Tech Watch Column: Government Information Mashups to the People!

Sonnet Ireland, University of New Orleans
Rebecca Blakeley, McNeese State University

Available at: https://works.bepress.com/sonnetireland/2/
Mashups...we've all heard the chattering about this relatively new technology, but what is it really? Simply put, a mashup is a webpage or application that combines data or functionality from two or more sources to create a new service. They often use open APIs (Application Programming Interfaces) and data sources (the data comes in standard formats such as Atom, RSS, XML, Text/CSV) to produce a new presentation of the information, usually not available from the original source. Usually a mashup is created to consolidate information with a more usable interface and is a great way to make information easier to find.¹ Ultimately, it’s about taking raw information and presenting it in a new, useful way. This can involve collecting information from one or more sources, and selecting, extracting, arranging, and combining it to fostering new understanding. Sounds like what librarians do too, doesn't it? We'll explain just why we should consider creating and using mashups with our library user communities and include some examples in the context of government information and data. But first, let's go over a few terms that will be used throughout this column. The Free Government Information website (FGI) provides a glossary of technology terms that you can refer to for more information as well.²

- **API** - a formal definition of how computer programs can interact with a particular dataset, database, website, or other cache of information. Webpages present information for human consumption; APIs present information for consumption by computer programs. An API defines a set of commands, functions, and protocols that programmers can use to create software to select, extract, acquire and process data from the providing source. When an information producer provides an API, it is allowing others to program new, different interfaces, views, and presentations of that information. Users benefit because they are no longer limited to a single view of the information but can view alternative presentations, customized for particular
communities, interests, or purposes. Mashups often use APIs to combine data from different sources (e.g., census data, crime data, and maps to create city information at Everyblock.com, dc.everyblock.com/crime/) to re-use and re-purpose what would otherwise be separate "stove-pipes" of information and static documents. Typically, APIs draw data in real time from the data-provider.

- **Raw Data** - information stored in a structured format or database suitable for processing by computer, but typically not suitable for direct presentation to a human without first being processed. For example, a database of all population census information is “raw” data that can be processed to produce many different tables and maps and charts. APIs usually deliver “raw data.”

- **Bulk Data** - The term "bulk data" simply refers to large quantities of raw data available for transferring in bulk from producer to consumer. While a library might use an API to draw data in real time from a data provider into a webpage, it might download data in bulk onto its own servers in order to present that data in ways that go beyond the definitions provided by an API. A library might, for example, download a file of raw census data and load it into statistical software for processing on library computers. Last year Tim O'Reilly wrote about how bulk data downloads are a breakthrough in terms of government transparency, and since the launch of Data.gov many government agencies are releasing bulk data in various formats (XML, CSV, etc.) and many individuals, organizations, and companies are creating unique mashups with this data.

- **XML** - Extensible Markup Language (XML) is "specification for creating custom markup languages intended for use on the Web". It can be used with HTML webpages and, like HTML, it allows text to be tagged or "marked up." However unlike HTML, which focuses on appearance, XML allows content to be tagged to signify meaning. For example, it can describe items accessed when a webpage loads. Simply put, XML allows you to create a...
database of information without having an actual database. XML is generally human-readable and therefore more easily preserved than proprietary, binary formats. It is also designed so that it can be easily parsed by computers and so is often called 'machine-actionable' because it makes it possible to more easily re-use and re-purpose content. There is an XML version of the *Federal Register* that has already prompted new uses of it at FedThread.org, which uses the government-provided XML to create *functionality* (collaborative annotation, customized feeds, advanced searching, etc.) that the government does not provide.

Now that we understand a bit more about what a mashup is and what it is comprised of, why should we consider using and creating mashups for our library user communities? What are some examples of mashups in the context of government information and data that we could use as models to create our own in order to creatively and effectively serve our users? Librarians help navigate the maze of government information and data, but how can we as librarians use and create mashups to engage our patrons? When we have rich collections of digital information and data from a variety of sources, it will be possible for us to develop rich interfaces (i.e. mashups) and combinations of information from those different sources.

When the private sector, or non-profit organizations such as Govtrack (www.govtrack.us) or the Sunlight Foundation (www.sunlightfoundation.com/) and its Sunlight Labs (www.sunlightlabs.com/), create a useful online government information service (i.e. MapLight.org or Transparencydata.com), they do so by acquiring the digital information in bulk and creating a mashup that makes correlations, making the data more useful, not just simply linking to a PURL or static website. The Gulf Coast Oil Spill Map, "How Big is the Deepwater Horizon Oil Spill" (paulrademacher.com/oilspill/) created by Paul Rademacher using the Google Earth API and data from the Louisiana Governor's Office of Homeland Security & Emergency Preparedness website (gohsep.la.gov/oilspill.aspx) and NOAA-NESDIS (response.restoration.noaa.gov/) is an example of what can be done when you acquire data to create a mashup. Libraries could do the same. Even mobile phone apps such as the ecofindeRRR for iPhone (www.ecofinderapp.com/) could be a useful model for those interested in developing simi-
lar local government data mashup apps for mobile phone use by their user communities. Libraries could facilitate with others such as local government and community groups, fellow "techie librarians" and "not-so-techie librarians" who specialize in government information, or their IT departments to create mashups that would benefit the community's specific information needs or interests and promote civic engagement.

There are libraries creating mashups for their user communities and patrons. For example, McNeese State University's Frazar Memorial Library Archives & Special Collections Department has created the Google Map mashup “The Fire of 1910 in Lake Charles, LA” (tinyurl.com/248zk97) using the Google MyMaps creation feature (maps.google.com/). Are there any government information librarians creating mashups that focus on government information and data? We could not find any, which is why we decided to write this column and gather practical resources and tools for us to use when we attempt to create our own government information related mashups! For example, Rebecca was inspired by the RTKNet: Right to Know Network’s TRI Database (www.rtknet.org/db/tri/) so she has been brainstorming ideas about creating a simple mashup for her user community with a focus on toxicological, pollution, and industry data in certain areas of Southwest Louisiana, which is a subject that patrons research the most in her government documents department.

You do not have to be an expert at programming or a "techie librarian" with mad coding skills to create a mashup. There are many online tutorials and "how-to" books that are available to learn from, as well as mashup creating tools ranging from easy-to-use to the more advanced.9 Programmable Web.com (www.programmableweb.com/howto) published a tutorial to help get you started if you have no idea where to begin. First, decide on a subject and where your data will come from (since this can help determine what APIs you'll use), via various government information resources such as Data.gov (www.data.gov) or government agency websites.10 Then you will need to extract data using tools such as Dapper (www.dapper.net/open/) and store that data. You may need to decide if you need a server, depending on the amount of data you need to store or how large your mashup will be. If you are not comfortable with creating a complex mashup with a larger amount of data from scratch, but would rather build a simpler mashup using a smaller amount of data, there are various mashup creating tools that are very easy to use. Mashable.com is a great website to...
find updates, news and reviews of the latest mashup making tools (both free and fee-based) such as Yahoo! Pipes (pipes.yahoo.com/pipes/), Google Mashup Tools (code.google.com/), and Intel Mash Maker (mashmaker.intel.com/web/). If you are more interested in data mapping tools that go beyond the capabilities of Google MyMaps, the Sunlight Foundation Blog has a great list of resources and tutorials for making map mashups, including such tools as Maker! (maker.geocommons.com/) and Mapspread (mapspread.com/).

All of these tools, though they are all created with the purpose of helping people create mashups, answer a wide variety of different needs and skill sets. If you would like to collaborate, strategize, or learn more about mashups, let’s meet up at the GODORT wiki’s Birds of a Feather (BOAF) page (wikis.ala.org/godort/index.php/Birds_Of_A_Feather) and ALAConnect's Government Information Interest Group (connect.ala.org/node/72764). The more we learn to use these mashup creating tools and create mashups, especially those with a government information focus, the greater the power we have in helping our library user communities efficiently find, access, and correlate information, putting power into the people's hands.

References


3. Ibid.


8. Ibid.

9. Nicole C. Engard, Library Mashups: Exploring New Ways to Deliver Library Data, (Med-

