborate**
Cheaper	For the Commission, placing the PDF version of the book on the Repository will be more expensive than just printing the 5000 copies of the book.	For the Commission, costs will increase with every new feature and functionality, new data set, new meta data requirements, etc. Supplier costs will increase with new meta data requirements. User costs will decrease due to better access.	Same for the Commission. Supplier costs will increase due to new meta data requirements but decrease due to automated loading. User costs will decrease due to better access and tools.
Faster		Need to encourage development of more comprehensive meta data. New roles for users and suppliers in governance.	Relationships may continue to change based on need to encourage the development of different or more extensive data sets.
Better	Users will have access to the indicator data faster than if they had to wait for the book to be printed and delivered.	Add: users will be able to conduct analyses faster than if they had to re-enter data from the book or get it from the source. Data analysis tools provided by the Repository will eliminate the need to download data for use in separate tools. Data and data manipulation tools will be online and readily available for use.	Add: Users will have more timely access to the data as suppliers will load data directly to the Repository upon its availability.

Cost Estimation Tool

Given the design and performance assessments that you made with the *Program Design Tool*, the last key question is what level of investment in a new program to recommend. Is the elaborate version the best level of investment or is it too expensive given the benefits it will deliver? What about a moderate or only a modest investment at first? It may be that the cost and performance assessments support taking no action at all. But how would you know?

The *Cost Estimation Tool* assists in this phase. It identifies the cost categories associated with the design, its implementation, and on-going operation. This tool allows for the costs of the program to be specified for modest, moderate, and elaborate designs. A comparison of the costs at each level to the benefits you identified earlier helps you choose the right level of investment for achieving your goals.

People often tend to underestimate the cost of developing and delivering effective information access programs. This may be due to the fact that there are so many different kinds of expenses that it is easy to forget some. Often planners have less than perfect information and so avoid costing out the parts that are not fully understood. We find that it is critical to identify explicitly as many of the costs as possible, even when you cannot be absolutely exact or certain about every amount. To make these estimations as straightforward as possible for the three levels of service you have described, we have constructed a model cost worksheet. The cost worksheet can be a useful tool for planning the evolution of your service. A worksheet should be completed to represent the costs for various levels of aspiration: modest, moderate, and elaborate. In this way, you can assess explicitly what the start-up and on-going costs might be for these three different versions. Sometimes it makes a great deal of sense to undertake substantial one-time investments in aiming for an elaborate level of service objectives from the very beginning. In other situations, enormous first-year costs can be daunting, so more modest investments may be more feasible. The point here is to be able to compare the costs of at least three alternative plans as thoroughly and explicitly as possible.

How to Use the Cost Estimation Tool

Using the *Cost Estimation Tool* well depends on a complete program design worksheet. The elements shown on the *Cost Estimation Tool* represent the kinds of work to be done to create the program you specified with the *Design Tool*. The cost categories therefore do not link to specific design elements, but rather to the kinds of work that will need to be done to create the design. For each of the three levels of aspiration, review the cost worksheet to identify relevant cost categories (i.e., kinds of work) and relevant specific cost elements. Then enter your best estimate of the relevant costs in the cells of the worksheet. The cost categories on the worksheet are:

- Project leadership
- Project management
- Organizational readiness
- Access and tools for staff and other internal users
- Support services
- Access site development and maintenance
- Content development and maintenance
- Distribution
- Host of site infrastructure

In each category, the worksheet allows for “one-time” and “annual” costs. One-time costs are incurred during development and implementation only, while annual costs recur for as long as the service continues to be delivered.

Most categories also partition costs in a second way: costs for human resources and costs for other purposes. You are likely to find that the human resource costs dwarf other costs. In making your estimates, you should account for

all the staff time necessary to plan, launch, and operate the service. For example, if you plan to move to a new distribution mechanism that requires staff training, you should consider two costs in your estimates: the cost to buy, or develop and deliver, the training program and the cost of staff time needed to attend the training classes.

Some of these activities may be done by in-house staff, others may be outsourced. When outsourcing, be sure to include staff costs for identifying potential vendors, selecting vendors, and managing contracts. On the other hand, if you plan to do all activities in-house, but don't have appropriate staff skills, you should include costs for recruiting and hiring staff or for retraining existing staff.

The full *Cost Estimation Tool* that divides these categories into detailed subcategories is located in the Appendix 4. The example displays the summary information for our hypothetical CHWDR project.

COST CATEGORIES

Project Leadership Include in this category the necessary costs of creating the basic program philosophy, policy structure, and governance framework. This could entail the active attention of a senior executive sponsor. You may also want to form an expert governing board or expand the responsibilities of an existing board.

Project Management The costs of completing and executing the complete program design should be estimated in this section. Include the cost of an overall project manager, the staff work needed to flesh out the basic program design, and develop the protocols and procedures for project management (such as roles and responsibilities, regular reports, and problem-solving techniques). This staff will also develop the implementation plan for the new program.

A variety of activities can support this process. Specific activities appropriate for planning and managing programs that are moving from traditional models to electronic models can be found in the CTG publication *Delivering Government Services on the World Wide Web: Recommended Practices for New York State*. More general suggestions may be found in CTG's *Making Smart IT Choices: A Handbook*. Both can be accessed by clicking on "Publications" on the CTG home page, www.ctg.albany.edu. Other relevant information and tools for developing access programs can be found in *Building State Government Digital Preservation Partnerships: A Capability Assessment and Planning Toolkit, Version 1.0*, which deals with the related issue of preserving information. Resources and tools for open government are being developed or are available as part of the *Opening Government* project.

Organizational Readiness This category includes the resources necessary to get the organization to the point where implementation of the program is feasible. It includes training, demonstration projects, and other educational activities designed to help management and staff become more familiar with the new service models, techniques, and technologies they will use. These will likely include process analysis, stakeholder and user analysis, and change management.

Access and Tools for Staff, Suppliers, and Users Your program will probably involve providing access to information, metadata, and value-added tools over the Internet. It is likely that users will need to be able to browse the Web, receive and respond to electronic messages, communicate with one another, and have access to specially developed internal Intranet applications as well as other Web-based resources. The users might be the public at large, agency staff, staff in other agencies and affiliated provider agencies, business partners, or clients.

In order to be successful, it may be necessary to provide hardware, software, Internet access, and training on both general software and on your particular service. For services aimed at the general public, user access is typically not subsidized by the project and general Web competency is all that is required. For users within the agency or those who participate through an affiliates or subscription service, the cost of necessary components should be included

when planning the project. You may have this infrastructure in place, or it might need to be further developed for a successful project. While this equipment might serve other purposes in addition to this specific project (such as general office automation and communication), if it is necessary for the success of the project, its cost must be considered. How much of the cost charged to the project varies depending on circumstances.

Any portion of this category may be outsourced, in which case you need to include the costs of research, procurement, and contract management.

Support Services Users and suppliers of the system will need training and help-desk support to make effective use of the resources. The sophistication and complexity of your service and the variability among users and suppliers will make big differences to your costs in this category. Again, depending on the type of relationship you have with users and suppliers, the support costs may be borne entirely by your organization or shared in some way with other parties.

The training and day-to-day support may be developed and provided in-house or outsourced to a separate organization.

Access Program Development and Maintenance
Developing an electronic access program entails, at a minimum, converting information into a form that can be delivered to users by Web servers. This is an area in which the support tools are rapidly changing, and project plans need to be reassessed frequently to keep up with the current technology.

An application involving two-way communication may add expenses. The use of Web 2.0 tools or older technologies such as instant messaging, electronic mail, or Web-accessible forms may involve specialized. programming. Staff may be needed on an ongoing basis to reply to users via various methods and handle the additional work generated through the Web service. Developing the technical infrastructure may involve security planning, technical evaluation of alternatives, specialized programming, linking of search engines and databases to Web pages, and on-site hosting of the website.

As the *Tool* shows, there are many different kinds of work in this category. In a large organization, these different functions may be carried out by separate staff members; in a smaller organization, one person may perform several of these functions.

Content Development and Maintenance Content development may involve in-house staff. It could be entirely handled by the data suppliers, or outsourced, or be some combination of these methods. Under any of these methods, however, editorial control and quality standards should be the responsibility of the access provider.

This category includes a number of technology-oriented activities to create and manage databases. It also entails a great deal of hands-on work to ensure data quality, consistency, and timelines. This demands that staff develop good working relationships with data suppliers.

If your program design includes value-added services, this is the place to estimate their costs. These may involve data manipulation, integration, analysis, packaging, and so on.

Promotion and Distribution Once your program design is set and your information resources are prepared, you will need to let potential users and additional suppliers know about the service. This could include the preparation of brochures, educational publications, advertising in selected media outlets, presentations, or other activities. You will also need to develop processes for handling requests and for preparing and distributing standard and customized products.

Host of Site Infrastructure To support a Web-based service, a system containing a Web server and space to store the information if a custodial approach is adopted must be available, usually on a 24 hour a day, seven days a week basis to support a website. Operating system and applications software are also needed. Advanced applications may require additional equipment and more

expensive software to support e-mail access, process forms, link to databases, perform searches, or generate customized HTML pages for individual users. Remember that advanced applications also demand advanced (and more costly) skills. These Web hosting activities may be outsourced to a vendor or may be implemented by connecting your organization's Web server to the Internet. While simple informational Web pages, forms, and electronic mail can typically be outsourced, more advanced two-way applications often require development of a custom Web server application and a dedicated host to provide that service.

Again, several roles and responsibilities are defined in the Tool. They are usually considered specialty areas and are generally carried out by professionals with different skills. A very small organization may combine some of these roles or outsource most of them for cost effectiveness.

Cost Estimation Tool applied to the Children's Health and Well-being Data Repository

EXAMPLE

Information Access Programs Cost Estimation Worksheet						
	MODEST		MODERATE		ELABORATE	
	First Year		First Year		First Year	
		Annual		Annual		Annual
Project Leadership	25,000	10,000	50,000	20,000	75,000	35,000
Project Management	30,000	30,000	100,000	100,000	175,000	175,000
Organizational Readiness **Annual costs will reduce over time*	25,000	15,000*	50,000	25,000*	75,000	32,500*
Access & Tools for Staff, Suppliers, and Users	20,000	5,000	25,000	5,000	25,000	5,000
Support Services	10,000	10,000	50,000	50,000	80,000	80,000
Access Program Development and Maintenance	20,000	10,000	100,000	80,000	200,000	150,000
Content Development and Maintenance	20,000	20,000	125,000	50,000	200,000	75,000
Distribution & Promotion	10,000	5,000	20,000	15,000	20,000	15,000
Host of Site-Infrastructure	50,000	40,000	50,000	40,000	75,000	50,000
INFRASTRUCTURE AND OTHER SUBTOTAL	52,500	36,250	142,500	96,250	231,250	154,375
HUMAN RESOURCES SUBTOTAL	157,500	108,750	427,500	288,750	693,750	463,125
GRAND TOTAL	210,000	145,000	570,000	385,000	925,000	617,500

Analysis of Costs and Benefits for the Children's Health and Well-being Data Repository

EXAMPLE

By comparing the costs of the various levels of implementation (shown in the *Cost Estimation Tool*) to their expected benefits (specified in the *Program Design Tool*), planners can begin to see where more or less spending makes sense.

In our CHWDR example, the modest level of implementation costs the least amount, but the figure is still substantial ($210,000 in the first year). What benefits would that amount of investment buy? According to the *Program Design Tool*, not much. The modest program design bears all the costs of a Web infrastructure, but produces only an electronic version of the old printed document. The Commission could accomplish the same thing for almost nothing by getting another agency that already has a relevant website to host the electronic document and by working with search engines and related sites to link to it.

During the design discussions, the modest plan seemed to be "the least worth doing," but after the costs were estimated, that was clearly not the case. The least worth doing is actually the new idea of having someone else host an electronic version of the paper document. This very small step might be a good short term "design" because it would quickly deliver some benefits to users while a Web repository of the underlying data is prepared.

The elaborate program design includes many sophisticated and complex features, including analysis tools for users and an array of value-added services that the Commission will need to develop. These features are expected to generate significant benefits. The cost of this plan is high, though—nearly a million dollars in the first year. Moreover, the plan represents an enormous demand for skills and technologies that do not exist at all at the Commission today. Both the cost and the degree of change argue against this design, at least as the first step in implementation.

The moderate program design appears to make the investment in a Web repository worthwhile. It provides for ad hoc queries of the data sets that underlie the tables in the printed book and gives users the ability to compare data across different sources to answer their own questions. Improved data loading techniques should streamline operations and maintenance. The initial cost is close to half a million dollars, but the annual costs are a more affordable $385,000. It would give the Commission a solid ability to use the Web first for children's issues, and over time for the other human service topics that fall within its mission. This seems to be the best choice to use as a foundation for a detailed system design and budget. During this detailing process, the Commission is likely to uncover some additional costs, but it may also find ways to cut or get better value out of its planned investment. The *Program Design* and *Cost Estimation Tools* are a start, not a substitute, for these essential next steps.

Conclusion

Access programs have the potential to ensure that the digital information maintained by government agencies and other organizations will be available and useable for the widest variety of contemporary and future uses. If designed well, these programs allow users to readily locate, use, analyze, and compare relevant data to answer questions that are important to them. While such programs can make life simpler for information users, they are not easy or simple to design. If your organization is about to become an access provider or revise an existing access program, these guidelines should help you deal with the complex work ahead:

- Start with a candid description of your planned program and its context.
- If necessary and/or appropriate, assess the public value of your proposed initiative(s) using the PVAT or similar approach.
- Assess users, uses, content, supplier, structure, and context dimensions.
- Diagnose the interactions among these dimensions and the options they offer for action.
- Identify whether a custodial, distributed, or combined model makes sense for your access program.
- Design a program at several levels of aspiration.
- Estimate the costs of these alternative designs.
- Analyze the relative costs and benefits of different design configurations.
- Confirm that you have chosen the optimal basic program model – custodial, distributed, or mixed – if not, reassess your situation for a different model that is a better fit with your capabilities and constraints.

Having taken these steps, you will be well-prepared to select the best program design for your situation, communicate it to stakeholders, develop detailed plans and budgets, and begin the difficult but rewarding work of implementation.

APPENDICES

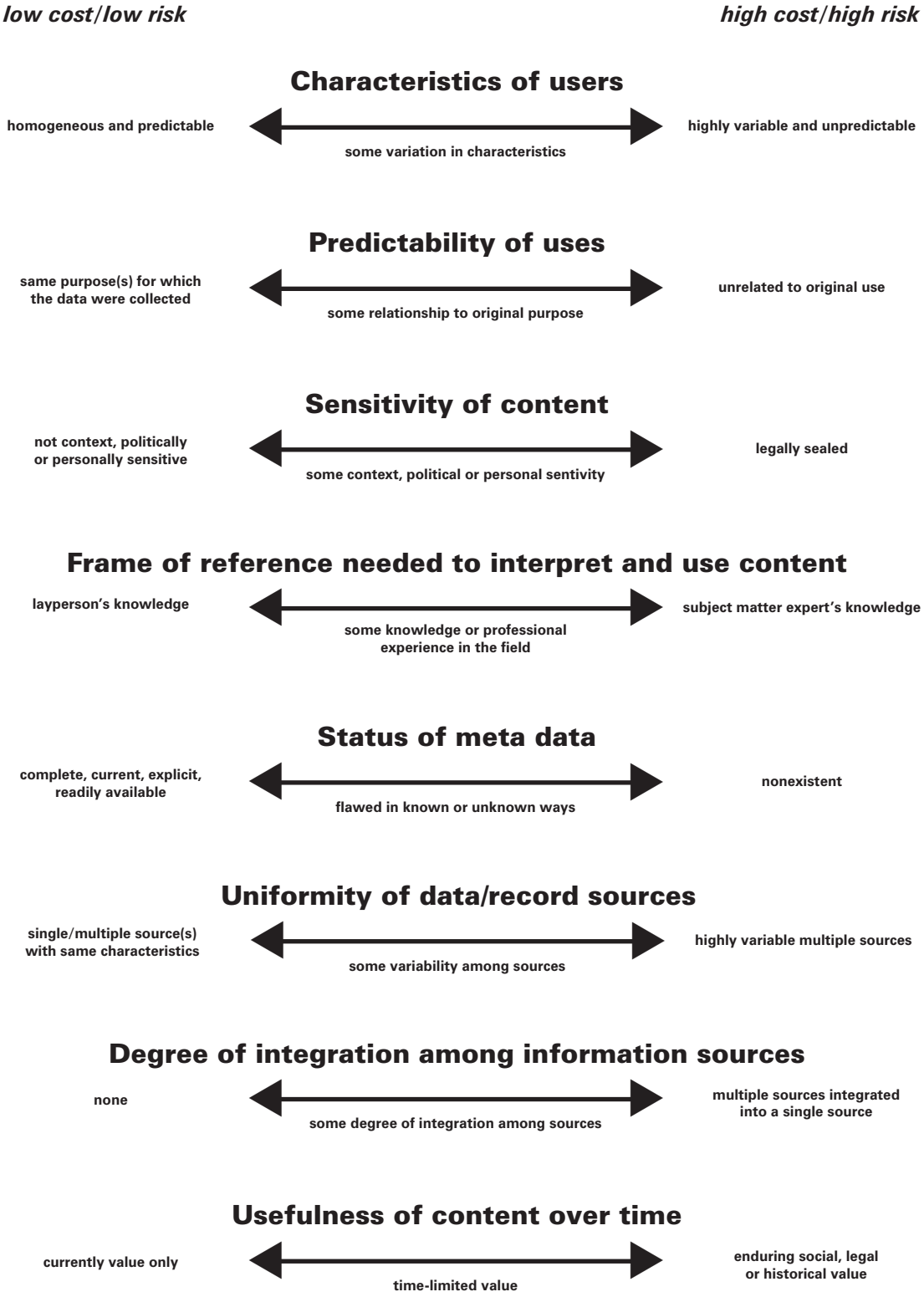
1. PROFILE OF CHARACTERISTICS

2. DIAGNOSTIC TOOL

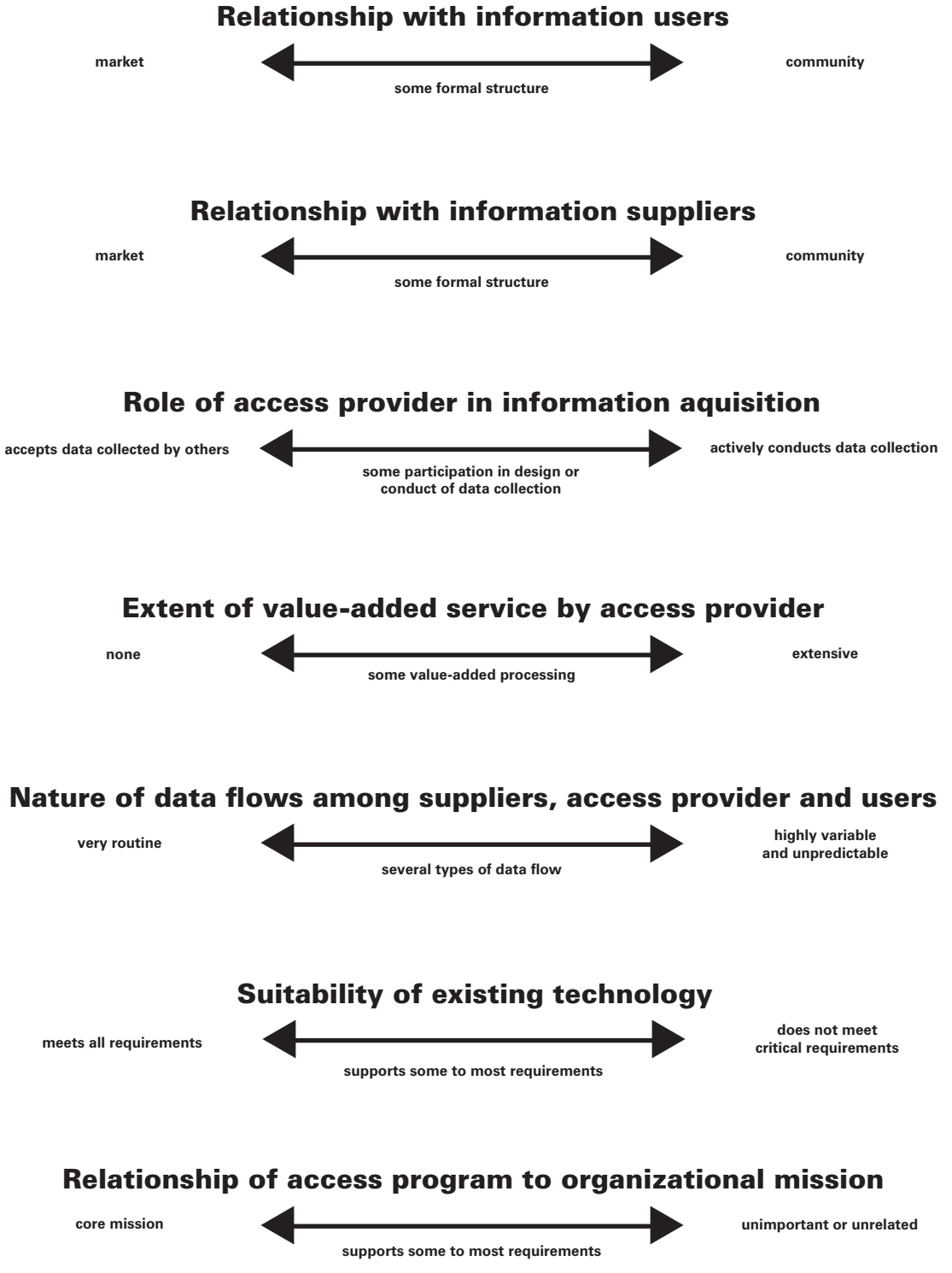
3. PROGRAM DESIGN TOOL

4. COST ESTIMATION TOOL

Appendix 1: Assessment Tool



low cost/low risk *high cost/high risk*



Appendix 2: Diagnostic Tool

The Diagnostic Tool – Use, User, Content, and Supplier Dimensions				
	Dimension	**Nature of Dimension**		**Source and Nature of Constraint or Flexibility**
		Key Constraint	**Adjustable**	
User, Use, Content, and Supplier Dimensions	Characteristics of users			
	Predictability of uses			
	Sensitivity of content			
	Frame of reference needed to interpret and use content			
	Status of meta data			
	Uniformity of data/record sources			
	Degree of integration among data/record sources			
	Usefulness of content over time			
Diagnosis of Dimension Interdependencies				
	Priorities			
	Trade-offs			
	New options			
	Other			

The Diagnostic Tool – Use, User, Content, and Supplier Dimensions				
	Dimension	**Nature of Dimension**		**Source and Nature of Constraint or Flexibility**
		Key Constraint	**Adjustable**	
Organizational structure and content dimensions	Structure of relationships with information suppliers			
	Structure of relationships with information users			
	Involvement of access provider in original information collection			
	Extent of value-added service by access provider			
	Nature of information flows			
	Suitability of existing technology			
	Relationship of access program to overall mission			
	Usefulness of content over time			
Diagnosis of Dimension Interdependencies				
	Priorities			
	Trade-offs			
	New options			
	Other			

Appendix 3: Program Design Tool

Program Model Tool, Part 1			
Implications for:	**Custodial**	**Mixed**	**Distributed**
Structure of relationships with data suppliers			
Structure of relationships with data users			
Involvement of access provider in original data collection			
Extent of value-added service by access provider			
Nature of data flows			
Suitability of existing technology for the envisioned access program			
Relationship of access program to overall mission			
Analysis			

Program Design Tool, Part 2 – Features and Functionality			
Features and functionalities	**Modest**	**Moderate**	**Elaborate**
Users, uses, suppliers and content			
Who are your customers?			
What will customers be able to do?			
What information sources will be included and what are their characteristics?			
How extensively will the sources be integrated?			
What meta data will be provided?			
What security and confidentiality measures must be implemented?			
Program structure and organizational context			
How will customers get access to these services?			
Program structure and organizational context			
How will information flow from the data providers through the access program to users?			
How will relationships with data providers be managed?			
What involvement will you have in original data collection?			
What value-added services will you provide to users?			
What technologies are needed?			
What activities will be outsourced?			

Program Design Tool, Part Part 3 - Benefits			
	Modest	**Moderate**	**Elaborate**
Cheaper			
Faster			
Better			

Appendix 4: Cost Estimation Tool

Information Access Programs Cost Estimation Worksheet						
	MODEST		MODERATE		ELABORATE	
	First Year	Annual	First Year	Annual	First Year	Annual
Project Leadership						
Human Resources						
Develop program policy structure						
Full program design						
Project sponsorship activities						
Governance board						
Other						
Project Management						
Human Resources						
Overall project manager						
Develop program management procedures						
Develop program implementation plan						
Support staff						
Other						
Organizational Readiness						
Training for technology awareness, process analysis, and change management						
Other						
Human Resources						
Staff time in training						
Planning for process and policy changes						
Process and procedural changes - Internal						
Process and procedural changes - External						
Other						
Access & Tools for Staff, Suppliers, and Users						
Hardware for staff						
Software for staff						
Network access and software for users						
Network access and software for suppliers						
Other vendor services						
Other						

Human Resources						
Start-up process for equipment procurement						
Establish and manage vendor and ISP contracts						
Other						
Support Services						
Contracted support services						
Other						
Human Resources						
Support services for suppliers						
Support services for users						
Establish and manage vendor contracts						
Develop and deliver user training						
Develop and deliver supplier training						
Help desk						
Other						
Access Program Development and Maintenance						
Hardware for developers						
Software for developers						
Security infrastructure						
Network and Internet access for developers						
Other vendor services						
Other						
Human Resources						
Start-up process for equipment procurement						
Establish and manage vendor contracts						
Develop and deliver staff training						
Staff time in training						
Interface design and development						
Webmaster						
Editorial review						
Website design and development						
Programming						
Database administration						
Other management support						
Other clerical support						
Other						

Content Development and Maintenance						
Hardware						
Software						
Network access for developers						
Other vendor services						
Other						
Human Resources						
Start-up process for equipment procurement						
Relationship management with source organizations						
Data acquisition process management						
Data integration						
Data manipulation, enhancement, quality control						
Development and delivery of staff training						
Staff time in training						
Content review						
Programming						
Database administration						
Other management support						
Other clerical support						
Other						
Distribution & Promotion						
Brochures, publications, advertising						
Postage and shipping						
Human Resources						
Request processing						
Product preparation						
Public information preparation						
Targeted marketing						
Other						
Host of Site-Infrastructure						
Hardware						
Software						
Network and internet access						
Other vendor services						

Other						
Human Resources						
Front-end research and technical evaluation						
Start-up process for equipment procurement						
Establish and manage vendor and ISP contracts						
Development and delivery of staff training						
Staff time in training						
Security infrastructure						
Network and systems administration						
Web server management						
Operations support						
Clerical support						
Other						
INFRASTRUCTURE AND OTHER SUBTOTAL						
HUMAN RESOURCES SUBTOTAL						
GRAND TOTAL						

About

THE CENTER FOR TECHNOLOGY IN GOVERNMENT

The mission of the Center for Technology in Government (CTG) at the University at Albany/SUNY is to foster public sector innovation, enhance capability, generate public value, and support good governance. We carry out this mission through applied research, knowledge sharing, and collaborative problem solving at the intersection of policy, management, and technology.

The results generated by each CTG project add to a growing knowledge base designed to support the work of both government professionals and academic researchers. Our guides, reports, and tools are freely available on our publications page: www.ctg.albany.edu/publications.

THE AUTHORS

Theresa Pardo

Director

Dr. Theresa A. Pardo is director of the Center for Technology in Government (CTG) and a member of the faculty of Rockefeller College of Public Administration and Policy and the College of Computing and Information at the University at Albany, State University of New York.

Under Dr. Pardo's direction, CTG is developing a public value assessment framework for open government initiatives. This work, funded by the U.S. National Science Foundation (NSF), is being carried out in partnership with numerous state and federal government agencies. Her most recent NSF-funded effort is to develop a data interoperability framework for the North American Free Trade Agreement (NAFTA) region. In addition to funding from the NSF, Dr. Pardo's work has been funded by the U.S. Department of Justice, the U.S. Library of Congress, the U.S. National Archives and Records Administration, the United Nations, SAP, Microsoft Corporation, the Organisation for Economic Co-operation and Development (OECD), and New York State, among others.

Dr. Pardo serves as a member of several national and international boards, including the Steering Committee of the National Gap Analysis on Homeland Security, the Digital Government Society of North America, Government Information Quarterly, the U. S. Government Accountability Office (GAO) Executive Council for Information and Technology Management, and the International Conference on Theory and Practice of Electronic Governance (ICEGOV). She is also a Senior Adviser to the State Information Center, P.R. China and has served as a member of the jury panel for the Sultan Qaboos Award for Excellence in eGovernment in Oman.

Sharon Dawes

Senior Fellow

Sharon is Senior Fellow at the Center for Technology in Government (CTG) and Professor Emerita of Public Administration and Policy and Informatics at the University at Albany, State University of New York. As the founding Director of CTG, from 1993-2007, she led the Center to international prominence in the field of digital government research. Her research interests are cross-boundary information sharing and collaboration, international digital government research, and government information strategy and management. Her current focus is international digital government partnerships including an international community building program funded by the US National Science Foundation (NSF) and partnerhip-building efforts in Asia, Europe and North America. A fellow of the US National Academy of Public Administration, she was elected the first President of the Digital Government Society of North America in 2006. She serves on advisory committees for NSF, the National Archives and Records Administration (NARA), and the United Nations University International Institute for Software Technology (UNU-IIST). Under her leadership, CTG received several prestigious national awards including an Innovations in American Government Award, sponsored by the Ford Foundation.

Anthony Cresswell

Deputy Director

As deputy director, Dr. Cresswell works with a variety of government, corporate and university partners to conduct applied research projects on the policy, management, and technology issues surrounding information use in the public sector.

He began working at CTG as a senior research fellow in 1994 and most recently served as interim director. One of his major contributions has been his focus on return on investment for government information technology and addressing the core issue of determining public value. In addition, Dr. Cresswell's efforts have been directed at problems of interorganizational information sharing, knowledge networks, and IT impacts on practice.

Dr. Cresswell has been at the University at Albany since 1979, with faculty appointments in Educational Administration and Information Science. Prior to arriving at the University at Albany, Dr. Cresswell served on the faculties of Northwestern University and Carnegie-Mellon University, and as a Faculty Advisor in the US Office of Management and Budget. In addition to research and teaching in the US, he has worked in information system and policy analysis projects in Africa, Asia, the Middle East, and Caribbean.

THE EDITOR

Alan Kowlowitz

Government Fellow

Retired from state service, Alan has brought his 32 years of experience with the New York State Archives and the Office for Technology (OFT) to CTG as a Government Fellow. Alan is applying his expertise and deep knowledge of NYS government and its critical challenges to identifying key themes across past projects, taking the lead on repackaging past reports, and researching related topics to help address emerging issues in digital government.

During his tenure at both the State Archives and OFT, Alan was involved with CTG projects in various capacities. While at the State Archives he co-authored and was principle State Archives participant in the Models for Action project. While at OFT he served on the Advisory Group for Gateways project and cooperated with CTG on many other e-Government initiatives.

Alan was on the staff of the State Archives between 1979-1999 where he helped establish and then manage that institution's electronic records program. While at the State Archives, he also assisted OFT in drafting New York State's Electronic Signatures and Records Act (ESRA). Between 2000-2004, Alan served on the OFT team that developed the ESRA regulations and guidelines and established the State's e-Government/e-Commerce Program. During his tenure with both the State Archives and OFT, Alan has had extensive experience working with local governments on electronic records and e-government issues.

From 2004-2010, Alan served in OFT's Security and Risk Management Office where he developed organizational security policies and standards covering areas from Identity and Access Management to wireless networks as well as overseeing the agency's Business Continuity Program. Alan served on the NYS CIO Council's Identity and Access Management Work Group, where he developed New York State's Identity Trust Model and Enterprise Identity Management (EIM) Governance Authority policy. Before leaving OFT, Alan completed a major project to revise the agency's security policies and standards to bring them into line with International Standard Organizations security standards. He is a Certified Information Security Manager (CISM).

UNIVERSITY AT ALBANY

State University of New York

Center for Technology in Government

187 Wolf Road, Suite 301
Albany, NY 12205

PH: 518-442-3892
FAX: 518-442-3886
EMAIL: info@ctg.albany.edu
www.ctg.albany.edu