

NNIP AND OPEN DATA

Thursday afternoon, 1:15 p.m. – 3:00 p.m.

NNIP is concluding its activities supported by the John D. and Catherine T. MacArthur Foundation to explore the implications of the Open Data movement for NNIP partners locally and for the network as a whole. This movement and NNIP share the common goal of increasing the amount of relevant data that are available to residents and organizations, whether to hold governments accountable, to improve citizen services, or to inform community decision-making. Since we began this exploration together in Detroit in 2010, many additional local governments have committed to releasing data to the public. And more partners are playing important roles in promoting open data – from participating in code-a-thons to advising local governments on data strategy.

The purpose of this session is to help local partners consider how their organization and the NNIP network can productively engage with governmental and non-governmental advocates and practitioners of open data.

Seema Iyer from the Baltimore Neighborhood Indicators Alliance at the Jacob France Center-University of Baltimore will moderate the session and open with remarks based on their open data work. Kathy Pettit from the Urban Institute will then give a short review of our accomplishments during the MacArthur grant.

The first part of the panel will focus on open data policy and data release. Emily Shaw, National Policy Manager from the Sunlight Foundation, will present the evolution of their local policy guidelines, thoughts on addressing concerns about liability confidentiality, and the growing role of non-governmental organizations publishing data. As an example from within the network, Jeff Matson will relate the story of open data policy adoption in the Twin Cities area.

The second part of the session will cover what organizations are doing locally to leverage progressive open data policy to support community-based groups and benefit low-income neighborhoods. Dan O'Neil, the Executive Director of the Smart Chicago Collaborative, will describe the unique mix of missions and audiences of his organization, and how they connect technologists to community groups to address neighborhood challenges. Jeff Matson will conclude by describing the goals and structure of the new Twin Cities initiative CURATech (<http://www.cura-tech.org/>), a new civic technology incubator funded by the McKnight Foundation.

There will be time for discussion after each section of the session.

Questions for Discussion

- What are your local government's policies on open data (if any)? Are there ways to support coalitions to promote new or improved open data policies?
- What is the mix of organizations in your area that are working on open data policy or practice? Where do you see gaps?
- What are ways that NNIP partners can take advantage of the progress on open data to benefit low-income neighborhoods?
- What are ways that partners can help community have a greater voice in open data?
- What work related to open data should UI be doing on behalf of the network (e.g. communications, training, documenting NNIP partner examples)?

Ten Principles for Opening Up Government Information

<https://sunlightfoundation.com/policy/documents/ten-open-data-principles/>

1. **Completeness:** Datasets released by the government should be as complete as possible, reflecting the entirety of what is recorded about a particular subject. All raw information from a dataset should be released to the public, except to the extent necessary to comply with laws regarding the release of personally identifiable information. Metadata that defines the data should also be included, along with explanations for any calculated fields. Doing so will permit users to understand the data available and examine them at the greatest possible level of detail.
2. **Primacy:** Datasets released by the government should be primary source data. This includes the original information collected by the government, details on how the data was collected and the original documents recording the collection of the data. In this way, users can verify that information was collected properly and recorded accurately.
3. **Timeliness:** Datasets released by the government should be available to the public in a timely fashion. Whenever feasible, information collected by the government should be released as quickly as it is gathered and collected. Priority should be given to data whose utility is time sensitive. Real-time information updates would maximize the utility the public can obtain from this information.
4. **Ease of Physical and Electronic Access:** Datasets released by the government should be as accessible as possible, with accessibility defined as the ease with which information can be obtained, whether through physical or electronic means. Barriers to physical access include requirements to visit a particular office in person or requirements to comply with particular procedures (such as completing forms or submitting FOIA requests). Barriers to automated electronic access include making data accessible only via submitted forms or systems that require browser-oriented technologies (e.g., Flash, Javascript, cookies or Java applets). By contrast, providing an interface for users to download all of the information stored in a database at once (known as "bulk" access) and the means to make specific calls for data through an Application Programming Interface (API) make data much more readily accessible. (An aspect of this is "findability," which is the ability to easily locate and download content.)
5. **Machine readability:** Machines can handle certain kinds of inputs much better than others. For example, handwritten notes on paper are very difficult for machines to process. Scanning text via Optical Character Recognition (OCR) results in many matching and formatting errors. Information shared in the widely-used PDF format, for example, is very difficult for machines to parse. Thus, information should be stored in widely-used file formats that easily lend themselves to machine processing. (When other factors necessitate the use of difficult-to-parse formats, data should also be available in

machine-friendly formats.) These files should be accompanied by documentation related to the format and how to use it in relation to the data.

6. **Non-discrimination** : "Non-discrimination" refers to who can access data and how they must do so. Barriers to use of data can include registration or membership requirements. Another barrier is the uses of "walled garden," which is when only some people can obtain the data. Full non-discriminatory access to data means that any person can access the data at any time without having to identify or justify him/herself.

7. **Use of Commonly Owned Standards**: Commonly owned (or "open") standards refers to who owns the format in which data is stored. For example, if only one company manufactures the program that can read a file where data is stored, access to that information is dependent upon use of the company's processing program. Sometimes that program is unavailable to the public at any cost, or is available, but for a fee. For example, Microsoft Excel is a fairly commonly-used spreadsheet program which costs money to use. Freely available alternative formats often exist by which stored data can be accessed without the need for a software license. Removing this cost makes the data available to a wider pool of potential users.

8. **Licensing**: The imposition of "Terms of Service," attribution requirements, restrictions on dissemination and so on acts as barriers to public use of data. Maximal openness includes clearly labeling public information as a work of the government and available without restrictions on use as part of the public domain.

9. **Permanence**: The capability of finding information over time is referred to as permanence. Information released by the government online should be sticky: It should be available online in archives in perpetuity. Often times, information is updated, changed or removed without any indication that an alteration has been made. Or, it is made available as a stream of data, but not archived anywhere. For best use by the public, information made available online should remain online, with appropriate version-tracking and archiving over time.

10. **Usage Costs**: One of the greatest barriers to access to ostensibly publicly-available information is the cost imposed on the public for access--even when the cost is de minimus. Governments use a number of bases for charging the public for access to their own documents: the costs of creating the information; a cost-recovery basis (cost to produce the information divided by the expected number of purchasers); the cost to retrieve information; a per page or per inquiry cost; processing cost; the cost of duplication etc. Most government information is collected for governmental purposes, and the existence of user fees has little to no effect on whether the government gathers the data in the first place. Imposing fees for access skews the pool of who is willing (or able) to access information. It also may preclude uses of the data that in turn generates business growth and tax revenues.



Sunlight Foundation

<http://sunlightfoundation.com/>

<http://sunlightfoundation.com/policy/local/>

The Sunlight Foundation is a nonpartisan nonprofit founded in 2006 that uses the power of the Internet to catalyze greater government openness and transparency. They do so by creating tools, open data, policy recommendations, journalism and grant opportunities to dramatically expand access to vital government information to create accountability of their public officials. Their vision is to use technology to enable more complete, equitable and effective democratic participation. Their overarching goal is to achieve changes in the law to require real-time, online transparency for all government information, with a special focus on the political money flow and who tries to influence government and how government responds. And, while their scope began with only a focus on the U.S. Congress, they now are defining open government on the local, state, federal and international level.

They believe that information is power, or, to put it more finely, disproportionate access to information is power. Indeed, they are committed to improving access to government information by making it available online, indeed redefining "public" information as meaning "online."

Local Policy Initiative:

The Sunlight Foundation supports government openness as a fundamental democratic principle. Open data policies represent a critical means for achieving that objective. The Foundation launched a new initiative in 2013 to look at municipal data disclosure and related state issues. They actively monitor and support existing open data and accountability initiatives, research case studies, support best practices and propose new policies.

The Sunlight Foundation supports the full range of benefits that can come from local implementations of open data policies. However, in line with our focus on transparency, they are especially interested in seeing open data policies contribute to better understanding of the money that that could influence government decision-making, improved knowledge of government services and transactions, and improved access to government processes and decision-makers for all citizens.

Their local government policy page has helpful resources related to crafting policy, navigating the current local #opengov landscape, and more.

Smart Chicago

<http://www.smartchicagocollaborative.org/>

Smart Chicago is a civic organization devoted to improving lives in Chicago through technology. We work on increasing access to the Internet, improving skills for using Internet, and developing meaningful products from data that measurably contribute to the quality of life of residents in our region and beyond.

Smart Chicago was founded in part by our municipal government and nurtured by some of its most venerable institutions. The founding partners are the City of Chicago, the John D. and Catherine T. MacArthur Foundation, and The Chicago Community Trust. As a funding collaborative, the organization helps to bring together municipal, philanthropic, and corporate investments in civic innovation.

They have a host of current projects and partnerships, and are actively seeking to connect ideas and resources in all areas of philanthropy in Chicago. Here's an incomplete list of the things they care about:

- Open: they focus on technologies and processes that are open
- Access: This is their genesis. They were born to address the digital divide and they administer a significant amount of federal broadband money around helping people gain access to the Internet. We will care about this until the problem is solved
- Chicago: This is their middle name— the focus of their money, their time, and their passion. Since Chicago makes up the bulk of Cook County and a significant portion of the State of Illinois, they care about how they all work together to solve common problems
- Everyone: That means they spend their money in ways that scale, with a focus on the entire city. They care about the needs of low and middle-income people. Forgetting no one, always pushing for inclusion, serving all. Having said that, everyone means everyone— in order to have a meaningful network, they need to include all
- Markets: They mean this in both the University of Chicago way (markets drive innovation, economics are rational and knowable, and they can understand humans via markets) and the Chicago technology venture capital / investment way (sustainable models matter more than cool ideas)
- Experience: They care about how people experience their city, their data, their loved ones. They seek to connect people to services, people, and things in ways that make sense to them and generate new activity that helps others in ways they couldn't imagine if they sat down to plan it out