

JLU TAG

Elain Bai, Yanhui Li, Weiqi Zhang

Jilin University

2023.11.15



- ① Project description
- ② Data resources
- ③ Knowledge Graph
- ④ Problems and Solutions
- ⑤ QA

- ① Project description
- ② Data resources
- ③ Knowledge Graph
- ④ Problems and Solutions
- ⑤ QA

Contributors

- Data Scientist: Elain Bai
- Domain Expert: Elain Bai, Yanhui Li
- Knowledge Engineer: Elain Bai, Weiqi Zhang
- Project Manager: Elain Bai
- Teacher: Rui Zhang

Pixiv

#Kiana_Kaslana

Honkai Impact 3
#崩坏3rd

#kiana_kaslana

Honkai Impact 3
#崩坏3rdKiana Kaslana
#キアナ・カスラナ

#kiana

Project description

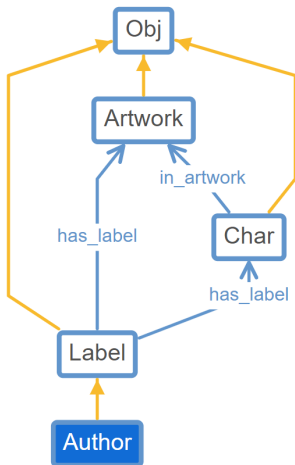
- This project aims to use knowledge graph technology to organize the tags of illustrations on websites such as Pixiv and Lofter to provide a better user experience. We build a knowledge graph to provide a structured representation of information about illustration works, characters, etc., and provide basic data for classification.
- Usage Scenarios:
Users can browse the illustrations on the above-mentioned websites more conveniently and find the content they are interested in through searches combined with the knowledge graph.

- 1 Project description
- 2 Data resources
- 3 Knowledge Graph
- 4 Problems and Solutions
- 5 QA

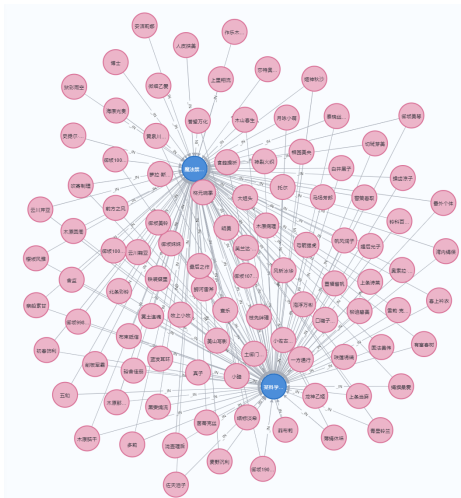
Data resources

- First of all, the largest illustration website in the world is Pixiv. We can crawl the illustration tag data from this website. We then crawled data about characters and titles from encyclopedia sites like Moegirl and Bangumi.
- We first determine the crawling target, then choose the crawler tool to crawl, and finally store the crawled data in JSON form.

- ① Project description
- ② Data resources
- ③ Knowledge Graph
- ④ Problems and Solutions
- ⑤ QA



KG



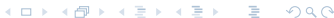


127.0.0.1:23648/search/柯南

您可能在找:

作品 [名侦探柯南](#): 别名 名探偵コナン, {描述}.

角色 [江户川柯南](#): 无别名, 是作品 [名侦探柯南](#) 中的



← → ↺ ⓘ 127.0.0.1:23648/search/爱丽丝

您可能在找:

角色 [天童爱丽丝](#): 无别名, 是作品 [蔚蓝档案](#) 中的角色, {描述}.

角色 [爱丽丝](#): 无别名, 是作品 [Fate](#) 中的角色, {描述}.

角色 [爱丽丝](#): 无别名, 是作品 [明日方舟](#) 中的角色, {描述}.

角色 [爱丽丝·玛格特罗依德](#): 别名 爱丽丝, 是作品 [东方 Project](#) 中的角色, {描述}.

角色 [爱丽丝菲尔·冯·爱因兹贝伦](#): 别名 爱丽丝菲尔, 是作品 [Fate](#) 中的角色, {描述}.

角色 [橘爱丽丝](#): 无别名, 是作品 [偶像大师](#) 中的角色, {描述}.

- 1 Project description
- 2 Data resources
- 3 Knowledge Graph
- 4 Problems and Solutions**
- 5 QA

Anti-crawling

pixiv事務局 official



[pixiv]過度訪問警告

親愛的用戶，您好。

我們確認到您有違反pixiv使用條款之行為。

■pixiv使用條款: <https://policies.pixiv.net/>

■pixiv本站方針: <https://www.pixiv.net/terms/?page=guideline>

第14條 禁止行為

19.超出一般範圍並造成本服務伺服器過度負擔之行為，或是造成本服務的營運或網路、系統障礙等行為。

請停止過度訪問。若此行為未得到改善，有可能停止或刪除您的帳號。

若有疑問，請由下列網址聯繫客服。

▼聯繫客服

[https://www.pixiv.net/support.php?
mode=inquiry&service=pixiv&type=6](https://www.pixiv.net/support.php?mode=inquiry&service=pixiv&type=6)

感謝您的閱讀及配合。

Huge amount of data

- The amount of data on the Pixiv, Moegirl and Bangumi is very large, and processing large-scale data may affect the efficiency and speed of the project.

System limitations

← → ↻ ⓘ 127.0.0.1:23648/uid/eb58eb0963c64062b5030663e77b49df

uuid eb58eb0963c64062b5030663e77b49df:

类型: 作品

名称(別名): 名偵探柯南、名探偵コナン

相关链接:

角色: 远山银司郎、朗姆、中森银三、佐藤正义、千间降代、黑田兵卫、铃木园子、登米、冲矢昴、东尾玛丽亚、茱蒂·斯泰琳、萩原研二、板仓卓、宫野厚司、怪盗基德、橘境子、浅井成实、白马探、胁田兼则、赤木量子、黑羽盗一、小岛元太、伊达航、羽田浩司、工藤新一、松田阵平、小岛元次、诸伏高明、黑羽快斗、安德雷·卡迈尔、鲛崎岛治、妃英理、远山和叶、沼渊己一郎、羽田秀吉、米原樱子、目暮十三、田布施阳菜、安室透、REI、中森青子、詹姆斯·布莱克、黑影君、佐藤美和子、山村操、宫本由美、伊织无我、吉田步美、服部静华、吐、宫野艾莲娜、荻野彩实、琴酒、羽贺响辅、本堂瑛祐、绫小路文麿、阿笠博士、灰原哀、世良真纯、工藤优作、风见裕也、萩原千速、维斯布朗德、和田阳奈、木之下茉纱绘、新出智明、库拉索、泽田弘树、白鸟任三郎、毛利兰、毛利小五郎、铃木绫子、楠田陆道、越水七海、三池苗子、小林澄子、卡尔瓦多斯、盐田平八郎、赤井务武、诸伏晃一、江户川柯南、贝尔摩德、山村美砂绘、小田切敏郎、松本清护、降谷、榎本梓、服部平藏、若狭留美、工藤新一、大和敢助、宫野明美、铁诸羽

- 1 Project description
- 2 Data resources
- 3 Knowledge Graph
- 4 Problems and Solutions
- 5 QA**

Thanks!