

WikiCite - Wien, May 23 2017

The Primary Sources Tool

Curating citations for Wikidata statements



WIKIMEDIA
FOUNDATION

Marco Fossati - User:Hjfocs

The story so far



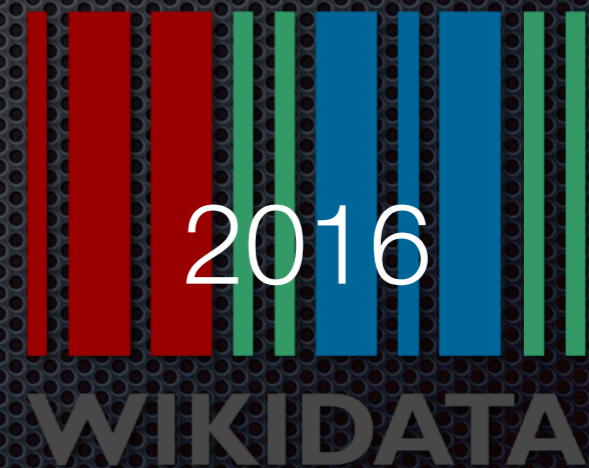
Leonardo da Vinci

Mathematician

Leonardo di ser Piero da Vinci, more commonly Leonardo da Vinci or simply Leonardo, was an Italian polymath whose areas of interest included invention, painting, sculpting, architecture, science, music, ...
[Wikipedia](#)

Born: April 15, 1452, [Anchiano](#)

Died: May 2, 1519, [Clos Lucé, Amboise, France](#)



genre



pop rock

edit

▶ 1 reference



alternative rock

+ approve claim

▼ 3 references

+ approve reference rej

reference URL

<http://www.nipponproject.com/de/artists.php?letter=K>

+ approve reference rej

reference URL

<http://www.last.fm/user/m4machine>

+ approve reference rej

reference URL

<http://www.nipponproject.com/fr/artists.php?letter=K>

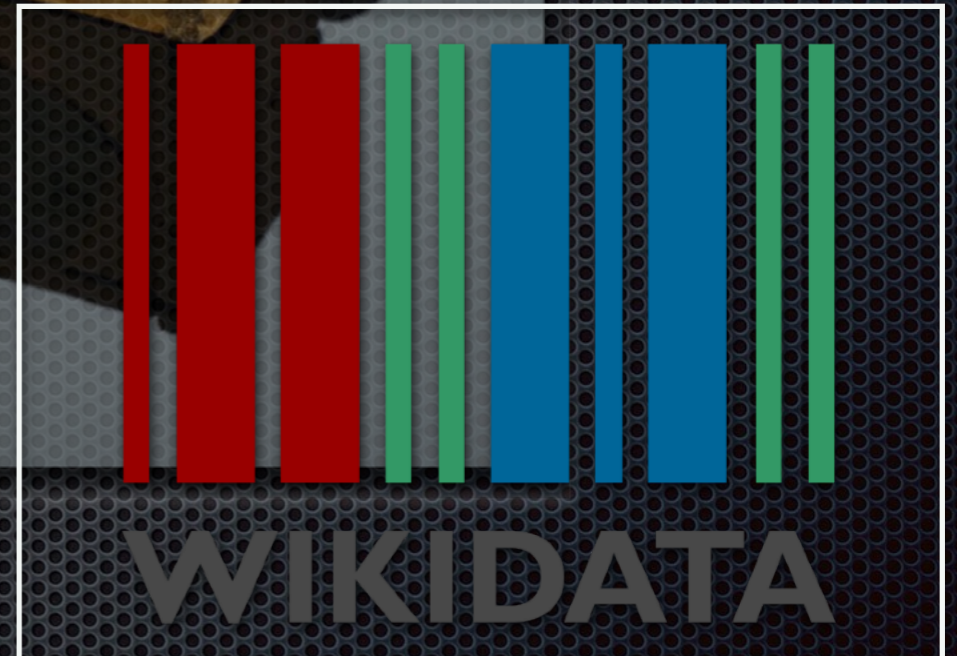
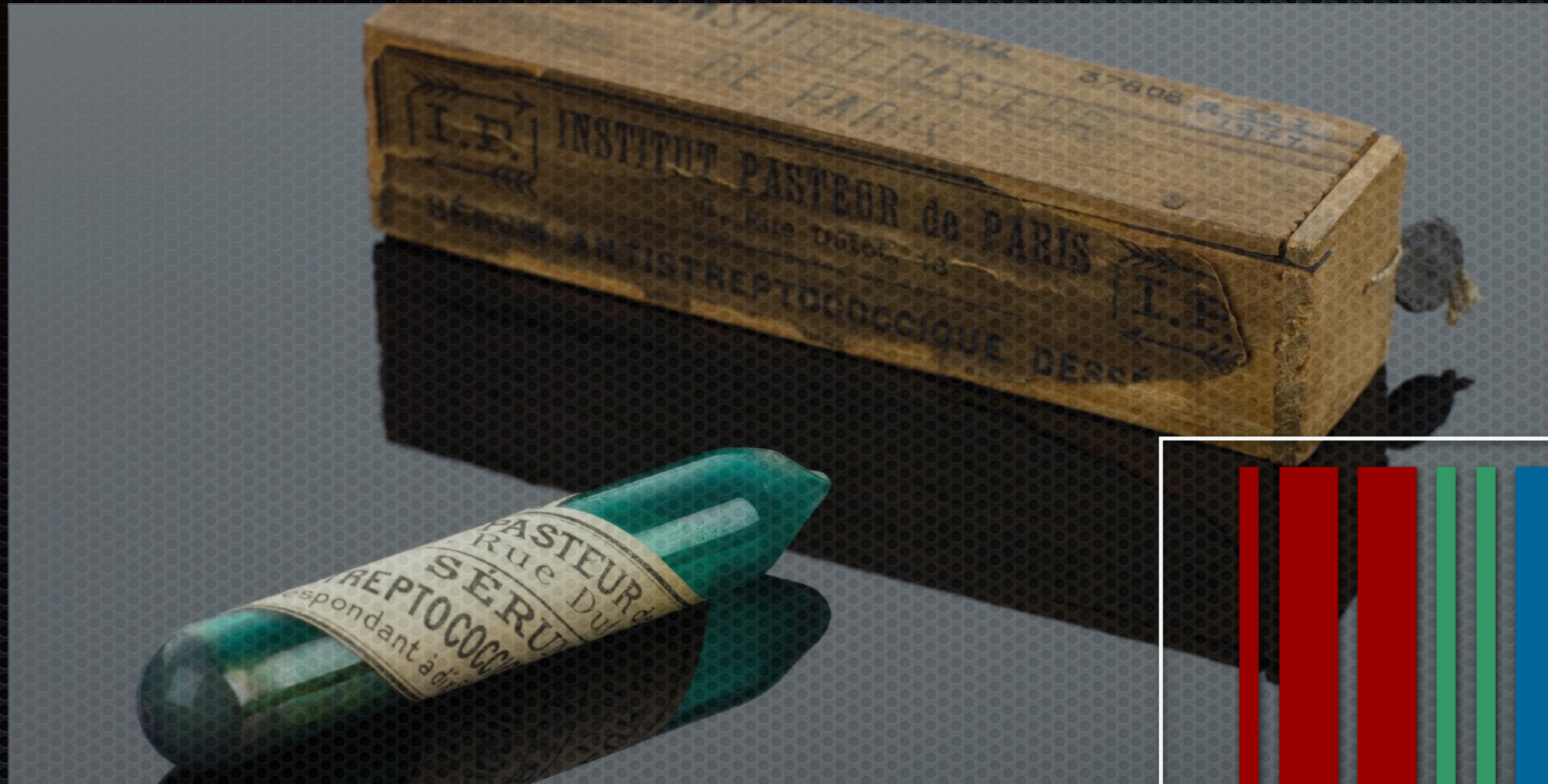
Primary sources tool

Wikidata reliability

- ✦ Need for trust on the content addition process
- ✦ Lack of validation against third-party resources
- ✦ Few references to external authoritative sources
- ✦ A reference for each statement is not ensured

Data-driven solution

- Harvest factual information from a corpus of reliable Web sources
- Generate Wikidata-compliant statements
- Validate existing ones via references to such sources
- Suggest new (referenced) statements



Outcome: StrepHit

a project funded via

Wikimedia Individual Engagement Grants

StrepHit progress

- 2016: beyond Freebase & Knowledge Vault
- 2017: next level of the primary sources tool

genre



pop rock

edit

▶ 1 reference



alternative rock

+ approve claim

▼ 3 references

+ approve reference rej

reference URL

<http://www.nipponproject.com/de/artists.php?letter=K>

+ approve reference rej

reference URL

<http://www.last.fm/user/m4machine>

+ approve reference rej

reference URL

<http://www.nipponproject.com/fr/artists.php?letter=K>

Time for an uplift!
IEG renewal
kick-off

Primary sources tool

Frustrating workflow

1. No filtering capabilities
2. No reference previews
3. No extracted fact highlighting
4. Too many page reloads
5. Duplicated data

1. Filtering, take 1

Browse Primary
Sources

▲ back to top ▲

MusicBrainz artist ID

country of citizenship

date of birth

date of death

genre

occupation

place of birth

place of death

ary
p ▲
izenship

0 references

singer

0 referen

actor

3 refer

referenc

Preview

referenc

Preview

reference URL

Preview

The screenshot shows the website China.org.cn with a navigation menu and a main article titled "Top 100 Chinese celebrities 2011" by Xu Lin, dated June 23, 2011. Below the article title is a table titled "2011 Forbes China Celebrities List".

Comprehensive Ranking	Name	Occupation	Comments
1	Andy Lau	Singer, Actor	Played in five films; held "unforgettable" concerts
2	Jay Chou	Singer, Actor	Published a new album <i>The Era</i> ; held 33 concerts
3	Faye Wong	Singer, Actress	Successful came back to the stage; held 13 concerts, almost full house
4	Jackie Chan	Actor	As director, actor and producer; having a stable position in the

+ approve claim reject claim

approve reference reject reference

approve reference reject reference

[06/23/content_22845446_11.htm](#)

+ approve reference reject reference

http://www.szdaily.com/content/2010-12/20/content_5185916.htm

name in kana

シャオ・シェンヤン

1 reference

2. Reference preview

← → ↻ Sicuro | https://www.wikidata.org/wiki/Q1018752#

occupation

actor

+ approve claim 🗑️ reject claim

▼ 3 references

+ approve reference 🗑️ reject reference

reference URL http://video.xbox.com/en-us/movie/a-woman-a-gun-and-a-noodle_shop/7fbd386f-e0f3-4921-

63	Xiao Shenyang (Shen He)	Actor, Singer	Played in the film <i>Just Call Me Nobody</i> ; His song was popular among the public
----	--------------------------------	---------------	---

+ approve reference 🗑️ reject reference

reference URL http://www.china.org.cn/top10/2011-06/23/content_22845446_11.htm

+ approve reference 🗑️ reject reference

reference URL http://www.szdaily.com/content/2010-12/20/content_5185916.htm

+ approve reference 🗑️ reject reference

+ add

name in kana

シャオ・シェンヤン


edit

▶ 1 reference

+ add

edit

+ add



Item Discussion

Xiaoshenyang (Q1018752)

Chinese actor
Xiao Shen-Yang

Main page

3. Extracted fact preview

Italian Condottiero
Francesco 03 Ordelaffi

In more languages [Configure](#)

Language	Label
English	Francesco
Italian	Francesco
French	Francesco
Sardinian	No label

All entered languages

Statements

instance of

sex or gender

given name

- Main page
- Community portal
- Project chat
- Create a new item
- Item by title
- Recent changes
- Random item
- Query Service
- Nearby
- Random Primary Sources item
- Help
- Donate

Tools

- What links here
- Related changes
- Special pages
- Permanent link
- Page information
- Concept URI

[Primary sources filter](#)

Sources

- ▲ back to top ▲
- child
- country of citizenship
- date of birth
- noble family
- occupation

Primary sources filter

Close

Source language:

Domain of interest:

Property:

Dataset:

Load

Subject	Property	Object	Reference	Action
ZIP code	country	United States of America	Source	Approve Reject
SW postcode area	country	United Kingdom	Source	Approve Reject
LA postcode area	country	United Kingdom	Source	Approve Reject
NN postcode area	country	United Kingdom	Source	Approve Reject
NP postcode area	country	United Kingdom	Source	Approve Reject
SO postcode area	country	United Kingdom	Source	Approve Reject
10048	country	United States of America	Source	Approve Reject
BA postcode area	country	United Kingdom	Source	Approve Reject
BH postcode area	country	United Kingdom	Source	Approve Reject
BR postcode area	country	United Kingdom	Source	Approve Reject
BS postcode area	country	United Kingdom	Source	Approve Reject
CA postcode area	country	United Kingdom	Source	Approve Reject

▼ 0 references

Primary sources filter Filtering, take 2

What do you think?



Marco Fossati - User:Hjfoqs

All images CC BY-SA 4.0, except the Google Knowledge Graph screenshot, the Google and Freebase logos, registered trademarks of Google Inc. © 2015 Google Inc. All rights reserved.