Enabling Innovation Measurement at the Sub-National Level

A WIPO toolkit

WIPO

Enabling Innovation Measurement at the Sub-National Level

A WIPO toolkit

Gaétan de Rassenfosse
École polytechnique fédérale de Lausanne

Sacha Wunsch-Vincent
IP and Innovation Ecosystems Sector
World Intellectual Property Organization

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

Acknowledgments

This background study has been carried out as part of the World Intellectual Property Organization's (WIPO) Intellectual Property (IP) and Innovation Ecosystems Sector (IES) work to support member states with development of their own innovation indices. For inputs relating to national projects, thanks go to Mr. Camilo Rivera, National Planning Department, Colombia; Mr. Ashok Sonkusare, National Institution for Transforming India Aayog, India; Ms. Sun Yunjie, Chinese Academy of Science and Technology for Development, China; Mr. Minh Hoang and Mr. Hung Vo Nguyen, Viet Nam Institute for Science, Technology and Innovation, Ministry of Science and Technology, Viet Nam; and Mr. Marcos Pinto, Ministry of Science, Technology and Innovation, Brazil. Thanks go to WIPO's Regional and National Development Sector (RNDS) for support, in particular Ms. Yaning Zhang, and to the Division for Asia and the Pacific and Mr. Andrew Ong.

Introduction

Published annually since 2007, the Global Innovation Index (GII) is a leading benchmarking tool for business executives and policymakers seeking insight into the state of innovation around the world. The core of the GII is a ranking of world economies' innovation capabilities and results.

Not only is innovation a matter of national innovation policy, but it also concerns all levels of political decision-making, including the city, regional and provincial levels. Next to the GII, interest is growing among member states to develop complementary and mutually reinforcing sub-national innovation indices. China, Colombia and India already publish their own indices, drawing on the GII framework. Other countries are following suit.

In particular, during a visit to the World Intellectual Property Organization (WIPO) in November 2021, the President of Viet Nam requested the support of WIPO in the development of Viet Nam's own provincial index. To respond to the demand of member states in this field, WIPO is proceeding in two ways:

1. organizing workshops on the exchange of best practices and
2. providing this background study on sub-national innovation indices.

Pursuant to point 1, in June 2022 a workshop took place assessing (i) the main rationales of having sub-national innovation indices, including the interaction with the GII; (ii) the technical approach used, including the index structure, the variables that are identical or different to the GII, the data collection methods and the aggregation techniques; (iii) the results and impacts of the said indices and (iv) related next steps (Annex 1).[1]

Pursuant to point 2, this background study addresses three points:

- First, it reviews the applicability of the GII framework to the development of sub-national innovation metrics and indices. The key question relating to this point is: Which of the GII metrics are available at the city, regional, provincial or other levels?
- Second, this study reviews existing sub-national innovation indices of WIPO member states who have pioneered this field, notably China, Colombia, the European Union (EU), India and Viet Nam (Annex 2). Key questions are: How have these countries approached this task? What metrics have they identified, within and beyond the GII arsenal? Which of these metrics are interesting for other countries and WIPO to study?
- Third, it reviews which future innovation metrics would lend themselves to the measurement of innovation at the sub-national level, in particular those exploiting "big data" and new computational methods. A key question is: Which of these data are promising in the context of benchmarking innovation at the city, provincial or regional level?

1 See www.linkedin.com/posts/sachawunschvincent_colombia-china-india-activity-6939934275085475840-HSS5?utm_source=share&utm_medium=member_desktop.

We do not attempt to provide a blueprint for sub-national innovation indices via this study. Providing a framework that applies to all countries is a vain task. The level of political autonomy of regions and data availability differs widely across countries. Rather, with the results presented here, we intend to provide a strong stimulus to ongoing national initiatives. These results can improve the journey towards better innovation measurement. We hope that these results do justice to the requests of member states for related assistance.

1. Applicability of the GII framework to the development of sub-national innovation metrics and indices

This first part reviews the applicability of the GII framework to the development of sub-national innovation metrics and indices. The key question relating to this point is: Which of the GII metrics are available at the city, regional, provincial or any other sub-national level?

Indicators often are not relevant below the national level for a vast number of countries (although admittedly, there can be exceptions). In that case, the indicator is marked in grey. blue-to-orange color code indicates how complex (or costly) it would be to implement the indicator at the regional level. The blue color indicates a high cost, and the orange color indicates a low cost. We provide a range of colors because data readiness inevitably differs across countries.

We acknowledge that the complexity of collecting data for some countries may fall outside the ranges reported below. The illustration summarizes our best effort at identifying the cost, with a view to providing a quick overview of the indicators most susceptible to being deployed at the regional level. However, countries wishing to develop sub-national innovation indices must perform a feasibility analysis considering their specific political and socioeconomic contexts and data availability.

Table 1. **Complexity analysis for GII pillar 1**

	Indicator	How straightforward
1. Institutions	1.1.1 Political and operational stability	
	1.1.2 Government effectiveness	
	1.2.1 Regulatory quality	
	1.2.2 Rule of law	
	1.2.3 Cost of redundancy dismissal	
	1.3.1 Policies for doing business	Country-specific
	1.3.2 Entrepreneurship policies and culture	

The quality and functioning of "institutions" (GII pillar 1) are *usually* national matters, such that the indicators may not lend themselves very well to sub-national innovation indices (Table 1). That said, some regional governments have considerable power when it comes to economic affairs, having the possibility to influence ease-of-doing-business policies. Finally, regions in a country may exhibit marked cultural differences and may have different attitudes toward entrepreneurship. However, building such an indicator is complex as it requires a dedicated survey.

Table 2. Complexity analysis for GII pillar 2

	Indicator	How straightforward
2. Human capital and research	2.1.1 Expenditure on education, % GDP	
	2.1.2 Government funding/pupil, secondary, % GDP/cap	
	2.1.3 School life expectancy, years	
	2.1.4 PISA scales in reading, maths and science	
	2.1.5 Pupil–teacher ratio, secondary	
	2.2.1 Tertiary enrolment, % gross	
	2.2.2 Graduates in science and engineering, %	
	2.2.3 Tertiary inbound mobility, %	
	2.3.1 Researchers, FTE/mn pop.	
	2.3.2 Gross expenditure on R&D, % GDP	
	2.3.3 Global corporate R&D investors, top 3, mn USD	
	2.3.4 QS university ranking, top 3	

Many indicators in the "human capital and research" (GII pillar 2) category should be readily available at the regional level in governments' databanks (Table 2). Some indicators, such as tertiary inbound mobility, make more sense in an international context than in a regional context.

Table 3. Complexity analysis for GII pillar 3

	Indicator	How straightforward
3. Infrastructure	3.1.1 ICT access	
	3.1.2 ICT use	
	3.1.3 Government's online service	
	3.1.4 E-participation	
	3.2.1 Electricity output, GWh/mn pop.	
	3.2.2 Logistics performance	
	3.2.3 Gross capital formation, % GDP	
	3.3.1 GDP/unit of energy use	
	3.3.2 Environmental performance	
	3.3.3 ISO 14001 environmental certificates/bn PPP$ GDP	

We expect government online service and e-participation to exhibit more heterogeneity across countries than within countries (Table 3 as part of GII pillar 3). Many regulations concerning digital government services are likely at the national level – although regions may have different degrees of implementation. The environmental performance indicator requires a dedicated study, hence the high complexity.

Table 4. Complexity analysis for GII pillar 4

	Indicator	How straightforward
4. Market sophistication	4.1.1 Finance for startups and scaleups	
	4.1.2 Domestic credit to private sector, % GDP	
	4.1.3 Loans from microfinance institutions, % GDP	
	4.2.1 Market capitalization, % GDP	
	4.2.2 Venture capital investors, deals/bn PPP$ GDP	
	4.2.3 Venture capital recipients, deals/bn PPP$ GDP	
	4.2.3 Venture capital received, value, % GDP	
	4.3.1 Applied tariff rate, weighted avg., %	
	4.3.2 Domestic industry diversification	
	4.3.3 Domestic market scale, bn PPP$	

Tariff rates are subject to bilateral (or multilateral) agreements and are predominantly a matter of national law (Table 4 as part of GII pillar 4). The indicator related to loans from microfinance institutions is missing for many countries in the GII, such that breakdown at the regional level may, often, not be feasible.

Table 5. Complexity analysis for GII pillar 5

	Indicator	How straightforward
5. Business sophistication	5.1.1 Knowledge-intensive employment, %	
	5.1.2 Firms offering formal training, %	
	5.1.3 GERD performed by business, % GDP	
	5.1.4 GERD financed by business, %	
	5.1.5 Females employed w/advanced degrees, %	
	5.2.1 University–industry R&D collaboration	
	5.2.2 State of cluster development and depth	
	5.2.3 GERD financed by abroad, % GDP	
	5.2.4 Joint venture/strategic alliance deals/bn PPP$ GDP	
	5.2.5 Patent families/bn PPP$ GDP	
	5.3.1 Intellectual property payments, % total trade	
	5.3.2 High-tech imports, % total trade	
	5.3.3 ICT services imports, % total trade	
	5.3.4 FDI net inflows, % GDP	
	5.3.5 Research talent, % in businesses	

Indicators in the “business sophistication” category are the most challenging to replicate at the regional level (Table 5 as part of GII pillar 5). Some indicators are complex to replicate because they require a new survey, such as the university–industry research and development (R&D) collaboration. Others are not so relevant at the regional level. Regarding the state of cluster development, we note that small regions may have no clusters. Joint-venture and strategic alliances usually involve multinational enterprises (MNEs) that may span several regions in a country. Thus, unless an MNE is headquartered solely in one region, it is difficult to delineate the impact of such deals.

Table 6. Complexity analysis for GII pillar 6

	Indicator	How straightforward
6. Knowledge and technology outputs	6.1.1 Patents by origin/bn PPP$ GDP	
	6.1.2 PCT patents by origin/bn PPP$ GDP	
	6.1.3 Utility models by origin/bn PPP$ GDP	
	6.1.4 Scientific and technical articles/bn PPP$ GDP	
	6.1.5 Citable documents H-index	
	6.2.1 Labor productivity growth, %	
	6.2.2 New businesses/th pop. 15–64	
	6.2.3 Software spending, % GDP	
	6.2.4 ISO 9001 quality certificates/bn PPP$ GDP	
	6.2.5 High-tech manufacturing, %	
	6.3.1 Intellectual property receipts, % total trade	
	6.3.2 Production and export complexity	
	6.3.3 High-tech exports, % total trade	
	6.3.4 ICT services exports, % total trade	

Intellectual property (IP) data are readily available and can be allocated with little effort to regions using applicant or inventor addresses (Table 6 as part of GII pillar 6). Constructing indicators based on scientific articles requires mapping institutions to regions, which is also doable with little effort. Building a regional indicator becomes quite challenging when data are proprietary, such as on software spending.

Table 7. **Complexity analysis for GII pillar 7**

	Indicator	How straightforward
7. Creative outputs	7.1.1 Intangible asset intensity, top 15, %	
	7.1.2 Trademarks by origin/bn PPP$ GDP	
	7.1.3 Global brand value, top 5,000, % GDP	
	7.1.4 Industrial designs by origin/bn PPP$ GDP	
	7.2.1 Cultural and creative services exports, % total trade	
	7.2.2 National feature films/mn pop. 15–69	
	7.2.3 Entertainment and media market/th pop. 15–69	
	7.2.4 Printing and other media, % manufacturing	
	7.2.5 Creative goods exports, % total trade	
	7.3.1 Generic top-level domains (TLDs)/th pop. 15–69	
	7.3.2 Country-code TLDs/th pop. 15–69	
	7.3.3 GitHub commit pushes received/mn pop. 15–69	
	7.3.4 Mobile app creation/bn PPP$ GDP	

Creative output indicators are also challenging to replicate at the regional level (Table 7 as part of GII pillar 7). For instance, one would need to observe users' locations to build the GitHub commit pushes indicator. This could be feasible using users' IP addresses but would be imperfect information. Regarding national feature films, the cinema industry is usually highly geographically concentrated in a country and not spread throughout the territory. This makes the indicator poorly adapted to a regional indicator.

2. Insights from existing sub-national innovation indices: China, Colombia, the European Union, India and Viet Nam

In this second part of the study, we review existing sub-national innovation indices of WIPO member states who have pioneered this field, notably China, Colombia, the European Union (EU), India and Viet Nam. Key questions are: How have these countries approached this task? What metrics have they identified, within and beyond the GII arsenal? Which of these metrics are interesting for other countries and WIPO to study?

We draw on indices that exist and are under development in select WIPO member states, notably China, Colombia, India and Viet Nam. These indices were presented at the June 7, 2022 workshop (see Annex 1).

- The "China Regional Innovation Capability Index" is developed by the Chinese Academy of Science and Technology for Development, in the People's Republic of China. It is one of several Chinese regional innovation indices.
- The "Índice Departamental de Innovación para Colombia" (Colombian Departmental Innovation Index) is developed by the National Planning Department, Colombia. It has 108 indicators and is based on the conceptual model of the GII.
- The "India Innovation Index (III)" is developed each year by the National Institution for Transforming India (NITI) Aayog, India, with the Institute for Competitiveness. The III has 66 indicators. NITI is working with states and union territories in improving their rankings in the index, with the understanding that improved regional III ranks will also lead to improving the Indian innovation performance in the GII.
- The "Provincial Innovation Index (PII)" is being developed by the Viet Nam Institute for Science, Technology and Innovation, Ministry of Science and Technology, Viet Nam. It exists in an unpublished pilot version in 2023. The pilot is implementing a government resolution to create a national version of the WIPO GII. Currently, the PII pilot has been implemented in 20 cities and provinces (out of 63), and the nationwide scaleup is planned for late 2023.

In addition, we review the European Regional Innovation Scoreboard (RIS) developed by the European Commission, which follows the method of the European Innovation Scoreboard. The RIS assesses innovation systems across 240 regions of 22 countries of the European Union (EU).

Tables 8–14 review these five indices and contrast them to the GII. In the following tables, the original GII indicators are marked in blue in the column labeled Indicator. Elements in white in the column denote indicators not in the GII but used in at least one other sub-national index. The blue color in the columns labeled China, Colombia, EU, India and Viet Nam indicates that the corresponding indicator, or a close equivalent, is also used in the GII. The green indicates that a country relies on an indicator not used in the GII in 2023.

Table 8. Indicators used in GII pillar 1 across selected regional indices

	Indicator	China	Colombia	EU	India	Viet Nam
1. Institutions	**1.1.1** Political and operational stability	-		-	-	
	1.1.2 Government effectiveness	-		-	-	-
	1.2.1 Regulatory quality	-		-	-	-
	1.2.2 Rule of law	-		-	-	
	1.2.3 Cost of redundancy dismissal	-	-	-	-	-
	1.3.1 Policies for doing business	-	-	-		
	1.3.2 Entrepreneurship policies and culture	-	-	-	-	-
	STI policy development and implementation	-	-	-	-	
	Business entry costs	-		-	-	
	Proactivity of provincial authorities	-	-	-	-	
	Public administration reform	-	-	-	-	
	Fair competition	-	-	-	-	
	Percentage of court cases pending	-	-	-		-
	Charge sheeting rate	-	-	-		-
	Cases investigated/Total cases reported	-	-	-		-
	Rate of cognizable crime	-	-	-		-
	Police personnel per lakh of pop.	-	-	-		-
	Violation of press freedom	-		-	-	-
	% of employed population affiliated to health and pensions funds and occupied population	-		-	-	-
	Tax payment index per year	-		-	-	-
	Property Registration Index	-		-	-	-

Note: Green cell if indicator implemented by that country and not in the GII. Blue cell if indicator from the GII.

The majority of indicators capturing “institutions” (pillar 1 in the GII, see Table 8) are not implemented in the sub-national innovation indices in question. Viet Nam and India have introduced their own indicators capturing institutional features. In Viet Nam, the “Science, Technology and Innovation (STI) policy development and implementation” indicator also includes entrepreneurship policies.

Not surprisingly, all sub-national indices have adopted indicators capturing facets of “human capital and research” (pillar 2 in the GII, see Table 9). In particular, all countries have education- and R&D expenditure-related indicators, although there are differences in the exact nature of the indicators used. We also note the presence of novel indicators tailored to each country’s political and socioeconomic contexts. For instance, the III measures the number of schools with information and communication technology (ICT) labs, whereas the EU RIS seeks to capture the population’s digital skills and the number of ICT specialists.

When it comes to “infrastructure” (pillar 3 in the GII, see Table 10), most sub-national innovation indices capture ICT access or use, although we note that the indicator may not always be a composite index. For instance, in terms of ICT use, the Chinese index captures the fraction of mobile internet users in the population. Most sub-national innovation indices reviewed also capture components of pollution or sustainability, but each with a different flavor.

Table 9. Indicators used in GII pillar 2 across selected regional indices

	Indicator	China	Colombia	EU	India	Viet Nam
2. Human capital and research	**2.1.1** Expenditure on education, % GDP	-	-	-		
	2.1.2 Government funding/pupil, secondary, % GDP/cap	-	-	-	-	-
	Expenditure on higher and technical education/lakh of pop.	-	-	-		-
	2.1.3 School life expectancy, years	-		-	-	
	Education index (literacy and school enrolment)	-	-	-	-	
	Synthetic Index of "Educational quality"	-		-	-	-
	Schools having STEM/STEAM education, %	-	-	-	-	
	Accolades in STEM Activities/1000 students	-	-	-		-
	Schools with ICT labs	-	-	-		-
	% of schools having science/technology tinkering/ innovation labs	-	-	-		-
	NER (net enrolment ratio) in school education, age 6 to 11 years	-	-	-		-
	Net coverage rate in secondary education, age 11 to 14 years	-		-	-	-
	Net coverage rate in middle education, age 15 to 16 years	-		-	-	-
	Secondary school level completition rate	-	-	-		-
	2.1.4 PISA scales in reading, maths and science	-		-		-
	2.1.5 Pupil–teacher ratio, secondary	-	-	-		
	Pupil–teacher ratio, higher education	-	-	-		-
	2.2.1 Tertiary enrolment, % gross			-		-
	Tertiary education graduates		-		-	-
	Performance of tertiary education students in state tests	-		-	-	-
	2.2.2 Graduates in science and engineering, %	-		-	-	-
	2.2.3 Tertiary inbound mobility, %	-		-		-
	Enrolment in PhD per lakh of pop.	-	-	-		-
	Doctoral graduates/100 th pop.		-	-	-	-
	Master's and doctorate scholarships	-		-	-	-
	2.3.1 Researchers, FTE/mn pop.			-	-	
	Enterprise R&D researchers, % total R&D researchers		-	-	-	-
	R&D personnel/10 th pop.		-	-	-	-
	Innovation intermediary employees/100 th pop.		-	-	-	-
	2.3.2 Gross expenditure on R&D, % GDP			-		
	R&D expenditures public sector, %GDP	-	-		-	-
	Nb. of R&D institutions funded by the state per lakh of pop.	-	-	-		-
	Nb. of private R&D units per lakh of pop.	-	-	-		-
	Expenditure on S&T, % local budget		-	-		
	R&D instrument and equipment expenditure/R&D FTE		-	-	-	-
	New fixed assets, % scientific research and technical services		-	-	-	-
	2.3.3 Global corporate R&D investors, top 3, mn USD	-	-	-	-	-
	2.3.4 QS university ranking, top 3	-	-	-		-
	Lifelong learning	-	-			-
	Digital skills	-	-		-	-
	Non-R&D innovation expenditures in SMEs	-	-		-	-
	Innovation expenditures per person employed in innovative SMEs	-			-	-
	ICT specialists	-	-		-	-

Notes: Viet Nam may revise the indicator "schools having STEM/STEAM education" to "students participating in STEM/STEAM competition/total students."
Green cell if indicator implemented by that country and not in the GII. Blue cell if indicator from the GII.

Table 10. Indicators used in GII pillar 3 across selected regional indices

	Indicator	China	Colombia	EU	India	Viet Nam
3. Infrastructure	**3.1.1** ICT access	-		-		
	3.1.2 ICT use			-		
	3.1.3 Government's online service	-	-	-		
	3.1.4 E-participation	-	-	-	-	-
	Digital trust and security services	-		-	-	-
	Digital government	-		-	-	-
	3.2.1 Electricity output, GWh/mn pop.	-		-	-	-
	3.2.2 Logistics performance	-		-	-	-
	3.2.3 Gross capital formation, % GDP	-	-	-		-
	3.3.1 GDP/unit of energy use			-	-	-
	Public investment in fixed capital	-		-	-	-
	General infrastructure	-	-	-	-	
	Industrial zones, %	-	-	-	-	
	Location entropy of equipment manufacturing industry		-	-	-	-
	3.3.2 Environmental performance	-		-	-	-
	Environment quality index		-	-	-	-
	Environmental pollution control index		-	-	-	-
	Environmental governance (seriousness in protection, air and water quality)	-	-	-	-	
	Air emissions by fine particulates	-	-		-	-
	3.3.3 ISO 14001 environmental certificates/bn PPP$ GDP	-		-	-	-
	Number of projects/proposals approved for environment clearance	-	-	-		-
	Number of common facility centers	-	-	-		-
	Intensity of business spending on R&D (%)	-		-	-	-

Note: Green cell if indicator implemented by that country and not in the GII. Blue cell if indicator from the GII.

Table 11. Indicators used in GII pillar 4 across selected regional indices

	Indicator	China	Colombia	EU	India	Viet Nam
4. Market sophistication	**4.1.1** Finance for startups and scaleups	-	-	-	-	-
	Incubators		-	-		-
	4.1.2 Domestic credit to private sector, % GDP	-		-		
	Ease to obtain credit	-		-	-	-
	Credit given to companies to innovate	-		-	-	-
	Private investment in fixed capital in the industry (% of GDP)	-		-	-	-
	4.1.3 Loans from microfinance institutions, % GDP	-	-	-		
	4.2.1 Market capitalization, % GDP	-		-	-	-
	4.2.2 Venture capital investors, deals/bn PPP$ GDP	-	-	-	-	-
	4.2.3 Venture capital recipients, deals/bn PPP$ GDP	-	-	-		-
	4.2.4 Venture capital received, value, % GDP	-	-	-	-	-
	4.3.1 Applied tariff rate, weighted avg., %	-	-	-	-	-
	4.3.2 Domestic industry diversification	-		-	-	-
	4.3.3 Domestic market scale, bn PPP$	-		-	-	-
	Export destination market diversification	-		-	-	-
	Performance of provincial S&T development fund	-	-	-	-	
	Firms providing professional S&T services/thousand firms	-	-	-	-	
	Services providers/thousand firms	-	-	-	-	

Note: Green cell if indicator implemented by that country and not in the GII. Blue cell if indicator from the GII.

"Market sophistication" (pillar 4 in the GII, see Table 11) is not well covered, except in the sub-national innovation indices of India and Viet Nam. Outside of India, the lack of venture capital information in most indices is surprising in light of the general policy focus on entrepreneurial activities.

Table 12. Indicators used in GII pillar 5 across selected regional indices

	Indicator	China	Colombia	EU	India	Viet Nam
5. Business sophistication Services providers/thousand firms	**5.1.1** Knowledge-intensive employment, %	-	(blue)	(blue)	(blue)	-
	NGOs involved in knowledge intensive areas	-	-	-	(green)	-
	5.1.2 Firms offering formal training, %	-	(blue)	-	(blue)	(blue)
	5.1.3 GERD performed by business, % GDP	(blue)	-	-	-	-
	5.1.4 GERD financed by business, % GERD	-	(blue)	(blue)	-	(blue)
	Enterprise tech. acquisition and tech. transformation expenditure	(green)	(green)	-	-	-
	Firms with R&D activities, % of total firms	(green)	-	-	-	(green)
	Companies with innovation activities, %	-	-	-	-	(green)
	Firms operating in industrial zones, %	-	-	-	-	(green)
	Digital economy (composite index)	-	-	-	-	(green)
	Total number of online services transaction/1000 pop.	-	-	-	(green)	-
	E-commerce sales, % GDP	(green)	-	-	-	-
	Percentage of subsidies or benefits transfered through digital platform	-	-	-	(green)	-
	Number of bank accounts/lakh of pop.	-	-	-	(green)	-
	Share of manufacturing & services as a % of GSDP	-	-	-	(green)	-
	5.1.5 Females employed w/advanced degrees, %	-	(blue)	-	(blue)	-
	5.2.1 University–industry R&D collaboration	-	(blue)	-	-	(blue)
	Companies that cooperate with international organizations (%)	-	(green)	-	-	-
	Public–private co-publications	-	-	(green)	-	-
	5.2.2 State of cluster development and depth	-	-	-	(blue)	-
	SMEs, cooperatives and individual businesses in industrial zones, %	-	-	-	-	(green)
	ISO certified companies, % of total firms	-	(green)	-	-	(green)
	5.2.3 GERD financed by abroad, % GDP	-	(blue)	-	-	-
	Direct foreign investment from abroad	-	(green)	-	-	-
	5.2.4 Joint venture/strategic alliance deals/bn PPP$ GDP	-	-	-	-	-
	Innovative SMEs collaborating with others	-	-	(green)	-	-
	5.2.5 Patent families/bn PPP$ GDP	-	-	-	-	-
	5.3.1 Intellectual property payments, % total trade	-	-	-	-	-
	5.3.2 High-tech imports, % total trade	-	(blue)	-	-	-
	5.3.3 ICT services imports, % total trade	-	-	-	-	-
	5.3.4 FDI net inflows, % GDP	-	-	-	(blue)	(blue)
	5.3.5 Research talent, % in businesses	-	-	-	-	(blue)
	Investment in ICT of companies that introduce new organizational methods	-	(green)	-	-	-
	Value of contract deals in domestic technical markets	(green)	-	-	-	-
	Industrial Specialization Index	-	(green)	-	-	-

Note: Green cell if indicator implemented by that country and not in the GII. Blue cell if indicator from the GII.

All sub-national innovation indices include several indicators capturing aspects of "business sophistication" (pillar 5 in the GII, see Table 12). Similar to other categories, there is little overlap in the indicators used across countries.

Table 13. Indicators used in GII pillar 6 across selected regional indices

	Indicator	China	Colombia	EU	India	Viet Nam
6. Knowledge and technology outputs	Sales of new-to-market and new-to-firm innovations				-	-
	6.1.1 Patent applications by origin/bn PPP$ GDP			-		
	Nb. of invention patents/10 th pop.		-	-	-	-
	6.1.2 PCT patents by origin/bn PPP$ GDP	-	-		-	-
	Patent requests in the ICT sector per million inhabitants	-		-	-	-
	Patents granted in the last three years per million inhabitants	-		-	-	-
	6.1.3 Utility models by origin/bn PPP$ GDP	-		-	-	
	Plant varieties by origin/10,000 pop.			-	-	
	6.1.4 Scientific and technical articles/bn PPP$ GDP			-		
	International scientific co-publication	-	-		-	-
	6.1.5 Citable documents H-index	-		-	-	-
	Most cited publications	-	-			-
	Product process innovators SMEs	-	-		-	-
	Business process innovators SMEs	-			-	-
	Number of grassroot innovators per lakh of pop.	-	-	-		-
	6.2.1 Labor productivity growth, %			-		-
	Labor productivity of high-tech industry		-	-	-	-
	Labor productivity of knowledge intensive service industry		-	-	-	-
	Capital productivity, ratio of gross product to capital inputs		-	-	-	-
	6.2.2 New businesses/th pop. 15–64	-		-		
	Number of S&T enterprises and the equivalents/1000 firms	-	-	-	-	
	Startups	-	-	-		
	Strictly innovative companies (%)	-		-	-	-
	Innovative companies in broad sense (%)	-		-	-	-
	Employment in innovative SMEs	-			-	-
	6.2.3 Software spending, % GDP	-	-	-	-	-
	Added value of information transmission, software and ICT services, % GDP		-	-	-	-
	6.2.4 ISO 9001 quality certificates/bn PPP$ GDP	-		-	-	-
	6.2.5 High-tech manufacturing, %			-		-
	6.3.1 Intellectual property receipts, % total trade		-	-	-	-
	6.3.2 Production and export complexity	-	-	-	-	-
	6.3.3 High-tech exports, % total trade			-		-
	6.3.4 ICT services exports, % total trade	-		-		-
	Operating income of high-tech industry,% industrial operating income		-	-	-	-
	Added value of knowledge intensive service industry, % GDP		-	-	-	-
	High-tech industry profit margin		-	-	-	-
	Turnover of technology output/10 th pop.		-	-	-	-
	Strictly innovative companies (%)	-		-	-	-
	Innovative companies in broad sense (%)	-		-	-	-

Notes: Viet Nam is considering not using the indicator "scientific and technical articles."
Green cell if indicator implemented by that country and not in the GII. Blue cell if indicator from the GII.

"Knowledge and technology outputs" (pillar 6 in the GII, see Table 13) has the most direct mapping between the GII indicators and the sub-national innovation indices' indicators. Twelve GII indicators are found in at least one, but often two or three, sub-national innovation indices. All sub-national innovation indices include data on patent filings (either national patent applications or via the Patent Cooperation Treaty, PCT) and scientific publication data, again with different flavors across countries. Interestingly, non-GII indicators are sometimes used in multiple sub-national innovation indices, including innovation survey data on, for example, innovative sales and startups.

Table 14. **Indicators used in GII pillar 7 across selected regional indices**

	Indicator	China	Colombia	EU	India	Viet Nam
7. Creative outputs	**7.1.1** Intangible asset intensity, top 15, %	-	-	-	-	-
	7.1.2 Trademarks applications by origin/bn PPP$ GDP	-				
	7.1.3 Global brand value, top 5000, % GDP	-	-	-	-	-
	7.1.4 Industrial design applications by origin/bn PPP$ GDP	-				
	Geographical indications by origin	-	-	-		
	Orange economy production (% GDP of the department)	-		-	-	-
	7.2.1 Cultural and creative services exports, % total trade	-	-	-	-	-
	7.2.2 National feature films/mn pop. 15–69	-	-	-	-	-
	7.2.3 Entertainment and media market/th pop. 15–69	-	-	-	-	-
	7.2.4 Printing and other media, % manufacturing	-	-	-	-	-
	7.2.5 Creative goods exports, % total trade	-		-	-	-
	7.3.1 Generic top-level domains (TLDs)/th pop. 15–69	-	-	-	-	-
	7.3.2 Country-code TLDs/th pop. 15–69	-	-	-	-	-
	7.3.3 GitHub commit pushes received/mn pop. 15–69	-	-	-	-	-
	7.3.4 Mobile app creation/bn PPP$ GDP	-	-	-	-	-
	Digital enterprises	-		-	-	-
	Software records	-		-	-	-

Note: Green cell if indicator implemented by that country and not in the GII. Blue cell if indicator from the GII.

When it comes to "creative outputs" (pillar 7 in the GII, see Table 14), the only indicators that are available in sub-national innovation indices relate to IP data. As mentioned in the previous section, "creative outputs" indicators are challenging to replicate at the regional level due to data availability.

Finally, some regional sub-national indices have proposed novel indicators (marked in blue) which are of interest both to the development of indices in other countries and for consideration in future editions of the GII, with the below caveats applying. Before going into details, one general point applies.

One difference between the GII and some sub-national indices is the reliance on innovation surveys in some of the latter – in particular in the case of the EU indices. Innovation surveys modeled on the Organisation for Economic Co-operation and Development (OECD)/Eurostat Oslo Manual 2018 are administered to a representative sample of firms across all sectors, yielding survey responses from said firms on their innovation activities, their nature and related specificities (including on collaboration) and related impacts (including on turnover).[2] These surveys yield the percentage of firms that claim to innovate and to collaborate, and they also assess the innovation impacts on their bottom line (for details, see the European Union community innovation survey; see also Box 3 in GII 2011). These surveys are a major breakthrough in innovation measurement, in particular for following innovation activities and impacts at the country level over time. For example, one question the surveys might address is: How has firm innovation changed in France over the last five years?

Cross-country comparability of these results is meaningful for a relatively cohesive group of countries – such as within the EU – in which the survey is administered in comparable and representative ways.

Yet, despite progress over the last 10–15 years, the availability of these innovation surveys at the international level is lmited and the comparability of results across countries is often poor. These limitations make the use of these survey data at the global level in the GII next to impossible in 2023. While we hope that more and more innovation surveys will be carried out at the national level, the likelihood of these results being sufficiently available (minimum 60–80 countries) and sufficiently comparable across these countries in 2023 or soon after is rather low, and hence not an avenue that the GII team is currently pursuing. That does not diminish the appeal of these surveys for sub-national innovation indices at all, however. If the sampling and sectoral coverage are satisfactory, the metrics are very useful.

2 www.oecd.org/innovation/inno/inno-stats.htm.

That said, other indicators employed in these sub-national indices deserve mention:

- In pillar 1 (**institutions**), a number of indicators specifically relate to (i) the state of innovation policy development and implementation (rather than policies in general), (ii) more fine-grained indicators capturing business entry, but also competition, (iii) press freedom (an indicator formerly present in the GII) and (iv) more granular data on the efficacy of the court and property rights system. While not easily available, indicators on (i) and (ii) would be particularly worthwhile in the GII context. The first is challenging to implement, as there is no universal framework for assessing innovation policies and none for monitoring their solid execution. As to the second, the GII team has been particularly interested in identifying new metrics relating to competition for several years, mostly for pillar 4 "business sophistication". A previous GII indicator – since dropped – exploited survey data to assess the state of competition in economies generally. It would be preferable to generate solid quantitative indicators of competition based on recognized data such as the Herfindahl–Hirschman Index (an indicator of the extent of competition among firms), and possibly even sectoral variations, as applicable. As some innovation sectors are growing more concentrated, with barriers to research entry and catching up increasing in many cases, this is an important research field for the future.
- In pillar 2 (**human capital and research**), the following fields are interesting: (i) better metrics on the quality of education (beyond GII indicator 2.1.4 on Program for International Student Assessment (PISA) results), (ii) data on the extent of Science, technology, engineering and mathematics (STEM) education and related achievements and (iii) various indicators approximating to what extent innovative thinking is taught at school (for example, the presence of innovation and tinkering labs, see the III which is pioneering this field). We find the latter to be particularly worthwhile in the GII context – addressing the question of how well schools prepare entrepreneurial and innovative mindsets – but also particularly challenging from an international measurement perspective. From a coverage point of view, the easiest solution could be for exercises such as the OECD PISA to cover such topics more explicitly, alongside existing efforts along these lines (PISA 2022, for example, already included an additional test of creative thinking, and see Chapter 3, GII 2014).
- In pillar 3 (**infrastructure**), only a few indicators emerge for possible consideration, primarily (i) variations of public investment data and (ii) a few ICT-related indicators, of which "digital trust" is an interesting one. Encouragingly, most indices reviewed mirror the GII in terms of encapsulating environmental soundness indicators on the innovation input (and not output) side. The GII and countries building indices could consider which of these green metrics are pertinent to their measurement ambitions.
- In pillar 4 (**market sophistication**), again, only a few additional indicators emerge, notably (i) variations of indicators related to credit availability and (ii) other funding mechanisms. One indicator of the number of incubators is particularly relevant but possibly hard to collect at the international level. Differentiating the quantity from the quality of incubators would also be a persistent challenge. Other metrics relate to the presence of public and sub-national innovation funds, another pertinent measurement ambit. This would be key at the international level, in particular, if impact-related metrics were available (disbursement of funds, impact over time in terms of firm entry, patenting, licensing or other performance metrics).
- In pillar 5 (**business sophistication**), the range of new indicators is large, although a lot of these indicators are covered elsewhere in the GII in different or similar forms (for example, foreign direct investment, International Organization for Standardization (ISO), industry diversification, specialization and others). One set of indicators concerns innovation survey data discussed earlier and likely is not of immediate relevance to the GII but clearly of relevance to sub-national innovation indices. Other metrics of interest are: (i) a wide set of indicators on e-commerce and online services and firm e-readiness, (ii) data on industrial zones and (iii) data on innovative small and medium-sized enterprises (SMEs) and their collaboration efforts.
- In pillar 6 (**knowledge and technology outputs**), a number of novel indicators either stem from innovation surveys (for example, firms with new-to-market innovations, or the share of SMEs that claim to be process innovators) or are derivates or variations of GII indicators. These are relevant mostly on the scientific publication and patenting front (for example, patents in the ICT sector or international scientific co-publication), but also with respect to turnover, profit margins or employment in innovative and/or knowledge-intensive firms.

In particular, the GII could study the availability of plant variety data produced at WIPO,[3] or the meaningfulness of international co-patenting and co-publishing data. The latter has previously been studied for GII use but has been discarded thus far. A really important field that deserves more work relates to new metrics on startups, innovative SMEs and other such metrics. The GII includes two related data points as of the GII2023 version (one on new firms and one on unicorns), but much work remains to generate internationally reliable data on this front. It is challenging to distinguish truly innovative and R&D-intensive firm creations with a positive medium- to long-term outlook from business registrations, which include all forms of new businesses (including mom-and-pop shops, hair salons and so on) and which thus might not be representative of innovative startups. This is a key research agenda for the near future, and as of 2023 work has started to this effect.

- In pillar 7 (**creative outputs**), the indicators going beyond the GII are most limited. Only one indicator – geographical indications (GIs) – is a noteworthy departure from the GII, and even that one is – despite WIPO data collections in the field – debatable as to its use for large cross-country data comparisons, mainly because the use of GIs is not equally spread across countries, and because the volumes are typically very low.

In sum, the sub-national studies under review innovate in terms of introducing and trialing innovative innovation indicators, more fit for the regional or city level, and possibly more available and fit for the specific sub-national country ambit in question.

Other countries and the GII framework will benefit from a closer study of these variables. The litmus test: Are these metrics available for a large group of countries (minimum 60) at high quality, and will these be produced consistently for years?

In turn, the GII team will take particular interest in data relating to competition, entrepreneurial and innovation culture (skills taught at schools and societal attitudes), innovation finance, startups and generally high-growth research-intensive young firms (so-called gazelles), and possibly new IP indicators, in particular plant varieties.

3 See also the WIP IP Indicators Report 2022 at www.wipo.int/edocs/pubdocs/en/wipo-pub-941-2022-en-world-intellectual-property-indicators-2022.pdf.

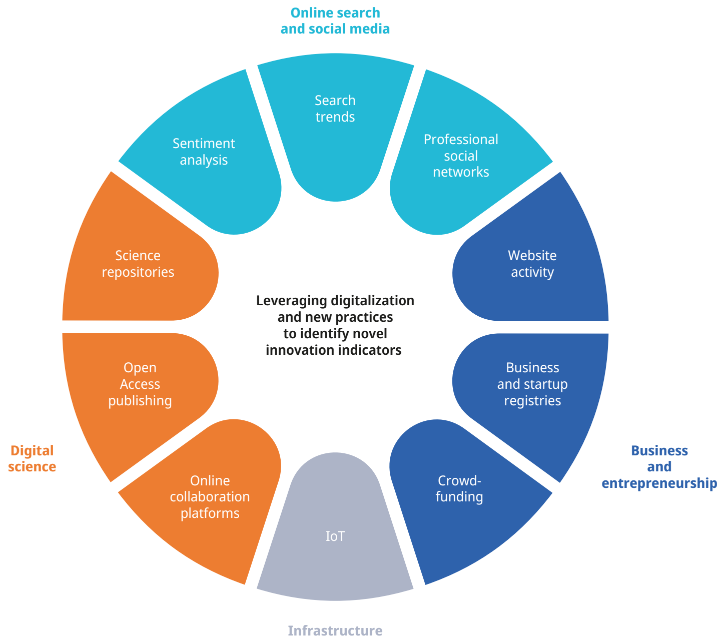
3. Future innovation metrics exploiting big data and new computational methods

The third part of this study assesses which future innovation metrics would lend themselves to the measurement of innovation at the sub-national level, in particular those leveraging advances in big data processing and digitalization, and those reflecting changing innovation practices.[4]

A key question is: Which of these metrics are promising in the context of gauging innovation at the city, provincial or regional level? What is required to move these innovation metrics from theory into practice? Who are potential early adopters? What are realistic timelines?

We group the metrics into four broad novel approaches capturing the data sources and generation process, as summarized in Figure 1, covering four broad categories, namely (i) online search and social media, (ii) business and entrepreneurship, (iii) infrastructure and (iv) digital science metrics.

Figure 1. Four broad novel approaches to innovation metrics



Before delving into these four groups of novel innovation metrics, it is worth starting with a general point. Most of these data have one thing in common: they contain details (such as an address, a geo-code or a web-related tag) that allow you to "map" the innovation activity or innovator to a particular geographic location. This works through unit record data such as patents, scientific publications, company reports or any other official document containing

4 See also NESTA (2008) and GII 2018, Annex 1, Box 1, www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2018.pdf for an earlier overview.

precise addresses. For instance, Bergquist *et al.* (2017) have geolocalized PCT patent and scientific publication data – as part of the WIPO GII – to build a ranking of top GII 100 science and technology clusters. Also, de Rassenfosse *et al.* (2019) have geolocalized worldwide patent data, allowing researchers to build their own follow-up indicators and aggregations. It also works through internet-based activities, which can be geolocalized – as described below – yielding new possibilities to precisely identify and cluster innovation activity.

Online search and social media

The advent of online search and social media platforms has introduced new opportunities to measure and understand innovation at the regional level. In this section, we will explore three aspects of online search and social media that have emerged as novel indicators of innovation: search trends, social media data and human capital as inferred from social media platforms.

Search trends

Internet search trends have become an essential source of real-time information on innovative activities and interests. Platforms such as Google Trends and Bing provide data on the popularity of specific search terms entered into their search engines over time and across regions. This information can be utilized to track interest in emerging technologies, identify innovative ideas gaining traction and estimate fluctuations in activities related to innovation.

Several studies have demonstrated the potential of search trends data for predicting economic activity and innovation. For example, Austin *et al.* (2021) showed how Google Trends could have been used to predict the fall and subsequent recovery in the gross domestic product (GDP) of selected countries during the early stages of COVID-19. Similarly, Woo and Owen (2019), Wolosko (2020), and Kohns and Bhattacharjee (2023) used Google Trends data to investigate various aspects of economic activity. While there are limitations to search trends data, such as the availability of search terms and potential biases introduced by search algorithms, the real-time nature of this information makes it a valuable complement to traditional sources of innovation data.

Furthermore, search trends data can be used to identify regional differences in the interest and adoption of specific innovations. By comparing the search behavior of users in different regions, researchers can measure the diffusion of new ideas and technologies. In turn, this information can help policymakers and industry leaders develop targeted strategies for promoting innovation in specific areas.

Social media data

Social media platforms such as Twitter, Facebook and LinkedIn have become increasingly relevant for capturing the pulse of innovative activities and trends. These platforms generate massive amounts of data, offering insights into public sentiment, the emergence of new ideas and the diffusion of innovative technologies.

One approach to leveraging social media data for innovation analysis is sentiment analysis. Sentiment analysis involves extracting and interpreting the emotional content of a text, providing insights into public perception and attitudes towards various topics. Pindado and Barrena (2021) used sentiment analysis on Twitter data to understand social perceptions of emerging food trends across different regions. They employed a density-based clustering technique on 7,014 tweets and used grid maps to examine regional variations. Similarly, Gibbons *et al.* (2019) and Zhang and Moe (2021) used sentiment analysis on Twitter and Facebook data, respectively, to explore consumer sentiments towards brands and products. By applying sentiment analysis to social media data, researchers can gain a deeper understanding of public opinion and acceptance of innovative ideas and technologies. To the extent that social media data are geolocalized, individual posts can be aggregated at the sub-national level to derive novel innovation metrics.

Another approach to harnessing social media data for innovation measurement is the analysis of social networks and connections. For example, collaborations, partnerships and knowledge sharing among individuals and organizations can be studied by examining connections and interactions on platforms such as LinkedIn. These insights can provide valuable information on the flow of knowledge and the strength of innovation ecosystems within regions.

Professional social networks

Human capital is a critical component of innovation ecosystems, and online platforms such as LinkedIn and Glassdoor offer unique opportunities to study the education, skills and professional experiences of the workforce within a geographic area. By analyzing this information, researchers can better understand the potential for innovation and the role of human capital in driving regional growth.

LinkedIn, for example, can provide data on the distribution of skills, professional experiences and education levels among the workforce in a region (Baruffaldi *et al.*, 2017). This information can be used to assess the availability of the necessary talent pool to support innovative industries and identify areas where skill gaps may exist. Additionally, analyzing the movement of highly skilled individuals, such as researchers or entrepreneurs, between regions can provide insights into the flow of innovative ideas and knowledge.

Glassdoor, on the other hand, can offer insights into job market trends, including innovation-related jobs. By examining this information, researchers can identify patterns in job creation, compensation and workplace culture that may influence the attractiveness of a region for innovative businesses and talent (Hershbein & Kahn, 2018). In contrast to existing indicators, which primarily focus on R&D personnel, Glassdoor data offer a window into a broader category of creative jobs, in the design or artistic sectors for instance.

In addition to using LinkedIn and Glassdoor data to study the education, skills and experiences of the workforce, researchers can also leverage these platforms to gain insights into the dynamics of the job market within regions. By analyzing job postings, it may be possible to identify patterns in the demand for specific skills or qualifications, as well as trends in job creation and growth within specific sectors or regions.

Business and entrepreneurship

Website activity

The advent of digital technology has transformed the way businesses communicate and share information with their customers, partners and investors. Corporate websites offer a rich source of data for researchers interested in tracking innovation and entrepreneurial activity at the regional level. By analyzing the content of these websites, it is possible to gain insights into the strategies, technologies and collaborative networks that underpin innovative ecosystems.

For instance, Kinne and Axenbeck (2020) and Kinne and Lenz (2021) proposed a web mining framework for creating a consistent and repeatable mapping of innovation ecosystems. By examining the textual and relational content of corporate websites, the researchers were able to identify patterns of innovation activity, collaboration and technology adoption across a wide range of businesses. This approach can help overcome some of the limitations associated with traditional measures of innovation, such as coverage, granularity, timeliness and cost.

In addition to analyzing website content, researchers can also track the launch of new products and services through press releases, blog posts and other forms of online communication. By monitoring these announcements, it is possible to identify trends in the development and adoption of innovative technologies, as well as the industries and regions in which they are gaining traction.

Business and startup registry

Business registration records represent a valuable source of data for tracking entrepreneurial activity and assessing the quality of startups within a given region. By analyzing the characteristics of newly registered businesses, researchers can gain insights into the quantity and quality of entrepreneurial activity, as well as the factors that contribute to the growth and success of innovative regions.

One pioneering approach in this area was developed by Scott Stern and colleagues from MIT, who used business registration records to separate the quantity of entrepreneurial activity (the number of startups) from its quality (the likelihood that a startup will achieve growth through an Initial Public Offering, IPO, or significant acquisition). By examining the characteristics of startups at the time of registration, such as the firm's name and filing of a trademark or patent, the researchers were able to estimate the probability of a growth outcome and identify clusters of high-growth startups within specific regions.

Building on this approach, the Startup Cartography Project (SCP) by Andrews *et al.* (2022) developed four distinct entrepreneurial ecosystem statistics: the Startup Formation Rate (SFR), the Entrepreneurship Quality Index (EQI), the Regional Entrepreneurship Cohort Potential Index (RECPI) and the Regional Ecosystem Acceleration Index (REAI). These metrics enable researchers to quantify and compare the performance of entrepreneurial ecosystems at the regional level, providing valuable insights for policymakers and industry leaders seeking to foster innovation and economic growth.

In addition to administrative records, the use of comprehensive business databases such as Orbis, provided by Bureau van Dijk, can be instrumental in developing novel innovation indicators. Orbis contains information on close to 450 million companies worldwide, including financials, ownership structures and corporate events. By analyzing this wealth of data, researchers can identify trends and patterns that provide insights into the innovation landscape of a specific region or industry. One such application is the measurement of high-tech, high-growth firms, which can serve as a proxy for the overall innovative potential of a region.

High-tech, high-growth firms are companies that operate within technology-intensive industries and demonstrate rapid growth in terms of revenue, employee count or market share (see Coad *et al.*, 2014). Identifying and quantifying these firms can provide a useful indicator of a region's ability to foster and support innovative businesses. By using data from providers such as Orbis, researchers can filter companies based on their industry classification, location and growth metrics to pinpoint high-tech, high-growth firms within a specific region.

Crunchbase and AngelList are two prominent online platforms that provide valuable data on startups, investors and other key players in the entrepreneurial ecosystem. By leveraging the information available on these platforms, researchers can gain a more nuanced understanding of the dynamics of business creation, funding and growth in different regions.

For instance, Crunchbase is a comprehensive database that contains information on millions of startups, investors and funding rounds. By analyzing its data, researchers can track the growth and development of startups, identify trends in investment activity and assess the performance of venture capital firms and other investors in supporting innovative ventures (see, for example, Marra *et al.*, 2017). In addition to offering a wealth of data on individual companies, Crunchbase also provides aggregate statistics on startup activity, funding rounds and exits, enabling researchers to benchmark the performance of sub-national innovation ecosystems against national and global standards.

Crowdfunding

Crowdfunding platforms such as Kickstarter and Indiegogo have emerged as popular alternatives to traditional sources of financing for entrepreneurs and innovators. By tracking the number, types and success of crowdfunding campaigns, researchers can gain insights into the dynamics of emerging entrepreneurial ecosystems.

Yu and Fleming (2022) introduced a novel metric for evaluating entrepreneurial ecosystems based on crowdfunding activity. By scraping data on current and past Kickstarter campaigns, the researchers were able to track the performance of startups and gain insights into the types of projects and ideas that resonate with potential investors and backers. This approach offers several advantages over traditional measures of innovation, such as real-time data, greater granularity and the ability to capture the earliest stages of idea development.

In addition to tracking the performance of individual campaigns, crowdfunding data can also be used to analyze trends and patterns in the types of projects that receive funding, the industries in which they operate, and the regions where they are based. For example, researchers can investigate the distribution of crowdfunding campaigns across different sectors and geographies, the factors that influence their success rates and the role of local innovation ecosystems in supporting entrepreneurial activity.

AngelList, on the other hand, is a United States of America platform that connects startups with investors, job seekers and other resources to help them grow and succeed. By examining the data available on AngelList, researchers can gain insights into the composition of the entrepreneurial community in a given region, the types of startups that are attracting investment and the networks and relationships that underpin the local innovation ecosystem. Furthermore, AngelList data can be used to analyze the demand for talent in the startup sector, providing insights into the skills and expertise that are most sought after by innovative companies.

Digital science

The digital revolution has reshaped the landscape of scientific research and development activities, creating new opportunities for collaboration, knowledge sharing and innovation. By examining various facets of the digital science ecosystem, it is possible to develop novel innovation metrics that can offer valuable insights into the level of scientific activity and collaboration taking place within a given sub-national level.

Online collaboration platforms

Online collaboration platforms such as GitHub, GitLab and Bitbucket have become essential tools for researchers and developers, enabling them to collaborate on projects, share code and manage version control. By providing an open, transparent environment for collaboration, these platforms have the potential to accelerate the pace of scientific discovery and foster a more inclusive, decentralized research ecosystem (Dabbish *et al.*, 2012).

GitHub, for example, is a web-based platform that hosts over 200 million repositories and facilitates collaboration among millions of users. Researchers can leverage GitHub to collaborate on projects, share code and track changes in real time, helping to streamline the research process and reduce duplication of effort. Similarly, GitLab and Bitbucket offer comparable functionality, allowing researchers to work together on projects, manage version control and share their findings with the wider scientific community.

By mining the activity and collaboration patterns on these platforms (Kalliamvakou *et al.*, 2014), researchers can develop metrics to assess the level of innovation in ICT-related technologies. For example, examining the number of repositories, contributors and collaborations on these platforms can provide insights into the extent of development activities taking place and the pool of software scientists in a region.

Open-access publishing

Open-access publishing has emerged as a powerful force in the scientific community, promoting greater access to knowledge and enabling researchers to build upon the work of their peers. By making research findings freely available to all, open-access publishing has the potential to accelerate the pace of scientific discovery, facilitate collaboration and reduce barriers to entry for researchers from underrepresented groups (Tennant *et al.*, 2016).

The rise of open-access publishing has transformed the way scientific knowledge is disseminated, making research findings more accessible to a wider audience and promoting greater transparency and reproducibility in scientific research. Analyzing trends in open-access publications and the citation patterns of these articles can provide valuable insights into the impact and reach of innovative research within a given region (Piwowar *et al.*, 2018). Additionally, tracking the number of preprints shared on platforms such as arXiv, bioRxiv, SocArXiv and the Social Science Research Network (SSRN) can help to assess the speed of knowledge dissemination and the degree of openness in the research community (Kaiser, 2017).

Science repositories

Science repositories play a critical role in the digital science ecosystem, providing researchers with access to a wealth of data, software and other materials that can be used to advance their work. By examining the patterns of usage, contribution and collaboration within these repositories, it is possible to develop novel innovation metrics that can offer insights into the level of scientific activity and knowledge exchange taking place within a given sub-national region.

Data repositories, such as Dryad, Figshare and Zenodo, enable researchers to share their datasets with the wider scientific community, ensuring that their data can be accessed, reused and built upon by others (Vines *et al.*, 2013). Analyzing the number of datasets contributed, the number of citations received by these datasets and the geographical distribution of contributors can help assess the level of data sharing and collaboration in a specific research area or region.

Software repositories, such as GitHub, GitLab and Bitbucket, serve as platforms for sharing and archiving software tools, libraries and applications essential for scientific research. By analyzing the number of repositories, contributors and collaborations on these platforms, as well as the degree of software reuse, researchers could develop metrics to assess the level of knowledge exchange within a given region or research community (Ojanperä *et al.*, 2019). Additionally, tracking the adoption of specific software tools and libraries can provide insights into the diffusion of innovative methodologies and approaches across different research areas and regions.

Infrastructure

The emergence of smart cities and the internet of things (IoT) has significantly impacted the way urban planning, infrastructure development and resource management are approached (Bibri & Krogstie, 2017). Driven by the integration of digital technologies, data analytics and connected devices, these developments create more efficient, sustainable and resilient urban environments. Consequently, new trends and data sources have arisen that can be employed to develop novel innovation metrics at the regional level.

Smart cities utilize a multitude of interconnected sensors and devices to collect, analyze and manage data on various aspects of urban life, such as traffic patterns, energy consumption and waste management. These data streams offer valuable insights into the efficiency and effectiveness of urban infrastructure and the adoption of innovative technologies and practices within a given region. For example, the widespread use of smart meters and energy management systems can serve as an indicator of a region’s commitment to sustainable energy practices.

Moreover, the IoT has given rise to an ever-expanding network of connected devices and systems that generate vast amounts of data on various aspects of daily life. From wearable fitness trackers to connected appliances, these devices collect data on individual behaviors, preferences and consumption patterns. When aggregated at the regional level, this data can offer valuable insights into the adoption and diffusion of innovative technologies and practices, as well as the overall digital maturity of a given region.

By leveraging smart city and IoT data, policymakers and stakeholders can develop novel innovation metrics that capture the unique characteristics and dynamics of regional innovation ecosystems. For instance, the density and diversity of IoT devices in a region could serve as an indicator of the region’s technological infrastructure and readiness for innovation. Similarly, the extent to which a region’s citizens and businesses are adopting smart city solutions can provide insights into the region's commitment to digital transformation and innovation.

In sum, in addition to unit records such as patents allowing us to combine addresses with maps, these four novel big data branches could be exciting sources of innovation metrics and indices in the future.

Conclusion

The objective of this WIPO toolkit is to enable more and better innovation measurement at the sub-national (regional, provincial or city) level, thereby also supporting the many initiatives of WIPO member states to sub-national innovation indices.

The first part of this study sets out which GII indicators are fit for purpose at the sub-national level, on a scale ranging from easy to impossible.

The second part of this study reviews the existing sub-national innovation indices of China, Colombia, the EU, India and Viet Nam to understand their set up, their overlaps with the GII or among each other, and the fields and metrics that look particularly new or promising. To the best of our knowledge, and despite their prominence, it is the first comparison of such sub-national innovation indices. Our hope is that the comparison will allow for mutual learning across these countries and also across the many countries that are envisaging the development of innovation indicators or indices at the national or sub-national level. This review has also inspired a number of measurement fields that should be of high priority in the development of future GII editions and that should be pursued in the coming months and years.

The third part provides a taxonomy for novel big innovation data built through geolocation codes and processes, based on either traditional innovation records, such as patents or publications, or novel online-related data feeds or the general ability to geolocalize the innovator or the innovation process via automatized and computational means. Realistically, these measures are mostly not yet widely deployed in systematic large-scale innovation metrics at the country, cluster, regional or city level. Rather they are often at the experimental stage of academic research and workshops. Few of the existing sub-national indices have ventured to deploy such computationally based tools (outside the WIPO Global Science and Technology (S&T) Cluster Ranking). As a result, this line of work bears much potential.

These three parts are intended to improve the journey towards better innovation measurement. We hope that these results do justice to the requests of member states for related assistance.

Annex 1: Workshop – Global Innovation Index

Sharing of experiences in the creation and implementation of regional innovation indices June 7, 2022 (13.30–16.30 CEST)

Objective of the workshop

To allow selected members states to share their experiences in the creation of regional innovation indices.

The main points of interests are:

1. The main rationale of having set up such an index, including the intended purpose and interaction with the Global Innovation Index (GII) or a national innovation index.
2. The technical approach used, including the index structure, the variables that are identical or different to the GII, the data collection methods, related problems and aggregation techniques. Here are the relevant benchmark documents of the GII, the data used and the four steps to devise indices as reference. The fact that these indices are regional in scope and done for one country only raises data opportunities and challenges. It would be interesting to find out how the regional indices deviate from international innovation indices and what novel data sources are tapped.
3. The results and impacts of the regional innovation index, also (possibly) in relation to the GII.
4. Next steps or plans as regards the regional innovation index, and related data collections.

Welcome (13:30–13:40 CEST)

Mr. Marco M. Alemán, Assistant Director General, IP and Innovation Ecosystems Sector (IES), World Intellectual Property Organization (WIPO)

Ms. LE Thi Tuyet Mai, Ambassador Extraordinary and Plenipotentiary, Permanent Representative of Viet Nam to the UN, WTO and other International Organizations in Geneva

Introduction on the GII and regional innovation indices (13:40–13:55 CEST)

Mr. Sacha Wunsch-Vincent, Head, and Ms. Lorena Rivera Léon, Economist, Composite Indicator Research Section (CIRS), Department for Economics and Data Analytics (DEDA), IES, WIPO

Ms. Michaela Saisana, Head of Unit, Monitoring, Indicators and Impact Evaluation Unit, Directorate-General Joint Research Centre (JRC), European Commission

Presentation and discussion of regional innovation indices (13:55–16:20 CEST)

Chair: Ms. Vanessa Behrens, GII Project Manager, CIRS, DEDA, IES, WIPO

"Índice Departamental de Innovación para Colombia," Mr. Camilo Rivera, Director of Innovation and Entrepreneurial Development, National Planning Department, Colombia (20 minutes)

"India Innovation Index," Mr. Shri Neeraj Sinha, Sr. Adviser (S&T) and Mr. Ashok A. Sonkusare, Dy. Adviser (S&T), National Institution for Transforming India (NITI) Aayog, India (20 minutes)

"China's Regional Innovation Capability Index," Ms. SUN Yunjie, Associate Researcher, Chinese Academy of Science and Technology for Development, People's Republic of China (20 minutes)

Viet Nam"Regional Innovation Index," Mr. Minh Hoang, President, and Mr. Hung Vo Nguyen, Researcher, Viet Nam Institute for Science, Technology and Innovation, Ministry of Science and Technology, Viet Nam (20 minutes)

Comments and conclusions (16:30 CEST)

Mr. Carsten Fink, Chief Economist and Mr. Jack Gregory, Innovation Data Analyst, CIRS, WIPO

Comments: Mr. Marcos Pinto, Director of Science, Technology and Digital Innovation, Ministry of Science, Technology and Innovation, Brazil (5 minutes)

Comments: Representatives of the Government of Chile (5 minutes)

Annex 2: Regional innovation indices

- "Índice Departamental de Innovación para Colombia," Mr. Camilo Rivera, Director of Innovation and Entrepreneurial Development, National Planning Department, Colombia https://ocyt.org.co/wp-content/uploads/2022/04/IDIC_2021_Documento.pdf
- "China's Regional Innovation Capability Index," Chinese Academy of Science and Technology for Development, People's Republic of China
- "European Union Regional Innovation Scoreboard," European Commission, Brussels https://research-and-innovation.ec.europa.eu/statistics/performance-indicators/regional-innovation-scoreboard_en
- "India Innovation Index," National Institution for Transforming India (NITI) Aayog, India www.niti.gov.in/sites/default/files/2022-07/India-Innovation-Index-2021-Web-Version_21_7_22.pdf
- Viet Nam "Provincial Innovation Index (PII)," Viet Nam Institute for Science, Technology and Innovation, Ministry of Science and Technology, Viet Nam (unpublished pilot version), https://vietnamnews.vn/opinion/1479577/viet-nam-takes-local-approach-to-innovation-based-growth.html

References

Andrews, R. J., Fazio, C., Guzman, J., Liu, Y., & Stern, S. (2022). The startup cartography project: Measuring and mapping entrepreneurial ecosystems. *Research Policy, 51*(2), 104437.

Austin, M. P. A., Marini, M. M., Sanchez, A., Simpson-Bell, C., & Tebrake, J. (2021). Using the Google Places API and Google Trends Data to develop high frequency indicators of economic activity. *International Monetary Fund Working Paper* 21/295.

Baruffaldi, S. H., Di Maio, G., & Landoni, P. (2017). Determinants of PhD holders' use of social networking sites: An analysis based on LinkedIn. *Research Policy, 46*(4), 740–750.

Bergquist, K., Fink, C., & Raffo, J. (2017). Identifying and ranking the world's largest clusters of inventive activity. *WIPO Economic Research Paper* no. 34.

Bibri, S. E., & Krogstie, J. (2017). Smart sustainable cities of the future: An extensive interdisciplinary literature review. *Sustainable Cities and Society, 31*, 183–212.

Coad, A., Daunfeldt, S. O., Hölzl, W., Johansson, D., & Nightingale, P. (2014). High-growth firms: Introduction to the special section. *Industrial and Corporate Change, 23*(1), 91–112.

Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012, February). Social coding in GitHub: Transparency and collaboration in an open software repository. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (pp. 1277–1286).

Gibbons, J., Malouf, R., Spitzberg, B., Martinez, L., Appleyard, B., Thompson, C., & Tsou, M. H. (2019). Twitter-based measures of neighborhood sentiment as predictors of residential population health. *PloS One, 14*(7), e0219550.

Hershbein, B., & Kahn, L. B. (2018). Do recessions accelerate routine-biased technological change? Evidence from vacancy postings. *American Economic Review, 108*(7), 1737–1772.

Kaiser, J. (2017). How biologists pioneered preprints – with paper and postage. *Science, 357*(6358), 1348.

Kalliamvakou, E., Gousios, G., Blincoe, K., Singer, L., German, D. M., & Damian, D. (2014, May). The promises and perils of mining GitHub. In *Proceedings of the 11th Working Conference on Mining Software Repositories* (pp. 92–101).

Kinne, J., & Axenbeck, J. (2020). Web mining for innovation ecosystem mapping: A framework and a large-scale pilot study. *Scientometrics, 125*(3), 2011–2041.

Kinne, J., & Lenz, D. (2021). Predicting innovative firms using web mining and deep learning. *PloS One, 1*6(4), e0249071.

Kohns, D., & Bhattacharjee, A. (2023). Nowcasting growth using Google Trends data: A Bayesian structural time series model. *International Journal of Forecasting*, *39*(3), 1384-1412.

Marra, A., Antonelli, P., & Pozzi, C. (2017). Emerging green-tech specializations and clusters: A network analysis on technological innovation at the metropolitan level. *Renewable and Sustainable Energy Reviews, 67*, 1037–1046.

NESTA (2008). Measuring Innovation. *Policy Briefing* MI/25, available at https://media.nesta.org.uk/documents/measuring_innovation.pdf.

Ojanperä, S., Graham, M., & Zook, M. (2019). The digital knowledge economy index: Mapping content production. *Journal of Development Studies, 55*(12), 2626–2643.

Pindado, E., & Barrena, R. (2021). Using Twitter to explore consumers' sentiments and their social representations towards new food trends. *British Food Journal, 12*3(3), 1060–1082.

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of open access articles. *PeerJ, 6*, e4375.

de Rassenfosse, G., Kozak, J., & Seliger, F. (2019). Geocoding of worldwide patent data. *Scientific Data, 6*(1), 260.

Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, C. H. (2016). The academic, economic and societal impacts of open access: An evidence-based review. *F1000Research, 5*, 632.

Vines, T. H., Andrew, R. L., Bock, D. G., Franklin, M. T., Gilbert, K. J., Kane, N. C., & Yeaman, S. (2013). Mandated data archiving greatly improves access to research data. *The FASEB Journal, 27*(4), 1304–1308.

Woloszko, N. (2020). Tracking activity in real time with Google Trends. *OECD Economics Department Working Papers* 1634.

Woo, J., & Owen, A. L. (2019). Forecasting private consumption with Google Trends data. *Journal of Forecasting, 38*(2), 81–91.

Yu, S., & Fleming, L. (2022). Regional crowdfunding and high-tech entrepreneurship. *Research Policy, 51*(9), 104348.

Zhang, K., & Moe, W. (2021). Measuring brand favorability using large-scale social media data. *Information Systems Research*, 32(4), 1128–1139.

Cover: Getty Images/adventtr / WIPO Reference RN2023-27EN, DOI: 10.34667/tind.48577

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

wipo.int