

From Creation to Caution: The Effect of Generative AI on Online Arts Market

Sijie Lin

University of Toronto

August 27, 2024

Motivation: Copyright Concerns Cause Pushback Against Generative AI

- ▶ The power of generative AI lies in its extensive training on a substantial volume of data, much of which consists of copyrighted materials.
- ▶ Multiple copyright lawsuits across different industries:
 1. Andersen v. Stability AI Ltd.: Artists against AI companies
 2. Doe v. GitHub: Programmers against GitHub
 3. Authors Guild v. OpenAI: Authors including George Martin (Game of Thrones) sue OpenAI
- ▶ Anti-AI protests on online art platforms: DeviantArt, ArtStation, LOFTER

How do copyright concerns impact the decision of creators?

Motivation: Why Is This Question Important?

1. Relevant to knowledge spillover

- ▶ May discourage future human innovation due to restricted access to existing content (Murray&Stern 2007, Williams 2013, Galasso&Schankerman 2015, Nagaraj 2018, Biasi&Moser 2021)
- ▶ Could harm future productivity of AI: AI models can collapse if it is trained on AI-generated content (Shumailov, Shumaylov, Zhao, Gal, Papernot, Anderson 2023)

2. Contribution to literature:

- ▶ Existing literature: IP protection↓, incentive for innovation↓, trade secrecy↑
 - ▶ Arrow 1962, Moser 2012, Giorcelli&Moser 2020
- ▶ This paper: IP protection↓, incentive for innovation **unchanged**, trade secrecy↑

This Paper

How do copyright concerns impact the decision of creators?

Find an empirical setting to answer this question

- ▶ DeviantArt, a leading online arts platform
 - ▶ Artists display and sell artworks
 - ▶ Companies (ad, games, etc) recruit employees
 - ▶ One of the largest platforms
- ▶ Nov 11, 2022: DeviantArt introduced DreamUp, an AI image generator

“Confused artists discover their work will be used for AI training by default.”

—— *Ars Technica*, Nov 11, 2022

This Paper

How do copyright concerns impact the decision of creators?

Find an empirical setting to answer this question

- ▶ DeviantArt, a leading online arts platform
 - ▶ Artists display and sell artworks
 - ▶ Companies (ad, games, etc) recruit employees
 - ▶ One of the largest platforms
- ▶ Nov 11, 2022: DeviantArt introduced DreamUp, an AI image generator

“Confused artists discover their work will be used for AI training by default.”

— *Ars Technica*, Nov 11, 2022

Why choose this platform?

1. Earliest copyright concerns on online art platform
2. Much attention on ongoing lawsuit (Andersen v. Stability AI Ltd.)

Main Findings

1. Diff-in-diff: 22% decline in publication volume of non-AI digital artists
2. Multi-homing artists: only withhold artworks on DeviantArt, not on Instagram
3. No evidence of quality change in published artworks

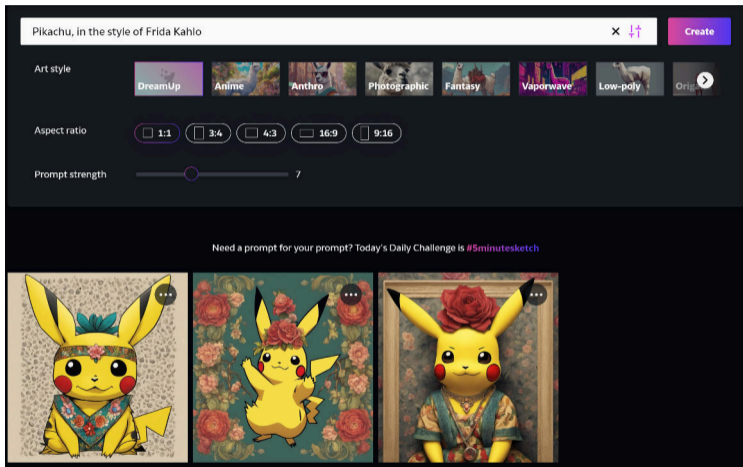
Timeline

There are 3 other well-known AI image generators: Stable Diffusion, Midjourney and DALL-E 2.

1. 12 July 2022: Midjourney image generation platform first entered open beta
2. 22 August 2022: Stability AI announced the public release of stable diffusion
3. 28 September 2022: DALL-E 2 was opened to anyone, and the waitlist requirement was removed
4. **11 November 2022: DreamUp (based on Stability AI) was introduced on DeviantArt**
5. 30 November 2022: ChatGPT released

Features of AI Image Generators

- ▶ Can specify style of a particular artist
- ▶ Time-efficient
 - ▶ 60 seconds for 3 artworks
- ▶ Cheap
 - ▶ < 10 cents per prompt



Data

Data

6835 artists from daily featured section on DeviantArt

1. Information on DeviantArt

- ▶ Artists demographics
- ▶ History of publication: publish date; number of views, downloads, favorites, comments; description and tags
- ▶ Other platforms they are using

2. Information on Instagram

- ▶ Obtain data of professional/business accounts
- ▶ History of publication: publish date; number of likes, comments; description and tags


Dbn of Multi-homing Artists	Artists%
Instagram	63.7%
Twitter	48.4%
Facebook	36.4%
YouTube	21.3%
Tumblr	18.1%
Fraction of Multi-homing artists	77%

Data: Identify AI Artworks

Title, Tags, Description

Support my work and get exclusive perks [View Subscriptions](#)

Add to Favourites | Comment

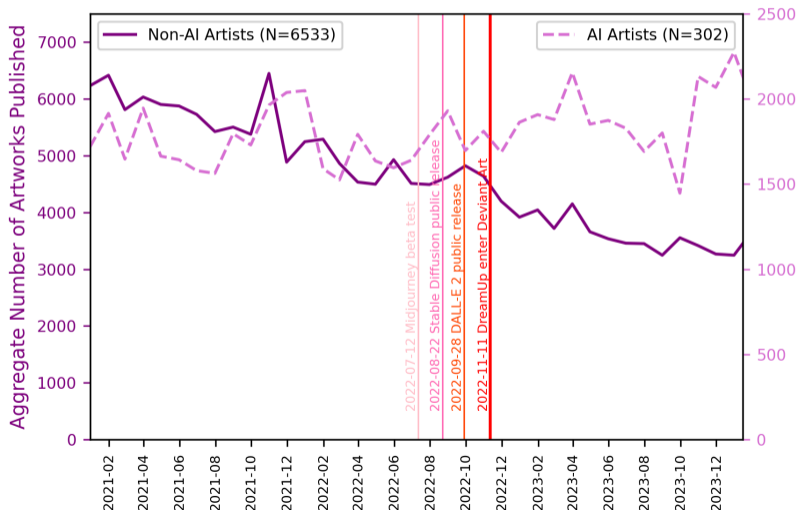
 **Midjourney 4967**
by **Javier-LLuesma** PRO [+ Watch](#)

★ 41 Favourites | 💬 2 Comments | 👁 1.7K Views

ai digitalart digitalartwork digitalillustration digitalpainting exclusive
premium prompt superior lluesma aiart artworkdigital midjourney
midjourneyart midjourneyartwork Less

Data: Time Trends

AI artists increase publication, while Non-AI artists decrease



Identification Strategy

Identification Strategy: Difference-in-Differences

Control Group:

Non-AI Artists specialize in **Artisan Crafts**

- ▶ Usually hand-made
- ▶ Jewelries, dolls, cross stitch, etc.
- ▶ **Less exposed to AI**



Treatment Group:

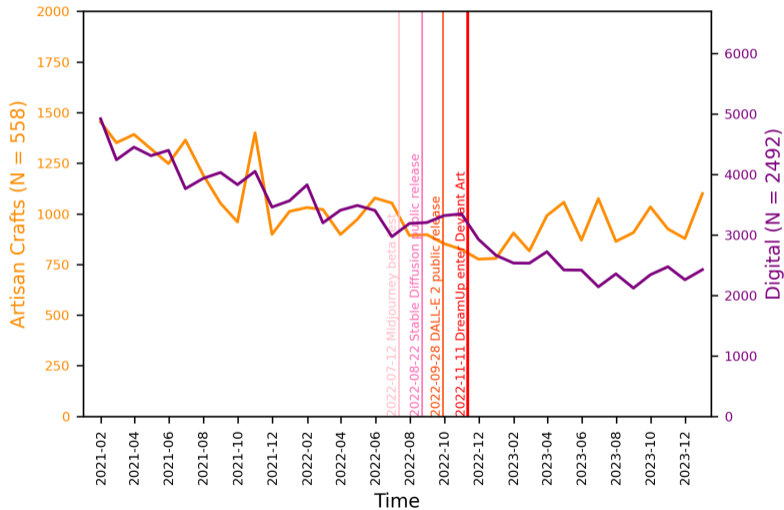
Non-AI Artists specialize in **Digital Arts**

- ▶ Usually made with Adobe Photoshop, Procreate on drawing tablets or iPad
- ▶ Dragons, fantasy, wallpapers, etc.
- ▶ **More exposed to AI**



Identification Strategy: Difference-in-Differences

Similar Trends Before Shock



Results

Result 1: 22% Decline of Publication Volume on DeviantArt

$$Artwork_{it} = \beta_0 + \beta_1 Post_t \times Treated_j + \delta_i + \delta_t + \epsilon_{it}$$

Table 1: Effect on Artist Publication Volume

Sample: Dep Var:	All Users Artworks on DeviantArt		Instagram Users Artworks on DeviantArt		Instagram Users Artworks on Instagram	
	(1) PPML	(2) OLS	(3) PPML	(4) OLS	(5) PPML	(6) OLS
$Post_t \times Treated_j$	-0.25*** (0.09)	-0.21* (0.12)	-0.29** (0.15)	-0.31 (0.20)	-0.09 (0.08)	-0.04 (0.41)
Pre-Treatment Mean		1.50		1.70		3.65
Artist FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N(Artist-Month)	109,800	109,800	31,968	31,968	31,968	31,968
N(Artists)	3,050	3,050	888	888	888	888
Adjusted R ²	0.53	0.58	0.44	0.33	0.49	0.59

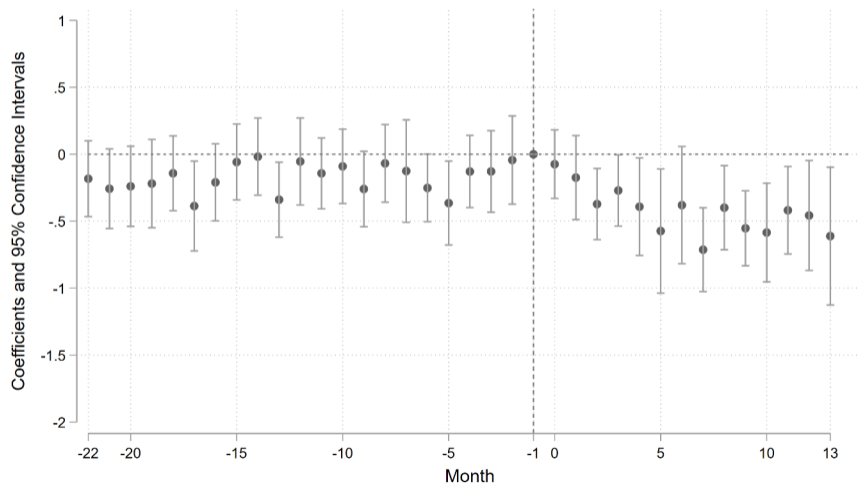
Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level.

Likely to be an **underestimation**: By the time I started collecting the data, some artists have already deactivated their accounts. robustness check

Result 1: 22% Decline of Publication Volume on DeviantArt

Pre-Trend

$$Artwork_{it} = \sum_t \beta_t Treated_i \times Month_t + \delta_i + \delta_t + \epsilon_{it}$$



Not Producing or Not Disclosing?

Example of hermit-homeboy

hermit-homeboy
Escape Reality
220 Watchers | 14 Deviations | 39.9K Pageviews

Home Gallery Favourites Posts About

Latest Deviations

NO TO AI GENERATED IMAGES

Deviation Spotlight

About hermit-homeboy

Artist // Professional // Film & Animation

Canada Deviant for 18 years @lenmanalo.com

Hq / Film

BY 800

INSTAGRAM

hermithomeboy Follow

163 posts 265 followers 939 following

Glen · he/him

hermithomeboy

Layout Artist | Game Artist
Vancouver ↔ Toronto
@richcamporado
store.steampowered.com/app/505460/Foxhole

POSTS REELS TAGGED

Grid of posts including: a character in a purple suit, a woman in a purple dress, a Snoopy illustration, a character in a purple suit, a close-up of a face, and a green-tinted portrait.

Result 2: No Reduction on Instagram, Only on DeviantArt

$$Artwork_{it} = \beta_0 + \beta_1 Post_t \times Treated_i + \delta_i + \delta_t + \epsilon_{it}$$

Table 2: Effect on Artist Publication Volume

Sample: Dep Var:	All Users		Instagram Users		Instagram Users	
	Artworks on DeviantArt		Artworks on DeviantArt		Artworks on Instagram	
	(1)	(2)	(3)	(4)	(5)	(6)
	PPML	OLS	PPML	OLS	PPML	OLS
$Post_t \times Treated_j$	-0.25*** (0.09)	-0.21* (0.12)	-0.29** (0.15)	-0.31 (0.20)	-0.09 (0.08)	-0.04 (0.41)
Pre-Treatment Mean		1.50		1.70		3.65
Artist FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N(Artist-Month)	109,800	109,800	31,968	31,968	31,968	31,968
N(Artists)	3,050	3,050	888	888	888	888
Adjusted R ²	0.53	0.58	0.44	0.33	0.49	0.59

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level.

Recap

Copyright concerns \Rightarrow *knowledge disclosure* \downarrow
 \Rightarrow *volume of knowledge spillover* \downarrow

Recap

Copyright concerns \Rightarrow *knowledge disclosure* \downarrow
 \Rightarrow *volume of knowledge spillover* \downarrow

Copyright concerns \Rightarrow *quality of knowledge disclosed?*
 \Rightarrow *quality of knowledge spillover?*

Recap

Copyright concerns \Rightarrow *knowledge disclosure* \downarrow
 \Rightarrow *volume of knowledge spillover* \downarrow

Copyright concerns \Rightarrow *quality of knowledge disclosed?*
 \Rightarrow *quality of knowledge spillover?*

Are artists withholding high-quality artworks from DeviantArt?

Are They Withholding High-Quality Artworks from DeviantArt?

Compare Quality of “Only-Instagram” and “Also-DeviantArt”



$$y_{ijt}^{Ins} = \beta_1 Post_t \times Matched_j + \beta_2 Matched_j + \mu_i + \mu_t + \epsilon_{ijt}$$

Match Artworks Across Platforms

1. For a given artist, each pair ($Artwork^{Ins}$, $Artwork^{DA}$), calculate similarity score based on **title**, **date**, **description**, **tags**
2. $m \times n$ matrix of similarity scores
3. Match artworks using Hungarian Algorithm
4. Randomly sample 150 artworks and manually check: 85% correct

Result 3: No Evidence of Withholding High-Quality Artworks

Compare Quality of “Only-Instagram” and “Also-DeviantArt”

$$y_{ijt}^{Ins} = \beta_1 Post_t \times Matched_j + \beta_2 Matched_j + \mu_i + \mu_t + \epsilon_{ijt}$$

Dep Var	<i>Likes^{Ins}</i>		<i>Comments^{Ins}</i>	
	(1) PPML	(2) OLS	(3) PPML	(4) OLS
<i>Post_t × Matched_j</i>	0.08 (0.06)	164.11 (139.15)	-0.04 (0.05)	-0.55 (1.10)
<i>Matched_j</i>	-0.03 (0.03)	-48.88 (86.03)	0.08*** (0.03)	1.33* (0.68)
Pre-Treatment Mean		2058.08		15.06
Artist FE	Y	Y	Y	Y
Month FE	Y	Y	Y	Y
N(Artworks)	87,769	87,840	87,836	87,840
Adjusted R ²	0.77	0.34	0.61	0.17

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level.

validate measure on instagram

Summary

How do copyright concerns related to AI training data impact the decisions of content creators?

1. How do copyright concerns disincentivize artists from disclosing their artwork?
 - ▶ Diff-in-Diff: 22% reduction in publication volume
2. Reduce production or reduce disclosure?
 - ▶ Using data of multi-homing artists: only reduce publication on DeviantArt, not on Instagram
3. Are artists withholding high-quality artworks?
 - ▶ Compare “only-Instagram” and “also-DeviantArt” artworks
 - ▶ Find no evidence of artists withholding high-quality artworks

Thank You!

All comments and suggestions are welcomed!

sijie.lin@mail.utoronto.ca

Literature

1. Copyright concerns associated with generative AI

- ▶ Theory: ↓ monopoly profits of original creators, welfare implication under different copyright regimes, data availability (Gans 2024; Yang&Zhang 2024)
- ▶ Empirical: ↓ availability of training data (Huang, Fu&Ghose 2023; Peukert, Abeillon, Haese, Kaiser&Staub 2024)
- ▶ **This paper: ↓ disclosure, not in production; AI adopters ↑ publication volume by 55%-60%**

2. Effect of piracy on revenue of information products

- ▶ ↓ sales due to displacement (Hui&Png 2003; Rob&Waldfogel 2006, 2007; Zentner 2006)
- ▶ ↑ sales due to word-of-mouth (Aguiar&Martens 2016; Givon, Mahajan&Muller 1995; Peukert, Claussen&Kretschmer 2007; Oberholzer-Gee&Strumpf 2007; Blackburn 2004)
- ▶ **This paper: even if AI art does not divert consumers' attention away from them, non-AI artists still withhold artworks**

3. Impact of copyright protection on innovation and knowledge diffusion

- ▶ ↑ prices, ↓ knowledge diffusion (Reimers 2019)
- ▶ ↑ quantity, ↑ quality of new products (Giorcelli&Moser 2020)
- ▶ **This paper: volume of innovation remains unchanged, knowledge diffusion ↓**

Appendix: Difference in Differences

Summary Statistics

Panel A: Artists Used in Difference-in-Differences								
	Artisan Crafts Artists				Digital Art Artists			
	Mean	Min	Max	sd	Mean	Min	Max	sd
Monthly Pre-Period Artworks	1.97	0	369	11	1.50	0	474	5.26
Profile Pageviews	1.04e+05	639	3.25e+06	2.30e+05	3.23e+05	799	5.38e+07	1.69e+06
Followers	1964	11	6.31e+04	4724	7053	11	6.76e+05	2.32e+04
<i>Views</i> ^{DA} per Artwork	5598	15	1.09e+06	4.06e+04	1.81e+04	14	5.49e+06	7.41e+04
<i>Downloads</i> ^{DA} per Artwork	5.85	0	1253	34	18	0	2.25e+04	118
<i>Favourites</i> ^{DA} per Artwork	39	0	2486	89	189	0	1.09e+04	432
<i>Comments</i> ^{DA} per Artwork	2.40	0	151	6.49	6.21	0	2887	15
<i>N(Artist)</i>			558				2492	

Notes: Use panel from January 2021 to December 2023.

Table 3: Summary Statistics

Appendix: Difference in Differences

Summary Statistics

Panel B: Artists Used in Instagram Difference-in-Differences								
	Artisan Crafts Artists				Digital Art Artists			
	Mean	Min	Max	sd	Mean	Min	Max	sd
Monthly Pre-Period Artworks	1.78	0	67	4.88	1.70	0	271	5.04
Profile Pageviews	1.37e+05	1668	3.25e+06	3.35e+05	3.64e+05	799	1.04e+07	9.98e+05
Followers	2853	24	6.31e+04	7200	1.00e+04	33	3.43e+05	2.66e+04
<i>Views</i> ^{DA} per Artwork	1.18e+04	35	1.09e+06	5.40e+04	2.72e+04	34	1.64e+06	9.30e+04
<i>Downloads</i> ^{DA} per Artwork	19	0	955	60	25	0	4089	128
<i>Favourites</i> ^{DA} per Artwork	90	0	2486	144	263	0	1.09e+04	523
<i>Comments</i> ^{DA} per Artwork	2.49	0	90	4.57	7.88	0	348	16
<i>Likes</i> ^{Ins} per Artwork	787	0	2.23e+05	4315	2015	0	1.29e+06	8985
<i>Comments</i> ^{Ins} per Artwork	13	0	1.17e+04	73	14	0	7906	63
<i>N(Artist)</i>		170				718		

Notes: Use panel from January 2021 to December 2023.

Table 4: Summary Statistics

Appendix: Difference in Differences

Robustness Check

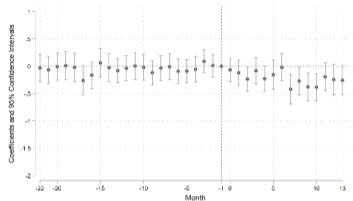
Table 5: Effect on Artist Publication Volume

	Baseline Estimation		Winsorize 99% of Dep Var		Drop 1% Largest SD. Artists	
	(1)	(2)	(3)	(4)	(5)	(6)
	PPML	OLS	PPML	OLS	PPML	OLS
$Post_t \times Treated_i$	-0.25*** (0.09)	-0.21* (0.12)	-0.17** (0.07)	-0.17** (0.07)	-0.27*** (0.09)	-0.30*** (0.11)
Pre-Treatment Mean		1.50		1.34		1.48
Artist FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N(Artist-Month)	109,800	109,800	109,800	109,800	108,936	108,936
N(Artists)	3,050	3,050	3,050	3,050	3,027	3,027
Adjusted R^2	0.53	0.58	0.46	0.50	0.49	0.37

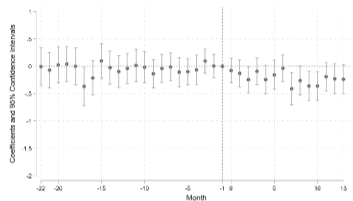
Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level.

Appendix: Difference in Differences

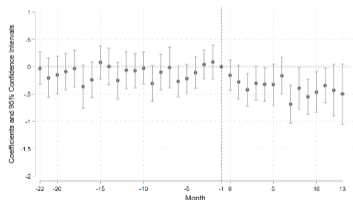
Robustness Check



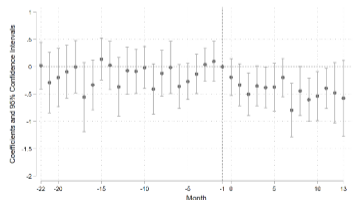
(a) Winsorize PPML



(b) Winsorize OLS



(c) Drop 1% largest SD PPML
main did



(d) Drop 1% largest SD OLS

Are They Withholding High-Quality Artworks from DeviantArt?

High Performance Correlation Across Platforms

$$y_{ijt}^{Ins} = \beta y_{ijt}^{DA} + \delta_i + \delta_t + \epsilon_{ijt}$$

Dep Var	<i>Likes^{Ins}</i>				<i>Comments^{Ins}</i>			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Favorites^{DA}</i>	2.73*** (0.56)				0.01*** (0.00)			
<i>Comments^{DA}</i>		74.76*** (13.32)				0.50*** (0.06)		
<i>Downloads^{DA}</i>			1.75 (1.08)				0.01 (0.01)	
<i>Views^{DA}</i>				0.01*** (0.00)				0.00*** (0.00)
Artists FE	Y	Y	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y	Y	Y
N(Artworks)	22,690	22,690	22,690	22,690	22,690	22,690	22,690	22,690
<i>Adjusted R²</i>	0.46	0.46	0.44	0.45	0.50	0.51	0.48	0.49

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level.

compare quality on instagram