



# HSE GLOBAL CITIES INNOVATION INDEX

Methodology Discussion



**RUSSIAN CLUSTER OBSERVATORY**  
Moscow 2022



# AGENDA

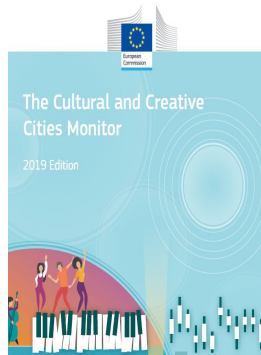
- Challenges of measuring innovation at the city level
- Theoretical framework for measuring innovations used by the Russian Cluster Observatory
- HSE Global Cities Innovation Index 2020
  - System of indicators
  - Sample of cities
- HSE Global Cities Innovation Index 2022
  - Amended system of indicators
  - Changing the approach to city sampling
  - Approaches to identifying agglomerations
  - Publication and patent analysis methodology
- Q&A session
- Questions for discussion

# CHALLENGES OF MEASURING INNOVATION AT THE CITY LEVEL



No unanimous position among researchers on the content and measurement methods of cities' innovation development trends

Formation of isolated ratings for technological and digital development, creative potential and infrastructure



Lack of reliable data sources for international comparisons of cities by their innovation development

The use of a small number of indicators reflecting the results of scientific activity (patents and publications), or reliance on unverifiable expert assessments and surveys



No unified concept among countries on what constitutes a city/agglomeration

Comparison of "convenient" cities (London, New York, Tokyo, Paris, etc.), ignoring other real competitors in the field of innovation (i.e. Silicon Valley)



**Result:** lack of a comprehensive vision of the objective comparative advantages of innovation centers → opportunities to significantly improve the quality of strategic planning and offer more specific tactical solutions for city managers

# OUR APPROACH IS BASED ON THE CONCEPT OF THE SUPERSTAR ECONOMY



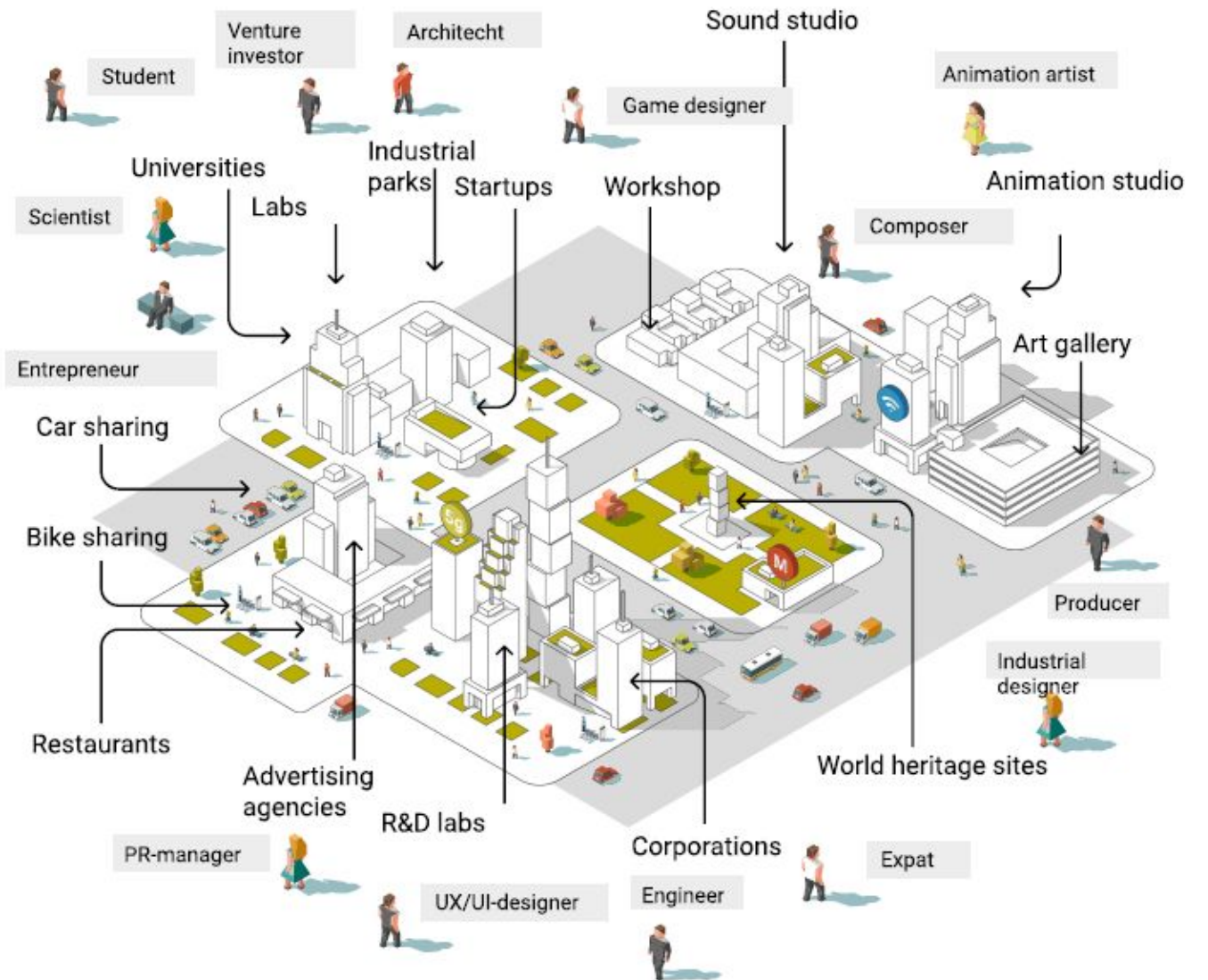
1. Relatively small numbers of people earn enormous amounts of money (Rosen, 1981). These people may be called A-list (Caves, 2000) or superstars
2. Superstar people tend to be concentrated in superstar countries
3. Superstar countries enjoy most of the gains from progress, with other countries being increasingly left behind (Korinek and Xuan, 2018)
4. The rise of human superstars is just beginning (Korinek and Xuan, 2018)

# CITY INNOVATIONS SHOULD BE MEASURED USING AN ECOSYSTEM APPROACH



Innovation ecosystem of a global city involves creators of products and technologies attracted by **recognized leaders of the knowledge economy (superstars)**

Last but not least, there is an **advanced infrastructure and friendly environment in such cities**





## HSE GCII 2020

# DATA COLLECTION PRINCIPLES

- **Transparency** – use of open international databases
- **Verifiability** – refusal to use "internal" data of city administrations on various aspects of innovative development, inaccessible to a wide range of users
- **Comparability** – the data used allows for the most objective comparison of cities
- **Objectivity** – rejection of opinion polls or expert interviews



## Technological Development

Fortune Global 500 Innovation 1000  
 Crunchbase StartupBlink  
 Web of Science PatStat Global  
 QS THE ARWU



## Creative Industries

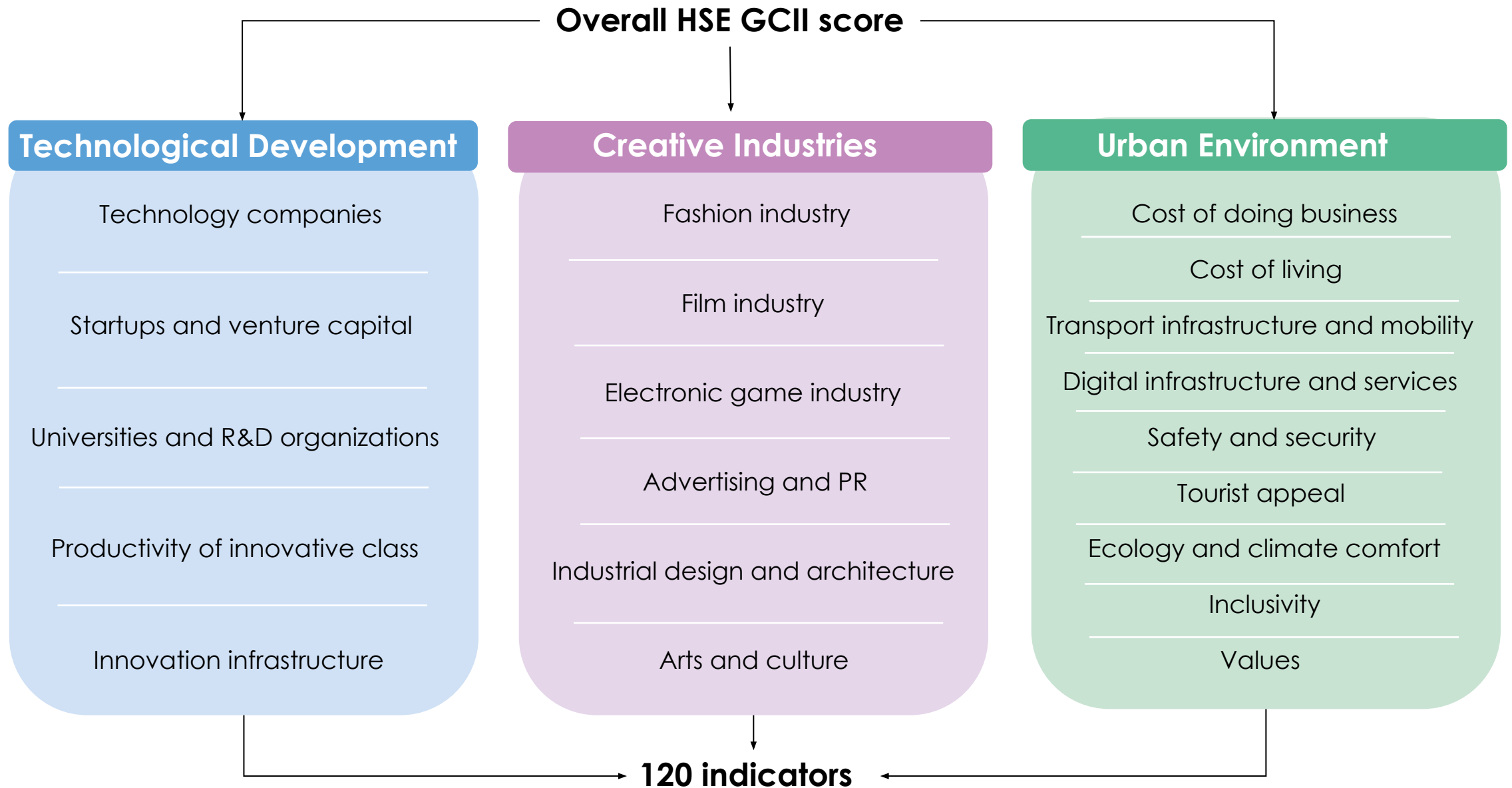
FARFETCH Fashion United IMDb Reddot  
 Effie Awards Spotify Provoke Media  
 Steam The Game Awards Spotify if  
 World Architecture Festival (WAF) Artprice  
 Cannes Lions Pritzker Prize Wikipedia



## Urban Environment

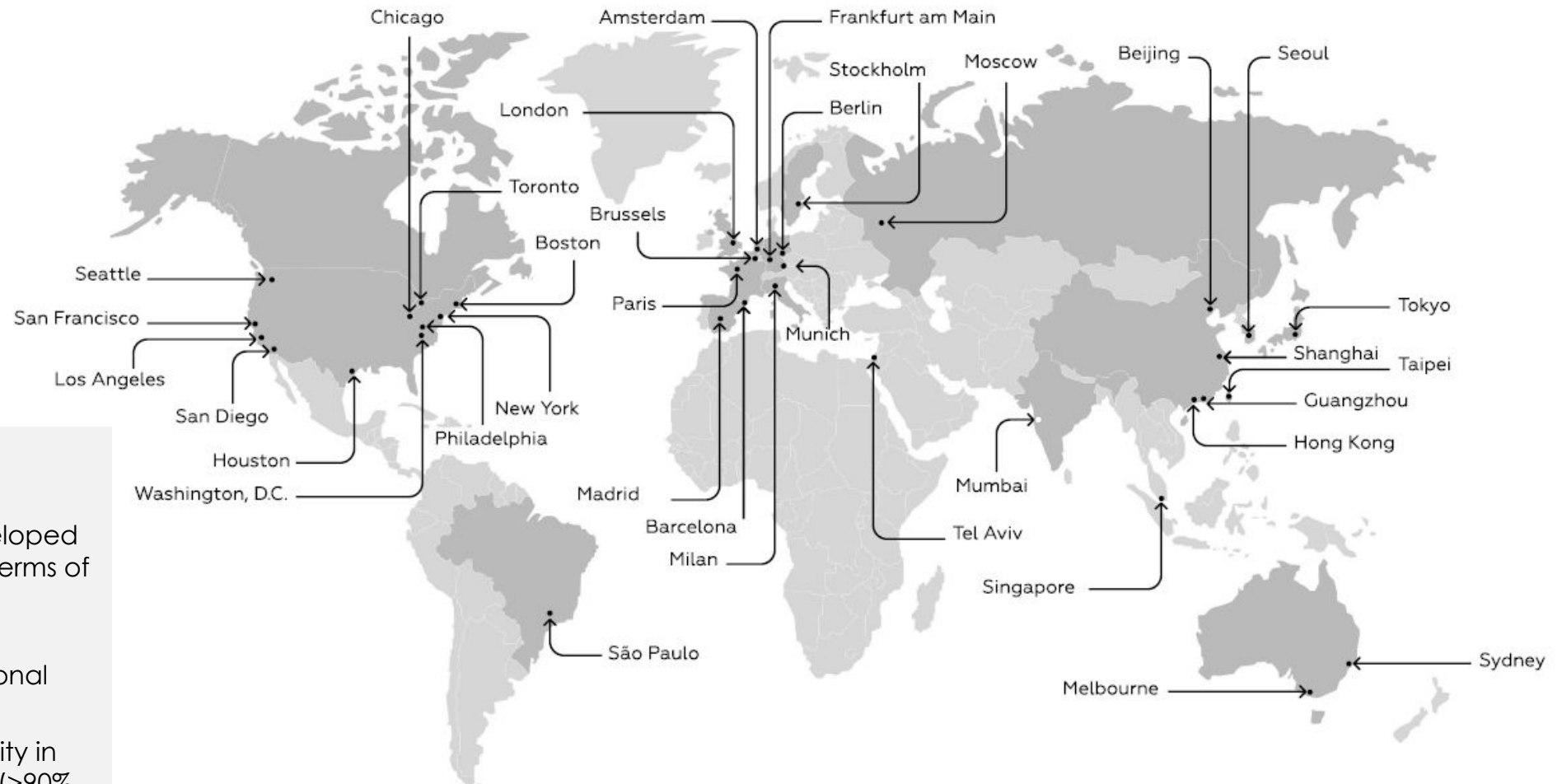
PassportEuromonitor TripAdvisor  
 OpenFlights WiFi Map Numbeo  
 Nomad List STC Database World Metro Database  
 World Value Survey

# SYSTEM OF INDICATORS





# SAMPLE: 36 CITIES THAT ARE THE WORLD'S LEADING CENTERS OF INNOVATION



## SELECTION CRITERIA

1. Leading cities in the key developed and developing countries in terms of the number of patents and publications
2. Presence of a city in international rankings on relevant topics
3. Completeness of data on a city in research information sources (>90% of indicators)

Each indicator's absolute score was normalized using formula (1) or formula (2), depending on the indicator's effect on the overall GCII index

**The values of sections, sub-indices and the integral HSE GCII are calculated through indicators**

$$x_i^{\text{norm}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

a higher value of the indicator corresponds to a greater innovative attractiveness

$$x_i^{\text{norm}} = \frac{x_{\max} - x_i}{x_{\max} - x_{\min}} \quad (2)$$

a higher value of the indicator corresponds to a lower innovative attractiveness

**section score**

$$y_i^{\text{norm}} = \frac{\sum_1^n x_i^{\text{norm}}}{n}$$

**sub-index score**

$$z_i^{\text{norm}} = \frac{\sum_1^m x_i^{\text{norm}}}{m}$$

**HSE GCII score**

$$w_i^{\text{norm}} = \frac{\sum_1^k x_i^{\text{norm}}}{k}$$

$x_i$  is a city's indicator score

$x_{\max}$  is the highest indicator score for all cities in the sample

$x_{\min}$  is the lowest indicator score for all cities in the sample

$i$  is the number of a city

$y_i^{\text{norm}}$  is the section score normalized using formula (1) or (2)

$z_i^{\text{norm}}$  is the subindex score normalized using formula (1) or (2)

$w_i^{\text{norm}}$  is the integral HSE GCII score is the arithmetic mean of all the normalized indicators in the index  
 $n, m, k$  is the number of indicators in the sections; subindex; overall index

# HSE GCII: 2020 RESULTS



1 – 10

11 – 20

21 – 30

31 – 36

New York		1	Singapore		11	Washington, D.C.		21	Houston		31
London		2	Berlin		12	Chicago		22	Seattle		32
Tokyo		3	Hong Kong		13	Munich		23	Philadelphia		33
Beijing		4	Sydney		14	Milan		24	Frankfurt am Main		34
Paris		5	Barcelona		15	Taipei		25	Tel Aviv		35
Los Angeles		6	Stockholm		16	São Paulo		26	Brussels		36
San Francisco		7	Boston		17	Guangzhou		27			
Moscow		8	Toronto		18	Amsterdam		28			
Shanghai		9	Melbourne		19	San Diego		29			
Seoul		10	Madrid		20	Mumbai		30			

# NOT JUST A RANKING OF CITIES, BUT A TOOL FOR POLICYMAKERS



## Beijing

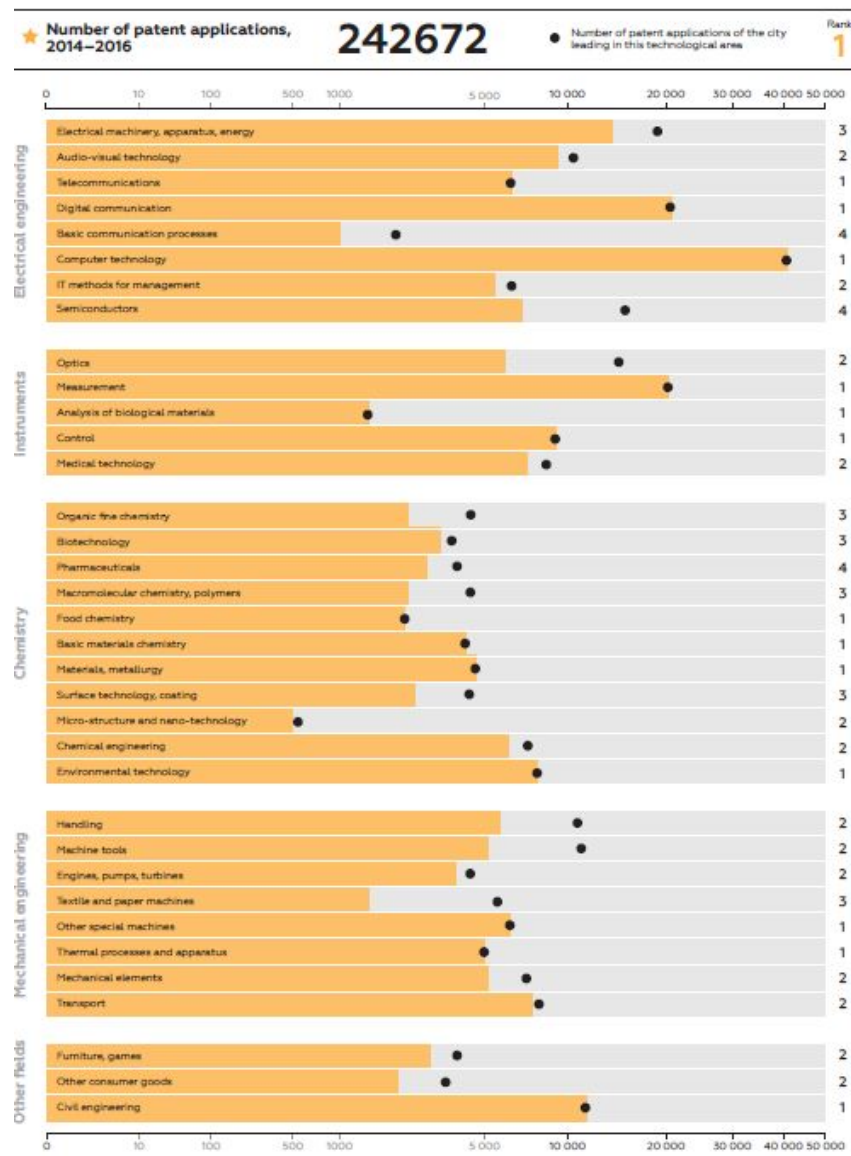
4

4

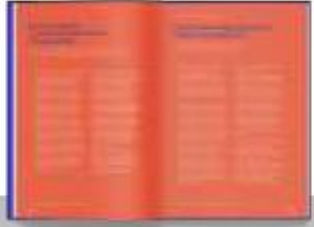
## Beijing

Technological Development	Rank 2	Creative Industries	Rank 6	Urban Environment	Rank 11
Technology companies	2	Fashion industry	23	Cost of doing business	13
Largest companies	2	Largest fashion companies	15-23	Estimated tax	15-17
Representative offices of largest companies	4	Fashion brands	26-28	Office space rental cost	32
R&D departments of largest companies	3	Fashion designers	12-14	Co-working space rental cost	18
Revenues of largest companies	1	Film industry	16	Salary	5-6
R&D expenditure of largest companies	3	Highest-rated films (audience)	14	Cost of living	7
Startups and venture capital	4	Highest-rated films (critics)	10-11	Food prices	8
Startups	18	Award-winning films in international film festivals	6	Apartment rental cost	9
Unicorn startups	1	Popular filming locations	18-19	Cost of living for an expat	9
Business angels	26	Electronic game industry	14-17	Cost of living for a local resident	7
Innovation support funds	4	The Game Awards winners	7-18	Transport infrastructure and mobility	1
Venture capital investment	5	Developers of most popular computer games	10-36	Air traffic	6
Juveniles and R&D organizations	3	Developers of most popular mobile games	11-20	Metro network length	1
Leading universities	2-3	Companies participating in video game trade shows	11-20	Bike-sharing stations	1
Leading R&D organizations	1	Advertising and PR	5	Commute time	30
Highly cited researchers	1	Leading advertising agencies	4	Digital infrastructure and services	14
Nobel Prize laureates and Fields Medal winners	12-16	Leading PR agencies	19-20	Wireless internet	25
Students	2	Industrial design and architecture	9	5G deployment	7-12
International students	9	Leaders of industrial design and architecture	6-7	Use of open source data	21-23
Domestic faculty staff	3	Internationally recognized architects	13-36	Car-sharing	23-27
International faculty staff	9	Arts and culture	4	Safety and security	27-29
Leading business schools	11-22	Internationally recognized artists	20-36	Homicide rate	15-17
Productivity of innovative class	1	Most influential people in contemporary art	6	Road traffic mortality rate	32-34
Population with higher education	18	Most collectible living artists	1-2	Natural disaster risk	26-28
Patent activity	1			Tourist appeal	6
Publication activity	1			Tourists	28
Innovation infrastructure	32			International hotels	2
Clusters	21-36			Culture and entertainment	9
Technology and science parks	8-9			Ecology and climate comfort	34
Incubators	28-30			Environmental pollution level	36
Co-working spaces	29-32			Annual mean air temperature	21-22
				Distance to the seacoast	28
				Inclusivity	32
				Multiculturalism	30
				Gender balance	29
				Values	16
				Strength of personality	24
				Tolerance	11
				Interpersonal trust	24
				Institutional trust	2
				Public attitudes towards science and technology	31

Benchmarking against the leading city



# BEST PRACTICE CASES



What unites  
Brooklyn  
and Outer London?



Barcelona – cluster  
capital of the world



From intricate  
calligraphy to cool  
baby monitor



Moscow goes digital  
on its way to the  
smart future



The Land of the Rising  
Technologies



Technological “front men”:  
how can one company  
propel the city to the top



The “dream factory”  
is branching out



A single innovator  
from everywhere:  
multicultural Toronto  
awaiting talents

Presented only in the Russian version of the HSE GCII 2020



# HSE GCII 2022

# CURRENT CHANGES



1. Improved system of indicators
2. More reasonable approach to city sampling
3. Transition from cities to agglomerations
4. Formation of long and short lists of the HSE GCII (under discussion)

# CHANGES IN THE SYSTEM OF INDICATORS: STATISTICAL AUDIT



## HSE GCII 2020 database statistical audit

1. Coefficients of kurtosis and skewness
2. Sensitivity of the integral rank of the city to changes in individual indicators
3. Correlations between indicators, sub-indices and the integral HSE GCII



## The indicators excluded

1. Revenues of largest companies
2. Domestic faculty staff
3. International faculty staff
4. Fashion designers from Big 4 Fashion weeks
5. Highest-rated films (critics)
6. Gender balance
7. A number of "Values" indicators

Note: Sections 3.5 "Safety", 3.7 "Ecology and climate comfort", 3.9 "Values", as well as a number of indicators of the "Inclusion" section are also subjects to exclusion. The decision on these sections will be made after additional statistical analysis of the updated database



# CHANGES IN THE SYSTEM OF INDICATORS: EXPERT DISCUSSIONS



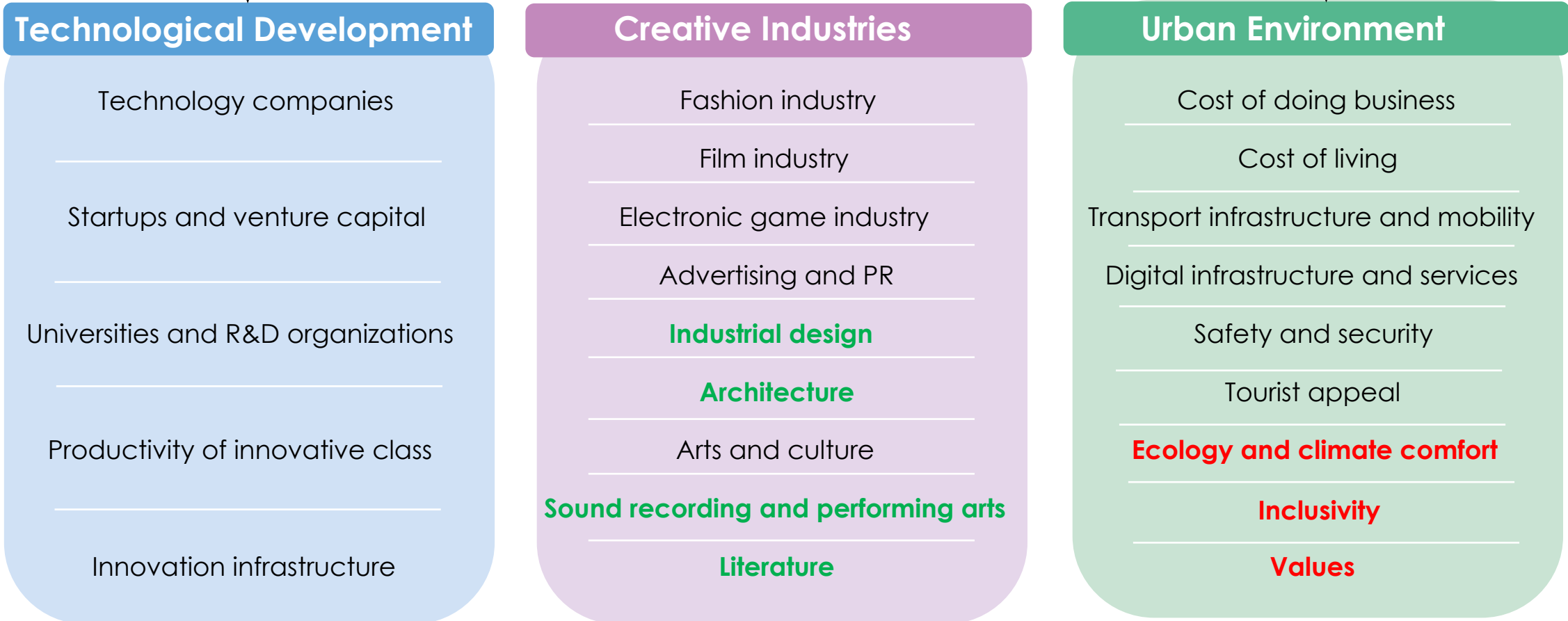
**Based on the results of discussions with experts (creative communities, researchers, etc.), the following changes have been made:**

1. New source for "Technology companies" section (R&D scoreboard 2500 instead of Fortune Global 500 and Global Innovation 1000)
2. New indicator "R&D Intensity of innovative companies"
3. The "Unicorns" indicator uses an additional data source – CBInsights (in addition to Crunchbase)
4. In the "Creative industries" sub-index new sections "Sound recording and performing arts" and "Literature" were added. Section "Architecture and Industrial Design" was divided into two: "Architecture" and "Industrial Design"

# SYSTEM OF INDICATORS



Overall HSE GCII score



Excluded  
Included

# SELECTION OF INDICATORS ON THE BASIS OF WHICH THE RESEARCH SAMPLE IS FORMED



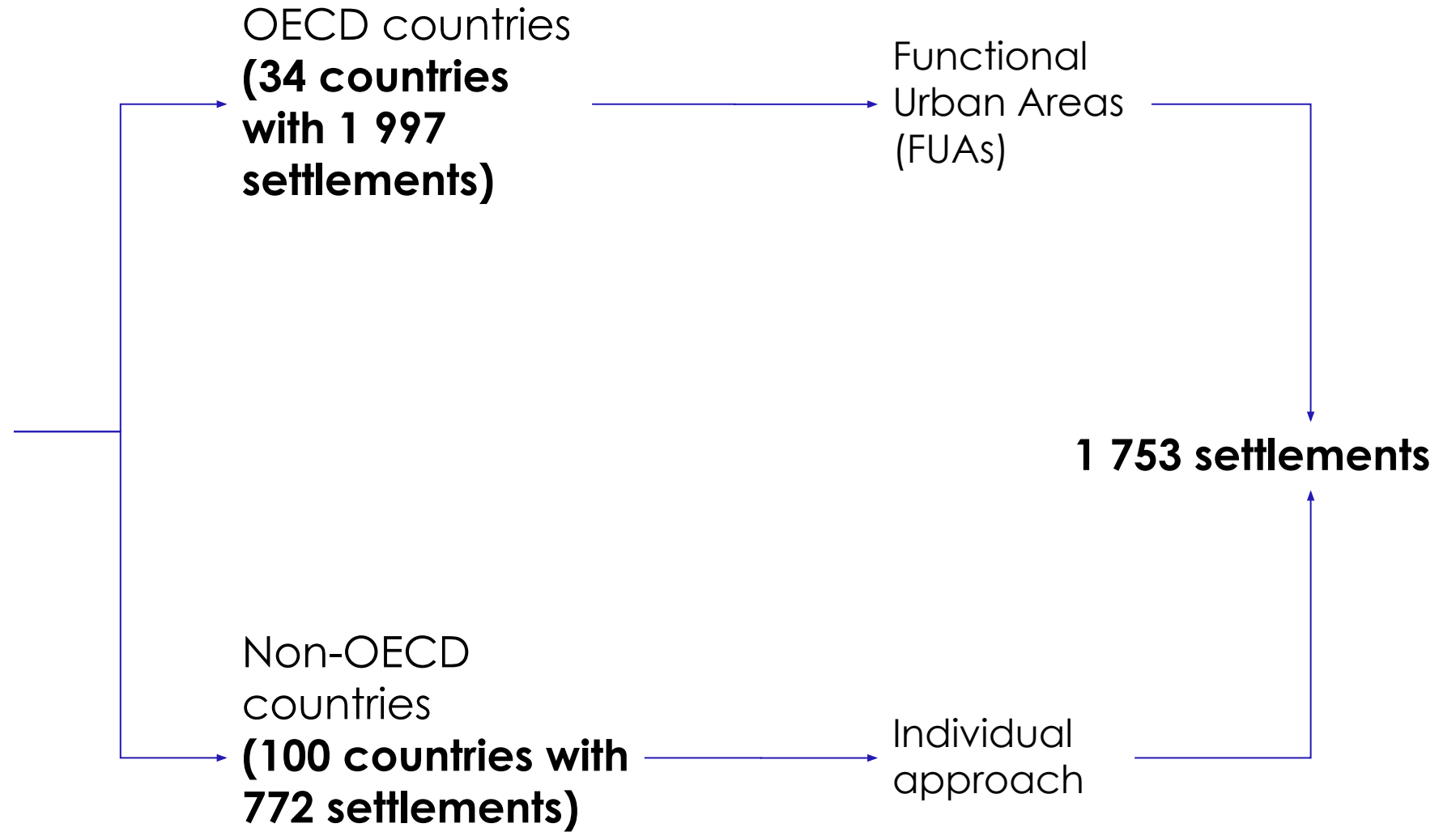
## A set of HSE GCII 2022 indicators collected for all possible settlements:

1. Largest companies (2 500 observations)
2. Unicorns (1 302 observations)
3. Leading universities (2 051 observations)
4. Highly cited researchers (6 332 observations)
5. Nobel Prize laureates and Fields Medal winners (384 observations)
6. Leading business schools (185 observations)
7. Technology and science parks (306 observations)
8. **All "Creative Industries" sub-index indicators.** For example, "Fashion brands" (2 589 observations), "Highest-rated films" (913 observations), "Cybersport Tournaments" (343 observations), "Most traded living artists" (206 observations)

# SAMPLE



Collection of data on the selected indicators identified **2 769** unique settlements in **134** countries

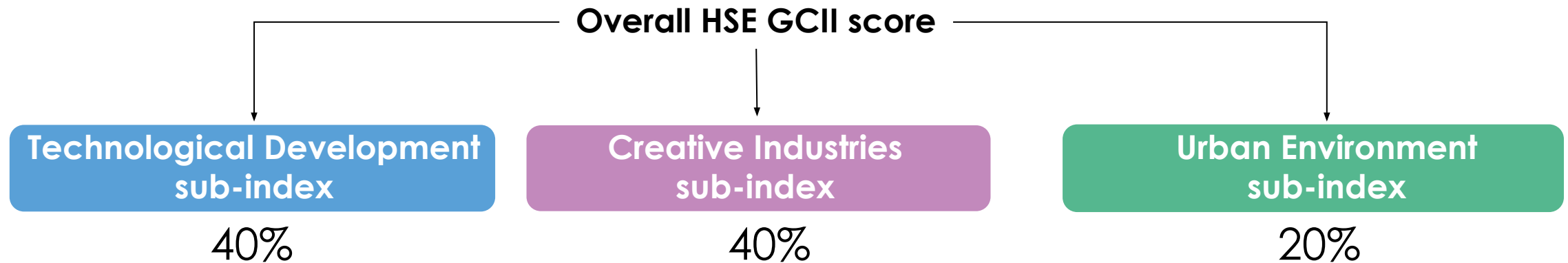


# METHODOLOGY (UNDER DISCUSSION)

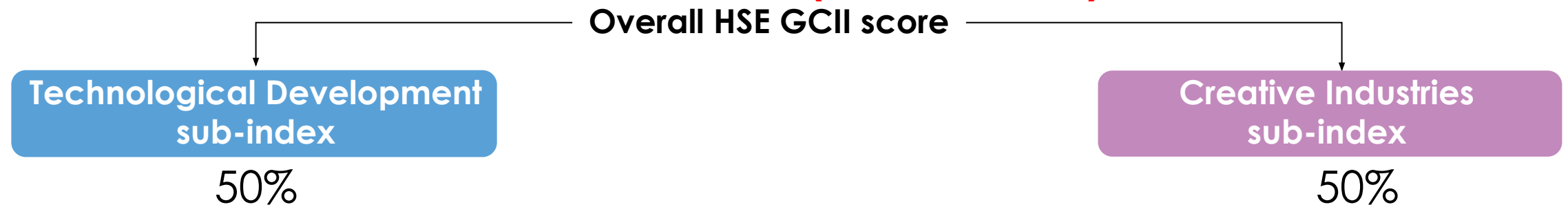


1. Calculation of the integral index through sub-indices
2. Defining weights for sub-indices

## FOR SHORT-LIST CITIES (204 settlements)



## FOR LONG-LIST CITIES (1753 settlements)



# PUBLICATION ACTIVITY ANALYSIS: BASIC PRINCIPLES

Database for analysis: Scopus

Time period: 2010-2021

Document types included: articles, reviews, proceeding papers, books, book chapters, letters, notes

In most cases we run the search of names of cities, towns, villages (and other settlements) included in the studied sample of agglomerations in AFFILCITY() search field in Scopus (i.e. among city names automatically detected by Scopus).

Example for Beijing:

PUBYEAR > 2009 AND PUBYEAR < 2022

AND DOCTYPE ( ar OR re OR cp OR ch OR bk OR le OR no )

AND AFFILCITY ( "Beijing" OR Peking OR "Langfang" OR "Zhuozhou" )

AND AFFILCOUNTRY ( China )



Scopus

Search Sources Scival

1,482,249 document results

PUBYEAR > 2009 AND PUBYEAR < 2022 AND DOCTYPE( ar OR re OR cp OR ch OR bk OR le OR no ) AND AFFILCITY("Beijing" OR peking OR "Langfang" OR "Zhuozhou") AND AFFILCOUNTRY ( china )

# SEARCH FOR PUBLICATIONS OF A SPECIFIC CITY: HOW DOES IT WORK



Take Beijing as example:

Scopus Query search: PUBYEAR > 2009 AND PUBYEAR < 2022 AND DOCTYPE ( ar OR re OR cp OR ch OR bk OR le OR no ) AND AFFILCITY ( "Beijing" OR Peking OR "Langfang" OR "Zhuozhou" ) AND AFFILCOUNTRY ( China )

Bai Yanan marked the affiliation as: University of Chinese Academy of Sciences, **Beijing**, 100049, China

**Beijing name of city was correctly determined by AFFILCITY() Scopus search field** and this publication was counted as publication if Beijing.

We do not know if Bai Yanan is really working in University of Chinese Academy of Sciences in Beijing. Possibly affiliation with this university is only formal for this author.

Nevertheless we count all publications where Beijing name of city was determined by AFFILCITY() Scopus search field as publications of Beijing.

Sc

*KSII Transactions on Internet and Information Systems* • Open Access • Volume 15, Issue 12, Pages 4345 - 4363 • 31 December 2021

## Privacy-preserving and Communication-efficient Convolutional Neural Network Prediction Framework in Mobile Cloud Computing

Bai, Yanan<sup>a, b</sup> ✉ ; Feng, Yong<sup>a</sup>; Wu, Wenyuan<sup>a</sup>

 [Save all to author list](#)

<sup>a</sup> Chongqing Key Laboratory of Automated Reasoning and Cognition, Chongqing Institute of Green and Intelligent Technology, Chinese Academy Sciences, Chongqing, 400714, China

<sup>b</sup> University of Chinese Academy of Sciences, Beijing, 100049, China



Scopus

[Q Search](#) [Sources](#) [Scival](#)

1,482,249 document results

PUBYEAR > 2009 AND PUBYEAR < 2022 AND DOCTYPE(ar OR re OR cp OR ch OR bk OR le OR no) AND AFFILCITY("Beijing" OR peking OR "Langfang" OR "Zhuozhou") AND AFFILCOUNTRY(china)

# PUBLICATION ACTIVITY ANALYSIS: METHODOLOGICAL ASPECT OF SEARCH OF CITIES



When we have **two cities with the same names located in different countries** (e.g. Cambridge, Massachusetts, USA and Cambridge, East of England, UK) we use AFFIL() search field and country name restriction.

When we have **two cities with the same names located in different states of the USA** (e.g. Wilmoington, Massachusetts, USA (Boston agglomeration) and Wilmington, Delaware, USA (Philadelphia agglomeration)) we use AFFIL() search field and state name restriction.

See the example of Boston and Philadelphia agglomerations:

```
Pubyear > 2009 and Pubyear < 2022 and doctype(ar OR re OR cp OR ch or bk or le OR no) And ( Affilcity( "Waltham" OR "Medford" OR "Andover" OR "Bedford" OR "Chestnut Hill" OR "Framingham" OR "Watertown" OR "Billerica" OR "Beverly" OR "Lowell" OR "Somerville" OR "Walpole" OR "Wellesley" OR "Acton" OR "Braintree" OR "Chelmsford" OR "Danvers" OR "Maynard" OR "Natick" OR "Newton" OR "North Reading" OR "Westford" OR "Woburn" ) OR Affil("Cambridge" OR "Boston" OR "Marlborough") and (U.S. OR US or USA or "United States") ) OR Affil("Wilmington" and (MA or Massachusetts) ) ) AND affilcountry( "United States" )
```

```
Pubyear > 2009 and Pubyear < 2022 and doctype(ar OR re OR cp OR ch or bk or le OR no) And( affilcity( "Philadelphia" OR "Princeton" OR "Malvern" OR "Camden" OR "Exton" OR "King of Prussia" OR "New Castle" OR "Swarthmore" OR "Ambler" OR "Pennsauken" OR "Wynnewood" OR "Audubon" OR "Collegeville" OR "Conshohocken" OR "Ewing Township" OR "Glen Mills" OR "Plymouth Meeting" OR "Radnor" OR "Yardley" OR "Pottsgrove" OR "Solebury Township" OR "Bryn Mawr" OR "Glassboro" OR "Haverford" ) OR affil("Wilmington" and (DE or Delaware)) OR affil(Wayne and (PA or Pennsylvania)) ) AND affilcountry("United states" )
```

We take **all variants of names of key cities of agglomeration on English and national languages** (as well as transliterated names)

Example of Munich, Germany:

```
Pubyear > 2009 and Pubyear < 2022 and doctype(ar OR re OR cp OR ch or bk or le OR no) And affilcity("Munich" OR München OR Munchen OR Minga OR "Garching" OR "Gilching" OR "Unterföhring" OR "Landsberg am Lech" OR "Martinsried" OR "Neubiberg" OR "Planegg" OR "Stockdorf" OR "Taufkirchen" OR "Oberhaching" OR "Olching" ) AND affilcountry(Germany )
```





# PATENT ANALYSIS: SEARCHING STRATEGIES

Database: PatStat Global

Time period: 2010-2021

We counted patent applications which, unlike patents granted, provide an up-to-date picture of the current situation without significant time lag.

## **Step 1. Searching by city**

City search in the address and in the name of the organization, excluding double counting

Country code restriction

## **Step 2. Searching by alternative city names**

person\_address like '%Hangzhou%' or '%Hangchow%' or '%Hángzhōu%'



## PATENT ANALYSIS: ASSESSMENT OF THE REPRESENTATIVENESS OF THE DATABASE

Calculation of target benchmarks for cities, clusters (**Global Innovation Index** databases), and relevant countries (**World Intellectual Property Organization** (WIPO) database) using available data.

Comparison of results with the database for the **HSE Global Cities Innovation Index 2020**.



# Q&A

# QUESTIONS FOR DISCUSSION



1. How do you rate the completeness of the CrunchBase and Startupblich databases for China? Are there more used databases on the Chinese venture capital market?
2. Have you used any additional data sources to assess the representativeness of patent data for GIHI?
3. Could you, please, explain the methodology of Patent collaboration network centrality of the GIHI?
4. We would like to discuss the composition of the Chinese agglomerations we are considering
5. What do you think should be the distribution of weights between the three sub-indices?
6. Should we use any other normalization methods besides minimax?
7. New sources of information and fresh ideas for the case studies describing the best practices of innovative development of Chinese cities, accompanied by HSE GCII 2022 data
8. We found that Chinese cities set R&D spending and patents as indicators in their development strategies. Do Chinese cities have innovation and science policy powers?