

4th ASEAN IP Register Regional Coordinators Meeting

Topic 8 ASEAN IP Register – Data Quality: Business Requirements, Challenges and Priorities

*WIPO IP Office Business Solutions
May 6, 2024*

Data Quality

Business Requirements

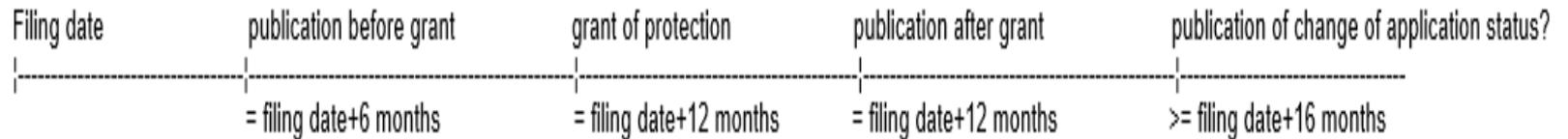
- Completeness of data coverage and minimum dataset
- Accuracy of data
- Timely data updates
- Standardization of data attributes as appropriate
- Structured data quality feedback to ease data correction and to improve data quality at source
- Authentic source of ASEAN IP collections for sharing with other IP offices, organizations and authorized third parties

Data Coverage Issues

- Data coverage issue observed on May 5, 2024

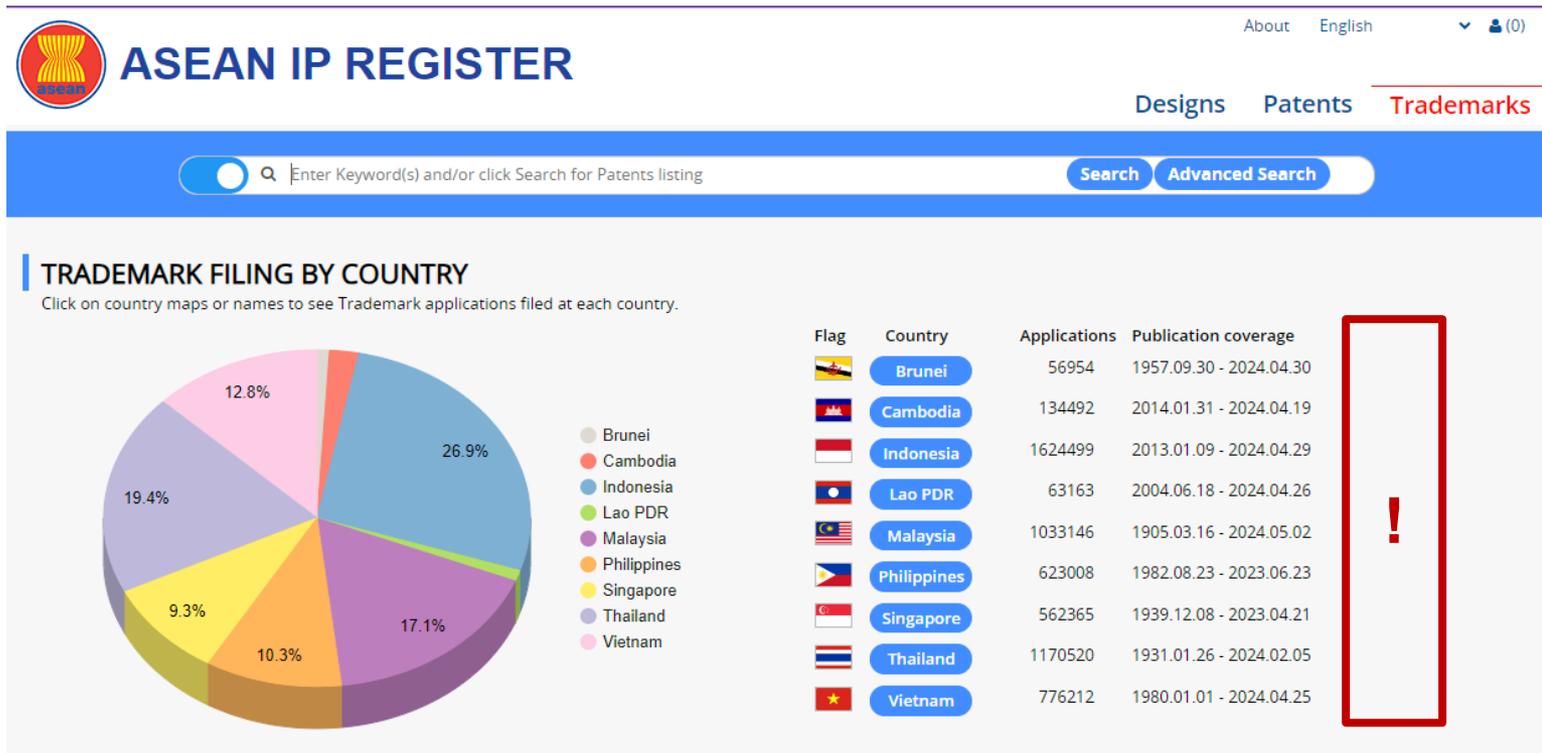
Trademark Collections	ASEAN IP Register	WIPO Global Brand	ASEAN TMView
Brunei	56'954	56'889	52'147
Cambodia	134'492	134'526	132'641

- No information about data availability
 - either published and registered applications or only registered applications, versus application filing counts
 - availability timeline



Data Accuracy Issues

No update indicators



Delay in data processing

Missing Application Data Attributes

- Standardization of application types (Madrid, national trademarks, PCT, PCT-NPE, national conventional patents, re-registration patents, utility models, Hague, national industrial designs)
- PCT national entry dates
- Standardization of file statuses
 - Application state (active, not active, terminated)
 - Key application events (published or granted)
- Publication date versus date for public inspection and gazette information as appropriate

Challenges

Data Coverage and Completeness

- Additional data coverage information
 - Implement WIPO Standard ST.37
(authority file = list of published applications/documents)

Example

```
...  
EP,2363052,A1,20110907,W, ABST-U, DESC-U, CLMS-U <CRLF>  
EP,2363053,A2,20110907,M, ABST-en, DESC-N, DESC-N<CRLF>  
EP,2540632,A1,20130102,P, ABST-N, DESC-N, CLMS-N<CRLF>  
EP,2540632,B1,20151202, , ABST-en ABST-fr ABST-de, DESC-en, CLMS-en<CRLF>
```

- Collect offices feedback on technical feasibility ?
(questionnaire)

Challenges

Data Coverage and Completeness

- Additional data coverage information
 - Implement WIPO Standard ST.37
(authority file = list of published applications/documents)

Example

```
...  
EP,2363052,A1,20110907,W, ABST-U, DESC-U, CLMS-U <CRLF>  
EP,2363053,A2,20110907,M, ABST-en, DESC-N, DESC-N<CRLF>  
EP,2540632,A1,20130102,P, ABST-N, DESC-N, CLMS-N<CRLF>  
EP,2540632,B1,20151202, , ABST-en ABST-fr ABST-de, DESC-en, CLMS-en<CRLF>
```

- Collect offices feedback on technical feasibility ?
(questionnaire)

Challenges

Additional Data Attributes

- New release of the extraction tool, on-premise version of WIPO Publish
- Data improvement for missing PCT-NPE data
 - abstracts, IPC, priority details, PCT application filing dates, drawings
 - English titles and/or abstracts
- Data improvement for national patents
 - English machine-translation of titles and abstracts when needed

Challenges

Standardization of Data Attributes

- Standardized Application File events based on WIPO Standards (ST.27 patents, ST.61 trademarks, ST.87 industrial designs), with the minimum key events Application filed, Application published, Application amended, IP right granted, IP right maintained, IP right ceased, IP right terminated
- Standardized Mark Category if needed (ST.96) Trademark, Mark, Service mark, Collective mark, Collective service mark, Certification mark
- Indicator of natural person in stakeholder details

Challenges

Frequency and Timely Data Updates

- Data sharing
 - before grant
 - as granted
 - In case of file status change

Challenges

Data Quality Feedback

- Collect offices feedback about the expected format and structure about data quality, describing
 - Unspecified file status and mark category
 - Missing NPE date
 - Unspecified indicator of natural person in stakeholder details
 - Missing IPC, title, abstract, priority details
- Possibility to retrieve data quality through API
 - (api specification to be shared in yaml for comments by IP Offices)

Prioritization of Improvements in 2024

- Addition of data coverage information and some useful data attributes
- Upgrade of the data exchange format to ST96 V7 or later
- Data extraction tools (e.g. on-premise version of WIPO Publish) to get additional application data attributes
- Cloud hosting to improve the scalability and timeliness of data sharing and processing
- Introduction of a few statistic charts
- Some recommendations of data quality manuals

Prioritization of Improvements in 2025

- Technical architecture to improve system performance
- User experience for data exploration
- Data extension to support IP portfolio service
- Machine translation service for non-English titles and abstracts
- Extension to support other types of IP rights
- Other recommendations of data quality manuals
- AI-based image search

Other Improvements

- Suggestions
- Discussion