

Biometrics

COMP 388-002/488-002 Computer Science Topics

Daniel Moreira
Fall 2025



LOYOLA
UNIVERSITY CHICAGO

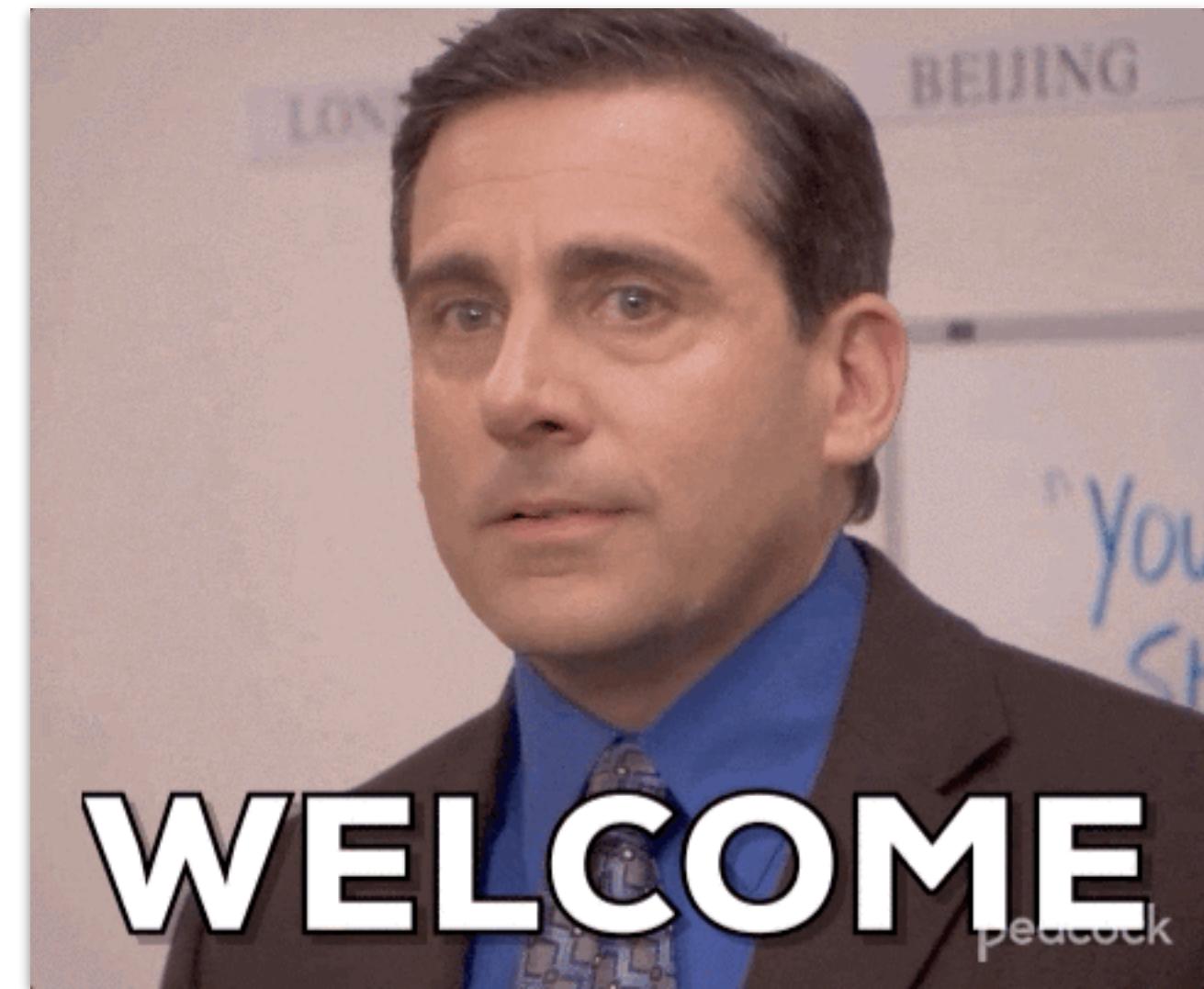
Welcome

COMP 388-002/488-002 Computer Science Topics Biometrics

Daniel Moreira

Contact: dmoreira1@luc.edu and Teams

Office: Doyle Center 310



Course Hours

Lectures: MON and WED, 4:15 to 5:30 PM, Information Commons 112

Office Hours: TUE evenings and FRI mornings

Doyle Center 310 or Zoom, by appointment (<https://tinyurl.com/yv76kpb>)

Communication

Sakai: <https://sakai.luc.edu/x/wYe0XI>

Webpage: <https://danielmoreira.github.io/teaching/biometrics-aut25/>

Today we will...

Get to know what is ahead of you
in the course.

About me

Computer Scientist

PhD from the University of Campinas (Brazil)

Theme: Sensitive-Video Analysis

Loyola University Chicago

Assistant Professor

Joined on August 15, 2022



Research

Media Forensics, Biometrics, Machine Learning, Computer Vision

Webpage: <https://danielmoreira.github.io>

(see following slides)



LOYOLA
UNIVERSITY CHICAGO

A photograph of a man with long hair and a beard, wearing a red baseball cap with the letters "ROY" on it. He is looking upwards and slightly to his left with a thoughtful expression. The background is a blurred outdoor scene with green and yellow colors.

Media Forensics

<https://danielmoreira.github.io/project/medifor/>

The Problem

CNET

SCI-TECH

Spy reportedly used AI-generated photo to connect with targets on LinkedIn

A fake account had links to politically connected figures in Washington, the Associated Press reports.

BY STEVEN MUSIL | JUNE 13, 2019 5:13 PM PDT

Connect



Katie Jones

Russia and Eurasia Fellow

Center for Strategic and International Studies (CSIS) ·
University of Michigan College of Literature, Science...
Washington · 49 connections

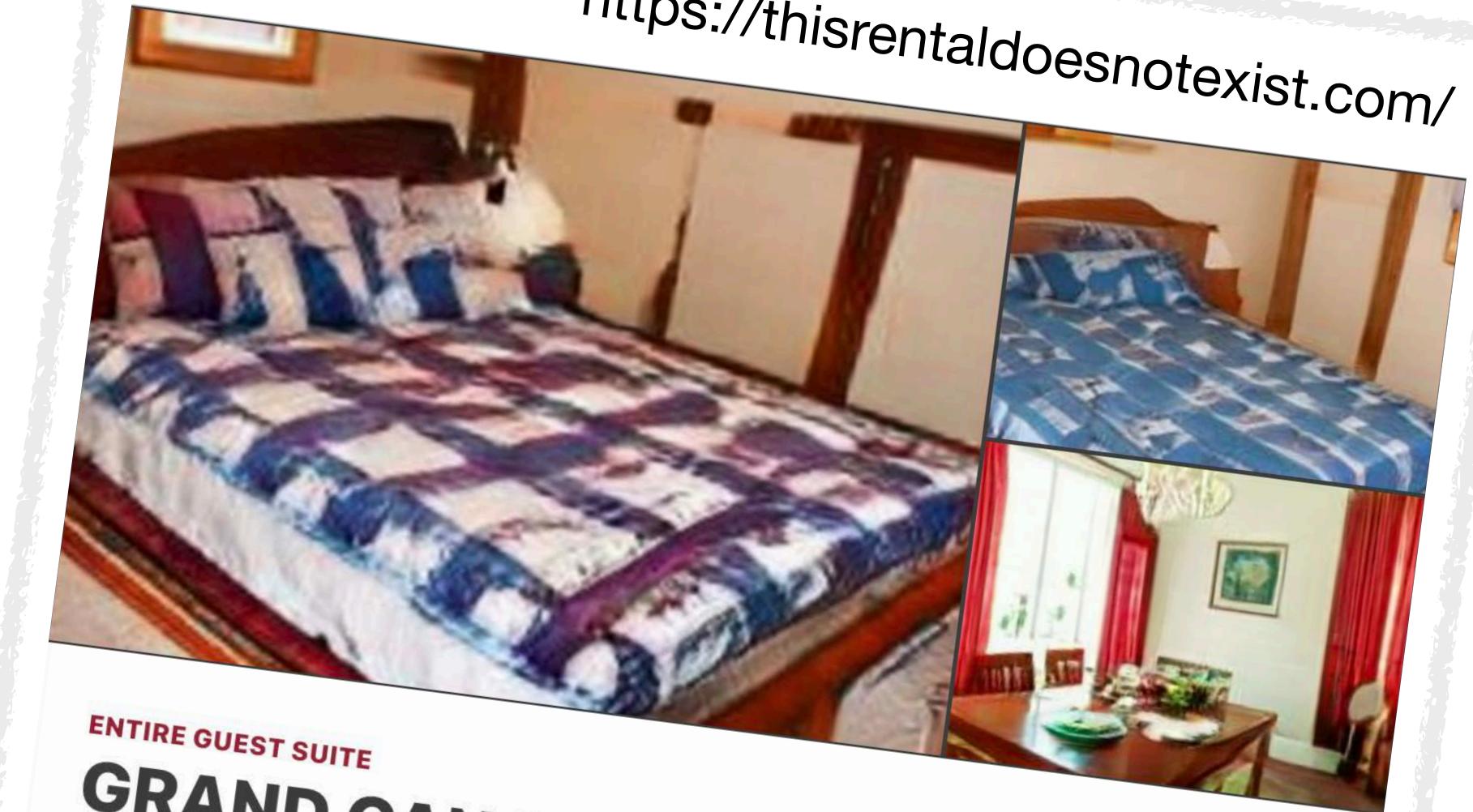
Adobe

**Crafting new images with
photo manipulation.**



6

<https://thisrentaldoesnotexist.com/>



ENTIRE GUEST SUITE
**GRAND CANAL TOUR VIEW 3 BED 1/2
BATH**

[https://www.youtube.com/
watch?v=p7-B8S734T4](https://www.youtube.com/watch?v=p7-B8S734T4)



LOYOLA
UNIVERSITY CHICAGO

A composite image. On the left, a man wearing a red baseball cap with 'ROYAL' and 'S' on it is shown from the chest up. On the right, a portrait of Kurt Cobain, the lead singer of Nirvana, looking slightly upwards and to his right.

The Notorious B.I.G.
NY scene rapper

HANGING OUT?

Kurt Cobain
Grunge scene musician

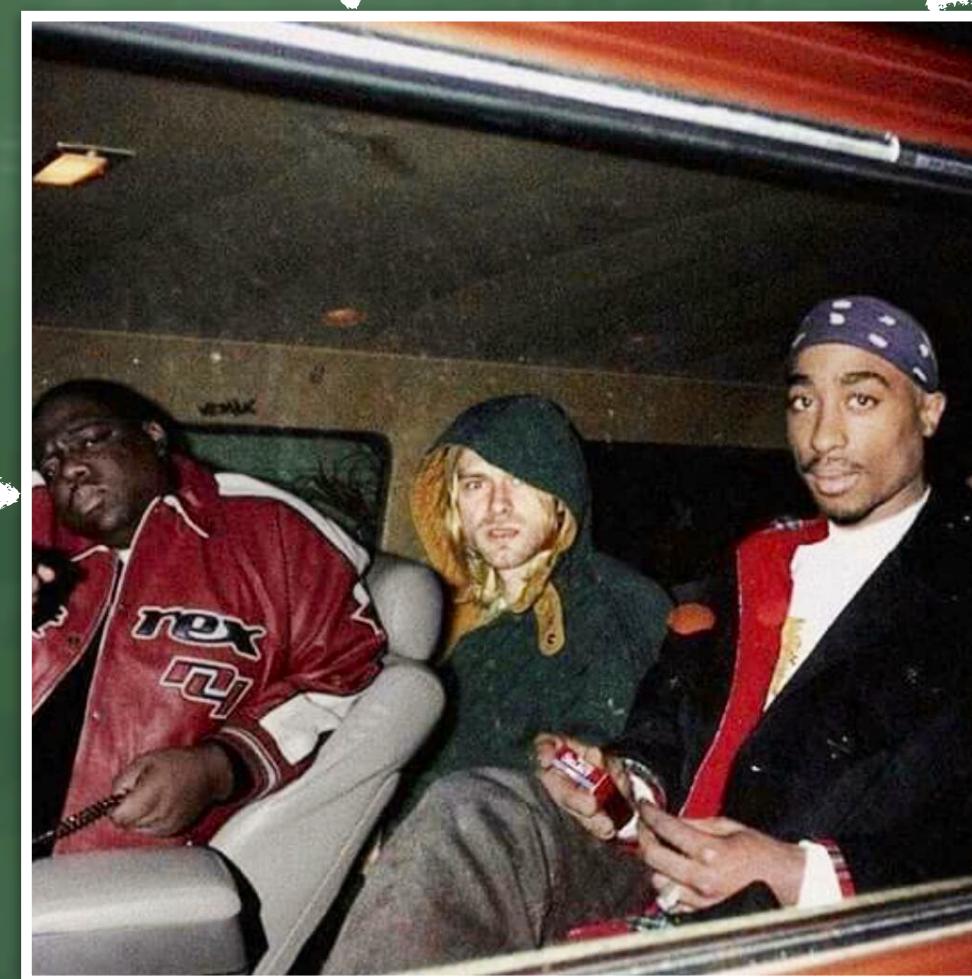
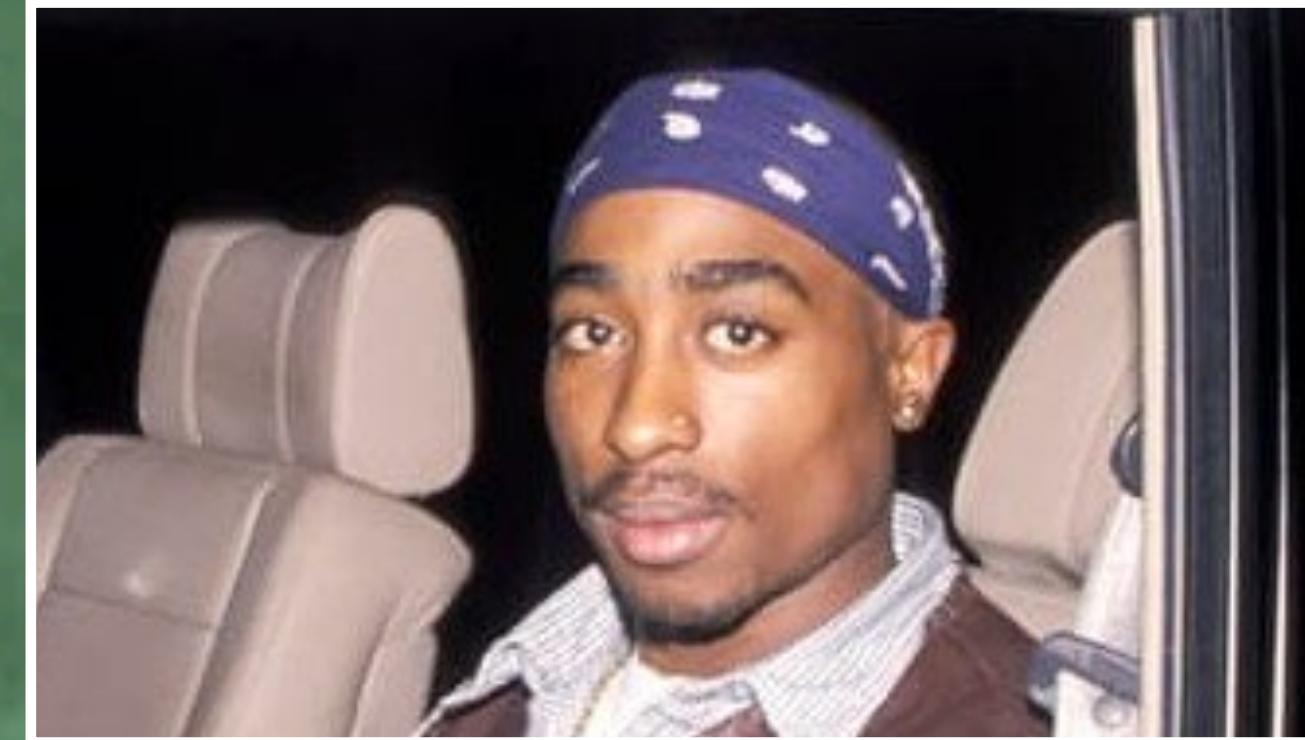


A collage of three images. On the left, a man in a red baseball cap and a red jacket with 'RED' and 'DODGERS' on it. In the center, a woman with long blonde hair wearing a green jacket. On the right, a man wearing a baseball cap and a red jacket.

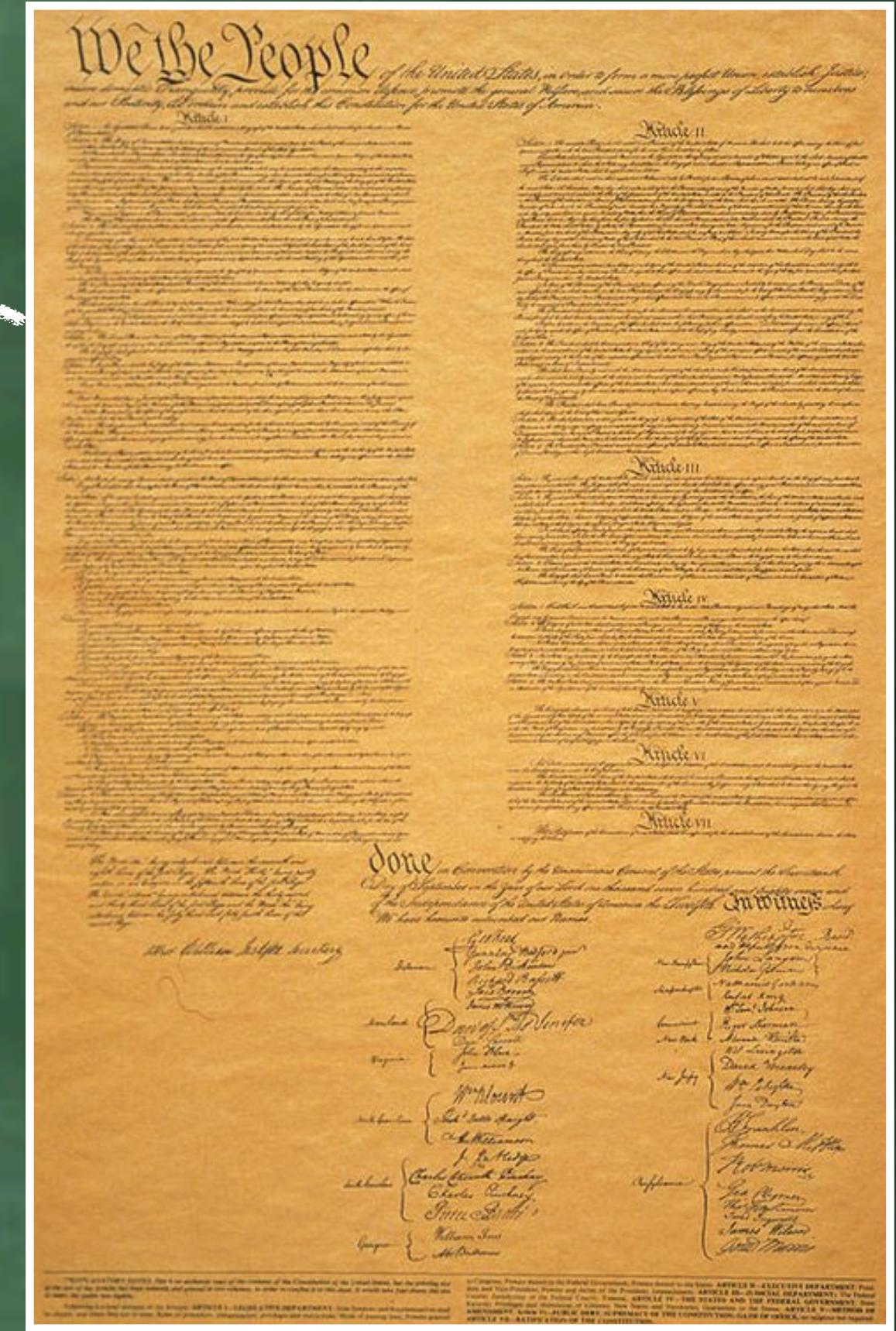
Tupac Shakur
LA scene rapper



Provenance Graph











Scientific Integrity

<https://danielmoreira.github.io/project/sciint/>



Jesse Springer
(@SpringerToons)

The Problem

nature

Retraction Note | Published: 15 April 2015

Retraction Note: Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1

Nature 521, 110 (2015) | Cite this article

27k Accesses | 34 Altmetric | Metrics

1 The Original Article was published on 09 October 2002

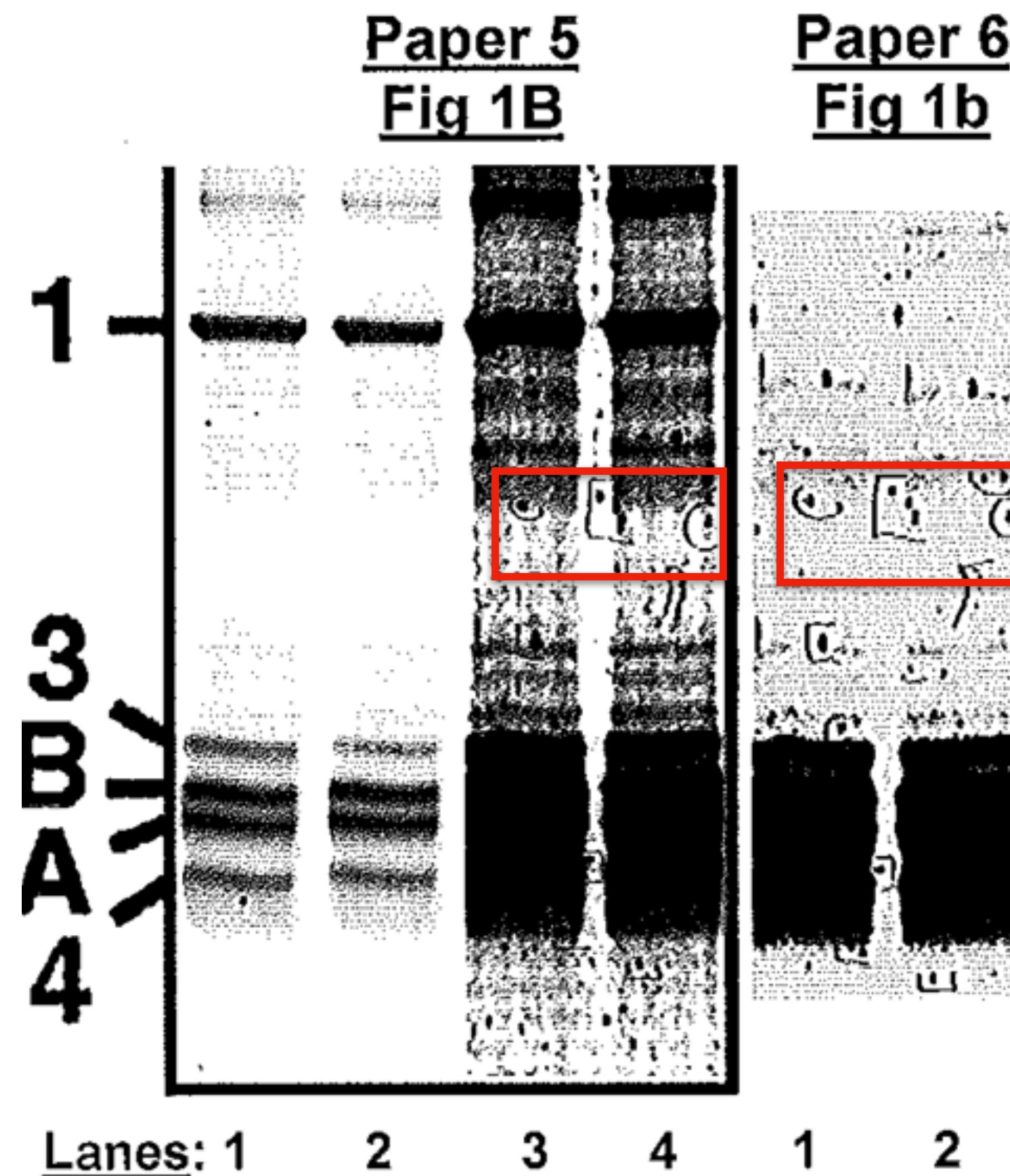
The authors and the University of [REDACTED] wish to retract this letter owing to inappropriate image manipulation in the published figures. The figure panels affected are Figure 1b, d, Figure 2b, e, Figure 3a and Figure 4d. *Nature* has not received a response from [REDACTED] to approve this retraction.

Paper 5 Fig 1B

Paper 6 Fig 1b

1 2 3 4 1 2

Lanes: 1 2 3 4 1 2



CORRECTIONS & AMENDMENTS

RETRACTION

doi:10.1038/nature14421

Retraction: Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1

Nature 419, 857–862 (2002); doi:10.1038/nature01126

The authors and the University of [REDACTED] wish to retract this Letter owing to inappropriate image manipulation in the published figures. The figure panels affected are Figure 1b, d, Figure 2b, e, Figure 3a and Figure 4d. *Nature* has not received a response from [REDACTED] to approve this retraction.

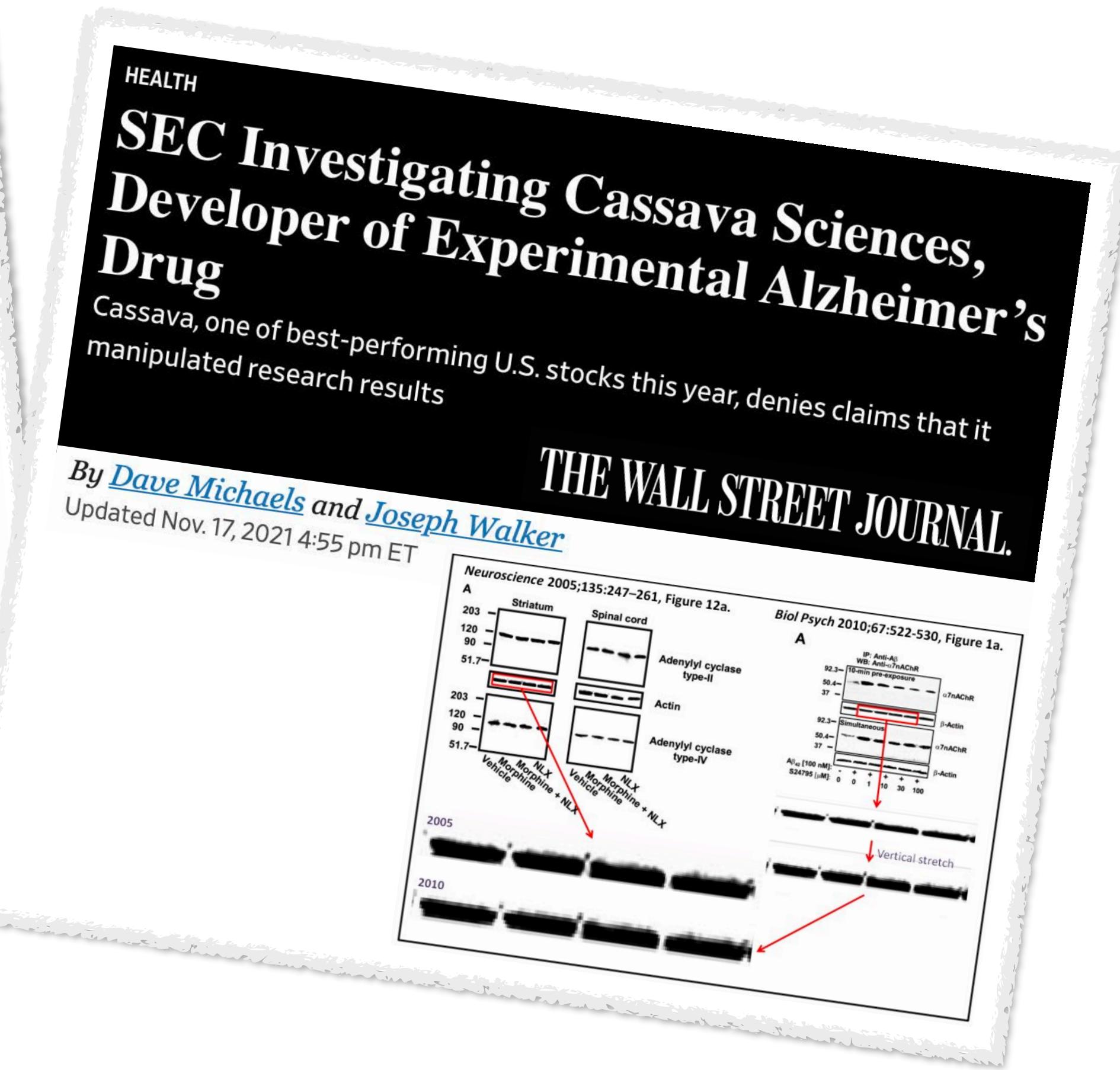
Additional information

The online version of the original article can be found at [10.1038/nature01126](https://doi.org/10.1038/nature01126)



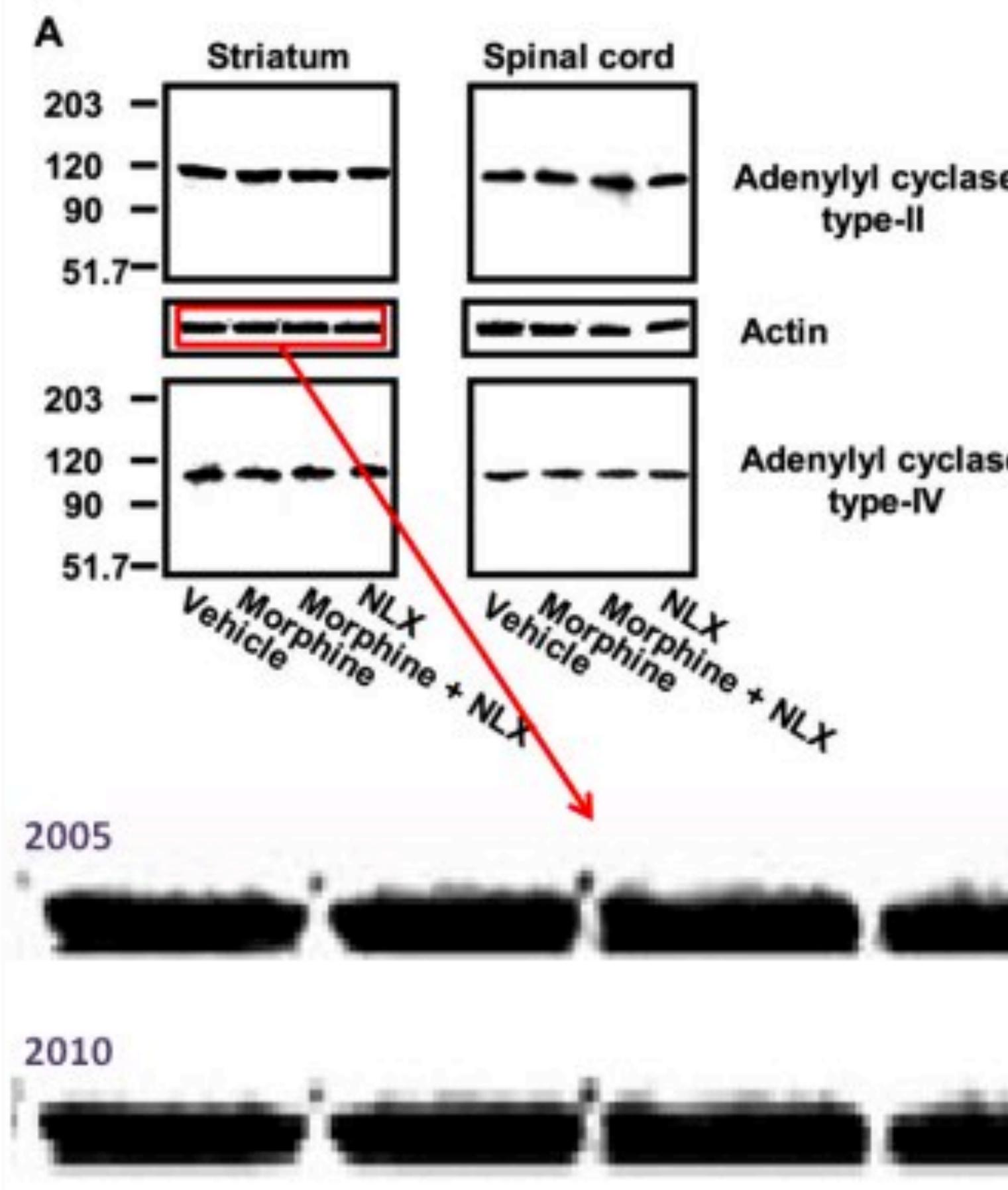
LOYOLA
UNIVERSITY CHICAGO

The Problem

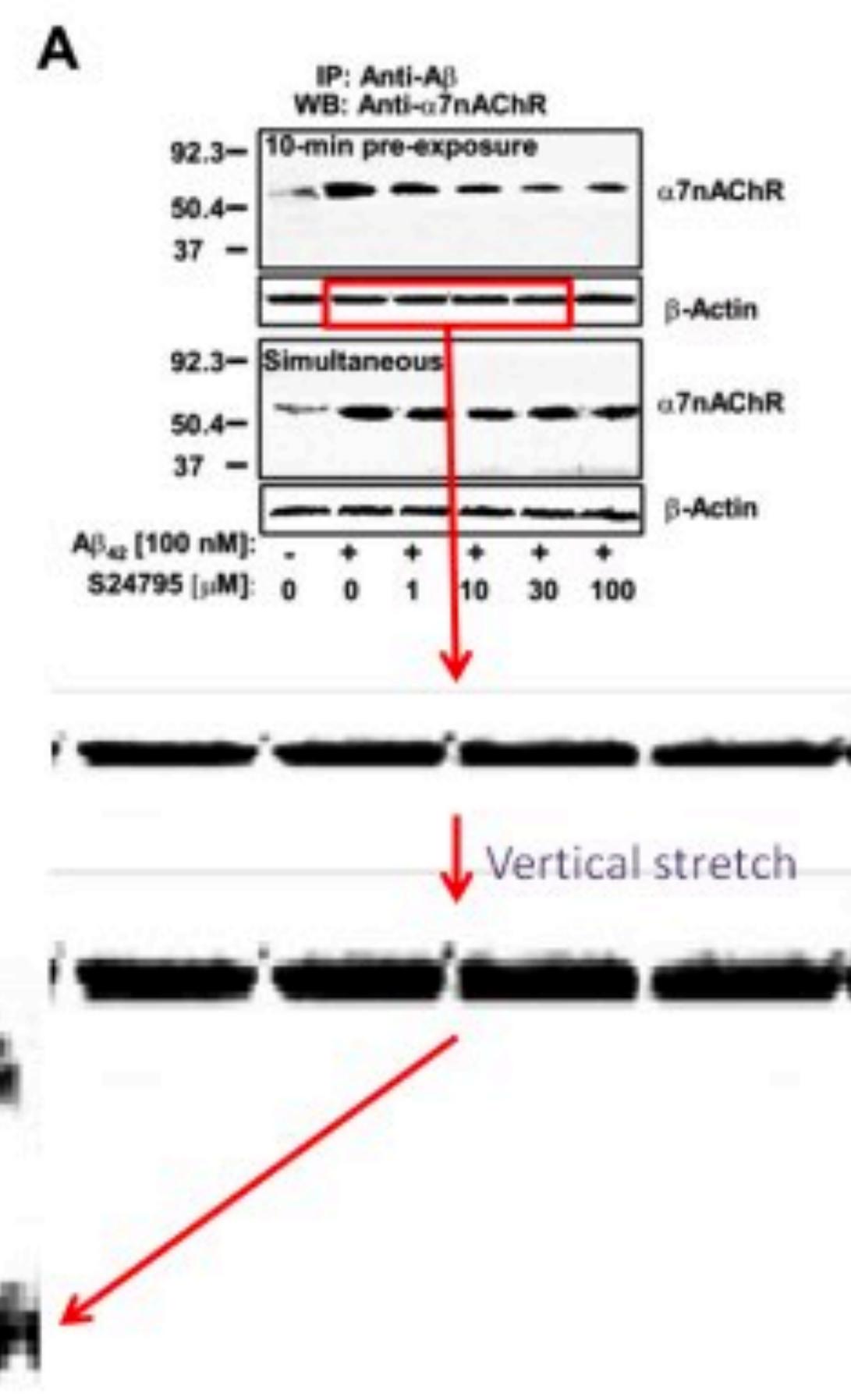


LOYOLA
UNIVERSITY CHICAGO

Neuroscience 2005;135:247–261, Figure 12a.



Biol Psych 2010;67:522-530, Figure 1a.



HEALTH

SEC Investigating Cassava Sciences, Developer of Experimental Alzheimer's Drug

Cassava, one of best-performing U.S. stocks this year, denies claims that it manipulated research results

THE WALL STREET JOURNAL.

By [Dave Michaels](#) and [Joseph Walker](#)

Updated Nov. 17, 2021 4:55 pm ET

The Problem

nature
Retraction Note | Published: 15 April 2015
Retraction Note: Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1
Nature 521, 110 (2015) | Cite this article
27k Accesses | 34 Altmetric | Metrics

The Original Article was published on 09 October 2002

The authors and the University of [REDACTED] wish to retract this letter owing to inappropriate image manipulation in the published figures. The figure panels affected are Figure 1b, d, Figure 2b, e, Figure 3a and Figure 4d. *Nature* has not received a response from [REDACTED] to approve this retraction.

Elisabeth Bik @MicrobiomDigest

Retraction of a paper by authors of The Third Military Medical University, Zhengzhou University, and @JohnsHopkins, 5 months after posting this to @PubPeer. pubpeer.com/publications/4...

3:17 PM · Oct 2, 2022 · Twitter Web App

3 Retweets 41 Likes

Paper 5 Fig 1B: Lanes 1, 2, 3, 4. Paper 6 Fig 1b: Lanes 1, 2.

Lanes: 1, 2, 3, 4 (Paper 5); 1, 2 (Paper 6)

(f) Figure 3f: Sham (WT), ICH (WT). (f) Figure 4f: Sham (WT), ICH (WT). Labels: AntiCD25, CD28SA.

SEC Investigating Cassava Sciences, Developer of Experimental Alzheimer's Drug
Cassava, one of best-performing U.S. stocks this year, denies claims that it manipulated research results

By [Dave Michaels](#) and [Joseph Walker](#)
Updated Nov. 17, 2021 4:55 pm ET

THE WALL STREET JOURNAL.

Neuroscience 2005;135:247-261, Figure 12a.
Biol Psych 2010;67:522-530, Figure 1a.

Red arrows point to specific bands in the blots.

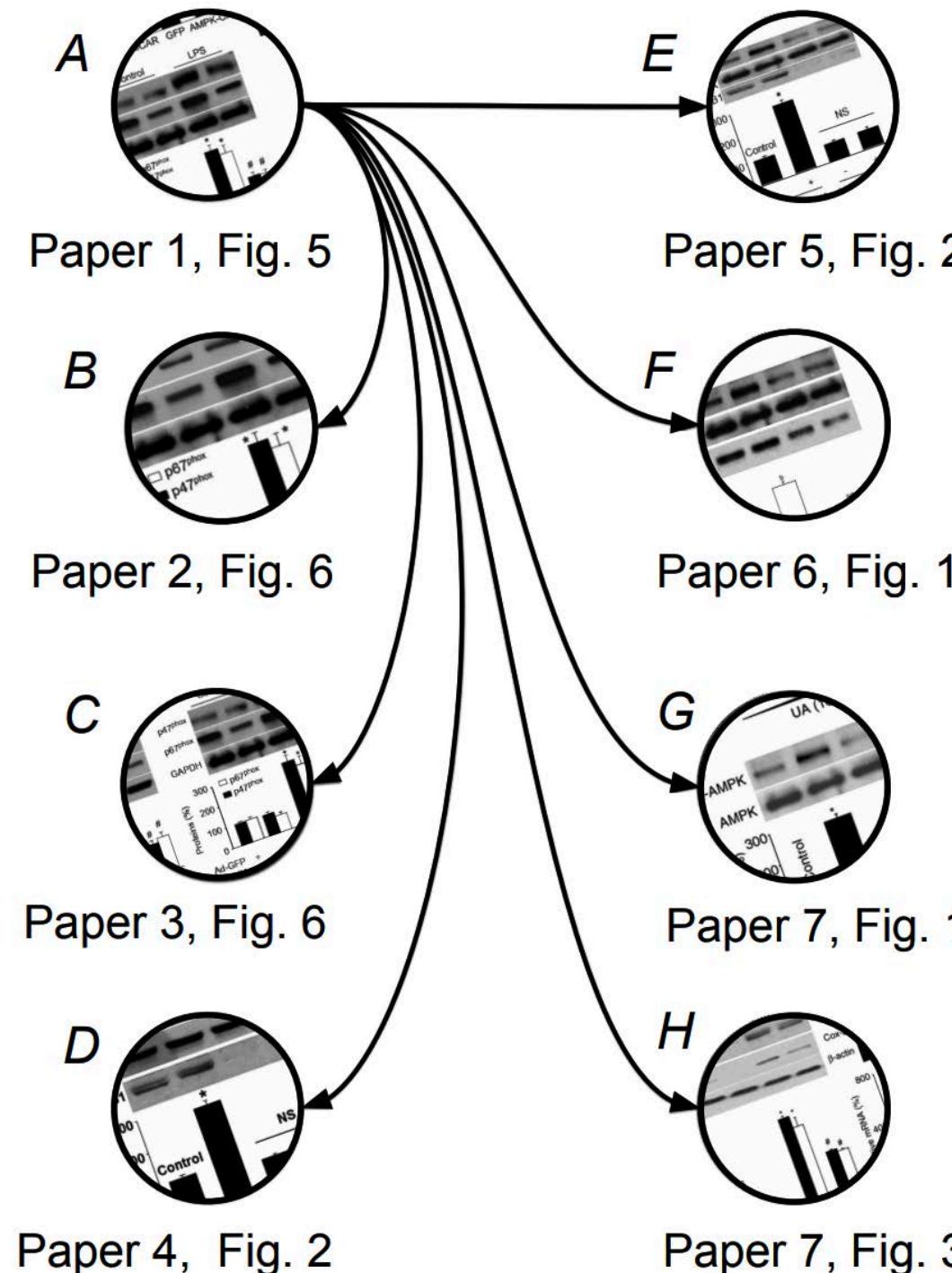


LOYOLA
UNIVERSITY CHICAGO

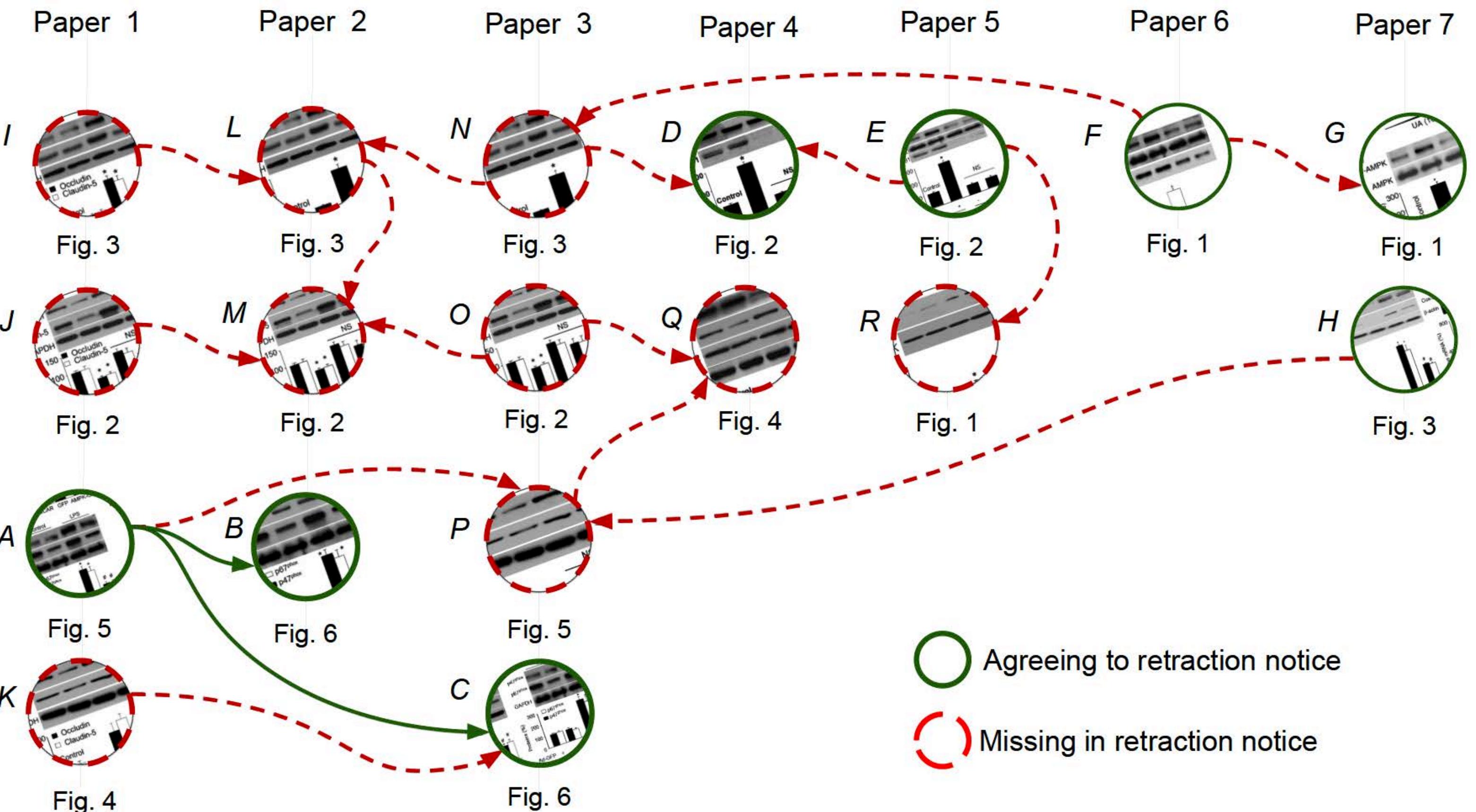
Provenance Analysis

Ground truth

(according to retraction notice)



Our findings

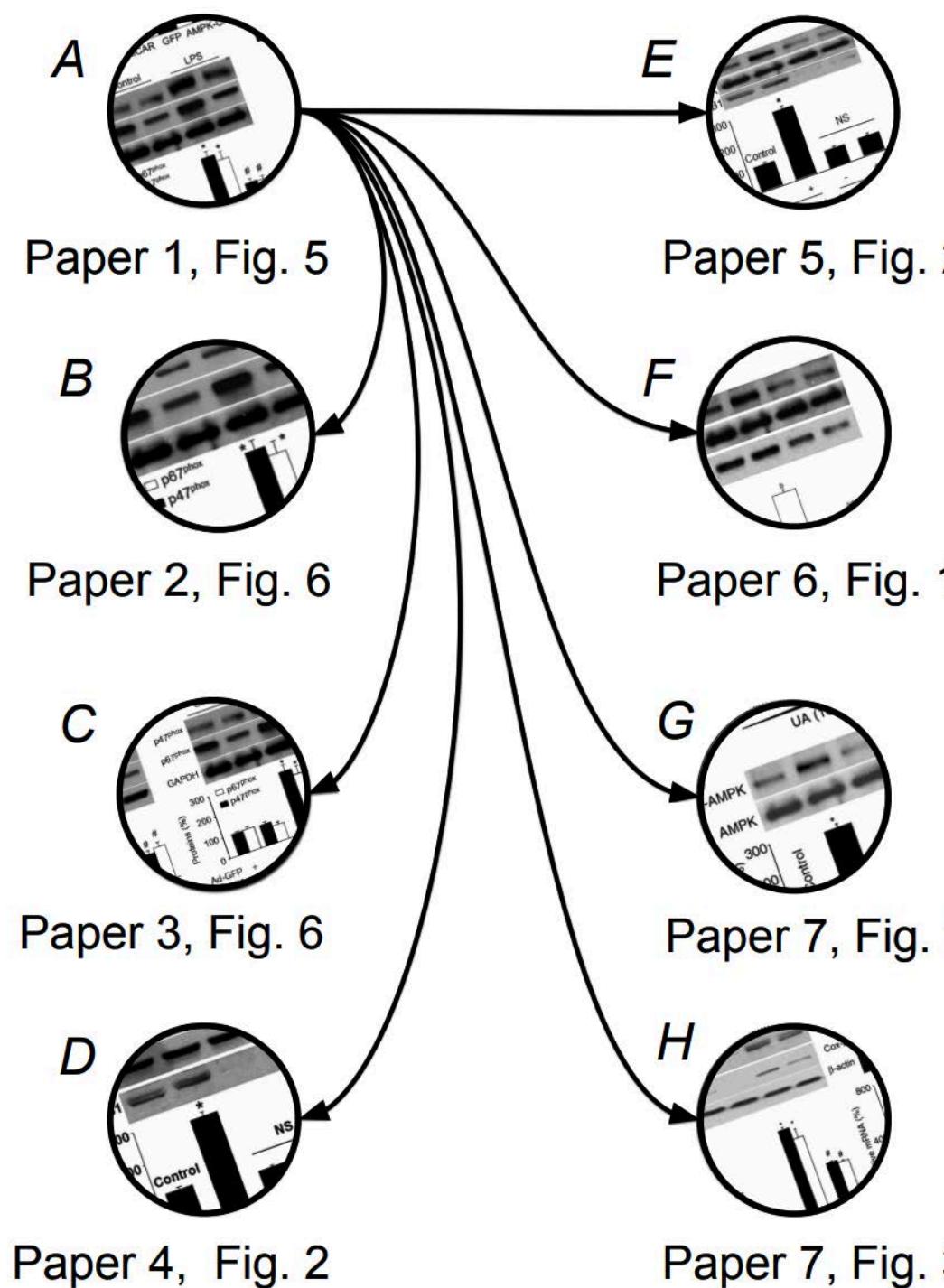


- Agreeing to retraction notice
- Missing in retraction notice

Provenance Analysis

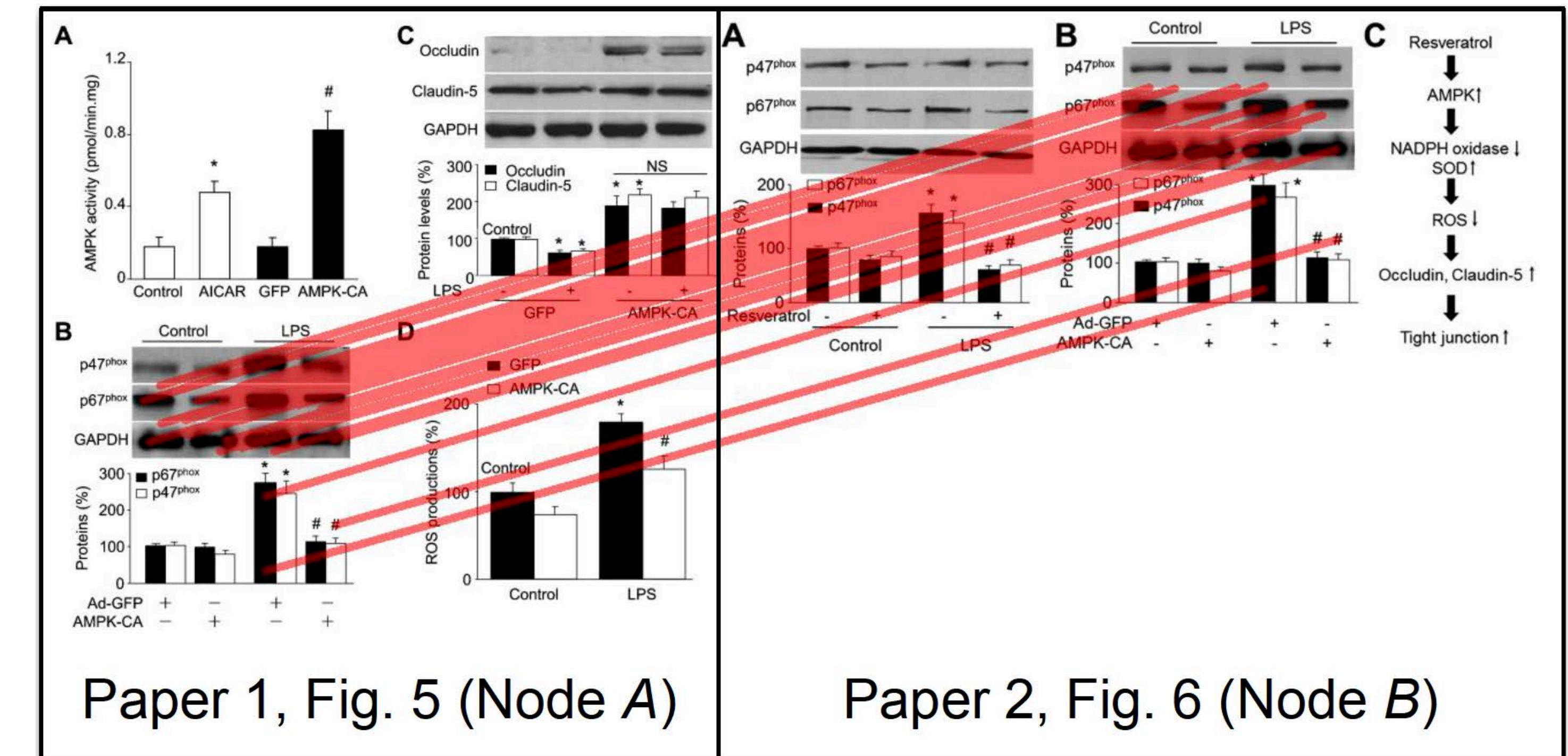
Ground truth

(according to retraction notice)



Our findings

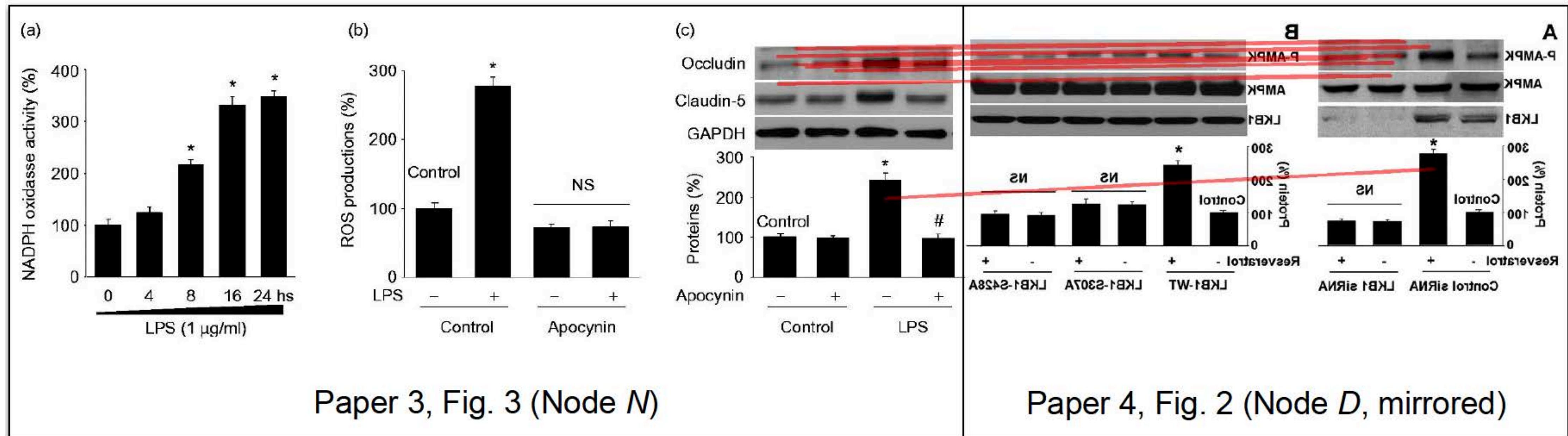
(in accordance with retraction notice)



Provenance Analysis

Our findings

(not reported in the retraction notice)



10.1371/journal.pone.0190562

Paper Mills

<https://tinyurl.com/4ukfbka7>

The collage includes several pieces of content:

- A news article from "Science" titled "PAPER TRAIL" with the subtitle "In the latest twist of the publishing arms race, firms churning out papers are taking to bribing journal editors". It features a photo of a person at a desk and a timestamp "18 JAN 2024 · 2:00 PM ET · READ MORE".
- A large graphic with the text "AT LEAST TENS OF MILLION DOLLARS FLOW TO THE PAPER MILL INDUSTRY".
- A tweet from Elisabeth Bik (@MicrobiomDigest) dated Jun 19, 2023, with the text: "Wow, this is pretty blatant. Author has gotten acceptance letter for their paper from the journal 'Molecular Biology and Evolution', and now authorship positions are being sold on @facebook for 5K or more. #PaperMill".
- A search bar with the text "These are organized crime rings that are committing large-scale fraud."

<https://tinyurl.com/3xtmedrb>

Paper Mills

Retraction Watch

**Want to earn \$10k per month?
Join the “journals mafia”**

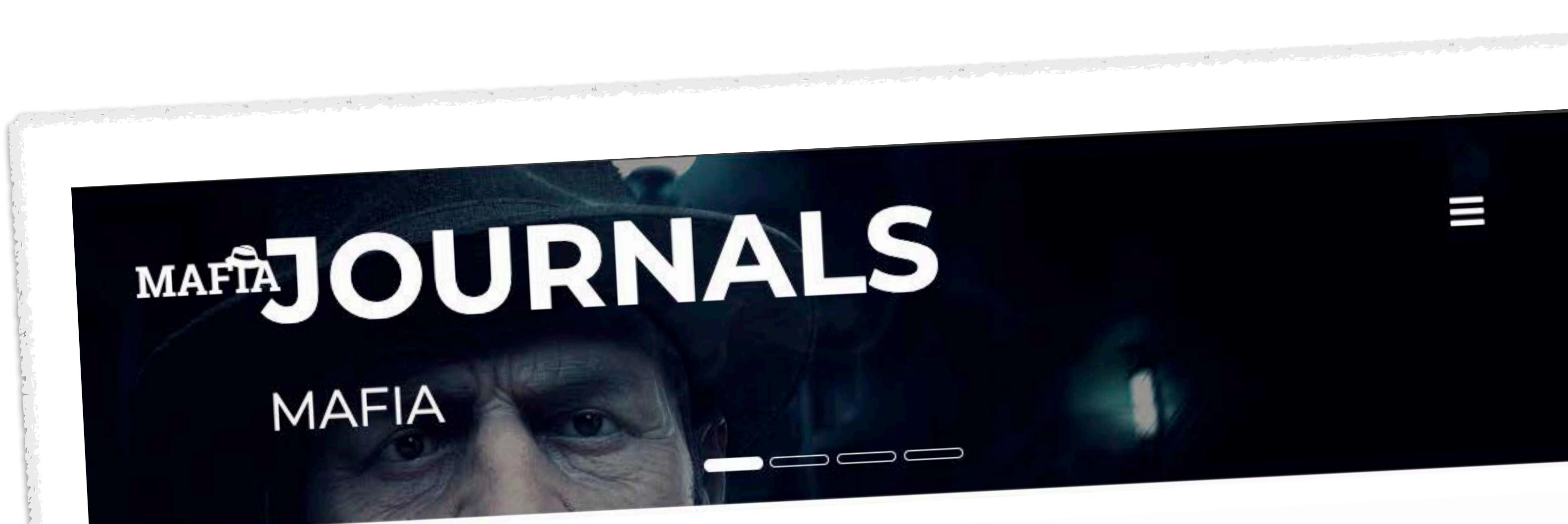
<https://tinyurl.com/4m3t9umb>

*The authors pay to publish the articles,
the company shares the profits with journals
that publish the paper.*



LOYOLA
UNIVERSITY CHICAGO

Paper Mills



Are you an editor of
the journal or a
member of editorial
board?

It is necessary to publish articles.
The same work that you do, but you can get more
money doing this with us.
The profit is from 1,000 up to 10,000 dollars per a month.

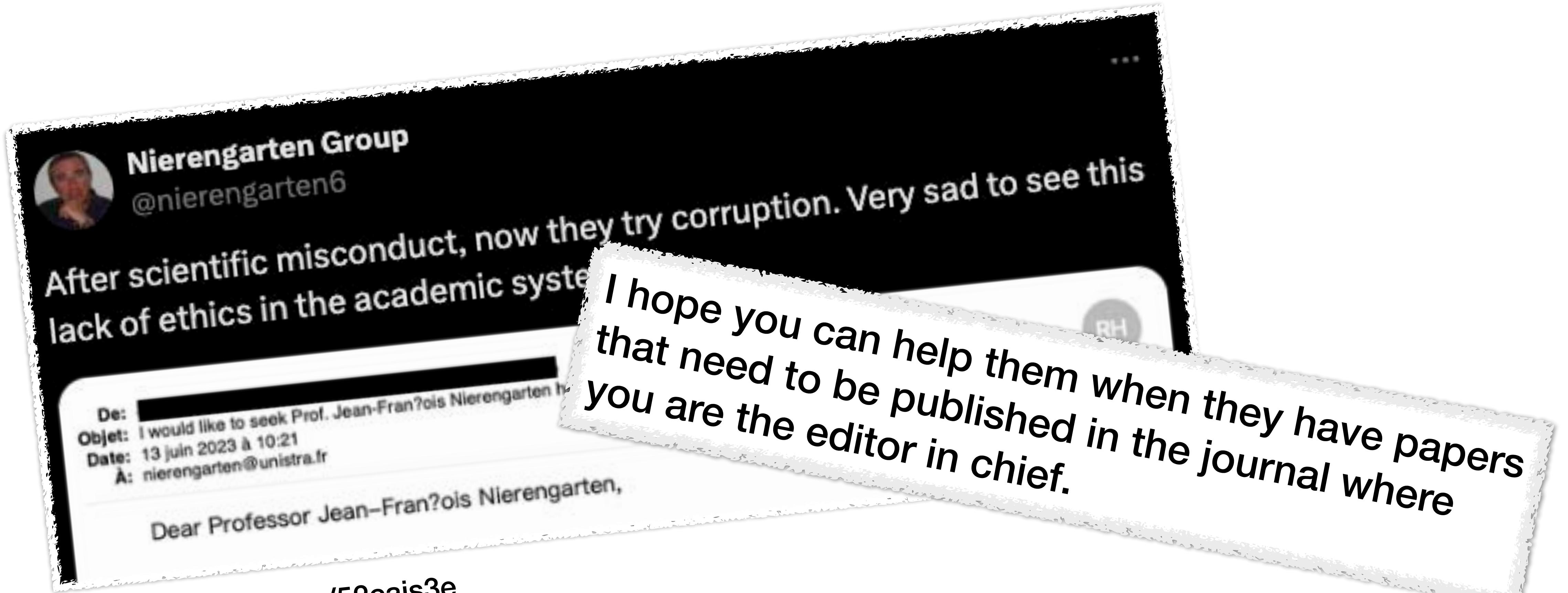
10.000 \$ PER
MONTH

PUBLISHING
BUSINESS

A dramatic, low-key lighting close-up of Marlon Brando's face as Don Corleone. He is shown from the chest up, wearing a dark suit and tie. His expression is serious and contemplative. The background is dark and out of focus.

I'm gonna make them an offer they can't refuse.

Paper Mills



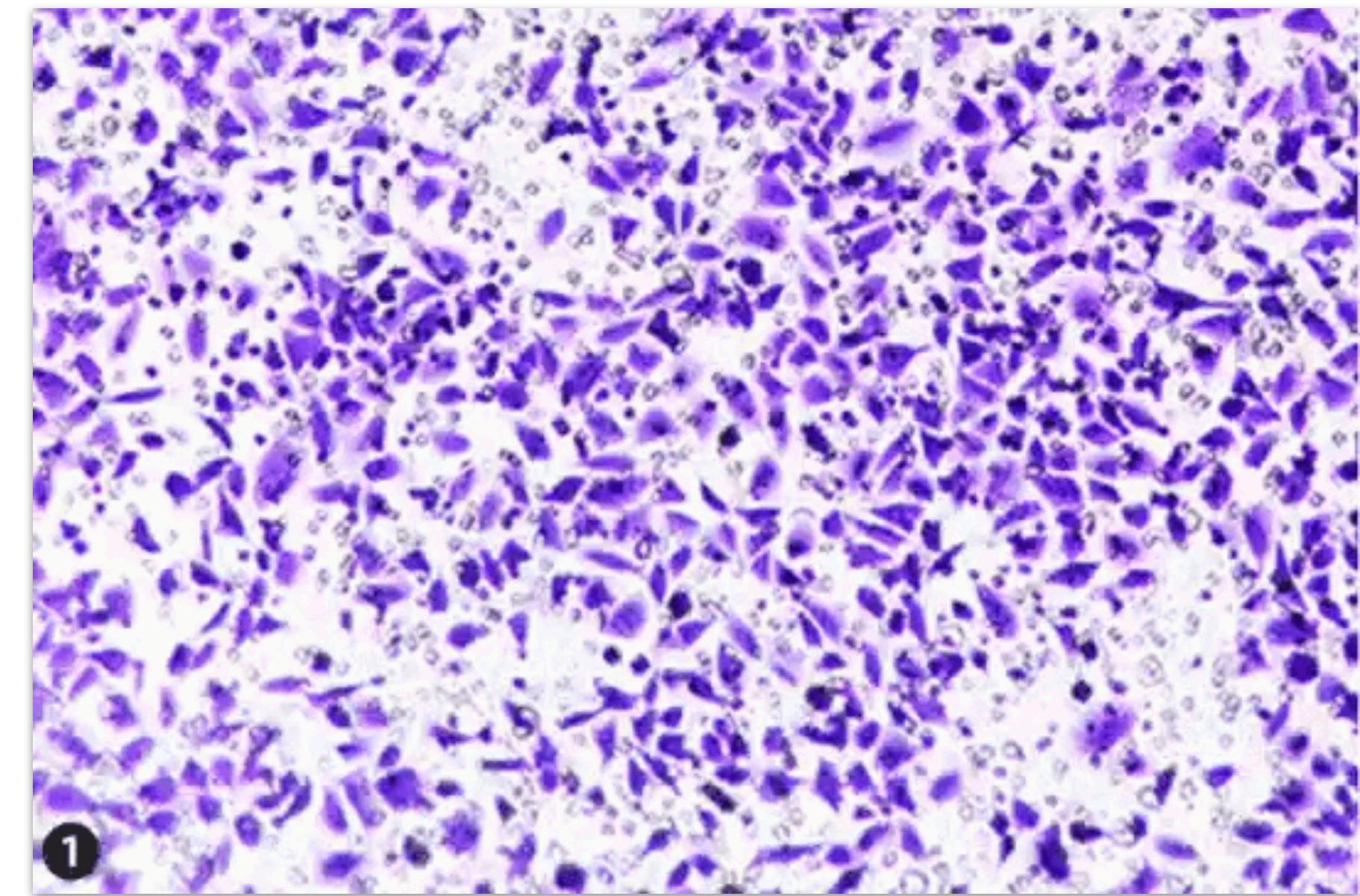
<https://tinyurl.com/52cajs3e>

Paper Mills

Problems with Images?



<https://tinyurl.com/2p8ecptc>



<https://tinyurl.com/4znhmpa6>

Paper Mills

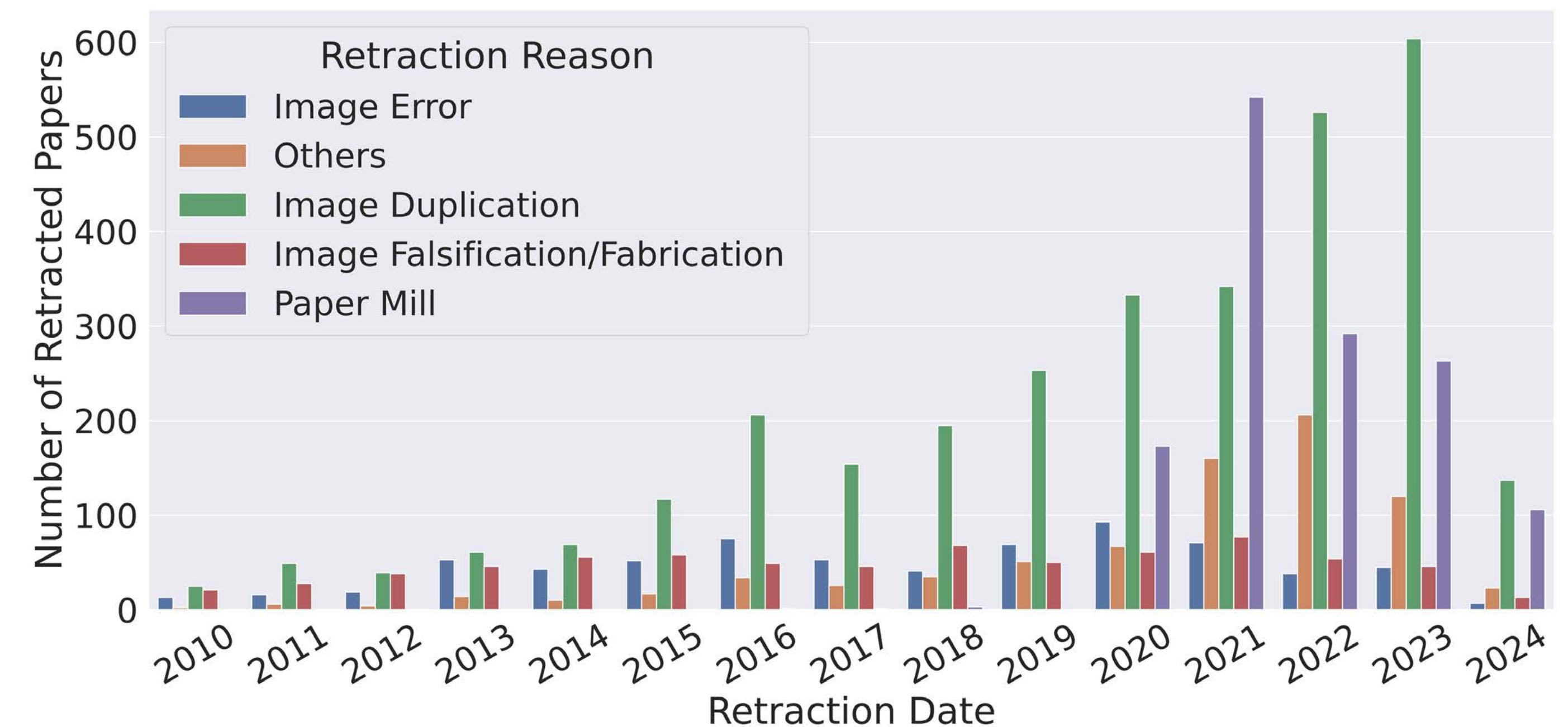
Problems with Images?

PLOS ONE

**Unveiling scientific articles from
paper mills with provenance analysis**

Cardenuto, Moreira, and Rocha, 2024

<https://tinyurl.com/2fx3p6tj>



Paper Mills

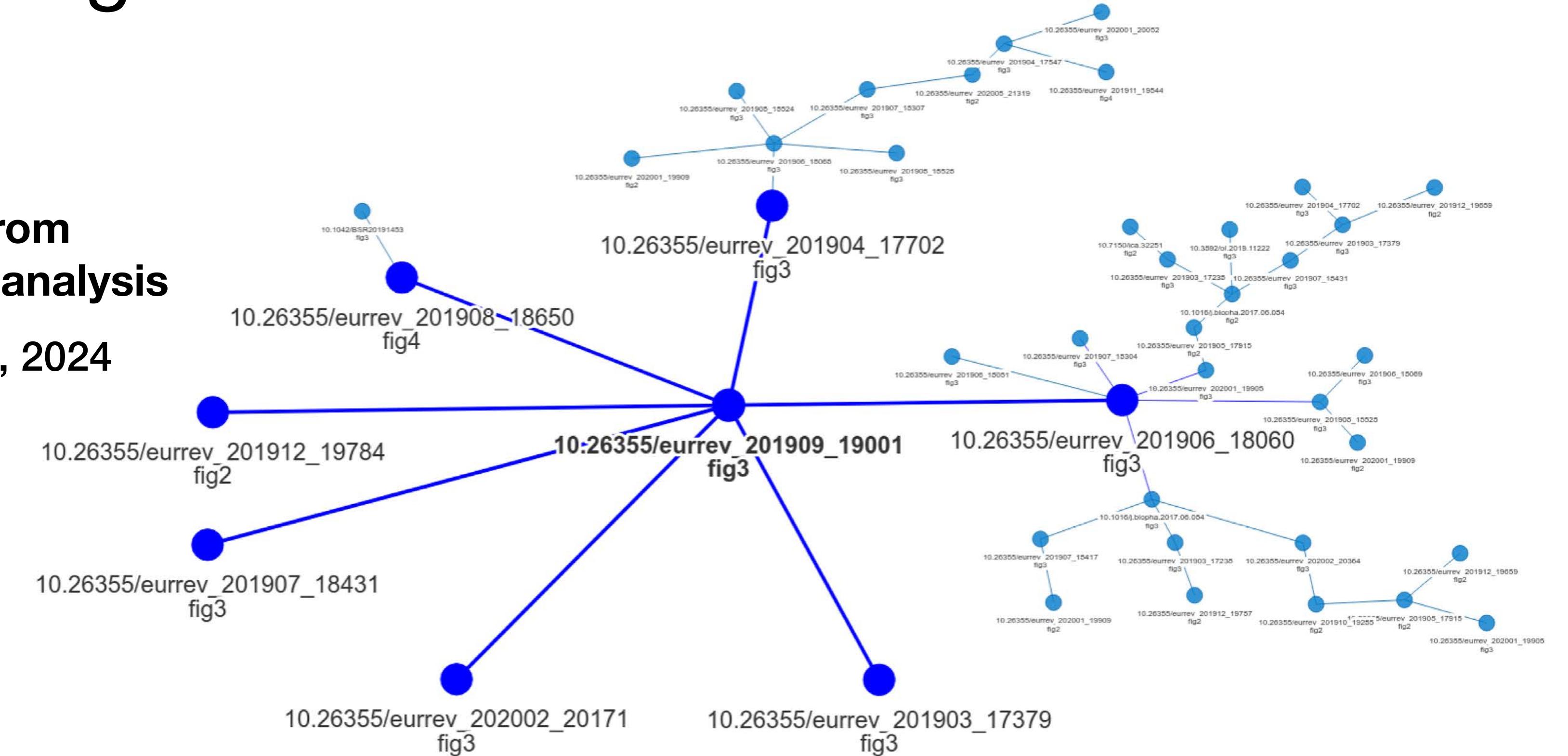
Problems with Images?

PLOS ONE

**Unveiling scientific articles from
paper mills with provenance analysis**

Cardenuto, Moreira, and Rocha, 2024

<https://tinyurl.com/2fx3p6tj>





We can use AI...

AI-generated Content



ChatGPT



DALL-E 3



Midjourney



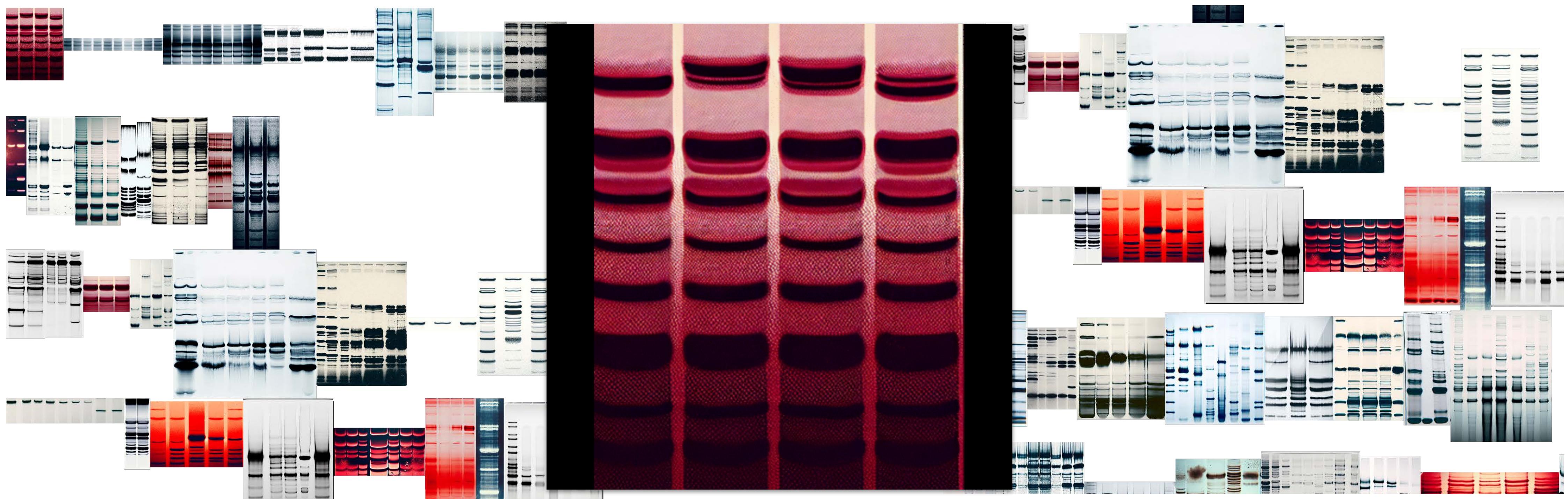
Hugging Face



DALL-E 3

AI-generated Content

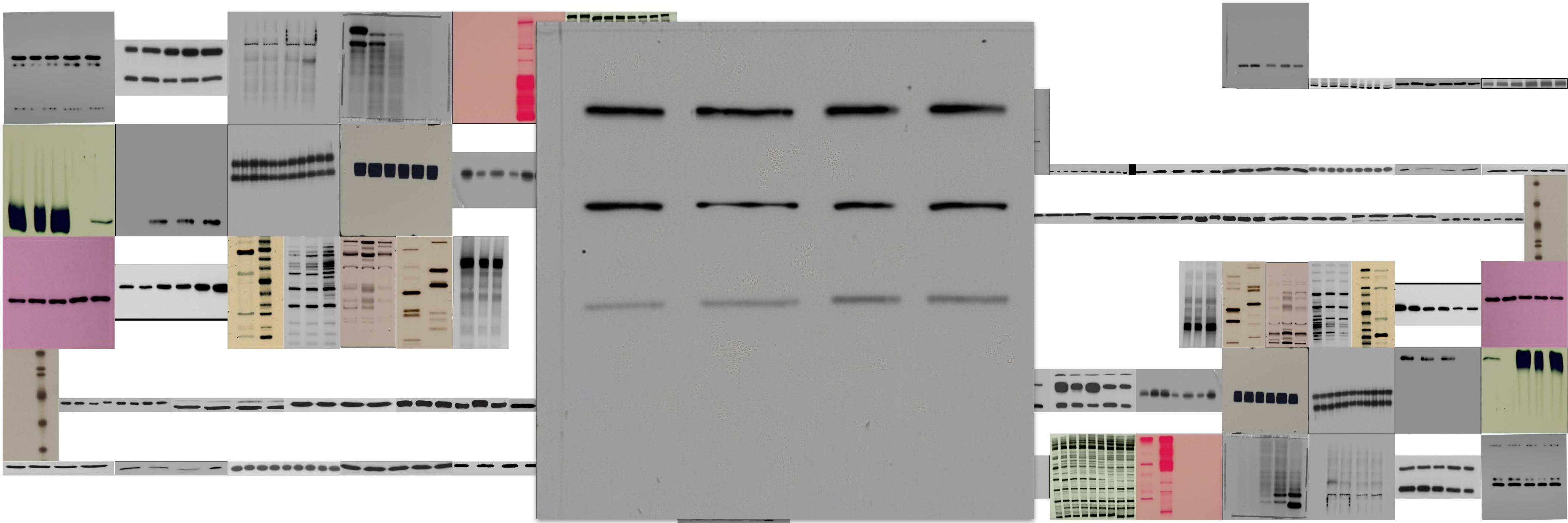
AI-generated Western blots!



DALLE-3

AI-generated Content

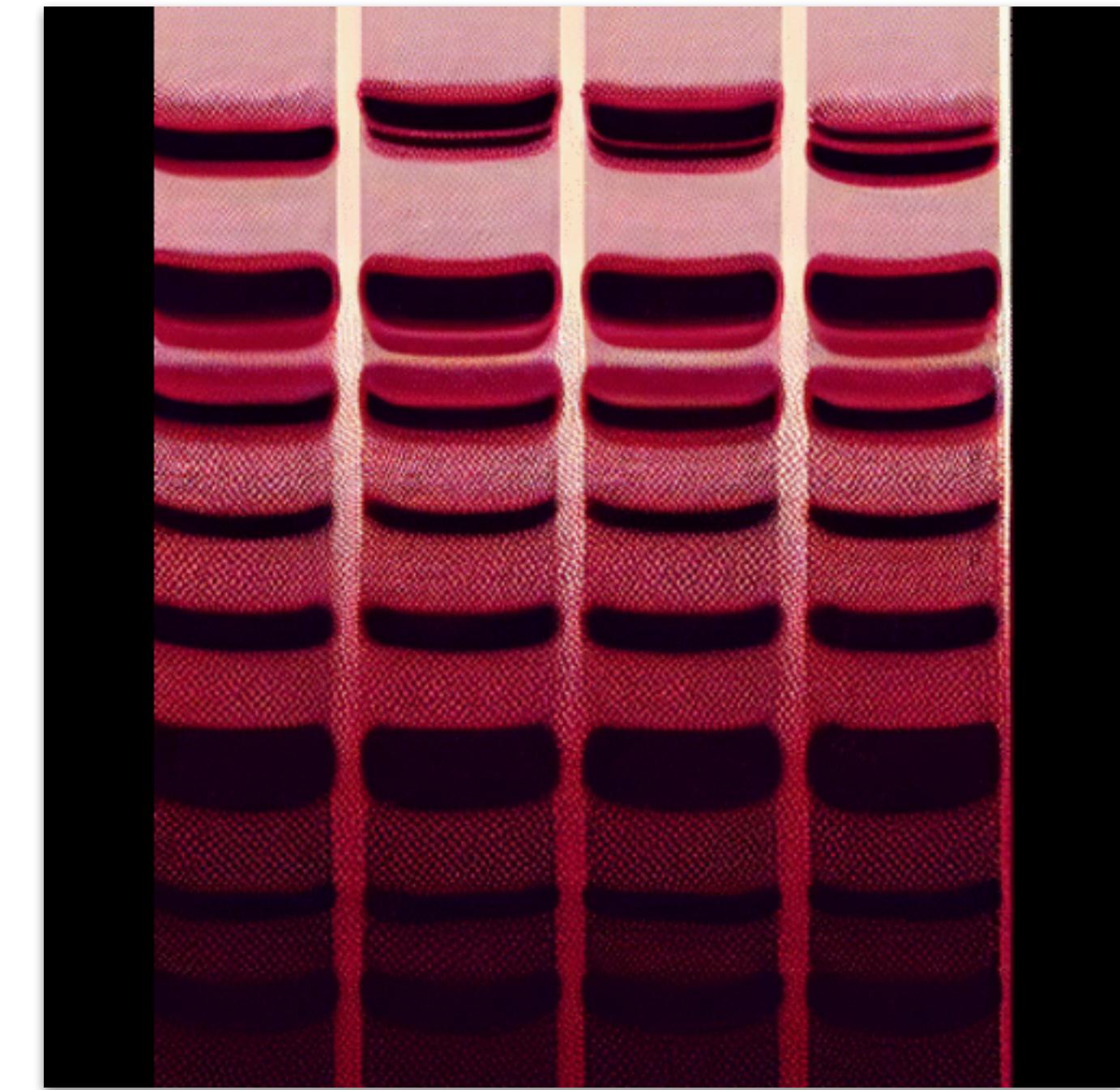
AI-generated Western blots!



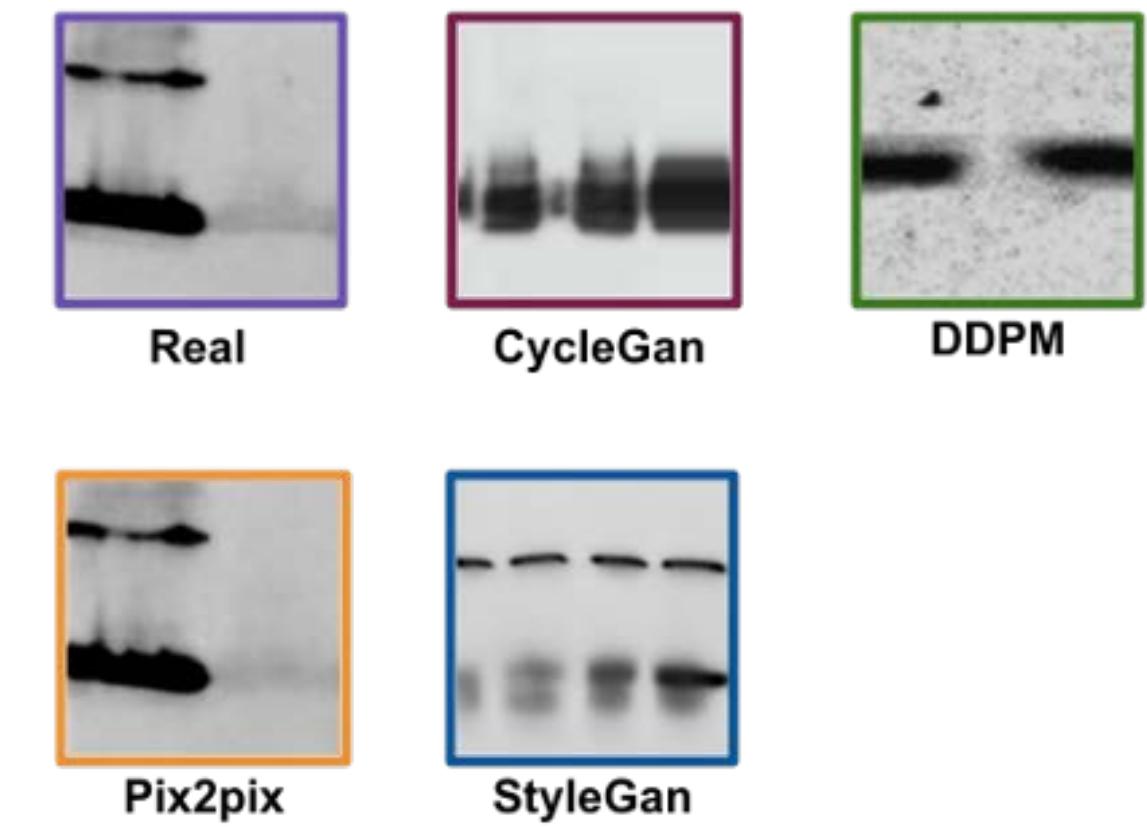
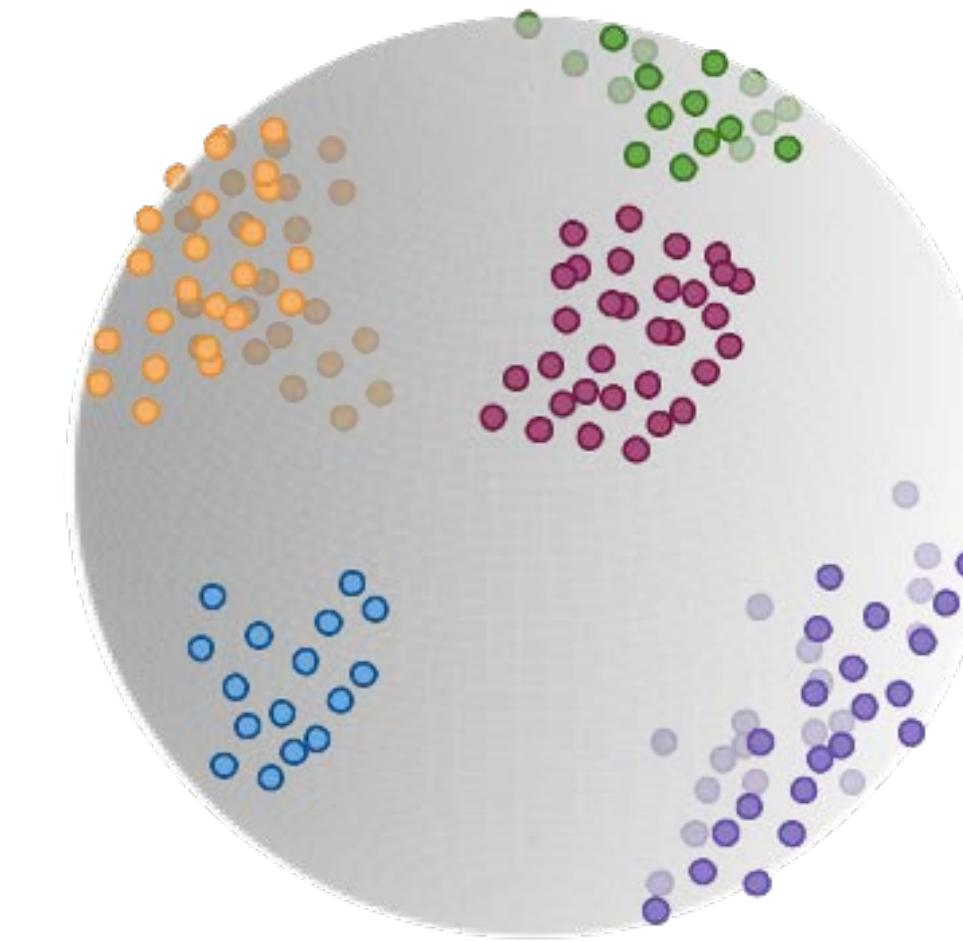
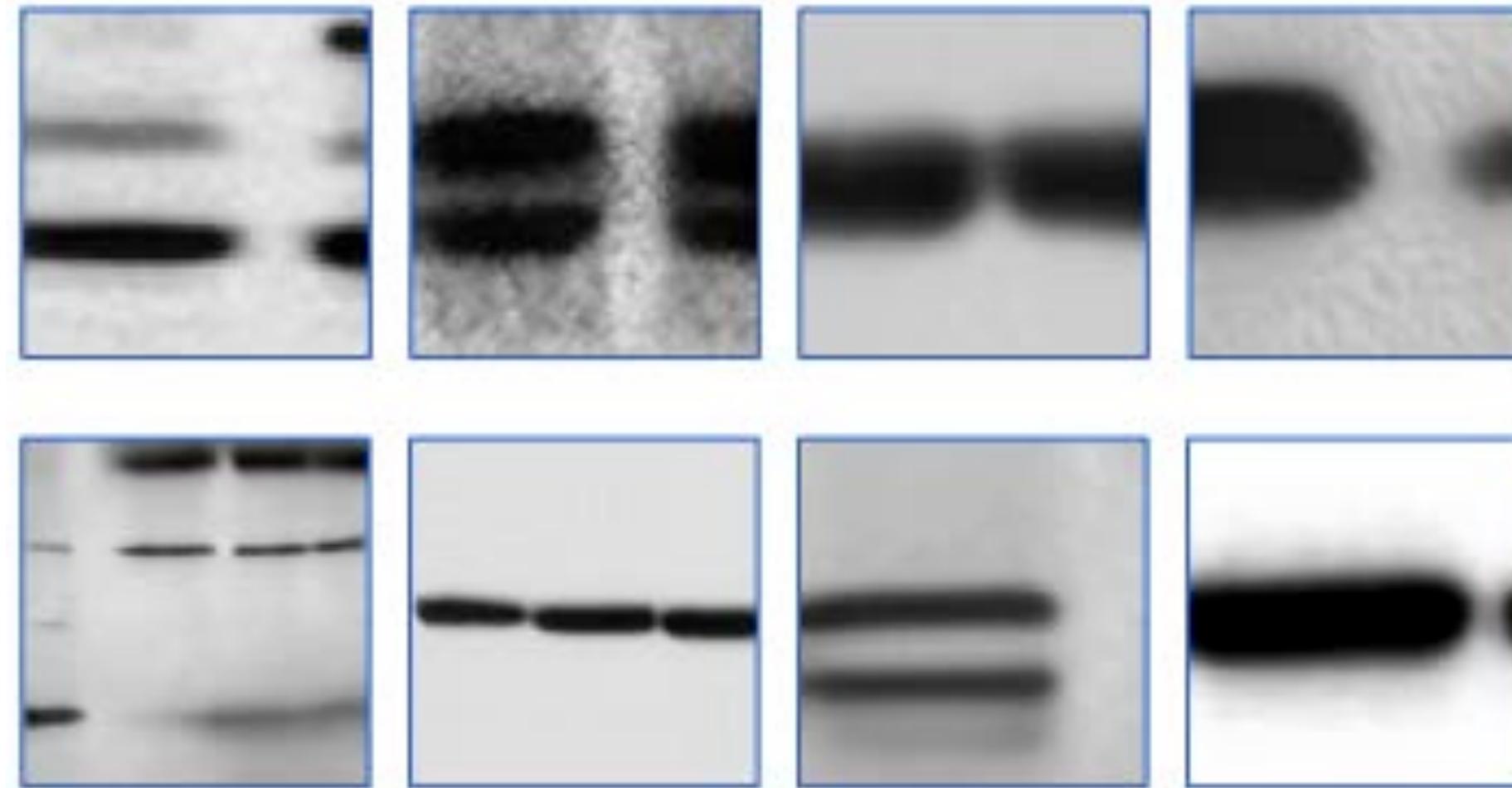
Stable Diffusion

AI-generated Content

What can we do about it?



Synthetic Image Attribution



Mandelli et al.
Forensic Analysis of Synthetically Generated Western Blots
IEEE Access

IEEE WIFS 2024

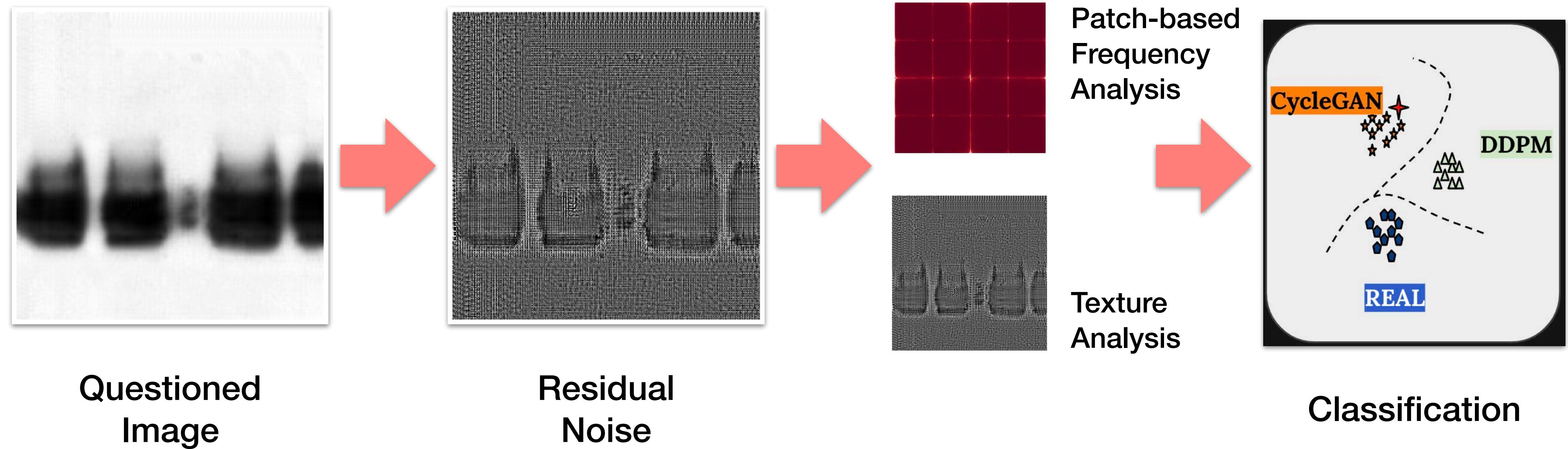


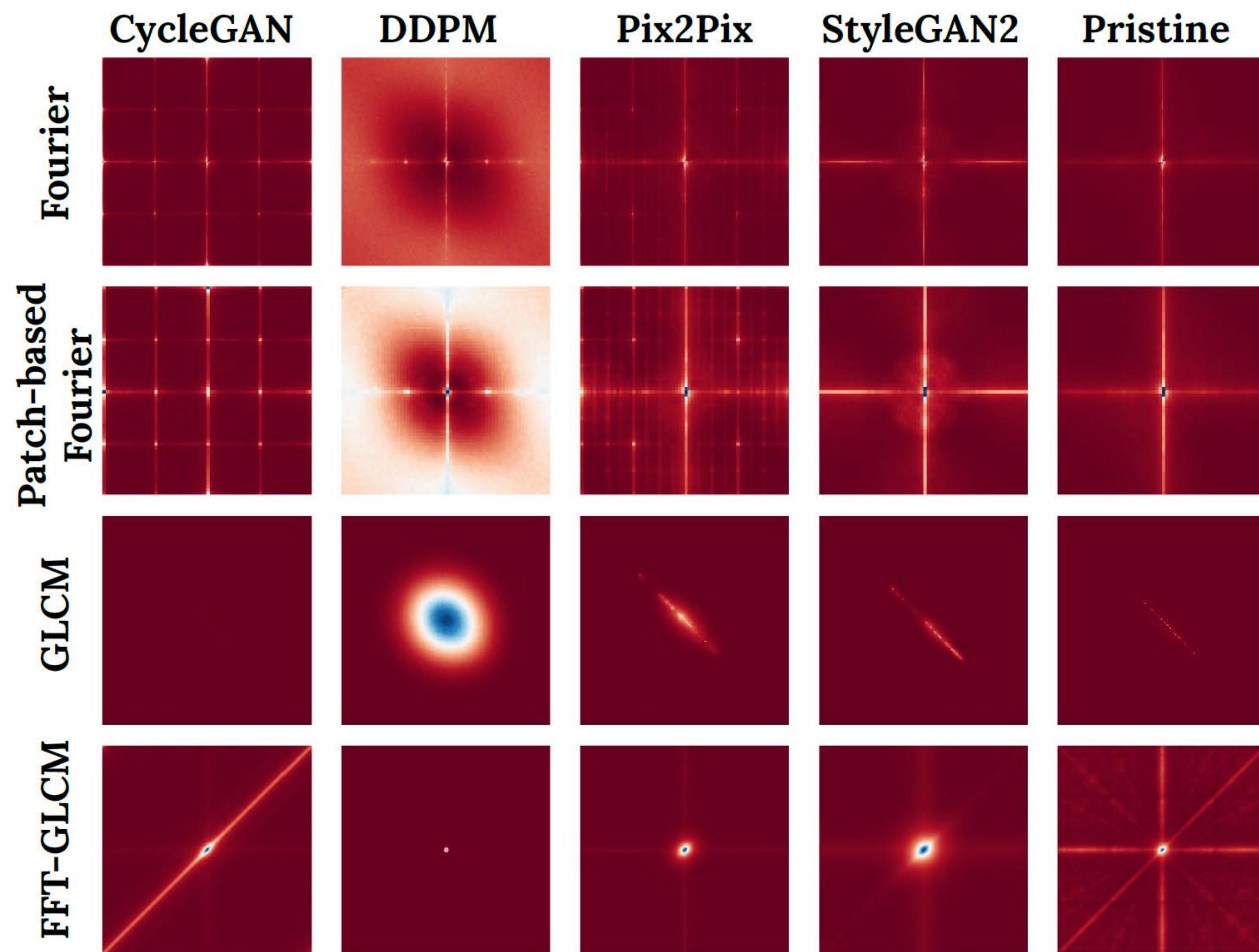
**Explainable Artifacts for
Synthetic Western Blot
Source Attribution**

Cardenuto, Mandelli, Moreira,
Bestagini, Delp, and Rocha
(Unicamp, Polimi, Purdue, and LUC)

<https://tinyurl.com/4fwa7syb>

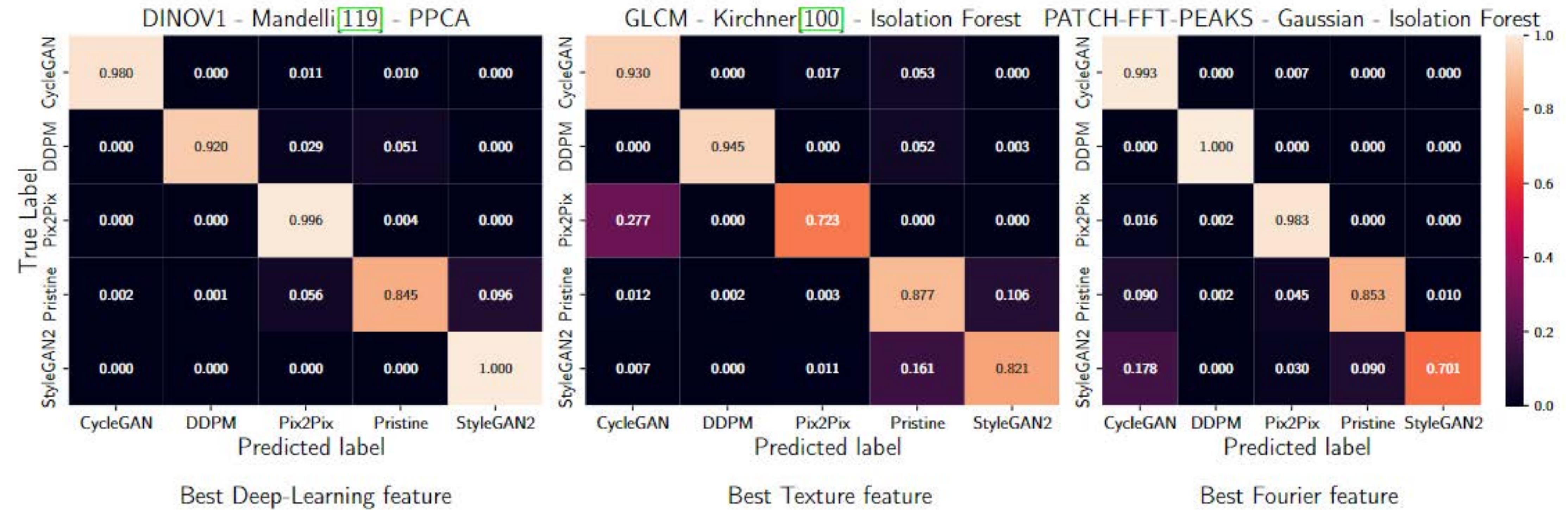
Explainable Artifacts





Experiments and Results

Source Attribution



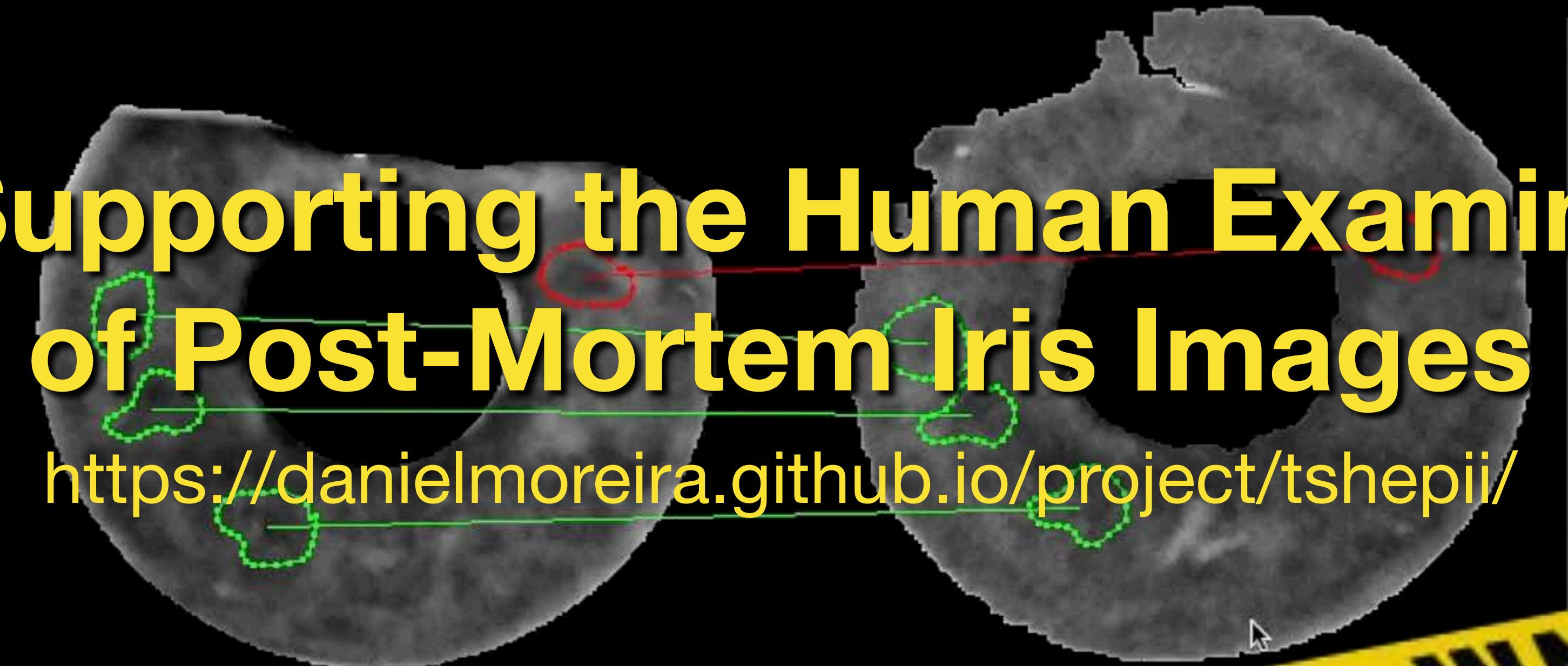
Load irises

Load examination

Save examination

Save report

Quit program



Tool Supporting the Human Examination of Post-Mortem Iris Images

<https://danielmoreira.github.io/project/tshepii/>

The Problem

BIOMETRICS

Interpretable Iris Recognition



How to convince people
who do not possess image
processing expertise?

Load irises

Load examination

Save examination

Save report

Quit program



Tool Supporting the Human Examination of Post-Mortem Iris Images

<https://danielmoreira.github.io/project/tshepii/>

The screenshot displays several functional panels:

- Human-Interpretable Features:** A section with three rows of checkboxes and counters for 'TSHEPII', 'SURF', and 'Crypts'. Each row includes a 'Show Matched' checkbox and a 'Show Unmatched' checkbox, followed by a counter table and a status message like 'out of 0'.
- Non-Human-Interpretable Features:** A section for 'Gabor Filters' with a threshold of 0.4461 and 'BSIF Filters' with a threshold of 0.4216.
- Manual Annotation:** A panel for 'Annotate...' with color-coded buttons for 'Matching Regions' (green) and 'Non-Matching Regions' (red). It also includes checkboxes for 'Show Matching Regions' and 'Show Non-Matching Regions'.
- Global match score:** A simple text field showing the value 0.4216.

How about you?

Background

9 graduate and 15 undergraduate students

What is your major?



[forms.gle/
wMNBSY2NwV5LHSbo9](https://forms.gle/wMNBSY2NwV5LHSbo9)



How about you?

Expectations

Given your (future) career, what are your course expectations?



[forms.gle/
wMNBSY2NwV5LHSbo9](https://forms.gle/wMNBSY2NwV5LHSbo9)

How about you?

Accommodation Needs

Please reach out to me in private ASAP.
We'll make things work.



Today's Attendance

Please fill out the form

forms.gle/wMNBSY2NwV5LHSbo9



LOYOLA
UNIVERSITY CHICAGO

About the topic

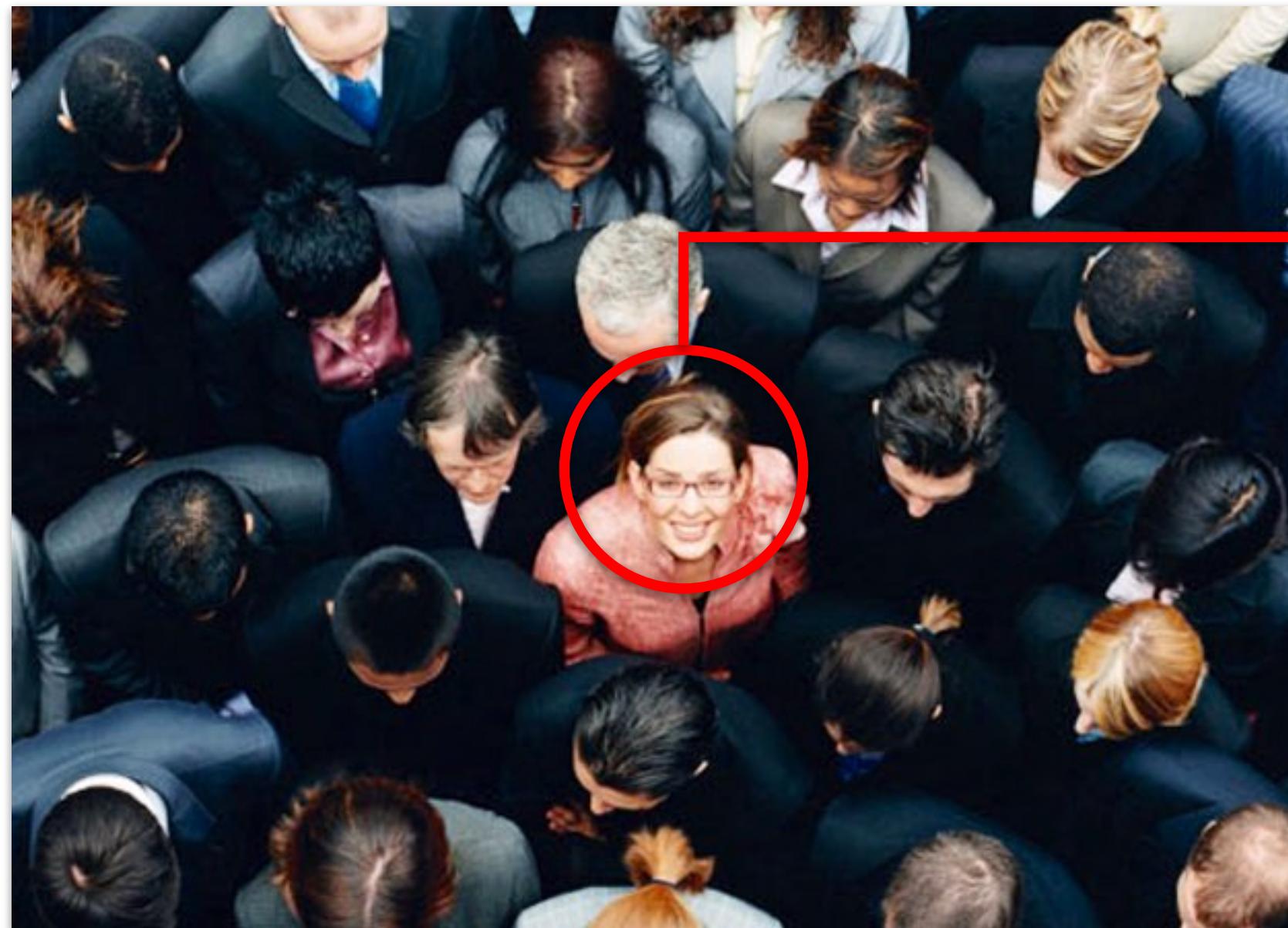
Biometrics

What comes to your mind?



[forms.gle/
wMNBSY2NwV5LHSbo9](https://forms.gle/wMNBSY2NwV5LHSbo9)

What is Biometrics?



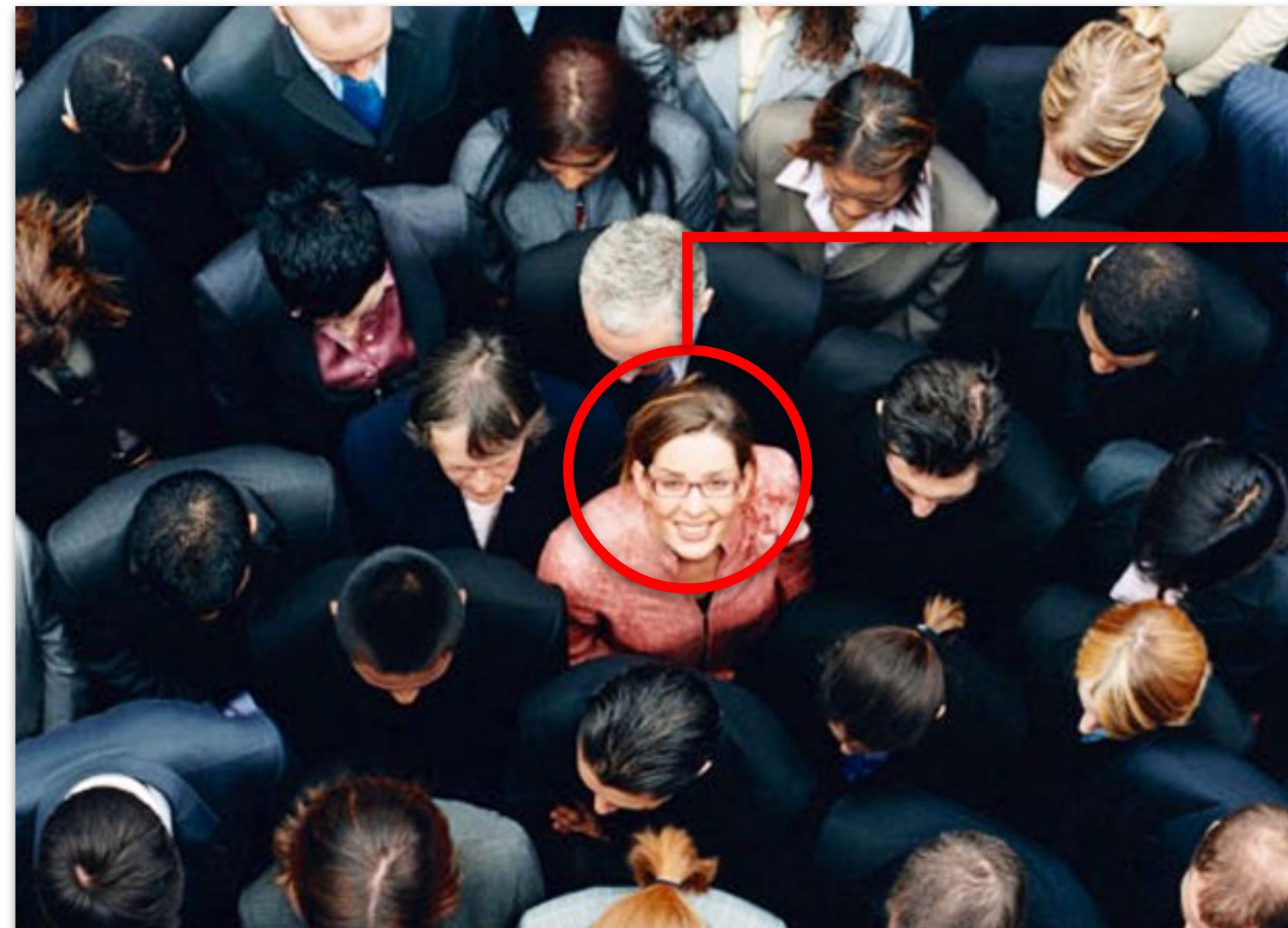
8.2 billion people
Who is this person?
Is this person Jane Doe?

What is Biometrics?



- 8.2 billion people**
 - Who is this person? (*Identification*)
 - Is this person Jane Doe? (*Verification*)

What is Biometrics?



- **8.2 billion people**
 - Who is this person? (*Identification*)
 - Is this person Jane Doe? (*Verification*)

Biometrics aims at *identifying* or *verifying* the claimed identity of an individual based on their *physical*, *chemical*, or *behavioral* traits.

What is Biometrics?



In this course, we aim at
computer-aided Biometrics.

We'll focus on **software solutions**
rather than hardware.

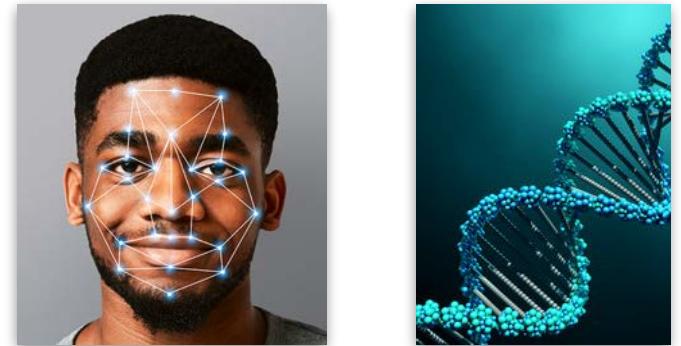
But we'll get to use some
devices, I promise.

What is Biometrics?

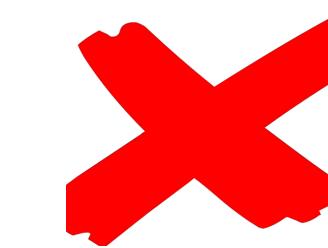
Identity verification through:



A unique trait
of yours.



Not something
you have.



Not something
you know.



Why use Biometrics?

Consumers prefer biometric authentication to traditional passwords, Visa says

⌚ Jan 6, 2020 | [Chris Burt](#)

CATEGORIES [Biometrics News](#) | [Financial Services](#)

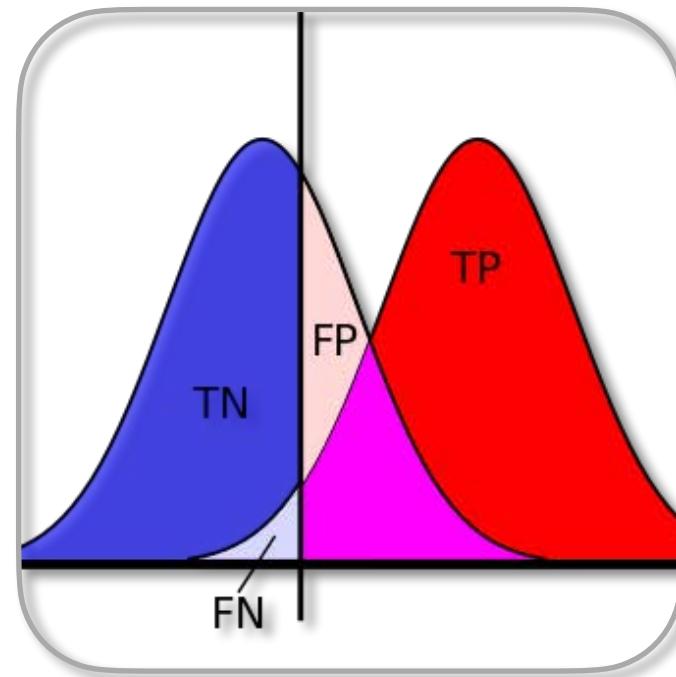


Almost 70 percent of U.S. shoppers did not go through with an online purchase because they either forgot the password, couldn't log in or couldn't receive a one-time passcode, according to research conducted by [Visa](#), while another report from Verizon found that as many as 80 percent of data breaches are caused by compromised and weak passwords.

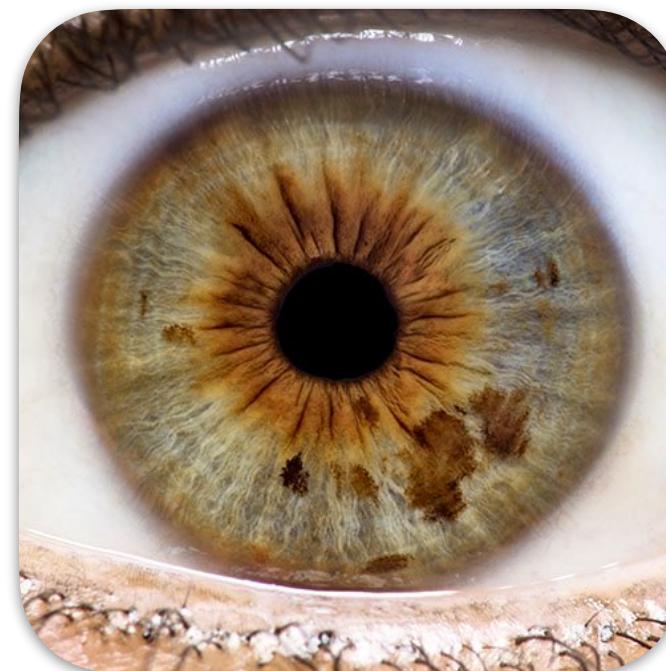
<https://www.biometricupdate.com/202001/consumers-prefer-biometric-authentication-to-traditional-passwords-visa-says>

Course Overview

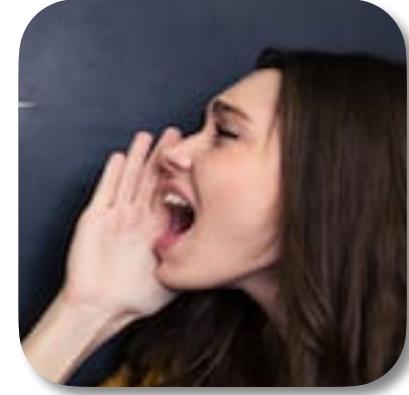
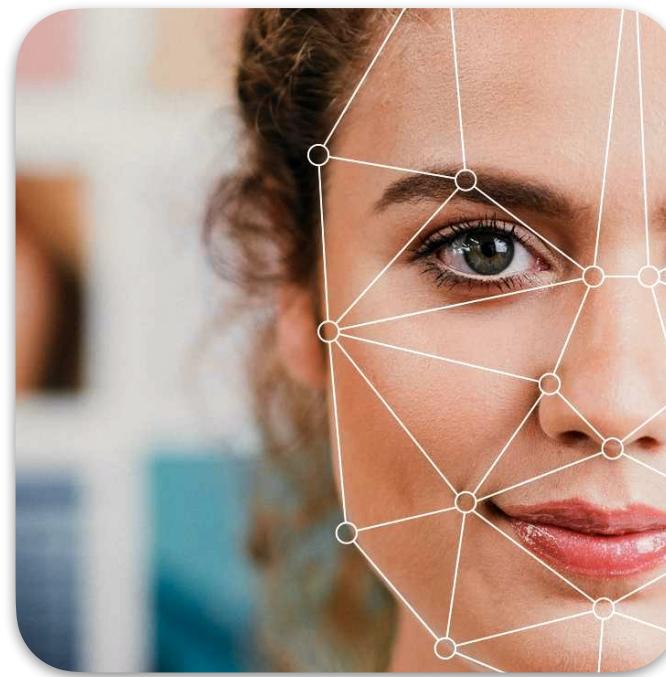
Content



Basics
Concepts
Metrics
Metric implementation



Core Traits (3)
Concepts
Baseline implementation
Data collection
Evaluation
Attacks
Assignments



Alternative Traits and Fusion Concepts



Invited Talks (2)
State of the art
Future work



LOYOLA
UNIVERSITY CHICAGO

Course Overview

Structure (*tentative*)

23 lectures

4 in-class coding days with data collection

2 invited talks

Workload

4 assignments

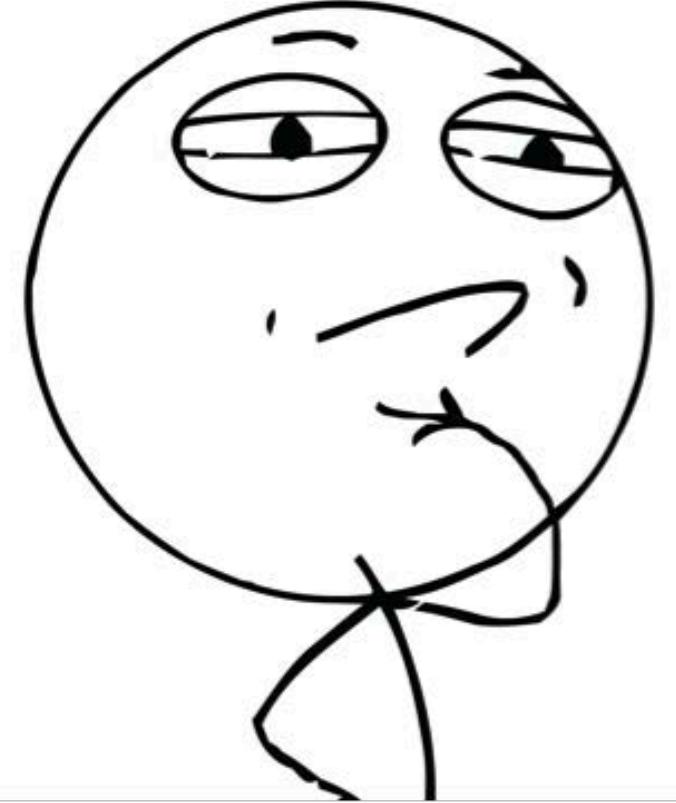
2 exams (midterm and final)

1 project with written report and presentation



Course Overview

CHALLENGE CONSIDERED



Schedule

- 08/25 - Syllabus, Course details.
- 08/27 - Basics I, Biometrics, traits, and systems.
- 09/01 - Labor Day, *no classes*.
- 09/03 - Reading Activity, *no sync class*.
- 09/08 - Basics II, Biometric systems, errors, and metrics.
- 09/10 - 1st Coding Class, Metrics' implementation.
- 09/15 - Fingerprint Recog. I, History and features.
- 09/17 - Fingerprint Recog. II, Acquisition and enhancement.
- 09/22 - Fingerprint Recog. III, Minutiae detection.
- 09/24 - Fingerprint Recog. IV, Data collection.
- 09/29 - 2nd Coding Class, Fingerprint recognition.
- 10/01 - Midterm Preparation, Recap, and project discussion.
- 10/06 - Fall Break, *no classes*.
- 10/08 - Midterm Exam, *good luck*.
- 10/13 - Face Recog. I, Why faces and faces vs. other traits.
- 10/15 - Face Recog. II, Acquisition and enhancement.
- 10/20 - Face Recog. III, Description and matching.
- 10/22 - Face Recog. IV, Deep learning face recognition.
- 10/27 - 3rd Coding Class, Face recognition.
- 10/29 - Iris Recog. I, Why irises and irises vs. other traits.
- 11/03 - Iris Recog. II, Acquisition and enhancement.
- 11/05 - Iris Recog. III, Description and matching.
- 11/10 - 4th Coding Class, Iris recognition.
- 11/12 - Other Traits, Alternative traits and Soft Biometrics.
- 11/17 - Multibiometrics, Data fusion.
- 11/19 - Feature Indexing, Index building and feature querying.
- 11/24 - 1st Invited Talk, *more info soon*.
- 11/26 - Thanksgiving, *no classes*.
- 12/01 - 2nd Invited Talk, *more info soon*.
- 12/03 - **Project presentations, show time**.
- 12/08 - Final Exam, *good luck*.

Course Overview

Grading

	Undergraduate	Graduate
Assignments (4)	40%	25%
Exams (2)	50%	40%
Project	10% (extra)	25%
Participation	10%	10%
On the News	1% (extra)	1% (extra)

A [96, 100)	B+ [88, 92)	C+ [76, 80)	D+ [64, 68)
A- [92, 96)	B [84, 88)	C [72, 76)	D [60, 64)
B- [80, 84)		C- [68, 72)	F (0, 60)

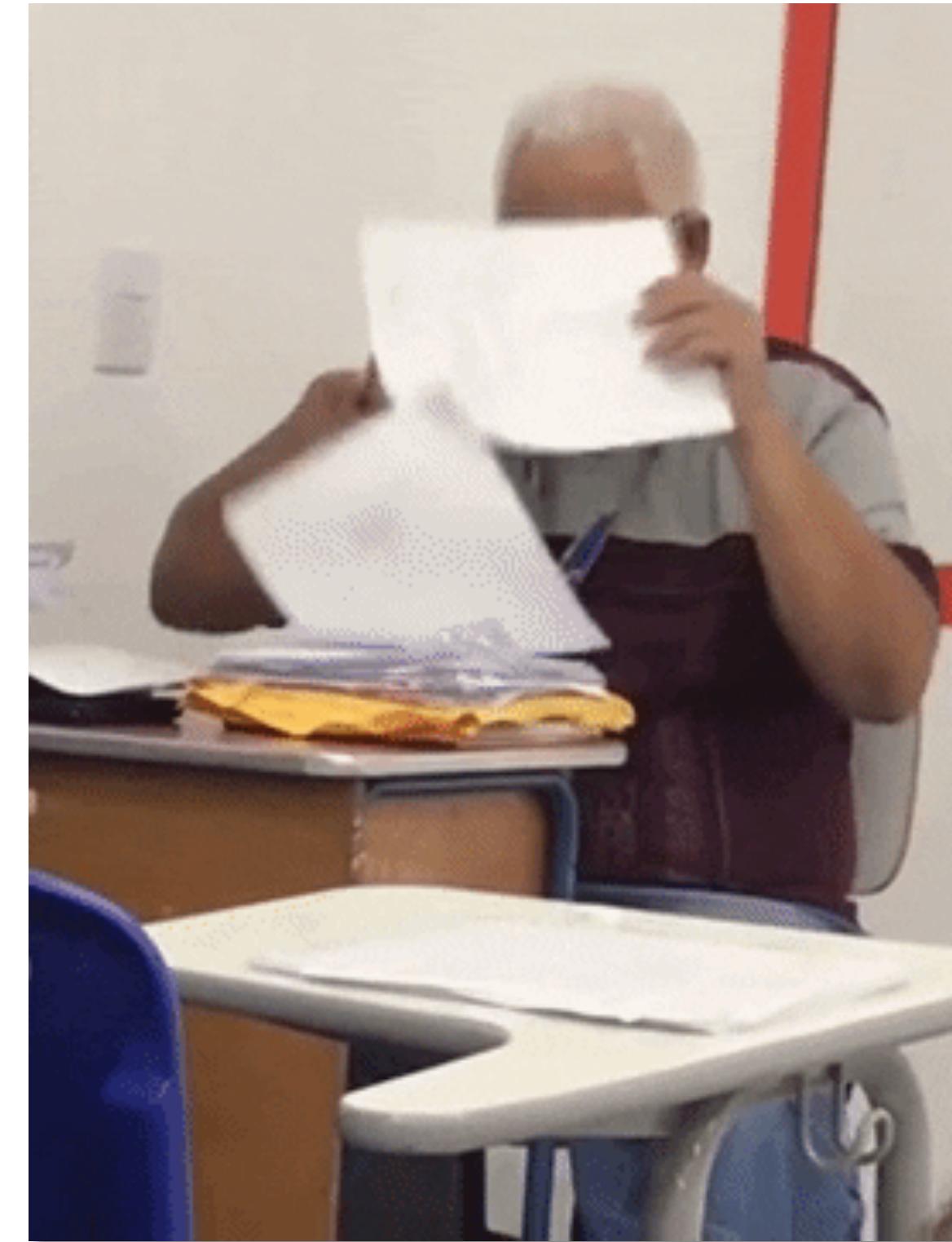


Course Overview

Grading

	Undergraduate	Graduate
Assignments (4)	40%	25%
Exams (2)	50%	40%
Project	10% (extra)	25%
Participation	10%	10%
On the News	1% (extra)	1% (extra)

A [96, 100)	B+ [88, 92)	C+ [76, 80)	D+ [64, 68)
A- [92, 96)	B [84, 88)	C [72, 76)	D [60, 64)
B- [80, 84)		C- [68, 72)	F (0, 60)



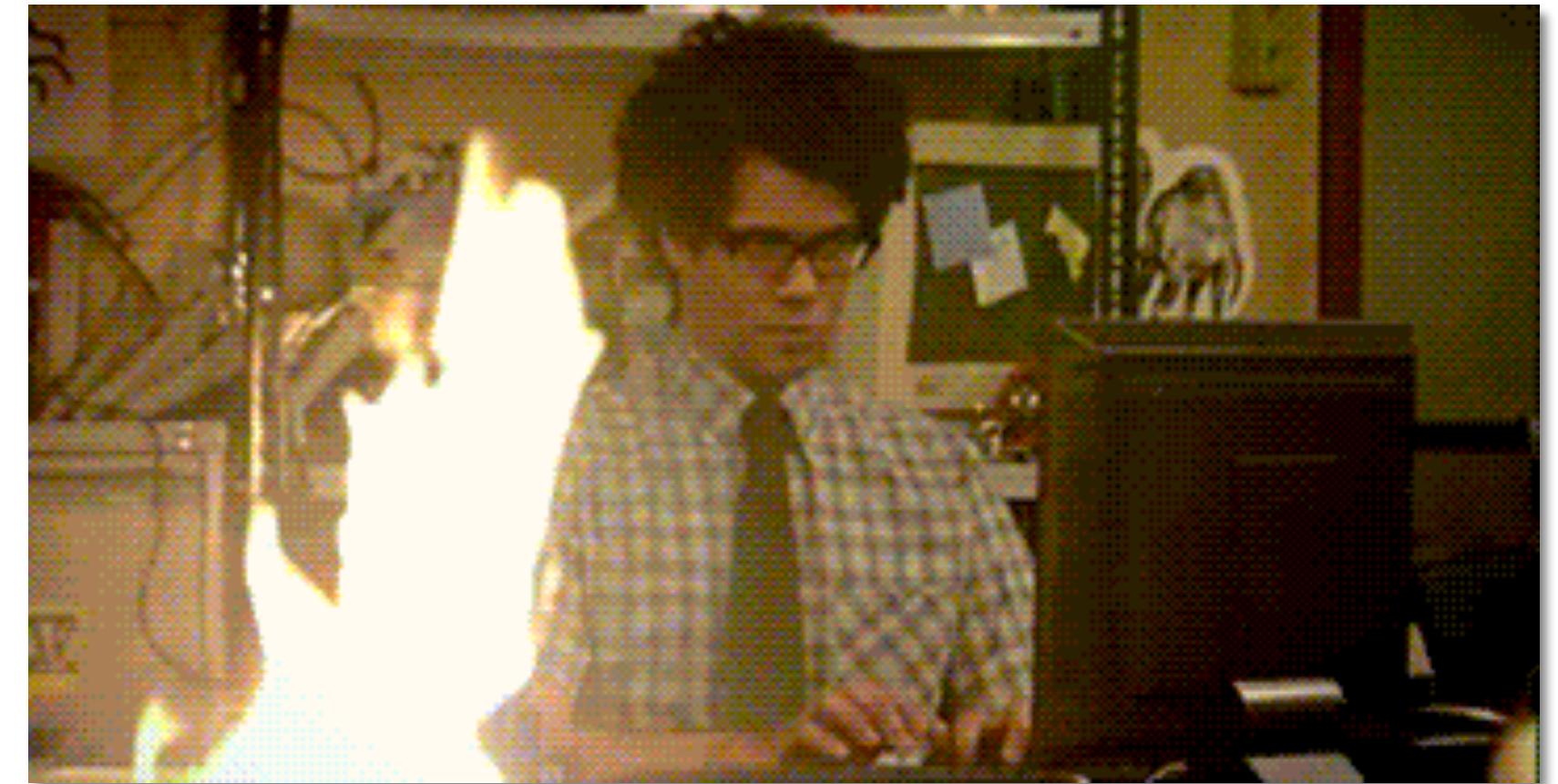
Course Overview

Assignments

Individual take-home activities

Submission through Sakai

Late policy: -10% of the maximum possible grade for each day of delay.



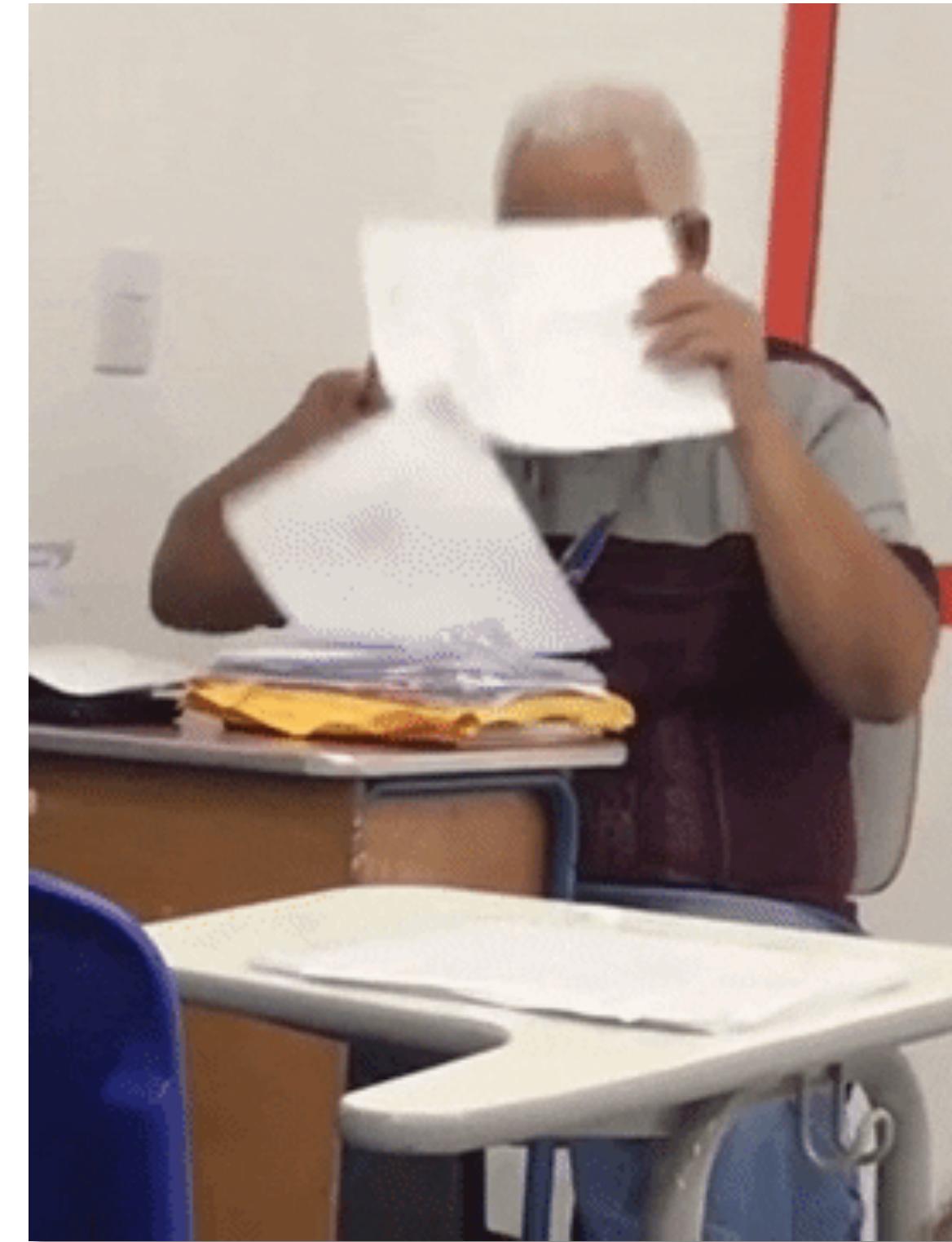
Assignment 1	Assignment 2	Assignment 3	Assignment 4
Metric Collection	Fingerprint Recognition	Face Recognition	Iris Recognition

Course Overview

Grading

	Undergraduate	Graduate
Assignments (4)	40%	25%
Exams (2)	50%	40%
Project	10% (extra)	25%
Participation	10%	10%
On the News	1% (extra)	1% (extra)

A [96, 100)	B+ [88, 92)	C+ [76, 80)	D+ [64, 68)
A- [92, 96)	B [84, 88)	C [72, 76)	D [60, 64)
B- [80, 84)		C- [68, 72)	F (0, 60)

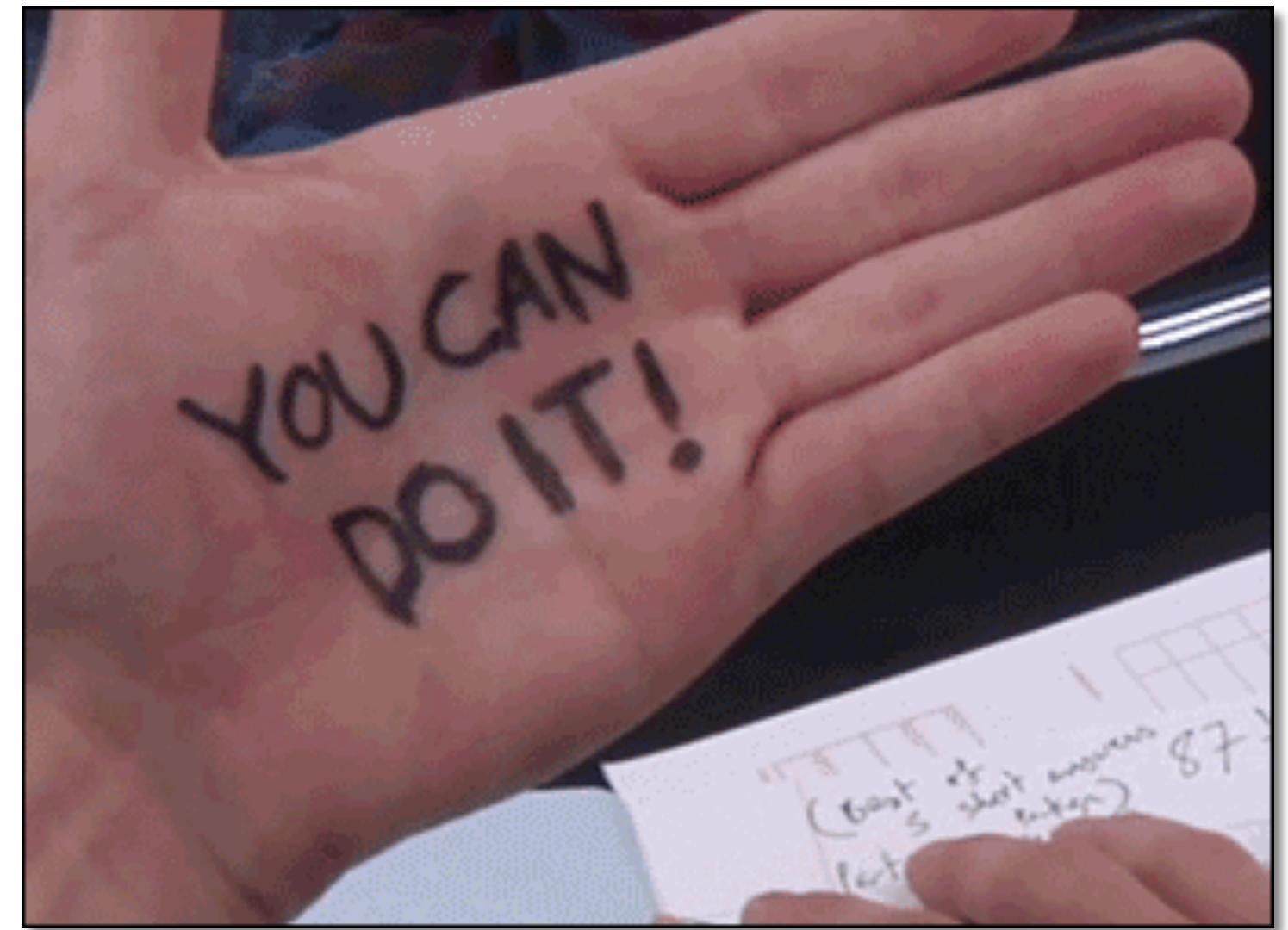


Course Overview

Exams

10/08: in-class written midterm
12/08: in-class written final

One-page cheat sheet is allowed.



Course Overview

Exams

Style example.

[Question 1] (2 points)

Suppose you were hired by a bank company to coordinate the deployment of an access management system to control the entrance of authorized people into the many vaults spread among different branches. The bank directors have heard about Biometrics but are not certain about the benefits of using it. They think using simple access cards and long passwords is as effective and much cheaper than using a biometric system. If it is your duty to change their mind, **what would you say to convince them?**

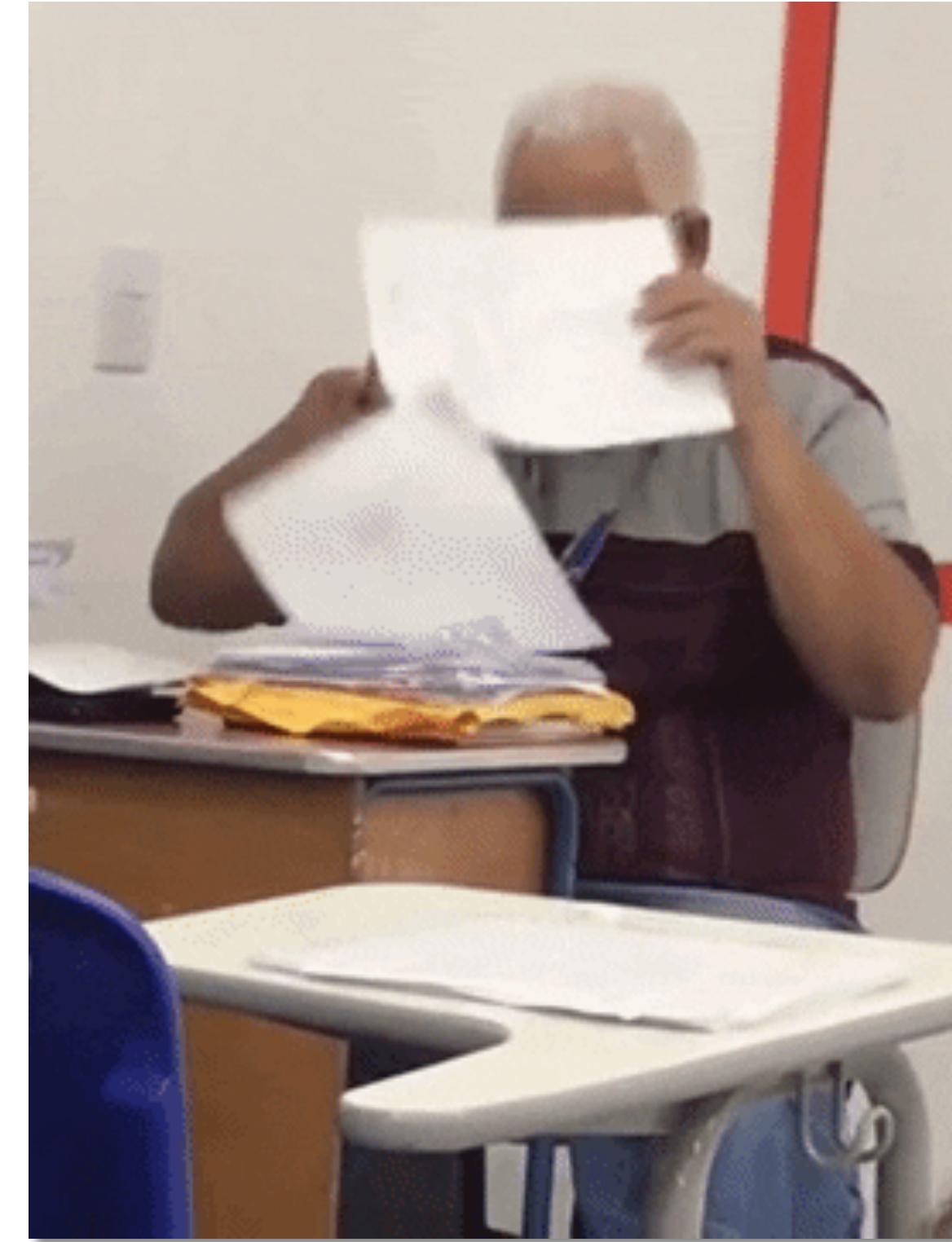
Using biometrics would be a much safer system, since it uses a physical or chemical trait, rather than something that can be stolen as easily as an access card. A password could also be given to somebody else or brute-force searched to produce attacks. Furthermore, it would be more convenient for the authorized people, as forgetting a long password or losing an access card would not be a problem. (Biometrics uses a trait you always have on you). Also, problems like typos, and card damages are more likely to happen than losing a fingerprint, iris, or face.

Course Overview

Grading

	Undergraduate	Graduate
Assignments (4)	40%	25%
Exams (2)	50%	40%
Project	10% (extra)	25%
Participation	10%	10%
On the News	1% (extra)	1% (extra)

A [96, 100)	B+ [88, 92)	C+ [76, 80)	D+ [64, 68)
A- [92, 96)	B [84, 88)	C [72, 76)	D [60, 64)
B- [80, 84)		C- [68, 72)	F (0, 60)



Course Overview

Project

Work alone or in groups.

Provide a written report and perform a presentation.

Optional for undergraduate students (it will grant extra points).



Course Overview

Project

Possible Topics

Presentation attack
(performance, detection, and mitigation)
of fingerprint, face, or iris recognition.

Implementation of recognition of traits
other than fingerprints, face, and iris.

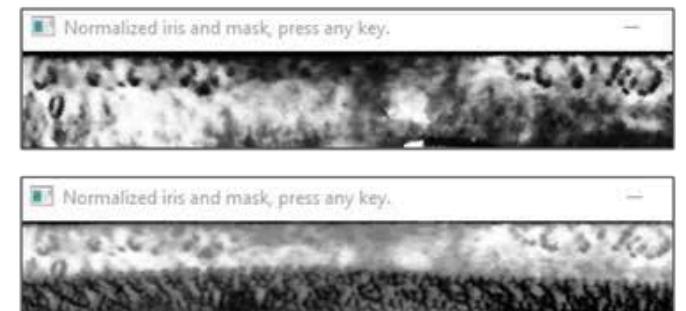
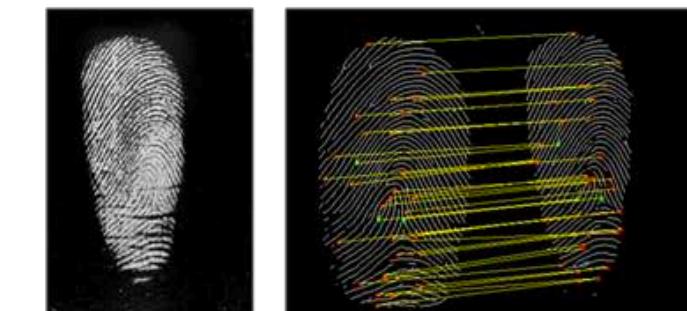


Course Overview

Project

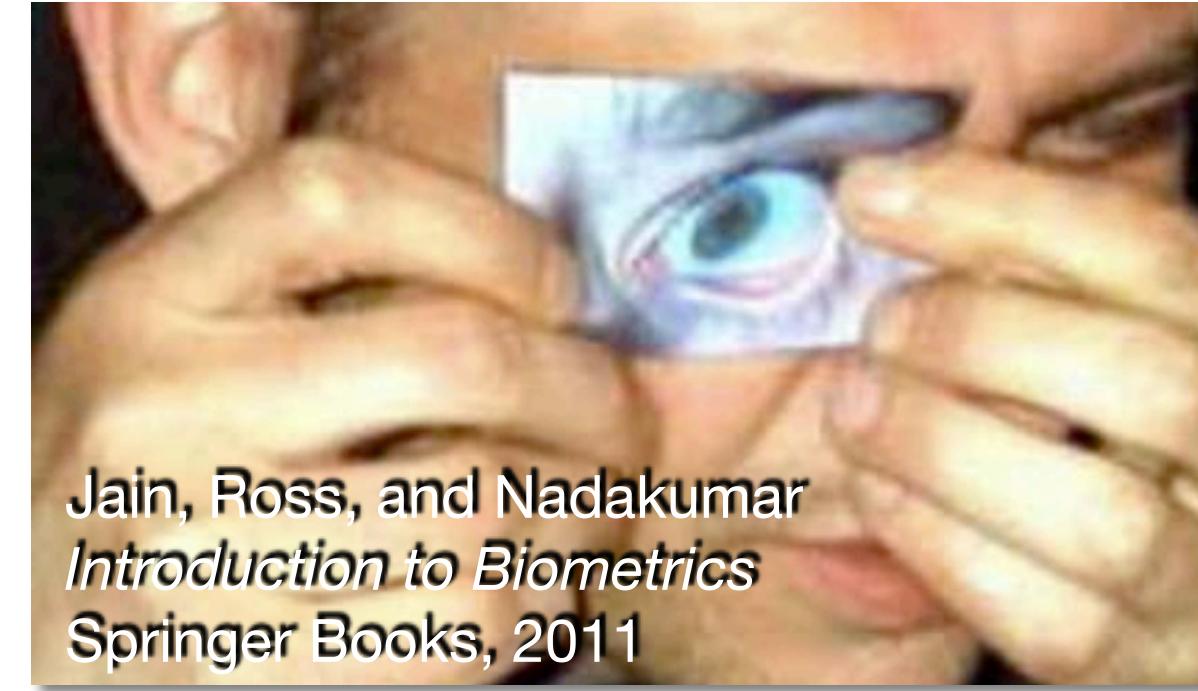
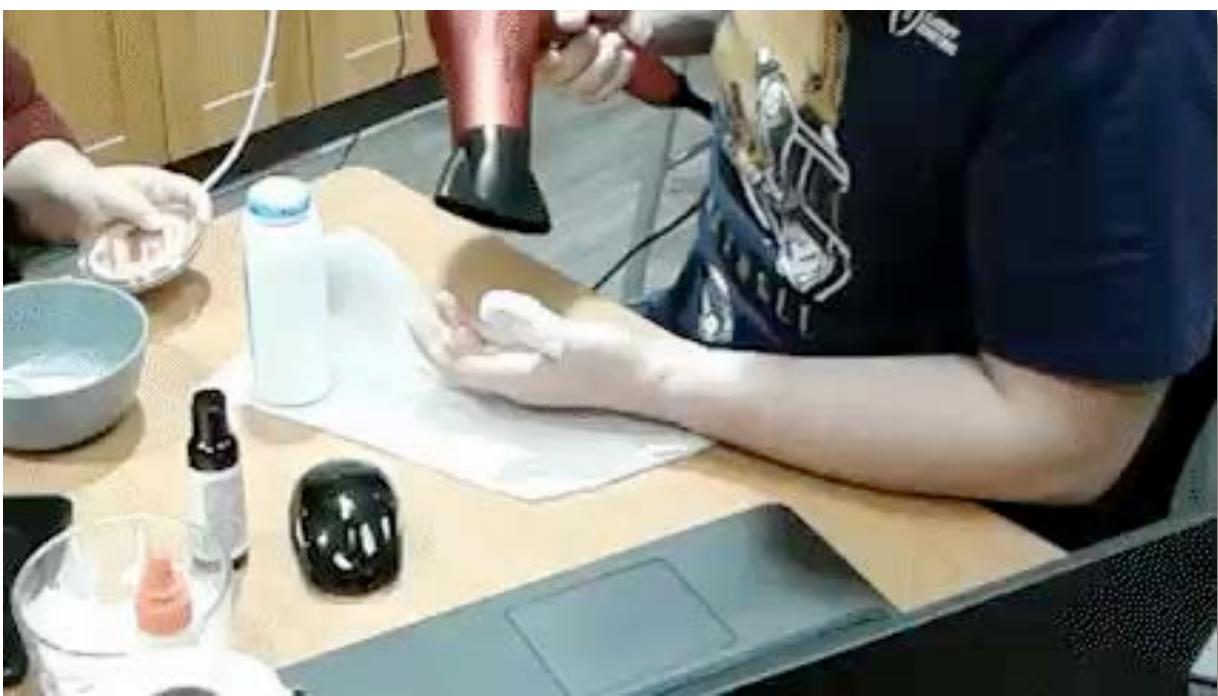
Possible Topics (*continued*)

Implementation of a complete class attendance system.



Presentation and implementation of state-of-the-art scientific publications.

Discussion about the ethical aspects of Biometrics and surveillance.



Jain, Ross, and Nadakumar
Introduction to Biometrics
Springer Books, 2011



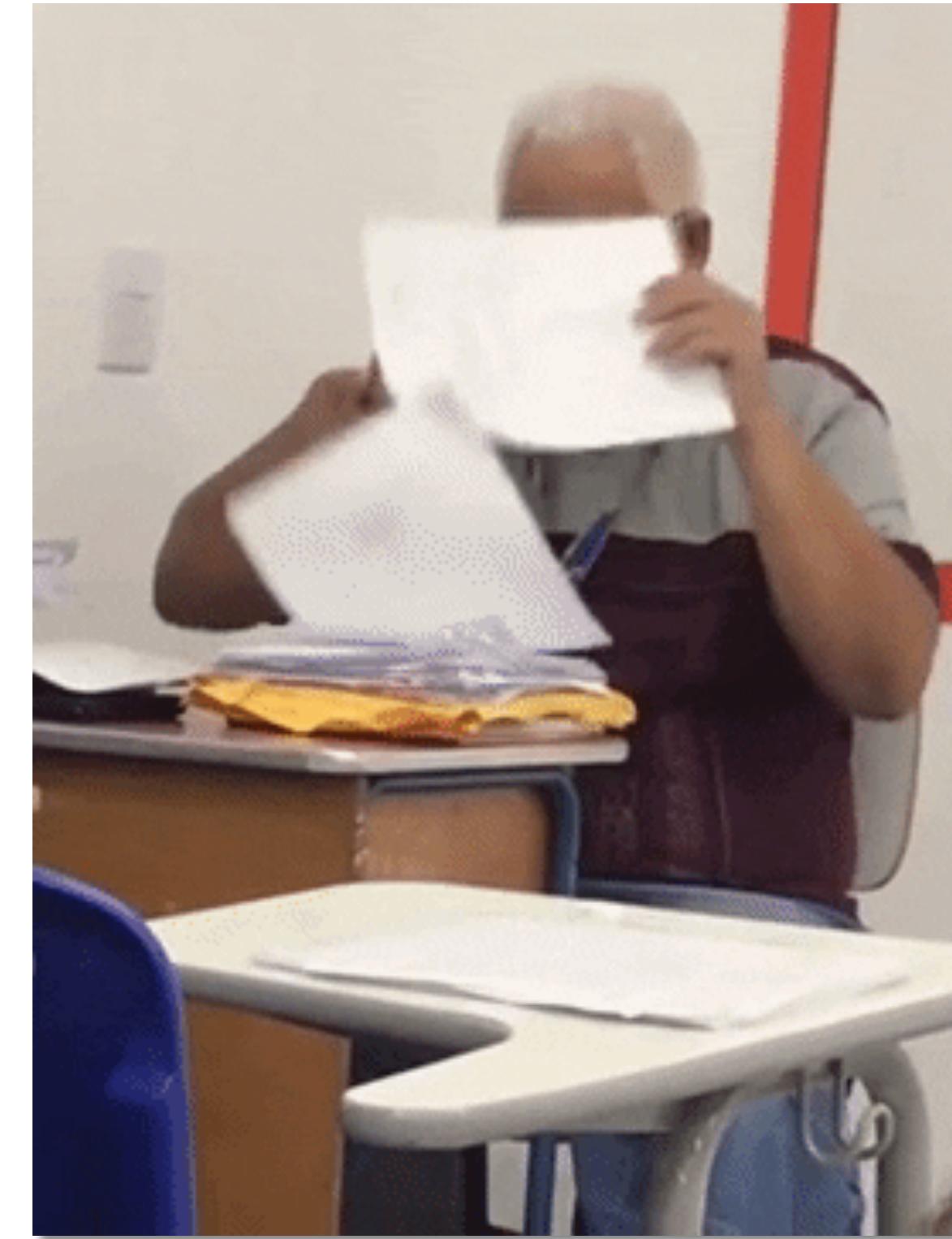
[https://twitter.com/haleshannon/
status/1441134345675153408](https://twitter.com/haleshannon/status/1441134345675153408)

Course Overview

Grading

	Undergraduate	Graduate
Assignments (4)	40%	25%
Exams (2)	50%	40%
Project	10% (extra)	25%
Participation	10%	10%
On the News	1% (extra)	1% (extra)

A [96, 100)	B+ [88, 92)	C+ [76, 80)	D+ [64, 68)
A- [92, 96)	B [84, 88)	C [72, 76)	D [60, 64)
B- [80, 84)		C- [68, 72)	F (0, 60)



Course Overview

Participation

Class Attendance

Every presence counts.

It is possible to get extra points based on interest and proactivity.



Course Overview

Participation

Today-I-missed Statements

Submit **on Sakai** after every class.

Answer one of

What is your biggest question after class?

or

What was the most interesting point you learned today?



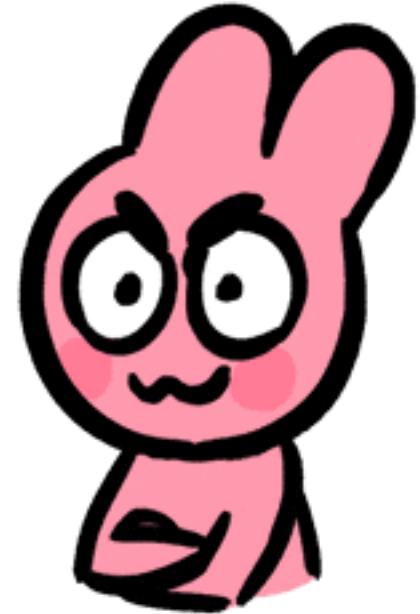
Course Overview

Participation

Grace Cards

Life happens, each student has 3 cards.

Cards can be used to excuse absences or late-delivered work (up to 9 days of delay).

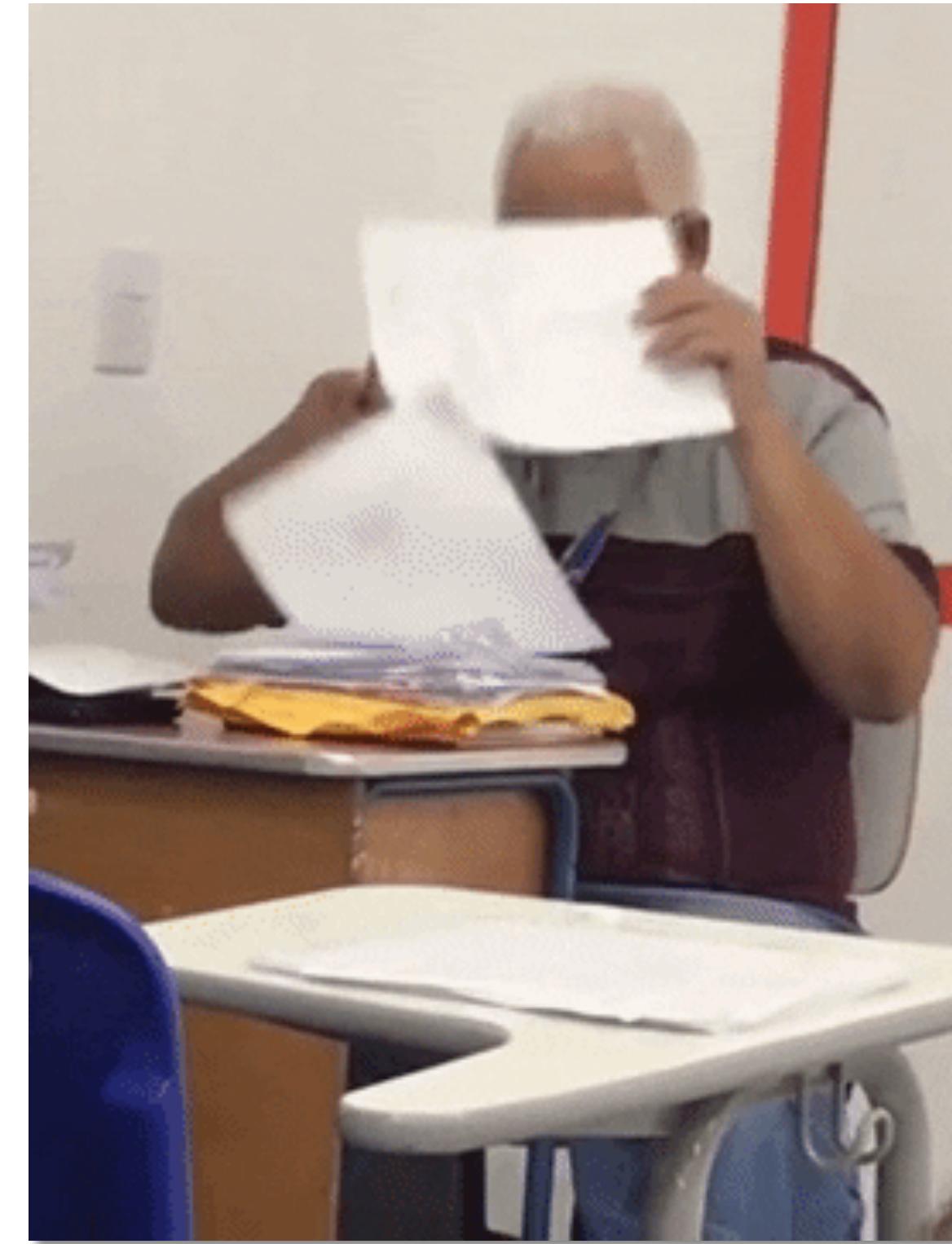


Course Overview

Grading

	Undergraduate	Graduate
Assignments (4)	40%	25%
Exams (2)	50%	40%
Project	10% (extra)	25%
Participation	10%	10%
On the News	1% (extra)	1% (extra)

A [96, 100)	B+ [88, 92)	C+ [76, 80)	D+ [64, 68)
A- [92, 96)	B [84, 88)	C [72, 76)	D [60, 64)
B- [80, 84)		C- [68, 72)	F (0, 60)

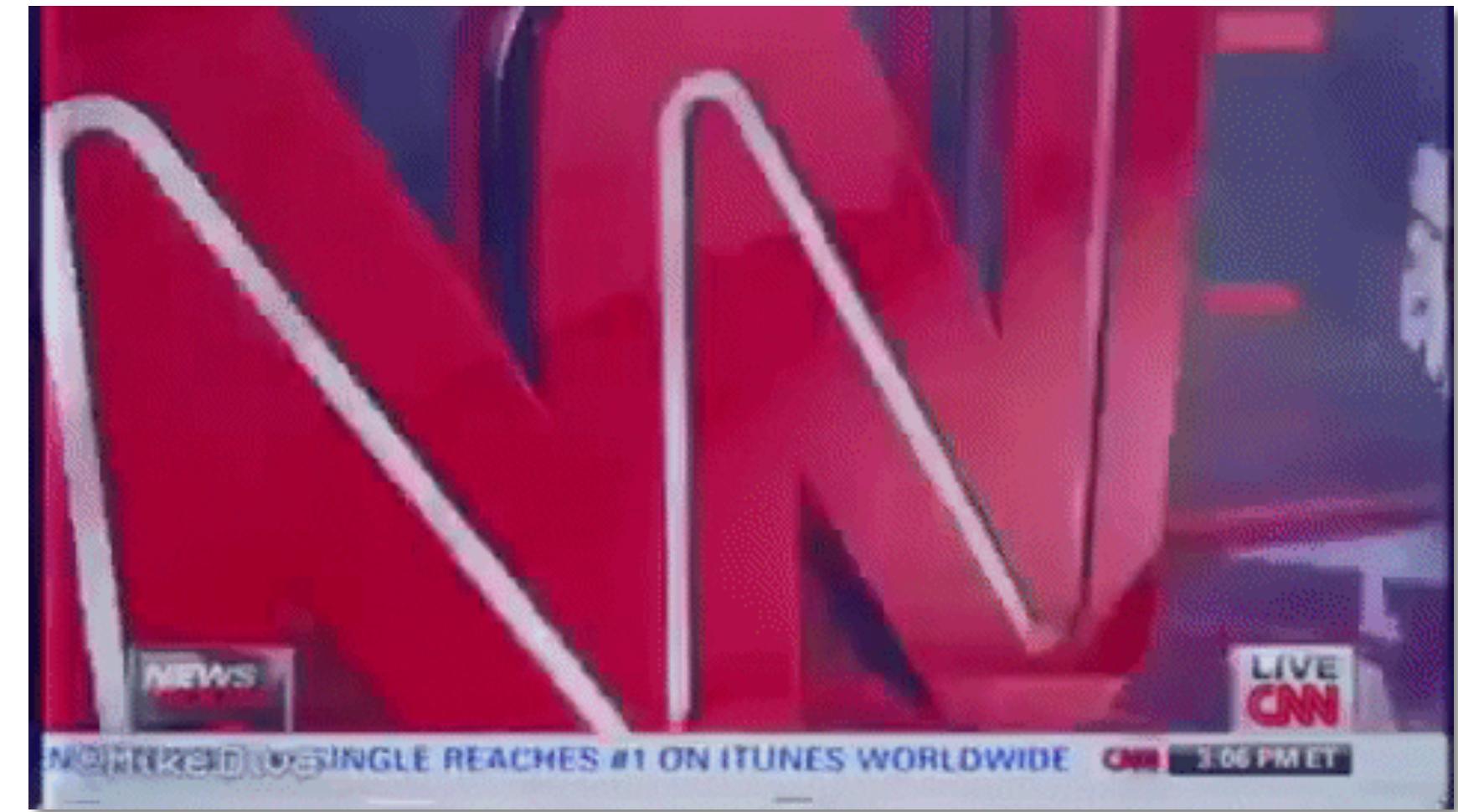


Course Overview

Biometrics on the News

Share with us news you find related to Biometrics.

Get extra points for doing that.



Course Overview

Grading

	Undergraduate	Graduate
Assignments (4)	40%	25%
Exams (2)	50%	40%
Project	10% (extra)	25%
Participation	10%	10%
On the News	1% (extra)	1% (extra)



Code of Honor

Please refer to <https://tinyurl.com/25nvdb8>.
Break it and get an F.

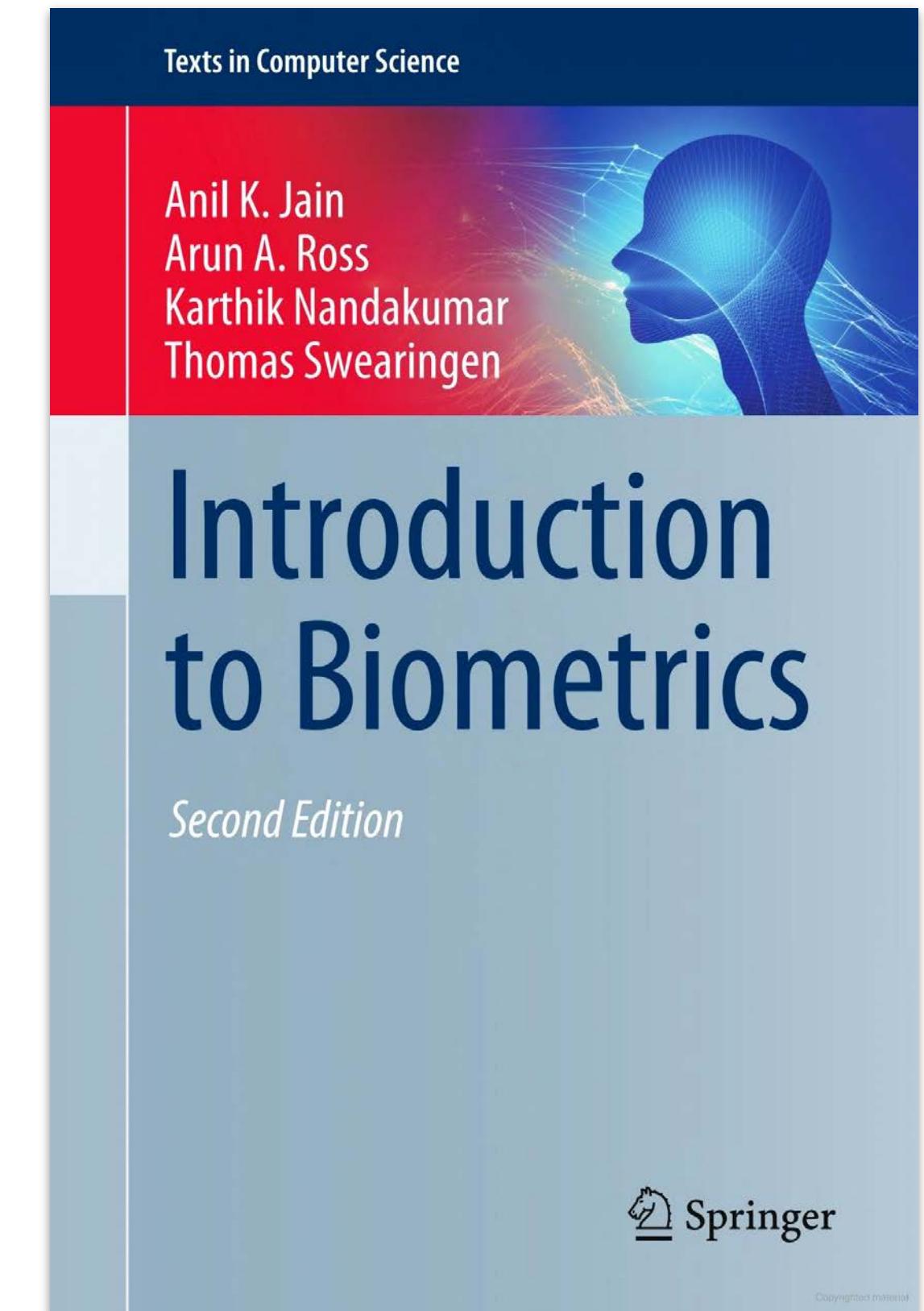


LOYOLA
UNIVERSITY CHICAGO

Course Overview

Bibliography

Jain, Ross, Nandakumar, and Swearingen
Introduction to Biometrics, 2nd Edition
Springer Books, 2025



Course Overview

Pre-requisites

Essential

Programming, basic prob & stats,
and data structures



Desired

Python, Numpy, Matplotlib, and OpenCV

Not sure?

Please talk to me in private.

Course Overview

Data Collection

We'll collect only **our own biometric data** (instructor's and students').

Our data **will only be used** for the purpose of the course.

Our data **will not be shared** with anybody outside the course.

Our data **will be deleted** after the course.



During assignments, folks in need of other publicly available biometric databases are welcome to contact me, so we can take care of privacy and copyright issues.

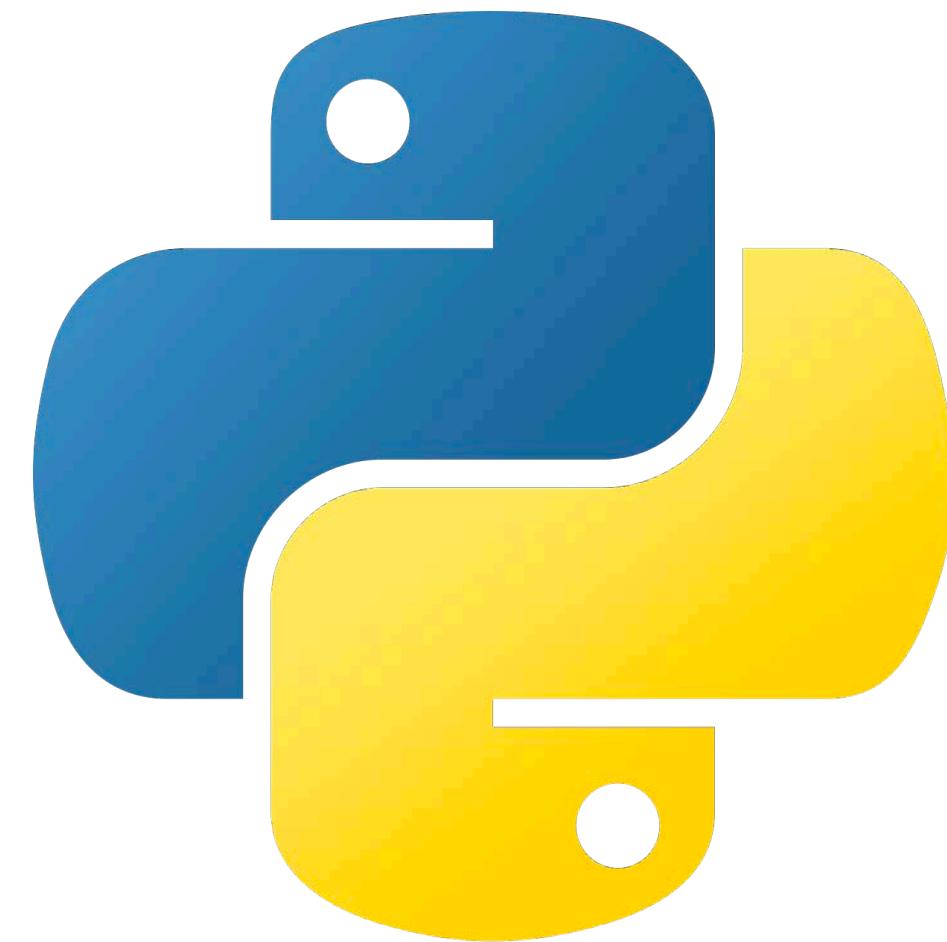
Course Overview

Why Python?

High-level enough

General-purpose enough

Good code readability



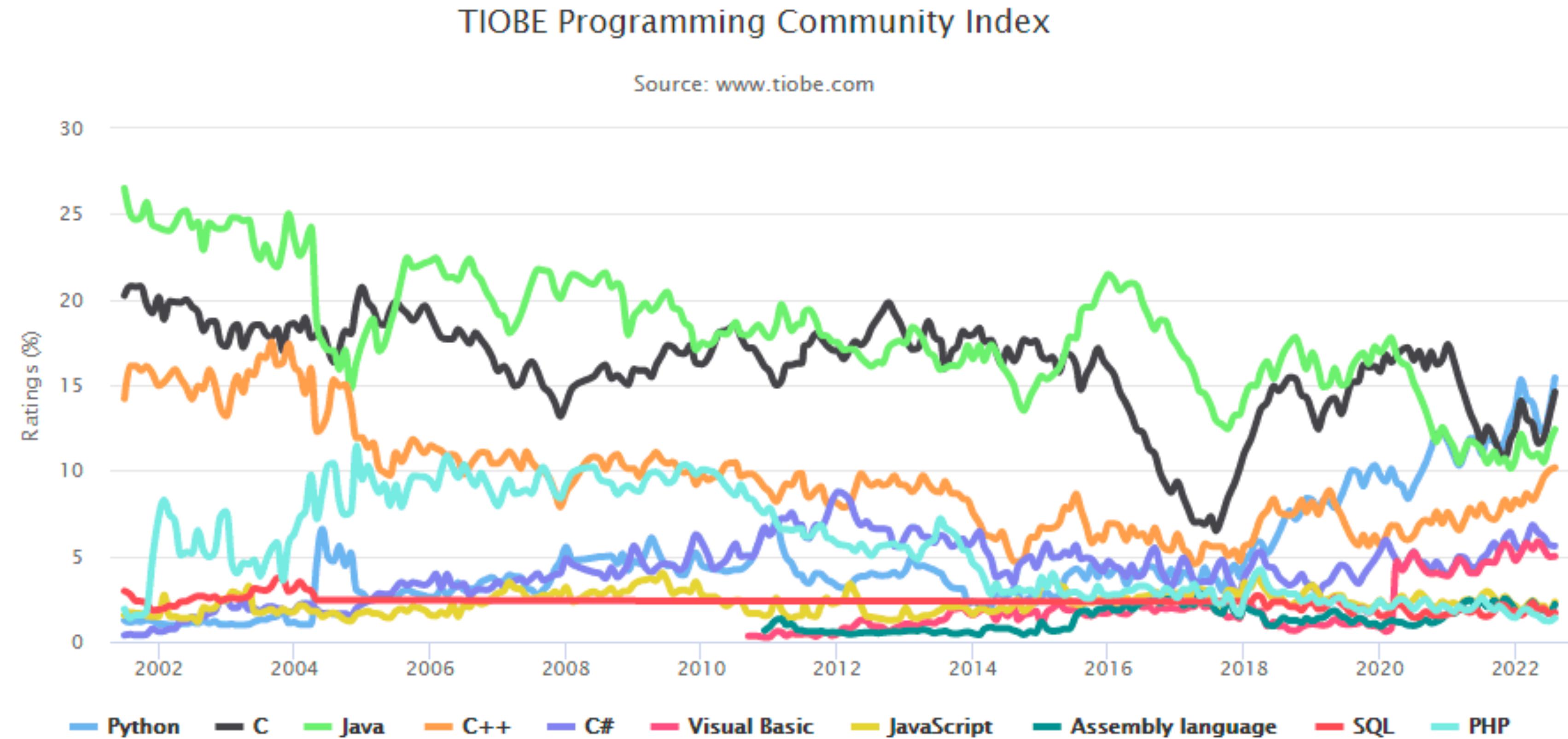
High productivity in data processing
(easy to manipulate strings, lists, and dictionaries).

Large supporting community.

Good libraries supporting scientific computing
(e.g., Numpy, Scipy, Matplotlib).

Course Overview

Why Python?
Increasing popularity



Course Overview

Jupyter Notebooks

Interactive computing

Adoption of the *notebook* interface:
Multiple cells for (1) explanation,
(2) coding, and (3) output of results.



LOYOLA
UNIVERSITY CHICAGO

Course Overview

Jupyter Notebooks

explanation

coding

output

PyCon 2018: Using pandas for Better (and Worse) Data Science

GitHub: <https://github.com/justmarkham/pycon-2018-tutorial>

```
In [1]: import matplotlib.pyplot as plt  
import pandas as pd  
pd.__version__  
  
Out[1]: '0.24.1'
```

Dataset: Stanford Open Policing Project ([video](#))

```
In [2]: # ri stands for Rhode Island  
ri = pd.read_csv('police.csv')
```

```
In [3]: # what does each row represent?  
ri.head()
```

```
Out[3]:
```

	stop_date	stop_time	county_name	driver_gender	driver_age_raw	driver_age	driver_race	violation_raw	violation	search
0	2005-01-02	01:55	NaN	M	1985.0	20.0	White	Speeding	Speeding	
1	2005-01-18	08:15	NaN	M	1965.0	40.0	White	Speeding	Speeding	
2	2005-01-23	23:15	NaN	M	1972.0	33.0	White	Speeding	Speeding	
3	2005-02-20	17:15	NaN	M	1986.0	19.0	White	Call for Service	Other	
4	2005-02-20	17:15	NaN	M	1986.0	19.0	White	Call for Service	Other	

Course Overview

Jupyter Notebooks

The diagram illustrates the structure of Jupyter Notebooks. It features a central box containing code snippets and dataset information, with three boxes on the left labeled 'explanation', 'coding', and 'output' connected by arrows pointing towards the central box.

PyCon 2018: Using pandas for Better (and Worse) Data Science

GitHub: <https://github.com/justmarkham/pycon-2018-tutorial>

```
In [1]: import matplotlib.pyplot as plt  
import pandas as pd  
pd.__version__  
  
Out[1]: '0.24.1'
```

Dataset: Stanford Open Policing Project ([video](#))

```
In [2]: # ri stands for Rhode Island  
ri = pd.read_csv('police.csv')  
  
In [3]: # what does each row represent?  
ri.head()  
  
Out[3]:
```

	stop_date	stop_time	county_name	driver_gender	driver_age_raw	driver_age	driver_race	violation_raw	violation	searched
0	2005-01-02	01:55	NaN	M	1985.0	20.0	White	Speeding	Speeding	
1	2005-01-18	08:15	NaN	M	1965.0	40.0	White	Speeding	Speeding	
2	2005-01-23	23:15	NaN	M	1972.0	33.0	White	Speeding	Speeding	
3	2005-02-20	17:15	NaN	M	1986.0	19.0	White	Call for Service	Other	

Course Overview

Google Colab

<https://colab.research.google.com/>

You'll need a Google account.

Select “New Notebook” on the bottom right of the form.

Do your first “Hello World!”



Course Overview

Local Installation?

Please come to office hours
(tinyurl.com/yv76kpb).

Local installation of Python
and Jupyter.

Any operating system.
Use your CPU.

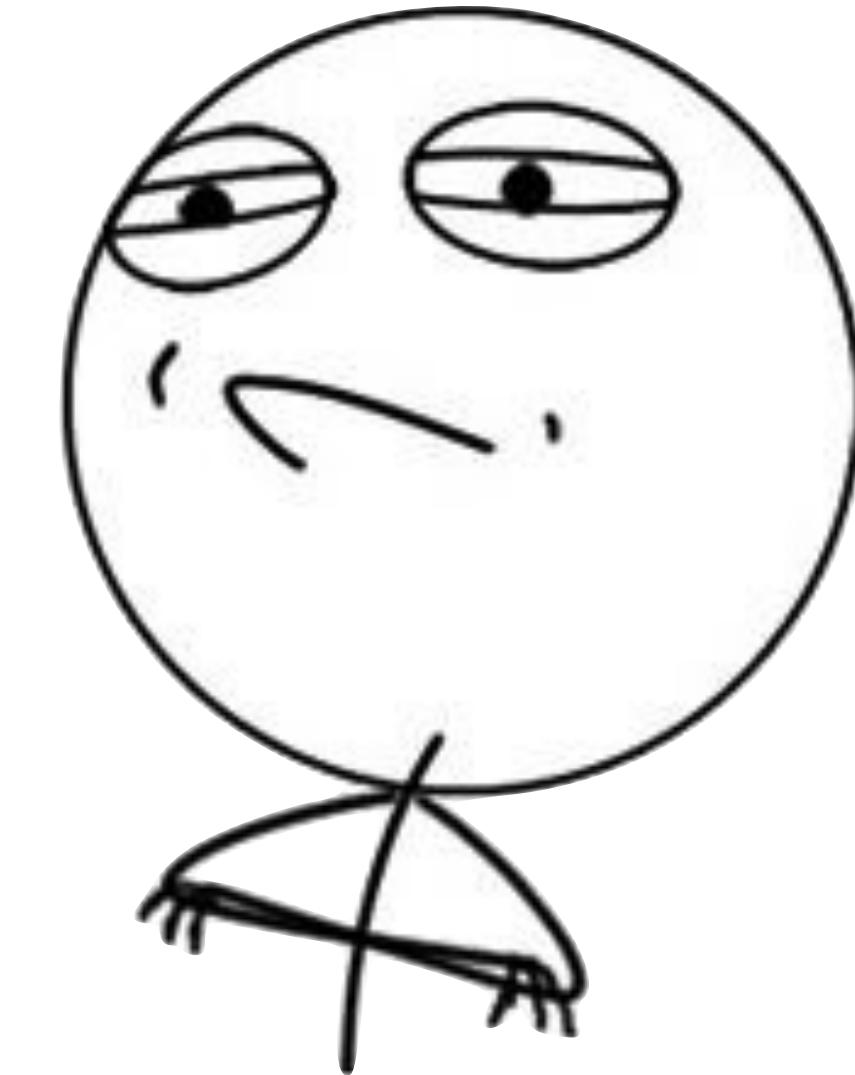


Your Next Tasks

Relax

Any issues? Please come and talk to me.

CHALLENGE ACCEPTED



Sakai is up!

Please visit it as soon as possible.

Important announcements will be made there.

Start filling out your *Today-I-missed* Statement

Please visit sakai.luc.edu/x/BCJs8K.