



中国科学院自动化研究所

Institute of Automation, Chinese Academy of Sciences

NFT 图文数据的跨模态检索

NFT界的Google搜索



汇报人：王树勋

目录

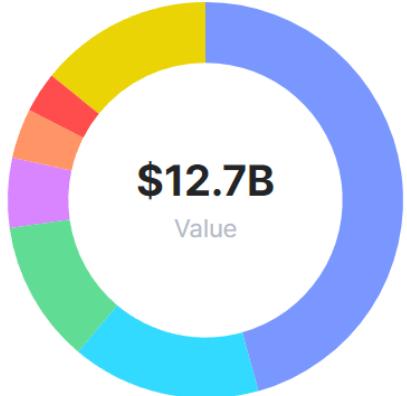
Contents



- 01 | *Background*
- 02 | *CONSTRUCTING NFT 1000*
- 03 | *Experiment & Demo*
- 04 | *Future Work*

Background

Top Sales



Top Collections

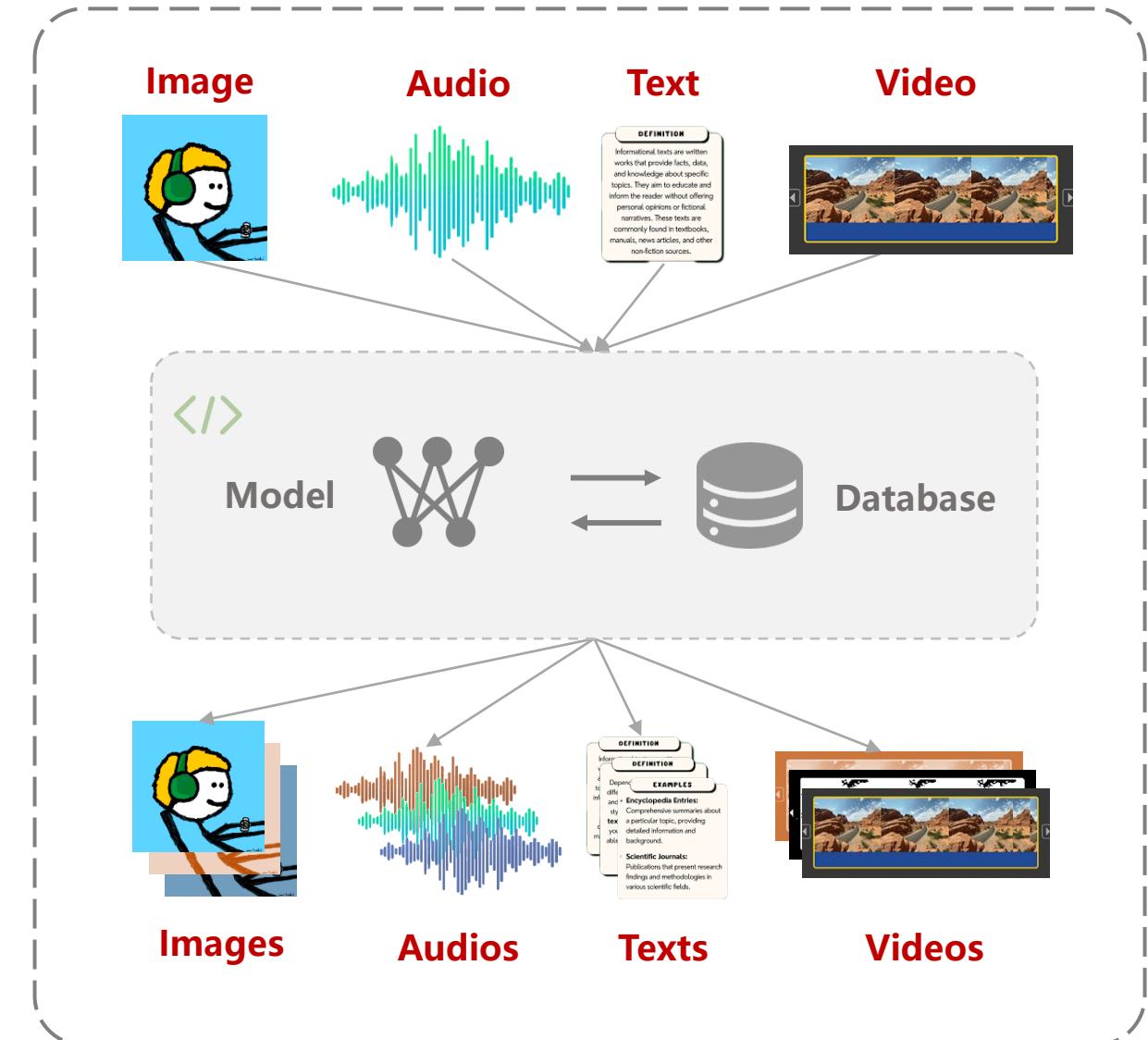
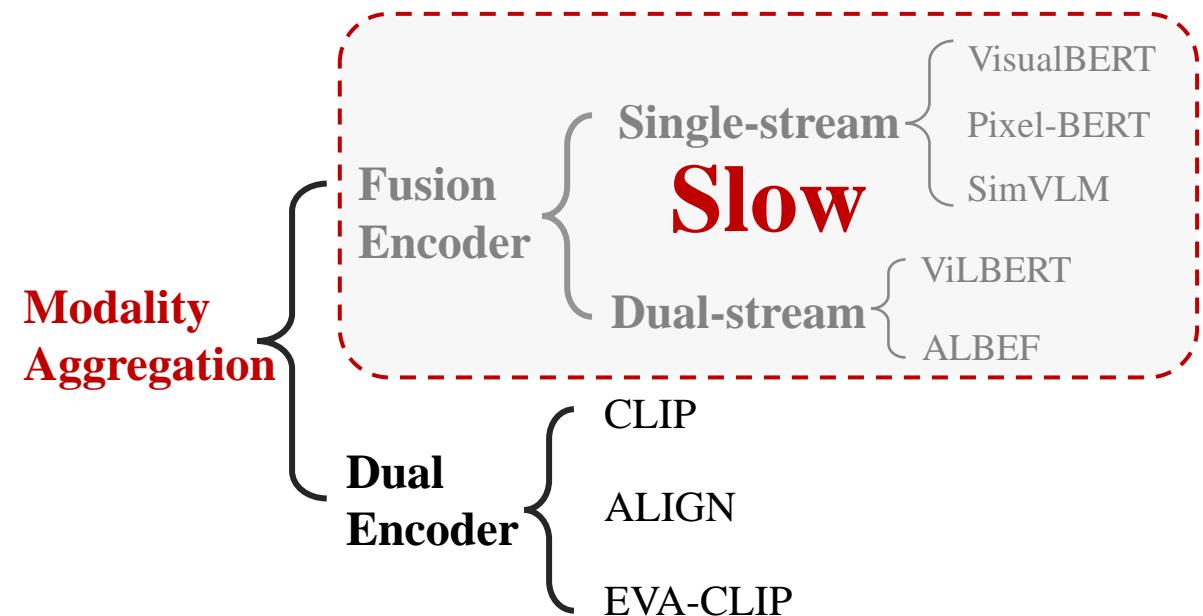
Top collections by market cap and other key indicators. ⓘ

#	Collection	Market Cap ⓘ	Volume All ⓘ	Floor Price ⓘ	
1	Bored Ape Yacht Club	◆ 115.43K	◆ 1.57M	◆ 11.55	\$ 36.777 亿
2	CryptoPunks	◆ 400K	◆ 1.09M	◆ 39.99	\$ 27.835 亿
3	Mutant Ape Yacht Club	◆ 39.47K	◆ 1.07M	◆ 2.02	\$ 25.232 亿
4	Azuki	◆ 39.4K	◆ 744.79K	◆ 3.73	\$ 16.150 亿
5	Otherdeed for Otherside	◆ 8.77K	◆ 687.49K	◆ 0.19	\$ 14.530 亿
6	CLONE X - X TAKASHI MURAKAMI	◆ 9.01K	◆ 415.8K	◆ 0.456	\$ 9.915 亿
7	Moonbirds	◆ 4.94K	◆ 362.69K	◆ 0.494	\$ 8.396 亿
8	Wrapped Cryptopunks	◆ 7.13K	◆ 352.71K	◆ 48.5	\$ 6.413 亿
9	Pudgy Penguins	◆ 92.94K	◆ 323.71K	◆ 10.17	\$ 7.574 亿
10	Doodles	◆ 16.5K	◆ 305.25K	◆ 1.65	\$ 7.336 亿

Background

Cross-Modal Retrieval

Cross-modal retrieval aims to use **the representation of data in one modality as a query** to retrieve related data in **another modality**. A well-trained AI model encodes semantically similar text, audio, or visual inputs into nearby points in a shared vector space. The target output is then obtained by calculating the "distance" between these feature vectors.



Background

High Similarity

BoredApeYachtClub



Cool Cats



Akutars



3Landers



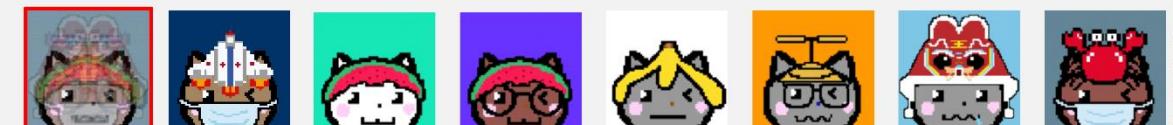
PudgyPenguins



Mfers



BanaCat



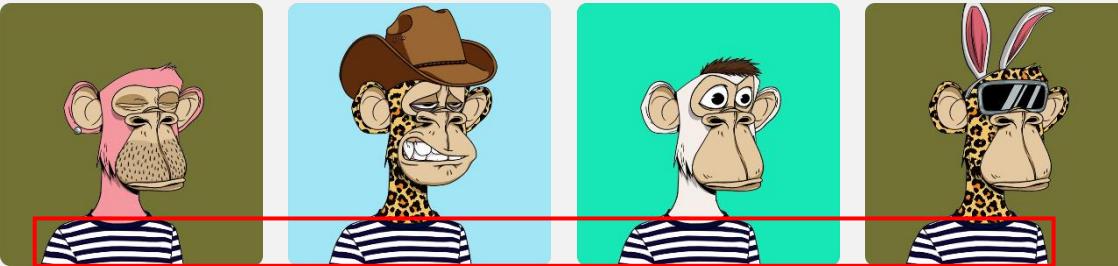
Different Collections



Background

Fixed-layer permutation and combination

Clothes: Navy striped tee



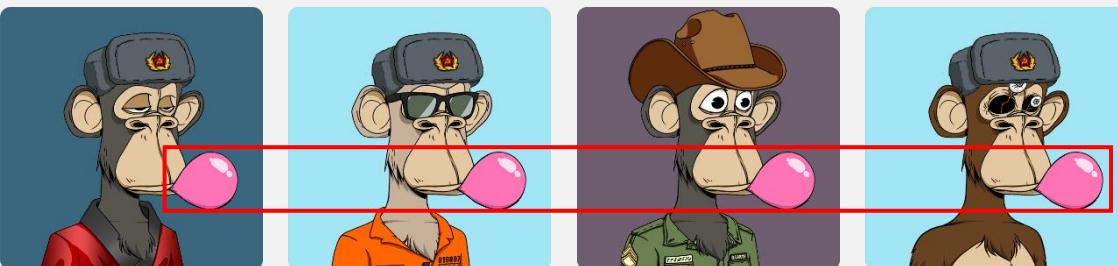
(a)

Eyes: 3d



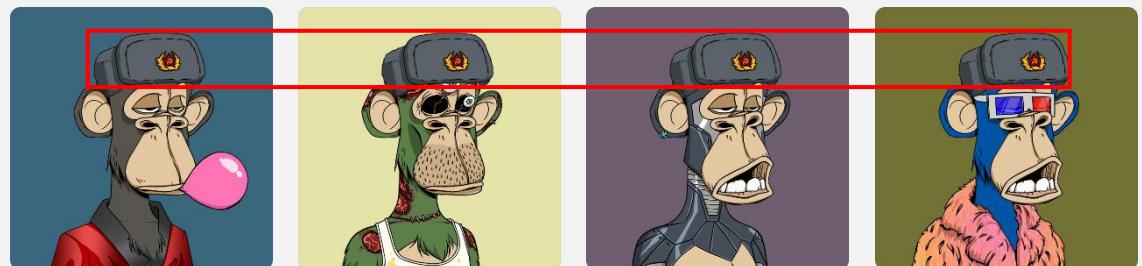
(b)

Mouth: Bored bubblegum



(c)

Hat: Commie hat



(d)

Background

Abstract Description

Adam Bomb Squad



#1

Superlative Secret Society



#6

```
{ "attributes": [ {"trait_type": "Character", "value": "Color Adam"}, {"trait_type": "Bomb Type", "value": "Classic"}, {"trait_type": "Colorway", "value": "Color Red"}, {"trait_type": "Distribution", "value": "General Release"}, {"trait_type": "Year", "value": "2018"}, {"trait_type": "Spark", "value": "Yes"}, {"trait_type": "Branded", "value": "No"}, {"trait_type": "Background Name", "value": "Camo"}, {"trait_type": "Background Colorway", "value": "Perfect Camo Pink"}, {"trait_type": "Background Year", "value": "2014"}, {"trait_type": "Background Branded", "value": "No"}, {"trait_type": "Error Terror", "value": "No"}, {"trait_type": "Double Trouble", "value": "No"}, {"trait_type": "Broken Background", "value": "No"} ] }
```

```
{ "attributes": [ {"trait_type": "角色", "value": "色彩亚当"}, {"trait_type": "炸弹类型", "value": "经典"}, {"trait_type": "颜色方案", "value": "红色"}, {"trait_type": "分布", "value": "通用发行"}, {"trait_type": "年份", "value": "2018"}, {"trait_type": "火花", "value": "是"}, {"trait_type": "品牌化", "value": "否"}, {"trait_type": "背景名称", "value": "迷彩"}, {"trait_type": "背景颜色方案", "value": "完美迷彩粉"}, {"trait_type": "背景年份", "value": "2014"}, {"trait_type": "背景品牌化", "value": "否"}, {"trait_type": "错误恐怖", "value": "否"}, {"trait_type": "双重麻烦", "value": "否"}, {"trait_type": "破损背景", "value": "否"} ] }
```

```
{ "attributes": [ {"trait_type": "Background", "value": "Chummy"}, {"trait_type": "Body", "value": "Captivating"}, {"trait_type": "Necklace", "value": "Obedient"}, {"trait_type": "Muffler", "value": "Folksy"}, {"trait_type": "Head", "value": "Lyrical"}, {"trait_type": "Orbs", "value": "Contemptible"}, {"trait_type": "Foreground", "value": "Restrained"}, {"trait_type": "Galaxy Glitter", "value": "Unfoolable"} ] }
```

```
{ "attributes": [ {"trait_type": "背景", "value": "友好"}, {"trait_type": "身体", "value": "迷人"}, {"trait_type": "项链", "value": "顺从"}, {"trait_type": "围巾", "value": "朴实"}, {"trait_type": "头部", "value": "抒情"}, {"trait_type": "眼球", "value": "可鄙"}, {"trait_type": "前景", "value": "克制"}, {"trait_type": "星河闪烁", "value": "难以愚弄"} ] }
```

CONSTRUCTING NFT1000

Clarifying the Targets



Ethereum Top1000 PFP ERC-721 NFT collections

CONSTRUCTING NFT1000

Download



Multi-Platform, Multi-Process, Multi-Threaded Rapid

Download of All Resources in the Target Collection

1250 collections 14TB

Base_URI + source_index

Smart Contract Address :

0xBC4CA0EdA7647A8aB7C2061c2E118A18a936f13D

baseURI :

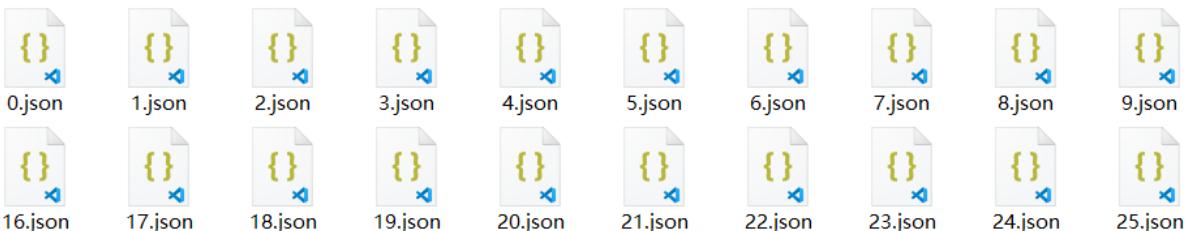
ipfs://QmeSjSinHpPnmXmspMjwiXyN6zS4E9zccariGR3jxcaWtq/

Metadata_Source :

ipfs://QmeSjSinHpPnmXmspMjwiXyN6zS4E9zccariGR3jxcaWtq/0

"image":

"ipfs://QmRRPWG96cmgTn2qSzjwr2qvfNEuhunv6FNeMFGa9bx6
mQ"



Construct NFT1000

Filter & standardize

~~Collection with completely duplicated media data~~



~~Projects with less than 500 NFT tokens~~

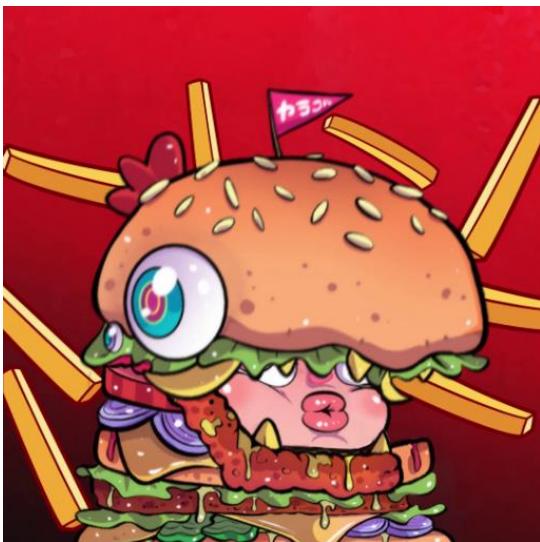
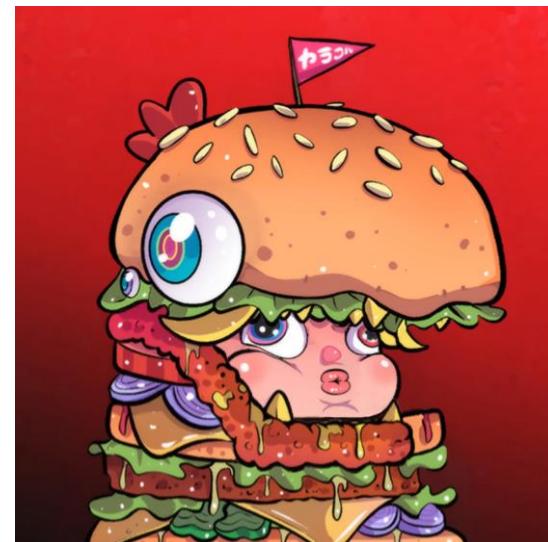
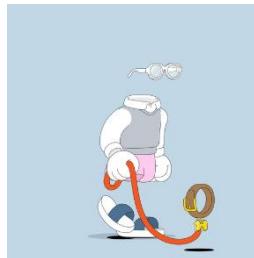
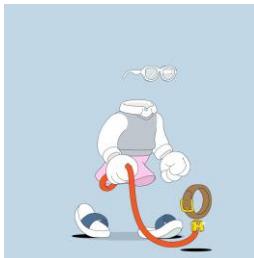


~~Projects lack JSON files or where the JSON files contain no substantive semantic information.~~

JPG, JPEG, SVG, PNG, WebP, GIF, MP4

PNG Width : 512 pixels
Height : adaptive

14TB → 1.75TB



CONSTRUCTING NFT1000

Caption Extraction

```
subject +  
{  
    "attributes": [  
        {  
            "trait_type": "Earring",  
            "value": "Silver Hoop"  
        },  
        {  
            "trait_type": "Background",  
            "value": "Orange"  
        },  
        {  
            "trait_type": "Fur",  
            "value": "Robot"  
        },  
        {  
            "trait_type": "Clothes",  
            "value": "Striped Tee"  
        },  
        {  
            "trait_type": "Mouth",  
            "value": "Discomfort"  
        },  
        {  
            "trait_type": "Eyes",  
            "value": "X Eyes"  
        }  
    ]  
}
```

Prompt Engineering



Meta
LLAMA 13B



Caption Template

A picture of { subject },
containing [{ value }
{ trait_type } for { value }
{ trait_type } in attributes.items]

1. Meet the futuristic Ape with a silver hoop earring, wearing a striped tee and an orange background. Its fur is metallic and its X-shaped eyes give it a robotic appearance. Its mouth looks uneasy, conveying discomfort.

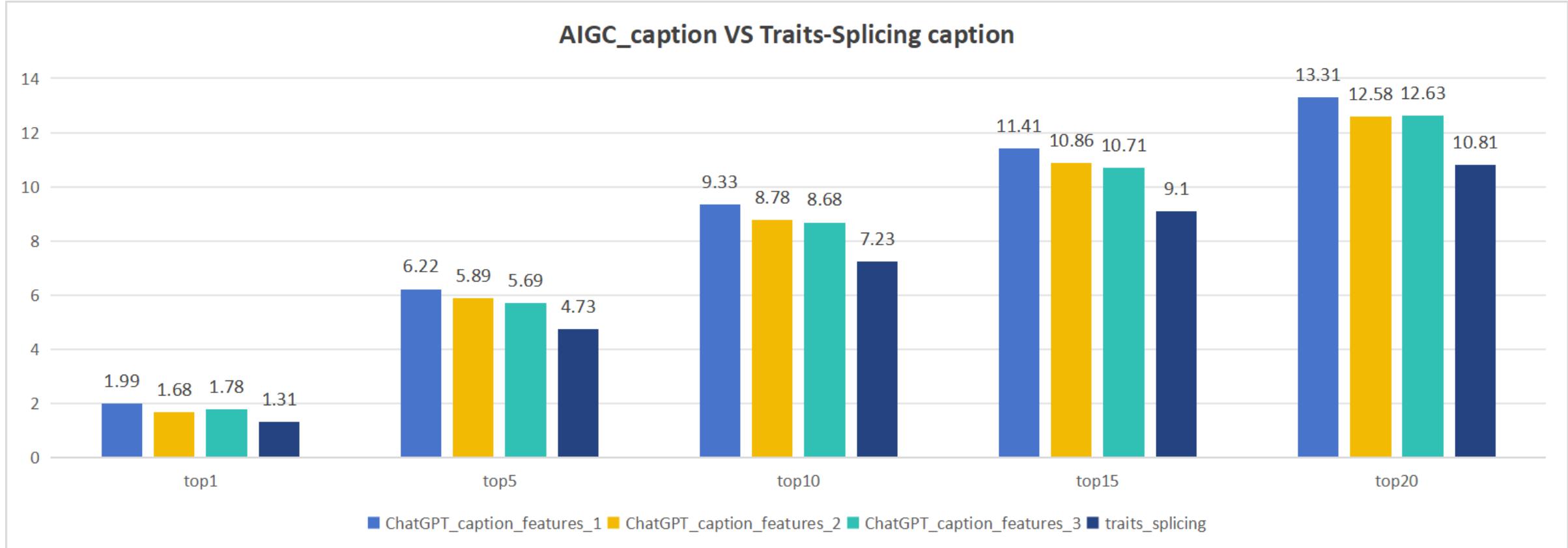
2. Observe the striking Ape with a silver hoop earring, donning a striped tee against an orange backdrop. Its metallic fur combined with X-shaped eyes grants it a futuristic look. Its mouth appears uneasy, perhaps due to discomfort.

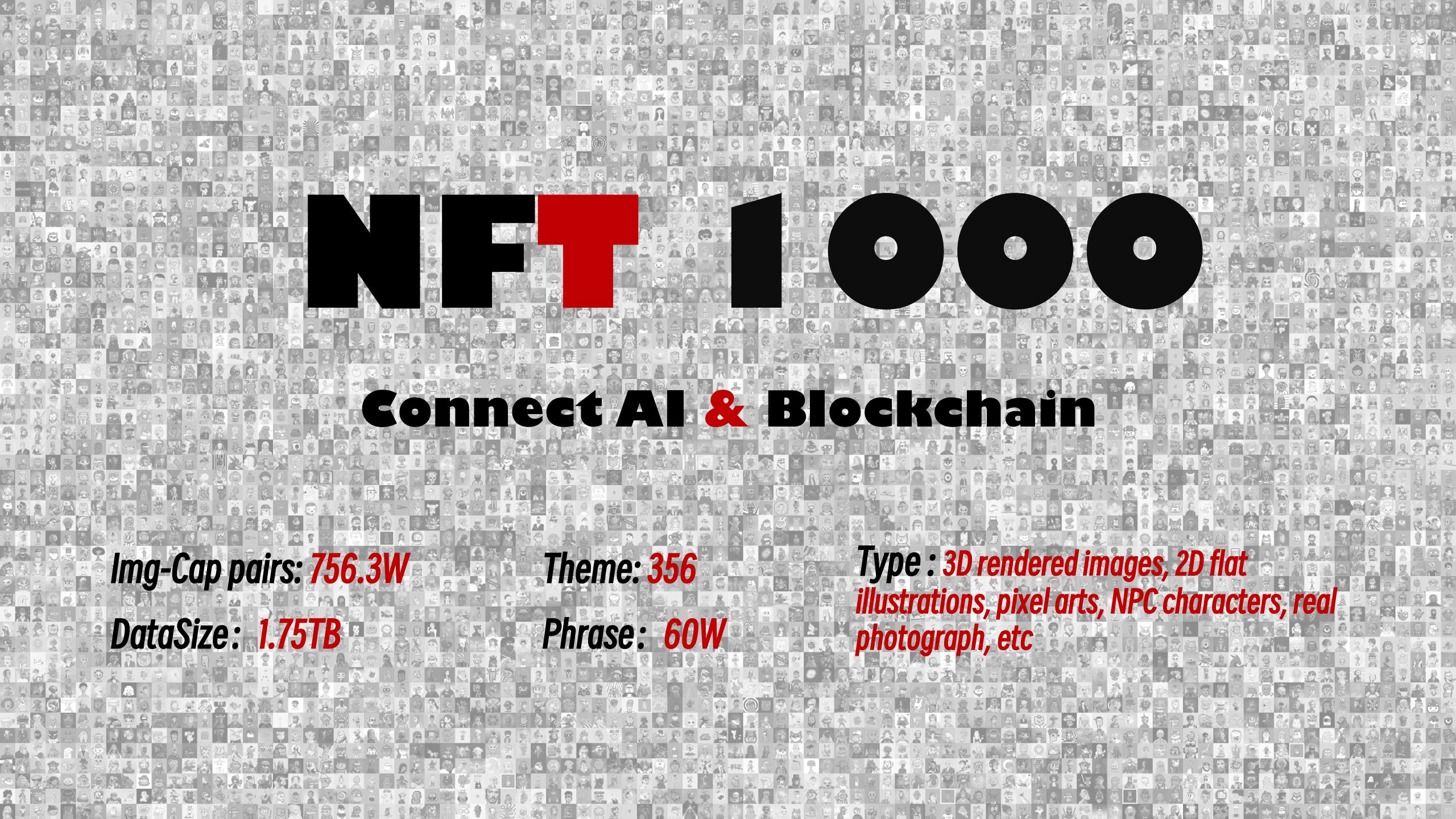
3. Introducing the unique Ape, decorated with a silver hoop earring and a striped tee on an orange background. Its fur has been replaced with robotic parts, and its X-shaped eyes give it an innovative edge. Its expression shows discomfort.

A picture of BoredApeYachtClub, containing Silver Hoop Earring Orange Background Robot Fur Striped Tee Clothes Discomfort Mouth X Eyes Eyes.

CONSTRUCTING NFT1000

Caption Extraction





NFT100

Connect AI & Blockchain

Img-Cap pairs: 756.3W

DataSize : 1.75TB

Theme: 356

Phrase: 60W

*Type : 3D rendered images, 2D flat
illustrations, pixel arts, NPC characters, real
photograph, etc*

CONSTRUCTING NFT1000

Data Partition

NFT 100 mini

Training Set:

Pro_num: **80**

Img-Cap pairs: **40000**

Validation set:

Pro_num: **5**

Img-Cap pairs: **2500**

Test Set:

Pro_num: **15**

Img-Cap pairs: **7500**

NFT 1000 mini

Training Set:

Pro_num: **800**

Img-Cap pairs: **794698**

Validation set:

Pro_num: **50**

Img-Cap pairs: **49738**

Test Set:

Pro_num: **150**

Img-Cap pairs: **147615**

NFT 1000

Training Set:

Pro_num: **800**

Img-Cap pairs: **6178249**

Validation set:

Pro_num: **50**

Img-Cap pairs: **383916**

Test Set:

Pro_num: **150**

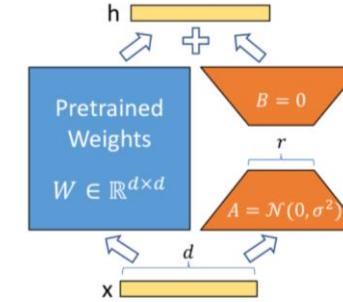
Img-Cap pairs: **1000838**

Experiments

Generation

1. Train LORA

Train the LORA plugin based on the SD model for a specific project



2. Generate Metadata

Determine which attributes make up the target NFT, enumerate a series of values for each attribute, and first construct a JSON file through permutations and combinations at the attribute level.



3. Base Model + LORA + Prompts

Synthesize the use of the basic diffusion model, LORA, and related prompts to generate a series of high-quality, coherent images



Azuki_SD 1.5

(a)



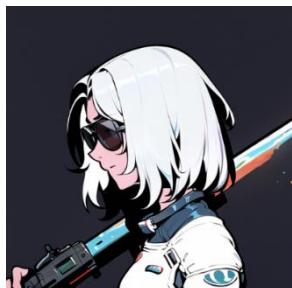
Prompt: 1boy,Blonde Swept Back Hair,Red Stripes Face Paint Face,Red Perfecto Jacket Clothing,Suspicious Eyes,Chuckle Mouth,Leather Katana Offhand,Red Background,<lora:azuki:1>

(b)



Prompt: 1gir,IViolet Flowy Hair,Sloth Kimono Clothing,Calm Eyes,Smile Mouth,Guitar,Offhand Dark Blue Background,<lora: azuki:1>,
(masterpiece:1.2),best quality,masterpiece,highres,original,extremely detailed wallpaper,perfect lighting,(extremely
detailedCG:1.2),drawing,paintbrush,

(c)

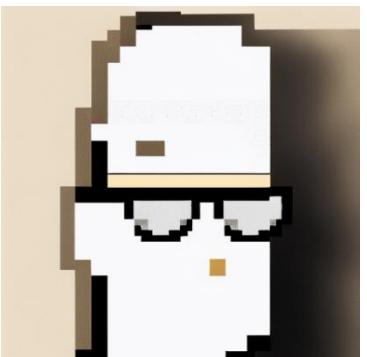
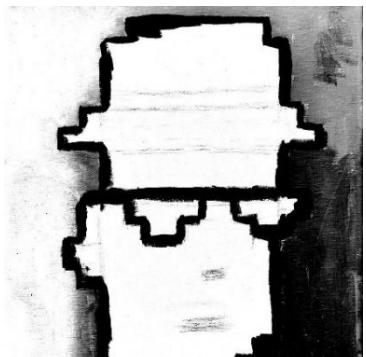
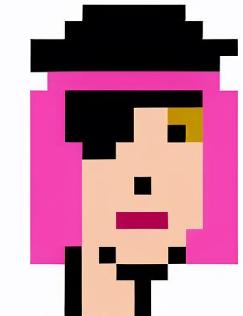


Prompt: 1girl,Water Hair,(space_suit_1.2),Striking Eyes,Frown Mouth,holding gun,3Dglasses,glass space helmet, space_background,<lora:azuki:1>

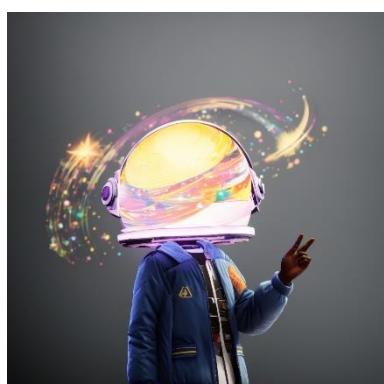
Azuki_toonyou



CRYPTOPUNKS

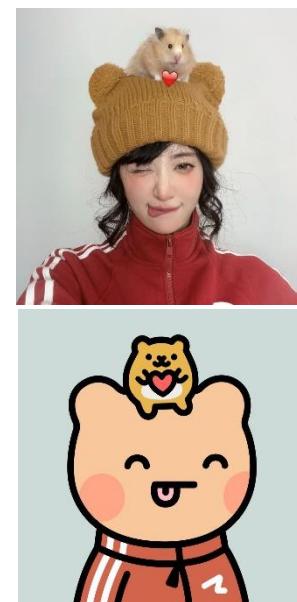


Akutars



Future Work

Recommendation



Kirara
@0xKirara

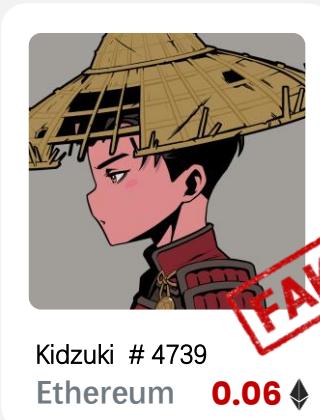
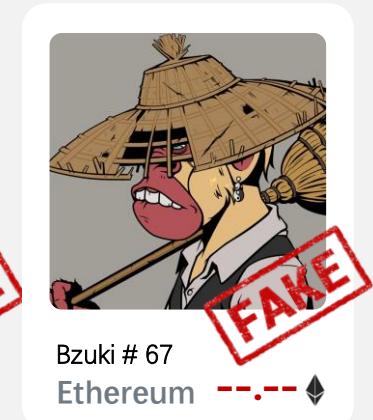
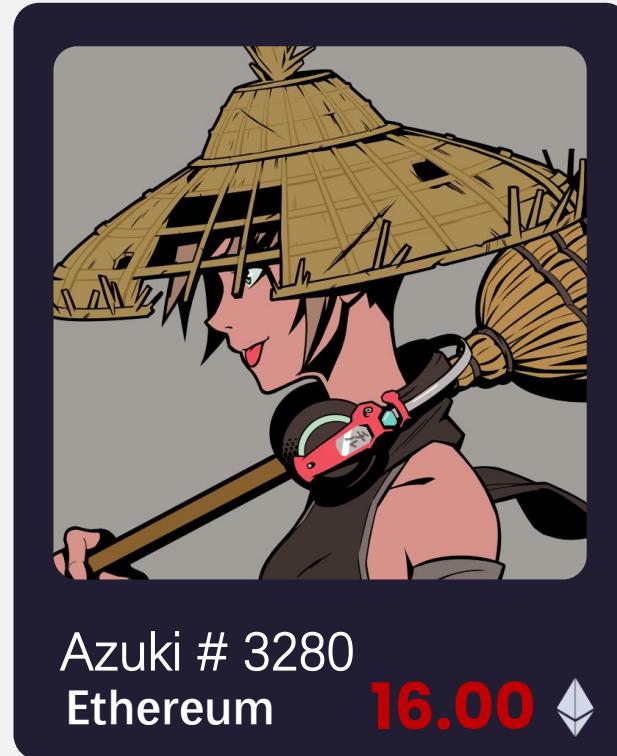
Blockchain is Not a Lawless Land: NFT Copyright Authentication

« NFT Fraud Detection Company Yakoa Secures \$4.8 Million in Seed Funding »

Scott Gray, Founder of Genie and Head of NFT Products at Uniswap Labs, believes that “in a decade, most of the world’s intellectual property will be stored on blockchains. Their mediums will range from virtual worlds, to marketplaces, social, and publishers. Yakoa is protecting creators and their communities from fraudulent infringements. We’re going to see the world’s largest brands and NFT collections leveraging Yakoa to protect their IP across all experiences.”

website: <https://www.yakoa.io/>

Report link: <https://medium.com/@yakoa/yakoa-raises-4-8m-to-combat-intellectual-property-fraud-in-web3-973efcc8087d>



Future Work

Depth Vertical

- 1. Expanding the NFT Dataset:** We aim to broaden our scope by continuing to collect outstanding projects from public blockchains such as Solana, Klaytn, and others, with a target of reaching 100 million image-text pairs.
- 2. Training Multimodal Large Language Models:** We will train large-scale multimodal language models, integrating Retrieval-Augmented Generation (RAG) techniques to offer personalized NFT recommendations through intelligent Q&A interactions.
- 3. Developing a Full-Chain NFT Search Engine:** Our goal is to create a comprehensive search engine tailored specifically for the NFT domain. This engine will search off-chain data and then return to the blockchain for final tracking, ensuring a seamless and efficient user experience.
- 4. Image-Based Search and Copyright Infringement Detection:** We will implement image-based search capabilities and develop tools for detecting NFT piracy, providing robust solutions to safeguard intellectual property.
- 5. NFT Generation:** We plan to train generative models specifically designed for NFT data, facilitating the creation of new NFTs and enhancing creative possibilities within the ecosystem.



ACM Multimedia 2024

« NFT1000: A Cross-Modal Dataset For Non-Fungible Token Retrieval »

Accept!

Openreview: <https://openreview.net/forum?id=xUtNrKH8iB>



中国科学院自动化研究所

Institute of Automation, Chinese Academy of Sciences



谢谢大家！