

# UCSF Industry Documents Library Collection Development Policy

## Introduction

The purpose of the UCSF Industry Documents Library (IDL) Collection Development Policy is to: 1) define the scope of the materials to be collected and preserved in IDL; 2) outline the process of acquiring these materials; and 3) provide detailed criteria for collecting decisions.

### *Industry Documents Library Mission Statement:*

The mission of the UCSF Industry Documents Library is to identify, collect, curate, preserve, and make freely accessible internal documents created by industries and their partners which have an impact on public health, for the benefit and use of researchers, clinicians, educators, students, policymakers, media, and the general public at UCSF and internationally.

## Research Community

The [UCSF Industry Documents Library](https://www.industrydocuments.ucsf.edu) (IDL) is an online portal to aid investigation about corporate practices within and across industries which have an impact on public health. IDL currently brings together materials created by the tobacco, opioid, drug, chemical, food, and fossil fuel industries and by individuals and organizations working with industries. These collections are an important resource for researchers, lawyers, journalists, advocates and the general public as they document industry strategies and tactics around scientific research, marketing campaigns, regulatory strategies, and public health policy formulation.

The Industry Documents Library is freely available to all and can be accessed at any time on the web at <https://www.industrydocuments.ucsf.edu>.

## Collecting Scope

The UCSF Industry Documents Library collects documents produced by industries and their partners which illustrate efforts to influence policies and regulations meant to protect public health.

IDL's primary emphasis has focused on the tobacco industry, with expansion into other industries, including opioid, drug, chemical, food, and fossil fuel industries. Additional industries may be added pending availability of relevant document sets, funding opportunities, and relevance to UCSF research priorities involving public health and public health policy.

IDL concentrates its efforts and resources on collecting digital copies of primary source documents publicly disclosed in litigation, including correspondence, memos, internal reports, meeting minutes, charts, pamphlets, advertisements, depositions, trial transcripts, photographs, illustrations, social media posts, webpages, and research data. IDL does not generally collect journal articles, publications, newspapers, or other secondary sources readily available elsewhere.

The primary responsibility of the Industry Documents Library is to serve the research needs of individuals interested in public health.

## Current Collections

As of July 2024, the Industry Documents Library includes documents in six major collections:

- Truth Tobacco Industry Documents (TTID): over 17 million documents
- Opioid Industry Documents Archive (OIDA): over 3 million documents
- Chemical Industry Documents Archive (CIDA): over 13,000 documents
- Drug Industry Documents Archive (DIDA): over 5,700 documents
- Food Industry Documents Archive (FIDA): over 173,000 documents
- Fossil Fuel Industry Documents Archive (FoFIDA): over 1,100 documents

**The Truth Tobacco Industry Documents Library** (formerly known as the Legacy Tobacco Documents Library) was created in 2002 by the [UCSF Library](#) in cooperation with the [UCSF Center for Tobacco Control Research and Education](#). The archive was built to house and provide permanent access to tobacco industry internal corporate documents produced during litigation between U.S. states and the seven major tobacco industry organizations and other sources. These internal documents give a view into the workings of one of the largest and most influential industries in the United States.

**The Opioid Industry Documents Archive (OIDA)** is a collaboration between the UCSF Industry Documents Library and [Johns Hopkins University](#). It was [launched in March 2021](#) and is envisioned as a groundbreaking digital archive of opioid industry documents that advances understanding of the root causes of the U.S. opioid epidemic, promotes transparency and accountability, and informs and enables evidence-based research and investigation to protect and improve public health. The documents are internal corporate records publicly disclosed from ongoing opioid litigation brought by local and state governments and tribal communities against opioid manufacturers, wholesalers, distributors, and pharmacies. These lawsuits argue that opioid manufacturers and distributors pursued manipulative and misleading marketing strategies, cast doubt on the addictiveness of their drugs, and disregarded the significant risks to health, leading to a national opioid overdose epidemic and public health crisis. Other documents are collected through public records requests, journalistic investigation, and other sources. The full scope of collecting activities for the Opioid Industry Documents Archive is outlined in the [OIDA Collection Development Policy](#).

**The Drug Industry Documents Archive** was created by the UCSF Library in collaboration with faculty members C. Seth Landefeld, MD and Michael Steinman, MD. Originally established to house documents from an off-label marketing lawsuit against Parke-Davis (*United States of America ex rel. David Franklin vs. Parke-Davis*), the archive has grown to include documents from additional sources illustrating how the pharmaceutical industry, academic journals and institutions, continuing medical education organizations and regulatory/funding agencies operate in ways that influence public health.

**The Chemical Industry Documents Archive** was created by the UCSF Library in 2017 in collaboration with the UCSF [Program on Reproductive Health and the Environment \(PRHE\)](#). It contains significant collections documenting internal practices of companies and regulatory agencies which impact the public's exposure to toxic chemicals.

**The Food Industry Documents Archive** launched in November 2018 as a collaboration between the UCSF Library and the [Philip R. Lee Institute for Health Policy Studies](#), with an initial focus on the sugar industry. The scope of this collection is expected to expand into other areas of food industry science and policy.

**The Fossil Fuel Industry Documents Archive** was established in June 2019 with a foundational collection of documents provided by the [Climate Investigations Center](#), which include internal memos, reports, correspondence and scientific studies from major fossil fuel corporations and related organizations which detail the industry's research and reaction to anthropogenic climate change beginning in the 1950s to the present.

## Collection Sources

Industry documents collections are obtained through:

- Litigation: IDL collects industry documents which have been publicly disclosed after being produced during legal discovery or as evidence submitted during trial. We both solicit and receive these documents via court order (as in the *US v. Philip Morris* racketeering case), attorneys (plaintiff's attorneys, class action lawyers, state attorneys general), expert witnesses, and PACER (Public Access to Court Electronic Records) downloads.
- Existing archives: IDL also enters into partnership with libraries and archives to digitize selected in-scope collections from their holdings, including company records and professional papers of academics or individuals that consulted with or received funding from industry.

- Records requests: Freedom of Information Act (FOIA) and state public records requests.
- Research files: collected or donated by academics, authors, investigative journalists, and others.
- Websites and social media accounts of companies, front groups, trade associations and other organizations.

In general, IDL does not collect documents that:

- Are outside of its collecting scope as defined in this Policy;
- Come with a substantial processing cost without a related funding source;
- Are readily available and accessible in other archives;
- Are not authorized for public disclosure due to permanent legal restrictions such as a protective order, confidentiality agreement, or other formal settlement terms.

## Collection Development for Web Archives

In addition to digitized and digital documents, the Industry Documents Library also collects and curates websites and webpages which are relevant to IDL's mission and to the industries documented in the collections. The [UCSF Industry Documents Library Web Archive](#) includes information from webpages, websites, and social media platforms.

IDL selectively collects and maintains web archives which complement the document collections as outlined below:

- Websites and webpages published by major companies in each industry (ex: [www.philipmorrisusa.com](http://www.philipmorrisusa.com)) or major sponsors/lobbyists
- Websites and webpages identified as relevant to specific research projects, such as e-cigarette marketing

- Websites and webpages related to a relevant and time-based event, such as industry-related ballot campaigns, regulation, or public policy measures. Currently the geographic scope of campaign-related web archives is limited to California.

## Budget

The Industry Documents Library does not purchase any documents or document sets. IDL funds are primarily obtained through grants, and are spent on digitization, indexing, and document storage costs to enable access to the collections.

## Intellectual Freedom and Censorship

The mission of the UCSF Industry Documents Library is to identify, collect, curate, preserve, and make freely accessible internal documents created by industries and their partners which have an impact on public health, for the benefit and use of researchers, clinicians, educators, students, policymakers, media, and the general public at UCSF and internationally. The IDL is wholly committed to the principles of universal access to information and intellectual freedom, and makes its collections freely available to all so that researchers can exercise their judgement and expertise in interpreting the documents in the broader context of available information and scientific inquiry. The presence of any particular document does not imply IDL or UCSF's endorsement or judgement of its content.

## Collecting Priorities for 2024-2025

1. Truth Tobacco Industry Documents
  - Process collections to support UCSF research relating to flavored tobacco products and e-cigarettes, including millions of documents disclosed in litigation against JUUL.
  - Assess collecting options for documents related to the marijuana/cannabis industry.
2. Opioid Industry Documents Archive

- With Johns Hopkins University, continue to serve as official repository for documents publicly disclosed as a result of the Insys and Mallinckrodt bankruptcy settlements and McKinsey settlement.
  - Manage the ingest, processing, and public release of documents disclosed under the Teva and Allergan settlements, the Endo bankruptcy settlement, and the Publicis Health settlement.
  - Continue working with offices of state attorneys general to collect, process, preserve, and make public additional opioid industry documents disclosed in ongoing settlements.
  - Work with partners and stakeholders to identify other opioid document collections.
3. Drug Industry Documents Archive
- Explore opportunities to secure additional collections and funding.
4. Chemical Industry Documents Archive
- Obtain additional collections and funding to support research needs of the UCSF Program on Reproductive Health and the Environment (PRHE) and other community stakeholders, including documents obtained in litigation related to talc and asbestos.
5. Food Industry Documents Archive
- Obtain additional collections and funding to support research needs of the UCSF Institute for Health Policy Studies (IHPS) and other community stakeholders.
6. Fossil Fuel Industry Documents Archive
- Monitor litigation against oil and gas companies and continue collecting relevant documents that become publicly available.

## Document Contributions

The IDL partners with libraries, archives, organizations, researchers, journalists, attorneys, and others to share and make publicly available digital copies of documents illustrating industry practices which impact public health.

Document contributions accepted are subject to the policies of the Regents of the University of California, with the understanding that the university is a custodian of the documents upon receipt and reserves the right to determine retention, location, cataloging, digitization and other considerations relating to use or disposition.

Those providing physical or digital documents to the Industry Documents Library agree to grant the IDL and its users full permission to digitize, download, enhance, describe, excerpt, re-use, and publish online all contributed documents and their contents. The IDL credits the contributing individual or organization in all document description, as appropriate.

The IDL welcomes inquiries regarding contribution of documents via the IDL website: <https://www.industrydocumentslibrary.ucsf.edu/about/donate/>.

## Collection Evaluation Methods

### Collecting Criteria

- Relevance: topical or evidential relevance to public health research and publications.
- Research (aka informational) value: documents which contain information about important people, companies, trade associations, products and issues.
- Evidential value: documents that by their very existence provide insight about the corporation or organization that created them.
- Uniqueness: whether these documents exist elsewhere, particularly online.
- Value: long-term value to public health research and public health policy.
- Scope: usefulness across industries and across disciplines.
- Timeliness: the issue represented is under public discussion, or ripe for discussion.
- Legal risk: copyright, privacy and confidentiality, litigation status.
- Cost: of processing compared with research value.

A detailed IDL Appraisal Checklist is available at the end of this document.



## Collaboration with Other Repositories

The Industry Documents Library partners with other libraries and archives to digitize and provide access to collections for the benefit of all users. Collections are obtained from other repositories with their full knowledge and permission, and are appropriately credited on each collection page.

## Deaccessioning and Disposition Policy

The Industry Documents Library may remove and deaccession documents or collections determined to be out of the scope of our collecting policy.

If deaccessioning of documents or collections is necessary, the Industry Documents Library is guided in this process by archival best practices, including the Society of American Archivists' (*SAA Guidelines for Reappraisal and Deaccessioning*; *SAA's Core Values Statement and Code of Ethics*); and the *ACRL Code of Ethics for Special Collections Librarians*, established by the Rare Books and Manuscripts Section (RBMS) of the Association of College and Research Libraries (ACRL). IDL will do its best to adhere to any donor specifications or wishes provided at the time of accession when considering the disposition of any documents from the collection.

## Takedown Policy

### **Copyright and Intellectual Property**

The Industry Documents Library makes every effort to ensure that it has appropriate rights to ingest and provide access to documents and collections. Please contact IDL staff with any questions or concerns regarding the copyright of specific documents, or to make a takedown request. Takedown requests will be acknowledged within ten business days and reviewed for further action. The Library may work with UCSF's Office of Legal Affairs to make determinations about copyright and appropriate use.

### **Protection of Personal Information**

Most of the documents in the IDL were created as internal records by corporate entities, including companies in the tobacco, opioid, drug, chemical, food, and fossil fuel industries. Many of these documents are subject to public disclosure by order of the Court, under the terms of settlement agreements reached in litigation, or have been released through Freedom of Information Act (FOIA) or public records requests.

These documents may contain personal information (such as addresses, phone numbers, or social security numbers) related to patients, employees, or other individuals, which may be considered protected data under applicable state and federal law. UCSF and its partners take great care to ensure that any protected information is identified and redacted before documents are made public, while also ensuring that data of crucial relevance to these cases, and to transparency and accountability, remain open.

“Protected information” includes:

- Personally Identifiable Information (PII) such as social security number, driver’s license number, or home address
- Protected Health Information (PHI) such as medical record number, date of birth, or diagnostic image
- Other sensitive personal information as defined by applicable state and federal privacy law;
- Information regarding company employees’ personal matters unrelated to the company or its products, including but not limited to emails discussing vacation or sick leave, family, or other personal matters.

#### **Information which is NOT redacted:**

Some information remains unredacted according to document disclosure terms in specific settlement agreements, including:

- Under the document disclosure terms of Mallinckrodt’s settlement, “‘Confidential personal information’ does not include the names of Mallinckrodt’s officers, directors, employees, agents, or attorneys.”<sup>1</sup>
- Under the document disclosure terms of McKinsey’s settlement, “‘Confidential personal information’ does not include the names of officers, directors, employees, agents, or attorneys of McKinsey, Purdue, Endo, Johnson & Johnson, or Mallinckrodt, or of a government agency.”<sup>2</sup>
- Under the document disclosure terms of Teva’s settlement, “‘Confidential personal information’ does not include the names of Teva’s officers, directors, employees, consultants, agents, or attorneys or of prescribers or of officials of a government agency.”<sup>3</sup>
- Under the document disclosure terms of Allergan’s settlement, “‘Confidential personal information’ does not include the names of Allergan’s officers, directors, employees, consultants, agents, or attorneys or of prescribers or of officials of a government agency.”<sup>4</sup>

If protected information which has *not* been redacted is found in a public document, the following steps are taken:

1. immediately remove the document from public access, pending review;
2. conduct a full review of the document under guidance from University legal and privacy experts;

---

<sup>1</sup> Order Granting Certain Debtors’ Motion for Injunctive Relief Pursuant to 11 U.S.C. §105 With Respect to the Voluntary Injunction. In re: Mallinckrodt PLC, et al., v. State of Connecticut, et al. United States Bankruptcy Court for the District of Delaware, Case No. 20-12522, Document 196-1.

<sup>2</sup> Assented-To Motion for Entry of Judgment. In re: Commonwealth of Massachusetts v. McKinsey & Company, Inc, United States. Suffolk County Superior Court.

<sup>3</sup> Final Teva Global Settlement Agreement and Exhibits. August 29, 2023. Retrieved from <https://nationalopioidsettlement.com/wp-content/uploads/2023/08/Final-Teva-Global-Settlement-Agreement-and-Exhibits-8.29.23.pdf>.

<sup>4</sup> Final Allergan Global Settlement Agreement and Exhibits. August 29, 2023. Retrieved from <https://nationalopioidsettlement.com/wp-content/uploads/2023/08/Final-Allergan-Settlement-Agreement-8-29-23.pdf>.

3. determine if the document does contain protected information and if that information has been inadvertently released;
4. if the document contains protected information, take steps to mediate, including alerting IT Security;
5. determine if:
  - a) the document should be redacted and published in redacted form;
  - b) the document does not require redaction and can be re-published in its original form;
  - c) the document contains wholly personal or protected information and should be removed from the archive.
6. Provide a written response and explanation if requested.

Please contact UCSF Industry Documents Library staff at [industrydocuments@ucsf.edu](mailto:industrydocuments@ucsf.edu) with any concerns regarding personal information in the collections.

## Industry Documents Library Collection Appraisal Checklist

1. What is the provenance of the records? Are the materials from an archival collection, litigation, FOIA request, or something else?
2. Is original order maintained? If archival materials, is original folder order retained? If litigation materials, are Bates numbers present?
3. What is the date range and extent of the collection?
4. What do the records document? What issues, people, companies, or trade associations are most obviously represented?
5. What is the relationship of the records to the current collection?
6. How do the records fit the goals of the Library and UCSF's mission?
7. How do the records fit the goals of the IDL collection policy?
8. Is the information or record duplicated elsewhere in the collection?
9. What is the storage medium? Is the information documented elsewhere in an easier-to-use format?

10. What are the costs of acquiring and preserving the records?
  - a. Identify any salient technical issues, such as poor OCR, need for extensive redaction, etc.
  - b. Identify document boundary definition needs: do the files need to be separated or combined into discrete documents?
  - c. Determine level of indexing: document-level or page-level? In-house or at vendor? Estimate cost.
11. How do legal guidelines apply to the records? Assess copyright risk: check for reprints/published material, number and identity of potential copyright holders, commercial value, likelihood of receiving significant attention. Can 17 U.S. Code Section 107 (Fair Use) or Section 108 (Reproduction by Libraries and Archives) be applied? Has permission been granted from holding library or archives?
12. How do records retention schedules apply to the records?
13. What restrictions apply to the records? Assess privacy, confidentiality, and personal safety issues: are there Social Security Numbers, individual postal or personal email addresses, or phone numbers? Is Protected Health Information (PHI) present? Is the source a whistleblower?
14. What archival value (informational, evidential, intrinsic, historical) do the records have?
15. What use is anticipated for the records?

## Policy Implementation and Revision Dates

The Industry Documents Library Collection Development Policy was drafted on December 13, 2017 and last updated on August 13, 2024. The Policy will be reviewed annually and revised as needed.

## References

RBMS Executive Committee, ACRL Board of Directors. (2020). *ACRL Code of Ethics for Special Collections Librarians*. Retrieved from Rare Books and Manuscripts Section, Association of College and Research Libraries, Division of the American Library Association: [https://rbms.info/standards/code\\_of\\_ethics/](https://rbms.info/standards/code_of_ethics/)

Society of American Archivists. (2020, August). *Core Values Statement and Code of Ethics*. Retrieved from Society of American Archivists:  
<https://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics>

Technical Subcommittee on Guidelines for Reappraisal and Deaccessioning. (2017, May). *Guidelines for Reappraisal and Deaccessioning*. Retrieved from Society of American Archivists:  
[https://www2.archivists.org/sites/all/files/GuidelinesForReappraisalDeaccessioning\\_2017.pdf](https://www2.archivists.org/sites/all/files/GuidelinesForReappraisalDeaccessioning_2017.pdf)