

# Digitizing Historical Medical Collections At the National Library of Medicine



Walter Cybulski

Preservation & Collection Management Section

[cybulskw@mail.nlm.nih.gov](mailto:cybulskw@mail.nlm.nih.gov)

# The National Library of Medicine has a Congressional mandate to preserve the biomedical literature.



*Edwin Smith Surgical Papyrus* (17th century B.C.). Earliest recorded reference to the brain.



**Imhotep** (fl. 27<sup>th</sup> century BC) reputed to be the author of the Smith papyrus. It was said that he extracted medicine from plants. The papyrus addresses a variety of injuries, including 27 injuries to the head.

# NLM's electronic version of the Smith papyrus replicates the "look and feel" of a scroll.



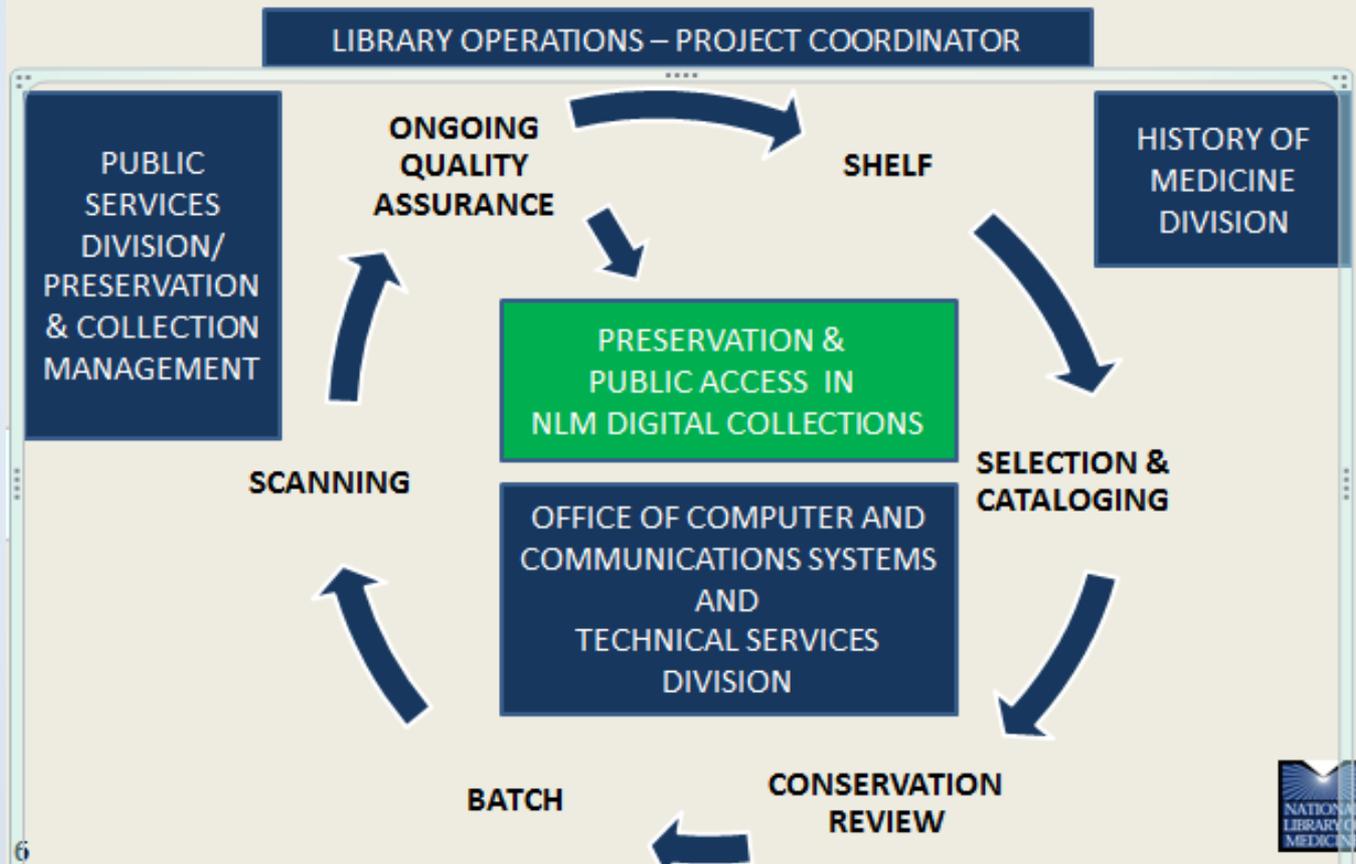
<http://archive.nlm.nih.gov/proj/ttp/flash/smith/smith.html>

**In order to connect today's researchers to the past, there must be a past to connect to.**



# Digitization Projects Workflow

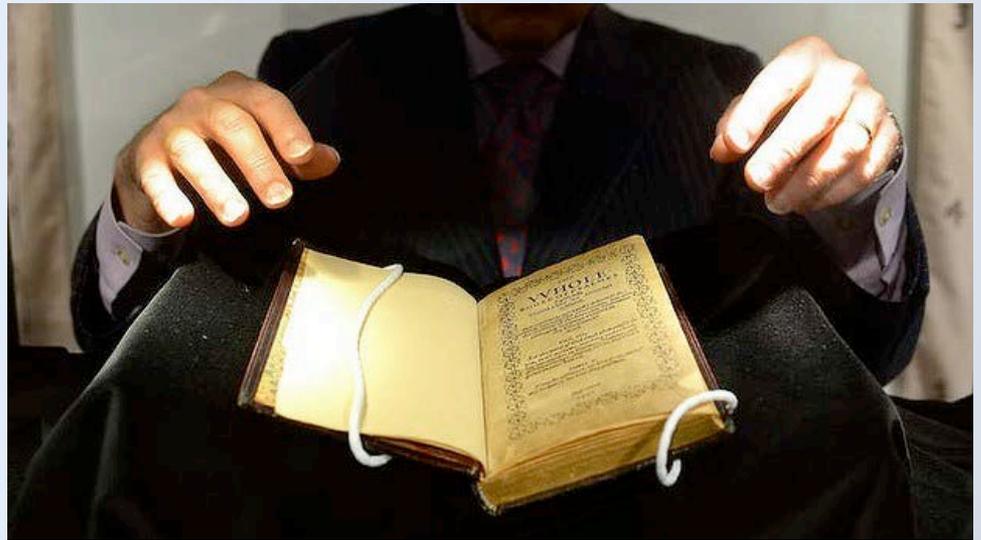
## DAILY PRODUCTION CYCLE



# Collection materials are handled with appropriate care.



Risks of damage are minimized from shelf to scanner and back.





## The Kirtas KABIS III Robotic Scanner, B3 Level, NLM.

Foot pedal for manual scanning.

Cushioned Floor mat for operator comfort.



**2-terabyte  
data backup  
drive.**

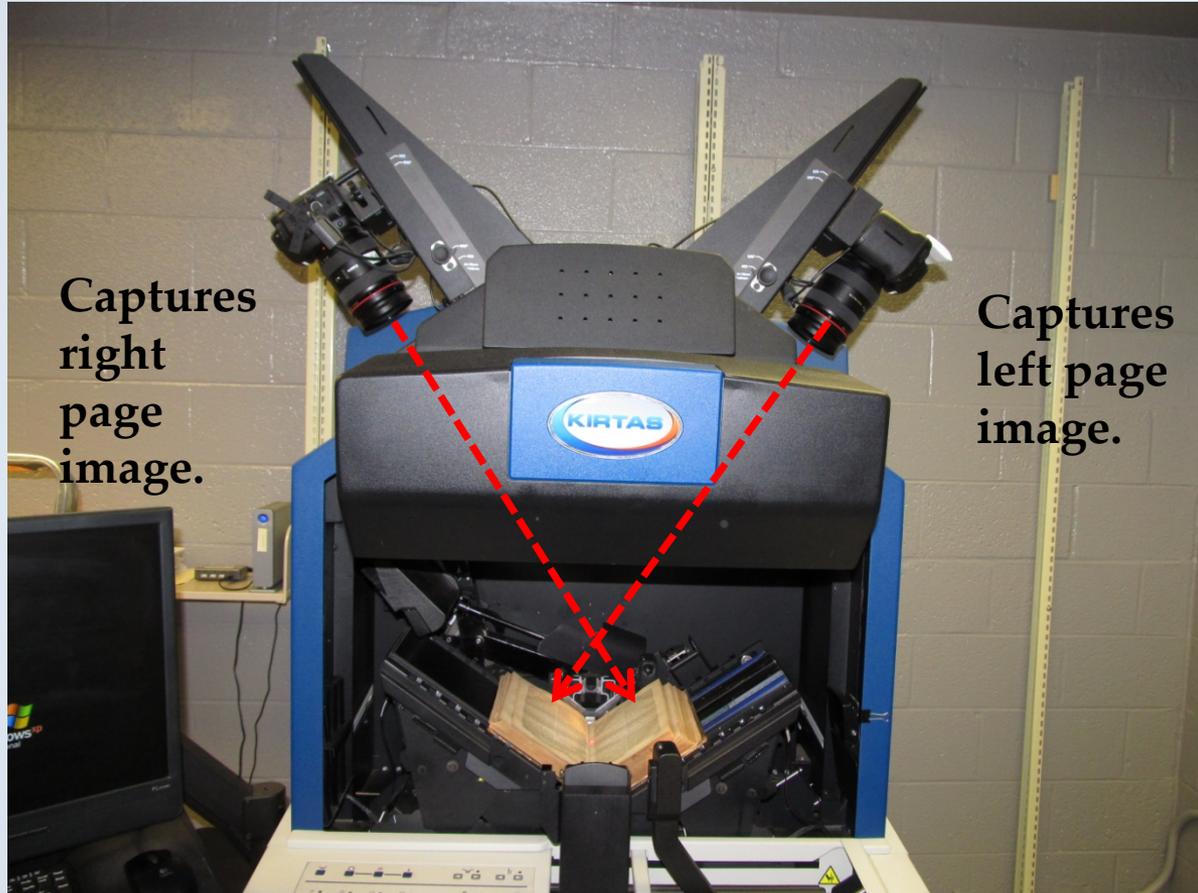


**Backup  
power  
supply.**



**KABIS III server and 2 PCs.**

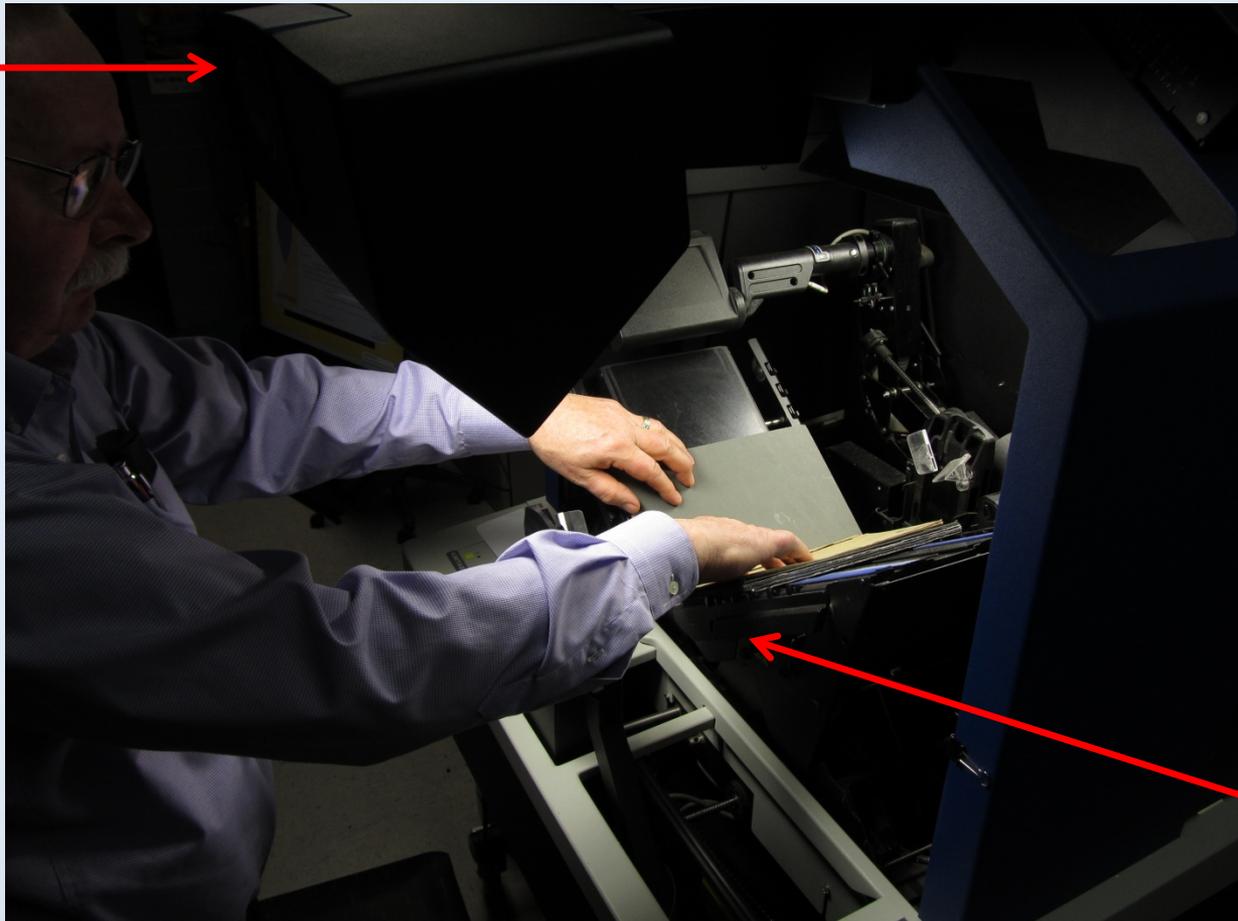
# KABIS III Cameras: 1 Click, 2 Pix



Two Canon® 21.1 megapixel EOS digital SLR cameras with 100 mm Macro lenses.

# Scanner Operation Ergonomics

LIGHT HOUSING



REPEATED BENDING FORWARD



REPEATED REACHING IN TO POSITION BOOK

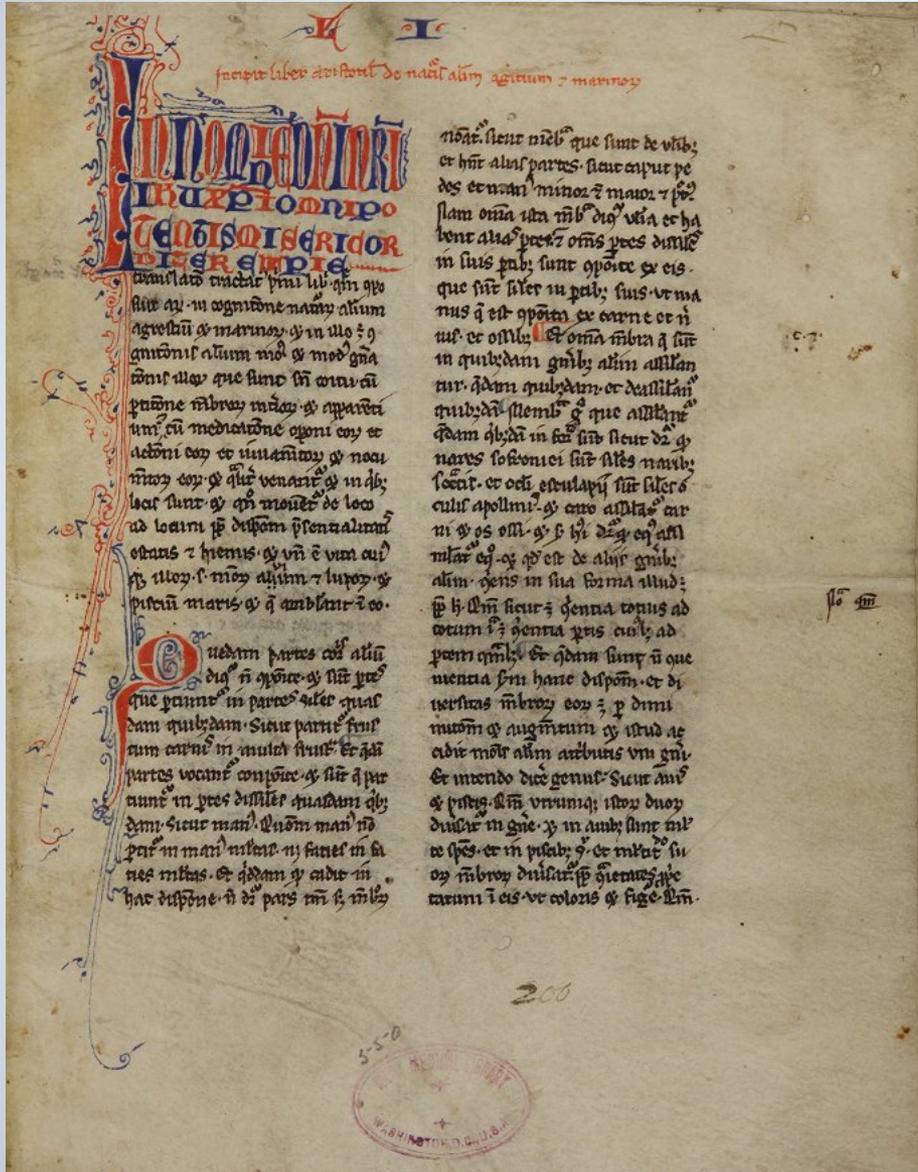




**Manuscripts** written onto animal skin (vellum) are very durable.

Vellum sheets can be difficult to scan.



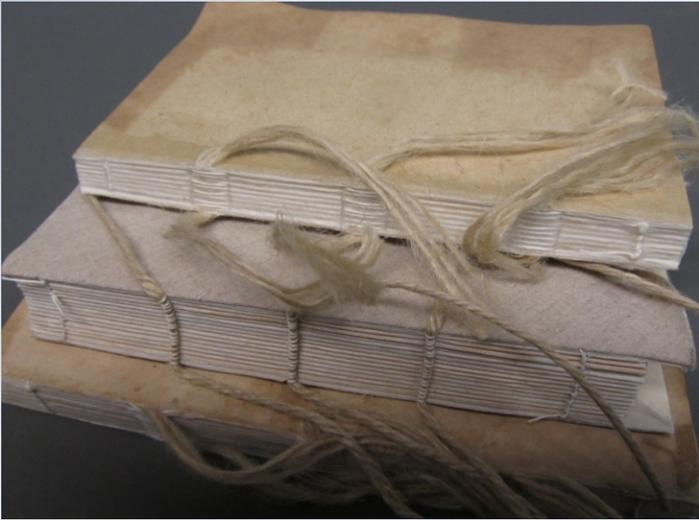


## Aristotle's *De animalibus*

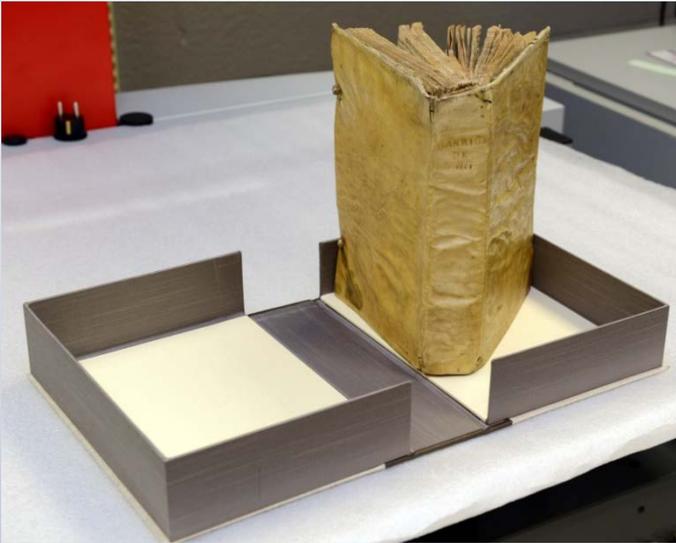
### *Liber Aristotelis de natura animalium agrestium et marinarum*

*"Aristotle's book about the nature of wild and marine animals"*

Manuscript codex created in France, ca. 1235-1245?



**Conservation repair.**



**Custom-built enclosure.**



**Book placed in cradle for scanning.**



Barrios, Juan de (born 1562)

*Verdadera medicina, cirugía, y astrologia: en tres libros dividida*

Con licencia y priuelegio ympreso en Mexico : Por Fernando Balli, 1607.

Tulley, 1701.

Stimber A N 1

# ALMANACK

For the Year of our Lord 1701.

Being First after Leap-year, & from the Creation, 5650.  
And from the Discovery of America by  
Chr. Columbus, 209. And of the Reign of our Gra-  
cious Sovereign, K. WILLIAM the Third (which  
began Febr. the 13th, 1688, 9.) the 13th year.

Wherein is Contained the Lunations, Courses,  
Spring Tides, Planets, Aspects and Weather,  
the Rising and Setting of the SUN, together  
with the Sun & Moons place, and time of Full  
Sea or High Water, with an account of the  
Eclipses, & other matters useful & necessary.

The Vulgar Notes of this Year are,  
Golden Number 11 } Cycle of the Sun 2  
The Epact 1 } Dominic. Letter E

Calculated for and Fitted to the Meridian of Boston  
in New-England, where the North Pole is  
Elevated 42 gr. 30 min. But may indiffe-  
rently serve any part of New-England.

By JOHN TULLEY.

Licensed by Authority.

Boston, Printed by B. Green, & J. Allen Sold at the  
Printing-House at the South End of the Town 1701.

of the Eclipse as it will appear in the East after Sun set.

The beginning 5 hours  
The middle 6 hours  
The End 8 hours  
The whole time of  
first to last will be  
The digits Eclipsed 6



30 min. } In the  
47 min. } Evening  
4 min. }  
Continuance from  
2 hours & 34 mi  
& 6 min. 35 Sec

John Tulley

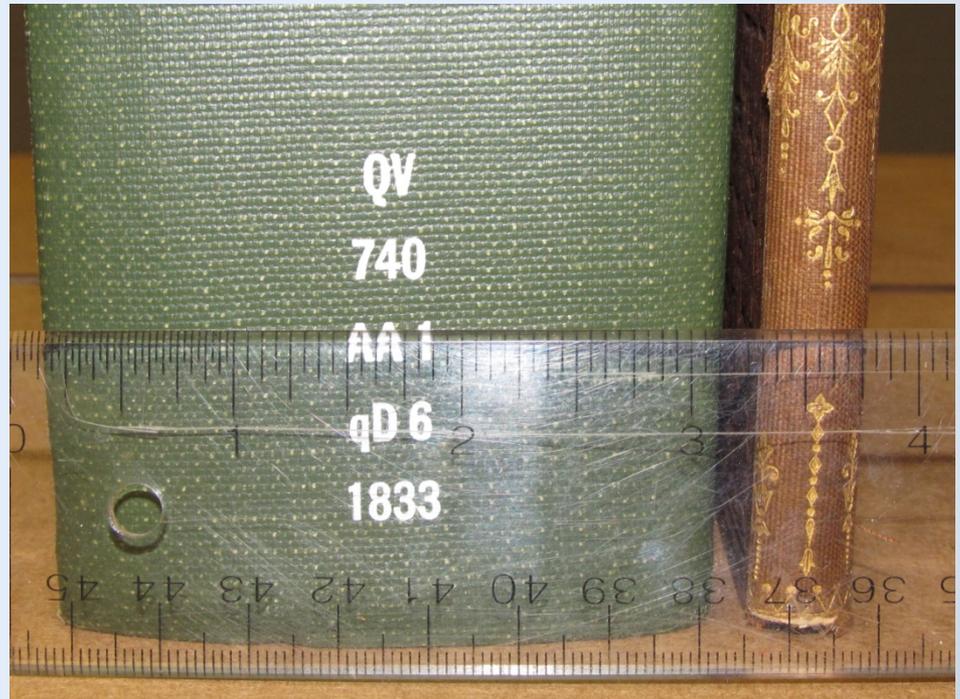
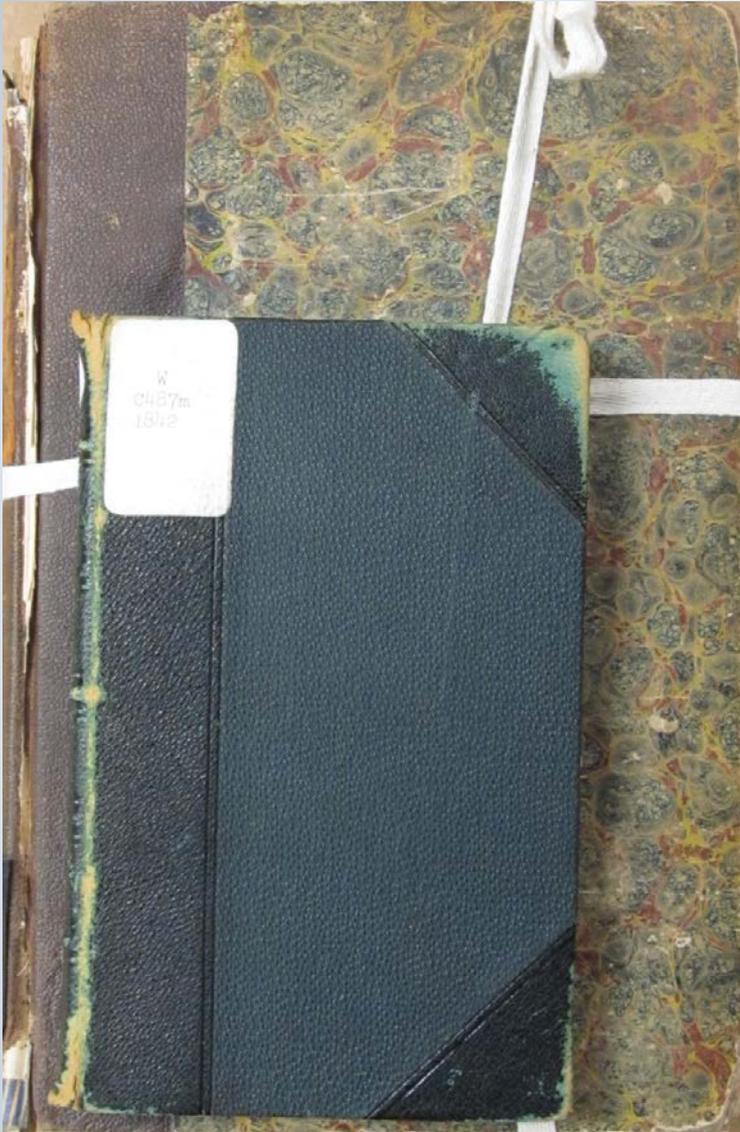
An almanack for the year of our Lord 1701

Boston : Printed by B. Green, & J. Allen  
sold at the printing-house at the south  
end of the town, 1701 [i.e., 1700?]

I'm safe while moving in my proper Sphere,  
In Plowing, Planting, there's no Treason there.

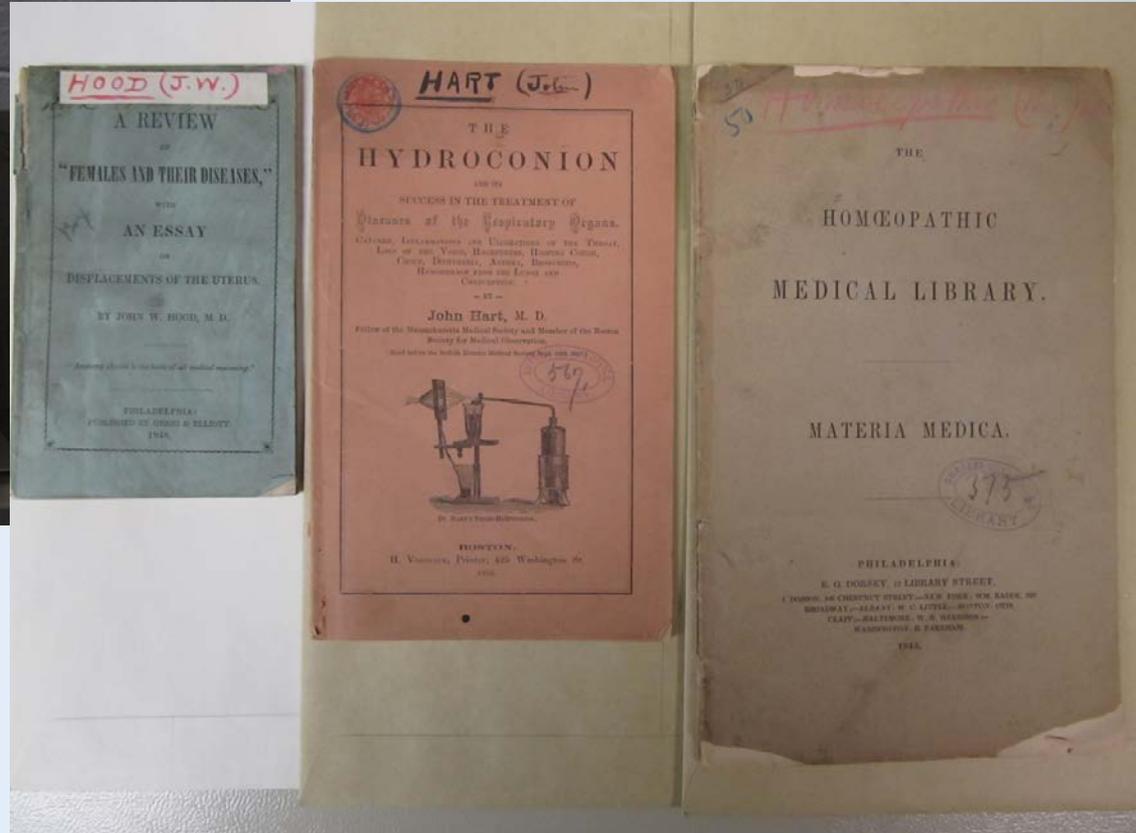


**Books vary in page height, page width and volume thickness.**



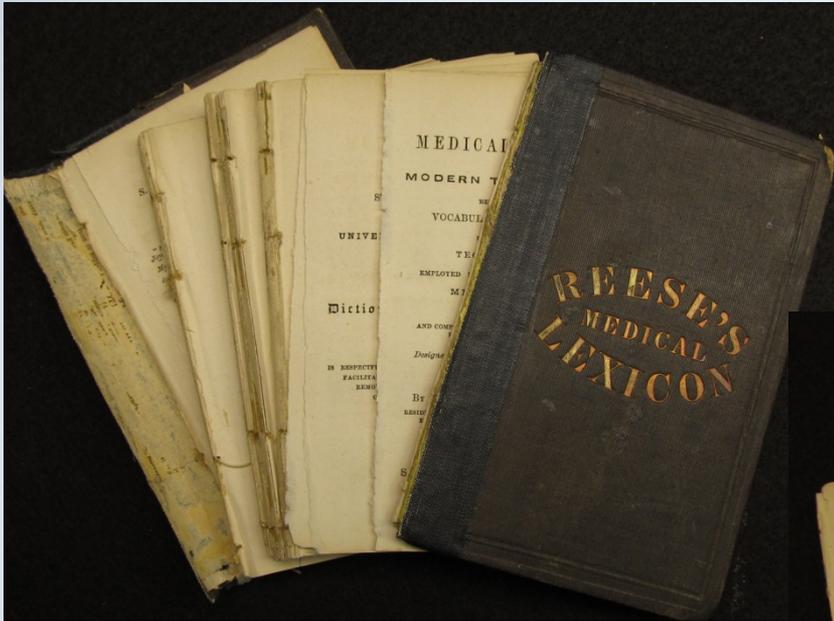


All pamphlets are scanned by hand.

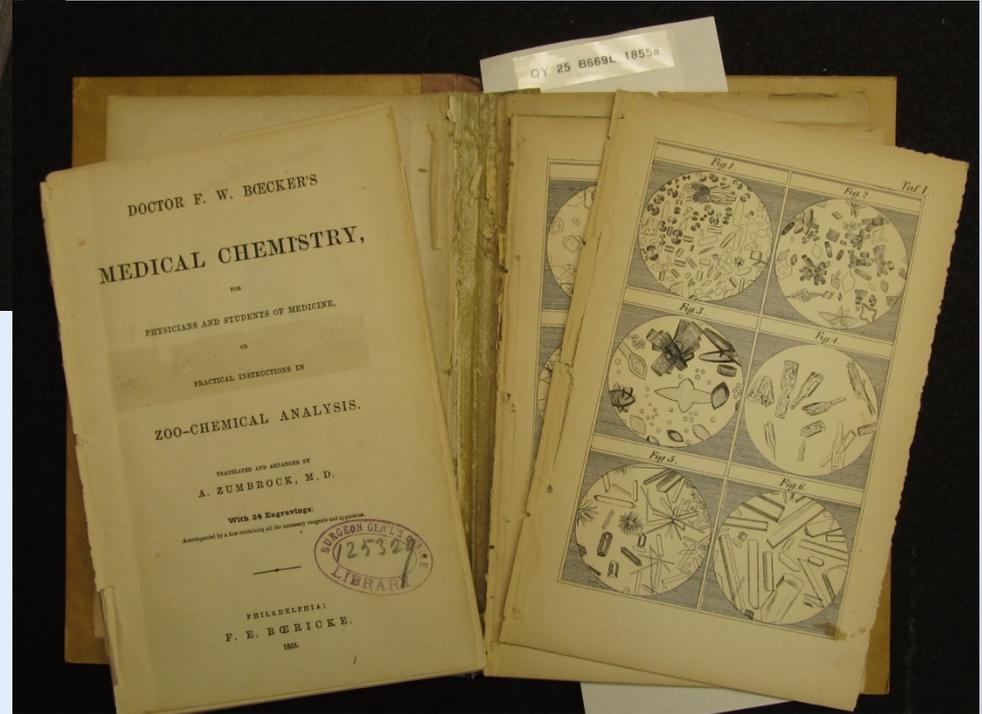


# Scanning Challenges

Broken bindings; loose and detached pages.



Brittle paper.



# Other Factors that Determine Method of Scanning (Robotic or Manual)

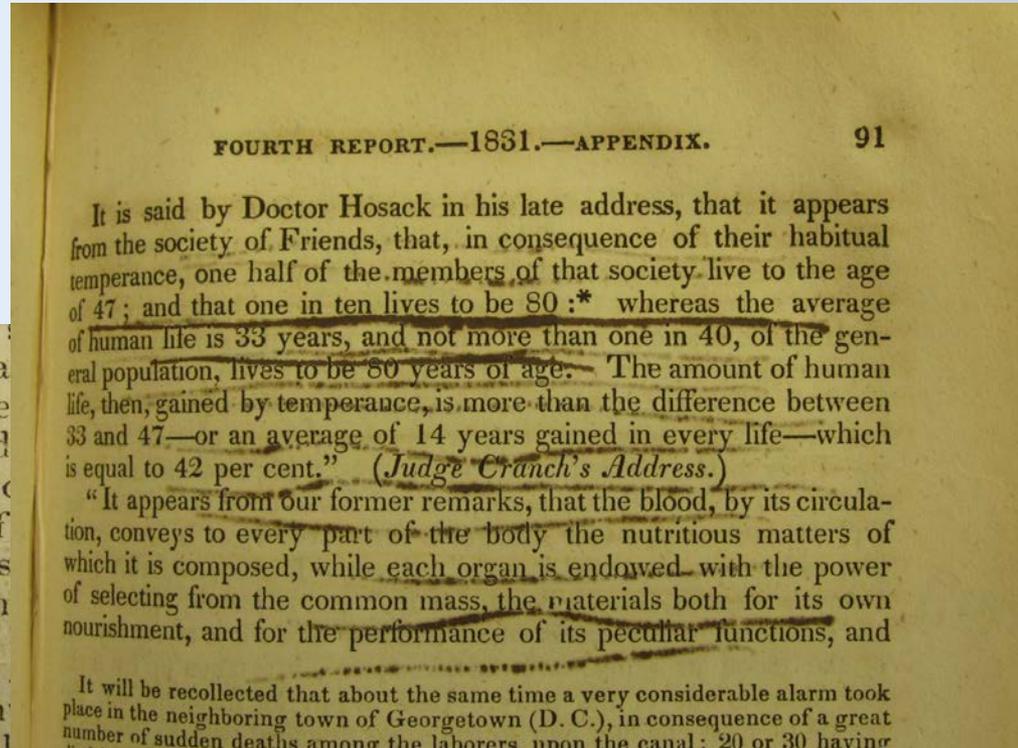
- Paper thickness or thinness.
- Pages unevenly trimmed.
- Tight binding (loss of text along inner margins).



# Printing Defects and Other Anomalies.

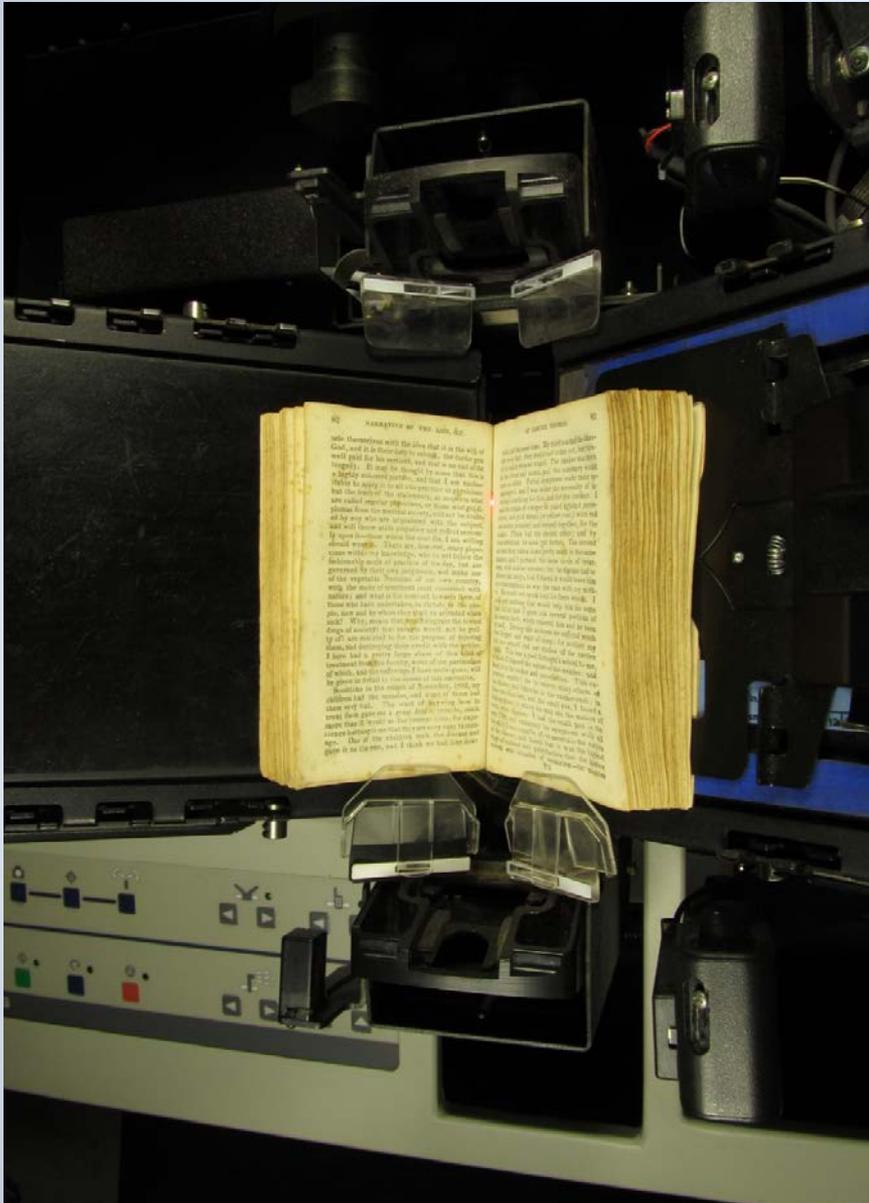
Page creased in printing.

... and those who sell it are all accessory, and as  
... held responsible at the divine tribunal. There wa  
... when the owners did not know the dangerous and de  
... qualities of this article—when the facts had not been d  
... and published, nor the minds of men turned to the subjec  
... they did not know that it caused such a vast portion of  
... and wretchedness of the community, and such wide-s  
... resolution to the temporal and eternal interests of n  
... though it then destroyed thousands, for both worlds,  
... the men who sold it, was comparatively small.  
... they sin against light, pouring down upon them with un  
... rightness; and if they know what they do, and in full view o  
... consequences, continue that work of death—not only let the  
... poison go out, but furnish it, and send it out to all who are dis  
... posed to purchase,—it had been better for them, and better fo  
... many others, if they had never been born. For,  
1. It is the selling of that, without the use of which, nearly al  
... the business of this world was conducted, till within less than  
... three hundred years; and which of course is not *needful*.  
2. It is the selling of that, without the use of which, nearly al

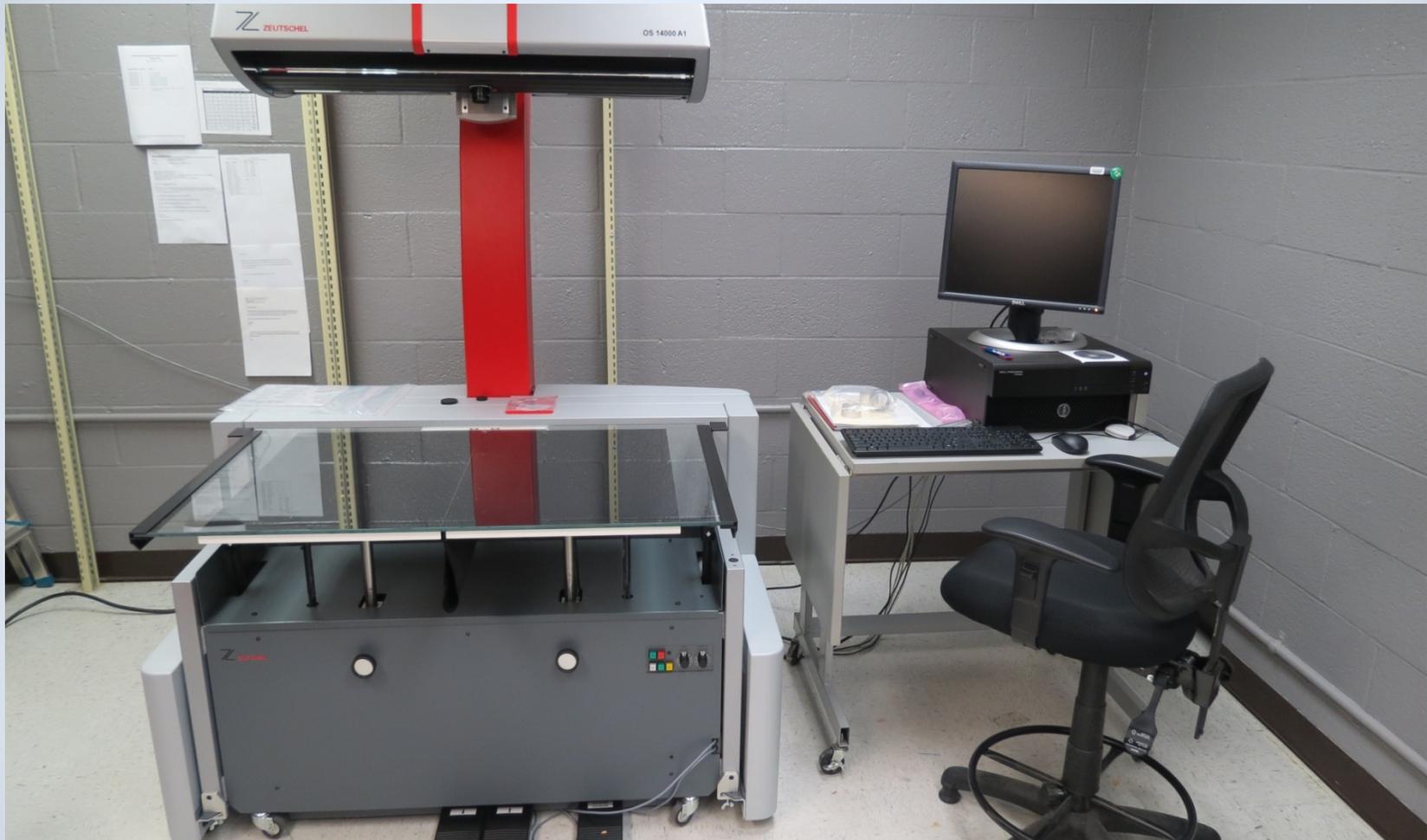


Heavy ink underlining.

**Small books require constant re-positioning when scanned.**



# Zeutschel OS14000 A-1 Overhead Scanner



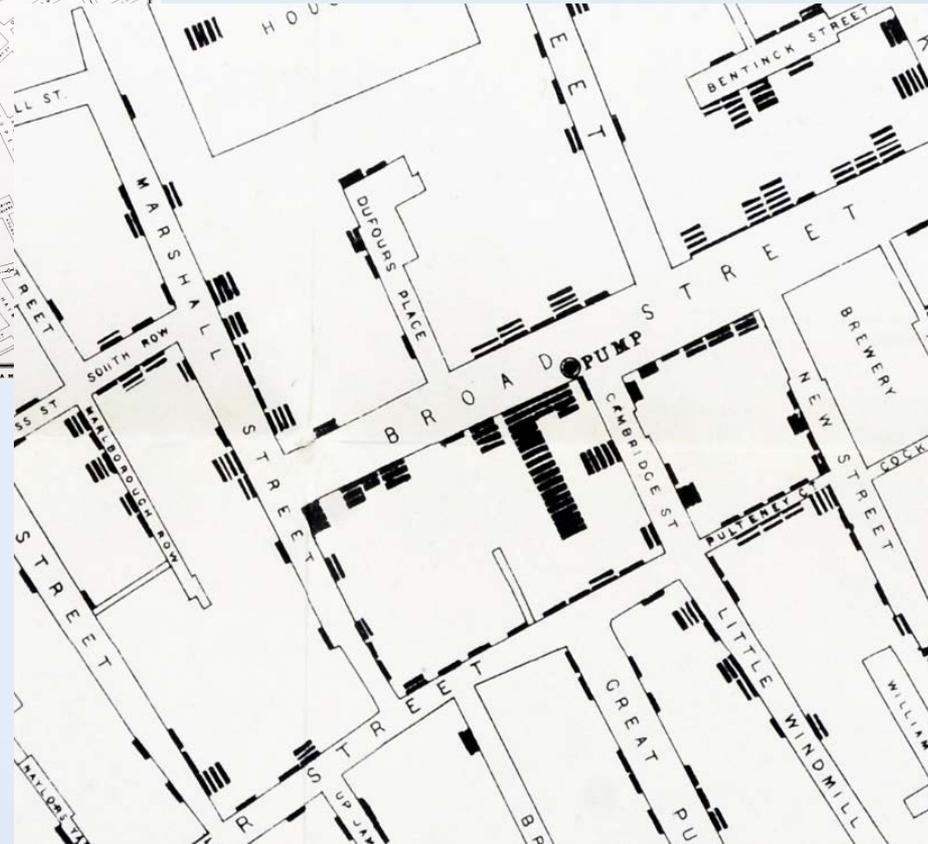




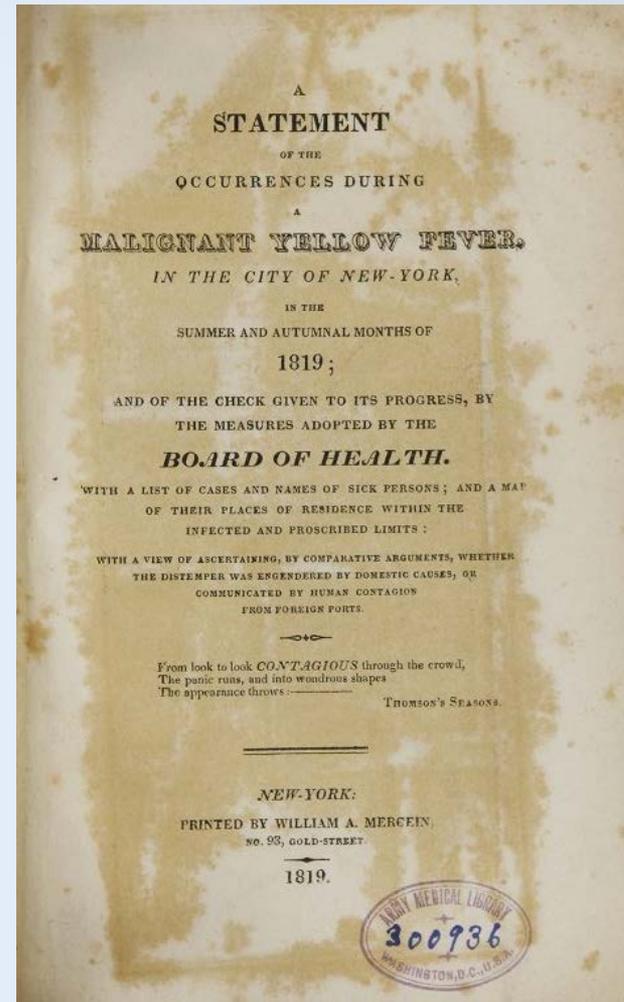
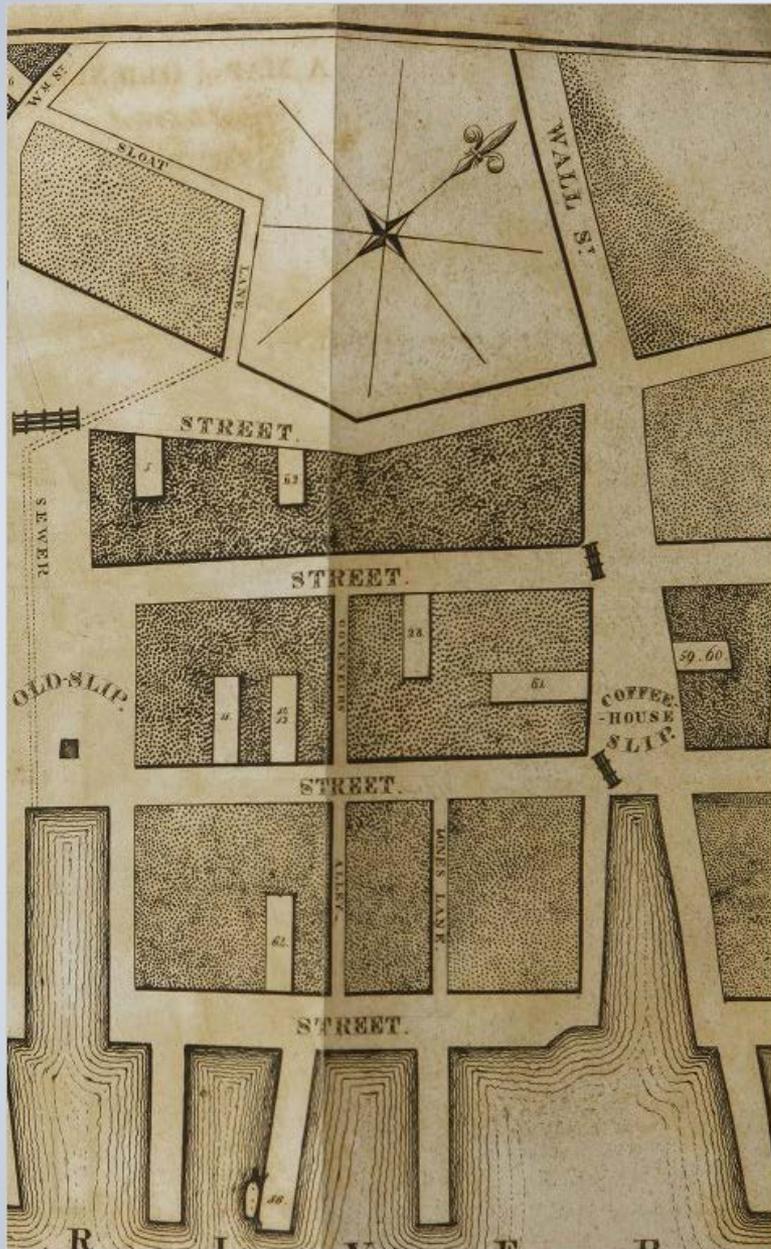
Snow, John (1813-1858)

On the Mode of Communication  
of Cholera

London: J. Churchill, 1849

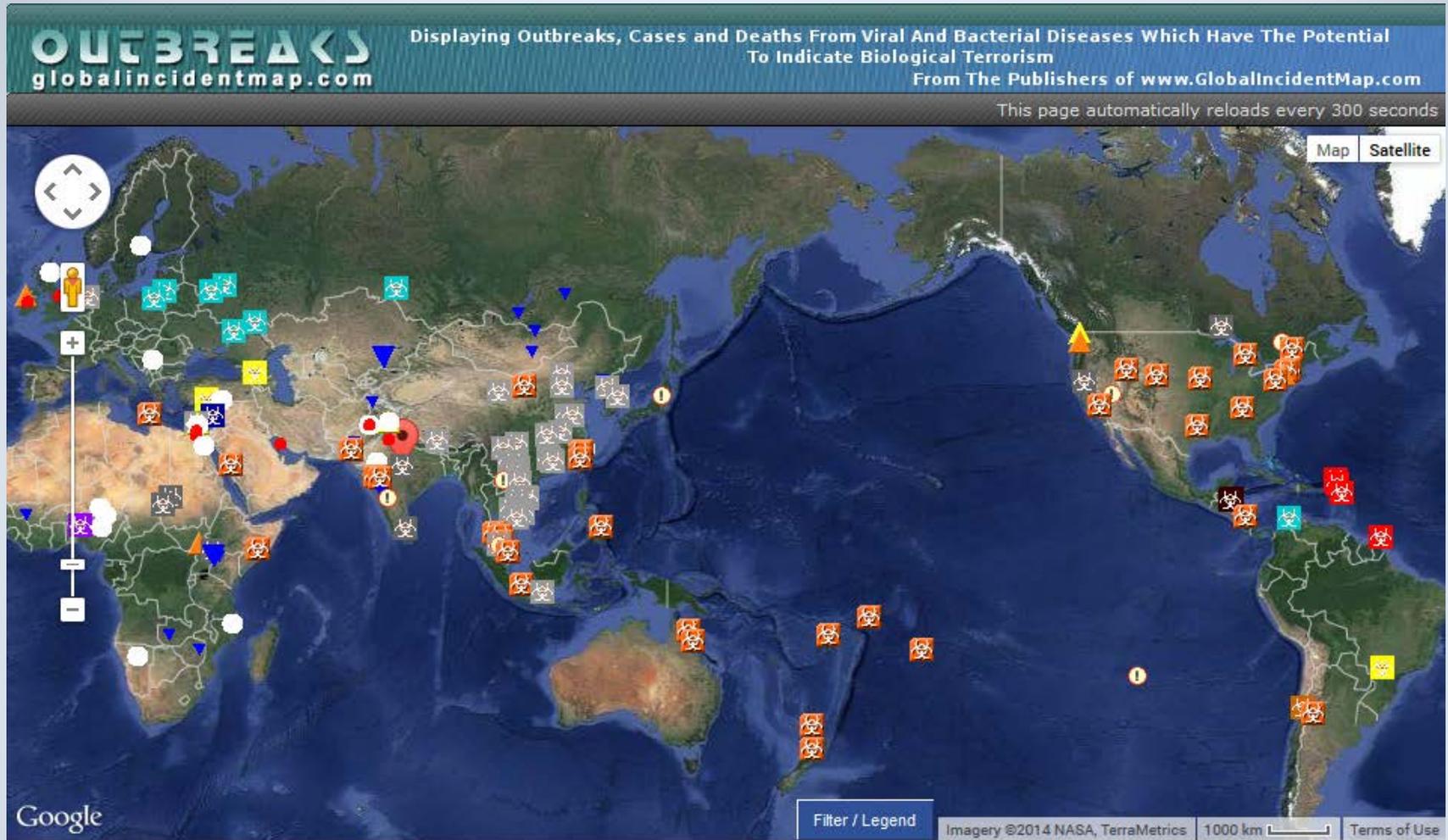


<http://resource.nlm.nih.gov/0050707>

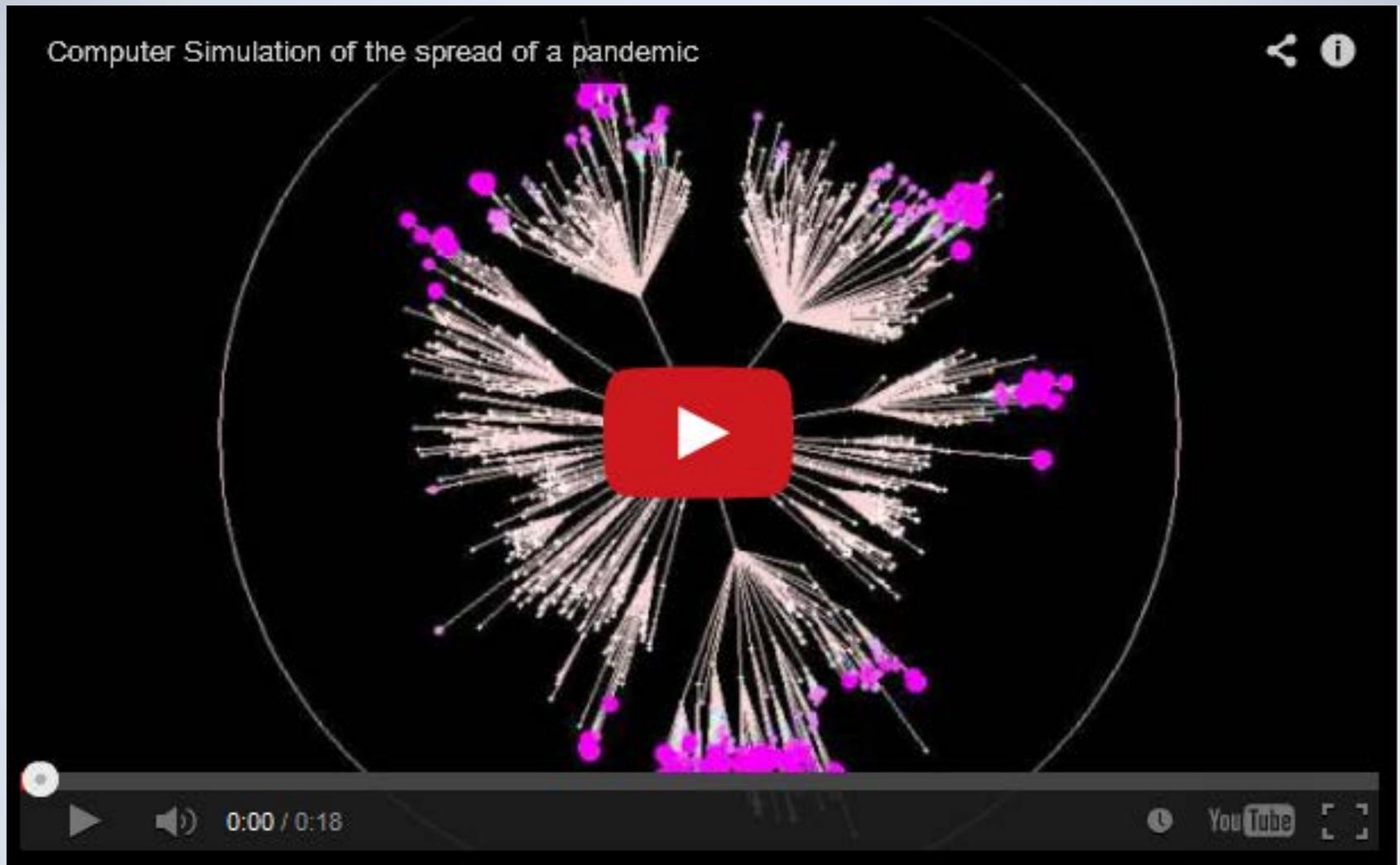


Pascalis-Ouvière, Félix, 1762-1833.  
A statement of the occurrences during  
a malignant yellow fever: in the city of  
New-York, in the summer and  
autumnal months of 1819.





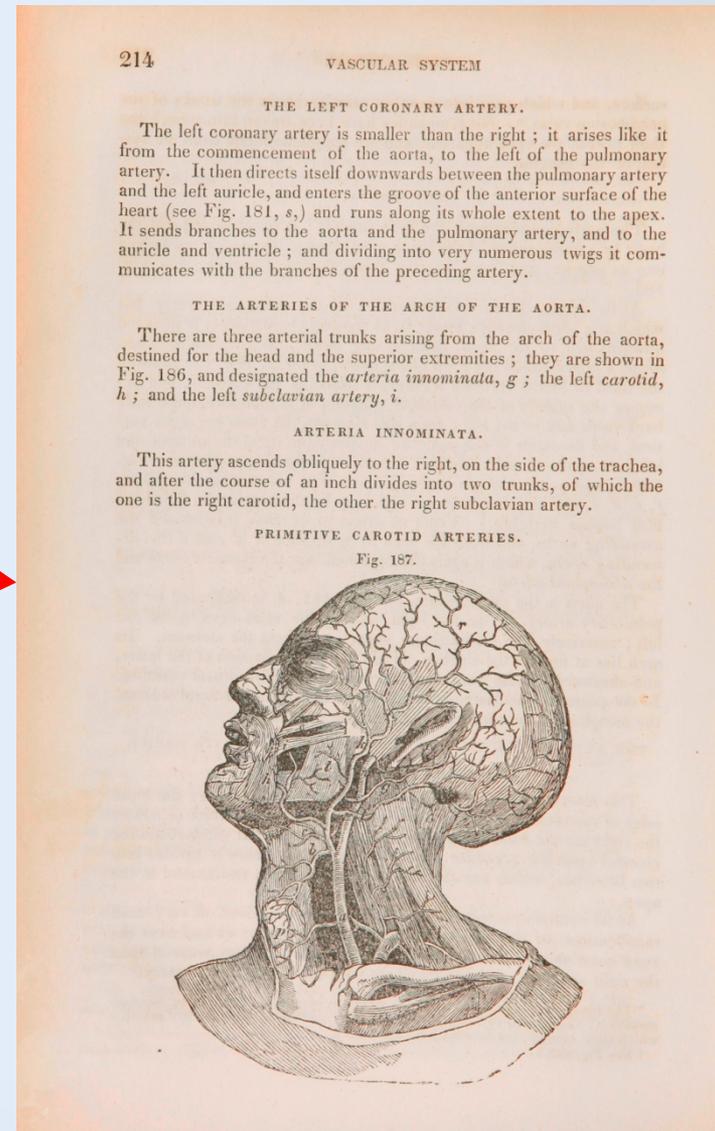
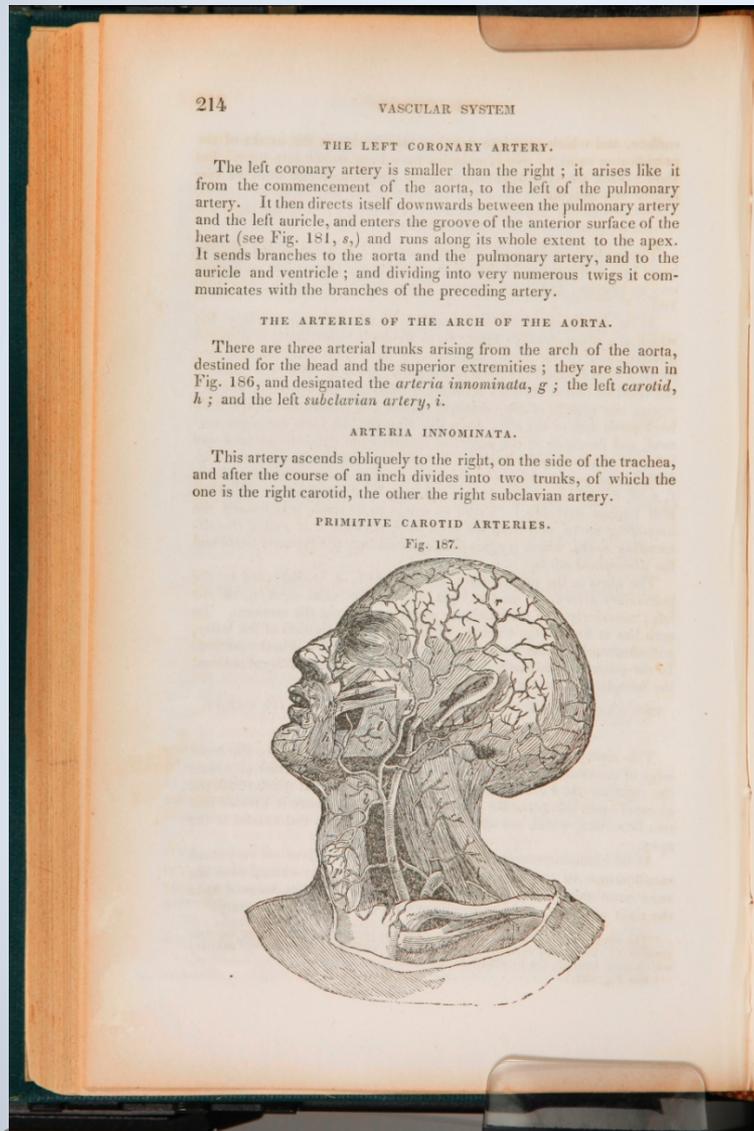
Outbreaks, cases and deaths from viral and bacterial diseases which have the potential to indicate biological terrorism. [www.GlobalIncidentMap.com](http://www.GlobalIncidentMap.com)



**Outbreak! Watch How Quickly An Epidemic Would Spread Across The World**

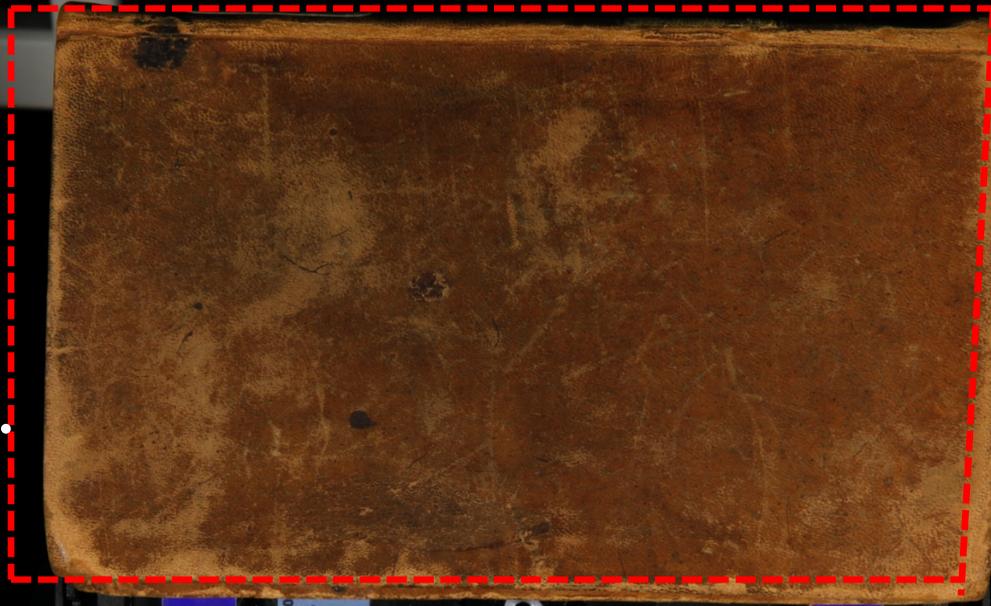
<http://www.fastcoexist.com/3023416/outbreak-watch-how-quickly-an-epidemic-would-spread-across-the-world>

# Post-Scan Image Processing

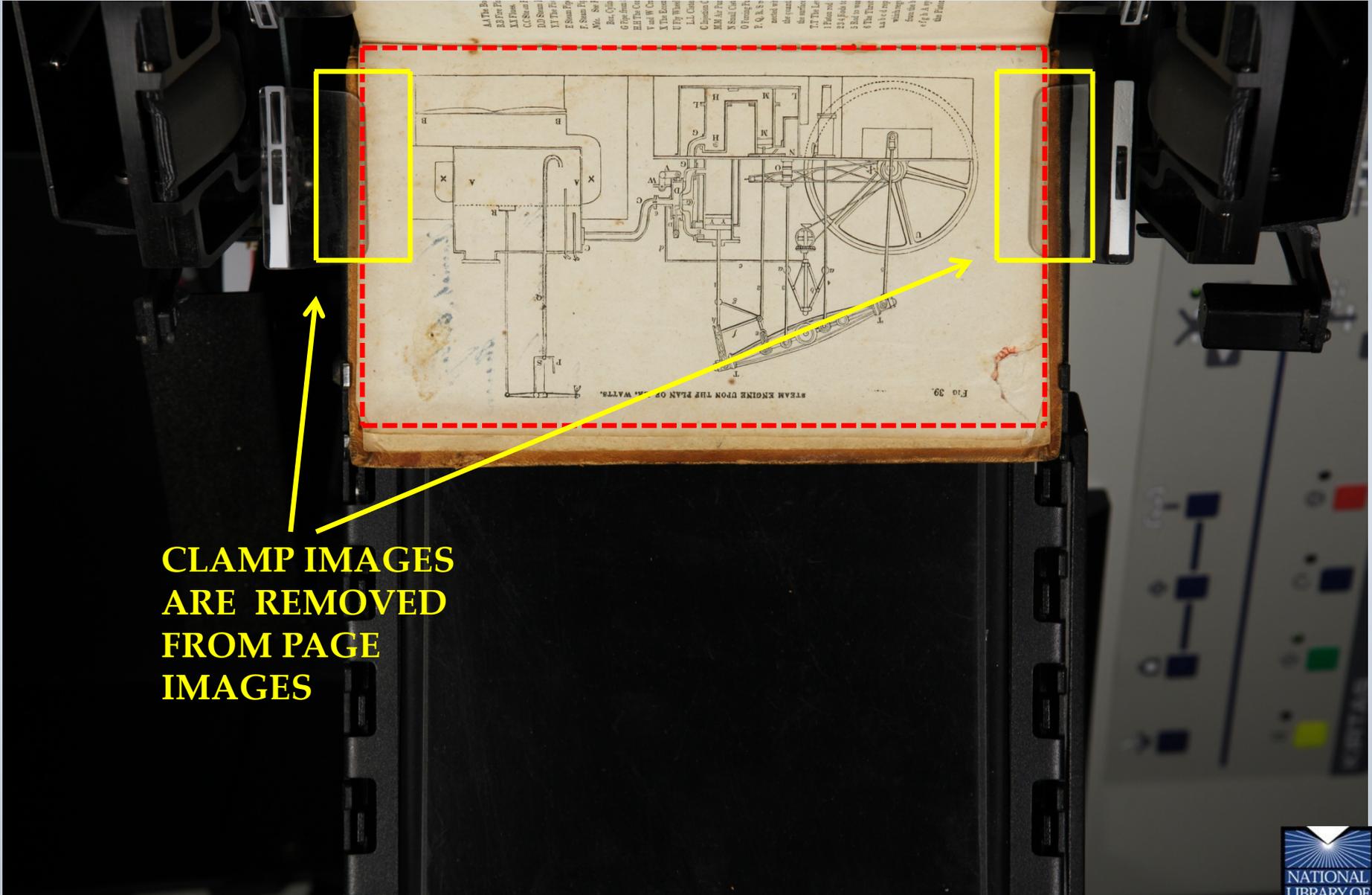


# Automated Image Cropping.

Images are cropped to remove background.



# Automated Clamp Removal



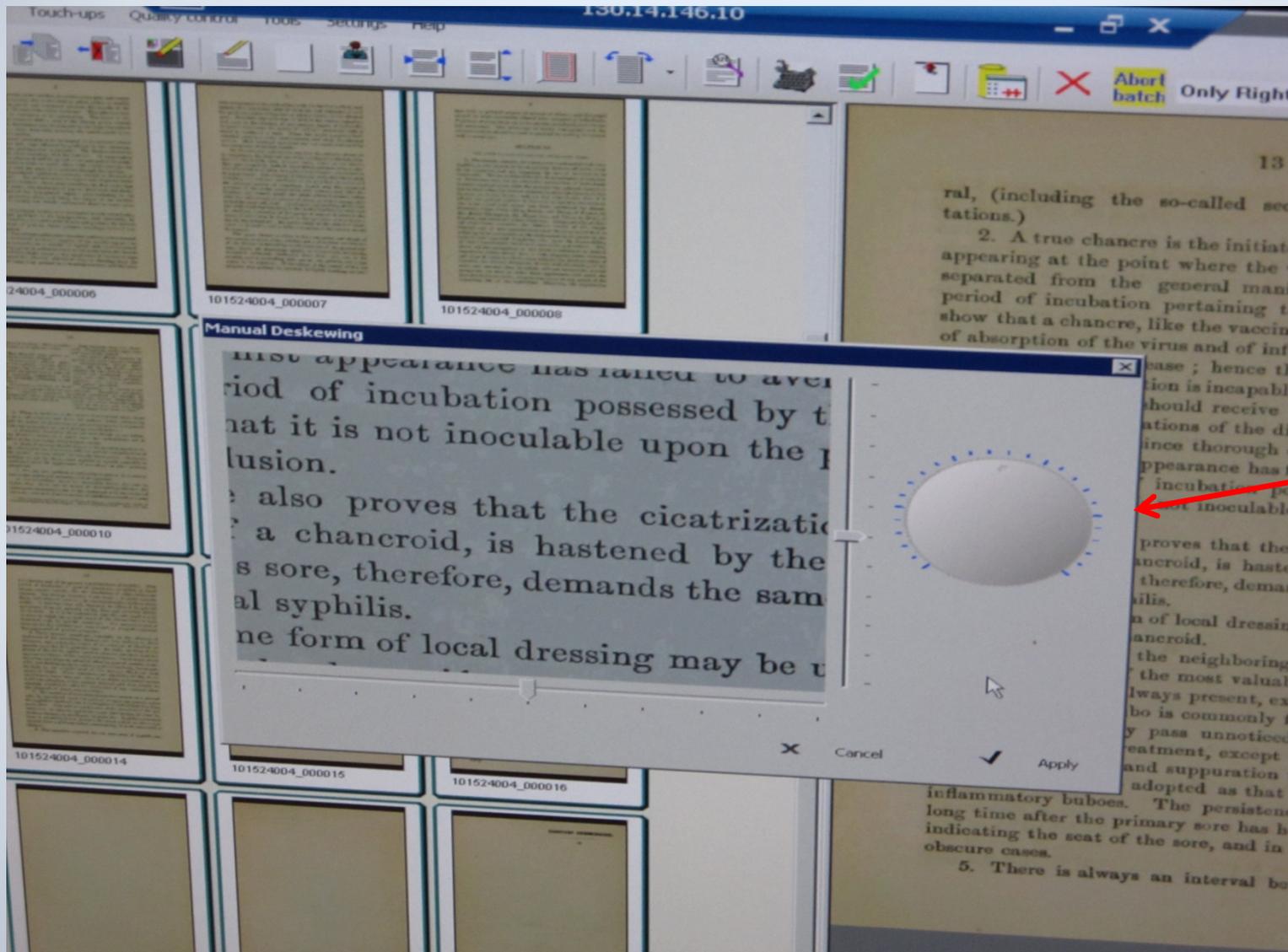
**CLAMP IMAGES  
ARE REMOVED  
FROM PAGE  
IMAGES**

# Text Skewed on Page

11

to inculcate a doctrine, which it might not suit his taste to practice. But in a vast empire, held by the frail tenure of opinion and especially, where the current of religious prejudices, Brahmin as well as Moslem, ran strongly against intoxication, it was soon found necessary, from imperious motives of policy, rather than of health, to discourage every tendency towards the acquisition of such dangerous habits. Happily what was promotive of our interest, was preservative of our health, as well as conducive to our happiness. And the general temperance in this respect, which now characterizes the Anglo Asiatic circles of society, as contrasted with An-

# De-skewing: Operator's View



Dial  
turned to  
rotate text.

# Quality Control

| CODE                          | UID      | Batch Number   | Received Date                           | Title                    |
|-------------------------------|----------|--|---|--------------------------|
| MHL                           | 2174045R | HMDBATCH0313   | 01/26/2012                              | Introductory diseases of |
| <b>IMAGES</b>                 |          | <a href="#">56 file(s) to view</a>   |   |                          |
| <b>METADATA</b>               |          | <a href="#">56 file(s) to view</a><br><a href="#">1 file(s) to view</a><br><a href="#">1 file(s) to view</a> |   |                          |
| ALTO                          |          |  |   |                          |
| METS                          |          |  |   |                          |
| MARC                          |          |  |   |                          |
| <b>OCR</b>                    |          | <a href="#">56 file(s) to view</a>   |   |                          |
| <b>PDF</b>                    |          | <a href="#">1 file(s) to view</a>  |   |                          |
| <b>WORD</b>                   |          | <a href="#">1 file(s) to view</a>  |   |                          |
| <b>Normalization</b>          |          | <a href="#">View normalization log file</a>  |   |                          |
| <b>Errors Found by System</b> |          |  |   |                          |
| <b>0 Bytes</b>                |          | <b>1 error found</b>   | <a href="#">View error log file</a>     |                          |
| <b>Checksum</b>               |          | 0 error found  |   |                          |
| <b>XML Schema/Syntax</b>      |          | <b>2 errors found</b>  | <a href="#">View error log</a>          |                          |
| <b>Image File Format</b>      |          | 0 error found  | <a href="#">View all FITS log files</a> |                          |
| <b>Number of Files</b>        |          | 0 error found  |   |                          |
| <b>Filename</b>               |          | 0 error found  |   |                          |

# Output Files



- JPEG for each page  
24-bit RGB color, 400 dpi



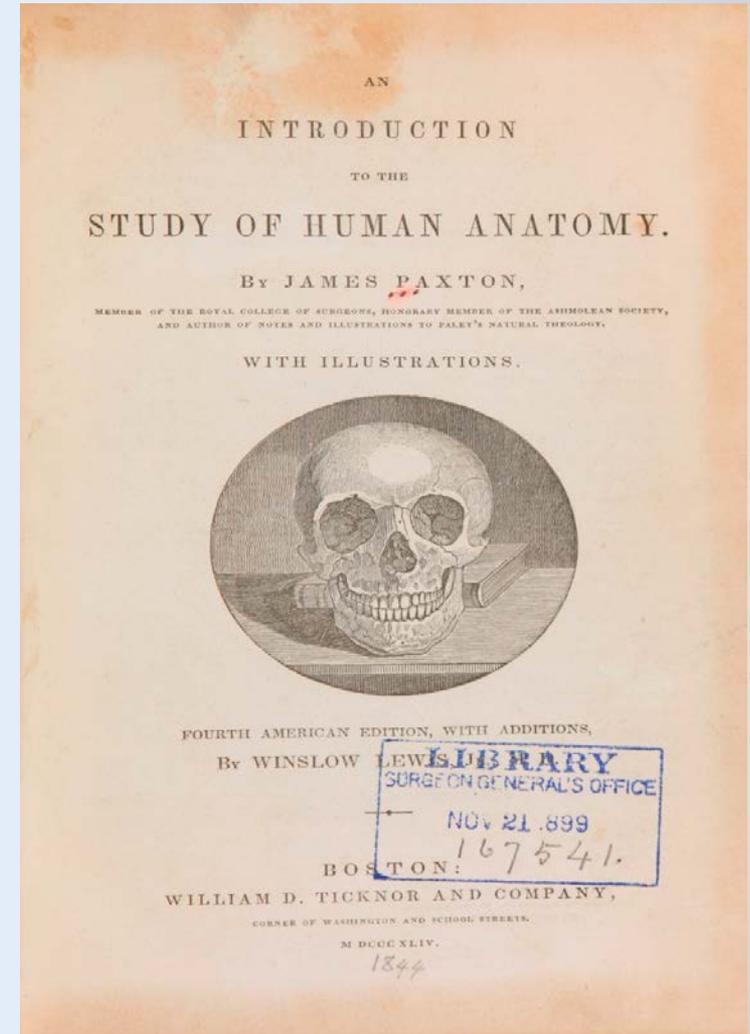
- Single PDF for entire book



- OCR .txt file for each page



- Metadata (technical and descriptive data required for preservation and future use) for all files



# Ingest into NLM Digital Collections

## Digital Projects

### Historical and contemporary sources for research and education



**Medicine in the Americas 1610-1920** features American medical books made possible in part through NLM's participation in the [Medical Heritage Library](#), a digital curation collaborative among the world's leading medical libraries.



**Films and Videos** encompasses historical public health and training films, with searchable transcripts.



**Profiles in Science** contains archival collections of twentieth-century leaders in biomedical research and public health.



**Historical Anatomies on the Web** features images from important anatomical atlases.



**Cholera Online** contains English-language monographs dating from 1817 to 1900 addressing the cholera pandemics of that period.



**Turning the Pages** allows users to leaf through medical masterpieces from past centuries.

### Online Exhibitions



**DNA to Beer: Harnessing Nature in Medicine and Industry** is an exhibition exploring some of the processes, problems, and potential inherent in technologies that use life.



**Surviving and Thriving: AIDS, Politics, and Culture** is an exploration of the rise of AIDS in the early 1980's and the evolving response to the epidemic over the 30 years that followed.



**Native Voices: Native Peoples' Concepts of Health and Illness** explores the interconnectedness of wellness, illness, and cultural life of Native peoples.



**Every Necessary Care and Attention: George Washington and Medicine** explores how, alongside the traditional demands of family life, and political and military leadership, George Washington focused considerable attention on the health and safety of his family, staff, slaves, and troops.



**"And there's the humor of it:" Shakespeare and the four humors** explores how Shakespeare used the four bodily humors to enliven his



Scanned to date:

10,983 books

2,218,423 images (pages and covers)

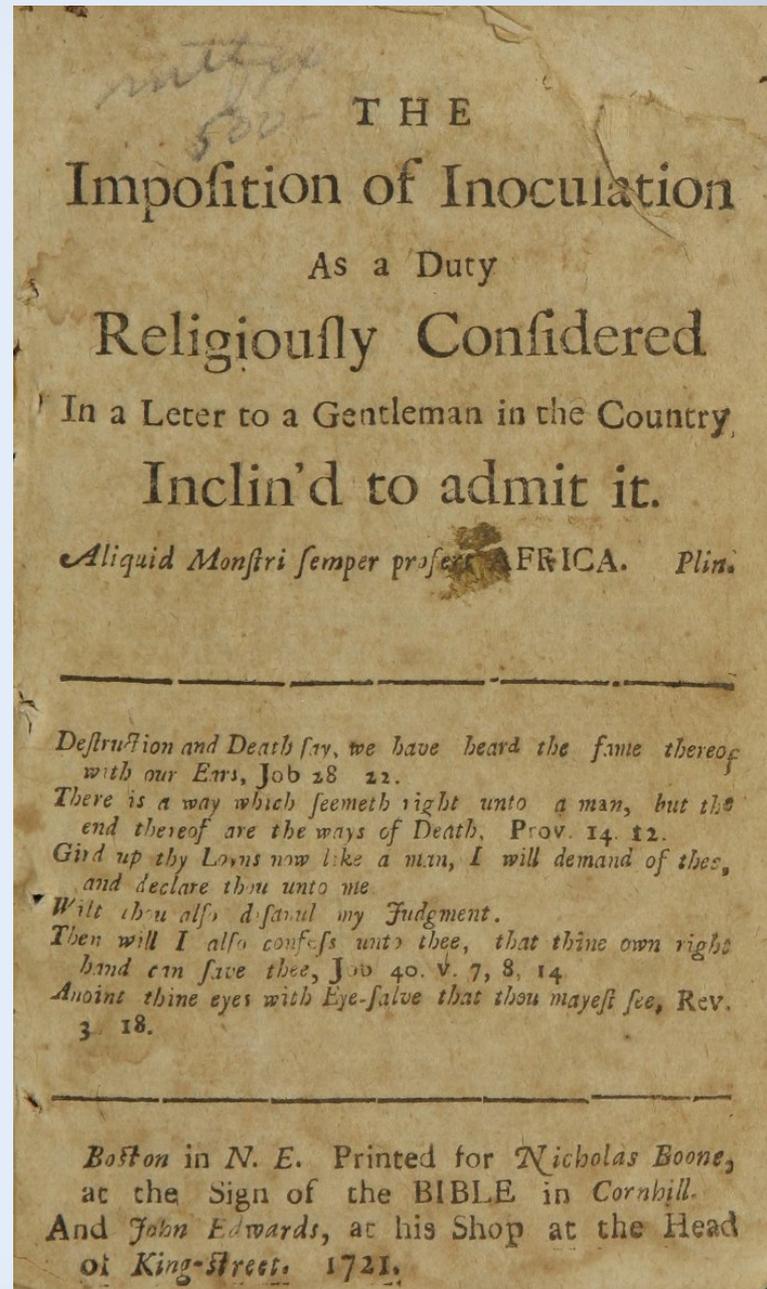
Per book average is just over 200 pages.



# Access to Scanned Titles via Google

[The imposition of inoculation as a duty religiously considered in a ...](#)

[.nlm.nih.gov/muradora/objectView.action?...nlm:nlmuid...](#)



# NLM collections on Internet Archive

INTERNET ARCHIVE [The imposition of inoculation as a](#) Search inside GO

A  
L E T T E R  
To a Gentleman in the Country.

S I R,

**A**T our last Meeting, we fell upon the common Topick, *Inoculation*. But the time not permitting us to come to any Resolution as to its *Lawfulness*. I have according to your Request, and my Promise at parting lent you my tho'ts thereon; and notwithstanding I know you to be a great Admirer of this new Practice; & with many (*Seduced by its Suppos'd success*) inclin'd to believe it lawtul. I am still perswaded you are Master of so much *Temper and Integrity*, as not to oppose the *Truth*, because it obligeth you to think otherwise than you have done, or requires of you to forsake a *Practice*, which tho' it may be attended with abundance of good Intencions has no other than a *Criminal Foundation* to support it.

Since *Errors in Judgment*, many times occasion *Errors in Practice*; it should be our endeavour to support our Judgment, with all the

Navigation icons: Home, Full Screen, Grid, Zoom In, Zoom Out, Previous, Next



