

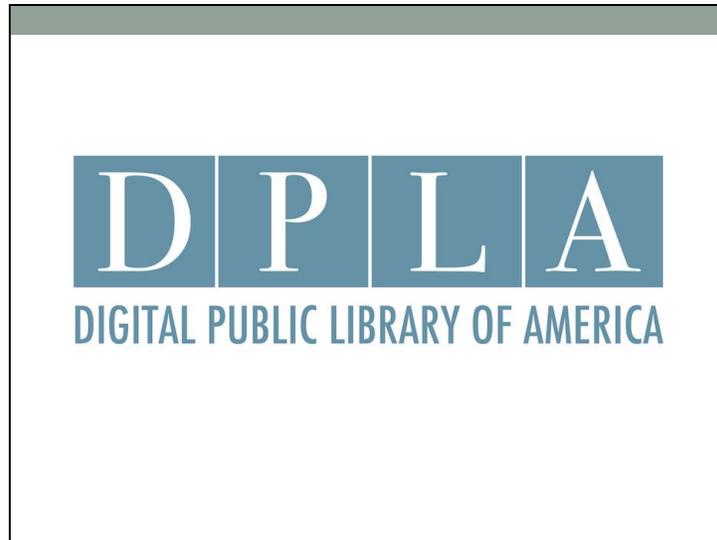
Slide 1

EMPIRE STATE DIGITAL NETWORK

New York Service Hub of the DPLA

Kerri Willette, ESDN Manager
Westchester Library System, November 20, 2014

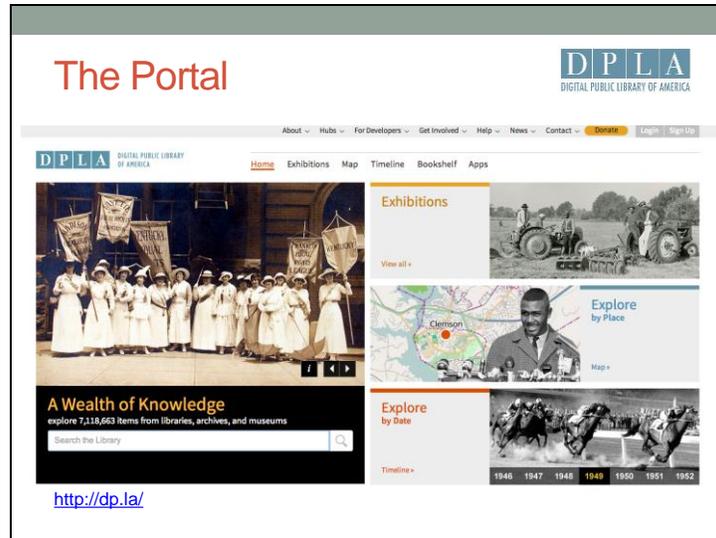
Slide 2



The Digital Public Library of America (DPLA) started as a concept back in 2010. The initial concept was to create an open, distributed network of online resources from libraries, universities, archives and museums in the US.

Just a couple years later, in April of 2013 DPLA launched it's site and services.

Slide 3

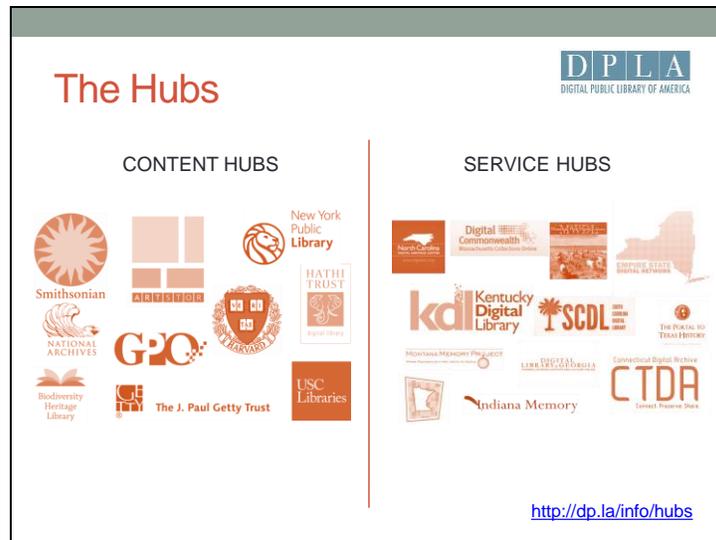


This is a screen grab of DPLA's portal. This is a publicly available website that provides:

- Access to over 7.1 million items – from libraries, archives, and museums around the US
- Options to browse by timeline, map, format, topic – they even have a virtual bookshelf
- Also have digital exhibitions curated by content partners and staff

In addition to this web interface, DPLA also provides a free and open Application Programming Interface (API) – a platform for developers and programmers. The API encourages creative reuse of this shared content.

Where exactly does the content come from?



Where do those 7.1 million items come from?

DPLA relies on a national network of partners that coordinate sharing of metadata from institutions across the US. Refer to this as the “Hub Model”.

Two types of Hubs: Content vs Services

Content hubs are large digital libraries, museums or repositories that work directly in a 1 on 1 relationship with the DPLA. To qualify as a content hub a partner institution must provide more than 250,000 unique metadata records to the DPLA.

On the right side, you see the Service Hubs. So, what are service hubs?

Service Hubs

“DPLA service hubs are state, regional, or other collaborations that host, aggregate, or otherwise bring together digital objects from libraries, archives, museums, and other cultural heritage institutions.”

<http://dp.la/info/hubs>

This is the exact wording from DPLA’s website describing what service hubs are.

For obvious reasons, DPLA cannot accommodate a 1 to 1 working relationship with every cultural heritage institution in each state like they do with the large content providers.

Service hubs are a crucial aspect of the scalability (and sustainability) of DPLA’s model.

DPLA relies on states and regions to coordinate centralized data hubs. And what they need is for each service hub to provide a single access point for DPLA to grab all of the contributor data from that state or region. And that’s the primary goal of the service hub. To bring all of that data in the state into a single stream for DPLA to grab.

So, let’s talk about how this is shaping in the state of NY...

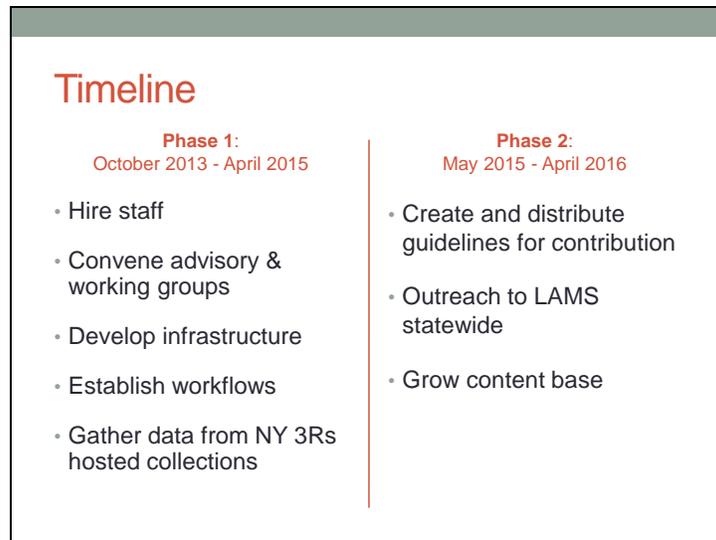


To begin, a little background.

As I mentioned, DPLA's website launched publically in April of 2013.

Right around that time - in May 2013 - the NY 3Rs association released a report called I2NY: Envisioning an info structure for NY. The report was the result of a 3Rs strategic planning initiative and it outlined 6 statewide collaborative initiatives. One of the 6 initiatives named was the creation of a New York service hub for the DPLA.

It was decided at this meeting that since the 3R's already play a key role in providing access to digital collections in NY, it made sense that the councils should coordinate and facilitate a statewide service hub. It was agreed then that Metropolitan New York Library Council (METRO) was best positioned to host the hub which so it was agreed that METRO would host and administer the hub in collaboration with the other 8 regional councils. METRO's board graciously agreed to provide a large portion of support for the hub and we were off and running.



One outcome from the initial project plan for ESDN was a 3 year timeline. The plan was divided into two phases.

As you can see, we are well on our way into Phase 1 which has focused primarily on laying the necessary groundwork to get the hub up and running. Phase 2 begins this coming May. That's when we will begin to open up participation more broadly throughout the state.

Staff

- Project Manager, Kerri Willette
- Metadata Specialist, Chris Stanton
- Technology Specialist, John Mignault

Three headshots of staff members are displayed in a row. From left to right: Kerri Willette, a woman with short blonde hair wearing a dark top; Chris Stanton, a man with short brown hair wearing a blue button-down shirt; and John Mignault, a man with short brown hair and glasses wearing a light blue button-down shirt.

Kerri Willette was hired in March 2014 as the project manager for ESDN. Then our metadata specialist, Chris Stanton joined us from the National Library of New Zealand in June. And, our Technology Specialist position, John Mignault started in July and he came to us from the New York Botanical Gardens in the Bronx.

Governance

- Advisory committee
- Metadata working group
- Regional outreach group

A photograph showing a group of approximately ten people seated around a large wooden conference table in a meeting room. They appear to be engaged in a discussion or presentation. A laptop is open on the table, and a screen in the background displays some information. The room has a whiteboard and a window in the background.

So in addition to hiring staff, one of the first steps toward putting a hub in place was establishing a governance structure.

Currently, governance of ESDN is structured in these three groups:

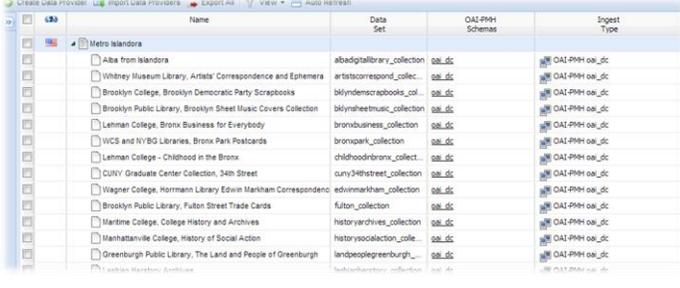
Advisory Committee is a group of stakeholders from all over the state. They advise ESDN on policy-level matters on things like strategic direction, program development, partnerships. This group had its first meeting on May 15th, 2014 and now meets quarterly.

Metadata Working group is more hands on in their work with us. This group is made up of content contributors and right now they're helping us iron out our workflows around data aggregation, but going forward they will provide guidelines and documentation to other contributors on matters like creating shareable metadata and complying to bibliographic standards. That kind of thing. The metadata group met for the first time this past July, and will also be meeting quarterly.

The **Regional Outreach group** is a standing group made up of 1 representative from each of the NY3Rs councils. This group will serve as ESDN's primary contact in each of the regions. When you begin hearing more about how to contribute to ESDN in phase 2, that information will be coming to you from your council's regional liaison to this group. Our first formal meeting of the regional outreach group was November 19th in Syracuse.

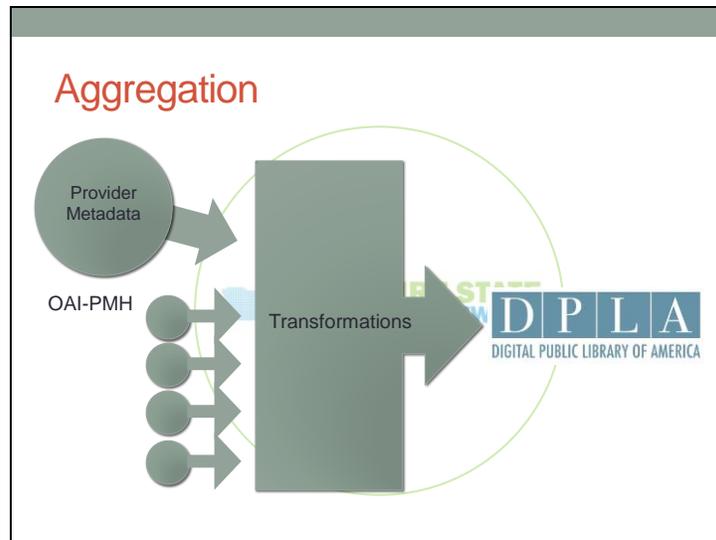
Technology & Workflows

- Software implementation
- Metadata ingest and remediation
- Sharing with DPLA



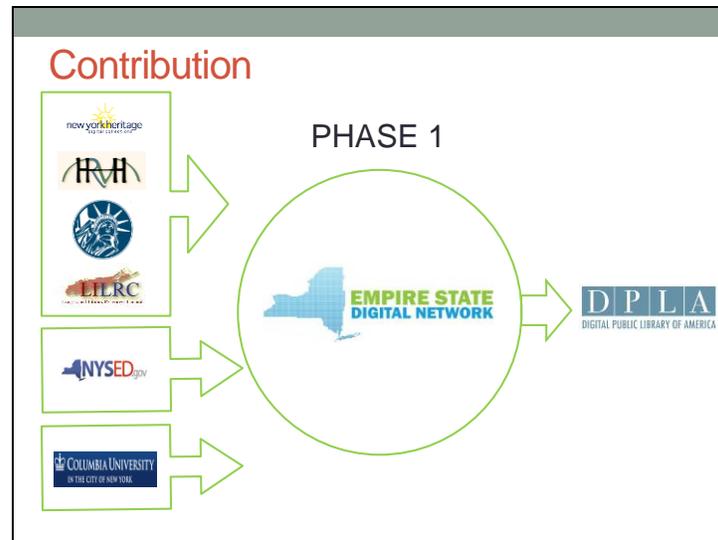
Name	Data Set	OAI-PMH Schemas	Ingest Type
Alba from Isadora	albadigitallibrary_collection	oai_dc	OAI-PMH oai_dc
Whitney Museum Library, Artists' Correspondence and Ephemera	artistscorrespond_collec...	oai_dc	OAI-PMH oai_dc
Brooklyn College, Brooklyn Democratic Party Scrapbooks	bklyndemscrabbooks_col...	oai_dc	OAI-PMH oai_dc
Brooklyn Public Library, Brooklyn Sheet Music Covers Collection	bklynshetmusic_collection	oai_dc	OAI-PMH oai_dc
Lehman College, Bronx Business for Everybody	bronxbusiness_collection	oai_dc	OAI-PMH oai_dc
WCS and NYBG Libraries, Bronx Park Postcards	bronxpark_collection	oai_dc	OAI-PMH oai_dc
Lehman College - Childhood in the Bronx	childhoodinbronx_collect...	oai_dc	OAI-PMH oai_dc
CUNY Graduate Center Collection, 34th Street	cuny34thstreet_collection	oai_dc	OAI-PMH oai_dc
Wagner College, Horrmann Library Edwin Markham Correspondence	edwinmarkham_collection	oai_dc	OAI-PMH oai_dc
Brooklyn Public Library, Fulton Street Trade Cards	fulton_collection	oai_dc	OAI-PMH oai_dc
Maritime College, College History and Archives	historyarchives_collection	oai_dc	OAI-PMH oai_dc
Manhattanville College, History of Social Action	historysocialaction_colle...	oai_dc	OAI-PMH oai_dc
Greenburgh Public Library, The Land and People of Greenburgh	landpeoplegreenburgh...	oai_dc	OAI-PMH oai_dc

That was staff and governance, let's talk a little bit about the technology, infrastructure and process we need to put into place to make all of this fly.



The goal is to harvest metadata from contributors, normalize it into a single stream, and then share that single stream of data with the DPLA.

As you might imagine this can be somewhat complex. Each digital collection system (like CONTENTdm or Islandora or Collective access) has its own protocols and quirks. Also, each local institution may follow slightly different metadata schemas and standards depending on local needs and practices. Our job is to ameliorate those differences between data sets so that we're delivering one consistent package of data to the DPLA.



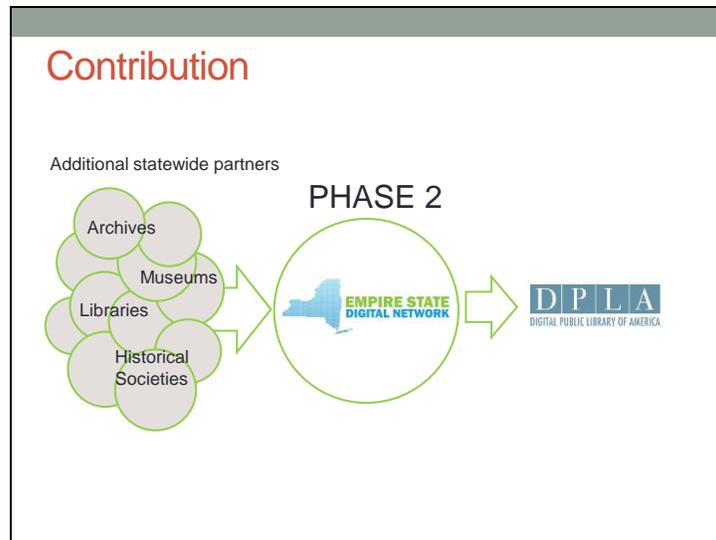
But for now – let’s go back to the project plan.

Phase 1 of our project timeline is really focused on working out the infrastructure and workflows necessary to make large scale, statewide aggregation feasible – and not just feasible but scalable and sustainable. Well, we couldn’t do that in theory. We needed some guinea pigs.

Since NY 3R’s already play a key role in providing access to digital collections throughout the state, it made sense ESDN would work directly with collection coordinators of existing hosted, regional/collaborative projects in the state.

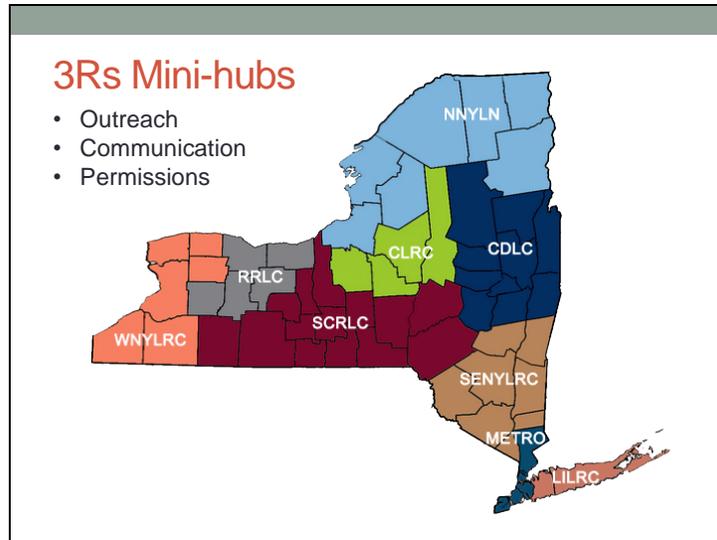
Additionally, we roped in the State Library and the State Archives and we wanted to work with at least one large, private institution in the state so we’ve also partnered with Columbia University for this first round.

These organizations agreed to provide us with sample data from their systems so that we can work through our technical processes – but even more importantly they’ve helped us establish our workflows for contribution. And to be clear – this has been a bit of trial by fire for all of us. Requiring much patience and good humor from all involved.



That's when we'll begin recruiting from outside that initial cohort. And we will absolutely be relying on regional outreach reps at that time to be our primary touch point with institutions and potential contributors in each region.

We do have a plan that we think is scalable, and phase 2 is just around the corner.



As you can imagine, there are simply too many individual organizations and potential providers in the state for it to make sense for our staff of 3 to coordinate communication and outreach to all of the new contributors going forward. Luckily, we won't need to!

This is the NY 3Rs councils map.

Here's the plan. In phase 2 of the ESDN project plan, all communication and contact with potential and existing contributors will be coordinated through the NY3R's regional councils – effectively, ESDN will be relying on the NY3R's councils as DPLA mini-hubs. Each council has appointed an ESDN Liaison (for you all in WNYLRC that representative will be Janelle).

Each liaison will be the primary contact for ESDN within a given region and will be taking the lead to communicate with contributors. I mentioned much earlier on in the presentation that our Regional Outreach Group will be meeting at the end of this month. That group is standing group that is made up of all the council liaisons to ESDN. That group will basically be the face of ESDN and your primary contact if and when you want to get involved.

What's next

Phase 1

- Initial ingest to DPLA in January 2015
- Ramp up for Phase 2 – publicity, documentation, clear lines of communication

Phase 2

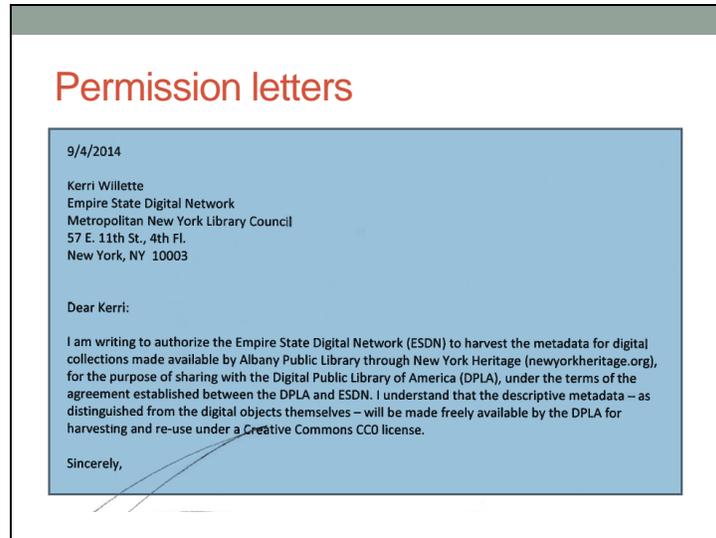
- Continue adding regionally hosted content
- Begin recruiting non-hosted partners
- Continue to grow New York content in DPLA

What should I do to get ready?

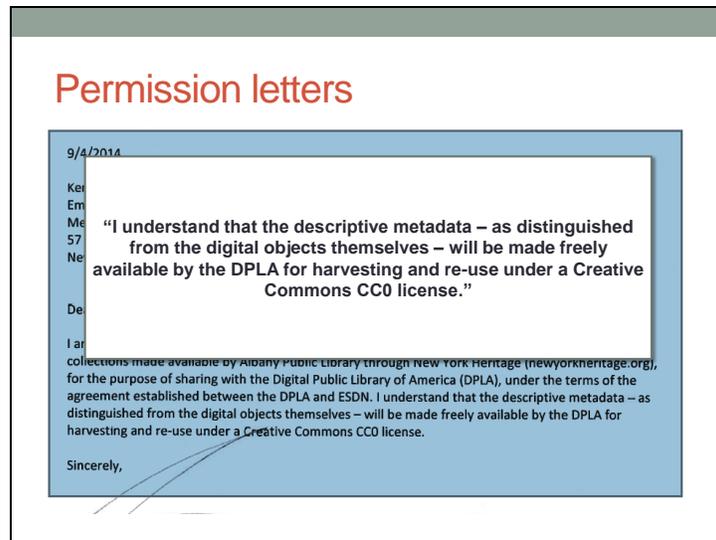
Metadata

- The more consistent your data is within a given collection, the easier it is for us to normalize and share.
- Follow standards consistently.
- Read a bit about shareable metadata.
- Review your rights statements.

Here are some basic principles. If you can address some of these things now, it will make the whole processes much easier and may even increase the click backs you receive from the DPLA website after you contribute your data.

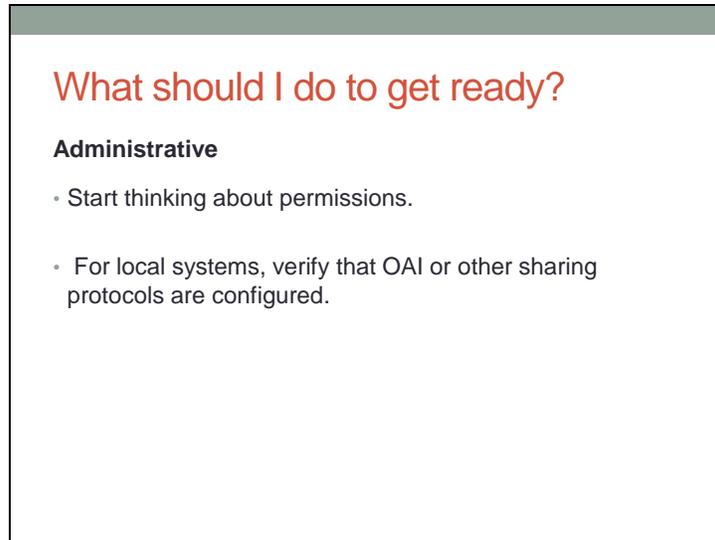


One of the key pieces of our contribution workflow which will be coordinated through the regional liaisons will be the acquisition of permissions letters from each contributing institution.



A Creative Commons CC0 license effectively places your metadata in the public domain. All metadata contributed to DPLA is available for download and reuse by anyone for any purpose.

One of DPLA's goals is to promote innovation and development within the cultural heritage community. Their commitment to open access goes beyond just making metadata searchable through their interface, it also includes making it available to developers and programmers to create new and transformative uses of our collective cultural heritage. This can only work if all of the metadata contributed to DPLA is unencumbered by copyright restrictions.



What should I do to get ready?

Administrative

- Start thinking about permissions.
- For local systems, verify that OAI or other sharing protocols are configured.

Again, these are things you can think about starting now, but once ESDN is fully functional and we begin open contributions more broadly, your council liaison will have information and resources for you to help get you step-by-step through the contribution process.

Follow Along

- Project website: <http://metro.org/esdn/>
- Follow: @NYSdigital

Kerri Willette
ESDN Manager
kwillette@metro.org