

# CONTRIBUTING SINGAPORE NAMED ENTITIES TO WIKIDATA

NKOS Consolidated Workshop 2023  
9 November 2023

**Chow Yaw Huah**  
**Glenn Hong**

**Knowledge Organisation Systems**  
**National Library Board Singapore**



**National Library Board**  
Singapore

# AGENDA

- **NLB's Named Entities**
- **Contributing to Wikidata**

# NLB'S KEY LINES OF BUSINESS

VISION: READERS FOR LIFE, LEARNING COMMUNITIES, KNOWLEDGEABLE NATION

MISSION: WE MAKE KNOWLEDGE COME ALIVE, SPARK IMAGINATION AND CREATE POSSIBILITIES



## Public Libraries

Network of 29 Public Libraries, including 2 partner libraries

Reading Programmes and Initiatives

Programmes and Exhibitions targeted at Singapore communities



## National Library

Preserving Singapore's Print and Literary Heritage, and Intellectual memory  
Reference Collections  
Legal Deposit (including electronic)



## National Archives

Transferred from NHB to NLB in Nov 2012

Custodian of Singapore's Collective Memory:  
Responsible for Collection, Preservation and Management of Singapore's Public and Private Archival Records

# LEVERAGING NLB'S ASSETS, BRAND AND REACH

We have achieved significant reach<sup>^</sup>.



**Over 3.5m**  
Library members  
(Singapore residents)



**14.8m**  
No. of visits to NLB premises



Overall Reach\* of  
**7 in 10 residents**

\* Singapore residents to have accessed NLB's physical touchpoints, digital platforms and/or content



**1.56m**  
No. of participants in NLB's programmes & exhibitions

Our collection is well-utilised.



**82.2m**  
No. of pageviews to NLB's websites  
(e.g. Archives Online, Infopedia) and  
usage of e-books and e-databases



**35.8m**  
Overall loans

<sup>^</sup> 3-year average

# LEVERAGING NLB'S ASSETS, BRAND AND REACH

## OVER 13M IN MULTIPLE FORMATS FOR MULTIMODAL DISCOVERY AND LEARNING

### Lending Collection

**Over 5m**  
print collection

**Over 1.7m**  
e-books and audio books

Search a **78**  
databases via the Singapore National Library Board (NLB) website

**Over 2.4m**  
music tracks

**Over 7,400**  
e-newspapers and e-magazines titles

Keen to pick up a new skill in design, business or more? **Over 8,000** e-learning courses

### Reference Collection

**Over 560,000**  
Singapore & SEA items

**Over 147,000**  
Chinese, Malay & Tamil Languages items

**Over 53,000**  
Arts items

**Over 62,000**  
Social Sciences & Humanities items

**Over 39,000**  
Science & Technology items

**Over 19,000**  
Rare Materials items

### Archival Materials

**Over 1.14m**  
Photographs

**Over 290,000**  
Government files & Parliament papers

**Over 190,000**  
Audiovisual & sound recordings

**Over 70,000**  
Maps & building plans

**Over 55,000**  
Speeches & press releases

**Over 35,000**  
Oral history interviews

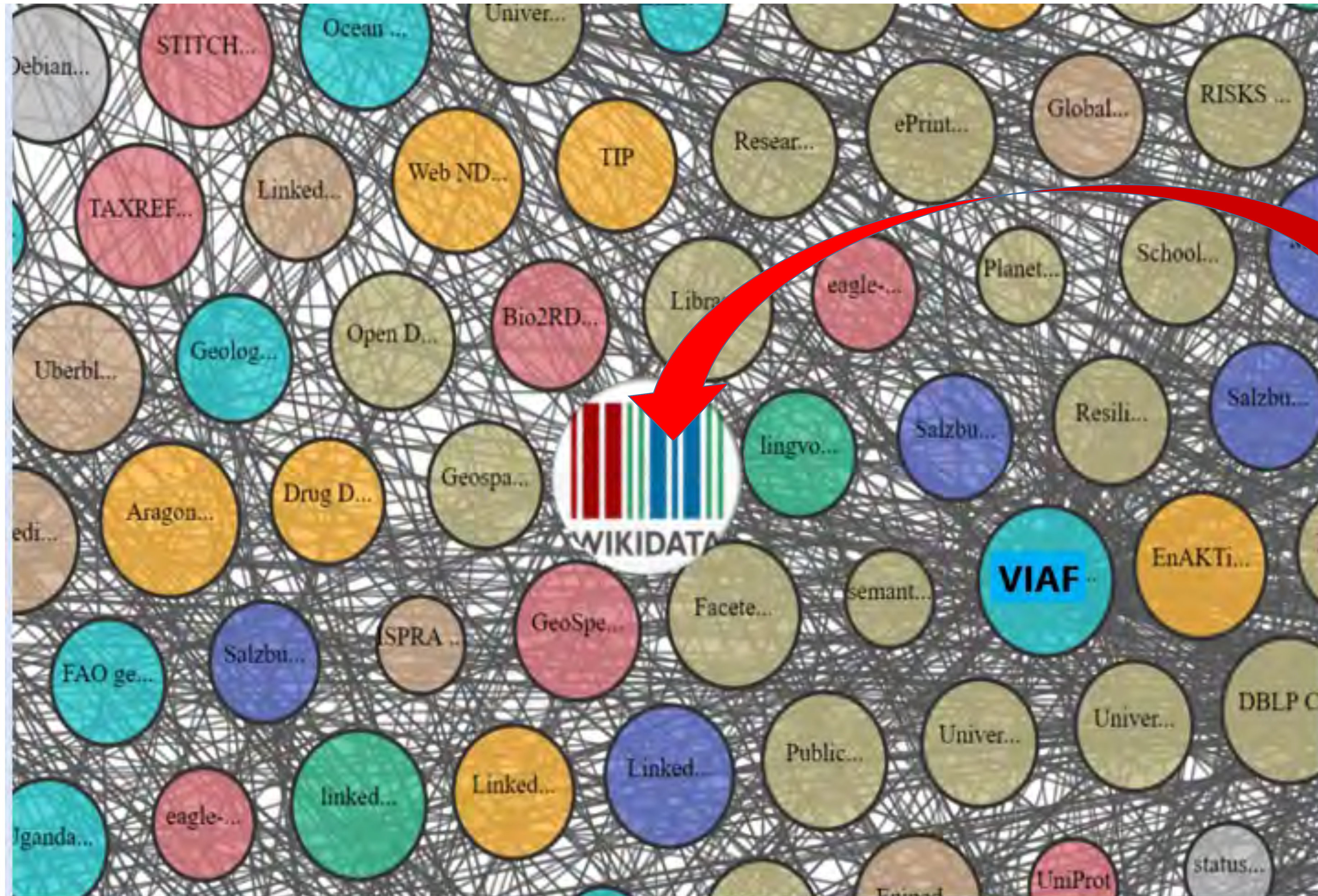
**Over 7,000**  
Posters

# WHAT ARE NLB NAMED ENTITIES ?

- **Unique** & **standardised** forms of person, organisation & place **names** used in our catalogue records
- **Consistent** and **unambiguous reference** to these entities in NLB's cataloguing of library materials
- To improve users' **search and discovery**
- Some examples of "**Named entities**":

<b>People</b> Lee, Kuan Yew	<b>People</b> Abisheganaden, Alex	<b>Place</b> Sultan Mosque	<b>Place</b> Central Fire Station
<b>People</b> Rajaratnam, S.	<b>People</b> Zubir Said	<b>Organisation</b> Raffles Institution	<b>Organisation</b> Singapore Film Society

# FROM NLB TO THE WIKIDATA CLOUD



**URI of Lee Kuan Yew**  
<https://eresources.nlb.gov.sg/IDDOC/NLBDM/vocab/w/W1ykB8puD4>

# Background on Wikidata

## Google Knowledge Graphs



A screenshot of a Google Knowledge Graph for Gregory Peck. It features a grid of five images of Peck at different stages of his career. Below the images, the name "Gregory Peck" is displayed in large text, followed by "American actor". A short biographical paragraph follows, mentioning his Academy Award win for "To Kill a Mockingbird". At the bottom, his birth and death information is listed.

**Gregory Peck**  
American actor

Eldred Gregory Peck was an American actor. He was one of the most popular film stars from the 1940s to the 1960s. Peck received five nominations for Academy Award for Best Actor and won once – for his performance as Atticus Finch in the 1962 drama film *To Kill a Mockingbird*. [Wikipedia](#)

**Born:** 5 April 1916, La Jolla, California, United States  
**Died:** 12 June 2003, Los Angeles, California, United States

## Wikidata

- (1) Structured & Linked Data
- (2) Human & machine-friendly
- (3) Used in various applications

Knowledge on the Go  
**Siri** has answers to all kinds of questions.

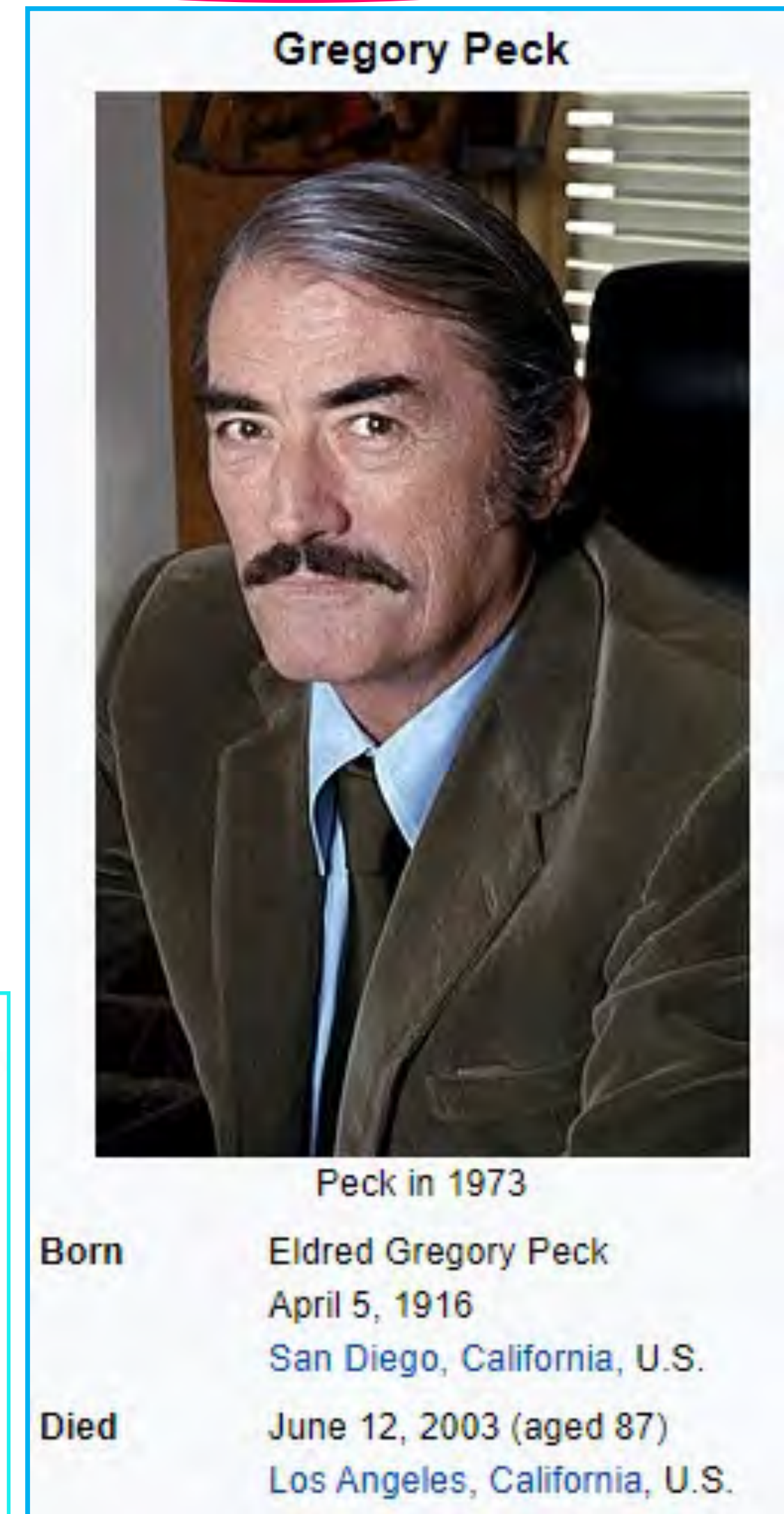
Quickly check facts, do calculations or translate a phrase into another language. It's as simple as asking. Even when you don't ask, Siri works behind the scenes like a personal assistant. A great example is the Siri watch face, which serves up events, news, suggestions and more when you need them.

## Inside the Alexa-Friendly World of Wikidata

*Virtual assistants do their jobs better thanks to Wikidata, which aims to (eventually) represent everything in the universe in a way computers can understand*



## Wikipedia Infoboxes



A screenshot of a Wikipedia infobox for Gregory Peck. It features a large portrait of Peck in a suit. Below the portrait, his name "Gregory Peck" is written. Underneath, it says "Peck in 1973". The infobox lists his birth and death information.

**Gregory Peck**

Peck in 1973

**Born** Eldred Gregory Peck  
April 5, 1916  
San Diego, California, U.S.

**Died** June 12, 2003 (aged 87)  
Los Angeles, California, U.S.



**Label** (for human)

**Identifier** (for machine)

Language	Label	Description	Also known as
English	Lee Kuan Yew	1st Prime Minister of Singapore (1923-2015)	LKY Harry Lee MM Lee Li Guangyao
Chinese	李光耀	新加坡建国之父、首任总理	新加坡国父 新加坡建国总理
Malay	Lee Kuan Yew	Perdana Menteri Singapura pertama	Lee Kwan Yew
Tamil	லீ குவான் யூ	சிங்கப்பூரின் முதல் பிரதம மந்திரி	

**Statements**

- instance of: human (2 references)
- date of birth: 16 September 1923 (Gregorian) (9 references)
- place of birth: Singapore (2 references)
- date of death: 23 March 2015 (statement is subject of: Death of Lee Kuan Yew) (4 references)
- place of death: Singapore General Hospital (2 references)

**Lee Kuan Yew**

From Wikipedia, the free encyclopedia

"LKY" redirects here. For other uses, see LKY (disambiguation).  
In this Chinese name, the family name is Lee (李).

**Lee Kuan Yew** GCMG CH SPMJ (born **Harry Lee Kuan Yew**; 16 September 1923 – 23 March 2015), often referred to by his initials **LKY**, was a Singaporean statesman and lawyer who served as Prime Minister from 1959 to 1990. As one of the founding members of the People's Action Party, Lee is recognised as the nation's founding father, credited with rapidly transitioning the country from a "developing third world country within a single generation" under his leadership.<sup>[4][5][6]</sup> Lee died of pneumonia on 23 March 2015, aged 91. In a week of national mourning, about 1.7 million Singaporeans and world leaders paid tribute to him at his lying-in-state at Singapore's Parliament House and at community tribute sites around the country.

After attending the London School of Economics, Lee transferred to and graduated from Fitzwilliam College, Cambridge, with starred-first-class honours in law. He became a barrister of the Middle Temple practised law until 1959. Lee co-founded the People's Action Party (PAP) in 1954 and was its first secretary-general until 1992, leading the party to eight consecutive victories in the general elections. Lee was prime minister in 1990 and was senior minister under his successor Goh Chok Tong until 2004, then minister mentor under Lee Hsien Loong until 2011. Lee held successive ministerial positions for 56 years, including the Tanjong Pagar seat until his death in 2015.<sup>[7][8]</sup>

Lee campaigned for Britain to relinquish its colonial rule, and eventually attained through a national referendum a merger with other former British territories in Southeast Asia to form Malaysia in 1963. However, ideological differences led to Singapore's separation, becoming a sovereign city-state two years later. With overwhelming parliamentary control at every election, Lee oversaw Singapore's transformation from a crown colony with a natural deep harbour into a developed country with a high-income economy. In the process, he forged a system of meritocratic, highly effective and anti-corrupt<sup>[9]</sup> government and eschewed populist policies in favour of long-term social and economic planning. He championed meritocracy<sup>[10]</sup> and multiracialism<sup>[11][12]</sup> as governing principles, making English the lingua franca<sup>[13]</sup> to facilitate trade with the world, whilst mandating bilingualism in schools to preserve students' mother tongue and ethnic identity.<sup>[13]</sup>

Lee's rule was criticised by various observers from liberal democracies of the West<sup>[14]</sup> for curtailing civil liberties such as freedom of the press, limits on public protests, restricting labour movements from bringing defamation lawsuits against some political opponents.<sup>[16]</sup> He argued that such disciplinary measures were "Asian values"<sup>[17][18][19]</sup> and that the measures were necessary for political stability and economic growth. Singapore underwent rapid development and experienced high economic growth under his leadership, and by the end of his tenure had become one of the most prosperous nations, with a highly developed economy, strong international trading links and a leading financial centre.<sup>[6][7]</sup>

**Contents** [hide]

- Family background
- Personal life
- Education
- Early life
- Early political career (1951–1959)
  - Fajar trial
  - Formation of the People's Action Party
  - Leader of the Opposition

Structured data from Wikidata displayed in Wikipedia Infobox

His Excellency  
**Lee Kuan Yew**  
GCMG CH SPMJ  
李光耀

Lee Kuan Yew in 2011

**1st Prime Minister of Singapore**

**In office**  
3 June 1959<sup>[1]</sup> – 28 November 1990

**President** *See list* [show]

**Deputy** *See list* [show]

**Preceded by** Lim Yew Hock (as Chief Minister)

**Succeeded by** Goh Chok Tong

**Member of Parliament for Tanjong Pagar**

**In office**  
2 April 1955 – 23 March 2015

**Constituency** Tanjong Pagar (Assembly) (1955–65)  
Tanjong Pagar SMC (1965–91)  
Tanjong Pagar GRC (1991–2015)

**Offices after 1965** [show]

**Offices before 1965** [show]

**Personal details**

**Born** Harry Lee Kuan Yew  
16 September 1923  
Singapore, Straits Settlements

**Died** 23 March 2015 (aged 91)  
Singapore General Hospital, Singapore

# WIKIDATA STRUCTURE - STATEMENTS

Wikidata page for Lee Kuan Yew (Q131248). The page shows the item name, description, and a table of labels in various languages. Below this is the 'Statements' section, which lists properties and their values. Three statements are highlighted with red boxes and lines:

- instance of** (Property) points to **human** (Value).
- date of birth** (Property) points to **16 September 1923** (Value).
- place of birth** (Property) points to **Singapore** (Value).

- Store triples in a form that can be queried:
- *item – property – value = subject – predicate – object*

**Item**  
(i.e. Concept)  
**ID represented by prefix "Q"**

**Property**  
(i.e. Predicate)  
**ID represented by prefix "P"**

**Value**

- **Item**
- **Identifier**
- **Literal** (date, free text)

## Example of "Item"

**human** (Q5)

common name of Homo sapiens, unique extant species of the genus Homo

human being | humankind | people | homosapiens | person | mankind | peoplekind | personkind | persons | humans | Human

# WHAT ARE WE TRYING TO DO ?

- Wikidata serves as a linking hub for URIs from different institutions globally
- NLB's identifiers (URIs) are linked to corresponding entities in Wikidata
- By contributing identifiers to Wikidata, Singapore entities are grouped in an equivalent relationship (owl:sameAs) with similar entities

**Deutsche National Bibliothek**

<b>Link zu diesem Datensatz</b>	<a href="http://d-nb.info/gnd/118997734">http://d-nb.info/gnd/118997734</a>
<b>Person</b>	Lee, Kuan Yew
<b>Geschlecht</b>	männlich
<b>Andere Namen</b>	Lee Kuan Yew Li Kuan Ju Ли Куан Ю Lee, Kuan-yew



The Library of Congress > Linked Data Service > LC Name Authority File (LCNAF)

Lee, Kuan Yew, 1923-2015

URI(s)  
- <http://id.loc.gov/authorities/names/n50039577>

Lee Kuan Yew (Q131248)

1st Prime Minister of Singapore (1923-2015)  
LKY | Harry Lee | MM Lee | Li Guangyao

WIKIDATA

Language	Label	Description	Also known as
English	Lee Kuan Yew	1st Prime Minister of Singapore (1923-2015)	LKY Harry Lee MM Lee Li Guangyao
Chinese	李光耀	新加坡建国之父、首任总理	新加坡国父 新加坡建国总理
Malay	Lee Kuan Yew	Perdana Menteri Singapura pertama	Lee Kwan Yew
Tamil	லீ குவான் யூ	சிங்கப்பூரின் முதல் பிரதம மந்திரி	

Library of Congress authority ID: [n50039577](#)

GND ID: [118997734](#)

National Library Board Singapore ID: [wW1ykB8puD4](#)



# WHY ARE WE DOING THIS ?

## Benefits for NLB:

- Interlinking between NLB's and other institutions' collections to increase global discovery of NLB's collection
- The traditional catalogue can be enhanced with relevant information from other knowledge bases

## Benefits for users:

- Greater access and discovery of information from other institutions

Sample catalog record - University of Wisconsin-Madison Libraries

Search > Catalog > No man is an island : a study of Singa...

No man is an island : a study of Singapore's Lee Kuan Yew

Minchin, James

☆ Save to Lists

× Information from the Web ⓘ

Lee, Kuan Yew, 1923-2015

BORN September 16, 1923  
DIED March 23, 2015

Abstract (Source: DBpedia)  
Lee Kuan Yew (born Harry Lee Kuan Yew; 16 September 1923 – 23 March 2015), often referred to by his initials LKY, was a Singaporean politician and lawyer who served as the first Prime Minister of Singapore from 1959 to 1990. As one of the founding members of the People's Action Party, Lee is recognised as the nation's founding father, credited with rapidly transitioning the country

LIBRARY LIBRARY OF CONGRESS

Everything Search Loc.gov

The Library of Congress > Linked Data Service > LC Name Authority File (LCNAF)

Lee, Kuan Yew, 1923-2015

URI(s)  
- <http://id.loc.gov/authorities/names/n50039577>

Instance Of  
- MADS/RDF PersonalName  
- MADS/RDF Authority  
- SKOS Concept ↗

Scheme Membership(s)  
- Library of Congress Name Authority File

Collection Membership(s)  
- Names Collection - Authorized Headings  
- LC Names Collection - General Collection

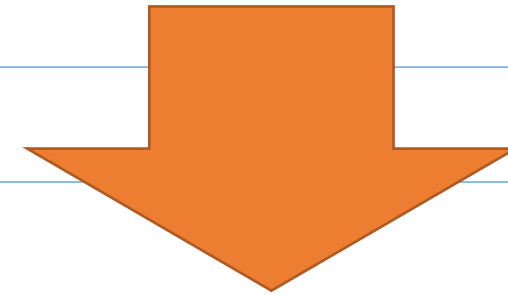
Identities on The Web  
<http://id.loc.gov/authorities/names/n50039577>  
<http://viaf.org/viaf/97994136>  
<http://www.wikidata.org/entity/Q131248>  
[http://dbpedia.org/resource/Lee\\_Kuan\\_Yew](http://dbpedia.org/resource/Lee_Kuan_Yew)

Subject Of Works  
83 resources filter Page 1 of 1 >  
21 shi ji chu ban she. Shen mai xin zhong de mi mi  
Xinjiapo yu Binlangyu de gu shi. 深埋心中的秘密 :  
新加坡与槟榔屿的故事  
Ang, Cheng Guan. Lee Kuan Yew's strategic  
thought  
Barr, Michael D. Lee Kuan Yew the beliefs behind  
the man  
Cai, Yulin, 1944- Li Guangyao shi dai vs hou Li  
Guangyao shi dai 李光耀时代vs后李光耀时代  
Centre for Liveable Cities, Singapore, chance of a

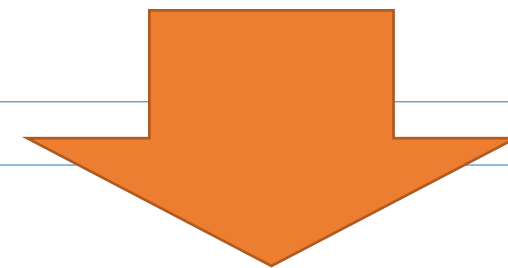
# HOW ARE WE DOING THIS ?



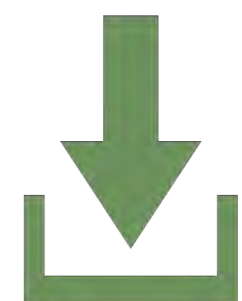
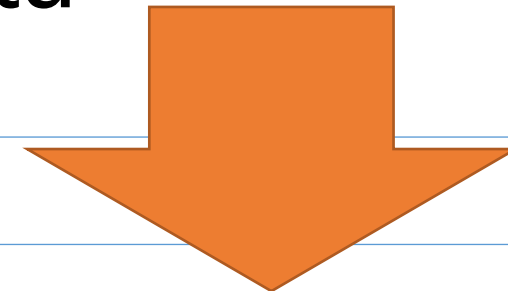
1. Create a new property for NLB ID



2. Identify a list of Singapore related entities that are available in Wikidata



3. Compare & match NLB entities with Wikidata



4. Insert NLB IDs into corresponding Wikidata entity descriptions



# CREATION OF A NEW PROPERTY FOR NLB ID

- Seek help from Wikidata community on how to proceed.
- Step by step guide in Wikidata page on proposing a property for authority control:  
[https://www.wikidata.org/wiki/Wikidata:Property\\_proposal/Authority\\_control](https://www.wikidata.org/wiki/Wikidata:Property_proposal/Authority_control)
- Data that is needed:
  - Label
  - Description
  - Data type
  - Domain
  - Example (i.e. Sample URI)
  - Formatter URL (i.e. prefix of URI)
- **Approval** by Wikidata community

# NATIONAL LIBRARY BOARD SINGAPORE ID

<https://www.wikidata.org/wiki/Property:P3988>

**National Library Board Singapore ID** (P3988)

Identifier of a person, organisation or place from the name authorities of National Library Board.  
Singapore National Library Board ID | NLB ID edit

**In more languages**  
Configure

Language	Label	Description	Also known as
English	National Library Board Singapore ID	Identifier of a person, organisation or place from the name authorities of National Library Board.	Singapore National Library Board ID
Chinese	No label defined	No description defined	
Malay	No label defined	No description defined	
Tamil	No label defined	No description defined	

All entered languages

**Data type**

External identifier

**Statements**

instance of

- Wikidata property for authority control for people edit  
0 references + add reference
- Wikidata property to identify organisations edit  
0 references + add reference
- Wikidata property for authority control for places edit  
0 references + add reference  
+ add value

country

- Singapore edit  
0 references + add reference  
+ add value

has quality

- VIAF component edit  
code: NLB  
0 references + add reference

# POPULATING IDENTIFIERS...MANUALLY

- Manual method: Look up NLB entities available in Wikidata and add in NLB identifiers item by item
- Started by looking up entities in Wikidata that match NLB's and insert identifiers of NLB entities
- Very time consuming, especially when working with huge amount of data

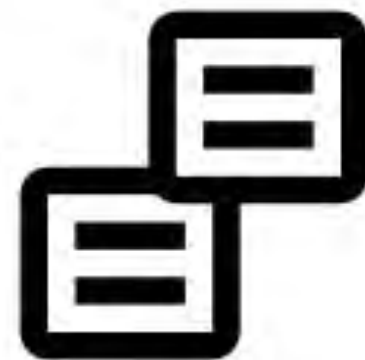


# SEARCHED FOR A TOOL...

- That can automate **MATCHING** of names in TTE with that in Wikidata in bulk
- That can automate **POPULATING** NLB IDs into the associated Wikidata items in bulk

# AUTOMATION FOR MATCHING AND POPULATING IN WIKIDATA

- Wikidata Data Donation: How to contribute data to Wikidata [https://www.wikidata.org/wiki/Wikidata:Data\\_donation](https://www.wikidata.org/wiki/Wikidata:Data_donation)
  - Mix'n'match
  - OpenRefine

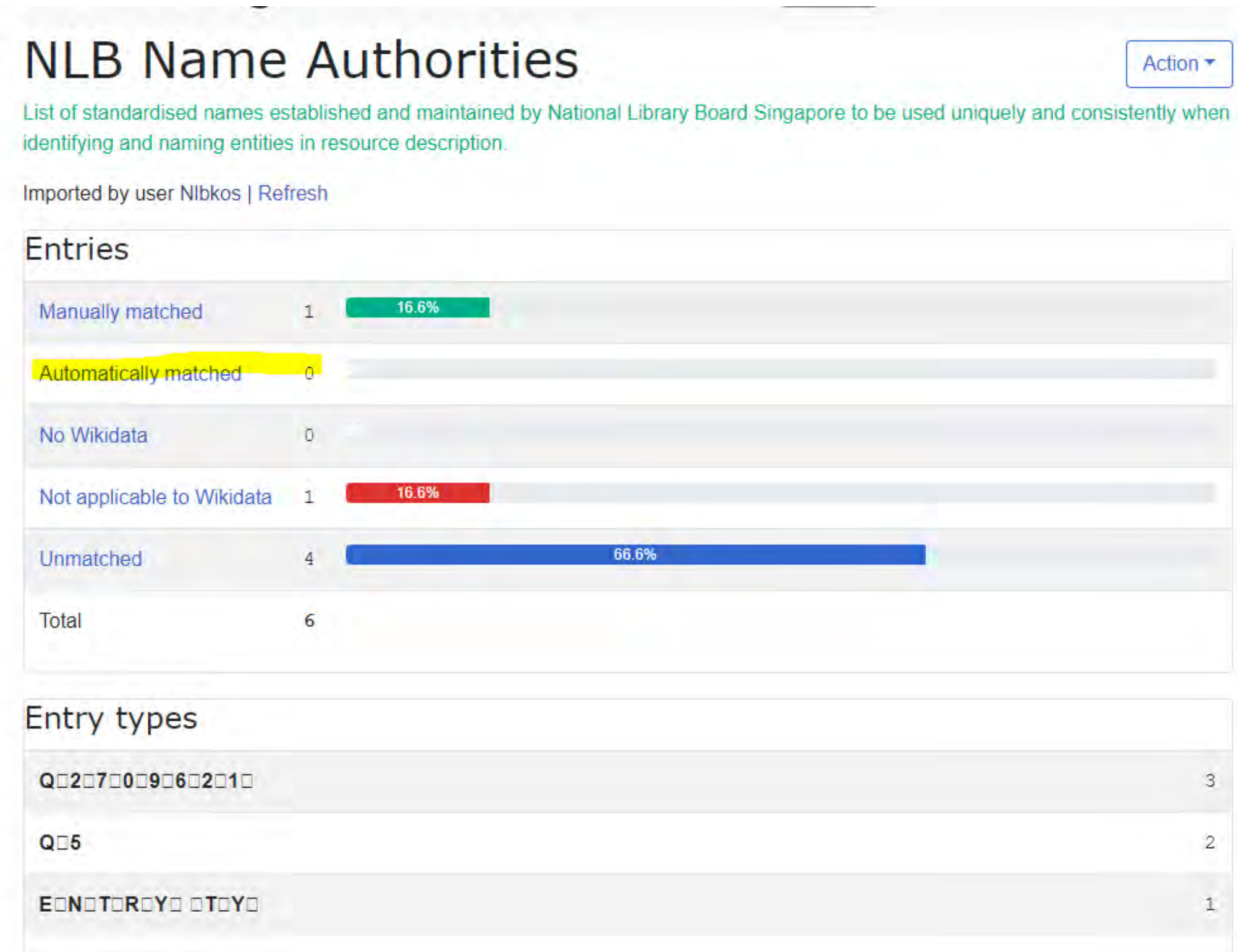


## Matching your data with other data sets

Many tools exist to augment or cross-check your data with data from other sources. One example is [BEACON](#) which can give you a list of your IDs matched against any other we have for the same items, e.g Art UK artists against VIAF. [Mix'n'match](#) and [OpenRefine](#) can be used to create connections between your database and Wikidata.

# Mix'n'match

- Match NLB entities with the same Wikidata item record
- Tested using 6 sample NLB entities
- Limitation:
  - Display error
  - Unable to match any names using automation
  - Unable to populate NLB ID in Wikidata



### NLB Name Authorities

List of standardised names established and maintained by National Library Board Singapore to be used uniquely and consistently when identifying and naming entities in resource description.

# <b>Pouiaou</b> Uobion	Places	Not matched
# <b>Singapore</b> Storiait	Places	Not matched
# <b>Foorot</b> Caanioning Hi	Places	Not matched
# "Lee, Kuaan Yeow"	People	Not matched

# EXPERIMENTATION OF ONLINE TOOLS TO AUTOMATE PROCESSES



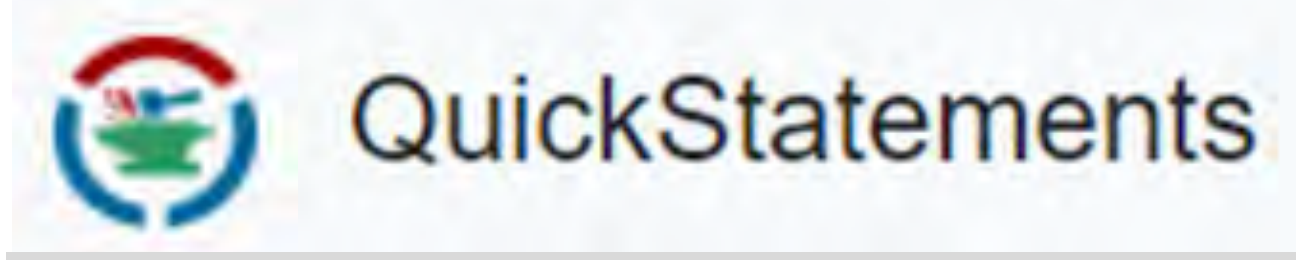
- Tried various online tools to **match** NLB named entities with Wikidata's and **populate** Wikidata with NLB's entity identifiers
- What are the pros and cons of each tool?
- Needed to combine 3 tools to complete the tasks



✓ For generating list of Singapore related Wikidata entities with NLB IDs

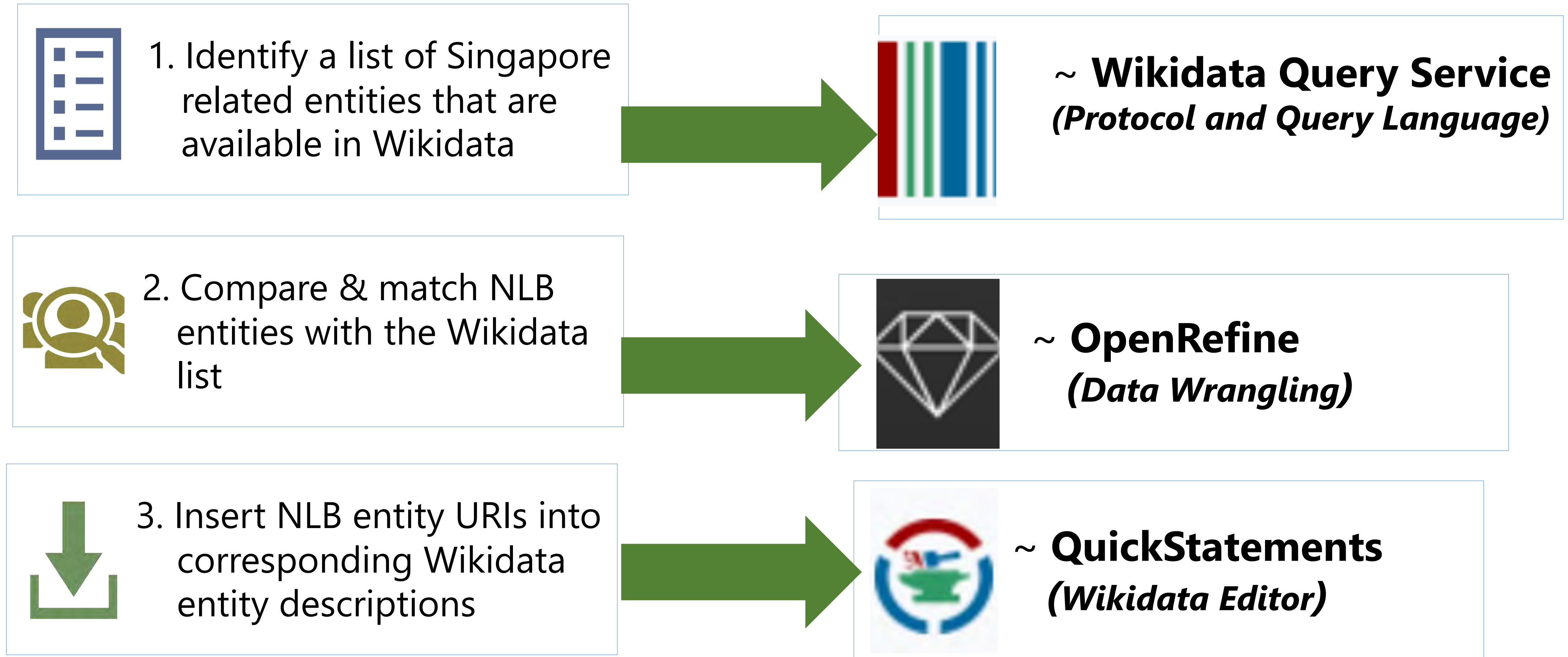


✓ For matching NLB named entities with that of Wikidata



✓ For populating NLB named entities IDs into Wikidata

# HOW DO THE 3 TOOLS WORK TOGETHER?



These 3 tools were found to be effective as well as time & labour-saving in completing the tasks

# WHAT IS OPENREFINE?

- More than cleaning up messy data
- Automatic bulk **MATCHING** of labels with those in Wikidata
- **MAPPING** attributes of named entities to corresponding properties in Wikidata
- Automatic **POPULATION** of data into Wikidata in bulk

# OPENREFINE – FINDING MATCHES IN WIKIDATA (PREPARATION)

English names in direct form for matching with Wikidata **Label**

Non-English names for matching with Wikidata **Aliases**

NLB ID to be uploaded to Wikidata

UID	Descriptor	Vocabulary	Name (DisplayName)	Name (DisplayName) - for matching	Chinese PT	Chinese PT - for matching	Description	Nationality	Place of Birth	VIAF ID	LDMS ID
18578086	Chua, Anthony	People	Anthony Chua	Anthony Chua			Artist working primarily with Chinese ink, influenced by both traditional Chinese ink and Western techniques. He is known for his abstract paintings and paintings of urban landscapes and old buildings. He won the Young Artist Award in 2001.	Singapore	Singapore	12148812785045211674	KyuloF2ZQjY
18662139	Chua, Chee Lay	People	Chua Chee Lay	Chua Chee Lay	蔡志礼	蔡志礼	Chinese linguist, educator and poet. He was also Lee Kuan Yew's Chinese language tutor.	Singapore		41574829	Igea6alsE4U
18658591	Chua, Enlai	People	Chua Enlai	Chua En Lai			Stage and television actor, and comedian. He is best known for his work on the television series, "The Noose" on MediaCorp Channel 5.	Singapore	Singapore	294519361	kLGJvhT2VZ8
18663237	Chua, Mui Hoong	People	Chua Mui Hoong	Chua Mui Hoong			Political journalist and opinion editor at Straits Times since 1991. She is the author of "A Defining Moment: How Singapore beat Sars" (2004) and "Singapore, Disrupted" (2018).			76033648	8qGZ1hb0mss

Additional Wikidata properties to assist OpenRefine to find the correct match:

- **VIAF ID** (P214)
- **Nationality** (Q231002) or **Place of birth** (Q1322263) that contain value **Singapore** (Q334)

# OPENREFINE – FINDING MATCHES IN WIKIDATA (RECONCILIATION)

144 rows Extensions: Wikidata

Show as: rows records Show: 5 10 25 50 rows first previous 1 - 50 next last

UID	Descriptor	Vocabulary	Name (DisplayN	Name (DisplayName) - for ma	Chinese PT	Chinese PT - for	Description	Nationality	Place of Birth	VIAF ID	LDMS ID
18578086	Chua, Anthony	_People	Anthony Chua	Anthony Chua <input checked="" type="checkbox"/> Anthony Chuang (92) <input checked="" type="checkbox"/> Anthony Chuang (92) <input type="checkbox"/> Create new item <a href="#">Search for match</a>			policy and political thought. Artist working primarily with Chinese ink, influenced by both traditional Chinese ink and Western techniques. He is known for his abstract paintings and paintings of urban landscapes and old buildings. He won the Young Artist Award in 2001.	Singapore	Singapore	12148812785045211674	KyuloF2ZQjY
18662139	Chua, Chee Lay	_People	Chua Chee Lay	Chua Chee Lay <input type="checkbox"/> Create new item <a href="#">Search for match</a>	蔡志礼	蔡志礼	Chinese linguist, educator and poet. He was also Lee Kuan Yew's Chinese language tutor.	Singapore		41574829	Igea6alsE4U
18658591	Chua, Enlai	_People	Chua Enlai	Chua Err Lai <input checked="" type="checkbox"/> Choose new match			Stage and television actor, and comedian. He is best known for his work on the television series, "The Noose" on MediaCorp Channel 5.	Singapore	Singapore	294519361	kLGJvhT2VZ8
18663237	Chua, Mui Hoong	_People	Chua Mui Hoong	Chua Mui Hoong <input checked="" type="checkbox"/> Create new item <a href="#">Search for match</a>			Political journalist and opinion editor at Straits Times since 1991. She is the author of "A Defining Moment: How Singapore beat Sars" (2004) and "Singapore.			76033648	8qGZ1hb0mss

Multiple potential matches for checking

No matches in Wikidata

Name matched with Wikidata label




# OPENREFINE – MAPPING PROPERTIES



144 rows Schema \* Issues 286 Preview

The Wikidata schema below specifies how your tabular data will be transformed into Wikidata edits. You can drag and drop the column names below in most input boxes: for example, the column 'Name (DisplayName) - for matching' can be dragged into the 'Label' box to generate labels for the Wikidata entity.

UID Descriptor Vocabulary Name (DisplayName) Name (DisplayName) - for matching Wikidata ID\_3 Wikidata ID\_2 Wikidata ID Chinese PT Chinese PT - for matching  
Description 4 Description 5 Description 6 Description 7 Description 8 Description 9 Description 10 Description 11 Description 12 Nationality Place of Birth VIAF


Name (DisplayName) - for matching  remove

**Terms**

Label	en	Name (DisplayName) - for mat... 	<input checked="" type="checkbox"/> override if present	<span>remove</span>
Description	en	Description 1 	<input checked="" type="checkbox"/> override if present	<span>remove</span>

+ add term

**Statements**

National Library Boa	LDMS ID 	<span>remove</span>	+ add qualifier
instance of	human	<span>remove</span>	+ add reference + add value

+ add statement  
+ add item

**Annotations:**



- Wikidata properties:** Points to the 'Terms' and 'Statements' sections.
- NLB named entities' attributes and their corresponding values:** Points to the 'LDMS ID' value in the 'National Library Boa' statement.
- Values - Wikidata entities:** Points to the 'human' value in the 'instance of' statement.

# OPENREFINE – PREVIEW OF MAPPED PROPERTIES BEFORE UPLOAD TO WIKIDATA

144 rows   Schema   Issues **286**   **Preview**   Extensions: Wikidata

This tab shows the first edits (out of 143) that will be made once you upload the changes to Wikidata. You can use facets to inspect the edits on particular items.

Ahmad Jaaffar Munasip (Q98807227)

Label (override if present)	Ahmad Jaaffar Munasip (English)
Description (override if present)	Malay author whose short story collection, "Jago yang Terlupa Dilupakan" (2011) won the Singapore Literature Prize Merit award
National Library Board Singapore ID (P3988)	 g6Alyd5kJ1E ▶ 0 references
instance of (P31)	 human (Q5) ▶ 0 references

Ahmad Nisfu (Q20450363)

Label (override if present)	Ahmad Nisfu (English)
National Library Board Singapore ID (P3988)	 u_P_5A5YFWg ▶ 0 references
instance of (P31)	 human (Q5) ▶ 0 references

# UPLOAD TO WIKIDATA – QUICKSTATEMENTS (1)

OpenRefine can also be used to upload data to Wikidata, but we found QuickStatements a better choice:

- ❑ OpenRefine has no alerts when the data fails to upload
- ❑ Unstable – OpenRefine succeeded in uploading some but not others



The screenshot shows the OpenRefine interface with a Wikidata extension menu open. The menu options are: Edit Wikidata schema, Manage Wikidata account, Import schema, Export schema, Upload edits to Wikidata (highlighted in yellow), and Export to QuickStatements. The main interface shows a table with 144 rows and 286 issues. The selected row is for Ahmad Jaaffar Munasip (Q98807227), with a label in English and a description in Malay. The National Library Board Singapore ID (P3988) is also visible.

# UPLOAD TO WIKIDATA – QUICKSTATEMENTS (2)

## What is QuickStatements?

- Use of text commands to edit and create Wikidata items

### i. Command sequence syntax

```
Q93217137      Len      "Pan Shou"
Q93217137      Den      "Accomplished Chinese calligrapher and poet who is known for his distinctive calligraphic style.
Born in China, he came to Singapore in 1929 at the age of 19. It was only after retirement from a successful career in
education that his second career in calligraphy and poetry flourished. He had been honorary adviser to the Chinese
Calligraphy Society of Singapore since 1977."
Q93217137      P31      Q5
Q93217137      P106     Q3303330
Q93217137      P1416    Q720430
Q93217137      P3988    "a3Jbw83C0gU"

Q93217140      Len      "Colin Schooling"
Q93217140      Den      "Businessman and softball player who represented Singapore in the 1968 Asian Men's Softball
Championship in Manila, Philippines and the Third Men's World Softball championship in 1972. He is treasurer of Singapore
Swimming Association, father of swimmer Joseph Schooling and nephew of Lloyd Valberg."
Q93217140      P31      Q5
Q93217140      P106     Q13388586
Q93217140      P1416    Q19577887
Q93217140      P3988    "511s_T1vM4k"
```

### ii. CSV file syntax

```
|qid,Len,Den,P31,P106,P1416,P3988
Q93217137,Pan Shou,Accomplished Chinese calligrapher and poet who is known for his distinctive calligraphic style. Born in
China, he came to Singapore in 1929 at the age of 19. It was only after retirement from a successful career in education
that his second career in calligraphy and poetry flourished. He had been honorary adviser to the Chinese Calligraphy
Society of Singapore since 1977.,Q5,Q3303330,Q720430,a3Jbw83C0gU
Q93217140,Colin Schooling,Businessman and softball player who represented Singapore in the 1968 Asian Men's Softball
Championship in Manila, Philippines and the Third Men's World Softball championship in 1972. He is treasurer of Singapore
Swimming Association, father of swimmer Joseph Schooling and nephew of Lloyd Valberg.,Q5,Q13388586,Q19577887,511s_T1vM4k
```

# CONVERT RECONCILED DATA FROM OPENREFINE TO QUICKSTATEMENTS

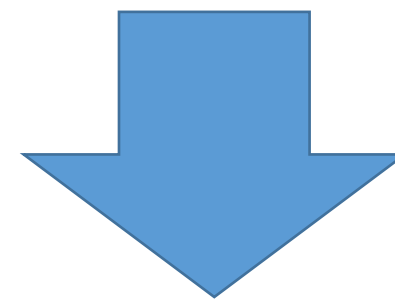
10 rows Schema Issues 1 Preview Extensions Wikibase

This tab shows the first edits (out of 1) that will be made once you upload the changes to Wikidata. You can use facets to inspect the

Moses Lim (Q6915857)

National Library Board Singapore ID (P3988) on\_r\_I12Qw0 0 references

- Manage Wikibase instances...
- Edit Wikibase schema
- Manage schemas...
- Manage Wikibase account...
- Upload edits to Wikibase...
- Export to QuickStatements**



Create new command batch for Wikidata as batch name

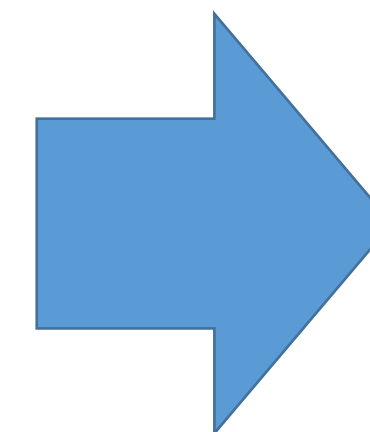
Q6915857 P3988 "on\_r\_I12Qw0"

Wikidata item ID

Wikidata property ID for NLB ID

NLB ID

Import V1 commands Import CSV commands



Moses Lim (Q6915857) edit

Singaporean actor

In more languages

Language	Label	Description	Also known as
English	Moses Lim	Singaporean actor	
American English	No label defined	No description defined	
Chinese	林益民	No description defined	
Malay	No label defined	No description defined	

All entered languages

Identifiers

VIAF ID 68624519 1 reference + add value

Library of Congress authority ID no2003083295 0 references + add reference + add value

National Library Board Singapore ID on\_r\_I12Qw0 0 references + add reference + add value

# CSV TO QUICKSTATEMENTS: POPULATE NLB ID IN EXISTING WIKIDATA ITEMS

Reconciled data in OpenRefine exported in csv format

	A	B
1	Name (DisplayName) (TTE)	National Library Board Singapore ID
2	Aw Boon Par	UJEENdBEshE
3	Darryl David	xcX_bdHv4H4

Wikidata item ID

Wikidata property ID for NLB ID

	A	B
1	qid	P3988
2	Q15995841	UJEENdBEshE
3	Q20979710	xcX_bdHv4H4

**Darryl David** (Q20979710)

Singaporean politician edit

**In more languages**

Language	Label	Description	Also known as
English	Darryl David	Singaporean politician	
Chinese	达瑞尔 大卫	新加坡政治人物	杰乐
Malay	No label defined	No description defined	
Tamil	No label defined	No description defined	

**Identifiers**

National Library Board Singapore ID xcX\_bdHv4H4 edit

+ 0 references

+ add reference

+ add value

Tool that converts csv file into Quickstatements

**Upload CSV file**

Select a CSV file.

Choose File No file chosen

Validator

**Result**

Q15995841	P3988	"UJEENdBEshE"
Q20979710	P3988	"xcX_bdHv4H4"

Batch on Wikidata by Nlbkos [Batches] Discuss/revert batch

Status: DONE 100% (2) of 2 done

1	done	Aw Boon Par [Q15995841]	ADD	Statement	National Library Board Singapore ID [P3988]: "UJEENdBEshE"
2	done	Darryl David [Q20979710]	ADD	Statement	National Library Board Singapore ID [P3988]: "xcX_bdHv4H4"

First Page 1 Last

All errors Init 10

# CSV TO QUICKSTATEMENTS: CREATE NEW ITEMS IN WIKIDATA

	A	B	C	D
1	Name (DisplayName)	Description	Vocab Class	LDMS URI
2	Pan Shou	Accomplished Chinese calligrapher and poet who is known for his distinctive calligraphic style. Born in China, he came to Singapore in 1929 at the age of 19. It was only after retirement from a successful career in education that his second career in calligraphy and poetry flourished. He had been honorary adviser to the Chinese Calligraphy Society of Singapore since 1977.	People	a3Jbw83C0gU
3	Colin Schooling	Businessman and softball player who represented Singapore in the 1968 Asian Men's Softball Championship in Manila, Philippines and the Third Men's World Softball championship in 1972. He is treasurer of Singapore Swimming Association, father of swimmer Joseph Schooling and nephew of Lloyd Valberg.	People	51ls_TlvM4k

Colin Schooling (Q93217140)

No description defined edit

**In more languages**

Language	Label	Description	Also known as
English	Colin Schooling	No description defined	
Chinese	No label defined	No description defined	
Malay	No label defined	No description defined	
Tamil	No label defined	No description defined	

All entered languages

**Statements**

instance of human edit

- 0 references

[+ add reference](#)

[+ add value](#)

Wikidata item ID (auto generated)

Wikidata Label in Eng

Wikidata Description in Eng

instance of (P31): human (Q5)

Wikidata property ID for NLB ID

	A	B	C	D	E
1	qid	Len	Den	P31	P3988
2		Pan Shou	Accomplis	Q5	a3Jbw83C0gU
3		Colin Schooling	Businessn	Q5	51ls_TlvM4k

**CREATE** Item *en*:Pan Shou

*en*:Accomplished Chinese calligrapher and poet who is known for his distinctive calligraphic style. Born in China, he came to Singapore in 1929 at the age of 19. It was only after retirement from a successful career in education that his second career in calligraphy and poetry flourished. He had been honorary adviser to the Chinese Calligraphy Society of Singapore since 1977.

instance of [P31]:human [Q5]

National Library Board Singapore ID [P3988]:"a3Jbw83C0gU"

---

**CREATE** Item *en*:Colin Schooling

*en*:Businessman and softball player who represented Singapore in the 1968 Asian Men's Softball Championship in Manila, Philippines and the Third Men's World Softball championship in 1972. He is treasurer of Singapore Swimming Association, father of swimmer Joseph Schooling and nephew of Lloyd Valberg.

instance of [P31]:human [Q5]

National Library Board Singapore ID [P3988]:"51ls\_TlvM4k"

**Result**

```

CREATE
LAST Len "Pan Shou"
LAST Den "Accomplished Chinese calligrapher and poet who is known for his distinctive calligraphic style. Born in China, he came to Singapore in 1929 a
LAST P31 Q5
LAST P3988 "a3Jbw83C0gU"

CREATE
LAST Len "Colin Schooling"
LAST Den "Businessman and softball player who represented Singapore in the 1968 Asian Men's Softball Championship in Manila, Philippines and the Third
LAST P31 Q5
LAST P3988 "51ls_TlvM4k"
    
```

# QUICKSTATEMENTS – ERROR ALERTS

- Unable to upload description into Wikidata item using QuickStatements if the character limit exceeds 250 characters.

1 **error!** Pan Shou [Q93217137] **ADD** Description en:Accomplished Chinese calligrapher and poet who is known for his distinctive calligraphic style. Born in China, he came to Singapore in 1929 at the age of 19. It was only after retirement from a successful career in education that his second career in calligraphy and poetry flourished. He had been honorary adviser to the Chinese Calligraphy Society of Singapore since 1977.

Description must be no more than 250 characters long



# WIKIDATA QUERY SERVICE

```
Wikidata Query Service Examples Help More tools
1 SELECT ?person ?personLabel
2 WHERE
3 {
4   ?person wdt:P31 wd:Q5,
5   ?person wdt:P19 wd:Q334.
6 SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
7 }
```

Instance of

Human

Singapore

Place of Birth

- Query data by typing in RDF triple statements (**Subject - Predicate - Object**)
- E.g. To obtain a list of "People born in Singapore found in Wikidata"
- Results are displayed in the form of columns of data
- Data can then be used to match against existing NLB named entities.

person	personLabel
<a href="#">Q4726467</a>	Paulyn Sun
<a href="#">Q4726552</a>	Aliff Shafaein
<a href="#">Q4726554</a>	Aliff Aziz
<a href="#">Q4734724</a>	Aloysius Yap

# WIKIDATA QUERY SERVICE: DUMB DOWN VERSION

## WIKIDATA QUERY BUILDER

### About this tool

The Wikidata Query Builder provides a visual interface for building a simple Wikidata query. It is ideal for users with little or no experience in [SPARQL](#), the powerful query language. The Query Builder doesn't offer SPARQL's full functionality, but you can always open your query in the Query Service, where you can view, edit or expand it via the link above the results.

[Feedback is welcome here.](#)

### Query

Find all items...

With Without      Include related values in the search (recommended)

and or

With Without      Include related values in the search (recommended)

**Wikidata Query Builder:** User friendly interface for building a simple Wikidata query even without SPARQL knowledge

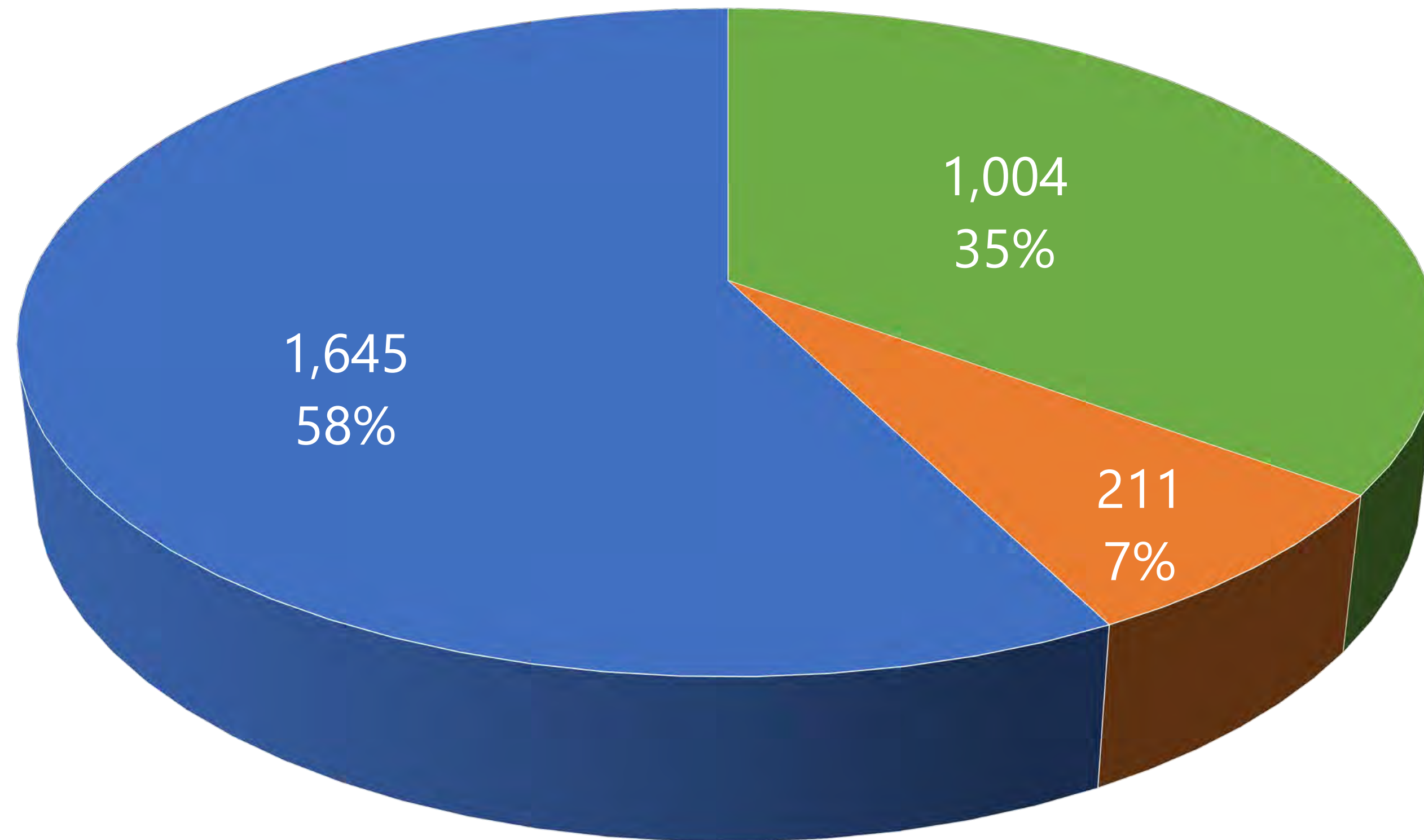
# WORKING WITH DATA FROM WIKIDATA QUERY SERVICE

- Query Service tested on querying Wikidata using the following two properties:
  - a) Place of birth (P19):** To retrieve the list of Wikidata people born in **Singapore**
  - b) Country of citizenship (P27):** To retrieve the list of Wikidata people who are **Singaporean**
- Based on the no. of Wikidata items with value "Singapore" in properties **country of citizenship** or **place of birth**, we need to find out:
  - 1) How many Wikidata items does not contain NLB IDs?
  - 2) How many of these items exist in NLB named entities and are "notable"?

# LIMITATIONS OF WIKIDATA QUERY SERVICE

- Unable to retrieve people whose place of birth or country of citizenship are outside Singapore, but nonetheless are significant to Singapore (e.g. Raffles)
- Longer time and effort matching using various Excel lists
- Much of the verification still require manual effort

# PERCENTAGE OF SUCCESSFUL MATCHES USING OPENREFINE VS WIKIDATA QUERY SERVICE



\*Based on total quantities of 2,860 NLB ID of People names submitted in FY20

- OpenRefine (Auto-matching)
- Wikidata Query Service (Using Excel to auto-match based on SPARQL results)
- Manual matching

# LIMITATIONS OF AUTO-MATCHING (1): DIFFERENT WIKIDATA ENTITIES WITH SAME LABEL

- Two different entities are assigned with the same label in Wikidata, but belong to **different entity types**
- E.g. Two “Bukit Timah”: <http://www.wikidata.org/entity/Q22970788> and <http://www.wikidata.org/entity/Q720209>
- Both descriptions refer to different types of entity:
  - <http://www.wikidata.org/entity/Q22970788> refers to the planning area
  - <http://www.wikidata.org/entity/Q720209> refers to the hill
- Solution: Manually assign correct NLB identifiers to the respective Wikidata entities – **Bukit Timah** (planning area) and **Bukit Timah** (hill).

**Bukit Timah** (Q22970788)

planning area and residential estate in Central Region, Singapore  
Bt Timah

▼ In more languages  
Configure

Language	Label	Description
English	Bukit Timah	planning area and residential estate in Central Region, Singapore
Chinese	武吉知马	No description defined
Malay	Bukit Timah	No description defined
Tamil	No label defined	No description defined

All entered languages

Statements

instance of planning area of Singapore

▼ 0 references

**Bukit Timah** (Q720209)

hill in Singapore  
Bukit Timah Hill

▼ In more languages  
Configure

Language	Label	Description
English	Bukit Timah	hill in Singapore
Chinese	武吉知马	No description defined
Malay	Bukit Timah	No description defined
Tamil	புக்கித் துமா	No description defined

All entered languages

Statements

instance of hill

# LIMITATIONS OF AUTO-MATCHING (2): ONE WIKIDATA ENTITY MATCHES WITH TWO NLB ENTITIES

- This occurs when the Wikidata entity type belongs to both organisations and places. E.g. hospital and hotel.
- Example: Goodwood Park Hotel exists both as an organisation and as a building in NLB named entities. In Wikidata, it is considered as a single entity, as the entity type “**hotel**” is a subclass of both **organisation** and **building**. This results in the Wikidata entity <http://www.wikidata.org/entity/Q5583726> matching with both NLB named entities **Goodwood Park Hotel (Place)** (i.e. the building) and **Goodwood Park Hotel** (i.e. the organisation).
- Solution: Prioritise the NLB identifier to assign to the corresponding Wikidata entity. In this example, identifier of **Goodwood Park Hotel (Place)** is given higher priority to be assigned to the Wikidata entity as the hotel building is well known as a national monument, rather than as an organisation.

Goodwood Park Hotel (Q5583726)

hotel in Singapore

[In more languages](#)  
[Configure](#)

Language	Label	Description
English	Goodwood Park Hotel	hotel in Singapore
Chinese	良木園酒店	No description defined
Malay	No label defined	No description defined
Tamil	No label defined	No description defined

[All entered languages](#)

Statements

instance of [hotel](#)

hotel (Q27686)

business enterprise that provides lodging in a single building paid on a short-term basis  
hotels

Statements

subclass of [horeca](#)  
0 references

subclass of [business](#)  
0 references

subclass of [building](#)

business (Q4830453)

organization involved in commercial, industrial, or professional activity

Statements

subclass of [organization](#)  
0 references

# LIMITATIONS OF AUTO-MATCHING (3): OTHER CHALLENGES

- Better luck with names in direct form than inverted names.
- Higher inaccuracy with generic names (mostly are non-authorized names).  
E.g. Tortoise Island (in Singapore, more popularly known as Kusu Island) is associated with an island in Australia
- Next best match if the name is not available in Wikidata.  
E.g. **Battery Road** (Singapore) matched with **6 Battery Road** (Singapore)
- VIAF ID: Only need to use the ID for matching, not the whole URI  
E.g. 11044086 **NOT** <http://viaf.org/viaf/11044086>
- Need to continuously experiment and refine the parameters to improve the matching. E.g. add in parameters such as country (for place names), place of birth and country of citizenship (for people's names); restricting the search to a particular type (e.g. human)



# WHAT HAS BEEN ACHIEVED ?

- Published NLB's Entity Page (Linked Data) in Wikidata
- Learnt and applied new Semantic Web tools (e.g. SPARQL; OpenRefine; Quick Statements)
- Engaged with online communities of practice, like Wikidata Community, for knowledge and support
- Linked more than 4,800 names of SG people, organisations and places in Wikidata since 2020

# LOST? HELP IS ON THE WAY...

- ❑ OpenRefine help page: <https://www.wikidata.org/wiki/Wikidata:Tools/OpenRefine>
- ❑ QuickStatement help page: <https://www.wikidata.org/wiki/Help:QuickStatements>
- ❑ Online tools available to work at Wikidata: <https://www.wikidata.org/wiki/Wikidata:Tools>
- ❑ Wikidata terminology: <https://www.wikidata.org/wiki/Wikidata:Glossary>
- ❑ Ask the community: [https://www.wikidata.org/wiki/Wikidata:Project\\_chat](https://www.wikidata.org/wiki/Wikidata:Project_chat)
- ❑ Important contacts:
  - **Pinto** (a.k.a. [Antonin Delpeuch](#)): Wikidata reconciliation interface and EditGroups tool (incl. QuickStatements). Talk to him via [https://www.wikidata.org/wiki/User\\_talk:Pinto](https://www.wikidata.org/wiki/User_talk:Pinto)
  - **Pigsonthewing** (a.k.a [Andy Mabbett](#)): Wikimedian in Residence, offer advice and assistance with Wikipedia, Wikidata, OpenStreetMap etc. Talk to him via [https://www.wikidata.org/wiki/User\\_talk:Pigsonthewing](https://www.wikidata.org/wiki/User_talk:Pigsonthewing)

# USING OPENREFINE FOR RECONCILIATION AND OTHER FUNCTIONS

- **OpenRefine Reconciliation Process** <https://youtu.be/LpIJf9O6JNw>  
*15min demo on the use of the reconciliation function with Wikidata in OpenRefine*
- **Reconciliation in OpenRefine (2 parts video by Owen Stephens)**  
<https://youtu.be/q8ffvdeyuNQ>  
<https://youtu.be/0tQPmfb6IFk>  
*More detail explanation on the reconciliation function in OpenRefine*
- **Community forum for OpenRefine**  
<https://forum.openrefine.org/>  
*Community that helps to provide solutions on issues encountered when working on OpenRefine*

# Thank you



*For enquiries, email us at  
[NLB\\_SNA\\_enquiry@nlb.gov.sg](mailto:NLB_SNA_enquiry@nlb.gov.sg)*