CIMR Research Working Paper Series

Working Paper No. 57

Mapping the distribution of Internet of Things competences across European regions

by
Margherita Russo[1], Annalisa Caloffi[2], Ana Colovic[3], Pasquale Pavone[1], Saverio Romeo[4], Federica Rossi[4]

15/12/2021

ISSN 2052-062X

[1]University of Modena and Reggio Emilia, Italy

[2]University of Florence, Italy

[3]NEOMA Business School, France,

[4]Birkbeck, University of London, UK

Mapping the distribution of Internet of Things competences across European regions

Margherita Russo, Annalisa Caloffi, Ana Colovic, Pasquale Pavone, Saverio Romeo, Federica Rossi

Margherita Russo*, University of Modena and Reggio Emilia, Italy, margherita.russo@unimore.it

Annalisa Caloffi*, Department of Economics and Management, University of Florence, Italy, annalisa.caloffi@unifi.it

Ana Colovic, NEOMA Business School, France, ana.colovic@neoma-bs.fr; corresponding author

Pasquale Pavone*, University of Modena and Reggio Emilia, Italy, pasquale.pavone@unimore.it

Saverio Romeo, Centre for Innovation Management Research -Birkbeck - University of London, and Founder ETO (Emerging Technology Observer), s.romeo@bbk.ac.uk

Federica Rossi*, Birkbeck, University of London, UK, f.rossi@bbk.ac.uk

* and CAPP - Research Centre for the Analysis of Public Policies, University of Modena and Reggio Emilia, Italy

Abstract

Digital transformation is a key strategic issue for countries and regions aiming to boost economic growth, job creation, technology development and innovation. With a focus on the Internet of Things (IoT) the paper maps the potential of IoT technologies across European regions, using textual analysis applied to the description of companies' activities. Results identify three categories of regions (IoT leaders, co-designers and suppliers) capturing their potential to harness opportunities in IoT, based on the variety of IoT competences that are present. This mapping can support regional policies, particularly in the context of smart specialization strategies building on IoT systems.

Keywords: Internet of Things, regional competences in IoT, text-mining, cluster analysis

JEL classification: O33, O32, O14, R12

1. Introduction

New digital technologies associated with Industry 4.0 – Internet of Things, cloud services, big data and analytics (GTAI, 2014; Hermann et al., 2016; Frank et al., 2019) - are transforming sectors, value chains and production systems (De Propris and Bailey, 202; Capello and Lenzi, 2021a, 2021b). These technologies are expected to shape the geography of innovation and knowledge production (Balland and Boschma, 2021) and create a range of opportunities for those regional economies that are able to exploit them.

Among the factors that can help regions to harness the potential of Industry 4.0 (henceforth: I4.0), the presence of a sizeable set of firms providing core elements of the technology, or parts of the value chain, is crucial: having firms with the required competences can allow the region to lead in technology development, to identify new applications of the technology leading to new potential markets, and to diversify into related technologies (Cooke and Schwartz, 2008; Boschma and Iammarino, 2009). In fact, firms' current competences strongly influence the possibility of further technological advancement along the same technological trajectory (Boschma et al., 2013; Heimeriks and Boschma, 2013; Kogler et al., 2013), as well as the potential for discovering new applications of the technology, which pave the way for market expansion (Capello and Lenzi, 2021a). They also provide the building blocks for further diversification into related technologies (Boschma and Frenken, 2011; Neffke et al., 2011), an aptitude that regional policies aim to leverage when they design their regional innovation smart specialisation strategies (Balland et al., 2019).

The territorial mapping of competences in I4.0 can highlight which contexts might have opportunities for growth in the new technological and market scenario, and which ones might instead experience difficulties in fitting in. Lacking knowledge about the competences of the firms present in their territory might prevent policymakers from designing comprehensive territorial-based innovation strategies aimed at enhancing local competitiveness in I4.0. It also limits the potential of policy instruments designed to develop specific competences, or to build on competences that are already present in the current ecosystems, both within and across regions (Cooke and Schwartz, 2008; Balland and Boschma, 2021; Capello and Lenzi, 2021a). While research has suggested that the competences needed to harness the new digital technologies are unevenly spread across European regions (Muscio and Ciffolilli, 2020), few attempts have been made to develop a regional mapping of competences of I4.0 (De Propris and Bailey, 2020; Capello and Lenzi, 2021a, 2021b). A crucial factor that makes it difficult to map the geographical distribution of competences in I4.0 technologies is that they do not fit existing industry classifications, a problem that is common to most emerging technologies (Feldman and Lendel, 2010).

Some scholars have approached the problem of mapping the geographical distribution of I4.0 knowledge by relying on patent data (Balland and Boschma, 2021;

Corradini et al., 2021). While these endeavours are certainly valuable, they also present some limitations. First, patents capture inventing activities rather than the development of actual products and production systems, hence they are most appropriate to uncovering where new knowledge about these technologies is generated, rather than the geographical location of design and production expertise (Capello and Lenzi, 2021a; 2021c). Second, it is well known that not all firms in the digital sector engage in patenting, either because they specialise in software development, where patenting opportunities are limited (Schohe et al., 2019), or because the solutions proposed to users are often customized, and therefore deprived of any patenting potential. Hence, mapping exercises based on patents exclude a large share of the firms in the sector. Third, emerging technologies are often multi-layered, combining hardware, software and platforms, which cannot be captured accurately by patent data. Instead of analysing patents, we look at production competences of firms (Capello and Lenzi, 2021a). Focusing on a specific part of the I4.0 technologies – the Internet of Things (IoT) – we develop an original methodology to map the regional distribution of IoT competences in Europe. Our objective is to identify firms' production competences by classifying description of their activities to a finer grain than their NACE code, drawing on the textual description of their activities. After identifying a set of IoT-related activities, we map their presence in 18 European countries. In addition to being important in its own right, geographic mapping allows us to identify leading regions, where the entire IoT value chain is present, co-designer regions, where companies have extensive IoT expertise, albeit concentrated in some particular activities, and supplier regions, where some scattered expertise is present.

The paper is structured as follows. In section 2, we discuss strands in literature that provide a background for our research, and then introduce IoT and its key components. In section 3, we provide an overview of technology mapping techniques applied to IoT, considering both bibliometric and big data approaches based on web scraping, Section 4 describes the data and the methodology and discuss the extent to which our methodology complements and extends current approaches. In section 5, we present the outcomes of the mapping exercise and propose a typology of regions based on the variety of IoT competences they feature. Finally, in section 6, we draw some conclusions and suggest avenues for future research. The Appendix presents supplementary material. A selection of figures in the text (marked with the symbol) can be browsed online using the Tableau Public navigation tool available at https://[anonymised, to be added]

2. Background: Competitiveness of regions and The Internet of Things

As it is well known, technologies are not place-neutral. They show important territorial roots both in the way they are created and in the way they develop over time. The impact that these technologies have on the territory influences the development possibilities of the regions (Perez & Soete, 1988).

Technological inventions and innovations arise in places with specific localised endowments of competences and resources (Jaffe et al., 1993; Capello & Lenzi, 2014; Boschma, 2017). The regional knowledge base is the collection of accumulated competences and resources present in the region; it results from past choices of firms, public policy, and the "information and communication ecology created by face-to-face contacts, co-presence and co-location of people and firms within the same industry and place or region" (Bathelt, Malmberg and Maskell, 2004, p. 38). Competences and resources accumulate over time through different forms of learning (Morgan, 1997) as well as spillovers via labour markets, formal, and informal networks (Maskell and Malmberg, 1999). Since they depend not only on explicit, but also on tacit and cumulative knowledge, competences are very specific to the local context (Frenken and Boschma, 2007; Buenstorf and Klepper, 2009). In turn, the future technological development of the region is dependent on the regional knowledge base, with a cumulative character (Martin & Sunley, 2006). Regions that show strong competences in certain technologies are more likely to develop new technologies, related to the previous ones, than regions that did not possess the same pre-existing knowledge base (Balland and Boschma, 2021). The concept of 'technological relatedness' has been used, in fact, to explain the technological trajectories of regions (Boschma, 2017). Scholars define relatedness "in terms of similarities between activities that capture the cognitive dimension of capabilities (and thus implies learning)" (Boschma, 2017, p. 352). Thus, activities are related when they require similar knowledge or input (Hidalgo et al., 2018). According to Breschi et al. (2003), relatedness occurs when knowledge is cognitively close, giving rise to interactive learning, and when the same knowledge is used in not just one, but in multiple technologies. Moreover, according to Boschma (2017), relatedness also encompasses complementarities – the need to draw together and combine different activities, technologies or products to accomplish specific goals (Broekel and Brachert, 2015).

The regional knowledge base also influences the way in which innovations and inventions - both those developed in the region and those created elsewhere - can be applied in the region, giving rise to new productive activities or enhancing existing ones (Capello and Lenzi, 2021a; 2021d). This cumulative nature of knowledge and technology diffusion at the regional level can lead to an uneven distribution of new technologies and their applications, with an increasing distinction between regions rich in knowledge and/or innovative productive activities and lagging-behind regions (De Propris and Bailey, 2020; Balland and Boschma, 2021; Capello and Lenzi, 2021a).

Evidence suggests that, also in the case of I4.0 technologies, existing competences and resources affect which new activities will be developed in regions (Hidalgo et al., 2018; Neffke et al., 2011; Ciffolilli & Muscio, 2018; Muscio and Ciffolilli, 2020). I4.0 technologies are more likely to spread within advanced manufacturing regions where there is a greater availability of technological competences related to previous technology waves (World Bank, 2017; Balland and Boschma, 2021). This is also true for competences in the IoT, although this technology area - being a bundle of different technologies - seems to present a less strong degree of concentration than other I4.0 technologies (Balland and Boschma, 2021).

IoT is an enabler of next-generation manufacturing, connecting physical objects to the Internet and allowing them to exchange information (McKinsey Global Institute, 2013; Rong et al., 2015; Trappey et al., 2017). Since the development of any IoT solution requires the integration of both hardware and software technologies, as well as additional services (de Sousa Jabbour et al., 2018; Li et al., 2017), and because the relevant competences are often distributed across several firms of different size, we expect different regions to host specific combinations of IoT competences, and thus to exhibit various levels of ability to harness the potential of this key enabling technology.

Although many key enabling technologies – as some consider the IoT to be (Küfeoğlu, 2021) – are highly complex and characterised by a high degree of uncertainty – in terms of the nature of the key players, the products and processes emergent from them, the viable marketing strategies and profitable business models (Srinivasan, 2008) – they are very important for regional development (Adner and Levinthal, 2002). In fact, having a strong knowledge base in key enabling technologies is linked to a greater regional economic performance (Laursen, 2000) and a greater number of new technological specialisations, which facilitates the branching out of the regional economy into new directions (Montresor and Quatraro, 2017).

However, regions differ widely in their ability to develop competences in a specific enabling technology (Evangelista et al., 2017). Even regions that are known for their high-tech capabilities are often very different one from the other, with high tech employment, patents and venture capital funding being concentrated in only a few industry segments in each region (Cortright and Mayer, 2001). This variety is even starker when we move beyond high tech regions and consider regions that are lagging in technological capabilities.

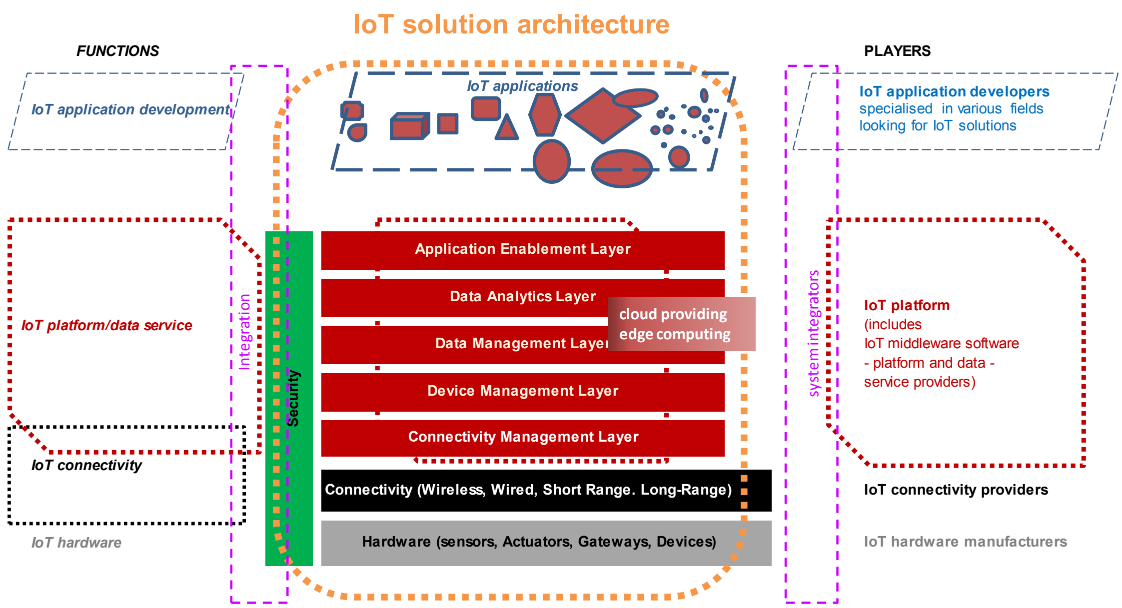
After describing the characteristics of the IoT (section 2.1), we will define the theoretical framework we use to classify regions according to their potential for IoT development (section 2.2).

2.1. The Internet of Things

IoT is a system that involves several complementary technologies, including software applications, connectivity and hardware components and devices, which rely on sensory, information-processing, communication and networking technologies to provide solutions for specific applications (Zhang and Chen, 2020). IoT technologies and applications emphasize strong connectivity, strong reliability, security, privacy, extremely low latency and the capacity to cope with a huge amount of data (Jiang et al., 2021; Kim, 2021; Sisinni et al., 2018). Like many other key enabling technologies, IoT solutions can be considered as a platform-based ecosystem (Tiwana et al., 2010; Teece, 2018), specifically targeting various types of applications. While each IoT solution is delivered through a complex set of devices and interconnected systems, and the specific features of each solution tend to be unique, IoT solutions architecture is made up of several conceptually distinct elements, or layers, that are present in all IoT solutions: software, connectivity and hardware. Adapted from Romeo (2016), a view of these elements, their functions and the various categories of players is presented in Figure 1. This view is widely accepted in the examination of IoT (Atzori et al.,2010; Chou, 2018; Navani et al.,2017; Razzaque et al. 2016; Sethi et al., 2017).

The software layers include IoT middleware software – platform and data – service providers. It is the IoT platform (in red) that embeds five essential layers: application enablement, data analytics, data management, device management, and connectivity management. The connectivity layer (in black) embraces wireless, wired, short-range, and long-range connectivity. The hardware layer (in grey) is the set of sensors, actuators, gateways, and computer and peripheral devices. The vertical functions of integration and security (in green) apply to all layers in specific ways. The implementation of IoT solutions also depends on IoT software developers that are specialised in various fields and are not totally bound within the IoT solutions architecture.

Figure 1 - IoT Solution architecture: functions and players, by layer



Source: Authors – Adapted from Romeo (2016).

IoT solutions are provided by an array of business organisations and service providers, each performing different but complementary functions that, together, can produce complex projects (Ikävalko et al., 2018; Ibarra et al., 2018). In general, IoT provision relies on complex value chains involving diverse competences across different domains, including, among other things, software engineering, telecommunications engineering, information networks management, and the manufacturing of hardware devices (Scully and Lueth, 2016; Romeo, 2020), without specific codes in the classification of economic activities, such as the NACE codes, that would allow its identification across industries.

2.2. Classifying regions according to their potential for IoT development

In the domain of emerging digital technologies such as IoT, the complexity and interdependence of tasks, skills and competences have frequently been put forward. Indeed, complex technological solutions often require an array of competences (Metallo et al., 2018). Following the arguments proposed by the literature on technological relatedness (Neffke et al 2011; Boschma, 2017; Balland and Boschma, 2021), particularly the notion that relatedness builds on complementarities, and not just on the similarities, between knowledge bases (Makri et al., 2010), it can be assumed that those regions that have a greater variety of competences in IoT at their disposal, should be better able to harness the technological and market opportunities created by these technologies, ensure the current regional competitiveness and secure the future trajectory of development in these technologies.

To map IoT competences across European regions, and critically assess the ability of various regions to harness IoT potential, we consider two dimensions capturing variety in IoT competences: variety of IoT-related industrial activities and variety of firm sizes.

Variety of industrial activities underpins greater potential to capture technological opportunities

There are a number of reasons why having greater variety of industrial activities in an enabling digital technology like IoT favours greater potential to harness opportunities for technological development in the same, or related, technological domains.

First, the co-location of firms with different industrial activities supports several mechanisms that have been identified as underpinning further technological development in the region (Boschma and Frenken, 2011). These are, in particular: (i) the greater interdependency among actors positioned along the same value chain, which drives the search for innovation throughout the region if new innovations are introduced in one part of the system; (ii) the technological complementarity among industries, which enables the introduction of major new innovations; (iii) supply relationships, where innovation on the supply side drives innovation down the value chain.

Second, an important effect of new technologies is that the “value” in the value chain migrates over time, so that a given technology’s potential to generate innovation rents may vanish from part of the value chain, rendering some firms’ business models obsolete (Srinivasan, 2008). Regions that have a greater variety of industrial activities within their borders are less likely to find themselves shut out of further technological development due to this shift in value chain opportunities.

Third, lagging regions are typically characterised by low levels of knowledge complexity and they also lack the diverse set of capabilities from which to derive their own complex technologies (Balland and Rigby, 2017). There are many instances of regions following trajectories founded in historical regional strengths that ultimately lead to ‘rigid specialisation’ and ‘lock-in’. The more varied the industrial activities around the enabling technology, the less likely this lock-in is to occur.

In our empirical analysis, we will therefore consider the variety of industrial activities related to the enabling technology in a region as an indication of that region’s potential to capture technological opportunities in that and/or in related technological domains.

Variety of company sizes underpins greater potential to capture market-related opportunities

Another element that can provide an indication of the potential for further development of the enabling technology in the region is the type of firms (in terms of their specialisation, size and value chains to which they belong) that provide it. We argue that a greater variety of firm sizes, including SMEs and large firms, favours greater potential to harness opportunities for market expansion. This is so for several reasons.

First, at the early stages in the technology evolution process, many technologies have limited if any functionality, and are suited only for limited, narrow applications. Over time, the nature of innovation in the technology changes to applications with growing commercial potential. In addition, as the technology develops, uncertainty around its performance decreases; the technology is applied to product applications with well-defined consumer benefits, and the performance–price ratios improve. Often, though not in all sectors or technological contexts, with growing commercial potential of the technology, large, established firms concerned about the effects of the new technology on their current business models begin to participate in the development of the technology (Abernathy and Utterback, 1978). In such contexts, large, well-established firms, become crucial players insofar as they can scale up the diffusion of the technology and grow new markets in distant locations and applications. Hence, as the technology matures, a combination of small and large firms is required to further entrench the industry within the region.

Second, an underlying mechanism that can strengthen the emergence of new industries is the creation of markets, for instance, through public procurement (Edler and Georghiou, 2007), including in the form of large infrastructure projects. These present a stable and sufficiently large market that enables regional firms to pursue new economic possibilities that can sustain innovation in the regional economy as a response to a new market creation (van den Berge et al., 2020). These projects are likely to require the involvement of large established firms, which are able to attract public funds to the region and thus strengthen the market for the enabling technology.

In our empirical analysis, we will therefore analyse the variety of sizes of firms involved in IoT in the region as an indication of the region’s potential to capture opportunities for market expansion.

3. An overview of technology mapping

Mapping the development of major new technologies, both spatially and according to other dimensions (such as scientific fields, technology classes, domains of application), is a challenge that has attracted increasing interest from academics, policymakers and industry alike (Daim et al., 2006; Jeong and Yoon, 2015; Pinto, 2009). Many mapping exercises rely on bibliometric analyses of publications or patents, which can be geo-referenced on the basis of inventors, applicators or authors' location to detect the dominant countries, regions or clusters in a particular technology (Daim et al., 2006; de Miranda Santo et al., 2006; Youtie et al, 2016; Ardito et al., 2018).

Focusing on Industry 4.0 technologies, the literature has underscored the uneven spatial distribution found in Europe. The most active regions are those that were already leaders in the third industrial revolution (Ménière et al., 2017; Ciffolilli and Muscio, 2018; Balland and Boschma, 2021). Using patent data, Balland and Boschma (2021) showed that the development of these technologies is path-dependent: as predicted by the relatedness framework (Boschma, 2017), new Industry 4.0 technologies tend to develop in places where there is a pre-existing base of related technologies. Large urban areas with the biggest research and technology transfer infrastructures such as London, Paris, Berlin and Madrid have emerged as leaders in artificial intelligence, quantum computers and other IT-related technologies. Technologies such as 3D printing have developed in the old manufacturing world's most innovative regions. On the other hand, Capello et al. (2020) identified a large number of regions involved in patenting application technologies using Industry 4.0 core and enabling technologies within specific application contexts. Corradini et al. (2021) conducted an analysis on patents filed between 2000 and 2014 to study the diffusion of four Industry 4.0 technologies across Europe; they find that cumulated regional technological capabilities, relatedness, technological search breadth and spatial proximity to Industry 4.0 invention, all play a role in explaining patenting activity in Industry 4.0 technologies at regional level. While these approaches are able to capture leading fields and locations of research and technological development, their scope is more limited when the mapping focus is not on the main inventors that create the new technologies, but on the large population of firms that propose other original solutions by integrating these technologies, or that develop elements of the technologies like software which are not always patentable.

Taking a 'demand-side' perspective, Capello and Lenzi (2021c) identify the regions where Industry 4.0 technologies are applied. Bringing together quantitative evidence and findings from case studies in six European countries, they focused on the following broad domains: technology invention, technology adoption in manufacturing sectors, technology adoption in services, and the ways in which Industry 4.0 transforms regions.

Other authors have relied on a 'big data' approach. For example, NIESR (2013) and

Nathan and Rosso (2015) mapped the UK's digital economy by extracting information from company websites about their products and services (using a set of predefined keywords), developing new categories of digital products and services, and reclassifying NACE sectors based on this new information, thereby uncovering a large number of digital companies. By scraping product directories and fan websites, Mateos-Garcia et al. (2014) created a more comprehensive list of UK video games companies than would have been possible by simply relying on NACE codes. The 'big data' approach overcomes the limitation of relying solely on outdated codes of economic activity. Additionally, it provides an up-to-the-minute picture of company activities, as it builds on up-dated information obtained online. On the other hand, this approach is probably insufficient to develop comprehensive mapping when the technology under scrutiny is complex and identification of individual keywords is insufficient to subsume all the relevant companies. In such cases, it might be reasonable to combine several sources (e.g., web scraping, text mining, sectoral codes), including expert judgement, to make sense of the spatial and organisational features of the new technology. With a focus on IoT, our methodology complements and extends current approaches in providing the identifications of IoT domains and the regional mapping of companies engaged in those domains.

4. Data and methodology

4.1 Data source

To perform our mapping exercise on IoT competences in European regions, we started from the Bureau van Dijk (BvD) Amadeus database. In recent years, in addition to the usual companies' balance sheet data, the database also provides textual information on the companies' activities, which are included in the field "overview".[1] We choose this database to leverage the opportunities offered by this descriptive field. The description of a company's activity, provided by the company itself on its own website, is potentially a very interesting source: we can assume that a company will pay attention to this channel of communication with potential customers, in order to showcase what it can do and what its strengths are, and that it will ensure that this information is accurate and current. Hence, this is a useful and up-to-date information that we propose to analyse using text mining techniques in order to detect specific information regarding IoT related activities. However, this information is not available for all the companies, but is more than adequate to identify the IoT activities associated to the various NACE codes, as we will describe in detail below.

[1] As described by BvD, this field is filled through a 'supervised' web scraping procedure on the firms' websites. The supervision consists in the assistance of one of the BvD consultants who indicates which fields of the websites are relevant and helps to classify the retrieved information (e.g., main products, main customers, company history).

4.2 Expert pre-selection: 4-digit NACE codes, 18 European countries

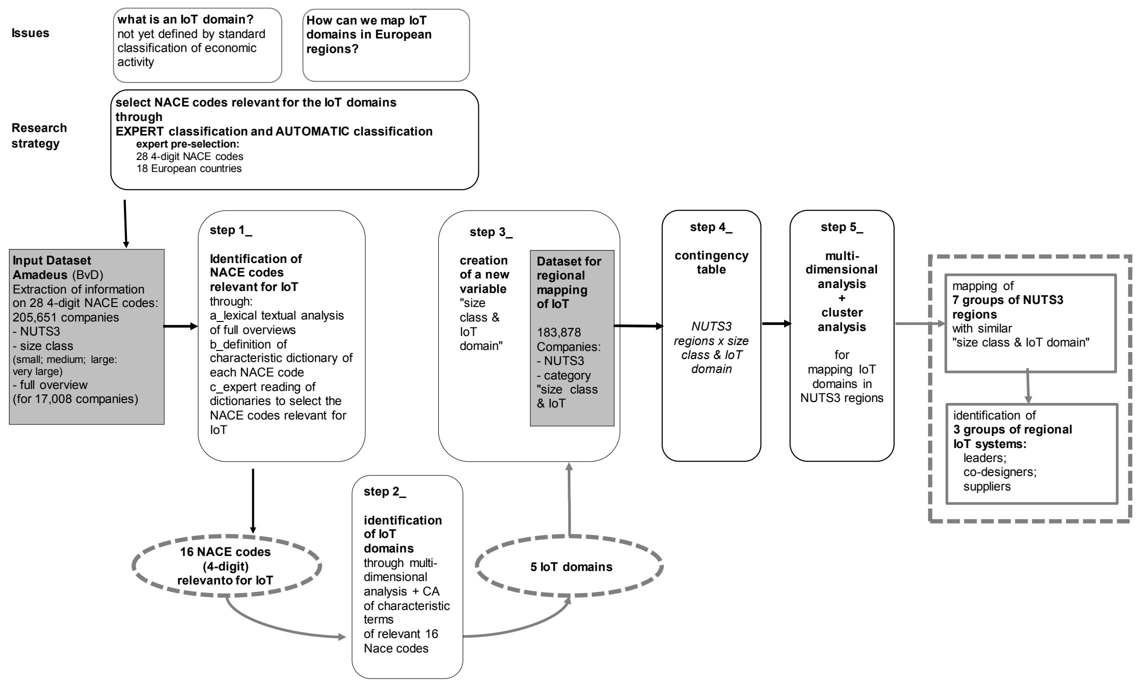
To extract data from the Amadeus dataset, we pre-selected a number of 4-digit NACE Rev.2 codes within the following divisions (excluding those codes that are completely unrelated to the production of IoT): 26-Manufacture of computer, electronic and optical products; 27-Manufacture of electrical equipment; 61-Telecommunications; 62-Computer programming, consultancy and related activities; and 63-Information service activities. Since there are no specific codes linked to the many diverse IoT-related activities in the NACE classification of economic activities, to do this first skimming we relied on the opinion of an expert[2] in IoT, who is also one of the authors of this paper. NACE codes do not reflect the complexity of the IoT solution architecture described in Figure 1. However, that architecture could be simplified to meet the nature of the NACE codes. An IoT solution architecture can be simplified in a three layers stack. The hardware layer composed of devices, sensors, and other machines. The connectivity layer encompasses all the telecommunications services needed. The software layer brings together a variety of components, from cloud to data analysis, all software-based. That has enabled us to associate the NACE codes to a specific part of the IoT solution architecture, for example: NACE code 61 (Telecommunications) is associated to the connectivity layer of an IoT solution. We identified a set of 28 codes listed in in Table A1 (Supplementary material). With respect to those codes, we compared the number of companies and employees in EU-27 countries plus UK, identifying a group of 18 countries with a significant presence of companies (Table A2).

In these countries, the number of companies operating in the preselected 4-digit NACE codes are 205,651. The field "full overview" is available only for 17,008 companies. As shown in Table A3 (Supplementary material), full overview was available for 97% of very large companies, 92% of large companies, almost 22% of medium-sized companies, and only 1% of small companies. Despite this limitation, we decided to use the data for our research analysis, assuming that small companies, each performing a very specific activity, are significantly represented in the sample of observations with full overview, while medium size and larger companies, which perform a greater variety of activities, both within each company and across the set of similar sized companies, are well represented by the information coverage in their full overview.

Our methodology is schematically represented in Figure 2, which highlights issues, research strategies, hypotheses, data and the five-step procedure we adopted to address the research questions at the core of this paper.

[2] The expert has an extensive industry experience in the IoT domain and is working with the largest consulting firms in the area of emerging technologies.

Figure 2 – Methodology: multi-step procedure adopted to map IoT domains in European regions



The multi-step procedure summarized in Figure 2 is described in detail below.

Step 1 Text analysis of full overview to identify the NACE codes-relevant to IoT domains

We created a corpus (collection of texts) of full overviews (available for 17,008 companies). It included 54,518 different terms (vocabulary of the corpus) for a total of 2,128,018 occurrences (size of the corpus). The lexical-textual analysis allowed us to extract 7,611 different active terms[3] (nouns and adjectives, including 3,781 multi-word expressions[4]) for a total of 635,384 occurrences (i.e., number of times in which active

[3] By active terms we mean all lexical forms (words) selected for the purpose of the analysis and thus contrasted with the terms we can call supplementary. In order to identify the semantic contents of the texts and to obtain groups of NACE codes on the basis of semantic contents, we therefore consider as active forms all words grammatically recognised as adjectives and nouns. The latter in particular represent the objects and subjects of texts and are therefore the central element of the message conveyed by a text. Moreover, in order to disambiguate the potentially ambiguous meaning of some words, we proceeded with a multiword expressions recognition, which, by linking the simple form to its qualification, allows both to improve the clustering process and to facilitate the interpretation process of the results.

[4] Multiword-Expressions (MWEs) represent all idiomatic nouns and technical-specialist terms, therefore representing the specialised terminology of a sector. The recognition of MWEs was performed by applying an information extraction model based on grammatical annotations and the search for recurrent syntactic structures (Pavone, 2018). The 20 most recurring MWEs we have found are: *information technology, data processing, consulting service, provision of computer, computer programming service, communication technology, financial service, computer hardware, software*

terms occur in the text). We then created an *Active terms × NACE codes* matrix (7611 × 28) in which we identified the characteristic dictionaries of the 28 pre-selected NACE codes by calculating the test-value[5] (Lebart et. al, 1998, p. 95) of each active term within each NACE code.

The expert reading of the characteristic dictionary of each NACE code allowed us to refine the initial list of codes and select 16 NACE codes (out of 28 identified previously) that are relevant to the IoT domain (list available in Table A4, Supplementary material).

Step 2 Identification of IoT domains through expert reading of characteristic NACE code dictionaries

In order to identify specific IoT layers and components in the IoT solution architecture (summarized in Figure 1), we used again a text-mining approach. We classified the 16 NACE codes with regard to their similarity in terms of IoT content, based on the characteristic terms emerging from the analysis of the full overviews. This classification was obtained through a correspondence analysis[6] of the *Relevant to IoT NACE Codes* × C*haracteristic Terms* (16 × 7,440) matrix. Through a cluster analysis applied to the first ten factors resulting from the correspondence analysis, we were able to group the NACE codes into clusters based on the similarity of distribution of terms[7]. This clustering phase constituted an unsupervised and unambiguous classification of the NACE codes, reflecting the semantic similarity between them, which could then be summarised in a category that was not defined *a priori* but was derived from the analysis.

This process allowed us to define five clusters of NACE codes that singled out specific IoT-related activities which we called 'IoT domains'. Table A5 (Supplementary material) summarises the key characteristics of these domains, which we have labelled on the basis of the main activities they encompass.

IoT domain 1 revolves around *software and data processing*; it includes competences relative to computer programming, designing computer systems, software development and software design, among others. These activities correspond to the 'IoT platform' layers in Figure 1 (highlighted in red). IoT domain 2's focus is on *telecommunications*. It includes know-how related to the provision of telecommunications services, satellites, broadband, radar station operations, satellite

---

development, supporting software, computer system, system integration, designing computer system, domestic market, project management, technology solution, data processing facility, provision of information, operation of clients, exceptional domain knowledge.

5 Test-value is a statistical criterion associated with the comparison of two portions (considered context and all the other contexts) within the framework of a hypergeometric law.

6 The correspondence analysis (Benzecri 1973, 1992, Greenacre 1984, 2016) is a factorial technique that can be used to obtain a reduced number of variables (or factors) on which to measure the similarity of a matrix, by examining row and column profiles.

7 We applied a mixed clustering (Lebart et al., 1998, p. 95) based on Ward's aggregation method (1963), with Euclidean distance.

tracking, networks, and different types of communication-related activities. These activities correspond to the connectivity layer (in black) in Figure 1. The next three domains correspond to different elements of the hardware layer (in grey) in Figure 1. In particular: IoT domain 3 includes competences related to *manufacturing of telecom equipment*, such as antennas, radio equipment, loudspeakers, cordless telephones, receivers and others; IoT domain 4 covers competences in *manufacturing of electronic components*, such as cables, fibre optics, connectors, wiring devices and others; lastly, IoT domain 5 comprises expertise in *manufacturing of measure instruments*, such as control instruments, sensors, precision tools, calibration, etc.

The classification of the 16 NACE codes (at 4-digits) allowed us to associate each company in the dataset to the five IoT domains. In this way we apply the classification made possible by the elaboration of descriptive texts of full overview to all the companies in the original dataset. Table 1 presents the data used for further steps in the analysis.

Table 1 – Number of companies belonging to the 16 identified NACE codes, by IoT domain and company size

Cluster IoT domains	'NACE code	Description	n. of companies		company size			
					Very Large	Large	Medium sized	Small
1 Software and Data Processing	2620	Manuf. of computers and peripheral equipment	**157682**	3872	62	197	812	2801
	6201	Computer programming activities		66514	343	1552	6166	58453
	6202	Computer consultancy activities		48575	307	1531	4448	42289
	6203	Computer facilities management activities		4886	72	228	510	4076
	6209	Other inform. techn. and computer service activities		24720	278	1169	2753	20520
	6311	Data processing, hosting and related activities		9115	107	530	991	7487
2 Telecommunica-tion	6110	Wired telecommunications activities	**12040**	3135	101	298	497	2239
	6120	Wireless telecommunications activities		2075	104	196	289	1486
	6130	Satellite telecommunications activities		252	27	55	40	130
	6190	Other telecommunications activities		6578	249	679	1039	4611
3 Manufacturing Telecom. Equip-ment	2630	Manuf. of communication equipment	**4880**	2831	62	231	732	1806
	2640	Manuf. of consumer electronics		2049	23	67	265	1694
4 Manufacturing Electronic Com-ponents	2731	Manuf. of fibre optic cables	**1202**	52		9	11	32
	2732	Manuf. of other electr. and electric wires and cables		350	25	55	124	146
	2733	Manuf. of wiring devices		800	9	95	295	401
5 Manufacturing Measur. Instrum.	2651	Manuf. of instruments and appliances for measuring, testing and navigation	**8074**	8074	156	638	2625	4655
Total			**183878**		**1925**	**7530**	**21597**	**152826**

Step 3: Creation of a new variable for the analysis

In step 3 we included the company class size in the analysis in order to take the variety of company sizes into account in the IoT mapping exercise. Combining each of the five IoT domains (identified in step 2) with each of the four company class sizes (small, medium-sized, large, very large), we obtained a new variable for the 183,878 companies in the dataset, namely, *IoT domain-and-Company size,* with 20 categories

Step 4: Construction of contingency table

Based on the new variable, we constructed a contingency table *NUTS3* × *IoT domain-and-Company size* (1095 × 20) to allocate companies of each NUTS3 region to the new variable[8].

Step 5 Clustering of NUTS3 regions according to the classes IoT domain-and-Company size

In the last step, we group the NUTS3 regions based on their similarities with respect to the new variable created in step 3 and weights provided by the contingency table of step 4. Grouping is obtained through a correspondence analysis and cluster analysis on the *NUTS3* × *IoT domain-and-Company size* (1,095 × 20) matrix[9].

Our clustering exercise allowed us to identify seven clusters of NUTS3 regions that we present in the following sections.[10]

5. Results and discussion

5.1. NUTS3 regions mapping

With regard to the characteristic categories in the cluster (details in Table A2, Supplementary material), we have labelled the seven clusters, identified from the mapping exercise as follows:

cl-1– Software service focus – Small companies
cl-2 – Mainly telecoms focus – SMEs
cl-3 – Telecoms and software services focus – Large companies
cl-4 – Entirely hardware focus – Mainly SMEs
cl-5 - Largely hardware focus with software services support – All sizes
cl-6 – Hardware focus with software and telecoms services support- Mainly small
cl-7 – Hardware and software services focus – All sizes

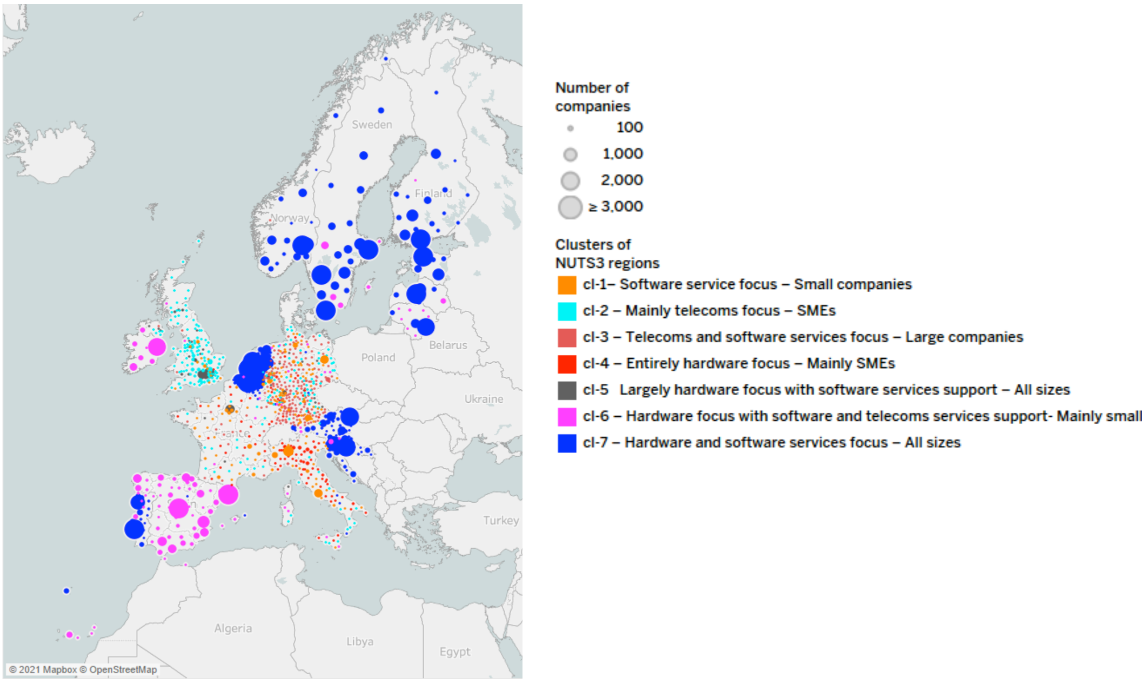
Figure 3 presents these results.

[8] Information on NUTS3 is present for 179,887 companies out of the 183,878 in the DB.

[9] Hierarchical clustering was implemented by applying the Ward method (Greenacre, 2016, p.120; Murtagh and Legendre, 2014; Ward, 1963) and chi-square distance.

[10] In Figure A2 we report the dendrogram of such cluster analysis. The optimal number of clusters is two, distinguishing between the NUTS3 regions characterised by the presence of small companies dedicated to software, data processing and telecommunications, and all the other NUTS3 regions. In order to obtain more detailed groups of territorial entities, seven groups of NUTS3 regions were selected.

Figure 3 - Map: Clusters of NUTS3 regions according to new *IoT domain-and-Company size* variable



Source: Authors elaboration based on data from Amadeus, downloaded on 30.09.2019

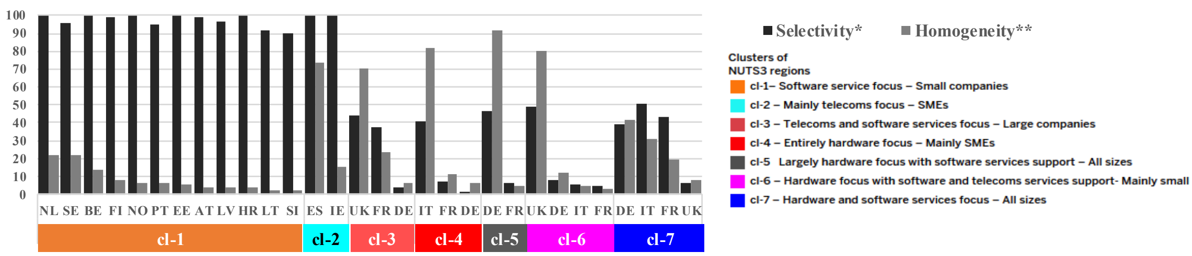
Hence, cl-1 (software service focus – small companies), cl-2 (largely telecoms focus – SMEs) and cl-4 (entirely hardware focus – mainly SMEs) each specialise in one of the relevant layers of the IoT solutions architecture, respectively, software services, telecoms (connectivity) and hardware. They are mainly composed of SMEs. Regions in cl-3 (telecoms and software services focus – large companies), cl-5 (largely hardware focus with software services support – all sizes) and cl-7 (hardware and software services focus – all company sizes) feature a greater diversity of competences, with strengths in at least two layers of the IoT system. Regions in cl-5 and cl-7 include, respectively, software services and hardware competences, the former mainly provided by a mix of SMEs and large companies, the latter by SMEs. Regions in cl-3 include connectivity and software services, mainly supplied by large companies. Regions in cl-6 (hardware focus with software and telecoms services support- mainly small companies) include all domains - hardware, connectivity and software - which are mainly provided by SMEs.

5.2. Focus on countries: selectivity and homogeneity in the cluster of NUTS3 regions

These results are explored in what follows with a focus on the pattern of regional mapping that emerges at country level. We summarise the patterns in terms of selectivity and homogeneity of the 18 countries with respect to the cluster of NUTS3 regions. The observations emerging from these explorations are then discussed to highlight the potential for regional IoT development.

In Figure 3 we can graphically observe that some countries are fully represented in clusters cl-2 and cl-1. Measurement of the extent to which a country is represented in each cluster is given by the selectivity index of the country, i.e., the specialization of the country in the cluster. A cluster homogeneity index indicates the weight of the country within the cluster. Both indexes are reported in Figure 3. For example, we can see that the Netherlands is represented within the cl-1 at 99.8%, and that its weight represents 22.3% of the cluster. Other countries present a greater variety of clusters (e.g., Germany, France, Italy and the Baltic countries), suggesting greater inter-regional diversity in IoT competence profiles within each country.

Figure 4– Focus on countries: selectivity and homogeneity in the clusters of NUTS3 regions



*Selectivity indicates the percentage of companies in the NUTS3 regions classified in the cluster.
**Homogeneity indicates the relative importance of the share of the country within the cluster.
Source: Authors elaboration based on data from Amadeus, downloaded on 30.09.2019

The specialised clusters, cl-1 (software service focus – small companies), cl-2 (largely telecom focus – SMEs) and cl-4 (fully hardware focus – mainly SMEs), are present in many regions of France, Italy, Spain, Portugal, Ireland, the Netherlands, Belgium, and the Scandinavian and Baltic countries, with many French and Italian regions presenting a specialisation in hardware, most Spanish and Irish regions in telecoms services, and most Scandinavian, Portuguese, Dutch and Belgian regions in software. These regions have less potential for further expansion of technological and market capabilities as they are strongly specialised in just one layer of the IoT architecture, with the main suppliers being SMEs. Most UK regions, and some French, German and Italian regions are in cluster cl-6 (hardware focus with software and telecoms services support- mainly small companies), which has greater potential to harness technological opportunities in IoT, though again the main suppliers are SMEs. Finally, many regions in Germany, France and Italy are in clusters cl-7 (hardware focus with software and telecoms services support – mainly small companies) and cl-5 (largely hardware focus with software services support – companies of all sizes), which have intermediate potential to harness technological opportunities, being specialised in two layers of the IoT architecture (hardware and software), with the latter including companies of all sizes. This offers greater potential to capture market-related opportunities. The map in Figure 3 shows that these competencies are concentrated in specific cities or areas within the regions. The number of companies differs widely as well, with some cities in the Iberian Peninsula (such as Barcelona, Madrid, Lisbon and Porto), Dublin in Ireland, and cities in the Netherlands, Belgium, Sweden and the Baltic countries exhibiting particularly large agglomerations of companies.

We previously argued that, while NUTS3 regions may have different combinations of IoT competences, and regions that concentrate on specific IoT layers potentially being highly competitive in their own area of specialisation, when it comes to the capacity to harness technological and market opportunities created by IoT, it is the regions that include a greater variety of competences and more diverse firm sizes (in particular, including a strong presence of large firms) that have the greatest potential. The presence of hardware in these areas, alongside other technological domains, points to the capacity to leverage the hardware technologies (considered more difficult to acquire) in order to expand into other technological IoT domains.

Table 2 provides a means to classify NUTS3 regions in terms of their potential for harnessing further technological and market-related opportunities in the field of IoT. IoT System Leaders are regions where the full IoT value chain and a mix of SMEs and large companies are present. These regions have the greatest potential to capture technological and market-related opportunities. IoT System Co-designers are regions where a substantial, but not an entire, IoT value chain is present. These are usually characterised by SMEs, although some large firms may also be present. Firms in these regions need to collaborate with other firms in other regions to provide complete IoT solutions (as complementarity cannot be achieved within the region), giving the regions' intermediate potential for harnessing technological and market-related opportunities. IoT System Suppliers are regions where only part of the IoT value chain has a significant presence, and where most of the firms are SMEs. Companies in these regions need to rely on extra-regional connections in order to participate in the provision of IoT solutions. Such regions have less potential for harnessing technological and market-related opportunities.

Table 2 – Potential for regional IoT development (high, intermediate, low) in the 18 countries

Results built on the relative importance of countries in the clusters of NUTS3 regions (based on the number of companies, homogeneity). In grey: countries with a homogeneity index below 20%

Regional IoT Systems	**Description**	**Clusters and countries (homogeneity index)** ***based on NUTS3 regions classification***
Leaders	The entire IoT value chain is present in the region. Higher potential for technological and market capabilities expansion	**cl-3** UK (70.28); FR (23.57); DE (6.15) **cl-5** DE (91.75); FR (5.10)
Co-designers	Substantial part of the IoT value chain is present in the region. Intermediate potential for technological diversification, lower potential for market capabilities expansion	**cl-6** UK (80.14); DE (12.33); IT (4.39); FR (3.13) **cl-7** DE (41.80); IT (31.01); FR (19.61); UK (7.58)
Suppliers	Limited part of the IoT value chain is present in the region. Lower potential for technological and market capabilities expansion	**cl-1** NL (22.32); SE (21.76); BE (13.99); FI (8.12); NO (6.38); PT (6.09); EE (5.97); AT (4.01); LV (3.80); HR (3.66); LT (2.15); SI (2.13) **cl-2**

		ES (73.81); IE (15.39) **cl-4** IT (81.83); FR (10.91); DE (6.23)

Source: Authors elaboration based on data from Amadeus, downloaded on 30.09.2019

5.3. Potential for the development of regional IoT systems

The analysis of our mapping exercise enabled us to establish a novel, comprehensive categorisation of regions with respect to their bundles of industrial activities as well as company sizes in IoT domains, and to evaluate their potential for harnessing the opportunities presented by IoT. As it is known, much of the potential of this technology comes from the fact that this can be considered a key enabling technology (Küfeoğlu, 2021). Regardless of the debate about whether this label can be applied to the IoT, there is no doubt that this type of technology has a very wide range of applications. The world's leading business consultancies have long identified this technology as the key to competing in the near future, not least because some of the technologies on which it is based are relatively mature (McKinsey, 2013). Moreover, these technologies have considerable potential to impact not only industries, products and services, but also people's lives (Espada et al., 2011). It is especially in a field such as this that important growth spaces open up, not only for those regions that are leaders in inventing these technologies, but also for those that are capable of imagining new ways of applying these technologies (Evangelista et al., 2017; Montresor and Quatraro, 2019).

Leading regions, possessing the whole mix of software and hardware, as well as the greatest organisational variety, may have important opportunities to identify such application possibilities. In co-designer and especially in supplier regions, more effort is required - by regional policymakers and others - to build growth potential around these technologies. However, suppliers can be islands of innovation that support the development of other innovation chains in the region.

6. Conclusion

Our study deploys an original mapping methodology combining NACE codes and text mining of descriptions of company activities to propose a more accurate depiction of regional strengths in individual IoT domains as well as in more complete value chains of IoT competences. While other mapping exercises have been published in recent years, scholars have tended to focus on activities related to Industry 4.0 more broadly, without conducting extensive and precise mapping of IoT in particular. Our study also differs from prior work insofar as we deployed an original methodology to respond to the challenges that arise when mapping new technologies, which are difficult to grasp within the existing classification of economic activities that are not usually suited to respond to new needs and cannot be significantly identified by using patent data.

The study contributes to theory by showing how the competence base of regions underpins their potential to develop and extend their technological bases in emerging digital technologies. Our findings have significant policy implications as they can support policymakers to identify which regions should be fostered as full IoT value-chain providers and which need to specialise further or, alternatively, diversify into new, related, complementary domains.

Our study has also some limitations, which open up avenues for further research. The present analysis is limited to 18 European countries, but IoT competences are found worldwide. Future research could expand the geographical scope of the analysis and additional sources of company information could also be explored since the Amadeus database presents some limitations.

Finally, this paper mainly focused on IoT systems on the 'supply side', mapping IoT competences at regional level, but the system also consists of the demand side, with IoT solutions tailored to customers' needs. Future studies might focus on the demand side to gain a deeper understanding of the IoT area as a whole. Moreover, we may guess that the "emergence" of competences in IoT domains stems from related competences already found in the region, and that these rely on demand from other regions. The system is thus made up of all these varying interconnections. While a preliminary analysis recently conducted by the European Commission et al. (2019) outlined the spatial distribution of demand for IoT solutions in European regions, as far as we know there has been no systematic investigation of all the entities in the system (which companies are present on the demand and the supply side, which regulatory agencies and policy actors) to date, nor of the relations between these entities.

References

Ardito, L., D'Adda, D., Messeni Petruzzelli, A. (2018) Mapping innovation dynamics in the Internet of Things domain: Evidence from patent analysis. Technological Forecasting and Social Change, 136, 317–330

Atzori, L., Iera, A., Morabito, G. (2010). The Internet of Things: A Survey. Computer Networks, 54(15), 2787-2805

Balland, P. A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. Regional Studies, 53(9), 1252-1268.

Balland, P. A., & Boschma, R. (2021). Mapping the potentials of regions in Europe to contribute to new knowledge production in Industry 4.0 technologies. Regional Studies, 1-15, https://doi.org/10.1080/00343404.2021.1900557.

Benzecri, J.P. (1973) L'analyse des données. L'analyse des correspondances. Paris: Dunod.

Benzecri, J.P. (1992) Correspondence Analysis Handbook. New York: Marcel Dekker.

Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. Economic geography, 85(3), 289-311.

Boschma, R. (2017) Relatedness as driver of regional diversification: A research agenda. Regional Studies, 51(3), 351-364.

Boschma, R., Frenken, K. (2011) Technological relatedness, related variety and economic geography. In P. Cooke, B. Asheim, R. Boschma, R. Martin, D. Schwartz, F. Todtling (eds) Chapter 14, The Handbook on Regional Innovation and Growth. Cheltenham: Edward Elgar.

Boschma, R., Minondo, A., Navarro, M. (2013) The emergence of new industries at the regional level in Spain: A proximity approach based on product relatedness. Economic Geography, 89(1), 29-51.

Breschi, S., Lissoni, F., Malerba, F. (2003) Knowledge-relatedness in firm technolog-ical diversification. Research Policy, 32(1), 69-87.

Broekel, T., & Brachert, M. (2015). The structure and evolution of inter-sectoral technological complementarity in R&D in Germany from 1990 to 2011. Journal of Evolutionary Economics, 25(4), 755-785.

Capello, R., & Lenzi, C. (2021a). Industry 4.0 and servitisation: Regional patterns of 4.0 technological transformations in Europe. Technological Forecasting and Social Change, 173, https://doi.org/10.1016/j.techfore.2021.121164.

Capello, R., & Lenzi, C. (2021b). 4.0 Technologies and the rise of new islands of innovation in European regions. Regional Studies, https://doi.org/10.1080/00343404.2021.1964698.

Capello, R., & Lenzi, C. (2021c). The regional economics of technological transformations. Industry 4.0 and Servitisation in European regions. London: Routledge.

Capello, R., & Lenzi, C. (2021d). 4.0 Technological Revolution and Economic Competitiveness: Unexpected Opportunities for Peripheral Areas. Scienze Regionali. Italian Journal of Regional Science, https://doi.org/10.14650/100722.

Chou, T. (2017). Precision – Principles, Practices, Solutions for the Internet of Things. McGraw Hill Education.

Ciffolilli, A., Muscio, A. (2018) Industry 4.0: national and regional comparative advantages in key enabling technologies. European Planning Studies, 26(12), 2323-2343.

Cooke, P., Schwartz, D. (2008) Creative regions: an introduction. In Cooke, P. and Schwartz, D. (Eds.), Creative regions: technology, culture and knowledge entrepreneurship. London and New York: Routledge.

Corradini, C., Santini, E., Vecciolini, C. (2021) The geography of Industry 4.0 technologies across European regions, Regional Studies, 55(10-11), 1667-1680.

Daim, T.U., Rueda, T., Martin, H., Gerdsri, P. (2006) Forecasting emerging technologies: Use of bibliometrics and patent analysis, Technological Forecasting and Social Change, 73, 981–1012.

de Miranda Santo, M., Massari Coelho, G., dos Santos, D.M., Fellows Filho, L. (2006) Text mining as a valuable tool in foresight exercises: A study on nanotechnology,

Technological Forecasting and Social Change, 73, 1013–1027.
De Propris, L., Bailey, D. (2020) (Eds) Industry 4.0 and regional transformations. London: Taylor & Francis.
de Sousa Jabbour, A. B. L., Jabbour, C. J. C., Foropon, C., Godinho Filho, M. (2018) When titans meet–Can industry 4.0 revolutionise the environmentally-sustainable manufacturing wave? The role of critical success factors. Technological Forecasting and Social Change, 132, 18-25.
Espada, J. P., Martínez, O. S., Lovelle, J. M. C., G-Bustelo, B. C. P., Álvarez, M. Á., García, A. G. (2011) Modelling architecture for collaborative virtual objects based on services. Journal of Network and Computer Applications, 34(5), 1634-1647.
Evangelista, R., Meliciani, V., & Vezzani, A. (2018). Specialisation in key enabling technologies and regional growth in Europe. Economics of Innovation and New Technology, 27(3), 273-289.
European Commission. Directorate General for Communications Networks, Content and Technology, The Joint Institute for Innovation Policy (JIIP), Joanneum Research Forschungsgesellschaft, Fundación TECNALIA RESEARCH & INNOVATION, VTT, KPMG. (2019) Study on mapping Internet of Things innovation clusters in Europe. Publications Office, LU.
Feldman, M. P., Lendel, I. (2010) Under the lens: the geography of optical science as an emerging industry, Economic Geography, 86(2), 147-171.
Frank, A. G., Mendes, G. H., Ayala, N. F., Ghezzi, A. (2019) Servitisation and Industry 4.0 convergence in the digital transformation of product firms: A business model innovation perspective. Technological Forecasting and Social Change, 141, 341-351.
Greenacre, M. (1984) Theory and application of Correspondence Analysis. London: Academic Press.
Greenacre, M. (2016) Correspondence Analysis in practice. New York: Chapman and Hall/CRC.
GTAI (Germany Trade & Invest) (2014) Industries 4.0-Smart Manufacturing for the Future.
Heimeriks, G., Boschma, R. (2013) The path- and place-dependent nature of scientific knowledge production in biotech 1986–2008. Journal of Economic Geography, 14, 339–364.
Hermann, M., Pentek, T., Otto, B. (2016) Design principles for industrie 4.0 scenarios. In: 2016 49th Hawaii International Conference on System Sciences, 2016, pp. 3928–3937.
Hidalgo, C., Balland, P. A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C. Kogler, D., Morrison, A., Neffke, F. Rigby, D., Stern, S., Zheng, S., Zhu, S. (2018) The principle of relatedness. In Springer proceedings in complexity. Springer. Paper originally presented at the International Conference on

Complex Systems (ICCS) (pp. 451-457). Cambridge, MA, USA, July 22-27.

Küfeoğlu, S. (2021). The Enabling Technology: Internet of Things (IoT). In Küfeoğlu , S (Ed) The Home of the Future, Springer, 109-140.

Ibarra, D., Ganzarain, J., Igartua, J.I. (2018) Business model innovation through Industry 4.0: A review. Procedia Manufacturing, 22, 4-10.

Ikävalko, H., Turkama, P., Smedlund, A. (2018) Enabling the mapping of Internet of Things ecosystem business models through roles and activities in value co-creation. Proceedings of the 51st Hawaii International Conference on System Sciences, http://hdl.handle.net/10125/50509

Jeong, Y., Yoon, B. (2015) Development of patent roadmap based on technology roadmap by analyzing patterns of patent development. Technovation, 39-40, 37–52.

Jiang, Z., Guo, Y., & Wang, Z. (2021). Digital twin to improve the virtual-real integration of industrial IoT. Journal of Industrial Information Integration, 22, 100196.

Kim, J. H. (2021) 6G and Internet of Things: a survey. Journal of Management Analytics, 1-17.

Kogler, D. F., Rigby, D. L., Tucker, I. (2013) Mapping knowledge space and technological relatedness in US cities. European Planning Studies, 21, 1374–1391.

Lebart, L., Salem, A., Berry, L. (1998) Exploring Textual Data. Dordrecht: Kluwer Academic Publishers.

Li, G., Hou, Y., Wu, A. (2017) Fourth Industrial Revolution: technological drivers, impacts and coping methods. Chinese Geographical Science, 27(4), 626-637.

Laursen, K. (2000) Trade specialisation, technology and economic growth. Edward Elgar Publishing,

Makri, M., Hitt, M. A., Lane, P. J. (2010) Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisi-tions. Strategic Management Journal, 31(6), 602-628.

Mateos-Garcia, J., Bakhshi, H., Lenel, M. (2014) A map of the UK games industry. London: Nesta.

McKinsey Global Institute (2013) Disruptive technologies: Advances that will transform life, business, and the global economy. Available at: https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/disruptive-technologies#

Ménière, Y., Rudyk, I., Valdes, J. (2017) Patents and the fourth industrial revolution. The inventions behind digital transformation. European Patent Office (EPO).

Metallo, C., Agrifoglio, R., Schiavone, F., Mueller, J. (2018) Understanding business model in the Internet of Things industry. Technological Forecasting and Social Change, 136, 298-306.

Murtagh, F., Legendre, P. (2014) Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? Journal of Classification, 31(3), 274–295.

Muscio, A., Ciffolilli, A. (2020) What drives the capacity to integrate Industry 4.0 technologies? Evidence from European R&D projects. Economics of Innovation and New Technology, 29(2), 169-183.

Navani, D., Jain, S., Nehra, S.M. (2017). The Internet of Things: A Survey of Architectural Elements, 13th International Conference on Signal-Image Technology and Internet-based Systems

Nathan, M., Rosso, A. (2015) Mapping Digital Businesses with Big Data: some early findings from the UK. Research Policy, 44(9), 1714–1733.

Neffke, F., Henning, M., Boschma, R. (2011) How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. Economic Geography, 87, 237–265.

NIESR (2013) Measuring the UK's digital economy using big data. Available at: https://www.niesr.ac.uk/sites/default/files/publications/SI024_GI_NIESR_Google_Report12.pdf

Pinto, H. (2009) The diversity of innovation in the European Union: Mapping latent dimensions and regional profiles. European Planning Studies, 17(2), 303-326.

Pavone, P. (2018) Automatic Multiword Identification in a Specialist Corpus. In A. Tuzzi (Ed.), Tracing the Life Cycle of Ideas in the Humanities and Social Sciences (pp. 151–166). Springer International Publishing.

Razzaque, M.A., Milojevic-Jervic, M., Palade, A., Clarke, S. (2016). Middleware for the Internet of Things: Survey. IEEE Internet of Things Journal, 3(1), 70-95.

Rong, K., Hu, G., Lin, Y., Shi, Y. Guo, L. (2015) Understanding business ecosystem using a 6C framework in Internet-of-Things-based sectors. International Journal of Production Economics, 159, 41-55.

Romeo, S. (2016) Looking for a Compass for Navigating the IoT Platform Landscape - IEEE Internet of Things, https://iot.ieee.org/newsletter/november-2016/looking-for-a-compass-for-navigating-the-iot-platform-landscape.html

Romeo, S. (2020) For a multi-stakeholder discussion on 5G in agriculture and rural area development.

Schohe, S., Appelt, C., & Goddar, H. (2019). Patenting software-related inventions in Europe. In Takenaka, T. (Ed) Research Handbook on Patent Law and Theory. London: Edward Elgar, 106-131.

Scully, P., Lueth, K. L. (2016) Guide to IoT Solution Development. IoT Analytics, whitepaper, September.

Sethi, P., Sarangi, S.R. (2017). Internet of Things: Architectures, Protocols, and Applications. Journal of Electrical and Computer Engineering, https://doi.org/10.1155/2017/9324035

Sisinni, E., Saifullah, A., Han, S., Jennehag, U., Gidlund, M. (2018) Industrial internet of things: Challenges, opportunities, and directions. IEEE Transactions on Industrial Informatics, 14(11), 4724-4734.

Teece, D. (2018) Profiting from innovation in the digital economy: Enabling technologies, standards, and licensing models in the wireless world. Research Policy, 47(8), 1367-1387.

Tiwana, A., Konsynski, B., Bush, A. A. (2010) Research commentary—Platform evolution: Coevolution of platform architecture, governance, and environmental dynamics. Information Systems Research, 21(4), 675-687.

Trappey, A. J., Trappey, C. V., Govindarajan, U. H., Chuang, A. C., & Sun, J. J. (2017). A review of essential standards and patent landscapes for the Internet of Things: A key enabler for Industry 4.0. Advanced Engineering Informatics, 33, 208-229.

Ward, J. H. (1963) Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association, 58(301), 236–244

Youtie, J., Porter, A., Shapira, P., Newman, N. (2016) Lessons from Ten Years of Nanotechnology Bibliometric Analysis, OECD Blue Sky Conference 2016.

Zhang, C., Chen, Y. (2020) A review of research relevant to the emerging industry trends: Industry 4.0, IoT, blockchain, and business analytics. Journal of Industrial Integration and Management, 5(01), 165-180.

Mapping the distribution of Internet of Things competences across European regions

SUPPLEMENTARY MATERIAL

Appendix

Table A1 – NACE Rev.2 codes at 2, 3 and 4 digits, used to identify the sample of companies in our IoT competences database

Codes	**Manufacturing**
C261*	Manufacture of electronic components and boards
C2611	Manufacture of electronic components
C2612	Manufacture of loaded electronic boards
C2620	Manufacture of computers and peripheral equipment
C2630	Manufacture of communication equipment
C2640	Manufacture of consumer electronics
C2651	Manufacture of instruments and appliances for measuring, testing and navigation
C2652	Manufacture of watches and clocks
C2731	Manufacture of fibre optic cables
C2732	Manufacture of other electronic and electric wires and cables
C2733	Manufacture of wiring devices
	Services
J61*	Telecommunications
J6110	Wired telecommunications activities
J6120	Wireless telecommunications activities
J6130	Satellite telecommunications activities
J6190	Other telecommunications activities
J620*	Computer programming, consultancy and related activities
J6201	Computer programming activities
J6202	Computer consultancy activities
J6203	Computer facilities management activities
J6209	Other information technology and computer service activities
J63*	Information service activities
J6310	Data processing, hosting and related activities; web portals
J6311	Data processing, hosting and related activities
J6312	Web portals
J639*	Other information service activities
J6391	News agency activities
J6399	Other information service activities n.e.c.

*This level of NACE classification was considered as some companies in the dataset have no specific 4-digit classification in this class

Source: Eurostat, Ramon, Metadata Download, Statistical Classification of Economic Activities in the European Community, Rev. 2 (2008) (NACE Rev. 2)

Table A2 - Number of companies in the pre-selected NACE codes (List in Table A1), in European countries

In bold: the 18 countries under analysis

	selected 4digit NACE codes (see Table A1)		
European countries	**industry 26-27**	**services 61-63**	**Total**
United Kingdom	**5,963**	**178,213**	**184,176**
Germany	**6,759**	**111,090**	**117,849**
Poland	2,900	90,582	93,482
Italy	**5,039**	**88,131**	**93,170**
France	**2,284**	**87,632**	**89,916**
Netherlands	**1,449**	**73,201**	**74,650**
Spain	**2,333**	**50,959**	**53,292**
Sweden	**1,545**	**41,763**	**43,308**
Czechia	3,247	35,956	39,203
Hungary	1,211	31,715	32,926
Belgium	**393**	**28,859**	**29,252**
Romania	756	18,088	18,844
Slovakia	1,154	16,841	17,995
Austria	**556**	**15,618**	**16,174**
Denmark	535	13,187	13,722
Portugal	**334**	**12,836**	**13,170**
Greece	427	11,925	12,352
Bulgaria	417	10,992	11,409
Norway	**271**	**10,808**	**11,079**
Ireland	**24**	**10,650**	**10,674**
Finland	**549**	**7,757**	**8,306**
Slovenia	**362**	**7,340**	**7,702**
Switzerland	1,232	5,303	6,535
Latvia	**181**	**5,807**	**5,988**
Croatia	**474**	**5,160**	**5,634**
Lithuania	**137**	**5,450**	**5,587**
Estonia	**116**	**4,401**	**4,517**
Serbia	823	2,824	3,647
Luxembourg	10	2,070	2,080
Cyprus	0	1,272	1,272
Iceland	14	1,205	1,219
North Macedonia	20	1,144	1,164
Malta	6	1,017	1,023
Bosnia and Herzegovina	65	956	1,021
Turkey	0	0	0

Source: Authors' elaboration Eurostat, Annual detailed enterprise statistics - 2017

Table A3 - Companies by size, for the pre-selected NACE codes (List in Table A1), with and without the 'full overview' field

	Full overview yes/no *absolute values*			**Full overview yes/no** *percentage values (by column)*			**Full overview yes/no** *percentage values (by row)*		
Company size	no	yes	Grand Total	no	yes	Grand Total l	no	yes	Grand Total
Very Large	68	2186	2254	0.04	12.85	1.10	3.02	96.98	100.00
Large	677	7885	8562	0.36	46.36	4.16	7.91	92.09	100.00
Medium sized	19484	5523	25007	10.33	32.47	12.16	77.91	22.09	100.00
Small	168414	1414	169828	89.28	8.31	82.58	99.17	0.83	100.00
Total	**188643**	**17008**	**205651**	**100.00**	**100.00**	**100.00**	**91.73**	**8.27**	**100.00**

Source: Authors' elaboration based on data from Amadeus, downloaded on 30.09.2019

As shown in the Table above, Full Overview was available for 97% of very large companies, 92% of large companies, almost 22% of medium-sized companies, and only 1% of small companies, making a total of 17,008 companies, or 8.3%. Despite this limitation, we decided to use the data for our research analysis, assuming that small companies, each performing a very specific activity, are significantly represented in the sample of observations with full overview, while larger companies, which include a greater variety of activities, both within each company and across the set of similar sized companies, are well represented by the information coverage in their full overview.

Table A4 List of 16 NACE codes resulting from Step 1 - Text analysis of full overview to identify the NACE codes-relevant to IoT domains

NACE code	NACE Description
2620	Manufacture of computers and peripheral equipment
2630	Manufacture of communication equipment
2640	Manufacture of consumer electronics
2651	Manufacture of instruments and appliances for measuring, testing and navigation
2731	Manufacture of fibre optic cables
2732	Manufacture of other electronic and electric wires and cables
2733	Manufacture of wiring devices
6110	Wired telecommunications activities
6120	Wireless telecommunications activities
6130	Satellite telecommunications activities
6190	Other telecommunications activities
6201	Computer programming activities
6202	Computer consultancy activities
6203	Computer facilities management activities
6209	Other information technology and computer service activities
6311	Data processing, hosting and related activities

Table A5 - First 50 characteristic terms of each IoT domain

[IoT domain number – *expert label* – characteristic terms sorted by decreasing order of test-value]

IoT domain 1– Software & data processing

software, computer, data processing, information technology, provision of computer, computer programming service, consulting service, management, computer hardware, modifying, advice, consulting, solution, supporting software, writing, provision, expertise, computer system, programming, designing computer system, software development, data processing facility, operation of clients, planning, technical, communication technology, exceptional domain knowledge, computer programming, information, financial service, service, training, computer system design, provision of information, support, finance, computer software design, streaming service, application hosting, mainframe facility, professional, modification of custom, hosting, technology solution, insurance, custom software, application service provisioning, variety of additional service, business intelligence, consultancy.

IoT domain 2 - Telecommunication

provision of telecommunication, voice, telecommunication service, satellite, transmission facility, broadband, radar station operation, satellite tracking, communication telemetry, network, data communication, satellite system, facility, communication service, internet, communication, telecommunication, fixed, data service, internet access, microwave, service, mobile, carrier, provision of communication, transmitting telecommunication, broadband internet, receiving telecommunication, call, internet service, transmission of voice, mobile phone, landline, terrestrial communication system, wireless broadband, distance, single technology, telecommunication applications, worldwide telecom, internet network, local phone, satellite broadband internet, satellite terminal station, reselling, maintaining switching, traditional local telephone service, competitive local telephone service, digital tv, provision of telephone, combination of technology.

IoT domain 3 – Manufacturing of telecom equipment

mechanical accessory, coaxial, gearbox, manufacture, antenna, equipment, radio, product, data communication equipment, loudspeaker, accessory, manufacture of communication, cordless telephone, receiving antenna, system, router, telephone answering machine, receiver, switching equipment, amplifier, transmitter, television, cable television equipment, wire telephone, television studio, broadcasting equipment, audio, mobile communication equipment, transmitting, communication equipment, telephone, pager, gateway, cellular phone, production, alarm, bridge, wireless communication equipment, component, intercom, television broadcast, speaker, mast, telecommunication product, fire, military, lan modem, sale of communication, detector, device.

IoT domain 4 – Manufacturing of electronic components

convenience, seal, strict, terminal block, trunking, fiber optics, aluminium, cable, connectors, switch, manufacture, electrical, wiring device, conductor, copper, lamp, electric, plug, socket, cord, product, voltage, relay, insulated, insulated wire, power cable, power, transformer, electrical equipment, component, cable product, outlet, lighting, coaxial cable, fuse, voltage cable, circuit breaker, rubber, production, cutout, connector, telecommunication cable, instrumentation cable, bare, panel, cabinet, accessory, control cable, electric wire, cable assembly.

IoT domain 5 - Manufacturing of measure instruments

nautical system, measuring, manufacture, instrument, sensor, measurement, equipment, product, temperature, gauge, meters, navigation, pressure, water, laboratory, control, valve, instrumentation, gas, manufacture of instrument, system, detection, production, measuring instrument, appliance, precision, calibration, analyzer, industrial, controlling device, navigational, electromedical, control instrument, component, thermometer, tester, probe, laser, heating, navigating, machine, weighing, physical property testing equipment, mechanical, aeronautical, meter, vibration, metrology, analytical instrument, test equipment.

Table A6– Characteristic categories of variable *IoT domain-and-Company-size*, by cluster of NUTS3 regions

Characteristic categories are sorted in decreasing order of test-value within each cluster

Cluster of NUTS3 regions & Characteristic categories	NUTS3 regions in the cluster	Test-value	% of category in the cluster	Number of companies in each category in the dataset
Cluster 1	**149**			
1 Soft. Data process._Large		47.59	20.53	5162
5 Manuf. Measur. Instr._Mediumsized		37.12	23.02	2602
5 Manuf. Measur. Instr._Small		33.99	15.77	4559
3 Manuf. Telecom. equip._Small		31.95	17.13	3427
3 Manuf. Telecom. equip._Mediumsized		29.64	30.94	989
1 Soft. Data process_Mediumsized		27.94	8.11	15379
5 Manuf. Measur. Instr._Large		21.74	28.21	631
1 Soft. Data process_VeryLarge		20.33	19.14	1139
3 Manuf. Telecom. equip._Large		15.33	29.55	291
5 Manuf. Measur. Instr._VeryLarge		11.43	30.52	154
4 Manuf. Electron. Comp._Small		11.30	15.36	560
2 Telecommunication_Large		10.46	10.35	1208
4 Manuf. Electron. Comp._Mediumsized		8.66	13.74	422
3 Manuf. Telecom. equip._VeryLarge		7.64	27.71	83
2 Telecommunication_VeryLarge		6.82	10.74	475
4 Manuf. Electron. Comp._Large		5.36	14.10	156
2 Telecommunication_Mediumsized		4.07	5.48	1789
Cluster 2	**227**			
5 Manuf. Measur. Instr._Small		64.55	26.85	4559
3 Manuf. Telecom. equip._Small		51.26	24.54	3427
1 Soft. Data process._Large		23.84	9.09	5162
2 Telecommunication_Large		18.20	14.24	1208
5 Manuf. Measur. Instr._Mediumsized		16.51	8.99	2602
5 Manuf. Measur. Instr._Large		13.25	14.42	631
3 Manuf. Telecom. equip._Mediumsized		11.65	10.21	989
4 Manuf. Electron. Comp._Small		8.69	10.18	560
1 Soft. Data process_VeryLarge		8.33	7.29	1139
3 Manuf. Telecom. equip._Large		7.89	12.71	291
2 Telecommunication_Mediumsized		7.33	5.70	1789
2 Telecommunication_VeryLarge		5.61	7.58	475
3 Manuf. Telecom. equip._VeryLarge		5.08	15.66	83
5 Manuf. Measur. Instr._VeryLarge		4.90	11.04	154
Cluster 3	**228**			
5 Manuf. Measur. Instr._Mediumsized		66.83	36.82	2602
5 Manuf. Measur. Instr._Small		48.09	17.24	4559
5 Manuf. Measur. Instr._Large		21.17	21.87	631
3 Manuf. Telecom. equip._Small		15.70	6.57	3427
3 Manuf. Telecom. equip._Mediumsized		14.85	11.32	989
5 Manuf. Measur. Instr._VeryLarge		10.18	21.43	154
1 Soft. Data process_Mediumsized		9.62	3.04	15379
1 Soft. Data process._Large		7.80	3.58	5162
3 Manuf. Telecom. equip._Large		4.68	6.87	291
4 Manuf. Electron. Comp._Mediumsized		3.97	5.21	422
1 Soft. Data process_VeryLarge		3.77	3.69	1139
4 Manuf. Electron. Comp._Small		3.65	4.46	560

Cluster 4	**95**			
4 Manuf. Electron. Comp._Mediumsized		31.92	41.23	422
4 Manuf. Electron. Comp._Small		27.94	28.21	560
3 Manuf. Telecom. equip._Small		23.54	7.24	3427
3 Manuf. Telecom. equip._Mediumsized		21.38	13.65	989
5 Manuf. Measur. Instr._Mediumsized		19.14	6.76	2602
5 Manuf. Measur. Instr._Small		17.06	4.52	4559
4 Manuf. Electron. Comp._Large		15.57	30.77	156
3 Manuf. Telecom. equip._Large		12.45	15.12	291
5 Manuf. Measur. Instr._Large		10.33	7.61	631
1 Soft. Data process._Large		8.89	2.58	5162
1 Soft. Data process_Mediumsized		4.94	1.51	15379
1 Soft. Data process_VeryLarge		3.32	2.28	1139
Cluster 5	**54**			
1 Soft. Data process._Large		67.22	26.21	5162
2 Telecommunication_Large		42.30	37.50	1208
1 Soft. Data process_VeryLarge		34.35	30.20	1139
2 Telecommunication_VeryLarge		28.42	41.26	475
3 Manuf. Telecom. equip._Small		24.28	11.29	3427
1 Soft. Data process_Mediumsized		15.07	4.62	15379
5 Manuf. Measur. Instr._Small		14.13	6.47	4559
2 Telecommunication_Mediumsized		14.06	9.33	1789
3 Manuf. Telecom. equip._Large		6.47	10.65	291
5 Manuf. Measur. Instr._Large		4.78	6.18	631
3 Manuf. Telecom. equip._Mediumsized		3.63	4.65	989
Cluster 6	**105**			
2 Telecommunication_Small		54.97	37.86	8030
2 Telecommunication_Mediumsized		22.81	35.27	1789
1 Soft. Data process_Mediumsized		17.41	18.73	15379
Cluster 7	**237**			
1 Soft. Data process._Small		173.43	85.55	131125
Total	***1095***			***192974***

Source: Authors elaboration based on data from Amadeus, downloaded on 30.09.2019

Table A7 – NUTS3 by IoT systems (leaders, co-designers, suppliers) and cluster

For each cluster: only NUTS3 with at least 2% weighting are listed

IoT Systems	Cluster of NUTS3 by IoT domains&size	NUTS3	% of NUT3 in the Cluster
leaders	cl-3	UKI31	15.67
		FR105	11.04
		FR101	10.13
		UKI32	6.58
		UKJ11	5.45
		UKJ25	4.72
		UKI42	3.70
		UKI43	3.44
		UKI44	3.25
		UKJ37	2.97
		DE712	2.58
		UKD33	2.51
		UKI33	2.27
		UKI75	2.16
	cl-5	DED21	2.95
		DE113	2.61
		ITC33	2.11
co-designers	cl-6	UKH12	4.70
		UKH23	4.30
		UKJ14	2.63
	cl-7	ITC4C	11.68
		ITI43	6.82
		DE300	5.73
		DE212	4.23
		ITC11	3.63
		DE600	3.11
		DE21H	2.59
suppliers	cl-1	SE110	10.52
		FI1B1	4.57
		EE001	4.37
		SE232	3.56
		NL329	3.38
		PT170	3.25
		LV006	2.68
		NL310	2.64
		SE224	2.47
		BE100	2.41
		NO011	2.32
	cl-2	ES300	25.41
		ES511	13.90
		IE061	9.70
		ES523	3.91
		ES618	2.19
	cl-4	ITC4D	5.25
		ITH55	4.99
		ITH36	4.73
		ITC47	4.58
		ITC41	4.01
		ITC46	4.01
		ITH32	3.86
		ITH31	3.60
		ITH54	3.40

ITH34	2.88
ITF47	2.83
ITH53	2.57
ITH52	2.16
ITI32	2.06

Figure A1 –Dendrogram NACE2 codes classification, matrix <28 NACE codes × 7440 Terms>

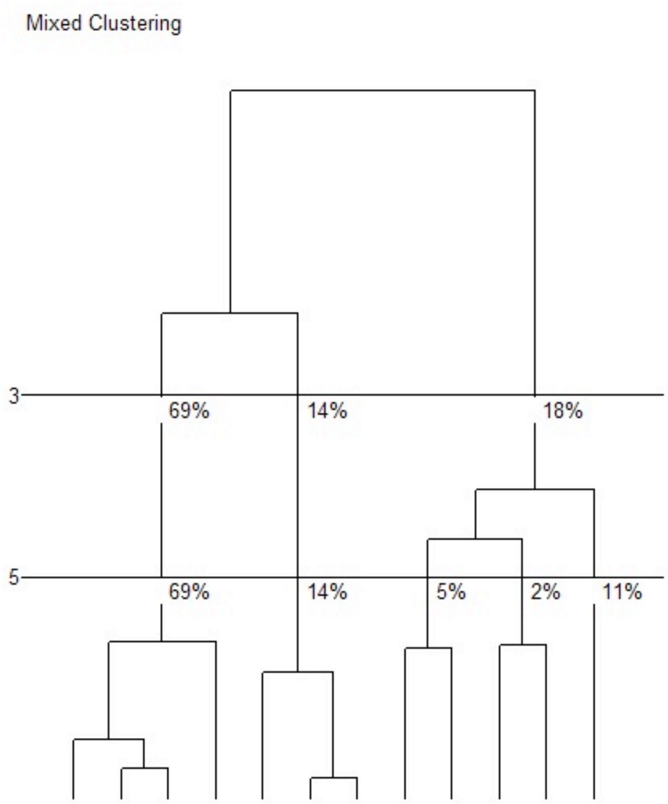


Figure A2 – NUTS3 regions classification dendrogram of the hierarchical cluster analysis, <NUTS3 × IoTdomains-and-Company size> matrix (1095 x 20)