BMJ Group

What we owe the future
What happened, what happens next?

Ian Mulvany BMJ CTO
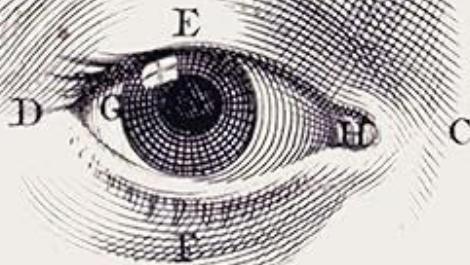
20 May 2025

Icebreaker

Introduction

Inspirations

MICHAEL
STREVENS

The
Knowledge
Machine

How an Unreasonable Idea
Created Modern Science

'Riveting... crystal-clear... unparalleled'
WALL STREET JOURNAL

POLICY FORUM

Large AI models are cultural **and social technologies**

Implications draw on the history of transformative information systems from the past

By Henry Farrell[1], Alison Gopnik[2,3], Cosma Shalizi[3,4], James Evans[3,5]

Debates about artificial intelligence (AI) tend to revolve around whether large models are intelligent, autonomous agents. Some AI researchers and commentators speculate that we are on the cusp of creating agents with artificial general intelligence (AGI), a prospect anticipated with both elation and anxiety. There have also been extensive conversations about cultural and social consequences of large models, orbiting around two foci: immediate effects of these systems as they are currently used, and hypothetical futures when these systems turn into AGI agents - perhaps even superintelligent AGI agents. But this discourse about large models as intelligent agents is fundamentally misconceived. Combining ideas from social and behavioral sciences with computer science can help us understand AI systems more accurately. Large Models should not be viewed primarily as intelligent agents, but as a new kind of cultural and social technology, allowing humans to take advantage of information other humans have accumulated.

The new technology of large models combines important features of earlier technologies. Like pictures, writing, print, video, Internet search, and other such technologies, large models allow people to access information that other people have created. Large Models — currently language, vision, and multi-modal - depend on the fact that the Internet has made
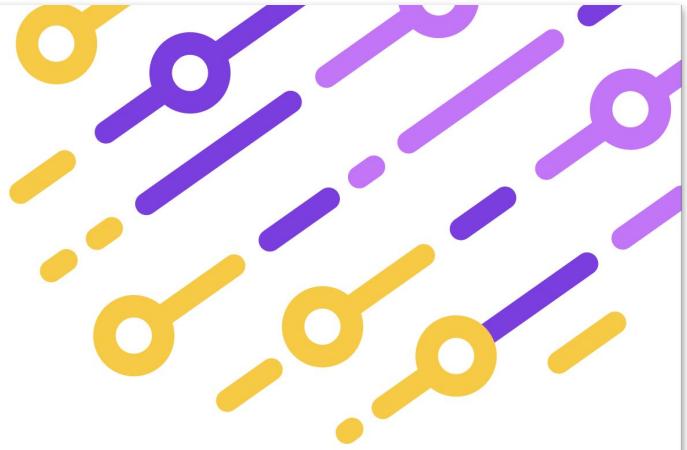
type of cultural and social technology. They are analogous to such past technologies as writing, print, markets, bureaucracies, and representative democracies. Then we can ask the separate question about what the effects of these systems will be. New technologies that aren't themselves cultural or social, such as steam and electricity, can have cultural effects. Genuinely new cultural technologies, Wikipedia for example, may have limited effects. However, many past cultural and social technologies also had profound, transformative effects on societies, for good and ill, and this is likely to be true for Large Models.

These effects are markedly different from the consequences of other important general technologies such as steam or electricity. They are also different from what we might expect from hypothetical AGI. Reflecting on past cultural and social technologies and their impact will help us understand the perils and promise of AI models better than worrying about superintelligent agents.

SOCIAL & CULTURAL INSTITUTIONS

For as long as there have been humans, we have depended on culture. Beginning with language itself, human beings have had distinctive capacities to learn from the experiences of other humans and these capacities are arguably the secret of human evolutionary success.

be thought of as a kind of technology (1). In the modern era, markets, democracies, and bureaucracies have been particularly important. The economist Friedrich Hayek argued that the market's price mechanism generates dynamic summaries of enormously complex and otherwise unfathomable economic relations (2). Producers and buyers do not need to understand the complexities of production: all they need to know is the price, which compresses vast swathes of detail into a simplified but usable representation. Election mechanisms in democratic regimes focus distributed opinion toward collective legal and leadership decisions in a related way. The anthropologist James Scott argued (3) that all states, democratic or otherwise, have managed complex societies by creating bureaucratic systems that categorize and systematize information. Markets, democracies, and bureaucracies have relied on mechanisms that generate lossy (i.e., incomplete, selective, and uninvertible) but useful representations well before the computer. Those representations both depend on and go beyond the knowledge and decisions of individual people. A price, an election result, or a measure like gross domestic product (GDP) summarizes large amounts of individual knowledge, values, preferences and actions. At the same time, these social technologies can also themselves shape individual knowledge

STM

Trusted Identity in Academic Publishing

Part 2: The Researcher Identity
Verification Framework
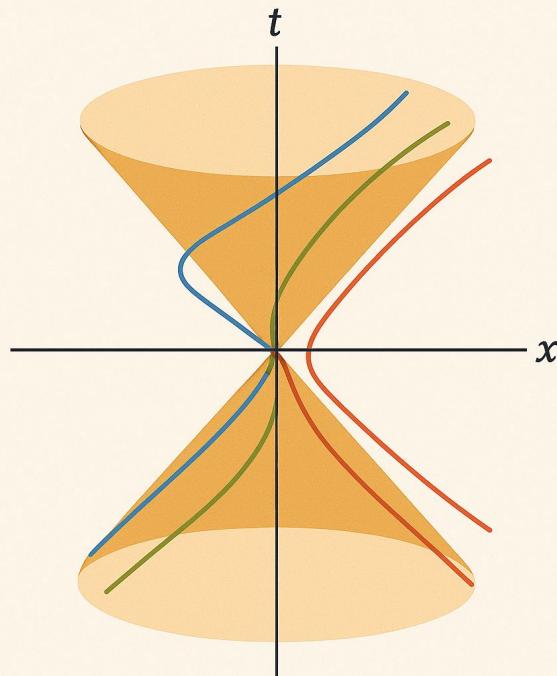
Richard Northover, Hylke Koers, Aaron Wood, Adam Sewell,   Andy Heard, Helen King,
Jacob Kendall-Taylor,   Jennifer Wright,   Joris van Rossum,   Kevin Lawson,   Liv Davies,
Lucy Loftus,   Matthew Salter,   Phil Reimann,   Ralph Youngen,   Sam Parker,   Tim Lloyd
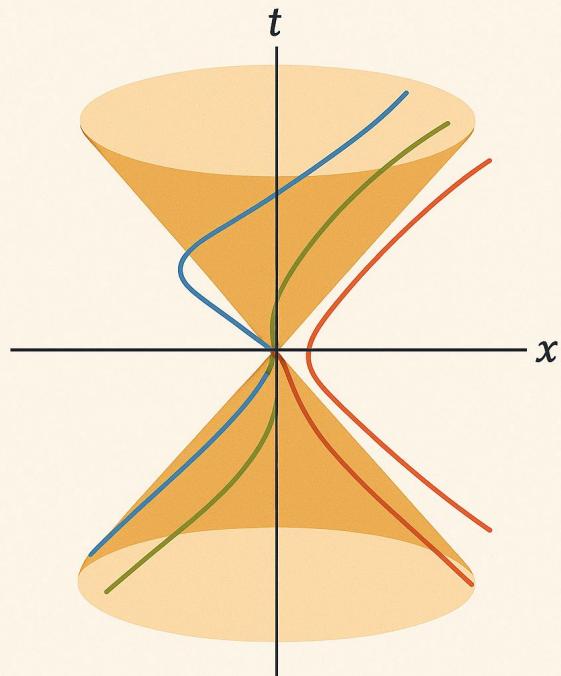
# What should you do?

Future light cone

Future light cone

SCHOLARSHIP
NEEDS
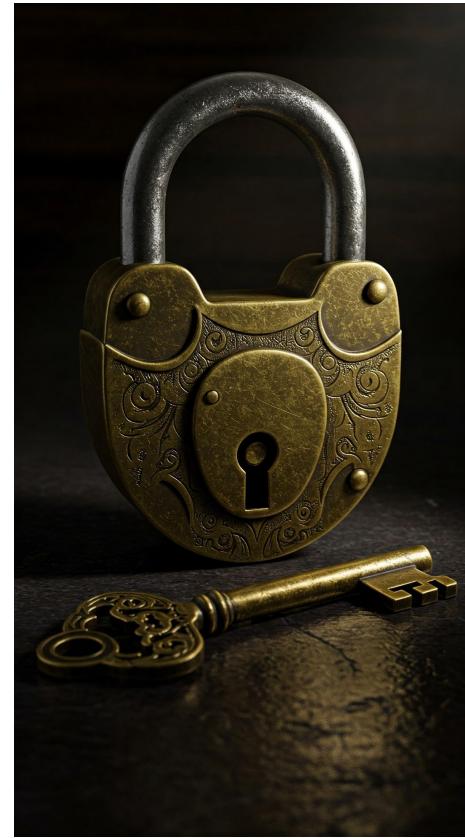
YOU

data

identity

trust

DRYAD

Making research data **discoverable, reusable, and citable** through:

- Curated data repository
- Persistent identifiers (DOIs)
- Open access publishing
- Standards-compliant metadata

A not-for-profit community committed to the open availability and routine reuse of all research data

ORCID Trust Markers: Enhancing Research Integrity

What Are Trust Markers?

- Validated information in ORCID records
- Added by trusted member organizations
- Visible indicators of research credibility
- Part of ORCID's distributed trust model

Key Trust Markers

- Validated affiliations (institutions)
- Validated works (publishers)
- Validated funding (funders)
- Validated peer review (journals)

ORCID Record Summaries: Making Trust Visible

Author Profile

Jane Doe

0000-0001-2345-6789

University of Science

Works

27 validated

3 self-asserted

Funding

8 validated

1 self-asserted

Peer Review
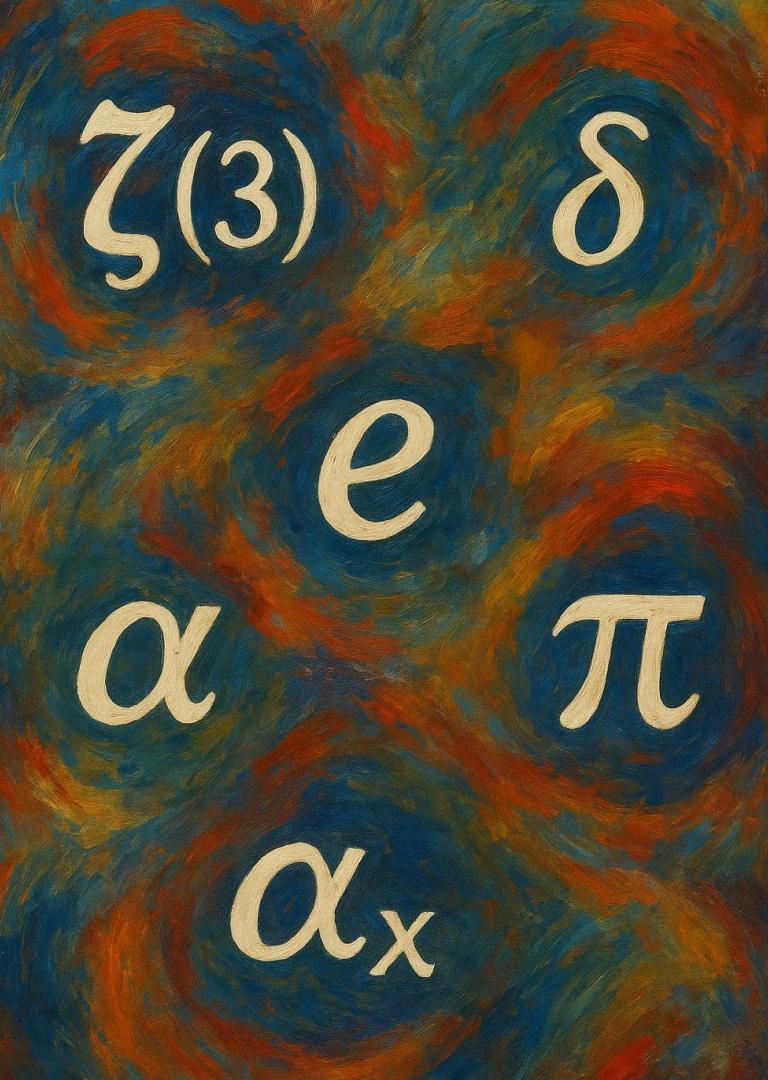
15 validated

Key Benefits

- Quick assessment of credentials
- Clear visual trust indicators
- Streamlined integrity checks
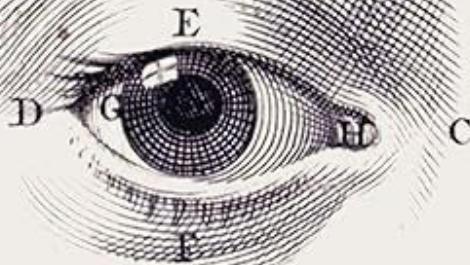
"Building trust in the scholarly ecosystem through verified connections"

What stays the same?

How does all this work?

MICHAEL STREVENS
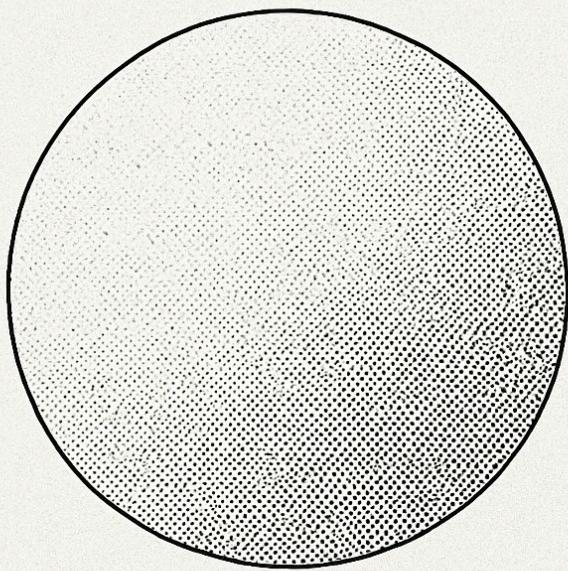
The *Knowledge* *Machine*

How an Unreasonable Idea
Created Modern Science
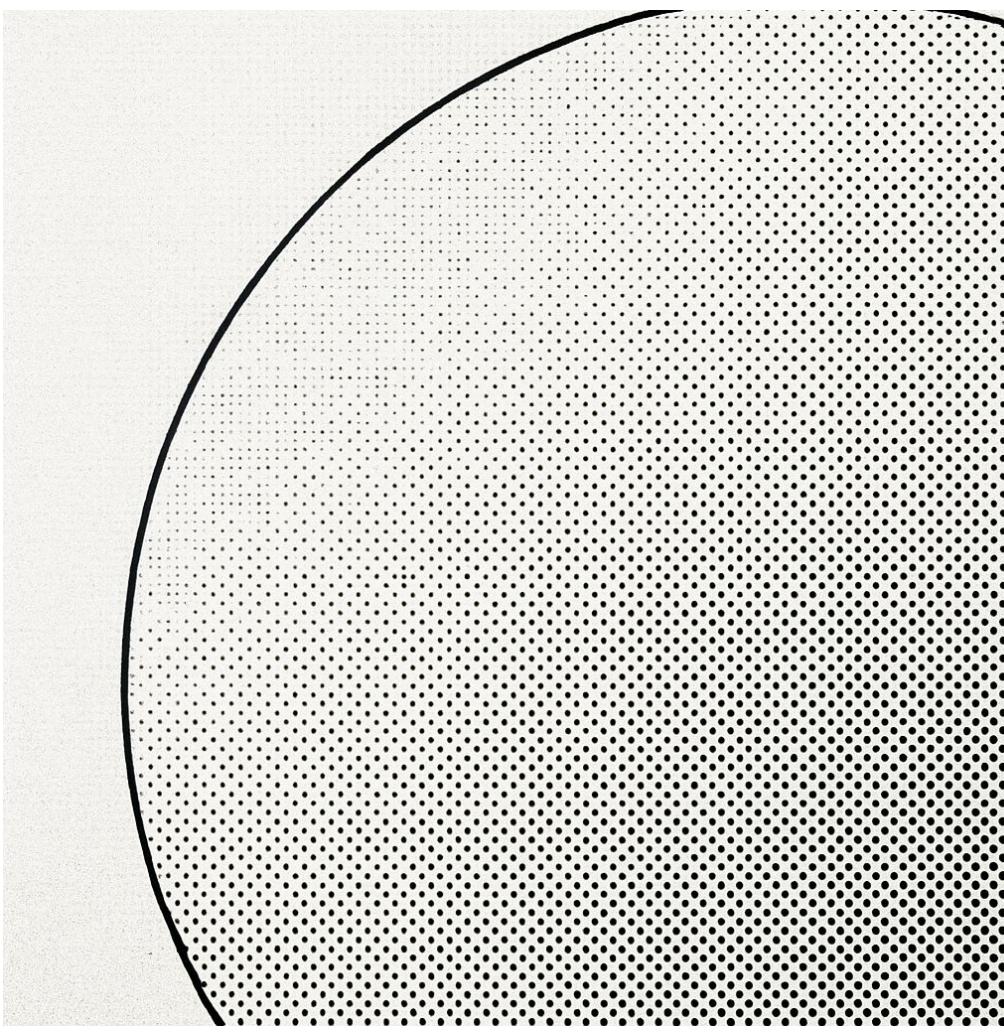
'Riveting... crystal-clear... unparalleled'
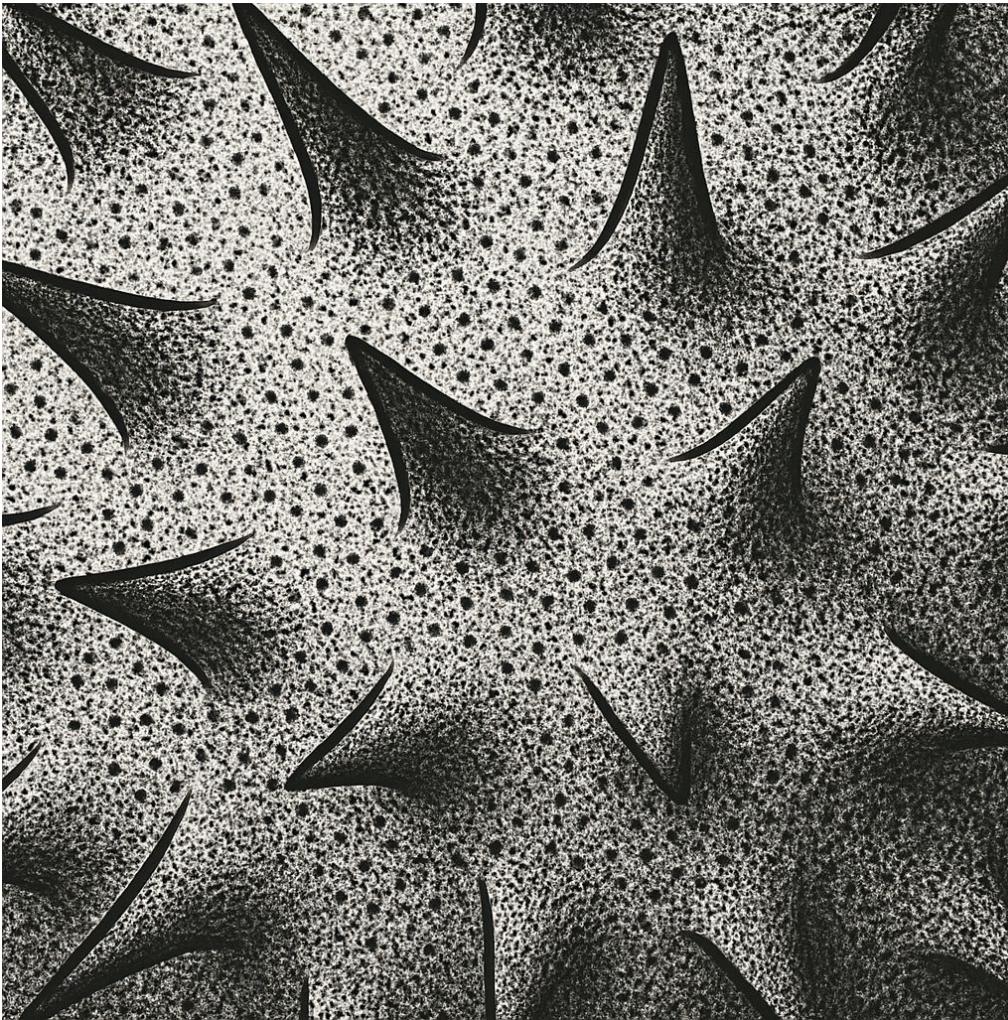
- Iron Rule

- Baconian Convergence
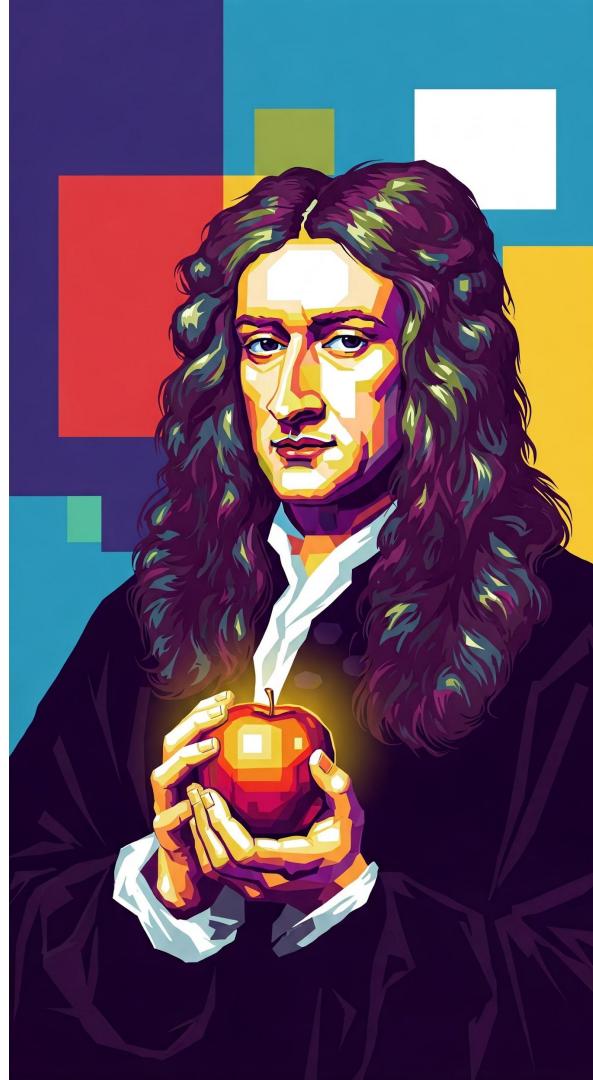
- Competition

- Accreditation

- Difference between private and public narrative

- The imperative to not mess with the system

I have not as yet been able to discover the reason for these properties of gravity from phenomena, and I do not feign hypotheses. For whatever is not deduced from the phenomena must be called a hypothesis; and hypotheses, whether metaphysical or physical, or based on occult qualities, or mechanical, have no place in experimental philosophy. In this philosophy particular propositions are inferred from the phenomena, and afterwards rendered general by induction.

- Isaac Newton - 1713

3200 BCE
First writing system in Babylon

2025 CE
Digital information and AI era

3000 BCE
3200 BCE

2000 BCE

1000 BCE

0 CE

1000 CE

2000 CE
1660 CE
2025 CE

1660 CE
Scientific society founded in London

https://claude.ai/public/artifacts/b4cdebd5-8ab6-4418-81de-3b54db2ef37c

1660
First scientific society for sharing research

2012
Unique digital identifiers for researchers
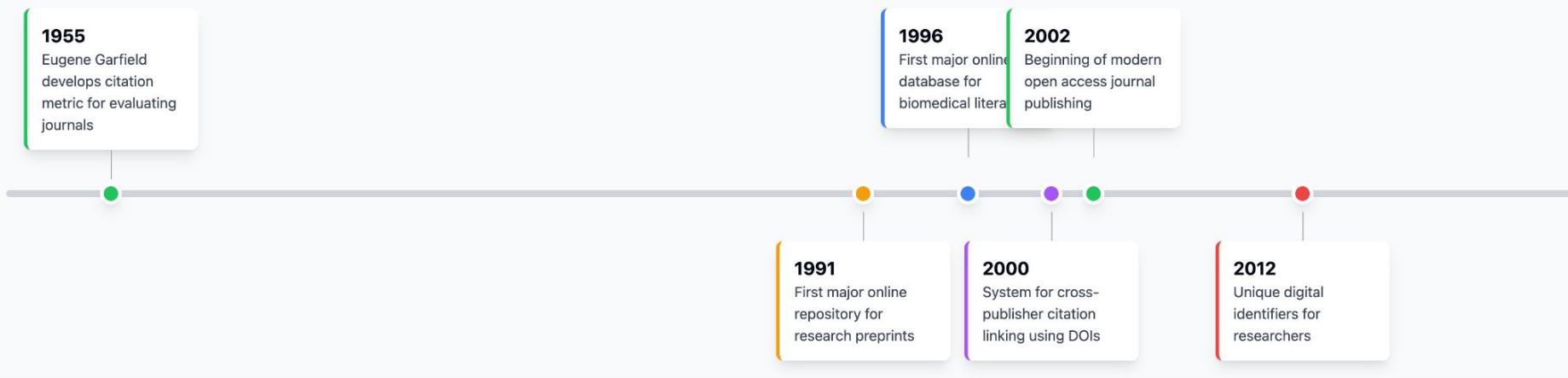
1955
Eugene Garfiel[d] citation metric[s] journals

2002
Beginning of modern open access journal publishing

1660   1700   1750   1800   1850   1900   1950   2000   2025
1660   1955   1999 2002 2012

https://claude.ai/public/artifacts/fafb
b714-d641-4e89-810a-a48e166a62
16

Academic Publishing Timeline

1955
Eugene Garfield develops citation metric for evaluating journals

1996
First major online database for biomedical litera

2002
Beginning of modern open access journal publishing

1991
First major online repository for research preprints

2000
System for cross-publisher citation linking using DOIs

2012
Unique digital identifiers for researchers

https://claude.ai/public/artifacts/e31df9ab-af4e-4828-ba17-17979289f056

Childhood Acute Lymphoblastic Leukemia (ALL)

5-Year Survival Rates, 1975-2020

https://www.cancer.gov/types/childhood-cancers/child-adolescent-cancers-fact-sheet

Weekly measles cases by rash onset date

2023–2025* (as of May 8, 2025)

≡4 NEW YORK Watch 24/7

TRENDING   Streaming News 24/7   Newark ...

NEWARK AIRPORT

Newark Airport passengers warned of possible measles exposure by NJ health officials

By **NBC New York Staff** and Associated Press •
Published May 15, 2025 • Updated on May 15, 2025 at 3:18 pm

Log in or create a free profile to save articles

The New Jersey Department of Health is warning travelers about a potential measles exposures after someone with a case of the illness traveled through Newark Liberty International Airport.

The infected person was a non-New Jersey resident

nbcnewyork.com

What stays the same?

connect claims to evidence

humans drive the process
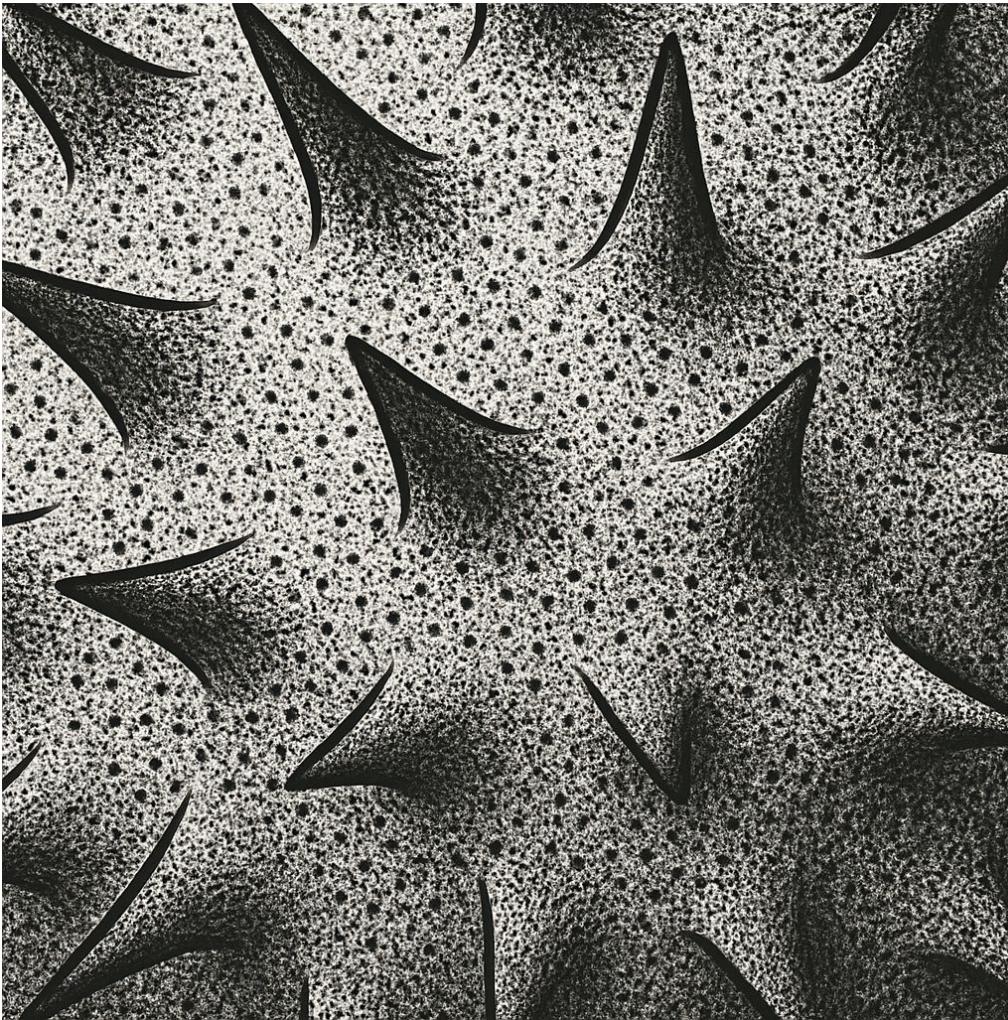
they need reward

the system is better than its parts

the system needs protection

Socio Technical System

A cultural and social technologies is a system or tool that enables humans to access, process, reorganise, transform, and restructure information accumulated by other humans, facilitating large-scale coordination within society

AN INTRODUCTION TO
CYBERNETICS

by

W. ROSS ASHBY
M.A., M.D.(Cantab.), D.P.M.
Director of Research
Barnwood House, Gloucester

NEW YORK
JOHN WILEY & SONS INC.
440 FOURTH AVENUE
1956

What should we do?

The **fundamental accretive and compound process of our knowledge** is the most important thing.

Making that work is the game, so to speak.

Many industry efforts, while not unimportant, are attempts to break past this fundamental truth.

We have had:

- Open access
- Altmetrics
- Impact

We are now in:
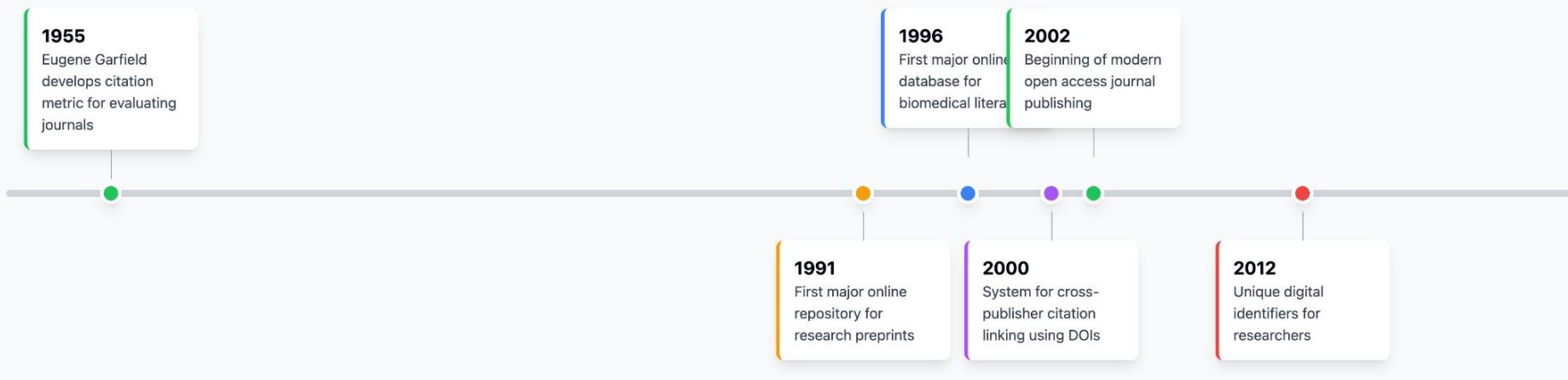
- Research integrity era

- Create as high a quality scholarly record as we possibly can

- Treat the participants of this enterprise with as much respect as we can, because going toe to toe against reality can be brutal

We have a problem!

Academic Publishing Timeline

1955
Eugene Garfield develops citation metric for evaluating journals

1996
First major online database for biomedical litera...

2002
Beginning of modern open access journal publishing

1991
First major online repository for research preprints

2000
System for cross-publisher citation linking using DOIs

2012
Unique digital identifiers for researchers

https://claude.ai/public/artifacts/e31df9ab-af4e-4828-ba17-17979289f056

XML

APC

Impact Factor

Decision points about the future

Identity

Who's Bob?

anonymous

known

private

public

https://www.cabinet.ox.ac.uk/oldenburgs-correspondence-0

anonymous

known

private

public

STM
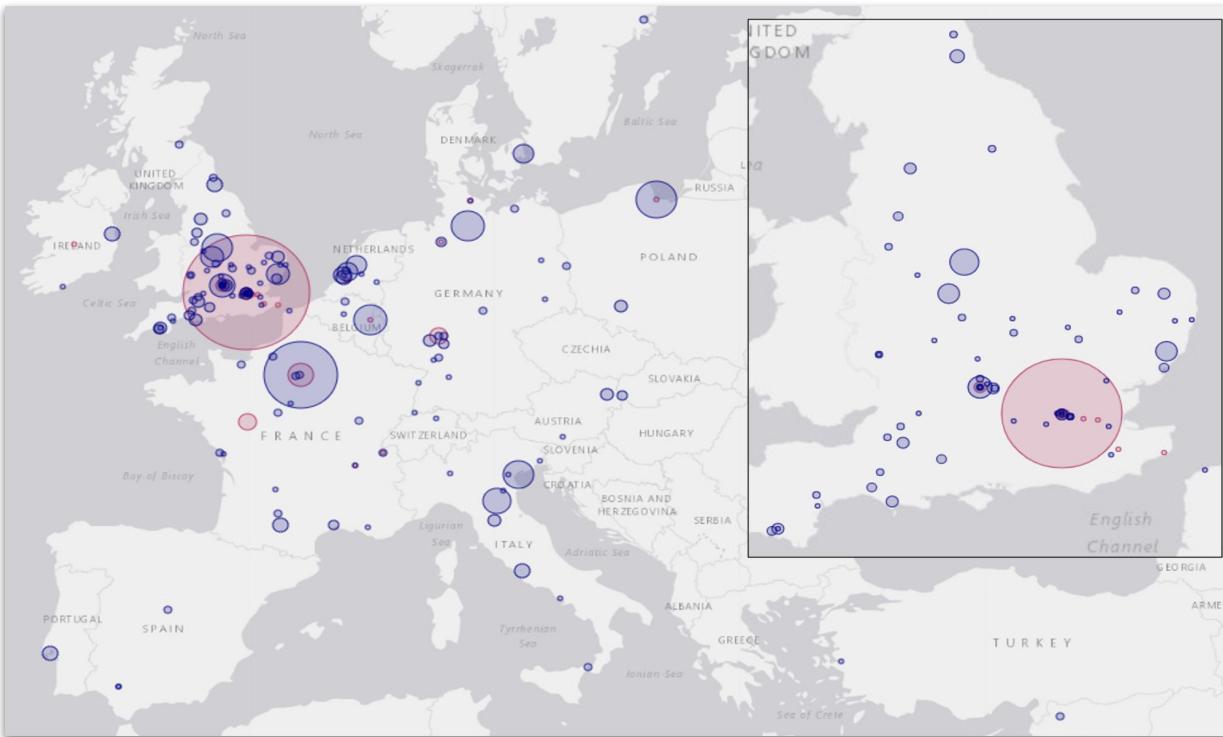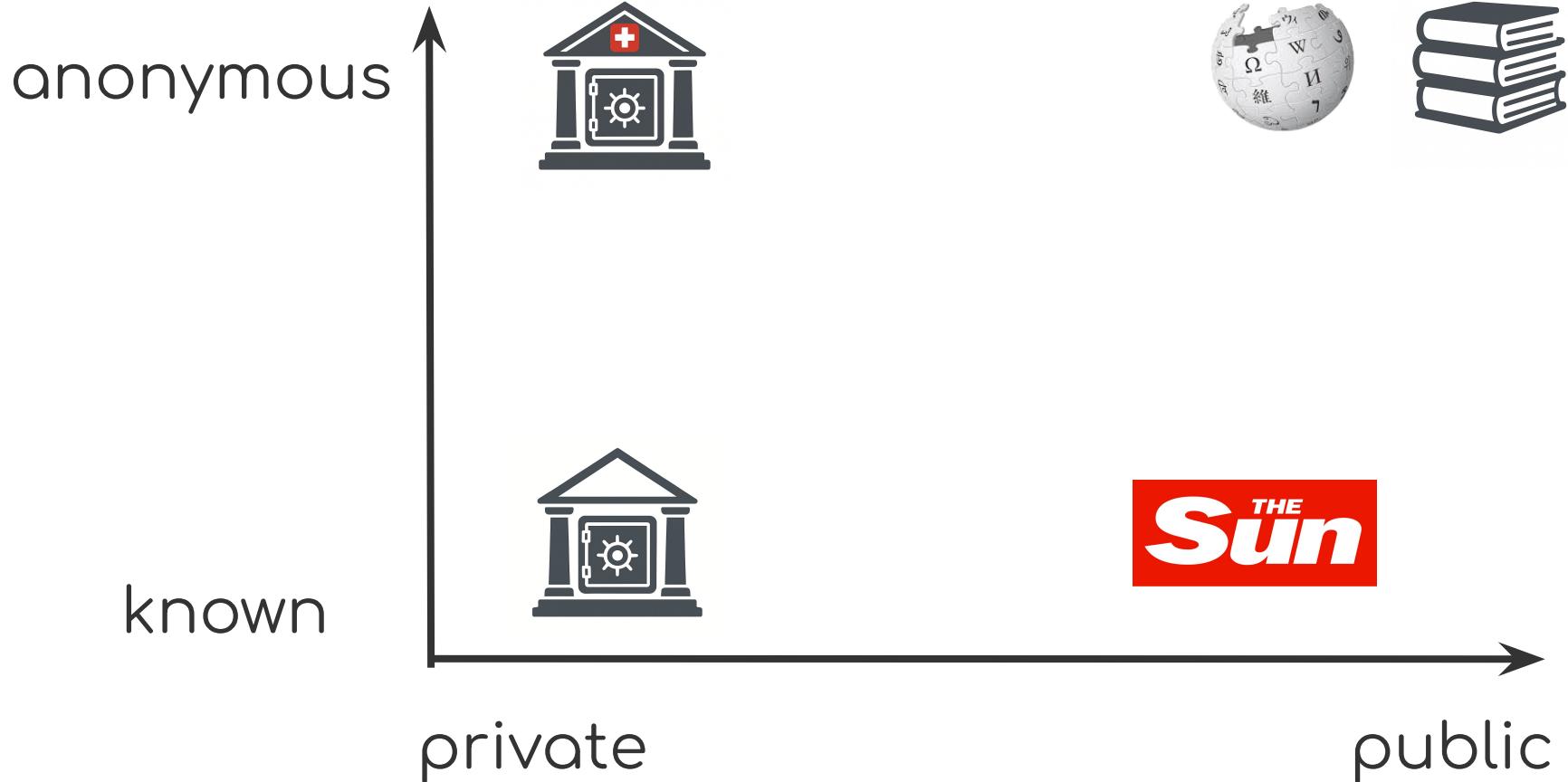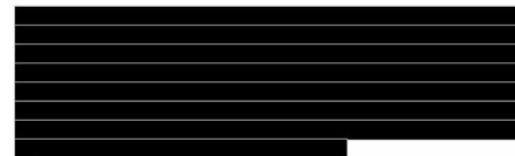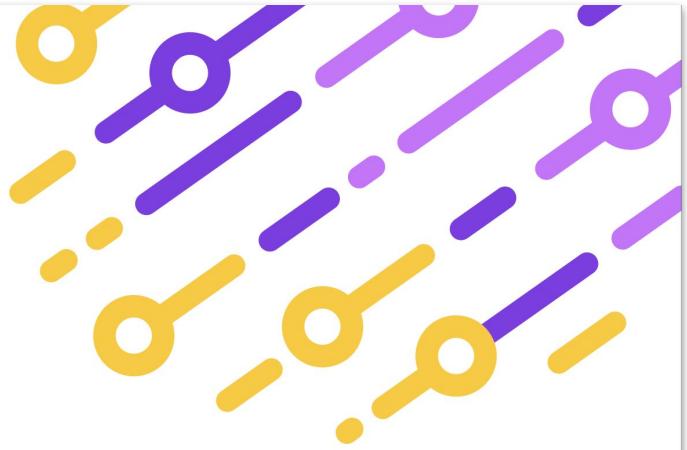
Trusted Identity in Academic Publishing

MARCH 2025

Part 2: The Researcher Identity
Verification Framework

Richard Northover, Hylke Koers, Aaron Wood, Adam Sewell,   Andy Heard, Helen King,
Jacob Kendall-Taylor,   Jennifer Wright,   Joris van Rossum,   Kevin Lawson,   Liv Davies,
Lucy Loftus,   Matthew Salter,   Phil Reimann,   Ralph Youngen,   Sam Parker,   Tim Lloyd

Draft for community review   www.stm-assoc.org

Researcher Identity Verification Framework

1 **Assessment**
Determine risk level and minimum trust required

2 **Verification**
Offer inclusive methods to meet trust requirements

3 **Evaluation**
Judge verification results against required trust level

4 **Action**
Allow or redirect based on verification outcome

5 **Feedback**
Measure outcomes to maintain accountability

Key Verification Considerations

🌐 Availability
Not all methods are available globally; provide multiple options to avoid excluding legitimate researchers

⏱️ Deferred Verification
Allow users to complete their primary task before verification when appropriate

🔄 Reverification
Balance security and convenience when determining how long to remember verified status

🔒 Security
Each verification method must be tied to a unique identifier to prevent multiple accounts
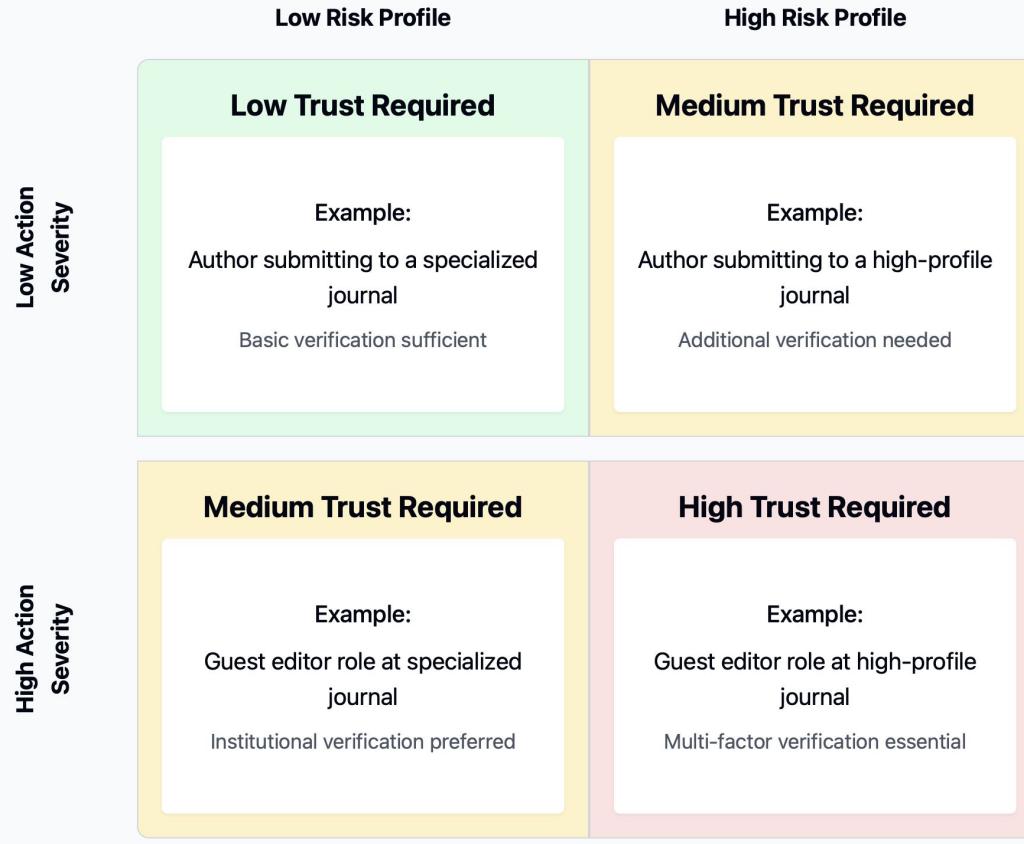
🔗 Multi-dimensional
Combine independent verification methods for stronger trust assurance

Risk Assessment Framework

Balancing Action Severity & Journal Risk Profile

	Low Risk Profile	High Risk Profile
Low Action Severity	**Low Trust Required** Example: Author submitting to a specialized journal Basic verification sufficient	**Medium Trust Required** Example: Author submitting to a high-profile journal Additional verification needed
High Action Severity	**Medium Trust Required** Example: Guest editor role at specialized journal Institutional verification preferred	**High Trust Required** Example: Guest editor role at high-profile journal Multi-factor verification essential

Evidence of individual identity	Evidence of academic participation		
	None	**Some**	**Good**
Good	**TRUST VIA ACCOUNTABILITY** Good confidence in identity alone means accountability despite lack of credibility	**GOOD TRUST** Good confidence in a user's identity and some that they are a genuine researcher	**MAXIMUM TRUST** Good confidence in a user's identity and that they are a genuine researcher
Some	**LOW TRUST** Some confidence in identity, but no supporting academic evidence	**MEDIUM TRUST** Some confidence in a user's identity and some that they are a genuine researcher	**GOOD TRUST** Trust provided by the good research evidence is limited by evidence of identity
None	**NO TRUST** No confidence in a user's identity or evidence that they are a genuine researcher	**NO TRUST** Some evidence of credibility but it can't be reliably linked to an identified individual	**NO TRUST** Good evidence of credibility but it can't be reliably linked to an identified individual

Verification Method Matrix

Identity Evidence	Academic Evidence		
	None	**Some**	**High**
High	**Document / Payment Card Verification** Trust via accountability despite lack of credibility	**Institutional affiliation + MFA** Good confidence in identity plus some academic evidence	**ORCID + trust markers + MFA** Maximum trust through verified identity and academic history
Some	**Non-institutional IDP + MFA** Limited trust due to uncertain identity	**Institutional affiliation** Medium trust through partial verification	**ORCID + trust markers** Good trust limited by identity verification
None	**Opaque email address, Non-institutional IDP** No verification of identity or academic credentials	**ORCID without trust markers** Unverifiable academic claims	**Manual Verification, Vouching** Requires human intervention to establish trust

Verification Methods & Trust Levels

A comprehensive approach to researcher identity verification

Trust Levels & Verification Methods

3 - High Trust

- ORCID + trust markers + MFA
- Manual Verification
- Vouching

2 - Medium Trust

- Institutional affiliation via IDP + MFA
- ORCID + trust markers

1 - Basic Trust

- Institutional affiliation via IDP/email
- Document verification
- Payment Card verification

0 - Insufficient Trust

- ORCID without trust markers
- Opaque email address
- Non-institutional IDPs

Challenges in Researcher Verification

1 **Verification could frustrate legitimate users**
Additional steps may create friction in user journeys

2 **Changes relationship between researchers and publishers**
Must preserve manuscript evaluation based on merit, not identity

3 **Affiliation doesn't guarantee research legitimacy**
Not all affiliates are legitimate researchers

4 **Won't stop all bad actors**
Some may bypass verification systems

5 **Email preference conflicts**
Users prefer personal emails over institutional ones

6 **Trust has an expiration date**
Institutional relationships change over time

7 **Uncertain trust thresholds**
Difficulty determining sufficient verification levels

8 **Universal self-service verification challenges**
Access barriers to verification methods

A more expansive vision
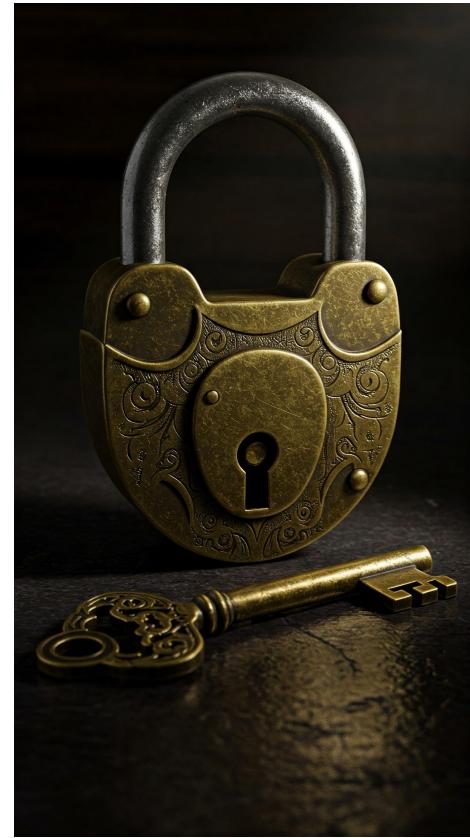
data

identity

trust

Final thought

Future light cone