

C. PCT 1434

December 10, 2014

Madam,
Sir,

Report on Characteristics of International Search Reports

1. This Circular is addressed to your Office in its capacity as an International Searching Authority under the Patent Cooperation Treaty (PCT). In that capacity, and as a member of the quality subgroup of the Meeting of International Authorities under the PCT, your Office is invited to comment on the annexed draft report on characteristics of international search reports with a view to identifying indicators of what should be the focus of further work by the International Authorities.

Background

2. At the fourth informal meeting of the quality subgroup held in Tel Aviv in February 2014, prior to the twenty-first session of the Meeting of International Authorities, the subgroup discussed a paper analyzing the responses from International Authorities to Circular C. PCT 1398, dated November 4, 2013, presenting various characteristics of international search reports. This discussion by the quality subgroup is summarized in paragraphs 28 to 35 of Annex II to document PCT/MIA/21/22. Paragraphs 33 to 35 of that document outline future work in this area recommended by the Subgroup as follows:

“33. The Subgroup recommended that the International Bureau should continue to prepare this report in future years, but should seek to improve the presentation and accessibility to underlying information, including measures to allow the information on breakdowns by technical area to be more easily accessible.

/...

“34. The Subgroup also recommended that Authorities should share their findings from the report to assist other Authorities in their own analysis and to inform the International Bureau in understanding the ways in which the report was used in order to present the information in a way better suited to the needs of Authorities. The Subgroup’s e-forum was likely to be the appropriate place for this.

“35. The Subgroup finally recommended that interested Authorities should contact the European Patent Office to conduct bilateral discussions of aspects of the report with a view to preparing a set of practical examples for discussion by the Subgroup of information which could be found within the data which could provide useful information to Authorities.”

Updated Report

./. 3. Annex I to this Circular presents a report on characteristics of international search reports, updating the data presented in Circular C. PCT 1398 up to the end of the fourth quarter of 2012. Further explanation of the description of the data and definitions used is included in Annex II to this Circular.

4. The characteristics which are the subject of the report are entirely unchanged from the previous report, save that item 1.7 (Percentage of PCT search reports with D citations) is omitted since it was agreed that it was of no value. That characteristic nevertheless remains in the package of data from which the annexed charts are taken because, in the absence of agreement on how the other characteristics should be improved, it is easier to update the entire data set than to remove the low value items.

5. The characteristics reflect international search reports issued up to the end of 2012. More recent information is not available since the citation information is obtained from the PATSTAT database, which requires information to be extracted and checked largely manually from published international search reports and 2012 is the most recent year for which that information is sufficiently close to complete.

6. To improve accessibility of the information, the data has been made available to Authorities not only in the zip package containing the same data types as previous reports, but in the form of two Excel files posted on the quality subgroup forum – one covering the main characteristics covering several years and a second covering the technical breakdowns for 2012. These files (which require Excel for Windows 2010 or later) allow the various characteristics to be compared in different ways – probably most importantly, the technical breakdown shows how results for some of the characteristics vary across different technical fields within an International Authority.

Issues for the Quality Subgroup

7. It is apparent that none of these characteristics provide any direct measure of quality of the international search reports. Moreover, it is not clear how any meaningful measure of quality could be made using the data currently available. On the other hand, the trends in the characteristics over time, or differences between different fields of technology within an Authority may provide useful pointers to areas which could be the subject of investigation.

8. While several of the characteristics (such as those relating to sources of applicants or languages of citations) are clearly extremely closely tied to the circumstances of each individual Authority, for some of the characteristics, comparison between Authorities may be of interest, if only to note whether any trends are in the same or different directions.

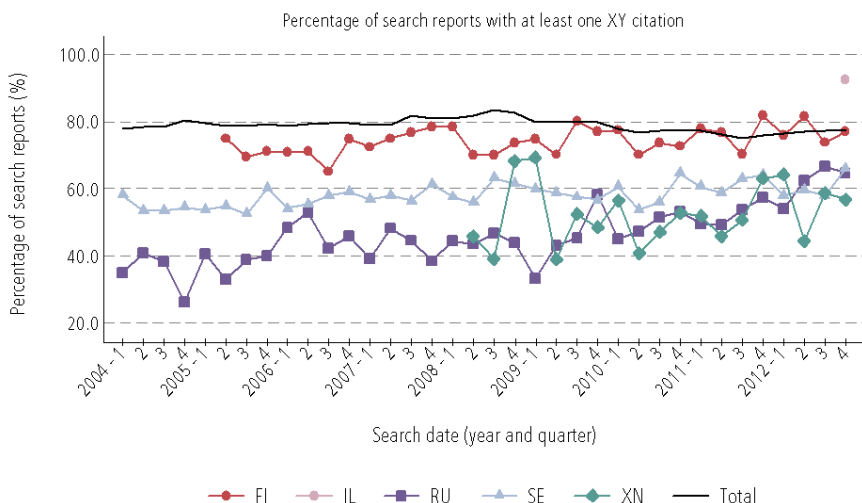
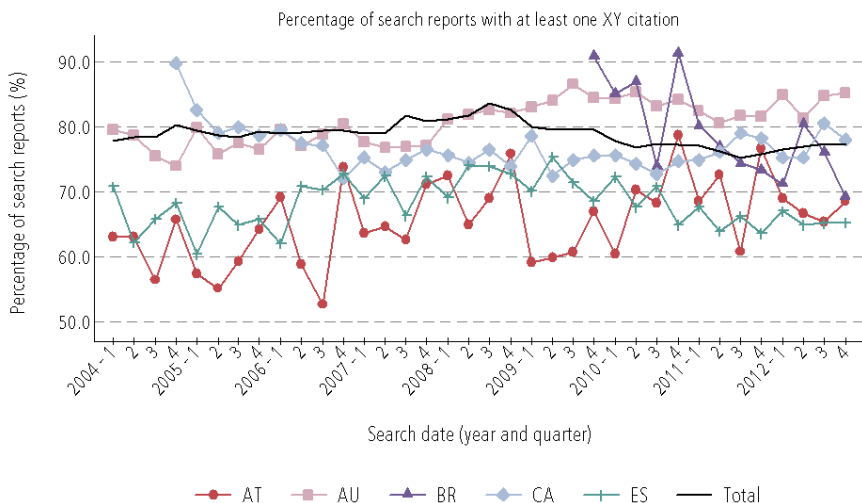
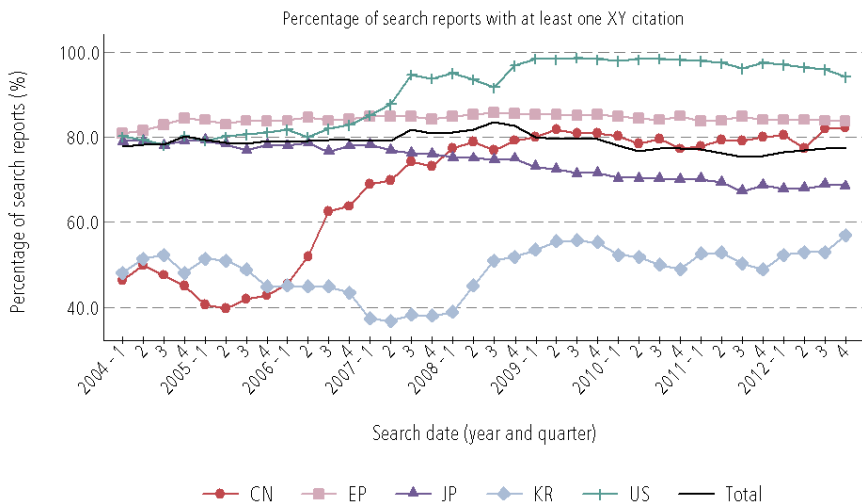
/...

CHARACTERISTICS OF INTERNATIONAL SEARCH REPORTS

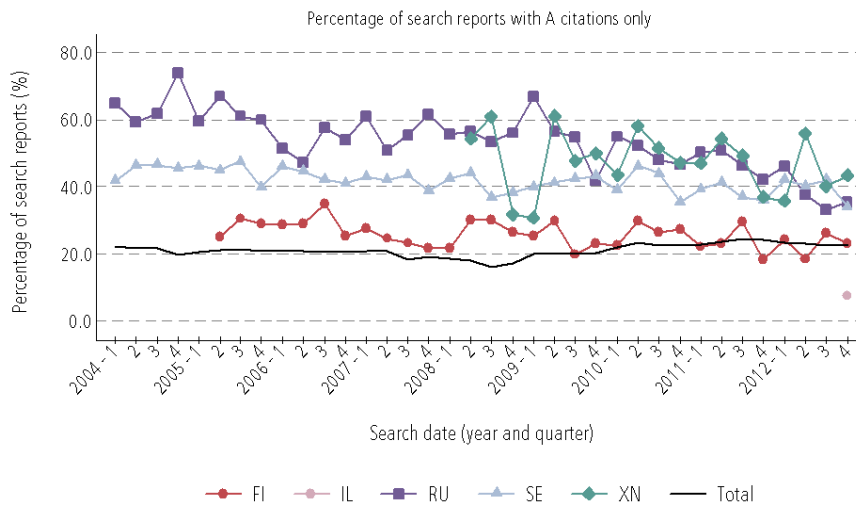
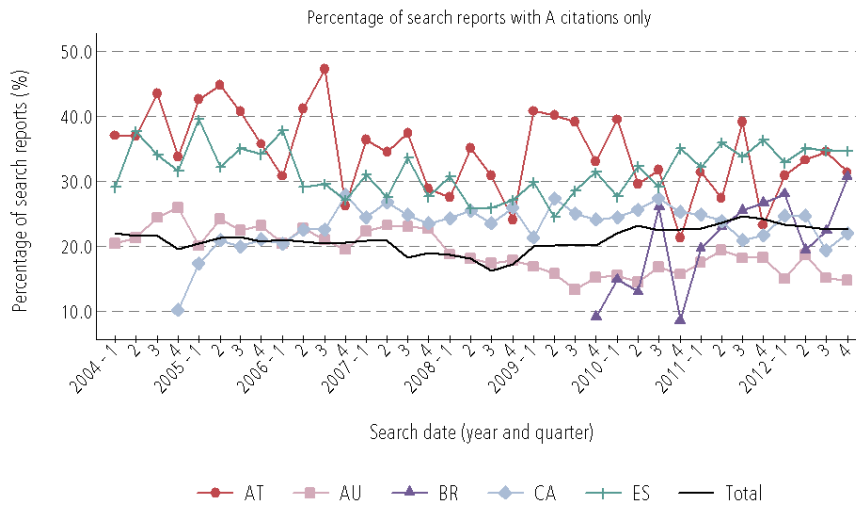
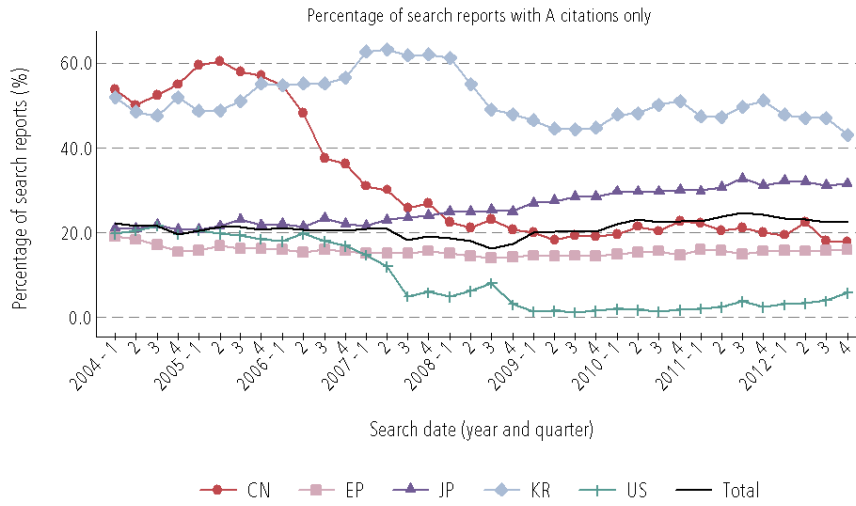
Table of Contents

1.1 – Percentage of PCT search reports with at least one X, Y or E citation.....	2
1.2 – Percentage of PCT search reports with A citations only.....	3
1.3 – Percentage of PCT search reports with P or E citations.....	4
1.4 – Percentage of citations in the category of P or E	5
1.5 – Percentage of citations in the category of P and X.....	6
1.6 – Percentage of PCT search reports with Y citations and without X citations.....	7
1.8 – Percentage of PCT search reports with O/T/L citations	8
1.9 – P/E Citation Breakdown for PCT Search reports (2012)	9
1.10 – Percentage of PCT search reports with at least one X or Y citation by top applicant's origin (2012).....	11
2.1 – Average number of citations per search report	12
2.2 – Average number of NPL citations per search report	13
2.3 – Average number of patent literature citations per search report.....	14
2.4 – Average number of X/Y patent literature citations per search report	15
2.5 – Percentage of NPL citations in all citations	16
2.6 – Percentage of NPL citations in the category of X or Y in all citations	17
2.7 – Percentage of PCT search reports with at least one X/Y NPL Citation.....	18
2.8 – Percentage of NPL citations in all citations	19
2.9 – Distribution of patent & non patent literature citations (2012).....	20
3.1 – Percentage of patent citations in non-official languages	21
3.2 – Percentage of patent citations by top publication authorities (2012)	22
3.3 – Percentage of patent citations by processing authorities (i.e. publication authorities except WO) (2012)	23
3.4 – Percentage of PCT search reports by top applicant's origin (2012)	24
Data set description	1
Data issues.....	1
Definition of concepts.....	1
Technology breakdown	1
Applicant origin.....	2
XY rate (Searches with XY citations).....	2
Citation category availability	2
A only rate	3
Y no X rate	3
Search date.....	3
Patent Literature/Non-Patent Literature	3
Non official language.....	3
Publication Authority (of citation)	4
Processing Authority (of citation).....	4

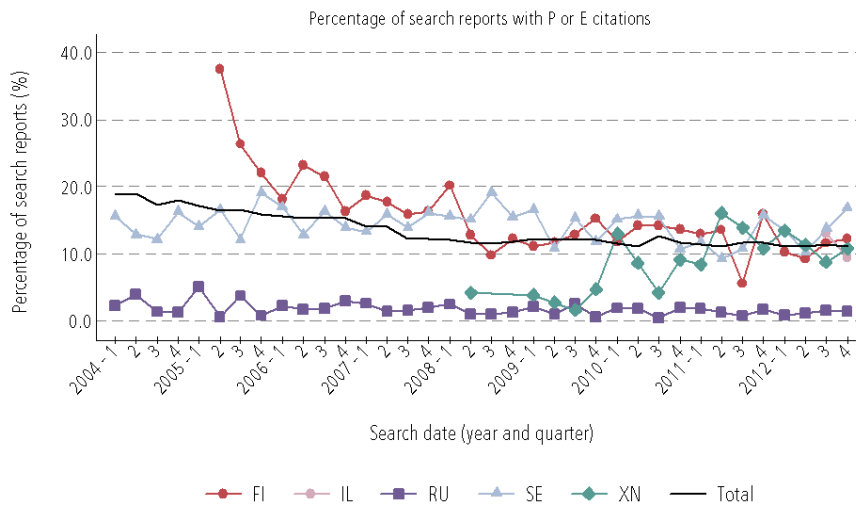
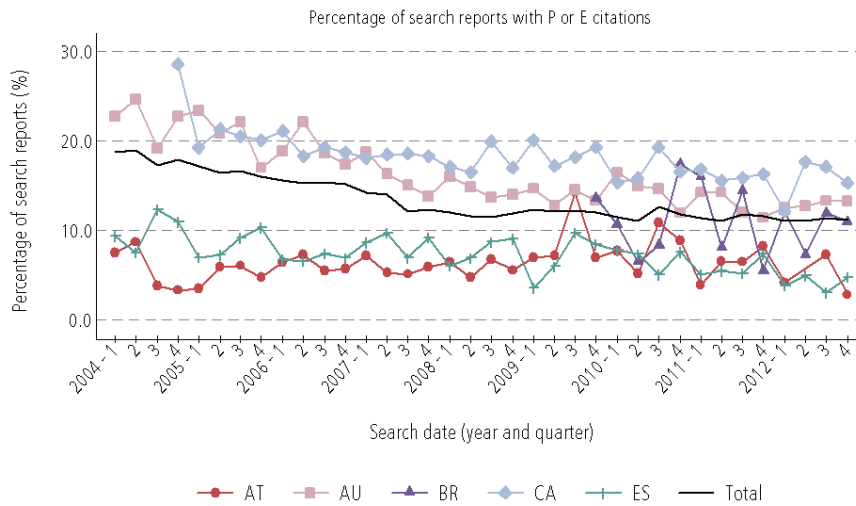
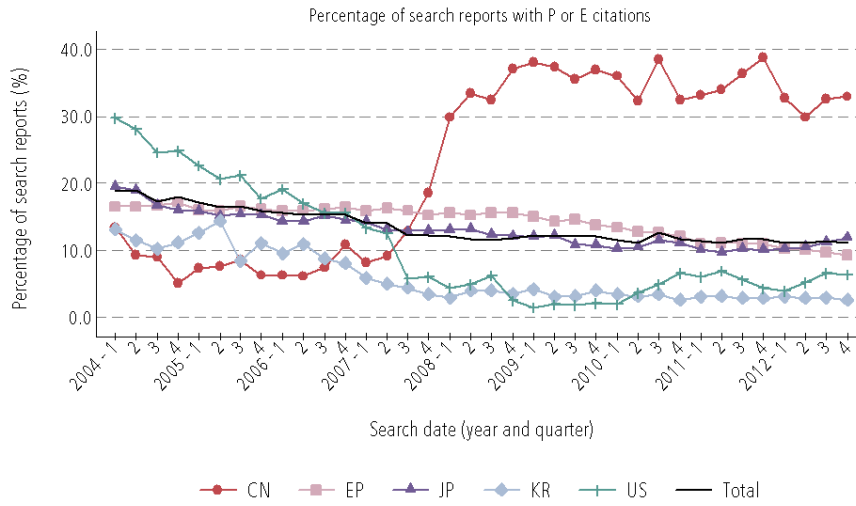
1.1 – PERCENTAGE OF PCT SEARCH REPORTS WITH AT LEAST ONE X, Y OR E CITATION



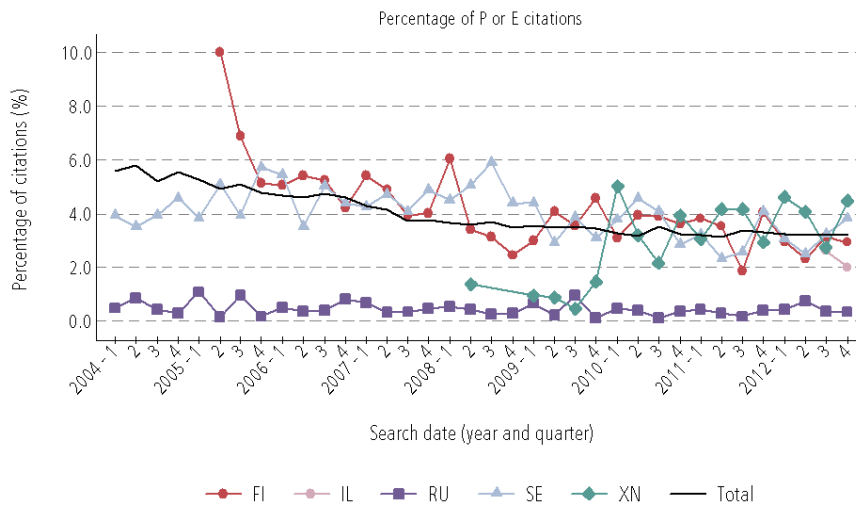
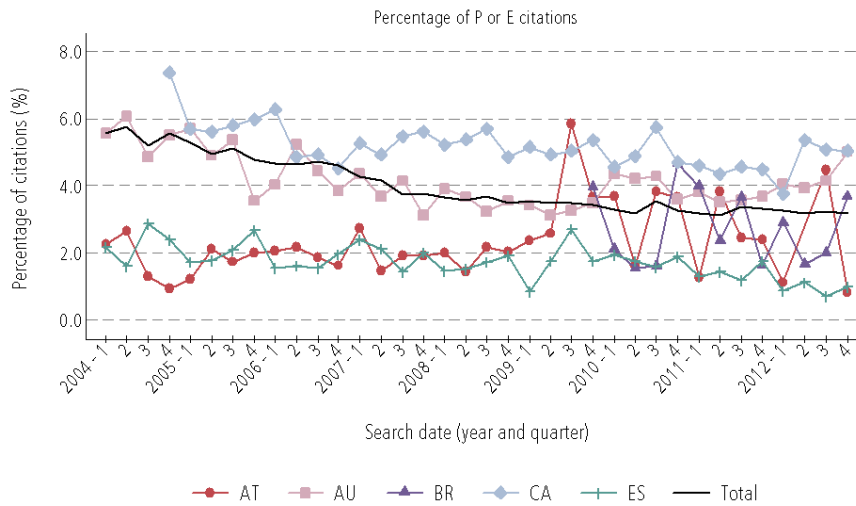
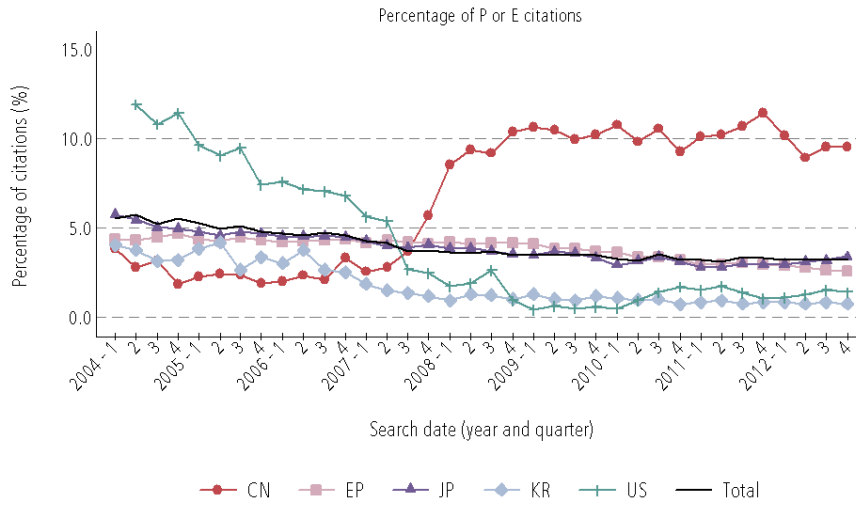
1.2 – PERCENTAGE OF PCT SEARCH REPORTS WITH A CITATIONS ONLY



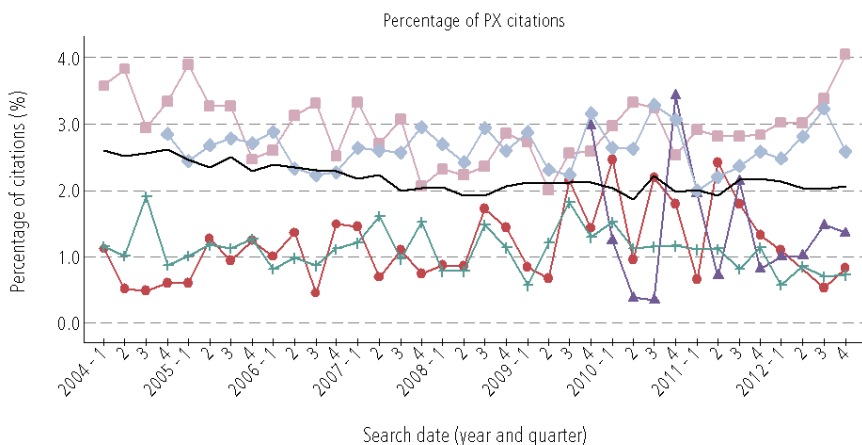
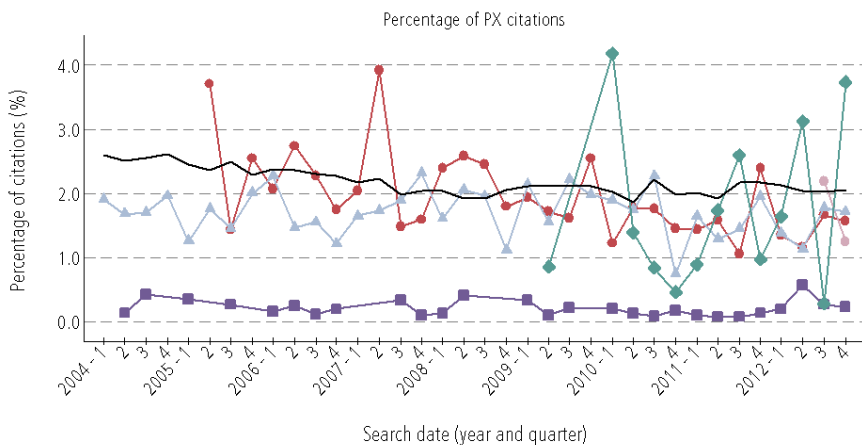
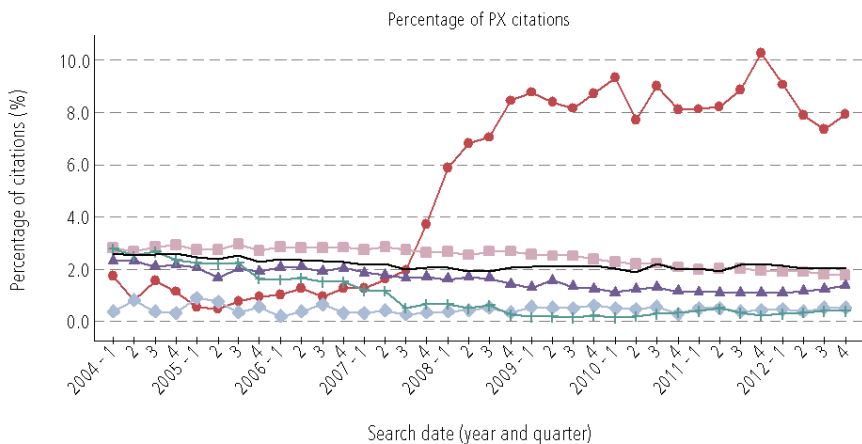
1.3 – PERCENTAGE OF PCT SEARCH REPORTS WITH P OR E CITATIONS



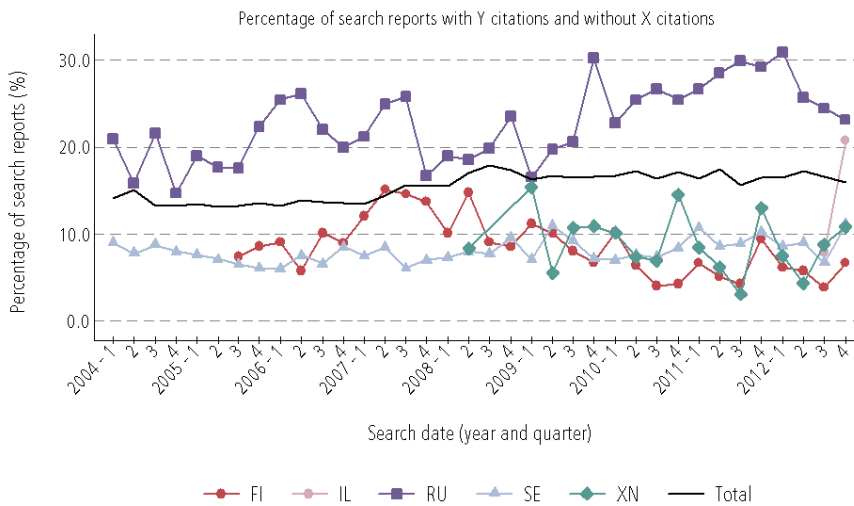
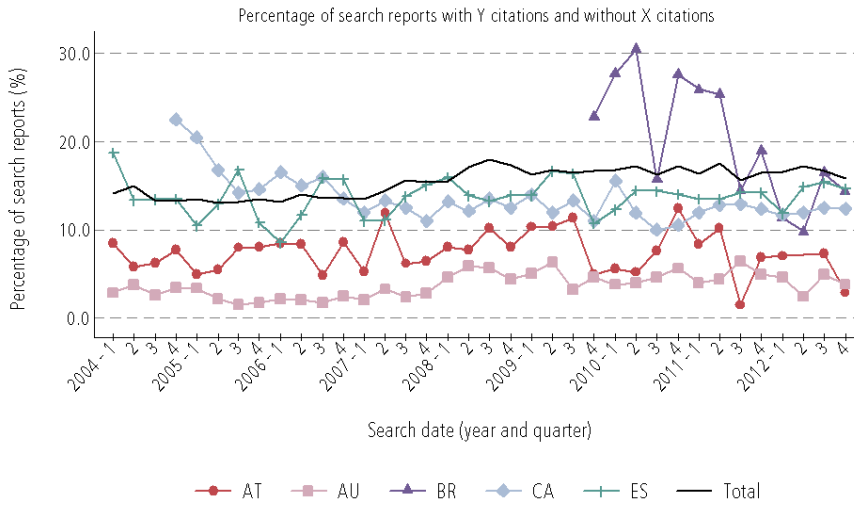
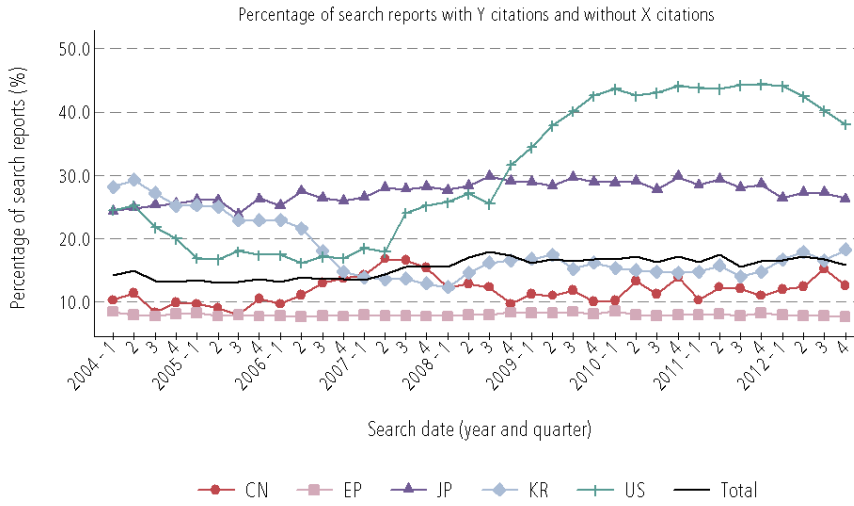
1.4 – PERCENTAGE OF CITATIONS IN THE CATEGORY OF P OR E



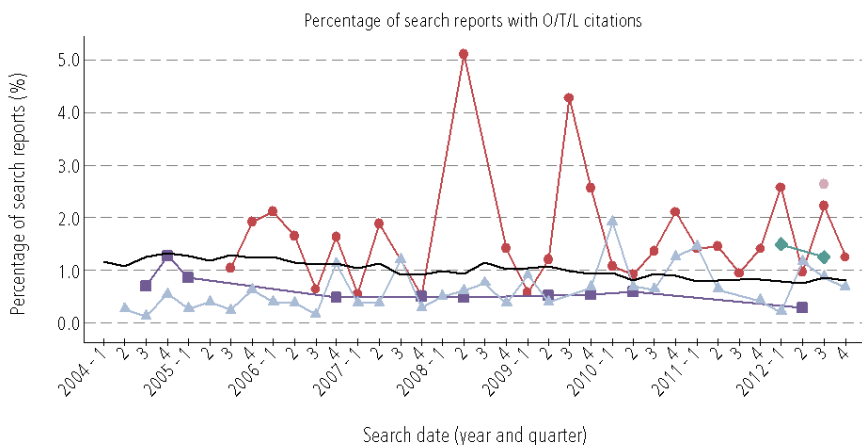
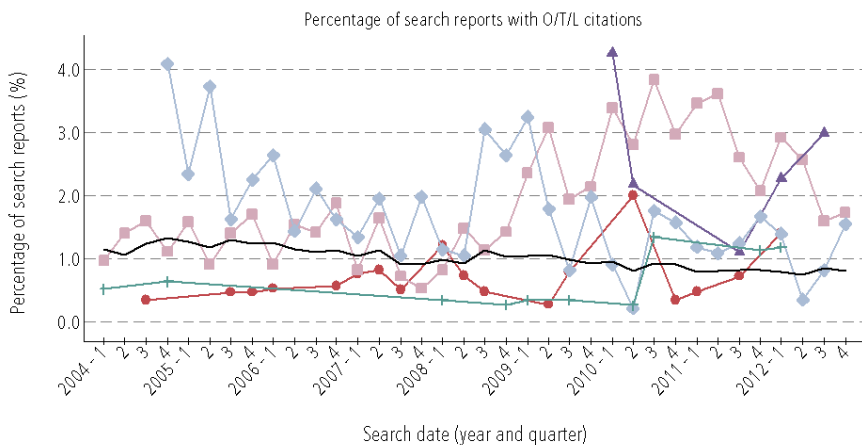
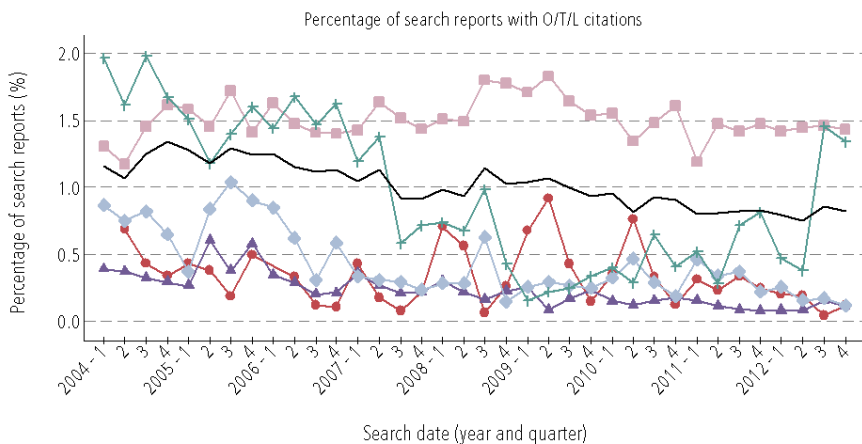
1.5 – PERCENTAGE OF CITATIONS IN THE CATEGORY OF P AND X



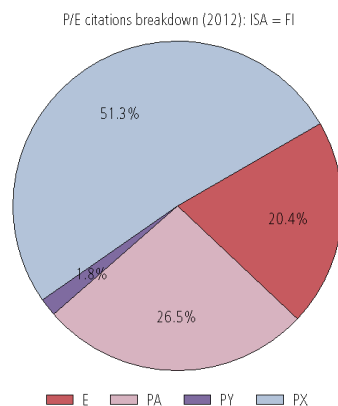
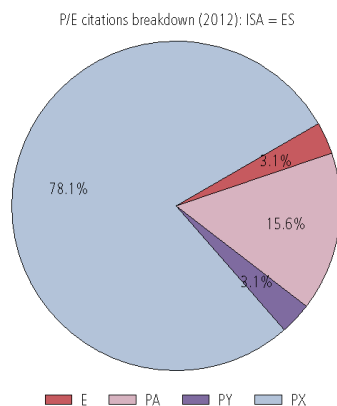
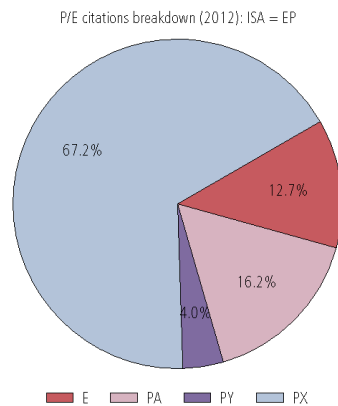
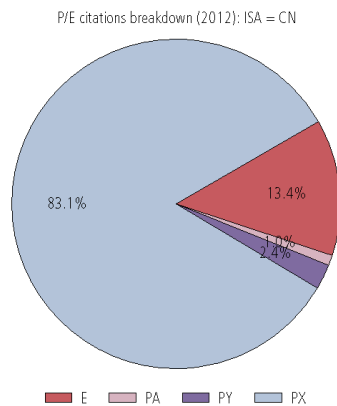
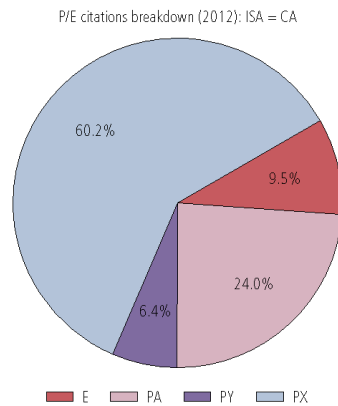
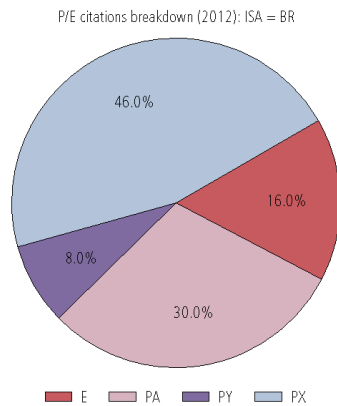
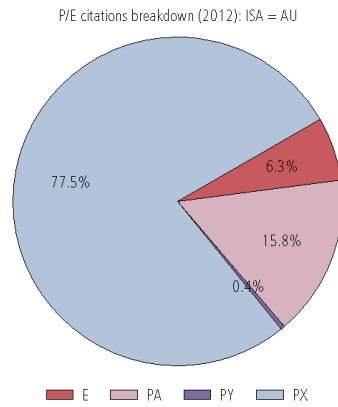
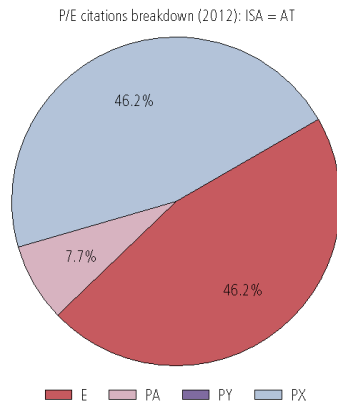
1.6 – PERCENTAGE OF PCT SEARCH REPORTS WITH Y CITATIONS AND WITHOUT X CITATIONS



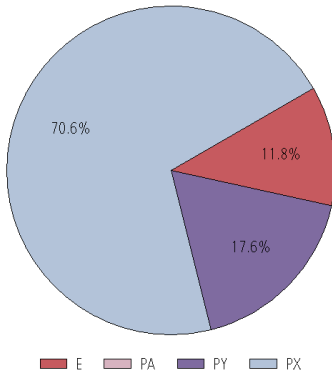
1.8 – PERCENTAGE OF PCT SEARCH REPORTS WITH O/T/L CITATIONS



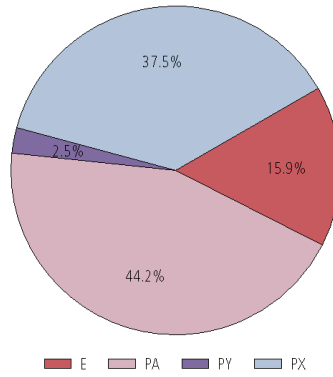
1.9 – P/E CITATION BREAKDOWN FOR PCT SEARCH REPORTS (2012)



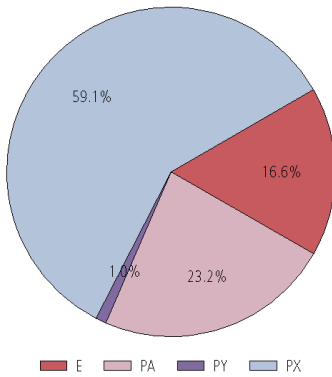
P/E citations breakdown (2012): ISA = IL



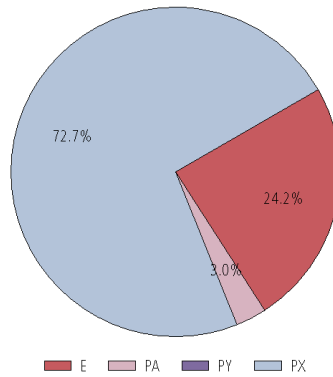
P/E citations breakdown (2012): ISA = JP



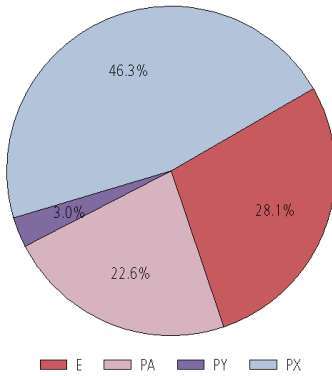
P/E citations breakdown (2012): ISA = KR



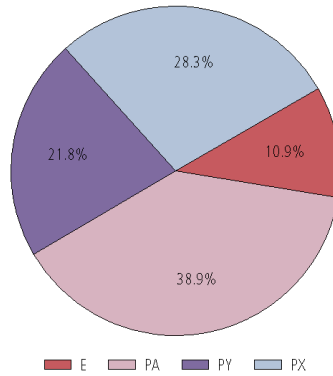
P/E citations breakdown (2012): ISA = RU



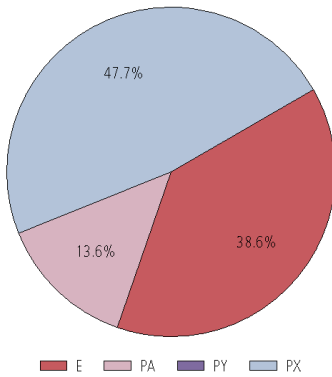
P/E citations breakdown (2012): ISA = SE



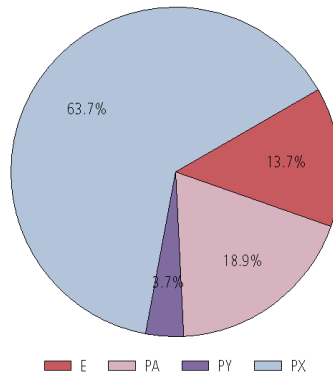
P/E citations breakdown (2012): ISA = US



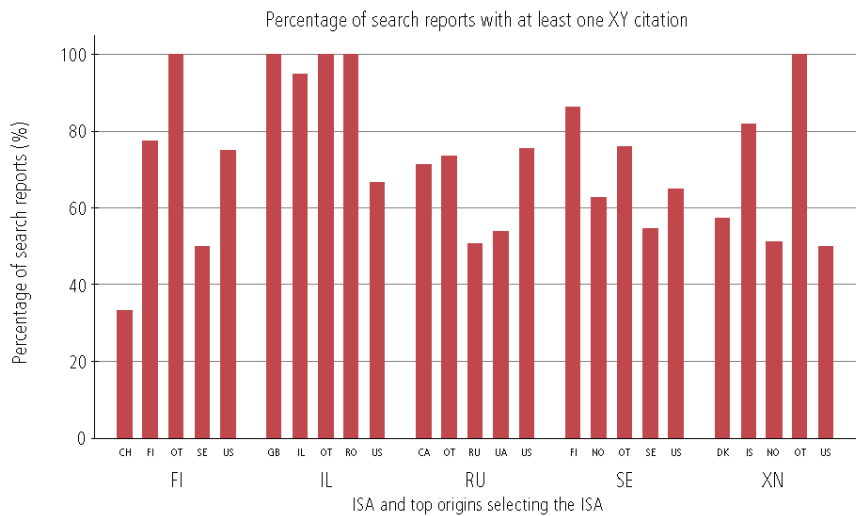
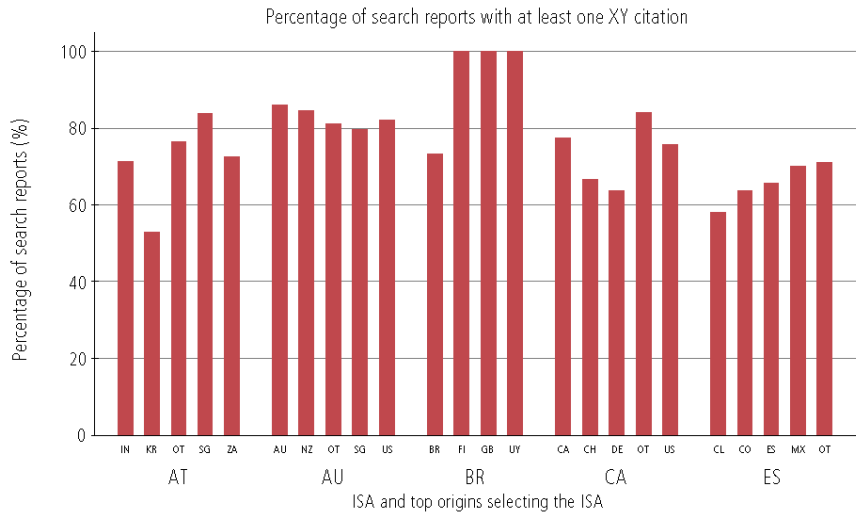
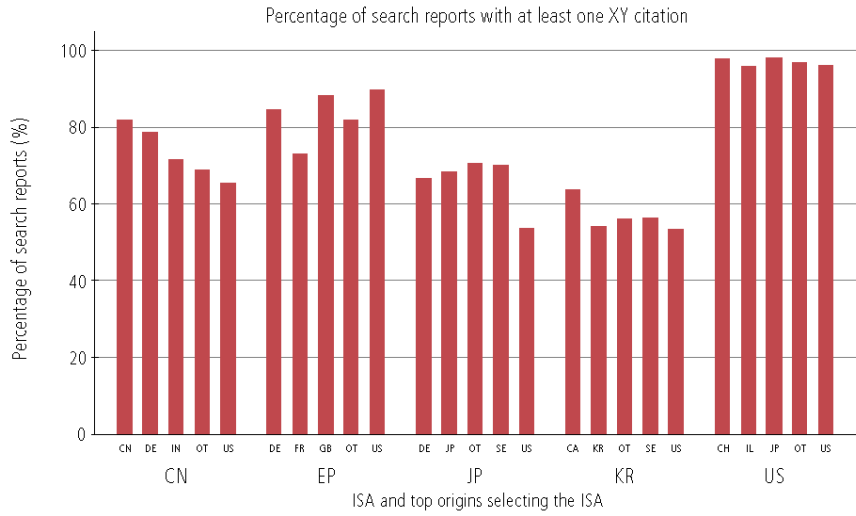
P/E citations breakdown (2012): ISA = XN



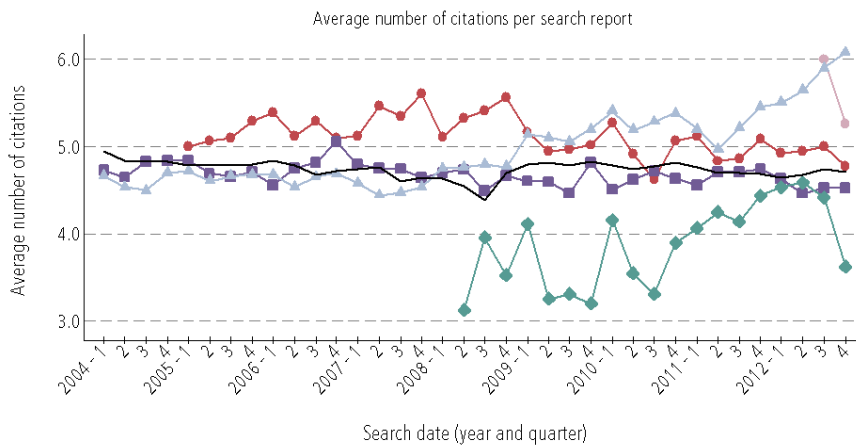
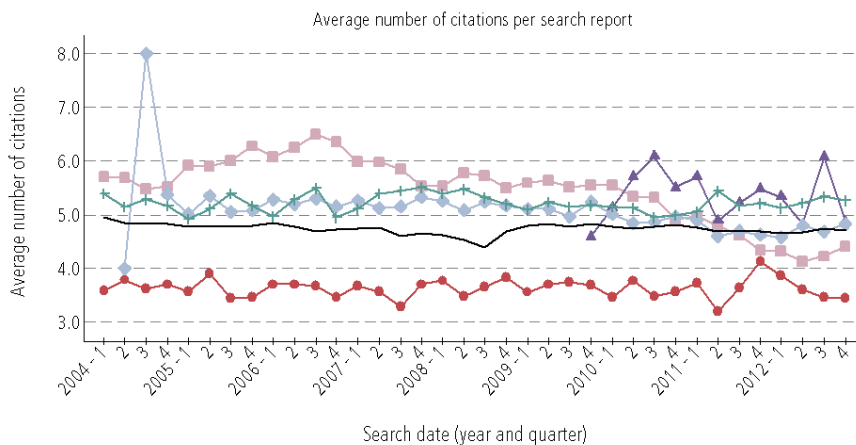
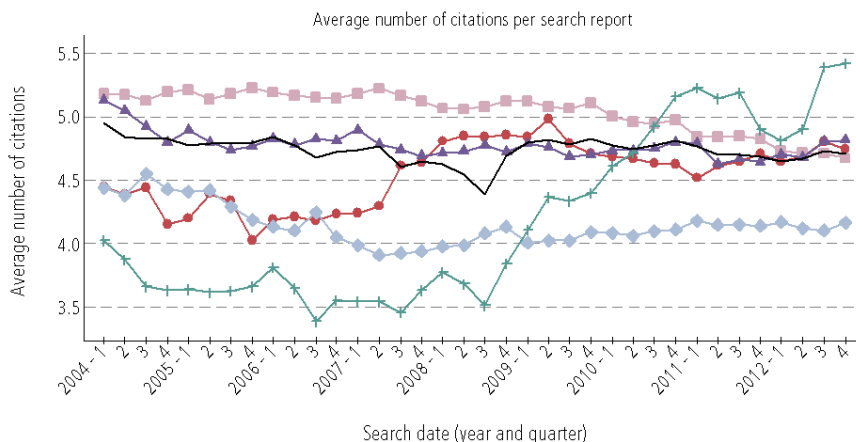
P/E citations breakdown (2012): Total



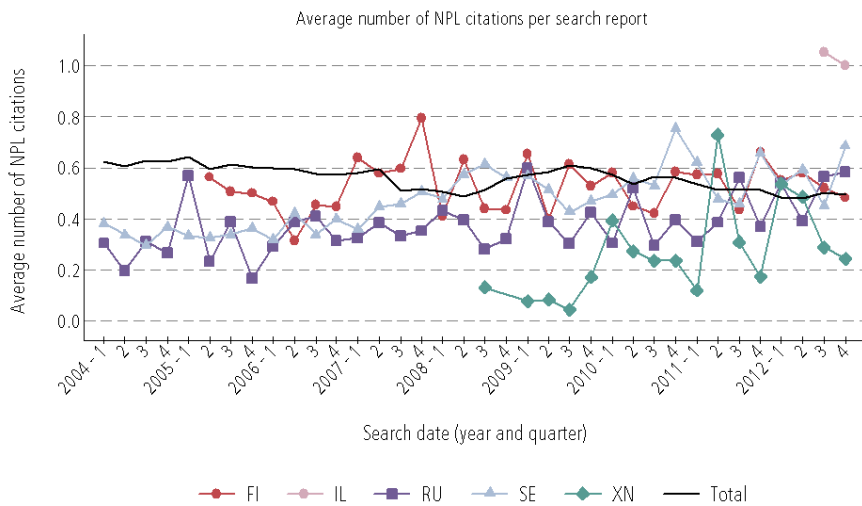
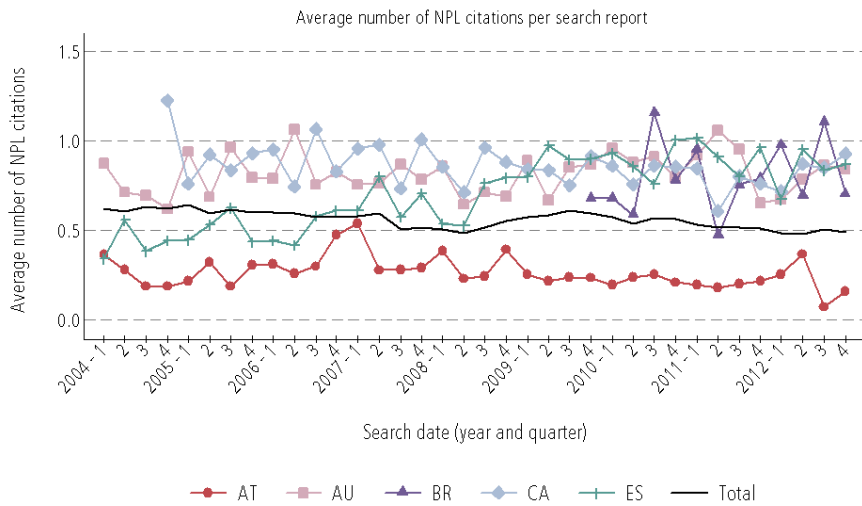
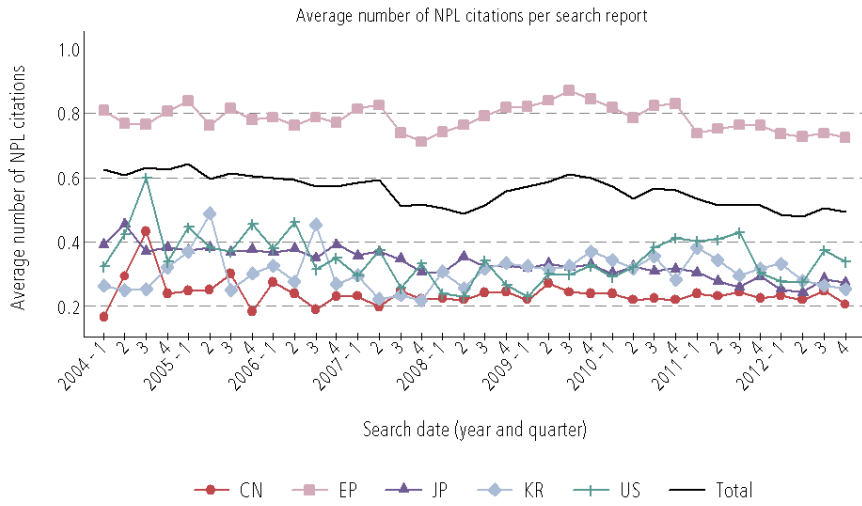
1.10 – PERCENTAGE OF PCT SEARCH REPORTS WITH AT LEAST ONE X OR Y CITATION BY TOP APPLICANT'S ORIGIN (2012)



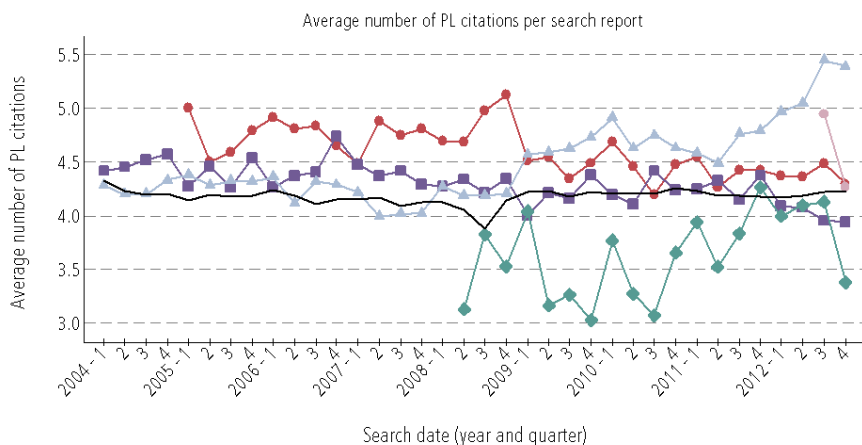
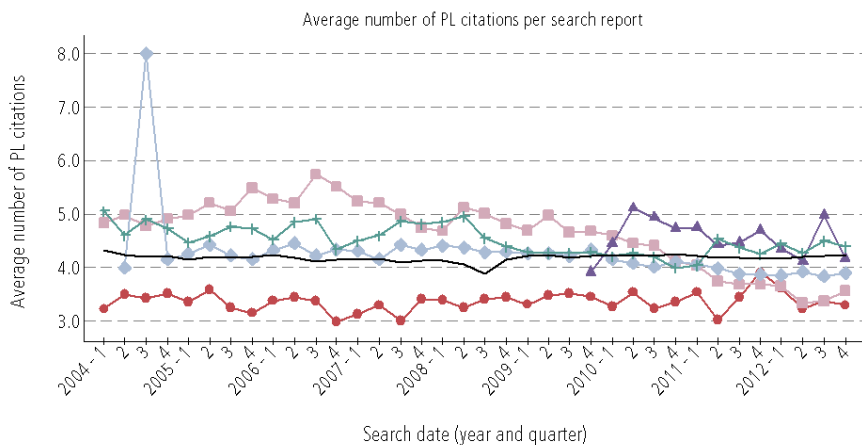
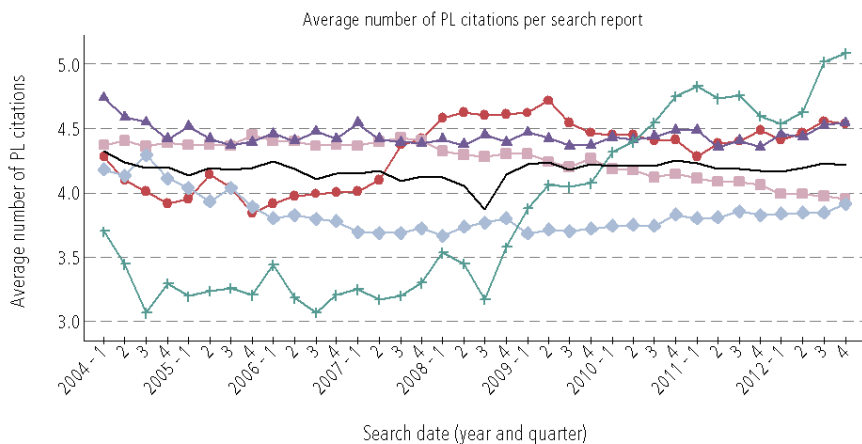
2.1 – AVERAGE NUMBER OF CITATIONS PER SEARCH REPORT



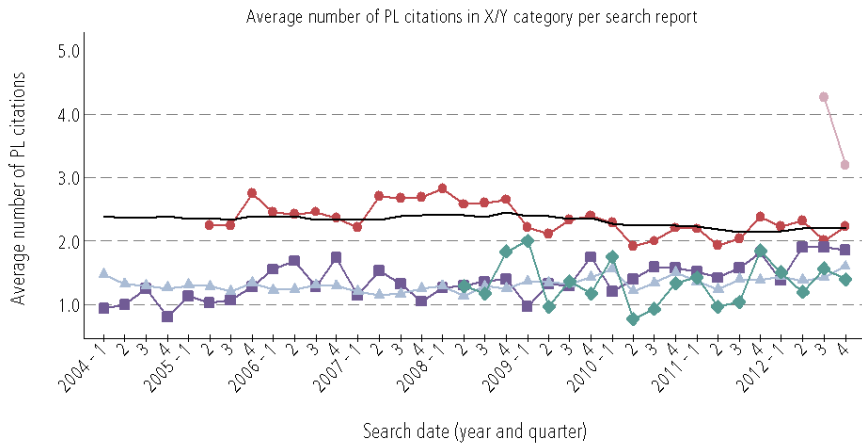
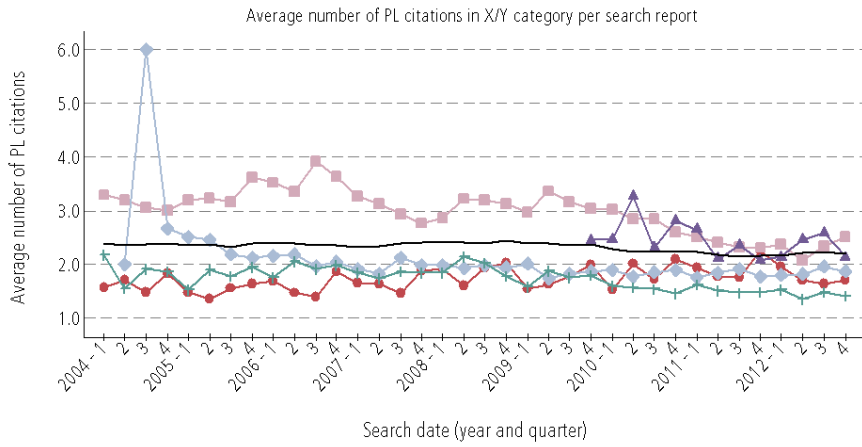
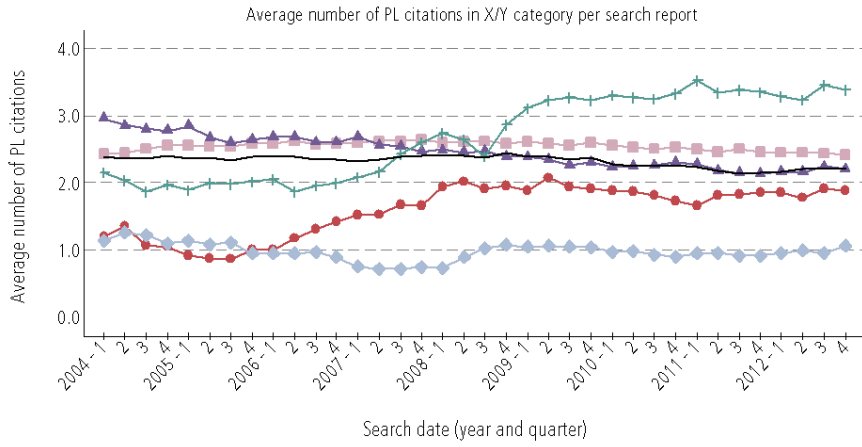
2.2 – AVERAGE NUMBER OF NPL CITATIONS PER SEARCH REPORT



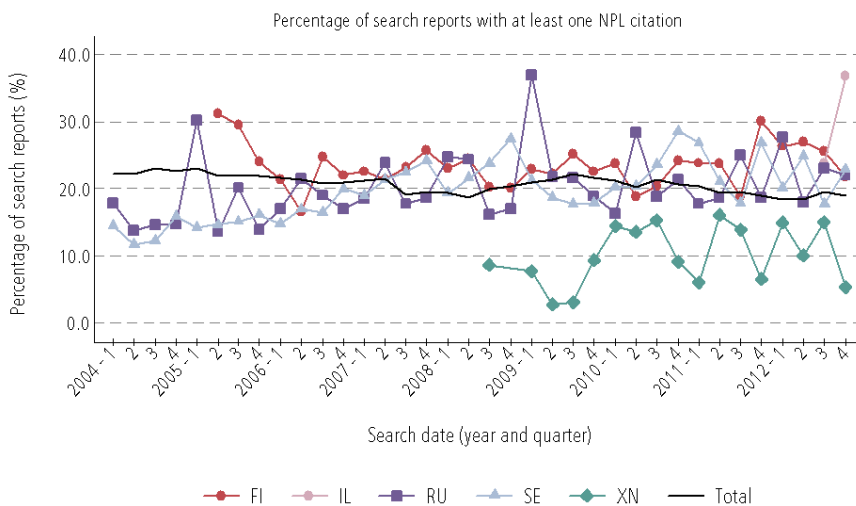
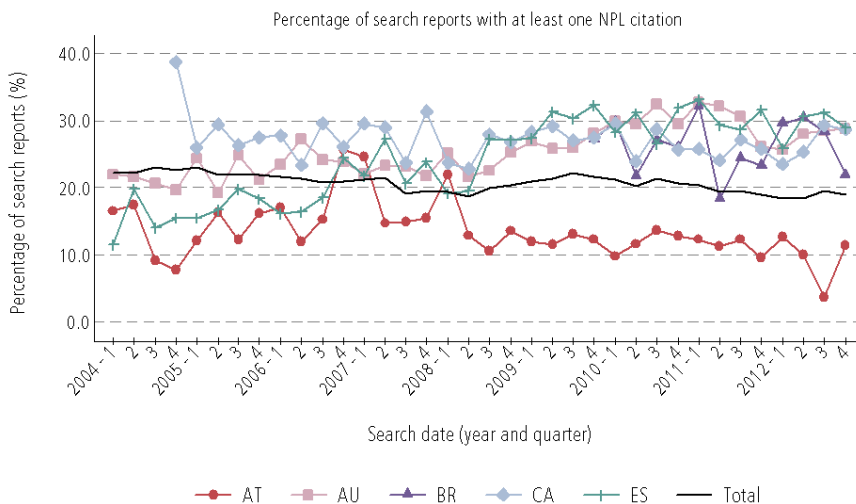
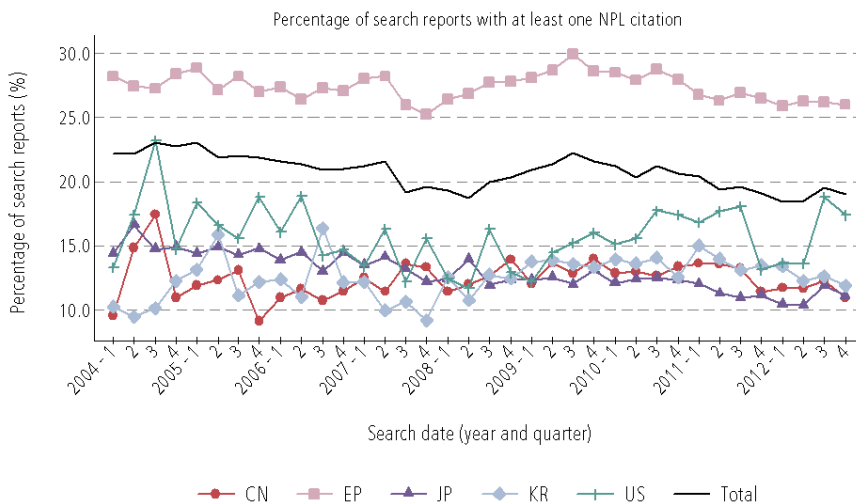
2.3 – AVERAGE NUMBER OF PATENT LITERATURE CITATIONS PER SEARCH REPORT



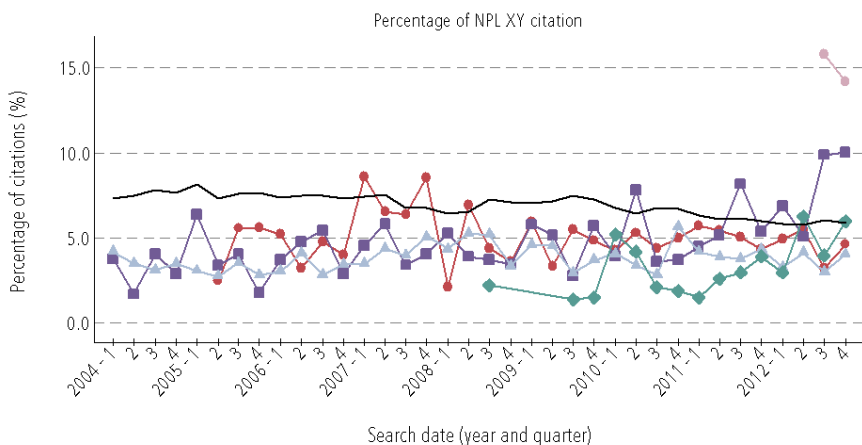
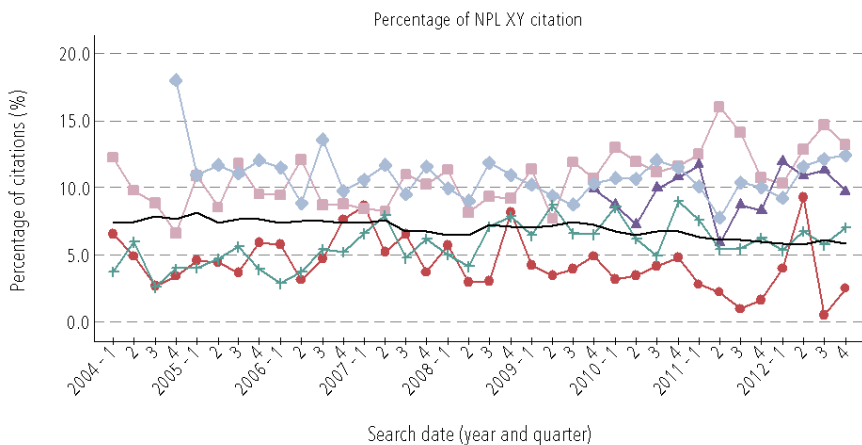
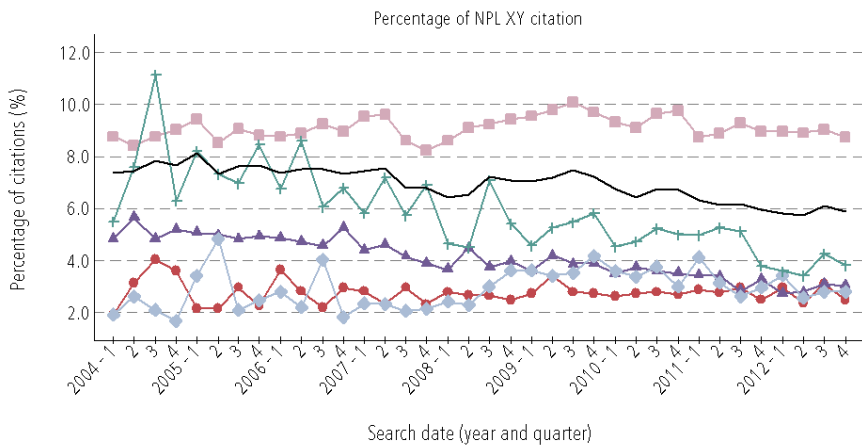
2.4 – AVERAGE NUMBER OF X/Y PATENT LITERATURE CITATIONS PER SEARCH REPORT



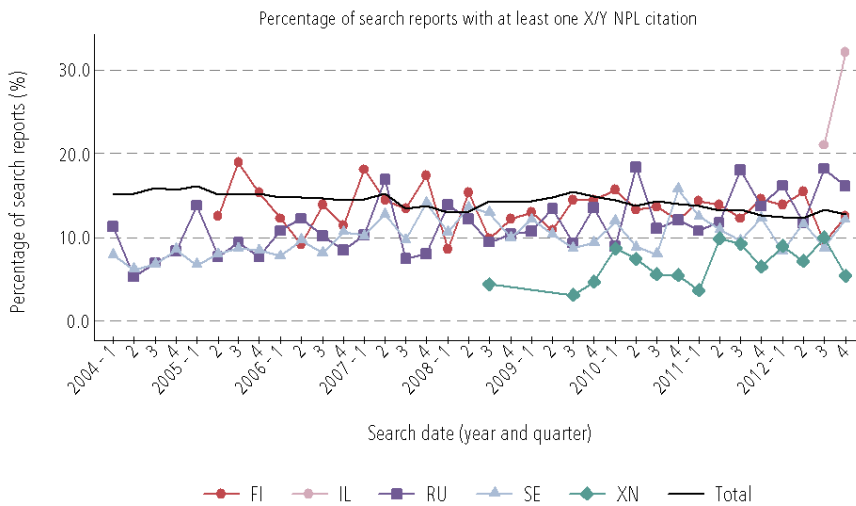
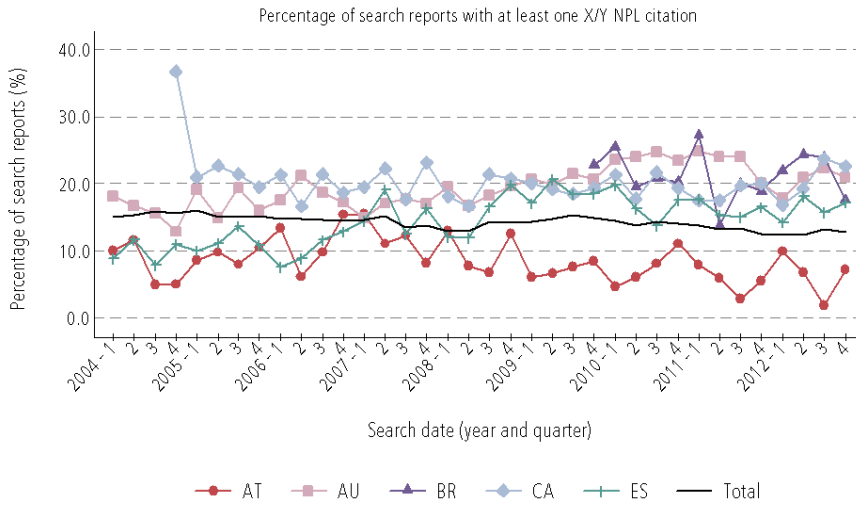
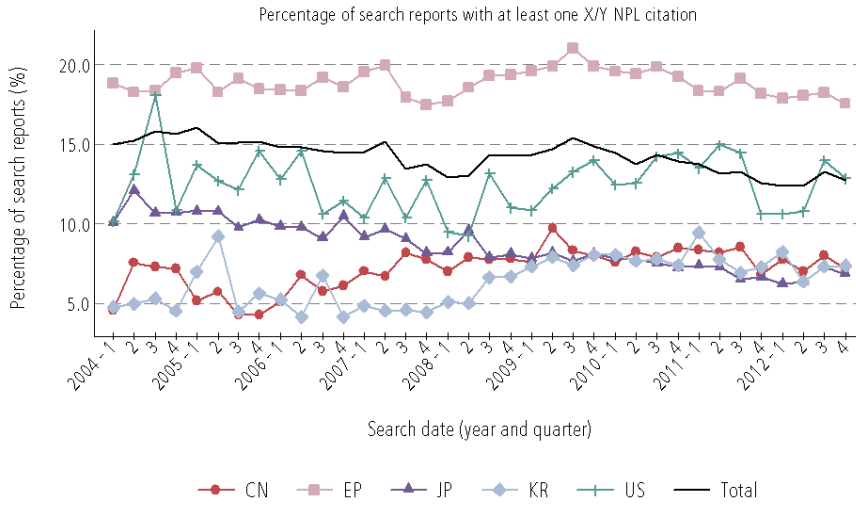
2.5 – PERCENTAGE OF NPL CITATIONS IN ALL CITATIONS



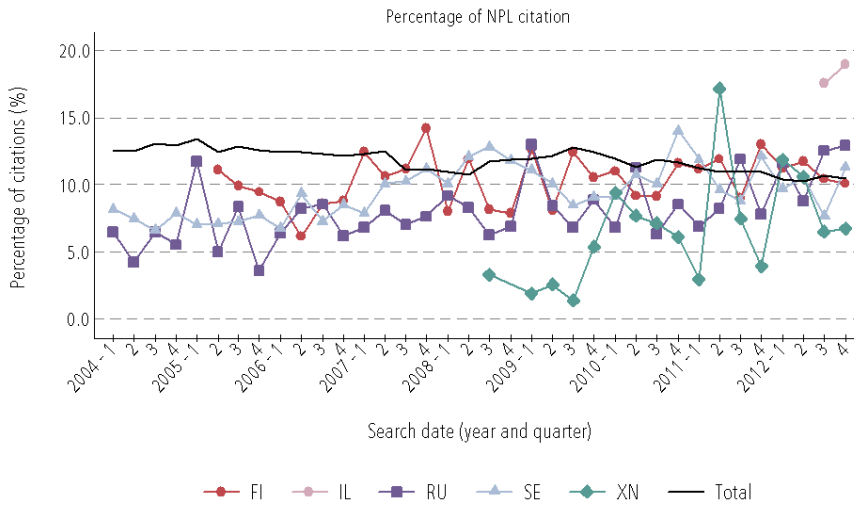
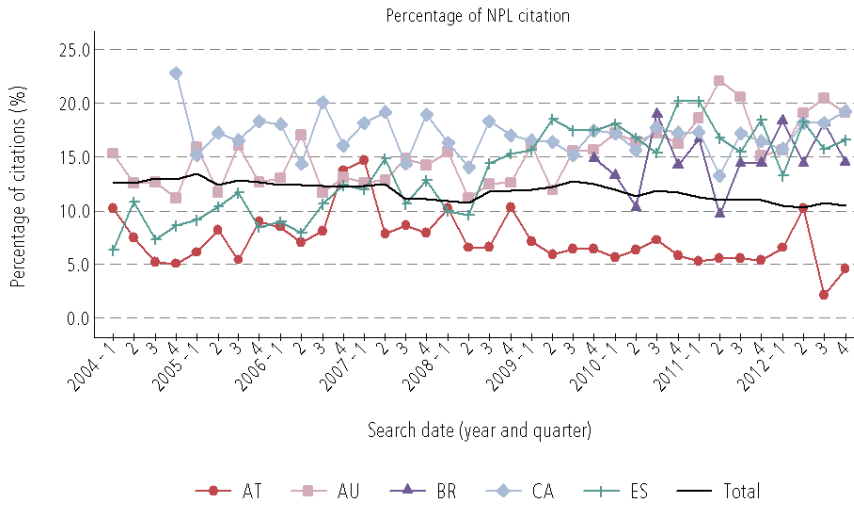
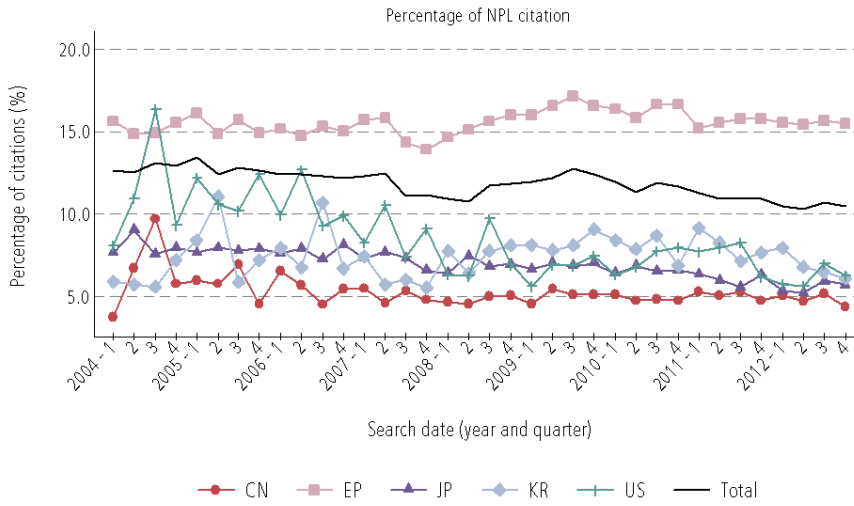
2.6 – PERCENTAGE OF NPL CITATIONS IN THE CATEGORY OF X OR Y IN ALL CITATIONS



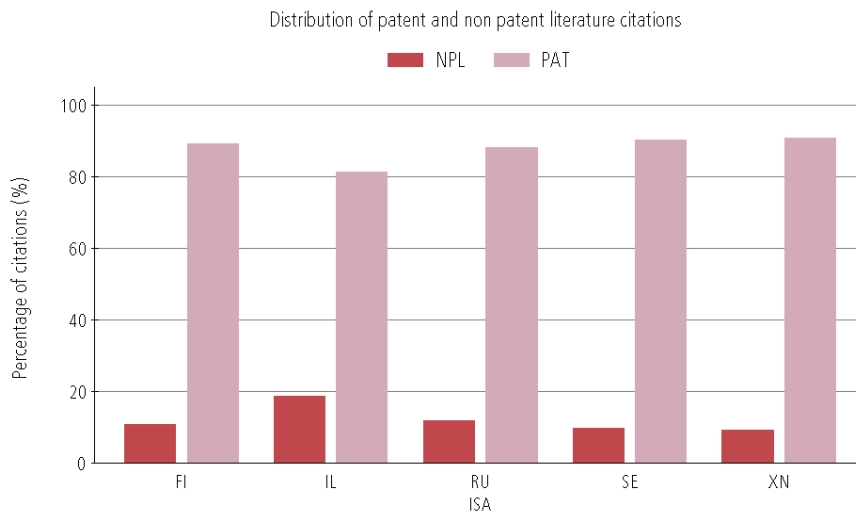
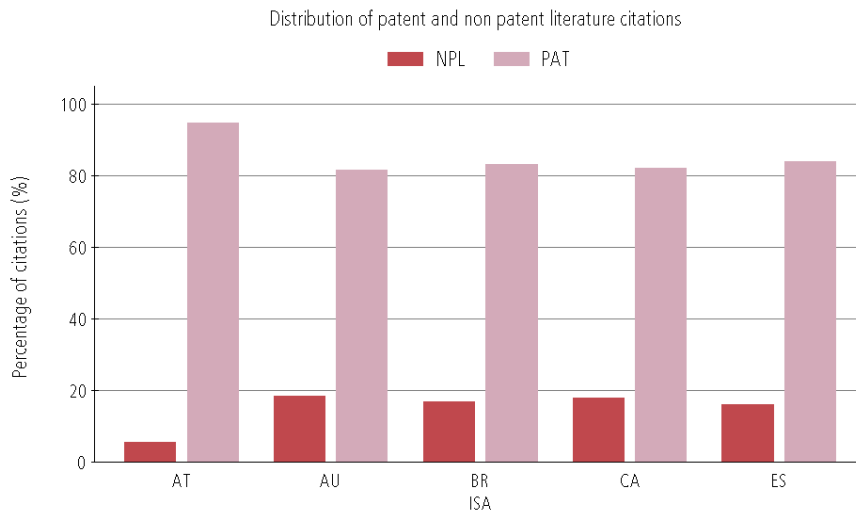
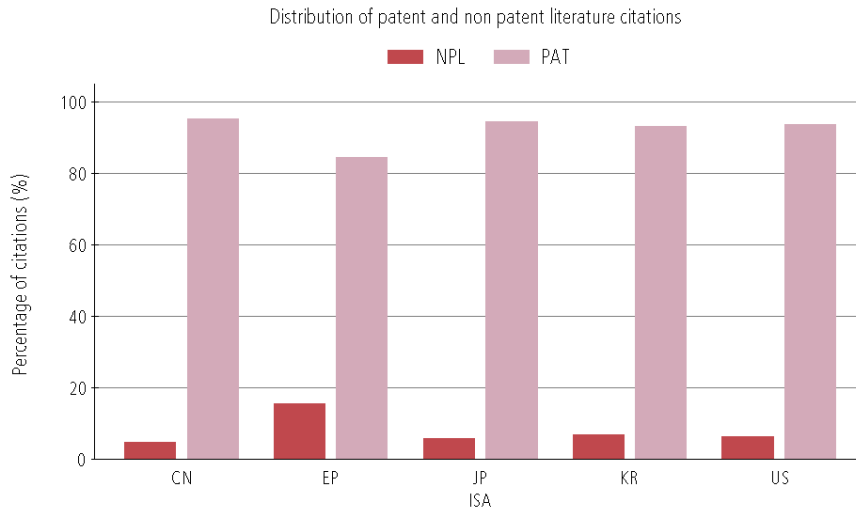
2.7 – PERCENTAGE OF PCT SEARCH REPORTS WITH AT LEAST ONE X/Y NPL CITATION



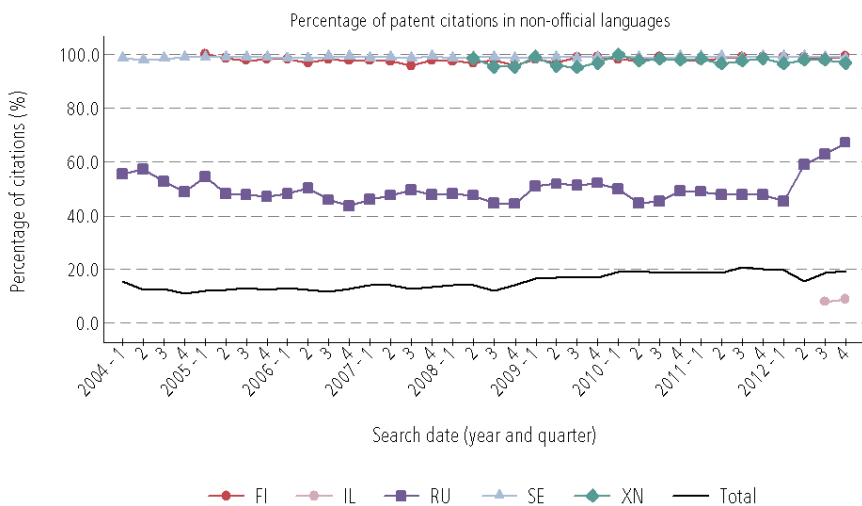
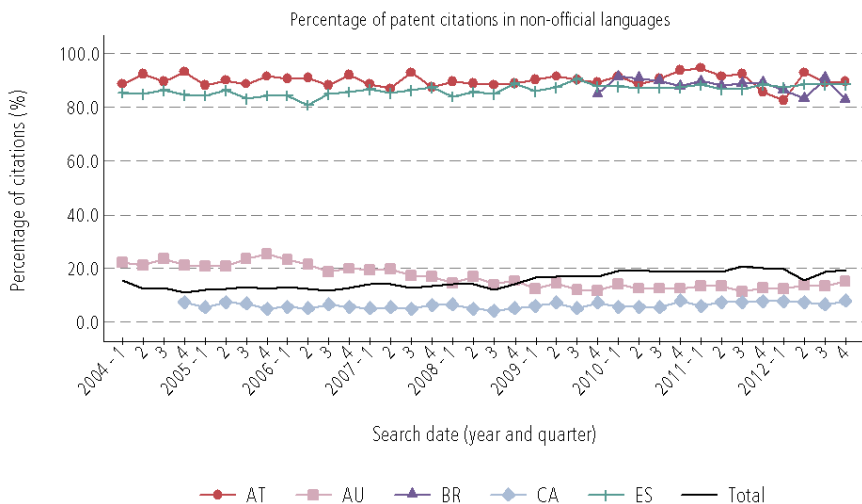
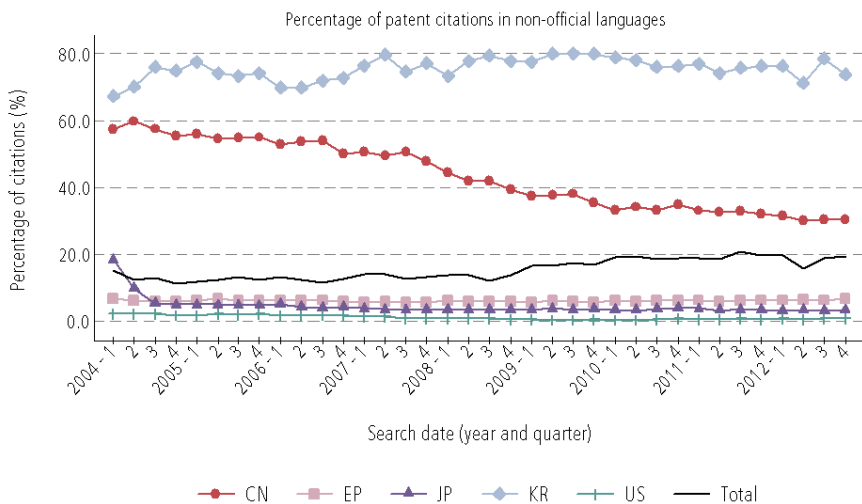
2.8 – PERCENTAGE OF NPL CITATIONS IN ALL CITATIONS



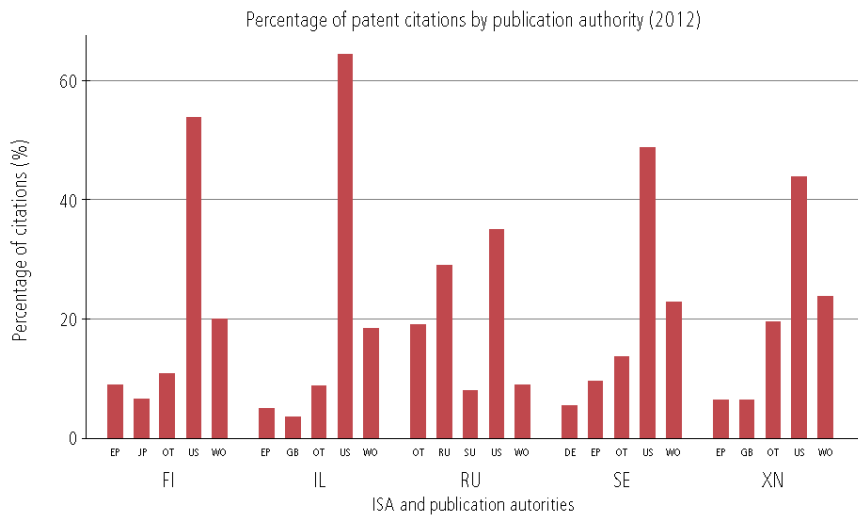
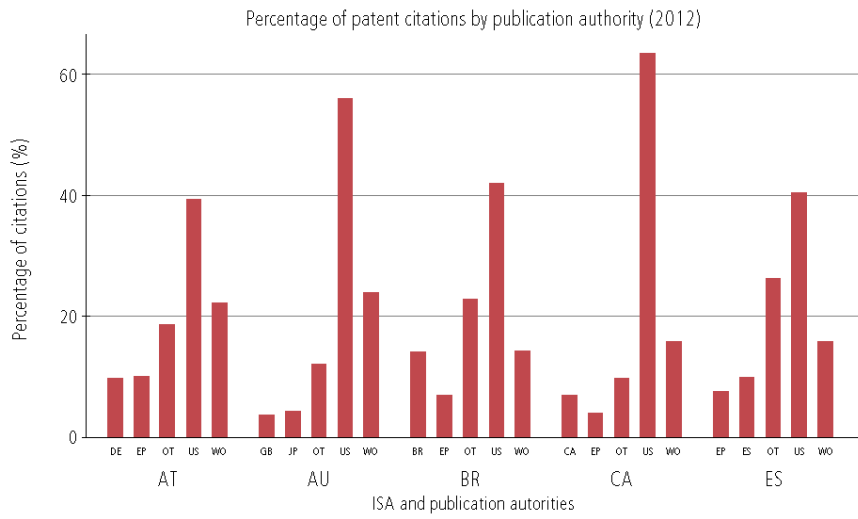
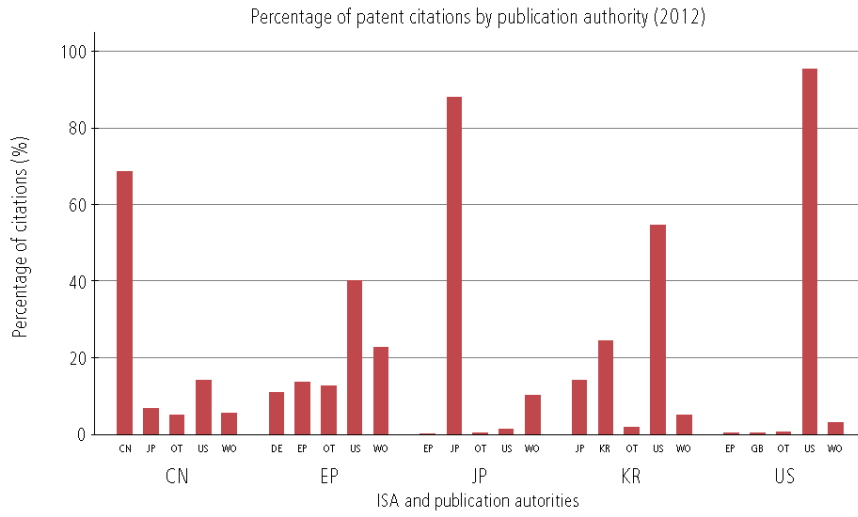
2.9 – DISTRIBUTION OF PATENT & NON PATENT LITERATURE CITATIONS (2012)



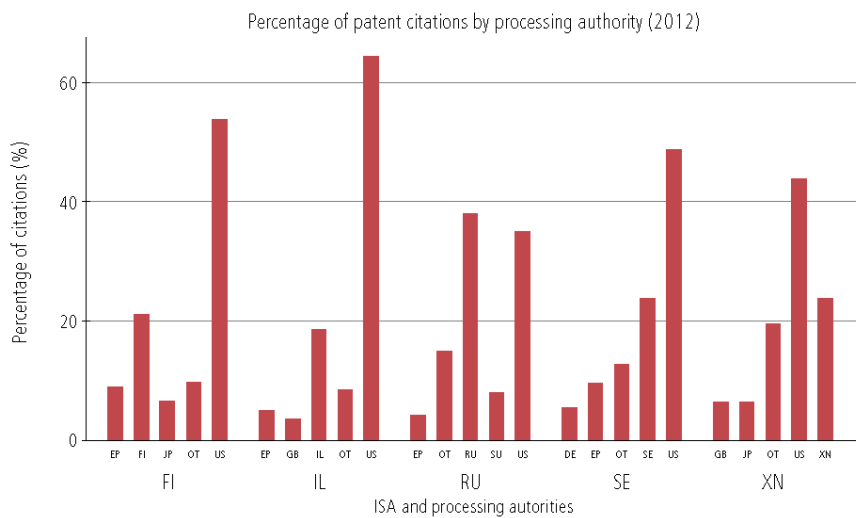
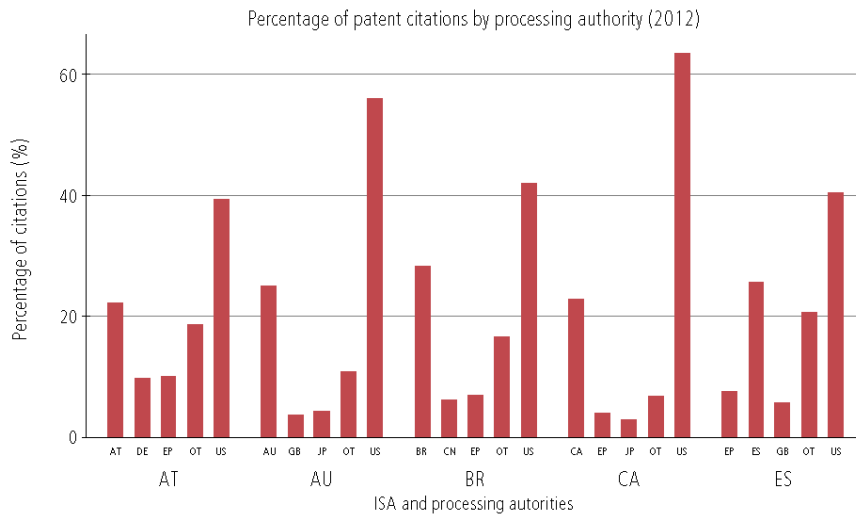
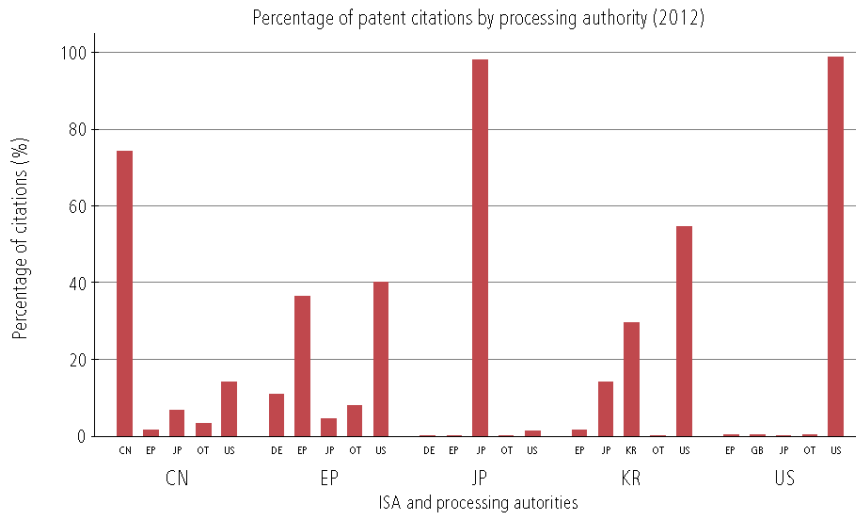
3.1 – PERCENTAGE OF PATENT CITATIONS IN NON-OFFICIAL LANGUAGES



3.2 – PERCENTAGE OF PATENT CITATIONS BY TOP PUBLICATION AUTHORITIES (2012)

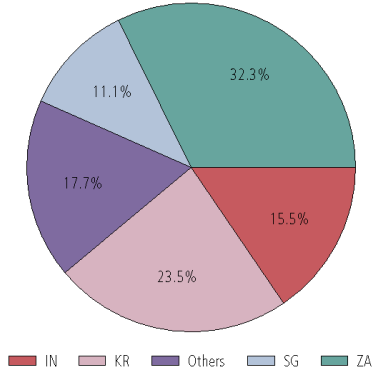


3.3 – PERCENTAGE OF PATENT CITATIONS BY PROCESSING AUTHORITIES (I.E. PUBLICATION AUTHORITIES EXCEPT WO) (2012)

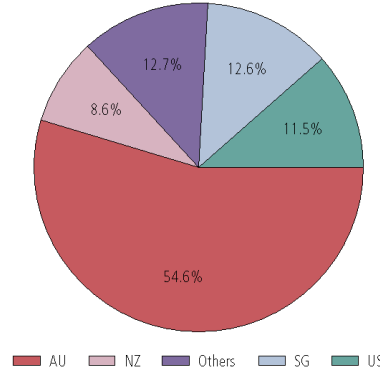


3.4 – PERCENTAGE OF PCT SEARCH REPORTS BY TOP APPLICANT’S ORIGIN (2012)

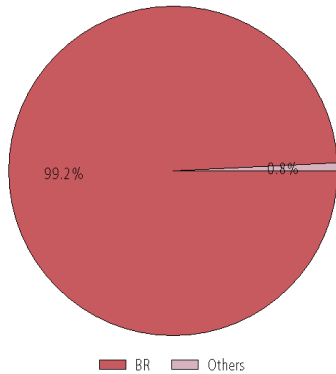
Percentage of search reports by applicants' origin (2012): ISA = AT



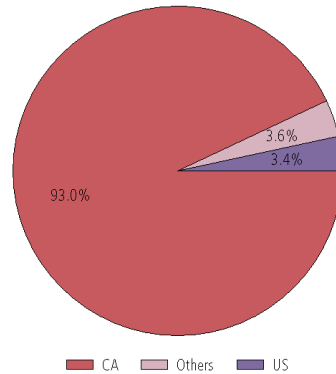
Percentage of search reports by applicants' origin (2012): ISA = AU



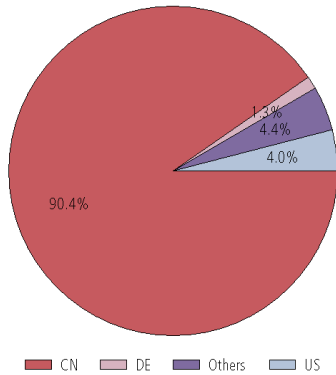
Percentage of search reports by applicants' origin (2012): ISA = BR



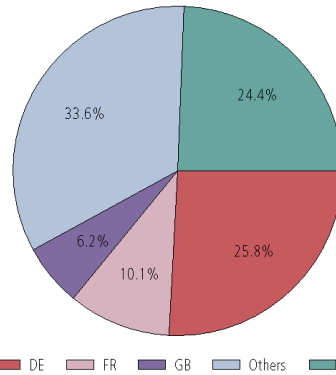
Percentage of search reports by applicants' origin (2012): ISA = CA



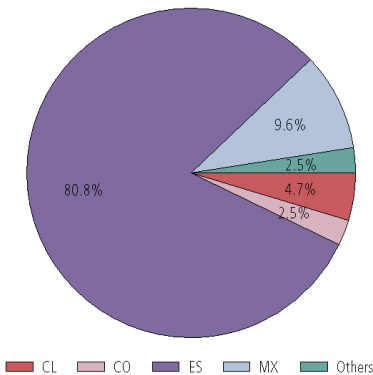
Percentage of search reports by applicants' origin (2012): ISA = CN



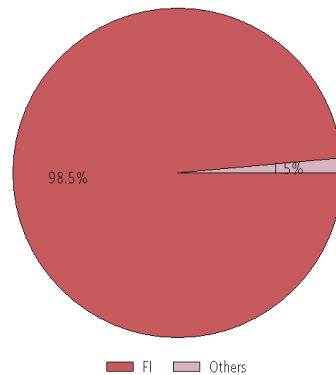
Percentage of search reports by applicants' origin (2012): ISA = EP



Percentage of search reports by applicants' origin (2012): ISA = ES

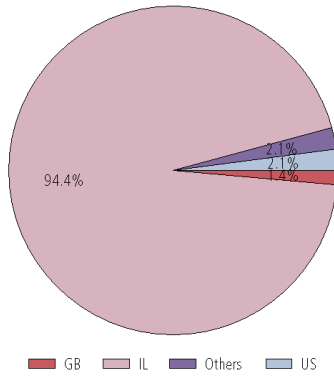


Percentage of search reports by applicants' origin (2012): ISA = FI

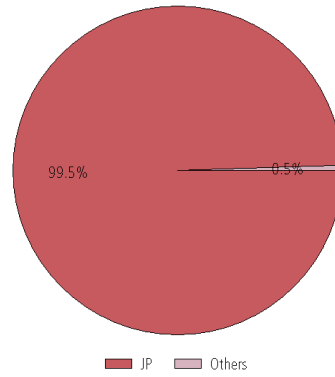


Annex I of Circular C. PCT 1434
page 25

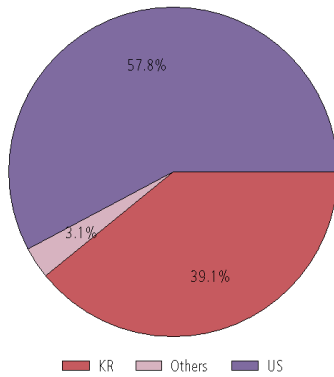
Percentage of search reports by applicants' origin (2012): ISA = IL



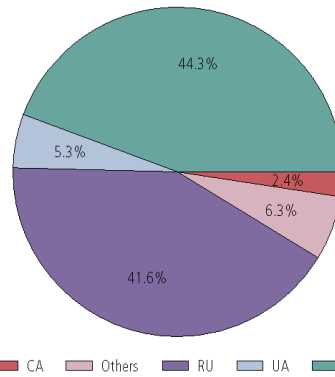
Percentage of search reports by applicants' origin (2012): ISA = JP



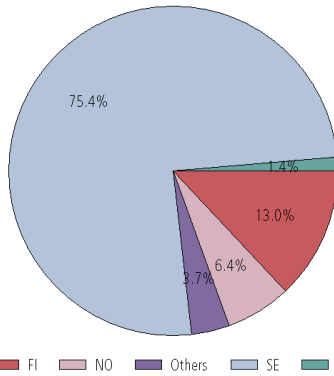
Percentage of search reports by applicants' origin (2012): ISA = KR



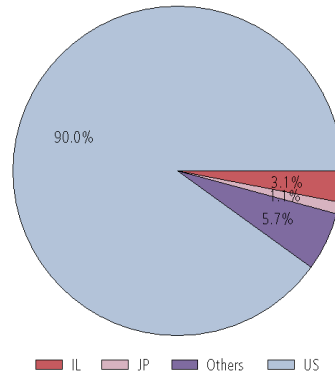
Percentage of search reports by applicants' origin (2012): ISA = RU



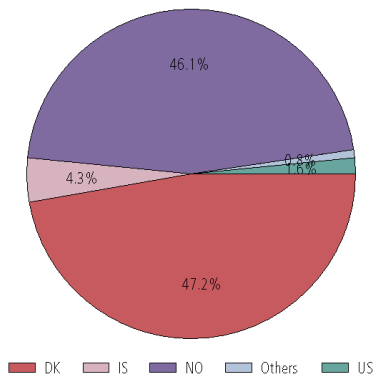
Percentage of search reports by applicants' origin (2012): ISA = SE



Percentage of search reports by applicants' origin (2012): ISA = US



Percentage of search reports by applicants' origin (2012): ISA = XN



[Annex II follows]

Descriptions and Definitions

DATA SET DESCRIPTION

- The data source is the European Patent Office's PATSTAT database for all citation information. Bibliographic information for international applications is taken mainly from the PATSTAT database, supplemented by information from WIPO internal databases where information could be provided which was not available from PATSTAT.
- The data provided is based on published PCT searches.
- Statistics are presented by search date up to 2012 Q4, meaning the date on which an international search report was transmitted to the International Bureau (since this information is available more consistently than the actual date of search).
- The date ranges for statistics take into account data availability. This is constrained by procedural latency such as time to publication, as well as cut-off dates for database extracts.
- No filing date constraint is applied.

DATA ISSUES

- Applications with no citation recorded are removed, as this generally means that no meaningful international search was carried out for these applications.
- A small number of patent citations are without category codes.
- In case of citn_origin = 5 (documents cited during international search), those citations are considered; otherwise, citations with citn_origin = 0 (documents cited during search) are considered. Citations with other citn_origin codes are removed.
- NPL citations with no category assigned and with ID >= 900000000 are removed, as they don't seem to be in the original search reports.
- All citation category codes recorded in the database for the valid citations are considered.
- Citation language codes for national patent documents are those recorded in the Patstat database, citation language codes for PCT documents are assigned using WIPO's PCT database as they are more reliable. The language codes are further cleaned up according to information of the authorities who publish those documents.
- No attempt has been made to determine the language of publication of non-patent literature documents.

DEFINITION OF CONCEPTS

Technology breakdown

- Technology sector and field are derived from the IPC classes assigned in the international phase search report or publication.
- The grouping into technology sector and field is based on a concordance provided by WIPO. (http://www.wipo.int/ipstats/en/statistics/patents/pdf/wipo_ipc_technology.pdf).
- This technology breakdown includes 8 technology sectors (Electrical engineering, Energy technology, Instruments, Mechanical engineering, Micro-structural and nano-technology, Other fields and Semiconductors), which are further broken down into 35 technology fields.

- Multiple IPC classes are often assigned to applications. For the present statistics, fractional counting method is applied, that is, an international application and all citations in its search report are evenly distributed to multiple technological fields when multiple fields are associated with it.
- IPC class information is not available for approximately 1% of applications.

Applicant origin

- In general this is the State in which the first-named applicant is resident (overall, this gives a more useful indication of origin of the application than the receiving Office because the International Bureau and regional Offices work for many States, whereas some States do not themselves operate a receiving Office).
- “Unknown” code is used for a small percentage of applications.

XY rate (Searches with XY citations)

- XY rate refers to share of search reports where at least one citation is in the category of X or Y.
- In addition the use of an E citation is counted as XY if it can be assumed that the E citation is prejudicial to novelty. This is the case unless the E category is assigned in combination with A.

Citation category availability

- PATSTAT does not contain all citation categories for each citation. The database contains one citation category per group of categories for each citation. The category groups are defined as follows:

Group 1	X Y A
Group 2	P E
Group 3	D
Group 4	O T L

- Only one category from the same group is selected. The category selected is determined from the order in the table above. In this way a citation will be categorized as X if the citation categories in the search report are XY for this citation. Priority for selecting the letter is according to the ranking of categories left to right within the groups above, rather than the order of their appearance within the citation in the international search report (that is, X will be shown even if the search report lists Y category claims first).
- A maximum of 3 categories is recorded.
- Citation Category Examples:

<i>Search report citation</i>	<i>Citation categories present in PATSTAT</i>
X, Y, A, P	X, P
Y, A, P, E	Y, P
Y, X, O, T	X, O
X, P, D, O	X, P, D

- This means for example in row 1 above the Y nor the A is not stored in PATSTAT.

- In practice it is therefore possible to determine whether a search has at least one X or Y citation. It is also possible to correctly count the number of X citations.
- In approximately 20% of cases it is not possible to correctly count the number of Y categories used, although it is possible to count the use of Y without an X.
- EPO data has been refined with an additional internal data source.

A only rate

- A-only rate refers to the share of search reports where no citation is in the category of X, Y or E.

Y no X rate

- Y no X rate refers to the share of search reports where at least one citation is in the category of Y and there is no X citation.

Search date

- The date when the search report is transmitted to WIPO (the actual date of search is not available in all cases).

Patent Literature/Non-Patent Literature

- Citations in PATSTAT are categorized into patent literature and non-patent literature.
- A citation is considered patent literature if it relates to patent abstracts provided by various providers.
- Less information is available for NPL citations. For example, the language of a NPL citation is not available.

Non official language

- This is used for counting patent citations that are not in an official language of the respective ISA:

<i>ISA</i>	<i>Official language</i>
AT	German
AU	English
BR	Portuguese
CA	English
CA	French
CN	Chinese
EP	German
EP	English
EP	French
ES	Spanish
FI	Finnish
JP	Japanese
KR	Korean
RU	Russian
SE	Swedish
US	English
XN	Danish
XN	Icelandic
XN	Norwegian

- The statistics are based on the actual official languages of the Office, but can easily be redefined to reflect any set of core languages which an Offices considers to be useful in assessing how effective its processes may be at discovering prior art beyond those languages.

Publication Authority (of citation)

- This is the patent organization who published a citation document.
- It is normally a national patent office, a regional office such as the EPO, or WIPO.

Processing Authority (of citation)

- Generally processing authority is assigned from the publication authority of the citation.
- For WO publications, the international search authority is chosen to indicate which office processed the cited patent publication. This gives an indication of the nature of the publication which will be more useful for some purposes than simply the number of WO citations, which may be in any of 10 languages.

[End of Annex II and of Circular]