

Pilot Evaluation and Return on Investment Analysis

26 November 2025

Prepared for

Open Access Ebook Usage Data Trust (OAEBUDT)



A C&E Report

For more information, contact:

Laura Ricci, Senior Consultant | lr Ricci@ce-strategy.com

Michael Clarke, Managing Partner | mclarke@ce-strategy.com

Clarke & Esposito | www.ce-strategy.com

CONTENTS

1. Executive Summary.....	3
2. Introduction.....	5
2.1 About the Data Space.....	5
2.2 Project Methodology	6
3. Project Participants.....	7
4. Pilot Integration	8
4.1 Process Feedback and Suggested Improvements	10
5. Pilot Goals and Outcomes	13
5.1 Reasons for Participation.....	13
5.2 Participant Goals and Progress	14
5.3 Additional Benefits Not Anticipated.....	17
6. Return on Investment (ROI) Assessment	17
6.1 Pilot ROI – Results	17
6.2 ROI Expectations.....	17
7. Findings and Recommendations.....	19
7.1 Establish incentives for future data space adoption	19
7.2 Understand and establish success criteria	27
7.3 Recommendations for future development.....	28

1. EXECUTIVE SUMMARY

The OA Book Usage Data Trust (OAEBUDT) is an international effort to facilitate the direct data exchange and benchmarking of open and proprietary usage data about open access (OA) books. The primary goal of the OAEBUDT effort is to address the challenge of aggregating and curating OA book usage data by enabling community-governed sharing of quality, interoperable ebook data.

As part of the “Advancing to Launch by Developing IDS Governance Building Blocks” project, the OAEBUDT project team launched a narrow international data space (IDS) pilot in 2025 to test the viability of exchanging COUNTER-compliant usage statistics through the IDS. Clarke & Esposito (C&E) was engaged to complete a Return on Investment (ROI) study of pilot participants, to capture the benefits of the IDS integrations, and to summarize recommendations to the project team regarding common pain points and opportunities for future integrations.

The pilot served as a proof of concept that succeeded in:

- Demonstrating the feasibility of the data exchange protocol
- Providing tangible experience with data space integrations, especially for participants’ non-technical staff
- Developing foundational technology that can be built upon for future integrations

Due to the limited nature of the pilot, participants did not complete “live” integrations and therefore did not make any enduring changes to existing workflows or capabilities. As a result, no participants saw a *quantifiable* ROI at this pilot stage.

In qualifying potential areas to generate ROI in future, participants did not see the data space as likely to generate significant savings in either time or resources (at least in the short term), or to contribute directly to revenue growth. Rather, the biggest anticipated source of ROI of the data space cited by participants was through *new capabilities* such as enabling controlled data sharing between parties (especially where there is no formal relationship in place), or providing greater visibility and auditability of data being shared.

Reflecting on the pilot itself, many participants saw the initial investment as worthwhile, especially to the extent that it creates an incentive to improve ebook usage data reporting capabilities among counterparts throughout the industry.

Participants noted three important success criteria for enabling these new capabilities and achieving ROI in future:

1. **There must be a critical mass of participants exchanging data with one another.** Attaining this critical mass will rely on a sufficiently low barrier to entry (both technical and financial), especially for small- to medium-size organizations. There must also be clear benefits for participation, so OAEBUDT must overcome its “chicken or the egg” problem – the need to demonstrate benefits to attract participants, while needing enough active participants to demonstrate a meaningful benefit.
2. **There must be improvements to provider data quality and standardization.** Increased access to poor data will provide only limited benefit. While the data space will not dictate reporting standards or practices, it can (and should) provide incentives to improve data quality shared through the IDS.
3. **There must be clarity about ongoing costs and sustainability.** Members of the pilot noted that the data space requires a clearer trajectory for sustainability, and more definition for how costs will be distributed between participants. Small- and medium-size organizations are particularly concerned about the risk of high costs imposed for participation, which would significantly impact their ability to generate a positive ROI.

Participants noted that potential benefits of data space participation are not limited to OA books, but rather could be recognized for any product under any business model. Expansion of the data space beyond OA books providers may in fact be a necessary condition given the relatively small size of the market and the limited role that OA books play within the scholarly publishing industry’s largest organizations.

Our recommendation to OAEBUDT to support future data space expansion is to more clearly articulate these potential benefits to target recruitment of one or more “critical node” organizations. The most likely organizations to act as a “critical node” are data providers with the ability to provide a robust set of data (both in terms of data volume and data quality), a strong direct *and* indirect network of stakeholders to act as potential data recipients, and the wherewithal to provide financial support for the data space in both the short and long term.

2. INTRODUCTION

2.1 About the Data Space

The OA Book Usage Data Trust (OAEBUDT) is an international effort to facilitate the direct data exchange and benchmarking of open and proprietary usage data about open access (OA) books. The primary goal of the OAEBUDT effort is to address the challenge of aggregating and curating OA book usage data by enabling community-governed sharing of quality, interoperable ebook data.

The project is adopting the International Data Space (IDS) model developed by the International Data Space Association (IDSA), which works with the values of:

- Increased interoperability
- Trust through secure and transparent exchange
- Multiparty data governance through usage controls and community-based accountability measures

The “Advancing to Launch by Developing IDS Governance Building Blocks” project, sponsored by the Mellon Foundation, aims to formalize the data space's community-governance mechanisms, quantify data space participation benefits in terms of Return on Investment (ROI), and understand the full operational costs related to international data space for OA book usage.

This project was launched in July 2022, led by Principal Investigators Christina Drummond (University of North Texas), Prodromos Tsiavos (OpenAIRE), and Yannick Legré (OPERAS), and concludes in late 2025.

As part of this project, the OAEBUDT project team launched a narrow IDS pilot to test the viability of exchanging COUNTER-compliant usage statistics through the IDS.

The pilot began in February 2025 with the creation of an IDS-compliant data space by OAEBUDT's technical vendor, [Think-It](#). Following creation of the data space, pilot participants each integrated with the pilot for the purposes of either *sending* or *receiving* data to or from other participants.

Clarke & Esposito (C&E) was engaged to complete an ROI study of pilot participants, to capture the benefits of the IDS integrations, and to summarize recommendations to the project team regarding common pain points or areas of opportunity to be used in future integrations.

2.2 Project Methodology

This project was conducted in the following stages:

- **Project Discovery:** A review of key aspects of data space participation to identify key hypotheses for exploration.
- **Pre-integration interviews** (*occurring at the point the pilot participant began active development of their data space integration*): A 45-60 minute interview with each pilot participant, to capture key goals and expectations of pilot participants.
- **Post-integration interviews** (*following successful integration*): A 45-60 minute follow-up interview with each participant to capture feedback, document time and resources needed, gauge time / cost savings opportunities, and determine the degree to which integration met participants' expectations. In addition, interviewees were invited to complete an *optional* pre-interview questionnaire in advance of post-integration interviews to help guide the interview discussion.

For both pre- and post-integration interviews, C&E employed a *semi-structured* format for interviews, touching on specific questions but not following a specific script. This format provided consistency in coverage of important topics but also enabled a degree of flexibility for deeper exploration of productive avenues of discussion as well as the ability to adapt questions to the role and specific circumstances of the interviewee.

The results of pre-and post-integration interviews are summarized in the accompanying **Case Studies Report** which will be made available on the [OAEBUDT project Zenodo](#).

In this **Pilot Evaluation and Return on Investment Analysis Report**, C&E captures key themes and benefits of the IDS integrations across all participants, and summarizes recommendations to the project team regarding common pain points and opportunities for future integrations.

3. PROJECT PARTICIPANTS

There were six original participants in the initial pilot, of whom five completed a test integration with the data space to exchange data with other participants. They represented a variety of different stakeholder types (publishers, aggregators, platform providers, analytics services) involved in the supply chain for OA books.

For the purposes of the pilot, participants played one of two roles:

- **Data Providers** who supplied usage data to another participant through the data space
- **Data Recipients** who received usage data from a provider through the data space

In one case, the participant participated as both a data provider and data recipient, across multiple exchanges).

The table below provides a brief summary of each participant, organization type, size, and pilot role (data provider or data recipient).

TABLE 1. Pilot Participants			
Participant Name	Organization Type	Size	Pilot Role
JSTOR	Digital Library	Large	Data provider
LibLynx	Analytics Services	Small	Data recipient
punctum books	Publisher	Small	Data recipient
Ubiquity Press	Publisher / Aggregator / Publishing Services	Large	Data provider
Michigan Publishing	University Press / Platform Provider	Medium	Data provider Data recipient

Knowledge Unlatched (KU) was also a pilot participant but did not complete a direct integration with the data space. Instead, Michigan Publishing provided data intended for KU to mutual partner LibLynx.

Participants in the pilot had a range of OA book programs and activities.

- Only one participant (punctum books) was 100% focused on OA books publishing. All other providers support OA books alongside other publication types (primarily journals) and business models (open access and paid-access).

- Three participants have launched new programs or models to support OA books:
 - **JSTOR**: Path to Open, in which it hosts books on a paid-access basis for three years, after which point books are converted to fully open access
 - **Michigan Publishing**: Fund to Mission, a collective action model in which the University of Michigan Press commits to publishing 75% of its annual scholarly books output as fully OA assuming annual funding targets are met.
 - **Ubiquity Press**: Ubiquity and parent company De Gruyter host the University Press Library Open (UPLO), a large online collection of free to read academic books by leading university presses
- LibLynx, as a technology provider, supports OA book publishers among their clients, but its usage analytics services are largely format- and business-model agnostic.

4. PILOT INTEGRATION

Think-It developed and launched the IDS-compliant data space between February and June 2025. Pilot integrations by each participant took place between June-August 2025.

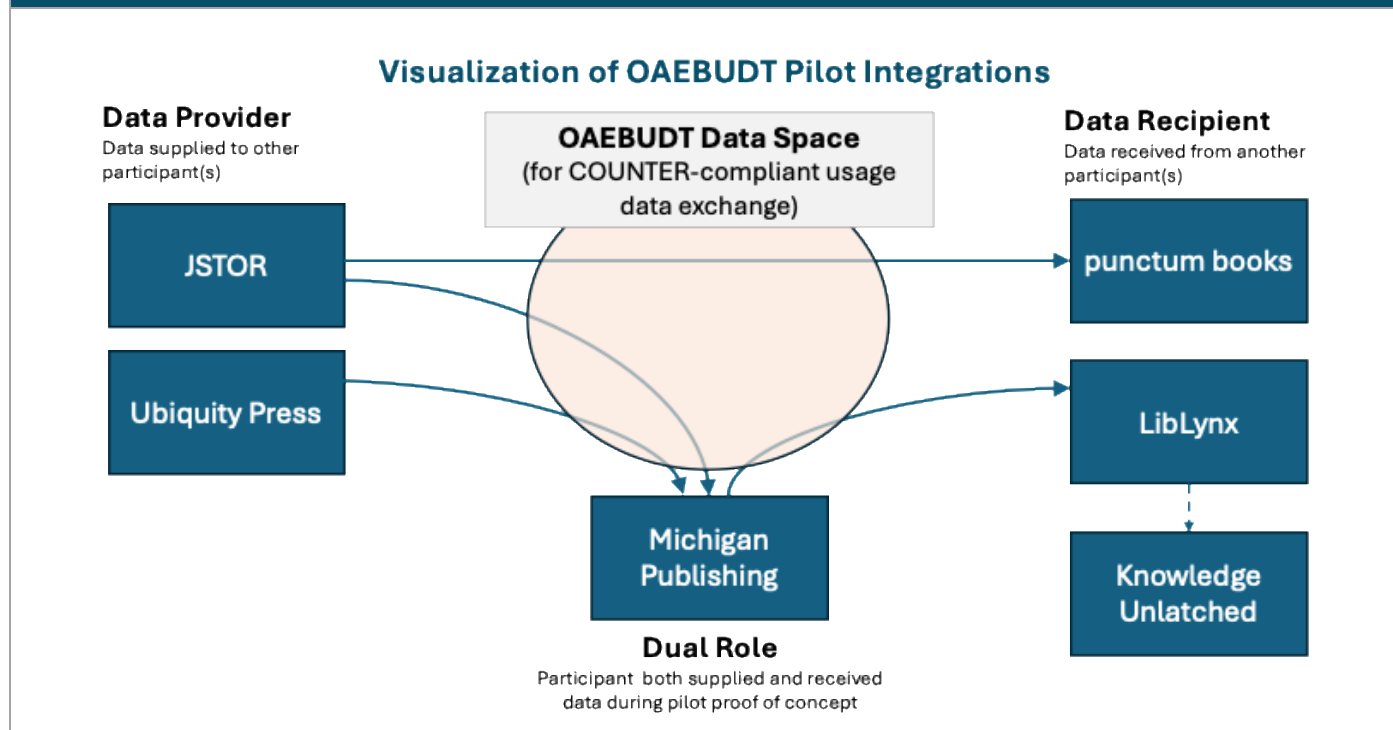
Each integration involved the exchange of one COUNTER-compliant report via the data space, with both parties following appropriate authentication steps to either initiate data flow (data providers) or to view the data made available to them (data recipients).

The following table summarizes the exchanges conducted during the pilot proof of concept.

TABLE 2. Data Space Pilot Exchange	
Data Provider	Data Recipient
JSTOR	punctum books
JSTOR	Michigan Publishing
Ubiquity Press	Michigan Publishing
University of Michigan Publishing	LibLynx (<i>on behalf of</i> Knowledge Unlatched)

An illustration of these exchanges is provided below in **Figure 1**.

FIGURE 1. Data Space Pilot Exchange



There were several limitations to note with pilot integrations:

- **Integrations were conducted as a proof-of-concept, not as longer term initiatives.** No participants continued to use the data space or maintained a “live” connection following their one-time data exchange.
- **Integrations were supported by manual oversight.** Several participants noted that additional work would be necessary to further automate the processes tested, including:
 - Negotiating report parameters or configuring report formats to match the specifications of the receiving party
 - Error alerts to support troubleshooting of each discrete authentication step
 - Use of a more scalable API client, ideally one developed and maintained specifically for the data space by OAEBUDT
- **“Live” integrations will require additional staff / stakeholder involvement.** For example, several pilot integrations were overseen and conducted by non-technical staff. Participants generally expected that a “live” integration would require additional support and technical development work not reflected in this pilot.

With these caveats in place, the table below illustrates the time/effort required for each of the five pilot participants completing integrations.

TABLE 3. Pilot Integration Time and Effort Required per Participant		
Participant	Time/Effort Required	Notes
JSTOR	2 FTE, 2 weeks	Time and effort reflects an “minimal viable product” (MVP) integration
LibLynx	16.5 hours	Did not involve LibLynx technical development team
punctum books	“minimal”	No technical work required for pilot
Ubiquity Press	12.5 hours	Additional work would be required for “live” integration, but the work completed in this project would be foundational to future integrations
Michigan Publishing	15 hours	No technical work conducted for pilot; “live” integration may need to involve support by third party services vendor (ScholarlyIQ)

In all, integrations ranged from “minimal” work (interpreted in this context to be less than five hours) required at the low end, up to ~2 weeks (with 2 FTE assigned) at the high end. The level of effort required did not appear to be influenced by whether the participant was playing the role of data provider or data recipient.

4.1 Process Feedback and Suggested Improvements

Pilot participants were generally positive about the integration experience, and all were able to successfully exchange data through the data space as a proof of concept. In pre-integration interviews, very few participants had a set budget for the integration – the primary exception being JSTOR, who had estimated 2 weeks of work for 2 FTE. Following the integration, JSTOR reported that this estimate was roughly correct, with some initial challenges for JSTOR’s development team offset by a re-scoping of the project.

The following table summarizes feedback provided by participants about the process and recommendations to guide future integrations.

TABLE 4. Participant Feedback on Data Space Pilot Integration

Integration step	Description	Recommended Improvements / Future Developments
Data Space Agreement	<ul style="list-style-type: none"> All parties completed a data space agreement prior to integration. Participants reported the agreement required minimal time to execute. 	<ul style="list-style-type: none"> One participant (Michigan Publishing) works with a third-party services provider (ScholarlyIQ, or SiQ) which distributes usage reports on its behalf. SiQ was not involved in the pilot due to the complications of striking the appropriate data space agreement. Another data space participant (LibLynx) provides usage processing services for clients, and did not fit neatly into any of the pre-defined roles anticipated in the data space agreement. Further development of roles may be necessary to better map to relationships specific to the scholarly publishing industry.
Technical documentation and instructions	<ul style="list-style-type: none"> Participants were generally happy with the provided documentation. It was described in terms such as “complete”, “thorough”, and “easy to follow.” Documentation was appropriate for non-technical staff, which is necessary to support participants without deep technical expertise. 	<ul style="list-style-type: none"> JSTOR’s development team found the “how-to” videos easier to use and more effective than the provided documentation. JSTOR also recommended that documentation include additional context and description of use cases, to help orient development teams unfamiliar with the IDS concept.

TABLE 4. Participant Feedback on Data Space Pilot Integration

Integration step	Description	Recommended Improvements / Future Developments
Technical setup	<ul style="list-style-type: none">• The pilot reflected a “minimum viable product” (MVP) setup based on cloud-based architecture designed to provide a proof of concept for the data exchange.	<ul style="list-style-type: none">• Two participants noted that the Postman API service was sufficient for the purposes of the pilot, but a “live” integration would require more robust architecture.• Creation of a user interface (UI) was out of scope for the pilot, but several participants noted the importance of this UI to support smaller organizations without deep technical expertise. (The UI was seen as less critical for integrations led by software engineers / developers.)
Project coordination and planning	<ul style="list-style-type: none">• Participants coordinated directly for the purposes of the pilot, sometimes conducting live videoconference walkthrough sessions supported by Think-It.	<ul style="list-style-type: none">• Participants anticipated a “live” integration would need to be streamlined. For example:<ul style="list-style-type: none">○ Failure notifications / alerts○ Confirmation of successful file delivery○ Defining report parameters and reaching agreement on data to be shared• Participants also welcomed more coordinated project management for future integrations.

5. PILOT GOALS AND OUTCOMES

In pre-integration interviews, participants discussed the goals and priorities of OAEBUDT development. In post-integration interviews, participants then commented on observed progress and benefits, both anticipated and unanticipated.

5.1 Reasons for Participation

Participants were asked to prioritize and provide commentary on a list of potential motivations to participate in the data space pilot. The table below summarizes the relative importance of each motivating factor across interviews.

TABLE 5. Data Space Pilot Participant Motivations	
Potential Motivating Factor	Average Rank (from 1-5; 1 = highest, 5 = lowest)
Improved accessibility of usage data	1.0
Improved data quality / accuracy	2.3
New business opportunities	2.7
Reduced labor costs for data management	3.3
Improved decision-making capabilities	4.0

Improved accessibility of usage data was universally the most important motivation across participants, and was chosen as the top reason to participate by all interviewees. These concerns typically stemmed from the complicated nature of ebook distribution, in which books can appear on multiple platforms, with each providing varying levels of visibility into usage data. In the words of one interviewee, *“My #1 concern is that these reports fully capture the usage out there, across all platforms, in a way that makes them accessible to all of us in the industry working with ebooks and distributing them.”*

Improved data quality and accuracy was also important to participants, even though data standardization is out of scope for the OAEBUDT data space itself. According to the [Participant Rulebook for the OA Book Usage Data Trust - International Data Space \(OAEBUDT-IDS\)](#), data providers are responsible for disclosing their adherence to defined trust indicators (accuracy, generation transparency, reliability, frequency, and granularity), but there is no minimum standard required for participation in the data space. For most participants, a project which simply *incentivizes* the culture and capabilities of better data quality and standardization is worth pursuing. (We note, however, that the provision of

poor-quality or inconsistent data by data providers will limit the potential utility to recipients – recipients can only do so much with meager data – so it is to OAEBUDT’s benefit to actively encourage data quality improvements among its pool of providers.)

Interestingly, **reduced labor costs for data management** was among the lowest priorities for integration. From the perspective of an ROI study, this suggests pilot participants did *not* expect to achieve significant time (and therefore cost) savings from participation in the data space. Similarly, **pursuit of new business opportunities** was of mixed importance: some providers considered it a high priority, while others saw it as the least important reason to participate.

Decision-making capabilities were the lowest-ranked priority among respondents, but this was typically because other criteria would typically need to be met first – specifically better data accessibility, and improvements to data accuracy or quality – before any impact to business or publishing strategy might be felt. In other words, the data space must first create a meaningful ecosystem of data exchange to have a practical effect on decision-making over the long term.

5.2 Participant Goals and Progress

In pre-integration interviews, participants were asked to describe their goals for pilot participation; progress towards these goals was then gauged in post-integration interviews.

Across all participants, these goals may be placed into a few broad categories:

- **Increasing access to usage data** (especially among participants acting as data providers)
- **Improving reporting standards and capabilities** (shared across both providers and recipients)
- **Increasing the efficiency and scalability of usage reporting** (from data providers, especially those hosting content on behalf of other parties)
- **Supporting development of industry infrastructure** (an ‘altruistic’ motivation; articulated to some degree by all participants but most often a concrete goal among data recipients)

The following table illustrates the degree to which the pilot was able to achieve each of these broad aims noted above, and what further action from the data space might be necessary to achieve full realization of each goal.

TABLE 6. Participant Goals, Progress, and Further Actions Needed

Goal Theme	Goals Identified by Participants	Reported Progress Towards Goal from OAEBUDT Pilot	Necessary Criteria to Achieve Goal in Future Phases
Increase access to usage data	<ul style="list-style-type: none"> • Broaden visibility and transparency of usage data • Maintain access to a broader set of usage data sources • Support current constituents • Improve reporting capabilities for partner presses 	The pilot did not demonstrate an expansion of access to usage data for any parties, as all connections tested involved existing partners.	<p>Forge new data sharing connections.</p> <p>Achieving this goal would require a high volume of participation, <i>especially</i> among parties without a clear mechanism for connection via existing channels.</p>
Improve reporting standards and capabilities	<ul style="list-style-type: none"> • Support better standardization of usage data • Incentivize improvements to ebook usage data reporting • Unlock potential of improved usage analytics 	While the data exchange does provide visibility into data quality and provenance, data standardization is <i>not a requirement</i> of the data space. As a result, all pilot data exchanges were to each data provider's existing standards.	<p>Incentivize consistency across providers.</p> <p>Without clear incentives for data providers, the data space will have limited ability to directly improve standards. Given the importance of data quality and accuracy among participants, it could be valuable for the data space to consider what role it can play in supporting greater standardization.</p>

TABLE 6. Participant Goals, Progress, and Further Actions Needed

Goal Theme	Goals Identified by Participants	Reported Progress Towards Goal from OAEBUDT Pilot	Necessary Criteria to Achieve Goal in Future Phases
Increase efficiency and scalability of usage reporting	<ul style="list-style-type: none"> • Share usage information more efficiently • Increase ability to scale data exchange • Reduce the labor cost of partner usage reporting 	Given that the pilot was not a "live" integration, there were no efficiencies immediately gained. Participants noted that adoption of the data space would be in parallel to existing reporting infrastructure until a critical mass of partners used the data space as their primary method of exchange.	Increase participation. Greater "live" participation in the data space will increase its ability to provide efficiency gains. Note this is a "chicken or the egg" challenge: for the data space to provide benefit, a critical mass of providers must participate, but providers will not participate unless they can be sure of the data space's benefit.
Supporting development of industry infrastructure	<ul style="list-style-type: none"> • Establish a proof of concept of data transfer protocols • Support open infrastructure for OA books • Support for publicly accessible usage sharing infrastructure • Support parallel OA book infrastructure initiatives (such as the Books Analytics Dashboard) 	The pilot did achieve success in providing a proof of concept of the data space concept as well as providing direct experience of the data space with pilot participants.	Build around a clear and practical use case. OAEBUDT infrastructure and data transfer protocols will only be useful to the extent by which they can support active workflows and relationships. Coordination with similar infrastructure initiatives (such as open source multi-provider book analytics dashboards) would provide a practical and visible use case for the data space.

5.3 Additional Benefits Not Anticipated

Participants also identified an additional benefit of the pilot not originally anticipated:

- **Auditability of usage exchange.** Part of the core functionality of the data space is to provide a record of what data was exchanged and with whom. This is a capability many providers are not set up for today: data are provided to multiple systems via self-service dashboards, automated systems, emails, or a mix of these, making it difficult to maintain visibility into fulfillment of data requests.

The value of this new capability is expected to increase with scale (meaning, the more data exchanged, the more valuable it is to have a record of those exchanges). It is also not limited to the sharing of usage data – an audit trail of exchanges can be valuable for any type of data exchange executed through the data space (one example noted was keeping track of which books have been shared for AI training, with which partners, for what period, and for what purposes).

6. RETURN ON INVESTMENT (ROI) ASSESSMENT

6.1 Pilot ROI – Results

Because of the limited scope of this pilot, it resulted in no changes to existing workflows, staffing, services, or capabilities; as such, participants were not in a position to see an ROI (whether positive or negative) from the pilot.

This is not to say that the pilot was unsuccessful. Several participants noted that their expectation was that the pilot a) was an investment in infrastructure, and b) was intended for the benefit of the whole community. In pre-integration interviews, **all participants were clear that investment in the pilot was not anticipated to generate an immediate return.**

6.2 ROI Expectations

As the OAEBUDT project continues to evolve, stakeholders in the data space will need to articulate the benefits of participation and help potential members develop a business case justifying further involvement and investment. The pilot has succeeded in better illustrating the potential avenues that this business justification might take based on concrete initial experience.

The table below summarizes the elements necessary to quantify the ROI offered by the data space, and the progress made towards defining these elements in the pilot phase.

TABLE 7. Return on Investment Progress

ROI Element	Description	Participant Feedback
Level of Investment	Participant's anticipated investment required to complete "live" integration with the data space	<p>The level of investment needs clarification.</p> <ul style="list-style-type: none"> Though pilot participation typically required low investment, "live" integration was expected to require much greater time / effort. Participants were generally unclear on total anticipated cost (i.e., future technical requirements, and membership / participation costs).
Return: Cost Savings	Participant's expectations of their ability to save costs through data space integration	<p>Cost savings is likely a long-term goal.</p> <ul style="list-style-type: none"> There may be a long timescale for any anticipated efficiencies. Data providers noted that they will need to support the data space alongside previously-existing workflows unless there is near-universal adoption.
Return: Revenue Growth	Participant's expectations of their ability to generate new revenue from existing products or services through data space integration	<p>This is not a (direct) path to increasing revenue.</p> <ul style="list-style-type: none"> Data Providers did not directly connect participation in the data space with increases in revenues. It is the hope among participants that increasing the visibility and utility of OA book usage data will help to provide "proof points" for further investment in OA models generally.
Return: Creation of New Opportunities	Participant's anticipation of whether the data space can create new commercial prospects or other strategic benefits (either directly to the organization or indirectly via industry-wide progress)	<p>Opening up new opportunities is the clearest path to creating benefit.</p> <ul style="list-style-type: none"> Most of the anticipated benefits of the data space relate to capabilities that are not possible today, which the data space could provide to participants. For example, several participants described potential benefits of accessing previously unavailable data sources, or gathering new information about where and when data is being shared.

In summary, the pilot has provided clarity about the potential benefits of the data space, which are largely realized through unlocking new capabilities and relationships – but there needs to be further clarity on anticipated costs and contributors for organizations to fully assess potential ROI.

That said, this assessment also reveals a fundamental paradox in developing the business case for new data-sharing capabilities: pilot participants all anticipated growing demand for enhanced data access via the IDS, but this demand remains largely hypothetical. This creates a classic “chicken or the egg” problem that complicates traditional ROI calculations: Participants expect to receive meaningful strategic benefit (and spread upkeep costs more widely) if the data space model for sharing data is widely adopted. But until the data space model is widely adopted, the benefits are low and the sustainability of the data space itself appears questionable.

Also complicating matters is that OAEBUDT infrastructure was initially developed with an intentional focus on the narrow use case of OA book usage (despite the potential for extensibility to other publication types and types of data exchange in the future).¹ Compounding the issue is that many of the OA book programs are themselves in their early stages, and pursuing relatively experimental business models; this makes it difficult to establish a baseline of expected “return.” The financial sustainability of OA book models—and by extension, the infrastructure supporting them—depends on multiple variables beyond data usage and accessibility: market adoption, pricing strategies, author behavior, funding mandates, and publisher appetite for experimentation and risk. Investment in OAEBUDT is one of many choices a publisher can make in support of its OA books program, in the hopes that the cumulative effect will result in a more sustainable ecosystem. This is likely why so many pilot participants considered “investment in infrastructure” a worthy pursuit; many providers still are navigating the path to viability of OA books *writ large*.

7. FINDINGS AND RECOMMENDATIONS

7.1 Establish incentives for future data space adoption

OAEBUDT would benefit from the involvement of early adopters who can provide clear and concrete case studies about how access to new information provided through the data space positively influenced customers or publishing partners with tangible business

¹ Note: In October 2025 the OAEBUDT Board of Trustees approved a new mission and vision for the initiative under the name “Scholarly Communications Trusted Dataspace”, signifying the expansion of support beyond OA books. As this study was conducted on a pilot among OA books market participants, we have maintained our focus on the ROI analysis for OA books in this report.

impact. A priority for future pilots should be to recruit data providers and data recipients who can together establish a clear use case for connection via the data space, to move past the hypothetical and illustrate how these new capabilities – once created – would impact the industry *in practice*.

Based on feedback from pilot participants, the most powerful incentives for joining the data space would be 1) to establish data sharing connections that would otherwise be complicated, or 2) to establish a highly desirable connection with one core provider (a “critical node”).

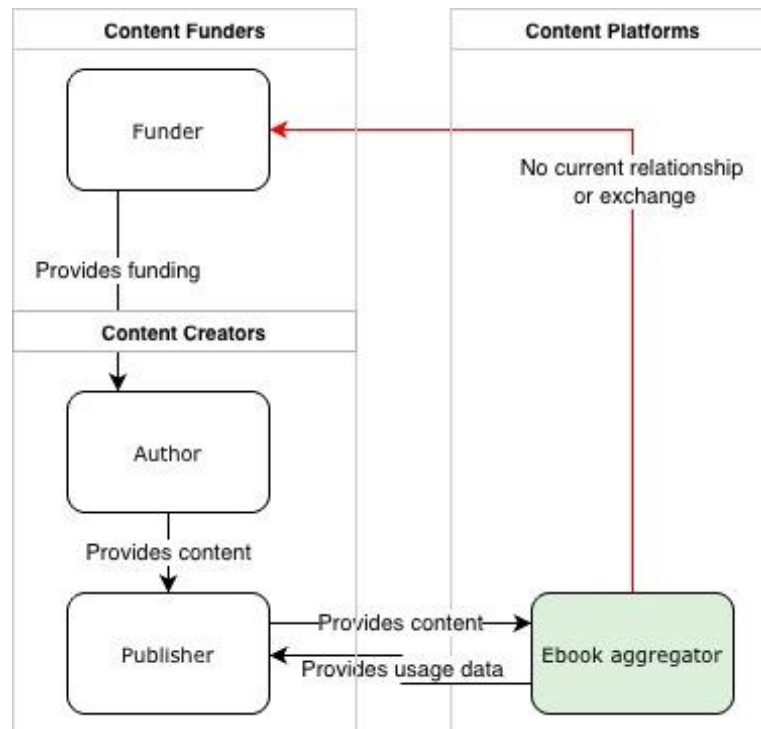
Incentive 1: Create data sharing opportunities between parties who have no existing relationship

In C&E’s [OA Books Supply Chain Mapping](#) report produced for the "Developing a Data Trust for Open Access Ebook Usage" project (the initiative preceding the OAEBUDT), relationships and data exchange between providers were explored and documented. The connections shown in that report represent *existing relationships*, where data exchange (however informal) is already taking place.

One of the strongest benefits identified by pilot participants is the opportunity for the data space to facilitate controlled exchange of data between parties *without* existing relationships in place.

As one example: As shown in **Figure 2** below, there is no current mechanism for an ebook aggregator to share usage data of relevant titles with organizations funding the publication of OA books.

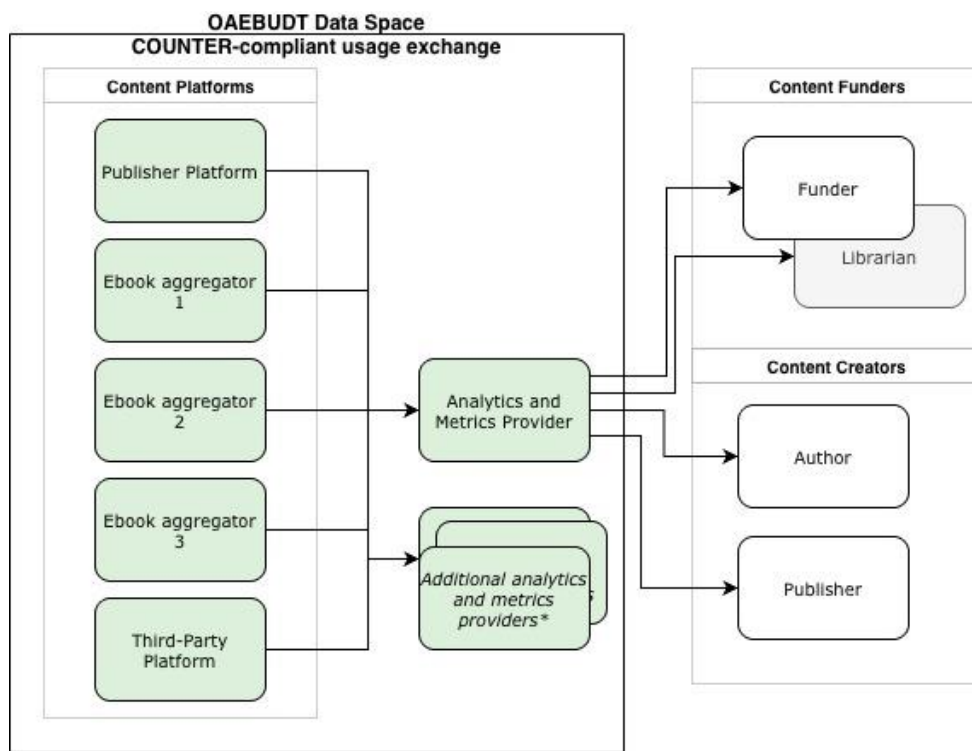
FIGURE 2. Example of OA Books Supply Chain Stakeholders with No Direct Relationship



A formal relationship for data sharing would require careful navigation of data ownership and entitlement. Could the data space provide a better mechanism for a funder to receive information about usage of the titles it has “opened”?

This does not necessarily need to be a direct connection between aggregators and funders. As shown below in **Figure 3**, one way the data space might address this issue is to simplify the connection between analytics and metrics providers (including those providing information for a funder audience) with content platforms (ebook aggregators, publisher platforms, other repositories).

FIGURE 3. Example of OAEBUDT Opportunity to Facilitate Data Exchange through Analytic and Metrics Providers



**Note: Though only one data exchange is shown in the figure above, data space infrastructure would not be limited to one path. The inclusion of additional analytics and metrics providers signifies the potential to support an unlimited number of providers in these roles.*

Interviewees noted parallel initiatives to build OA ebook usage dashboards, including the [Book Analytics Service](#) operated by OAPEN. The OAEBUDT project itself originated from a previous initiative to explore an OA ebooks usage dashboard. Would focusing on this potential use case provide a concrete benefit that would incentivize participation in the data space?

Given the high value of *receiving* usage data, it is relatively straightforward to make the case for integration as a data recipient. But, building a successful network requires the participation of *at least one usage data provider* (no matter how many data recipients are willing to onboard, the data must come from somewhere). The value of *sending* usage data can be harder to articulate clearly, especially when there is investment involved. Effective recruitment relies on identification of the most important (and untapped) connections for data providers, and a clear rationale for why the data space is a unique and compelling solution for building those connections.

Incentive 2: Establish a connection to a “critical node” participant

One way to address the “chicken or the egg” problem is for one strong use case to emerge in which connection through the data space is an essential requirement. There is similar precedent in the adoption of certain ebook standards: the ONIX metadata exchange standard was not widely used until Amazon, acting as a “critical node” for publishers, began to require it.

A “critical node” participant is an organization that can 1) generate a high volume of usage data with a broad network of direct and indirect connections with whom it might share that data, and 2) act as a strong enough ‘first mover’ to incentivize other organizations to participate in the data space. In the Amazon example, many strong organizations *could have* encouraged the adoption of ONIX among book publishers. It wasn’t until Amazon used its market power to *require* use of the ONIX standard that it became a lingua franca of book publishing. In other words, Amazon’s insistence on the use of ONIX started a flywheel of momentum which gradually built upon itself. The data space protocol, while not a standard, requires a similar flywheel of momentum.

To achieve critical mass, OAEUDT likely needs *at least one* “critical node” participant to play that first-mover role – not only in terms of technical support for building integrations and enabling live data exchange, but also in terms of financial support while the data space builds up to self-sustaining levels of membership. Multiple organizations working together for a similar purpose could collectively act as a “critical node” – for example, if a set of publishers and presses collectively developed a way to streamline reporting to a common funder, or a strong data / analytics provider required platforms and aggregators to provide data through the data space (as shown in the example of Figure 3). The role the “critical node” needs to play is to provide a concrete reason for another organization to prefer the data space above all other simpler alternatives, where the value of this connection overcomes the ROI hurdle of an initial integration.

The data space will need to appeal in particular to data providers with a **strong network** of direct and indirect connections, and the **market power** to incentivize data providers and recipients to connect via the data space (as opposed to current, simpler usage data sharing protocols). There are many roles within the supply chain which might support a “critical node” participant, as outlined in the table below (in no particular order):

TABLE 8. Potential Data Space “Critical Node” Participants

Supply Chain Role	Data Space Role	Potential Criteria for “Critical Node”	Example Providers
Funder	Data recipient	<ul style="list-style-type: none"> Organizations who are responsible for funding a large volume of OA works have an interest in monitoring the performance of those works. They also do not have direct relationships with the publishers or platforms who host the works they fund. Funders (or a group of funders) may wish to encourage these parties to provide information about content usage through the data space. It is possible some funders will be willing to complete their own integration with the data space to act as a data recipient. (This is most likely for organizations like Knowledge Unlatched or JISC, which fund content but also provide content platforms and usage dashboards.) Most funders will prefer a third-party analytics and metrics service to provide usage data for funded works rather than to build the infrastructure directly. 	<ul style="list-style-type: none"> Knowledge Unlatched JISC
Publishers with proprietary content platforms	Data provider	<ul style="list-style-type: none"> Publishers with their own proprietary platforms have control over the usage data generated by their content as well as over the technical roadmap to complete a data space integration. For the data space to make sense to these providers, they must have a meaningful number of technically-adept partners interested in receiving data. This might include large funders, distribution partners, or analytics/metrics providers. Specific use cases must be defined further to make a compelling case for OAEBUDT’s ROI. 	<ul style="list-style-type: none"> Taylor & Francis / Routledge Springer Nature <p><i>If expanding beyond OA books, additional candidates may include Elsevier, Wiley, Wolters Kluwer</i></p>

TABLE 8. Potential Data Space “Critical Node” Participants

Supply Chain Role	Data Space Role	Potential Criteria for “Critical Node”	Example Providers
Third-party platform provider	Data provider	<ul style="list-style-type: none"> • Third-party platforms host a significant amount of content on behalf of publisher partners, but do not have direct relationships with other potential recipients of the usage data they generate (i.e., their publisher partners’ partners). • There must be in return a meaningful number of technically-adept data recipients to connect with. (Or, potentially, a single “critical node” recipient could work, like an analytics / metrics service, so long as both parties can, over time, expand the network of total connections.) • Platform providers already have made significant investments in usage data reporting capabilities, so the data space must provide a unique benefit to other services. 	<ul style="list-style-type: none"> • Silverchair • Atypon • HighWire

TABLE 8. Potential Data Space “Critical Node” Participants

Supply Chain Role	Data Space Role	Potential Criteria for “Critical Node”	Example Providers
Aggregator	Data provider	<ul style="list-style-type: none"> As with third-party platforms, aggregators host content from multiple sources and generate a significant amount of usage of that content. However, typically they are not the primary “home” for that content (in contrast to publisher platforms or third-party platforms). The data space is not likely to provide a compelling case for data sharing within existing partnerships. Many aggregators work with small publishers who will not have the technical capabilities to integrate into the data space, or who will prefer to maintain current data sharing processes. Are there <i>new</i> connections with those publishers’ stakeholders that aggregators might be able to forge through the data space? 	<ul style="list-style-type: none"> EBSCO Ebook Central / ProQuest Project MUSE JSTOR <p><i>If expanding beyond OA books, additional candidates may include Ovid, ClinicalKey, ResearchGate</i></p>
Analytics and Metrics Services	Data recipient	<ul style="list-style-type: none"> Analytics and metrics services seek to gather information from a wide variety of sources. There is a clear advantage to these services in expanding their ability to gather data securely and efficiently across multiple parties. To succeed as a “critical node”, the analytics and metrics service must exert enough of a draw on data providers to motivate connection through the data space protocol. Enough of these data providers must themselves have the capability to integrate as well, and must cumulatively represent a meaningful proportion of overall engagement. (Usage metrics will have limited value to users if key information is missing.) 	<ul style="list-style-type: none"> Book Analytics Service Sensus Impact Digital Science Dimensions <p><i>If expanding beyond OA books: Altmetrics, Clarivate Web of Science, Elsevier Scopus</i></p>

7.2 Understand and establish success criteria

According to interviewees, to increase the appeal of data space integration and the potential return on investment for future participants, it will also be important to achieve the following prerequisites:

TABLE 9. Data Space Success Criteria

Success Criteria	Description
Low barrier to entry	Pilot participants frequently expressed concern about potential obstacles to integration for their data recipient / data provider counterparts. Many stakeholders within the OA books and other scholarly communications ecosystems have limited technical capabilities or wherewithal to undertake development work without a strong business incentive. OAEBUDT will find the most success if it can ensure that integration comes with a low barrier to entry (financially, technically, and otherwise) to overcome any potential objections from eligible participants.
Data quality	There is a wide range in quality of the usage data available from ebook providers throughout the industry, a frustration expressed by pilot participants. For the data space to be worth the effort of integration, the data received from data providers must be of good enough quality to be useful . Incomplete or insufficient data will create a disincentive for data recipients to participate.
Standardization across data sources	Similarly, data recipients may see very little utility in access to multiple data sources that cannot be consistently interpreted or combined. This is a common problem across ebook providers and many other scholarly publishing platforms. While the data space is intended to be neutral, with data quality indicators for each data source laid out in the participant rulebook but no minimum quality requirements, OAEBUDT should not ignore its potential to play a powerful convening role among data providers to improve the standardization of their data and improve its utility .
Cost / scope definition	At the time of writing, the sustainability model for the next phase in the OAEBUDT remains in development. With the OAEBUDT Board's recent reorientation toward a "Scholarly Communications Trusted Dataspace," this next phase will require further articulation of the implications and benefits within this larger remit: What are the other compelling problems that a data space can solve? Though a wider swathe of the scholarly publishing industry are now potential participants, these participants will still require clarity on the potential uses of the protocol, the benefits of the data space over alternatives, and an understanding of the short- and long-term costs to weigh in their cost / benefit analyses.

7.3 Recommendations for future development

Based on the learnings from the OAEBUGDT pilot and the opportunities it suggests for future development, we recommend the following actions to the OAEBUGDT team and stakeholders:

1. Clarify expectations about ongoing costs and sustainability.

Members of the pilot noted that the data space requires a clearer trajectory for sustainability, and more definition for how costs will be distributed between participants. The pilot enabled a greater understanding of the costs for the data space itself, but questions remain about the costs of technical work on the participants' side (which was not fully quantified by any pilot participant) as well as the perennial question of supporting ongoing infrastructure maintenance. Small- and medium-size organizations were particularly concerned about the risk of potentially high costs of participation, which would significantly impact their ability to justify the investment (especially given the many other financial pressures on scholarly publishing). Large organizations were equally concerned about ensuring a low barrier to entry for potential data sharing counterparts.

A short- to medium- term sustainability plan that relies on grant funding or subsidizes the start-up costs for smaller organizations could help to get over the initial hurdle. To achieve a *self-sustaining* critical mass of participants, the data space must clarify what participants can expect as the full costs of implementation and participation, while building a plan for the long term to protect against these investments going to waste.

2. Actively support improvements to provider data quality and standardization.

The data space will not dictate reporting standards or practices, and will instead provide a rubric for data providers to express the quality of the data shared through the IDS. However, data recipient participants can expect little benefit from increased access to poor data.

We recommend OAEBUGDT partner with similar initiatives (or else itself play a convening role) to increase the quality of usage data and improve the business case for efforts to share that data more broadly.

3. Onboard participants (both data providers and recipients) with at least one “critical node” incentivizing connection via the data space.

OAEBUGDT must overcome the “chicken or the egg” problem of articulating these benefits without a critical mass of participants already established. Recruitment of

at least one strong “critical node” participant to act as a first mover and incentivize others in its network will maximize the chances that this critical mass is reached. Most likely this critical node will play a data provider role – there must be at least one strong provider in place for the data space to function as intended – but *any* network which provides enough mutual connections should be explored and encouraged.

4. Expand beyond OA book usage to additional products and data types.

Focusing exclusively on the OA books ecosystem limits the data space’s utility to a relatively new market, with only a small number of potential “critical nodes”. Given the technical infrastructure is data-agnostic, there is likely more opportunity from expanding the universe of potential participants to any providers generating high volumes of usage data with a broad network of potential stakeholders. We recommend in its next phase of development that OAEBUDT continue to do the work to define data sharing needs and use cases outside of the OA books ecosystem, and clearly articulate why and how the data space can provide a unique solution for clear and compelling problems.