

Request for Information

Regional Open Data Portal

December 15, 2014

Please note updated due date of 1/20/2015 (changed 1/15/2015)

University of Pittsburgh Center for Social and
Urban Research

3343 Forbes Ave.

Pittsburgh, PA 15260

Confidentiality Statement

Your response to this Request for Information will be kept confidential, and will only be shared with individuals involved in the open data portal technology selection process. Our process may involve individuals not affiliated with the University of Pittsburgh, City of Pittsburgh, and Allegheny County.

Submission Details

Submission Deadlines

Electronic submissions in response to this Request for Information are strongly encouraged. Please submit as an email attachment addressed to Bob Gradeck, at rmg44@pitt.edu. Use the subject line "Submission for Regional Open Data Portal." Submissions are due by 4:00 pm eastern time January 19-20th, 2015. We will confirm receipt via email, however we encourage respondents to call if an e-mail receipt was not received. No submissions will be reviewed until the submission deadline has passed.

Please note: The University of Pittsburgh is closed December 24th through January 5th. No confirmations will be made during that time.

Submission Delivery Address

The delivery address to be used for all submissions is:

Bob Gradeck

University of Pittsburgh, University Center for Social and Urban Research

3343 Forbes Ave.

Pittsburgh, PA 15260

Voice: 412-624-9177

Email: rmg44@pitt.edu

Submission Questions and Clarifications

You may contact Bob Gradeck (contact information listed above) if you have any questions or require clarification on any topics covered in this Request For Information. Please contact him with questions no later than January 13. Questions will be answered in writing by January 15, 2015, and shared with all prospective respondents via the UCSUR Website (www.ucsur.pitt.edu), or at <http://bit.ly/1yA80pD>

Please note: The University of Pittsburgh is closed December 24th through January 5th. No responses to questions will be made over that time period.

Submissions

Electronic submissions in response to this Request for Proposal are encouraged. Please submit as an email attachment addressed to Bob Gradeck, at rmg44@pitt.edu. Use the subject line "Response to RFI: Regional Open Data Portal." Submissions are due by 4:00 pm eastern time January 19, 2015. We will confirm receipt via email, however we encourage respondents to call Bob Gradeck at 412-624-9177 if they do not receive an immediate reply. No submissions will be reviewed until the submission deadline has passed.

Please note: The University of Pittsburgh is closed December 24th through January 5th. Submissions made during that time frame will not receive a response until the University reopens.

Please adhere to the following document standards/layout:

- File format: Adobe Portable Document Format (.pdf) or MS Word-compatible (.doc or .docx)
- Submissions are limited to no more than 8 pages in length, not including a cover page. Appendices containing information about your company or organization can be included as a separate attachment, with no length limitations.
- Please use an 8 ½ x 11 inch page size, 12 point Arial font, 1.5 line spacing, 1 inch margins.

Please organize content using the following headers (underlined):

1. Basic Company Information: Basic information about your organization/company
2. Portal Features and Benchmarks: List data portal features available at the following schedule benchmarks - Assume work to start in February, 2015:
 - a. One month following the start of work
 - b. Three months following the start of work
3. Statement of Interest: Please indicate your interest in both working from a formal scope of services, and/or as an hourly consultant.
4. Responses to Questions: Please respond to our 13 questions by providing only the question number - no need to re-state the question. Where possible, try and keep responses to two paragraphs or less per question.
5. Appendices

Introduction and Executive Summary

The Regional Data Center (RDC) at the University of Pittsburgh's Center for Social and Urban Research is being created to support regional initiatives through the provision of timely and useful information and related services. Political fragmentation and the lack of effective public information systems in many communities provide an opportunity to build shared systems for managing the work of government, civic organizations, and academic institutions. Such systems allow for the establishment of effective data standards, provide publicly accessible and timely information in real-time, create a legal framework for information sharing, and increase the efficiency and effectiveness of local government. Shared systems and common standards will also increase opportunities for cross-community, cross-jurisdictional, and cross-sector collaboration. We want Pittsburgh to serve as a global model for collaborative data infrastructure to support research, analysis, decision making, and community engagement.

The RDC will be home to the Pittsburgh region's open data portal, and will begin to host datasets from various governments, public authorities, university research projects, along with non-profit and community organizations. The first users of the data portal include the City of Pittsburgh and Allegheny County. The RDC's portal is modeled on "federated" community data portals operating in Philadelphia and Colorado. Through this RFQ, we are seeking to launch an open data portal early in 2015.

What makes the RDC truly unique among open data initiatives is the level of support to be offered to data providers and data users to build an information ecosystem in the region. Services provided to data publishers include training and technical assistance in extracting data from internal systems, data transformation, and loading to the portal. Workshops will also be offered on metadata creation, data licensing, and data standards. The RDC will also promote the value of open data among potential data

providers, including government staff and elected officials. Support to data users include training in the use of the data portal, visualization tools, support to learning and volunteer communities, data user guides, technical assistance, and special events.

Overview & Background

An agile development philosophy will guide development of the open data portal, where the product will be improved over time by learning what features and capabilities are needed by users and stakeholders. We intend to build the capacity to manage the portal in-house using project staff (to be hired in early 2015), but are looking to accelerate the process of launching a viable open data portal by February, 2015 through this RFQ. While the federated portal will initially feature data from the City of Pittsburgh and Allegheny County, it must be designed to scale over time as more organizations share their information through the portal.

Through this RFI, we are not only looking to learn more about vendors/consultants that can help us establishing an open data portal. We are also seeking to learn more about how various open data portal technologies can be deployed to meet the goals and objectives of the RDC, and how these decisions will impact the structure and staffing of the project. We welcome your input into our approach to open data, and encourage you to offer advice and challenge our assumptions in your response.

Detailed Specifications

We are looking to include the following functionality in early iterations of our open data portal. While not all are essential at the launch date, we'd like to know which ones are possible at launch, and which are possible at a later date. In response to this RFQ, we'd like to know if it is feasible to: 1. provide all listed core functionality at launch, and 2. provide additional functionality as listed three months following the start of work.

Core functionality at launch (minimum viable product), 1 month following start of work

- Dataset keyword search
- Metadata integration
- Data license integration
- Ability to view offerings by an individual data source/publisher
- View data table in browser
- Bulk data download

Additional functionality 3 months following the start of work - implementation no later than two months following launch.

- API integration
- SQL / ODBC access
- Data visualization capability beyond tabular view (charts, maps) using a variety of data visualization tools and platforms (open source and proprietary)
- Ability to administer groups, and restrict dataset publication by group
- “Home” pages for each publishing organization
- Users can provide feedback through the site to data creators and other users to inform them about data errors and other feedback

While we are not yet asking for any specific pricing information, please indicate if you are able to 1. work as part of a formal contract tied to a scope of services (turnkey approach), or 2. on an hourly consulting basis, assisting in-house staff in preparing the open data portal for launch.

Additional Questions

In addition to a response related to configuring an open data site with the features listed above, we would welcome a brief response to each of the following questions. Please try and limit your responses to one or two paragraphs if at all possible.

1. *Please describe your company's specific expertise, capabilities, and qualifications in designing, developing, and configuring open data portals. Feel free to include links to completed projects.* In response to this question, we're asking specifically about your work relevant to open data portals. Appendices containing information about your company or organization can be included as a separate attachment, with no length limitations.
2. *What open data portal platform is a good choice for the Regional Data Center? How are we likely to view particular strengths and weaknesses of the portal technology you suggest we use? What communities should we view as a model deployment?*
3. We have the ability to host the site at the University of Pittsburgh's Network Operations Center, but are not required to use that option. *What hosting options should we consider, and what should we take into account beyond cost and reliability?*
4. *How challenging will it be for the RDC to implement server operating system and open data portal technology upgrades if the site is hosted in-house at the University? Are there any actions we should take into account to minimize the burden of these system updates?*

5. We're looking to hire staff for the RDC. *What types of skills and experience are necessary to 1. manage the portal's server environment, and 2. integrate various data visualizations with the portal?*

6. The Regional Data Resource Center intends to assist data providers with integrating Extract Transform Load (ETL) processes with the portal, including automated processes for loading data and protecting sensitive or personally-identifiable information. *What models exist for integrating ETL processes with the portal? What experience does your organization have in developing ETL processes?*

7. *If we adopt your portal technology, who owns the software, and are there limitations to the types of modifications we can make? Is the software open-source? Are there any ongoing licensing costs or user fees involved? Are we able to switch vendors or server environments without a penalty?*

8. We are very interested in measuring the performance of the RDC. *What types of use statistics are available through the data portal software?*

9. At scale, we expect the RDC open data platform to support visualizations that meet the needs of a variety of users. Some of these visualizations may be specifically developed for a single audience (e.g. one government department). Others may serve the public. We are interested in using a variety of open-source and proprietary data visualization tools, many of them directly integrated with the RDC data repository. *How have those using your platform integrated a variety of external data visualization tools? Are these tools directly integrated with the portal itself?* Please include examples.
10. *As most of our GIS departments use ESRI geodatabase file formats, what processes are needed to integrate geospatial data from sources using ESRI's file geodatabase and shapefile formats? How can your company help us organize and make geodata available in these formats for use?*
11. *What types of data format interoperability exists on the site? What common formats are accepted on import, and what formats are available upon export? Are there any limitations to file size (memory and record length), file formats, or incorporation of real-time data feeds?*
12. The RDC is currently evaluating a series of metadata schema. *Is there a particular one you prefer? Is it possible to customize/adapt a metadata schema of our choosing, or is a particular framework baked-in? How easy is it for a data provider to create metadata using your technology platform?*
13. *What level of effort is needed to design an effective, developer-friendly API using your technology?* Please share links to a few tools using APIs from your preferred portal technology.

Selection Criteria and Timeline

The selection of a consultant will be made by the University of Pittsburgh in close consultation with partners at the City of Pittsburgh and Allegheny County. UCSUR may also solicit advice from prospective members of the RDC's Advisory Board.

Based on responses received, we may choose to either solicit bids as part of a formal bid process, or to speak directly with qualified consultants.

Decision-making criteria for the project will include (in no particular order):

- Fit with overall RDC project goals and structure
- Vendor experience and capability
- Customizability
- Scalability
- Startup costs
- Operating Costs
- Functionality (w/ focus on user needs)
- Reliability
- Usability

After reviewing submissions, UCSUR may contact respondents to schedule a presentation, conversation, or request more information.

Our timeline is very aggressive, and we intend to move quickly and launch a minimally viable product before the end of March, 2015.