

# DATA CURATION NETWORK

## Curation + Community: About the Data Curation Network

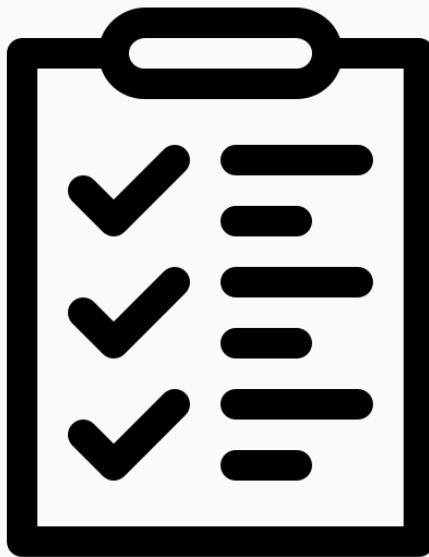


Mikala Narlock, Director, Data Curation Network

Presentation to  
Landesinitiative für Forschungsdatenmanagement [fdm.nrw](https://www.fdm.nrw)  
May 2023

# Agenda

- Why data curation?
- Overview of the DCN
  - How we were started
  - How we're structured
  - How we're funded



# DATA CURATION NETWORK



## Mission

---

Trusted, community-led  
network of curators  
advancing open research  
by making data

**Ethical. Reusable. Better.**

# Why data curation?

Research data have value beyond their original purpose.

But....

# Data can be messy, incomprehensible, and lack context

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	0.296	-0.55	-0.243	-3.04	-2.885	0.071	1.847	-0.521	0.107	-3.528	-3.089	2.45	-1.46	2.404	-0.857	0.354	-0.923	1.487	-2.06
2	0.354	-0.487	-0.193	-2.974	-2.933	-0.029	1.951	-0.506	0.061	-3.643	-3.162	2.131	-1.438	2.482	-0.796	0.435	-1.075	1.448	-2.03
3	0.438	-0.449	-0.099	-2.979	-2.901	-0.09	2.051	-0.501	-0.005	-3.647	-3.34	1.713	-1.428	2.519	-0.681	0.548	-1.258	1.372	-2.10
4	0.519	-0.431	-0.018	-3.042	-2.831	-0.13	2.107	-0.452	-0.027	-3.616	-3.42	1.322	-1.434	2.535	-0.564	0.624	-1.438	1.351	-2.22
5	0.696	-0.37	0.023	-3.065	-2.832	-0.187	2.202	-0.415	-0.072	-3.648	-3.504	0.831	-1.403	2.529	-0.467	0.718	-1.651	1.315	-2.30
6	0.939	-0.332	0.083	-3.089	-2.815	-0.233	2.314	-0.342	-0.119	-3.603	-3.648	0.38	-1.333	2.551	-0.367	0.773	-1.856	1.243	-2.39
7	1.188	-0.295	0.171	-3.08	-2.753	-0.295	2.431	-0.197	-0.186	-3.598	-3.619	-0.07	-1.272	2.594	-0.263	0.867	-1.978	1.137	-2.48
8	1.503	-0.284	0.279	-3.129	-2.746	-0.363	2.52	-0.116	-0.298	-3.487	-3.5	-0.608	-1.175	2.591	-0.121	0.904	-2.033	1.019	-2.60
9	1.826	-0.288	0.36	-3.183	-2.743	-0.496	2.59	-0.012	-0.316	-3.318	-3.456	-0.989	-1.077	2.574	0.034	0.946	-2.054	0.915	-2.68
10	2.153	-0.289	0.369	-3.162	-2.632	-0.615	2.653	0.197	-0.345	-3.249	-3.388	-1.33	-0.985	2.529	0.16	0.989	-2.125	0.805	-2.75
11	2.59	-0.244	0.359	-3.205	-2.51	-0.761	2.761	0.412	-0.416	-3.204	-3.296	-1.58	-0.896	2.494	0.318	0.99	-2.223	0.734	-2.83
12	2.97	-0.196	0.319	-3.218	-2.463	-0.944	2.933	0.643	-0.421	-3.143	-3.089	-1.746	-0.836	2.486	0.426	1.002	-2.274	0.687	-3.00
13	3.269	-0.222	0.297	-3.148	-2.454	-1.045	3.051	0.904	-0.356	-2.983	-2.829	-1.813	-0.74	2.485	0.58	1.003	-2.314	0.659	-3.05
14	3.512	-0.266	0.274	-3.157	-2.429	-1.147	3.119	1.116	-0.286	-2.783	-2.595	-1.927	-0.659	2.505	0.75	0.954	-2.351	0.609	-3.15
15	3.684	-0.271	0.289	-3.214	-2.396	-1.255	3.052	1.222	-0.227	-2.627	-2.292	-2.081	-0.595	2.456	0.915	0.918	-2.353	0.549	-3.19
16	3.824	-0.275	0.233	-3.289	-2.4	-1.262	2.996	1.39	-0.16	-2.475	-2.019	-2.286	-0.449	2.324	1.062	0.856	-2.367	0.54	-3.25
17	3.889	-0.294	0.186	-3.295	-2.303	-1.306	2.961	1.545	-0.083	-2.293	-1.825	-2.461	-0.368	2.243	1.196	0.764	-2.4	0.463	-3.26
18	3.896	-0.295	0.158	-3.289	-2.266	-1.383	2.93	1.645	-0.095	-2.195	-1.606	-2.573	-0.234	2.222	1.344	0.701	-2.392	0.432	-3.29
19	3.838	-0.283	0.152	-3.286	-2.273	-1.352	2.876	1.615	-0.074	-2.086	-1.385	-2.672	-0.063	2.182	1.556	0.622	-2.289	0.405	-3.24
20	3.712	-0.338	0.139	-3.328	-2.23	-1.302	2.778	1.637	-0.007	-1.971	-1.328	-2.759	0.078	2.144	1.789	0.566	-2.074	0.358	-3.25
21	3.526	-0.363	0.125	-3.387	-2.198	-1.275	2.604	1.624	-0.019	-1.814	-1.35	-2.817	0.217	2.032	1.996	0.562	-1.771	0.293	-3.33
22	3.343	-0.393	0.112	-3.466	-2.107	-1.235	2.412	1.645	-0.008	-1.744	-1.385	-2.886	0.373	1.929	2.219	0.573	-1.502	0.225	-3.38
23	3.17	-0.397	0.04	-3.434	-1.975	-1.199	2.27	1.612	0.025	-1.692	-1.416	-2.868	0.528	1.793	2.473	0.514	-1.197	0.207	-3.46
24	3.011	-0.384	-0.04	-3.374	-1.829	-1.116	2.122	1.604	0.014	-1.607	-1.452	-2.679	0.624	1.66	2.757	0.527	-0.949	0.173	-3.46
25	2.758	-0.369	-0.067	-3.294	-1.669	-1.005	1.963	1.586	-0.099	-1.53	-1.546	-2.495	0.732	1.585	3.029	0.597	-0.721	0.163	-3.44
26	2.467	-0.313	-0.129	-3.247	-1.425	-0.874	1.744	1.469	-0.191	-1.427	-1.618	-2.295	0.804	1.482	3.248	0.663	-0.519	0.17	-3.39
27	2.172	-0.267	-0.23	-3.282	-1.103	-0.738	1.466	1.408	-0.267	-1.316	-1.674	-2.177	0.827	1.373	3.46	0.756	-0.4	0.178	-3.32
28	1.897	-0.224	-0.304	-3.319	-0.821	-0.656	1.237	1.361	-0.354	-1.228	-1.716	-2.019	0.877	1.216	3.548	0.872	-0.293	0.187	-3.32
29	1.643	-0.174	-0.372	-3.24	-0.6	-0.544	1.066	1.295	-0.409	-1.167	-1.694	-1.944	0.995	1.05	3.599	0.934	-0.215	0.234	-3.27
30	1.391	-0.133	-0.456	-3.215	-0.378	-0.465	0.91	1.237	-0.462	-1.151	-1.73	-1.808	1.111	0.924	3.666	1.023	-0.194	0.243	-3.18
31	1.152	-0.131	-0.533	-3.195	-0.257	-0.381	0.792	1.159	-0.466	-1.102	-1.711	-1.622	1.162	0.819	3.66	1.102	-0.133	0.244	-3.08
32	0.893	-0.159	-0.614	-3.167	-0.204	-0.363	0.65	1.048	-0.518	-1.1	-1.734	-1.394	1.251	0.679	3.67	1.171	-0.044	0.278	-3.01
33	0.635	-0.144	-0.643	-3.13	-0.092	-0.352	0.517	0.933	-0.579	-1.149	-1.763	-1.301	1.317	0.551	3.664	1.223	0.043	0.266	-2.93
34	0.415	-0.151	-0.681	-3.169	-0.075	-0.292	0.429	0.861	-0.619	-1.156	-1.785	-1.176	1.403	0.405	3.572	1.291	0.119	0.318	-2.86

# Digital file formats are constantly at risk



# What is data curation?

“Data curation is the active and ongoing management of data through its lifecycle and interest and usefulness to scholarship, science, and education; curation activities enable data discovery and retrieval, maintain quality, add value, and provide for re-use over time.”

- University of Illinois' Graduate School of Library and Information Science



# What is data curation?

“Data curation is akin to work performed by an art or museum curator. Through the curation process, data are organized, described, cleaned, enhanced, and preserved for public use, much like the work done on paintings or rare books to make the works accessible to the public now and in the future.”

- ICPSR (Inter-university Consortium for Political and Social Research)

# Levels of Curation

Distributed as deposited

Level 0

## Record Level Curation

perform brief metadata checks for increased **findability (F)**



## File Level Curation

review files arrangement and perform file format conversions for increased **accessibility (A)**



## Documentation Level Curation

review documentation and request/add missing information for increased **reusability (R)**

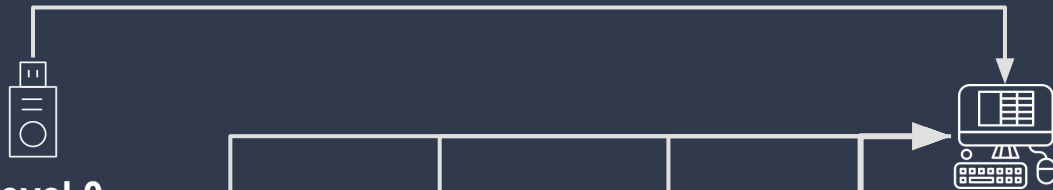


## Data Level Curation

open files and review data contents and may annotate /edit the data for accuracy or **interoperability (I)**



Level 4



# Who assists in curation?

Data curation involves

- Automated curation actions taken by **repositories**
- Decisions and actions taken by the **curator**
- Engagement with the **researcher**

[Definitions of Data Curation Activities used by the DCN](#)

# Overview of the DCN



Ethical. Reusable. Better.

**DATA CURATION NETWORK**

[datacurationnetwork.org](http://datacurationnetwork.org)



## Curation



Curate data as a cross-institutional network of nearly 50 individual experts



## Education



Offer professional development opportunities for an emerging data curator professional community



## Primers



Create and openly share data curation best practices



## Research



Support informal research teams addressing a specific topic: big data, human subjects, racial justice...

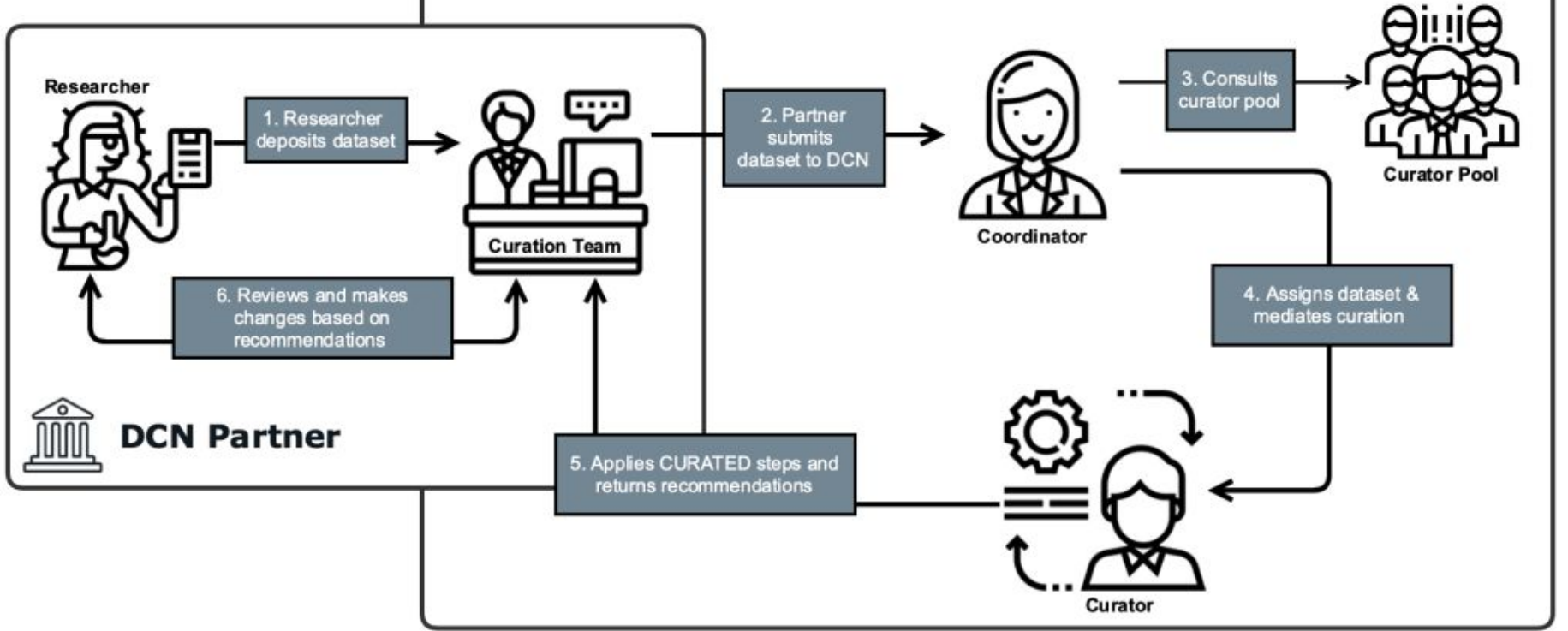


## Community



Build community at annual events for discussion, training, and networking.

# DATA CURATION NETWORK



# The CURATE(D) Steps

- C** **Check** files and read documentation.
- U** **Understand** the data (or try to), if not...
- R** **Request** missing information or changes.
- A** **Augment** metadata for findability.
- T** **Transform** file formats for reuse.
- E** **Evaluate** for FAIRness.
- (D)** **Document** your curation activities



[z.umn.edu/curate](https://z.umn.edu/curate)



# Data Curation Primers



Grant # RE-85-18-004018

- Concise and iterative resources
- 30 primers released so far on Github!

## DATA CURATION NETWORK

### Confocal Microscopy Data: A Primer for Curators

#### Executive summary

Topic	Description
File Extensions	.ism (Zeiss - proprietary) <sup>1</sup>
	.dzi (Zeiss - proprietary)
	.nd2 (Nikon NIS-Elements 2 - proprietary)
	.lif (Leica - proprietary)
	.oib (Olympus - proprietary)
MIME Type	.zip
	.tiff (open source - recommended for archiving) <sup>2</sup>
Structure	Image/tiff

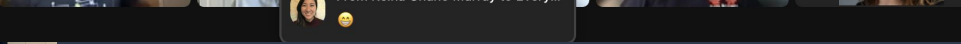
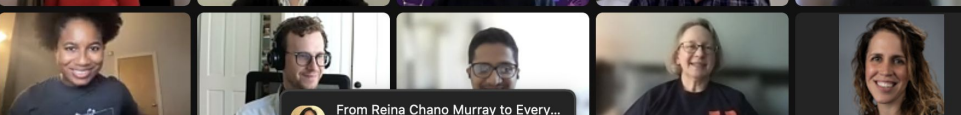
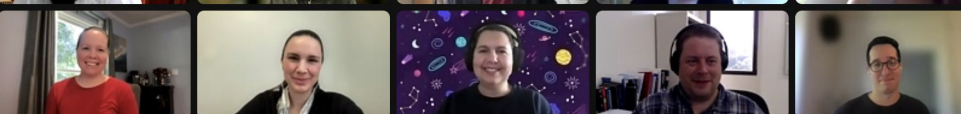
Free and  
open!

**Peer-reviewed concise resources** to assist the data curator with recommendations for curating specific formats or curation topics!

Ivey, Susan; Koshoffer, Amy; Sneff, Gretchen; Wang, Huajin. (2019). Confocal Microscopy Images Data Curation Primer. [Data Curation Network GitHub Repository](#).

## Radical interdependence is more than cooperation –

- It's depending on others **outside of your organization** to “do the work.”
- It's about trust, vulnerability, and accountability, not from a punitive perspective, but from an intrinsic desire to show up for and serve one another based on **clear expectations, commitments, and the shared goal of advancing the field.**
- It's about pooling our resources, our knowledge, and our interests, to better accomplish our work and **advocate for the role of data curators.**
- It's about recognizing the **interconnectedness** of the work we are all doing.



From Reina Chano Murray to Every...  
😊



Ethical. Reusable. Better.

**DATA CURATION NETWORK**

[datacurationnetwork.org](http://datacurationnetwork.org)

# How we started

[doi.org/10.3998/  
mpub.12782791](https://doi.org/10.3998/mpub.12782791)

# THE ART, SCIENCE, AND MAGIC OF THE DATA CURATION NETWORK

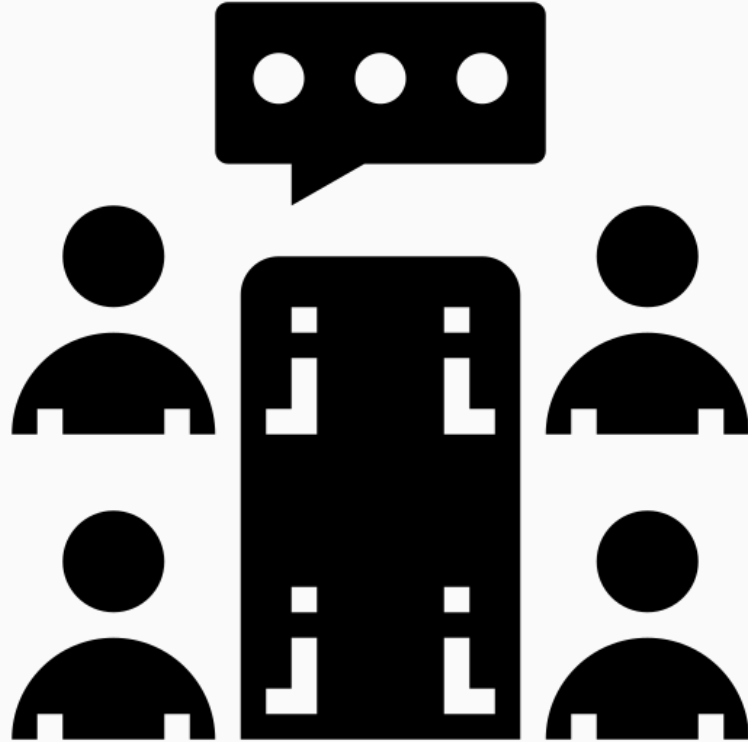
A RETROSPECTIVE ON CROSS-INSTITUTIONAL COLLABORATION



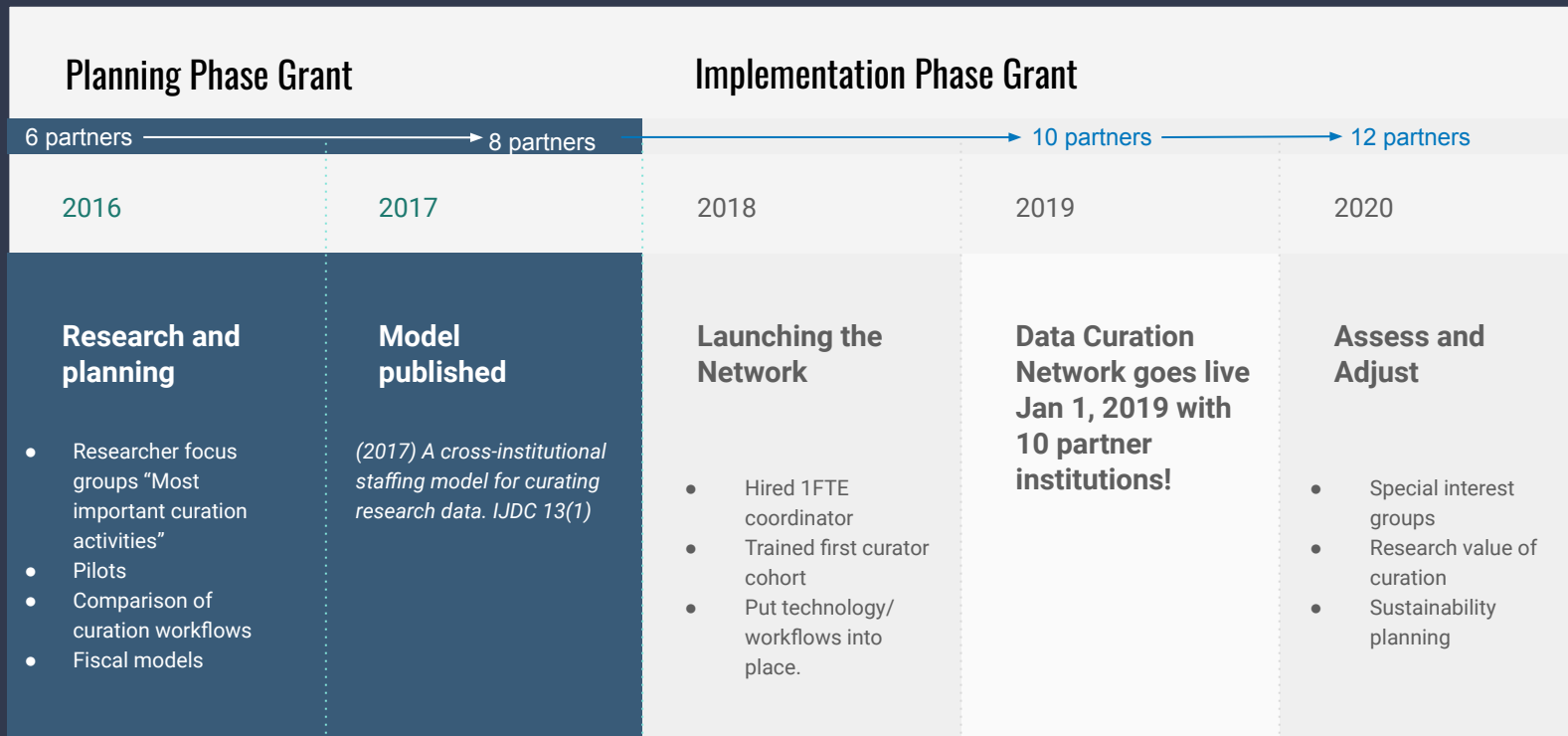
**DATA CURATION NETWORK**

## What if...

- Six individuals at six institutions
- Similar challenges in launching curation services
- They wondered – what if we build a shared workflow? What if we could design our services to they depended on each other?



# Development Timeline

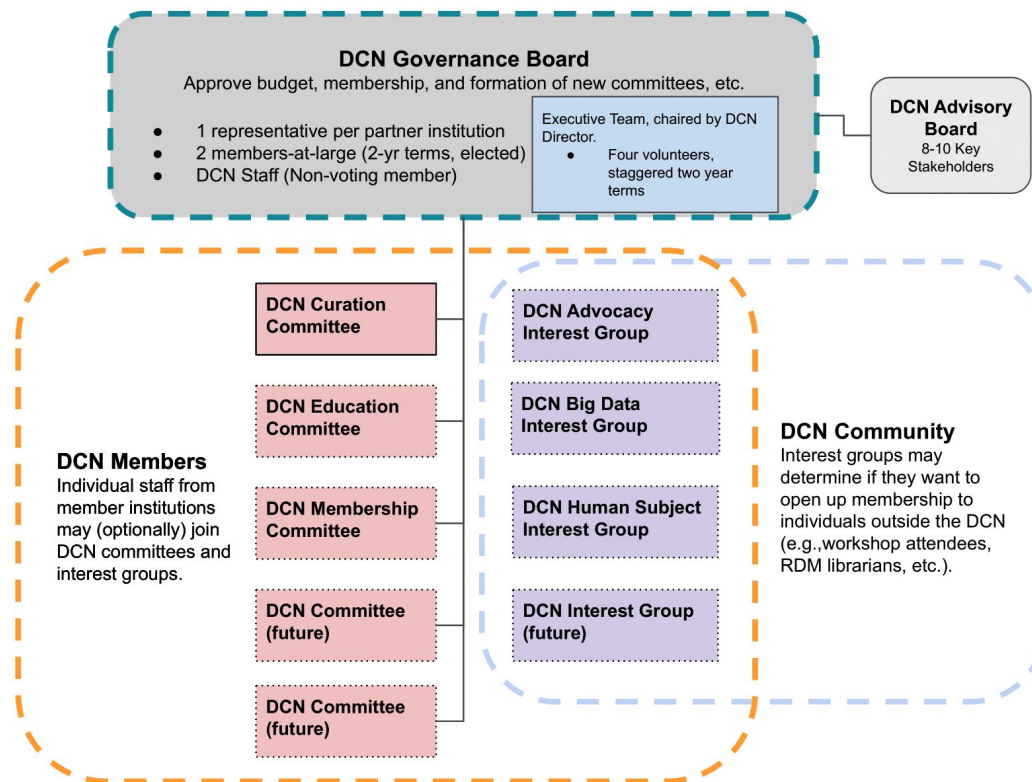


# Important notes

- Started with shared challenge and opportunity
- Full time project coordinator and staff member– essential
- Slow, thoughtful growth to ensure community connections and trust



# How we're structured

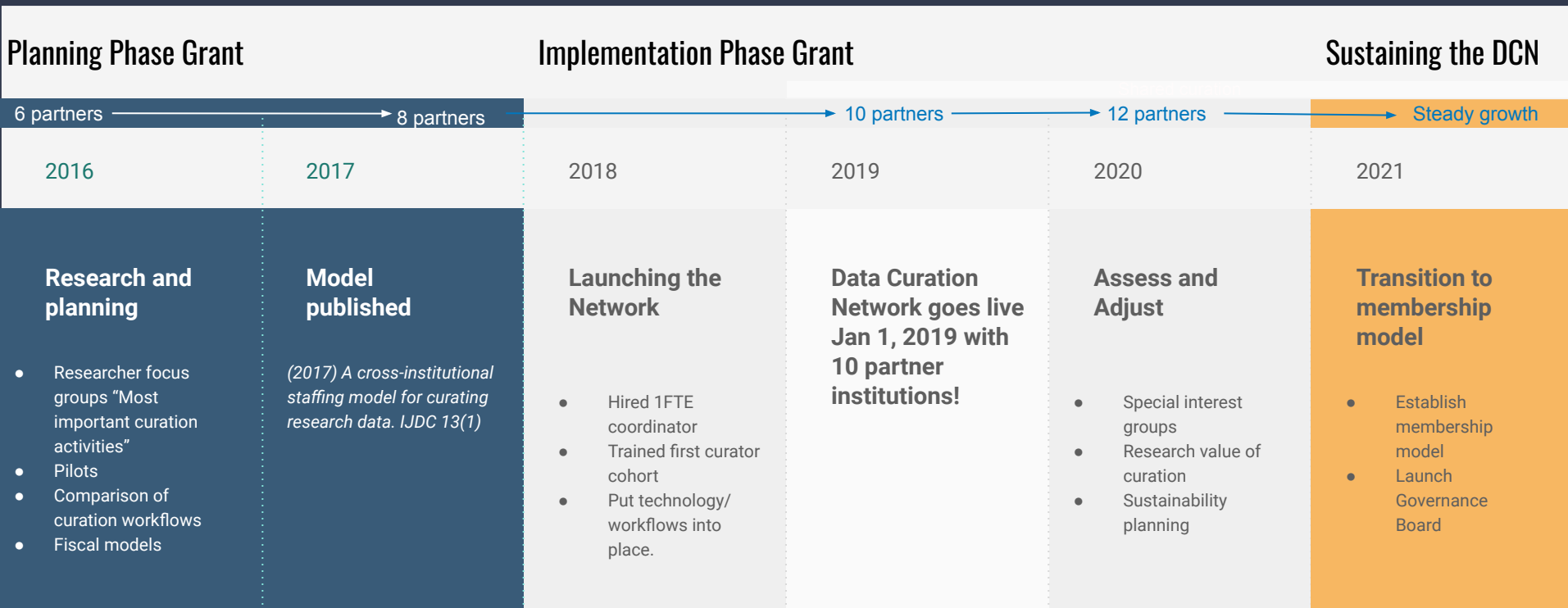


# Strategic Framework

- Collaboratively developed
- Goals for 2022-2025
- Balancing continued growth and development with available resources
- <https://hdl.handle.net/11299/226182>

# How we're funded

# Towards Sustainability





# DCN's Membership Model

## Sustainer Institution (\$10,000)

This is the highest level of participation. Sustainer institutions contribute curator staff to the Network and have full participation in governance.

- Radical interdependence: Participate in shared curation model (up to 200 hours)
- Representation in DCN Governance
- Travel support (\$2k) to annual All Hands Meeting (AHM)
- Logo on website

## Ambassador (\$2,500)

Ambassadors invite DCN instructors to teach our 2-day in-person CURATED workshop for local/regional staff.

- Host a DCN workshop
- DCN community of practice
- Invited to present virtually at AHM
- Logo on website

# Some thoughts...

- Community is essential
- Community and trust must be nurtured and grown
- Slow growth can be essential



Please reach out with any questions!

[mnarlock@umn.edu](mailto:mnarlock@umn.edu)