Distributed Usage Logging Update
Count Every Download

Chris Shillum
VP Identity and Platform Strategy, Elsevier
Chair, DUL Working Group
6 Nov 2018

Counting Every Download

- Reality: content lives on multiple platforms

- What if we were able to collaborate and count all these downloads from anywhere?

- And make them visible to publishers and their customers, and for better metrics?

readcube

MENDELEY

Kopernio

Academia.edu
share research

Dimensions

Semantic Scholar
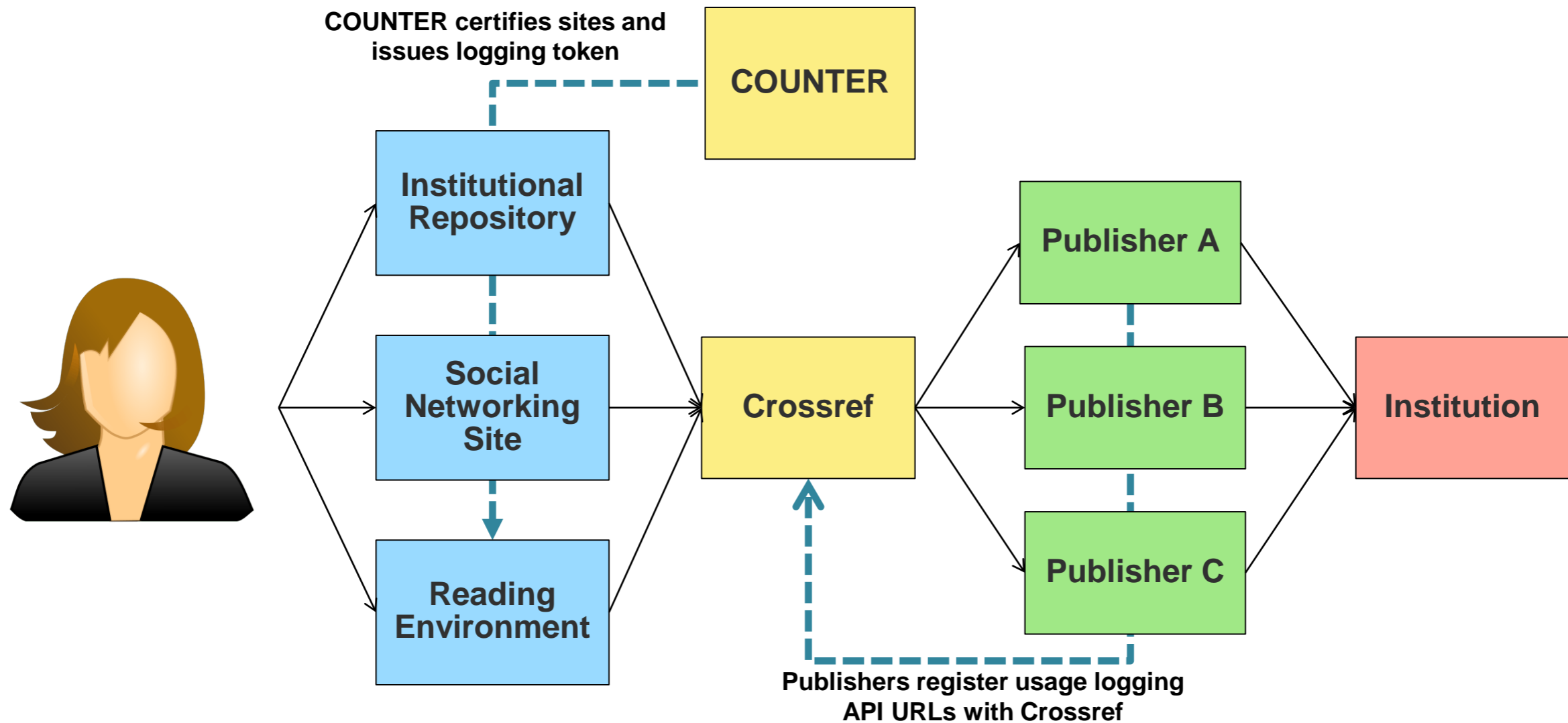
¹findr

WEB OF SCIENCE™

ResearchGate

The Problem

- Researchers have more ways to access scholarly content and new ways to share and collaborate on non-publisher platforms:
  - **Scholarly Collaboration Networks** – Academia.edu, ResearchGate, Mendeley, etc.
  - **Reading environments and tools** – ReadCube, EndNote, etc.
  - **Aggregator platforms** – EBSCOhost, IngentaConnect, etc
  - **Institutional Repositories**

- As more content is accessed through these non-traditional platforms, assessing true usage is extremely difficult.
  - Most libraries evaluate their subscribed content using standard publisher-provided COUNTER reports that provide only a partial picture,

- This leads to inaccurate assessments of the value of content under both subscription and open access business models.
  - **Publishers** are unable to demonstrate the full value of their content to library customers.
  - **Authors** are unable to get a full picture of usage of their articles
  - **Institutions** do not have a complete picture of usage when making purchasing decisions

The Idea

Build on the Crossref infrastructure to create a framework which allows usage information to flow from the point of usage (the alternative platforms) to the publishers, from where the data can be aggregated and incorporated into existing COUNTER usage reporting streams.

1. Researchers access articles on site of choice	2. Sites look up usage logging endpoint in Crossref metadata	3. Sites log usage via generic API Including DOI, IP address, Institutional ID	4. Publishers include third-party site usage in COUNTER reports sent to customers

Taking the Initiative Forward

Collaboration between multiple industry sectors is essential for success

Role of COUNTER

- Define semantics of usage logging messages
- Determine criteria for participation in the scheme
- Define CoP and oversee compliance auditing process

Role of CrossRef

- Define syntax of usage logging messages
- Build and operate technical infrastructure
- Define technical API specs
- Provide training and documentation on technical integration

Role of Platform Vendors

- Integrate with logging API
- Send usage events via API to Publishers
- Adhere to COUNTER defined CoP

Role of Publishers

- Integrate with logging API
- Receive usage events from API
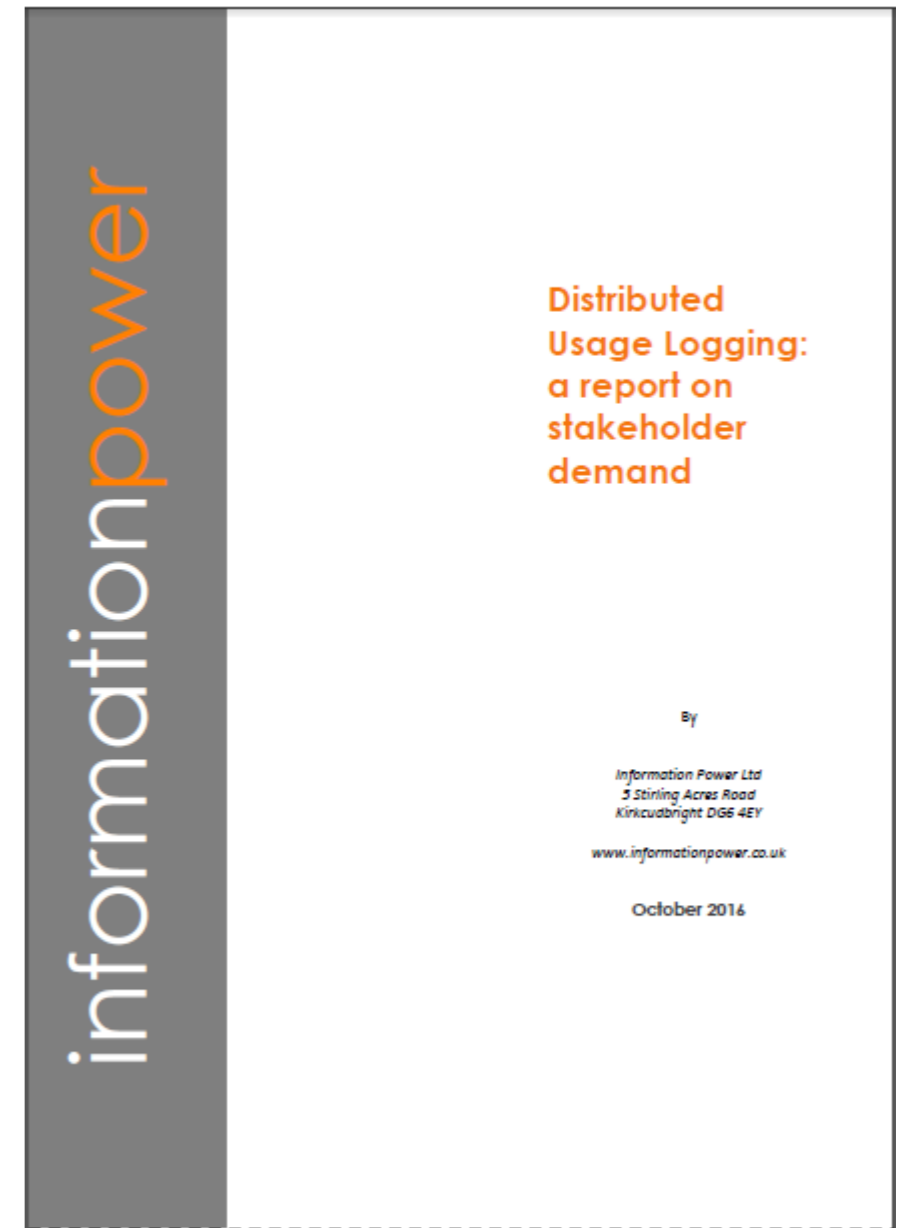- Incorporate into existing COUNTER-compliant usage reporting stream

Progress to Date: MVP/Proof of Concept – Early 2017

- Working group formed
- Initial definition of the type of "usage" data within scope agreed
- Focus groups, interviews and survey conducted to understand demand for DUL among different stakeholder groups
- Requirements drafted for including non-publisher usage sources in COUNTER reports
- Analysis and recommendation made for preserving user privacy
- DUL envelope and message specifications drafted
- Recommendations made for message authentication
- Proof of concept test to demonstrate full end-to-end transaction pipeline with validation credentials

Stakeholder Demand for Distributed Usage Logging

- Face-to-face focus group, a webinar and an online survey in fall of 2016

- Over 200 respondents, including libraries, library consortia, IRs, publishers and aggregators

- Over 88% of survey respondents said they would be interested in reports on usage on SCNs and aggregator platforms

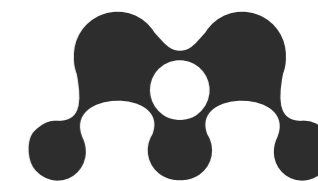- Majority (70%) wanted this information to be reported separately to usage on publisher platforms

Distributed Usage Logging: a report on stakeholder demand

By

Information Power Ltd
5 Stirling Acres Road
Kirkcudbright DG6 4EY

www.informationpower.co.uk

October 2016

https://www.projectcounter.org/distributed-usage-logging-report-stakeholder-demand/

Progress to Date: Beta – 2018

Core group of participants are architecting and implementing the full Distributed Usage Logging transactional framework:

- Finalized technical specifications and security infrastructure for expressing and verifying the authenticity of DUL messages
- Released final policy for processing and reporting on DUL data in COUNTER COP5
- Developed technical infrastructure for testing the registration and discovery of logging end-points
- Released final DUL reference manual consolidating all technical documentation
- Enabled publisher registration of DUL URL endpoints
- Set up a non-production handle system for testing
- Communicated updates on the initiative to the broader community

Participant	Role	Coverage
Mendeley	Producer	Downloads from Mendeley private libraries and public groups
Digital Science	Producer	Downloads from ReadCube and Papers private groups Accesses via Dimensions and Access Anywhere
Atypon	Consumer	Content from Atypon publishers including Wiley and ACS
Elsevier	Consumer	Elsevier content
Springer	Consumer	Springer Content

DUL Protocol design

Feature	Benefit
• Providers may send either disaggregated usage events and pre-aggregated usage snippets	• Light-weight implementation for **provider platforms** that don't have a pre-existing COUNTER usage reporting pipeline
• Events are sent directly from provider to receiving publisher	• Centralized infrastructure is **light-weight** responsible only for receiver end-point lookup
• Messages include sufficient information to apply COUNTER du-dupe rules whilst preserving user privacy	• **Receiving publishers** can re-use existing COUNTER reporting pipelines
• Protocol includes message authentication using industry standard tokens and PKI infrastructure	• Prevents gaming and injection of **fraudulent usage** into publisher reports

DUL Incorporated into COUNTER COP5

- DUL-captured usage may appear on Master Reports.

- DUL-captured usage captured that appears on Master Reports MUST be reported under the platform name where the transaction occurred.

- An organization that supplies usage transactions using DUL MUST include their platform identifier with each transaction and their platform MUST be registered with COUNTER.

- Reporting usage through DUL is OPTIONAL.

- The publisher receiving transactions through DUL is responsible for performing COUNTER processing to eliminate double-clicks, eliminate robot/crawler or other rogue usage, and perform the actions to identify unique item and unique title.

- Publishers that plan to include usage reported through DUL in their COUNTER Master Reports are responsible for ensuring that DUL-reported usage is included in the audit.

Sample Report

Report_Name	Title Master Report
Report_ID	TR
Release	5
Institution_Name	Sample University
Institution_ID	isni=1234567890
Metric_Types	*as selected*
Report_Filters	*as selected*
Report_Attributes	as selected
Exceptions	
Reporting_Period	2017-01-01 to 2017-06-30
Created	2017-05-25
Created_By	Publisher Platform Alpha

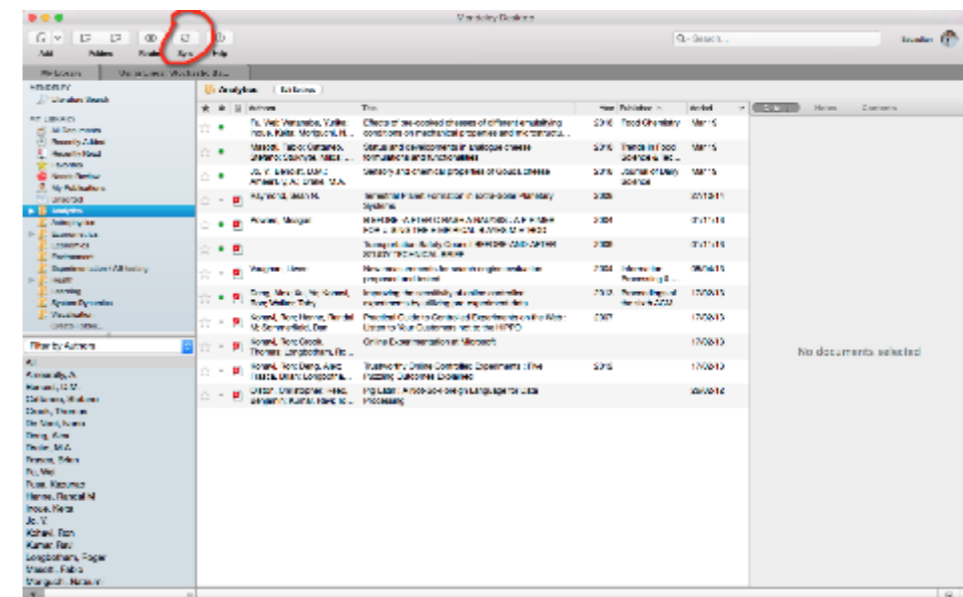| Title	Publisher	Publisher_ID	Platform	DOI	Proprietary_ID	ISBN	Print_ISSN	Online_ISSN	URI	Data_Type	Section_Type	YOP	Access_Type	Access_Method	Metric_Type	R
Journal Six	Gamma	1234_gam	PPα										Controlled	Regular	Unique_ Item_Requests	
Journal Six	Gamma	1234_gam	Institutional Repository A										Free_to_read	Regular	Unique_ Item_Requests	
Journal Six	Gamma	1234_gam	Sharing Platform Aleph										Controlled	Regular	Unique_ Item_Requests	

User journeys captured as DUL events – Mendeley Example

Mendeley: "Open PDF"

→ User opens Mendeley Desktop Library

→ User opens PDF by double-clicking on a document with PDF icon OR

→ by switching to another tab with an already opened PDF

Mendeley Library: Download PDF

→ User opens Mendeley Desktop OR user clicks "Sync" button

→ PDF is downloaded in the background

Benefits

For Libraries

- Make a fuller assessment of the value of content under both OA and subscription models
- Make more informed collection development decisions

For authors

- Gain a fuller picture of the usage of their work
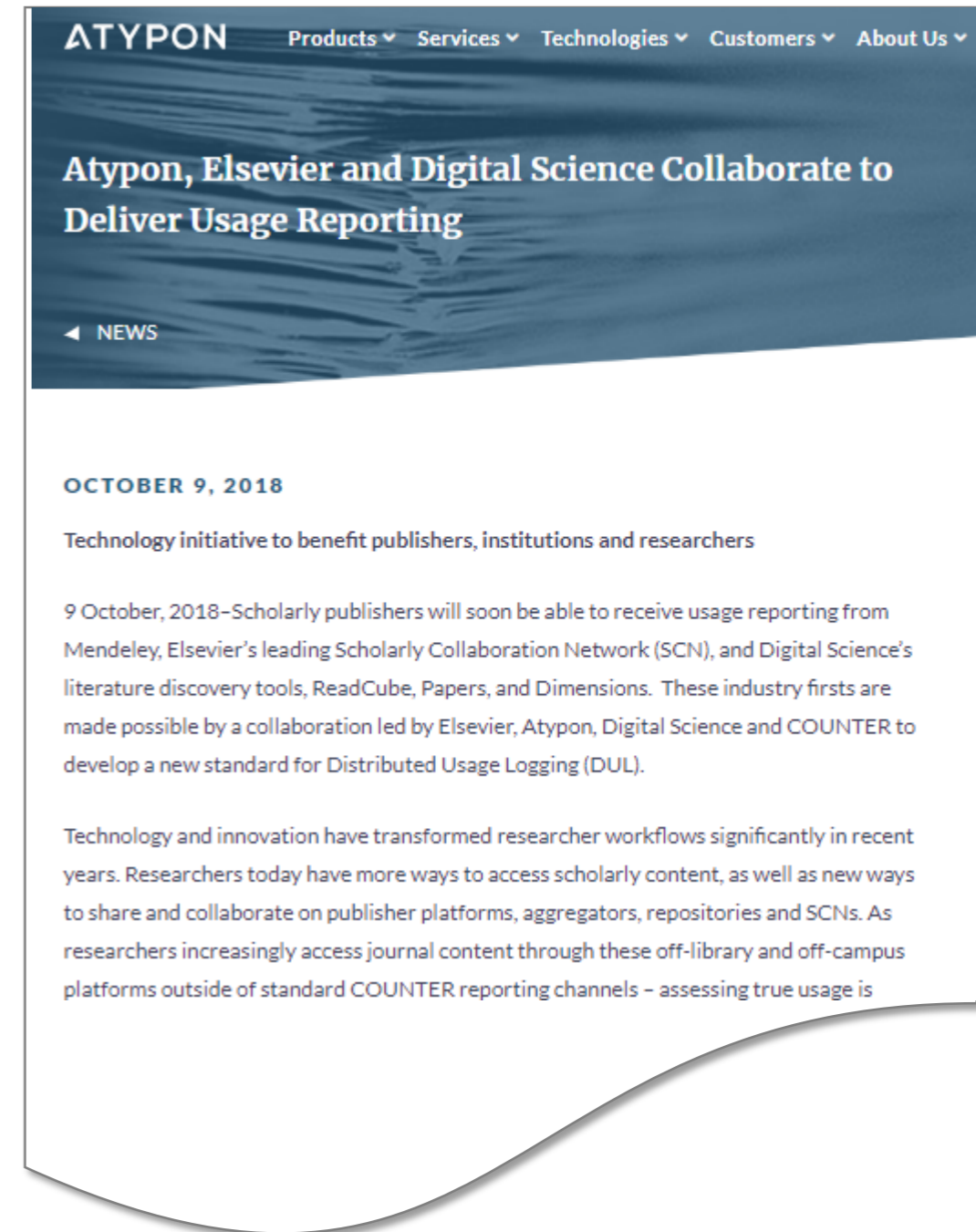
For Publishers

- Provide customers with a wider range of usage information in a COUNTER compliant way
- Provide authors with more comprehensive insights into the consumption of their articles
- Gain better insight into how a journals are consumed to inform portfolio development strategies

For platform providers

- Provide flexibility to users to consume content in their preferred context
- Demonstrate value to libraries and authors
- Fully participate in the scholarly comms ecosystem and gain support from content providers

Next Steps

- Atypon, Elsevier and Digital Science have announced their intention to productionize the service in early 2019

- Crossref Board due to take a decision on future participation at their November meeting. Considering three options:
  1. Full ownership, including outreach, policy settings, etc.
  2. Provide technical infrastructure only
  3. No further involvement

https://www.atypon.com/news/atypon-elsevier-and-digital-science-collaborate-to-deliver-usage-reporting/

13

?

Questions