

Open Data Consortium Proposes Geodata Transaction Requirements

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

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! (Fig 1, typical geodata 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.

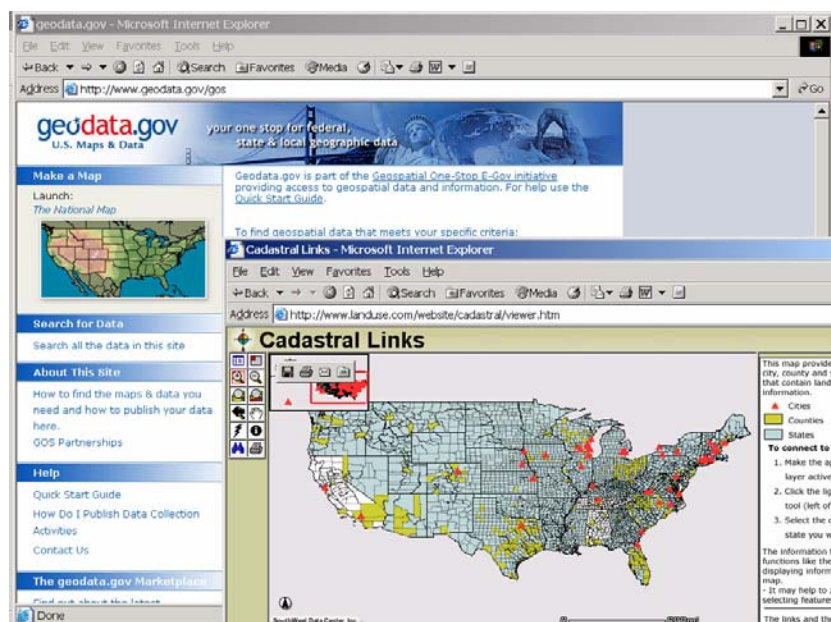


Figure 1, Geospatial One Stop data portal

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. Therefore, a vast amount of public geodata remains invisible; their metadata are unpublished in catalogs and data portals.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 for anonymous web-based distribution. Perhaps more local governments would make their geodata available if an automated process could

"negotiate" 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, and with the Open Data Consortium (ODC - www.OpenDataConsortium.org) to conceptualize the workflow and functional requirements for automating the data distribution process.

Automated data distribution must take account of the specific restrictions and requirements each public agency 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. 2) which was developed for the Bay Area Regional GIS Council's homeland security data server project (see www.BAAMA.org/bargc).

		USE CATEGORIES			
		(1) Internal Use Only	(2) Blitmap display via web	(3) Free distribution to third parties	(4) Free distribution to third parties via Internet
USER TYPES	(A) Emergency Service Provider	A1	A2	A3	A4
	(B) Government Agencies or their delegated agents	B1	B2	B3	B4
	(C) Other Public or Educational Institutions	C1	C2	C3	C4
	(D) Data Contributors	D1	D2	D3	D4
	(E) Public Domain	E1	E2	E3	E4

Fig. 2, 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 some kind of standardization of their 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 initiative to formulate a model data distribution policy.

This author organized the ODC initiative, which was supported by the USGS 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 standard for automating data distribution through geodata portals because it enables a wide latitude of policy alternatives within a standardized framework. 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.

Our formulation of requirements for automating a geodata distribution portal began by describing the workflow components (Fig 3).

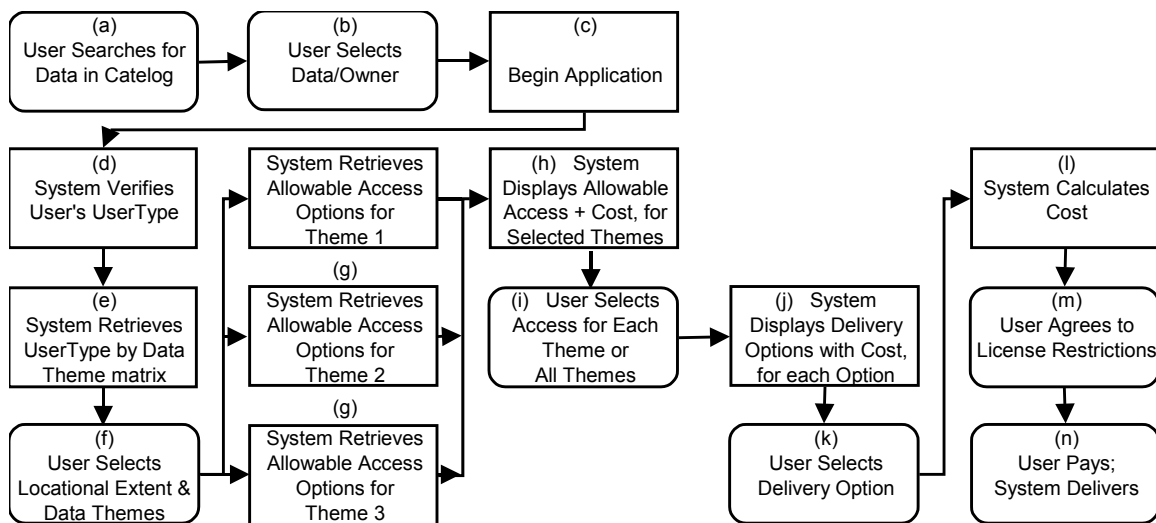


Fig 3, Automated Geodata Distribution Workflow

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 contacts the geodata owner's web-based policy.

The system queries the user and verifies the category of UserType he/she belongs to [component (d)]. Note that the technical method for UserType verification is to be specified by the OGC as part of its 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 number of UserTypes to choose from. A given data owner may choose to differentiate among 12 standard categories, defined in the model policy, or may clump them together into three, or even two, subgroups. 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.

The first "sieve" of UserType permissions that the automated portal transaction activates is the presentation of geodata themes, as well as specific map features and data attributes, that a given UserType is allowed to view or acquire [component (e)]. The automated system does this according to a UserType by Data Theme specification matrix that each data owner completes for its data. Each agency has the option of defining the degree of access it wishes to give to each type of user, for each of its geodata features and attributes.

The user then selects the locational extent and identifies which data themes or features to see [component (f)], and the system responds by presenting the user options for accessing the geodata, following a matrix completed by the data owner (Fig 4) [component (g)].

for Parcel GeoData		UserType										
		3.1 a All subcategories of <u>Value Provider</u> 3.1	3.1 b Emergency Service Providers	3.1 c Government Agencies or agents	3.1 d Data Contributors	3.1 e Members	3.20 <u>Data Redistributor</u>	3.3 a All subcategories of <u>Data User</u> 3.3	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)
5.1 a	Access Rights 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

Fig 4, 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

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. 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)].

Finally, 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. The user then pays (if necessary) and the system delivers.

This requirements definition enables the flexibility necessary to engage many more local agencies in using geodata portals to distribute their data, while protecting their rights. Greater detail is in the specification paper, available at the ODC website.

[Author's Bio]

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. He organized the Open Data Consortium (www.OpenDataConsortium.org) 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.