Contents

▼ ADVISOR INTRODUCTIONS
From Your Managing Editor: Aggregating Access to Digital Depositories
by George Machovec .......................................................... 3
Advisor Review Elements .................................................. 4
Advisor Rating Guidelines .................................................. 4

▼ ADVISOR REPORTS FROM THE FIELD
Mobile Apps for Libraries: The DiRT on Digital Humanities Apps
by Jolanda-Pieta (Joey) van Arnhem ........................................ 53

▼ ADVISOR INTERVIEWS
Wiley and ReadCube: An Interview with Todd Toler, Vice President, Digital Product Management at John Wiley & Sons
by Dennis Brunning .......................................................... 57

▼ ADVISOR SCOREBOARD
TCA Scoreboard ............................................................... On the Web

Advisor Reviews in This Issue

▼ ADVISOR REVIEWS—COMPARATIVE REVIEW
OAIster on EBSCO Discovery Service, FirstSearch, and OAIster.worldcat.org
by David J. Gaines and Jody Condit Fagan ........................... 5

▼ ADVISOR REVIEWS—STANDARD REVIEWS
Books at JSTOR
by Susan Maret .................................................................. 13
Congress.gov
by Katy Stein Badeaux ...................................................... 17
Data-Planet
by Jennifer Starkey ............................................................ 21
Discogs.com
by Joseph Hartnett ............................................................. 26
Gale’s Biography in Context
by Lizah Ismail ................................................................. 34
Independent Voices
by Donna Gunter .............................................................. 39
Kanopy Streaming
by Robert Flatley .............................................................. 42
PlumX
by Merinda McLure ........................................................... 46
U.S. Intelligence on the Middle East, 1945–2009
by Anna L. Shparberg ....................................................... 49

Advertisers
Accessible Archives ............................................................. 59
Annual Reviews ................................................................. 60
Cold Spring Harbor Laboratory Press ................................... 61
Gale Cengage Learning ....................................................... 63
Modern Language Association ............................................ 64
World Scientific Publishing Co. .......................................... 62

Subscription information on pages 16 and 33. Scoreboard on the Web.
Abstract
Data-Planet is the largest single repository of harmonized data, sourced from a wide variety of producers including the U.S. Government, intergovernmental bodies, and private organizations. Two distinct interfaces give Data-Planet the flexibility to serve users with basic data needs and/or complex research projects. Data-Planet makes vast amounts of data easier to find through a single point of access, makes data comparable across data sources, allows data downloading and exporting, and enables analysis with built in visualization tools that do not require specialized software or knowledge. Although the advanced interface takes persistence to master, the product ultimately streamlines researchers’ access to datasets and statistical information.

Pricing
Data-Planet offers a choice of two pricing models for academic customers: an annual subscription and a one-time purchase with perpetual rights. Annual pricing is based on institutional FTE of students, faculty and staff (approximately $0.70 per FTE for 2015) with a price cap ($16,800 for 2015) and a minimum in place. The one-time purchase model provides access in perpetuity to the basic content for a fee that is larger up front with a lower annual maintenance fee to access the platform and content updates. Matt Dunie, President of Data-Planet described this model in an email correspondence with the reviewer in January 2015:

“[t]he deliverable is the license to use the data sheets in perpetuity for all of the data added to the system for up to and through the current year. Updates for future years are priced at a nominal fee going forward, and include both updated data and continued access to the interfaces. Access to the interfaces is included for one year. Updates to the content, and continued access to the interfaces, and updated data sheets are provided for approximately 10% of the First Year price, but there is no obligation for the customer to continue with updates if they do not want to or have the budget. The Perpetual Rights model seems to resonate with our customers because we are able to immediately deliver the DataSheet meta-data with DOI links as well as provide access to the interfaces and update the data continuously throughout the term.”

There are a small number of partner organizations that do not allow Data-Planet to sell perpetual access to their content: Dave Leip’s Political Atlas, China Data Center, British Bankers Association, and XIGNITE. There are also several premium add-ons to DataSets Basic which have an additional cost to cover the higher royalties the company pays to the data producers. Currently available add-ons are: EASI Market Planner (consumer spending data), China Data Center (broad subject coverage of China), and International Stocks and Precious Metals.

Current licenses to DataSets Basic include access to both interfaces: Statistical Datasets and Statistical Ready Reference, as well as MARC records (or XML or text files) for about 150,000 datasheets.

Product Description
Data-Planet is a product of Conquest Systems, Inc., the company that designed the platform behind Lexis-Nexis Statistical Datasets (no longer available to libraries.) The Lexis-Nexis product was subsequently purchased by ProQuest in 2010 which distributed it to library subscribers as ProQuest Statistical Datasets until 2012. The agreement between ProQuest and Conquest Systems ended, and since 2013 it has new branding as Data-Planet and is sold directly by Conquest Systems to academic and public library customers. Lexis-Nexis remains the distributor to companies and organizations outside of the library marketplace.

Data-Planet is the largest single repository of statistical data that is harmonized to the greatest extent possible to facilitate comparison across datasets. It contains 18.9 billion data points, and 3.9 billion time series. It includes both U.S. and international data in a wide range of subjects. The core product, DataSets Basic, includes access to all sources in Data-Planet; three premium add-ons are also available at additional cost: EASI Market Planner, China Data Center, and International Stocks and Precious Metals. Beyond serving as a data repository, Data-Planet has developed two different and complementary interfaces as access point, both of which come standard with a license to the core product. The Statistical Datasets interface is the flagship product, which contains the advanced analytical tools and has numerous visualization options for users to customize the output. The Ready Reference interface provides entry point to searching 180,000 datasheets, aimed at users seeking quick answers. Each datasheet includes a preset graph of a single table with a graph of one time series, a brief description of the time series and the parent data set, basic options to download or refine the level of detail, and a link to the equivalent table in the Statistical Datasets interface.

Data-Planet has appeal for users in a wide array of academic and professional fields, particularly in the social sciences and business. The dual-access to Statistical Datasets and Statistical Ready Reference offers flexibility that will meet the needs of different types of users, from undergraduates through faculty researchers. Researchers and advanced students are the primary audience for Statistical Datasets, as the interface is powerful and complex. Librarians will also value having a single point of access for use in teaching information literacy and in answering reference questions involving statistical information.
releases and commits to having the most recent data available as close as possible to the time that new data is released from the source organizations.

Data-Planet has agreements with two discovery platforms, Ex Libris Primo and EBSCO Discovery. At this time Data-Planet has provided metadata to these platforms that links back to the approximately 180,000 datasheets in Statistical Ready Reference. Data-Planet does not have an agreement with Summon at this time, although Data-Planet President Matt Dunie mentioned that it is a possibility for the future.

**Critical Evaluation**

**CONTENT**

Data-Planet contains datasets from reputable producers that represent a broad range of geographic areas and topics. The selection criteria for inclusion considers five factors: customer requests, authority of the source organizations, quality of the dataset (documentation, sampling, data collection methods, and analysis procedures), fit within the scope of the 16 subject areas, and timeliness and regularity of data releases from the source organizations.

The 16 subject areas are: Agriculture & Food; Industry, Business, & Commerce; Banking, Finance, & Insurance; International Relations & Trade; Criminal Justice & Law Enforcement; Labor & Employment; Education; Military & Defense; Energy Resources & Industries; Natural Resources & Environment; Government & Politics; Population & Income; Health & Vital Statistics; Prices, Consumption, & Cost of Living; Housing & Construction; Transportation & Traffic.

There are approximately 70 source organizations listed at the time this review was written. About half are U.S. government agencies such as the Bureau of Labor Statistics, Census Bureau, and National Center for Education Statistics. Of datasets not published by the U.S. government, some come from other freely available sources, and some originate from source organizations to which Data-Planet pays royalties. International content is very strong, with datasets from major intergovernmental bodies including OECD, World Bank, United Nations, IMF, and EuroStat (from the European Commission). In addition to assuring that data originates from quality sources, subscribers can be confident in the processing of datasets which does not alter or “clean up” the data in any way. As new datasets are brought into the system, the programmers and statisticians map the data to standardized fields, to the extent possible within the limitations of the source material. The way that the data is collected and described by the source organization determines the extent to which it can be harmonized. Users can also be confident that they are using the most up-to-date information available, since Data-Planet does rolling content releases and commits to having the most recent data available as close as possible to the time that new data is released from the source organizations.

Data-Planet has agreements with two discovery platforms, Ex Libris Primo and EBSCO Discovery. At this time Data-Planet has provided metadata to these platforms that links back to the approximately 180,000 datasheets in Statistical Ready Reference. Data-Planet does not have an agreement with Summon at this time, although Data-Planet President Matt Dunie mentioned that it is a possibility for the future.

**STATISTICAL READY REFERENCE**

Statistical Ready Reference is the simpler of the two interfaces, and provides a searchable access point to 180,000 datasheets. The landing page offers a large basic search box with options to limit to a topic category, a place, and a year (Figure 1). Users can also jump to any of the 16 subject categories to browse. Each datasheet includes a single time series graph and descriptive information about the data and its parent dataset. Depending on the chart, users may have some basic options for customizing the view. When possible, users can select the granularity of geographic areas such as state or county-level, change the order or data points to ascending or descending order or alphabetical order (such as by state), or zoom in to a selected range of years. Each datasheet also has a short list of subject terms (on the right below the chart) which retrieve topically related datasheets. At the bottom of the page, users will find a formatted source citation with a DOI that returns to the static chart as it was created at the time of viewing. Data-Planet also uses stable URLs which go back to the chart as it stood when first accessed. There is a link at the bottom of every datasheet that is entered through a DOI or stable URL that generates an updated chart using the most recent available data. The Statistical Ready Reference datasheets all have links to View in Statistical Datasets which takes users to the other interface and provides the more complete functionality of that tool. Datasheets have standard output options of PDF, Excel, and CSV file (there are additional advanced
output options in the Statistical Datasheets interface). Figure 2 shows a sample datasheet.

Libraries serve users with a variety of needs, from basic to complex research, and the existence of dual interfaces is a testament to Data-Planet’s responsiveness to their customers. Statistical Ready Reference takes a more streamlined approach, there is much less to take in on a page, and the user is required to make very few decisions. It is a good complement to the Statistical Datasets interface which is evaluated in the next section.

STATISTICAL DATASETS
The Statistical Datasets interface is the flagship product, and former customers of Lexis-Nexis Statistical Datasets and ProQuest Statistical Datasets will recognize the three-panel layout. Data-Planet calls this layout I-C-Results, Indicator (left), Criteria (top) and Results (bottom). Figure 3 shows the landing page. The interface is built with HTML5 rather than Java, and works well on all standard browsers and almost any device. This interface was created with sophisticated researchers in mind, and does take a bit of time and persistence to learn how to use effectively.

BROWSING
The landing page begins by displaying a time series related to current events such as gas prices, unemployment rate, or national debt, in the In The News category of the indicator panel. Most beginning users will gravitate toward keyword searching but as users become more familiar with the database, they will appreciate the precision of browsing sources and subjects in the indicator panel. The + sign expands out subject and source hierarchies to four or five levels of depth. When the user selects an indicator it populates the criteria panel and the results panel with the default view, and options for further refinement appear at the top such as years, geographic limiters, and refinements that are unique to the dataset being viewed.

College enrollment data from the NCES, for example, can be viewed in aggregate for the USA, by year, by state, by institution, by gender, by race, and any combination of those criteria. Each time a refinement is selected, it updates the chart in the view window. Users can even select multiple choices within one criteria window (ctrl+click) to create a graph, for example, of the undergraduate enrollment over time of several universities for comparison. One minor shortcoming of the
Data-Planet Review Scores Composite: ★★★★☆ 1/8

The maximum number of stars in each category is 5.

<table>
<thead>
<tr>
<th>Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>★★★★</td>
</tr>
<tr>
<td>User Interface/Searchability</td>
<td>★★★</td>
</tr>
<tr>
<td>Pricing</td>
<td>★★★★ 1/2</td>
</tr>
<tr>
<td>Contract Options</td>
<td>★★★★★</td>
</tr>
</tbody>
</table>

Data-Planet is the largest single source of harmonized data available. Much of the content originates from freely available sources, but the value is enhanced by the single point of access across datasets with built-in visualization tools.

There are two distinct interfaces with different audiences in mind. The more advanced interface, Statistical Datasets, is sophisticated and takes some persistence to learn to use rather than being intuitive. The search function does not always lead to satisfying results, but it may be simply due to the nature of searching a huge number of datasets rather than the search algorithm.

The pricing is determined based on institutional FTE, and offers flexibility and choice of subscription or one-time purchase. The pricing appears fair, and a good value for institutions with data needs.

The contract is consistent with industry standards and specifically allows for use of data in course packs and in scholarly publications. The newly announced program for Data-Planet to serve as a data repository to help libraries solve problems in RDM (Research Data Management) is promising.

In some ways, searching data is different than searching textual information so one would expect searching to work differently. Despite the fact that Data-Planet records use 37 fields of metadata, the search experience doesn’t take advantage of that very deeply. The only refinement that accompanies the search box is Topic Category, and a few facets appear after a search is run: subject terms, source, level of detail, and category. On the whole, the search experience is less satisfying than some other aspects of Data-Planet’s design. The relevance ranking doesn’t always bring up the best results first, but of course this is true of any kind of keyword search. A keyword search for the terms “college enrollment Ohio” retrieved a table from the Census Bureau with the following title “B14004. Sex by College or Graduate School Enrollment by Type of School by Age for the Population 15 Years and Over from the American Community Survey Summary File: 5-Year Estimates Dataset shown in # Persons.” The title would certainly discourage some users (the title was supplied by the Census Bureau not by Data-Planet) but still, one would hope that the search would have led to the NCES dataset rather than Census Bureau’s American Communities Survey estimates. Changing the keyword choice from “college” to “university enrollment Ohio” brought up completely different results; the first page of hits were enrollment data from individual universities. Each one led to the appropriate dataset from NCES, with the dataset highlighted in the indicator panel. From here it was fairly easy to unselect a single institution and get back to the aggregate state-level data. A keyword search for the phrase “subprime mortgages” came up with a table called “Rate spread 3% to 4% from the Home Mortgage Loan Applications Dataset shown as count” pertaining to a single lender Macquarie Mortgages USA as the first result. For the average undergraduate user the search is more likely to overwhelm and confuse than to solve a problem, and even for the persistent user the search is not always effective.

Data-Planet has invested a great deal of time in creating LibGuides to help lead users to the best content more quickly. Presently the LibGuides number 246 and cover every state and country as well as a handful of topics such as oil production and international debt. This large effort signifies an acknowledgement that finding the right data isn’t always efficient by browsing or searching, and second, that Data-Planet is responding to library needs and is using a tool that is familiar to libraries and patrons with the goal of saving users’ time. The extensive collection of published guides is available at <http://data-planet.libguides.com/>.

The indicator panel at the left is also the menu from which to select multiple datasets to view simultaneously. This is where the power of data harmonization makes it possible to click the mouse a few times to make a chart that would otherwise require a great deal of effort in a spreadsheet or knowledge of specialized statistical software. It is useful to note that the editors have collected 15 frequently consulted key economic indicators for quick access and additional time series over lays that show periods of recession, democratic or republican presi dencies, and house or senate majority. Selecting multiple datasets is not without quirks however; only the first dataset that was selected can use the filters and refinements. It is not possible to use the criteria panel to refine more than one dataset. Instead the system detects the level of geographic detail common to all selections and automatically uses that to limit the criteria. It is also not obvious how to select more than one dataset and it’s not even clear that it’s possible to make multiple selections, but savvy users will figure out that ctrl+click makes it work.

**SEARCHING**

In some ways, searching data is different than searching textual information so one would expect searching to work differently. Despite the fact that Data-Planet records use 37 fields of metadata, the search experience doesn’t take advantage of that very deeply. The only refinement that accompanies the search box is Topic Category, and a few facets appear after a search is run: subject terms, source, level of detail, and category. On the whole, the search experience is less satisfying than some other aspects of Data-Planet’s design. The relevance ranking doesn’t always bring up the best results first, but of course this is true of any kind of keyword search. A keyword search for the terms “college enrollment Ohio” retrieved a table from the Census Bureau with the following title “B14004. Sex by College or Graduate School Enrollment by Type of School by Age for the Population 15 Years and Over from the American Community Survey Summary File: 5-Year Estimates Dataset shown in # Persons.” The title would certainly discourage some users (the title was supplied by the Census Bureau not by Data-Planet) but still, one would hope that the search would have led to the NCES dataset rather than Census Bureau’s American Communities Survey estimates. Changing the keyword choice from “college” to “university enrollment Ohio” brought up completely different results; the first page of hits were enrollment data from individual universities. Each one led to the appropriate dataset from NCES, with the dataset highlighted in the indicator panel. From here it was fairly easy to unselect a single institution and get back to the aggregate state-level data. A keyword search for the phrase “subprime mortgages” came up with a table called “Rate spread 3% to 4% from the Home Mortgage Loan Applications Dataset shown as count” pertaining to a single lender Macquarie Mortgages USA as the first result. For the average undergraduate user the search is more likely to overwhelm and confuse than to solve a problem, and even for the persistent user the search is not always effective.

Data-Planet has invested a great deal of time in creating LibGuides to help lead users to the best content more quickly. Presently the LibGuides number 246 and cover every state and country as well as a handful of topics such as oil production and international debt. This large effort signifies an acknowledgement that finding the right data isn’t always efficient by browsing or searching, and second, that Data-Planet is responding to library needs and is using a tool that is familiar to libraries and patrons with the goal of saving users’ time. The extensive collection of published guides is available at <http://data-planet.libguides.com/>.
ANALYTICAL TOOLS

The analytical tools that come standard in the Statistical Datasets interface are one of Data-Planet’s strengths. The Results panel has numerous options to customize data views. The default view shows a trend line for a selected time series, and buttons at the top can toggle to a map view, a pie chart, and a ranked bar chart. All options are not available for every time series; it depends on the format and content of the dataset. In the upper left of the results panel is an option to show/hide data panel on the right, which opens a mini-spreadsheet displaying the individual data points. There are further custom options to sort the columns and even run calculations in new columns, much like the functionality of a spreadsheet. The map view works for most data, since most datasets have clear geographical delineations, and users can customize the number of gradations on the map scale (up to eight) and the color choices.

Just as with the datasheets, each time series displays descriptive metadata, subject tags, DOIs, and citation information just below the chart image. Output formats for Statistical Datasets are: Excel, PDF, delimited text, SAS, SPSS, XML, Shapefile (GIS), and RefWorks direct export.

ADDITIONAL CONSIDERATIONS

Data-Planet can be characterized as a customer-centric organization. Customers are invited to request particular datasets to add to the platform. The staff is eager to offer demonstrations and is quick and thorough in responding to questions. One interesting new development in 2015 is that Data-Planet is beginning to work with libraries and institutions on RDM (Research Data Management) solutions. Data-Planet can ingest local datasets created by researchers at subscriber institutions. Access to the local data can be configured for private use by the researcher alone or made available to all users at the institution. It is too new to evaluate this service fully, but it seems promising at a time when colleges and universities are grappling with how to manage researchers’ data.

COMPARISONS

There are no products that directly compete with Data-Planet, but there are a number of resources specializing in access to statistical information that are available to libraries. There is considerable change and development happening in the market for data and analytics products at this time, with established firms and startups partnering, changing hands, and testing out new products. ProQuest Statistical Insight is a well-established and wide ranging statistical database but it is a very different kind of product. Statistical Insight indexes published statistical documents and reports, and offers full-text PDF downloads of these reports sources from a variety of producers, primarily government agencies. While ProQuest indexes at the table level, the data is fixed in static PDF documents. There is no customization possible or downloading of datasets, so users are not able to work with the data directly; PQ Statistical Insight has statistical information that has been digested and analyzed, and not for raw data. It is likely to work well for undergraduates who want statistical information, but who are not themselves doing any data manipulation or analysis. Most products in the data marketplace are not providing static data sources and are moving toward analytics, visualization, and customization. ProQuest International Datasets is a new product that presents the opportunity for users to customize and manipulate data, with a focus on demographic information about countries and cities of the world. It is a rebranding of a partner product called DataMarket, an Icelandic company that was recently purchased by a Swedish company, Qlik (headquartered in the U.S.), which specializes in data and analytics. There has been no word from ProQuest yet about what that acquisition means for the International Datasets product. Statista is a newer company providing access to statistical information, specializing in business, industry and marketing information. Statista does not concentrate on historical data and time series data, but on current data and trends. Within academia Statista is best suited for business, marketing and entrepreneurship students. It is available to libraries but the user base seems to be largely private companies and individuals. Social Explorer is a mapping tool that contains historical census data and demographic data down to the block level, displayed as a customizable map (not as downloadable datasets), and Policy Map is a cloud-based GIS product focused on micro-level data covering a broad array of topics, with a focus on current data. Each of these products has some aspect that overlaps with Data-Planet, but they all have different strengths and purposes. Visualization tools are becoming more popular and more sophisticated, and libraries should anticipate new developments and new products in the coming years. Overall, Data-Planet rates highly in comparison, with its broad coverage of subjects, focus on time series data, provision of raw data that can be downloaded or viewed using the analytical tools, and the overall number of data sources available.

Contract Provisions

The contract is consistent with industry standards and allows for use of data in course packs and in scholarly publications. Subscriptions include MARC records (or XML or text files) for approximately 150,000 records. Usage data for logins, searches, and requests are provided in Excel spreadsheets upon request, with monthly or annual periodization.

Authentication

Authentication is typically by IP range for unlimited simultaneous users and authentication through remote access servers is allowable.

Author’s References


About the Author

Jennifer Starkey is a Reference & Instruction Librarian at Oberlin College Library in Oberlin, Ohio, where she serves as liaison to the History, Economics, and German departments. She earned her MLS at the School of Information at the University of Michigan.