

Proceedings of the Workshop Exploring National Infrastructure for Public Access Usage and Impact Reporting

Edited by: Christina Drummond and Charles Watkinson

Proceedings of a workshop sponsored by the National Science Foundation and held in conjunction with the Coalition for Networked Information's Spring Meeting at the Sheraton Denver in Denver, Colorado on April 2, 2023

Table of Contents

Part I Workshop Objectives	2
Part II Methodology and Invited Speakers	3
Part III Ideation and Discussion Outputs	6
A. Observations on the State of Publicly Accessible Usage Data Analytics.....	6
B. Applying FAIR and CARE Principles to Usage Data	6
C. US-based Public Access Usage Reporting Infrastructure Interoperability.....	9
D. Facilitating Cross-Platform Research Usage and Impact Analytics across Scholarly Output Types.....	13
Part IV Recommended Priority Action Areas.....	16
Action Area 1 Stakeholder Education, Engagement, and Communication	17
Action Area 2 Develop and Learn from a Minimum Viable Product.....	18
Action Area 3 Explore Related Values and Principles.....	19
Action Area 4 Determine who "Owns" Usage Data, i.e. Who Can Give Permission for Use and Reuse.....	20

Acknowledgements: The organizers extend their gratitude to the organizations that made this workshop possible, including the National Science Foundation (NSF Award 2315721) and the Coalition for Networked Information whose staff extended their negotiated conference rates and logistical expertise.

Contributors listed in alphabetical order by last name: Tina Baich, Christina Drummond, Emily Farrell, Lisa Hinchliffe, Kristi Holmes, Martin Halbert, Patricia Hswe, Joe Karaganis, Ken Klingenstein, Katherine Klosek, Sharla Lair, John Lenahan, Mikala Narlock, Wendy Queen, Laura Ricci, Chris Shillum, Niels Stern, Charles Watkinson, and Maurice York.

Part I | Workshop Objectives

Invited international experts and leading scholarly cyberinfrastructure representatives joined workshop organizers Christina Drummond and Charles Watkinson for an eight-hour facilitated workshop on April 2, 2023. Together they aimed to:

- identify the challenges preventing cross-platform public and open scholarship impact analytics at scale,
- explore open infrastructure opportunities to improve the findability, accessibility, interoperability, and reuse i.e. “FAIRness” of usage data, and
- identify what’s needed to scaffold America’s national infrastructure for scholarly output impact reporting in light of a) the August 2022 Office of Science and Technology Policy (OSTP) “Nelson Memo” regarding “Ensuring Free, Immediate, and Equitable Access to Federally Funded Research,” and b) the European Open Science Cloud Core and Interoperability Framework.

Participants were encouraged to consider the challenges related to impact reporting and storytelling for research outputs ranging from data, articles, and books to simulations, 3D models, and other multimedia.

The workshop objectives shared in advance of the meeting with participants were:

- *identify what’s needed to scaffold America’s national infrastructure for scholarly output impact reporting,*
- *develop recommendations for national infrastructure and investment, and*
- *prioritize and begin to map out what activities we need to undertake next to support these recommendations.*¹

Aspirational impacts of the workshop shared by organizers with the facilitator pre-event included:

- increasing relationships and understanding among usage and impact metrics related stakeholders and infrastructures,
- informing of future collaborations and investments,
- exploring of the potential for providing economies of scale through shared national data infrastructure that interoperates at the global scale while providing an onramp for domestic stakeholders, and
- increasing awareness among US stakeholders of disruptive sovereign data governance related innovation specific to controlled data exchange via data spaces instead of current data-harvesting models.

¹ Christina Drummond, Charles Watkinson, & Katherine Skinner. (2023). Workshop Agenda | Exploring National Infrastructure for Public Access Usage and Impact Reporting. Zenodo. <https://doi.org/10.5281/zenodo.8209812>

Part II | Methodology and Invited Speakers

Workshop activities were designed by professional facilitator Dr. Katherine Skinner under the direction of co-organizers Drummond and Watkinson.

Pre-conference Logistics: The event was organized to maximize knowledge sharing while fostering open conversation and recommendation generation. Pre-event communications were sent to participants by Drummond through a series of three emails. Select participants were invited to prepare 5-minute talks to summarize a particular perspective on usage analytics or data-focused infrastructure.

Participant Preparation: Confirmed participants were asked to dedicate 60 minutes to event preparation, which entailed reviewing the following materials and consideration prompts:

- The workshop summary² and agenda noting participants and speakers³
- The [2021 OA Books Supply Chain Mapping Report](#) authored by Michael Clarke and Laura Ricci⁴.
 - *Participant prompt: Do the issues identified in this report for book usage and impact analytics extend to other scholarship outputs?*
- Six pre-recorded five-minute talks invited from the following presenters on the quoted topics:
 1. “How data collaboratives can bring public and private actors together to create new public goods while addressing the issues and concerns of all involved,” presented by Stefaan Verhulst of NYU GovLab.
Participant prompt: How are data collaboratives facilitating multi-party data brokerage through shared governance structures?
 2. “Systemic challenges facing COUNTER usage reporting standards adoption,” presented by Tasha Mellins-Cohen of Project COUNTER.
Participant prompt: What, if anything, is needed beyond COUNTER to improve usage and impact data interoperability?
 3. “IRUS-UK approach to federating institutional repository usage data reporting,” presented by Jo Lambert of JISC and IRUS-UK.
Participant prompt: What of the IRUS approach could be extensible to other types of usage statistics?
 4. “FAIR data brokerage and clearinghouse via the EOSC Core Interoperability Framework,” presented by Paolo Manghi of OpenAIRE and European Open Science Cloud (EOSC).
Participant prompt: How could the EOSC Core Interoperability Framework approach inform multi-platform public-private usage and impact analytics exchange in the US?
 5. “The Open Research Funders Group’s work and interests in usage and impact reporting” presented by Greg Tananbaum of the Open Research Funders Group and Higher Education Leadership Initiative for Open Scholarship (HELIOS).
Participant prompt: How are open research funders informing the demand for impact analytics?
 6. “Introduction to the National Secure Data Service (NSDS) Demonstration Project” presented by Heather Madray of the National Center for Science and Engineering Statistics.
Participant prompt: How could the NSDS inform multi-platform public-private usage and impact analytics exchange?

² Drummond, Christina, & Watkinson, Charles. (2023). Project Summary | Workshop | Exploring National Infrastructure for Public Access Usage and Impact Reporting. Zenodo. <https://doi.org/10.5281/zenodo.8209816>

³ Christina Drummond, Charles Watkinson, & Katherine Skinner. (2023). Workshop Agenda | Exploring National Infrastructure for Public Access Usage and Impact Reporting. Zenodo. <https://doi.org/10.5281/zenodo.8209812>

⁴ Clarke, Michael, & Ricci, Laura. (2021). OA Books Supply Chain Mapping Report. Zenodo. <https://doi.org/10.5281/zenodo.4681725>

Links to the following optional reference materials were shared in advance:

- The [August 2022 Office of Science and Technology Policy \(OSTP\) “Nelson Memo”](#) regarding “Ensuring Free, Immediate, and Equitable Access to Federally Funded Research”, authored by Dr. Alondra Nelson
- The [February 2021 European Open Science Cloud \(EOSC\) Interoperability Framework Report](#) from the EOSC Executive Board Working Groups for FAIR and Architecture, which defines four “layers” of interoperability: technical, semantic, organizational, and legal.
- The [April 2022 EOSC Interoperability Framework \(EIF\)](#) presentation by Michelle Williams of GÉANT
- The [Trustworthy Architecture for a Data Economy](#) infographic published by the International Data Spaces Association (IDSA)

TABLE 1: INVITED "5 MINUTE" TALKS

* Denotes a presentation delivered in person

Topic	Presenter	Title	Affiliation
National Secure Data Service Demonstration Project	Heather Madray	Data Access, Confidentiality and Quality Assessment (DACQA) Program Director	National Secure Data Service Project
How data collaboratives can bring public/private actors together to create new public goods while addressing the issues and concerns of all involved.	Stefaan Verhulst	Director	GovLab Data Program
Systemic challenges facing COUNTER usage reporting standards adoption	Tasha Mellins-Cohen	Executive Director	COUNTER
Institutional Repository Usage Statistics UK (IRUS-UK) approach to federating institutional repository usage data reporting	Jo Lambert	Head of Licensing Intelligence and Analytics, Digital Resources	JISC/IRUS-UK
FAIR data brokerage and clearinghouse via the EOSC Core Interoperability Framework	Paolo Manghi	Chief Technical Officer	OpenAIRE / EOSC
ORFG's work and interests in usage and impact reporting	Greg Tananbaum	Director	Open Research Funders Group / HELIOS
Overview of the usage and impact data landscape*	Laura Ricci	Senior Consultant	Clarke & Esposito
Advancing an European Research Infrastructure Consortium (ERIC) to support SSH scholarly Open Access publication*	Niels Stern	Executive Director	OAPEN / DOAB
Facilitating US consortia access to global open scholarship infrastructure*	Sharla Lair	Senior Strategist, OA and Scholarly Communications Initiatives	LYRASIS
Make Data Count and DataCite experiences with usage data* (<i>recording unavailable</i>)	Kristi Holmes	MDC Advisory Board	DataCite / Make Data Count
Leveraging the EU IDS model for cross-platform OA book usage data exchange*	Christina Drummond	Executive Director, OA Book Usage Data Trust	OA Book Usage Data Trust / UNT

On-site Workshop Facilitation: PIs Watkinson and Drummond co-hosted the 8:00am-4:00pm April 2nd, 2023 workshop, providing set-up and notetaking support to workshop facilitator Skinner.

Stakeholders were seated at tables with peer perspectives, with responses via post-it and marker color coded to designate stakeholder affinity groups for post-event reporting. At the beginning of the session, the group agreed to a workshop Code of Conduct and Chatham House Rules.

Information Sharing and Ideation: Five invited “5-minute” talks were presented in-person to orient attendees and build upon the pre-recorded talks.⁵ Pre-recorded and in-person speakers are listed in Figure 1.

Facilitation focused the first half of the day on partner activities with affinity and full-group discussion. Discussion prompts included:

- 1) FAIR Principles, CARE Principles, and Usage Data Sharing
 - a) *Is the way we currently exchange and leverage usage and impact metrics FAIR?*
 - b) *Does the way we currently exchange usage and impact metrics support CARE?*
 - c) *Why are FAIR and CARE challenging for usage and impact metrics across scholarship outputs and types?*
- 2) US-based public access usage reporting interoperability given EU activities
 - a) *What can we learn from and lean on that has already been created in the EU?*
 - b) *Do we need a complementary US-based infrastructure?*
 - c) *How can EU-based work and the US-based work best align? What mechanisms might we need to aid this, and who can provide them?*
- 3) Cross-platform usage and impact analytics across scholarly output types
 - a) *What do we need for this facilitation?*
 - b) *How do we leverage R&D from existing efforts?*
 - c) *What is needed to ensure trans-national infrastructure interoperability?*

Each topical discussion began with capturing personal reflections on post-its, followed by partner sharing and affinity group discussion. Ideas were collated and summarized prior to full group report outs. (See [Part III](#)) At the end of each discussion round, affinity groups were asked to identify recommendations for action. Over lunch, recommendations were pooled and shared for reflection. Full group discussion and recommendation prioritization surfaced action areas that merited immediate attention and deeper consideration. Four “action-area” topics were then developed further through “conversation café” style rotating discussion. Images were captured of all group ideation outputs, which were transcribed for inclusion herein. (Figures 1-8)

Knowledge Dissemination: Following the workshop, organizers Drummond and Watkinson immediately shared out the action areas that emerged. They joined European participant Niels Stern to provide an overview of the workshop’s findings at the co-located Coalition for Networked Information (CNI) Spring Member Meeting.⁶ They also summarized workshop findings through an Educause podcast recorded onsite in Denver.⁷

Resources generated through this NSF award were uploaded into a Zenodo Community dedicated to this effort for further distribution and broad discoverability.⁸

⁵ In-person presentations were recorded and archived via the University of Michigan <https://bit.ly/nsf-pa>.

⁶ CNI: Coalition for Networked Information. (2023, April 26). *Workshop Report Out: National Infrastructure for Public Access Usage and Impact Reporting* [Video]. YouTube. <https://youtu.be/it8X-BcztHE>

⁷ Drummond, C, Watkinson, C, and Bayne G. (2023) “Improving the FAIRness of Usage Data,” The CNI Interviews Podcast, Coalition for Networked Information/Educause. Available at <https://er.educause.edu/podcasts/the-cni-interviews-podcast/improving-the-fairness-of-usage-data> [May 1, 2023].

⁸ See: <https://zenodo.org/communities/2023-workshop-exploring-national-infrastructure-for-public-access-usage-and-impact-reporting>

Part III | Ideation and Discussion Outputs

A. Observations on the State of Publicly Accessible Usage Data Analytics

After the in-person presentations, participants shared initial observations on the current state of usage and impact metrics for publicly available scholarship. Ideas shared fell into three themes:

- 1) Complexity in the usage data supply chain for publicly accessible scholarship:
 - “Easy to count but hard to measure”
 - Supply chain roles, actors, and related gaps; it’s unclear which roles are played by whom
 - In the US, power is in the consortia which are structured as buying clubs for cost efficiency and focus
- 2) Unclear incentives to interoperate:
 - Unclear stakeholder incentives to participate and interoperate in a hybrid open/paywalled economy
 - Unclear legal risks and operational impacts of granular usage data exchange
- 3) Opportunities, such as:
 - Object-type agnostic work and focus
 - Impact metrics to illuminate the difference in paywalled vs. open access
 - Incorporating values and measures such as Humetrics HSS

These full-group reflections set the stage to consider the controlled, responsible exchange and use of usage data.

B. Applying FAIR and CARE Principles to Usage Data

After being reminded of the EOSC approach to FAIR data brokerage and the [CARE Principles for Indigenous Data Governance](#), namely Collective benefit, Authority to control, Responsibility, and Ethics, affinity groups were asked to consider three questions.

- 1) Is the way we currently exchange and leverage usage and impact metrics FAIR?
- 2) Does the way we currently exchange usage and impact metrics support CARE?
- 3) Why are FAIR and CARE challenging to implement for usage and impact metrics across scholarship outputs and types

Systemwide Challenges

Challenges were identified to applying the FAIR and CARE Principles to usage and impact metrics. (Table 2) A lack of awareness of how these principles apply to the usage data value was cited along with a lack of principles adoption given other priorities or a lack of stakeholder buy-in. Cross-cultural values and ethical principles related to research impact measurement was noted, with discussion surfacing the need to understand the “rights” associated with usage data access, ownership, and use, recognizing interpretations of such rights differ across nations.

“FAIR is an afterthought and CARE is unknown”

The impact of stakeholder competition on organic cooperation and trust was raised. Competition among players for grant opportunities within a fragmented funding environment was noted, as was competition among public and private services for patrons, users, and customers. While the value of usage and impact data analytics was recognized, multiple issues related to trust surfaced, such as the gamification of metrics, COUNTER non-compliance, and trust in how shared data would be applied or used downstream. A current lack of usage and impact data transparency, data provenance and universally understood definitions were noted as related issues impacting trust in usage data quality.

Participants recognized the complexity of the global legal landscape surrounding the global research ecosystem. Issues related to data accountability surfaced usage and impact ‘data ownership’ and ‘data sovereignty’ were not fully understood. Participants noted the difference between US and EU regulatory frameworks and the increasing state-by-state variance in the US related to data aggregation and data brokerage focused privacy regulation.

FIGURE 1: CHALLENGES FACING FAIR & CARE USAGE DATA | IDEATION RESULTS



Participants discussed how it was not clear which organizations had the authority to determine what's allowable and ethical in a legal landscape that touches upon related issues of copyright, licensing, privacy, and proprietary revenue generation interests.

A lack of clarity around the collective benefit of usage data exchange was noted, recognizing that value propositions and incentives would be necessary for more granular, controlled exchange of usage and impact data created among both public and private organizations. It was noted that for many, commercial or competitive interests resulted in usage and impact metrics being viewed as sensitive data that could not be made fully open or freely accessible as public data.

“Aggregators feeling uncertain whether they can share data with contributing publishers or purchasing/supporting libraries”

“Some companies stand in the way of sharing”

“Much data is in corporate hands subject to individual corporate rules and policies”

“Be FAIR and CARE to SHARE”

Actors positioned to assist with making usage FAIR and CARE to share

Suggestions of who could be engaged to overcome said challenges spanned:

- Individuals: researchers, policy makers, data stewards, plus the public
- Organizations: libraries, archives, publishers, funders, and the organizations providing professional development and education to data librarians and researchers, standards bodies, and both nonprofit and commercial players.
- Networks: library consortia, professional associations including libraries

Specific named networks included the Data Curation Network (DCN), the Research Data and Preservation Association (RDAP), and institutions participating in the Higher Education Leadership Initiative for Open Scholarship (HELIOS). The role of policymakers in setting standards and guidelines, and coordinating stakeholders was underscored.

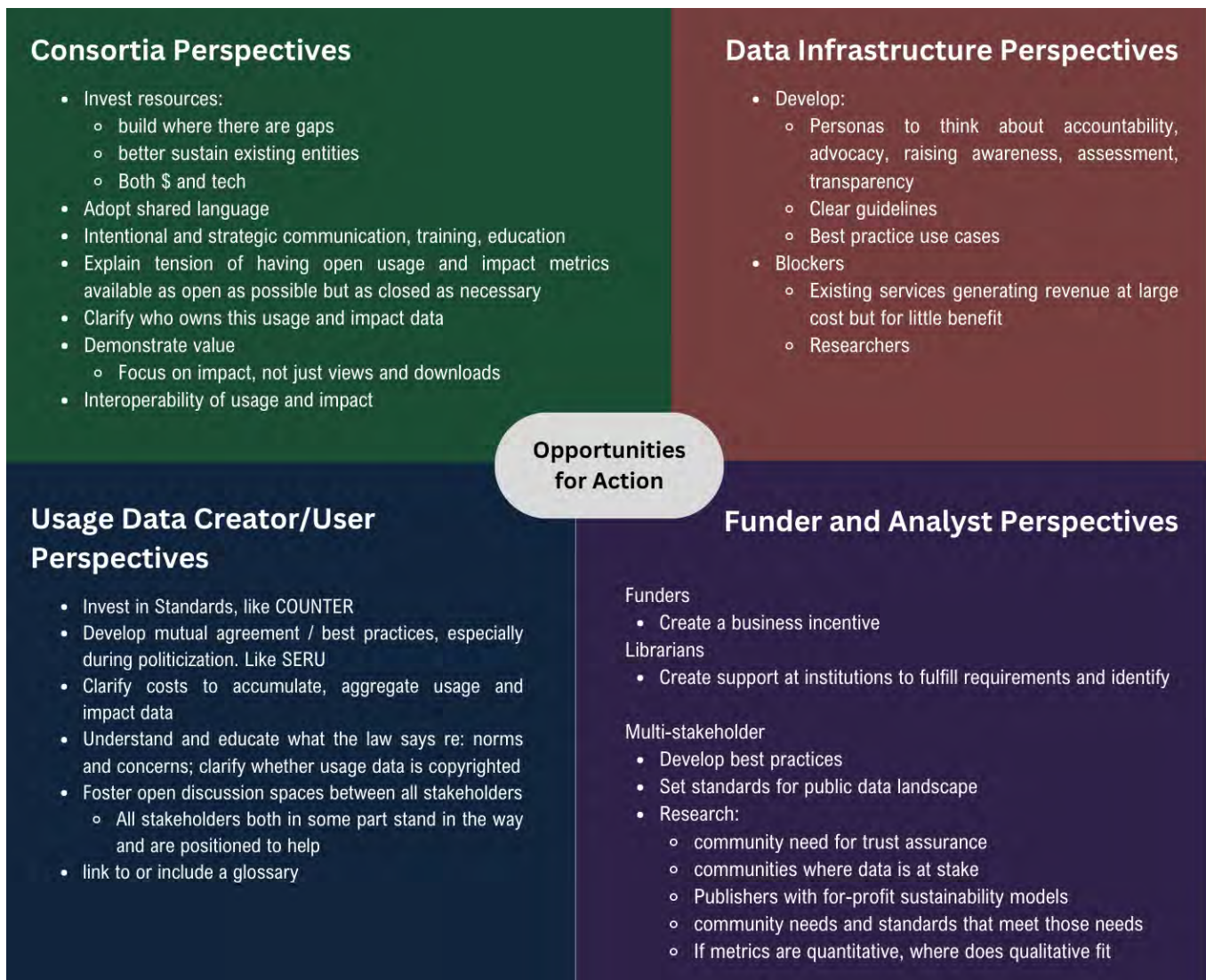
Opportunities for Action

Opportunities for research, coordination, and investment were ideated (Table 3). Investment recommendations spanned research, development and sustaining existing efforts like COUNTER. Funding agencies, consortia, and standards organizations were noted for their ability to create incentives and support institutions looking to improve their usage data practice. The need for communications efforts to raise awareness, develop shared vocabularies, and foster multi-stakeholder collaboration and trust was noted, with specific recommendations to develop strategic materials for stakeholder-focused education leveraging guidelines, personas, and use cases to establish value and increase awareness at the cross-sections of usage and accountability, assessment, and transparency.

The need to understand the legal and regulatory landscape pertaining to usage data was noted by two of four groups. Additional information on norms and concerns was sought to clarify what distinct stakeholder groups needed to increase trust in their usage data. The need to learn from a demonstration project such the OA Book Usage Data Trust was cited, to understand why public and commercially created usage and impact data may need to be exchanged in ways that are as open as possible but also as controlled as necessary to achieve public benefit.

“The definition of public data is fractured at best – much data is in corporate hands...subject to individual corporate rules and policies”

FIGURE 2: NEXT STEPS TO PURSUE FAIR AND CARE FOR USAGE | IDEATION RESULTS

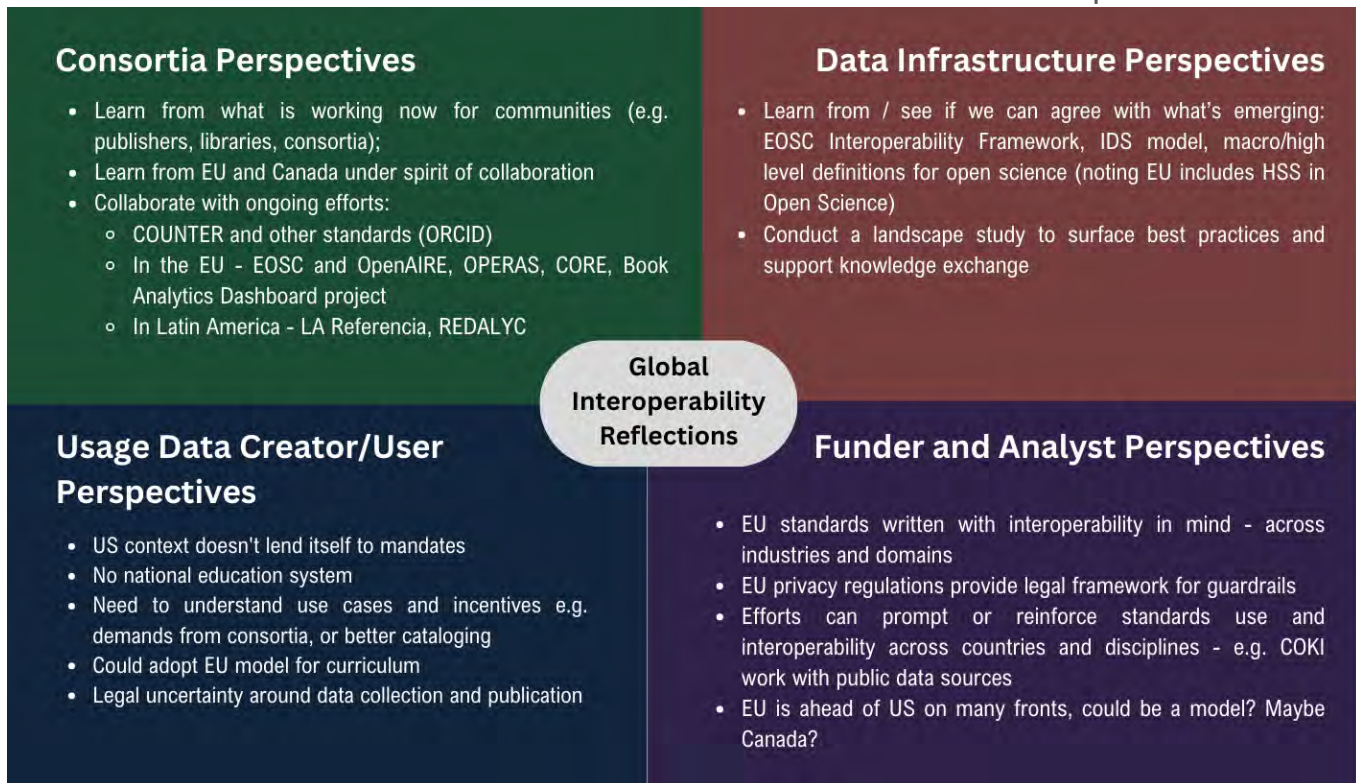


C. US-based Public Access Usage Reporting Infrastructure Interoperability

To explore how to improve usage reporting infrastructure interoperability, participants discussed three conversation prompts to then identify potential opportunities for action.

The first conversation was seeded with the prompt, “*What can we learn from efforts such as the European Open Science Cloud (EOSC) Interoperability Framework and European-engaged efforts to faster cross-platform data exchange?*” A shared desire to leverage what’s already working in a spirit of collaboration surfaced, noting in particular the EOSC Interoperability Framework, the International Data Spaces model and its related OA Book Usage Data Trust effort, and the macro-level European definitions of open science as including the humanities and social sciences. Specific efforts to engage included global standard efforts, such as COUNTER and ORCID, and organizational networks that support open science like OpenAIRE’s network of 35 national OA Desk nodes, the OPERAS research infrastructure consortia and La Referencia’s 12 national nodes in Latin America.

FIGURE 3: WHAT CAN WE LEARN FROM EFFORTS SUCH AS THE EUROPEAN OPEN SCIENCE CLOUD (EOSC) INTEROPERABILITY FRAMEWORK AND EUROPEAN-ENGAGED EFFORTS TO FOSTER CROSS-PLATFORM DATA EXCHANGE | IDEATION RESULTS



Groups were asked to consider the domestic applicability of such global approaches. Thoughts on the relevance of the European approaches to the US context resulted in recognition of the role library consortia and developed collaborative networks play in the US. The importance of identifying US specific goals, objectives, and organizations working on international usage and impact metrics was noted, with an eye towards identifying use cases and incentives for increased data sharing and standards adoption. Identifying US stakeholders actively engaged in usage and impact analytics efforts was as important as the need to explore what from European efforts could be adopted or adapted for the US. Engaging existing US parties engaged in multi-national data and science initiatives surfaced as a strategy to move quickly, given their existing networks and ability to collaborate across national and disciplinary boundaries.

“Talk to each other at all levels to ensure interoperability”
“Consider who can bring all of the people together in a bigger tent”
“Use global infrastructures, but include all relevant stakeholders”

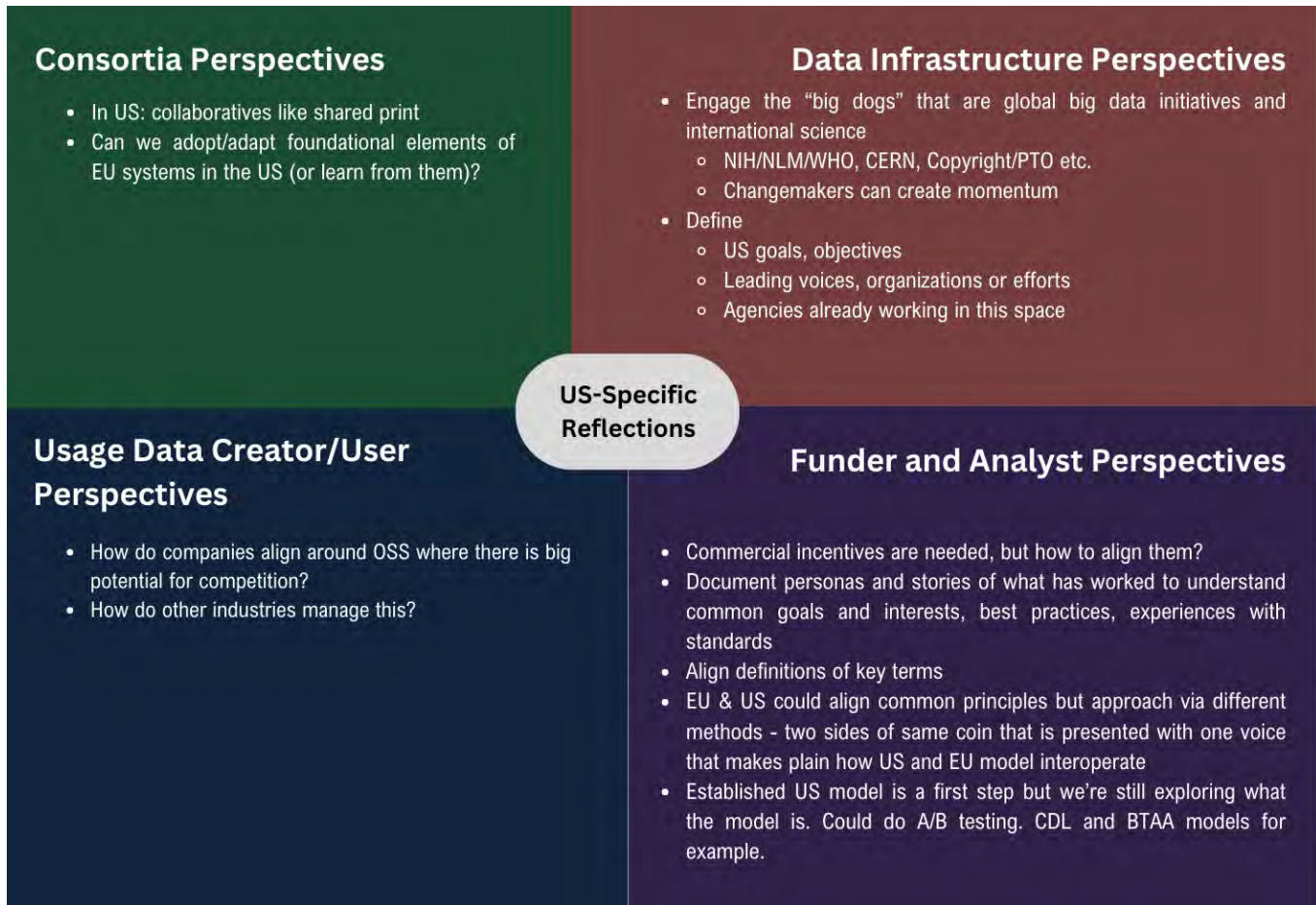
Participants were asked to consider how best to ensure international alignment and Trans-Atlantic coordination as solutions develop. Multiple groups noted the importance of high-level conversations for interoperability and framework development across disciplines, stakeholder types, output types and national contexts. The Research Data Alliance and World Wide Web Consortium were noted as potential global venues for policymakers and agencies to joint action.

“Explore how to incentivize data holders to share usage and impact data – funders could help, policymakers and funders have a role to play.”

“Policy incentives make it easy to do the right thing...”

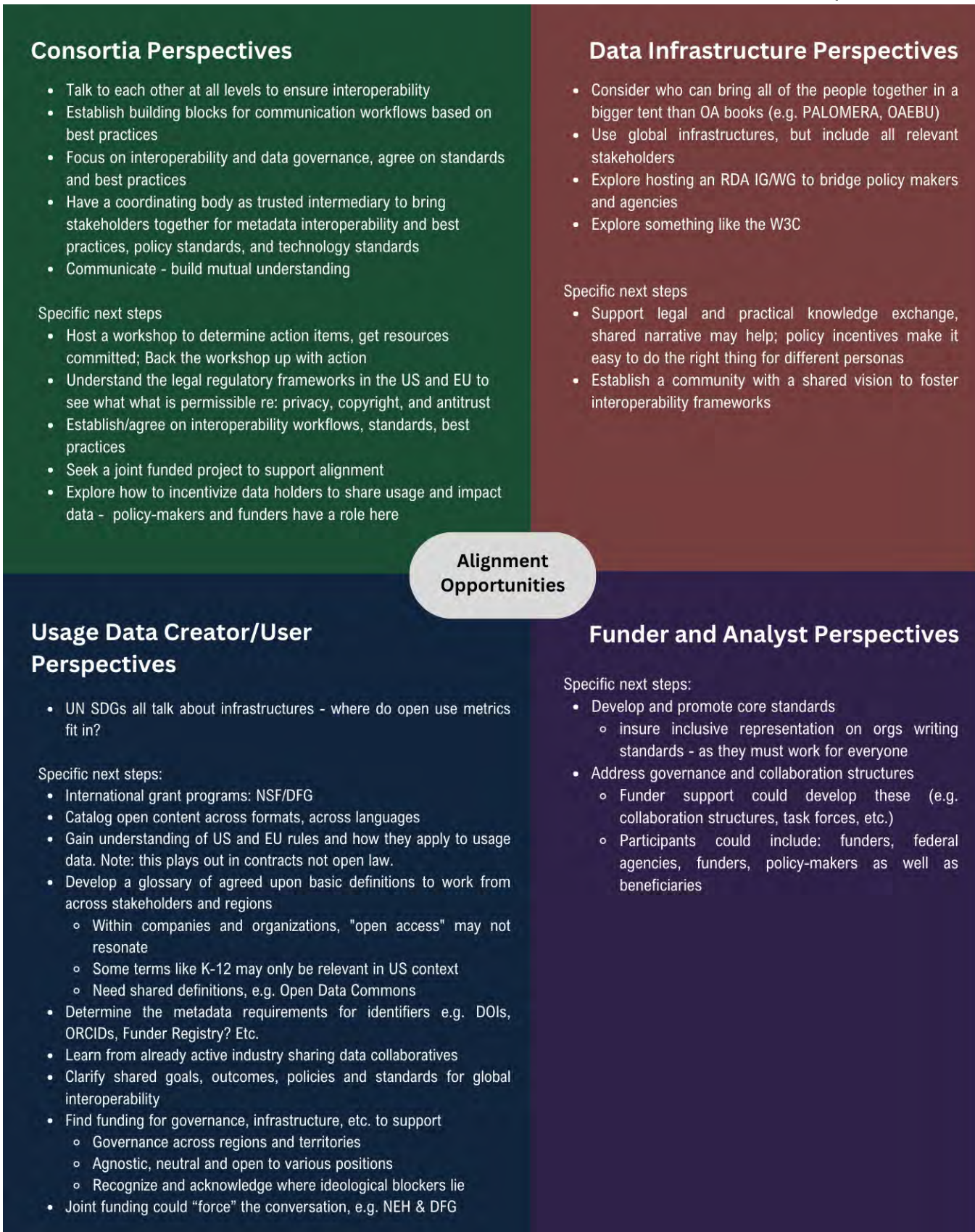
“Address governance and collaboration structures... Funder support could develop these (e.g., collaboration structures, task forces, etc.)”

FIGURE 4: THOUGHTS ON US-BASED INFRASTRUCTURE INTEROPERABILITY | IDEATION RESULTS



A trusted intermediary or coordinating body was suggested to foster multi-stakeholder alignment and research related to metadata interoperability, best practices, policy, and technology standards. However, the need to financially support such an effort was underscored, recognizing a need to be agnostic and neutral like the UN while being able to identify and support ideological and policy differences related to the public value and transparency of scholarship usage and impact metrics. Participants noted that jointly funded projects or workshops could determine action items and commit resources for identified action items. The importance of communications development was raised, with vision statements, value propositions, and user personas surfacing as needed items. Multiple groups noted how policymakers and funding agencies could incentivize action for different data holding stakeholders if they clearly understood what was needed. Potential action items ranged from developing a shared glossary or crosswalk of usage and impact data stewardship related terms to understanding metadata requirements and determining how existing data collaboratives address legal issues across transnational boundaries.

FIGURE 5: HOW CAN TRANS-ATLANTIC COORDINATION AMONG SUCH EFFORTS BE KEPT IN ALIGNMENT? | IDEATION RESULTS



D. Facilitating Cross-Platform Research Usage and Impact Analytics across Scholarly Output Types

In this final affinity group session, individuals discussed: 1) What is needed to facilitate cross-platform analytics, 2) What existing research can be leveraged, and 3) What is needed to ensure trans-national infrastructure interoperability as research and development continues on efforts such as the OA Book Usage Data Trust?

Action items identified for advancing the state of cross-platform analytics included understanding varied impact metrics across scholarship output types and disciplines, so that stakeholders could, “demonstrate (the) real impacts” of research on teaching, policy, patents, and public education.

“Strengthen / define what “counts” as an “access” for usage”

“Not just scholarly output types, but also sources of attention (where is usage coming from)”

“Are we in danger of looking backward? What is on the horizon...LLN authored works, multi-component works, multimodal works”

The idea of a crosswalk surfaced, noting differences in measures, citations, definitions, and schema for different scholarship outputs. Documenting the current state of interoperability and exchange between infrastructures surfaced as critical to informing transnational interoperability and infrastructure interconnectivity.

“...research is international – doesn’t help to have misaligned definitions of what’s asked for for tracking and reporting in the EU and US”

“Not a time to tear the house down; leverage what we have.”

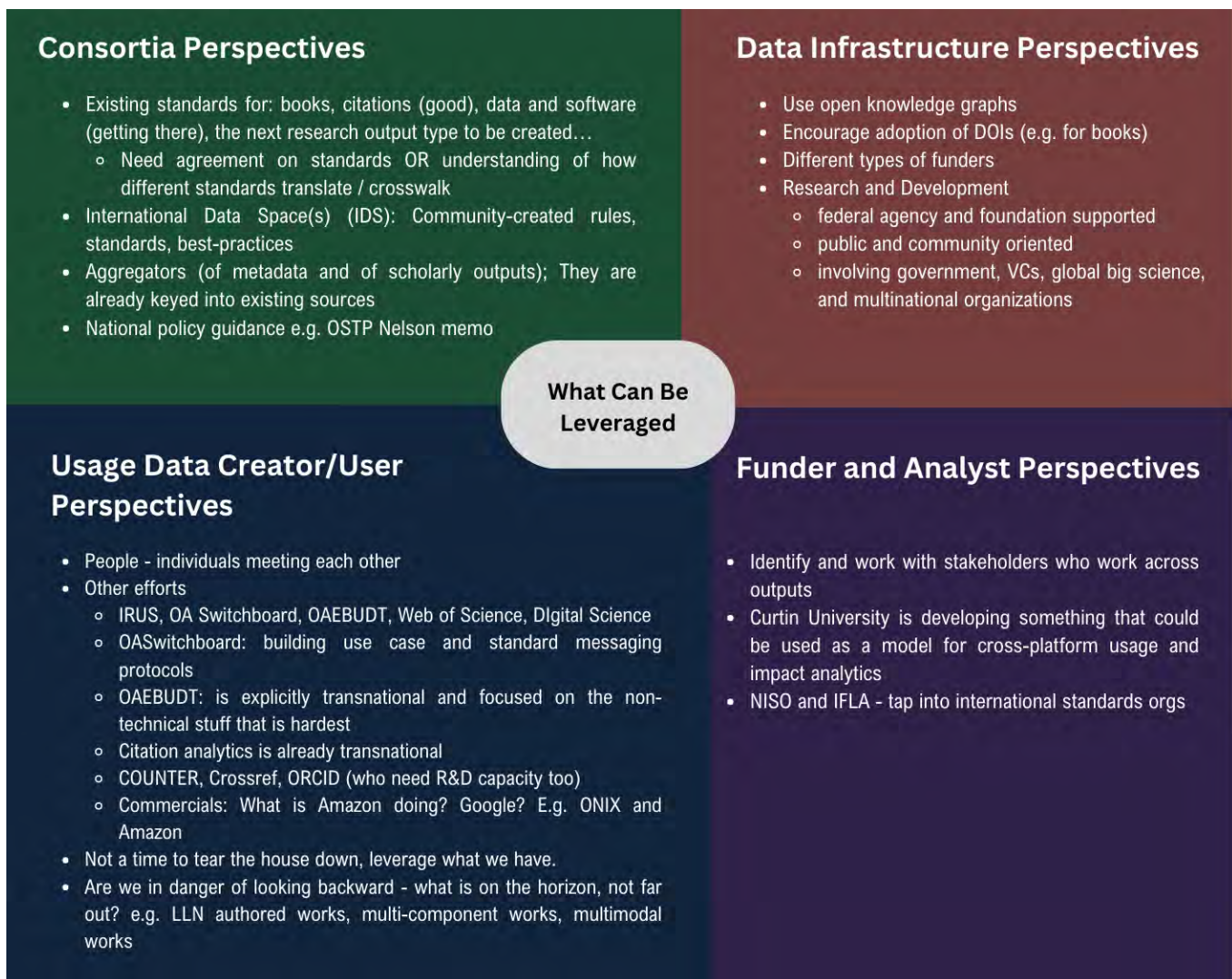
Existing resources to leverage were noted. Existing efforts to engage included those that presented, standards efforts, persistent identifier authorities, and organizational networks such as NISO and IFLA; as was the importance of learning from efforts that have ceased, such as the Distributed Usage Logging project. The concept of “scaling small” came up, noting how a simple proof of concept could foster iterative community development, alignment, and learning. Such an opportunity could also illuminate the costs associated with launching and sustaining a federated, shared data exchange infrastructure. The OA Book Usage Data Trust was cited as modeling this principle, intentionally developing an extensible International Data Space for usage and impact data by starting with the narrow use case of simplifying multi-platform exchange and aggregation of publicly accessible digital book and chapter views and downloads.

Multiple ways to foster interoperability were noted, from supporting networks of leaders and international alignment exercises, to developing key performance indicators for usage and impact analytics. Fostering collaboration across both public and commercial platforms was flagged as critical as financially supporting the infrastructure governing bodies capable of bringing parties together to stewards such infrastructure efforts.

FIGURE 6: OPPORTUNITIES TO IMPROVE CROSS-PLATFORM USAGE AND IMPACT ANALYTICS | IDEATION RESULTS



FIGURE 7: RESOURCES THAT CAN BE LEVERAGED



Ensuring Interoperability Going Forward

The necessity of funding to support international coordination and networked leadership was flagged, alongside the need to build on what exists already. COUNTER, the Distributed Usage Logging Project, DOIs, OA Switchboard and the OA Book Usage Data Trust were mentioned as efforts to leverage, in addition to the “big dogs” in science infrastructure who already operate transnationally such as instruments, repositories and the National Research and Education Network (NREN) internet consortia.

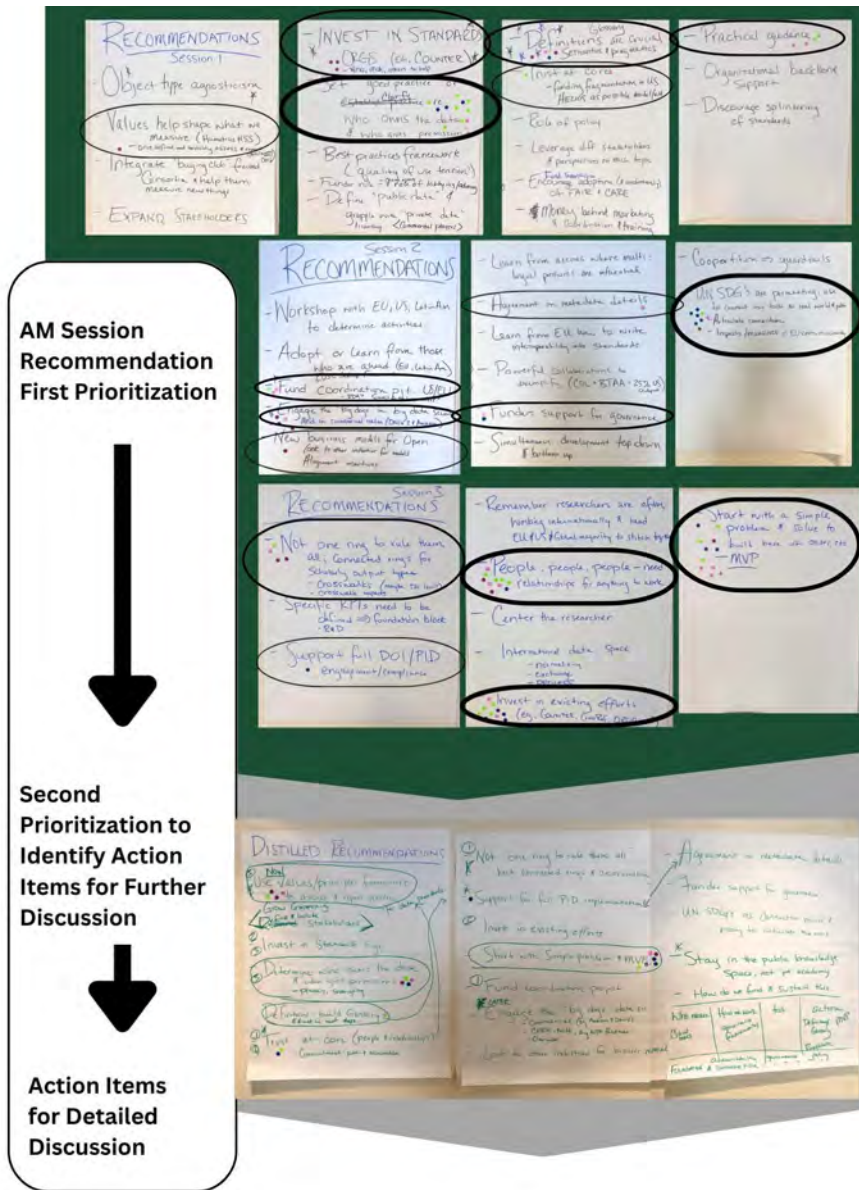
“Usage statistics we can do; impact is much harder.”

The value of communications and educational materials to support efforts across output types was noted, with participants recommending the development of case studies to illustrate the real-world impacts of usage and impact related analytics on publicly available scholarship. While noting the importance of starting with what’s easiest technically and legally to prove what’s possible, participants highlighted the need to stay open to “new and more meaningful standardized metrics.” Understanding what impact means for different outputs was identified as a prerequisite to getting more meaningful metrics beyond views and downloads.

Part IV | Recommended Priority Action Areas

During and after lunch, participants self-selected into new groups to cull and prioritize the opportunities for action raised in the am sessions. The four recommendations receiving the most votes became stations through which participants rotated to generate recommendations on specific activities and next steps.

FIGURE 8: ACTION ITEM PRIORITIZATION PROCESS



Individual recommendations from AM sessions were summarized by the facilitator for a priority vote from participants. Participants each had 3 votes to identify priority actions.

After the first distillation, participants received an additional single vote to identify areas ripe for action to discuss in detail during the final session.

The four action areas considered in-depth at stations were: 1) Stakeholder Engagement, Education and Communication, 2) Usage Data Ownership and Stewardship, 3) Minimum Viable Product Development, and 4) Values and Principles Development. Opportunities for action and recommendations for immediate next steps were captured for each (Figures 9-12)

Action Area: Education, Engagement, & Communication

Needed Activities

- Explore and document who we are as an ecosystem (define stakeholders) and how we work on OA impact (what, why)
- Confirm if we have a shared problem, e.g. aggregating usage data across platforms for publicly accessible scholarship of all types
- Articulate differences and attribute to contextual use (i.e. don't create a value hierarchy or right/wrong or prescriptive approach)
- Create a vocabulary and glossary to support a common understanding (and terms) across stakeholders
 - Gather existing definitions and analyze for overlap / conflict
 - Common sense and empirical check
 - Avoid definitional conflicts - crosswalk where necessary, determine what is yet to be defined: Metrics of usage and indicators of engagement and impact, technologies used and actions needed to have trust at the core
 - Recognize contributors and existing efforts
- Start with focus from MVP and existing standards (COUNTER, IDS)
- Communicate to educate
 - Clarify audience / purpose / context for communications and educational documentation or events
 - Identify and socialize the problem to center outreach and engagement

Recommended Next Steps

1. Develop shared glossary with term definitions and/or crosswalks to support diverse stakeholder collaborations
2. Document and/or define related values and principles
3. Engage and strengthen Persistent Identifier Organizations (PID Authorities)

Action Area: Develop and Learn from a Minimal Viable Product

Needed Activities

- Define the minimum loveable product that people will pay for to sustain
- Determine if we're focusing on "impact" or "use"
- Define and document the problem statement and solution
- Clarify how to measure it
- Pick the core data to start with

Recommended Next Steps

1. Host "competition" or RFI
2. Host stakeholder workshops
3. Contract dedicated "Product Manager" to move this forward
4. Develop straw model to go from problem definition to MVP

Note: Development is iterative. It could occur in 2 tracks simultaneously: one defining the straw data model, the other defining principles, legal constraints, and stakeholder incentives. Also - leverage what is already underway in other efforts; don't recreate the wheel.

Action Area: Explore Values and Principles

Needed Activities

- Explore who we are: how we work, technologies used, and actions that are needed to have trust at the core.
- Define relevant values and principles, leveraging existing research (e.g. POSI, OAEBUDT)

Recommended Next Step

1. Understand current research on value and principles application in existing related efforts.

Action Area: Determine Who “Owns” Usage Data i.e. Who Can Give Permission for Use and Reuse

Needed Activities

- Consider who could own it? Stakeholders involved: Authors, publisher, funder, reader, library, public, platforms
- Consider if data rights management should be framed as stewardship not ownership to align with the laws on data controlling and processing
- Understand the points where permissions are granted, and how data is distributed/stored
 - Raw data and processed data will have different overlays of permissions
- Determine what rights are in play
 - Document which aspects of data ownership are statutory (e.g. copyright), contract law, and/or norms by country;
 - Clarify and document who has legal authority to grant permission across platforms; Learn from other collaborative data initiatives
- Determine what is meant by “usage data” and “impact data”
 - e.g. Granular IP address data sharing would require stronger protections than usage counts; Define usage “data” (IP addresses or raw usage counts) terms of sharing and reuse; descriptive and prescriptive documentation; from downloads/views to highlighting/annotation and citation; recognize usage data definition will shift with innovation that redefines “use”
- Remember to incorporate principles and values, e.g. reader privacy and OA usage data anonymization

Recommended Next Steps

1. Commission research to document relevant legal (contractual & policy) and cultural norms
2. Develop model license language