

# Reproducibility in Practice: A Panel Discussion STM Innovations 3 December 2019, London



Rebecca Grant **SPRINGER NATURE**

Neil Chue Hong



Software  
Sustainability  
Institute

Leslie McIntosh



ripeta

David Smith



The Institution of  
Engineering and Technology

Chris Graf



**WILEY**

# Publishers are faced with a tremendous opportunity ...

- Reproducibility in the sciences is moving from an academic conversation to a reality – albeit slowly
  - The literature dates back to the mid-90s
  - Action has been limited, but we are on the cusp of step-wise advances
- The U.S. National Academies of Science created an AdHoc Committee in 2017 to address Reproducibility in Science
  - The AdHoc was in response to a request from the U.S. Congress
  - 234-page report delivered in June 2018

# Recommendations of the National Academies include several roles for publishers

- Publishers have a golden opportunity to be a key enabler of Reproducibility
  - “Publishers can have a strong influence on behavior”
  - “assess the reproducibility of a manuscript’s results before the manuscript is published”
  - “Shifting rewards and incentives will require thoughtful changes on the part of research institutions, working with funders and publishers”
  - The effectiveness of solutions “depends on whether they are clear, easy to follow, and harmonized across funders and publishers”
- Our panel today will look at the problem from several perspectives

Download the National Academies full report

<https://sites.nationalacademies.org/sites/reproducibility-in-science/index.htm>



## Reproducibility vs Replicability

**Reproducibility** is obtaining consistent results using the same input data, computational steps, methods, and code, and conditions of analysis.

**Replicability** is obtaining consistent results across studies aimed at answering the same scientific question, each of which has obtained its own data.

The National Academies of Sciences, Engineering and Medicine, Reproducibility and Replicability in Science: <https://www.nap.edu/catalog/25303/reproducibility-and-replicability-in-science>

# Reproducibility Badging Effort

- A NISO workgroup to define a recommended practice
- Group formed in March 2019
  - Group composed of Researchers, publishers, librarians
- Recommended practice (hopefully) out for public comment by mid-December 2019

# 17 workgroup members representing wide range of stakeholders



Association for  
Computing Machinery



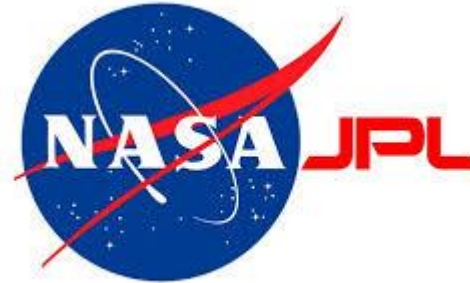
Sandia  
National  
Laboratories



RUTGERS  
THE STATE UNIVERSITY  
OF NEW JERSEY



ELSEVIER



ILLINOIS  
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

SPRINGER NATURE

NIST  
National Institute of  
Standards and Technology  
U.S. Department of Commerce





# Proposed Badge 1: Open Research Objects

- This badge signals that all relevant author-created digital objects used in the research (including data and code) are permanently archived in a public repository that assigns a global identifier and guarantees persistence.
- **Note:** This is akin to author-supplied supplemental materials, shared under a standard public license such as an OSI-approved license for code and a Creative Commons license or public-domain dedication for data and other materials.
- Corresponds to the ACM “Artifacts Available” badge, and to the combined COS “Open Data” and “Open Materials” (pertaining to digital objects) badges.

# Badge 2: Research Objects Reviewed

- This badge signals that all relevant author-created digital objects used in the research (including data and code) were reviewed according to the criteria provided by the badge issuer. The badge metadata should link to the award criteria.
- **Note:** A publication may be awarded the Research Objects Reviewed badge, while not being eligible for the Open Research Objects badge, and vice-versa.
- This badge corresponds to the ACM “Artifacts Evaluated” badge.
-

# Badge 3: Results Reproduced

- An additional step was taken or facilitated by badge issuer (e.g., publisher, trusted third-party certifier) to regenerate computational results, using the author-created research objects. (*Results Reproduced* assumes that that the research objects were also reviewed).

# Proposed Badge 4: Findings replicated

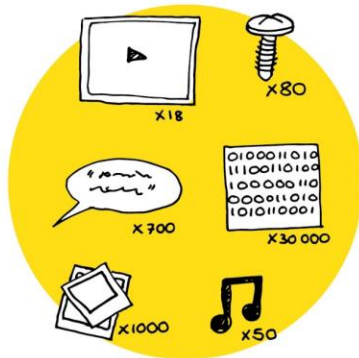
- This badge signals that an independent study, aimed at answering the same scientific question, has obtained consistent results leading to the same findings (potentially using new artifacts or methods). The badge links to the persistent identifier for that secondary publication. This badge is awarded by the publisher of the original work that is being badged.
- **Note:** Partially replicable findings, where an independent effort has been made to regenerate the findings, but has not succeeded, should be made visible in some way in the scholarly record, but we don't recommend a badge in this instance. -
- A replication study is published separately, linking to the original study via citation. Best practice for journals is to accept replication studies for publication, as a separate article type.
-

“Typically we remove huge chunks of what was done in order to communicate through formats that we have predetermined are the appropriate formats”

Kristen Ratan, Coko Foundation ALPSP Conference Sept 2017

<https://youtu.be/oIPWNebn-Yo?t=1h18m0s>

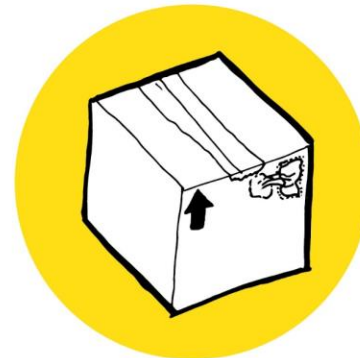
# TRADITIONAL MEDIA



Media is captured using traditional tools.



A linear programme is produced from the media.

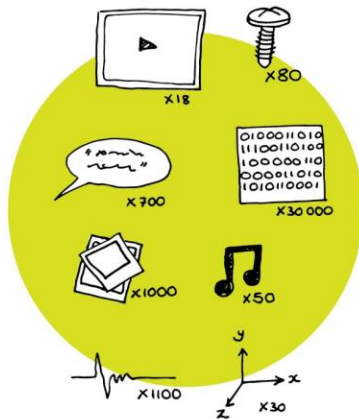


The programme is broadcast to everyone.

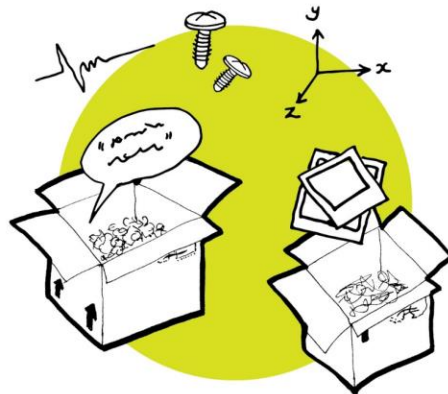


The same content is played back on all devices, resulting in compromises for some experiences.

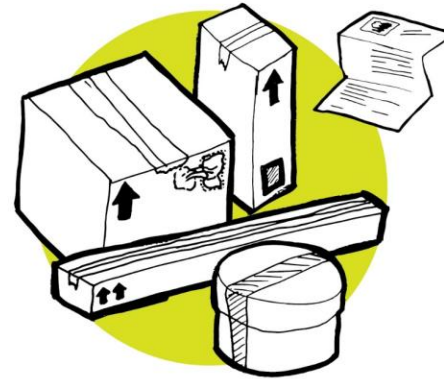
# OBJECT BASED MEDIA



Media is captured using new and traditional tools. Metadata is produced and recorded.



These are packaged as a collection of objects.



The objects are broadcast to everyone, accompanied by the metadata, which describes all the ways in which the objects can be assembled.



Individual devices in the home assemble objects according to the metadata, producing the best experience for the viewer in the context of their devices, environment, and preference.

**TECHNICAL & BUSINESS SUPPORT**

- Support
- Professional Services
- Optimization Guidance
- Partner Ecosystem
- Training & Certification
- Solutions Management
- Account Management
- Security & Billing Reports
- Personalized Dashboard

**MARKETPLACE**

- Business Apps
- Business Intelligence
- DevOps Tools
- Security
- Networking
- Databases
- Storage

**ANALYTICS**

- Data Warehousing
- Elasticsearch
- Business Intelligence
- Data Pipelines
- Hadoop/Spark
- Interactive SQL Queries
- Streaming Data Analysis
- ETL
- Streaming Data Collection

**DEV/OPS**

- One-click App Deployment
- Resource Templates
- Build and Test
- Application Lifecycle Management
- DevOps Resource Management
- Triggers
- Containers
- Analyze and Debug

**MOBILE SERVICES**

- API Gateway
- Single Integrated Console
- Identity
- Sync
- Mobile Analytics
- Mobile App Testing
- Targeted Push Notifications

**IoT**

- Rules Engine
- Device Shadows
- Device SDKs
- Device Gateway
- Registry
- Local Compute

**AI**

- Machine Learning
- Image Recognition
- Text to Speech
- Conversational Interface
- Deep Learning Frameworks

**ENTERPRISE APPS**

- Virtual Desktops
- Sharing & Collaboration
- Corporate Email
- App Streaming
- Communications

**HYBRID ARCHITECTURE**

- Data Integration
- Integrated Networking
- Integrated Identity & Access
- Integrated Resource & Deployment Management
- Integrated Devices & Edge Systems

**MIGRATION**

- Schema Conversion
- Exabyte-Scale Data Migration
- Application Migration
- Database Migration
- Server Migration

**APP SERVICES**

- Queuing & Notifications
- Email
- Workflow
- Transcoding
- Search

**INFRASTRUCTURE**

- Regions
- Availability Zones
- Points of Presence

**CORE SERVICES**

- Compute**  
VMs, Auto-scaling, Load Balancing, Containers, Virtual Private Servers, Batch Computing, Cloud Functions, Elastic GPUs, Edge Computing
- Storage**  
Object, Blocks, File, Archivals, Import/Export, Exabyte-scale data transfer
- Networking**  
VPC, DX, DNS
- Databases**  
Relational, NoSQL, Caching, Migration, PostgreSQL compatible
- CDN**

**SECURITY & COMPLIANCE**

- Identity Management
- Access Control
- Monitoring & Logs
- Assessment & Reporting
- Web Application Firewall
- Configuration Compliance
- Key Management & Storage
- Account Grouping
- Resource & Usage Auditing
- DDOS Protection

**MANAGEMENT TOOLS**

- Manage Resources
- Service Catalogue
- Configuration Tracking
- Monitoring
- Server Management
- Resource Templates