

Queen's in the Global Knowledge Commons

Data Day 2016, Queen's University

Martha Whitehead, Vice-Provost and University Librarian

Outline

What this is really about: research quality and impact

Open Science context of research data management

Broadening the concept of scholarly communications

Institutional RDM services emerging at Queen's

National context and Portage Network opportunities

International perspectives

My Lens: Information Access

Effective and sustainable scholarly communication

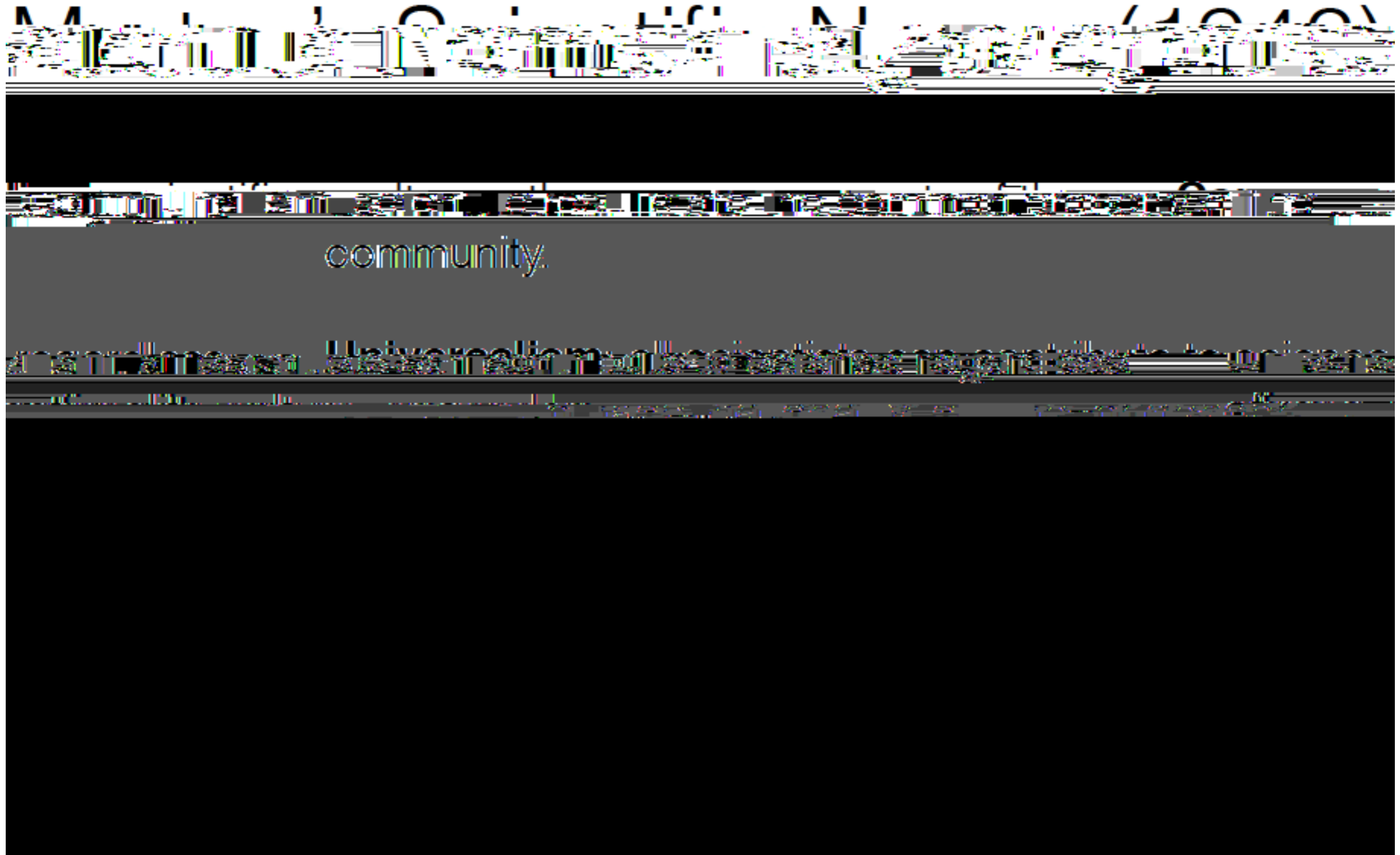
Public policy that enables information access

Respect for intellectual property

Protection of privacy

Preservation





Credit (this and next two slides): Victoria Stodden, *Defining the Scholarly Record for Computational Research*, CNI Annual Meeting, April 4, 2016

Skepticism -> Reproducibility



the Transactions of the Royal Society in the 1660's

Really Reproducible Research

David Benoit, 1993, *Journal of the American Statistical Association*, 88(423), 1093-1098

Introduction: John G. Norous

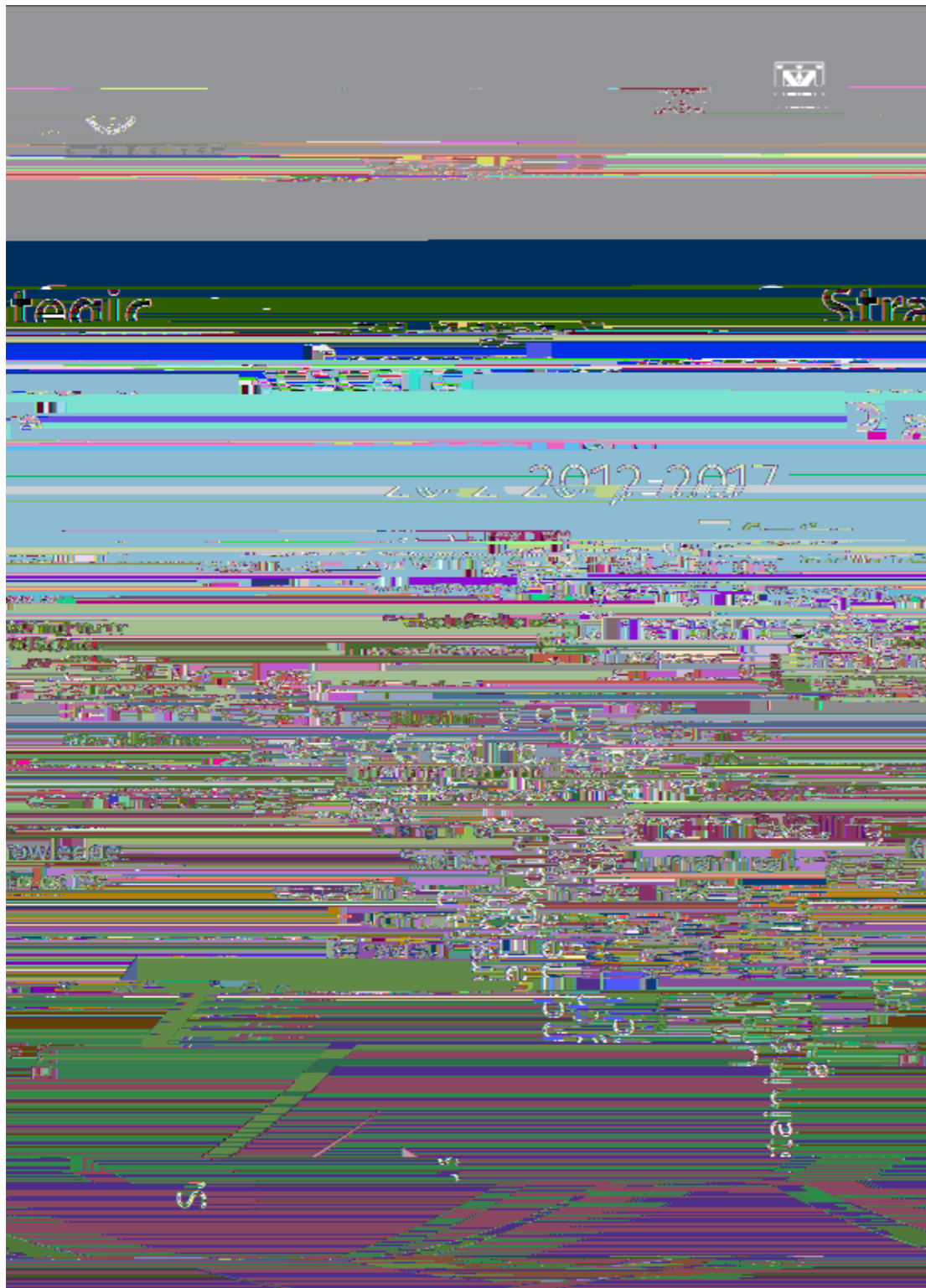
"The idea is: An article about computational science in a

journal of the profession should include a section that contains

the complete set of instructions and data which generated

the figures." David Benoit, 1993

Institutional Context



Research/IT/Library intersecting responsibilities

- Queens researchers must be able to participate fully in the research process as researchers and as librarians.
- We must integrate research and research practice.
- We must provide research and research practice.
- We must provide research and research practice.
- We must provide research and research practice.

Research Data Management Services

Library, guided by working group including University Research Services representatives, works closely **with researchers** to help with

- *Planning for data management*
data management plans and grant applications
- *Preparing your data*
coding, formatting, anonymizing, and more
- *Documenting your data*
describing your data so others can understand and use it
- *Archiving your data*
finding an appropriate repository or archive
- *Making your data discoverable*
promoting open access, indexing, assigning DOI's



Scholars Portal Dataverse

A Service of the Ontario Council of University Libraries

Scholars Portal Dataverse

A Service of the Ontario Council of University Libraries



SD Dataverse Network

Search this Dataverse Network

Advanced Search

The Ontario Portal Dataverse network is a repository for research data collected by institutions in Ontario. Access to data and associated documentation are freely available to all users. For more information on Ontario Portal Dataverse, visit [Ontario Portal Dataverse](#).

For more information on Ontario Portal Dataverse, visit [Ontario Portal Dataverse](#).

Create Dataverse

Studies

Dataverse

Call Us Now

RECENTLY RELEASED STUDIES

[Wilfrid Laurier University Library Research Data Repository](#) Aug 7, 2015

[on Canada](#)

[Public Data Collection program](#) Jul 20, 2015

[Canadian Career College Students Survey, 2008 \(Canada\) - Phase I, II](#)

[Open Education Dataverse](#) Aug 13, 2015

[Post-Secondary Education](#)

[Canada Millennium Scholarship Foundation](#) Aug 13, 2015

[Student Financial Review - 2011-2012 \(Canada\)](#)

Ontario Library Research Cloud



- Established to provide cost-effective long-term storage for digital assets
- Large-scale cloud storage service using low-cost technologies
- Storage nodes at multiple Ontario campuses
- Advanced text mining tools
- Utilizes industry-standard APIs and interfaces to maximize compatibility with existing library use cases, including institutional repositories

Accelerating the pace of Canadian research

The **Centre for Advanced Computing** (formerly HPCVL) at Queen's University is a consortium comprised of Carleton

Institutional concerns

How can we ramp up to support funder guidelines for data management?

Where is the infrastructure? (technical and human)

Isn't this a national / international concern?

Will researchers engage with RDM issues?

Bigger than any one institution

- Resources vary by institution and region, and nobody has it solved completely
- Managing research data is complex (diversity of format, metadata, analytical tools): researchers and service providers need help
- Researchers and their data cross institutional, regional and disciplinary boundaries
- Preservation storage is not yet addressed

*Research data management is an element
of national digital infrastructure*

Open Science

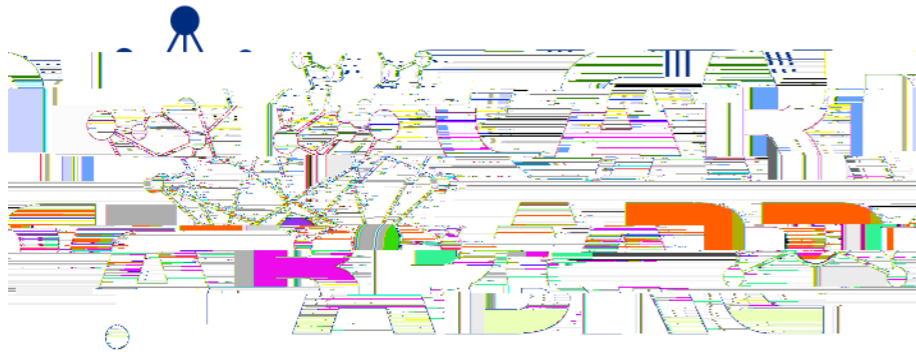
Canada's Action Plan on Open Government 2014-2016



will establish a government-wide approach to Open Science to increase access to federally-funded scientific

Tri-Agencies (CIHR, NSERC, SSHRC)

- Consultation in 2013 (Report: *Capitalizing on Big Data: Toward a Policy Framework for Advancing Digital Scholarship in Canada*)
- *Tri-Agency Statement of Principles on Digital Data Management*



portage

A dark, abstract graphic with a red and blue gradient background, featuring a small colorful grid icon in the top right corner. The graphic is composed of several overlapping rectangular shapes in shades of red, blue, and purple, creating a layered effect. The overall aesthetic is modern and digital.

LES PARTAGES POUR LES DONNÉES DE RECHERCHE
SHARED STEWARDSHIP OF RESEARCH DATA

What is Portage?

A national, library-based research data management network that coalesces initiatives in research data management to build capacity and to coordinate activities better.

2013 (Dec): Concept meeting at the ARC Hotel

2014 (Mar): Launch of Project ARC, a one-year pilot

2015 (May)

Portage in Canada's RDM ecosystem

Research data are a public good

Intelligent access: openness, with respect for privacy

Collaborative approaches: cost savings and shared expertise

Inclusiveness: aim to serve all researchers and create a more level playing field

Other RDM principles

Tri-agency Statement of Principles on Digital Data Management

- Canadian Institutes of Health Research
- Natural Sciences and Engineering Research Council of Canada
- Social Sciences and Humanities Research Council of Canada
- January 2016

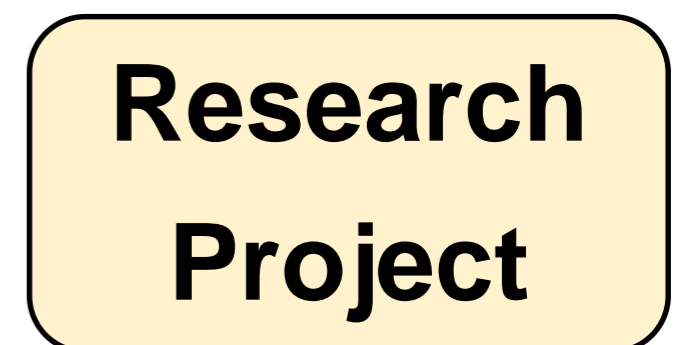
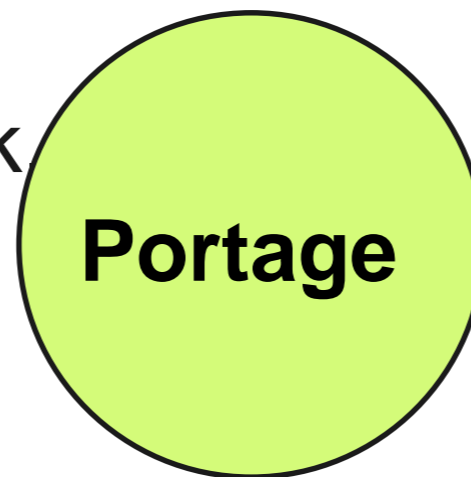
Research Data Management in Canadian Universities: A statement of principles

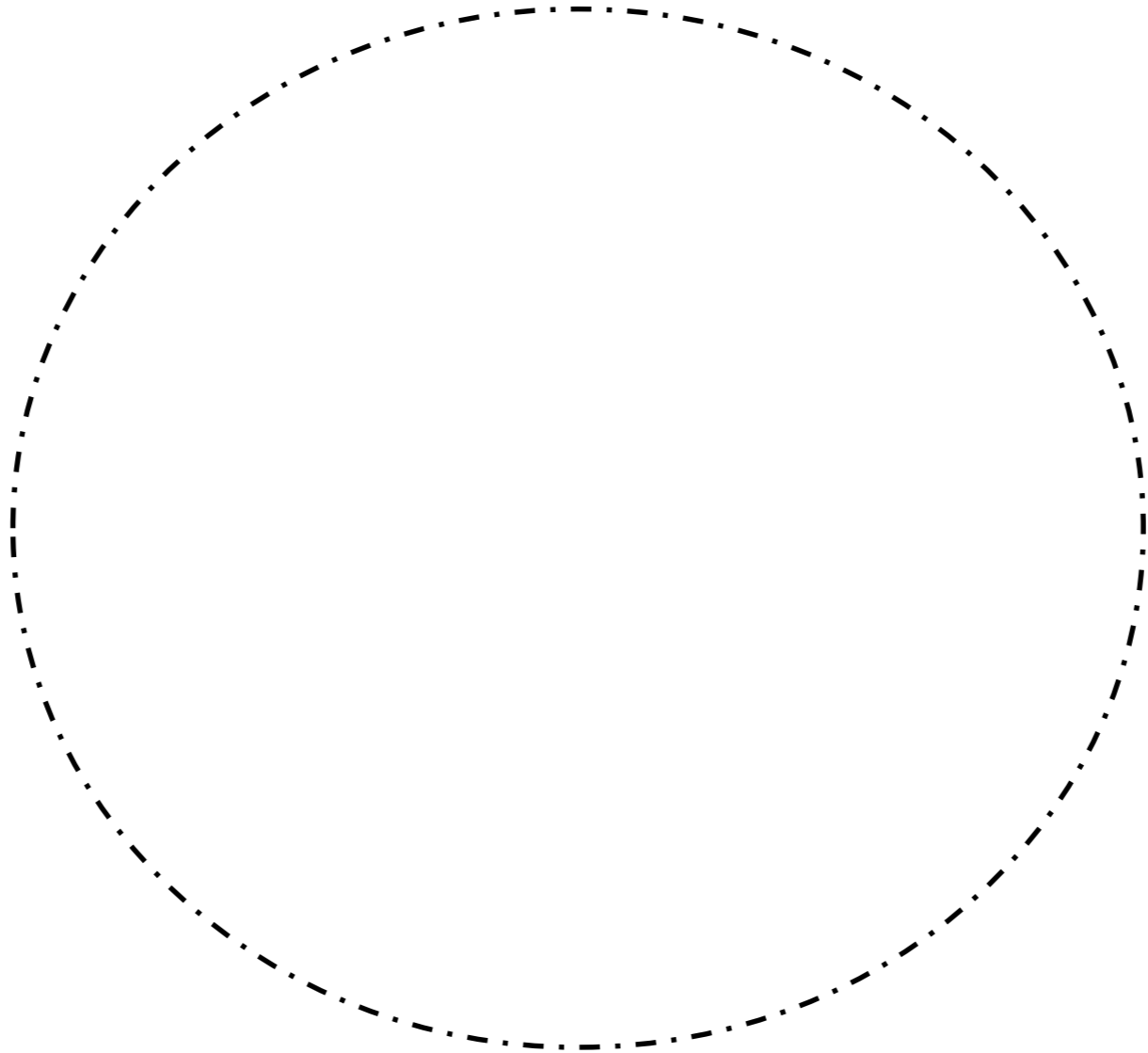
- Research Data Canada and Vice-Presidents of Research from the U15
- March 2016

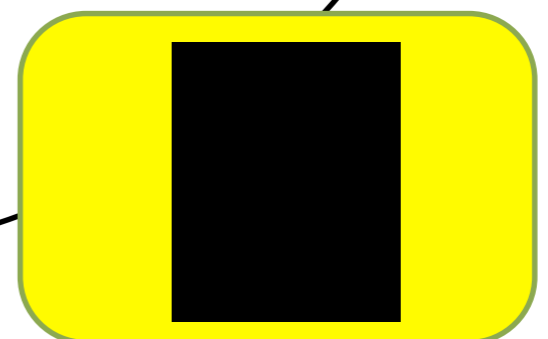
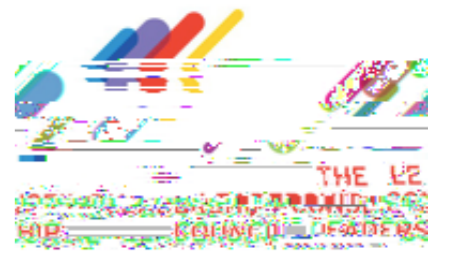
Is there a best level to collaborate?

Portage aims to provide infrastructure and services at the levels of the research project, the institution, and the network.

By network, we mean collaborating in regional and national partnerships.







Libraries

Researchers

Grad Studies

IT

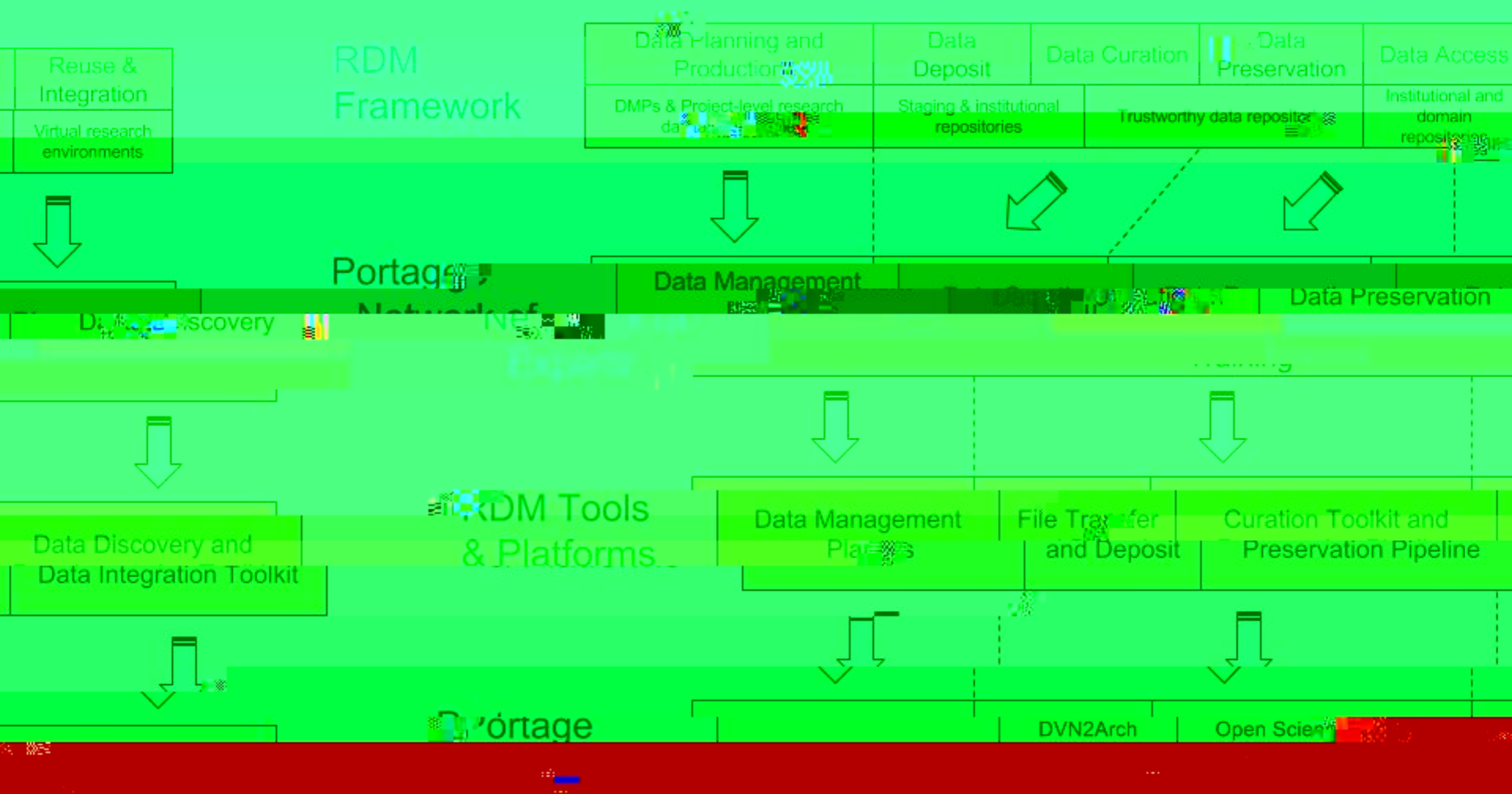
Ethics

RSO



Portage RDM Framework

Portage Network Research Data Management Framework



Portage Network of Expertise

Capitalize on the research data management expertise and services in libraries.

Expert Groups currently operational

Data Management Plan Expert Group

Data Preservation Expert Group

Data Discovery Expert Group

RDM Learning Expert Group

Expert Groups forthcoming

Data Curation Expert Group

Ethics and Data Expert Group

Shared stateful bits of research data

My plans

Grants

About

Help

Grants

Grants

Grants

successfully created

Plan was

My plans / Grants

Plan details

Plan Data Management

Grants

Plan name

My plan (FOI) (FOI) (FOI)

ID

Grant number

Principal Investigator/Researcher

Infrastructure platforms and tools

Develop RDM platforms and tools through partnerships and memoranda of understanding.

Current partnerships in place or nearly signed

DMP Assistant (UAL-CARL)

National Research Data Repository Prototype (CC-CARL)

Preservation Pipeline (CC-OCUL/SP-UTL-CARL)

Dataverse-to-Archivematica middleware (OCUL/SP-CARL)

Forthcoming

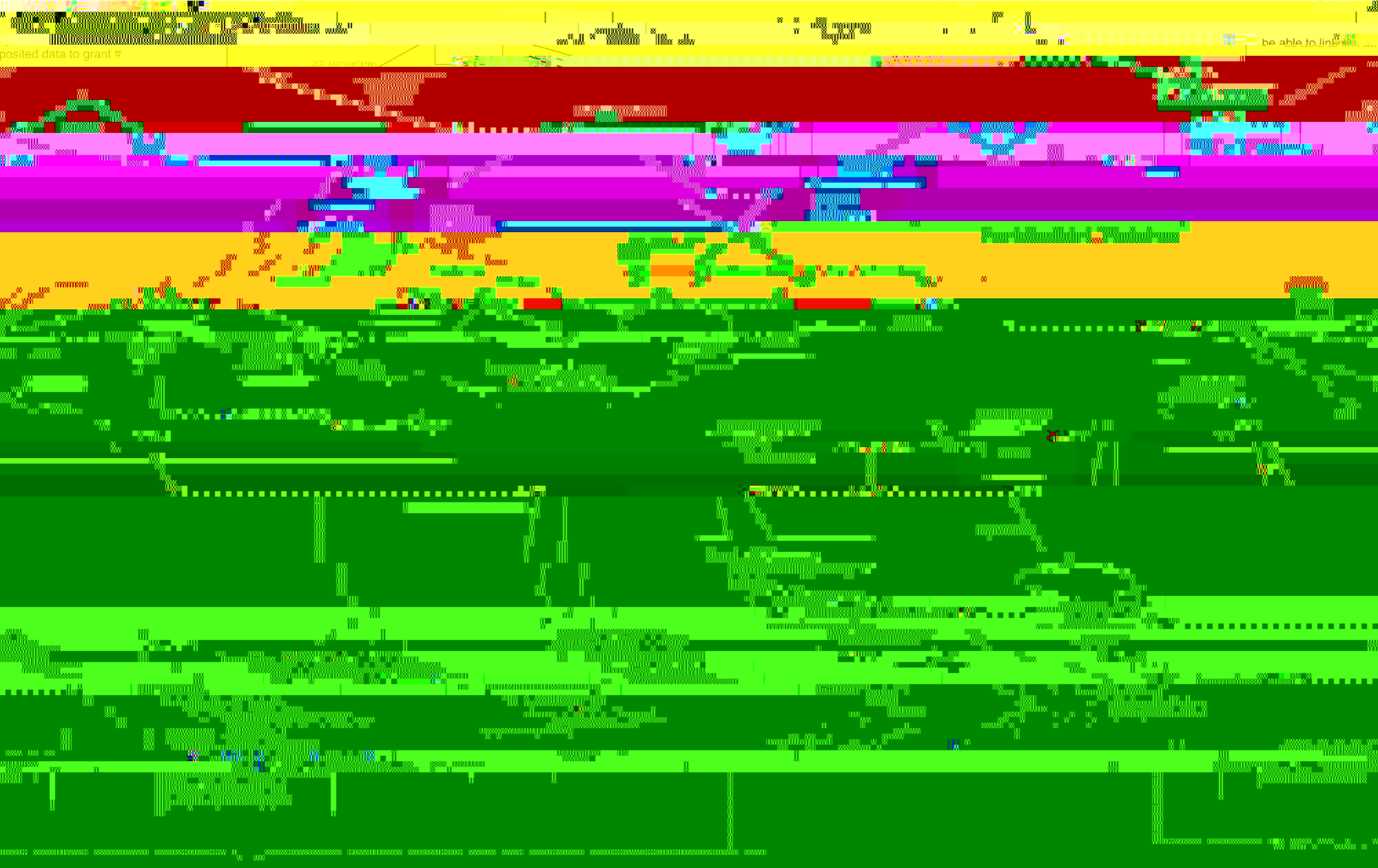
PREFER: Policy Format Register

Federated data discovery system

Data curators toolkit

to deposit data to the interoperability layer

deposited data to grant #



be able to link



•

•

Stakeholder engagement

Build a community of practice for research data management through well established working relationships among stakeholders across research sectors.

Regional library consortia

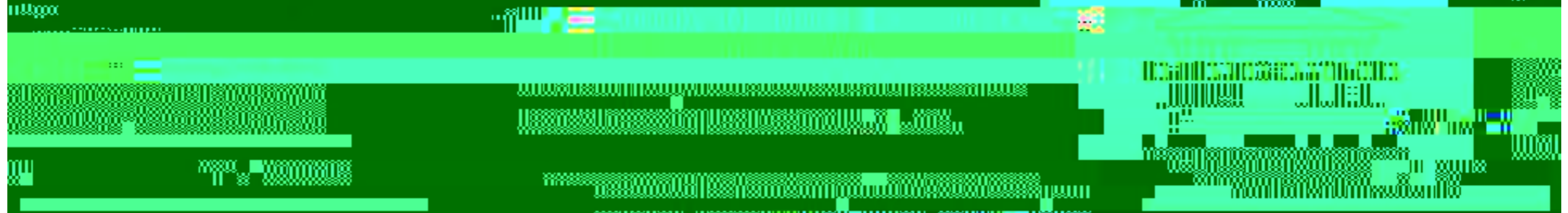
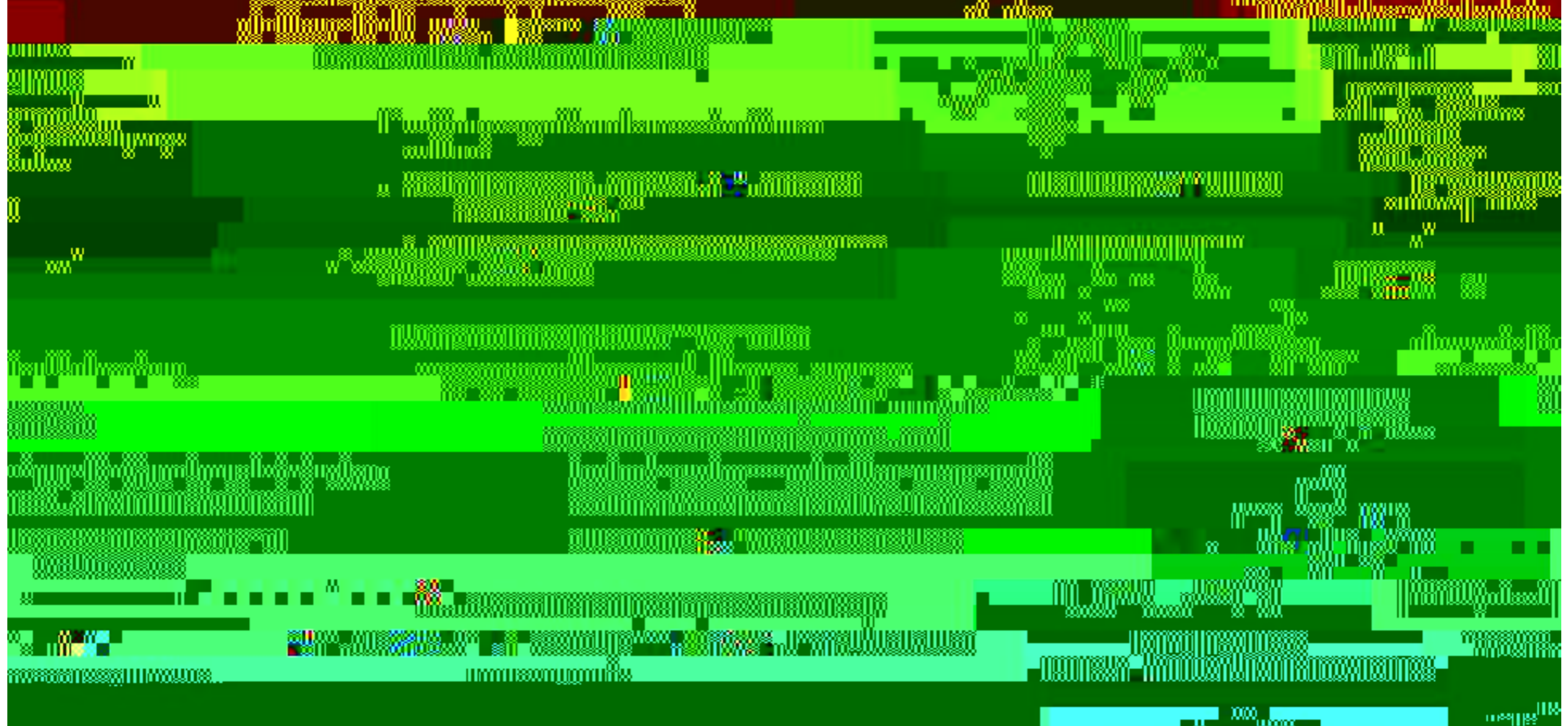
Individual academic libraries

Academic RDM stakeholders, e.g., VPR's, REBs, Scholarly Associations, Networks of Centres of Excellence

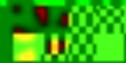
Other RDM stakeholders, e.g., Research Data Canada, Leadership Council for Digital Infrastructure



THE 100

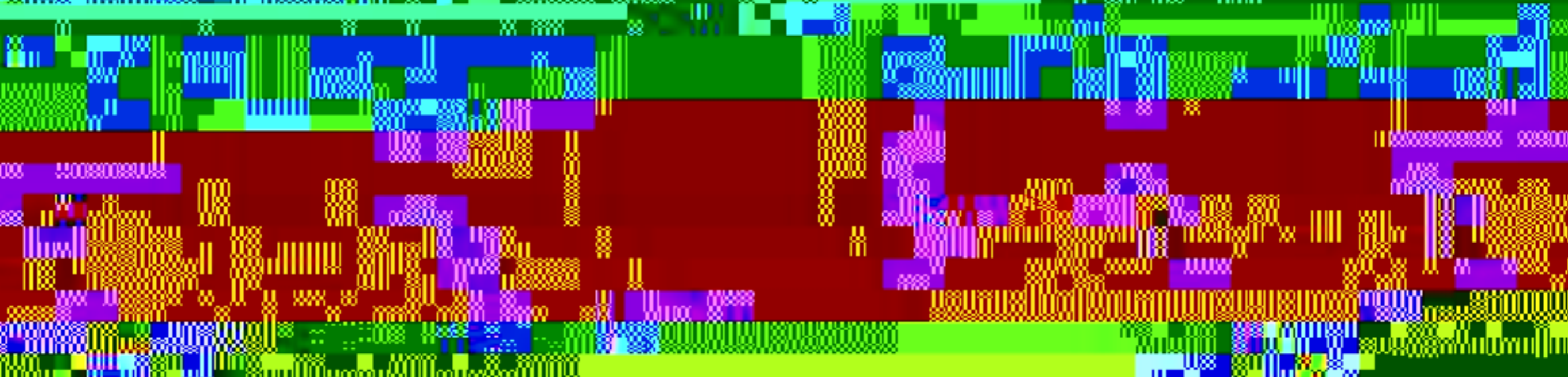
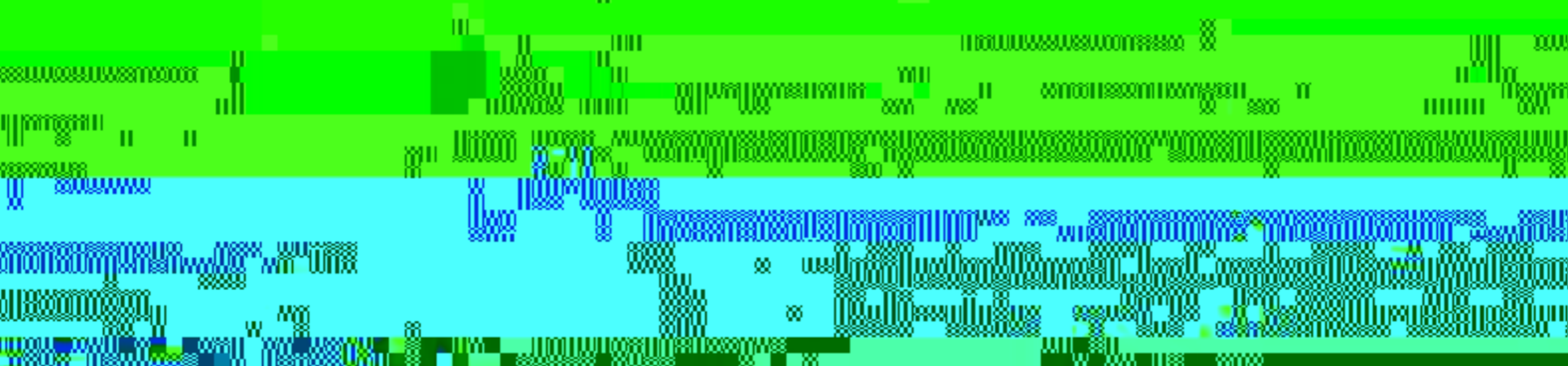


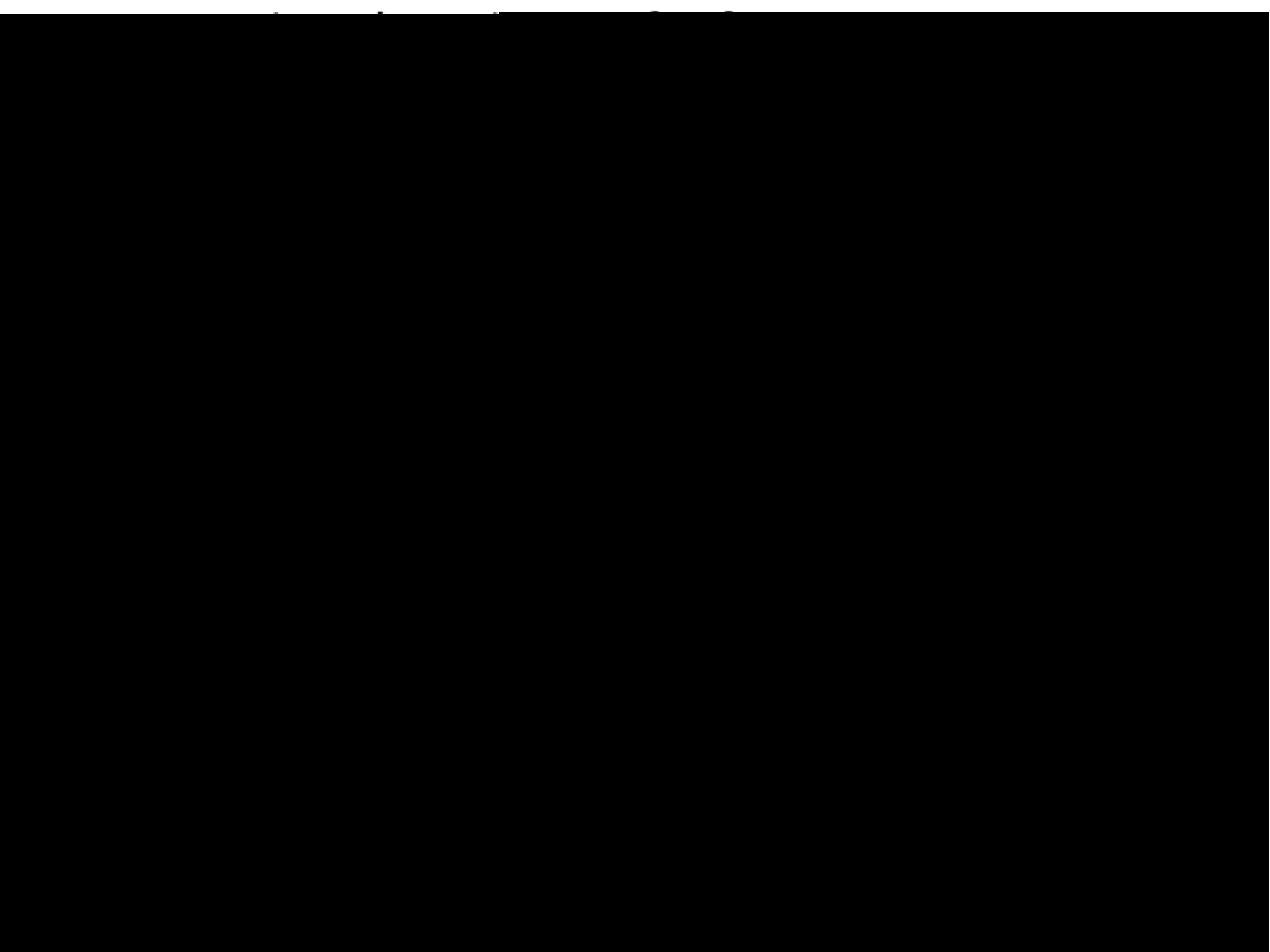
UK University

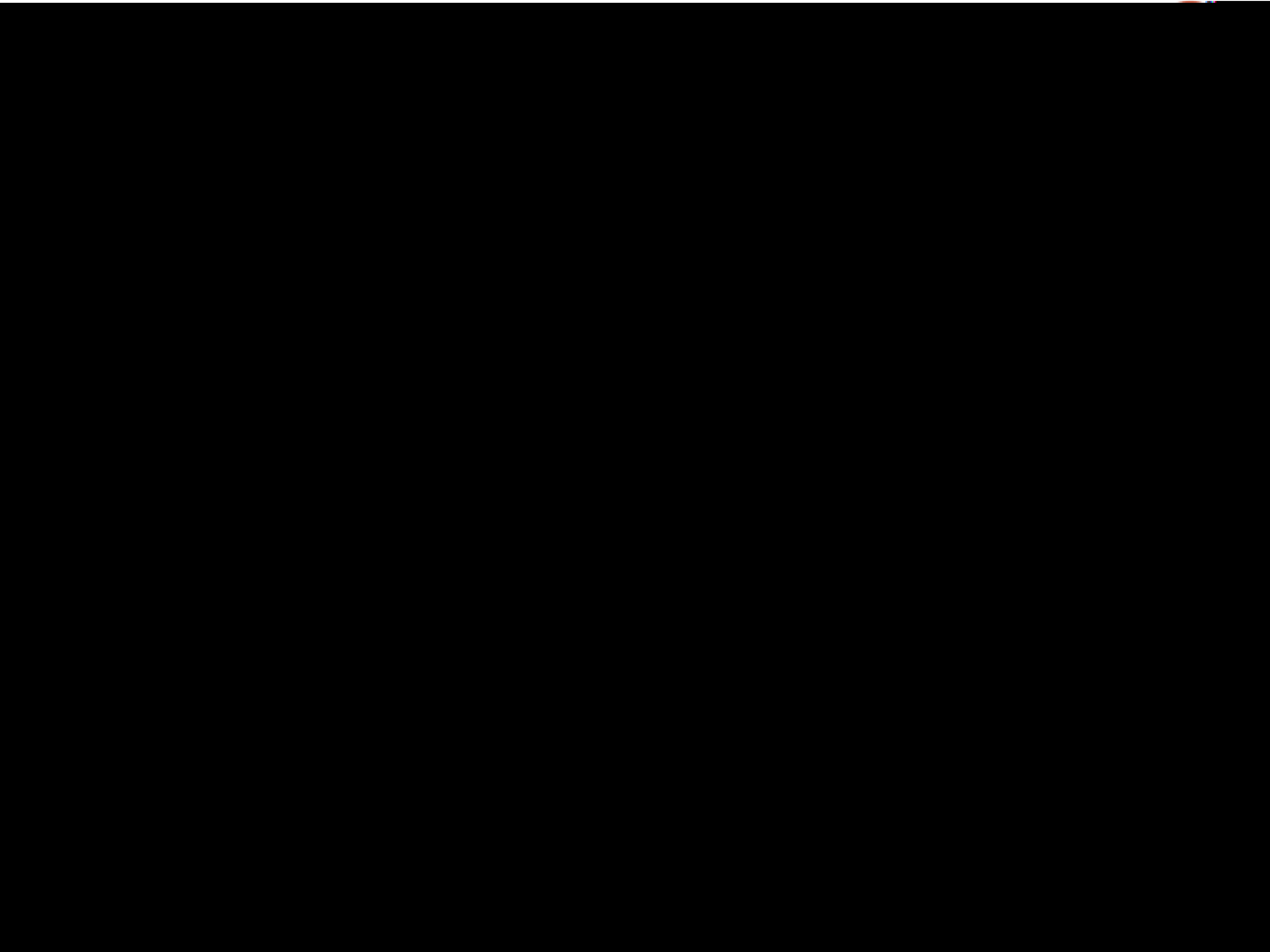


University of Exeter | University of York | University of Liverpool | University of Manchester | University of Bristol | University of Warwick | University of Nottingham | University of Leeds | University of Sheffield | University of Birmingham | University of Southampton | University of Oxford | University of Cambridge | University of Edinburgh | University of Glasgow | University of Aberdeen | University of Dundee | University of Stirling | University of Stirling | University of Stirling

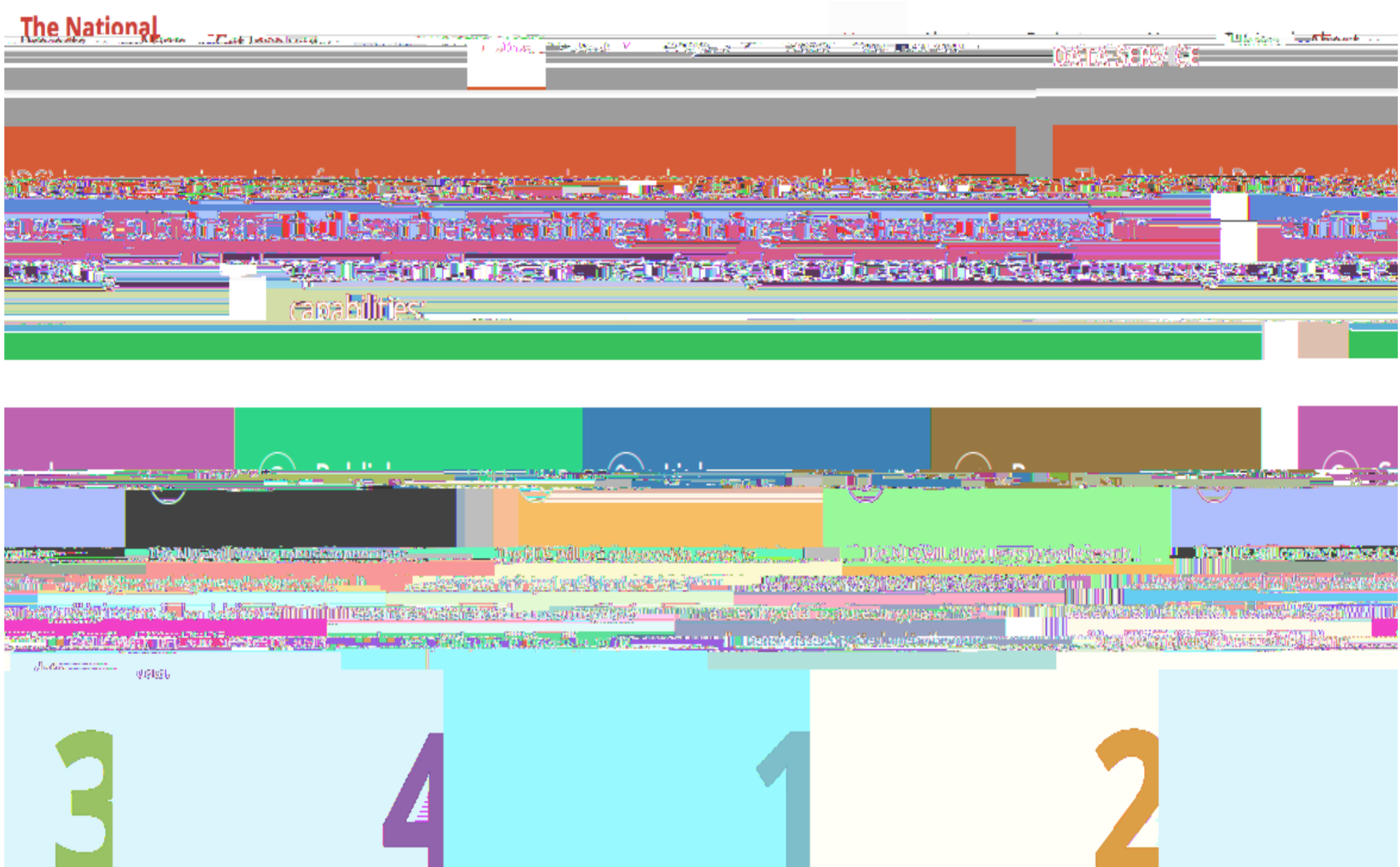
University of Exeter | University of York | University of Liverpool | University of Manchester | University of Bristol | University of Warwick | University of Nottingham | University of Leeds | University of Sheffield | University of Birmingham | University of Southampton | University of Oxford | University of Cambridge | University of Edinburgh | University of Glasgow | University of Aberdeen | University of Dundee | University of Stirling | University of Stirling | University of Stirling







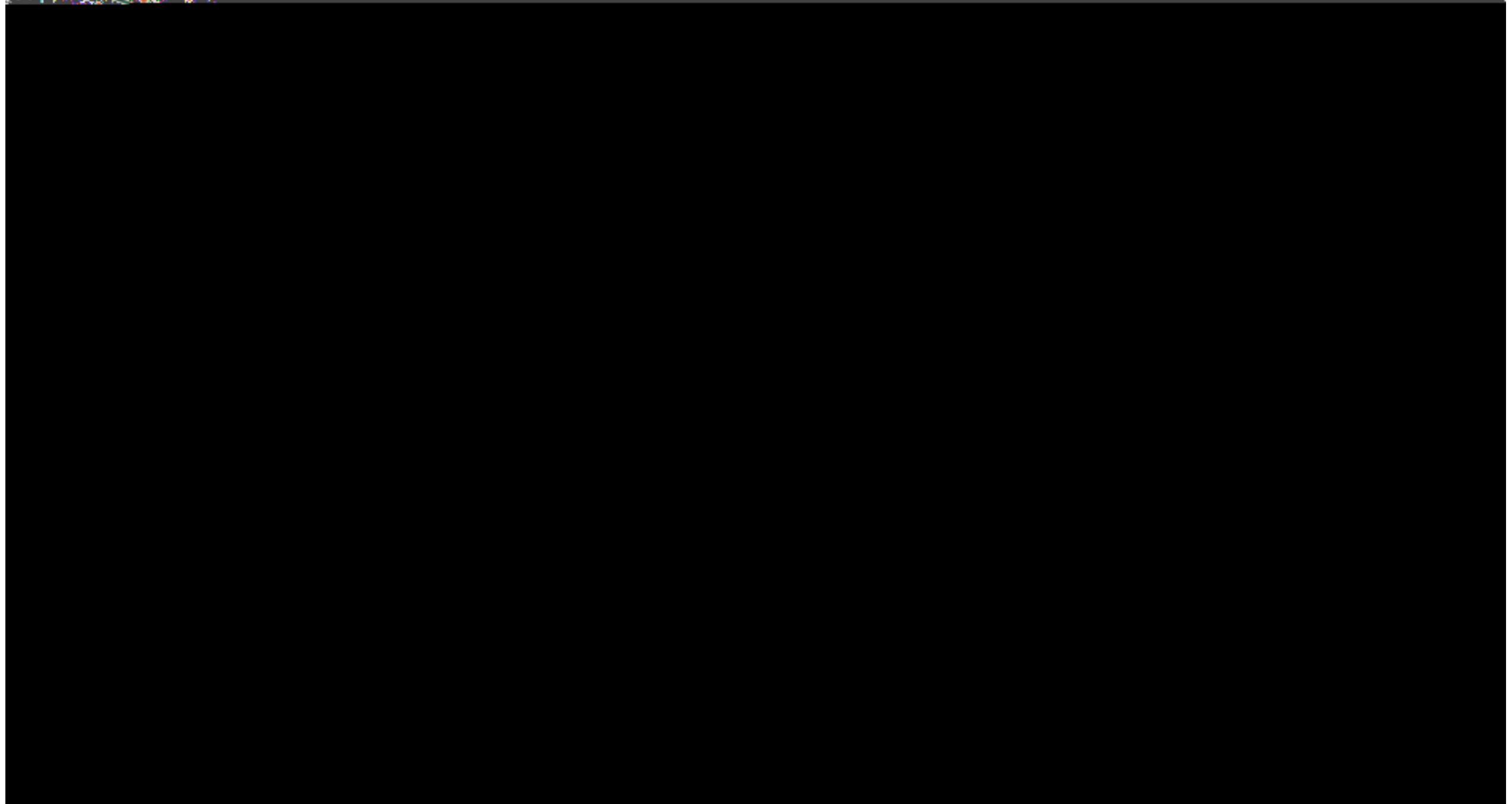
National Data Service (US)



NAME _____

DATE _____

MISSING DATA





EOOSC: Key requirements

- New modes of scholarly communication

- Innovative fit for purpose funding schemes for sustainable underpinning infrastructures

Cross disciplinary collaboration specific measures for review

Support for the transition from scientific insights towards societal innovation

- Support

- The EOOSC needs to be developed as an eco-system of infrastructures
- Key Performance Indicators should be developed for the EOOSC

- FAIR principles

Search



Group Home

Members

Me

Calendar

Google Home UI

Professional data publishers offering their services software and data processing services. Funding agencies (private and public) increasingly concerned with proper Data

processing services. Funding agencies (private and public) increasingly concerned with proper Data

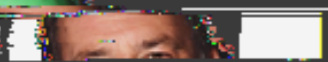
showcase a large number of data sets in a single place. This is a good idea. It is a good idea.

It is a good idea. It is a good idea. It is a good idea. It is a good idea. It is a good idea.

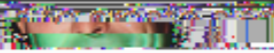
Search

handling a routine research activity. Downloading machine-readable data on the spot is a

It is a good idea. It is a good idea. It is a good idea. It is a good idea. It is a good idea.



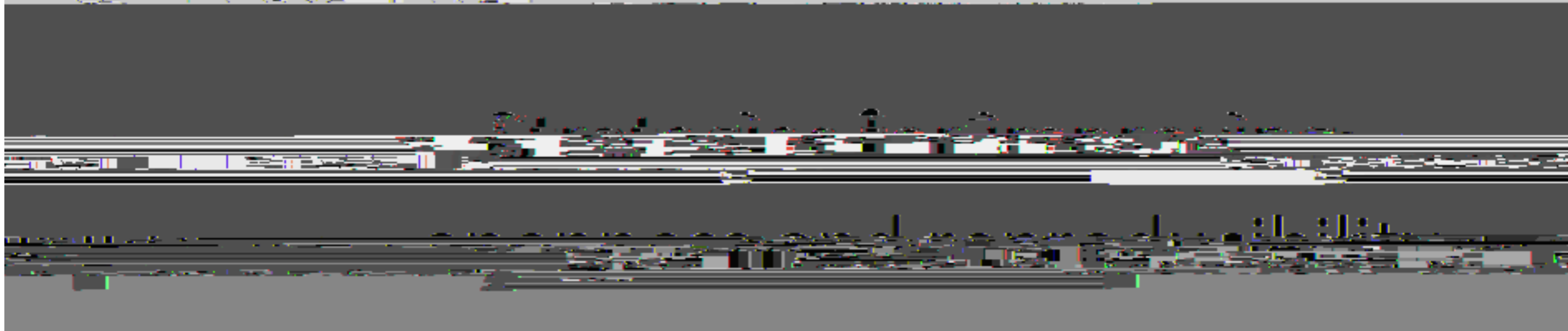
Shared Memories





FOR OPEN SCIENCE CENTER

OPEN SCIENCE

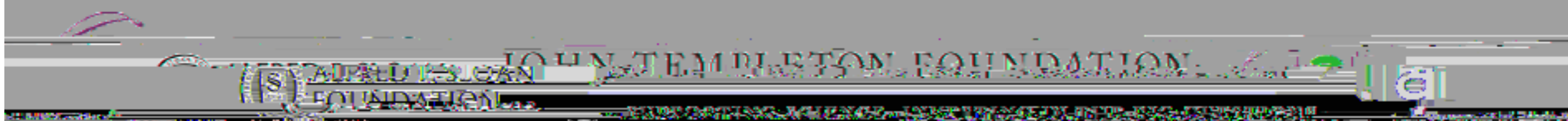


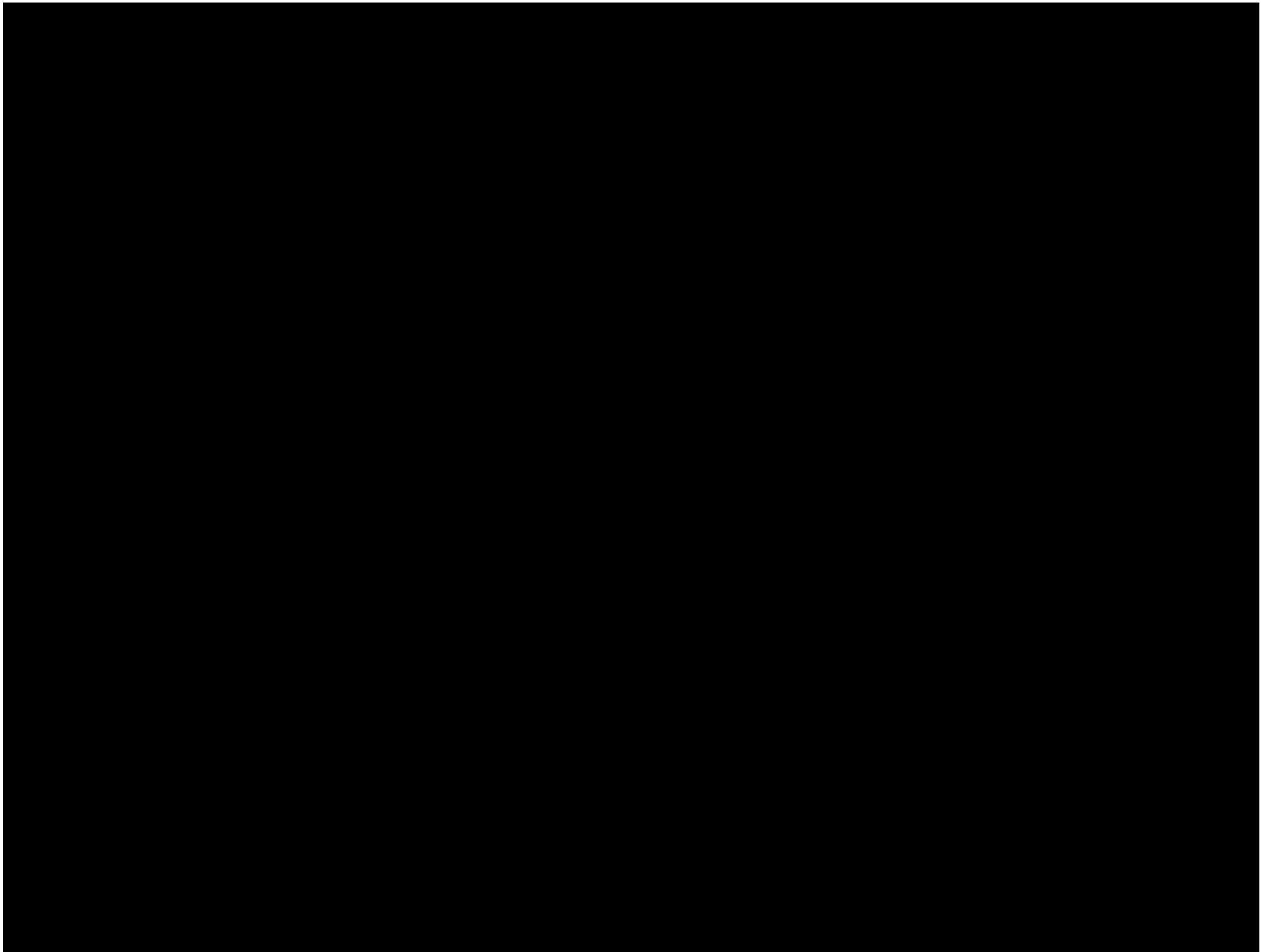
Sallans
Open Science

Andrew
Center for O

with a focus on

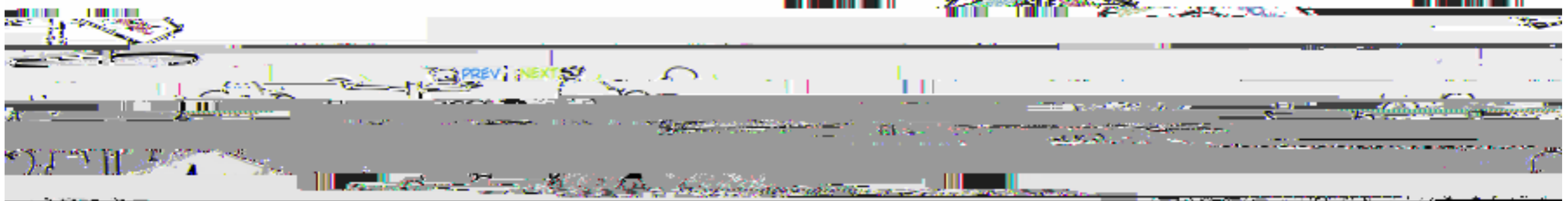
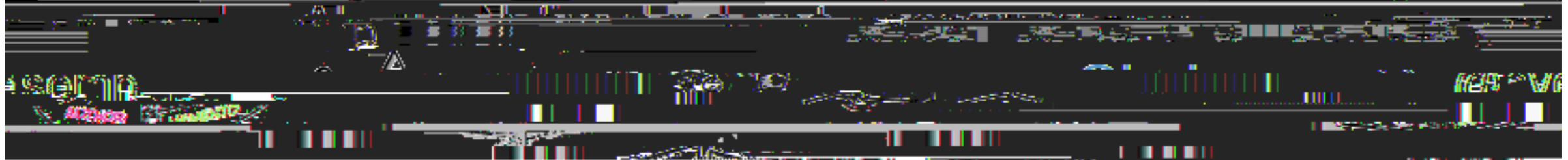
Open Science



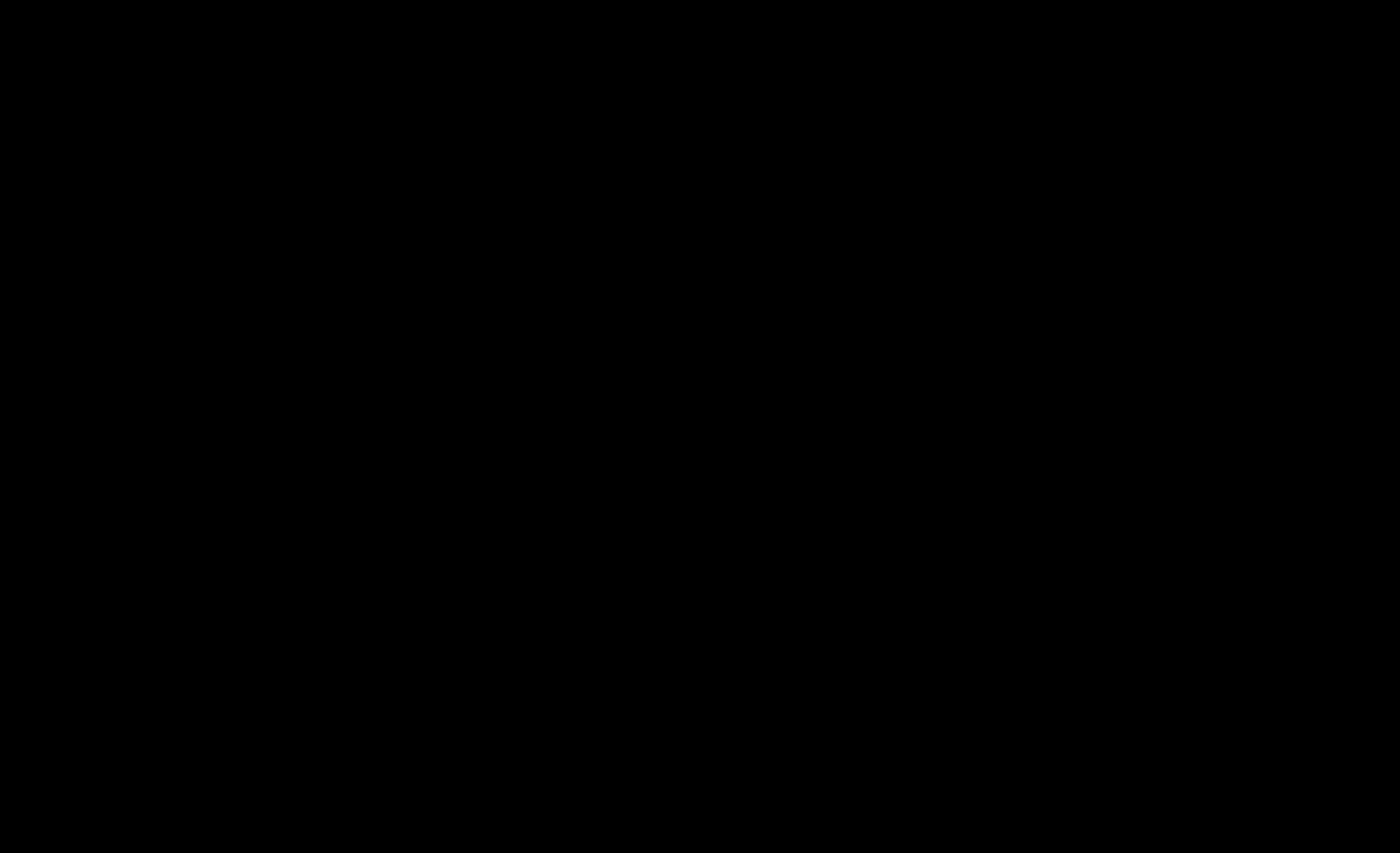
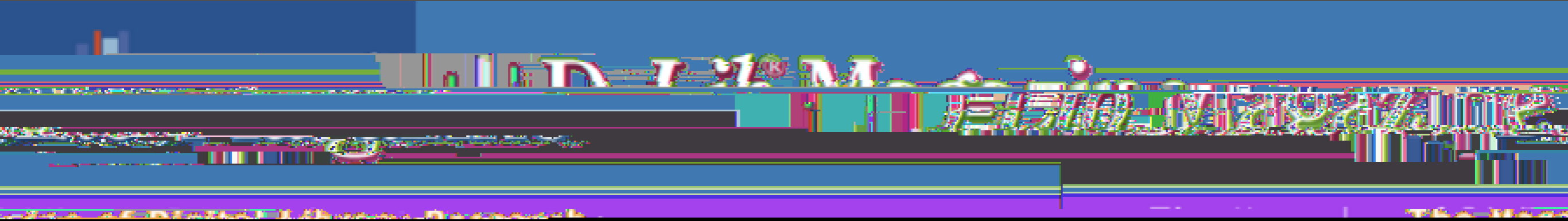


Robert van de Sompeel

Robert van de Sompeel



PREV | NEXT



repository, ...

• **The node at the center of the universe**

line interface for your node, expect others to use it

• Define a mach

system on the web. **Disseminate**
names

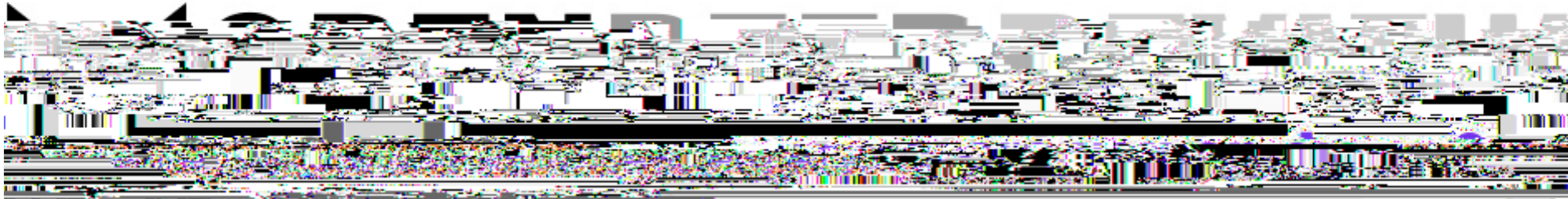
COAR Roadmap: Future Directions for Repository Interoperability

Addressing Strategic Challenges

Adoption of common metadata, identifiers (for authors, institutions, research funding organizations, publications), vocabularies and taxonomies

Interaction with other local infrastructures and external systems (administrative research systems, publishing platforms, e-learning systems, etc)





nes

Open Peer Review Module for Institutional Repositories

12:20-13:15 h.

Presentation of the project and its results

ARVO, DIGITAL.CSIC, IIA, IEO, SECABA-University of Granada

Co-organized with members of the OPRM Consortium (Open Scholar, ARVO, DIGITAL.CSIC, IIA, IEO, SECABA-University of Granada)

Open Peer Review Module for Open Access Journals

Pandelis Perikakis, Open Scholar

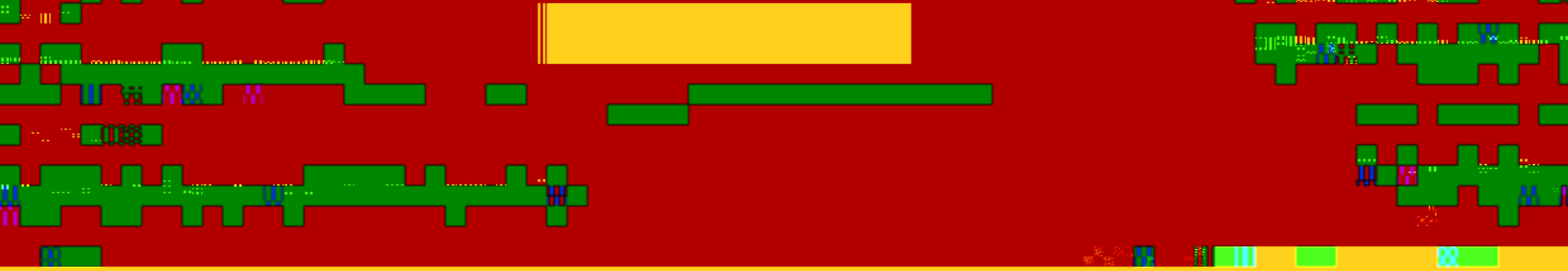
Presentation available at [DIGITAL.CSIC](#)

Open Peer Review Module for Open Access Journals

Emilio Lorenzo, ARVO Consultores

Presentation available at [ARVO](#)

cal details



In conclusion...

Key Aims

Open global ecosystem of infrastructures

New modes of scholarly communications

Modern reward and recognition practices

A global knowledge commons

And key questions for us to ponder

What services are best offered locally, nationally, internationally?

In developing services locally, how do we ensure international interoperability?

What can we do to explore new modes of scholarly communications and modern reward and recognition practices?

Thank you!

Questions? Thoughts?