



ECOSYSTEM HEALTH CHECK

THE GUIDE

How to select the right proxies for your ecosystem

THE GUIDE

HOW TO SELECT THE RIGHT PROXIES FOR YOUR ENTREPRENEURIAL ECOSYSTEM

This guide is part of the 'Ecosystem Health Check' toolkit, which is designed to help ecosystem builders in their work. It does so by providing validated insights that can be used to check whether an entrepreneurial ecosystem project is on the right track to achieve systemic growth and development.

The toolkit consists of:

- The FRAMEWORK – the theoretical foundations on which the tool is based;
- The GUIDE – a practical set of information, including two case studies on how to select the right metrics;

- The TOOL – an interactive dashboard that can help to track data and monitor performance over time.

Support:

The 'Ecosystem Health Check' toolkit has been developed as part of the Credit Suisse – Swisscontact (CSSC) Initiative, funded by the Credit Suisse Financial Inclusion Initiative (FII) and implemented by Swisscontact in collaboration with Startup Heatmap Europe.

Authors:

Swisscontact

Benjamin Meier, Global Project Manager CSSC,
Swisscontact, Switzerland
beny.meier@swisscontact.org | [LinkedIn](#)

Matthias Grawehr, Project Assistant,
CSSC, Swisscontact, Switzerland
matthias.grawehr@swisscontact.org | [LinkedIn](#)

Ariane Appel,
CSSC, Swisscontact, Switzerland
ariane.appel@swisscontact.org

Startup Heatmap Europe

Thomas Kösters,
President Startup Heatmap Europe, Germany
thomas@europeanstartupinitiative.eu | [LinkedIn](#)

Erica Santini,
Head Research Startup Heatmap Europe, Italy
erica@europeanstartupinitiative.eu | [LinkedIn](#)

Support:

This publication is a product of the Credit Suisse – Swisscontact (CSSC) Initiative, funded by the Credit Suisse Financial Inclusion Initiative (FII) and implemented by Swisscontact in collaboration with Startup Heatmap Europe.

BACKGROUND INFORMATION ON THE ORGANIZATIONS:

SWISSCONTACT:

[Swisscontact](#) is a leading organization for the implementation of international development projects. We promote inclusive economic, social and environmental development to make an effective contribution towards sustainable and widespread prosperity in developing and emerging economies.

In cooperation with Credit Suisse, Swisscontact launched the [CSSC](#) Initiative that aims to identify best practice approaches to the promotion of entrepreneurship and ecosystem building in emerging markets and to share our insights with the global community.

Swisscontact has been pioneering the use of social network analysis (SNA) to map and better understand the dynamics within entrepreneurial ecosystems. Through our research, we provide guidance on how to grow ecosystems and make sure information travels fast and effectively. While the SNA studies provide in-depth analysis on ecosystem dynamics, the ecosystem health framework presented here has been designed as a comprehensive self-assessment tool for ecosystem builders to measure the health and equity of an entrepreneurial ecosystem over time.

For more information on our SNA approach and in-depth reports on Kigali (Rwanda), Kampala (Uganda) and Phnom Penh (Cambodia), visit www.sna-mapping.org.

STARTUP HEATMAP EUROPE:

The [Startup Heatmap](#) aggregates data on start-up ecosystems in Europe and beyond. It offers a versatile dataset of >100 variables at city level, ranging from investments and job creation to meetup activities and purpose orientation. The unique focus on the city as a unit of analysis positions the Heatmap as an ideal partner to ecosystem builders in regional development, municipalities and international collaboration.

The [DEEPSEA Ecosystem Accelerator](#) program has trained more than 180 ecosystem leaders on three continents in a multi-week program. The team behind consists of practitioners, investors and entrepreneurs. Together with Swisscontact, their mission is to provide ecosystems all around the world with a self-assessment tool to evaluate the health of and growth opportunities for their start-up communities.

LIST OF CONTENTS

1. INTRODUCTION	6
2. TWO CASE STUDIES SHOWING HOW TO USE EE HEALTH CHECK	7
2.1. Data Sources	7
2.2. Data Tool	9
3. THE CASE OF COLOGNE'S HIGH-GROWTH ECOSYSTEM	10
3.1. Community	10
3.1.1. Pool of Entrepreneurs	11
3.1.2. Entrepreneurship Support Network	11
3.1.3. Proxies selected for Community	12
3.2. Environment	12
3.2.1. Social Capital	12
3.2.2. Governance	13
3.2.3. Proxies selected for Environment	15
3.3. Performance of Cologne's ecosystem	15
4. THE CASE OF GUATEMALA'S INCLUSIVE AND EQUITABLE ECOSYSTEM	17
4.1. Community dimension	17
4.1.1. Pool of Entrepreneurs	17
4.1.2. Entrepreneurship Support Network	18
4.1.3. Proxies selected for Community	20
4.2. Environment	20
4.2.1. Social Capital	21
4.2.2. Governance	21
4.2.3. Proxies selected for Environment	22
4.3. Performance of Guatemala's ecosystem	22
5. CONCLUSIONS	24

1. INTRODUCTION

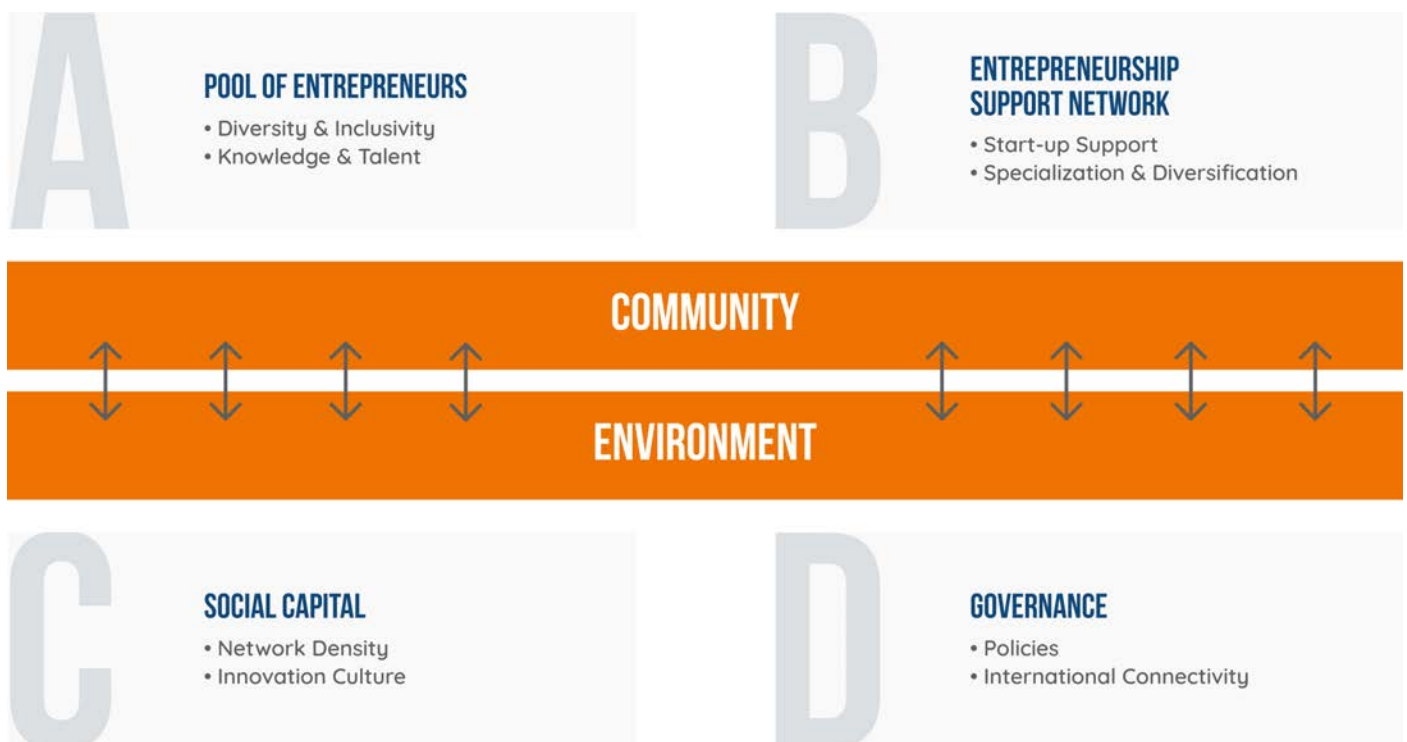
Monitoring conditions in entrepreneurial ecosystems (EE) and developing coherent interventions are vital steps to improve overall ecosystem health.

Ecosystem builders need to check whether an EE project is on the right track. To do so, they require metrics that help them to identify the incremental steps influencing the functioning of their EE. In this way, it is possible to adjust current interventions, suggest new strategies to achieve intermediate goals, or reroute the ecosystem onto a more sustainable development path.

However, tracking the development of ecosystems is difficult due to a lack of key performance indicators (KPIs) that clearly show the success achieved with specific interventions. Especially in early-stage ecosystems, traditional KPIs such as funding numbers are low and volatile, employment statistics are hard to obtain, and the opinions of ecosystem founders are difficult to assess.

Introducing standardized criteria to measure and diagnose ecosystems can help to create strong theoretical foundations that reduce the uncertainty facing ecosystem builders and increase the impact of their actions and resources, as well as the relevance of their work.

In our framework, we propose a gradual fragmentation of the ecosystem's components, allowing for the definition of metrics that are measurable and comparable over time. In an ideal world, it would be possible to measure the identified metrics. However, these metrics are generally not measurable using a reasonable amount of resources. Consequently, it is more convenient to use a proxy that provides reliable information about the metrics we are assessing. The components were selected in an iterative process, starting with a literature review and focus groups with ecosystem builders in Europe, Latin America and East Africa.



2. TWO CASE STUDIES SHOWING HOW TO USE THE EE HEALTH CHECK

Not all ecosystems are the same. In our framework, we explore the difference between ‘high-growth ecosystems’ (HGE) and ‘inclusive and equitable ecosystems’ (IEE). They represent the two broad directions that ecosystems can take, with one built around the logic of venture capital with all its upsides and downsides, and the other based on slower but organic growth. This distinction is in no way a judgement of the quality of the respective ecosystems; instead, it reflects the decision about which factors to examine when checking the health of an ecosystem in relation to its ambitions. We selected two cases as the basis to demonstrate the data collection process in both scenarios: Cologne represents an HGE, while Guatemala City’s focus is on developing a more inclusive and equitable ecosystem (IEE).

Cologne is the fourth-largest city in Germany in terms of population (1,085,664 inhabitants in 2019¹) and it is located in the North Rhine-Westphalia federal state. Its most important industrial sectors are the automotive industry, the chemical industry, mechanical engineering, metal-production and the food industry. It also has a well-established service sector focusing in particular on the insurance and media industries². According to a study by the Institut der deutschen Wirtschaft (German Economic Institute), the city has a strong ICT market, employing around 25,000 people and generating a turnover of approximately EUR 6 billion.

Guatemala City is located in Central America, and its metropolitan area has a population of approximately 2.9 million people³. It is the capital of Guatemala, which borders on Mexico, El Salvador and Honduras. The most relevant industrial sectors in Guatemala are textiles, furniture, petrol, sugar cane and refined sugar, flowers, fruit and vegetables, processed foods and chemical products⁴. The service sector is another important indus-

try in Guatemala, and it includes sectors such as tourism and hospitality, customer service, finance and banking, communications, and retail, which collectively account for about 60% of Guatemala’s total industrial activity⁵. Guatemala is an established near-shoring destination for the US market and for the call center industry, which taps into a pool of over 17,000 university graduates who are all required to be proficient in English when completing a degree⁶. One of Guatemala’s most famous entrepreneurs is the inventor of Captcha and the CEO of Duolingo Luis von Ahn, who left Guatemala to study in the US and has since established several multi-million dollar companies. In contrast, the local start-up scene is still in its infancy but promising initiatives by small venture firms, community groups and also some international corporates (e.g. Telefonica) are entering the local start-up scene.

2.1. DATA SOURCES

Before looking at data collection, it is important to say a few words about methodologies and which sources of data you can use. Information about your ecosystems might be collected from a variety of sources using a range of tools. In this context, the first distinction to be made is between primary and secondary data.

Primary data are collected from primary sources such as interviews, self-administered surveys and questionnaires, experiments, etc. Secondary data are collected from research publications such as the Global Entrepreneurship Monitor or the Startup Heatmap. Of course, secondary data are originally gathered in the form of primary data and then used by other parties as secondary data.

There are advantages and disadvantages to both types of data, as set out in Table 1 below.

1. https://ec.europa.eu/eurostat/databrowser/view/urb_cp0p1...

2. <http://www.startupregion.koeln/industry.html>

3. <https://www.worldscapitalcities.com/capital-facts-for-guat...>

4. <https://www.bizlatinhub.com/es/industrias-grandes-hacer...>

5. <https://www.worldatlas.com/articles/the-biggest-industries-in...>

6. <https://www.workforce.com/news/guatemala-sees-rapid-growt...>

	ADVANTAGE	DISADVANTAGES
PRIMARY DATA	<ol style="list-style-type: none"> 1. These data offer a high degree of accuracy and reliability. 2. For some questions, secondary data are not available. 3. No additional precautions are needed when using these data. 	<ol style="list-style-type: none"> 1. Collecting these data can take a lot of time and potentially also a lot of (financial) resources. 2. Trained personnel are sometimes needed to collect these data. 3. Collecting primary data may not always be possible due to a lack of access to the research target.
SECONDARY DATA	<ol style="list-style-type: none"> 1. These data are easy to collect. 2. The use of these data saves time and money. 3. It is possible to rely on the scientific expertise of the provider. 4. Secondary data may be more reliable than primary data, e.g. due to necessary corrections in the dataset. 	<ol style="list-style-type: none"> 1. It is sometimes hard to see what is behind the numbers and how they were calculated. 2. The purpose of the original data collection may have been different. 3. You are dependent on the publication frequency of a third party and cannot know if research will be continued.

Table 1: Types of data you can use.

There are advantages and disadvantages to both types of data, as set out in Table 1.

Since the measurement of your EE is not a simple undertaking and many metrics, e.g. innovation culture or the quality of policies, cannot be assessed directly, you will probably have to use indirect measurements referred to as 'proxies'. Proxy data are used to study a situation, phenomenon or condition for which no direct information is available. It is possible to use two types of proxy:

- INPUT-related proxies in EE refer to a set of proxies measuring actions and instruments that are in place to achieve a set of goals.
- OUTPUT-related proxies measure the possible outcomes of tangible and intangible mechanisms that are in place in the system.

We propose using a combination of the two types of proxy to make the measurement process more balanced. This also means that the type of proxy that is most readily available can be used for practical reasons.

In order to identify the best proxy describing our metric, we checked the quality and timeliness of the various data sources and addressed questions including: Can we trust the secondary data? Are these data available with sufficient frequency or are they only updated once every three or four years? If you think that your secondary data at city level are not reliable and you cannot find alternative data, you can use a sample. Remember that the size of the sample is very important. It has to be large enough to present an accurate reflection of the reality and also be suitable in order to be generalized at the ecosystem level.



2.2. DATA TOOL

To guide the measurement of the health of your EE, we have developed an online dashboard that guides every step of the process.

You can find the dashboard [here](#). To start using the tool, you first need to create your personal copy of the dashboard.

Steps to complete the data tool::

- Select proxies
- Select timeline
- Select benchmarks

Thanks to this framework, the measurement of your EE using secondary data can be a good option, especially if you choose to measure your ecosystem on an annual basis. By following the correct steps, you will spend around three hours completing all the boxes in the tool based on annual secondary data. However, if you would also like to collect primary data and add more granularity with regard to the timeline of your data, you can do

so on a quarterly basis. We suggest doing this with your statistical unit, an external advisor or, at the very least, a data-savvy member of your team (this might be you) who can collect primary data on a regular basis.

In the following section, we present both the achievements and the different challenges that can emerge during the data collection process for the two different EE.

We tested the framework under real-life conditions and provided two impartial analysts with the framework, the list of metrics and the proxies that we felt adequate to measure the ecosystem health dimensions (see Table 1 in the Annex). We did not explain to the analysts the meaning of the framework, nor did we suggest any collection strategy.

After the first round of data collection, which took an average of three hours, we explored the different results and issues that emerged. In the next subsection comprising the different steps in this process, we will show you how to avoid some of the pitfalls that we discovered and offer various examples of data collection strategies.

3. THE CASE OF COLOGNE'S HIGH-GROWTH ECOSYSTEM

3.1. COMMUNITY

We start by measuring the health of the COMMUNITY (see Figure 1) that will be mapped by looking at two different dimensions that are defined by two metrics:

1. Pool of Entrepreneurs
 - 1.1. Diversity & Inclusivity
 - 1.2. Knowledge & Talent
2. Entrepreneurship Support Network
 - 2.1. Start-up Support
 - 2.2. Specialization & Diversification

3.1.1. Pool of Entrepreneurs

a. Diversity & Inclusivity

The first metric for the Community dimension is Diversity & Inclusivity, which allows you to understand whether the EE is able to embed formal and informal mechanisms fostering inclusion, knowledge sharing, and the emergence and legitimization of new ideas.

In the case of Cologne, we need to choose the right proxy from the list of proxies referring to HGE. The list contains two options to measure Diversity & Inclusivity: Share of female founders and share of foreign-born founders.

At first glance, we might have a preference for one or the other proxy, and there are definitely arguments in favor of either of them – or some local contexts that might even call for further creative ideas for proxies for this metric. The most important thing is to identify adequate proxies that measure how inclusive the local community is to newcomers and outsiders and to en-

sure that you can gather accurate data. In our case, we did not have a preference for either of the two options and we therefore looked at other possible data sources to help reach our decision.

We encourage you to use more than one proxy to estimate the different metrics. However, if not enough data are available, using one proxy is still an option.

If you have sufficient data, you can create an index by, for example, taking the average of the proxies used.

At city level, we did not find reliable data for foreign-born founders. Instead, we found good data on female founders on the Startup Heatmap website⁷, where we checked the set of information available for the city of Cologne. Based on this source, we discovered that from a sample of 111 founders in Cologne in 2020, 6.48% were female. To have an alternative source, we also looked for a sample that would give us more detailed information. The options for this type of smaller but more accurate sample could be the portfolio of one or several accelerator programs or an online survey. In this case, we selected Startplatz Accelerator⁸, which is one of the more established start-up programs in Cologne and has made quite a lot of data about its portfolio available online. This sample gave us an insight into the background of the founders and indicated their gender.

However, the Startplatz Accelerator has 49 start-ups and it is not necessarily possible to derive general insights from it. Consequently, we selected the share of female founders from total founders in Cologne as the proxy for Diversity & Inclusivity.

7. <https://www.startupheatmap.eu/Cologne/>

8. <https://www.startplatz.de/accelerator/>

b. Knowledge & Talent

The second metric Knowledge & Talent describes the density of the talent pool triggering potential network effects and increasing the innovative potential of the ecosystem.

In the list of proxies specific to HGE, we found three options: Number of developers per capita; share of engineers among founders; and number of tech start-ups per capita.

The Atomico State of European Tech Report provides data on the number of developers per city on an annual basis⁹. According to the Atomico website, the number of developers was 162,252 in 2019 and 165,900 in 2018. Since the population of Cologne totaled 1,085,664 in 2019 and 1,080,384 in 2018 according to Eurostat, the number of developers per capita was 15.22% (165,252/1,085,664) in 2019 and 15.36% (165,900/1,080,384) in 2018.

As we feel this source is both reliable and sufficiently up to date, we have not looked for alternatives and have selected it as our proxy. If your ecosystem is not listed on such a database, you could also analyze the portfolio of a specific accelerator and examine the skills profiles of the founders based on their CVs.

3.1.2. Entrepreneurship Support Network

a. Start-up Support

	City level	Sample
Number of start-ups per capita	Good data, available each month	Mediocre (1 accelerator) but only available for 2020
Share of mentors with start-up experience	Mediocre	Good data (1 accelerator with 58 mentors) but only available for 2020
Share of accelerated start-ups receiving follow-on funding	Not available	Mediocre (1 accelerator) but only available for 2020

Table 3: Description of quality of data from data collection in Cologne for Start-up Support.

9. <https://2018.stateofeuropeantech.com/chapter/europes-got-talent..>

10. "Exits" in the startup terminology means the sale of a company, either by founders selling their shares to individual investors (M&A) or going public on the stock market (IPO).

The first metric for the Entrepreneurship Support Network dimension is the actual Start-up support, defined in our framework as the tangible or intangible assets of the EE supporting entrepreneurship.

In the list of proxies specific to HGE, we found four options: Number of start-ups per capita; share of companies with exits¹⁰ from total companies receiving investments; share of mentors with start-up experience¹¹; and share of accelerated start-ups receiving follow-on funding.

The city of Cologne has a database of start-ups on the website <https://startup-map.cologne/>¹², which can be easily divided by the population of Cologne. A total of 991 start-ups are listed, which equates to approximately 0.9 start-ups per 1,000 inhabitants. While this proxy might seem very simple, it is in fact one of the hardest to measure, as it is difficult to trust the data regarding the total number of all start-ups. The question is: Does it really capture all of the companies founded in Cologne?

To have an alternative, we looked again at the information published by the Startplatz Accelerator¹³ that lists all 58 of the mentors online. Based on their CVs that we located on the Internet, we were able to calculate the second possible proxy – the share of mentors with start-up experience. Of the 58 mentors, 31 have clear previous start-up experience (53.4%), which is defined as having previously started their own company.

11. Start-up experience refers to mentors who have founded a company before.

12. https://startup-map.cologne/companies.startups/f/all_locations/all..

13. <https://www.startplatz.de/accelerator/>

Since 58 mentors is not a small sample from which to derive general insights at the ecosystem level, we selected the share of mentors with start-up experience as the proxy for Start-up Support (see Table 3).

b. Specialization & Diversification

The second metric for the Entrepreneurship Support Network dimension is Specialization & Diversification, which describes the level of specialized but also lateral support formats allowing for the exploitation of specialized knowledge and cross-fertilization between various knowledge domains.

In the list of proxies specific to HGE, we found three options: Investments based on growth stages or sectors, and the diversity of firms across sectors.

As the city of Cologne has a detailed investment database, we can find good data for the period from 2012 to 2020 for investments based on verticals on its website. It is therefore obvious that this proxy should be selected for 2020 (i.e. 37 million of start-up investments for 7 sectors) and 2019 (i.e. 67 million for 12 sectors). A more difficult task is to turn these figures into a proxy that shows the level of specialization and diversification in one clear number. The solution is to apply the Herfindahl-Hirschman Index (HHI).

The Herfindahl-Hirschman Index measures the concentration among values of a frequency distribution, thus showing how diversified a set of inputs is. If all investments were made in one sector, the HHI would be 1. If all investments were made in diverse sectors, (N) it would be 1/N.

For Cologne, the HHI was 0.234 in 2020 and is 0.295 in 2019.

3.1.3. Proxies selected for Community

Based on the previous considerations, we want to

present the list of proxies selected for this case study to measure Community.

a. Pool of Entrepreneurs

a.1. Diversity & Inclusivity: share of female founders from total founders in Cologne (6.48%)

a.2. Knowledge & Talent: The number of developers (15.22%)

b. Entrepreneurship Support Network

b.1. Start-up Support: share of mentors with start-up experience (53.40%)

b.2. Specialization & Diversification: Investments based on verticals (0.234)

3.2. ENVIRONMENT

At this point, we start to measure the health of the ENVIRONMENT (see Figure 1), which will be mapped by looking at two different dimensions that are each defined by two metrics:

a. Social Capital

a.1. Network Density

a.2. Innovation Culture

b. The Governance

b.1. Policies

b.2. International Connectivity

3.2.1. Social Capital

a. Network Density

The first metric for the Environment dimension is Network Density, which mirrors the opportunities for stakeholders to collaborate and connect within the EE.

In the list of proxies specific to HGE, we found two options: Participants in tech meetups per capita and number of tech meetups per capita.

14. <https://startup-map.cologne/heatmaps/funding/cities..>

The website [meetup.com](https://www.meetup.com) has become a go-to tool for community organizers around the world. While not all community activities are organized via the platform, it provides a good indication of the general level of network density. We can search for the number of tech meetups by selecting Cologne as a city and using 'tech' as a keyword¹⁵. The list shows recurring meetups with

	City-Level	Sample
Participants in tech meetups per capita	Not available	Good data every month
Number of tech meetups per capita	Not available	Good data every month

Table 4: Description quality of data from data collection in Cologne for Network Density

their number of followers. It is possible to obtain information about the number of participants, the topic of the meetups, the location and the date of the event. These data are available over time and it is updated continuously. We therefore believe that the data available on [meetup.com](https://www.meetup.com) provides a good sample for our proxies.

We have selected participants in tech meetups per capita as our proxy. The 38 tech meetup groups in Cologne have a total of 22,754 regular followers. Since Cologne's population totaled 1,118,789 in 2020, there are 20.35 meetup participants per 1,000 inhabitants.

b. Innovation Culture

The second metric for the Social Capital dimension is Innovation Culture, which refers to the set of soft rules in place that prevent moral hazards and opportunistic behavior in open innovation practices.

In the list of proxies specific to HGE, we found three options: Sum of foreign investments per capita; total VC investments per capita; and number of deals EUR > 1 million per capita. Once again, we were pleased to find good data for the period from 2014 to 2020 on the website of the Cologne Startup Map¹⁶.

Consequently, we decided to select as the proxy the

sum of foreign investments per capita. This proxy describes the perception of the innovation culture in Cologne among foreigners, i.e. do they trust the formal and informal institutions of the ecosystem to turn their investments into future returns? This would not be the case if there were a lack of formal protection of innovation or a toxic environment where innovation could not

thrive. We have considered investments coming from outside Germany, since the attraction of foreign interest is of decisive importance in a high-growth ecosystem. Specifically, in 2017 the population was 1,082,000 and foreign investments reached EUR 22.5 million, resulting in an index of 20.80. Based on the same method, the index rose to 36 in 2018, dropped to 13.26 in 2019 and rose again slightly to 16.98 in 2020.

3.2.2. Governance

a. Policies

The first metric for the Governance dimension is Policies, which can be defined as the fundamental institutions and reliable governance system that foster a healthy EE. This metric makes it possible to determine whether entrepreneurship is supported by regulations that reduce corruption, ease business processes and increase partnerships.

In the list of proxies specific to HGE, we found two options: Ratings for ease of doing business and share of start-ups with no international headquarters.

In terms of the share of start-ups with no international headquarters, we found only mediocre data at city level.

15. <https://www.meetup.com/it-IT/fort=recommended..>

16. <https://startup-map.cologne/heatmaps/funding/cities/cologne/indus..>

	City level	Sample
Ratings for ease of doing business	Good data for 2018 and 2020	Not available
Share of start-ups with no international headquarters	Mediocre	Not available

Table 5: Description of quality of data from data collection in Cologne for policies.

Consequently, we checked for data at city level regarding ratings for ease of doing business. Both the Global Entrepreneurship and Development Institute (GEDI) dataset and Startup Heatmap provide this kind of data. However, the GEDI dataset only shows information for 2018. We therefore decided to use as our proxy the ratings supplied by Startup Heatmap in 2020 and 2018. These data indicate that in 2020, 70% of founders rated the ease of doing business in Cologne as high, compared to only 42% in 2018¹⁷.

b. International Connectivity

The second metric for the Governance dimension is International Connectivity, which describes the ability of the EE to connect internationally. The more powerful the brand of a nation or region, the more it will be able to attract international investments and company founders.

In the list of proxies specific to HGE, we found three options: Number of international top start-ups with offices in the city, number of international conference participants in the city, and ecosystem brand and visibility.

We found a good data sample for the number of international top start-ups with offices in the city based on a list of the top 100 start-ups supplied by Startup Europe in 2018¹⁸. After the identification of the top start-ups in Europe, we checked manually how many of them have branches in Cologne. Our findings were that only Hello-Fresh from Berlin also has an office in Cologne. As this data had not been updated since 2018, we looked for an alternative.

We also explored a second option for the proxy by focusing on two local conferences: PIRATE Summit and StartupCon¹⁹. The Startup Heatmap Europe offers data about the number of international followers based on an analysis of their Facebook followers. PIRATE Summit had 52.8% international followers, while the Share for StartupCon is only 9.49%. Once again, however, the data are only available for 2018.

Finally, an annual measure of the brand perception of Cologne's start-up scene by Startup Heatmap Europe called the 'TrustScore' has been available since 2016. It is based on an annual representative survey about the

	City-Level	Sample
Number of international top start-ups with offices in the city	No data available	Good data on top 100 start-ups from 2018
Number of international conference participants in the city	No data available	Mediocre (2 conferences); only available for 2018
Ecosystem brand and visibility	Good data available annually	No data available

Table 6: Description quality of data from data collection in Cologne for International Connectivity.

17. <https://startup-map.cologne/heatmaps/funding/cities/col..>

18. <https://startupeuropepartnership.eu/sep-elite-tech-scaleup-100/>

19. <http://startupsandplaces.com/conference-explorer/>

founders' favorite start-up hubs in Europe. Cologne's TrustScore for the four years listed was as follows: 74.14% (2016), 64.66% (2017), 87.93% (2018) and 81.90% (2019).

Since only one proxy is available at city level over a period of several years, we decided to use the Heatmap TrustScore as the metric for International Connectivity (see Table 6).

3.1.4. Proxies selected for Environment

Based on the previous considerations, we want to present the list of proxies selected for this case study to measure Environment:

- a. The Social Capital
 - a.1. Network Density: Participants in tech meetups per 1,000 inhabitants (20.35)
 - a.2. Innovation Culture: Foreign investments per capita (16.98)
- b. The Government
 - b.1. Policies: Ratings for ease of doing business (70%)
 - b.2. International Connectivity: Heatmap TrustScore (8.90%)

3.3. PERFORMANCE OF COLOGNE'S ECOSYSTEM

The selection of the final proxy marks the completion of the data collection process, meaning that you can proceed to enter information in the data tool. What is still missing is the definition of the benchmarks used to calculate the performance results per metric. It is therefore vital to use realistic and adequate benchmarks. When defining your benchmarks, you can have clear objectives about what you want to achieve and, in this case, the results should be checked against your goal. However, if you do not have clear objectives, you can look at a comparison ecosystem that performs considerably better than yours in a metric and add its actual numbers as a benchmark. An even better approach is to take the median of a group of comparison ecosystems. However, these data might be difficult to find – or not available at all. It is therefore also possible to add an estimate of a realistic improvement in your ecosystem's performance. We indicated the name of our benchmark in the table below to give an idea of how to proceed.

Looking at the benchmarks (Table 7), we see a healthy performance by Cologne's ecosystem in the Community dimension, with only the Diversity & Inclusivity metric significantly underperforming. In contrast, the Environment dimension lags behind its potential – in particular for the metrics of Network Density, Innovation Culture and International Connectivity.

Metric	Proxy	Cologne	Benchmark	Benchmark Details	Result
Diversity & Inclusivity	Share of female founders	6.48%	15.50%	European average	41.81%
Knowledge & Talent	Share of developers per capita	15.22%	16%	Frankfurt, highest share per capita in Germany	95.13%
Start-up Support	Share of mentors with start-up experience	53.40%	50%	Estimate of a good performance	106.80%
Specialization & Diversification	HHI based on sector investments	0.234	0.2	Estimate of a good performance	117%
Network Density	Tech meetup participants per 1,000 inhabitants	20.35	60.4	Munich, highest share per capita in Germany	33.69%
Innovation Culture	Total of international investments per capita	16.98	36	(All-time high in Cologne)	47.17%
Policies	Ease of doing business rating	70%	68%	Average of top 20 hubs in Europe	102.94%
International Connectivity	Heatmap TrustScore	81.90%	87.93%	(All-time high in Cologne)	93.14%

Table 7: Results of the Ecosystem Health Check for Cologne's HGE.

4. THE CASE OF GUATEMALA'S INCLUSIVE AND EQUITABLE ECOSYSTEM

4.1. COMMUNITY DIMENSION

In line with the previous case study, we start by measuring the health of the Community (see Figure 1), which can be mapped by looking at two different dimensions that are defined by two metrics:

1. Pool of Entrepreneurs
 - 1.1. Diversity & Inclusivity
 - 1.2. Knowledge & Talent
2. Entrepreneurship Support Network
 - 2.1. Start-up Support
 - 2.2. Specialization & Diversification

4.1.1. Pool of Entrepreneurs

a. Diversity & Inclusivity

In the case of Guatemala, we need to choose the right proxy from the list of proxies referring to Inclusive and Equitable EE (IEE).

In the list, we found three options to measure Diversity & Inclusivity: Share of female founders, share of founders from minorities/marginalized communities, and share of female PhD graduates.

These three proxies serve the purpose of measuring how inclusive the local community is vis-à-vis newcomers and outsiders. As we have highlighted in the previous cases, we might prefer one or the other proxy. The most important thing is to identify adequate proxies in terms of the quality and timeliness of the various data sources.

We encourage you to use more than one proxy to estimate the different metrics. However, if not enough data are available, using one proxy is still an option.

If you have sufficient data, you can create an index by, for example, taking the average of the proxies used.

Also in this case, the data collection process can be implemented considering city-level data or, alternatively, by selecting a sample that we can collect ourselves.

We found good data on female founders on the Global Entrepreneurship Monitor website, where we have checked the set of information available for Guatemala. Based on this source, we discovered that in 2019, 45% of entrepreneurs in the country were female. Since data at city level are not available, we have to assume that data at country level more or less mirror the data for Guatemala City. Further, we do not yet have data for 2020. An important caveat is the definition of entrepreneurs used by the Global Entrepreneurship Monitor, which describes them as “nascent entrepreneurs or owner-managers of a ‘new business’”. This definition is very broad and can include all types of self-employment.

To have an idea of the numbers for a narrower definition, we decided to look at 77 start-ups with headquarters in Guatemala City listed on Crunchbase. Although these data focuses solely on tech businesses, we believe that it is still a relevant proxy for Guatemala. Nevertheless, it should be noted that the tech sector is an area where women are chronically underrepresented, and an increase in this space will very likely also lead to an increase in other sectors. Taking the Crunchbase data, the share of female founders was 18% in 2020.

20. https://gem.ufm.edu/wp-content/uploads/2020/11/Reporte_2019...

21. <https://www.gemconsortium.org/wiki/1154>

Considering that data at city level are not available from GEM, as well as the sample size and the fact that data are updated regularly on Crunchbase, we selected the share of female founders in the sample of 77 start-ups in Guatemala as the proxy for Diversity & Inclusivity (see Table 8).

b. Knowledge & Talent

The second metric Knowledge & Talent describes the density of the talent pool triggering potential network effects and increasing the innovative potential of the ecosystem.

In the list of proxies specific to IEE, we found three options: Share of start-ups with tech product, start-up skills, and Share of ambitious entrepreneurs²².

Data referring to the last two proxies are supplied at country level by international organizations that update this information annually, albeit with gaps (see Table

9). Specifically, data on start-up skills are provided by GEDI²³, and the share of ambitious entrepreneurs is still available on the Global Entrepreneurship Monitor website²⁴. For the latter, data are available for 2018 where this share is 75%.

In view of the weakness of the data mentioned above, we decided to focus on the 77 start-ups with headquarters in Guatemala City listed on Crunchbase. This sample offered supplementary insights into the products of these start-ups. Using the description of the start-ups, it is possible to identify the technological content of their products.

In line with the previous metric, and considering the sample to be large enough to derive general insights at ecosystem level, we selected the share of start-ups with tech product as the proxy. The relevant figure was 10% in 2020.

	City level	Sample
Share of female founders	Mediocre, but data are available for the period to 2019	Good data (77 Start-ups), Updated daily
Share of founders from minorities/ marginalized communities	No data available	No data available
Share of female PhD graduates	Mediocre	No data available

Table 8: Description of quality of data from data collection in Guatemala for Diversity & Inclusivity.

	City level	Sample
Share of startups with tech product	No data available	Good data (77 Start-ups)
Start-up skills	Mediocre	No data available
Share of ambitious entrepreneurs	Mediocre	No data available

Table 9: Description of quality of data from data collection in Guatemala for Knowledge & Talent.

22. Ambitious entrepreneurs refer to the adult population in a region that has the ambition to grow a new business to a size of 20 employees within 5 years.

23. <https://thegedi.org/tool/>

24. https://gem.ufm.edu/wp-content/uploads/2020/11/Reporte_20..

4.1.2. Entrepreneurship Support Network

a. Start-up Support

The first metric for the Entrepreneurship Support Network dimension is the Start-up Support defined in our framework as the tangible or intangible assets of the EE supporting entrepreneurship.

In the list of proxies specific to IEE, we found three options: Share of mentors with start-up experience, number of profitable social enterprises, share of people with an entrepreneurial attitude²⁵, and support programs for entrepreneurs in non-wealthy districts.

In terms of the share of mentors with start-up experience, it is not possible to find the relevant data. Even if the number of profitable social enterprises can be determined, the information is out of date, since the last update relates to 2015. Consequently, we looked for other proxies measuring the entrepreneurship support network of the ecosystem. The Global Entrepreneurship Monitor provides data for 2019²⁶ about the share of people with an entrepreneurial attitude.

We therefore explored different options to measure the start-up support network. For example, geographic diversification of entrepreneurial resources is critical to create access points for underrepresented entrepreneurs. Equally, an inclusive ecosystem should provide services not only for high-growth businesses but – as we indicated in our framework – also for dynamic and sustainable enterprises. We looked for these data in the Booklet Mapeo Ecosistema de Emprendimiento²⁷, a comprehensive listing of providers of support in Guatemala City. The picture shows that out of 79 support programs for entrepreneurs, 43% target high-growth businesses, while 25% are focused on social enterprises, 20% on dynamic enterprises and only 11% on sustainable enterprises. Looking at the second indicator, i.e. the geographical distribution of support programs, we found that 85% are located in business districts or around the wealthier districts (Zones 10, 14, 16, or Antigua), creating an additional barrier for unrepresented entrepreneurs to

access the resources of the Guatemalan ecosystem. We believe that the geographical indicator is the best alternative proxy to measure the inclusiveness of the support network, which is 15% in the case of Guatemala (with support services being offered in areas with a lower socio-economic status).

b. Specialization & Diversification

The second metric for the Entrepreneurship Support Network dimension is Specialization & Diversification, which describes how diverse the support system is. For example, does the ecosystem mainly support one vertical (e.g. AgriTech) and regularly succeed in this sector, while start-ups in other sectors fail to grow?

In the list of proxies specific to IEE, we found two options: Share of programs focusing on specialized verticals (sector, technology, interest), and diversity of projects based on business models (for-profit and high-growth vs. NGO and non-profit) and revenue stages.

We have no data available at city level or in a sample for the latter. However, using our sample of 77 start-ups with headquarters in Guatemala City listed on Crunchbase, it is possible to categorize start-ups by looking at their purpose and technological advancement. We have therefore identified four categories in this context: Social low-tech start-ups (11.69%), social high-tech startups (3.90%), non-social high-tech start-ups (6.49%), and low-tech and non-social start-ups (77.92%).

The Herfindahl-Hirschman Index (HHI) measures the concentration among values of a frequency distribution, thus showing how diversified a set of inputs is. If all investments were made in one sector, the HHI would be 1. If all investments were made in diverse sectors, (N) it would be 1/N.

For Guatemala, the HHI in 2020 was 0.63. This figure reflects quite a high concentration in one category, i.e. low-tech and non-social start-ups.

25. It refers to people considering entrepreneurship a desirable career option.

26. https://gem.ufm.edu/wp-content/uploads/2020/11/Reporte_2019_

27. Available on www.startupguatemala.org

4.1.3. Proxies selected for Community

Based on the previous considerations, we want to present the list of proxies selected for this case study to measure Community.

1. Pool of Entrepreneurs
 - 1.1. Diversity & Inclusivity: Share of female founders (18%)
 - 1.2. Knowledge & Talent: Share of start-ups with tech product (10%)
2. Entrepreneurship support network
 - 2.1. Start-up Support: Support programs for entrepreneurs in non-wealthy districts (15%)
 - 2.2. Specialization & Diversification: Diversity of projects based on business models (0.63)

4.2. ENVIRONMENT

We start at this point to measure the health of the Environment (see Figure 1), which will be mapped by looking at two different dimensions that are defined by two metrics:

1. The Social Capital
 - 1.1. Network Density;
 - 1.2. Innovation culture;
2. The Governance
 - 2.1. Policies;
 - 2.2. International Connectivity.

4.2.1. Social Capital

a. Network Density

The first metric for the Environment dimension is Network Density, which reflects the possibilities for stakeholders to collaborate and connect within the EE.

In the list of proxies specific to IEE, we found two options:

Number of recurrent meetups per capita and number of entrepreneurial communities.

As stated in the previous case, the website meetup.com has become a go-to tool for community organizers around the world. While not all community activities are organized via the platform, it provides a good indication of the general level of network density. We can search for the number of meetups by selecting Guatemala as a city²⁸. The list shows recurring meetups with their number of followers. It is possible to obtain information about the number of participants, the topic of the meetups, the location and the date of the event. These data are available over time and are updated continuously. We therefore believe that the data available on meetup.com are a good sample for our proxies.

Consequently, we have selected the number of meetups per capita as our proxy. Since we have eight meetup groups in Guatemala City and the population of the city in 2020 was 2,935,000, the number of meetups per capita is 0.00027.

b. Innovation Culture

The second metric of the Social Capital dimension is Innovation culture, which refers to the set of soft rules in place that prevent moral hazards and opportunistic behaviors in open innovation practices.

In the list of proxies specific to IEE, we found three options: Share of research expenditure on business enterprise, foreign direct investments, and innovation global ranking.

In terms of the share of research expenditure on business enterprise, the UNESCO website²⁹ supplies R&D expenditure at country level. However, even if the information is updated every year, there is a gap, and we only have data covering the period up to 2018. The same applies for foreign direct investment. A better proxy might be the Global Innovation Index (GII)³⁰, which provides an indication of the innovation performance of 131 countries and economies around the world. Its 80 indicators cover a broad range of innovation-

28. <https://www.meetup.com/es/find/?allMeetups=false&keywords=entre..>

29. <http://uis.unesco.org/en/country/gt?theme=science-technology..>

30. https://www.wipo.int/global_innovation_index/en/

relevant aspects, including the political environment, education, infrastructure, and business sophistication. Based on this proxy, Guatemala ranked 106th among the 131 economies included in the GII for 2020. This compared to a ranking of 107th in 2019 and 102th in 2018.

It is important to note that the ranking can deteriorate even if Guatemala improves, as it also depends on the performance of other countries. Data related to rankings are not an optimal solution and, when possible, it is preferable to have the actual figures. However, where better data are lacking, the ranking can be taken as the proxy.

4.2.2. Governance

a. Policies

The first metric in the Governance dimension is Policies, which can be defined as the fundamental institutions and reliable governance system that fosters entrepreneurial activity. This metric makes it possible to understand whether entrepreneurship is supported by regulations that reduce corruption, ease business processes and increase partnerships.

In the list of proxies specific to IEE, we found two options: Ratings for ease of doing business and the corruption index.

We first checked for data on ratings for ease of doing business at city level. Both the GEDI dataset and The World Bank provide this kind of data. However, the GEDI dataset presents information only for 2018, and The World Bank only for 2019³¹. We therefore decided to use the latter source of information as our proxy. Guatemala was rated 96th among 190 economies for ease of doing business, according to the latest World Bank annual rankings that relate to 2019. This reflected an improvement from 98 in the previous ratings.

Through Transparency International, we also have information on the corruption ranking for 2019: Guatemala is the 146th least corrupt nation out of 180 countries, according to the 2019 Corruption Perceptions Index reported by Transparency International³².

In such cases, the choice of proxy is based on the type of information that we prefer to capture, considering the metric. In this specific case, we selected the first proxy: Ratings for ease of doing business. This proxy offers a more comprehensive view regarding the regulatory environment conducive to starting and operating a local firm. Indeed, The World Bank takes account of 41 indicators for 10 topics on doing business³³.

b. International Connectivity

The second metric for Governance is International Connectivity, which describes the ability of the EE to connect internationally. The more powerful the brand of a nation or region, the more it will be able to attract international investments and founders.

In the list of proxies specific to IEE, we found two options: Share of companies with sales in foreign markets and share of start-ups that moved to the ecosystem from another country.

As with the former proxy, we found data on the GEDI website³⁴. Specifically, Guatemala's score for variable internationalization is 0.12. This information has not been updated since 2018, and it is comparable to a ranking – making the analysis process more complex. Consequently, we looked for an alternative.

We also explored the second option for the proxy by focusing on the information published by Crunchbase. In this case, we took a sample of the 17 companies most recently listed on Crunchbase, which shows there are no foreign nationals among their CEOs. This gives us an indication that Guatemala is not a destination for foreign entrepreneurs to set up their business. Unfortunately, we did not find any data on outbound connectivity, as the share of international sales of start-ups from Guatemala was not available.

Since only one proxy is available for 2020, we decided to use the sample of 17 companies listed on Crunchbase as the metric for International Connectivity (see Table 10).

31. <https://www.doingbusiness.org/en/rankings>

32. <https://www.transparency.org/en/countries/guatemala>

33. <https://openknowledge.worldbank.org/bitstream/handle/10986/32436..>

34. <https://thegedi.org/tool/>

	City-Level	Sample
Share of young companies with sales in foreign markets	Mediocre	No data available
Share of start-ups that moved to the ecosystem from another country	Mediocre	Good data (sample of 17 companies)

Table 10: Description of quality of data from data collection in Guatemala for International Connectivity.

4.2.3. Proxies selected for Environment

Summarizing the previous considerations, we want to show you the list of proxies we selected for this case study for measuring Community.

1. The Social Capital
 - 1.1. Network Density: Number of meetups per capita (0.00027)
 - 1.2. Innovation Culture: Global Innovation Index (106th)
2. The Governance
 - 2.1. Policies: Ratings for ease of doing business (96th)
 - 2.2. International Connectivity: Share of foreign entrepreneurs (0%).

4.3. PERFORMANCE OF GUATEMALA'S ECOSYSTEM

The selection of the final proxy marks the completion of the data collection process, meaning that you can proceed to enter information in the data tool. What is still missing is the definition of the benchmarks used to calculate the performance results per metric. It is therefore vital to use realistic and adequate benchmarks. When defining your benchmarks, you can have clear objectives about what you want to achieve and, in this case, the results should be checked against your goal. However, if you do not have clear objectives, you can look at a comparison ecosystem that performs considerably better than yours in a metric and add its actual numbers as a benchmark. An even better approach is to take the median of a group of comparison ecosystems. However, these data might be difficult to find – or not available at all. It is therefore also possible to add an estimate of a realistic improvement in your ecosystem's performance. We indicated the name of our benchmark in the table below to give an idea of how to proceed.

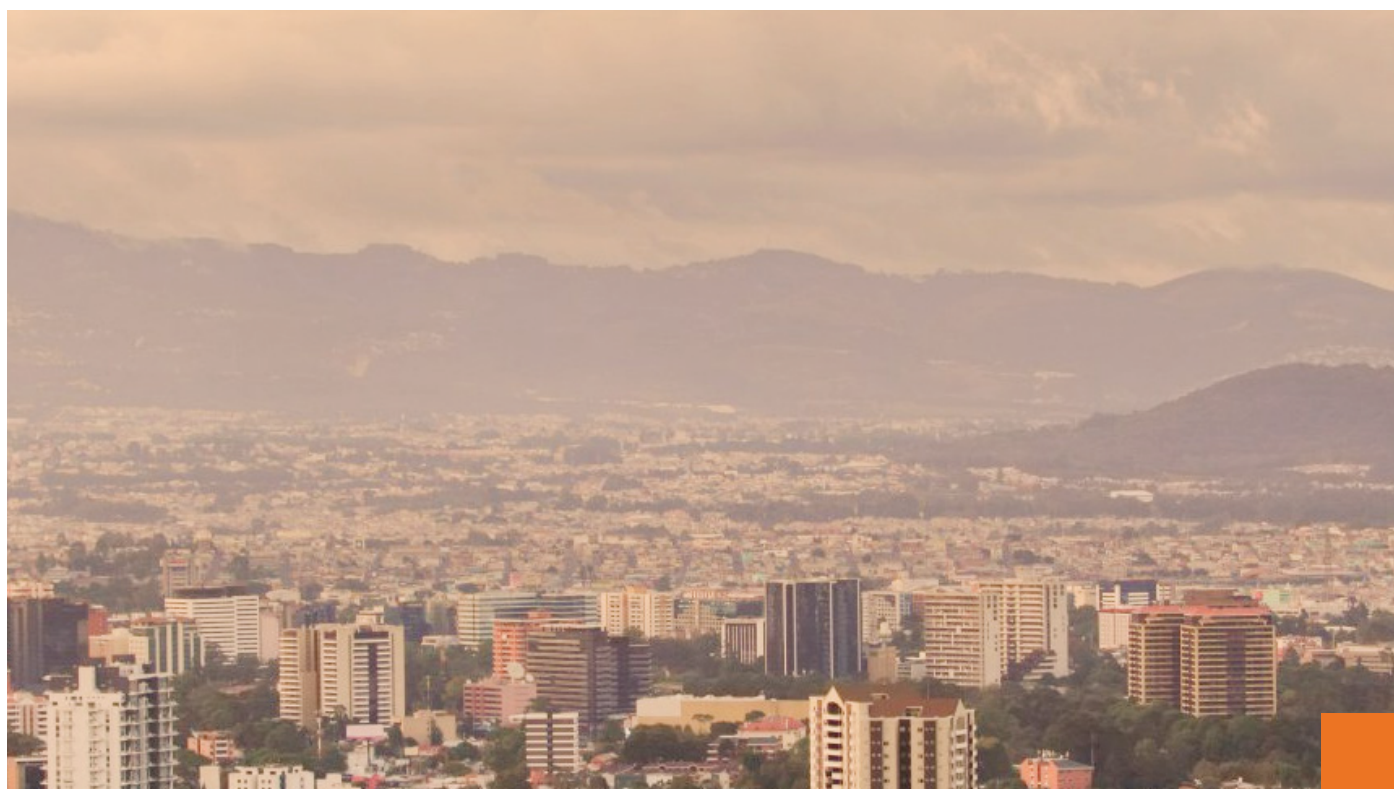
Metric	Proxy	Guatemala	Benchmark	Benchmark Details	Result
Diversity & Inclusivity	Share of female founders	18%	15.5%	Average % of female founders among 550 accelerated start-ups tracked in Latin America	116%
Knowledge & Talent	Share of startups with tech product	10%	30%	Estimate	33%
Startup Support	Share of support services being offered in areas with a lower socio-economic status	15%	30%	Estimate	50%
Specialization & Diversification	HHI based on business model	0.63	0.3	Estimate	210%

Metric	Proxy	Guatemala	Benchmark	Benchmark details	Result
Network Density	Number of meetups per capita	0.00027	0.011	18.323 ³⁵ meetup followers in Quito, Ecuador	2%
Innovation Culture	Innovation global rank	106	50	Estimate	n/a
Policies	Ease of doing business Rating	96	50	Estimate	n/a
International Connectivity	Share of foreign entrepreneurs	0	6%	Foreign founders in accelerators in Brazil	n/a

Table 11: Results of the Ecosystem Health Check for Guatemala's IEE - Part 2.

Looking at the benchmarks (Table 11), we believe that a key point to make Guatemala's ecosystem more open and accessible is to reduce geographical barriers by being closer to entrepreneurs in terms of physical proximity and to focus on network density. Having a relationship with other founders and feeling part of a community are factors that are strongly associated with enhanced start-up performance.

Further, Guatemala should continue to work on strengthening its entrepreneurial pool. Given its market size and proximity to the US, Guatemala has an opportunity to position itself as a test market for US or European social ventures that are looking to determine their product market fit. However, if Guatemala wants to attract more of these founders, it should also improve on policy aspects to facilitate setting up a new business (or subsidiary).



35. <https://www.meetup.com/de-DE/find/tech/?allMeetups=false...>

5. CONCLUSIONS

Thanks to the standardized criteria to measure and diagnose ecosystems, you can track conditions in your EE and develop coherent interventions.

Starting with the different metrics of the framework and the list of proxies proposed in the Annex (Table 1), you can define your strategy to map your EE over time. You can use primary or secondary data, and collect them on an annual or quarterly basis. We strongly recommend taking a coherent approach and avoiding the different challenges presented for the two different EE.

As you can see from the two case studies, it is important to identify proxies that are measurable and comparable over time. However, in some cases this is a complex undertaking. We have presented two cases where the diagnosis is founded on metrics measured by one proxy. We would encourage you to use more than one proxy to estimate the different metrics. However, if not enough data are available, using one proxy is still an option. If you have sufficient data, you can go further and create an index by, for example, taking the average of the used proxies used.

Some of the selected proxies, especially for the Guatemala case, can be improved. Specifically, the ranking is not an optimal solution for checking the health of your EE's metrics, such as Innovation Culture and Policies. As rankings depend on the status of other EE, it may seem that you are improving but the achievement of a higher ranking could also simply mean that the performance of others is declining. We therefore recommend collecting primary data and using surveys to estimate the suggested proxies. While the framework's overall robustness when using 'guestimates' might be somewhat lower, our experience shows that it is still useful and relevant. The process of defining the metrics and benchmark in a collaborative manner is already a first step towards improving the ecosystem. Having a joint understanding of

the overall objective and a collective agreement on the future vision is key. This makes it possible to jointly work towards a common end-goal. Hopefully, data quality will improve over time, as there is more willingness from the various participants to measure and share it. Having a shared vision and clear objective will unlock the collective power of all ecosystem players and lead to a situation where the whole is indeed greater than the sum of its parts.



ANNEX

Structure	Dimension	Metric	Proxy	
			HGE	IEE
COMMUNITY	Entrepreneurial Landscape	Diversity & Inclusivity	<ol style="list-style-type: none"> 1. Share of female founders 2. Share of Foreign-Born founders 	<ol style="list-style-type: none"> 1. Share of female founders 2. Share of founders from minorities/marginalized communities 3. Share of female PhD graduates
		Knowledge & Talent	<ol style="list-style-type: none"> 1. Number of developers per capita 2. Share of engineers among founders 3. Number of tech start-ups per capita 	<ol style="list-style-type: none"> 1. Share of start-ups with tech product 2. Start-up Skills 3. Share of ambitious entrepreneurs
	Entrepreneurship Support Network	Startup Support	<ol style="list-style-type: none"> 1. Number of start-ups per capita 2. Share of accelerated start-ups receiving follow-on funding 3. Share of companies with exits from total companies receiving investments 4. Share of mentors with start-up experience 	<ol style="list-style-type: none"> 1. Share of mentors with startup experience 2. Number of profitable social enterprises 3. Share of people with an entrepreneurial attitude
		Specialization & Diversification	<ol style="list-style-type: none"> 1. Investments based on stages 2. Investments based on sectors 3. Diversity of firms based on sectors 	<ol style="list-style-type: none"> 1. Share of programs focusing on specialized verticals (sector, technology, interest) 2. Diversity of projects based on business models (for-profit and high-growth vs. NGO and non-profit) and revenue stages

Table 1: Overview of proxies for the two Entrepreneurial Ecosystems - Part 1

Structure	Dimension	Metric	Proxy	
			HGE	IEE
ENVIRONMENT	Social Capital	Network Density	1. Participants in tech meetups per capita 2. Number of tech meetups per capita	1. Participants in tech meetups per capita 2. Number of start-up communities /entrepreneur communities
		Innovation Culture	1. Total of foreign investments per capita 2. Total VC investments per capita 3. Number of deals EUR > 1 million per capita	1. Share of research expenditure on business enterprise 2. Innovation global ranking 3. Foreign direct investment
	Governance	Policies	1. Ratings for ease of doing business 2. Share of start-ups with no international headquarters	1. Rankings for ease of doing business 2. Number of pro-business initiatives
		International Connectivity	1. Number of international top start-ups with offices in the city 2. Number of international conference participants in the city 3. Ecosystem brand and visibility	1. Share of young companies with sales in foreign markets 2. Share of start-ups that moved to the ecosystem from another country

Table 1: Overview of proxies for the two Entrepreneurial Ecosystems - Part 2

Source	Description	Link
Atomico State of European Tech Report	Annual report aggregating data and insights on European tech ecosystems covering many aspects from investments to job creation	https://2020.stateofeuropeantech.com/
Crunchbase	Global start-up investment database	https://www.crunchbase.com
Craft.co	Global company analysis database tracking the growth of start-ups	https://craft.co/
Dealroom	European focused start-up investment database	https://www.dealroom.co/
European Startup Monitor	The annual European Startup Monitor is based on a survey among European founders and their opinions	http://www.europeanstartupmonitor2019.eu/
GEDI – The Global Entrepreneurship and Development Index	US-based think tank collecting data and producing indexes on the development of start-up ecosystems worldwide	https://thegedi.org/
GALI – Global Accelerator Learning Initiative	Longitudinal on more than 23,000 ventures that applied to participating accelerator programs between 2013 - 2020	https://www.galidata.org/
Startup Heatmap Europe	Collection of 1st and 3rd party data on start-up ecosystems for 100 cities in Europe	www.startupheatmap.eu
Startup Europe Partnership	A collection of analysis on European scale-ups powered by the Startup Europe.	https://startupeuropepartnership.eu/reports/
Startup Blink	Platform for crowdsourced list of start-ups and accelerators worldwide. Also publishes ecosystem reports.	https://www.startupblink.com/
Stackoverflow Developer Survey	Global Survey of roughly 65,000 developers conducted by stackoverflow	https://insights.stackoverflow.com/survey/2020#developer-profile--survey-respondents
Seed DB	A global dataset of accelerators and participating startups based on crunchbase data	https://www.seed-db.com/accelerators
VC4Africa	Open data on African startups searchable by tag, sector, country and fundraising status	https://vc4a.com/blog/2015/09/28/announcing-open-vc4a-venture-data/
World Bank Ease of Doing Business	Study by The World Bank on regulations governing the economic activities in countries around the world	https://www.doingbusiness.org/en/rankings
The Global Startup Ecosystem Report	Annual Report published by consultancy firm Start-up Genome on top ecosystems in the world	https://startupgenome.com/report/gser2020

Table 2: List of Sources - Part 1

Source	Description	Link
Global Entrepreneurship Monitor	Annual survey and report on entrepreneurship framework conditions around the world by the Kauffman Foundation	https://www.gemconsortium.org/
Decoding Digital Talent	Unregular global survey among digital talent for their favorite destinations by Boston Consulting Group	https://www.bcg.com/publications/2019/decoding-digital-talent
Teleport	Aggregator of city level data ranging from quality of life indices to crime rates.	https://www.teleport.org
EBAN – European Business Angel Network	Annual Statistics Compendium on Business Angel activity in Europe	https://www.eban.org/knowledge-center/industry-reports/
QS University Ranking	Ranking of Top Universities in various categories	https://www.topuniversities.com/university-rankings/world-university-rankings/2021
CBinsights	Annually updated list of active unicorns (tech startups evaluated > 1bn USD) worldwide	https://www.cbinsights.com/research-unicorn-companies
GHTorrent	Access to the Github developer platform REST API offering information on developer communities around the world	https://ghtorrent.org/faq.html
Ookla	Global Internet Speed Index	https://www.speedtest.net/global-index
Invest Europe	European private equity and venture capital investments database	https://www.investeurope.eu/research/activity-data/
UBI Global	Ranking of university based incubators worldwide	https://ubi-global.com/publications/
Angel Capital Association	North American Business Angel association publishing activity data of business angels in the US and Canada	https://www.angelcapitalassociation.org/research/

Table 2: List of Sources - Part 2