The Vanishing Record

COVID Origins, Hidden Emails, ResearchGate, and the Battle Over the Scientific Archive

Douglas C. Youvan

2026

Copyright and Reader's Note

Copyright 2026 Douglas C. Youvan. All rights reserved.

This book is a documentary investigation of scientific memory: how emails, federal records, research papers, platform files, metadata, and archived web pages can become evidence in a public controversy. It distinguishes sharply among established facts, official allegations, the author's personal account, reasonable inference, and open questions.

Nothing in this book should be read as an allegation of criminal wrongdoing by any person who has not been charged or convicted. In particular, Anthony S. Fauci has not been charged in the David Morens criminal case, and this book does not claim that he caused, requested, or participated in any action by ResearchGate concerning the author's publications. Likewise, the fact that investors or institutions have relationships with one another does not establish coordination, motive, or causation.

The manuscript was completed after the August 18, 2026 guilty plea of David M. Morens. Events after August 22, 2026 are outside its documentary horizon.

How to Read the Evidence Labels

The central discipline of this book is classification. A controversy becomes impossible to understand when memory, allegation, inference, and proof are blended into a single narrative. The following labels are therefore used throughout.

DOCUMENTED means that the proposition is supported by a primary public record, an official publication, an archived page, a contemporaneous document, or another source whose existence can be independently checked.

ADMITTED OR ADJUDICATED means that a fact has moved beyond allegation because it was admitted in a guilty plea, found by a court, or otherwise formally established in a legal proceeding.

OFFICIAL ALLEGATION means that the proposition appears in an indictment, subpoena, congressional report, investigative letter, or other formal government document but has not necessarily been adjudicated.

AUTHOR'S ACCOUNT means that Douglas C. Youvan reports the event from personal experience or from his own records. Such an account is evidence of what the author observed, but it is not automatically proof of another person's motive.

INFERENCE means that the proposition follows plausibly from known facts but requires additional evidence before it should be treated as established.

OPEN QUESTION identifies a factual issue that could potentially be resolved by records that are not presently public: platform audit logs, moderation tickets, private communications, server histories, preservation records, or compelled production.

These labels are not decorative. They are the argument of the book.

Contents

Copyright and Reader's Note
How to Read the Evidence Labels
Chapter 1: The Day the PDFs Disappeared
Chapter 2: Two Million Reads
Chapter 3: The Archive That Survived
Chapter 4: An Email from 2002
Chapter 5: Before COVID Had a Name
Chapter 6: From Chow to Rottier to Baric
Chapter 7: NIAID and the Architecture of Funding
Chapter 8: The Gates-Fauci Global-Health Era
Chapter 9: ResearchGate
Chapter 10: The COVID-Origin War
Chapter 11: FOIA
Chapter 12: David Morens
Chapter 13: The Guilty Plea
Chapter 14: The Missing Record Problem
Chapter 15: The Private Archive Problem
Chapter 16: 5,700 Missing PDFs
Chapter 17: What Compulsory Process Could Answer
Chapter 18: Evidence, Association, and Conspiracy
Chapter 19: The Scientific Record as Evidence
Chapter 20: Who Owns Scientific Memory?
Epilogue: The Record Must Outlive the Argument
Appendix A: Working Timeline
Appendix B: The 2002 NIH Correspondence Provenance
Appendix C: ResearchGate Record-Preservation Questions
Appendix D: Selected Publications from the Surviving Corpus
Appendix E: Primary Sources and References
About the Author

Chapter 1: The Day the PDFs Disappeared

On the weekend of July 18-19, 2026, a large scientific corpus ceased to be publicly available in the form in which readers had known it on ResearchGate. The titles did not necessarily vanish. Publication pages and traces could remain. What disappeared from public access, according to my records and observation, were the full-text PDF files themselves on a scale that was impossible to mistake for the accidental loss of one paper or one upload.

> **AUTHOR'S ACCOUNT:** Approximately 5,700 PDFs were affected. I received no warning that such an account-wide action was coming. I had not received the sequence of strikes that ResearchGate's public policies describe. I received no contemporaneous explanation identifying a violated rule, a copyright complainant, a governmental request, an automated enforcement category, or any other reason for what happened. I later sent numerous emails to ResearchGate and to people connected with the company. I received no substantive answer explaining the mass removal.

The number matters, but not because a large number proves a sinister motive. It matters because scale changes the nature of an event. A single missing file can be a broken upload. Twenty missing files can be a rights problem. Thousands of full texts becoming unavailable in a narrow interval invite an account-level explanation. In a database, an account-level event normally leaves an account-level trace: a moderation action, a batch process, a policy code, a support ticket, a database operation, a visibility change, or a human authorization.

The absence of that trace from the user's view is not the same as the absence of the trace itself.

ResearchGate's own help materials show that the platform is technically capable of changing full-text visibility at scale. A user can make all public files private at once. The company can also remove, restrict, disable access to, or otherwise act on content under its policies. Its Appeals Policy says that when it takes qualifying action against content or an account, the user will be told what action was taken and why, subject to stated exceptions. Its Strikes Policy describes warnings and strikes, and its copyright materials describe procedures through which publishers or rights holders can report allegedly infringing material.

Those policies create the first hard question of this book. If the July event was routine enforcement, what routine enforcement was it?

There are several innocent or ordinary possibilities. The files could have been subjected to a copyright workflow. They could have been made private by a bulk system event. An account-level integrity rule could have been triggered. A product migration or database failure could have changed visibility. A platform employee could have applied an enforcement category. Some or all of the files could have been treated differently because they were uploaded versions of published works. Even a mistake is possible.

The important point is that none of those explanations requires a theory about politics, Fauci, COVID origins, Bill Gates, Peter Thiel, virology, or government pressure.

The second hard question arises because the corpus was not random. It contained mathematics, artificial intelligence, biology, theology, political analysis, media criticism, historical writing, and a very large quantity of pandemic-era material. Among the pandemic papers were writings concerning Anthony Fauci, NIAID, gain-of-function research, Ralph Baric, Peter Daszak, EcoHealth Alliance, coronavirus origins, biodefense, and the politics of scientific consensus.

That content mix does not prove selective suppression either. In fact, if all subject areas lost public full texts together, that would initially favor an account-wide explanation over content-specific targeting. A useful investigation therefore begins by resisting the most emotionally satisfying conclusion.

The chronology, however, is difficult to ignore. On April 28, 2026, the U.S. Department of Justice announced the indictment of David M. Morens, a former senior adviser in NIAID's Office of the Director. DOJ alleged that Morens had participated in a scheme to evade Freedom of Information Act requests and federal records

requirements concerning communications about coronavirus research grants. The indictment was an allegation at that point, not a conviction.

Then came the ResearchGate event in July.

On August 18, 2026, Morens pleaded guilty to conspiracy to commit offenses and to defraud the United States. According to DOJ's description of his plea, Morens admitted participating in a scheme involving the use of his personal Gmail account to keep official communications outside government systems, exchange non-public NIH information, assist efforts concerning a terminated bat-coronavirus grant, and hide communications from public view in anticipation of FOIA requests. The guilty plea transformed a portion of the public-records controversy. One no longer had to speculate about whether at least one senior NIAID official had deliberately tried to keep relevant federal communications outside the ordinary public-record system. In the Morens case, the answer became yes.

> **ADMITTED OR ADJUDICATED:** Morens pleaded guilty to a federal conspiracy charge tied to evading FOIA and the Federal Records Act in connection with coronavirus-research-grant communications.

That fact does not connect Morens to ResearchGate. It does not connect NIAID to ResearchGate. It does not connect Fauci to ResearchGate. It does not connect any investor to the disappearance of my PDFs.

It does something subtler and more important: it changes the background probability that record preservation itself is a serious subject rather than a rhetorical flourish.

A person writing about COVID origins in 2021 could reasonably be accused of treating missing records as a speculative obsession. A person writing after a guilty plea concerning deliberate concealment of federal records is in a different evidentiary environment. The question is no longer whether documentary systems can be manipulated. The question is which systems were manipulated, by whom, under what authority, and with what consequences.

This book therefore begins with restraint. The ResearchGate event and the Morens crime are not presented as parts of a single operation. They are presented as two different events that force the same methodological question: how does a society reconstruct truth when the records through which truth must be reconstructed are unstable?

That is the vanishing-record problem.

A scientific paper is usually treated as a container of propositions. An email is treated as communication. A server log is treated as technical residue. A platform profile is treated as a convenience. But once a controversy becomes historical and legal, every one of those objects changes category. The paper becomes evidence that an argument existed at a particular date. The email becomes evidence of who was told what and when. The server log becomes evidence of who changed access. The platform profile becomes evidence of what the public could see before and after an intervention.

The most important information may not be in the paper at all. It may be in the metadata surrounding the paper.

That is why the July weekend matters. It is a beginning, not a conclusion.

A reconstruction before an accusation.

The first obligation after a mass digital disappearance is to freeze the scene conceptually. That means writing down exactly what the user could see before memory begins to reorganize events. Which browser was used? Was the researcher logged in? Did the publication pages load? Did download buttons disappear, or did links return errors? Were the PDF icons gone while titles remained? Was the change visible to an anonymous visitor as well as to the account owner? Did mobile and desktop views agree?

These details can seem trivial, yet they distinguish several technically different events. If the files remained visible to the account owner but not to anonymous visitors, privacy status may have changed. If publication

pages remained but the full-text control disappeared, the platform may have detached or restricted the file. If the publication pages themselves vanished, a different object class was affected. If old direct PDF URLs still worked, the user interface may have changed without the files being removed. Each pattern points toward different internal records.

A serious chronology should therefore use multiple clocks. The human clock records when I noticed the event. The platform clock, if ever produced, will record when state changes occurred. The search-engine clock may show when indexes stopped resolving PDFs. The Internet Archive clock may preserve the last public capture. Email timestamps show when I began asking what happened. These clocks need not align because crawlers and caches lag behind live systems.

Why the lack of warning is evidence but not proof of motive.

My account is that no meaningful warning preceded the event. That is relevant because ResearchGate's present policies describe notification in connection with specified enforcement actions. But a notification mismatch can arise in many ways. An email address may be stale. A bulk operation may use a different workflow from a single-item removal. A notice may be suppressed by a legal restriction. An automated job may fail to trigger the normal message. A product migration may fall outside the enforcement framework. A notice might even have been sent and lost in an email system.

This is why the correct evidentiary question is not simply, 'Did I receive an email?' It is, 'What does ResearchGate's system record about notification?' If a notification object exists with a timestamp and destination, that can be compared with mailbox records. If no notice object exists, that tells a different story. If the action record contains a field explaining why notice was not sent, the mystery narrows further.

The danger of retrospective compression.

Once the Morens guilty plea occurred one month later, the July event acquired a new emotional context. Human memory naturally compresses nearby events into a causal sequence: indictment, disappearing archive, guilty plea. A book written responsibly must resist that compression. The April indictment was public. A company employee could have known about it; there is no evidence one did. The July removal affected my account; there is no evidence it was related. The August plea established Morens's conduct; it did not retroactively reveal ResearchGate's motive.

This is a recurring problem in historical writing. Later knowledge changes the meaning we assign to earlier events. A memo that looked routine before a scandal looks ominous afterward. Sometimes the new interpretation is justified. Sometimes it is hindsight bias. The solution is contemporaneous documentation. What did I write to ResearchGate before the guilty plea? How did I describe the event before learning the new legal facts? Those messages are more probative of my original observation than a later reconstruction shaped by subsequent news.

The preservation opportunity.

The event also created an unexpected advantage. Because so much of the corpus survived elsewhere, it became possible to perform a before-and-after study at unusually large scale. A handful of missing papers invites anecdote. Thousands permit statistics. We can classify the archive by subject, publication date, DOI status, publisher status, file size, upload origin, and readership. We can ask whether journal-published PDFs behaved differently from self-authored preprints. We can ask whether public-domain or wholly author-owned works disappeared along with publisher-controlled material. We can ask whether files generated by the author himself were treated differently from articles bearing third-party copyright.

A copyright hypothesis makes predictions. If publisher rights were the primary trigger, the set of affected files should correlate with publisher-controlled versions and known rights complaints. If essentially all PDFs, including original preprints with no outside rights holder, disappeared together, a simple publisher-takedown explanation becomes less complete.

An anti-spam or account-integrity hypothesis makes different predictions. It would more naturally operate across subject matter and copyright status. A technical migration hypothesis might affect a class of file objects or upload dates. A politically selective hypothesis would predict topic-related differences or internal discussion.

The archive is therefore not merely the thing that was harmed. It is the experimental dataset for determining what kind of event occurred.

Chapter 2: Two Million Reads

ResearchGate built much of its identity around visibility. A scientific paper that once sat on a departmental shelf or behind a journal paywall could be found, followed, recommended, downloaded, and counted. The company introduced its Reads metric as a way to help researchers understand attention to their work. ResearchGate's own history describes this as part of a broader effort to modernize scientific reputation and dissemination.

> **AUTHOR'S ACCOUNT:** Before the mass disappearance of the full texts, my ResearchGate corpus had accumulated roughly two million reads. That figure should not be confused with two million unique human readers, two million complete readings, or two million endorsements. Platform metrics are never that simple. But it establishes that the account was not dormant. It was a highly visible, heavily accessed body of material within the platform's own measurement system.

That distinction is essential. Attention is not quality. A controversial paper can be widely read because readers reject it. A provocative title can generate clicks. Automated traffic can distort metrics. ResearchGate has changed its metrics over time. None of those caveats makes a large cumulative readership meaningless. The number remains relevant to materiality.

If a researcher has ten obscure documents and a platform changes their visibility, the historical effect is limited. If thousands of documents with sustained public traffic disappear from direct access, the event becomes part of the history of scientific communication. It matters even if the platform had a perfectly lawful and ordinary reason for acting.

The scale also raises a preservation question. ResearchGate was never designed as a national archive. It is a private company. Users accept terms of service. Features change. Files may be removed for rights reasons. Accounts may be restricted. Nothing about a high read count converts a private platform into a public repository with a perpetual duty to host every file.

Yet scholarly behavior can turn a private platform into a de facto archive even when the legal relationship says otherwise.

Researchers cite ResearchGate pages. Search engines index them. Readers rely on them when publisher links fail. Students discover papers there. Authors use platform-generated DOIs for certain uploads. ResearchGate itself has repeatedly described its mission in terms of making research open and accessible. In 2014 the company celebrated allowing researchers to upload work and make certain outputs permanently citable with a DOI. Such language encourages archival expectations even when the underlying service remains revocable.

The collision between expectation and legal reality is one of the themes of this book.

My own use of ResearchGate became unusually extensive. Over time the account became not simply a profile but an index into a very large AI-assisted writing program. The corpus ranged across scientific history, computational work, mathematics, genetics, artificial intelligence, theology, political analysis, and pandemic controversies. The sheer speed of production was unconventional and, to many readers, objectionable. That objection is itself part of the record. A platform may reasonably develop policies to handle spam, low-quality uploads, mass automation, copyright problems, or other forms of scale.

But the proper response to scale is still a documented process.

One can imagine a perfectly mundane internal note: account exceeds policy threshold; bulk full-text visibility change applied; notice sent. If such a record exists, it would immediately narrow the controversy. One can also imagine a copyright enforcement sequence showing specific complaints and associated actions. That too would be informative. The absence of an explanation to the user does not prove that ResearchGate lacks an explanation internally.

This is where the two-million-read figure becomes operationally relevant. A sophisticated platform often has escalation pathways for high-visibility accounts. Large accounts can produce more support traffic, more legal risk, more moderation complexity, and more reputational exposure. An investigator would therefore ask whether the account was handled by an automated system alone or escalated to employees with authority to make exceptional decisions.

> **OPEN QUESTION:** Did ResearchGate internally classify my account as a high-impact, high-volume, high-risk, copyright-sensitive, politically sensitive, or otherwise exceptional account before the July 2026 action?

The answer should not be guessed. It should be found in records.

A read metric also has another evidentiary function. It can help reconstruct timing. If historical screenshots, exported profile pages, emails, or archived HTML contain cumulative read counts, those numbers can provide time-stamped snapshots of the corpus before and after the disappearance. Changes in per-paper metrics could help identify whether pages were deleted, whether full texts were merely made private, and whether statistics survived.

ResearchGate's own help page explains that deleting an entire publication page can remove associated statistics, whereas removing a public full text can leave the publication page in place. That creates a testable forensic distinction. Did the titles and statistics remain while the PDF access disappeared? If so, the event may have been a visibility or full-text action rather than destruction of the publication records themselves.

Language matters. I have often described what happened as deletion because from the public reader's perspective the PDFs were gone. A legal or technical investigation should be more exact. A file can be deleted from storage, detached from a page, made private, blocked by rights management, hidden from public access, quarantined, or retained in backup while unavailable to users. Those states are not equivalent.

The book therefore uses 'disappearance' as the neutral umbrella term and reserves 'deletion' for cases in which the underlying file was actually removed.

This may sound pedantic. It is not. In digital forensics, the difference between inaccessible and destroyed can determine whether the record is recoverable. It can also determine what legal process can obtain.

The public story of a vanished archive begins with what a reader saw. The evidentiary story begins with what the database did.

What a read actually means.

Platform metrics invite both inflation and dismissal. A number such as two million looks authoritative because it is precise enough to be counted but vague enough to be misunderstood. ResearchGate's Reads metric has changed over time and should be interpreted according to the definitions in force when counts were accumulated. A 'read' is not equivalent to a journal citation, a verified download by a unique scientist, or a complete reading of a paper. It measures platform-defined engagement.

For this book, its value is not prestige. Its value is evidence that the corpus was visible enough to leave traces. A high-traffic account is more likely to be cached by search engines, referenced by other users, captured in screenshots, and represented in analytic emails. It may also be more likely to trigger automated integrity systems or manual review.

Metrics can therefore help reconstruct the platform's own understanding of the account. Did ResearchGate send periodic emails celebrating milestones? Did it rank the account in field-specific statistics? Were particular papers unusually active? Did the platform ever warn that volume was abnormal? Those communications can show whether ResearchGate treated the account as legitimate research activity before July 2026.

Hyper-prolific AI-assisted publishing as a platform stress test.

There is a larger story here that has nothing to do with COVID. Generative AI changed the economics of scholarly text production. A single researcher can now draft, revise, format, and disseminate work at a scale previously associated with departments or institutions. Existing scholarly platforms were designed around much slower human production rates.

That mismatch creates real governance problems. How should a platform distinguish a genuine high-volume research program from automated spam? How should it handle duplicate versions, machine-generated abstracts, or thousands of uploads in a short period? Should there be rate limits? Quality screening? Provenance labels? Different indexing rules? These are reasonable questions, and a platform can answer them without reference to political content.

My corpus sits near the frontier of that problem. It is therefore important not to assume that political sensitivity is the only plausible explanation for unusual platform treatment. An account with thousands of AI-assisted papers is itself unusual enough to attract automated or policy attention.

At the same time, unusual does not mean disposable. If a platform allows the material to accumulate, assigns metrics, creates publication pages, supports DOIs, and distributes it to readers, then an abrupt account-wide change should be explainable. The novelty of AI-assisted scholarship strengthens the case for transparent platform governance because the rules are still being invented.

Attention versus endorsement.

Some of my pandemic writings are intentionally provocative. A high read count may reflect controversy rather than agreement. That is not a weakness for the archival argument. Historically influential texts are often influential because people argue about them.

A platform's proper role in preserving the record should not depend on ideological approval. If a document violates law or policy, action may be warranted. If it is simply unpopular or wrong, preservation can still serve criticism. Future scholars should be able to examine why a claim attracted attention and how it changed over time.

This becomes particularly important when a platform metric itself is used to characterize an author. Removing the content while retaining the metric, or removing the metric while retaining the title, produces different historical distortions. A complete reconstruction should therefore capture both content and attention data.

The quantitative audit.

A rigorous audit could calculate at least five distributions: reads by subject category, reads by year, reads by publication type, reads by DOI status, and reads by removal state. If the most-read COVID papers show a distinct treatment pattern, that would be notable. If the distribution of affected material mirrors the distribution of the corpus as a whole, subject-specific targeting becomes less likely.

The audit could also identify outliers. A paper with tens of thousands of reads may have been handled differently from a paper with ten. If high-read items were preferentially reviewed, the mechanism may have been visibility-driven rather than topic-driven. Again, internal records would be needed to know why, but external statistics can guide the request.

The point is to move from the rhetorical force of 'two million reads' to a reproducible analysis of what those reads can tell us about the system.

Chapter 3: The Archive That Survived

The most important decision I made about my writing was not an argument. It was redundancy.

ResearchGate was only one surface on which the papers appeared. Copies survived elsewhere, particularly in Google Drive and later in broader archival efforts. That redundancy changes the meaning of the July 2026 event. Without it, the disappearance of thousands of full texts from a platform could have meant the practical loss of the corpus. With it, the event instead becomes an opportunity to compare states: what ResearchGate once exposed, what other archives retained, and what can still be reconstructed.

> **DOCUMENTED:** My connected Google Drive contains surviving copies of many publications that were previously available through ResearchGate. For example, the archive contains *Fermi Paradox, Viropaths, and Depopulation*, *Shadow Academia: Unveiling the Concealed Narratives of Global Health Crises*, *Biden's Presidential Pardon - Fauci Can Still Face Civil Lawsuits*, and *When the Witnesses Fear for Their Lives - Fauci Staff, Presidential Pardons, and a Hypothetical Biological-Release Testimony Crisis*. It also contains broader index materials describing a corpus of roughly six thousand AI-assisted preprints.

Those titles demonstrate why preservation matters. Some are restrained. Some are provocative. Some contain allegations or hypotheses that require careful separation from proven facts. A later reader should be able to see what was actually written rather than rely on a critic's summary or an author's later memory.

An archive preserves error as well as insight.

That is one of its virtues.

A healthy scientific record does not retroactively cleanse itself of bad predictions, overheated language, or positions that later became embarrassing. Historical understanding requires access to the mistakes too. When documents disappear, both defenders and critics lose the ability to audit what a person said at the time.

The Internet Archive serves a different preservation function. It captures public web pages rather than a researcher's private file store. A Wayback snapshot can establish that a page was publicly accessible at a particular time, subject to the limitations of web archiving. It can also preserve old links long after the original domain changes owners or disappears.

That became important for the Biological Manhattan Project, a site associated with my early-2000s concerns about biodefense, NIH, and dual-use biotechnology. A 2022 paper in my surviving corpus explicitly told readers that an old email concerning Fauci and the weaponization of anthrax had already been published on the now-archived Biological Manhattan Project website. Its reference pointed to a Wayback capture of biomanpro.org.

The claim is important not because a later paper proves the contents of the old email. It does not. It is important because it provides a contemporaneous chain of provenance within my own publication history: in 2022 I pointed readers back to an archived site rather than announcing the existence of the correspondence for the first time after the COVID controversy had matured.

> **DOCUMENTED:** The 2022 paper *Fermi Paradox, Viropaths, and Depopulation* contains a reference to a September 20, 2003 Wayback capture of `http://www.biomanpro.org/` and states that the email had been published on the archived Biological Manhattan Project site.

A separate archived path, `biomanpro.org/text/nih.htm`, has also been identified in the Wayback record as an NIH correspondence page captured in 2003. The exact HTML should be preserved independently and subjected to ordinary document authentication if it is ever used in a formal legal setting.

The point is methodological. When one platform loses a file, another platform's copy can prove content. When an old domain disappears, a web archive can prove public existence. When a publication cites the old domain years before a new controversy, it can strengthen provenance. When a Google Drive file contains an old

ResearchGate cover page with DOI and upload date, it can preserve metadata that the live platform no longer exposes.

No single archive is enough.

Cloud storage can be closed. Domains lapse. Platforms change terms. Search engines de-index. File formats become obsolete. Accounts die with their owners. A robust historical record therefore needs independent layers: local storage, cloud storage, web archives, institutional repositories, DOI registries, public archives, and where appropriate printed copies.

This is not paranoia. It is ordinary archival engineering.

The COVID-origins controversy demonstrates why such engineering matters. The debate depends heavily on early emails, grant documents, sequence databases, laboratory records, drafts, Slack messages, private accounts, meeting notes, and the exact dates at which scientific views were formed or revised. A missing email can alter a chronology. A missing grant attachment can hide a proposed experiment. A deleted webpage can erase evidence of what an institution publicly claimed before changing its language.

The same logic applies to my own corpus. If I argue that ResearchGate removed thousands of full texts, the proper response is not to ask readers to trust my memory. It is to reconstruct the before-and-after state using surviving files, archived pages, search indexes, screenshots, emails, and platform metadata.

The archive that survived is therefore not merely a backup. It is the evidentiary control.

Redundancy as scientific method.

Redundant storage is usually discussed as disaster recovery. In a controversial archive it also functions as replication. If the same file appears independently in Google Drive, an old ResearchGate download, an Internet Archive item, and a local disk, those copies can be hashed and compared. Identical hashes demonstrate that the byte sequence is the same. Different hashes prompt a version analysis.

This matters because a critic can reasonably ask whether an author altered an old paper after a controversy arose. A cryptographic digest calculated today cannot prove what a file contained in 2022 unless an older independent copy exists. But an Internet Archive capture or a file with independently verifiable historical provenance can anchor the content. Multiple copies with matching digests strengthen the chain.

The archive as a map of evolving claims.

My writing changed during the pandemic. Early pieces sometimes used more categorical language than I would use after later evidence emerged. Other hypotheses were revised or abandoned. A preserved corpus makes those changes visible. This is scientifically healthier than presenting only the latest polished view.

A historical archive should therefore preserve versions rather than overwrite them. Version 1.0 and version 1.1 may differ by only a paragraph, but that paragraph can matter if it changes an allegation or adds a source. File naming, modification times, and checksums should be retained in an index.

The same principle should apply to government and platform documents. When ResearchGate changes its Strikes Policy, the previous version remains relevant to actions taken under it. When NIH revises guidance, investigators need the version that governed at the time of a grant decision. Web archives are especially valuable for policy history because live pages often display only the current rule.

Archive provenance is stronger than screenshot culture.

Modern political disputes rely heavily on screenshots. They are easy to share and easy to misunderstand. A screenshot can crop out a date, omit context, or be altered. It can nevertheless be useful as a visual artifact if paired with stronger provenance.

For the Biological Manhattan Project, the ideal evidence bundle would include the archived URL, Wayback timestamp, raw HTML, embedded image or attachment URLs, and any surviving local copy. A screenshot can then serve as a convenient illustration, not the primary proof.

For ResearchGate, an exported HTML page or platform-generated PDF can preserve metadata more reliably than a photograph of a screen. Browser developer tools can reveal object identifiers and network requests, though those techniques should be used lawfully and without circumventing access controls.

Public archives versus personal archives.

A personal Google Drive is excellent for preservation but weak as public provenance because the owner controls it. A public Internet Archive item has different evidentiary properties because the archive is independently operated and timestamps deposits. Neither is perfect. A comprehensive preservation strategy uses both.

Public archives also create a social guarantee. If a controversial author dies, loses account access, or becomes incapacitated, the corpus can remain available. That matters for historical work. A private archive known only to the author may be lost even if every byte is technically intact.

The archival lesson of ResearchGate itself.

ResearchGate's mission language often emphasized openness and accessibility. The July event taught me to separate discovery platforms from preservation institutions. A platform optimized for current engagement is not necessarily designed for permanent custody. Its commercial incentives, legal obligations, and product architecture can change.

The correct response is not to abandon platforms. It is to use them as distribution layers while maintaining independent preservation layers.

In that sense, the disappearance may ultimately improve the archive. It forced a shift from assuming persistence to engineering persistence.

Chapter 4: An Email from 2002

One of the strangest features of the present controversy is temporal distance. The old correspondence is not from 2020 or 2021. It predates SARS-CoV-2 by almost two decades.

In August 2002, the United States was still living in the immediate shadow of the September 11 attacks and the anthrax-letter attacks that followed. The federal government was rapidly expanding biodefense programs. Money was moving into pathogen research, diagnostics, vaccines, detection systems, and emergency preparedness. At the same time, scientists and policy officials were confronting an old problem in a new form: the same biotechnology that could protect a population could also be misused.

My own professional life had already intersected with NIH's small-business programs. Kairos Scientific and related work had exposed me to the culture of federal biomedical funding and commercialization. Against that background, I sent correspondence to NIH officials in August 2002 concerning national security and NIH small-business funding opportunities.

> **AUTHOR'S ARCHIVAL RECORD:** The surviving provenance identifies an August 22, 2002 correspondence involving NIH officials including Jo Anne Goodnight and Catherine Lewis, with Anthony S. Fauci copied at an NIAID email address. Subject language associated with the thread includes 'national security at NIH?' and 'NIH SBIR/STTR Funding Opportunities.'

The fact that Fauci was copied is narrow evidence. It establishes that the message was addressed in a way that included his account. It does not establish that he opened it, read it, agreed with it, remembered it, acted on it, or even personally saw it. Any book that converts 'cc' into 'notice received and understood' would be overstating the record.

Yet a copied official can still matter historically. It places a concern inside the communication network of an institution at a specific time. If the concern related to dual-use biotechnology, biodefense, or security vulnerabilities in scientific funding and review, the email becomes part of the history of what issues were circulating before later controversies.

This is especially true because the correspondence appears to have been published publicly on the Biological Manhattan Project website by the early 2000s. That publication history makes it difficult to portray the document as something invented after COVID-19. The significance lies in chronology, not prophecy.

There is a temptation to read old warnings through the lens of everything that happened later. That temptation should be resisted. A 2002 concern about biological weapons does not predict SARS-CoV-2. A warning that open scientific systems can be exploited does not prove that a later laboratory incident occurred. A criticism of peer review from a security perspective does not establish that any particular grant was dangerous.

Historical documents should first be read in their own time.

In 2002, the anthrax attacks had demonstrated that biological material could become an instrument of terror and that attribution could be extraordinarily difficult. The attacks also exposed institutional discomfort with the overlap between legitimate research, biodefense, intelligence, and criminal investigation. Federal agencies had to decide how much biological information should remain open, how to vet researchers, how to classify dangerous techniques, and how to fund rapid innovation without creating new vulnerabilities.

Small-business programs added another layer. SBIR and STTR mechanisms were designed to move innovation into small companies, where government could stimulate technologies that large institutions might not pursue. That structure is valuable, but it also broadens the network of people handling sensitive concepts, materials, and applications. A security-minded critic might reasonably ask whether grant review itself could expose dangerous ideas to too many people or whether public solicitations could reveal strategic interests.

The old email therefore belongs in this book for three reasons.

First, it demonstrates that my engagement with NIH and biodefense questions did not begin with COVID politics.

Second, it shows how digital preservation can rescue context that would otherwise be lost. The original domain is gone as a living site, but the archived page and later references create a provenance chain.

Third, it provides a miniature version of the book's central problem. An email exists at the intersection of personal memory, institutional communication, web publication, and archival capture. Each layer tells us something different. None should be substituted for the others.

If this correspondence were ever provided to investigators, the proper evidentiary package would include the full archived HTML; complete email headers if available; original attachments; the Wayback capture metadata; contemporaneous copies from local backups; any replies; and an affidavit or declaration explaining custody and provenance. A screenshot alone would be weak. A chain of independent digital artifacts would be stronger.

This is how the book treats memory: as a lead to records, not a replacement for them.

What would make the 2002 correspondence legally interesting?

An old email becomes relevant to an investigation only through the issue it tends to prove. Its age alone is not significance. If the message merely copied Fauci on a routine funding announcement, it would be historically trivial. If it described specific biodefense vulnerabilities, warned about misuse of a funding mechanism, identified particular people or programs, or prompted a response, its value could increase.

The first task is therefore transcription without interpretation. Every line should be captured exactly from the archived source. Misspellings, headers, routing information, and quoted earlier messages should be preserved. Only then should a narrative summary be written.

The second task is authentication. A Wayback capture provides strong evidence that a web page was publicly archived at a recorded date, but legal authentication may still require testimony or records explaining how the capture was made. Local email files, if any survive, could provide full headers and Message-ID values. Recipient archives might independently contain the same message.

The third task is contextualization. Who were Jo Anne Goodnight and Catherine Lewis in 2002? What responsibilities did they have in NIH's small-business programs? What role did Fauci have as NIAID director? Was the subject within his institute's remit? Was the message sent directly to him or copied as part of a broad distribution? Did anyone reply?

The fourth task is materiality. Even a genuine email can be legally irrelevant if it has no connection to the conduct under investigation. A prosecutor would not care simply because a famous name appears in the CC line.

Why publication on biomanpro.org matters.

Publishing the correspondence on a public website in the early 2000s creates an unusual provenance advantage. It shows that the document was not kept secret and unveiled only after COVID became politically salient. The archived website can establish that the author's concern existed in that historical period.

This can protect against a common weakness in retrospective memoir: the tendency to reinterpret old experiences as warnings that were more precise than they actually were. The archived text can show what the warning really said.

If the old wording turns out to be less dramatic than later memory, that should be acknowledged. If it is more specific, that should also be acknowledged. The archive gets the final word.

The anthrax context.

The 2001 anthrax-letter attacks transformed U.S. biodefense policy. The attacks killed five people, sickened others, closed government and postal facilities, and generated a massive investigation. They also demonstrated

the difficulty of attribution. For years, uncertainty surrounded the source. The episode sharpened national attention on biological threats and accelerated investments in preparedness.

Within that context, discussions of NIH national security were not exotic. The government was actively trying to expand research capacity while worrying about how scientific openness could be exploited. This is the environment in which a 2002 concern about dual-use funding should be read.

Fauci's position in the chronology.

Anthony Fauci had already been NIAID director for many years by 2002. He was therefore a senior figure in infectious-disease policy long before the pandemic made him a household name. The fact that his name appears in a pre-COVID correspondence can feel uncanny only if one forgets that infectious-disease networks are persistent. The same senior officials, institutions, and scientific communities often remain active across decades.

This persistence is precisely why old records can matter. Institutional memory spans crises: HIV/AIDS, bioterrorism, SARS, influenza, MERS, Ebola, and COVID-19. Policies and relationships built during one crisis can shape responses to the next.

The email belongs in that long history. It should not be made to bear more weight than its text can support.

Chapter 5: Before COVID Had a Name

The intellectual prehistory of the COVID-origin controversy begins long before Wuhan. It begins with the normalization of increasingly powerful biological manipulation.

Modern molecular biology was built on the proposition that mechanisms can be understood by altering them. Mutate a gene and observe the phenotype. Change a receptor and watch binding. Delete a regulatory sequence and measure expression. Introduce a plasmid and trace function. In most fields this experimental logic is uncontroversial because intervention is how causality is learned.

Viruses create a special problem. Their relevant traits include host range, transmissibility, immune escape, replication efficiency, pathogenicity, and tissue tropism. Experiments that alter these traits can generate knowledge with obvious medical value. The same experiment can also create a pathogen with properties that did not previously exist in the laboratory stock.

The phrase 'gain of function' became politically overloaded because it covers more than one scientific idea. In the broadest sense, any mutation that adds or increases a biological function can be described that way. In policy debates, the phrase often refers more narrowly to research expected to enhance dangerous properties of potential pandemic pathogens. Different agencies and commentators have used different definitions at different times. Much public argument therefore consisted of people using the same phrase while referring to different categories of experiment.

The deeper issue is dual use.

Dual-use research produces information or capabilities that can serve both beneficial and harmful purposes. A method for increasing vaccine yield can potentially help production. A method for altering host range can help model emergence. A method for reconstructing a virus can preserve research capacity if a natural sample is lost. Each can also lower barriers to misuse.

After the 2001 anthrax attacks, the United States dramatically expanded biodefense. That expansion brought money, facilities, personnel, and political attention. It also created incentives. New funding streams attract new research programs. New programs create communities. Communities create norms. Over time, an experiment that once would have sounded extraordinary can become a standard technique inside a specialized field.

This is not unique to virology. Weapons laboratories, nuclear physics, cryptography, artificial intelligence, and cyber security all demonstrate the same institutional phenomenon: once a capability becomes technically possible and strategically relevant, prohibition becomes difficult because knowledge itself is distributed.

The scientific establishment responds with governance: biosafety levels, institutional biosafety committees, select-agent rules, DURC review, grant conditions, publication review, export controls, and professional norms. Critics respond that these mechanisms are fragmented and often depend on the same scientific community whose incentives favor research continuation.

The conflict is not simply science versus anti-science. It is a conflict among scientific values: openness, reproducibility, curiosity, safety, accountability, and precaution.

That distinction matters for interpreting my earlier writings. Some of them used language far sharper than the language I use here. *Fermi Paradox, Viropaths, and Depopulation*, written in 2022, argued from the possibility that repeated high-risk pathogen work creates a cumulative civilizational risk. It linked that risk to a Fermi-paradox-style question: if technological civilizations routinely develop biological capabilities that can destroy them, perhaps self-extinction is one answer to cosmic silence.

That is a speculative thesis, not an established scientific result. But embedded within the speculation was a serious governance question: what happens when the number of laboratories, experiments, engineered traits, and opportunities for error increases faster than the reliability of containment and oversight?

The mathematics of cumulative risk is unforgiving. If an event has a small probability per trial but the number of trials becomes very large, the probability of at least one failure rises. The exact probability for pathogen escape is not known well enough to support precise calculation, and experiments differ radically in hazard. Still, the general principle is elementary.

A civilization can therefore create systemic risk without any individual actor intending catastrophe.

This is an important alternative to conspiracy thinking. Catastrophic outcomes do not require a central planner. They can emerge from incentives, institutional fragmentation, publication competition, national rivalry, funding abundance, and the human tendency to normalize previously exceptional practices.

That is why the record matters. To govern risk, society needs to know what experiments were proposed, funded, performed, modified, reported, and discussed internally. It needs accurate grant records. It needs laboratory notebooks. It needs incident reports. It needs emails. It needs sequence provenance. It needs conflict-of-interest disclosures. It needs records of who raised concerns and how institutions responded.

When those records are missing, the scientific question and the governance question become inseparable.

The pandemic did not create that problem. It exposed it.

Risk without intent.

One reason biological-risk debates become conspiratorial is that people assume a catastrophe must have either a natural cause or an evil planner. Complex systems offer a third category: dangerous outcomes produced by ordinary incentives.

A young investigator needs publications. A laboratory needs grants. A funding agency wants preparedness. A nation wants strategic capability. A journal wants novelty. A pharmaceutical company wants products. Each actor can pursue a locally rational goal while the global system accumulates risk.

This resembles financial systemic risk. No mortgage broker had to intend a global financial crisis for correlated incentives to produce one. In biotechnology, no scientist has to desire a pandemic for research systems to create opportunities for accident.

The governance question is therefore not, 'Are virologists evil?' It is, 'What feedback mechanisms prevent individually rational choices from producing unacceptable aggregate risk?'

Biosafety levels are necessary but not sufficient.

Public discussion often treats BSL numbers as if they were simple hazard scores. In reality, biosafety levels describe combinations of practices, equipment, facility design, and containment appropriate to categories of work. The relevant level depends on the agent and procedure. The same organism may be handled differently depending on whether the work involves culture, animals, aerosols, or noninfectious material.

A high-containment laboratory reduces risk but does not make risk zero. Human error, equipment failure, protocol drift, waste handling, transportation, and maintenance all matter. Conversely, not every coronavirus experiment requires maximum containment. The technical details of a procedure determine the hazard.

This is why records of actual experiments matter more than slogans about BSL-2 or BSL-4.

The publication dilemma.

Dual-use research also creates a conflict over publication. Science advances through disclosure, but disclosure can spread capabilities. Restricting publication may slow harmful replication while also slowing countermeasure development and independent scrutiny.

The United States has experimented with advisory mechanisms for dual-use research of concern, but no system can eliminate the tension. Once a technique is widely known, suppressing one paper may have little effect. Before a technique becomes common, publication decisions can matter more.

AI and automated biology intensify this dilemma. The barrier between a sequence and a physical construct continues to shrink as synthesis, design software, and laboratory automation improve. The record of what was done becomes more important even as the capability to reproduce it becomes more distributed.

Why pre-COVID history should reduce polarization.

Looking backward can depersonalize the debate. Host-range modification, reverse genetics, and viral engineering did not begin with Fauci, Daszak, Baric, or Wuhan. They emerged from decades of molecular biology. The relevant policy problem is therefore larger than any current political target.

If the only lesson of COVID is 'punish the people on my side's enemy list,' the next risk will survive untouched. The more durable lesson concerns incentives, transparency, oversight, international standards, and record integrity.

Before COVID had a name, the technical capabilities and governance tensions already existed. That fact should move the debate away from personalities and toward systems.

Chapter 6: From Chow to Rottier to Baric

A useful history of controversial virology should not begin in 2019. Techniques that later became politically charged were built incrementally, often for ordinary scientific reasons, by researchers who could not have anticipated the specific pandemic debate that would surround their fields decades later.

My surviving archive contains a paper titled *Tracing the Origins of Gain-of-Function Research to Polio (Chow 1991) and Coronavirus (Rottier 2000)*. The title reflects an attempt to push the historical frame backward: away from the personalities of the COVID era and toward the evolution of experimental methods.

The first lesson of that history is definitional. A virologist changing a viral protein to test receptor usage may not describe the work in the policy vocabulary later attached to 'gain of function.' A retrospective critic may nevertheless view the experiment as part of the technical lineage through which host range and pathogenic traits became manipulable. Both descriptions can be true if they operate at different levels.

This is why historical reconstruction must distinguish technique from policy category.

Poliovirus research in the late twentieth century included sophisticated manipulation of viral genomes, antigenic sites, host interactions, and vaccine-related phenotypes. Such work contributed to basic understanding and to public health. It also demonstrated that a virus could be deliberately modified in ways that changed biologically important behavior. That fact was not scandalous; it was the purpose of experimental genetics.

Coronavirus research followed a similar trajectory. By 2000, work associated with Peter Rottier and colleagues had shown that coronavirus tropism and host interactions could be experimentally altered through molecular manipulation. Again, the scientific significance was obvious. Coronaviruses were not immutable natural objects. Their entry machinery could be studied by changing it.

As reverse genetics improved, researchers no longer had to wait for nature to provide every viral variant. Genetic material could be assembled, exchanged, deleted, inserted, or rescued into viable virus under laboratory conditions. Ralph Baric's laboratory at the University of North Carolina became especially important in the development and application of coronavirus reverse-genetics systems. These systems enabled detailed study of replication, pathogenicity, cross-species potential, antiviral susceptibility, and vaccine design.

> **DOCUMENTED:** Coronavirus reverse genetics became a mature experimental technology before the emergence of SARS-CoV-2. Baric and other coronavirus researchers used such systems to study viral function and emergence. The existence of these methods is not evidence that SARS-CoV-2 was engineered or that any particular laboratory caused the pandemic.

That final sentence must be repeated because technical capability and historical causation are different propositions.

A laboratory may possess the capability to make a class of modification without having made the virus that later appears in nature. A scientist may publish work that demonstrates host-range manipulation without being responsible for a later outbreak. A funding agency may support coronavirus research without being responsible for every experiment performed under every downstream collaboration. The history establishes capacity and institutional interest. It does not by itself establish origin.

Yet capacity is not irrelevant. Origin investigations necessarily ask what viruses existed, what samples were held, what manipulations were feasible, what proposals were submitted, what experiments were performed, and what records survive. A laboratory-origin hypothesis cannot be evaluated without understanding the technical landscape.

This is one reason the early dismissals of laboratory-origin discussion were so damaging. If a hypothesis is treated as socially disqualifying before the relevant records are available, evidence gathering becomes entangled with reputation management. Scientists may fear that asking a question will be interpreted as endorsing a

political faction. Journalists may avoid lines of inquiry because they have been labeled misinformation. Institutions may communicate defensively because they perceive existential reputational risk.

The result can be epistemic lock-in.

The opposite danger is equally real. Once laboratory origin becomes politically attractive, every old experiment can be reinterpreted as a step in a hidden plan. Ordinary technical lineage is converted into a conspiracy genealogy. That is not sound reasoning either.

The more defensible approach is to build a capability timeline and keep it separate from the event timeline.

A capability timeline asks: when did scientists learn to alter viral host range? When did reverse genetics permit assembly of infectious coronaviruses? When did chimeric constructs become routine enough to support systematic experiments? When did animal models with human receptors become available? When did sequence synthesis become cheap enough that digital information could substitute for physical transfer?

An event timeline asks: what samples were collected, where were they stored, what grant aims were active, what experiments were actually performed, what illnesses occurred, what sequence data appeared, and what communications took place around the outbreak?

The two timelines intersect only where evidence shows they intersect.

My earlier writing often used the first timeline to criticize the governance of the second. The strongest version of that critique is not that every dangerous technique caused COVID-19. It is that scientific and regulatory institutions had decades of warning that viral traits could be manipulated, and therefore pandemic preparedness should have included rigorous, auditable records of high-risk research.

The Morens case makes that governance issue harder to dismiss. If a senior NIAID adviser deliberately kept grant-related communications outside official systems in anticipation of FOIA, then the documentary layer needed to assess controversial research was itself compromised.

That is why a 1991 polio paper or a 2000 coronavirus paper belongs in a book about disappearing records. The technical history tells us what scientists learned to do. The archival history tells us whether society retained enough information to know when, where, and why they did it.

Technique is not accusation.

The most misleading sentence in origin debates often has the form, 'Scientists could do X, therefore X happened.' Technical feasibility is a prerequisite for some hypotheses, but it is rarely sufficient evidence. A forensic investigator separates capability, opportunity, and event evidence.

Capability evidence shows that a technique existed. Opportunity evidence shows that relevant materials, people, and facilities were in a position to use it. Event evidence connects a specific technique to the object under investigation through records, sequence signatures, laboratory documentation, witnesses, or other direct traces.

Coronavirus reverse genetics supplies capability evidence. Proposals and grant aims can supply opportunity evidence. Neither automatically supplies event evidence for SARS-CoV-2.

This hierarchy should apply equally to natural-origin reasoning. Demonstrating that zoonotic spillover is common and biologically plausible supplies capability, in the broad sense. Identifying susceptible animals and market conditions supplies opportunity. Establishing a direct chain from an infected animal to early human cases would supply stronger event evidence.

A fair investigation demands symmetrical standards.

The significance of experimental lineage.

Even when it does not establish origin, technical lineage matters for governance. If the scientific community has known for decades how to alter host range, immune recognition, or pathogenic traits, regulators cannot claim

that modern dual-use problems appeared without warning. The sophistication of oversight should track the sophistication of the tools.

This is why older papers can matter more as policy history than as forensic evidence. They establish the trajectory by which manipulation became routine. They help answer whether governance evolved at the same rate.

Baric as a case study in scientific ambiguity.

Ralph Baric's work is frequently invoked by both critics and defenders of coronavirus research. To admirers, reverse-genetics platforms are precisely the technologies needed to understand emerging viruses and develop countermeasures. To critics, the same platforms exemplify how laboratories can create chimeric or altered viruses with uncertain risk.

Both descriptions can be grounded in the same methods. The disagreement concerns risk-benefit judgment and relevance to specific events.

That makes Baric an instructive figure for the book's broader theme. A scientific fact can be stable while its institutional meaning is contested. The record must therefore preserve enough detail for later readers to reconstruct the experiment without inheriting the political label attached to it.

A future technical appendix.

A later edition of this book could profitably include a neutral technical appendix listing representative milestones in viral genetics: infectious clone systems, reverse genetics, receptor substitution, chimeric spike experiments, animal-adaptation protocols, pseudovirus systems, and sequence synthesis. Each entry could specify what the technique enables, what it does not imply, and which records would be needed to connect it to an origin hypothesis.

That would turn a politicized vocabulary into an experimental map.

Chapter 7: NIAID and the Architecture of Funding

Federal biomedical research is not a single laboratory with a single chain of command. It is a network of institutes, program officers, review panels, grants, contracts, universities, nonprofit organizations, subawards, principal investigators, collaborators, and foreign partners. That complexity matters because public arguments often compress the network into a misleading sentence such as 'Fauci funded the research' or 'NIH did the experiment.'

A funding agency can influence science powerfully without directly controlling every pipette.

The National Institute of Allergy and Infectious Diseases is one of the institutes within NIH. For decades it has supported research on infectious disease, immunology, vaccines, therapeutics, emerging pathogens, HIV, influenza, and many other subjects. Its director sets priorities and represents the institute publicly, but individual grants move through administrative and scientific processes involving many people.

A principal investigator submits an application. Scientific reviewers evaluate significance, approach, investigators, environment, and other criteria. Program and grants-management staff oversee the award. Institutions accept compliance obligations. Subawards may transfer portions of the work to other organizations. Progress reports describe activity. Changes in scope may require approval. Publications emerge downstream.

That architecture distributes both responsibility and information.

It also creates documentary seams.

A grant file may contain the formal aims but not every informal scientific conversation. An email may reveal what program officials thought about a proposal but not what reviewers saw. A subaward agreement may describe work at a partner institution while laboratory notebooks contain the actual experiments. A progress report may summarize results without preserving failed attempts. A private Gmail account can contain official business that the government system does not initially capture.

The public record is therefore assembled from layers that were created for different purposes.

This matters especially for the EcoHealth Alliance coronavirus-grant controversy. DOJ's 2026 Morens materials describe a NIAID-funded grant called *Understanding the Risk of Bat Coronavirus Emergence*. According to DOJ, the grant was awarded to an organization whose principal investigator made a subaward to the Wuhan Institute of Virology. NIH later terminated the grant amid the intense origins controversy. Morens and others then engaged in communications about restoring funding and responding to the lab-leak narrative.

The book deliberately follows DOJ's co-conspirator and company labels where the plea materials do, rather than expanding criminal attribution by inference. Public reporting and congressional documents may identify people associated with the underlying grant, but an evidence-disciplined narrative does not need to convert every association into a criminal allegation.

The importance of the grant lies partly in how it made funding architecture visible to the general public. Many citizens learned for the first time that a U.S. institute could fund a domestic organization that in turn supported research at a Chinese laboratory. That arrangement is not inherently improper. International collaboration is normal in infectious-disease research because pathogens do not respect borders and because field samples often originate where outbreaks occur.

But international subawards create oversight challenges. The primary grantee may be responsible for monitoring work it does not physically perform. U.S. officials may depend on reports from intermediaries. Foreign partners may operate under different legal systems, biosafety cultures, transparency norms, and political pressures. If a dispute arises, access to underlying records may be limited.

The governance problem becomes more severe when scientific definitions themselves are contested. If an agency says funded work did not meet a particular policy definition of enhanced potential pandemic pathogen

research while critics describe the same experiments as gain of function in a broader sense, the public hears contradiction where the underlying dispute may partly be definitional.

That does not mean definitions are merely semantic. Policy definitions determine what review mechanisms are triggered. They determine which experiments receive additional scrutiny. They can therefore influence safety.

A prosecutor, inspector general, congressional investigator, or historian examining such a system needs more than published papers. The relevant evidence includes grant applications, scoring statements, progress reports, prior approvals, correspondence about scope, biosafety documentation, meeting notes, internal legal interpretations, and records of who was informed about potential problems.

The Morens guilty plea matters because it shows that some official communications concerning this funding controversy were intentionally routed through a private Gmail account. DOJ states that the conspirators used that account to exchange non-public NIH information, discuss efforts to influence NIH funding, edit letters to NIH leadership, and back-channel information to a senior NIAID official.

> **ADMITTED OR ADJUDICATED:** In his guilty plea, Morens admitted participation in a conspiracy to keep federal records outside the ordinary government record system in anticipation of public-record requests.

This is not a technical violation with no epistemic consequence. FOIA is one of the mechanisms through which outsiders reconstruct how public institutions made decisions. If official business is deliberately moved to private systems to defeat retrieval, the resulting historical record becomes biased toward what officials were willing to leave discoverable.

That is a form of selection bias imposed on history.

A scientist would never accept a dataset if subjects could secretly remove the observations most likely to produce an unfavorable result. Yet public-record analysis can face exactly that problem when officials control which communications remain in official systems.

The solution is not to assume that every missing email contains a scandal. The solution is preservation, recovery, and transparent accounting of gaps.

NIAID's architecture of funding is therefore also an architecture of records. Understanding the science requires understanding both.

The problem of delegated oversight.

Funding networks often operate through delegation. NIH delegates performance to a grantee institution. The grantee delegates a portion of work through a subaward. Institutional biosafety committees review work locally. Program officers receive reports. Foreign collaborators may operate under their own national rules.

Delegation is efficient because no central agency could directly manage every experiment. It also diffuses accountability. When something goes wrong, each layer can point to another: the agency relied on the grantee, the grantee relied on the subrecipient, the subrecipient complied with local procedures, and everyone relied on the definitions in force.

This is a classic principal-agent problem. The principal funds work but lacks complete information about the agent's behavior. Reporting requirements exist to reduce that information gap. If reporting is incomplete or informal communications migrate outside official systems, the gap widens.

Grant language as evidence.

Grant proposals deserve careful reading because they occupy a peculiar status. They describe intended work, often in ambitious language designed to win funding. Not every proposed experiment is ultimately performed. A proposal can therefore prove intention or contemplated scope without proving execution.

Progress reports move closer to execution evidence but can still summarize selectively. Publications prove that some work was completed but may omit failed, preliminary, or unpublished experiments. Laboratory records are closer still.

This hierarchy is important when public arguments quote a dramatic phrase from a proposal and treat it as proof that a specific experiment occurred. The correct question is what downstream records show.

Oversight incentives.

Program officials face their own incentives. Successful grants produce publications and progress. Terminating work can generate conflict. Admitting that an award was poorly supervised can harm an institution. None of these incentives proves misconduct, but they are reasons to design review systems that do not depend entirely on self-report.

Independent audits, clear definitions, mandatory disclosure of material scope changes, and durable communication systems can reduce these conflicts.

Why records should follow the money.

Financial systems routinely preserve audit trails because money is fungible and incentives matter. High-risk research should receive at least comparable documentary discipline. If federal funds support work through multiple entities, the record should allow an investigator to trace what was authorized, paid for, performed, modified, and reported.

The phrase 'follow the money' is sometimes used conspiratorially. In public administration it has a simpler meaning: funding paths define responsibility paths. Each transfer should have associated records.

The Morens case shows why informal back channels can undermine that architecture. Even if the formal grant file is complete, important discussions about restoring or defending the award can exist elsewhere. The administrative record and the real decision process can diverge.

Chapter 8: The Gates-Fauci Global-Health Era

Bill Gates and Anthony Fauci became symbols in the political imagination of the COVID era, often in ways that obscured the mundane institutional reality of their relationship. The documented connection is neither secret nor surprising. For years before COVID-19, the Gates Foundation and NIAID occupied overlapping territory in global health: vaccines, HIV, infectious disease, research infrastructure, product development, and access in low-resource countries.

> **DOCUMENTED:** In December 2010, the World Health Organization, UNICEF, NIAID, and the Bill & Melinda Gates Foundation announced the Decade of Vaccines Collaboration. The Gates Foundation's announcement named Fauci, then NIAID director, among the leaders involved in the collaboration. The initiative aimed to coordinate vaccine discovery, development, delivery, and access.

> **DOCUMENTED:** In October 2019, NIH and the Gates Foundation announced a collaboration in which each planned to invest at least $100 million toward developing affordable gene-based cures for sickle cell disease and HIV. NIH's announcement quoted Fauci describing the collaboration as an ambitious effort to use NIH's global HIV research infrastructure in pursuit of a scalable cure.

Those are direct institutional ties. There is no need to embellish them.

Philanthropic foundations have become major actors in biomedical science because they can commit large sums outside ordinary appropriations cycles. They can focus on diseases neglected by wealthy markets. They can fund technology platforms, product development, clinical partnerships, surveillance, and global delivery. Governments bring regulatory authority, national research infrastructure, public accountability, and much larger long-term budgets. Partnerships between the two are therefore predictable.

The same structure also raises governance questions. When a private foundation becomes deeply integrated into public-health strategy, traditional categories blur. A foundation is not a government agency, yet its funding priorities can shape public institutions. Its executives are not elected, yet they may participate in global policy networks. Its investments can be mission-driven and beneficial while still concentrating agenda-setting power.

Those are legitimate subjects of political analysis without implying secret control.

The ResearchGate connection is separate but documented. In June 2013, ResearchGate announced a $35 million financing round led by private investor Bill Gates and Tenaya Capital, with participation from other investors. ResearchGate later published a photograph of Gates visiting its Berlin headquarters. The company described Gates as an investor and quoted him praising the potential of digital tools to improve science.

Again, no inference is required: Gates invested in ResearchGate.

But investment is not operational control. Public investment announcements do not tell us whether an investor held a board seat, possessed veto rights over moderation, received information about individual users, or had any role in content decisions. In the absence of such evidence, a book should not imply that a shareholder relationship explains a moderation event thirteen years later.

The same caution applies to Founders Fund. ResearchGate announced in February 2012 that Founders Fund led its Series B financing, with partner Luke Nosek leading the investment and joining the board. Peter Thiel co-founded Founders Fund, making his name relevant to the firm's history, but the evidence supports the statement that Founders Fund invested. It does not automatically support the statement that Thiel personally directed ResearchGate.

Why include these relationships at all?

Because institutional maps are useful when they are treated as maps rather than verdicts.

A map tells investigators where records may exist. Investors may have board materials. Board members may receive risk reports. Foundations may communicate with agencies. Companies may develop policies in response to legal or reputational pressure. Public-private partnerships create channels through which information travels.

The existence of a channel does not prove that a particular message traveled through it.

That distinction is the difference between network analysis and conspiracy theory.

The COVID era encouraged both extremes. One side often treated any mention of elite institutional relationships as irresponsible insinuation. The other treated the mere existence of overlapping networks as proof of a coordinated plan. Both approaches are epistemically weak.

A better method asks four questions in sequence.

First, does the relationship exist?

Second, what formal powers or obligations does the relationship create?

Third, is there evidence of communication about the event under investigation?

Fourth, did that communication cause or materially influence the event?

The Gates-Fauci relationship passes the first question easily. The Gates-ResearchGate investment relationship also passes the first. The ResearchGate-founders' biomedical backgrounds pass the first. None of those facts, at present, answers the third or fourth question concerning my July 2026 full-text disappearance.

This is why compulsory records could matter. If ResearchGate's internal communications contain no relevant external contact, the network hypothesis loses force. If they contain a request from a rights holder or a routine automated policy action, the explanation may be ordinary. If they contain unusual executive escalation tied to named COVID-origin papers or outside pressure, then a different evidentiary question arises.

The map tells us where to look. It does not tell us what we will find.

Philanthropy and public power.

The Gates Foundation's role in global health deserves neither demonization nor exemption from scrutiny. Large philanthropy can do things governments find difficult: make long-horizon bets, fund neglected diseases, convene international partners, and absorb technological risk. It can also concentrate agenda-setting power in private hands.

The governance question is not whether philanthropy is good or bad. It is how public institutions maintain accountability when private capital becomes integral to public strategy.

Formal agreements, advisory roles, co-funding arrangements, and public announcements create a record. Informal influence is harder to measure. A rigorous institutional history therefore distinguishes money, governance rights, scientific collaboration, and personal relationships.

Investment is not control.

The ResearchGate example illustrates another common confusion. Venture investment can carry board rights, information rights, preferred shares, or other contractual powers, but public press releases rarely disclose the complete governance package. Saying that Gates invested in ResearchGate is safe because ResearchGate said so. Saying that Gates controlled its moderation would require evidence about corporate governance and the specific decision.

Similarly, Founders Fund's investment makes Peter Thiel part of the broad financial history because he co-founded the firm, but the 2012 ResearchGate announcement specifically identifies Luke Nosek as leading the investment and joining the board. Precision prevents a network map from turning into a celebrity map.

Why the network remains worth mapping.

Even without proving causation, financial and institutional relationships tell an investigator where relevant records may have been distributed. Board members may receive high-level reports. Major investors may be informed about crises. Philanthropic partners may exchange policy documents with government agencies. These are potential repositories of evidence.

A good map therefore includes dates. A relationship that ended years before an event may have little operational significance. A board seat active at the time of an event is different from a historical investment. Temporal accuracy is as important as relational accuracy.

The 2019 collaboration as a pre-pandemic marker.

The October 2019 NIH-Gates Foundation announcement about gene-based cures for HIV and sickle cell disease is useful because it demonstrates a robust institutional relationship immediately before the pandemic, in a context unrelated to SARS-CoV-2. It shows that NIAID and the foundation were already accustomed to coordinating major biomedical initiatives.

That fact should be interpreted at face value. It explains why Gates and Fauci appear in the same global-health network. It does not explain ResearchGate's 2026 action.

The distinction between contextual relevance and causal relevance is one of the book's recurring disciplines.

Chapter 9: ResearchGate

ResearchGate was born from a frustration familiar to scientists: research is produced inside communities but published through systems that often make communication slow, fragmented, and inaccessible. In 2008, physicians Ijad Madisch and Sören Hofmayer developed the idea for a scientific network and brought in computer engineer Horst Fickenscher as the third founder.

ResearchGate's own tenth-anniversary history describes the founding in unusually personal terms. Madisch and Hofmayer were medical doctors; Fickenscher supplied the technical expertise. The company grew from ten thousand members in 2008 to millions within a few years.

Madisch's scientific background is particularly relevant to the atmosphere of this book because his doctoral work was in virology, including human adenoviruses. Hofmayer's biomedical work also included virology and immunology-related research. Those backgrounds are interesting. They are not evidence of a COVID-related motive in 2026.

The distinction is worth making explicit because coincidence is psychologically powerful. A researcher critical of virology sees his work disappear from a company founded by biomedical scientists with virology experience, financed in part by Gates, whose foundation collaborated with Fauci. The mind naturally draws a line through the nodes.

A prosecutor cannot draw that line without communications.

ResearchGate's funding history shows how rapidly the company became connected to major capital. Benchmark invested in 2010. Founders Fund led the Series B in 2012. Bill Gates and Tenaya Capital led a $35 million investment announced in 2013. Later financing included the investment arm of Wellcome Trust and Goldman Sachs Investment Partners, among others. ResearchGate itself has published these milestones proudly.

The investor list tells us that sophisticated parties believed the platform could become important infrastructure for science.

That belief turned out to be justified. By aggregating profiles, publications, citations, questions, projects, and metrics, ResearchGate occupied a space between social network, bibliographic index, repository, and professional identity service. It was not a traditional publisher, but it became part of the publication ecosystem.

This hybrid status created recurring copyright conflict. Publishers objected when users uploaded articles in versions that contractual rights did not permit. ResearchGate negotiated arrangements with publishers and created systems to check uploads against rights information. Its policies allow removal or restriction of content in response to intellectual-property complaints.

> **DOCUMENTED:** In a 2018 cooperation announcement with publishers, ResearchGate said that when participating publishers identify content they believe infringes copyright, ResearchGate removes it expeditiously and informs the scientist who uploaded it. The company also stated that the arrangement was not intended to let publishers preemptively block or filter research uploads.

That historical statement is relevant to my account because I report receiving no notice explaining the disappearance of thousands of full texts.

There are several possibilities. The July event may not have been a publisher-removal process covered by that statement. The policy may have changed. Notices could have failed in delivery. An exceptional legal restriction could have prevented disclosure. A bulk visibility action might not have been classified as removal. Or the platform might simply have failed to follow its ordinary notification practice.

Without internal records, one cannot know.

ResearchGate's present Appeals Policy is clearer. It says users may appeal certain actions removing, restricting, or disabling access to content or accounts. It states that when ResearchGate takes qualifying action, it will notify

the user of what action was taken and why, including the rule or guideline believed to have been violated. The policy also says that removed content cannot necessarily be technically restored, although the user may be allowed to upload it again if the action was erroneous.

Its Strikes Policy likewise contemplates warnings, strikes, content actions, and account disabling. Copyright matching by the platform's automated system does not automatically count as a strike, although later removal following a publisher infringement report may do so.

These distinctions matter because my own account history, as I remember and preserve it, does not fit a simple three-strikes narrative.

> **AUTHOR'S ACCOUNT:** I had no warning that thousands of public full texts were about to become unavailable. I did not receive a sequence of strikes explaining an escalating enforcement process. After the event, repeated inquiries produced no substantive explanation.

That discrepancy is enough to justify an audit question. It is not enough to prove bad faith.

ResearchGate is also a German company operating within a European legal environment. Its terms, data-protection obligations, Digital Services Act responsibilities, and corporate structure may affect what records exist and what forms of legal process can reach them. The company also has a U.S. corporate presence. Any real subpoena strategy would therefore require lawyers to determine which entity possesses or controls the relevant records and what legal mechanism applies.

This book is not a subpoena. It is a map of what a subpoena would need to ask.

The most valuable ResearchGate evidence may be almost embarrassingly technical: an internal event record containing an account identifier, timestamp, operation, reason code, actor identifier, and ticket number. A single database row can be more probative than a hundred pages of speculation about investors.

That is the paradox of digital power. The public sees a mystery. The server may see a field.

A platform between repository and publisher.

ResearchGate's ambiguous identity is central to understanding user expectations. It does not ordinarily peer-review an uploaded paper as a journal would. It does not function exactly like an institutional repository either. It aggregates bibliographic metadata, facilitates sharing, hosts full texts in some cases, provides social features, and tracks attention.

This hybrid position creates asymmetrical expectations. Researchers may treat it as an archive because their papers remain there for years. Publishers may treat it as a distribution platform subject to copyright rules. ResearchGate may treat itself as an intermediary governed by terms of service. Readers may treat it as a search engine.

When content disappears, each group interprets the event through a different model.

The copyright wars.

ResearchGate's history includes sustained conflict with major academic publishers over user-uploaded full texts. This context is essential because a mass full-text action can have a copyright explanation even when publication metadata remain visible. Rights to share a preprint, accepted manuscript, or final publisher PDF can differ.

The user's ownership of the underlying ideas does not necessarily confer the right to distribute a publisher-formatted version. Conversely, wholly original preprints may present no third-party rights issue.

A forensic review of my affected corpus should therefore label file version type wherever possible. If the same policy action affected both author-created preprints and publisher PDFs, that pattern would help discriminate among explanations.

Platform rules and due process.

Private platforms are not courts, but procedural fairness still matters. Notice gives the user a chance to understand and correct behavior. Specific reasons reduce arbitrary enforcement. Appeals create a mechanism for error correction. Consistency makes policy predictable.

ResearchGate's Appeals Policy explicitly frames these mechanisms in fairness terms. That public commitment is relevant even if contractual rights remain limited.

The absence of a usable explanation is particularly problematic for a mass action because the user cannot know what behavior to change. If the cause was copyright, which files? If it was upload volume, what threshold? If it was content, which rule? If it was a legal restriction, what can be disclosed?

Silence prevents remediation.

The corporate-record question.

An investigation would also need to determine ResearchGate's internal structure in 2026: which entity operated the service, where user data were stored, where moderation employees worked, which systems held logs, and what retention schedules applied. The visible website brand is not necessarily the legal custodian of every record.

This is why subpoena strategy belongs to lawyers with jurisdictional expertise. The book can identify the questions but should not pretend that serving one office address automatically reaches all data.

Chapter 10: The COVID-Origin War

The origin of SARS-CoV-2 became one of the most politically contaminated scientific questions of the twenty-first century. That contamination did not arise only because the evidence was difficult. It arose because the implications of each hypothesis were enormous.

If the virus emerged through natural zoonotic spillover, the policy lesson emphasizes wildlife interfaces, animal surveillance, markets, agricultural systems, and early outbreak detection. If the virus emerged through a research-related accident, the policy lesson shifts toward laboratory safety, research governance, grant oversight, international transparency, and the wisdom of certain experimental programs. If manipulation contributed, still more difficult questions arise about intent, design, and accountability.

These hypotheses are not morally equivalent in public imagination, and that fact distorted the epistemology from the beginning.

In early 2020, political figures used the possibility of laboratory origin as a geopolitical weapon against China. At the same time, scientists worried that speculation could fuel xenophobia and undermine public health. Those concerns were real. But a hypothesis can be politically abused and still deserve scientific investigation.

The phrase 'lab leak' also concealed multiple hypotheses. A naturally occurring virus collected in the field could infect a laboratory worker. A virus could adapt during passage in cells or animals. A recombinant construct could escape. A research sample could contaminate another process. An infected researcher could transmit before symptoms. These scenarios have different evidentiary signatures.

Natural origin likewise encompasses multiple pathways: direct bat-to-human spillover, intermediate host, market amplification, or an outbreak that began elsewhere and became visible through a market cluster.

A rigorous origin inquiry should therefore compare models rather than slogans.

The problem is that some of the most important evidence was unavailable. Early patient data were incomplete or contested. Access to Chinese laboratory records was limited. Wildlife sampling did not immediately reveal a direct progenitor. Some sequence databases became unavailable. International investigations were constrained by diplomacy. U.S. funding records became politically contested. Private communications among influential scientists later revealed a level of uncertainty and discussion not always visible in contemporaneous public statements.

This does not prove laboratory origin. It proves that the information environment was imperfect.

My own pandemic-era writing often approached that imperfection adversarially. *Shadow Academia: Unveiling the Concealed Narratives of Global Health Crises* described a parallel ecosystem of independent researchers, whistleblowers, citizen investigators, and scientists operating outside conventional publishing channels. It treated groups such as DRASTIC as examples of decentralized inquiry able to pursue questions that mainstream institutions had initially marginalized.

That paper should be read critically. Some claims circulated in alternative networks without sufficient evidence. Some anonymous researchers made errors. Some lines of investigation became entangled with ideological communities. Decentralization protects dissent, but it also weakens quality control.

The same duality exists in traditional academia. Peer review filters some errors while enforcing social norms. Institutional reputation supplies expertise while creating conflicts of interest. Journals slow publication but force formal argument. Social media moves quickly but rewards certainty and outrage.

The origin controversy became a competition between imperfect epistemic systems.

This is where record access becomes decisive. When evidence is open, rival communities can check one another. When evidence is hidden, each community fills the gaps with its priors.

A missing record is therefore not neutral. It increases the influence of ideology.

FOIA became central precisely because it offered outsiders a way to force disclosure from U.S. agencies. Journalists, nonprofit groups, congressional investigators, and private citizens sought emails concerning early origin discussions, grant oversight, EcoHealth Alliance, gain-of-function terminology, and communications among officials and scientists.

The resulting releases sometimes altered public understanding. They showed private uncertainty where public statements had sounded settled. They exposed debates about wording and publication. They also generated misleading interpretations when isolated email fragments were removed from context.

FOIA is not a truth machine. It is a record-access mechanism.

Its value depends on the records being there.

That sentence leads directly to David Morens.

Origin science and narrative science.

The pandemic created two parallel research problems. Origin science asked what biological and epidemiological process produced SARS-CoV-2. Narrative science asked how institutions, media, governments, and online communities came to accept, reject, or stigmatize particular explanations.

The two became entangled because scientific actors were also participants in the narrative. A virologist evaluating origin evidence might have grant relationships, professional friendships, or fears about public misunderstanding. A journalist might use scientific consensus as a heuristic. A politician might select whichever hypothesis harmed an adversary.

This does not invalidate expert opinion. It means that the sociology of expertise becomes part of the historical record.

Consensus as a moving object.

Scientific consensus is often spoken of as though it were a vote taken once. In reality, consensus is a distribution of confidence that evolves as evidence changes. Early in an outbreak, uncertainty should be high. Public communication, however, often demands simple statements. The pressure to simplify can make a provisional judgment look final.

When later records reveal that experts privately entertained possibilities they publicly downplayed, trust can suffer even if their public position was reasonable given the evidence. The problem becomes one of disclosure: how much uncertainty should experts communicate during a crisis?

The cost of stigmatizing hypotheses.

Labeling a hypothesis a conspiracy theory can reduce misinformation when the hypothesis truly depends on unsupported secret coordination. It can also suppress legitimate investigation if the label is applied to a testable causal scenario such as an accidental laboratory infection.

Once a label becomes socially costly, researchers may avoid the topic, grant reviewers may penalize proposals, and journalists may self-censor. This can create an information cascade in which apparent consensus partly reflects reputational incentives.

Again, the opposite risk exists. A hypothesis can become politically rewarded and attract low-quality evidence. The solution is not to make every idea respectable. It is to keep evidentiary standards explicit.

Why missing records amplify polarization.

When primary evidence is unavailable, communities rely on trust. One group trusts public-health institutions; another trusts dissident researchers; another trusts intelligence assessments; another trusts political leaders. Each interprets ambiguity through institutional loyalty.

Open records can move debate away from identity. A grant document can be read by everyone. An email chain can be placed in context. A sequence database can be independently analyzed. Transparency does not guarantee agreement, but it narrows the range of plausible narratives.

The Morens case is therefore not merely a scandal about email etiquette. It is evidence that one mechanism for reducing polarization - access to authentic government records - was intentionally frustrated.

Chapter 11: FOIA

The Freedom of Information Act is often described as a transparency law, but that description understates its epistemic role. FOIA is one of the instruments through which citizens can test the public story of government against the government's own records.

The statute does not guarantee that every requested document will be produced. Agencies may withhold information under exemptions protecting national security, personal privacy, law-enforcement interests, trade secrets, privileged deliberations, and other protected categories. Requests can be broad, burdensome, delayed, disputed, litigated, redacted, or denied. Agencies can make mistakes. Requesters can make unreasonable demands.

But FOIA depends on a prior condition so basic that it is easy to overlook: the record must exist in a retrievable system.

If an official intentionally conducts public business on a private account and then deletes the messages to avoid retrieval, the dispute is no longer merely over which FOIA exemption applies. The information architecture itself has been manipulated.

That is why federal records law matters alongside FOIA. Government employees do not own federal records in the same way they own private correspondence. Records documenting official activity belong to an institutional system designed to preserve accountability, continuity, and history.

The Morens case sits at the intersection of those regimes.

Before the criminal case, congressional investigators had already focused on Morens's private-email practices. A June 2023 congressional letter stated that documents suggested he may have used personal email to avoid transparency and FOIA and may have deleted federal records. A 2024 subpoena sought records from his personal accounts. Congressional hearing materials later described tens of thousands of pages produced in response to investigative demands.

At that stage, the allegations were politically contested. Congressional investigations are not criminal convictions. Partisan committees can overstate evidence, and witnesses can dispute interpretations.

The April 2026 indictment changed the status of the matter but still did not prove guilt. DOJ alleged that Morens conspired with others to evade FOIA and records requirements, deliberately used personal Gmail for official business, and concealed or destroyed records.

The August guilty plea changed it again.

> **ADMITTED OR ADJUDICATED:** According to the U.S. Attorney's Office for the District of Maryland, Morens pleaded guilty on August 18, 2026 to conspiracy to commit offenses and to defraud the United States. DOJ said the conspiracy involved evading FOIA and the Federal Records Act in connection with communications concerning coronavirus-research grants.

The legal significance belongs to the courts. The historical significance is immediate.

A public controversy that depended on email disclosure now includes an admitted scheme to keep certain emails from disclosure.

This creates a problem of unknown missingness. In statistics, missing data can sometimes be treated as random. If observations disappear for reasons unrelated to the variable being studied, standard techniques may estimate around the gaps. If observations are missing specifically because of their content, the bias becomes much more difficult.

Record concealment designed to defeat FOIA is the documentary equivalent of data missing not at random.

The communications most likely to be hidden are precisely those the participants believe would be requested or damaging. That does not mean every hidden message proves wrongdoing. It means the surviving official mailbox is not necessarily an unbiased sample of the underlying activity.

For historians, this is devastating. A chronology reconstructed only from government accounts may systematically omit the conversations participants intentionally moved elsewhere.

For journalists, it changes source strategy. One must search personal accounts, recipient accounts, attachments, drafts, calendar records, and third-party systems.

For prosecutors, it changes intent analysis. The language used to discuss avoiding disclosure can become evidence of consciousness of legal obligations.

For scientists, it should be a warning. Public trust in research oversight depends not only on whether the underlying science was sound but on whether the administrative record was honest.

The Morens case also demonstrates that email is not just speech. In bureaucratic systems, email performs governance. Officials send drafts, transmit non-public information, influence funding decisions, coordinate responses, and brief superiors. Moving such work off-system can change what future investigators see.

This is why the book resists a narrow 'gotcha' reading of FOIA. The issue is not whether embarrassing emails should be exposed for sport. The issue is whether the public record is allowed to be selectively engineered by the people whose conduct the record may later be used to evaluate.

The same conceptual question applies outside government, even where FOIA does not.

ResearchGate is a private platform. Its internal moderation records are not ordinarily public records. A user cannot simply file FOIA for its audit logs. But the evidentiary structure is parallel: a private company controls records that could explain a public event. If those records are preserved, lawful process may answer the question. If they are routinely deleted, the historical explanation can become irretrievable.

Government transparency law and private-platform preservation therefore meet in the same place: retention.

FOIA as an adversarial replication mechanism.

In science, replication allows an independent group to test a published result. In public administration, FOIA performs a rough analogue by allowing outsiders to test the official description of events against internal records. The analogy is imperfect but useful. A press release is the published result; the underlying email, draft, calendar entry, and memorandum are the raw data.

This is why officials who regard FOIA merely as an annoyance misunderstand its function. Requests can certainly be abusive or burdensome, but the law exists because self-report by institutions is insufficient for democratic accountability.

Search architecture matters.

Even when employees comply in good faith, FOIA searches are constrained by technical systems. Agencies must identify custodians, date ranges, search terms, repositories, and record types. A poorly designed search can miss relevant material without anyone intentionally concealing it.

Personal-account use makes this harder because the agency may not possess the mailbox. Retrieval may depend on the employee preserving and forwarding records. When the employee's intent is to avoid disclosure, that weakness becomes exploitable.

Investigators can compensate by searching recipient accounts. If Morens wrote from Gmail to an NIH employee's official address, the recipient copy may survive in federal systems. Attachments may exist in document-management tools. Calendar invitations may reveal meetings even when messages are gone. This is why records investigations are network problems rather than single-mailbox problems.

Redaction is not disappearance.

Public debates often treat a heavily redacted page as equivalent to a deleted page. The distinction matters. A redacted record at least establishes that a record exists, has a date, a sender, a recipient, and a legal basis asserted for withholding portions. That structure can be challenged in litigation.

A deliberately uncreated or destroyed record is harder because there may be no object to contest.

FOIA and scientific privilege.

Scientists sometimes resist disclosure because preliminary discussion can be messy. They worry that opponents will cherry-pick abandoned hypotheses, jokes, or uncertainty. That concern is legitimate. But the solution cannot be to conduct official business outside record systems. Better solutions include context, appropriate legal exemptions, and careful public interpretation.

Transparency without context can mislead. Secrecy without accountability can corrupt. Governance must manage both risks.

Why the Morens facts should not be generalized beyond their scope.

A guilty plea by one official does not prove that NIAID as a whole systematically destroyed records. It does not prove that every relevant email is missing. It does not prove that senior officials joined the scheme. It does not prove a particular COVID-origin hypothesis.

The plea proves what the plea establishes. Its broader significance is that record integrity cannot be assumed merely because official systems exist.

Chapter 12: David Morens

David M. Morens was not a household name during most of the pandemic. He was, however, positioned close to the center of NIAID's leadership structure. DOJ describes him as a senior adviser in the Office of the Director from 2006 through 2022. His role included advising senior staff on policy and infectious-disease issues and helping gather and interpret information relevant to NIAID leadership.

That position matters because advisers can serve as connective tissue. They may not sign every grant or make every final decision, but they communicate across scientific, administrative, and external networks. They can understand what senior officials need, translate technical questions, maintain relationships with grantees, and help shape responses.

The criminal case did not arise from a claim that Morens performed a dangerous laboratory experiment. It arose from records and influence.

DOJ's April 28, 2026 indictment alleged that after NIH terminated a bat-coronavirus grant, Morens and others worked to help restore the funding and counter the laboratory-origin narrative. In anticipation that communications would be sought through FOIA, the indictment alleged, they agreed to use Morens's personal Gmail account rather than his official NIH account.

The indictment further alleged that the private account was used to exchange non-public NIH information, discuss efforts to influence NIH, edit drafts of letters to NIH leadership, and back-channel information to a senior NIAID official. DOJ also alleged an illegal-gratuities component involving gifts and a proposed scientific commentary.

At indictment, all of that remained allegation. The presumption of innocence applied.

On August 18, Morens entered a guilty plea to the conspiracy charge. DOJ's account of the plea repeats key elements of the concealment scheme and says he admitted participation.

A guilty plea is not the same as a full trial record. Some counts may not be adjudicated in the same way. Plea agreements can narrow issues. Sentencing had not yet occurred when this book was written. The exact identities, liability, and conduct of unindicted people should not be assumed beyond the public record.

Still, the plea is a watershed because intent is central.

Many government-record controversies can be explained by carelessness: an employee uses a personal phone, an auto-delete rule operates, records are misplaced, retention settings are confusing, or a message is accidentally omitted from a search. Here, DOJ's description of the admitted conspiracy centers on intentional concealment in anticipation of FOIA.

That intent converts a records-management problem into a public-integrity problem.

One of the most revealing features of such cases is the language people use when they believe they are communicating privately. Official memoranda are written for institutional audiences. Private messages can reveal motives, frustrations, jokes, contempt, fear, loyalty, and tactical thinking. For that reason, they are also easy to overinterpret. Sarcasm can look like confession. Bravado can look like policy. A joke can look like instruction.

Context is therefore essential.

The correct method is to place each message in a chain: what came before, who received it, what action followed, and whether independent records corroborate its meaning.

This is where a records-concealment case becomes especially difficult. If the chain itself has missing links, the most dramatic surviving sentence may be misleading in either direction.

The Morens case has another implication for the broader COVID-origin debate. It shows that reputation management and scientific communication were not completely separate from grant administration. DOJ says the conspirators sought both to address funding and to counter the lab-leak narrative. That overlap does not prove that the natural-origin hypothesis was false. It does show that institutional interests and public narrative were being discussed within the same network.

Scientists often imagine that scientific truth floats above institutional incentives. In reality, scientists work inside institutions with grants, careers, reputations, partnerships, and political exposure. Those interests do not automatically corrupt science, but they create reasons to preserve records precisely so outsiders can distinguish legitimate scientific judgment from institutional self-protection.

The central question is not whether scientists have interests. Everyone does. It is whether the evidentiary system allows those interests to be audited.

Morens's conduct, as admitted, interfered with that audit.

This book therefore treats him neither as a symbol of universal corruption nor as an isolated eccentric. He is a concrete example of how a small number of choices by a well-placed official can affect the discoverability of an entire controversy.

The lesson is architectural: systems of accountability must be designed on the assumption that some participants will sometimes prefer opacity.

The adviser role and plausible deniability.

Organizations often rely on advisers precisely because they can work across formal boundaries. They gather information, maintain relationships, and help senior leaders understand complex domains. This flexibility is valuable. It can also create ambiguity about authority.

When an adviser communicates with an outside grantee, is he speaking personally, scientifically, or institutionally? When he relays information to a senior official, is he making a formal recommendation or simply keeping the official informed? Those distinctions matter in legal analysis.

The Morens case demonstrates why role descriptions, performance expectations, and actual communication patterns should be examined together. DOJ states that the matters conducted through the private account fell within his senior-adviser role and constituted federal records. That characterization defeats the idea that the messages were merely private scientific gossip.

Back channels are not inherently unlawful.

Every institution has informal communication. A phone call, personal relationship, or off-site conversation can speed coordination. The legal problem arises when official records obligations are intentionally defeated or when the informal channel is used for unlawful purposes.

The phrase 'back channel' can therefore sound more sinister than it is. The relevant questions are what information traveled, why the channel was chosen, and whether the communication was required to be preserved.

Scientific commentary and institutional influence.

DOJ's indictment alleged an unusual overlap between gifts, official assistance, and a proposed scientific commentary concerning natural origin. The guilty-plea release also describes an illegal-gratuities conspiracy. The details should be taken from the court record rather than embellished by partisan summaries.

The episode nonetheless highlights a general problem: scientific articles can have policy and reputational value beyond their scientific content. A commentary supporting a particular origin narrative can influence journalists, grant decisions, and institutional reputation. That makes authorship and conflicts of interest relevant.

What Morens does not tell us.

Morens cannot answer whether ResearchGate's action was routine. He cannot answer what happened inside a Berlin-based platform. The fact that his guilty plea occurred a month after my full texts disappeared creates chronology, not nexus.

The temptation to make him the bridge between every record controversy should be resisted. If a bridge exists, it must be documentary.

Chapter 13: The Guilty Plea

August 18, 2026 provides the fixed point around which this manuscript was assembled.

The U.S. Attorney's Office for the District of Maryland announced that David Morens, age seventy-eight, of Chester, Maryland, pleaded guilty in federal court to a conspiracy charge stemming from a scheme to evade FOIA and the Federal Records Act in connection with communications about coronavirus-research grants. He had served as a senior adviser within NIAID's Office of the Director from 2006 to 2022.

According to DOJ, the admitted conspiracy developed after NIH terminated the grant *Understanding the Risk of Bat Coronavirus Emergence*. DOJ's account says that Morens and other participants worked on efforts relating to the terminated grant and to the public narrative surrounding a possible laboratory origin of COVID-19.

The conspirators, DOJ says, agreed in writing to hide their communications from public view by using Morens's personal Gmail account instead of his NIH account because they anticipated FOIA requests.

That sentence is the center of gravity.

It is possible to debate gain-of-function definitions. It is possible to debate whether a grant met a policy threshold. It is possible to debate whether SARS-CoV-2 emerged naturally or through research. It is possible to debate the meaning of ambiguous emails.

It is no longer speculative, however, that Morens participated in an admitted plan to keep some relevant official communications out of ordinary government record systems.

DOJ also stated that the conspirators used the private account to exchange non-public NIH information and to back-channel information to a senior NIAID official. The public release does not license this book to assign criminal liability to that senior official. Receiving information, being discussed, or being the intended recipient of a back channel is not the same as joining a conspiracy.

> **DOCUMENTED LIMIT:** Anthony Fauci has not been charged with wrongdoing in the Morens case. This book does not state otherwise.

Why, then, does the guilty plea alter the meaning of older arguments about hidden records?

Because evidence has directionality. Before the plea, a critic alleging deliberate record concealment had to overcome the possibility that suspicious email practices were being misread. After the plea, deliberate concealment by Morens is an admitted fact. That does not validate every criticism of NIAID. It validates a narrower concern: investigators were right to take the integrity of the record seriously.

The same evidentiary humility should govern the ResearchGate problem.

I know what I observed: thousands of public full texts became unavailable. I know what I did not receive: a warning, a strike sequence, or a substantive explanation. I know that the corpus included extensive COVID-origin and NIAID criticism. I know that ResearchGate had prominent biomedical and technology investors and medically trained founders. I know that one investor, Bill Gates, was part of a long-standing global-health ecosystem that included NIAID and Fauci.

None of those facts establishes why ResearchGate acted.

The Morens plea teaches exactly why one should not fill that gap with imagination. If the truth is hidden in communications, the proper response is to obtain the communications.

The criminal case also identifies who had institutional interest in records: the FBI and HHS Office of Inspector General investigated; Assistant U.S. Attorneys Joseph R. Baldwin and Bijon A. Mostoufi were identified by DOJ as prosecutors. For a person considering whether his own archive contains potentially relevant evidence, this provides a concrete rather than abstract audience.

What would such investigators care about?

Not that I dislike a platform decision. Not that a prominent investor has relationships in global health. Not that a founder studied viruses.

They would care about contemporaneous records, provenance, intent, and nexus.

A 2002 email may matter if it establishes that specific biodefense concerns were communicated to relevant officials. A 2022 paper may matter if it proves that the email's archival location was publicly identified years before the current prosecution. A July 2026 platform event may matter if internal records show that it was triggered by someone connected to an investigation, a government request, a named participant, or subject-specific pressure. If the internal records instead show routine platform enforcement unrelated to COVID, the event becomes legally uninteresting to that investigation.

That is falsifiability.

A good evidentiary hypothesis contains the conditions under which it dies.

The thesis 'ResearchGate suppressed my papers because of Fauci' is presently too strong because it cannot be supported by the public record. The hypothesis 'ResearchGate possesses internal records that can establish whether any external or subject-specific pressure contributed to an unexplained mass visibility event' is both modest and testable.

The guilty plea is therefore not the conclusion of this book. It is the methodological turning point.

Why a plea is different from leaked email interpretation.

For years, public understanding of the Morens controversy depended on fragments: congressional letters, released messages, testimony, and competing partisan interpretations. A guilty plea changes the evidentiary status because the defendant accepts criminal responsibility under a formal judicial process.

That does not make every allegation in an indictment true. Plea documents and the count of conviction define what has been admitted. A careful book should therefore cite the plea release and, where available, the plea agreement or statement of facts rather than treating the original indictment as entirely adjudicated.

Sentencing remains future.

At the time this manuscript was completed, sentencing was scheduled for November 12, 2026. The court had not yet imposed sentence. That date matters because future editions should update the procedural history rather than freezing the case at the plea stage.

The public-interest significance of intent.

The most consequential admitted feature is not that an employee used Gmail. Millions of people mix work and personal systems. It is the alleged and admitted purpose of hiding communications from public view in anticipation of FOIA. Purpose turns an otherwise mundane technical choice into evidence of obstruction of accountability.

This distinction is useful beyond government. If ResearchGate made a routine policy decision and failed to explain it, that is one category of problem. If internal records showed a deliberate decision to conceal the reason because disclosure would expose improper pressure, that would be another. Intent changes the meaning of the same outward event.

The prosecutor's lens.

A prosecutor asks different questions from an author. What statute applies? What is the evidence of agreement? What records authenticate intent? What witness can explain the system? Which overt acts occurred? What jurisdiction exists? Is the evidence admissible? Can alternative explanations be excluded beyond a reasonable doubt?

My ResearchGate concerns are nowhere near that evidentiary threshold. They are leads. The correct way to present them to investigators is concise: here is the event, here is the preserved archive, here are the unanswered record questions, and here is why the timing may justify checking for a nexus.

Overstatement would make the material less useful.

Chapter 14: The Missing Record Problem

Every archive contains gaps. The problem becomes dangerous when the pattern of the gaps is unknown.

Consider a laboratory notebook with pages 1 through 100 except pages 47 through 52. Perhaps the pages were blank. Perhaps a solvent spill destroyed them. Perhaps an investigator removed them because they contained failed experiments. Perhaps they contained the only positive result. Until the reason for the gap is known, the absence becomes part of the evidence.

Digital archives create more subtle gaps because deletion is not visually obvious. A mailbox can look complete while selected messages have been removed. A cloud system can retain an attachment after the message is gone. A platform can preserve a publication page while detaching the file. A backup can hold material that the live database no longer exposes. Search indexes can preserve snippets of content that no longer exists at the source.

The result is an archaeological landscape of partial persistence.

The COVID-origin controversy contains many examples of this problem in principle: inaccessible laboratory databases, contested sequence availability, private email use, deleted messages, withheld or redacted government records, and records held by institutions outside U.S. jurisdiction.

Some missingness is lawful and justified. Patient privacy matters. National-security information may require protection. Proprietary research can have legitimate confidentiality. Diplomacy sometimes depends on private communication. No serious transparency regime publishes everything.

The issue is accountable missingness.

An accountable system can say: this category exists, this many records were withheld, this legal basis applies, this retention schedule governs, and this review process is available. An unaccountable system simply produces silence.

Silence invites narrative substitution.

People are pattern-seeking organisms. When a record is missing, they do not leave the space empty. They fill it with whatever larger story already feels plausible. To one audience, the missing record proves conspiracy. To another, concern about the missing record proves paranoia.

Both reactions are epistemic shortcuts.

The correct response is a missingness inventory.

For a government investigation, that means identifying expected record classes: official email, personal email used for official business, text messages, attachments, calendars, draft documents, grant-management systems, meeting notes, agency chat platforms, external-recipient mailboxes, and backups. Investigators then compare what should exist with what can be recovered.

For ResearchGate, the same logic applies at the platform level. Expected records might include account audit logs, content-object histories, moderation tickets, policy reason codes, automated classifier outputs, copyright complaints, external legal requests, support correspondence, employee messages, database change logs, backup snapshots, and notice-delivery logs.

> **OPEN QUESTION:** Which of these records ResearchGate actually retains, for how long, and under which corporate entity is presently unknown to me.

That uncertainty is why preservation is time-sensitive. Logs often have shorter retention periods than user content. An email may live for years; an application event log may live for months. A backup may rotate. A support ticket may persist while ephemeral chat disappears.

Anyone seeking legal review should therefore distinguish a preservation request from a demand for disclosure. Preservation says, in effect: do not destroy potentially relevant records while lawful process is considered. Production asks the company to provide them under an appropriate authority.

This is not legal advice, and the correct mechanism depends on jurisdiction, procedural posture, and who is requesting the records. The conceptual point is that the order matters. A perfect subpoena served after routine retention periods have expired can produce nothing.

The missing-record problem also changes how authors should manage their own evidence.

Original files should be preserved in read-only copies. Hashes can be computed. Email should be exported with full headers rather than printed to PDF and treated as original. Screenshots should be accompanied by URLs and timestamps. Web pages should be archived through independent services. A spreadsheet should index title, DOI, file name, size, creation date, and cryptographic digest. Copies should exist in more than one physical and organizational location.

Preservation is not only for institutions one distrusts. It is protection against one's own future memory.

The longer a controversy lasts, the more personal recollection hardens into narrative. A preserved file can correct the author too.

Deletion, concealment, and non-creation.

Three forms of missingness should be separated. Deletion removes an existing record. Concealment leaves the record somewhere but defeats expected discovery. Non-creation means official activity occurs in a way that never generates the required record in the proper system.

The Morens conduct described by DOJ implicates more than one category because official business conducted on personal Gmail can both evade creation in the official system and facilitate later deletion.

ResearchGate's event may involve none of these in the same sense. A visibility change can leave every byte intact while making the public experience one of disappearance. The language should follow the technical evidence.

Data provenance as a graph.

A useful model is to treat every important record as a node in a graph. Edges connect copies, references, replies, attachments, citations, and derived documents. A deleted node may still be reconstructed from neighbors.

An email from A to B may survive in B's mailbox. An attached proposal may survive in a grant repository. A webpage may survive in Wayback. A paper may survive in a reader's download directory. A DOI may preserve metadata after the file disappears.

This graph approach turns archival recovery into a systematic search for redundancy.

Unknown unknowns.

The hardest records are those whose existence is not known. A FOIA requester can ask for emails containing known terms, but a private chat using code words may evade keyword searches. An undisclosed platform moderation category may not be visible externally. An unknown outside complainant cannot be searched by name.

Investigators deal with this by requesting structured record categories, not only keyword hits. For example, all moderation events affecting an account during a date range may reveal a reason code that suggests new search terms.

The investigation should therefore proceed iteratively: obtain metadata, learn the system, refine the request.

Preserving absence.

Even an unsuccessful search should be documented. If an archive search finds no ResearchGate notice in a mailbox, preserve the search terms, account, date range, and method. A negative result without methodology is weak. A reproducible negative result is evidence, though never absolute proof that no message existed.

This discipline prevents the author from turning 'I did not find it' into 'it never existed.'

Chapter 15: The Private Archive Problem

FOIA creates a democratic expectation that government records can, within limits, be requested. No equivalent general law requires a private scholarly platform to give the public its internal moderation history.

This creates a paradox. Modern scientific memory increasingly depends on privately governed infrastructure: cloud drives, journal platforms, preprint servers, social networks, bibliographic databases, code hosts, domain registrars, search engines, and identity systems. These companies can become essential to public knowledge without becoming public institutions.

ResearchGate illustrates the paradox unusually well.

Its mission has repeatedly been framed in terms of connecting scientists and making research accessible. Its scale means that a ResearchGate profile can function as a researcher's public bibliography. Its metrics can influence perceptions of impact. Its pages are indexed by search engines. Its full texts can be the copy readers actually use when institutional subscriptions fail.

Yet access remains contractual and technical, not constitutional.

A private platform can change product features. It can enforce copyrights. It can remove spam. It can respond to lawful orders. It can redesign a metric. It can close an account under its terms. It can exit a market. Unless a specific law or contract says otherwise, the public's reliance on the platform does not create a perpetual archival guarantee.

That is not an indictment of ResearchGate. It is a structural weakness in contemporary scholarly communication.

The same weakness exists elsewhere. A social-media company can erase a decade of posts. A cloud-storage provider can terminate an account after a billing failure. A preprint server can alter screening rules. A journal can change access policies. A domain can expire. A hosting company can disappear.

The historical record has become distributed across organizations whose primary responsibility is not historical preservation.

Traditional libraries solved this problem through ownership. They acquired physical copies that could remain even if a publisher ceased business. Digital licensing often replaces ownership with access. A library may subscribe to a database without possessing the underlying files. Researchers may link to a platform rather than store a version they are legally permitted to preserve.

The result is fragile memory.

My July 2026 experience is an extreme example because of scale, but it is not conceptually unique. Thousands of papers were publicly reachable through one interface and then were not. The survival of copies elsewhere prevented the event from becoming an archival catastrophe.

That survival should not be accidental.

Scientific culture needs a preservation norm distinct from a publication norm. Publication asks whether a work should be disseminated. Preservation asks whether future readers should be able to reconstruct what was disseminated at a particular time. Even material later retracted, corrected, rejected, or criticized can have archival value.

This is especially true for controversial periods. Propaganda, error, preliminary hypotheses, and failed arguments may be historically important precisely because they show what people believed.

A mature system can preserve without endorsing.

Internet Archive's Wayback Machine embodies this principle for the web. A capture is not a seal of truth. It is a record that a page appeared in a particular form. Likewise, an archive of ResearchGate PDFs would not certify the scientific correctness of each paper. It would preserve the corpus for audit.

There is a legal tension here. Copyright owners have rights. Privacy law imposes limits. Defamatory or unlawful material raises separate problems. Security-sensitive information may need restriction. Preservation cannot mean indiscriminate public replication.

But those complexities strengthen the case for professionally designed archives rather than weakening it.

The private archive problem is ultimately a governance problem: who decides what future historians can see?

When the answer is 'whatever happens to remain on commercial servers,' we have no archival policy at all.

The Digital Services Act context.

European platform regulation increasingly imposes transparency and procedural obligations on online intermediaries. ResearchGate's own legal-help materials reference appeal rights associated with the European Union's Digital Services Act for covered users. The exact application of those provisions to a U.S.-based user and a particular 2026 content action is a legal question, not something this book resolves.

The broader point is that platform governance is moving toward more formalized notice, explanation, and appeal systems. That trend recognizes a basic reality: private content decisions can have public consequences.

Scholarly platforms are unusually sensitive.

A social-media post may be ephemeral by design. A scientific paper carries different expectations. It may be cited years later. It may be part of a priority dispute. It may contain data needed for replication. It may become evidence in litigation or policy review.

This does not mean a platform must host unlawful or infringing material forever. It means that removal procedures should consider scholarly provenance. At minimum, metadata and reason codes should survive even when public access to the file does not.

The case for exportability.

A platform can reduce archival dependence by allowing users to export a complete machine-readable record: publication metadata, file identifiers, upload dates, statistics, DOI links, and visibility status. Such an export would not force the platform to keep hosting content, but it would allow researchers to maintain continuity elsewhere.

Interoperability is therefore an archival safeguard.

The difference between moderation and erasure.

Moderation addresses present access. Erasure affects future recoverability. A platform can remove a harmful or infringing file from public view while retaining an internal evidentiary record of the decision. This is analogous to a library withdrawing a book from circulation while preserving catalog history.

The public debate often collapses these functions. Critics demand permanent availability; platforms respond with removal. A better architecture would separate access control, evidence retention, and metadata preservation.

Private power and public dependency.

The deeper issue is dependency. When millions of researchers build professional identities on a private platform, the platform's decisions shape the public record even if users voluntarily accepted the terms. This is not unique to ResearchGate. It is the general problem of privately owned public infrastructure.

The solution is not necessarily regulation alone. It is redundancy, open standards, institutional repositories, public archives, and user-controlled exports. No single company should be the only place where a scientific record can be reconstructed.

Chapter 16: 5,700 Missing PDFs

The number 5,700 is imprecise in the way many large personal archives are imprecise. A researcher can count file objects, publication pages, duplicates, versions, supplements, and platform records differently. The proper forensic project is therefore not to defend the number rhetorically but to turn it into a reproducible inventory.

> **AUTHOR'S ACCOUNT:** On the weekend of July 18-19, 2026, approximately 5,700 ResearchGate-hosted full-text PDFs associated with my account ceased to be publicly available. The corpus had accumulated roughly two million reads. I received no advance warning, no known sequence of strikes, and no substantive explanation despite later outreach.

A prosecutor or civil investigator would immediately decompose that statement.

Which account identifier?

How many publication pages existed before the event?

How many had public full texts?

How many publication pages remained afterward?

Did the PDFs become private, become detached, become blocked, or disappear from storage?

Were DOI-bearing uploads treated differently from conventional journal articles?

Were files uploaded by coauthors treated differently from files uploaded by me?

Were all subject categories affected at the same time?

Did any subset disappear before the main batch?

Were there copyright notices for any subset?

Did read counts remain?

Did Google or Bing caches continue to resolve the full-text links?

Did the platform's API or page source expose hidden file identifiers after public links vanished?

These questions can convert a story into a dataset.

The first reconstruction should begin with the surviving Google Drive archive. File metadata show that many copies were ingested into Drive around July 18, 2026, while their internal or source modification dates are earlier. That timing may be important because it suggests an archival consolidation occurring around the same weekend. The exact reason for those Drive creation timestamps should be documented rather than assumed.

The archive contains a *Network Map of Keywords and Titles to ~6,000 Youvan AI-Assisted Preprints on ResearchGate*. Such index documents are especially valuable because they can serve as snapshots of the platform corpus near the period of disappearance.

The second reconstruction should use surviving ResearchGate publication pages. Even when full texts are gone, a landing page may retain title, author, date, DOI, abstract, citation count, and historical reads. Those fields allow matching to local copies.

The third layer is external indexing. Search engines, DOI registries, scholarly aggregators, and the Internet Archive may preserve metadata or snippets. A web cache should not be treated as proof of file storage, but it can establish that a public page or link existed.

The fourth layer is correspondence. Every email to or from ResearchGate around the event should be exported with headers. The absence of warning can itself be documented by searching the mailbox for ResearchGate

sender domains and policy terms. If ResearchGate later asserts that notices were sent, message-delivery logs could be compared with the mailbox evidence.

The fifth layer is screenshots. Screenshots are weak when isolated but powerful when they show account-level counts, error messages, or visibility changes tied to a date and URL. Their value increases if original image metadata survive.

Only after this reconstruction should one classify the event by subject matter.

This step is critical. If essentially every full text across mathematics, theology, computer science, biology, political writing, and pandemic material changed state in the same batch, the most parsimonious explanation is account-level action. Subject-matter motive would require additional evidence.

If, however, a COVID-origin subset had distinct treatment - earlier removal, different notices, special flags, or separate escalation - that would justify more targeted inquiry.

The book already knows that the archive contained a significant pandemic cluster. Search of the surviving Drive turns up titles directly concerning Fauci, global health, gain-of-function research, COVID origins, media narratives, and scientific transparency. It also contains papers on the historical lineage of viral manipulation and on the relationship between scientific institutions and public narratives.

A cluster is not evidence of targeting. It is a reason to tag the dataset.

This is where machine-readable analysis becomes useful. Titles and abstracts can be classified by topic. Removal timing can be compared across categories. If one category has a statistically distinct pattern, investigators gain a lead. If not, an attractive theory can be discarded.

Evidence should be allowed to disappoint us.

The same applies to the two-million-read figure. If historical records show a different number, the manuscript should be corrected. If the metric includes large automated components, that should be stated. If only a subset of the reads belonged to the affected files, that should be separated.

A documentary book earns trust by showing where the author's own memory can be tested.

The surviving corpus provides that opportunity.

Turning the corpus into a forensic table.

The ideal reconstruction has one row per ResearchGate publication record and columns for every independently observable attribute. At minimum: title, authors, ResearchGate URL, ResearchGate publication identifier, DOI, publication type, original publication date, date added to ResearchGate, full-text file name, file size, local hash, Drive location, last known public date, first known unavailable date, last observed read count, subject category, and notes about copyright status.

A second table should record evidence objects: screenshots, emails, Wayback captures, Google search snippets, platform notifications, exported PDFs, and local browser history. Each evidence object can then point to one or more publication rows.

This database would allow a prosecutor or civil lawyer to understand the event without navigating thousands of files manually.

Sampling before full reconstruction.

Because 5,700 records are labor-intensive, the first pass can use a stratified sample. Choose papers from five or six subject categories and multiple years. Include heavily read and lightly read papers, journal PDFs and self-authored preprints, DOI and non-DOI items, COVID and non-COVID material. Determine exactly what state each item entered.

If the sample reveals a uniform pattern, the full project can be designed around that mechanism. If it reveals heterogeneity, the heterogeneity itself becomes the lead.

The difference between account disappearance and corpus disappearance.

It is also possible for a platform account to remain active while full-text privileges change. The user may still have a profile, followers, citations, and titles. That differs from account deletion. The distinction affects what records a platform might preserve and what readers experienced.

This is why language such as 'ResearchGate deleted me' can be rhetorically understandable but technically imprecise. The more precise description is stronger: publication metadata remained while thousands of public full texts became inaccessible, if that is what the evidence shows.

The role of the July archive transfer.

Many Drive copies show creation or ingestion timestamps around July 18, 2026. That coincidence should be documented carefully. If I was actively consolidating the corpus because I recognized a threat to access, the transfer itself becomes a contemporaneous response to the event. If the transfer was automated or unrelated, that is different.

The archive should preserve not only the files but the process by which they arrived. Batch manifests, transfer logs, and folder dates may help establish that the corpus existed in substantially the same form at that time.

What would count as a smoking gun?

The phrase is overused, but the standard can be defined. A truly dispositive record would be something like an internal ResearchGate message naming the account, identifying an outside requester or political subject, and directing a bulk action for that reason. Short of that, evidence will likely be cumulative: timing, reason codes, ticket narratives, unusual escalation, and external communications.

Conversely, a bulk-operation record explicitly tied to a general platform policy with no relevant external contact would strongly support an ordinary explanation.

The purpose of reconstruction is to make either outcome possible.

Chapter 17: What Compulsory Process Could Answer

There is a limit to open-source investigation. A public web search can tell us what ResearchGate says its policies are. It can identify founders, investors, corporate entities, and press releases. It can show which publication pages remain visible. It cannot reveal a private moderation ticket, an internal Slack thread, a legal request, or the database event that changed thousands of file states.

That is the domain of compulsory process.

The precise legal instrument would depend on who is investigating, the legal theory, the location of the records, and the entity with custody or control. A federal grand-jury subpoena, administrative subpoena, search warrant, civil discovery request, preservation demand, mutual legal-assistance process, or German/EU legal mechanism are not interchangeable. This book does not pretend to choose among them.

It can, however, specify the evidence categories.

The first category is **account audit history**. A well-designed service records administrative changes to a user account: restriction, suspension, bulk visibility change, policy flag, or deletion. Investigators would want timestamps, actor identifiers, reason codes, and linked tickets.

The second is **content-object history**. Each publication or full-text file may have its own state transitions. Those records can show whether thousands of objects were modified by one batch job or through separate actions.

The third is **moderation and Trust & Safety records**. These may include user reports, internal reviews, policy categorizations, escalation notes, automated signals, and human decisions.

The fourth is **copyright enforcement**. Investigators would want notices from publishers or rights holders, automated match records, takedown demands, correspondence concerning ownership, and whether any complaint triggered an account-wide action.

The fifth is **external legal or governmental requests**. This category should be framed neutrally and broadly enough to test the hypothesis rather than assume it: communications from government agencies, courts, law-enforcement entities, contractors, or other official intermediaries concerning the account or specified publications.

The sixth is **executive and board escalation**. If a high-profile account was discussed above routine moderation channels, records may exist in email, messaging systems, board materials, risk reports, or legal summaries.

The seventh is **notice-generation and delivery logs**. ResearchGate's policies contemplate notifying users about certain actions. If the platform believes notice was generated, its systems may show the recipient address, template, timestamp, delivery status, bounce information, or suppression reason.

The eighth is **backups and retention records**. If the public database no longer contains the relevant state, snapshots may show what existed immediately before and after the July event.

The ninth is **employee access logs**. For a narrow period, records may identify which employees accessed the account or moderation ticket. Such records should be sought only where legally justified; they can implicate employee privacy and security.

The tenth is **communications containing defined search terms**. Search terms might include my name, account identifiers, relevant DOIs, particular publication titles, and subject terms such as NIAID, NIH, Fauci, Morens, EcoHealth, Daszak, Baric, COVID origins, laboratory origin, gain of function, FOIA, misinformation, or legal request.

The purpose of such terms is not to imply that every named person contacted ResearchGate. It is to ensure that if such a communication exists, it is discoverable.

The most valuable result could be negative.

Imagine the production shows a routine copyright event: an automated or publisher-driven process identified a large class of files, a bulk action was triggered, no government or investor communication existed, and notice failed because of a technical problem. That would substantially close the political branch of the inquiry.

Imagine instead that the records show an account-integrity rule unrelated to content: upload volume or automation exceeded a platform threshold. Again, much speculation would die.

Imagine a mistaken bulk action. The story becomes one of platform governance and accountability rather than censorship.

Only if records show external pressure, subject-specific escalation, or deliberate concealment of the reason would a more serious hypothesis gain support.

A subpoena is therefore not an accusation. Properly used, it is a mechanism for replacing imagination with records.

This is exactly the lesson of the Morens case. Investigators did not have to infer the content of a private Gmail account forever. They obtained records, reconstructed communications, and brought a case that ultimately produced a guilty plea.

The parallel is methodological, not factual.

Preservation before production.

Digital investigations are races against retention schedules. A preservation request can be narrowly tailored to prevent routine destruction without requiring immediate disclosure. The categories should include logs and ephemeral messaging precisely because those may disappear sooner than user content.

The request should identify the account unambiguously through user ID, registered email addresses, profile URLs, and known publication IDs. Names alone are risky because they can be misspelled or duplicated.

Date windows.

A useful preservation window would not begin on the day the files vanished. It should extend far enough backward to capture planning or escalation and far enough forward to capture internal discussion after user complaints. April through August 2026 is an obvious initial interval because it includes the Morens indictment, the July ResearchGate event, subsequent outreach, and the August guilty plea. A broader window may be justified if earlier moderation occurred.

The date window should not imply that the Morens indictment caused anything. It is simply a defensible period for reconstructing the platform event in context.

External-request logs.

Large platforms often maintain processes for legal demands, government requests, intellectual-property complaints, and law-enforcement contacts. Even if the substantive request is privileged or restricted, metadata may show that a request occurred.

Investigators would want requestor identity, date, legal basis, targeted account or content, disposition, and associated internal ticket numbers. If no request exists, that negative record is valuable.

Corporate messaging systems.

Internal decisions may live in Slack, Microsoft Teams, email, Jira, Zendesk, Salesforce, custom moderation dashboards, or other systems. A subpoena that asks only for email can miss the actual decision trail. Counsel should first understand the company's system architecture or request records across communication platforms.

Automated decision systems.

If an algorithm or rule engine triggered the event, investigators need model or rule outputs, policy version, threshold, and human-review history. A reason code without the rule definition may be meaningless. Conversely, production of source code is not always necessary; a clear rule and event log may answer the question.

Minimization and proportionality.

A legitimate investigation should avoid sweeping up irrelevant employee or user data. Search terms, date limits, and account identifiers can narrow production. The goal is to understand one mass action, not to inspect the entire company.

This proportionality also strengthens the request. A narrowly defined evidentiary need is easier to justify than a fishing expedition built around investor associations.

Chapter 18: Evidence, Association, and Conspiracy

Conspiracy theories often begin with real facts.

A company really did receive investment from a prominent person. A scientist really did know another scientist. An agency really did fund a grant. An email really was sent. A platform really did remove material. The error comes in the connective tissue.

Association becomes coordination. Coordination becomes intent. Intent becomes causation. Causation becomes criminal conspiracy, all without the intermediate evidence.

The opposite error is common among institutional defenders. Because associative reasoning can be abused, they treat relationships as irrelevant. They dismiss network maps as guilt by association even when networks are precisely how institutions operate.

Both errors can be avoided by using an evidence ladder.

Level 1: Coexistence. Two people or institutions operate in the same broad ecosystem.

Level 2: Relationship. A documented professional, financial, organizational, or personal link exists.

Level 3: Communication. Evidence shows that the parties exchanged information relevant to the event.

Level 4: Coordination. Evidence shows that the parties aligned actions or plans.

Level 5: Causation. Evidence shows that the coordination materially caused the event under investigation.

Level 6: Unlawful conspiracy. Evidence establishes an agreement to pursue an unlawful objective or lawful objective through unlawful means, under the elements of the applicable offense.

Most internet arguments jump from Level 2 to Level 6.

This book refuses that jump.

Take Bill Gates and ResearchGate. Level 2 is documented: Gates invested in ResearchGate in 2013. That does not establish Level 3 concerning my account in 2026.

Take Gates and Fauci. Level 2 is documented through years of global-health collaboration involving NIAID and the Gates Foundation. That does not establish a communication about ResearchGate.

Take ResearchGate founders and virology. Biomedical and virology backgrounds are documented for key founders. Expertise is not an instruction to remove content.

Take my 2002 email. Fauci being copied creates a communication-path fact. It does not prove that he read or acted on the message.

Take Morens and federal-record concealment. Here the ladder progresses further because DOJ obtained evidence, charged conduct, and Morens pleaded guilty to a conspiracy offense. That legal record permits claims that would be irresponsible in the ResearchGate context.

The distinction is not merely legalistic. It protects inquiry.

If every question about institutional networks is dismissed as conspiracy theory, investigators will miss real coordination. If every network is treated as proof of conspiracy, investigators will drown in false positives.

The scientific method offers a better model: generate hypotheses, define discriminating evidence, seek data capable of falsification, and update.

For ResearchGate, one hypothesis is routine copyright enforcement. Another is account-level anti-spam or integrity enforcement. Another is a technical error. Another is policy enforcement unrelated to the political content. Another is subject-specific moderation. Another is external pressure. Another is a combination.

Each predicts different records.

Copyright enforcement predicts rights-holder notices and matching events.

Account-integrity enforcement predicts automated thresholds and policy codes.

Technical error predicts incident logs and broad effects beyond one account.

Subject-specific moderation predicts review of particular publications and topic-related internal discussion.

External pressure predicts communications from an outside party linked in time to the action.

A serious investigation ranks hypotheses by evidence, not emotional appeal.

This is also how a responsible author writes about people who are still alive. Defamation risk is not the only reason for caution. Intellectual integrity requires that a reader be able to see exactly where the record ends.

The subtitle of this book includes 'COVID Origins' and 'Hidden Emails' because those are documented contexts. It includes 'ResearchGate' because the platform event is documented in the author's account and surviving archive. It does not include a claim that ResearchGate joined a COVID conspiracy because that has not been established.

The absence of such a claim does not weaken the book. It gives the unanswered question sharper edges.

Bayesian discipline.

The evidence ladder can be expressed probabilistically. A documented relationship may modestly increase the plausibility of communication because connected people are more likely to communicate than strangers. But if the prior probability of improper coordination is low, the relationship alone may not move the posterior probability very far.

Direct communication about the event has much greater evidentiary weight. A message instructing action has more still. This intuitive Bayesian structure explains why networks are leads but not conclusions.

Competing narratives as models.

Instead of asking which story feels right, assign each story predictions.

Routine copyright enforcement predicts rights-holder records, version-specific targeting, and standardized notices.

AI-volume enforcement predicts account-scale thresholds, automation signals, and similar treatment of unrelated subject matter.

Technical failure predicts system incident reports and potentially other affected users.

Political content moderation predicts topic labels, escalation around particular papers, and perhaps policy references concerning misinformation or harmful content.

External pressure predicts a communication or request followed by action.

The models may overlap. A rights complaint could trigger a human review that then uncovers an account-integrity issue. An automated flag could lead to executive escalation. The goal is not to force one cause but to reconstruct causal sequence.

The ethics of naming.

Public figures can be named when their documented roles are relevant. But naming can create an implication even when the text contains a disclaimer. The structure of a chapter therefore matters as much as individual sentences.

If a book places Gates, Fauci, Madisch, Hofmayer, Morens, and ResearchGate on one page under a heading such as 'The Network,' readers may infer coordination. The manuscript instead places each relationship in its own evidentiary category and repeatedly states what is not established.

This is not timidity. It is analytic design.

Conspiracy as a legal concept.

In criminal law, conspiracy generally requires an agreement to pursue an unlawful objective or lawful objective through unlawful means, with jurisdiction-specific elements. Casual association, shared interests, parallel behavior, or even mutual benefit do not automatically satisfy those elements.

The Morens case uses conspiracy in this legal sense because a criminal defendant pleaded guilty. Applying the same word to ResearchGate without comparable evidence would collapse legal proof into rhetoric.

That collapse would undermine the very evidentiary standard the book argues for.

Chapter 19: The Scientific Record as Evidence

Scientists are trained to think of the record primarily in terms of knowledge: papers, datasets, methods, notebooks, and results. Lawyers and historians see additional dimensions.

A paper is a time-stamped statement.

A draft is evidence of evolution in thought.

An email is evidence of communication.

A grant is evidence of proposed work and institutional authorization.

A progress report is evidence of what a grantee represented to a funder.

A server log is evidence of access or change.

A DOI record is evidence of persistent identification.

A web capture is evidence of public availability.

A platform metric is evidence of measured attention under a defined system.

A deletion event is evidence of a state transition, although not necessarily of motive.

These categories matter because scientific controversies often involve retrospective claims about what was known. Someone says the possibility was never considered. An email shows it was discussed. Someone says a paper was not available. A Wayback capture shows it was public. Someone says a grant did not contemplate a type of work. An attachment can confirm or contradict the characterization.

The adversarial system is uncomfortable for scientists because it treats ordinary professional materials as potential evidence. Yet the same principle is familiar in research ethics. Laboratory notebooks must be preserved because results need traceability. Clinical trials are registered because selective reporting distorts inference. Data provenance matters because an analysis without provenance cannot be reproduced.

Public policy needs comparable provenance.

The pandemic exposed a mismatch between the standards scientists demand for experimental data and the casualness with which institutional communications are sometimes managed. A researcher may document a reagent lot number while discussing a major funding controversy on a personal Gmail account.

The Morens case makes the mismatch literal.

The answer is not total surveillance of scientists. Private deliberation has value. People need room to explore ideas, make mistakes, and speak candidly. But when a government employee conducts official business, records obligations attach for a reason.

Similarly, private scholarly platforms should not be forced to expose every moderation discussion to the public. Employee privacy, security, legal privilege, and anti-abuse systems require confidentiality. But when thousands of scholarly full texts change state, a platform should have durable internal provenance sufficient to explain the action to lawful authorities and, ordinarily, to the affected user.

Provenance is the compromise between transparency and chaos.

My own archive is now being reorganized with this principle in mind. Papers should be linked to original files, publication dates, DOIs where available, platform pages, archive captures, and cryptographic hashes. Email should retain headers. Websites should have capture dates. Indexes should be machine-readable.

This transforms a personal corpus into a historical collection.

The distinction may become increasingly important as AI-assisted writing expands. When one person can produce thousands of documents, conventional bibliographic expectations break down. Platforms may develop new policies. Readers may question whether volume itself is meaningful. Archives may face storage and curation burdens.

Those questions are legitimate. But mass production makes metadata more important, not less.

If a corpus is too large to read linearly, future analysis will depend on indexing, clustering, provenance, version control, and automated comparison. A vanished file among five thousand becomes visible only if there is an inventory.

The scientific record of the AI era will therefore be partly a record about records.

Chain of custody for scholars.

Chain of custody is usually associated with police evidence, but scholars handling controversial digital records can borrow the concept. Who possessed the file? When was it copied? Was it modified? What software processed it? Is the original preserved?

A simple protocol can answer these questions. Keep an untouched source directory. Calculate SHA-256 hashes. Work on copies. Record acquisition date and source URL. Preserve metadata in CSV or JSON. If conversion is necessary, retain both original and converted formats.

Email as a forensic object.

Printing an email to PDF destroys much of its forensic richness. The raw message may contain Message-ID, routing headers, authentication results, MIME boundaries, and attachment hashes. Those fields can help establish authenticity and relationships among messages.

For the 2002 correspondence, any surviving `.eml`, `.mbox`, or server export would be more valuable than a screenshot. The archived web page remains useful as evidence of public publication but is not a substitute for the original message object.

The evidentiary status of AI-assisted text.

AI-assisted papers raise a new question: which statements are the author's and which were generated by a model? For historical purposes, the published artifact remains the author's publication if he released it under his name, but provenance can help readers understand the drafting process.

Future archives may need to preserve prompts, model versions, revision histories, and human edits for certain forms of scholarship. That level of detail is not always necessary, but it can become important when a contested sentence is attributed to an author.

Corrections should be additive, not destructive.

When a paper contains an error, the best archival practice is to issue a corrected version while preserving the prior version with clear labeling. Silent replacement destroys the history of correction.

The same should apply to this book. If later records show that the ResearchGate event had a mundane cause, a future edition should say so explicitly while preserving the first edition as a record of what was known on August 22, 2026.

Evidence can mature.

A document's meaning can change when new corroboration appears. The 2002 email may be merely historical today. If future litigation concerns prior notice of a specific risk discussed in that email, its relevance could increase. Conversely, if the full text turns out not to concern the issue remembered, its significance decreases.

Archival preservation keeps those future evaluations possible.

Chapter 20: Who Owns Scientific Memory?

Ownership is the wrong word for memory, but it reveals the conflict.

The author may own copyright in an unpublished manuscript. A publisher may hold publication rights. A platform may control access to the hosted copy. A university may own laboratory notebooks under institutional policy. A federal agency may own the official record. A cloud provider may control the physical infrastructure. A web archive may preserve a copy under its own legal framework. A reader may have downloaded another copy years earlier.

No single actor owns the historical fact that the document existed.

Yet any one actor can sometimes make that fact harder to recover.

This distributed control is a defining feature of twenty-first-century science. The old scholarly record was concentrated in journals and libraries. The new record is scattered across email, repositories, preprint servers, code platforms, cloud drives, social media, collaborative documents, messaging systems, lab databases, grant portals, and commercial networks.

The benefit is redundancy. The risk is fragmentation.

The ResearchGate disappearance illustrates both. A single platform event removed public access to a huge set of full texts, according to my account. Redundant storage meant the corpus survived. The event therefore did not erase the writings from existence, but it did change their discoverability and break a public access path used by readers.

The Morens case illustrates a different form of fragmentation. Official communications were moved to a private account in an admitted effort to evade public-record systems. The messages may still have existed somewhere, but their location defeated the normal retrieval path.

The two episodes have no demonstrated operational connection. Conceptually, however, they reveal the same vulnerability: control of location can become control of discoverability.

Discoverability is power.

A fact buried in an inaccessible archive is not equivalent, in practical social terms, to a fact indexed by a search engine. A paper whose title remains but whose full text is unavailable has a different public life. An email held in a personal account has a different accountability status from an email captured by federal retention systems.

The future of scientific integrity therefore depends on more than preventing fabricated data. It depends on maintaining durable pathways from claims to underlying records.

This suggests several reforms.

Publicly funded research should have clear, enforceable preservation rules for grant-related communications and data.

High-risk biological research should use auditable systems that make scope changes, biosafety review, incident reporting, and material transfers traceable.

Private scholarly platforms should provide users with exportable inventories and meaningful explanations for account-wide content actions, subject to legitimate legal restrictions.

Researchers should maintain independent archives rather than treating social platforms as permanent repositories.

Web archives and public libraries should be recognized as critical scientific infrastructure.

DOI persistence should be paired with content-preservation strategies rather than mistaken for preservation itself.

And historians should begin collecting the digital ephemera of scientific controversies while those records still exist.

The central danger is not that one villain will erase science. It is that a distributed system will allow important records to vanish through a mixture of intention, policy, neglect, technical change, litigation, privacy constraints, corporate failure, and ordinary entropy.

The vanishing record is not a single conspiracy.

It is a structural condition.

Scientific memory as a commons.

A commons is a resource used by many people whose preservation depends on governance. Scientific memory fits that description. No one scientist, publisher, platform, or government owns the total record, yet everyone benefits when it remains reconstructable.

Commons fail through both overuse and neglect. In digital scholarship, neglect may take the form of dead links, obsolete formats, account closures, paywall migrations, and missing metadata. Overuse may appear as overwhelming volumes of low-quality content that make curation difficult. AI accelerates both pressures.

The answer is layered stewardship.

Authors preserve originals. Institutions maintain repositories. Publishers preserve version-of-record material. Public archives capture the web. Platforms expose export tools. DOI agencies maintain identifiers. Libraries curate collections. Regulators enforce records duties where public accountability requires them.

No layer is sufficient, but together they create resilience.

Memory and power.

Control over archives has always been political. States preserve some records and destroy others. Churches, corporations, universities, and families curate their own histories. Digital technology did not invent archival power; it distributed it among new actors.

What changed is speed. A database operation can alter access to thousands of documents in seconds. A retention policy can erase years of logs automatically. A search-engine ranking change can make a record practically invisible without deleting it.

Power over memory increasingly operates through software defaults.

The role of independent archives.

The Internet Archive is important precisely because it is institutionally separate from the sites it captures. Independence creates evidentiary value. A company can change its own page, but it does not directly control an older Wayback capture.

Scientific institutions should support more such independence. Critical records should not depend on the goodwill or longevity of the institution whose conduct they may later be used to evaluate.

A positive program.

The book's conclusion need not be cynical. The same digital tools that make erasure easy also make preservation cheap. Storage costs have collapsed. Hashing is trivial. Distributed mirrors are possible. Metadata standards can be automated. AI can help classify large archives.

The problem is therefore less technological than institutional. We know how to preserve. We have not decided consistently that preservation is part of scientific integrity.

The vanishing record is a warning, but it is also a design problem we can solve.

Epilogue: The Record Must Outlive the Argument

This book was written at an unstable moment. David Morens had pleaded guilty only days earlier. His sentencing had not occurred. Investigations and political arguments surrounding Anthony Fauci and the COVID-origin question continued. ResearchGate had not given me a substantive explanation for the disappearance of thousands of public full texts. The surviving archive had not yet been transformed into a complete forensic database.

That incompleteness is part of the point.

Books about active controversies often pretend to stand at the end of history. They arrange evidence as though every uncertainty has already been resolved. This manuscript does the opposite. It freezes a moment and labels the unknowns.

We know that Morens pleaded guilty to a federal conspiracy involving evasion of FOIA and federal-record requirements in connection with coronavirus-research-grant communications.

We know that ResearchGate was founded by medically trained researchers and a computer engineer; that it received major investments from Benchmark, Founders Fund, Bill Gates and others; and that it developed into important infrastructure for scholarly communication.

We know that NIH/NIAID and the Gates Foundation had extensive documented global-health collaborations years before COVID-19.

We know that my own archive contains extensive pandemic-era criticism of Fauci, NIAID, gain-of-function research, and the politics of scientific consensus.

We know that an older thread of my writing about biodefense reaches back to an NIH correspondence published through the Biological Manhattan Project website and preserved in web archives.

We know, from my account and surviving records, that approximately 5,700 public full texts ceased to be publicly available through my ResearchGate corpus during the weekend of July 18-19, 2026, after the account had accumulated roughly two million reads.

We do not know why ResearchGate took or experienced that action.

We do not know whether the action was a copyright process, account-integrity process, technical event, policy enforcement, legal restriction, external request, or something else.

We do not know whether any government official, investor, foundation, publisher, or politically interested third party communicated with ResearchGate about my account.

We do not know whether the old 2002 correspondence has relevance to any present investigation beyond its historical value.

Those unknowns are not defects to hide. They are the research program created by the book.

A future edition should be willing to become less dramatic. If ResearchGate eventually produces a clear mundane explanation supported by logs, that explanation belongs at the center of the revised narrative. If the 5,700 figure changes after a machine count, use the corrected number. If a remembered email proves less significant when the original is recovered, downgrade it. If the Morens case develops new adjudicated facts, update them. If a hypothesis fails, record the failure.

Truth is not served by preserving a theory. It is served by preserving the record that allows theories to die.

That principle extends beyond COVID. Artificial intelligence is accelerating the production of scholarship, commentary, code, images, and data. The quantity of intellectual material will soon make traditional archival habits inadequate. We will need automated inventories, hashes, provenance graphs, version histories, and

independent mirrors. We will need systems that can distinguish removal from deletion, correction from rewriting, and privacy from erasure.

The paradox is that memory has never been easier to store and never been easier to alter at scale.

A printed book is slow and static, which is now an advantage. This edition records what could be established as of August 22, 2026. Its electronic sources may change. Websites may disappear. Policies may be revised. Court cases will move. But the text establishes a baseline from which later evidence can be compared.

The final obligation is therefore simple: preserve first, argue second.

The record must outlive all of us.

Appendix A: Working Timeline

1991 - Viral manipulation as established experimental practice. My later historical writing points to early poliovirus work associated with Chow as part of a longer technical lineage in which viral properties could be altered experimentally. The relevance is methodological history, not proof concerning COVID-19.

2000 - Coronavirus host-range and molecular manipulation. A coronavirus paper associated with Peter Rottier and colleagues, indexed at PubMed PMID 10627550, became one of the technical milestones cited in my later work on host-range modification. Again, capability does not establish the origin of SARS-CoV-2.

September 11-October 2001 - Terrorist attacks and anthrax letters. The United States enters a major biodefense expansion. National-security concerns around biological research become part of ordinary federal policy rather than a fringe topic.

August 22, 2002 - NIH correspondence. My archival record identifies correspondence concerning NIH national security and SBIR/STTR funding involving NIH officials including Jo Anne Goodnight and Catherine Lewis, with Anthony S. Fauci copied. The exact preserved text and headers should control any interpretation.

2003 - Biological Manhattan Project pages captured. The Internet Archive captures biomanpro.org. A later reference in my 2022 paper points to a September 20, 2003 capture, and an NIH correspondence path at `biomanpro.org/text/nih.htm` is associated with a 2003 Wayback capture.

2008 - ResearchGate founded. ResearchGate's corporate history says physicians Ijad Madisch and Sören Hofmayer developed the idea and were joined by computer engineer Horst Fickenscher.

2010 - Benchmark investment. ResearchGate's history identifies Benchmark as an early institutional investor.

December 2010 - Decade of Vaccines Collaboration. The Gates Foundation announces a collaboration involving WHO, UNICEF, NIAID and the foundation. Anthony Fauci is named among the participating global-health leaders.

February 22, 2012 - Founders Fund investment. ResearchGate announces a Series B led by Founders Fund partner Luke Nosek. Nosek is described as joining the ResearchGate board.

2012 - Public gain-of-function debate. Fauci publishes and speaks publicly in defense of the potential benefits of certain research despite acknowledged risks. This period becomes important in later debates about pandemic-pathogen research governance.

June 4, 2013 - Bill Gates invests in ResearchGate. ResearchGate announces a $35 million financing round led by Bill Gates and Tenaya Capital, with additional participation.

November 2013 - Gates visits ResearchGate in Berlin. ResearchGate publishes an account and photograph of Gates visiting the company's headquarters.

2015 - Additional ResearchGate financing. ResearchGate's own history says Wellcome Trust's investment arm and Goldman Sachs Investment Partners participated in its Series D round.

2018 - ResearchGate publisher cooperation. ResearchGate announces arrangements with scientific publishers concerning copyright enforcement and says that in covered cases it removes reported content expeditiously and informs the scientist who uploaded it.

October 23, 2019 - NIH-Gates Foundation gene-cure collaboration. NIH announces that it and the Gates Foundation plan to invest at least $100 million each toward accessible gene-based cures for sickle cell disease and HIV. Fauci is quoted in the NIH announcement concerning the HIV component.

2019-2020 - SARS-CoV-2 emerges and the origin controversy begins. Natural spillover and research-related scenarios become politically and scientifically contested.

2020 - EcoHealth-related grant controversy. NIH terminates the NIAID-funded grant described by DOJ as *Understanding the Risk of Bat Coronavirus Emergence* amid the dispute over Wuhan laboratory research and COVID origins.

2020 onward - FOIA litigation and records requests. Journalists, nonprofits, congressional investigators and others seek NIH and NIAID communications related to origins, grants and influential scientists.

February 18, 2022 - *Fermi Paradox, Viropaths, and Depopulation*. My paper is uploaded to ResearchGate. The surviving PDF later reports more than 40,000 reads on its ResearchGate cover page and explicitly references the archived Biological Manhattan Project site in connection with an old Fauci-related email.

June 29, 2023 - Congressional inquiry to Morens. A House Select Subcommittee letter states that records in its possession suggest Morens may have used personal email to avoid FOIA and may have deleted federal records. At this stage these are investigative allegations.

2024 - Morens congressional investigation intensifies. Investigators subpoena personal-account records and conduct hearings. Public debate continues over the interpretation of Morens's messages and their relationship to NIAID leadership.

January 2024 - *Shadow Academia*. My paper *Shadow Academia: Unveiling the Concealed Narratives of Global Health Crises* describes independent pandemic-origin researchers and criticizes perceived failures of transparency in conventional institutions.

April 28, 2026 - Morens indicted. DOJ announces federal charges including conspiracy and records-related offenses. An indictment is an allegation and Morens remains presumed innocent at that stage.

June 2026 - Additional pandemic/institutional papers in the archive. Surviving Drive materials show continued writing on Fauci, staff testimony, platform attention metrics, and related institutional questions.

Weekend of July 18-19, 2026 - ResearchGate full-text disappearance. According to my account, approximately 5,700 publicly available PDFs associated with my ResearchGate corpus cease to be publicly accessible without advance warning or an identified strike sequence. Many archive copies have Drive ingestion timestamps around this weekend.

July-August 2026 - Outreach to ResearchGate and related contacts. I send numerous requests for explanation and receive no substantive response explaining the mass event.

August 18, 2026 - Morens guilty plea. The U.S. Attorney's Office for the District of Maryland announces that Morens pleaded guilty to conspiracy to commit offenses and to defraud the United States in a scheme involving evasion of FOIA and the Federal Records Act in connection with coronavirus-research-grant communications.

August 22, 2026 - Documentary cutoff for this edition. The ResearchGate cause remains unresolved. The manuscript is assembled as an evidence-preservation project rather than a conclusion of motive.

November 12, 2026 - Scheduled Morens sentencing. Future at the time of this edition. Later editions should update the outcome from the court record.

Appendix B: The 2002 NIH Correspondence Provenance

The 2002 correspondence deserves a precise provenance statement because it is easy for later controversy to distort old material.

What is presently supported. The surviving archive and earlier recovery work identify an August 22, 2002 NIH-related email thread associated with subject language including 'national security at NIH?' and 'NIH SBIR/STTR Funding Opportunities.' The thread involved NIH officials including Jo Anne Goodnight and Catherine Lewis, and Anthony S. Fauci was copied at an NIAID address.

What that proves. It supports the proposition that an email was addressed through a network that included Fauci's account. It may also demonstrate that I was raising biodefense or national-security questions with NIH in 2002, depending on the exact recovered text.

What it does not prove. A CC line does not prove that Fauci personally opened the email, read it, remembered it, agreed with it, or acted on it. It does not prove that a later event resulted from the warning. It does not establish foreknowledge of COVID-19.

Public archival provenance. My February 2022 paper *Fermi Paradox, Viropaths, and Depopulation* states that I had already published the email on the archived Biological Manhattan Project website. Its reference 13 points to a Wayback capture of `http://www.biomanpro.org/` dated September 20, 2003. An additional Wayback path identified in later research is `http://www.biomanpro.org/text/nih.htm`, captured in 2003.

Recommended evidence bundle. If the correspondence is ever provided in a formal investigation, preserve: the raw Wayback HTML; Wayback capture metadata; all linked images or attachments; any local original email files; full email headers; any replies or preceding quoted messages; file hashes; and a short declaration explaining how the records were located and preserved.

Authentication priority. The best evidence would be the original electronic mail object from a sender or recipient mailbox. The archived webpage is powerful corroborative evidence of public historical existence but should not be described as the original email file unless it actually contains the raw message.

Historical context. The correspondence occurred within a year of the 2001 anthrax attacks, during a major U.S. expansion of biodefense and NIH-related small-business innovation programs. Concerns about dual-use research and biological national security were therefore part of the policy environment at the time.

Research agenda. Future work should search old backups, exported mailboxes, Kairos Scientific records, NIH correspondence, web mirrors, and recipient archives for the original message and any response. Exact text should replace remembered paraphrase wherever the two differ.

Appendix C: ResearchGate Record-Preservation Questions

This appendix is written as an investigative checklist, not a legal demand. Counsel or investigators would determine the appropriate legal process, jurisdiction, scope and wording.

Account identification

- What unique ResearchGate user ID corresponds to Douglas C. Youvan's account?
- What email addresses, profile URLs and historical usernames were associated with it?
- Were multiple accounts ever merged or linked?
- Was the account assigned any internal high-volume, integrity, risk, rights, moderation or legal status?

The July 2026 event

- What administrative or automated actions affected the account between April 1 and August 31, 2026?
- Did one bulk operation change the visibility of thousands of files?
- What exact operation was performed: delete, detach, make private, block, restrict, quarantine, demote, de-index or another state?
- What timestamp did the operation begin and end?
- How many file objects and publication pages were affected?
- Did any objects fail or receive a different state?
- What software job, employee action or moderation workflow initiated the operation?

Reason and policy

- What reason code or policy category was recorded?
- Which version of the relevant policy was in force?
- Was the action based on copyright, platform integrity, spam, automation, community guidelines, legal process, misinformation policy, safety policy, account volume, technical migration or another category?
- Was the action standard for comparable accounts?
- Were other accounts affected by the same rule during the same period?

Notice and appeal

- Did ResearchGate generate a notice to the user?
- What address was used?
- What was the template or text of the notice?
- Was delivery attempted, successful, bounced, suppressed or blocked?
- Did an internal rule permit action without notice?
- Was an appeal case created or made available?
- How were the user's subsequent messages classified and routed?

Copyright

- Which files, if any, were reported by publishers or rights holders?
- Who submitted each report?
- What legal or contractual basis was asserted?
- Did any copyright complaint result in an account-level rather than item-level action?
- Were wholly author-created preprints and unpublished manuscripts affected by the same operation?
- Were automated copyright matches involved, and if so, did those matches count as strikes?

External communications

- Did ResearchGate receive any request, complaint, inquiry or communication from a U.S. government agency concerning the account or specified publications?

- Did it receive communications from NIH, NIAID, HHS, CDC, DOJ, FBI, OSTP, White House offices or contractors concerning the account?
- Did it receive communications from any person or organization connected to EcoHealth Alliance, COVID-origin research, or named researchers discussed in the affected corpus?
- Did any investor, board member, foundation, publisher or partner inquire about the account?
- If such communications existed, did they cause or influence the July action?

These questions are deliberately neutral. A 'no' answer is as important as a 'yes.'

Internal communications

- Search email and internal messaging for the user ID, account email, name, key publication IDs and high-read COVID-origin titles.
- Search terms should include ResearchGate-specific action terminology as well as NIAID, NIH, Fauci, Morens, Daszak, EcoHealth, Baric, COVID origins, lab leak, gain of function, FOIA, government request, legal request, copyright, misinformation, escalation and bulk removal.
- Identify tickets, channels or case IDs revealed by initial searches and use them to retrieve complete threads.

Audit and database records

- Preserve account audit logs.
- Preserve publication-object and file-object state histories.
- Preserve administrator and moderator access events relevant to the account.
- Preserve automated classifier outputs or rule-engine events that contributed to the decision.
- Preserve the policy/rule version and threshold used by any automated system.

Backups and historical states

- Identify database snapshots covering the weeks before and after July 18-19, 2026.
- Determine whether the files themselves remain in backup or private storage.
- Determine whether file hashes and object IDs persist after public access is removed.
- Preserve retention schedules showing when relevant logs or messages would ordinarily be deleted.

Corporate custody

- Identify which ResearchGate legal entity controls user-content records.
- Identify which entity controls moderation and legal-request records.
- Identify data-storage jurisdictions and relevant service providers.
- Identify custodians with responsibility for trust and safety, copyright, legal compliance and account integrity during the relevant period.

Investigative endpoint

The inquiry should be designed to produce a causal sequence: trigger -> review -> decision -> technical action -> notice -> response. Once that sequence is documented, speculative network analysis should yield to the records.

Appendix D: Selected Publications from the Surviving Corpus

The following titles were located in my surviving Google Drive archive during preparation of this book. Their inclusion here documents subject matter and survival; it does not endorse every factual claim contained within them.

Fermi Paradox, Viropaths, and Depopulation - February 2022. The surviving ResearchGate-formatted PDF identifies Douglas C. Youvan and Charles Juvon as authors and includes a reference to the archived Biological Manhattan Project website. The paper uses a speculative Fermi-paradox framework to discuss cumulative risks from dangerous biological research.

Shadow Academia: Unveiling the Concealed Narratives of Global Health Crises - January 2024. A paper about independent researchers, alternative scientific networks, pandemic origins, media narratives, gain-of-function debates and scientific transparency. Its claims should be distinguished from independent verification.

Tracing the Origins of Gain-of-Function Research to Polio (Chow 1991) and Coronavirus (Rottier 2000) - a historical paper connecting earlier experimental virology to later policy controversies.

The Silent Controversy - Gain-of-Function Research and the Disbanding of MIT's Department of Applied Biological Sciences - a paper connecting personal institutional history at MIT with concerns about viral manipulation and research governance.

The Dual-Edged Sword - Navigating Ethical and Scientific Frontiers at MIT - related historical and ethical analysis of dual-use scientific work.

Red Team Training of AI Is Analogous to Gain-of-Function For Viruses - an analogy between adversarial AI development and biological enhancement research, illustrating how the gain-of-function concept migrated into my later AI safety writing.

Biden's Presidential Pardon - Fauci Can Still Face Civil Lawsuits - a legal/political analysis preserved in Drive. The title reflects an argument; it is not evidence that any particular civil claim is valid.

When the Witnesses Fear for Their Lives - Fauci Staff, Presidential Pardons, and a Hypothetical Biological-Release Testimony Crisis - a hypothetical scenario paper from June 2026. The word 'hypothetical' is essential to how its claims should be read.

Quotes From Notable People on the COVID-19 Pandemic and the Operation Warp Speed Vaccine - a compilation that includes sections on Fauci, Ralph Baric and Nobel laureate Sir Richard Roberts. Quotation compilations require independent verification of original sources before legal use.

Perception Warfare and the Pandemic Trigger - Reframing COVID-19's Origin as a Strategic Information Operation - a paper using an information-warfare framework. Its interpretive thesis should not be confused with established origin evidence.

The Global Bioweapons Network - broader analysis of institutions, laboratories and biological-security concerns.

Exploring Historical Context and Future Predictions - An AI-Guided Dialogue - a dialogue-format work that demonstrates the experimental AI-assisted nature of portions of the corpus.

AI Faked the Read - A Transcript Case Study in the Collapse of Preprint-Server Attention Metrics - a later paper questioning the reliability and meaning of platform attention metrics, relevant to interpreting large ResearchGate read counts.

Network Map of Keywords and Titles to ~6,000 Youvan AI-Assisted Preprints on ResearchGate - especially important as an index artifact that may help reconstruct the pre-disappearance corpus.

These surviving documents demonstrate that the affected body of work was heterogeneous. It contained controversial pandemic material, but also methodological, historical, AI, mathematical, theological and other writings. A full forensic inventory should classify all categories before drawing conclusions about selectivity.

Appendix E: Primary Sources and References

The list below emphasizes primary and institutional sources used to establish the factual framework. Contemporary commentary and the author's own papers are treated separately.

United States Department of Justice, U.S. Attorney's Office, District of Maryland. "Former Senior NIAID Official Pleads Guilty to Charges Connected to Concealing Federal Records During COVID-19 Pandemic." August 18, 2026. https://www.justice.gov/usao-md/pr/former-senior-niaid-official-pleads-guilty-charges-connected-concealing-federal-records

This is the principal source for the Morens guilty plea, his position at NIAID, the described purpose of the conspiracy, the use of personal Gmail, the coronavirus-grant context, investigating agencies, prosecutors and maximum penalty.

United States Department of Justice, Office of Public Affairs. "Former Senior NIAID Official Indicted for Concealing Federal Records During COVID-19 Pandemic." April 28, 2026. https://www.justice.gov/opa/pr/former-senior-niaid-official-indicted-concealing-federal-records-during-covid-19-pandemic-0

Used for the pre-plea procedural history and the distinction between allegations in the indictment and later admitted conduct.

ResearchGate. "Appeals Policy." Updated February 19, 2024. https://help.researchgate.net/hc/en-us/articles/22440088156433-Appeals-Policy

Used for ResearchGate's public description of qualifying content/account decisions, user notification, rule explanations and appeal procedures.

ResearchGate. "Strikes Policy." Last updated February 20, 2024. Available through ResearchGate's legal-policy archive: https://help.researchgate.net/legal/archive-of-researchgate-terms-and-policies

Used for ResearchGate's public description of warnings, strikes, content actions and account disabling.

ResearchGate. "How to make content private or remove it." https://help.researchgate.net/hc/en-us/articles/14293099743121-How-to-make-content-private-or-remove-it

Used to distinguish making full texts private, removing public full texts and deleting publication pages.

ResearchGate. "Copyright checks on ResearchGate." https://help.researchgate.net/hc/en-us/articles/14295499851665-Copyright-checks-on-ResearchGate

Used for the distinction between automated copyright matches and strikes, and for publisher-identified infringement processes.

ResearchGate. "Celebrating 10 years of ResearchGate." May 23, 2018. https://www.researchgate.net/blog/celebrating-10-years-of-researchgate

Primary corporate history for the founding team, growth milestones and investment chronology.

ResearchGate. "Founders Fund Invests in ResearchGate, The Professional Network For Scientists." February 22, 2012. https://www.researchgate.net/press-newsroom/founders-fund-invests-in-researchgate

Used to identify Founders Fund's Series B role, Luke Nosek's leadership of the investment and his announced board role.

ResearchGate. "Bill Gates and Tenaya Capital invest in ResearchGate to advance scientific innovation through open science." June 4, 2013. https://www.researchgate.net/blog/post/bill-gates-and-tenaya-capital-invest-in-researchgate-to-advance-scientific-innovation-through-open-science

Primary source for the announced $35 million financing and Gates's investment relationship with ResearchGate.

ResearchGate. "Bill Gates @ ResearchGate." November 14, 2013. https://www.researchgate.net/blog/bill-gates-researchgate

Corporate account of Gates's visit to ResearchGate's Berlin headquarters.

ResearchGate. "ResearchGate welcomes cooperation agreement with three major scientific publishers." May 2018. https://www.researchgate.net/press-newsroom/researchgate-cooperation-agreement-with-scientific-publishers

Used for ResearchGate's historical statement concerning publisher reports, removal and notification of scientists.

Bill & Melinda Gates Foundation. "Global Health Leaders Launch Decade of Vaccines Collaboration." December 2010. https://www.gatesfoundation.org/ideas/media-center/press-releases/2010/12/global-health-leaders-launch-decade-of-vaccines-collaboration

Primary source documenting collaboration among WHO, UNICEF, NIAID and the Gates Foundation and naming Fauci among the participating leaders.

National Institutes of Health. "NIH launches new collaboration to develop gene-based cures for sickle cell disease and HIV on global scale." October 23, 2019. https://www.nih.gov/news-events/news-releases/nih-launches-new-collaboration-develop-gene-based-cures-sickle-cell-disease-hiv-global-scale

Primary source documenting the planned NIH and Gates Foundation investments and quoting Fauci on the collaboration.

Congressional investigative materials. House Select Subcommittee on the Coronavirus Pandemic correspondence and hearing records concerning David Morens, 2023-2024. Public materials include a June 29, 2023 letter and subsequent hearing documents archived through Congress.gov and House Oversight resources. These records are used as evidence of the investigation and allegations at the time, not as substitutes for criminal adjudication.

Internet Archive, Wayback Machine. Captures of `biomanpro.org`, including a September 20, 2003 site capture cited in the author's 2022 paper and an identified 2003 capture of `biomanpro.org/text/nih.htm`. https://web.archive.org/

Used for provenance of the Biological Manhattan Project web archive. Exact archived pages should be preserved separately for evidentiary use.

Youvan, Douglas C., and Charles Juvon. *Fermi Paradox, Viropaths, and Depopulation.* Preprint, February 18, 2022. DOI shown in the surviving ResearchGate-format PDF: 10.13140/RG.2.2.15213.46563. Surviving copy in the author's archive.

This source is used principally for evidence of what the author published in 2022 and for its archival reference to the Biological Manhattan Project. Controversial assertions within it are not treated as independently established facts by this book.

Youvan, Douglas C. *Shadow Academia: Unveiling the Concealed Narratives of Global Health Crises.* January 13, 2024. Surviving copy in the author's archive.

Used as evidence of the author's earlier framing of decentralized pandemic research and perceived transparency problems. Factual claims within the paper require independent verification.

Youvan corpus index and surviving Drive copies. The connected archive includes a network map of roughly six thousand AI-assisted ResearchGate preprints and numerous surviving publication PDFs. These materials provide a basis for reconstructing the corpus but should be converted into a machine-readable inventory for rigorous quantitative analysis.

About the Author

Douglas C. Youvan is a biophysicist and author whose career has spanned academic research, biotechnology, computational work and independent publishing. His recent work has focused heavily on AI-assisted scholarship, mathematical experimentation, scientific provenance and the preservation of large digital research corpora.

This book is part investigation, part archival protocol, and part argument that scientific integrity must include the integrity of the record itself.