

TM-Link

Data Dictionary

Document Version

Version	Date	Author	Comment
0.1.1	2019-04-24	Min Kim	Initial draft for the public release
0.1.2	2019-10-14	Min Kim	Minor edits to the text.

Please contact IPDataPlatform@ipaustralia.gov.au for queries and suggestions.

Table of Contents

1	Dataset	3
1.1	Description	3
1.2	Data Coverage	3
1.3	Known Issues	3
1.4	Schema	4
1.5	Variables	4
2	Data Sources	7
2.1	Australia	7
2.2	Canada	7
2.3	European Intellectual Property Office	8
2.4	New Zealand	9
2.5	United States of America	9
2.6	United Kingdom	10

1 Dataset

1.1 Description

TM-Link is an international dataset IP Australia and Swinburne University have developed in collaboration. The dataset provides information from various jurisdictions, modelled under a common schema for greater accessibility to researchers and analysts. TM-Link also links¹ together similar trade marks from different countries based on common information, such as similar trade mark phrases and applicant names. These links identify families of international trade marks, which provide a new and unique insight into international branding trends and export behaviours.

IP Australia and Swinburne University are looking to continually develop TM-Link to become a core part of the global IP data landscape. If you have any suggestions or requests to model any additional data points, or improve the current accuracy of the data please let us know via email to ipdataplatfom@ipaustalia.gov.au.

TM-Link is available for bulk download via data.gov.au. The dataset can also be accessed and analysed on the IP Data Platform (www.ipdataplatfom.com), an analytical environment that enables researchers and policy makers to investigate and interrogate data to make decisions and to create economic and policy insights.

1.2 Data Coverage

Application country/Filing Office	Min filing date	Max filing date	Count
AU	1906-07-02	2018-01-23	1,722,244
CA	1872-12-14	2016-02-12	1,442,348
EM	1996-04-01	2016-03-03	1,300,992
GB	1876-01-01	2018-01-28	1,280,258
NZ	1874-10-03	2015-10-19	518,077
US	1823-03-06	2018-03-26	9,124,765

1.3 Known Issues

Issues

¹ Petrie S, Kollmann T, Codoreanu A, Thomson R, Webster E (2019) *International Trademarking and Regional Export Performance*, available at SSRN: <http://ssrn.com/abstract=3445244>

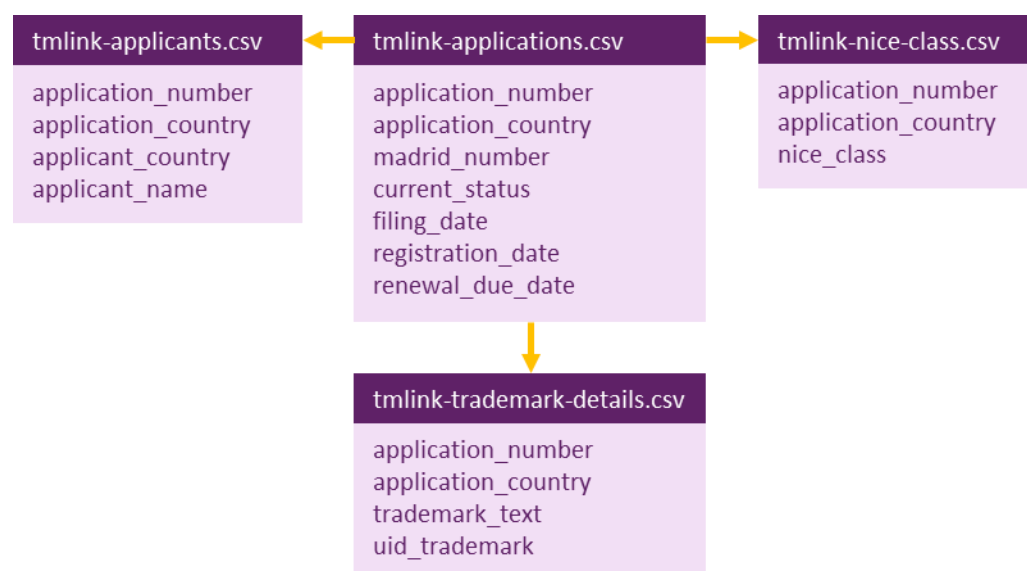
Issues
The data available for NZ did not contain the country information for the applicants.
The data available for GB and CA did not contain Madrid numbers.
Each IP offices use different set of definitions for the application status. These status descriptions have not been standardised yet.

Found an issue with TM-Link? Please let us know via email to

ipdataplatfrom@ipaustalia.gov.au.

1.4 Schema

The TM-Link dataset consists of four tables: tmlink-applications, tmlink-applicants, tmlink-trademark-details, and tmlink-nice-class. The tables can be joined using **application_number** and **application_country** as keys.



1.5 Variables

1.5.1 Applications

Column	Description/Comment
application_number	Key Unique number for the trade mark application as used by the application office.
application_country	Key The country/office where the trade mark is being filed.

Column	Description/Comment
madrid_number	Madrid number for the trade mark application.
current_status	These have not been standardised yet.
filing_date	When the application is filed.
registration_date	When the trade mark becomes registered.
renewal_due_date	Renewal due date or expiry date.

1.5.2 Applicants

Column	Description/Comment
application_number	Key Unique number for the trade mark application as used by the application office.
application_country	Key The country/office where the trade mark is being filed.
applicant_country	The country of the owner/applicant of the trade mark.
applicant_name	The name of the owner (person or entity) of the trade mark.

1.5.3 Trade mark Details

Column	Description/Comment
application_number	Key Unique number for the trade mark application as used by the application office.
application_country	Key The country/office where the trade mark is being filed.
trademark_text	The text component of the trade mark.
uid_trade mark	A unique identification number that matches similar trade marks. Generated by Swinburne University through Machine Learning techniques.

1.5.4 Nice Class

Column	Description/Comment
application_number	Key Unique number for the trade mark application as used by the application office.
application_country	Key The country/office where the trade mark is being filed.

Column	Description/Comment
nice_class	Nice classes 1~45.

2 Data Sources

2.1 Australia

Data sourced from data.gov.au IP Government Open Data (IPGOD) 2018 dataset. [Link](#)

TMLink Column	Source Table	Source Column
application_number	all tables	tm_number
madrid_number	IPGOD203	ir_number
trademark_text	IPGOD222	trademark_text
nice_class	IPGOD204	class_code
current_status	IPGOD203	status_code_desc
filing_date	IPGOD203	lodgement_date
registration_date	IPGOD203	sealing_date
renewal_due_date	IPGOD203	renewal_due_date
application_country	None	<i>Set as AU</i>
applicant_country	IPGOD202	country <i>missing data as AU</i>
applicant_name	IPGOD202	cleanname

2.2 Canada

The CA dataset has been provided by CIPO to Swinburne University (and IP Australia) through an FTP server.

TMLink Column	Source Table	Source Column
application_number	all tables	Trade mark ApplicationNumber ST13A pplicationNumber
madrid_number	None	None
trademark_text	cipo_Trade mark.csv	tm_text
nice_class	cipo_GoodsServicesBag_GoodsS ervicesClassification.csv	ClassNumber
current_status	cipo_Trade mark.csv	MarkCurrentStatusCode

TMLink Column	Source Table	Source Column
filing_date	cipo_Trade mark.csv	ApplicationDate
registration_date	cipo_Trade mark.csv	RegistrationDate
renewal_due_date	cipo_Trade mark.csv	ExpiryDate
application_country	None	<i>Set as CA</i>
applicant_country	cipo_Trade mark.csv	Applicant Con Pos Pos CountryCode
applicant_name	cipo_Trade mark.csv	Applicant Contact Name EntityName_en <i>or</i> Applicant Contact Name EntityName_fr

2.3 European Intellectual Property Office

The EM dataset has been provided by EUIPO to Swinburne University (and IP Australia) through an FTP server.

TMLink Column	Source Table	Source Column
application_number	euipo_TradeMark.csv	TradeMark ApplicationNumber
madrid_number	None	None
trademark_text	euipo_WordMarkSpecification.csv	MarkVerbalElementText
nice_class	euipo_GoodsServicesDetails ClassDescription.csv	ClassNumber
current_status	euipo_TradeMark.csv	MarkCurrentStatusCode
filing_date	euipo_TradeMark.csv	TradeMark ApplicationDate
registration_date	euipo_TradeMark.csv	TradeMark RegistrationDate
renewal_due_date	euipo_TradeMark.csv	ExpiryDate
application_country	None	<i>Set as EM</i>
applicant_country	euipo_Applct_Applicant.csv	AddressCountryCode
applicant_name	euipo_Applct_Applicant.csv	OrganizationName

2.4 New Zealand

Data obtained by Swinburne University.

TMLink Column	Source Table	Source Column
application_number	iponz_appln.csv	tm_no
madrid_number	iponz_appln.csv	intl_reg_no
trademark_text	iponz_appln.csv	tm_text
nice_class	iponz_appln.csv	nice_class
current_status	None	None
filing_date	iponz_appln.csv	submission_date
registration_date	iponz_appln.csv	filing_date
renewal_due_date	None	None
application_country	None	<i>Set as NZ</i>
applicant_country	None	None
applicant_name	iponz_appln.csv	all_owners

2.5 United States of America

Data sourced from uspto.gov Trademark Case Files Dataset. [Link](#)

TMLink Column	Source Table	Source Column
application_number	all tables	serial_no
madrid_number	uspto_case_file.csv	ir_registration_no
trademark_text	uspto_case_file.csv	mark_id_char
nice_class	uspto_intl_class.csv	intl_class_id
current_status	uspto_case_file.csv	cfh_status_cd
filing_date	uspto_case_file.csv	filing_dt
registration_date	uspto_case_file.csv	registration_dt
renewal_due_date	uspto_case_file.csv	renewal_dt
application_country	None	<i>Set as US</i>

TMLink Column	Source Table	Source Column
applicant_country	uspto_owner.csv	owner_ctry_code <i>missing data as US</i>
applicant_name	uspto_owner.csv	owner_name_clean

2.6 United Kingdom

Data sourced from gov.uk trade mark data release. [Link](#)

TMLink Column	Source Table	Source Column
application_number	OpenDataDomestic2018.txt, OpenDataIR2018.txt	Trade mark
madrid_number	None	None
trademark_text	OpenDataDomestic2018.txt, OpenDataIR2018.txt	Mark Text, Mark
nice_class	OpenDataDomestic2018.txt, OpenDataIR2018.txt	Class
current_status	OpenDataDomestic2018.txt, OpenDataIR2018.txt	Status
filing_date	OpenDataDomestic2018.txt, OpenDataIR2018.txt	Filed, Date of Designation
registration_date	OpenDataDomestic2018.txt, OpenDataIR2018.txt	Registered, Protected
renewal_due_date	OpenDataDomestic2018.txt, OpenDataIR2018.txt	Renewal due date, Expired
application_country	None	<i>Set as GB</i>
applicant_country	OpenDataDomestic2018.txt, OpenDataIR2018.txt	Country <i>missing data as GB</i>
applicant_name	OpenDataDomestic2018.txt, OpenDataIR2018.txt	Name