

GEODATA ACCESS WHILE MANAGING DIGITAL RIGHTS: THE OPEN DATA CONSORTIUM'S REFERENCE MODEL

Bruce Joffe, GISP
Open Data Consortium organizer
GIS Consultants
510-238-9771
GIS.Consultants@joffes.com

Todd Bacastow
College of Earth and Mineral Sciences
Penn State University
814-863-0049
bacastow@psu.edu

INTRODUCTION CONTEXT: GEODATA DIGITAL RIGHTS MANAGEMENT

As geographic content becomes widely available over ubiquitous networks, it becomes easier to distribute, share, copy, integrate, and alter geospatial data. When you go to data portals like Geospatial One Stop (www.geodata.gov) or the Geography Network (www.geographynetwork.com), most of the geodata reported in these catalogs is freely available to the public. Just click and download!



Figure 1, Geospatial One Stop data portal

But there is an even larger amount of geodata that you don't see in these catalogs: public agency geodata that are not freely available. Increasingly, many public agencies, as well as private data providers, desire to protect their rights over the usage of their data. Many local governments and utilities make their geodata available only to a limited group of users. Some make their geodata available only for a price. And some don't make their geodata available at all because they haven't yet defined their data distribution policy.



Figure 2, Geography Network data portal

The current inability to confidently control the description, trading, protection, monitoring, and tracking of intellectual property rights has been a barrier to broader adoption of web-based geospatial data distribution. Therefore, a vast amount of public geodata remains unavailable; their metadata are unpublished in catalogs and data portals.

Is it possible to encourage more public agencies to register their geodata in metadata catalogs? Can more public agencies be encouraged to make their data available through data portals?

Perhaps the answer is "yes" if the geodata portals could handle transactions that are more complex than simply publishing via anonymous web-based distribution. Perhaps more public agency data creators would make their geodata available if an automated process could "negotiate" and establish their particular data distribution requirements with each of the users that contacts them via a geodata portal.

To move these questions forward toward answers, the USGS awarded a small grant to the GeoData Alliance (GDA - www.geoall.net) to investigate the digital rights management of geodata. GDA in turn contracted with the Open Geospatial Consortium (OGC - www.opengeospatial.org) to assess the status of geodata digital rights management, and with the Open Data Consortium (ODC - www.OpenDataConsortium.org) to conceptualize the workflow and functional requirements for automating the geospatial data distribution process. This cooperative work is part of the FGDC's geographic Digital Rights Management (geoDRM) initiative to develop digital rights management capabilities for geographic applications.

The general expression of DRM, or geoDRM, is that the holder of intellectual property (digital geodata) agrees to issue specific permissions (rights) to identified parties (users), within explicit constraints (time interval, number of usage sessions, geographic territory) and under certain obligations (payment, information tracking).

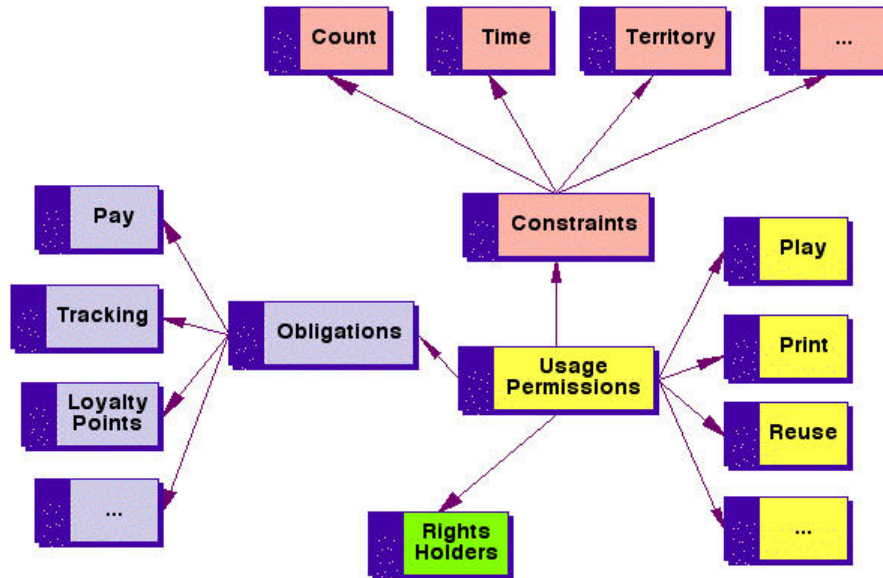


Figure 3, DRM rights expression model¹

Key permissions include rights for viewing (such as printing/plotting, on-screen display, interactive query, manipulation/analysis), transport (such as copy, move, loan, share), and deriving new data (through such actions as extraction, editing, embedding, adding value, and integration).

Automated data distribution must take account of the specific restrictions and requirements each data creator (owner) may have for its data, such as allowing different kinds of access to different types of users, for various geodata themes and features. The variety of possible restrictions is illustrated in a simplified matrix (Fig. 4) which was developed for the Bay Area Regional GIS Council's "homeland security data server" project (see www.BAAMA.org/bargc).

¹ Iannella, Renato, "Digital Rights Management Architectures", D-Lib Magazine, June 2001, <http://www.dlib.org/dlib/june01/06iannella.html>

		USE CATEGORIES			
		(1) Internal Use Only	(2) Einmalig Display	(3) Freiwillig Freigabe	(4) Freiwillig Distribution via Internet
USER TYPES	(A) Emergency Service Provider	A 1	A 2	A 3	A 4
	(B) Government Agencies or their delegated agents	B 1	B 2	B 3	B 4
	(C) Other Public or Educational Institutions	C 1	C 2	C 3	C 4
	(D) Data Contributors	D 1	D 2	D 3	D 4
	(E) Public Domain	E 1	E 2	E 3	E 4

Figure 4, User Types by Use Categories

As local governments contribute their geodata to this mutually beneficial emergency geodata server, they indicate which of five User Types can have access to their geodata, and they indicate the type of access allowed, from among four Usage Categories. Automating a general purpose geodata portal would be similar, although a bit more complicated.

Currently, nearly every public agency has its own, unique, data distribution policy. Yet, in order for an automated distribution portal to "represent" each agency's individual policy, those policies would have to follow standardized data distribution requirements. Fortunately, a standardized framework for individual distribution policies has already been defined, and endorsed, by 117 participants who worked through the Open Data Consortium's (ODC) initiative to formulate a model data distribution policy.

One of the authors (Joffe) organized the ODC initiative, which was supported by the USGS through the GDA, under a separate contract. The process consisted of 24 teleconference workshops in which the participants, who often represented strongly differing views about public geodata, were able to listen to each other and agree on a consensus-based policy.

The ODC model policy provides a useful reference model for automating data distribution through geodata portals because it enables a wide latitude of policy alternatives within a standardized framework. This Reference Model provides definitions and a structure for describing the implicit and explicit concepts and relationships used in the future development of standards relating to the management of digital rights in the geospatial domain. Those local governments that feel they must sell their geodata can do so in a way that is less obstructive than most agency's current policy restrictions, while those

agencies that offer their geodata for free, or simply for the cost of duplication, can also use the model policy to frame their concerns about liability, security, and privacy protection. The ODC data distribution policy model is available for inspection and comment at www.OpenDataConsortium.org.

DATA DISTRIBUTION USE CASES – WORKFLOW CONCEPT

The current geoDRM project envisions an automated (or semi-autonomous) process by which a prospective data user "negotiates" the terms and conditions (distribution rights) of specific data with the owner of that data. Successful negotiation (resulting in a license agreement) then enables the user to access or download the data. The technical mechanics of data discovery, distribution, and retrieval are referenced in this conceptual model, but they are not specified (OGC is working on the technical specification). The purpose of this concept paper is to define what the technology should enable, to define the management of the data rights that would allow the data discovery, distribution and retrieval to take place.

The generic process by which a public agency (generally, a local city or county, regional, state, or tribal government) allows access to (or actively distributes) its geospatial data is outlined as a series of Actions. Each action is comprised by a set of Options. The selection of a specific option for each action in a sequence of actions defines a Use Case. An automated negotiation process would handle each type of use case.

The combination of possible options through the sequence of actions yields approximately 100 potential Use Cases, based on the potentially different data Access Rights to be allowed for the set of potentially distinctive User Types. A user may be allowed up to 19 distribution methods for receiving the data. A data owner may allow a different combination of use cases for each type of data theme. Different owners may allow different combinations of use cases.

The very large number of possible use cases might be simplified by a set of matrices, one for each owner's data theme. One axis would identify the type of user while the other axis would identify the access rights allowed. The intersecting cell between a given user type and a given access right would indicate whether the subject data theme permitted such access for that User Type; it would also indicate the cost, if any, for the user's access to that data theme. This is illustrated in the simplified matrix (Fig 4, above), showing four possible Access Rights (labeled as Use Categories) for five potential User Types. Each owner would fill in the cells of this matrix, one matrix for each of their data themes.

This section reviews the sequence of actions and outlines the options possible in each action. The set of potential use cases for automation of the data distribution process can be derived from this combinatorial sequence (or set of matrices).

Sequence of Actions and Options

Figure 5 illustrates the sequence of actions leading to the definition of each specific data distribution use case, and its fulfillment. Action 0 references the process by which a data owner defines and encodes the distribution rights for each data theme under its control. It presumes that the owner uploads this information as metadata into an on-line geospatial data catalog, portal, or repository.

This distribution management metadata would be embedded with the metadata that is currently defined by the FGDC Metadata Standard, or similar specifications, which describes the contents, accuracy, quality, format, and currency of the data. If the data owner uploads its data into a data repository, the distribution metadata and data description metadata would be embedded with the geospatial data as well. The details of the data registration process are not described here, because they would depend upon the technical means being specified by OGC. It may be assumed that they correspond with the data discovery and fulfillment requirements, described in Actions 1 through 10.

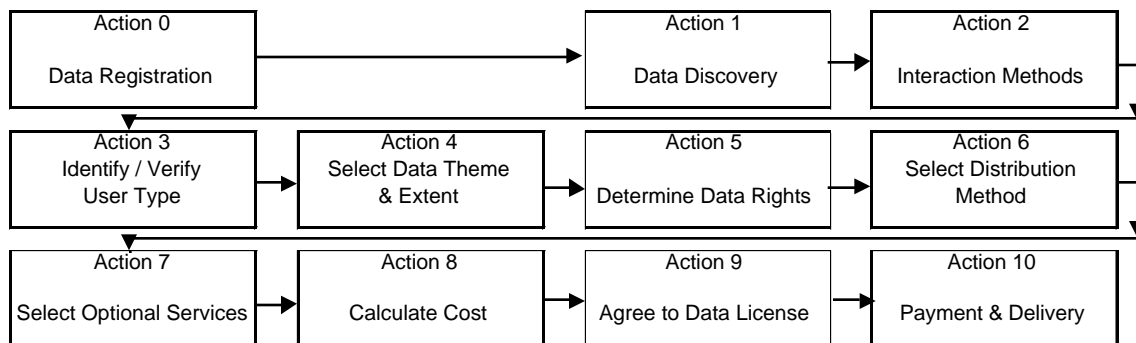


Figure 5, Data Distribution Workflow Actions

Actions 1 through 10 outline the process by which a potential user discovers and selects the data, locational extent, and distribution method; and the data owner verifies the user's identity, offers data access based on the User Type, receives payment and delivers the data.

Action 0 - Registration of Data

0.1 Data owner defines its data distribution policy. It is assumed that the owner's data distribution policy is based on the ODC model policy. Experience indicates that data owners generally act with one of the following intentions:

- a Owner wants to distribute its data widely; often uses website download or web-based query/report applications to do this.
- b Owner will comply with data requests; but requires direct interaction with the user (2.1a, 2.1b).
- c Owner resists distribution; barriers include cost, availability, or fulfillment.

Implementation of an automated Geodata Transaction Application (GTApp), based on these workflow specifications, would enable and enhance the owner's intentions, making it easier to distribute their geodata.

- 0.2 Owner enters its data distribution policy parameters for each data theme's metadata.
- 0.3 Owner uploads metadata into a data catalog or portal, or owner uploads metadata into a public repository. The uploaded metadata corresponds to the information being retrieved in the subsequent Action's matrices, used to determine a given users data rights and options.

Action 1 - Discovery of Data

The following capabilities already exist as on-line services.

- 1.1 User searches an on-line data catalog, portal, or repository for both location and subject of interest. The data catalog/portal is assumed to be linked to the NSDI network of catalogs.
- 1.2 User may read the metadata of data records selected by the search, and,
- 1.3 User may browse bitmap images of selected maps, data, or photoimagery.

Action 2 - Methods of Interaction

Upon finding the source (Owner) of the data that the user seeks, the user may communicate with the owner in one of three ways:

- 2.1 User communicates directly with the owner:
 - a Person-to-person communication via telephone or in-person appearance
 - b eMail communication to a person in the owner's organization
 - c web interface (i.e., automatic interaction)
- NOTE: These options also apply to users who communicate with 2nd party distributors of the owner's data.
- 2.2 User is passed through to the owner's website via the portal/catalog.
 - c web interface (i.e., automatic interaction)
 - 2.3 User "negotiates" directly with the portal or data repository, functioning as a 2nd party distributor.
 - c web interface (i.e., automatic interaction)

This concept paper is focused on the web interface (c) types of Use Cases.

Action 3 - Identify and Verify the User Type

User Types The presumption is that each type of user will qualify for a different set of distribution rights, costs, and restrictions. The ODC model policy recommends as few user distinctions as necessary. The ODC classifications are indicated below as numbers; other optional sub-classifications are indicated as letters.

- 3.1 Value Provider - Includes data sharing partners, cost sharing partners, agencies that offer services or data in return for using the owner's data, or agencies whose mission is integral to the owner's mission. This class is seen as

providing value to the owner, possibly in lieu of a data distribution fee. Optional, specific subcategory distinctions might include:

- a. All subcategories are included in 3.1
 - b. Emergency Service Providers
 - c. Government Agencies or their designated agents
 - d. Data Contributors
 - e. Members (agencies that have joined in a consortium together, for purposes like cost sharing or data sharing; usually defined by payment of a fee or annual subscription and a signing legal agreement)
- 3.2 Data Redistributor – Comprised of private companies that re-sell the owner's data, and public agencies that redistribute the owner's data. Public agencies not interested in controlling resale or royalty fee recovery may omit this class.
- 3.3 Data User - All other recipients, including other public agencies, private agencies and private companies that do not redistribute data, non-governmental and non-profit organizations, students, and private citizens. Optional, specific subcategory distinctions might include:
- a. All subcategories are included in 3.3
 - f. News Media
 - g. Educational and Research Institutions
 - h. Users within the Owner's legal jurisdiction
 - i. Trusted Users (identity has been verified)
 - j. Unknown Users (some owners don't care to make any distinctions)

Verification Methods Owners may want to make positive identification of their data users, for security reasons. Verification would include both initial registration and recognition of the user when he/she makes subsequent requests. This conceptual description does not specify the technical means (OGC is working on that specification), but suggests the following optional methods:

In-person: User presents personal identification to owner.

Notarization: User presents identification to a Notary Public and submits notarized affidavit to owner.

Digital Registration: User presents identification or notarization to a 3rd party registration service, and receives a coded digital signature.

Password ID: Registered or Verified User presents his/her User ID and password to log on to each on-line session at owner's data website

Computer ID: Registered or Verified User's computer automatically presents a coded ID number to log on to each on-line session at owner's website

None: Owner offers data without restriction or identification

Action 4 Select Data Theme and Extent

In a process similar to the search of a data catalog, the user indicates which data themes to be requested. The user may also indicate the locational extent, if a geographic subset is requested.

- 4.1 Locational Extent; recommended interfaces include:
 - a drawing a rectangle or general polygon on a map
 - b entering an address or Assessor Parcel Number, and a radius
 - c pointing to a mapped feature and indicating a radius
- 4.2 Mapped data themes are collections of mapped features.
 - a Themes would include the seven NSDI Framework data themes along with any others maintained by the owner. The two themes which frequently carry the most restriction and price are property parcels and orthoimagery, due to their cost to the owner.
 - b Some owners may provide the option for users to select any combination of mapped features to define their own map theme. In this case, the user would have the lowest access right allowed for any of the selected map features, according to his/her user type. It is suggested that an automatic system notify the user if specific features reduce the overall access level.
- 4.3 Associated tabular data to be requested/provided may include
 - a Metadata describing the characteristics of the mapped features/themes
 - b Descriptive attributes of the mapped features, normally associated with those features
 - c Records associated with the mapped features which come from data tables normally not automatically associated with those features.

An owner determines and specifies (in Action 0) which data themes, features, and attributes would be available to each types of user (Action 3), and also specifies the type of access rights allowed (Action 5). Some owners may even want to code specific data attributes (fields in data records) with differing access rights for specific user types.

Action 5 - Determine Data Rights for Usage Scenarios

Selected data rights may be conferred by the owner to allow the following usage scenarios, according to user type (Action 3) and data theme (Action 4). The ODC model recommends establishing as few data rights distinctions as necessary.

- 5.1 View - User looks at the data² or samples of the data, but does not acquire data.
 - a User may view the data, or query and view the data through the owner's on-line application interface; sees the result but can not retain it in user's computer.

² "Data" refers to maps, imagery and descriptive attributes of geospatial features unless otherwise noted.

- 5.2 Internal Use Only - for use only by the user, its employees, or its designated agents
 - b User may analyze the data through the owner's on-line GIS application interface, and may receive the results as:
 - .1 bitmap image of map and data (e.g., in .pdf format)
 - .2 vector graphics and flat file (comma separated values [csv] or Excel) of the data subset results
 - c User may receive owner's data or a subset of the data as:
 - .1 bitmap image of map and data (e.g., in .pdf format)
 - .2 vector graphics or native GIS format
 - .3 vector graphics or native GIS format, with related attribute data:
 - a) data in flat file format (csv or Excel)
 - b) data in relational table format, with schema
- 5.3 Creation of Derivative Products³ - includes (options a through c), plus:
 - d User receives owner's data and may create derivative products for internal use
 - e User receives owner's data and may create derivative products for external use
- 5.4 Data Update
 - f User updates the owner's database directly
- 5.5 Data Redistribution

The User functions as a 2nd party agent of the owner, distributing its data to 3rd party end users. The data redistributor would implement the same access method policies (options a through f), plus perhaps, provide some additional customer services (see Action 7). Additional scenario rights would include:

 - g User may redistribute copies or subsets of the owner's data to Members or to Trusted Users (defined in Action 3).
 - h User may redistribute copies or subsets of the owner's data to unknown users.

Action 6 Select Data/Distribution Method

Based on the type of user and data themes requested, the owner confers certain data Access Rights to the user (Action 5); the access rights allow certain distribution methods by which the user can receive the data. While several options are listed, note that many states' public record laws require public data to be distributed in the same format as it is used internally by public agency data owner.

- 6.1 View the data (no download)
 - a as a bitmap image (limited pan, zoom or feature identification)
 - b as vector graphics or native GIS format

³ "Derivative Products" are defined in the ODC model policy as works created by the Licensee which incorporate all or part of the Owner's data that remain identifiable and extractable. If the Owner's data can no longer be identified nor extracted, the Licensee's product is considered to be different from the Owner's data.

- c as vector graphics or native GIS format, linked with associated data
 - d as vector graphics and associated data that may be queried or manipulated through the owner's on-line application
- 6.2 Receive the data
- e Receive in digital hardcopy (tape or disk)
 - .1 as a bitmap image (e.g., a .pdf file)
 - .2 as vector graphics or native GIS format
 - .3 as vector graphics or native GIS format, linked with associated data
 - a) data in flat file format (csv or Excel)
 - b) data in relational table format, with schema
 - f Receive as controlled download to Members or Trusted Users; download from website, or FTP, or via eMail
 - .1 as a bitmap image (e.g., a .pdf file)
 - .2 as vector graphics or native GIS format
 - .3 as vector graphics or native GIS format, linked with associated data
 - a) data in flat file format (csv or Excel)
 - b) data in relational table format, with schema
 - g Receive as uncontrolled download to unknown users; download from website
 - .1 as a bitmap image (e.g., a .pdf file)
 - .2 as vector graphics or native GIS format
 - .3 as vector graphics or native GIS format, linked with associated data
 - a) data in flat file format (csv or Excel)
 - b) data in relational table format, with schema

Action 7 Select Optional, Additional Services

The data provider (either owner or redistributor) may offer additional customer services for which additional charges may apply; these could include:

- a Extracting subsets of the data
- b Reformatting the data
- c Delivery in various media to end users
- d Customer technical support
- e Notification of the availability of data updates
- f Automatic transmission of data updates to user

Action 8 Calculate Costs

The cost to the user for the data is determined according to the owner's policy, as entered (conceptually as a cost matrix) in Action 0. The cost matrix parameters include User Type by Data Access Right by Data Theme. Additional costs may accrue from the selection of allowable delivery method and the user's selection of optional services. The owner may also opt to charge no fee.

8.1 Calculated Fee

Owner may opt to charge a fee for each allowable Data Theme selected by the user, as indicated in the User Type by Access Right matrix. The costs of each data theme selected would be summed for a total cost to the user in each data request session.

8.2 Fixed Fee

Owner may opt to charge a fixed fee for any combination of allowable Data Themes, for each user's data request session.

8.3 Membership Subscription

Owner may opt to charge a membership fee for any allowable Data Themes, for all of the user's data request sessions. The membership fee may be a time-limited or count-limited subscription, or it may be perpetual. The user would be designated as a Member. In the case of a cost-sharing or data-sharing consortium, the owner may opt to vary the membership fee according to some "fair share" formula, such as relative population size or area of interest.

8.4 Data-Variable Fee

Some data Themes may be charged according to their quantity (e.g., number of parcels), or the area of the selected locational extent.

8.5 Time and Materials Payment

Optional services (Action 7) and Distribution Methods (Action 6) may be charged according to the owner's staff time expended and the cost of hardcopy materials and postage. This cost probably can not be automated.

8.6 Resale Royalty Payments

These payments pertain to 2nd party data redistributors. Payments may be calculated as:

- a a percentage of the data distributor's gross revenue from the resale of the owner's data, or
- b a transaction fee based on the number of sales, or
- c a single fixed fee

This cost probably can not be automated.

Action 9 Agree to Other Limitations

Data Access Rights not explicitly conferred to the User remain with the Owner. Additional restrictions or duties may be explicitly agreed to, before data access is granted. The following conditions could be included in a Data Use License:

- 9.1 Copyright - The Owner may assert its right to regulate the distribution of its data through its claim of ownership as a copyright. The Data Use License may include two requirements:

- a Acceptance of the following statement: "The data Owner asserts ownership of its data and all its portions. All title, ownership, and intellectual property rights which may exist or be created with the geospatial data shall remain with the Owner."
- b Require posting the Owner's copyright notices.
 - .1 All publication, including via the internet, using any of the Owner's data for release to the public or to others outside the Licensee's organization must include the following notice:
"Copyright, <year>, <name of Owner>"
 - .2 All publications, including via the internet, using geographic information derived from the Owner's data and identifiable therefrom, must include the following notice:
"Derived from data that is Copyright, <year>, <name of Owner>"

9.2 Liability Waiver

- a The Owner may require that the User (Licensee) accept liability waiver statements similar to the following :
 - .1 The Licensee accepts the data "as is", with no guarantee or warranty of accuracy, currency, completeness, or fitness for any use. The Licensee agrees to accept any and all data from the Owner on an "as is" basis. No oral or written information or advice given by the Owner shall create a warranty.
 - .2 While all due efforts will be made to assure that the data conforms to specifications of accuracy and completeness, neither party will make demands on the other if errors or omissions are found.
 - .3 The Licensee waives any and all responsibility of the Owner, explicit or implied, for any damage or liability caused through the use of this data in any way. The Licensee agrees to defend and hold the Owner harmless for any damages of any kind which may be caused by any errors or omissions in the data.
 - .4 The Owner shall not be liable for any occurrence or activity relating to the data, including: lost profits, the fitness of the data for a particular purpose, the installation of the data, or the results obtained from use of the data.
 - .5 This disclaimer shall survive the termination of the License Agreement.
- b The Owner may require that the Licensee displays the following disclaimer note on printed maps, digital web pages, or other reproductions utilizing the Owner's data:

"This is not a survey product. The information is derived from the <Owner's> Databases. The Owner does not assume any liability for damages arising from errors, omissions, or use of this information. Users of this data are advised to be aware of the locational accuracy, data

collection dates, compilation methods, and cartographic format. Users are advised to use this data appropriately."

9.3 Data Correction and Update - The Owner may require that the User (Licensee) agree to actions similar to the following statements:

- .1 If the Licensee detects errors in the data, Licensee shall notify the Owner of such errors in a format compatible with the Owner's data system.
- .2 If the Licensee corrects errors in the data, or updates its copy of the data with more current information, Licensee shall make these modifications available to the Owner in a format compatible with the Owner's data system.

Note: the user may be given access to post corrections directly into the owner's database, at the owner's discretion.

9.4 Usage Tracking - The Owner may require that the User (Licensee) accept the following conditions:

- .1 The Owner or its designated agents have the right to inspect the Licensee's workplace, and work products, to assure that its Data Rights are being observed.
- .2 The Owner or its designated agents have the right to interview the Licensee's employees, agents, clients or customers, to assure that its Data Rights are being observed.
- .3 If the Owner has granted the Licensee redistribution rights for a royalty fee (8.6a or 8.6b), the Owner or its designated agents have the right to inspect the Licensee's accounting records regarding sale of the Owner's data or its derived products.

Action 10 Payment & Delivery

Payment of the calculated cost (Action 8) may be made on-line, through standard credit card mechanisms, or off-line via standard procedures. The owner may decide whether to defer access to the data until payment has been received, or whether receipt of an electronic agreement to pay in the future is sufficient. The use of existing payment delivery services is assumed.

These Actions have described the options that comprise the variety of workflow "use cases" that an automated distribution system (called the Geodata Transaction Application) must enable. The following section describes the application methods through which the use cases are processed.

GEODATA TRANSACTION APPLICATION – COMPONENTS & METHODS

This section describes the types of computational methods required to implement the GeoData Distribution Workflow. We call this automated geodata distribution rights management process the Geodata Transaction Application (GTApp). The workflow

components, diagrammed below (Fig. 6), are labeled (a) through (n), and generally correspond with the Actions described in the previous section.

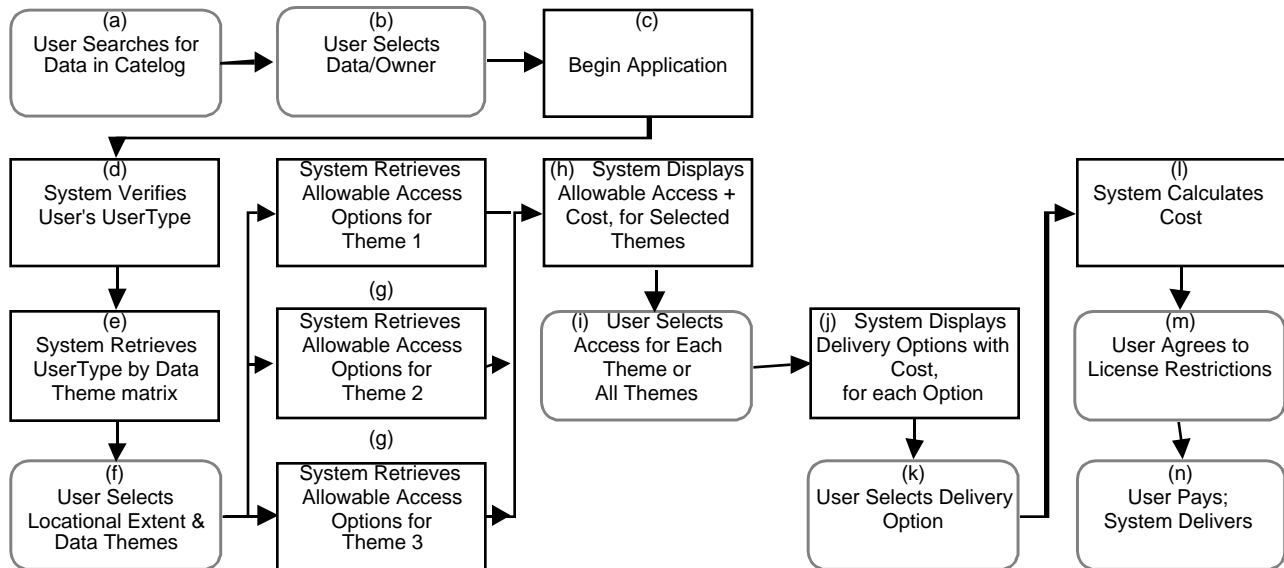


Figure 6, GTApp Components and Methods Workflow

Each component represents either a user's action or the proposed system's response, starting as the user finds and selects the geodata of interest [components (a) and (b)], and then the system retrieves the geodata owner's distribution policy.

Component (a) User Searches for Data in metadata catalogs⁴

This is described in Action 1; it is prerequisite, but outside of the scope of the current requirements statement.

Component (b) User Selects Data/Owner

The User selects one of the data themes of interest within the catalog. This automatically identifies the source agency, or owner of the data. This standard function of data catalogs is prerequisite, but outside of the scope of the current requirements statement.

Component (c) Application Begins

It is assumed that most data owners who have registered their metadata with an NSDI metadata catalog will use the catalog servers to implement the application that controls and enables this Geodata Transaction Application (GTApp). Control may

⁴ An example of a metadata catalog in the NSDI network of catalogs is the California Environmental Information metadata catalog, at <http://gis.ca.gov/catalog/>

be passed to separate application servers under the stewardship of the same agency responsible for the data catalog.

In other cases, control would be passed directly to the owner's application server which would be running an up-to-date version of the GTApp. The requirements for the GTApp are the same, regardless of the server or servers that implement it.

Some data owners may have set up highly controlled, query-report applications for their users.⁵ This option should be sent to the metadata catalog so that a prospective user has the ability to select going directly to the query/report application or to select going to the GTApp.

Component (d) System Verifies UserType

If the owner's policy is to distribute data to unknown users, this component would be skipped.

Initially, the user would have registered its UserType as described in Action 3. The user's identity would be registered either at each specific data owner's site, or at a registration service database. The user enters its UserType ID and password to authenticate identity.

Note that the technical method for UserType verification is to be specified by the OGC as part of the Digital Rights Management initiative. Each category of UserType may be allowed a specified degree of data access permission. Here is an example where data policy standardization is essential, so that there are a standard set of UserTypes to choose from. A given data owner may choose to differentiate among 12 standard sub-categories, defined in the model policy (Action 3), or may clump them together into three, or even two, categories. Each category of user (UserType) is distinguished by different permissions to view or obtain an owner's geodata. Different data owners may confer different combinations of permissions to each category of user.

Component (e) Consult the UserType by Data Theme matrix

This component informs the GTApp of which data themes are available to the subject UserType. It is unlikely that selected data themes would be unavailable to certain UserTypes, because even highly restricted data themes could be displayed in view-only mode. However, the option to completely restrict some data themes is offered.

⁵ An example of a data owner's query-report application is the City of Vallejo's Economic Development Information System, at <http://www.vallejosite.com/ed.asp?s=739>

UserType	Data Theme									
	Geodetic Control	Orthoimagery	Elevation	Transportation	Hydrography	Admin. Boundaries	Cadastral (parcels)	Land Use/Zoning	Land Cover	Geology/Soils
3.1 a. All subcategories of <u>Value Provider</u> 3.1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.1 b. Emergency Service Providers										
3.1 c. Government Agencies or agents										
3.1 d. Data Contributors										
3.1 e. Members										
3.2 <u>Data Redistributor</u>	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
3.3 a. All subcategories of <u>Data User</u> 3.3										
3.3 f. News Media	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.3 g. Educational and Research Institutions	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.3 h. Users within Owner's legal jurisdiction	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.3 i. Trusted Users (verified identity)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3.3 j. Unknown Users	N	A	A	A	A	A	A	A	A	N

Figure 7, UserType by Data Theme matrix

Y = Yes, access is granted

N = No, not granted

A = Access granted only under application control

In the example UserType by Data Theme matrix, all Value Providers are given access to all data themes (Y). Data Redistributors are not given access to cadastral parcels (N), and the access rights for Data Users are determined according to the subcategory of Data User. Unknown Users are only able to view the data through the owner's controlled application (A), which does not show the geodetic control nor geology themes.

This component would initiate a controlled view/query-report application for those UserTypes that did not have access rights to select and retrieve data.

Component (f) User Selects Locational Extent and Data Themes/Features

The system invokes the locational extent application, enabling the user to identify area of interest, using one of the methods outlined in Action 4.1

The name of the data theme that the user selected in the metadata catalog [Component (b)] would have already been passed to this GTApp. In addition, the user would be presented with a list of data themes available to its UserType [based on the UserType by Data Theme matrix, component (e)]. The user would be able to select additional themes. In a more sophisticated version of the GTApp, the user would be presented with a list of mapped features and data attributes from which to select as well.

Component (g) System Retrieves Allowable Access Options for each selected Theme

An owner may give different UserTypes different access rights, for each data theme. Each cell in an Access Right by UserType matrix (Fig.8) indicates which access rights each UserType has, and also indicates the cost of each access right, where appropriate. In cases where the owner wishes to distribute the data freely, the cost is "0".

Access rights to the selected geodata may vary from only viewing the data through an application program, to receiving the results of a query as a bitmap image, to receiving the results as a vector map, to receiving a vector map with attached data, to receiving the entire data file, or finally to be allowed to redistribute the data (outlined in Action 5). This variety of access rights may be conferred differently by each data owner to each category of user, and differently for each data theme.

If the owner chooses to charge a fee for the data, or different fees for different degrees of access to the data, or different fees for different data features or themes, the automated transaction process will present those options to the user [component (h)].

In the example Access Right by UserType matrix, the owner has assigned differing access rights to various types of Value Providers. The owner has assigned the same access rights to all "Other Users" except Unknown Users.

for Parcel GeoData		UserType										
		3.1 a All subcategories of Value Provider	3.1 b Emergency Service Providers	3.1 c Government Agencies or agents	3.1 d Data Contributors	3.1 e Members	3.20 Data Redistributor	3.3 All subcategories of Data User	3.3 f News Media	3.3 g Educational and Research Institutions	3.3 h Users within Owner's legal jurisdiction	3.3 i Trusted Users (verified identity)
Access Rights												
5.1 a	View Only through Application	0	0	0	M	0	0					0
5.20	Internal Use Only											
5.2 b	Receive results from Application:											
5.2 b 0.1	bitmap image (.pdf)	0	0	0	M	0	0					0
5.2 b 0.2	vector graphics + flat file	0	0	0	M	0	0					N
5.2 c	Receive geodata as:											
5.2 c 0.1	bitmap image of map and data	0	0	0	M	0	0					N
5.2 c 0.2	vector graphics or native GIS	0	0	0	M	0.01	0.01					N
5.2 c 0.3	vector graphics or GIS, + attributes:											
5.2 c 0.3a)	flat file format (csv or Excel)	0	0	0	M	0.03	D					N
5.2 c 0.3b)	relational table format,+ schema	0	0	0	M	0.03	D					N
5.30	Create Derivative Products:											
5.3 d	for internal use	0	0	0	M	0.03	0.03					N
5.3 e	for external use	0	N	N	N	0.03	N					N
5.40	Data Update	N	0	N	N	N	N					N
5.50	Data Redistribution											
5.5 g	to Members or to Trusted Users	N	N	N	N	0.03	N					N
5.5 h	to unknown users.	N	N	N	N	N	N					N

Figure 8, Access Rights by UserType

N = no access

M = membership fee

D = refer to a separate matrix for access to attribute data

[number] = price per parcel

0 = no fee for the data

The method for encoding each cell with the unit cost of the data theme access right is still to be determined. The cost for Members, for example, may be a subscription fee irrespective of the data themes chosen. Some owners may opt for a fixed fee, irrespective of amount or type of data selected for a transaction session. Others may charge on a unit rate: cost per parcel, or cost per area, or cost per megabyte. The matrix informs the system of the unit cost with coded values in the matrix cells.

The owner would describe its distribution policy for each data theme as a separate matrix.

If specific access rights were to be given for specific data attributes, or for specific mapped features, the cells in this matrix would reference a separate Cell No. by

Data Attribute matrix (not shown). In such a matrix, the vertical axis would contain the reference number for each cell in the Access Rights by UserType matrix that was indicated with a "D". The horizontal axis would contain all the data attributes for which access rights (including cost) were to be differentiated.

Component (h) System Displays Cost for Selected Themes

The system consults the owner's cost strategy (described in Action 8), which is stored in a reference matrix. It then creates a table for the user, indicating the various access options and costs for the data themes that the user has selected. This table is built from the information in the Access Right by UserType matrix, and is displayed as an Access Right Table (below). Only allowable options are displayed for a specific user.

	Access Rights and Cost for Known Data User of Parcel Data	Unit Cost	Total Cost
5.1 a	View Only through Application	0	0
5.20	Internal Use Only		
5.2 b	Receive results from Application:		
5.2 b 0.1	bitmap image (.pdf)	0	0
5.2 b 0.2	vector graphics + flat file	0	0
5.2 c	Receive geodata as:		
5.2 c 0.1	bitmap image of map and data	0	0
5.2 c 0.2	vector graphics or native GIS	0.01	\$12.50
5.2 c 0.3	vector graphics or GIS, + attributes:		
5.2 c 0.3a)	flat file format (csv or Excel)	D	see table
5.2 c 0.3b)	relational table format,+ schema	D	see table
5.30	Create Derivative Products:		
5.3 d	for internal use	0.03	\$37.50
	<u>Roads/Streets Data</u>		
5.1 a	View Only through Application	0	0
5.20	Internal Use Only		
5.2 b	Receive results from Application:		
5.2 b 0.1	bitmap image (.pdf)	0	0
5.2 b 0.2	vector graphics + flat file	0	0
5.2 c	Receive geodata as:		
5.2 c 0.1	bitmap image of map and data	0	0
5.2 c 0.2	vector graphics or native GIS	0.01	\$1.20
5.2 c 0.3	vector graphics or GIS, + attributes:		
5.2 c 0.3a)	flat file format (csv or Excel)	.02	\$2.40
5.2 c 0.3b)	relational table format,+ schema	D	see table
5.30	Create Derivative Products:		
5.3 d	for internal use	0	0

Figure 9, Costs of Allowable Access Rights, for specific User, for selected Data Themes

Component (i) User Selects Access for Each Theme or All Themes

The user interacts with the Access Right Table [Component (h)], to select which themes, and which type of access desired. The user also sees the unit cost for each type of access right available. When the cost is calculated on a per-unit basis, the system automatically calculates the full cost of that data theme, based on the area extent entered in Component (f).

Component (j) System Displays Delivery Options with cost for each option

Similar to the Access Right by UserType matrices [Component (g)], the system consults a UserType by Distribution Method matrix (Fig. 10) to determine which options to offer the user and what the cost of each option is. There is only one such matrix which covers all available data themes. The system displays the user's options, with costs, as a Delivery Option Table (not shown). The potential delivery options are outlined in Action 6.

The user selects among the delivery options that are available for his/her UserType, ranging from simply viewing the data, to receiving a digital hardcopy, or to immediately downloading the data. Security protections are implicit because the data owner specifies whether anonymous users can download their data, or whether that access is given only to verified categories of UserType.

		UserType											
		3.1 a	3.1 b	3.1 c	3.1 d	3.1 e	3.20	3.3 a	3.3 f	3.3 g	3.3 h	3.3 i	3.3 j
		All subcategories of Value Provider	Emergency Service Providers	Government Agencies or agents	Data Contributors	Members	Data Redistributor	All subcategories of Data User	News Media	Educational and Research Institutions	Users within Owner's legal jurisdiction	Trusted Users (verified identity)	Unknown Users
Distribution Method													
6.1	View the data (no download)												
6.1a	as a bitmap image	0	0	0	0	0	0	0	0	0	0	0	0
6.1b	vector graphics or native GIS format	0	0	0	0	0	0	0	0	0	0	0	0
6.1c	vector graphics or GIS,+ data	0	0	0	0	0	0	0	0	0	0	0	0
6.2	Send the data												
6.2e	Send data in digital hardcopy												
6.2e.1	as a bitmap image	25	0	25	25	25	25	25	0	25	25	25	N
6.2e.2	vector graphics or native GIS format	50	0	50	50	50	100	50	0	50	50	50	N
6.2e.3	vector graphics or GIS,+ data												
6.2e.3a)	data in flat file format (csv, Excel)	25	0	25	25	25	150	25	0	25	25	25	N
6.2e.3b)	relational table format,+ schema	50	0	50	50	50	200	50	0	50	50	50	N
6.2f	Send as controlled download												
6.2f.1	as a bitmap image	0	0	0	0	0	0	0	0	0	0	0	N
6.2f.2	vector graphics or native GIS format	0	0	0	0	0	100	50	0	0	0	50	N
6.2f.3	vector graphics or GIS,+ data												
6.2f.3a)	data in flat file format (csv, Excel)	0	0	0	0	0	150	75	0	0	75	75	N
6.2f.3b)	relational table format,+ schema	0	0	0	0	0	200	100	0	0	100	100	N
6.2g	Send as uncontrolled download												
6.2g.1	as a bitmap image	na	na	na	na	na	na	na	na	na	na	na	0
6.2g.2	vector graphics or native GIS format	na	na	na	na	na	na	na	na	na	na	na	N
6.2g.3	vector graphics or GIS,+ data												
6.2g.3a)	data in flat file format (csv, Excel)	na	na	na	na	na	na	na	na	na	na	na	N
6.2g.3b)	relational table format,+ schema	na	na	na	na	na	na	na	na	na	na	na	N

Figure 10, User Type by Distribution Method matrix

[number] = cost in \$ USD

N = distribution method not available for this user

Component (k) User Selects Delivery Options, may also select additional services.

The user interacts with the Delivery Option Table, indicating the method preferred.

In a more advanced implementation, the user might also select some of the additional services outlined in Action 7. However, initially, this option will not be automated.

Component (l) System Calculates Cost

The system refers to the owner's cost strategy table (described in Action 8), to know whether to calculate the costs or impose a fixed cost, or no cost at all.

Component (m) User Agrees to Other License Conditions

The owner may choose to require that the user agree to certain standard conditions or limitations concerning copyright, liability, data correction, or usage tracking. These conditions are described in Action 9. A standard form would appear giving the user the option of "accepting" or "canceling" the transaction. The license agreement would be transmitted or delivered to the user along with the requested data. Repeat users, or Members, would need to acknowledge this agreement only the first time, as the agreement remains in the owner's records.

Component (n) User Pays, System Delivers

Payment would be made through a standard web interface. It could be a credit card transaction, or it could be a debit to the user's existing account with the data owner. Upon payment, the owner's system transmits the data to the user according to the delivery method selected [Component (k)].

This GTApp requirements definition enables the flexibility necessary to engage many more local agencies in using geodata portals to distribute their data, while protecting their rights. The geoDRM initiative has the vision of advancing an interoperable capability for describing, identifying, trading, protecting, monitoring and tracking all forms of geodata rights usages.

Since 2002, the FGDC has been working with the GDA and the ODC to create the Model Data Distribution Agreement and to apply this in a standardized machine-readable form. The current effort increases the awareness and understanding of digital rights management in the geographic community. Future goals include prototyping an interoperable capability meeting the needs of a regional geographic organization. Thus far, the ODC has developed use cases and the GDA has convened a workshop on digital rights management, vetted the use cases, and selected a digital rights language for prototypical implementation.

ABOUT THE AUTHORS

Bruce Joffe, founder of GIS Consultants in Oakland, CA (<http://joffes.com/gis>), provides GIS implementation planning and management assistance to local governments and utilities. In 2001, he organized the Open Data Consortium to resolve the many contentious issues surrounding geodata distribution, through consensus-building communication among government, business, and academia. GIS Consultants continues assisting public agencies develop their geodata distribution policy.

Todd Bacastow, Assistant Professor of Geography at Penn State University, is Assistant Director of the Earth & Environmental Systems Institute. Todd is a past Chair of the GeoData Alliance. His experiences include working with numerous Federal and state agencies on Geographic Information System (GIS)-related issues; serving on technology planning committees and councils; and advising the Pennsylvania State Legislature on spatial technologies. He has recently completed research in peer-to-peer spatial systems.

BIBLIOGRAPHY

Bell, Tom W. *Fair Use v. Fared Use: The Impact of Automated Rights Management on Copyright's Fair Use Doctrine*, North Carolina Law Review 76 (1998): 557.

Iannella, Renato. *Digital Rights Management Architectures*, D-Lib Magazine, June 2001, <http://www.dlib.org/dlib/june01/06iannella.html>

Joffe, Bruce. *To Sell or Not to Sell: GIS's Budgetary Dilemma*. GeoInfo Systems magazine, September 1995, Advanstar Communications, Eugene, OR.

Joffe, Bruce,. *Open Data Consortium Proposes Geodata Portal Requirements*, ArcNews (Fall 2004).

Lessig, Lawrence. *The Law of the Horse: What Cyberlaw Might Teach*, Harvard Law Review, Fall 1999, available at <http://cyber.harvard.edu/lessig.html>.

Lessig, Lawrence. *Code and Other Laws of Cyberspace* (New York: Basic Books, 1999), esp. Chs. 4 and 10.

Paskin, Norman. *Digital Object Identifier*; Proceedings from Technical Association for the Graphic Arts 54th Annual Technical Conference. (April 14-17 2002).

Software & Information Industry Association. *The Value of Information, Building the Net: Trends Shaping the Digital Economy* (Washington, D.C.:SIIA, 2000), available at <http://www.trendsreport.net>.

Stefik, Mark. *Shifting the Possible: How Trusted Systems and Digital Property Rights Challenge Us to Rethink Digital Publishing*, Berkeley Technology Law Journal 12 (1997): 137.

U.S. Department of Commerce, Task Force - Working Group on Intellectual Property Rights. *Intellectual Property and the National Information Infrastructure: The Report of the Working Group on Intellectual Property Rights*, available at <http://www.uspto.gov/web/offices/com/doc/ipnii/>.

ADDITIONAL REFERENCES

A list of links to data policy articles and some data distribution websites is maintained on the Open Data Consortium website, at www.OpenDataConsortium.org click on "News/Links" and then click on "Web Links." Links to local government data policies are at www.OpenDataConsortium.org click on "Information Repository."

The following article links relate to GIS Data Policy:

Link to most state and federal statutes www.law.cornell.edu/statutes.html

On-line compendium of state open record laws
www.rcfp.org/cgi-local/tapping/index.cgi

Adelaide City Council
http://www.adelaidecitycouncil.com/council/publications/Policies/Spatial_Data_Policy.pdf
http://www.adelaide.sa.gov.au/council/publications/Policies/Spatial_Data_Policy.pdf

ANZLIC <http://www.anzlic.org.au/policies.html>

Florida National Areas Inventory http://www.fnai.org/PDF/GIS_policy.pdf

Office of Spatial Data Management <http://www.osdm.gov.au/osdm/policy.html>

West Virginia Department of Environmental Protection data policy
<http://129.71.240.42/gps/geospatial.html>

Additional articles, collated by Amirali Shaerzadeh :

Boulder County CO pricing policy

http://www.co.boulder.co.us/gis/cost_recovery/cost_pricing.htm

Canadian Data Policy study by KPMG

<http://cgdi.gc.ca/english/supportive/KPMG/KPMG.pdf>

Critique of Canadian data sales policy in Geo Place magazine

<http://www.geoplace.com/gw/1999/0699/699can.asp>

Digital Earth Site policy study

http://www.digitalearth.ca/pdf/DE_A_227.PDF

Durham NC data sales policy

http://www.ci.durham.nc.us/forms/gis_commercial_data_policy.pdf

Netherlands study of data policies

http://www.lmu.jrc.it/Workshops/8ec-gis/cd/papers/3_p_uw.pdf

New York State Office for Technology policy recommendations

http://www.oft.state.ny.us/policy/tp_976.htm

Revisions to US A-16 policy in GIS Monitor

<http://www.gismonitor.com/news/newsletter/archive/082902.php>

University of Maine research agenda for spatial databases

http://www.spatial.maine.edu/tempe/onsrud_2.html

West Virginia Department of Environmental Protection data policy

<http://129.71.240.42/gps/geospatial.html>