European Commission

FROM HORIZON 2020 TO HORIZON EUROPE

MONITORING FLASH

#2.1 DYNAMIC NETWORK ANALYSIS

November 2018

FROM FP6 TO HORIZON 2020

Key overview data

5 million
collaborations within FP6, FP7 and Horizon 2020 projects

5/1
ratio of new to maintained collaborations in the first four years of Horizon 2020

7,509
collaborative projects launched in the first four years of Horizon 2020

23,664
participants collaborated under the first four years of Horizon 2020

149
countries represented in collaborative projects under Horizon 2020

The size of the Horizon 2020 collaboration network is massive: since 2014, Horizon 2020 has funded more than 7,500 collaborative projects among 23,664 participants from 149 countries, which results in almost 1.5 million of one-to-one collaborations. For every maintained collaboration, there are almost 5 new ones. Since FP6, 5 million collaborations have been generated by the Framework Programmes.

A MASSIVE NETWORK FOR R&I ACTIVITIES

Key overall messages

- The most central countries in the network are also the largest ones: Germany, France, the UK, Italy, and Spain.
- When taking into account the size of countries, some “punch above their weight”: the most central country in Horizon 2020 is Finland, followed by Slovenia. While some new Member States like Cyprus, Estonia, Malta and Slovenia are as central as EU-15 countries, other EU-13 countries remain at the bottom of the connectivity rankings..
- Slovenia, Luxembourg, Croatia, Portugal and Cyprus show the most striking increases in terms of size-normalised centrality between FP7 and Horizon 2020, while the position of the UK and Hungary dropped.
- Although EU-15 participants have reduced their collaborations with EU-13 participants in FP7, this trend appears to be reverting in Horizon 2020.
- Geographical and cultural proximities between participants seem to play an important role in shaping the structure of the Horizon 2020 collaboration network.

Research and Innovation

Contact : Julien Ravet
RTD-A5-SUPPORT@ec.europa.eu

1 Introduction

A key EU Added Value of the EU Framework Programmes for Research and Innovation (R&I) consists in the creation of transnational and multidisciplinary networks (European Commission, 2017 and 2018). The Framework Programmes offer unique collaboration and networking opportunities between researchers. Collaborations within the network generate spillover effects and knowledge sharing while bringing the R&I effort in Europe closer to the critical mass required to tackle global societal challenges. The majority of the Horizon 2020 budget is spent on supporting such collaboration through collaborative R&I projects. However, to fully reap the benefits of collaborative R&I, it is important that the network remains open and easily accessible to new participants. In this context, a good understanding of the way researchers collaborate within the Programme is crucial.

The Horizon 2020 Interim Evaluation already provides insights into the collaborations between researchers based on publications[1] and project data. In particular, it suggests that collaboration patterns may have evolved between the 7th Framework Programme (FP7) and Horizon 2020. A previous study also examined the evolution of collaborations between the 6th Framework Programme (FP6) and FP7 (Science Metrix, 2015; with 40% of the projects completed in FP7).

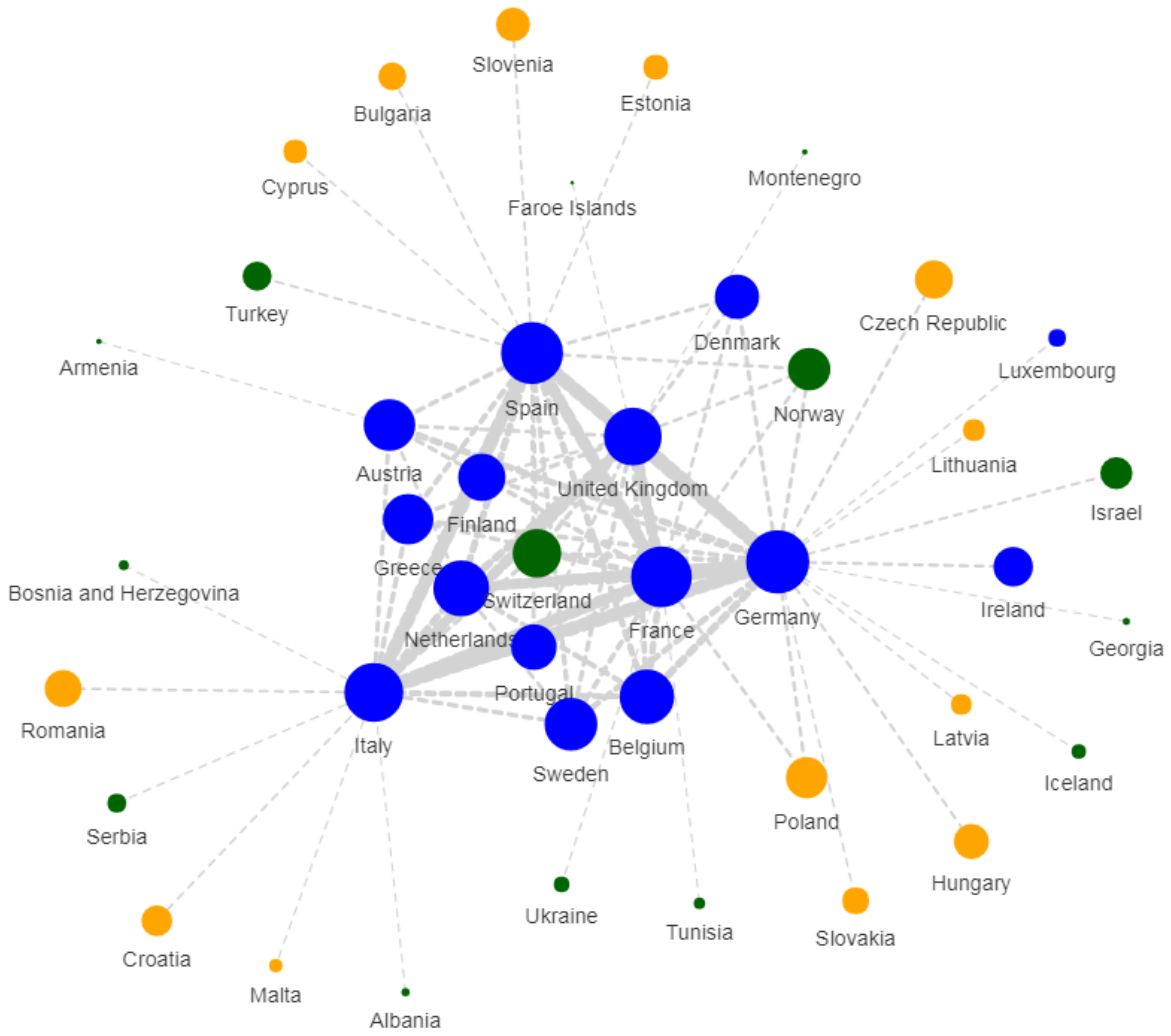
This monitoring flash further explores certain aspects of the collaborations between participants and provides additional evidence related the dynamic evolution of the network of participants to the Programme. While the complexity of such a large network can be examined from different angles, this flash focuses on cross-country collaborations. In particular, the analysis highlights how the situation of entities in participating countries has changed over the last decade.

The analysis is based on monitoring data of Horizon 2020 and its predecessor programmes, FP6 and FP7, covering the 2003-2017 period[2]. The data covers collaborative projects[3] launched during the first four years of implementation of Horizon 2020, and the full implementation of FP6 and FP7[4]. The data is stored in the Common Research Data Warehouse (CORDA), an internal database maintained by DG Research and Innovation. For this paper, country groups (i.e. EU-15, EU-13, associated countries and third countries) are based on the situation in Horizon 2020. This flash is based on the main results from a more extended study, Balland and Ravet (2018)[5].

[1] Elsevier (2017).
[2] Year of signature of the contract. Cut-off date for Horizon 2020 is 1/1/2018.
[3] Data include all evaluated calls for collaborative projects. Projects under Public-Public Partnerships, EIT's Knowledge and Innovation Communities (KICs) and direct actions of the Joint Research Centre are not included.
[4] Projects with incomplete data on signature date, duration and participant identifier were removed from the analysis (the final dataset includes about 99.1% of the initial dataset of collaborative projects).
[5] Balland, P.A., and Ravet, J. (2018). Dynamic Network Analysis of the EU R&I Framework Programme. European Commission report, forthcoming. More details, including on the technical aspects, are available in the study. Calculations were produced by Pierre-Alexandre Balland (Utrecht University and Massachusetts Institute of Technology).

2 Overview of the Horizon 2020 network

Figure 1 Structure of the H2020 Collaboration Network



Note: This graph represents the backbone of H2020. Nodes are countries, and links represent strong[6] connections based on Horizon 2020 projects. EU-15 countries are represented in blue, EU-13 countries are represented in orange and Associated Countries (AC) are represented in green. The size of the nodes expresses their centrality, and the width of the links expresses the number of collaborations. Source: CORDA data.

The size of the Horizon 2020 collaboration network is massive. Since 2014, Horizon 2020 has been funding a very large number of collaborative projects, which involved a very large network of collaborations between R&I stakeholders. Over 2014-2017, Horizon 2020 funded more than 7,500

[6] The links displayed on this graph with N actors combines the N-1 links of a maximum spanning tree (MST) and the N-1 strongest links of the overall network. The MST represents the backbone of a weighted network and is based on three rules. First, it keeps only N-1 links from a network with N actors. Second, rule #1 should be satisfied while keeping the strongest links. Third, rule #1 and #2 should be satisfied without creating any isolate in the network.

collaborative projects among 23,664 participants from 149 countries, which results in almost 1.5 million of one-to-one opportunities to collaborate[7].

The strongest connections that emerged out of Horizon 2020 are represented as a country-country graph in Figure 1[8]. The figure shows two types of connections: (i) the single strongest connection of each country to another country, and (ii) the top 40 strongest connections in the network. Centrality can be defined as the importance of a country in the network. This importance as such can have different meanings, hence different definitions, with the most straightforward definition being based on the number of connections of a country's participants in the whole network. The size of the nodes is proportional to the centrality of the country. **The core of the network is mainly composed of EU-15 participants. Germany, France, the UK, Italy, and Spain appear to be key players in the network of participations to Horizon 2020**.

EU-13 participants have a substantial number of collaborations with the largest players in the network, which are participants from EU-15 countries[9]. As a result, German participants are frequent partners of several EU-13 countries, such as Czech Republic, Hungary, Latvia, Lithuania and Slovakia. Croatia, Malta and Romania present strong ties with Italy, while Bulgaria, Cyprus, Estonia and Slovenia tend to connect with Spanish participants. Important collaborators of Polish participants are French participants.

3 Evolution of general characteristics of the network

The nature of the network has changed as follows since FP6:

- **The network of participations to the Framework Programmes seems to be very dynamic over time[10]**. Between FP6 and FP7, about new 1,226,970 connections between partners were created, while 166,508 connections were maintained and 772,822 were lost. Between FP7 and the first four years of Horizon 2020, 909,444 new connections were made, against 195,474 maintained and 1,198,004 lost.

- **The network is more dynamic for EU-13 countries than for EU-15 countries.** EU-13 countries participants seem to have a higher propensity to be involved in new collaborations than participants from EU-15 countries, which is especially true in Horizon 2020.

- **The network tends to be opening to less connected participants since FP6.** On average, participants are slightly less central[11] in the network in FP7 and Horizon 2020 compared to FP6. This might signal the entry of smaller players.

- **Participant acting as hubs (i.e. with high centrality) also seem to connect more likely with other types of participants (non-hubs, with low centrality)[12]**. This suggests that key actors in

[7] Before Horizon 2020, FP6 and FP7 funded respectively 5,912 and 12,493 collaborative projects, which correspond to 1,305,305 and 1,989,450 collaborations between participants.
[8] Third countries are excluded from this visualisation. They are analysed in Balland and Ravet (2018). The analysis is based on the network of participations without any threshold in the number of connections between two participants. Balland and Ravet (2018) use two sets of data for robustness: one without threshold and one based on connections in at least two projects in order to exclude one-off collaborations. The key messages are similar between the datasets. This Flash is based on data that include all connections.
[9] Top 5 EU-15 participants that present the largest numbers of collaborations with EU-13 participants are Fraunhofer (DE), CNR (IT), CNRS (FR), CEA (FR) and VTT (FI). Top 5 EU-15 participants that present the highest share of collaborations with EU-13 participants in their collaborations are ENEA (IT), NERC (UK), CINECA (IT), UoA (EL) and JUELICH (DE).
[10] Based on Jaccard coefficients (Ripley et al., 2016).
[11] Centrality can be defined as the importance of a node (here a participant) in the network. This importance as such can have different meanings, hence different definitions. Measures used here are degree centrality and eigenvector centrality.

the network have maintained a certain level of openness to other participants throughout the different programmes.

- **Network inequality coefficients[13] are particularly stable over time**. This suggests that a few organisations have many connections, while most organisations have only a few - this is a general tendency of real-world complex networks. This aspect of the network has not been reinforced over time.

- The average path length between participants has remained close to 3, meaning that **on average a participant can be connected to any other participant in the network within 3 connections ("degrees of separation")**. This measure is relatively small, indicating a highly-connected network in general. The average path length has not changed much over time.

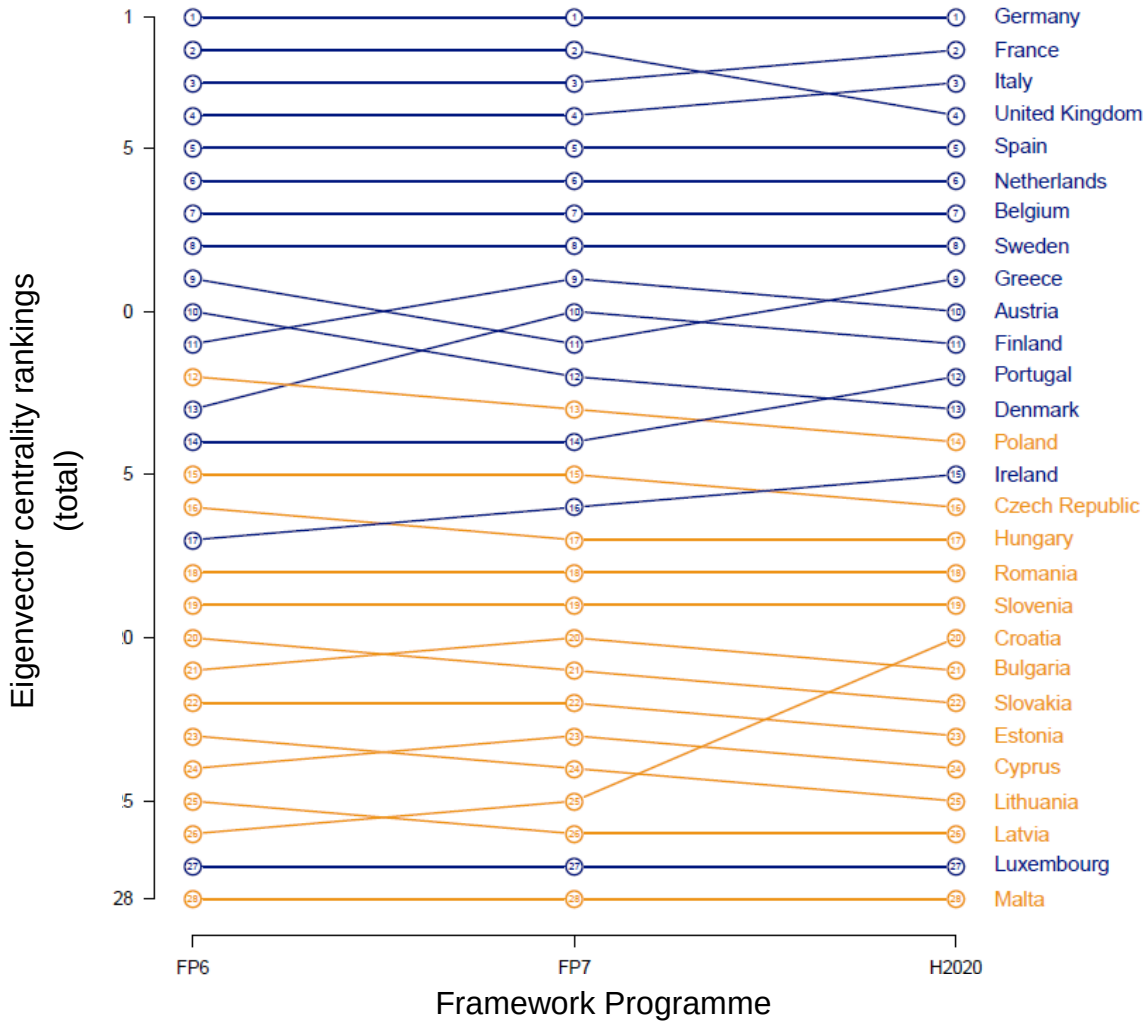
4 Evolution of the position of countries in the network

Which countries are becoming more central?

This section presents network indicators computed at the participant level, and averaged or aggregated at the country level. Figure 2 details the evolution of country rankings based on centrality measures (here eigenvector centrality coefficients[14]). **Germany is both the largest participant in the Framework Programme and the most central country in the network. After Germany, France and Italy are the most central countries in Horizon 2020.** While the UK was more central than France and Italy in FP6 and FP7, its central position worsened in Horizon 2020. Greece, Portugal and Ireland have improved their centrality in the network between FP7 and Horizon 2020 according to this ranking. The chart also confirms that participants from EU-15 countries tend to be more central than their EU-13 counterparts: the bottom of the chart is occupied by a majority of EU-13 countries. Only Croatia seems to have significantly improved its position since FP6.

[12] Based on assortativity coefficients (extent to which participants in the network associate with other participants with similar degree centrality). The Framework Programmes are disassortative networks in the sense that central nodes tend to connect to less central nodes.
[13] Network Gini coefficient of the degree distribution.
[14] This indicator measures the influence of a country in the network by examining whether participants of a country are linked to other important participants (i.e. participants with many connections).

Figure 2 Network positions of participants by EU country (ranking)

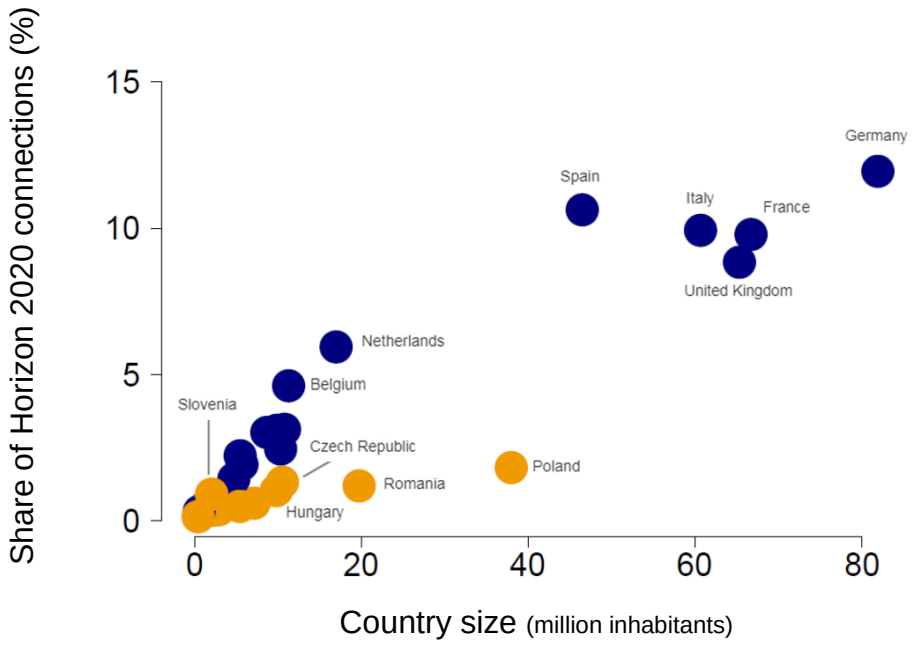


Source: Corda data.

However, these measures are absolute and are significantly influenced by country size[15]. Figure 1 and 2 illustrate this size effect in the network: **the most connected countries are also the largest ones.** This is also shown in Figure 3 (based on million inhabitants). The most connected country is Germany, with around 12% of the collaborations within the network involving German participants, followed by Spain (11%), Italy (10%), and France (10%). Overall, 79.3% of the collaborations involve participants from EU-15 countries against 9.8% for EU-13 countries (and respectively 6.6% for associated countries and 4.2% for third countries). Poland is the EU-13 country with most connections (1.8% of all connections).

[15] To ensure robustness, other variables describing country size have been tested, such as the national population of researchers (source: Eurostat). This does not affect the key messages from the analysis. However, using population reduces data noise over time and ensures reliability in the evolution of the ranking.

Figure 3 Country size and share of connections under Horizon 2020



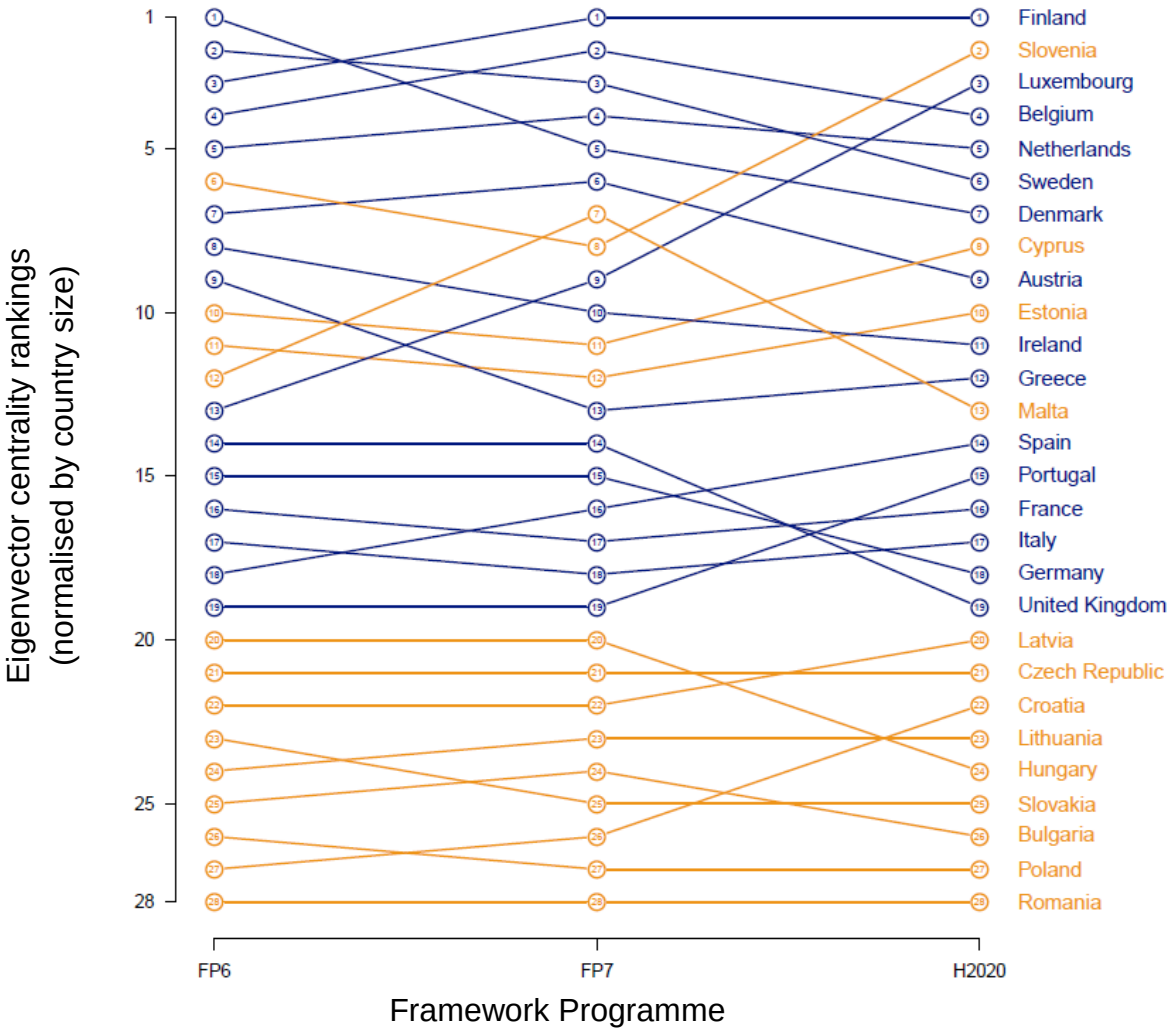
Source: Author's calculations based on CORDA data (Horizon 2020) and World Bank (country population).

Normalisation for size leads to an overall different picture. Figure 4 presents the evolution of the same centrality coefficients as in Figure 2 when normalising by country size (as measured by population). Different trends can be observed. The most central country, relative to its size, is actually Finland. Some EU-13 countries also appear to be very central in the network for their size: **Slovenia is now the second most central country in the network after normalisation for size effect**. This was not the case in previous programmes: Slovenia was ranked 5th in FP6 and 8th in FP7 in terms of centrality. This is the most striking increase observed within all EU countries. Luxembourg, the Netherlands, Belgium, Sweden and Denmark are next in terms of size-normalised centrality measure. Among EU-13 countries, Cyprus and Estonia also present strong centrality after normalisation.

Hence EU-15 and EU-13 groups are not homogenous groups, with some EU-13 countries being more central, relative to their size, than most EU-15 countries. The position of the UK and Hungary in this ranking dropped significantly between FP7 and Horizon 2020[16]. Still, several EU-13 countries are consistently found at the bottom of the ranking over the period.

[16] The position of Malta also decreased significantly over the same period, but it follows a significant increase in FP7 and the position of small countries is more volatile in the ranking.

Figure 4 Network positions of participants by EU country normalised by country size

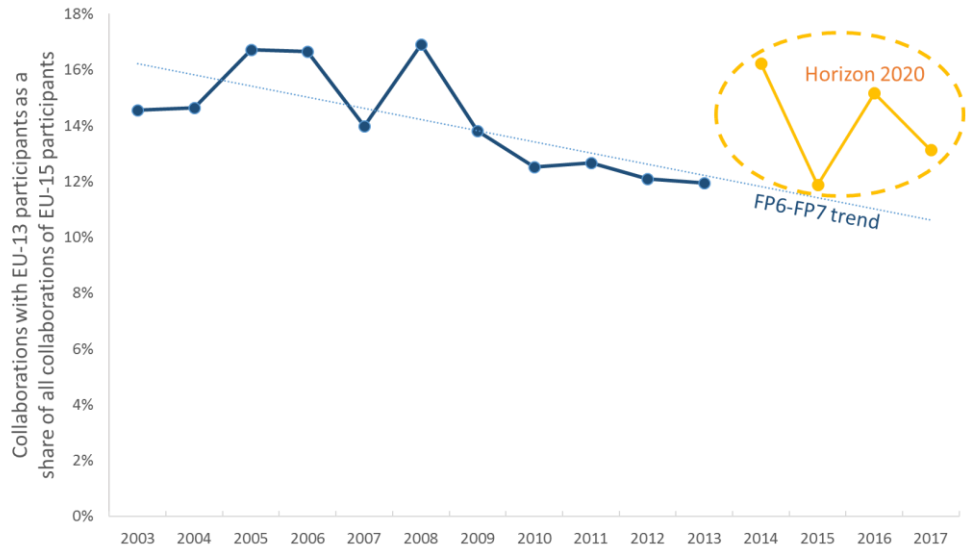


Source: CORDA data (Framework Programme) and World Bank (country population).

Opening of EU-15 to EU-13 participants

In terms of dynamic, while EU-15 participants seem to have reduced their collaborations with EU-13 participants between FP6 and FP7, they appear to have opened up to EU-13 participants in Horizon 2020. In FP6, the percentage of connections between EU-15 participants and EU-13 participants was 15.3% of all collaborations from EU-15 participants (Figure 5). While this percentage decreased to 13.1% during FP7, it increased again to 14.5% in Horizon 2020. Hence, while the opening of EU-15 countries to EU-13 countries seems to have worsened during FP7, the situation has improved with Horizon 2020. In parallel, the share of collaborations between EU-13 participants with other EU-13 has been stable since FP6.

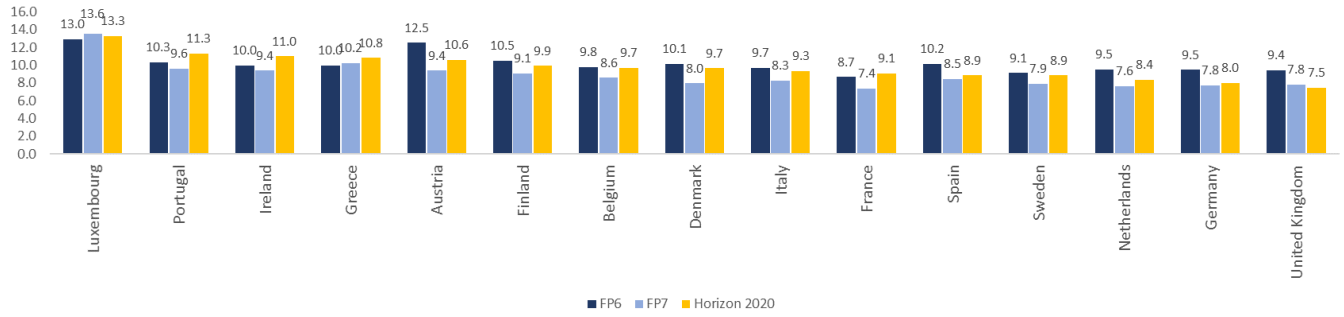
Figure 5. Connections with EU-13 participants as a percentage of all connections of EU-15 participants



Source: CORDA data.

The evolution of these collaborations between EU-15 and EU-13 countries are detailed for each EU-15 country in Figure 6[17]. **While there is a clear general decrease in the collaborations with EU-13 participants between FP6 and FP7, almost all EU-15 countries collaborate more often with EU-13 participants in Horizon 2020 compared to FP7**. The only exceptions are Luxembourg and the United Kingdom, which are also respectively the countries with the largest (13.3% in Horizon 2020) and the smallest share of connections (7.5%) with EU-13 participants. Since FP6, this trend has been continuously negative only for the UK and continuously positive only for Greece.

Figure 6 Connections with EU-13 countries as a percentage of all connections



Source: CORDA data.

The top EU-15 participants that present the largest numbers of collaborations with EU-13 participants in Horizon 2020 are Fraunhofer (DE), CNR (IT), CNRS (FR), CEA (FR) and VTT (FI). These participants are the most important actors in terms of bridging EU-15 and EU-13 countries. Top 5

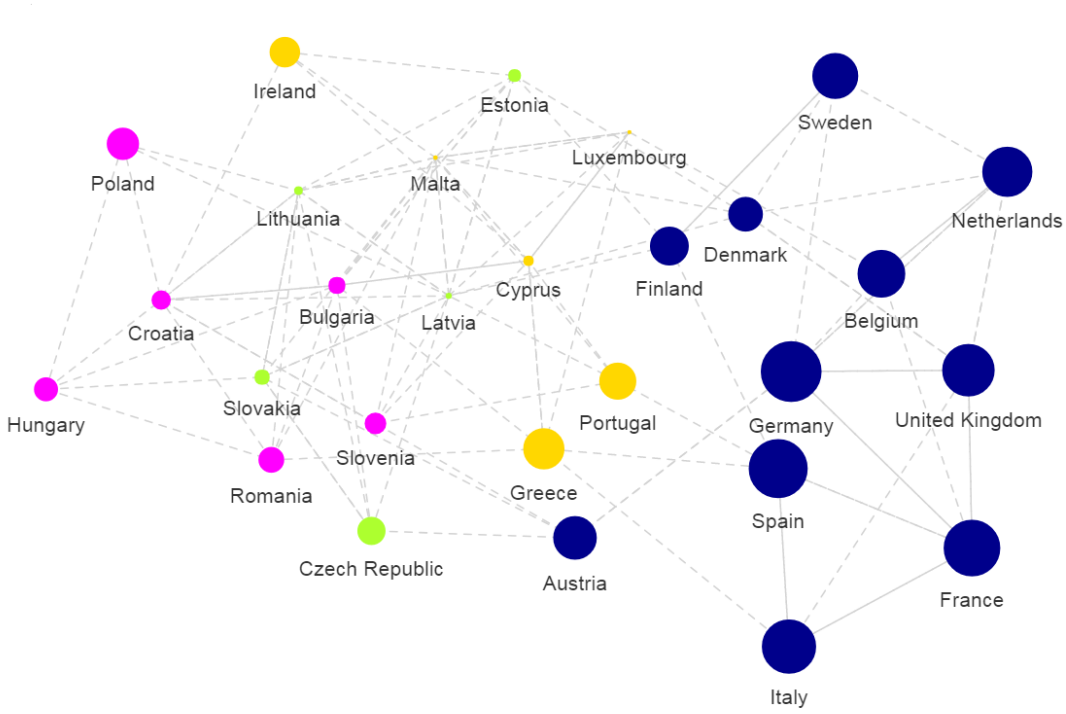
[17] The patterns in Figure 5 and Figure 6 are qualitatively similar. But due to collaborations within country groups, the aggregated values do not numerically correspond to the average of countries.

EU-15 participants that present the highest share of collaborations with EU-13 participants in their collaborations are ENEA (IT), NERC (UK), CINECA (IT), UoA (EL) and JUELICH (DE).

5 Preferences in collaborations

Figure 7 shows the country relatedness network, which expresses collaboration preferences between countries. To compute this relatedness, the number of connections between two countries is divided by the number of connections expected by chance[18], i.e. based on the amount of participations of both countries (Hidalgo et al. 2007, Balland et al. 2018). In Figure 7, the top four strongest connections of each country are represented. As a result, participants appear to show very specific preferences in their cross-country collaborations. Several clusters of countries can be observed[19]. Countries in a same cluster of strong preferences are represented by the same colour. Participants from Baltic countries, Czech Republic and Slovakia tend to collaborate more with each other than what would be expected statistically (green cluster). Cyprus, Greece, Ireland, Luxembourg, Malta, and Portugal form another group of preferred connections (yellow cluster). These two groups bridge to some extent the other two clusters, which are formed respectively by large EU-13 countries (pink cluster) and large EU-15 countries (blue cluster). Overall, **these preferences show that different forms of proximity, including cultural and geographical proximities tend to shape the structure of the Horizon 2020 network.**

Figure 7 Preferred connections – network relatedness (Horizon 2020)



Note: Colours based on community structure (Blondel et al, 2008). The top four strongest connections (after normalisation) of each country are represented. A plain link indicates that the connection is in the top four connections of both countries. A dashed link indicates that the connection is in the top four of one of both countries. Source: CORDA data.

[18] Relatedness is computed using the EconGeo software, implemented as a R package (Balland, 2017).
[19] Communities within the network are based on the multi-level modularity optimisation algorithm for finding community structure as described by Blondel et al. (2008).

6 Conclusions

The overall network of participants in Horizon 2020 shows **that the most central countries in the network are also the largest ones**: Germany, France, the UK, Italy, and Spain. This observation is expected as country size correlates with the number of participations in the Framework Programme and the number of collaborations between participants. **Participants from associated countries and third countries are on average less central than EU participants**, but these country groups are very heterogeneous. For instance, Switzerland and Norway are very important actors in the network.

When examining the evolution over time and normalising for this size effect, results show a different picture. **Some countries punch above their weight: when normalising by country size, the most central country in Horizon 2020 is Finland, followed by Slovenia. Slovenia, Cyprus, Estonia and Malta are as central as EU-15 countries.** Still, other EU-13 countries are found at the bottom of the ranking. Slovenia, Luxembourg, Croatia, Portugal and Cyprus show the most striking increases in terms of size-normalised centrality from FP7 to Horizon 2020, while the UK and Hungary dropped positions.

Between FP6 and FP7, EU-15 participants have been reducing to some extent their collaborations to EU-13 participants. However, **this trend has reverted in Horizon 2020**, as EU-15 countries appear to have opened up to EU-13 participants compared to FP7. This is consistent with the finding of the Monitoring Flash #1, which shows that there are indications that an increasing share of multi-partner Horizon 2020 projects involve at least one EU-13 participant, reversing a downward trend observed under FP7. Moreover, **the network of participations to the Framework Programmes appears to be very dynamic over time** and tends to be opening to less connected participants. These trends deserve further detailed attention, as well as a more frequent update of observations.

Participants appear to show very specific preferences in their cross-country collaborations. As result, **geographical and cultural proximities between participants seem to play an important role in shaping the structure of the Horizon 2020 collaboration network**.

Overall, these results show a network that is relatively open, albeit with some persistently peripheral countries. The analysis also presents encouraging trends regarding the openness of the network, in particular between FP7 and Horizon 2020. However, there is still room for improving the connectivity and centrality of several countries, especially countries with lower R&I performance. This calls for continuous emphasis and effort, in particular for these countries, to ensure the openness of the programme's networks to their entities. This could be achieved through support activities such as organising information/networking campaigns, boosting national capacity building, offering further opportunities to entities for accessing successful R&I projects and established networks, or by supporting matchmaking between potential participants informed by analytics and network affinities.

References

Balland, P.A. (2012). Proximity and the Evolution of Collaboration Networks: Evidence from Research and Development Projects within the Global Navigation Satellite System (GNSS) Industry. Regional Studies, 46 (6): 741-756.

Balland, P.A. (2017). Economic Geography in R: Introduction to the EconGeo Package. Papers in Evolutionary Economic Geography, 17 (09): 1-75.

Balland, P.A., Boschma, R., Crespo, J. and Rigby, D. (2018). Smart Specialization policy in the EU: Relatedness, Knowledge Complexity and Regional Diversification. Regional Studies, forthcoming.

Balland, P.A., and Ravet, J. (2018). Dynamic Network Analysis of the EU R&I Framework Programme. European Commission report.

Blondel, V.D., Guillaume, J.-L., Lambiotte, R. and Lefebvre, E. (2008). Fast unfolding of communities in large networks. J. Stat. Mech. P10008.

Elsevier (2017). Study on overall output of select geographical group comparators and related FP7- and Horizon 2020-funded publication output. European Commission report.

European Commission (2017). Interim Evaluation of Horizon 2020. Staff Working Document. SWD(2017)220.

European Commission (2018). Impact Assessment of Horizon Europe. Staff Working Document. SWD(2018)307.

Hidalgo, C.A., Klinger, B., Barabási, A.L., Hausmann, R. (2007). The product space conditions the development of nations. Science 317: 482-487.

Ripley R., Snijders T.A.B., Boda Z., Voros A., Preciado P. (2016). Manual for RSiena. Available at: http://www.stats.ox.ac.uk/~snijders/siena/RSiena_Manual.pdf.

Science-Metrix (2015). Study on Network Analysis of the 7th Framework Programme Participation. European Commission report