Developing the tools for a true global search

The world's most complete IP data collection

LIGHTHOUSE IP

Contents

The search for global coverage

Patent information specialists face constant challenges in their task of identifying and utilising the global patent literature. Both commercial and public-sector database providers are under pressure to create search tools which are fit for an ever-expanding variety of purposes. Simply "adding more country coverage" may seem an obvious, and desirable, objective but it may not be practicable for many reasons. Historically, data providers have been guided by customer demand and reliability of supply as the major factors determining their coverage expansion policies. This paper presents an evidence-based method for assessing the potential value in new national data streams.

It has been generally accepted for some time that a very large proportion of global patent application activity is focussed on the so-called IP5 authorities, that is, the patent-granting agencies for the United States (USPTO), Japan (JPO), Republic of Korea (KIPO) and China (CNIPA) and the regional Office covering 39 member states in Western and Central Europe (the EPO). The WIPO IP Statistics Data Center (www.wipo.int/ipstats/en/) can provide an overview of this dominance, by extracting the numbers of applications filed at each national office over the last 20 years. The graph at ***Figure 1 (next page)*** clearly shows the proportion of total filings which these five authorities capture; the bottom five areas relate to filings received each calendar year at the IP5.

It is clear that these major publishing and granting authorities cannot be excluded from any database which hopes to present an accurate picture of patenting activity. But what about the rest of the world? If we exclude the IP5, and concentrate on the next ten ranked authorities, a more interesting picture emerges.

The next ten authorities between them account for some 350,000-400,000 applications per year, roughly the same as Japan alone.

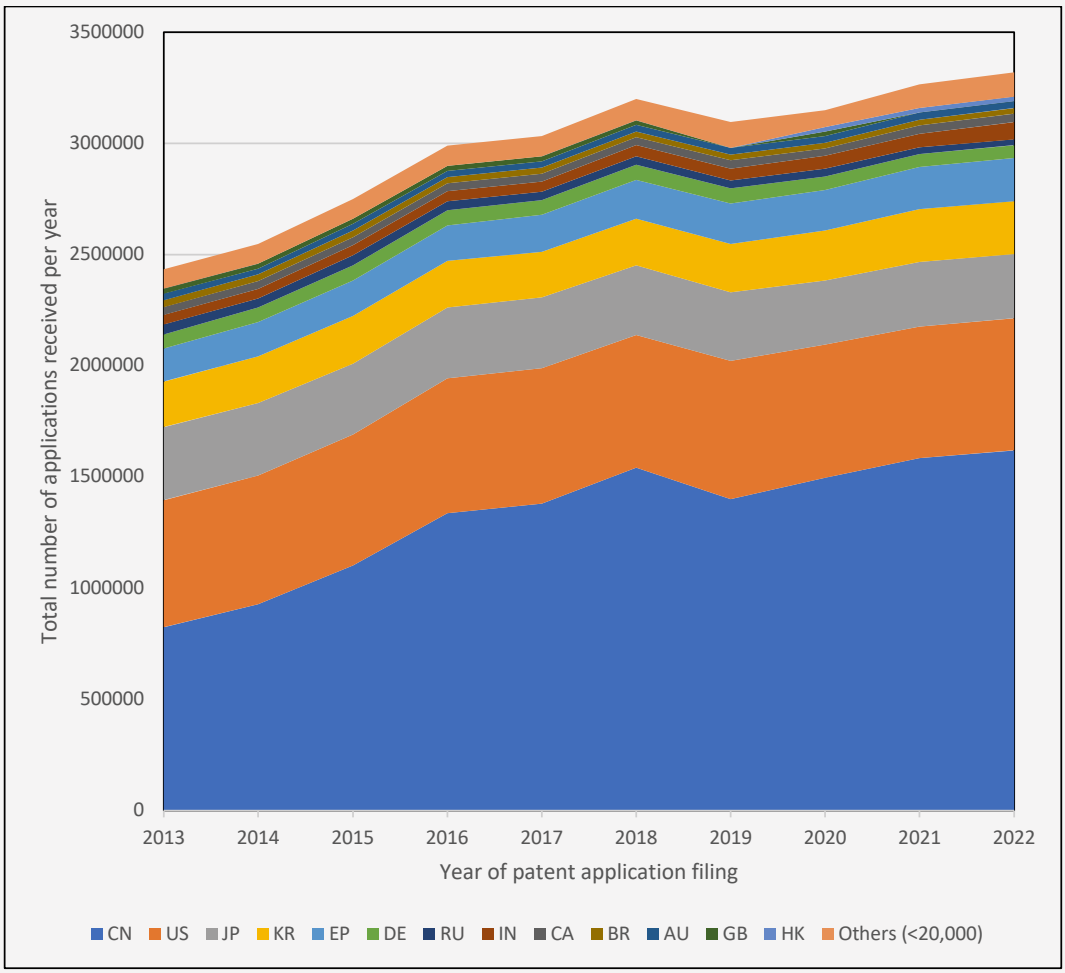

Figure 1: *Proportion of filings by year, 2013-2022*

If the statistics are presented as a normalised (100%) stacked column, it becomes easier to see that the proportion of filings lodged at each office has undergone some variation. ***Figure 2 (next page)*** shows the results; whilst some offices (e.g. Canada) represent a fairly stable 10% of applications received, India has increased from 12% to nearly 20% while Brazil has decreased from approximately 8.5% to 6%. These fluctuations can occur as a result of short- to medium-term economic factors, but a more serious concern for the database producer is whether there is any longer-term trend, representative of factors such as increased industrialisation,

domestic innovation or awareness of the importance of IP rights in a developing economy. If it is possible to predict which countries are likely to be the next "rising stars", it would help a producer to resource and to budget for a plan of controlled growth.

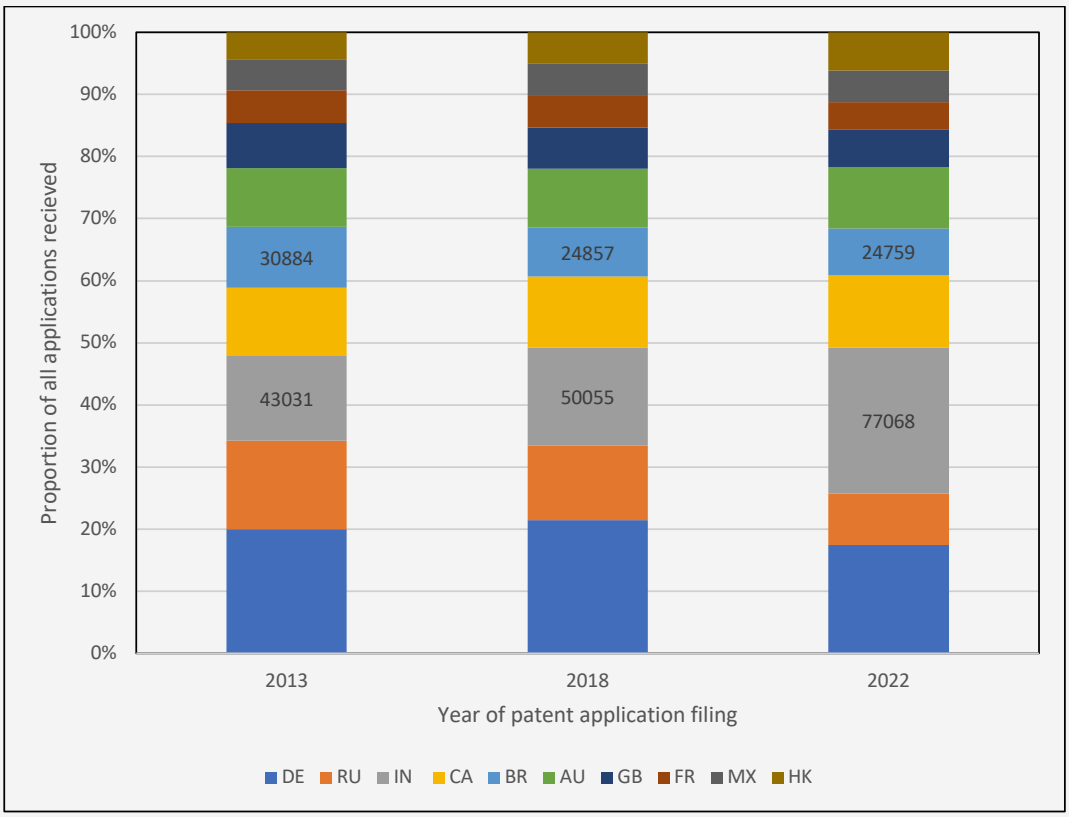

Figure 2: *Trends in volume of filing at major patent offices*

What is 'an active patent office'?

While it is true that consideration of numbers of applications received at an office can be of some assistance, the varied information needs within industry demand that we consider other factors as well. Although absolute or relative numbers of applications filed gives an indication of the perceived attractiveness of seeking IP rights from a given country, the legal status searcher is more interested in which rights are actually being **granted**. The patentability searcher, on the other hand, wishes to establish whether a new application is novel and inventive when assessed on the global scale; this is better measured by the rate of growth of the published state-of-the-art, notably in the form of **published unexamined** patent applications.

Even amongst the IP5, consideration of these three statistics (applications filed, patents granted and unexamined applications published) is instructive. In ***Table 1***, data for the latter value are derived from the EPO Espacenet Worldwide file; the remainder are from the WIPO IP Statistics Data Center.

	2022 applications (WIPO)	2022 grants (WIPO)	2022 published applications (DocDB)
CNIPA	1619268	798347	1738864
EPO	193610	81086	176599
JPO	289530	201420	322370
KIPO	237633	135180	227417
USPTO	594340	323410	402590
Totals		1539443	2867840
Ratio (published applications:grants)			**1.86:1**

Table 1: *IP5 contribution to the state-of-the-art*

The Table reminds us that, in broad terms, many patent offices are publishing two additions to the global state-of-the-art (public disclosures) for every one addition to the global database of substantive IP rights; the ratio averaged across all IP5 authorities is 1.86:1. Database producers are ultimately dependent upon the output (publishing streams) from patent offices as source material for both bibliographic and legal status databases. Unfortunately, WIPO statistics do not collect data on published national applications, and therefore can give no guidance as to whether a given office is contributing significant additions to the state-of-the-art by means of this route. Furthermore, we cannot analyse national publication data rates in a global context unless a country is already covered in a multi-country bibliographic database, which is not helpful if the objective of the research is to justify the effort of *starting to include* a new country.

The Global Innovation Index

In order to measure broader trends, we can turn to the Global Innovation Index (GII), produced in partnership with WIPO (globalinnovationindex.org/). Since 2010, the survey has included a consistent range of 130-140 countries, and the methodology has stabilised to focus on approximately 80 so-called 'indicators' of innovation capacity. To this extent, we can use the most recent 10-12 reports to detect some trends. This study considers two methods of reviewing the GII data:

a) Long-term improvements in overall GII ranking, and
b) Significant high scores in certain aspects of innovation outputs.

The background methodology of the GII is described in the Annexes of each report. The 2022 report states:

"The overall GII ranking is based on two sub-indices ...: the Innovation Input Sub-Index and the Innovation Output Sub-Index. ... Innovation outputs are the result of innovative activities within the economy. ...The overall GII score is the average of the Input and Output Sub-Indices, on which the GII economy rankings are produced."

The Output Sub-Index is calculated using a range of different indicators. Some of these relate specifically to the country's output of granted patents and scientific literature. ***Table 2 (next page)*** shows the definitions of the indicators of interest in this study, from the 2022 edition of the GII.

Long-term improvements in overall GII ranking

Each edition of the GII assigns an overall ranking to each economy/country surveyed. The authors of the report are at pains to emphasise that the absolute rank is not directly comparable from one survey to the next. However, each report does note instances of significant rise or fall in ranking and, to some extent, any long-term trends within groups of comparable economies or geographic regions.

GII Pillar, Section, Indicator	Title	Definition in GII 2022 report	Comment
6	Knowledge & Technology Outputs		One of two 'pillars' to the Output Sub-Index; the other is Creative Outputs.
6.1	Knowledge creation		Measurable statistics on how much the country is adding to the public disclosure of innovative activity.
6.1.1	Patents by origin (*)	Number of resident patent applications filed at a given national or regional patent office (2020 data)	Indicative of the use of the patent system by national applicants, irrespective of whether foreign patents are also sought (i.e. priority filings).
6.1.2	PCT applications by origin (*)	Number of Patent Cooperation Treaty (PCT) applications (2021 data)	Indicative of the use of the patent system by national applicants, in order to obtain protection in additional countries (i.e. equivalent filings).
6.1.3	Utility models by origin (*)	Number of resident utility model applications filed at the national patent office (2020 data)	Use of a second-tier protection system; more popular with SME's and national inventors seeking only local IP rights.
6.1.4	Scientific and technical articles (*)	Number of scientific and technical journal articles (2021 data)	Collated from 182 research areas, based on citations in Clarivate Web of Science (journals from the Science Citation Index Expanded (SCIE) and Social Sciences Citation Index (SSCI)).
6.1.5	Citable documents H-index	The H-index is the country's number of published articles (H) that have received at least H citations (2021 data)	H-index is tabulated from the number of citations received in subsequent years by articles published in a given year, divided by the number of articles published that year. It quantifies both journal scientific productivity and scientific impact.

(*) this indicator is normalised by dividing the basic count by the country's Gross Domestic Product, measured as purchasing power parity in billions of international dollars (bn PPP$ GDP).

Table 2: *GII 2022 definitions for Knowledge Creation indicators*

Figure 3 (next page) shows the GII rankings from 2011-2022 for the five countries which have shown the strongest upward trend in their overall position and have each succeeded in reaching the top 50% of the table (ranked 60 or above out of approximately 130 countries in the survey) at least once during the period. Their individual trend was assessed by computing the linear regression line which best fitted the plot of annual ranks; the absolute value of the slope of the line is indicative of the speed of rise up the table. Since low-numbered ranks are 'good' and high-numbered ranks are 'bad', all slopes are negative as countries rise up the rankings. The steepest overall riser is Iran (slope -5.60), while the remaining four countries (Philippines, Vietnam, India and Mexico) all achieved a slope of -2.0 or greater.

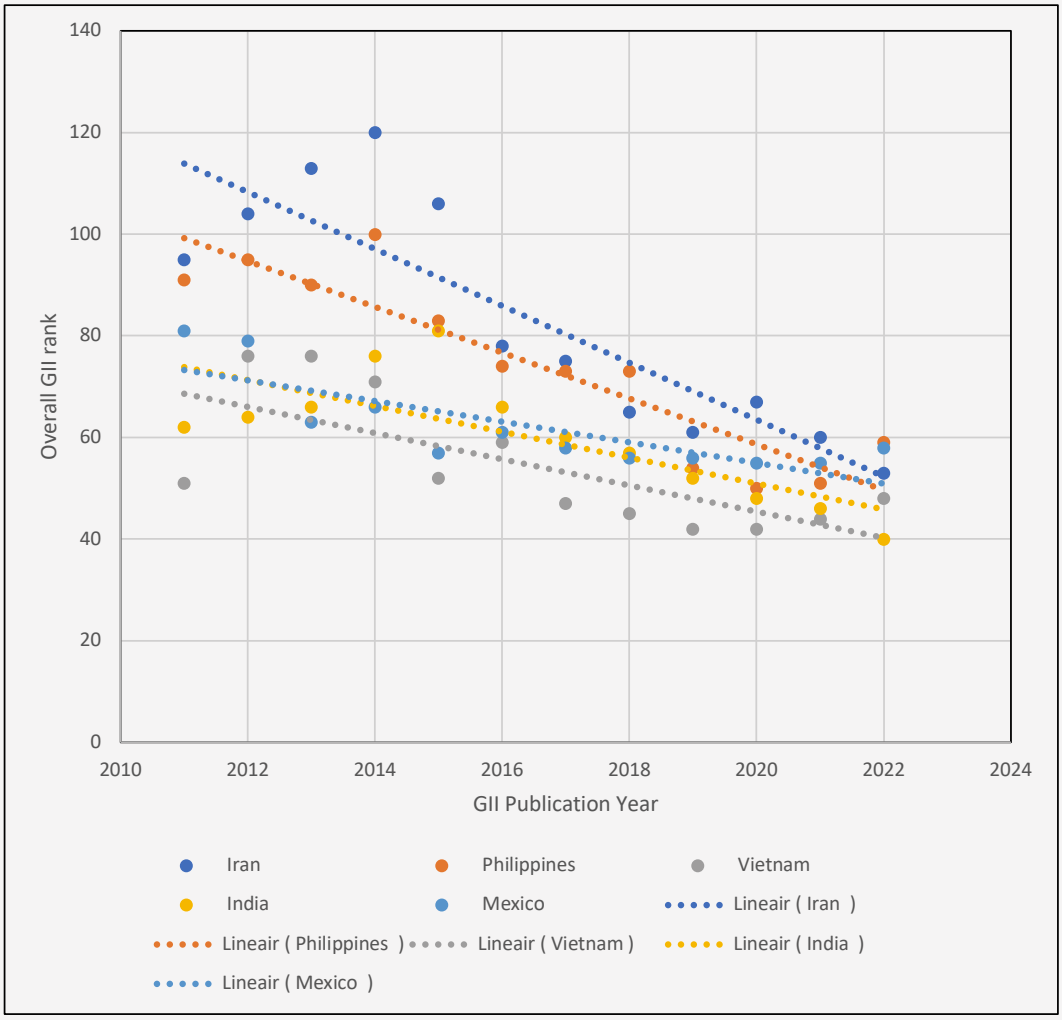

Figure 3: *Long-term improvers in GII score*

Innovation Output scores

As noted above, the overall ranking in GII is a composite of two separate indices, which represent Inputs (broadly, the factors which make an economy a good environment in which to conduct research & development) and Outputs (broadly, factors representing the fruits of innovative activity, including the benefits accruing due to intellectual property rights). Although the two indices are clearly

related, it is reasonable to suppose that there will be a time lag between improvements in the R&D standing of a country and any publicly-disclosed advances, such as published results or IP filings. Hence, if the Output Index rank is substantially higher than the overall ranking, it may be an early indicator that the country is beginning to benefit from investment in local innovation activity, and becoming a net contributor to the public state-of-the-art by disclosing home-grown research, rather than a consumer of foreign technology. More detail of the indicators which make up the Knowledge Creation section of the survey are shown in ***Table 2 (page 9)***.

The first step in the analysis consisted of identifying the countries with the largest disparity between their overall GII ranking and their corresponding Innovation Output rank. The four national authorities from the IP5 and all of the member, extension or validation states of the EPO were excluded from this analysis, in order to concentrate on countries which are currently less well covered in patent databases. Any countries with an incomplete data record over the past 10 years were also excluded. This yielded a list of 88 countries, with deviations in score ranging from +21 (Madagascar, overall rank 106, output sub-index rank 85) to -37 (Brunei Darussalam, overall rank 92, output sub-index rank 129).

To refine this set, the WIPO IP Statistics Data Center values for patents granted in each of the 88 countries was extracted (where available); the mean number of grants was just over 2,000 in 2022. Filtering the initial list to exclude any country with fewer than 1,000 patent grants reduced the list to 23 countries, of which the top 12 (those with ranking deviation score of 0 or greater) are shown in ***Table 3 (page 12)***. Many of these countries are already included in major patent databases, with the notable exception of Iran, which scores particularly high on the "deviation index". Interestingly, the five countries already identified as the long-term improvers from ***Figure 3 (page 10)*** are also present in ***Table 3 (page 12)***.

A second filter on the main results is also of interest. Reducing the cut-off point further, from 1000 to 100 or more patents granted in 2022, while maintaining the criterion of a positive "deviation index" (greater than zero) yields a selection of five smaller countries which can be considered as promising candidates to grow into the high fliers of tomorrow. This list is shown in ***Table 4.***
For completeness, the breakdown of scores in section 6.1 of the GII are collated for the countries from ***Table 3*** and ***Table 4***, at ***Table 5 (page 13).***

WIPO ST.3 Country Code	Country	Overall GII rank	Output sub-index rank	Deviation (Overall rank minus Output rank)	Patents granted (2022)
IR	*Iran*	53	38	15	2250
UA	Ukraine	57	48	9	1566
PH	*Philippines*	59	51	8	2004
VN	*Vietnam*	48	41	7	3868
AR	Argentina	69	62	7	1949
NG	Nigeria	114	107	7	1081
MX	*Mexico*	58	55	3	9698
IN	*India*	40	39	1	30490
BR	Brazil	54	53	1	23546
ID	Indonesia	75	74	1	9970
ZA	South Africa	61	61	0	11267
IL	Israel	16	16	0	5358

Note*: countries in italics featured as significant improvers, in the analysis at Figure 3*

Table 3: *GII ranked 'net exporters' and patenting authorities (>1000 grants)*

WIPO ST.3 Country Code	Country	Overall GII rank	Output sub-index rank	Deviation (Overall rank minus Output rank)	Patents granted (2022)
PK	Pakistan	87	69	18	189
LK	Sri Lanka	85	68	17	186
BY	Belarus	77	63	14	302
MN	Mongolia	71	64	7	118
EG	Egypt	89	83	6	495

Table 4: *GII ranked 'net exporters' and patenting authorities (100-999 grants)*

WIPO ST.3 Country Code	Country	Overall GII rank (2022)	Output Sub-Index rank	K & T Output rank (6)	Knowledge creation rank (6.1)	6.1.1	6.1.2	6.1.3	6.1.4	6.1.5
AR	Argentina	69	62	77	62	65	-	52	80	36
BR	Brazil	54	53	55	48	43	54	26	51	23
BY	Belarus	77	63	40	60	32	63	16	104	75
EG	Egypt	89	83	79	64	70	80	-	48	47
ID	Indonesia	75	74	78	92	80	100	30	128	57
IL	Israel	16	16	7	12	20	1	-	20	16
IN	India	40	39	34	46	28	52	-	78	21
IR	Iran	53	38	50	20	10	40	-	15	39
LK	Sri Lanka	85	68	66	88	60	59	-	110	69
MN	Mongolia	71	64	85	32	44	88	1	72	105
MX	Mexico	58	55	58	73	77	72	43	101	34
NG	Nigeria	114	107	123	99	81	99	-	107	61
PH	Philippines	59	51	41	69	75	76	15	122	55
PK	Pakistan	87	69	70	54 (*)	87	-	-	40	46
UA	Ukraine	57	48	36	29	29	49	1	97	51
VN	Vietnam	48	41	52	84	66	85	37	90	58
ZA	South Africa	61	61	56	52	72	44	-	45	31

(*) Data Minimum Coverage (DMC) requirements not met at this level of the analysis

Table 5: *Selected country rankings from the Knowledge Creation section of GII*

Notes/Key:

Pillar 6	Knowledge & Technology Outputs
Section 6.1	Knowledge creation
Indicator 6.1.1	Patents by origin/bn PPP$ GDP
Indicator 6.1.2	PCT applications by origin/bn PPP$ GDP
Indicator 6.1.3	Utility models by origin/bn PPP$ GDP
Indicator 6.1.4	Scientific and technical articles/bn PPP$ GDP
Indicator 6.1.5	Citable documents H-index

	Country strength (within income group)
	Country strength

	Second quartile rank
	First quartile rank

Conclusions

Combining data on IP publication rates with the broader measures from the GII survey has identified seventeen countries which have improved their standing as innovative R&D environments over the last decade and are showing signs of being net contributors to the state-of-the-art. Many of them show significant local or international strength in their Knowledge Creation indicators. This kind of evidence may assist database producers to make systematic plans for future country coverage expansions. On the searcher side, it illustrates the need for information specialists to be constantly aware of developments in the content of both commercial and public-sector search tools. One example is Diamond File from Lighthouse IP, a database which has sought from the beginning to establish both consistent and widespread coverage across the largest number of issuing authorities worldwide and currently makes available content from many of these identified so-called "minor sources".

About the author

Stephen Adams is the managing director of Magister Ltd., a UK-based consultancy specialising in patents information. Mr. Adams is a Qualified Patent Information Professional (number 20190044100092) and holds a B.Sc. in chemistry from the University of Bristol and an M.Sc. in Information Science from City University, London, as well as professional memberships of the Royal Society of Chemistry (RSC) and the UK's Chartered Institute of Library and Information Professionals (CILIP). He is the author of three editions of "Information Sources in Patents", the latest published in 2020 by Walter de Gruyter KG, contributed several book chapters and written numerous articles in the field of patent information, including over 25 refereed papers for the Elsevier journal "World Patent Information". His professional service includes the Editorial Advisory Board of "World Patent Information" between 2006-2020 and three terms on the Board of PIUG Inc., the International Society for Patent Information, as Director-at-Large (2002-2006) and Vice-Chair (2014-2016 and 2016-2018), as well as service on the management committee of the UK's Patent and Trade Mark Group over many years. He received the PIUG's Special Recognition Award in 2008 and the IPI Award in 2012 for outstanding contribution to patent information.

Tupolevlaan 81, 1119 PA Schiphol-Rijk
The Netherlands
Telephone +31 85 800 0024
info@lighthouseip.com

lighthouseip.com