

Appendix IV Global Innovation Index science and technology cluster methodology

Since 2016 the Global Innovation Index (GII) has sought to identify Science and Technology (S&T) clusters using a bottom-up approach. This approach disregards administrative or political borders and instead pinpoints those geographical areas showing a high density of inventors and scientific authors. The resultant clusters often encompass several municipal districts, sub-federal states, and sometimes even two or more countries.

The same methodology used in previous editions of the GI was employed in the compilation of this year's list of the top 100 GI S&T clusters worldwide (Bergquist and Fink, 2020: 43–63). It comprised:

- selecting inventors listed in published patent applications under WIPO's Patent Cooperation Treaty (PCT) spanning the period 2015 to 2019;
- selecting authors listed in scientific publications in the Web of Science's Science Citation Index Expanded (SCIE) covering the same period;
- geocoding inventor and author addresses and then applying the density-based spatial clustering of applications with noise (DBSCAN) algorithm to the geocoded inventor and author points.

The WIPO PCT patent dataset consists of approximately 1.1 million patent applications published between 2015 and 2019 containing 3.2 million inventor addresses. For the SCIE, the dataset comprises 9.1 million articles published during the same period containing 27.7 million listed author addresses.

The geocoding of addresses for this report is as follows. PCT inventor addresses were geocoded using the Environmental Systems Research Institute (ESRI) ArcGIS World Geocoder service.¹ When the ESRI address matches proved either insufficiently accurate or ambiguous, the city name in the address string was extracted and matched using records in the city level dataset from the GeoNames Gazetteer database.² This latter database gives the geolocation of cities around the globe and contains 48,000 geocoded cities. This same city matching approach was applied to all SCIE author addresses.

Overall, 96.4% of inventor addresses were geocoded at either the city level or a more accurate level, while 95.5% of scientific author addresses were geocoded at the city level. Annex Table 5 provides a summary of the geocoding results for the top 20 countries, which together account for the majority of inventor and scientific author addresses. As shown in the table, the coverage of geocoded addresses across all 20 countries is typically above 95%, only falling below 90% in one instance.

Addresses were clustered by applying the DBSCAN algorithm. This algorithm requires pre-defined radius and density parameters. As in previous years, a radius of 15 km and a density of 4,500 was applied. Equal weight was given to inventors and authors by expressing data points as a share of total inventor and author addresses, respectively. Given that the number of scientific articles far exceeds the number of patents, cluster identification based on the raw data points would have resulted in clusters shaped predominantly by the scientific author landscape.

The result was an initial list of 227 clusters. After review, neighboring clusters were merged if the edge of a cluster was within 3–5 km of another and where the co-author/co-inventor relationships were higher than they were for any other relationship with any other cluster or non-cluster points. A total of 22 clusters met these criteria, mergers reducing the overall number of clusters identified to 216.³

The remaining 216 clusters were then put into rank by counting the number of patents and scientific articles in a given cluster. Numbers were aggregated utilizing fractional counting, where counts reflect the share of a patent's inventors and an article's authors present in a particular cluster. In addition, mirroring the equal weighting approach described above, fractional counts are relative to the total numbers of patents and scientific articles.

To produce an intensity ranking, the European Commission's Global Human Settlement Layer (GHSL) population distribution data were matched geographically to the top 100 clusters identified in the overall ranking. Just as with inventor/author geocoded locations, this population data allowed us to define the total population of a cluster using a bottom-up approach. We chose to delimit a cluster's area as being all the space within 0.05 degrees of each inventor/author location. Overlaying the resultant cluster polygons on top of the population data and aggregating all points which lay within the polygon gave a total population estimate for each cluster.⁴ The clusters were then ranked by dividing the total S&T share by population.

Annex Table 3
Top 100 clusters, 2021

Cluster rank	Cluster name	Economy	Share of total PCT filings (%)	Share of total publications (%)	Total	Rank change
1	Tokyo–Yokohama	JP	10.78	1.61	12.40	0
2	Shenzhen–Hong Kong–Guangzhou	CN/HK	7.79	1.51	9.30	0
3	Beijing	CN	2.62	2.95	5.57	1
4	Seoul	KR	3.93	1.61	5.54	–1
5	San Jose–San Francisco, CA	US	3.69	1.03	4.72	0
6	Osaka–Kobe–Kyoto	JP	2.88	0.72	3.60	0
7	Boston–Cambridge, MA	US	1.44	1.47	2.91	0
8	Shanghai	CN	1.36	1.49	2.85	1
9	New York City, NY	US	1.11	1.54	2.66	–1
10	Paris	FR	1.26	1.02	2.28	0
11	San Diego, CA	US	1.77	0.38	2.15	0
12	Nagoya	JP	1.74	0.24	1.99	0
13	Washington, DC–Baltimore, MD	US	0.43	1.44	1.86	0
14	Los Angeles, CA	US	0.89	0.78	1.67	0
15	London	GB	0.42	1.21	1.63	0
16	Houston, TX	US	0.96	0.51	1.46	0
17	Seattle, WA	US	1.05	0.38	1.42	0
18	Nanjing	CN	0.21	1.07	1.28	3
19	Amsterdam–Rotterdam	NL	0.40	0.88	1.28	–1
20	Cologne	DE	0.73	0.53	1.26	–1
21	Hangzhou	CN	0.60	0.60	1.20	4
22	Daejeon	KR	0.87	0.29	1.16	0
23	Chicago, IL	US	0.50	0.64	1.14	–3
24	Munich	DE	0.74	0.36	1.09	–1
25	Wuhan	CN	0.24	0.82	1.05	4
26	Stuttgart	DE	0.82	0.21	1.03	0
27	Tel Aviv–Jerusalem	IL	0.66	0.35	1.01	–3
28	Taipei–Hsinchu	TW	0.29	0.69	0.97	–1
29	Singapore	SG	0.38	0.52	0.90	–1
30	Philadelphia, PA	US	0.31	0.58	0.89	1
31	Melbourne	AU	0.19	0.69	0.87	4
32	Moscow	RU	0.18	0.68	0.86	0
33	Xi'an	CN	0.08	0.77	0.86	7
34	Minneapolis, MN	US	0.58	0.27	0.85	–4
35	Stockholm	SE	0.54	0.31	0.84	–2
36	Eindhoven	BE/NL	0.76	0.07	0.83	–2
37	Sydney	AU	0.23	0.58	0.81	0
38	Raleigh, NC	US	0.27	0.54	0.80	–2
39	Chengdu	CN	0.15	0.62	0.77	8
40	Toronto, ON	CA	0.22	0.54	0.76	–1
41	Tehran	IR	0.02	0.74	0.75	2
42	Frankfurt Am Main	DE	0.47	0.28	0.75	–4
43	Brussels	BE	0.30	0.44	0.73	–2
44	Portland, OR	US	0.58	0.14	0.72	–2
45	Berlin	DE	0.31	0.40	0.71	–1
46	Madrid	ES	0.13	0.58	0.71	–1
47	Barcelona	ES	0.22	0.49	0.71	–1
48	Milan	IT	0.21	0.44	0.65	0
49	Istanbul	TR	0.28	0.36	0.64	2
50	Zürich	CH/DE	0.29	0.34	0.63	–1
51	Denver, CO	US	0.24	0.37	0.61	–1
52	Tianjin	CN	0.08	0.53	0.61	4
53	Qingdao	CN	0.28	0.32	0.60	16
54	Montréal, QC	CA	0.19	0.41	0.60	–2
55	Heidelberg–Mannheim	DE	0.36	0.23	0.59	–2
56	Copenhagen	DK	0.28	0.30	0.59	–2
57	Atlanta, GA	US	0.16	0.40	0.56	–2
58	Cambridge	GB	0.26	0.29	0.55	–1
59	Changsha	CN	0.06	0.48	0.54	7
60	Rome	IT	0.08	0.45	0.53	–2
61	Cincinnati, OH	US	0.37	0.15	0.52	–2
62	Bengaluru	IN	0.32	0.20	0.52	–2
63	Suzhou	CN	0.33	0.18	0.51	9
64	Delhi	IN	0.09	0.41	0.50	3
65	Dallas, TX	US	0.29	0.20	0.49	–3
66	São Paulo	BR	0.07	0.41	0.48	–5
67	Pittsburgh, PA	US	0.15	0.33	0.48	–3
68	Nuremberg–Erlangen	DE	0.33	0.14	0.47	–5
69	Chongqing	CN	0.09	0.38	0.47	8
70	Ann Arbor, MI	US	0.12	0.35	0.47	–5
71	Vienna	AT	0.14	0.30	0.44	–1
72	Oxford	GB	0.14	0.31	0.44	–1
73	Hefei	CN	0.07	0.37	0.44	6
74	Helsinki	FI	0.25	0.19	0.44	–6
75	Harbin	CN	0.02	0.40	0.42	5
76	Jinan	CN	0.07	0.34	0.41	6
77	Vancouver, BC	CA	0.13	0.27	0.41	–3
78	Lyon	FR	0.22	0.19	0.41	–2
79	Busan	KR	0.20	0.20	0.40	–4
80	Cleveland, OH	US	0.12	0.27	0.39	–7
81	Changchun	CN	0.02	0.37	0.39	6
82	Phoenix, AZ	US	0.23	0.16	0.39	–4
83	Hamamatsu	JP	0.33	0.04	0.37	2
84	Kanazawa	JP	0.32	0.05	0.37	7
85	Ottawa, ON	CA	0.18	0.19	0.37	–4
86	Brisbane	AU	0.11	0.25	0.36	–3
87	Bridgeport–New Haven, CT	US	0.12	0.24	0.36	–3
88	Austin, TX	US	0.20	0.15	0.35	–2
89	Ankara	TR	0.04	0.30	0.35	–1
90	Shenyang	CN	0.04	0.30	0.34	14
91	Hamburg	DE	0.17	0.17	0.34	–1
92	Lausanne	CH/FR	0.17	0.17	0.34	–3
93	Mumbai	IN	0.13	0.21	0.34	5
94	Lund–Malmö	SE	0.20	0.13	0.33	2
95	Manchester	GB	0.09	0.23	0.32	–2
96	St. Louis, MO	US	0.09	0.23	0.32	–2
97	Dalian	CN	0.06	0.26	0.32	13
98	Daegu	KR	0.16	0.16	0.32	3
99	Göteborg	SE	0.18	0.14	0.32	1
100	Warsaw	PL	0.04	0.28	0.32	–1

Source: WIPO Statistics Database, April 2021

Annex Table 4
Ranking of S&T intensity, 2015–2019

Intensity rank	Cluster name	Economy	PCT applications per capita ^a	Scientific publications per capita ^a	Total S&T share per capita ^a	Rank change
1	Cambridge	GB	6,051	54,840	1.27	0
2	Eindhoven	BE/NL	8,274	6,116	0.81	1
3	Ann Arbor, MI	US	2,137	49,399	0.80	2
4	Oxford	GB	2,899	54,032	0.79	-2
5	San Jose–San Francisco, CA	US	6,595	15,217	0.77	-1
6	Daejeon	KR	5,752	15,903	0.73	1
7	Boston–Cambridge, MA	US	3,898	32,690	0.72	-1
8	Seattle, WA	US	4,846	14,432	0.60	0
9	San Diego, CA	US	5,314	9,380	0.58	0
10	Raleigh, NC	US	1,850	30,887	0.52	1
11	Lund–Malmö	SE	3,551	19,940	0.50	-1
12	Kanazawa	JP	4,022	5,241	0.47	5
13	Munich	DE	3,210	12,759	0.44	2
14	Lausanne	CH/FR	2,756	21,535	0.44	-1
15	Stockholm	SE	3,042	14,369	0.42	-1
16	Göteborg	SE	2,425	16,374	0.38	0
17	Nuremberg–Erlangen	DE	2,762	9,619	0.38	2
18	Copenhagen	DK	1,929	17,279	0.38	2
19	Bridgeport–New Haven, CT	US	1,160	19,079	0.36	9
20	Pittsburgh, PA	US	1,146	21,186	0.36	2
21	Tokyo–Yokohama	JP	3,232	3,996	0.34	5
22	Portland, OR	US	3,031	6,022	0.34	-1
23	Helsinki	FI	2,240	14,230	0.33	-5
24	Ottawa, ON	CA	1,581	14,097	0.33	5
25	Zürich	CH/DE	1,710	16,534	0.33	-1
26	Stuttgart	DE	2,905	6,066	0.33	1
27	Hamamatsu	JP	2,891	2,780	0.32	5
28	Minneapolis, MN	US	2,462	9,426	0.31	-5
29	Washington, DC–Baltimore, MD	US	748	20,741	0.31	6
30	Heidelberg–Mannheim	DE	1,980	10,513	0.31	0
31	Cleveland, OH	US	958	17,401	0.29	2
32	Houston, TX	US	1,973	8,679	0.29	-1
33	Beijing	CN	1,442	13,441	0.29	3
34	Cincinnati, OH	US	2,227	7,612	0.28	0
35	Seoul	KR	1,920	6,502	0.25	2
36	Atlanta, GA	US	667	14,332	0.24	6
37	Nagoya	JP	2,162	2,513	0.23	2
38	Melbourne	AU	515	15,468	0.23	13
39	Sydney	AU	710	14,631	0.23	2
40	Osaka–Kobe–Kyoto	JP	1,956	4,037	0.23	7
41	Frankfurt Am Main	DE	1,439	7,006	0.22	8
42	St. Louis, MO	US	714	15,481	0.22	-2
43	Philadelphia, PA	US	806	12,710	0.22	5
44	Lyon	FR	1,305	9,074	0.22	2
45	Vancouver, BC	CA	776	13,157	0.22	-1
46	Denver, CO	US	932	11,651	0.21	-3
47	Brisbane	AU	611	11,857	0.21	8
48	Paris	FR	1,241	8,323	0.21	4
49	Chicago, IL	US	1,003	10,678	0.21	1
50	Austin, TX	US	1,443	8,939	0.20	-12
51	Shenzhen–Hong Kong–Guangzhou	CN/HK	1,759	2,818	0.19	6
52	Amsterdam–Rotterdam	NL	643	11,700	0.19	2
53	Nanjing	CN	320	13,467	0.18	13
54	Toronto, ON	CA	529	11,038	0.18	8
55	Berlin	DE	870	9,124	0.18	1
56	Vienna	AT	675	12,195	0.18	-3
57	Montréal, QC	CA	599	10,774	0.18	3
58	London	GB	499	11,827	0.18	0
59	New York City, NY	US	777	8,907	0.17	2
60	Brussels	BE	783	9,549	0.17	-1
61	Hangzhou	CN	907	7,524	0.17	7
62	Milan	IT	537	9,324	0.16	5
63	Barcelona	ES	549	9,970	0.16	1
64	Tel Aviv–Jerusalem	IL	1,130	4,980	0.16	-1
65	Rome	IT	248	12,266	0.15	0
66	Xi'an	CN	152	11,490	0.15	11
67	Los Angeles, CA	US	810	5,887	0.14	3
68	Cologne	DE	874	5,215	0.14	4
69	Phoenix, AZ	US	904	5,005	0.14	2
70	Qingdao	CN	691	6,541	0.14	14
71	Wuhan	CN	317	8,991	0.14	10
72	Dallas, TX	US	844	4,749	0.13	1
73	Changsha	CN	158	11,127	0.13	5
74	Singapore	SG	587	6,557	0.13	0
75	Hamburg	DE	780	6,471	0.13	-6
76	Madrid	ES	260	9,245	0.13	-1
77	Warsaw	PL	177	10,150	0.12	-1
78	Daegu	KR	690	5,622	0.12	n.a.
79	Changchun	CN	70	9,587	0.12	4
80	Tehran	IR	28	9,414	0.11	5
81	Shanghai	CN	595	5,388	0.11	1
82	Busan	KR	612	5,120	0.11	-3
83	Jinan	CN	205	8,349	0.11	3
84	Manchester	GB	340	7,375	0.11	-4
85	Harbin	CN	41	8,451	0.09	4
86	Hefei	CN	171	7,776	0.09	1
87	Taipei–Hsinchu	TW	288	5,731	0.09	1
88	Dalian	CN	203	6,895	0.09	n.a.
89	Chongqing	CN	166	6,098	0.09	4
90	Chengdu	CN	165	5,812	0.08	4
91	Suzhou	CN	594	2,771	0.08	0
92	Tianjin	CN	110	6,018	0.08	0
93	Moscow	RU	147	4,591	0.07	2
94	Ankara	TR	108	6,088	0.07	-4
95	Shenyang	CN	81	5,042	0.06	n.a.
96	Bengaluru	IN	288	1,469	0.04	1
97	Istanbul	TR	205	2,210	0.04	-1
98	São Paulo	BR	41	2,006	0.03	0
99	Delhi	IN	39	1,506	0.02	0
100	Mumbai	IN	68	942	0.01	0

Source: WIPO Statistics Database, April 2021.

Notes: ^a Per capita figures refer to 1,000,000 of population. n.a. indicates not applicable.

Annex Table 5
Summary of geocoding results

Country	Scientific publications			PCT applications				
	Number of addresses	City-level address accuracy (%)	Publications covered (%)	Number of addresses	Block-level address accuracy (%)	Sub-city level address accuracy (%)	City-level address accuracy (%)	Applications covered (%)
United States of America	6,182,602	96.88	98.16	854,454	94.42	5.29	0.14	99.87
China	4,055,364	98.86	99.40	552,389	86.81	0.06	8.53	95.47
Japan	1,155,048	92.06	95.38	566,043	31.60	27.42	39.11	98.51
Germany	1,324,151	97.36	98.19	262,762	97.45	0.50	1.70	99.81
Republic of Korea	765,479	94.63	96.95	231,499	0.08	0.96	79.62	87.33
United Kingdom	1,347,330	96.64	97.74	81,471	69.54	20.72	8.27	98.61
France	1,068,353	92.93	95.09	107,038	88.02	1.65	6.08	96.67
Italy	1,053,749	95.60	97.05	41,973	89.28	5.09	4.83	99.30
India	692,442	91.19	93.66	39,998	33.29	48.56	16.28	98.47
Canada	854,790	98.37	98.99	41,732	96.80	2.56	0.50	99.79
Spain	804,686	96.84	98.07	26,229	77.23	10.76	11.22	99.40
Australia	815,110	85.97	89.98	20,479	92	4.98	2.37	99.46
Netherlands	494,358	97.38	98.50	50,950	85.84	0.34	13.53	99.73
Brazil	614,712	98.60	99.55	9,423	83.13	11.50	4.76	99.65
Sweden	287,747	97.63	98.18	42,930	94.30	0.80	4.52	99.68
Russian Federation	370,048	98.96	99.24	14,083	88.35	5.28	5.25	99.50
Switzerland	318,693	90.68	92.40	36,586	90.90	2.36	3.60	97.92
Turkey	376,436	96.35	96.71	14,422	38.02	47.74	11.51	97.55
Iran (Islamic Republic of)	396,857	97.15	98.35	774	0.39	2.58	92.51	94.68
Israel	152,955	91.04	95.38	29,351	58.76	3.32	29.55	95.78

Source: WIPO Statistics Database, April 2021.

Note: Listed are the top 20 countries with the highest combined shares of scientific articles and patents. PCT inventor addresses were geocoded to the highest level of detail. Due to the far larger volume of scientific author addresses, these were geocoded only to city level. DEA is Data Envelopment Analysis.

Notes

- 1 ESRI ArcGIS World Geocoder service. <https://www.esri.com/en-us/arcgis/products/arcgis-world-geocoder>.
- 2 GeoNames. <http://geonames.org>.
- 3 The mergers were: Guangzhou with Shenzhen–Hong Kong; Hsinchu with Taipei; Matsudo with Tokyo–Yokohama; Jureselem with Tel Aviv; Istanbul Europe with Istanbul Asia; Rotterdam with Amsterdam; Irvine with Los Angeles; Boulder with Denver; Worcester with Boston–Cambridge; Dortmund with Cologne; Baltimore with Washington DC.
- 4 See Bergquist and Fink (2020: 61–63) for a more detailed description of how population data was matched to clusters: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2020.pdf.

Reference

K. Bergquist and C. Fink (2020). The top 100 science and technology clusters. In Dutta, S., B. Lanvin and S. Wunsch-Vincent (eds), *The Global Innovation Index 2020: Who Will Finance Innovation?* Ithaca, NY, Fontainebleau, and Geneva: Cornell University, INSEAD, and WIPO.