JRC SCIENTIFIC AND POLICY REPORTS

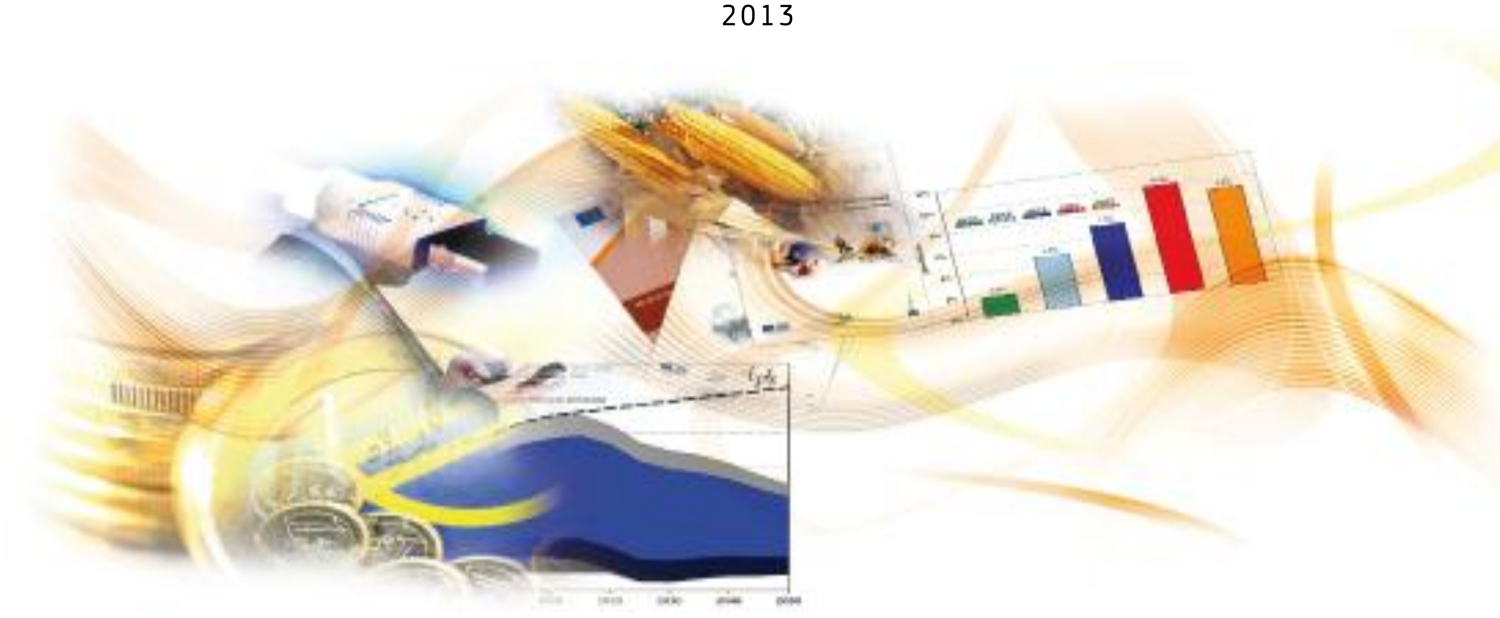
Identifying European Poles of Excellence: The Methodology

JRC-IPTS / DG CONNECT Joint Project nr 31786-2010-06

Authors: Giuditta De Prato, Daniel Nepelski

Editors: Marc Bogdanowicz, Zoe Kay

2013



Report EUR 26264 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

Contact information
Address: Edificio Expo. c/ Inca Garcilaso, 3. E-41092 Seville (Spain)
E-mail: jrc-ipts-secretariat@ec.europa.eu
Tel.: +34 954488318
Fax: +34 954488300

http://ipts.jrc.ec.europa.eu
http://www.jrc.ec.europa.eu

JRC85356

EUR 26264 EN

ISBN 978-92-79-34484-8 (pdf)

ISSN 1831-9424 (online)

doi:10.2791/35893

Luxembourg: Publications Office of the European Union, 2013

Printed in Spain

Acknowledgments

This analysis was produced by the Information Society Unit at the European Commission's Joint Research Centre - Institute for Prospective Technological Studies for DG Communications Networks, Content and Technology (Project Nr 31786-2010-06). It is part of the project: "European ICT Poles of Excellence" (EIPE) jointly funded by DG CONNECT and JRC-IPTS. The authors wish to thank and acknowledge the following experts and colleagues for their valuable input and comments: Zoe Kay (DG CONNECT), Jean-Paul Simon (JPS Consulting), Andrea De Panizza (OECD), Marc Bogdanowicz (JRC-IPTS), Paul Desruelle (JRC-IPTS) and all the participants at the EIPE Expert Workshop, Seville, 11-12 November 2010. Finally, thorough checking and editing of the text by Patricia Farrer is gratefully acknowledged.

How to cite this publication

Please cite this publication as:

De Prato, G. and Nepelski, D. (2013) 'Identifying European ICT Poles of Excellence. The Methodology', JRC Scientific and Policy Reports – EUR 26264 EN. Seville: JRC-IPTS.

Preface

The European ICT Poles of Excellence (EIPE) research project is a joint project of DG CONNECT and the JRC Institute for Prospective Technological Studies (Project Nr 31786-2010-06). The overall objectives of the EIPE project are to set the general conceptual and methodological conditions for defining, identifying, analysing and monitoring the existence and progress of current and future EIPE, in order to develop a clear capacity to distinguish these among the many European ICT clusters, observe their dynamics and offer an analysis of their characteristics.

The EIPE project spanned the period between 2010 and 2013. Over this time, it developed a tool based on a database of original ICT activity indicators, which was enriched with geographical information to allow localisation and aggregation at NUTS 3 level. The tool helps to answer such questions as:

- How is ICT R&D, innovation and economic activity distributed in Europe?
- Which locations are attracting new investments in the ICT sector?
- What is the position of individual European locations in the global network of ICT activity?

The EIPE project had four main steps (see Figure 1). First, European ICT Poles of Excellence were defined. Second, a statistical methodology to identify EIPE was elaborated. Third, the empirical mapping of EIPE was performed and fourth, an in-depth analysis of five NUTS 3 regions was undertaken. This work was documented in a series of five EIPE reports:

- Defining European ICT Poles of Excellence. A Literature Review,
- Identifying European ICT Poles of Excellence. The Methodology,
- Mapping the European ICT Poles of Excellence. The Atlas of ICT Activity in Europe,
- Analysing the European ICT Poles of Excellence. Case studies of Inner London East, Paris, Kreisfreie Stadt Darmstadt, Dublin and Byen Kobenhavn,
- Key Findings and Implications of the European ICT Poles of Excellence project.

Figure 1: Overview of the EIPE project

STEP	Defining European ICT Poles of Excellence	Methodology to identify EIPE	Mapping EIPE	Zooming-in at the European ICT landscape
INPUT	Literature Review Taking stock of existing initiatives and case studies	Elaboration of indicators Identification of data sources Composite indicators	4 composite indicators based on 42 indicators on ICT R&D, Innovation and Business for the whole Europe at NUTS 3 level	Detailed information on ICT activity in Inner East London, Paris, Kreisfreie Stadt Darmstadt, Dublin and Byen Kobenhavn
OUTPUT	Definition of European ICT Poles of Excellence	Methodology to identify EIPE	Atlas of European ICT activity	An in-depth analysis of 5 key ICT locations in Europe
EIPE Report	1	2	3	4

More information on the European ICT Poles of Excellence (EIPE) project can be found at: http://is.jrc.ec.europa.eu/pages/ISG/EIPE.html

Table of Contents

1 Introduction

ICT-related innovation is considered at the core of economic recovery, growth and productivity. On the one hand, technological progress in ICT-producing sectors is an important driver of growth, as evidenced for example by its role in the productivity acceleration observed in the late 90s in the US. On the other hand, ICT-enabled innovation in ICT–using sectors has provided the base for permanent and widespread growth-enhancing effects of ICT adoption throughout the economy.

The 2009 Commission's Communication entitled "A Strategy for ICT R&D and Innovation in Europe: Raising the Game"[1] proposes to reinforce Europe's industrial and technology leadership in ICT. Building on Europe's assets, for example its many ICT industrial clusters, the strategy seeks to step up the effort in ICT Research and Development and Innovation (R&D&I). In particular, the Communication anticipates a landscape where by 2020, "(…) *Europe has nurtured an additional five ICT poles of world-class excellence* (…)".

Further, the Commission's "Digital Agenda for Europe" (DAE)[2] organised around 7 pillars, meant to "reboot the EU economy" and to enable Europe's citizens and businesses to get the most out of digital technologies. In its Pillar V on Research and Innovation, the DAE aims at increasing the amount of resources invested in ICT R&D and at developing world class infrastructure and adequate funding to attract Europe's best minds to research.

Finally, the policy context of this study is also rooted in a strong geographical rationale, which is in turn motivated by the debates around the role of European regions in Innovation policies, giving rise to a rethinking of regional policies and regional funding of (ICT-related) innovation, (ICT) technology transfer and more generally technology-driven economic growth.

Taking into account the above, the EIPE study aims to identify ICT R&D&I-related activities which are geographically concentrated and which demonstrate high performance in ICT innovative activities. It also aims to help map the dynamics of ICT-related innovation and economic geography in Europe, pointing to the presence and possibly the emergence of agglomerated and globally performing ICT activities: i.e. the European ICT Poles of Excellence.

An additional challenge of the EIPE project was that this identification process had to be based only on the analysis of quantitative data, and built on a set of relevant criteria leading to measurable indicators. The present report documents the methodologies and data sources used for this purpose.

1 Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0116:FIN:EN:PDF

2 Available at: http://ec.europa.eu/digital-agenda/en/digital-agenda-europe

2 The methodology to identify EIPE

2.1 The definition of European ICT Poles of Excellence

The foundations of the EIPE project are rooted in a vast body of scientific literature that has described and analysed for almost a century[3] the spatial agglomeration of economic activities. More recently analysis of knowledge-intensive and ICT-related activities, at local and at global level has been included. This literature review was the subject of the first EIPE report which concluded with a definition of European ICT Poles of Excellence (Nepelski & De Prato, 2013).

This review shows that the concept of European ICT Poles of Excellence can be associated with many existing views and formulations (Districts, Clusters, Centres, etc.). Hence, the definition of European ICT Poles of Excellence builds on the main modern concepts and models emerging from the existing academic and non-academic literature. However, it differentiates the concept of Poles of Excellence from neighbouring concepts such as those of Industrial clusters, Innovative regions or Centres of excellence to name a few. None of these concepts, theoretical or empirical, fully meets the requirements and specificity which the EIPE study has determined for a location to be considered a European ICT Pole of Excellence.

As seen in the first EIPE report (Nepelski & De Prato, 2013), the increasing globalization of economic activity on the one hand, and the pervasive role of knowledge in the economy on the other, is affecting the spatial distribution of economic activity. These changes in the spatial distribution of economic activities are evident both in the mutating role of countries and in the emergence of different productive realities within countries. The concept of EIPE integrates those most recent aspects in its definition.

In particular, analysts observe two conflicting forces governing the globalisation: on the one hand, there is a trend towards the global expansion and redistribution of economic and knowledge-intensive activities, *and* towards the concentration of such activities in limited spatial areas or regions, in particular large metropolitan areas, on the other hand. Or what has been referred to as the paradox of the 'sticky places within a slippery space' (Dunning, 2002; Liu & White, 2001).

Being the above literature review and the resulting conceptual decisions taken in the project, the characteristics of a European ICT Pole of Excellence are organized within a two-dimensional scheme (see Figure 2) capturing simultaneously:

- a strong performance in knowledge (R&D and Innovation) and business activities,
- with all activities being investigated in terms of their characteristics: geographical agglomeration, and also international reach and centrality in global networks.

This view encompasses a certain affiliation with the concept of industrial clusters, but it clearly broadens the perspective on several essential points by:

- expanding the scope of the observed activities from business (production) to knowledge-related activities (R&D and Innovation), thus acknowledging the importance given today to the knowledge function in advanced economies.
- assessing the global internationalisation of the production and R&D&I activities.
- putting an additional emphasis on the network position of any individual location, the centrality in a network being taken as an indicator of the strategic role of the location in the global landscape of R&D&I and production activities.[4]

The above perspective takes on board and echoes two theoretical approaches that articulate the two dimensions of the EIPE measurement framework:

---

[3] Starting in particular with Marshal and Weber in the conceptualisation of industrial districts.

[4] Several authors consider such central position as giving access and control over resources or information that are crucial for the overall activity of an industry or a market. See for example: Becker 1970; Nepelski & De Prato (2012); Powell, Koput, & Smith-Doerr (1996).

- the CDM model, developed by Crépon, Duguet and Mairesse (1998):[5] this structural model aims to explain productivity by innovation outputs driven by R&D investments, reflecting interdependencies of knowledge and business activities.
- the 'buzz versus pipelines"[6] balance which takes into account the differentiated and mutual benefit of proximity ("the buzz") and global networking ("the pipelines") for agglomeration economies.

In addition, because of the study's policy context – ICT R&D and Innovation, the proposed definition is explicitly technology specific: it emphasises the observation of ICT as a technology, and also of its supply side, across Europe.

Hence, EIPE are defined as follows:

> **European ICT Poles of Excellence (EIPE) are geographical agglomerations of best performing Information and Communication Technologies production, R&D and innovation activities, located in the European Union, that exert a central role in global international networks.**

This definition needs now to be made operational, by determining the best available data sources, indicators and measurements that will help us to identify and locate EIPE in Europe on the basis of a set of quantitative observations.

According to the proposed definition presented above and the earlier reviewed empirical studies on the identification and performance assessment of spatially-bounded research, innovation and business activity (Nepelski & De Prato, 2013), the proposed framework for the selection of EIPE indicators accounts for the two important above-mentioned dimensions. Accordingly, this section prepares the ground for the selection of indicators that are relevant for the identification of ICT-related activities and that describe their agglomeration, internationalisation and networking characteristics.

2.2 The conceptual and operational framework

First, the definition of EIPE recognizes that R&D&I activities are interlinked with (industrial) business activity. More precisely, there is a mutual inter-dependency between R&D&I and business activities, which implies that these are often co-located.[7] This is to say that for an EIPE, neither of these activities is likely to exist in vacuum, but that instead they are embedded in common spatially-agglomerated industrial and business activities, supporting and forming the basis of inventive activity.

This indicates that in the process of selecting the indicators for identifying ICT Poles of Excellence it will be important to account for both the research and development performed in a given location, innovation as well as the business activity. In other words, **we will observe three activities** in the EIPE, echoing the CDM model and its acknowledgement of the innovation stage as an intermediary one between the R&D and the business activity on the market. In practice, R&D, innovation and economic activities will each be taken in account.

[5] Crépon B., Duguet E., Mairesse J. (1998) Research, Innovation and Productivity: an Econometric analysis at the firm level. NBER Woorking paper n° 6696.

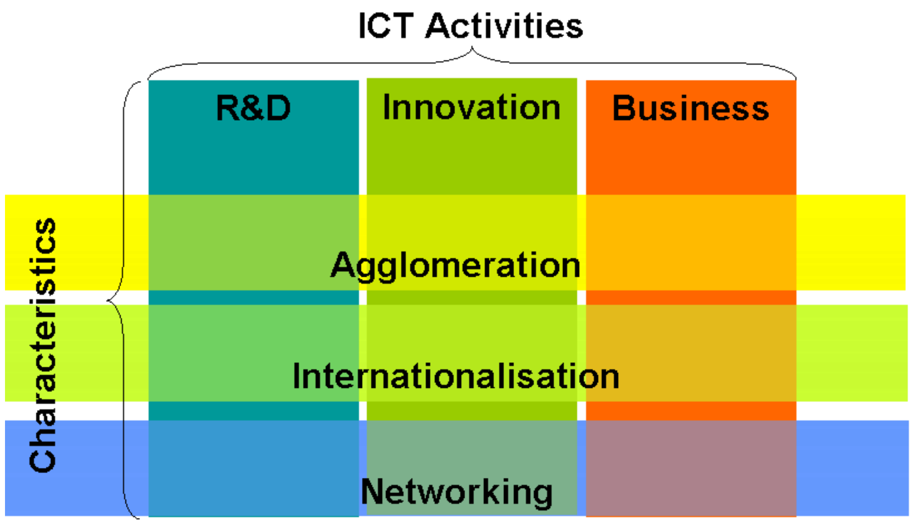
[6] See for example: Rodríguez-Pose A. (2012). Cluster vs. pipeline options in innovation policies: A tale from Norway. London School of Economics and IMDEA Social Sciences. Presentation at the "Innovation policies – theoretical and empirical approaches" Expert Workshop. JRC-IPTS. Seville, 21st May 2012.

[7] This point is supported by the fact that localised knowledge spillovers and agglomeration economies foster a local business system towards a specialisation in both production and technology (Paci et al., 1998). This is then reflected by strong co-location patterns of production and research units within close proximity (Defever, 2006; OECD, 2010). Consequently, such an agglomeration of R&D and business activity might include ICT-related research and development business units and institutes together with firms specializing in the production of ICT-related products and services.

Second, the definition points at **three important characteristics of these activities**, acknowledging the debate on the various aspects of proximity and global assets of a location, i.e. agglomeration, internationalisation and networking, where each can be observed (and measured) for each type of activity, i.e. R&D, innovation and business.

This approach creates a matrix of activities and their characterisation, as shown in Figure 2.

Figure 2: A visual approach to the definition of Poles of Excellence



The above mentioned activities are defined and approached in the EIPE project in the following way:

- **R&D activities**

Research and development comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge and the use of this stock of knowledge to devise new applications (OECD, 2002). R&D is often scientific or for the development of particular technologies and is frequently carried out as a corporate or governmental activity (OECD, 2008a). Thus, R&D activity is an investment in knowledge accumulation and in the development of technologies (Parham, 2009). The corresponding assets, i.e. the stocks of knowledge and technologies, are intangible assets whose values are largely unobservable. Thus, although the existence of a link between R&D, technical change, and economic growth is widely acknowledged, this link is difficult to quantify because the benefits from, or output of R&D (a critical component of the link) are not easily measured.

The main practical measurement option put forward has been to approximate the volume of knowledge assets by capitalised R&D expenditures (Fraumeni & Okubo, 2005). Thus, the typical measures of R&D include, for example, R&D expenditure, R&D employment or R&D facilities on the input side (OECD, 2002; Ojanen & Vuola, 2003). On the output side, as the results of R&D are used to reduce the costs or increase the output of final goods, ways of capturing the results of R&D include, for example, calculating the changes in the total factor productivity (Hall, Mairesse, & Mohnen, 2009; Lipsey & Carlaw, 2004), or by conducting firm level surveys investigating the introduction of product or process innovations. Existing data sources also allow us to derive science and technology indicators such as the technology balance, bibliometrics or the “technology intensity” of the products or industries concerned (OECD, 2002).

Subject to data sources availability and their compliance with the EIPE project needs (see Section 2.3), the EIPE project takes into account the above mentioned approaches to identify, measure and map ICT R&D activities across the whole of the EU and at a sufficiently low level of geographic aggregation. In particular, the EIPE project builds up a measurement of ICT R&D activity by observing the actual presence of ICT technology producers (universities, companies, R&D facilities), their R&D expenditures and bibliometric data.

- **Innovation activities**

Innovations comprise technologically new products and processes which have been implemented and significant technological improvements in products and processes (OECD, 2005b). Innovation is part of a business strategy based on turning ideas into value (OECD, 2010). An innovation has been implemented if it has been introduced on the market or used within a production process. Simply put, innovation is an output of the R&D process.

The great variation in innovation processes, in terms of their objectives, organisation, cost, use of research, and so on, also means variation in the problems and constraints with respect to the measurement of innovation. This is exemplified by the fact that firms invest in innovation to gain market share, reduce costs or, more generally, to become more productive. Thus, spending on innovation is greater and different to spending on R&D. To "make value out of ideas", firms invest in R&D *but also* in other tangible and intangible assets, e.g. patents, licences, trademarks, and acquisition of machinery, equipment and software. As a result, among the few available indicators of technology output, patent-based indicators are probably the most frequently used (Griliches, 1990; Hagedoorn & Cloodt, 2003; OECD, 2002). In addition, some outputs of innovation include the enhancement in human capital and new organisational structures. A typical way of capturing these types of innovation is through company-level surveys (OECD, 2005b). Other ways of measuring innovation that make use of the existing data sources rely on the fact that high economic dynamics are a key source and channel of technological and non-technological innovation. Thus, information on entry and exit of companies is exploited as a proxy of technological or commercial opportunities that are introduced on the market (OECD, 2010). These dynamics are exemplified by the economy's share of young fast-growing firms that fuel innovation, developing new goods or improving existing ones, services or processes (Veugelers, 2009). A related measure of such innovation activities is the financing which new enterprises receive, e.g. venture capital. Venture capital funds may invest in the later stages of product development or market launch and, hence, offer a way of measuring the outcome of the innovation process in its final stage.

Subject to data sources availability and their compliance with the EIPE project needs (see Section 2.3), the EIPE project takes into account the above mentioned approaches to identify, measure and map ICT Innovation activities across the whole of the EU and at a sufficiently low level of geographic aggregation. In particular, the EIPE project builds up a measurement of ICT Innovation activities by observing patenting patterns, overall investments in intangibles by ICT companies and venture capital financing of ICT start-ups.

- **Business activities**

These activities relate to the production of tangible and intangible goods and services that are produced and meet the needs of consumers in the market and encompass the aggregate economic activities of the all sectors of an economy.

At the most aggregate level, Gross Domestic Product (GDP) is the most commonly used measure of a country's or a region economic activity (Chiripanhura, 2010). More disaggregated data on business activity show business indicators such as the number of firms, employment, capital, turnover, value added, profits, and wages and salaries (ABS, 2012; Frankish, Roberts, & Storey, 2009).

The type of information on business activity can be divided into two types. The first is reflected in statistics compiled by government bodies or by other official organizations in the form of business registers. The second is data on small businesses and their owners produced by a wide range of non-official organizations that aim to capture the dynamism in the economy.

Subject to data sources availability and their compliance with the EIPE project needs (see Section 2.3), the EIPE project takes into account all the above mentioned approaches to identify, measure and map ICT business activities across the whole of the EU at a sufficiently low level of geographic aggregation. In particular, the EIPE project builds this measurement by observing the actual presence and development of ICT firms (headquarters and affiliates, employment data, turnover and investments).

Concerning the characterisation of these activities, they are defined in the following way:

- **Agglomeration characteristics**

Spatial proximity of similar and related firms and industries and the general tendency of people and economic activity to locate in large cities and economic core regions lead to agglomeration. This process can be seen as the outcome of a process involving two opposing types of forces, that is, agglomeration forces and dispersion forces (Masahisa Fujita & Thisse, 1996). As a result of the interplay between these forces, economic activity, population, employment and wealth are unevenly distributed. Residents, workers, and firms are typically agglomerated in urban areas, which gives rise to the notion of agglomeration economies (M. Fujita & Thisse, 2002; Krugman, 1991; Ottaviano & Thisse, 2004).

In general, measuring the agglomeration of economic activity across spatial units is not a straightforward task. There are two ways of approaching it. The first concerns the problem of spatial concentration of economic activity and aims to identify how agglomerated or dispersed certain activities or industries are (Kominers, 2013). Here typical measures of agglomeration are, for example, geographic concentration indicators such as the Herfindahl index (Barrios, Bertinelli, Strobl, & Teixeira, 2005; Dumais, Ellison, & Glaeser, 2002; Ellison & Glaeser, 1997; Puga, 2010). This measure permits us to observe the geographic distribution of a given activity (Devereux, Griffith, & Simpson, 2004).

The second way of looking at agglomeration relates to the issue of the relative size of economic activity in a certain place (Ciccone & Hall, 1996; World Bank, 2009), its intensity. This approach considers each unit of observation individually and does not relate it to the other spatial units. Intensity measures, for example, the amount of labour, human, and physical capital *relative* to physical space (Ciccone & Hall, 1996) or *relative* to the size of population (Perez & Sanchez, 2002; Wandel, 2010). These measures of agglomeration of activity and/or resources in spatial units allow us to compare inter-regional differences as regards the size of the spatial units (Broekel & Brenner, 2011; Koschatzky & Lo, 2007). All in all, it permits us to compare heterogeneous spatial units with each other with respect to the level of a given activity conducted within their borders and therefore is the most adequate for the EIPE project (see also 2.2).

Subject to data sources availability and their compliance with the EIPE project needs (see Section 2.3), the EIPE project takes into account the above mentioned approaches to identify, measure and map the agglomeration of R&D, Innovation and business activities across the whole of the EU at a sufficiently low level of geographic aggregation. In particular, the EIPE project builds up this measurement by observing the level of agglomeration of technology producers (universities, companies, R&D facilities), R&D expenditures and bibliometric data.

- **Internationalisation characteristic**:

Both R&D&I and economic activity have been going through an intensive process of internationalisation over the last three decades (Doz, Santos, & Williamson, 2001; Dunning, 1994). Multinational companies are increasingly building a new kind of competitive advantage by discovering, accessing, mobilising, and leveraging knowledge and markets from a number of locations across the globe. More and more, firms are locating their R&D centres outside the country where the company is headquartered. This type of spatial division of labour reflects the increasing transfer of sophisticated, knowledge-intensive activities to locations other than companies' domestic markets. The ICT sector and ICT-related activities are the prime example of this internationalisation process (Nepelski & De Prato, 2012; UNCTAD, 2005).

In recent literature, measuring the internationalisation characteristic is done by looking at the collaboration between scientists or innovators, the location of corporate R&D centres for R&D&I activity, the flow of foreign direct investments or the level of export and import for business activity (Dorrenbacher, 2000; OECD, 2005a).

Subject to data sources availability and their compliance with the EIPE project needs (see Section 2.3), the EIPE project takes into account the above mentioned approaches to identify, measure and map the internationalisation of R&D, Innovation and business activities across the whole of the EU and at a sufficiently low level of geographic aggregation. In particular, the EIPE project builds up this measurement by observing the level of internationalisation of ICT firms through the location of their R&D centres, their patented co-inventions and the location of their affiliates.

- **Networking characteristic**:

One of the consequences of companies' decisions concerning the location and internationalisation of R&D&I and economic activity is the amplification of R&D, innovation and production networking and the emergence of networks of activity that span the globe (De Benedictis & Tajoli, 2011; De Prato & Nepelski, 2012). An important implication of this process for both MNEs and locations of R&D activity is that being connected globally is increasingly recognized as a crucial determinant of the position of individual MNEs and locations in a global hierarchy (Cantwell & Janne, 1999; Malik, 2013; Meyer, Mudambi, & Narula, 2011).

The application of a network perspective that allows us to treat geographically-dispersed R&D&I and economic activity as a system of inter-linked activities becomes crucial in the process of determining the position in the network of locations engaged in these activities. In order to measure the networking characteristics of observed units, network analysis can be used. Network analysis tools provide a range of measures of the number and quality of connections between them. These allow us to capture their positions in the network e.g. core - periphery, and their role, e.g. broker - gatekeeper.[8]

Subject to data sources availability and their compliance with the EIPE project needs (see Section 2.3) the EIPE project takes into account the above mentioned approaches to identify, measure and map the network positions of R&D, innovation and business activities across the whole of the EU at a sufficiently low level of geographic aggregation. In particular, the EIPE project builds up this measurement by observing FP7 programme networks, collaboration in patented co-inventions, and location of business affiliates to identify the position of each region in those networks.

2.3 The methodological framework

Defining the ICT activity

Due to the policy-driven scope – ICT R&D and Innovation, the EIPE project concentrates on ICT-related activities, i.e. those concerning ICT technologies and the ICT sector. Thus, ICT activities are defined either through characterization of the technological activity or characterization of the firms performing it.

With respect to the characterization of firms performing ICT-related activities, e.g. R&D expenditures, venture capital funding, company location investment or the number of ICT companies, the NACE Rev 2 definition of the ICT sector is used (OECD, 2007). Following the NACE Rev 2 definition, the ICT sector is composed of the following sub-sectors:

[8] For network analysis methodology applied in the current project, see Annex 6.1.

Table 1: ICT sector definition based on NACE Rev. 2

NACE code	**Sector name**
ICT manufacturing	
261	Manufacture of electronic components and boards
262	Manufacture of computers and peripheral equipment
263	Manufacture of communication equipment
264	Manufacture of consumer electronics
268	Manufacture of magnetic and optical media
ICT services	
4651	Wholesale of computers, computer peripheral equipment and software
4652	Wholesale of electronic and telecommunications equipment and parts
5820	Software publishing
61	Telecommunications
62	Computer programming, consultancy and related activities
631	Data processing, hosting and related activities; web portals
951	Repair of computers and communication equipment

With respect to the technology, examples of the characterization used include:

- Computer science and engineering with respect to university faculties,
- Computer science with respect to scientific publications,
- ICT hardware and software with respect to R&D activity performed in R&D centres,
- ICT technological fields defined by the International Patent Classification (IPC) system with respect to patents.

Throughout the EIPE project, several approaches to specifying and defining the technology or the sector are taken, the selection criteria varying between data sources.

Choosing the spatial unit of observation

One of the central problems in the quantitative analysis of the geography of economic activity is the lack of data at regional level with a satisfactory level of granularity (Koschatzky & Lo, 2007). A "region" is defined as a tract of land with more or less definitely marked boundaries, which often serves as an administrative unit below the level of the nation state. A region is an attempt to group together populations or places with enough in common to comprise a logical unit for administrative purposes. It is a recognition that spatial differences require appropriate administrative structures (EUROSTAT, 2009).

Two types of regional division are usually recognized:

- normative regions reflect political will; their boundaries are fixed in terms of the remit of local authorities and the size of the region's population regarded as corresponding to the economically optimum use of the resources they need to accomplish their tasks;
- analytical (or functional) regions are defined in terms of particular analytical requirements; they categorise areas on the basis of specific geographical criteria, such as altitude or soil type, or by economic and social criteria, such as the homogeneity, complementarity or polarisation of regional economies.

In the EIPE project, we make use of the first type of regional classification, which follows Eurostat's "Nomenclature of Statistical Territorial Units" (NUTS) as a single, coherent system for dividing up the European Union's territory in order to produce regional statistics for the Community. For practical reasons connected with data availability and regional policy implementation, the NUTS classification is therefore based largely on the institutional divisions applied in the Member States. The NUTS regional classification system has a hierarchical structure and includes three categories: NUTS 1, 2, and 3. NUTS subdivides each Member State into a number of regions at NUTS 1 level.

Each of these is then subdivided into regions at NUTS level 2, and these in turn into regions at NUTS level 3.

The NUTS Regulation lays down the following minimum and maximum population thresholds for the average size of the NUTS regions:

Level	Minimum	Maximum	Number of units in the EU27
NUTS 1	3 Million	7 Million	97
NUTS 2	800 000	3 Million	271
NUTS 3	150 000	800 000	1303

The standard level of regional data availability provided by, for example, EUROSTAT is NUTS level 2. Only for certain variables, NUTS level 3 is also available, but generally this is the exception. For some statistics and some countries only NUTS level 1 data are available.

For the purposes of this study, the NUTS 3 level was chosen as the unit of analysis, as it allows us to collect and compare data in a harmonised and standardized way across the entire European Union. This unit of analysis gives us the (theoretical) opportunity to observe over 1300 spatially standardised areas across the EU, while avoiding the administrative boundaries of a NUTS level 2 analysis. It also facilitates the later re-aggregation towards functional regions, outside the limits of too constraining administratively predetermined boundaries.

However, because different data providers use different data formats in reporting the names of organisations, the categorisation of data, the location and geographic information (e.g. city, ZIP code, or less commonly NUTS 3 region), the geographical information provided (e.g. city or ZIP code), needed to be matched with its equivalent in the NUTS classification at level 3. The final objective was to provide consistent indicators, which were relevant for the pursued purposes, representative of the EU 27 and which could be (dis)aggregated to the desired level, i.e. NUTS 3 level.

Accounting for region's size

In order to account for size differences between the regions, normalization with respect to the number of inhabitants is made, using a modified version of the Balassa index. The aim of the Balassa index is to measure the relative specialization or, in a spatial context, the agglomeration level of a region's given characteristics, e.g. ICT employment (JRC-IPTS, 2007). The advantage of the Balassa index over a simple comparison of shares in total employment is that relative size is also taken into account here.

The original Balassa index compares a region's share of a specific variable, for example, employment in the ICT sector, with the region's share in the general variable, i.e. in this case the total employment. However, because of the focus of this project, which aims to map ICT-related activities and the lack therefore of a number of data at the general level, the normalisation is conducted with respect to a region's share in the overall population of the EU. This index calculation allows us to take into account the agglomeration of activities irrespective of the size of the region (measured by the population of the region).

Thus, the agglomeration measure based on the Balassa index as used in the EIPE project is:

$$Aggl_{ij} = \frac{X_{ij} / \sum_j X_{ij}}{Pop_j / \sum_j Pop_j}. \qquad (1)$$

where *X* is the value of indicator *i*. Population in region *j* is denoted as Pop_j.

Selecting and processing data sources

The choice of the spatial unit of observation, i.e. NUTS 3, and the policy-driven focus on ICT creates a double constraint on data. It implies that in most cases, there is no official data available to illustrate the activities and characteristics as defined for the purpose of the EIPE project. Hence, a number of data selected for the EIPE project come from non-official data sources, e.g. private databases. To guarantee the quality of the collected indicators, a range of the most reliable and recognized data providers were carefully tested and selected, such as Thomson Reuters for bibliometrics, Bureau van Dijk for company-level information, Dow Jones for venture capital data, etc. Their limitations are well known by practitioners, but they have been selected as, after testing, they were shown to be useful and relevant for the purpose of this project.

The eight primary data sources used in EIPE are the following: FP7 data on FP participation from EC DG Connect, REGPAT by OECD, QS World University Rankings by QS, Web of Science by Thomson Reuters, Design Activity Tool by IHS iSuppli, European Investment Monitor by Ernst & Young, ORBIS by Bureau Van Dijk, and VentureSource by Dow Jones.[9]

More details about these data sources can be found in Chapter 5.

Selecting indicators

A list of indicators for the EIPE project was carefully selected on the basis of the above-described framework of activities and their characteristics and the discussion on their empirical measurements. In this selection process, the following additional criteria were applied:

- Validity: an indicator must be able to capture a relevant dimension of the issues at stake. In order to ensure this, indicators whose use can be traced and validated by previous research literature were selected (see Section 2.2).
- Measurability: an indicator must be measurable. It needs to have a quantifiable dimension, reflecting its most relevant characteristic on a scale.
- Universality: as indicators must be sufficiently comparable across the spatial units of observation, they need to be implemented across the entire population of the spatial units, i.e. all EU 27 NUTS 3 regions, in order to allow comparison.
- Technological relevance: indicators need to reflect the technological purpose (ICT as technology or sector) pursued by the study.
- Reliability: preference was given to indicators issued by sources considered reliable, and to topics that allow factual observation (i.e. published patents or articles, company registers) rather than experts' opinions or surveys.

[9] Some secondary data sources were used such as the (ICT) industrial scoreboard (JRC-IPTS). They are not listed here as they were used as secondary tools to support the processing and extraction of data from the primary ones.

3 EIPE indicators

Table 2 offers a first schematic presentation of the organisation of the nine EIPE assessment criterion and their corresponding indicators around the three activities and their three characteristics, as observed in this study. It serves as a visual tool to capture the overall organization and labelling of the indicators.

Table 2: An overview of the groups of assessment criterion in EIPE

	ICT activities		
Characteristics	**R&D**	**Innovation**	**Business**
Agglomeration	Agglomeration of R&D activities ID: AgRD	Agglomeration of Innovation activities ID: AgIn	Agglomeration of Business activities ID: AgBuss
Internationalisation	Internationalisation of R&D activities ID: IntRD	Internationalisation of Innovation activities: ID: IntIn	Internationalisation of R&D activities: ID: IntBuss
Networking	Networking of R&D activities ID: NetRD	Networking of Innovation activities ID: NetIn	Networking of Business activities ID: NetBuss

Special care was taken to select a representative mix of indicators, representing each type of activity and its characteristics. However, due to the limited availability of data and potential indicators meeting the requirements of this study, it was impossible to maintain an equal number of indicators for each assessment criterion. This limits the diversity of information used to evaluate each individual assessment criterion. However, the final results of the EIPE project arise from an aggregation exercise on all 3 activities (composite indicator) and all of their corresponding individual assessment criteria, and therefore this final result is not strongly influenced by any of those individual criteria and the indicators that compose them.

The full selection of the 42 EIPE indicators meeting the characteristics specified by the definition, framework and criteria, can be found in Table 3. Those indicators and their characteristics are further described in the next Sections 3.1 to 3.3 of this report. A detailed description of some more advanced methodologies applied to elaborate several indicators is given in the Annexes.

Table 3: Overview of the EIPE indicators: the EIPE ID card

Activity	Characteristic	Name of Indicator	Indicator ID	Nr
R&D	Agglomeration	Universities ranked in the QS University Ranking	AgRD 1	1
		Academic ranking of a Computer Science faculty	AgRD 2	2
		Employer ranking of a Computer Science faculty	AgRD 3	3
		Citations ranking of a Computer Science faculty	AgRD 4	4
		R&D expenditures by ICT firms	AgRD 5	5
		FP7 funding to private organisations	AgRD 6	6
		FP7 participations	AgRD 7	7
		FP7 funding to SMEs	AgRD 8	8
		FP7 participations by SMEs	AgRD 9	9
		Location of ICT R&D centres	AgRD 10	10
		Ownership of ICT R&D centres	AgRD 11	11
		Scientific publications in Computer Science	AgRD 12	12
	Internationalisation	Outward ICT R&D internationalisation	IntRD 1	13
		Inward ICT R&D internationalisation	IntRD 2	14
	Networking	Degree in ICT R&D network	NetRD 1	15
		Closeness centrality in ICT R&D network	NetRD 2	16
		Betweenness centrality in ICT R&D network	NetRD 3	17
		Eigenvector centrality in ICT R&D network	NetRD 4	18
Innovation	Agglomeration	Investment in intangibles by ICT firms	AgIn 1	19
		Venture Capital financing to ICT firms	AgIn 2	20
		ICT patents	AgIn 3	21
	Internationalisation	International co-inventions	IntIn 1	22
	Networking	Degree in ICT innovation network	NetIn 1	23
		Closeness centrality ICT innovation network	NetIn 2	24
		Betweenness centrality ICT innovation network	NetIn 3	25
		Eigenvector centrality ICT innovation network	NetIn 4	26
Business	Agglomeration	Location of ICT Scoreboard Headquarters	AgBuss 1	27
		Ownership of ICT Scoreboard affiliates	AgBuss 2	28
		Location of ICT Scoreboard affiliates	AgBuss 3	29
		Location of ICT firms	AgBuss 4	30
		ICT employment	AgBuss 5	31
		Growth in ICT employment	AgBuss 6	32
		Turnover by ICT firms	AgBuss 7	33
		Growth in turnover by ICT firms	AgBuss 8	34
		New business investments in the ICT sector	AgBuss 9	35
	Internationalisation	Outward ICT business internationalisation	IntBuss 1	36
		Inward ICT business internationalisation	IntBuss 2	37
	Networking	In-degree in ICT business network	NetBuss 1	38
		Out-degree in ICT business network	NetBuss 2	39
		Closeness centrality in ICT business network	NetBuss 3	40
		Betweenness centrality in ICT business network	NetBuss 4	41
		Eigenvector centrality in ICT business network	NetBuss 5	42

3.1 ICT R&D activities indicators

3.1.1 Agglomeration of ICT R&D (AgRD)

The 12 indicators characterising the agglomeration of ICT R&D activity are listed and described in Table 4. They are presented together with a first indication of the data sources used and their time coverage. With 12 different measurements, these indicators cover a broad range of aspects related to inputs and outputs of R&D. In particular, they acknowledge the importance given in EIPE to the presence and the quality of major knowledge production organisations, such as universities (and their computer science departments), private and public research centres (in particular those of multinational companies), innovative SMEs,[10] and also R&D expenditures or bibliometric output.

Table 4: ICT R&D Agglomeration indicators (AgRD)

Indicator ID	AgRD 1	AgRD 2	AgRD 3	AgRD 4	AgRD 5	AgRD 6
Name of indicator	**Universities ranked in the QS University Ranking**	**Academic ranking of a Computer Science faculty**	**Employer ranking of a Computer Science faculty**	**Citations ranking of a Computer Science faculty**	**R&D expenditures by ICT firms**	**FP7 funding**
What does it measure?	Measures the number of universities in QS university ranking	Measures the performance of the Computer Science faculty according to the academic ranking of QS	Measures the performance of the Computer Science faculty according to the employer ranking of QS	Measures the performance of the Computer Science faculty according to the citations ranking of QS	Measures the average annual amount spent on R&D in the ICT sector	Measures the amount received for research in ICT R&D
Unit of measurement	Region's share in the total number of EU ranked universities to a region's share in the EU population	The highest rank of a Computer Science faculty in the academic ranking	The highest rank of a Computer Science faculty in the employer ranking	The highest rank of a Computer Science faculty in citations ranking	Region's share in the R&D expenditures by ICT firms in the EU to a region's share in the EU population	Region's share in the total EU FP7 funding to a region's share in the EU population
Definition of ICT dimension	None	Computer science faculty			Based on NACE Rev. 2 (see Table 1)	ICT areas of the FP7 programme (see Section 5.2)
Unit of observation	NUTS 3					
Source	QS World University Rankings by QS (see Section 5.1)				Company-level information: ORBIS by Bureau Van Dijk (see Section 5.7)	FP7 database by EC DG Connect (see Section 5.2)
Reference year(s) considered	2011				2005-2011	2007-2011

[10] In order to account for the differences of size of the NUTS3 regions, normalization with respect to the number of inhabitants is made. A modified version of the Balassa index is used, as defined in Section 2.3. This index aims to measure the *relative* agglomeration level of the given R&D activities in each region. The indicators for which this relative measure is used include indicators from AgR&D 5 to AgR&D 12.

(continued): ICT R&D Agglomeration indicators (AgRD)

Indicator ID	AgRD 7	AgRD 8	AgRD 9	AgRD 10	AgRD 11	AgRD 12
Name of indicator	**FP7 participations**	**FP7 funding to SMEs**	**FP7 participations by SMEs**	**Location of ICT R&D centres**	**Ownership of ICT R&D centres**	**Scientific publications in Computer Science**
What does it measure?	It measures the total number of ICT R&D FP7 projects to which organisations, located in the observed region, have participated to	It measures the total amount of ICT R&D FP7 funding given to SMEs located in the observed region	It measures the total number of ICT R&D FP7 projects to which SMEs, located in the observed region, have participated to	It measures the total number of ICT R&D centres located in the observed region	It measures the total number of ICT R&D centres owned worldwide by companies located in the observed region	It measures the total number of scientific publications , in the Computer Science area produced by organisations located in the observed region
Unit of measurement	Region's share in the total number of FP7 participations to a region's share in the EU population	Region's share in the total EU FP7 funding to SMEs to a region's share in the EU population	Region's share in the total number of FP7 SMEs participations to a region's share in the EU population	Region's share in the total number of R&D centres located in the EU to a region's share in the EU population	Region's share in the total number of R&D centres owned by EU firms to a region's share in the EU population	Region's share in the total number of publications in Computer Science to a region's share in the EU population
Definition of ICT dimension	ICT areas of the FP7 programme (see Section 5.2)			Based on HIS iSuppli classification of the major "semiconductors influencers" (see Section 5.4)		Computer Science as defined by Web of Science® classification of Research Areas
Unit of observation	NUTS 3					
Source	FP7 database by EC DG Connect (see Section 5.2)			ICT R&D centre location: Design Activity Tool by IHS iSuppli (see Section 5.4)		Bibliometrics: Web of Science by Thomson Reuters (see Section 5.3)
Reference year(s) considered	2007-2011			2012		2000-2012

Data on the agglomeration of ICT R&D is extracted from information available about:

- The performance of universities and computer science faculties across the world, as reported by the QS University Ranking. For a detailed description of the data source, see Section 5.1.
- Information about the funding and organizations participating in the European FP7 programme in the period 2007-2012. For a detailed description of the data source, see Section 5.2.
- The location and ownership of over 2,800 ICT R&D centres belonging to more than 170 multinational ICT companies across the world in 2012. For a detailed description of the data source, see Section 5.4.
- The scientific output, measured in terms of the number of publications in the computer science research area, of the research institutions in Europe for the period 2000-2012 from the Web of Science by Thomson Reuters. For a detailed description of the data source, see Section 5.3.

- Company-level information on R&D expenditures in the ICT sector for the period 2005-2011 in Europe stemming from the ORBIS database by Bureau Van Dijk. For a detailed description of the data source, see Section 5.7.

3.1.2 Internationalisation of ICT R&D (IntRD)

The 2 indicators characterising the internationalisation of ICT R&D activity are listed and described in Table 5. They are presented together with a first indication of the data sources used and their time coverage.

These two measures of the internationalisation of ICT R&D are defined as follows:

- Outward R&D internationalisation: Number of ICT R&D centres located abroad that are owned by companies' headquarters located in the observed region.
- Inward R&D internationalisation: Number of ICT R&D centres located in a region that are owned by foreign companies.

To address the issue of internationalisation of ICT-related R&D activity in NUTS level 3 spatial units across the EU, a distinction between in- and outward internationalization of R&D activities is made. Another way of addressing the issue of ICT R&D internationalisation would be to look at the FP7 data. However, due to its focus, this type of data would not allow us to take into account the global dimension of ICT R&D activity. Thus, the information contained in FP data is used to construct other indicators, e.g. R&D agglomeration or ICT R&D networking.

Table 5: ICT R&D Internationalisation indicators (IntRD)

Indicator ID	IntRD 1	IntRD 2
Name of indicator	**Outward ICT R&D internationalisation**	**Inward ICT R&D internationalisation**
What does it measure?	It measures the number of ICT R&D centres located abroad (outside the country) that are owned by companies' headquarters located in a region	It measures the number of ICT R&D centres located in a region that are owned by foreign companies
Unit of measurement	Region's share in the total number of R&D centres located abroad that are owned by companies' headquarters located in the EU to a region's share in the EU population	Region's share in the total number of R&D centres owned by foreign companies in the EU to a region's share in the EU population
Definition of ICT dimension	Based on HIS iSuppli classification of the major "semiconductors influencers" (see Section 5.4)	
Unit of observation	NUTS 3	
Source	ICT R&D centre location: Design Activity Tool by IHS iSuppli (see Section 5.4)	
Reference year(s) considered	2012	

Data on the internationalisation of ICT R&D is extracted from information available about the location and ownership of over 2,800 ICT R&D centres belonging to more than 170 multinational ICT companies across the world and for the period 2012. For a detailed description of the data source, see Section 5.4.

3.1.3 Networking in ICT R&D (NetRD)

Networking measures addressing the ICT R&D activity rely on the network analysis of the locations of FP7 programme participants. Based on the number of connections between regions and a subsequent analysis of these connections, a set of network indicators have been constructed: Degree, Closeness centrality, Betweenness centrality and Eigenvector centrality. For a detailed description of the methodology of network analysis and indicators, see 6.1.

The 4 indicators characterising the networking of ICT R&D activity are listed and described in Table 6. They are presented together with a first indication of the data sources used and their time coverage.

Table 6: ICT R&D Networking indicators (NetRD)

Indicator ID	NetRD 1	NetRD 2	NetRD 3	NetRD 4
Name of indicator	**Degree in ICT R&D network**	**Closeness centrality in ICT R&D network**	**Betweenness centrality in ICT R&D network**	**Eigenvector centrality in ICT R&D network**
What does it measure?	It measures the total number of connections a region maintains with other regions through organizations participating in common FP7 projects	It measures the average distance that each node is from all other nodes in the network	It measures the number of shortest paths in a network that traverse through that node	It measures the importance of a node in a network, based on the importance of its direct neighbours
Unit of measurement	Rank between 0 and 1.			
Definition of ICT dimension	ICT areas of the FP7 programme (see Section 5.2)			
Unit of observation	NUTS 3			
Source	FP7 database by EC DG Connect (see Section 5.2)			
Reference year(s) considered	2007-2011			

Data on the networking of ICT R&D is extracted from information available about the funding and organizations participating in the European FP7 programme in the period 2007-2011. For a detailed description of the data source, see Section 5.2.

3.2 ICT innovation activities indicators

3.2.1 Agglomeration of ICT innovation (AgIn)

The indicators characterising the agglomeration of ICT innovation activities are listed and described in Table 7. It offers a first indication of the data sources used and their time coverage.

To the extent allowed by the availability of indicators and data, the proposed indicators capture the input (investment in intangibles, and venture capital investments) and outputs (patenting activity) of innovation activities. With venture capital data, we aim to capture indirectly the dynamics of emerging new innovative companies: at the time of publication of this report, there was no serious European-wide collection of data on these dynamics. Similarly, patent counting and analysis has become one of the main acknowledged sources of information on innovation output across the world, particularly since the creation and divulgation of the EPO's PATSTAT database. The benefits of a patent-based approach are considered to largely compensate for the obvious limitations (Bergek & Bruzelius, 2010). For a detailed description of the methodology of patents fractional counting, see Annex 6.2.

Table 7 lists and describes the ICT Innovation Agglomeration indicators.[11]

[11] In order to account for the differences of size of the regions, a normalization with respect to the number of inhabitants is made. A modified version of the Balassa index is used, as defined in Section 2.3. Such index aims to measure the relative agglomeration level of given innovation activities in each region. The indicators for which this measures is used include AgIn 1, AgIn 2 and AgIn 3

Table 7: ICT Innovation Agglomeration indicators (AgIn)

Indicator ID	AgIn 1	AgIn 2	AgIn 3
Name of indicator	**Investment in intangibles by ICT firms**	**Venture Capital financing to ICT firms**	**ICT patents**
What does it measure?	Measures the average annual amount spent on intangibles in the ICT sector	Measures the amount of venture capital invested in the ICT sector	It measures the amount of ICT patent applications with inventors residing in the region
Unit of measurement	Region's share in the total investments in intangibles by ICT firms in the EU to a region's share in the EU population	Region's share in the total VC funding in to ICT firms in the EU to a region's share in the EU population	Region's share in the total number of ICT patents in the EU to a region's share in the EU population
Definition of ICT dimension	Based on NACE Rev. 2 (see Table 1)	Based on the Dow Jones classification of industry segments (see Section 5.8)	Based on the OECD definition of ICT patents following IPC taxonomy (OECD, 2008b)
Unit of observation	NUTS 3		
Source	Company-level information: ORBIS by Bureau Van Dijk (see Section 5.7)	Venture capital: VentureSource by Dow Jones (Section 5.8)	Patent data: REGPAT by OECD (Section 5.6)
Reference year(s) considered	2005-2012	2000-2012	2000-2009

Data on the agglomeration of ICT innovation is extracted from information available about:

- Company-level information on investments in intangibles by over 1,200 ICT firms located Europe wide in the period between 2005 and 2012 provided by ORBIS by Bureau Van Dijk. For a detailed description of the data source, see Section 5.7.
- Over 26,000 venture capital deals executed in Europe in the ICT sector between 2000 and 2012, data collected by Dow Jones. For a detailed description of the data source, see Section 5.8.
- Patenting activities of for over 5,000 regions in the period between 2000 and 2009. For a detailed description of the data source, REGPAT by OECD see Section 5.6.

3.2.2 Internationalisation of ICT innovation (IntIn)

The indicator characterising the internationalisation of ICT innovation activities is described in Table 8. This table offers a first indication of the data sources used and their time coverage.

To address the issue of internationalisation of ICT innovation activities, patent-based indicators are used. Despite a number of limitations (Bergek & Bruzelius, 2010), patent-based indicators have a long-standing tradition in serving to analyse the internationalisation patterns in the field of innovation studies (Bas & Sierra, 2002; Patel & Pavitt, 1991; Patel & Vega, 1999; Picci, 2010).

A detailed description of the methodology of constructing patent-based measures of internationalization can be found in Section 6.2. For a full description of the methodology of network analysis and indicators applied, see Section 6.1.

Table 8: ICT Innovation Internationalisation indicators (IntIn)

Indicator ID	**IntIn 1**
Name of indicator	**International co-inventions**
What does it measure?	It measures the number of international ICT patents, i.e. patents with at least two inventors residing in different countries, and attributes to the observed region the (fractional count) of those patents for which at least one inventor is residing in the region.
Unit of measurement	Region's share in the total number of international ICT patents in the EU to a region's share in the EU population
Definition of ICT dimension	Based on the OECD definition of ICT following IPC taxonomy (OECD, 2008b).
Unit of observation	NUTS 3
Source	Patent data: REGPAT by OECD (Section 5.6)
Reference year(s) considered	2000-2009

Data on the internationalization of ICT Innovation is extracted from the information available about patenting activities for over 5,000 regions for the period between 2000 and 2009. For a detailed description of the data source, REGPAT by OECD, see Section 5.6.

3.2.3 Networking in ICT innovation (NetIn)

The 4 indicators characterising the networking of ICT Innovation activity are listed and described in Table 9. They are presented together with a first indication of the data sources used and their time coverage.

Networking measures addressing ICT R&D activity rely on network analysis of the locations of co-inventors, who are based in different locations and jointly develop ICT inventions for which a patent application has been introduced.[12]

The relationship between different locations can be described as the total sum of co-inventions developed by inventors residing in different regions. According to (Guellec & Van Pottelsberghe de la Potterie, 2001), the total number of patents co-invented by residents of region *i* in collaboration with researchers in other regions is

$$CoInn_i = \sum_{j \neq i} CoInn_{ij} \,. \tag{2}$$

Based on the number of connections between regions and a subsequent analysis of these connections, a set of network indicators have been constructed: degree, closeness centrality, betweenness centrality and eigenvector centrality. For a detailed description of the methodology of network analysis and indicators, see Section 6.1.

The resulting ICT innovation networking indicators are listed in Table 9.

[12] As already said earlier, despite a number of limitations (Bergek and Bruzelius 2010), a patent-based methodology was proposed by Breschi, Cassi and Malerba (2007) and used since by De Prato and Nepelski (2012).

Table 9: ICT Innovation Networking indicators (NetIn)

Indicator ID	NetIn 1	NetIn 2	NetIn 3	NetIn 4
Name of indicator	**Degree in ICT innovation network**	**Closeness centrality in ICT innovation network**	**Betweenness centrality in ICT innovation network**	**Eigenvector centrality in ICT innovation network**
What does it measure?	It measures the total number of connections a region maintains with other regions through joint inventions	It measures the average distance that each node is from all other nodes in the network	It measures the number of shortest paths in a network that traverse through that node	It measures the importance of a node in a network, based on the importance of its direct neighbours
Unit of measurement	Rank between 0 and 1	Rank between 0 and 1	Rank between 0 and 1	Rank between 0 and 1
Definition of ICT dimension	Based on the OECD definition of ICT following IPC taxonomy (OECD, 2008b).			
Unit of observation	NUTS 3 for EU and TL3 for the remaining OECD countries			
Source	Patent data: REGPAT by OECD (Section 5.6).			
Reference year(s) considered	2000-2009			

Data on the ICT Innovation networking is extracted from the information available about global patenting activities for over 5,000 regions for the period between 2000 and 2009. For a detailed description of the data source, REGPAT by OECD see Section 5.6.

3.3 ICT business activities indicators

3.3.1 Agglomeration of business activities (AgBuss)

The indicators characterising the agglomeration of ICT business activities are listed and described in table 10. It offers a first indication of the data sources used and their time coverage.[13]

To the extent allowed by the availability of indicators and data, a mix of measures capturing business activities is proposed that, in addition, acknowledges the importance given to the business activities deployed by ICT multinationals and ICT firms in general, as well as to their turnover and employment size and growth.

[13] In order to account for the differences of size of the regions, a normalization with respect to the number of inhabitants is made. A modified version of the Balassa index is used, as defined in Section 2.3. Such index aims to measure the relative agglomeration level of given innovation activities in each region. The indicators for which this measure is used include from AgBuss 1 to AgIn 5, together with AgBuss 7 and AgBuss 9.

Table 10: ICT Business Agglomeration indicators (AgBuss)

Indicator ID	AgBuss 1	AgBuss 2	AgBuss 3	AgBuss 4	AgBuss 5
Name of indicator	**Location of ICT Scoreboard Headquarters**	**Ownership of ICT Scoreboard affiliates**	**Location of ICT Scoreboard affiliates**	**Location of ICT firms**	**ICT employment**
What does it measure?	It measures the number of ICT Scoreboard Headquarters located in the observed region	It measures the number of ICT Scoreboard affiliates owned worldwide by ICT Scoreboard Headquarters located in the observed region	It measures the total number of ICT Scoreboard affiliates located in the observed region	It measures the number of ICT firms located in the observed region	It measures the total employment in ICT firms in the observed region
Unit of measurement	Region's share in the total number of ICT Scoreboard Headquarters located in the EU to a region's share in the EU population	Region's share in the total number of ICT Scoreboard affiliates owned by EU ICT Scoreboard Headquarters to a region's share in the EU population	Region's share in the total number of ICT Scoreboard affiliates located in the EU to a region's share in the EU population	Region's share in the total number of ICT firms located in the EU to a region's share in the EU population	Region's share in the total employment by ICT firms located in the EU to a region's share in the EU population
Definition of ICT dimension	Based on NACE Rev. 2 (see Table 1)				
Unit of observation	NUTS 3				
Source	Company-level information: ORBIS by Bureau Van Dijk (see Section 5.7				
Reference year(s) considered	2008	2008	2008	2008	2005-2011

(continued): ICT Business Agglomeration indicators (AgBuss)

Indicator ID	AgBuss 6	AgBuss 7	AgBuss 8	AgBuss 9
Name of indicator	**Growth in ICT employment**	**Turnover by ICT firms**	**Growth in turnover by ICT firms**	**New business investments in the ICT sector**
What does it measure?	It measures employment growth in ICT firms in the observed region	It measures the average annual turnover by ICT firms in the observed region	It measures turnover growth in ICT firms in the observed region	It measures the number of new investments in the ICT sector in the observed region
Unit of measurement	Growth rate in %	Region's share in the total turnover by ICT firms located in the EU to a region's share in the EU population	Growth rate in %	Region's share in the total number of new investments in the ICT sector to a region's share in the EU population
Definition of ICT dimension	Based on NACE Rev. 2 (see Table 1)			
Unit of observation	NUTS 3			
Source	Company-level information: ORBIS by Bureau Van Dijk see Section 5.7)			European Investment Monitor by Ernst & Young (Section 5.5)
Reference year(s) considered	2005-2011	2005-2011	2005-2011	2000-2011

Data on the agglomeration of ICT business activities is extracted from information available about:

- Company level information on investments in intangibles by over 1,200 ICT firms located Europe wide for the period between 2005 and 2012 provided by ORBIS by Bureau Van Dijk. For a detailed description of the data source, see Section 5.7.
- Over 40,000 investment deals executed in Europe since 1997 to 2011, data collected by Ernst& Young. For a detailed description of the data source, see Section 5.5.

3.3.2 Internationalisation of ICT business activities (IntBuss)

The 2 indicators characterising the internationalisation of ICT business activities are listed and described in Table 11, which offers a first indication of the data sources used and their time coverage.

The measurement of the internationalization of business activity is proxied in EIPE by the information on the location of business affiliates owned by companies belonging to the (ICT) Industrial Scoreboard and the location of their respective Headquarters. To address the issue of internationalisation of ICT business activity in NUTS3 level spatial units across the EU, a distinction between in- and outward internationalization of business activities is made. These two measures of the internationalisation of ICT R&D are defined as follows:

- Outward business internationalisation; ownership of business affiliates located abroad,
- Inward business internationalisation: hosting of foreign-owned business affiliates.

Table 11: ICT Business Internationalisation indicators (IntBuss)

Indicator ID	**IntBuss 1**	**IntBuss 2**
Name of indicator	**Outward ICT business internationalisation**	**Inward ICT business internationalisation**
What does it measure?	It measures the number of affiliates located abroad (outside the country) that are owned by ICT Scoreboard Headquarters located in a region	It measures the number of affiliates located in a region that are owned by ICT Scoreboard Headquarters located abroad
Unit of measurement	Region's share in the total number of affiliates located abroad that are owned by European ICT Scoreboard Headquarters to a region's share in the EU population	Region's share in the total number of affiliates owned by foreign ICT Scoreboard Headquarters in the EU to a region's share in the EU population
Definition of ICT dimension	Based on NACE Rev. 2 (see Table 1)	
Unit of observation	NUTS 3	
Source	Company-level information: ORBIS by Bureau Van Dijk (see Section 5.7)	
Reference year(s) considered	2008	

Data on the internationalisation of ICT business activity is extracted from company level information provided by ORBIS by Bureau Van Dijk. For a detailed description of the data source, see Section 5.7.

3.3.3 Networking in ICT business activities (NetBuss)

The 4 indicators characterising the networking of ICT business activity are listed and described in Table 12. They are presented together with a first indication of the data sources used and their time coverage.

The measurement of the internationalization of business activity is proxied in EIPE by the information on the location of business affiliates owned by companies belonging to the ICT Scoreboard and the location of their respective headquarters.

In order to address the issue of networking in the context of business activity, the full network of international affiliates is re-created. A way of constructing a network of foreign affiliates is through the ownership and location relationship. A line between each pair of regions is drawn whenever a firm from one region owns an affiliate in another region, or vice versa. This way it illustrates the destination of expansion of multinational enterprises (MNEs), by tracking the existence of business relationships between regions. By doing this for all the regions owning and hosting subsidiaries of the MNEs, EIPE has created a unique map of ownership and location of business affiliates.[14]

Based on the number of connections between regions and a subsequent analysis of these connections, a set of network indicators have been constructed: degree, closeness centrality, betweenness centrality and eigenvector centrality. For a detailed description of the methodology of network analysis and indicators, see Section 6.1.

Table 12: ICT Business Networking indicators (NetBuss)

Indicator ID	**Net Bus 1**	**Net Bus 2**	**Net Bus 3**	**Net Bus 4**	**Net Bus 5**
Name of indicator	**In-degree in ICT business network**	**Out-degree in ICT business network**	**Closeness centrality in ICT business network**	**Betweenness centrality in ICT business network**	**Eigenvector centrality in ICT business network**
What does it measure?	It measures the total number of connections a region maintains with other regions whenever an ICT Scoreboard Headquarters located in that region owns an affiliate located in other regions	It measures the total number of connections a region maintains with other regions by hosting affiliates owned by ICT Scoreboard Headquarters located in other regions	It measures the average distance that each node is from all other nodes in the network	It measures the number of shortest paths in a network that traverse through that node	It measures the importance of a node in a network, based on the importance of its direct neighbours
Unit of measurement	Rank between 0 and 1				
Definition of ICT dimension	Based on NACE Rev. 2 (see Table 1)				
Unit of observation	NUTS 3 for EU and TL3 for the remaining OECD countries				
Source	Company-level information: ORBIS by Bureau Van Dijk (see Section 5.7)				
Reference year(s) considered	2008				

Data on the networking of ICT business activity is extracted from company-level information provided by ORBIS by Bureau Van Dijk. For a detailed description of the data source, see Section 5.7.

14 We focus our attention on bilateral relationships between regions and do not take into account loops, i.e. when a company's new investment and headquarter is located in the same region.

4 Composite indicators

The selected 42 indicators, their measurement and the resulting multiple rankings of 1,303 regions represent an abundance and diversity of information that seems impossible to analyse at first sight. In order to provide synthesised comparable results for further analysis and interpretation, the information contained in individual indicators has been aggregated, constructing a final composite EIPE indicator.

4.1 Normalization and rescaling of data

Most indicators are incommensurate with others, and have different measurement units. For example, the number of patent application is expressed per capita, while the share of ICT R&D centres owned by companies from a region and located abroad is expressed as a percentage of the total number of R&D centres owned by companies from a region.

To deal with this problem, indicators must be made comparable by bringing them to the same measurement scale, by transforming them into pure, dimensionless, numbers (OECD-JRC, 2008). This is the normalization process.

Normalization process

In order to normalise the data used in this study, a standardization method, i.e. z-scores, is used. This method is the most commonly used because it converts all indicators to a common scale with an average of 0 and a standard deviation of 1 (EC-JRC, 2005). The average of zero means that it avoids introducing aggregation distortions stemming from differences in indicators averages. The scaling factor is the standard deviation of the indicator across the units of observation, i.e. in the context of the current study of 1,303 NUTS 3 regions.[15]

In a more formal way, the normalized score of a raw score *x* is

$$z = \frac{x - \mu}{\delta}. \qquad (3)$$

where μ is the mean of observations across the regions and σ is the standard deviation across the regions. The quantity *z* represents the distance between the raw score and the mean population in units of the standard deviation.

The advantage of z-scores over other normalisation methods is that an indicator with extreme values will have intrinsically a greater effect on the composite indicator. This behaviour is desirable in the current study, as there is an intention to reward exceptional performance, i.e. above average results on few indicators is considered of higher value than average performance on many indicators.

Rescaling process

In the next steps, the normalized scores are further rescaled in order to avoid negative scores and to ensure the incorporation of the indicators variability in the results. This is done through the *minmax* rescaling procedure, whose formula is:

$$Nx_{rj} = \frac{x_{rj} - x_{j,\min}}{x_{j,\max} - x_{j,\min}} \times 100. \qquad (4)$$

where Nx_{rj} is the normalised and rescaled value of indicator *j* in the territorial unit *r*, x_{rj} is the normalised raw value of indicator *j* in the territorial unit *r*, $x_{j,\min}$ and $x_{j,\max}$ are the minimum and maximum values of indicator *j*.

[15] The intermediate results, i.e. the production of rankings for individual indicators, is done only for regions with indicator value greater than 0.

This method has found its way into a number of policy-oriented projects. For example, z-scores are used for the two composite indicators of the knowledge-based economy, published by the European Commission on Key Figures 2003-2004, for the environmental sustainability index developed at Yale University, and for the internal market index 2002 (EC-JRC, 2005).

4.2 European ICT Poles of Excellence Composite Indicator (EIPE CI)

An EIPE Composite Indicator (CI) is formed by compiling individual indicators into a single index, on the basis of an underlying model of the multi-dimensional concept that was introduced by in Chapter 2 of this report.

An important issue related to the construction of composite indicators is the one of weighting. Unfortunately, no agreed methodology exists to weight individual indicators (EC-JRC, 2005). In particular the context of the current study does not make the choice of a weighting scheme easy, as there is no theoretical framework that would say which indicator should be more influential than others. Considering this, equal weighting is given to the indicators to construct composite indicators.

The EIPE CI is composed of all indicators. Its construction is done in two steps. In a first step, composite sub-indicators are created, one for each of the activities: R&D, Innovation and Business. Three intermediate sub-indicators are organized along the three activities defined in Chapter 2, i.e.:

- **R&D sub-indicator** (R&D CI) comprises all relevant indicators included in Section 3.1 normalized and equally weighted.
- **Innovation sub-indicator** (Innovation CI) comprises all relevant indicators included in Section 3.2 normalized and equally weighted.
- **Business sub-indicator** (Business CI) comprises all relevant indicators included in Section 3.3 normalized and equally weighted.

For the sake of using the same scale, the values of the three sub-indicators are standardized with the MiniMax procedure, in order to present them on a scale from 0 to 100.

In the second step, all information is synthesised in one final EIPE CI by aggregating the values of the three earlier sub-indicators in this final one. Thus, sub-indicators values are equally weighted, i.e. each with 33% weight. As above, in order to present EIPE CI on a scale from 0 to 100, the values are standardized with the MiniMax procedure.

4.3 Sensitivity analysis

An important issue related to the construction of composite indicators is weighting. Unfortunately, no agreed methodology exists to weight individual indicators (EC-JRC, 2005). In particular the context of the current study does not make the choice of a weighting scheme easy, as there is no theoretical framework that could say which indicator is more influential than the others. Considering this, we have used equal weighting in the construction of the composite indicators.

The most debated problem in building composite indicators is the difficulty in assessing properly the relative importance of the sub-indicators (Saisana, Saltelli, & Tarantola, 2005). Experience shows that disputes over the appropriate method of establishing weights cannot be easily resolved. The two most commonly encountered difficulties when proposing weights to combine indicators into a single measure are that many published weighting schemes are either based on too complex multivariate methods or they are so simple that they have little meaning. For these reasons, it is necessary to check the sensitivity of the model proposed in this exercise to uncertainties that arise mainly from the choice of indicators and their weighting.

To test the overall robustness of the index and the effects of such variations on the value of the index, a sensitivity analysis is applied. Sensitivity analysis is the study of how the uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input (Saltelli, Tarantola, & Campolongo, 2000).

The weightage allocated to each sub-indicator is varied by between the three sub-indices in the following way:

- R&D CI: 20%, Innovation CI: 40% and Business CI: 40%,
- R&D CI: 40%, Innovation CI: 20% and Business CI: 40%,
- R&D CI: 40%, Innovation CI: 40% and Business CI: 20%.

The results of the subsequent rankings are compared to see whether the changes in weighting schemes affect the final ranking. Such test has been applied to the EIPE composite indicator and its results showed not to affect the final ranking in any significant way.

5 Data Sources

The following eight databases have been the primary data sources used to elaborate the indicators and measurements of EIPE:

1. QS World University Rankings by QS,
2. FP7 database by EC DG Connect,
3. Bibliometrics: Web of Science by Thomson Reuters,
4. ICT R&D centres locations: Design Activity Tool by IHS iSuppli,
5. European Investment Monitor by Ernst & Young,
6. Patent data: REGPAT by OECD,
7. Company level information: ORBIS by Bureau Van Dijk,
8. Venture Capital: Venture Source by Dow Jones.

In the following sections, each of the data source is described.

5.1 QS World University Rankings by QS

The rankings of Universities and Computer Science and Electronic Faculties originate from the QS World University Rankings®. It was formed in 2008 to meet the increasing public interest in comparative data on universities and organisations, and the growing demand for institutions to develop deeper insight into their competitive environment.[16]

The QS World University Rankings® currently considers over 2,000 universities in the world and evaluates over 700 of them, ranking the top 400. This list is used to build an indicator of the location of a ranked university in a region within the current project.

In addition, due to the fact the QS ranking includes 52 subject disciplines, one of which is Computer Science, additional faculty-level information is extracted for the purpose of the EIPE study.

To construct measures of faculty performance, the EIPE study used QS proprietary datasets to investigate its subject area at three levels, namely academic and employer reputation surveys and the Scopus data for the Citations per Faculty indicator. In detail, each of the faculty ranking pieces can be described in the following way:

- The **Academic reputation** survey is the centrepiece of the QS World University Rankings® since their inception in 2004. In 2010, it drew upon over 15,000 respondents to compile the results. In the survey, respondents are asked to identify the countries, regions and faculty areas that they have most familiarity with and up to two narrower subject disciplines in which they consider themselves expert. For each of the faculty areas they identify, respondents are asked to list up to ten domestic and thirty international institutions that they consider excellent for research in the given area. They are not able to select their own institution. The analysis places more emphasis on an international reputation than a domestic one – domestic responses are individually weighted at half the influence of international responses. This is a global exercise and recognizes institutions that have an international influence in these disciplines. Weightings are also applied to balance the representation by region.
- The **Employer reputation** survey considers the students' employability as a key factor in the evaluation of international universities and in 2010 drew on over 5,000 respondents to compile the results for the overall rankings. The employer survey works on a similar basis to the academic one only without the channelling for different faculty areas. Employers are asked to identify up to ten domestic and thirty international institutions they consider

[16] More information under: http://www.topuniversities.com (last accessed 01.02.2012)

excellent for the recruitment of graduates. They are also asked to identify from which disciplines they prefer to recruit. From examining where these two questions intersect, a measure of excellence in the given discipline is inferred. Employers seeking graduates from any discipline are weighted at 0.1 and those from a parent category (i.e. Social Sciences) are weighted at 0.25 relative to the weight of a direct response for the subject area. This analysis also places more emphasis on an international reputation than a domestic one, with domestic responses carrying half the weighting of international responses.

- **Citations per faculty** takes into account the size of an institution, and also observes its penetration into the global research landscape. The data for citations originate from Scopus by Elsevier E.V.[17] Papers in Scopus are tagged with an ASJC (All Science Journal Classification) code which identifies the principal foci of the journal in which they were published. When aggregated, these totals per faculty and their associated citations provide an indicator of volume and quality of output in the given discipline.

The main reason why this data source was selected for EIPE is that, in addition to the university ranking, it also offers the rankings described above by teaching subject, including Computer Science. This information allows us to observe the location of research and education in ICT activities at world- level.

This data source, though carefully selected from a range of data sources pursuing similar purposes, shows some limitations. The main constraint is that it offers only a limited number of universities, which does not allow us to cover the entire population of the European higher education institutions.

5.2 FP7 database by EC DG Connect

The Framework Programmes for Research and Technological Development, also called Framework Programmes or abbreviated to FP1, through to FP7, are funding programmes created by the European Union in order to support and encourage research in the European Research Area (ERA). FP7 spans through the period 2007 - 2013.

The analysis of the Framework Programme 7 programmes and participants is based on the database provided by DG Connect in November 2011 (and hence covers the period 2007-11). It is not available publically. In the current report, information on FP7 is used and concerns only the Information and Communication Technology (ICT) areas. The list of instruments through which projects were financed includes: CSA-ERA-PLUS, CSA-CA, CP-SICA-INFSO, CP-FP-INFSO-FET, CSA-SA, CP-IP, NoE, CP-CSA, CP-IP-INFSO-FET, CP-FP-INFSO, CP-FP, CSA-SA-INFSO-FET and CSA-CA-INFSO-FET.

The main reasons why this data source was selected for EIPE is that it offers a proxy for public R&D expenditures in ICT and allows us to observe the location of the R&D activity in ICT.

This data source, though carefully selected, shows some limitations. The main constraint is that it offers only a limited snapshot of EU-level publicly-financed ICT R&D in Europe. In particular, it does not cover national and regional expenditures in ICT R&D.

5.3 Bibliometrics: Web of Science by Thomson Reuters

The Web of Science is an online academic citation index provided by Thomson Reuters. It is designed to provide access to multiple databases, cross-disciplinary research, and in-depth exploration of specialized subfields within an academic or scientific discipline. As a citation index, any cited paper will lead to any other literature (book, academic journal, proceedings, etc.) which currently cites this work, or has done so in the past. In addition, literature which shows the greatest impact in a field covered by the Web of Science, or more than one discipline, can be selectively obtained. For example, a paper's influence can be determined by linking to all the other papers that have cited it. In this way, current trends, patterns, and emerging fields of research can be assessed. The Web of Science has indexing coverage from 1900 to the present.

[17] More information at: http://www.scopus.com (last accessed 01.02.2012).

Regarding its scope, the Web of Science encompasses over 11,000 journals selected on the basis of impact evaluations. This selection includes open-access journals and over 12,000 conferences each year (2009), spanning multiple academic disciplines. Coverage includes the sciences, social sciences, arts, and humanities, and it is also cross disciplinary. For the purpose of the EIPE exercise, journals classified in the Computer Science research area are considered.

The main reason why this data source was selected for EIPE is that it offers a comprehensive overview of scientific output throughout the world divided into individual research areas, which permits the inclusion of EIPE-relevant fields such as Computer Science. This information allows us to observe the location of ICT R&D activity.

This data source, though carefully selected from a range of data sources pursuing similar purposes, has some limitations. The main constraint is that it offers only limited possibilities with respect to the extraction of information at the level of, for example, authors. Instead, only aggregation of information at the institutional level is possible.

5.4 ICT R&D centre location: Design Activity Tool by IHS iSuppli

The data used for the purpose of identification of ICT R&D centre locations originates from the 2011 IHS iSuppli database, a company-level dataset dedicated to observing the internationalization of R&D. It includes a list of R&D centres belonging to a number of high-tech companies together with their exact location, and additional information on the type of R&D activity performed in these centres.

The data on R&D locations is collected by IHS iSuppli, an industry consultancy,[18] to map R&D locations and activities of companies considered to be the major semiconductor influencers, i.e. the main users of semiconductors or, in other words the largest manufacturers of applied electronic and microelectronic products.

In order to check how representative the sample is, we compared it to the R&D Scoreboard, a list of the top 2,000 R&D investors in Europe and the rest of the world,[19] and also with the list of companies filing their patents at the USPTO. The results revealed that the firms contained in the dataset represent nearly 30% of the 2008 R&D budget of all companies included in the R&D Scoreboard and more than 30% of all patent applications filed to the USPTO in 2009. This way we are assured that the sample is representative for the population of large high-tech multinational firms.

The main reason why this data source was selected for EIPE is that it offers relatively detailed unique information on the location and ownership of ICT R&D centres worldwide. This information allows us to observe the location of ICT R&D activities.

This data source, though carefully selected from a range of data sources pursuing similar purposes, shows some limitations. For example, the characteristics of the dataset do not allow the building of time series. Also, the information available from this data source concentrates on the number of R&D centres, their ownership and location, as detailed information on employment or R&D expenditures in those centres is not available at this level of granularity.

5.5 European Investment Monitor by Ernst & Young

The European Investment Monitor (EIM) is a unique monitor of foreign investment in Europe by companies from all over the world, but excludes investments in their home countries. Since 1997, data has been collected from all European countries and is published on a quarterly basis. As of 2011, it included over 40,000 observations.

The EIM is recognized as a comprehensive industry standard, tracking investment projects across Europe. It is a business information tool used by both professionals involved in corporate location

[18] More information at: http://www.isuppli.com (last accessed 01.02.2012).

[19] More information at: http://iri.jrc.ec.europa.eu/research/scoreboard_2010.htm (last accessed 01.02.2012).

strategy and inward investment issues and academic researchers (De La Tour, Glachant, & Ménière, 2011). It is a benchmark for government and private sector organizations wishing to identify trends in jobs and industries, business and investment.

The EIM identifies the project-based foreign inward investment announcements that are new, expanding, or co-located in an international context.[20] When the consulting group discovers a new project, they track it in order to determine its exact location at the city level. Only investments where at least 10 jobs are created are considered.

The basic description of each investment project described by the EIM data includes the name of the firm, the parent company name, the name and the origin country of the parent company, the sector and both the country and the city of location. It also includes the function of each investment (unit of production and different service activities, such as headquarters, research and development centres, logistics, or sales and marketing offices).

The data collected by the EIM enables to:

- Review developments and movements in the inward investment marketplace, identify emerging sectors, industries and clusters,
- Benchmark regions and develop location strategies,
- Undertake in-depth, wide-ranging data analysis; for example: Which is Europe's most popular location for headquarters investments? What is the scale and nature of investment from South Korea? Or what is Germany's market share of pharmaceutical investment?

The main reason why this data source was selected for EIPE is that it offers relatively detailed unique information on new investments in Europe and, due to the sector information included in the description, it permits the retrieval of ICT-specific investments. This information gives us a proxy for the dynamics of business activity in ICT.

This data source, though carefully selected from a range of data sources pursuing similar purposes, has some limitations. For example, as the EIM relies on data collection from the media, the main advantage of this source of information, i.e. being up-to-data and the speed of the information provision, can also be a disadvantage. This is related to the fact that not all investments are reported by the media and, hence, they will not be available from this source to the EIM.

5.6 Patent data: REGPAT by OECD

The OECD REGPAT database stores patent data, based on patent applications to the EPO and PCT filings, linked to more than 5 500 regions using the inventors/applicants addresses. This information has been linked to NUTS3 regions according to the addresses of the applicants and inventors. The data have been regionalised at a very detailed level so that more than 2 000 regions are covered across OECD countries. The selection of ICT patents follows the definition by OECD (OECD, 2008b).

When compiling or analysing indicators with regionalised patents, it is necessary to have some characteristics of patents and some rules in mind, so as to make the best use of the information and not misinterpret the indicators. The data from the REGPAT database, are constructed along the following principles:

- *Inventor v. owner region*: Patent data can be regionalised on the basis of the address of either the inventor or the holder. The inventor's address usually indicates where the invention was made, while the owner's address indicates where the holder has its headquarters. These two concepts have obviously different economic interpretations,

[20] The EIM excludes mergers and acquisitions or joint ventures (unless these result in new facilities, new jobs created), licence agreements, retail and leisure facilities, hotels and real estate investments, utility facilities including telecommunications networks, airports, ports or other, fixed infrastructure investments, extraction activities (ores, minerals or fuels), portfolio investments (i.e. pensions, insurance and financial funds), factory / production replacement investments (e.g. a new machine replacing an old one, but not creating any new employment), not-for-profit organisations.

especially as many patents are filed by large companies with several establishments located in different regions and countries.

- *Fractional v. whole counting*: Patents usually have several inventors and can have several owners. When regionalising patents, a patent with, say, inventors in two regions can be either attributed wholly to the two regions, or shared (with a total of shares of 100%) between the two regions. As a significant proportion of patents have inventors from different regions it is important to specify what rule is used, and when one is better, to use it. For instance, when comparing the performance of regions it is recommended to use fractional accounting, which i) attributes to each region its actual contribution to the invention; ii) when summed over all regions gives a total of 100%. On the other hand, when compiling an indicator like *share of patents with co-inventors from another region*, the use of whole counting, both at the numerator and the denominator, is recommended.
- *Priority year*: This is the year of first filing for a patent. It is the closest to the actual date of invention, and should therefore be used as the reference date when compiling patent indicators which aim to reflect technological achievements. Other dates, e.g. follow-up application, publication or grant, are dependent on administrative procedures and can be one to ten years after the invention and thus misleading when interpreting the data.

The methodology developed to identify regions on the basis of the addresses of patent inventors consists of an iterative procedure that matches postal codes and/or town names, identified in the addresses, with regions using a set of lookup tables (such as a postal code - NUTS3 correspondence).

The main reason why this data source was selected for EIPE is that it offers unique information on patenting activity at regional level across a number of countries, which allows us to extract information on ICT innovation activity at NUTS 3 level. The EIPE study uses this information as a proxy for innovative activity in ICT.

This data source, though carefully selected, shows some limitations, which, if not taken into account, can affect the results of the EIPE project or their interpretation. For example, REGPAT relies on the EPO and PCT filings, which considerably limits the number of patents that are considered.

5.7 Company-level information: ORBIS by Bureau Van Dijk

Company-level information is taken from the ORBIS database by Bureau Van Dijk. It contains comprehensive information on companies worldwide.

In order to meet the requirements of the EIPE project, while constructing the indicators on the number of employees, turnover, intangible and R&D expenditures at the NUTS 3 level, the following criteria were applied:

- Geographic coverage: EU 27;
- The ICT industry was defined according to the NACE Rev 2 definition of the ICT sector (OECD, 2007);
- Company status: Active companies;
- Type of entities: Industrial companies
- In order to avoid double-counting, separate searches were run using a filter on consolidation code. First, companies with consolidated accounts only and then companies with unconsolidated accounts only were selected.
- Time coverage between 2005 and 2011, the last available date.

Besides providing the company-level information that was used to count the number of firms or the employment, ORBIS was also used to map the organizational structure of the main ICT R&D investors by including their affiliates. This allowed to observe the internationalisation of ICT business activities and to construct the ICT business network.

In addition, the information on the location of business affiliates owned by companies belonging to the ICT Scoreboard and the location of their respective Headquarters originates from the Orbis database. The analysis presented in this report is based on company data from the 2009 EU industrial R&D Scoreboard 3 (henceforth the Scoreboard) in which R&D investment data, and economic and financial data from the last four financial years are presented for the 1,000 largest EU and 1,000 largest non-EU R&D investors in 2008. The Scoreboard covers about 80% of all company R&D investments worldwide. From the Scoreboard, we have extracted the sub-set of ICT sector companies, which we refer to in this report as the ICT Scoreboard. This dataset serves for the following analysis that aims to benchmark R&D investments of EU ICT companies against those of non-EU companies.

The construction of the ICT Scoreboard dataset followed two steps, which were carried out at IPTS as part of a larger project. First, the ICT sample was selected from the European R&D Investment Scoreboard 2009 and then it was merged with the BvD Orbis database.

The R&D Scoreboard collects information on R&D investment, sales, operating surplus, employment and capital expenditure (to be interpreted as a flow, the increasing of tangible assets) for the top 1,000 European groups and top 1,000 non-European groups, ranked according to the amount of nominal R&D investment. The period covered includes the four years from 2005 to 2008.

The merge with the database Orbis was done in order to collect the information on the individual shareholders that have relevant participations in group headquarters. We used a standard rule of thumb of direct or indirect share above 20%. As a result, in our database, the individual observation is a group, for which we have the R&D Scoreboard information together with information on up to a potential maximum of five shareholders, with their legal entity and details of the amount of shares. The ownership threshold was set at the level of 50.1%.

The main reason why this data source was selected for EIPE is that it offers unique and standardized information on company-level information for the ICT sector that can be regionalised and presented at the NUTS 3 level. This information offers a proxy for the economic and, to some extent, the innovative activity of ICT companies.

This data source, though carefully selected from a range of data sources pursuing similar purposes, has some limitations. The most important limitation is the geographical coverage and the incompleteness of the data collected. In addition, there are significant problems concerning the extraction of detailed information, e.g. on a firm's ownership structure.

5.8 Venture capital: VentureSource by Dow Jones

Dow Jones VentureSource provides comprehensive data on venture capital-backed and private equity-backed companies – including their investors and executives – in every region, industry sector and stage of development throughout the world.

According to Kaplan et al. (2002), who provide a detailed overview of this database and compare it with Venture Economics (an alternative source of information), the VentureSource data are generally more reliable, more complete, and less biased.

This database contains information on venture capital transactions, the financed companies and the financing firms. The data are largely self-reported by venture capital firms, but the database conducted several plausibility checks.

The selection of ICT companies was based on Dow Jones classifications and includes companies belonging to the following industry segments: Communications & Networks, Electronics & Computers, Information Services, Semiconductors, Software and Other IT.

This data source was selected for EIPE because it offers unique and standardized information on venture capital deals with all the detailed information concerning the financed and financing entities. In addition, it allows us to select deals that concern the ICT sector. This information can be used as a proxy for the funding of innovative products, particularly those in the commercialisation phase, and companies.

This data source, though carefully selected from a range of data sources pursuing similar purposes, has some limitations. VentureSource relies on the voluntary information provision by Venture Capital funds and companies. Thus, despite being up-to-date, there is no guarantee that it covers the entire universe of venture capital.

6 Annex: Technical issues

6.1 Definition and characteristics of a network structure

Design

A straightforward way of representing the existing networks linking NUTS3 regions is by drawing a line connecting two different regions whenever two organizations from these regions show a relation (i.e. participate in the same project of the FP7 programme (Cassi, Corrocher, Malerba, & Vonortas, 2008), share a patent, are part of the same business group, etc.). Thus, knowing the location of each organisation, it is possible to build a directed network. To establish the network and its structure, it is necessary to identify the set of nodes, *V*, as the regions where the observed organisations are located, and the set of arcs, *A*, as the bilateral relationships that exist whenever an organization from one region shows the given relation with an organization from a different region.[21]

Indicators

According to the above defined methodology, based on the number of connections between regions and a subsequent analysis of these connections, a set of network indicators were constructed.

Analysis

A network consists of a graph whose elements include two sets: a set of nodes (vertices), that correspond to the selected unit of observation, and a set of lines or relationships, that represent relations between units. Relationships relevant in the context of the EIPE project include, for example, location-ownership of R&D centres between regions and countries (see Section 3.1.3) or collaboration between inventors located in different regions and countries (see Section 3.2.3). A line can be directed – an arc, or undirected – an edge. In a formal way, a network

$$N = (V,L,W,P) \qquad (5)$$

consists of a graph $G = (V, L)$, where V is the set of nodes, A is the set of arcs, if the lines are directed, and E is the set of edges, if the lines are not directed, and $L = E \cup A$ is the set of lines. Additional information on the lines is given by the line value function W and on nodes by the value function P.

Density of a network

Regarding the structural properties of a network, the **density of a network** is, among others, a key indicator providing information about the network structure. The density of a network is the number of edges that is expressed as a proportion of the maximum possible number of connections. It is formally defined as

$$\lambda = \frac{m}{m_{\max}} \qquad (6)$$

where $m_{\max}$ is the total number of lines in a complete network, i.e. a network where all the nodes are connected to each other, given the same number of nodes. Thus, in practical terms, network density is a measure of the level of network connectedness.

Node's degree

In order to obtain further information on the structure of a network it is worthwhile to analyse network centrality. Centrality is an important concept in studying networks (Freeman, 1978). In conceptual terms, centrality measures how central an individual is positioned in a network. The

[21] In the following, we focus our attention on bilateral relationships between regions and do not take into account loops, i.e. when a company's R&D centre and headquarter are located in the same region.

most obvious way of capturing degree centrality of V_i is counting the number of its neighbours, i.e. its degree. The way to compute degree centrality is to count the number of nodes connected to V_i, i.e.:

$$C_i^d = \frac{d}{V-1} \tag{7}$$

If there is no information on the direction of edges, i.e. an un-directed network, the measurement of a node's position can be measured by the total number of connections with the node. Then, **a node's degree** is defined as:

$$k_i \equiv \sum_{j \neq i} a_{ij} \tag{8}$$

If there is information on the direction of edges, i.e. directed network, the measurement of a node's position can be disaggregated to account for the incoming and outgoing connections to and from the node. Then, **the in-degree and out-degree** are defined as:

$$k_i^{in} \equiv \sum_{j \neq i} a_{ij} \tag{9}$$

$$k_i^{out} \equiv \sum_{j \neq i} a_{ji} \tag{10}$$

where a_{ij} represents the directed link from V_i to V_j and a_{ji} the reverse relationship.

Nodes' centralities in a network can have large or small variance. On the one hand, a network, where few actors have much higher centrality than other actors is said to be strongly centralised. A typical example is a star network. On the other hand, if unit centrality measures have small variance, the centralisation of a network is low. Thus, in order to assess the level of **centralisation of the network**, we use a network degree centralisation defined as

$$C^d = \frac{\sum_{i=1}^{n} | C_i^d - C_i^{d*} |}{(n-2)(n-1)}, \tag{11}$$

where C_i^{d*} is the highest value of centrality measure in the set of units of a network (Freeman, 1978). Network centralisation index can take any value between 0, if all units have equal centrality value (cycle graph), and 1, if one unit completely dominates all other units (star graph).

Closeness centrality of a node

Except for the degree centrality defined in (7), within graph theory and network analysis, there are a number of other measures of the centrality of a vertex within a graph that show the relative importance of a vertex within the graph (Koschützki et al., 2005). In this we use of two additional most commonly applied measures, i.e. closeness centrality and betweenness centrality.

The closeness centrality of a node *i* is the number of the remaining nodes divided by the sum of all distances between that node and all the remaining ones, i.e.:

$$C_i^c = \frac{n-1}{\sum_{j \neq i}^{n-1} \partial_{ij}}. \tag{12}$$

At the aggregate level, **centrality closeness of a network** is defined as:

$$C^c = \frac{\sum_{i=1}^{n} | C_i^c - C_i^{c*} |}{(n-2)(n-1)/(2n-3)}, \tag{13}$$

where C_i^{c*} is the highest value of closeness centrality measure in the set of units of a network (Freeman, 1978). The index takes values between 0 and 1, whereas the closeness centrality of a star network is 1. In practical terms, closeness centrality is a measure of the average distance that each node is from all other nodes in the network.

Betweenness centrality of a node

The betweenness centrality of a node is the proportion of all geodesics distances between pairs of other nodes that include this vertex. Formally, the betweenness centrality of V_i can be expressed as:

$$C_i^b = \sum_{j \neq k} \frac{\partial^i_{jk}}{\partial_{jk}}, \tag{14}$$

where ∂_{jk} is the total number of shortest paths joining any two nodes V_k and V_j, and ∂^i_{jk} is the number of those paths that not only connect V_k and V_j, but also pass through V_i. The betweenness centrality of each node is a number between 0 and 1. This property of a network reflects the amount of control that a node exerts over the interactions of other nodes in the network (Yoon, Blumer, & Lee, 2006). The measure of betweenness centrality rewards nodes that are part of communities, rather than nodes that lie inside a community. Betweenness centrality reflects the shortest path between two others. Therefore, it can be regarded as a measure of gatekeeping and is considered to be a measure of strategic advantage and information control.

Similarly, the **network betweenness centralization measure** can be defined as:

$$C^b = \frac{\sum_{i=1}^{n} | C_i^b - C_i^{b*} |}{(n-1)}, \tag{15}$$

where C_i^{b*} is the highest value of betweenness measure among all nodes. This measure compares the variance of betweenness centrality in a network and takes as a reference a star graph (C^b=1). In such a graph, the node in the middle holds the highest betweenness centrality, i.e. a strategic position and the graph is highly unequal or highly centralized.

In practical terms, betweenness centrality is a measure of the number of shortest paths in a network that traverse through a node.

Eigenvector centrality of a node

Further measure of a node's position in the network used in this study is **eigenvector centrality** and relates to the quality of a node's connections and is. It assigns relative scores to all nodes in the network based on the principle that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. Google's PageRank is a variant of the Eigenvector centrality measure (Spizzirri, 2011). In practical terms, eigenvector centrality is a measure of the importance of a node in a network, based on importance of its neighbours expressed by the quality of their connections.

6.2 Patent data and patent-based internationalisation measures

Assigning patents to countries

Assigning patents to countries or regions relies upon the concept of fractional counting of patents. A fully detailed explanation of the methodology for patent counting is described in a JRC-IPTS co-authored article on the worldwide count of priority patents(de Rassenfosse, Dernis, Guellec, Picci, & van Pottelsberghe de la Potterie, 2013).

To help make the discussion as easy to follow as possible, a simple fictitious example is used. Three countries *i*, United Stated (US), France (FR), and Germany (DE), are considered that in a given year produce a total of P=3 patents. Column I in Table 13 indicates the nationality of the inventors and applicants that contributed to these three inventions.

In order to assign patents to countries, two alternative criteria may be chosen: either according to the nationality of the applicant(s), or of the inventor(s). The former defines the "applicant criterion" and the latter the "inventor criterion". Whenever an application has more than one inventor or applicant, some of them coming from different countries, patent assignment is carried out by resorting to fractional counts. So, for example, patent n. 1 counts as ½ German and ½ American according to the applicant criterion, and ½ American, ¼ German and ¼ French according to the inventor criterion.

Table 13: Fractional counts of three fictitious patents

I	II			III			IV	
	$Inv_{US,p}$	$Inv_{DE,p}$	$Inv_{FR,p}$	$App_{US,p}$	$App_{DE,p}$	$App_{FR,p}$	$\sum_{i=1}^{N} Inv_{ip}$	$\sum_{i=1}^{N} App_{ip}$
P=1: Inv: DE, FR, US, US	0.5	0.25	0.25				1	
P=1: App: DE, US				0.5	0.5	0		1
P=2: Inv: DE, DE, FR, FR	0	0.5	0.5				1	
P=2: App: FR, US				0.5	0	0.5		1
P=3: Inv: FR, US	0.5	0	0.5				1	
P=3: App: US, US				1	0	0		1
$Inv_i = \sum_{p=1}^{P} Inv_{ip}$	**1**	**0.75**	**1.25**				**3**	
$App_i = \sum_{p=1}^{P} App_{ip}$				**2**	**0.5**	**0.5**		**3**

In the following, $Inv_{i,p}$ represents the fraction of patent *p* attributed to country *i* according to the inventor criterion, and $App_{i,p}$ the analogous measure according to the applicant criterion (Picci, 2010).[22] Column II and III of Table 13 report these measures for the three patents. For each patent application, the sum of all the country's contribution according to the inventor criterion has to be equal to 1: for each patent, $Inv_{US,p}$ + $Inv_{DE,p}$ + $Inv_{FR,p}$ = 1, where the first subscript indicates the country, and the second the patent. These sums are indicated in Column IV of Table 13.

The total fractional assignment of the three patents to each country is simply equal to the sum of the individual assignments:

(1) $$Inv_i = \sum_{p=1}^{P} Inv_{ip}$$

[22] When considering the fictitious example, instead of the subscript i we will use the mnemonic symbol of the relevant country. Also, for clarity we omit in all cases a time subscript, that should always be present.

and:

(1') $App_i = \sum_{p=1}^{P} App_{ip}$

They are reported in the last two rows of Table 13. For example, Germany produced a total of 0.75 patents according to the inventor criterion, and of 0.5 patents according to the applicant criterion.

Patent-based measures of innovation internationalisation

The analysis uses measures of internationalisation that are based on the presence of inventors residing in different regions of the world among the list of inventors. An international patent application is defined in the analysis presented here as a patent application that includes at least two inventors residing in different countries. Using this methodology, we use the concept of internationalisation of innovation measured by international co-invention. This concept is used to construct a relative measure of international collaboration between inventors.

In practical terms, the strength of the relation between inventors in country *i* and *j* is expressed as the product of the attribution of that patent to the two countries:

(2) $InvInv_{ijp} = Inv_{ip} \cdot Inv_{jp}$

This measure attributes a greater weight to collaborations where the two countries have more similar weights. So, for example, the collaboration between the US and France is equal to ½ · ¼ = $^1/_8$ in patent n. 1 (where there are 1 French and 2 American inventors) and to ½ · ½ in patent 3 (where the total number of inventors, 2, is equally divided between the US and France. In fact, if *i* is different from *j*, $0 \le InvInv_{ijp} \le 1/4$, where the upper bound is reached when the total number of inventors is equally divided between two countries, and the lower limit applies when a patent is national.

The aggregate strength of the relation between the inventors of two countries is defined as the sum of the above, over all patents:

3) $InvInv_{ij} = \sum_{p=1}^{P} Inv_{ijp}$

Below, the values for all the combinations of the three patents in Table 13 are reported, where for clarity, instead of the indexes *i* and *j*, the acronyms of the countries are employed.

$$InvInv_{US,US} = 0.5 \cdot 0.5 + 0 \cdot 0 + 0.5 \cdot 0.5 = 0.5$$

$$InvInv_{US,DE} = 0.25 \cdot 0.5 + 0 \cdot 0 + 0 \cdot 0 = 0.125$$

$$InvInv_{US,FR} = 0.25 \cdot 0.5 + 0 \cdot 0 + 0.5 \cdot 0.5 = 0.375$$

The top part of Table 14 shows the values of these interactions for all three cases. Note that $Inv_{ij} = Inv_{ji}$ (the order of the countries is irrelevant). Using (1), it is easy to see that:

(4) $\sum_{i=1}^{N} InvInv_{ij} = Inv_j$ and $\sum_{j=1}^{N} InvInv_{ij} = Inv_i$

For example, as predicted by (4):

$$InvInv_{US,US} + InvInv_{US,DE} + InvInv_{US,FR} = 0.5 + 0.125 + 0.375 = 1 = Inv_{US}$$

These sums are reported for all three countries in the last column and in the last rows of the top part of Table 14, and correspond to the values reported in Table 13. They show that the country patent portfolio, assigned according to the inventor criterion, may be expressed as a sum of pairwise measures of country inventive collaboration ($InvInv_{ij}$).

The measure of applicant internationalisation is constructed along the same lines, and the following formulae hold:

(2') $$App_{ijp} = App_{ip} \cdot App_{jp}$$

(3') $$App_{ij} = \sum_{p=1}^{P} App_{ijp}$$

(4') $$\sum_{j=1}^{N} App_{ij} = App_i \text{ and } \sum_{i=1}^{N} App_{ij} = App_j$$

All computations for this case are shown in the middle part of Table 14. Note that $App_{ij} = App_{ji}$ (again, the order of the countries is irrelevant). Equation (4') allows us to express a country patent portfolio, according to the applicant criterion, as a sum of interactions between applicants in different countries. The values reported in the last column and row of the middle part of Table 14 correspond to those of Table 13.

A measure of Inventor-Applicant internationalisation is constructed similarly. The strength of the collaboration between inventors in country *i* and applicants in country *j*, for a single patent *p*, is defined as:

(5) $$Invapp_{ijp} = Inv_{ijp} \cdot App_{ijp}$$

Summing over patents provides a measure of the strength of the overall collaboration between country *i* inventors and country *j* applicants:

(6) $$Invapp_{ij} = \sum_{p=1}^{P} Inv_{ijp} \cdot App_{ijp}$$

These measures aggregate to the patent attributed to a country either according to the inventor, or to the applicant criterion, depending on whether the summation is over *i*, or over *j*:

(7) $$\sum_{j=1}^{N} Invapp_{ij} = Inv_i$$

(7') $$\sum_{i=1}^{N} Invapp_{ij} = App_j$$

The bottom part of Table 14 indicates all computations for our fictitious example. Note that $InvApp_{ij}$ generally differs from $InvApp_{ji}$.

The quantities defined in (3), (3') and (6) are the three measures of internationalisation of innovative activities. In order to provide a first description of the degree of internationalisation, *relative* measures of internationalisation are used which are expressed as a share of the total number of patents. It is straightforward to construct relative measures of (3) and (3'):

(8) $$Inv_{ij|i} = Inv_{ij} / Inv_i$$

and

(8') $$App_{ij|i} = App_{ij} / App_i$$

where $\sum_{j=1}^{N} Inv_{ij|i} = 1$ and $\sum_{j=1}^{N} App_{ij|i} = 1$.

There are in fact two conditional measures of inventor-applicant internationalisation, depending on whether the normalization is carried out with respect to the inventors of country *i*, or to the applicants of country *j*:

(9) $Invapp_{ij|i} = Invapp_{ij} / Inv_i$

(9') $Invapp_{ij|j} = Invapp_{ij} / App_j$

where $\sum_{j=1}^{N} Invapp_{ij|i} = 1$ and $\sum_{i=1}^{N} Invapp_{ij|j} = 1$.

Table 14: Computation of measures of internationalisation of three fictitious patents

$InvInv_{ij} = \sum_{p=1}^{P} Inv_{ijp}$	*j* = US	*j* = DE	*j* = FR	$\sum_{j=1}^{N} InvInv_{ij} = Inv_i$
i = US	0.5	0.125	0.375	**1**
i = DE	0.125	0.3125	0.3125	**0.75**
i = FR	0.375	0.3125	0.5625	**1.25**
$\sum_{i=1}^{N} InvInv_{ij} = Inv_j$	**1**	**0.75**	**1.25**	

$AppApp_{ij} = \sum_{p=1}^{P} App_{ijp}$	*j* = US	*j* = DE	*j* = FR	$\sum_{j=1}^{N} AppApp_{ij} = App_i$
i = US	1.5	0.25	0.25	**2**
i = DE	0.25	0.25	0	**0.5**
i = FR	0.25	0	0.25	**0.5**
$\sum_{i=1}^{N} AppApp_{ij} = App_j$	**2**	**0.5**	**0.5**	

$InvApp_{ij} = \sum_{p=1}^{P} InvApp_{ijp}$	*j* = US	*j* = DE	*j* = FR	$\sum_{j=1}^{N} InvApp_{ij} = Inv_i$
i = US	0.75	0.25	0	**1**
i = DE	0.375	0.125	0.25	**0.75**
i = FR	0.875	0.125	0.25	**1.25**
$\sum_{i=1}^{N} InvApp_{ij} = App_j$	**2**	**0.5**	**0.5**	

References

ABS. (2012). *Business Indicators*: Australian Bureau of Statistics.

Barrios, S., Bertinelli, L., Strobl, E., & Teixeira, A.-C. (2005). The dynamics of agglomeration: evidence from Ireland and Portugal. *Journal of Urban Economics, 57*(1), 170-188.

Bas, C. L., & Sierra, C. (2002). 'Location versus home country advantages' in R&D activities: some further results on multinationals' locational strategies. *Research Policy, 31*(4), 589-609.

Bergek, A., & Bruzelius, M. (2010). Are patents with multiple inventors from different countries a good indicator of international R&D collaboration? The case of ABB. *Research Policy, 39*(10), 1321-1334.

Broekel, T., & Brenner, T. (2011). Regional factors and innovativeness: an empirical analysis of four German industries. *The Annals of Regional Science, 47*(1), 169-194.

Cantwell, J. A., & Janne, O. (1999). Technological globalisation and innovative centres: the role of corporate technological leadership and locational hierarchy. *Research Policy, 28*(2-3), 119-144.

Cassi, L., Corrocher, N., Malerba, F., & Vonortas, N. (2008). Research Networks As Infrastructure For Knowledge Diffusion In European Regions. *Economics of Innovation and New Technology, 17*(7-8), 663-676.

Chiripanhura, B. (2010). Measures of economic activity and their implications for societal well-being. *Economic and Labour Market Review, 4*(7), 56-65.

Ciccone, A., & Hall, R. E. (1996). Productivity and the Density of Economic Activity. *National Bureau of Economic Research Working Paper Series, No. 4313*.

De Benedictis, L., & Tajoli, L. (2011). The World Trade Network. *The World Economy, 34*(8), 1417-1454.

De La Tour, A., Glachant, M., & Ménière, Y. (2011). *Innovation and international technology transfer: The case of the Chinese photovoltaic industry*. Paris: MINES ParisTech.

De Prato, G., & Nepelski, D. (2012). Global technological collaboration network. Network analysis of international co-inventions. *Journal of Technology Transfer*, 1-18.

de Rassenfosse, G., Dernis, H., Guellec, D., Picci, L., & van Pottelsberghe de la Potterie, B. (2013). The worldwide count of priority patents: A new indicator of inventive activity. *Research Policy, 42*(3), 720-737.

Devereux, M. P., Griffith, R., & Simpson, H. (2004). The geographic distribution of production activity in the UK. *Regional Science and Urban Economics, 34*(5), 533-564.

Dorrenbacher, C. (2000). *Measuring corporate internationalisation: a review of measurement concepts and their use*. Berlin: Wissenschaftszentrum Berlin.

Doz, Y., Santos, J., & Williamson, P. (2001). *From Global to Metanational: How Companies Win in the Knowledge Economy*: Harvard Business Press.

Dumais, G., Ellison, G., & Glaeser, E. L. (2002). Geographic Concentration as a Dynamic Process. *The Review of Economics and Statistics, 84*(2), 193-204.

Dunning, J. (1994). Multinational enterprises and the globalization of innovatory capacity. *Research Policy, 23*(1), 67-88.

Dunning, J. (2002). *Regions, Globalization, and the Knowledge-Based Economy*. Oxford: Oxford Scholarship Online Monographs.

EC-JRC. (2005). *Tools for Composite Indicators Building*. Ispra: EC-JRC.

Ellison, G., & Glaeser, E. L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy, 105*(5), 889-927.
EUROSTAT. (2009). *European Regional and Urban Statistics. Reference Guide*. Luxemburg.
Frankish, J., Roberts, R., & Storey, D. (2009). *Measuring Business Activity in the UK*. London: Institute for Small Business and Entrepreneurship.
Fraumeni, B., & Okubo, S. (2005). R&D in the National Income and Product Accounts: A First Look at Its Effect on GDP. In (pp. 275-322): National Bureau of Economic Research, Inc.
Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215-239.
Fujita, M., & Thisse, J.-F. (1996). *Economics of Agglomeration*: C.E.P.R. Discussion Papers.
Fujita, M., & Thisse, J.-F. (2002). *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge: Cambridge University Press.
Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature, 28*(4), 1661-1707.
Guellec, D., & Van Pottelsberghe de la Potterie, B. (2001). The internationalisation of technology analysed with patent data. *Research Policy, 30*(8), 1253-1266.
Hagedoorn, J., & Cloodt, M. (2003). Measuring innovative performance: is there an advantage in using multiple indicators? *Research Policy, 32*(8), 1365-1379.
Hall, B. H., Mairesse, J., & Mohnen, P. (2009). Measuring the Returns to R&D. *National Bureau of Economic Research Working Paper Series, No. 15622*.
JRC-IPTS. (2007). *Mapping the ICT in EU Regions: Location, Employment, Factors of Attractiveness and Economic Impact*. Seville.
Kaplan, S. N., Strömberg, P., & Sensoy, B. A. (2002). *How Well do Venture Capital Databases Reflect Actual Investments?*
Kominers, S. (2013). Measuring agglomeration. Boston: Harvard University.
Koschatzky, K., & Lo, V. (2007). *Methodological framework for cluster analyses*: Fraunhofer Institute for Systems and Innovation Research (ISI).
Koschützki, D., Lehmann, K. A., Peeters, L., Richter, S., Tenfelde-Podehl, D., & Zlotowski, O. (2005). Centrality Indices. In U. Brandes & T. Erlebach (Eds.), *Network Analysis: Methodological Foundations* (pp. 16–61): Springer-Verlag.
Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy, 99*(3), 483-499.
Lipsey, R. G., & Carlaw, K. I. (2004). Total Factor Productivity and the Measurement of Technological Change. *The Canadian Journal of Economics / Revue canadienne d'Economique, 37*(4), 1118-1150.
Liu, X., & White, S. (2001). An exploration into regional variation in innovative activity in China. *International Journal of Technology Management, 21*, 114-129.
Malik, P. (2013). Information and communication technologies industry in emerging India. In G. De Prato, D. Nepelski & J.-P. Simon (Eds.), *Dancing with the tigers: Asia in the global ICT innovation network*: Chandos Publishing.
Meyer, K. E., Mudambi, R., & Narula, R. (2011). Multinational Enterprises and Local Contexts: The Opportunities and Challenges of Multiple Embeddedness. *Journal of Management Studies, 48*(2), 235-252.
Nepelski, D., & De Prato, G. (2012). Internationalisation of ICT R&D: a comparative analysis of Asia, EU, Japan, US and the RoW. *Asian Journal of Technology Innovation, 20*(2), 219-238.

Nepelski, D., & De Prato, G. (2013). *Defining European ICT Poles of Excellence. A Literature Review*. Seville: JRC-IPTS.
OECD-JRC. (2008). *Handbook on Constructing Composite Indicators. Methodology and user Guide*. Paris: OECD-JRC.
OECD. (2002). *Frascati Manual 2002. Proposed Standard Practice for Surveys on Research and Experimental Development*. Paris: OECD Publishing.
OECD. (2005a). *Measuring Globalisation. OECD Handbook on Economic Globalisation Indicators*. Paris: OECD.
OECD. (2005b). *Oslo Manual. Guidelines for Collecting and Interpreting Innovation Data*: OECD Publishing.
OECD. (2007). *INFORMATION ECONOMY - SECTOR DEFINITIONS BASED ON THE INTERNATIONAL STANDARD INDUSTRY CLASSIFICATION (ISIC 4)*. Paris: OECD.
OECD. (2008a). *OECD Factbook 2008: Economic, Environmental and Social Statistics*. Paris: OECD.
OECD. (2008b). *Science, Technology and Industry Outlook*. Paris: OECD.
OECD. (2010). *Measuring Innovation: A New Perspective*. Paris: OECD.
Ojanen, V., & Vuola, O. (2003). *Categorizing the Measures and Evaluation Methods of R&D Performance: A State-of-the-art Review on R&D Performance Analysis*. Lappeenranta: University of Technology.
Ottaviano, G., & Thisse, J.-F. (2004). Agglomeration and economic geography. In J. V. Henderson, P. Nijkamp, E. S. Mills, P. C. Cheshire & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2563-2608): Elsevier.
Parham, D. (2009). *Empirical Analysis of the Effects of R&D on Productivity*: OECD Publishing.
Patel, P., & Pavitt, K. (1991). Large Firms in the Production of the World's Technology: An Important Case of "Non-Globalisation". *Journal of International Business Studies, 22*(1), 1-21.
Patel, P., & Vega, M. (1999). Patterns of internationalisation of corporate technology: location vs. home country advantages1. *Research Policy, 28*(2-3), 145-155.
Perez, M., & Sanchez, A. (2002). Innovatio and firm location in the Spanish medical instruments industry. In Z. J. Acs, H. de Groot & P. Nijkamp (Eds.), *The Emergence of the Knowledge Economy: A Regional Perspective (Advances in Spatial Science)*. Berlin: Springer.
Picci, L. (2010). The internationalization of inventive activity: A gravity model using patent data. *Research Policy, 39*(8), 1070-1081.
Puga, D. (2010). The Magnitude And Causes Of Agglomeration Economies. *Journal of Regional Science, 50*(1), 203-219.
Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society Series A, 168*(2), 307-323.
Saltelli, A., Tarantola, S., & Campolongo, F. (2000). Sensitivity Analysis as an Ingredient of Modeling. *Statistical Science, 15*(4), 377-395.
Spizzirri, L. (2011). Algebra in Geography: Eigenvectors of Networks.
UNCTAD. (2005). *World Investment Report 2005. Transnational Corporations and Internationalization of R&D*. New York and Geneva: United Nations.
Veugelers, R. (2009). A lifeline for Europe's young radical innovators. Bruegel Policy Brief 2009/01, March 2009 (pp. 8).

Wandel, C. (2010). *Industry Agglomerations and Regional Development in Hungary. Economic Processes during European Integration* (Vol. 42). Frankfurt am Main: Peter Lang.

World Bank. (2009). *World development report 2009 : reshaping economic geography*. Washington, DC: World Bank.

Yoon, J., Blumer, A., & Lee, K. (2006). An algorithm for modularity analysis of directed and weighted biological networks based on edge-betweenness centrality. *Bioinformatics, 22*(24), 3106-3108.

European Commission
EUR 26264 – Joint Research Centre – Institute for Prospective Technological Studies

Title: Identifying European Poles of Excellence: The Methodology

Authors: Giuditta De Prato, Daniel Nepelski

Luxembourg: Publications Office of the European Union

2013 - 51 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online)

ISBN 978-92-79-34484-8 (pdf)

doi:10.2791/35893

Abstract

The Commission Communication entitled "A Strategy for ICT R&D and Innovation in Europe: Raising the Game" proposes reinforcing Europe's industrial and technology leadership in ICT. Building on Europe's assets, the Communication anticipates a landscape where, by 2020, "(…) Europe has nurtured an additional five ICT poles of world-class excellence (…)". This study attempts to identify ICT R&D&I-related agglomeration economies in Europe that would meet world-level excellence, and to identify weak signals that would indicate the dynamics of a changing ICT-related economic geography in Europe. Both of those identification processes are based on quantitative data, built on a set of relevant criteria leading to measurable indicators. The study is developed around several tasks, the results of which are presented in a series of EIPE reports. This second EIPE report develops an empirical assessment strategy for the identification of ICT agglomerations in Europe. It defines measures and data sources used in the study.

LF-NA-26264-EN-N

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

Publications Office

doi:10.2791/35893
ISBN 978-92-79-34484-8 (pdf)