Australian Government

IP Australia

Overview of the Intellectual Property Government Open Data

IP Australia Economic Research Paper 02

Report commissioned by IP Australia and authored by:

Bradley Man

Suggested reference:

Man, Bradley. 2014. Overview of the Intellectual Property Government Open Data. IP Australia Economic Research Paper 02.

http://www.ipaustralia.gov.au/about-us/reports/economics_research_paper02

Acknowledgments:

Written with guidance from Benjamin Mitra-Kahn (Chief Economist, IP Australia), Adam Bialowas (Economist, IP Australia), Melanie O'Byrne (Senior Analyst, IP Australia) and Matthew Johnson (Economist, IP Australia)

ISBN 978-1925245-02-8 (Online)
ISSN 2203-661X (Online)

Contents

Glossary

ABN	Australian Business Number
ABR	Australian Business Registry
ACN	Australian Company Number
ANZSIC	Australian and New Zealand Standard Industrial Classification
ASIC	Australian Securities and Investment Commission
ATO	Australian Taxation Office
EPO	European Patent Office
IP	intellectual property
IPC	International Patent Classification
IPGOD	Intellectual Property Government Open Data
IPRIA	Intellectual Property Research Institute of Australia
MIIPA	Melbourne Institute–IP Australia
NBER	National Bureau of Economic Research
NCL	Nice Classification
OECD	Organisation for Economic Co-operation and Development
PATSTAT	EPO Worldwide Patent Statistical Database
PBR	plant breeder's rights
REGPAT	OECD 'regionalised' patent database
SMEs	small- and medium-sized enterprises
UKIPO	United Kingdom Intellectual Property Office
USPTO	United States Patent & Trademark Office
WIPO	World Intellectual Property Organization

Executive summary

This report serves as an overview of the Intellectual Property Government Open Data (IPGOD) produced by IP Australia.

The IPGOD provides a comprehensive, flexible and reliable dataset that is both of value to IP researchers and professionals, and will be used to support the development and implementation of policies associated with intellectual property in Australia.

The IPGOD goes back to the early 1900s and includes all inventions, brands, designs and plant breeder's rights filed in Australia up to 31 December 2013.

To illustrate the potential of the IPGOD, this report provides a number of examples which serve to illustrate the breadth of the data held within it. Particular attention is given to examples which make use of a series of unique identifiers which can be used to match IP administrative data with firm-level business characteristics for Australian entities.

The ability to match IP administrative data with firms and firm-level characteristics is an important feature of the IPGOD. The matched data enables far more detailed analysis of the potential impacts of government policies on specific industries or entities.

It is intended for the complete IGPOD to be updated on an annual basis, to coincide with the release of the annual Australian Intellectual Property Report in April.

The IPGOD is hosted on data.gov.au and can be downloaded or accessed via third-party software and websites through data.gov.au.

Introduction

This report is a companion to the Intellectual Property Government Open Data (IPGOD) released by IP Australia, which is available on data.gov.au. It is intended to provide an introductory overview to the IPGOD, and an insight into the wealth of information available within the IPGOD. It is hoped that the IPGOD will be become an invaluable resource for policy makers and researchers working in intellectual property.

The IPGOD includes over 100 years of data on all IP rights administered by IP Australia, comprising patents, trade marks, designs and plant breeder's rights. The data is highly detailed, including information on each aspect of the application process from application through to granting of IP rights. The IP data has been further augmented by a unique set of identifiers which allow IP rights to be linked to individual firms and firm-level business information.

The IPGOD project is part of a commitment to data across government more accessible to support evidence-based decision making. Under this paradigm, the IPGOD allows information to be freely shared and analysed, thus increasing the transparency of government activities leading to improved policy outcomes.

While much of the data included in the IPGOD is already publicly available, until now the data has not been available free of charge and in a linked form from a single source.

Patent offices around the world provide IP data in the form of publicly available registries such as Espacenet or the European Patent Register, hosted by the European Patent Office (EPO). In addition to these registries, patent data is also published in bulk format such as the EPO Worldwide Patent Statistical Database (PATSTAT) or the Organisation for Economic Co-operation and Development's Regionalised Patent Database (REGPAT). A paid subscription is typically required for access to these databases.

Analysis of intellectual property data from these sources, such as assignee and inventor information, can provide answers to high-level questions but cannot answer more complex questions about the impact of policy changes on specific entities. This requires linking IP rights information with firm-level business data. To date, the construction of matched datasets have been complicated by the difficulties in associating IP information with firm-level business information of their owners.

Whilst the use of surveys has been used to link intellectual property information with firm-level business data, there are concerns with the accuracy of the data obtained using this method. An alternative approach to constructing a firm-level dataset is to match the information from intellectual property databases with firm-level business data. Examples of matched firm-level datasets include the National Bureau of Economic Research (NBER) Patent Data project,[1] and the dataset constructed by Thoma et al. which links the EPO and the United States Patent & Trademark Office (USPTO) patent data to the Bureau van Dijk's Amadeus database.[2]

These firm-level datasets, however, only contain information on a single intellectual property right so do not give a complete picture of firms' use of IP. The Oxford Firm-Level Intellectual Property Database is

[1] See Hall, B. H., Jaffe, A. B. & Trajtenberg, M., 2001, "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER Working Paper 8498, http://www.nber.org/papers/w8498.pdf

[2] See Thoma, G. & Torrisi, S., 2007, "Creating Powerful Indicators for Innovation Studies with Approximate Matching Algorithms. A test based on PATSTAT and Amadeus databases", CESPRI-Bocconi University WP 211, http://ideas.repec.org/p/cri/cespri/wp211.html; or Thoma, G., Torrisi, S., Gambardella, A., Guellec, D., Hall, B.H., & Harhoff, D., 2010, "Harmonizing and Combining Large Datasets – An Application to Firm-Level Patent and Accounting Data", NBER Working Paper No. 15851, http://www.nber.org/papers/w15851.pdf

the only example of a firm-level dataset that contains matched information on both patents and trade mark applications, filed with the United Kingdom Intellectual Property Office (UKIPO).[3]

The IP Government Open Data represents a dataset spanning all intellectual property rights administered and maintained by IP Australia. The IP Government Open Data has been further augmented by the construction of tables linking IP data of Australian entities with firm-level business characteristics of their owner. It is the first dataset to cover all IP rights in a country, provided as easily accessible .csv files through an open platform, and is to be updated annually.

[3] See Helmers, C., Rogers, M. & Schautschick, P., 2011, “Intellectual Property at the Firm-Level in the UK: The Oxford Firm-Level Intellectual Property Database”, University of Oxford: Department of Economics Discussion Paper Series Number 546, http://www.economics.ox.ac.uk/materials/working_papers/paper546.pdf

1. Data

The Intellectual Property Government Open Data (IPGOD) contains information on all intellectual property (IP) rights administered by IP Australia. The IPGOD includes patents, trade marks, designs and plant breeder's rights and includes detailed information on each aspect of the process from application through to granting of IP rights. The data within the IPGOD has been further augmented with a set of identifiers which allows IP rights to be linked to individual firms and firm-level characteristics.

Organisation of the Intellectual Property Government Open Data

A feature of the IPGOD is that it draws upon all bibliographic information collected by IP Australia as part of the application process for granting IP rights. For instance, the dataset includes botanical information regarding the genus and species of new varieties of crops, administrative information surrounding the granting or refusal of extensions to patents on pharmaceuticals, and legal information relating to hearings and oppositions associated with the granting of trade marks. A brief overview of the information in the IPGOD may be found in Table 1 below.

A complete list of all of the available fields may be found in the IPGOD data dictionary at data.gov.au/organization/ip-australia.

Table 1: Summary of data available in the IPGOD

	Coverage
Patents	Application details, Oppositions, Claims, Extensions of term, Licensees, Court decisions, Re-assignments, Conversions, Terminations
Trade marks	Application details, Oppositions, Claimed interest, Endorsements
Designs	Application details, Change of ownership
Plant breeder’s rights	Application details, Genetic resource centre
Aggregate	IP numbers, Applicant ID, Geo-codes in Australia, ABN, ACN

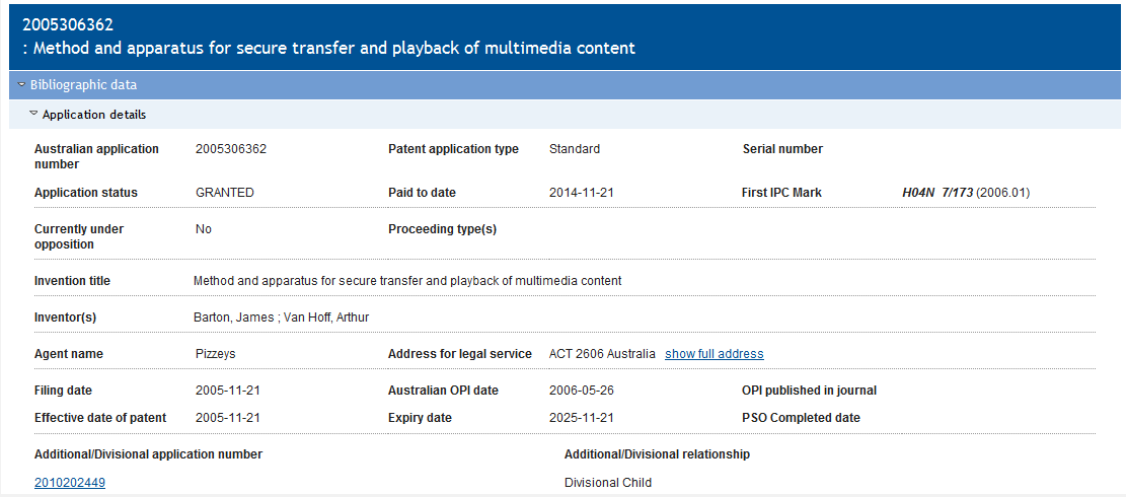
While much of the data included in the IPGOD is already publicly available, it is not usually in a form which makes it convenient for analysis. The IPGOD seeks to overcome this problem by presenting data held by IP Australia in a comma-separated-value (.csv) format which is easily read by data analysis programs such as Excel, STATA, R and SAS. Box 1 compares the accessibility of data obtained from IP Australia's AusPat database with the same information as presented in the IPGOD.

Box 1 - From filing to the IPGOD

When a patent application is filed, bibliographic information is published in the AusPat database and the Australian Official Journal of Patents. AusPat is an online database allowing access to bibliographic information and correspondence associated with Australian applications.

The bibliographic information shown in an AusPat entry, as in Figure 1, includes information such as the inventor(s), the technology associated with the application ('First IPC Mark') and key milestones associated with the application ('Filing date').

Figure 1: AusPat entry for a granted Australian patent application



2005306362
: Method and apparatus for secure transfer and playback of multimedia content

Bibliographic data

Application details

Australian application number	2005306362	Patent application type	Standard	Serial number	
Application status	GRANTED	Paid to date	2014-11-21	First IPC Mark	*H04N 7/173* (2006.01)
Currently under opposition	No	Proceeding type(s)			
Invention title	Method and apparatus for secure transfer and playback of multimedia content				
Inventor(s)	Barton, James ; Van Hoff, Arthur				
Agent name	Pizzeys	Address for legal service	ACT 2606 Australia show full address		
Filing date	2005-11-21	Australian OPI date	2006-05-26	OPI published in journal	
Effective date of patent	2005-11-21	Expiry date	2025-11-21	PSO Completed date	

Additional/Divisional application number	Additional/Divisional relationship
2010202449	Divisional Child

Bibliographic information has been extracted from AusPat and organised in bulk data format as part of the IPGOD. An excerpt of the patent data released as part of the IPGOD is shown in Figure 2. For example the application number is populated in the 'australian_appl_no' column, the first IPC mark is populated in the 'edited_ipc_mark_value' column and the IPC version is listed under 'ipc_class_version_date'.

Figure 2: Examples of the fields in the IPGOD

application_detail	australian_appl_no	ipc_mark_value	edited_ipc_mark	ipc_class_version_da	section_co	class_code	subclass_c	main_grou	subgroup_	full_mark_	ipc_mark_t	classificatic	classificatic	generating	classificatic	classification_source_co
1507271	2006246204	C07C 275/22	C07C275/22	1/01/2006	C	7	C	275	22	L	A	I	EP	B	H	
1507271	2006246204	C07D 239/42	C07D239/42	1/01/2006	C	7	D	239	42	L	A	I	EP	B	H	
1507271	2006246204	C07D 487/04	C07D487/04	1/01/2006	C	7	D	487	4	L	A	I	EP	B	H	
1507271	2006246204	C07C 313/06	C07C313/06	1/01/2006	C	7	C	313	6	L	A	I	EP	B	H	
1507271	2006246204	C07C 275/36	C07C275/36	1/01/2006	C	7	C	275	36	L	A	I	EP	B	H	
1507271	2006246204	C07F 9/22	C07F9/22	1/01/2006	C	7	F	9	22	L	A	I	EP	B	H	
1507271	2006246204	C07D 215/18	C07D215/18	1/01/2006	C	7	D	215	18	L	A	I	EP	B	H	
1507271	2006246204	C07D 235/14	C07D235/14	1/01/2006	C	7	D	235	14	L	A	I	EP	B	H	
1507271	2006246204	C07C 327/44	C07C327/44	1/01/2006	C	7	C	327	44	L	A	I	EP	B	H	
1507271	2006246204	C07C 275/18	C07C275/18	1/01/2006	C	7	C	275	18	L	A	I	EP	B	H	
1507271	2006246204	C07D 207/36	C07D207/36	1/01/2006	C	7	D	207	36	L	A	I	EP	B	H	

The IPGOD is organised into tables. These tables are split according to the intellectual property rights administered by IP Australia. All patents data start with a 1, trade marks 2, designs 3 and plant breeder's rights 4. Each set of IP rights has standard datasets associated with them (numbered 0–9) and specific datasets (numbered 20+) after the first digit. For example, table 101 is the general summary table for patent data, while 201 is the equivalent for trade marks. Table 2 below outlines the data structure of the IPGOD.

Table 2: Data structure in the IPGOD

	Title	1xx	2xx	3xx	4xx
		Patents	Trade marks	Designs	PBR
x01	Summary	101	201	301	401
x02	Firm information	102	202	302	402
x03	Applicant	103	203	303	403
x04	Technology	104	204	304	
x05	Inventor / further applicant	*	205		405
x06	Attorney	*	*	*	
x07	IP Australia process	107	207	307	
x08	Hearings / oppositions	108	208		
x09	Associated rights	109			

An asterix indicates that the data is not available in the 2014 IPGOD due to the Privacy Act. Changes have been made across the tables in order for the data to comply with the Privacy Act, so names associated with applications are not identified, except where the applicant has an Australian Business Number.

Firm-level data – MIIPA tables

Whilst analysis of intellectual property data such as assignee and inventor information can provide answers to high-level questions, the identification of specific industries or entities affected by policy changes can only be achieved with firm-level business data. Historically, the major barrier to creating such matched datasets has been the lack of common identifiers for applicants between IP data and firm-level business data, and the cost of linking the aforementioned data. The IPGOD overcomes this barrier by augmenting the IP data with a unique set of identifiers which allows IP rights to be matched to individual firms and firm-level characteristics.

The matched firm-level data within the IPGOD was constructed in collaboration with the Intellectual Property Research Institute of Australia and the Melbourne Institute of Applied Economic and Social Research. Consequently the tables which contain the matched data are referred to as Melbourne Institute–IP Australia (MIIPA) tables.

The MIIPA tables provide a range of information on applicants, and allow IP data to be matched to firm information and international applications. Each MIIPA table has one row per applicant for a given IP right, so a single patent with three applicants will have three rows in table 102. Each observation in the tables includes a unique name for the applicant, an applicant ID number which is consistent across all the MIIPA tables, a country of origin, a marker for whether the applicant is Australian, and the equivalent identification number in the PATSTAT database for patent applications. If an applicant is marked as Australian, the MIIPA tables further provide an entity marker and a marker to indicate if the entity is a large firm, small- and medium-sized enterprise or sole trader. Australian entities further include ABN or ACN data if found, and an ACN for the ultimate owner if the applicant is a subsidiary company of an Australian firm.

The IPGOD includes geo-spatial data such as the state/territory and postcode of the applicant, and a geo-code of the applicant address, as well as a marker indicating the quality of the geo-coding. Such data may be found in the Summary tables (x01).

The IPGOD data dictionary published on data.gov.au provides the full set of variables and data descriptions.

How the firm-level data, or MIIPA tables, was created

The information on patents, trade marks, designs and plant breeder's rights applications filed between 1990 and 2013 was obtained from IP Australia databases. The patent dataset only contains information on applications which are open for public inspection and have entered national phase as of 31 December 2013. A detailed discussion of the data processes and matching used in the construction of the MIIPA tables can be found in Julius and Rassenfosse (2014).[4]

With the raw dataset extracted from IP Australia databases, there are several stages of processing, which are briefly discussed below. A list of unique applicants, which are further assigned a unique identifier within the IPGOD, was generated using the OECD Harmonised Applicant Names database as a starting point.

ABNs and ACNs were matched to applicants identified as Australian entities using the ABR and ASIC databases. However there were some Australian entities that could not be matched to ABN or ACN. The ABN or ACN for these entities who had more than 15 applications across all IP rights were manually matched.

As Table 3 below shows, this matching has been conducted with a high success rate.

Identification of geographical details, such as the postcode, state/territory and geo-code were generated from each applicant's address. Note that there may be errors of a clerical nature associated with the postcodes, or as a result of postcode revisions by Australia Post. Where postcodes are not available, the state/territory from which the applicant originates was identified from the address. The address associated with each applicant was further geo-coded, to improve data matching in future releases.

Table 3: Summary of the MIIPA tables

	Number of records	Number of Australian records	Percentage of Australian entities matched with an ABN and/or ACN (%)
Patents	509,290	70,590 (14%)	94
Trade marks	1,032,560	716,278 (69%)	91
Designs	148,244	96,875 (65%)	88
PBR	7,651	4,299 (56%)	92

To assign the firm-size marker, the Bureau van Dijk company accounts database and Australian Business Database were matched to the ABN, and firms with more than 200 employees were given a marker. Subsidiary or sister companies of firms with more than 200 employees were also marked, to

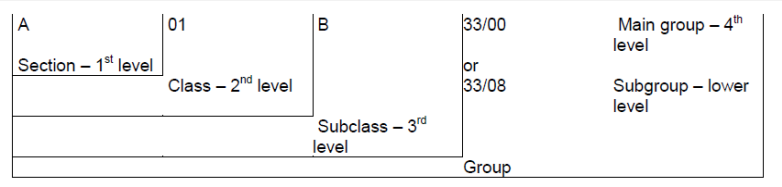
[4] See Julius, T. D. & Rassenfosse, G., 2014, "Harmonising and Matching IPR Holders at IP Australia.", Melbourne Institute Working Paper Series Working Paper No. 15/14. http:\\www.melbourneinstitute.com/downloads/working_paper_series/wp2014n15.pdf

give a set of large firms. The ABNs were then matched against the ABR to identify any sole traders, leaving three categories of matched entities: sole trader, large firms, and small- and medium -sized enterprises.

Box 2 – International Patent Classification (IPC)

To identify the technology to which the patent application pertains, each application is assigned an International Patent Classification mark. The IPC mark is a hierarchical system for the classification of patents according to the technology with which the patent is associated. There are eight high-level classifications known as IPC sections (sections A to H) and within each section there are further sub-divisions – known as classes, subclasses, main groups and subgroups – providing further information on the invention. A detailed explanation of the IPC hierarchy is shown in Figure 3.[A]

Figure 3: Hierarchy of the IPC Classification



For example, an application with an IPC mark of A63B 1/00 indicates the invention lies in human necessities (IPC section A), specifically in brushes (class 63, subclass B) relating to brush bodies and bristles moulded as a unit (group 1/00).

[A] See World Intellectual Property Organization, "Guide to the IPC (2014)", http://www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide_ipc.pdf

Box 3 –- Nice Classifications

The International Classification of Goods and Services for the Purposes of the Registration of Marks, otherwise known as the Nice Classification or NCL, is used to indicate the type of goods or services with which a trade mark application is associated.

Within the NCL there are 34 classes of goods and 11 classes of services in which a trade mark application may be filed. [A] Within each class there are further classifications providing additional details of the goods or services which the trade mark application pertains. [B]

Where a trade mark application contains multiple applicants and multiple NICE classifications, the MIIPA tables will contain an entry for each applicant for each NICE classification.

[A] See World Intellectual Property Organization, "International Classification of Goods and Services for the Purposes of the Registration of Marks (Nice Classification). Tenth Edition, Part II", http://tinyurl.com/pp6wqrt

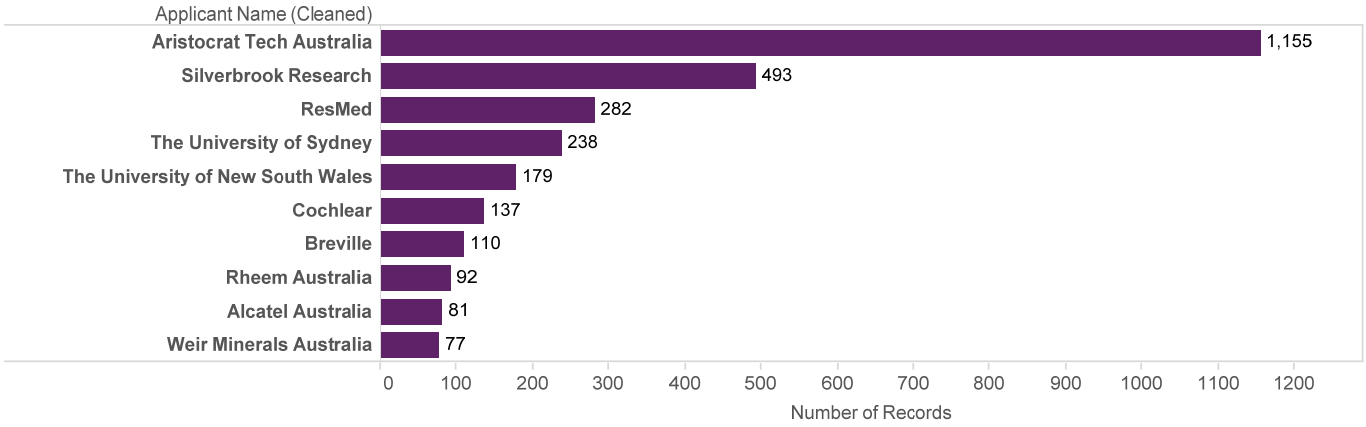
[B] See Nice Classification, http://tinyurl.com/qztox9w

2. Using the IP Government Open Data

Prior to the development of the MIIPA tables it was not possible to analyse the use of intellectual property at a firm level. Analysis of IP data, as published by the IPGOD, allows entities which are using intellectual property and their associated technologies to be identified. With the development of the IPGOD incorporating the MIIPA tables, it is now possible to analyse the use of intellectual property by different industrial sectors in Australia.

For example, Aristocrat Technologies Australia is the top filer of patents originating from New South Wales (Figure 4). These patents are primarily in the area of games and amusements (IPC A63F), and coin-actuated or like apparatus (G07F). This is consistent with their activity in the development of coin-operated gaming machines. Aristocrat Technologies Australia also filed trade mark applications in classes associated with games (NCL class 9) and electronic goods (class 28).

Figure 4: Top 10 patent applicants in NSW from 1990-2013



*Includes patents filed by Newsouth Innovations and Unisearch

In addition to understanding the use of intellectual property by the different industrial sectors in Australia, it is also possible to use the IPGOD to provide alternative technological boundaries around industries and growth sectors. The following sections provide a number of case studies highlighting the use of IPGOD, in particular the MIIPA tables.

Some of the examples in this section combine the IPGOD and Australian and New Zealand Standard Industrial Codes (ANZSIC), to show the location of IP activity by industry classification. This is one area where we are currently unable to publish firm-level data as the Australian Business Register (ABR) makes ANZSIC codes available as non-public data.[5] In discussions with the data managers at the ABR an agreement was reached which allows us to publish summary statistics, incorporating high level ANZSIC classifications (i.e. single digit) in this publication.

Which technologies are being patented by industries?

This section illustrates how the MIIPA tables with the Australian and New Zealand Standard Industrial Classification (ANZSIC) can be used to show the location of IP activity by industry classification.

Figure 5 plots the ANZSIC division codes against the IPC sections, highlighting the technologies in which the Australian entities in different industries are filing patents. Note that the Australian Taxation Office

[5] See https://abr.gov.au/For-Government-agencies/Accessing-ABR-data/ABR-data-explained/ABR-data/

could not identify the predominant activity of some Australian entities in the MIIPA tables. As a result those Australian entities have not been included in the figure below.[6]

Figure 5: Number of applicants by patent technology and ANZSIC divisions, 1990-2013

		IPC Section / IPC Section Code							
		Chemistry; Metallurgy	Electricity	Fixed Constructions	Human Necessities	Mechanical Engineering; Lighting; Heating; Weapons; Blasting	Performing Operations; Transporting	Physics	Textiles; Papers
ANZSIC Division	**Div. Code**	C	H	E	A	F	B	G	D
Manufacturing	C	677	636	2,321	2,737	1,215	2,450	1,283	81
Professional, Scientific and Technical Services	M	1,960	570	570	1,632	512	1,166	2,234	49
Wholesale Trade	F	108	302	445	810	272	566	275	21
Financial and Insurance Services	K	174	108	328	298	171	572	409	7
Education and Training	P	445	102	55	347	23	146	375	7
Rental, Hiring and Real Estate Services	L	107	57	284	197	178	355	146	10
Construction	E	47	61	520	82	148	151	75	1
Mining	B	294	6	168	8	75	134	31	1
Retail Trade	G	24	22	63	226	54	84	110	3
Health Care and Social Assistance	Q	110	7	5	209	4	32	83	
Other Services	S	34	11	59	72	41	133	70	
Arts and Recreation Services	R	127	12	14	119	22	37	65	
Information Media and Telecommunications	J	5	129	15	18	1	8	202	1
Administrative and Support Services	N	11	13	58	101	23	61	69	
Agriculture, Forestry and Fishing	A	35	3	25	118	15	107	16	
Electricity, Gas, Water and Waste Services	D	75	18	41	15	43	53	26	
Transport, Postal and Warehousing	I	12	4	24	26	10	118	32	1
Public Administration and Safety	O	28	8	38	40	24	21	63	1
Accommodation and Food Services	H	3	1	5	18		16	18	

Australian entities whose main activity is in manufacturing (division C) or delivering professional, scientific and technical services (division M) are the top filers of patents. Australian entities in the manufacturing division predominately file patents in technologies such as performing operations (IPC section B), human necessities (section A) and fixed construction (section E). The Australian entities filing patents in these technologies include Aristocrat Technologies Australia, AMCOR and ASSA ABLOY.

Australian entities in the professional, scientific and technical services division predominately file patents in the technologies such as human necessities (IPC section A), chemistry (section C) and physics (section G).

It is interesting to note that sole traders do not contribute to the filing of patents associated with enzymes and microorganisms. One possible explanation is that specialised equipment required for research in these fields is expensive and difficult to acquire.

Other Australian entities classified in the professional, scientific and technical services division filing patents in enzymes and microorganisms include Agriculture Victoria Services[7] and Grains Research &

[6] See Australian Taxation Office, 2013, “Business industry codes 2013”, https://www.ato.gov.au/Print-publications/Business-industry-codes-2013/

Development Corporation[8]. Both of these Australian entities are government-based organisations which have been established to plan and organise research in agricultural technologies. This is consistent with the significant number of PBR applications associated with these Australian entities.

Which goods are trade marks being sought for by industries?

In this example the ANZSIC codes have been matched against the trade mark classifications to highlight the type of goods and services for which firms are seeking trade mark protection. In Figure 6, Australian entities in the professional, scientific and technical services division dominate the filing of trade marks, which are localised in NCL classes 35, 42 and 9.

Figure 6: Number of applicants by Nice classifications and ANZSIC divisions, 1990-2013

		Nice Classification / Nice Class No.									
		Advertising, business management and administration	Apparatus and instruments	Education, training, entertainment, and sporting and cultural acti..	Scientific, industrial, design and legal services	Insurance, financial and real estate affairs	Paper, paper goods and office requisites	Building construction, repair and installation services	Foodstuff of plant origins	Pharmaceuticals and other preparations of medical purposes	Telecommunications
ANZSIC Division	Div. Code	35	9	41	42	36	16	37	30	5	38
Professional, Scientific and Technical Services	M	16,230	13,547	8,655	14,684	4,166	4,720	2,475	808	1,302	3,421
Financial and Insurance Services	K	7,841	4,356	4,911	4,024	16,296	4,221	1,890	1,440	1,029	2,303
Wholesale Trade	F	6,330	7,587	1,218	2,318	538	3,359	1,670	4,428	5,586	583
Manufacturing	C	3,660	7,260	1,574	2,523	373	3,645	2,434	6,742	4,415	555
Information Media and Telecommunications	J	4,184	4,884	5,376	2,486	783	4,326	384	114	46	5,803
Retail Trade	G	7,766	3,318	1,409	2,159	838	1,937	1,168	1,626	1,960	758
Other Services	S	3,129	1,436	4,699	2,110	2,160	2,141	1,296	399	265	392
Rental, Hiring and Real Estate Services	L	3,409	1,373	1,609	1,406	3,515	1,144	1,622	646	394	481
Administrative and Support Services	N	4,079	1,156	3,058	1,208	906	1,152	936	190	306	404
Arts and Recreation Services	R	1,395	1,398	5,745	545	385	2,258	98	137	74	331
Construction	E	1,565	1,254	577	1,618	1,378	469	4,972	75	66	275
Education and Training	P	1,316	1,139	5,498	1,049	282	1,681	71	42	110	195
Health Care and Social Assistance	Q	1,397	784	2,247	1,223	1,056	841	193	81	581	89
Public Administration and Safety	O	1,468	842	1,585	994	450	916	412	57	42	281
Accommodation and Food Services	H	920	222	1,202	995	267	458	130	1,367	60	43
Transport, Postal and Warehousing	I	915	445	401	496	478	460	329	122	31	299
Electricity, Gas, Water and Waste Services	D	768	452	430	747	367	250	674	5	15	179
Agriculture, Forestry and Fishing	A	485	116	190	303	143	116	57	306	248	9
Mining	B	283	156	112	462	63	52	408	21	28	54

Firms who invent but do not manufacture in Australia

Figure 5 shows that Australian entities associated with wholesale trade (division F) are filing significant volumes of patents in a diverse range of technologies. In fact, it is the third-highest area of focus for applicants in the MIIPA tables. Examining the patents filed by Australian entities associated with wholesale trade, there is a wide spread of technology which includes containers (B65D), fixed construction (E06B), kitchen equipment (A47J) and pharmaceuticals (A61K).

---

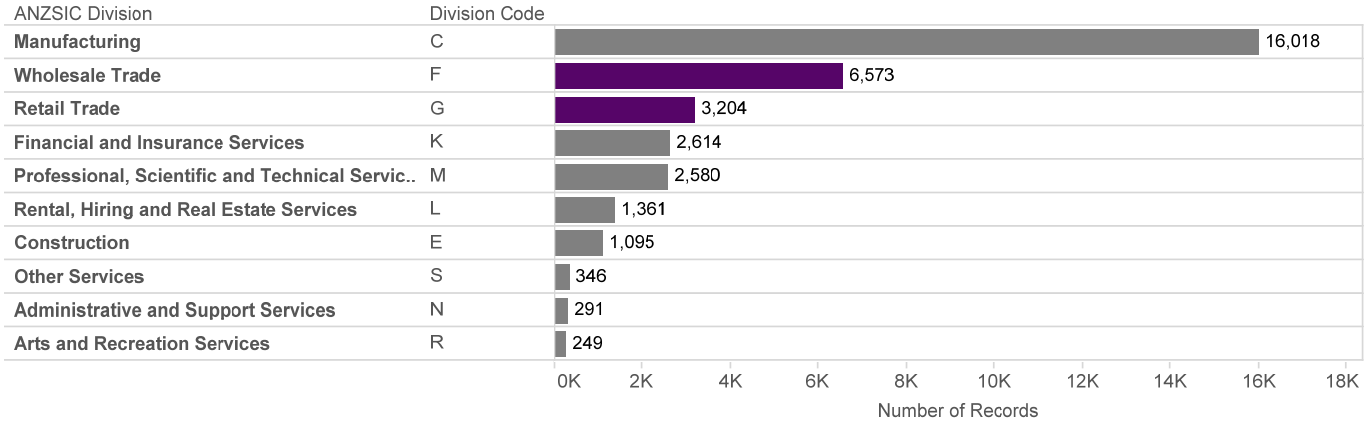
[7] See http://www.agvic.com.au/

[8] See http://www.grdc.com.au/About-Us

Australian entities are classified in the wholesale trade division if their predominant activity involves the purchasing and on-selling of goods. It is also important to note that Australian entities in the wholesale trade division are distinguished from Australian entities in the manufacturing division. They do not own the material inputs, and the production is outsourced or licensed to a third party.[9] Traditionally one might have thought of them as retailers importing products from abroad, but if these Australian entities are filing patents or designs, it suggests that they are innovating in Australia but outsourcing the production.

Beyond patents, one can look at design rights – often-used by companies like Apple – which represents a monopoly on the visual features that form the design of an item. Figure 7 shows that Australian entities in the wholesale division file more design applications compared to Australian entities in the retail trade division, who predominately sell goods from premises with goods on display.

Figure 7: Design applications by ANZSIC division, 1990-2013



Analysis of the design filings for the ANZSIC divisions suggests that Australian entities in wholesale trade (division F) are placed high on the global value-adding supply chain. Rather than being involved in the production of goods, these Australian entities prefer to develop new products (protected by patents and designs), outsource or license the production to a third party and then sell the product to consumers, which should imply a higher profit margin. This is consistent with the definition of wholesale trade where they do not own the material inputs but own the final outcome.

Technology boundaries around industries and growth sectors

The government has identified a number of sectors which reside at the centre of the government's industrial strategy.[10] Within these sectors a number of industries have been identified as having high growth potential such as the pharmaceutical industry.[11] Traditionally, one would use the ANZSIC sub-divisions to define the boundaries of the pharmaceutical sector. Bearing in mind that the ANZSIC is based on activity, one can instead use patent data to draw technological boundaries.

The patent portfolio for Schering-Plough, to select one firm in the MIIPA tables, includes patents associated with preparations for medical, dental or toilet purposes (IPC A61K). There are further

[9] See Australian Bureau of Statistics, "Australian and New Zealand Standard Industrial Classification (ANZSIC) 2006", ABS Cat. No. 1292.0, http://www.abs.gov.au/ausstats/abs@.nsf/mf/1292.0

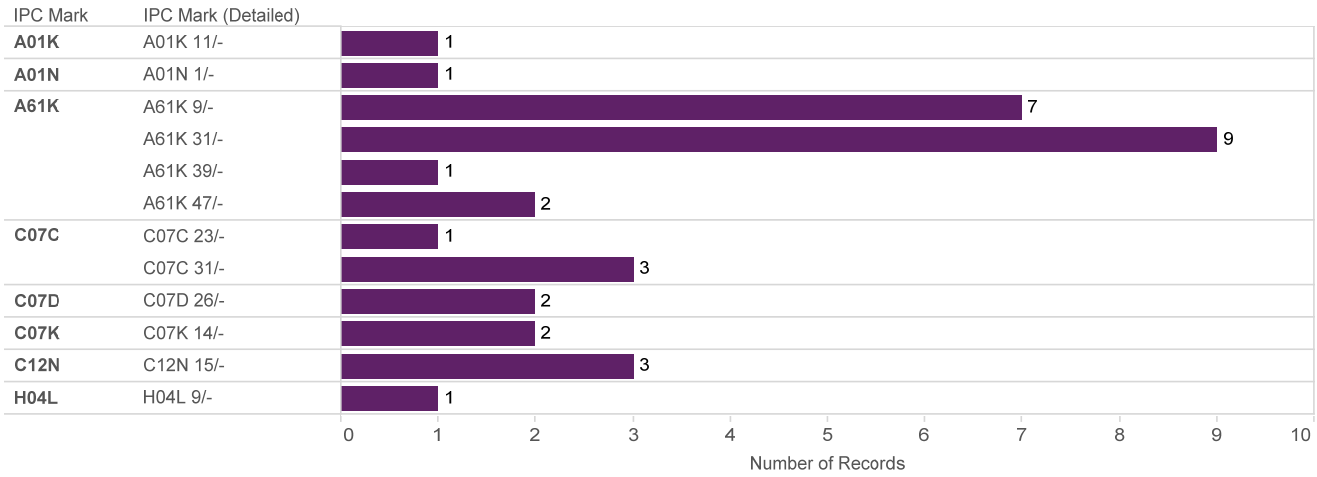
[10] See Liberal Party of Australia, 2013, "Our Plan, Real Solutions for all Australians, The direction, values and policy priorities of the next Coalition Government".

[11] See Department of Industry, "Review of South Australian and Victorian Economies", Appendix 5, http://www.industry.gov.au/AboutUs/CorporatePublications/ReviewofSouthAustralianandVictorianEconomies/Pages/default.aspx

classifications which define the nature of the active ingredients in these preparations. Figure 8 shows that the patents filed by Schering-Plough are directed towards compositions which contain active ingredients such as antigens or antibodies (A61K 39/00), or organic molecule (A61K 31/00).

Patents directed towards pharmaceutical compositions account for 48 per cent of the patents filed by Schering-Plough while 41 per cent are directed towards the active ingredients *per se,* which have the IPC marks C07D, C07C and C07K). Based on the organisation of the IPC scheme associated with pharmaceutical compositions and their active ingredients, it can be reasoned that the pharmaceutical industries are supported by entities which focus on the development of the active ingredients.

Figure 8: Schering-Plough patent applications, 1990-2013



Therefore, entities supporting the pharmaceutical industries would be those that are, for example, innovating in the development of non-cyclic and carbocyclic compounds (C07C), heterocyclic compounds (C07D) or biomolecules such as antibodies (C07K).

Having identified the technologies that support pharmaceutical companies, the MIIPA tables can identify the entities which are filing patents, and therefore innovating, in these technologies. The entities in the three technology classifications identified above include private firms, universities and publicly funded research organisations. These entities are unlikely to be found in the ANZSIC sub-divisions associated with pharmaceutical industries.

3. Closing remarks

This paper gives an overview of the IPGOD containing bibliographic information on patents, trade marks, design rights and plant breeder's rights. The data is available in a bulk format that allows the information to be easily searched, analysed and manipulated by a variety of data analysis programs such as Excel, STATA, R and SAS.

IP Australia has constructed a matched firm-level dataset, known as the MIIPA tables, within the IPGOD. The MIIPA tables contain both intellectual property data and firm-level business data spanning all of the intellectual property rights in Australia. The matching process was developed in collaboration with the Intellectual Property Research Institute of Australia.

Being the first matched dataset spanning all IP rights in a country represents a significant contribution in supporting evidence-based decision making around innovation and intellectual property. The development of the matched dataset not only allows the analysis of the use of IP rights on a firm level but also which industries are using IP rights.

To highlight the utility of the matched dataset, this paper has presented a number of case studies highlighting the patents or trade marks applications being filed by the industries as defined by the ANZSIC divisions, and providing alternative boundaries to industries and growth sectors.

The IPGOD, and the MIIPA tables contained therein, are merely a first step in a long process for the development of a large, flexible and reliable dataset supporting evidence-based decision-making around innovation and intellectual property policies.