Does Generative AI Infringe Copyright?

James Grimmelmann
Cornell { Tech | Law School }

AI2
April 10, 2024

Katherine Lee, A. Feder Cooper, & James Grimmelmann, *Talkin' 'Bout AI Generation: Copyright and the Generative AI Supply Chain,* Journal of the Copyright Society of the U.S.A. (forthcoming 2024)

Katherine Lee

A. Feder Cooper

(1) GETTY IMAGES (US), INC
(a company incorporated under the laws of the
(2) GETTY IMAGES INTERNATION
(a company incorporated under the law
(3) GETTY IMAGES (UK) LIMIT
(4) GETTY IMAGES DEVCO UK LII
(5) ISTOCKP
(a company incorporated
(6) THOMAS M. B
(a company incorporated under th

J. DOE 1 and J. DOE 2, individually and on
behalf of all others similarly situated,

Individual and Representative Plaintiffs,

v.

SARAH SILVERMA
CHRISTOPHER Go
RICHARD KADREY

S., P.M., N.G., R.F., J.D. and G.R.,
and on behalf of all others similarly

Plaintiffs,

THE NEW YORK TIMES COMPANY

Plaintiff,

v.

MICROSOFT CORPORATION, OPENAI, INC.,
OPENAI LP, OPENAI GP, LLC, OPENAI, LLC,
OPENAI OPCO LLC, OPENAI GLOBAL LLC,
OAI CORPORATION, LLC, and OPENAI
HOLDINGS, LLC,

Defendants.

GLE DEEPMIND, and

individual;

individual;

N, an individual;

presentative Plaintiffs,

-al

v.

PAUL TREMBLAY, an indi
MONA AWAD, an individu

Indi

OPENAI, INC., a D
Delaware limited par
limited liability corpo
limited liability company; OPENAI STARTU

META PLATFORMS, INC., a Delaware

PLAINTIFFS P.M., K.S., B.B., S.J., N.G., C.B.,
S.N., J.P., S.A., L.M., D.C., C.L., C.G, R.F., N.J.,
and R.R., individually, and on behalf of all others
similarly situated,

Plaintiffs,

v.

OPENAI, INC., a D
L.P., a Delaware lim
Delaware limited liab
Delaware limited liab
GP I, L.L.C., a De
STARTUP FUND I,
OPENAI STARTUP
limited liability comp

vs.

OPENAI LP, OPENAI INCORPORATED,
OPENAI GP, LLC, OPENAI STARTUP FUND
I, LP, OPENAI STARTUP FUND GP I, LLC,
OPENAI STARTUP FUND MANAGEMENT
LLC, MICROSOFT CORPORATION and DOES
1 through 20, inclusive,

SARAH ANDERSEN, an individual;
KELLY MCKERNAN, an individual;
KARLA ORTIZ, an individual,

Individual and Representative Plaintiffs,

v.

STABILITY AI LTD., a UK corporation;
STABILITY AI, INC., a Delaware
corporation; MIDJOURNEY, INC., a
Delaware corporation; DEVIANTART, INC.,
a Delaware corporation,

Defendants.

Does generative AI infringe copyright?

Obviously yes

Copyright in one slide

- "Copyright protection subsists … in original works of authorship fixed in any tangible medium of expression" 17 U.S.C. § 102(a)

- "[T]he owner of copyright … has the exclusive right[] … to reproduce the copyrighted work" 17 U.S.C. § 106(1)

Overview

The Common Crawl corpus contains petabytes of data, regularly collected since 2008.

Choose a crawl...   ⌄

The corpus contains raw web page data, metadata extracts, and text extracts.

Common Crawl data is stored on Amazon Web Services' Public Data Sets and on multiple academic cloud platforms across the world.

Learn how to **Get Started**.

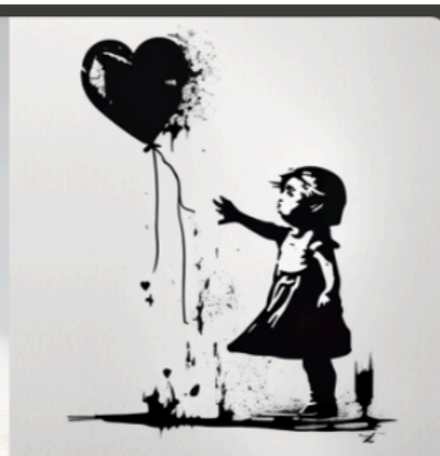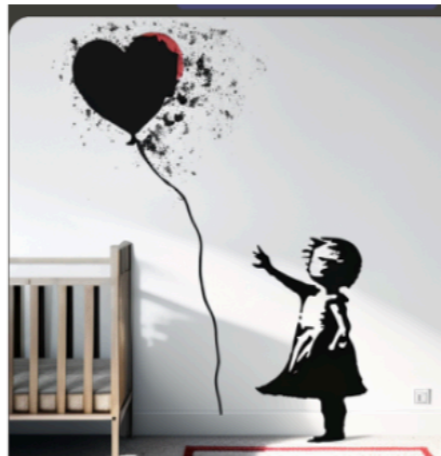## Access to the corpus hosted by Amazon is free.

You may use Amazon's cloud platform to run analysis jobs directly against it or you can download it, whole or in part.

You can search for pages in our corpus using the **Common Crawl URL Index.**

Check out the **Example Projects**, view **Use Cases**, or **Statistics** for our crawls.

Copies everywhere

- There are copies of millions of works

    - In the training datasets

    - In the outputs

    - In the trained models (which can be seen "with the aid of a machine or device," 17 U.S.C. § 101)

- All of these are infringing

- Or I should say, *potentially* infringing …
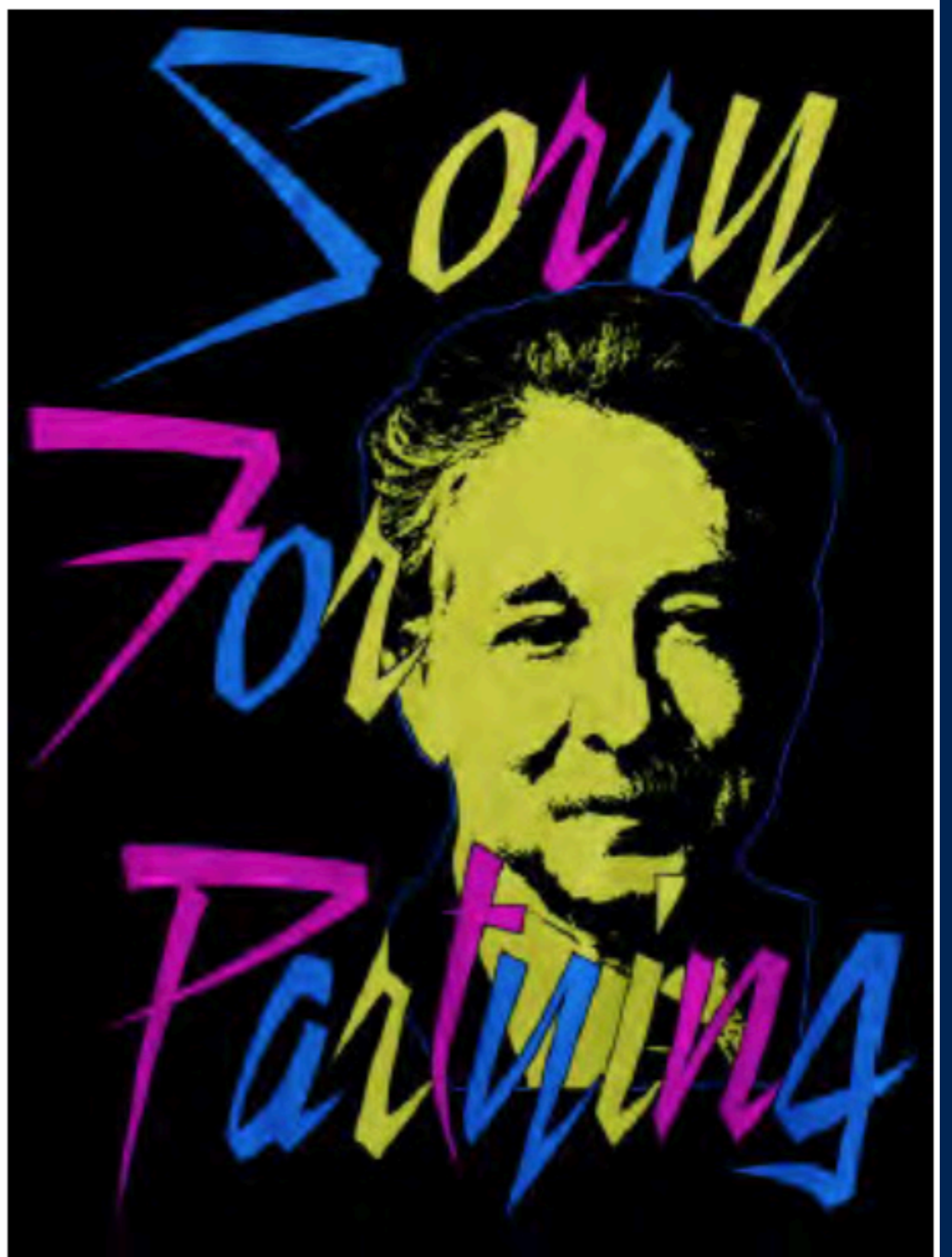
Does generative AI infringe copyright?

Obviously no

Fair use
17 U.S.C. § 107

"[The fair use of a copyrighted work … for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. … [T]he factors to be considered shall include—

(1) the purpose and character of the use, …

(2) the nature of the copyrighted work;

(3) the amount and substantiality of the portion used … ; and

(4) the effect of the use upon the potential market for or value of the copyrighted work."

Photo Credit

Sorry
For
Parting

BHP was founded in Melbourne in 1885. Billiton was founded in 1860. In June 2001, the two companies merged. In the 2003-2004 financial year, it had total revenue of $34.087 billion, a total market capitalization of $136.35 billion. It employed 35,000 people, making it the second-largest mining group company globally. It was listed on stock exchanges in Australia, London, and New York. The company is the world's third-largest supplier of iron ore. The company has more than 100 years of business relationships with China, including importing and exporting mineral products and steel, minerals, and land and sea oil exploration. Board Chairman Don Argus and CEO Charles Goodyear.

Over the past 100 years, from the trade of resource products, domestic and foreign joint ventures to the procurement of mining machinery and equipment in China, BHP Billiton has established a long-term and friendly all-round partnership with Chinese business partners. As the world's second-largest economy, China is critical to developing BHP Billiton. Over the past 40 years, BHP Billiton has supplied China with more than 2 billion tons of iron ore and coking coal, liquefied natural gas, copper concentrate, nickel, and other resource products in the process of China's rapid industrialization and urbanization, providing a solid resource guarantee for China's development. Like many multinational companies, BHP Billiton is a vital participant, witness, contributor, and beneficiary of China's reform and opening up and has forged an indissoluble relationship with China.

BHP Billiton Group is the world's largest production and distribution enterprise of various

Match Overview

79%

Currently viewing standard sources

EN View English Sources (Beta)

Matches

1	Submitted to University... Student Paper	23% >
2	Submitted to Queensla... Student Paper	11% >
3	Submitted to RMIT Uni... Student Paper	6% >
4	Submitted to University... Student Paper	6% >
5	Submitted to Flinders U... Student Paper	5% >
6	Submitted to University... Student Paper	4% >
7	Submitted to Swinburn... Student Paper	4% >

Google Books Ngram Viewer

artificial intelligence,machine learning,copyright

1800 - 2019   English (2019)   Case-Insensitive   Smoothing

(click on line/label for focus)

Fair use everywhere

- Outputs: transformative fair use

- Training data: nonexpressive fair use

- Models: nonexpressive *and* transformative fair use

- … but what if *some* of the outputs infringe?

Does generative AI infringe copyright?
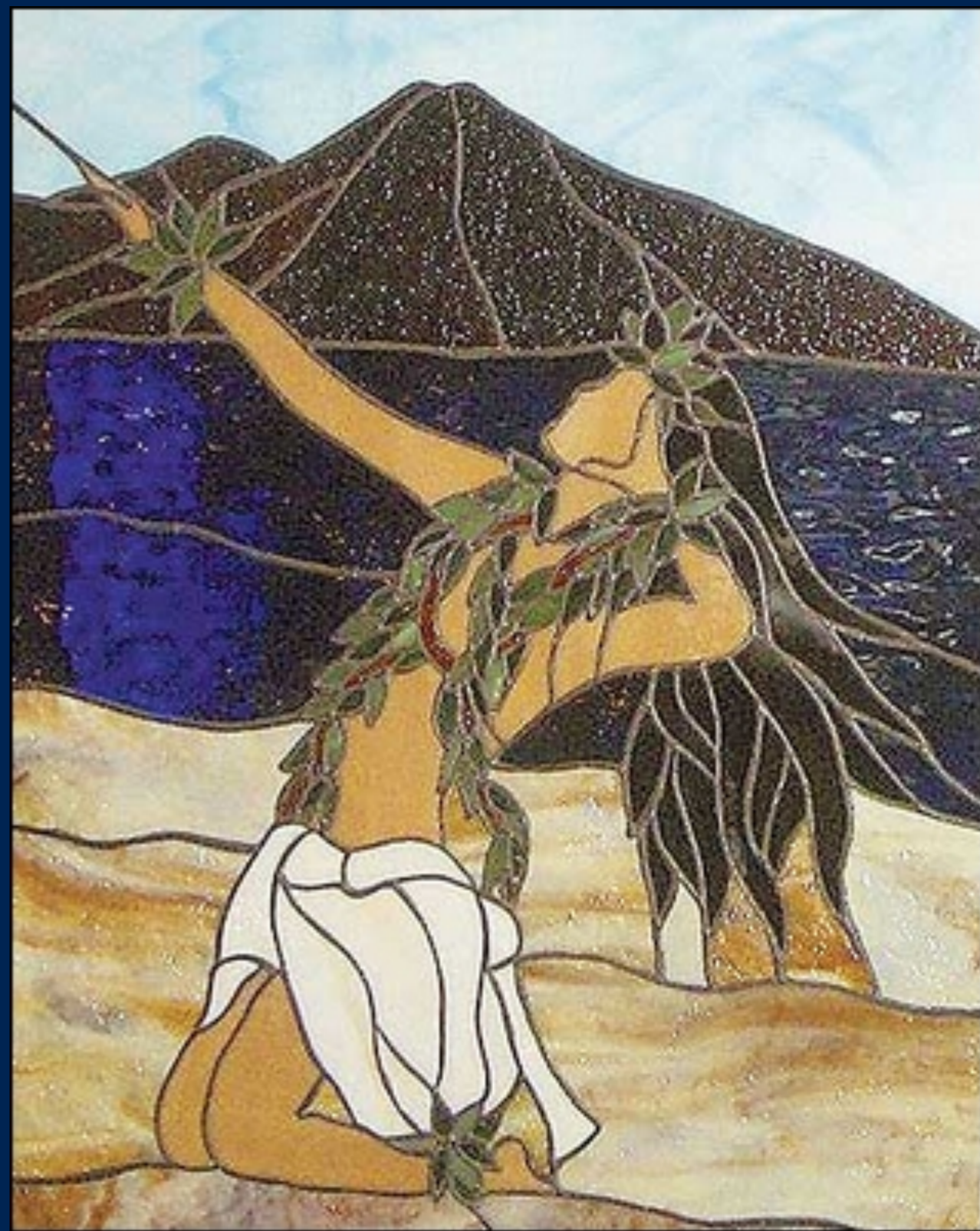
Obviously only by users

Why not notice and takedown?

- Where does the Copyright Act authorize it?

- How can copyright owners find infringing works?

- How can AI companies block infringing outputs?

Stop.

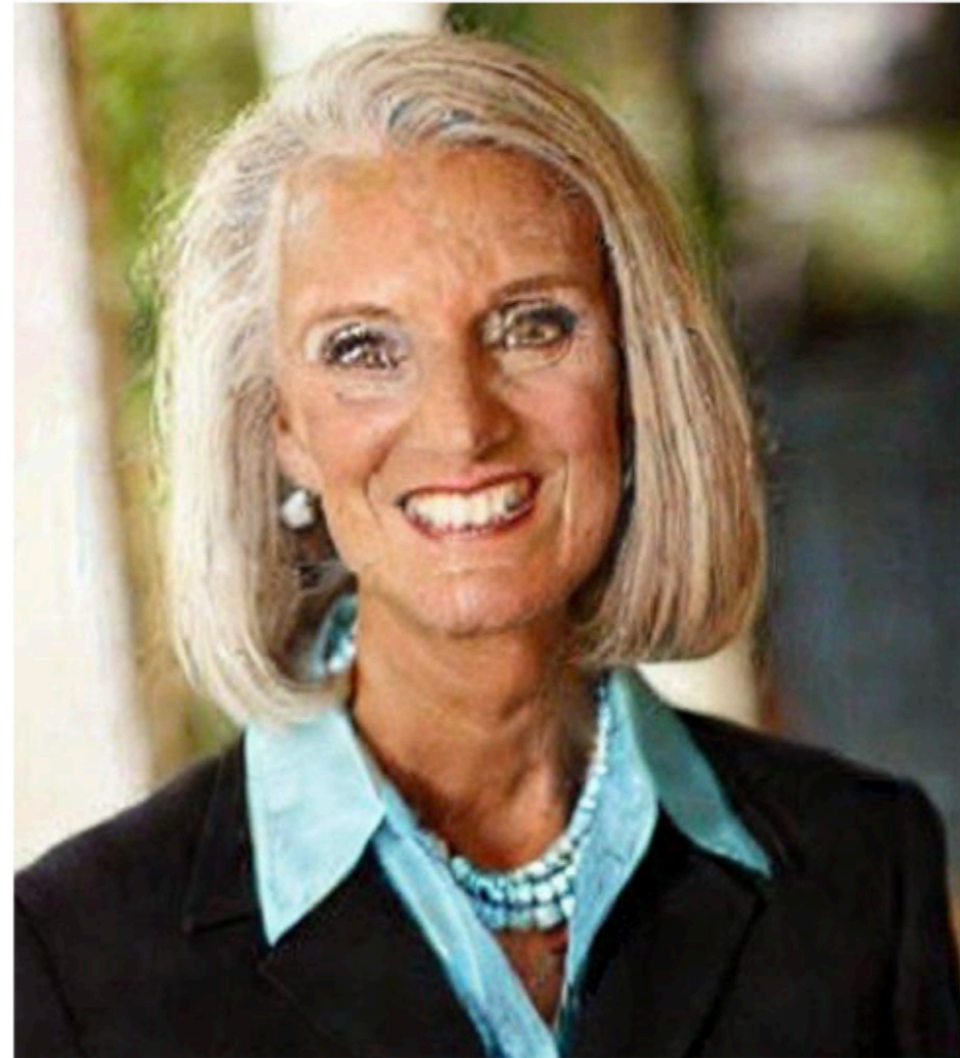# Does generative AI infringe copyright?
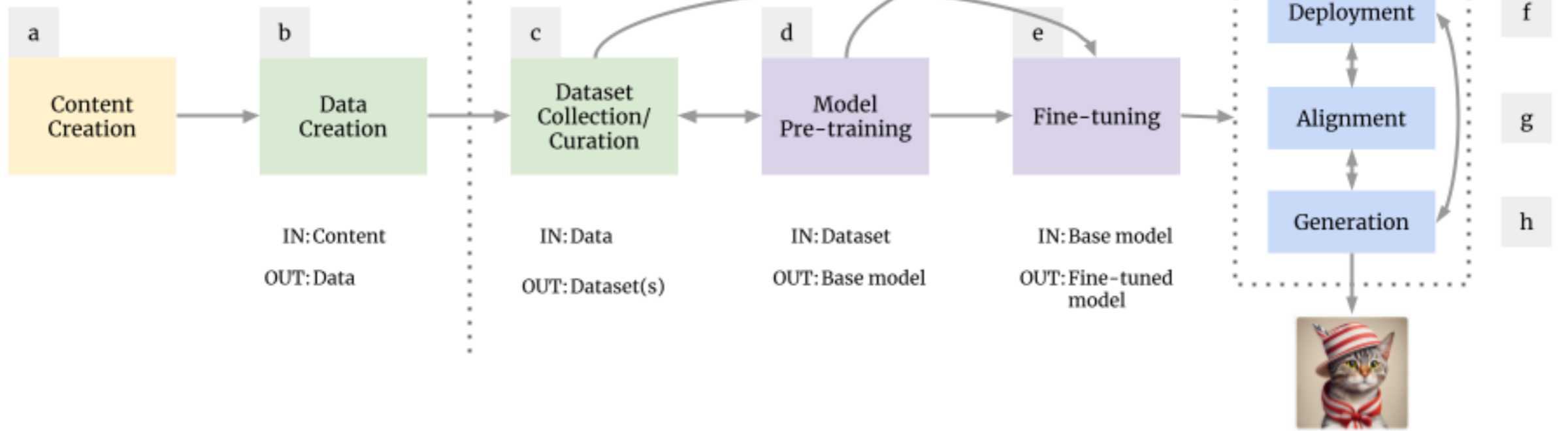
It depends

Training Set

Generated Image

Caption: Living in the light with Ann Graham Lotz

Prompt: Ann Graham Lotz

A diverse supply chain

- A dataset/model/app could …
    - … draw on multiple upstream ones
    - … be used by multiple downstream ones
- Every stage could be …
    - … public or private
    - … proprietary or open
    - … carried out by the same or different parties

Zooming out

Five futures

Strong copyright:

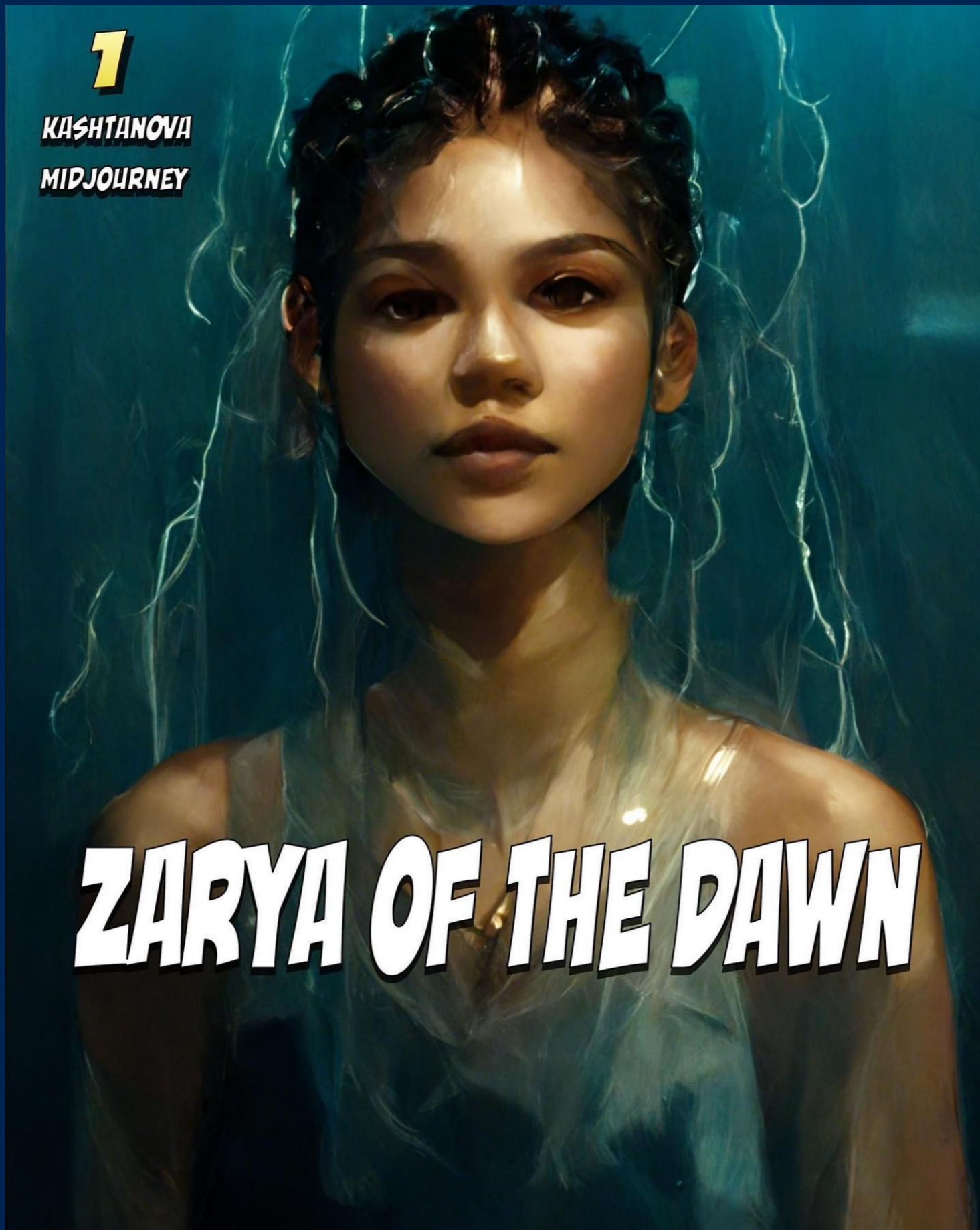   (1) Big Copyright wins again

    (2) The anticommons

Weak copyright:

    (3) Big Tech wins again

    (4) *You* get a model, and *you* get a model, and …

(5) Muddling through

(261)

Anno Octavo

Annæ Reginæ.

An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books, in the Authors or Purchasers of such Copies, during the Times therein mentioned.

Hereas Printers, Bookfellers, and other Perfons have of late frequently taken the Liberty of Printing, Reprinting, and Publishing, or caufing to be Printed, Reprinted, and Publifhed Books, and other Writings, without the Confent of the Authors or Proprietors of fuch Books and Writings, to their very great Detriment, and too often to the Ruin of them and their Families: For Preventing therefore fuch Practices for the future, and for the Encouragement of Learned Men to Compofe and Write ufeful Books; May it pleafe Your Majefty, that it may be Enacted, and be it Enacted by the Queens moft Excellent Majefty, by and with the Advice and Confent of the Lords Spiritual and Temporal, and Commons in this prefent Parliament Affembled, and by the Authority of the fame, That from and after the Tenth Day of April, One thoufand feven hundred and ten, the Author of any Book or Books already Printed, who hath not Transferred to any other the Copy or Copies of fuch Book or Books, Share or Shares thereof, or the Bookfeller or Bookfellers, Printer or Printers, or other Perfon or Perfons, who hath or have Purchafed or Acquired the Copy or Copies of any Book or Books, in order to Print or Reprint the fame, fhall have the fole Right and Liberty of Printing fuch Book and Books for the Term of One and twenty Years, to Commence from the faid Tenth Day of April, and no longer; and that the Author of any Book or Books already Compofed and not Printed and Publifhed, or that fhall hereafter be Compofed, and his Affignee, or Affigns, fhall have the fole Liberty of Printing and Reprinting fuch Book and Books for the Term of Four-

Ttt 2   teen