

FRAMEWORK TO EVALUATE INTERNET USE AND DIGITAL DIVIDE IN FIRMS

Michela Serrecchia *

*Istituto di Informatica e Telematica, CNR **
*Via G. Moruzzi, 1 - 56124 Pisa, Italy **

Maurizio Martinelli *

*Istituto di Informatica e Telematica, CNR **
*Via G. Moruzzi, 1 - 56124 Pisa, Italy **

ABSTRACT

This article illustrates a framework that aims to examine and analyze Internet use and the resulting Digital Divide in firms. The framework is based on the extraction of .it domain names registered by Italian companies. The analysis includes a phase of automatic gathering of information of the firms registering a domain name and a quantitative and qualitative evaluation of firms that are starting to use the Internet. This approach provides a tool that can support the improvement of Digital Divide analysis, referring both to the territorial area of competence of the firms concerned, and to their socio-economic characteristics. This methodology can also be applied to other countries for the benefit of the scientific community, government institutions and the public.

KEYWORDS

Digital Divide, Internet use, Firms, domain names.

1. INTRODUCTION

As in the rest of the world, Europe has also recognized the Internet as an exceptional catalyst of reference for the growth of creativity, collaboration and innovation. Not only individuals but also companies are starting to use the Web to exploit the great potential of the Internet. The advantages for businesses provided by the Internet are not only linked to the sale of products and services (direct advantages), but can also be indirect (Hansons, 2000). For example, among the most important of these are reduced costs, image consolidation, greater customer loyalty, and a wider diffusion of products offered by the company. They are referred to as "indirect" since they do not lead directly to sales and do not generate immediate profits. However, they are important since they will probably be the greatest benefits to businesses offered by the Internet. Based on these arguments, in this article we propose a framework that aims to optimize Internet use analysis and the potential Digital Divide, referring to the socio-economic characteristics of individual companies and the territorial area in which they operate. The Digital Divide can be defined as the difference between those who possess the material and cultural conditions to exploit the new technologies, and those who do not, or those who lack the crucial ability to adapt to the rapid continual change that characterizes the Internet today (Warschauer, 2001). The Digital Divide can occur either as a "local" (within a given country) or "global" (between developing and industrialized countries) phenomenon. So the Digital Divide can be considered at three levels of analysis: - the individual level, the organizational level, and the global level (Dewan and Riggins, 2005). Therefore, this paper proposes a methodology to analyze Internet use among businesses and the resulting Digital Divide (at organizational level), by using as an endogenous indicator the number of domain names. Using endogenous indicators (such as Internet hosts or domain names) compared to other exogenous indicators (such as questionnaires) has several advantages. First of all, the collection of

data, being carried out in an automatic way, ensures that the examined sample is huge and therefore representative of reality. In addition, endogenous indicators allow good geographical characterization of the phenomenon because they are based on data that enable differentiation of users on a national, regional and provincial level (Diez-Picazo, 1999).

The proposed framework therefore includes: 1) extraction of the firms with their intrinsic features in a specific territory; 2) verification that firms have or have not registered a domain name through the identification of their VAT number or numerical tax code (data that companies are obliged to provide upon registration of a domain name), and 3) analysis of these data and quantitative and qualitative evaluation of the individual profile of the firms that register a domain name. Our study aims to obtain for the first time, automatically, a global characterization of the companies that register a domain name. To achieve this objective a definite improvement is necessary in the provision of data, in order to overcome the limitations due to non-availability of comprehensive data. The availability of data, such as the legal status, the number of employees and the corporate capital of firms that register a domain name will make it possible to carry out a research capable of providing an analysis of the phenomenon and an interpretative framework. This can be a basis for reflection, consultation and intervention for the scientific community, government institutions and the public. However as it is not possible to identify precisely the characteristics of the firms that register domains automatically (businesses at the time of registration of a domain name are not required to give information about their characteristics, such as social capital, etc.), it is necessary to consider access to appropriate information coming from other sources. Consequently, during the first phase all information concerning the companies located in a specific area must be extracted in an automatic way so that the sample is exhaustive and also representative of reality. The identification of companies that are part of a specific territory by means of their VAT number or numerical tax code makes it possible to establish if these companies have registered a domain name. In this case, from the institutional databases, it is also possible to obtain a direct automatic extraction of the features of such firms, without the use of questionnaires, which would limit the provision of data, resulting in a distortion of the results. The last phase involves the analysis of Internet use and the potential Digital Divide, taking into account not so much the socio-economic features of the territories, as in previous studies (Serrecchia et al., 2007), but primarily the individual profile of the single companies, related to the social context in which they operate.

Our framework has been tested by evaluating the companies of the Italian state by extracting the list of firms from the database maintained by *Infocamere* (Infocamere is the informatics company of the Italian Chambers of Commerce that manages all information pertaining to companies in Italy) and the list of domains registered in the database of the ccTLD ".it" Registry managed by the Institute of Informatics and Telematics (IIT), of the CNR in Pisa. Access to these databases (Infocamere and .it Registry) allowed us to obtain a complete list, respectively, of all the firms and of all firms that register a domain name.

2. RELATED WORK

The web can help businesses grow. The Internet expands and diversifies business opportunities, allowing expansion in international markets. According to the "Internet Factor" report of the Boston Consulting Group ([http://www.fattoreinternet.it/pdf/Fattore% 20internet-2011.pdf](http://www.fattoreinternet.it/pdf/Fattore%20internet-2011.pdf)), companies active online achieved a turnover increase of 1.2%, compared to a decrease of 4.5% for offline companies. 34% of small and medium-sized enterprises (SMEs) that are online registered an increase in their staff, while 65% reported benefits in terms of productivity thanks to the Internet. If we look at exports, 14.7% of online SME turnover comes from international markets. On the contrary, offline firms have an international turnover of only 4%. Martins and Oliveira (2008), by using a data set for Portuguese firms, showed that the significant drivers of intra-firm diffusion are: firm size, work skills, technological capacities and outsourcing partner usage. Tan and Ouyang (2004) examined the diffusion and the impacts of e-commerce in China based on the results of a large-scale survey conducted in 10 countries, including China. The survey focused on three sectors -manufacturing, wholesale/retail, and banking/insurance-. Their results showed that the three surveyed sectors, manufacturing, wholesale/retail, and banking/insurance, appear to follow quite different paths in adopting e-commerce. There are also differences between large firms and SMEs regarding e-commerce diffusion strategies and impacts. Greenstein and Price (2007) analyzed the rapid diffusion of the Internet across the United States over the past decade for both households and firms. They highlighted different economic

perspectives and explanations for the Digital Divide, that is, unequal availability and use of the Internet. In a previous study (Serrecchia et al., 2007) we analyzed Internet diffusion considering .it domain names registered by the firms in Italian regions and provinces. We verified the existence of a Digital Divide in terms of geography (Macro-areas, regional and provincial) and an initial interpretation of this phenomenon was made. This was done by focusing on the factors that determine the Digital Divide at the regional level by taking into account economic, cultural, demographic, technological and related to education variables in the provinces and regions of reference.

3. OUR APPROACH

The objectives of the framework proposed in this article can be summarized as follows: 1) create an experimental methodology for the evaluation of Internet use and Digital Divide in firms, 2) provide a tool for analyzing the behavior of single firms in the use of the Internet and their characterization in terms of geography, and 3) provide a tool for the monitoring of firms using the Internet and those who do not use it, 4) give a support to actions aimed at reducing a possible Digital Divide. To achieve these goals three basic steps were followed: the extraction of all the firms to be analyzed, the association of these data with the data available in the Database of the .it Registry and the analysis of these data in order to produce a quantitative assessment of the firms who are new to Internet. Domain names registered under the ccTLD “.it” are extracted from the Database of registrations managed by the IIT-CNR of Pisa, using automatic and semiautomatic procedures. The .it Registry subdivides the domains into seven categories (individuals, firms, professionals, non-profit institutions, public bodies, other institutions and foreign entities). Particular attention was paid to the registration of domain names by firms. Furthermore, these institutions were classified according to their geographical location (at the level of macro area, regional and provincial areas). However, to study the socio-economic characteristics of the firms that register a domain name, it is necessary to integrate the data available at the IIT with specific data of the single firms. These data are not always extractable automatically from the database managed by IIT. In the past, this problem was partially solved by using semi-automatic and/or manual procedures through the use of the Telemaco Database, run by Infocamere. Through the Telemaco Database, it was possible to trace only the legal status of firms by using the VAT number and/or the name of the organization, data that companies provide the IIT with at the time of registration of their domains. However, this procedure was time-consuming and led to the analysis of only a small sample of firms who register domain names, causing possible distortion in the results. In this study, however, in order to automatically obtain the greatest amount of information on registering firms, the data provided by Infocamere were cross-checked with the information on firms available on the databases managed by IIT. In fact, IIT has information such as VAT number and/or the social security number of companies registering a .it domain name. Infocamere on the other hand, in addition to the VAT number and/or social security number of all businesses in the territory concerned, provides IIT with other information such as corporate capital, the business sector in which they operate, number of employees, legal status, the province in which the company operates, the current company activity status (active, in liquidation, bankruptcy, suspended, inactive). Through the numerical tax code and/or VAT number, all the additional information provided by Infocamere was then associated to the companies that register .it domains. The analysis was carried out by monitoring the companies registered with the Italian Chambers of Commerce in the provinces of Milan, Naples and Florence. This monitoring activity is carried out not only to learn about the socio-economic characteristics of companies, but also to know if there are inequalities concerning the registration of a domain name between firms that reside in one province or another.

3.1 Quantitative evaluation of data

Once a new database has been obtained, resulting from the association of information extracted from the two Infocamere and .it Registry databases, we subdivided the firms that register domain names according to their size, legal status and field of activity. For size we used as a proxy corporate capital and number of employees. In addition, businesses were subdivided also according to legal status in order to ensure that only companies are considered in the analysis, thus eliminating public bodies and non-profit organizations (i.e. associations, foundations, committees, etc.). Finally, the firms that register domain names were divided

according to their economic activity, in particular according to the ATECO 2007 code. The ATECO code is an alphanumeric identifier which ranks companies according to their economic activity. The new classification of economic activities, ATECO 2007, was adopted by ISTAT (National Institute of Statistics, the Italian public research institution) from January 1 2008. In this study, firms that register domain names were divided according to the following ATECO 2007 codes: **C** - Manufacturing Activities, **I** - Service activities of Accommodation and Catering; **J** - Information and communication services; **K** - Finance and Insurance, **M** - Professional, Scientific and Technical services, **N** - Rental Activities, Travel Agencies, Support Services to Companies

3. RESULTS

In this study, more than 700,000 data items related to the provinces of Milan, Florence and Naples were analyzed. Inactive firms, in bankruptcy and those suspended from their activity, were eliminated from our research. In this analysis, therefore, only active firms in the three provinces indicated above were considered. Table 1 shows the results of the evaluation of Internet use and the resulting Digital Divide, measured as the difference in the registration of domain names by firms. This evaluation was measured for each field of activity analyzed.

Table 1. Internet use and Digital Divide in active firms in the provinces of Florence, Milan and Naples

Fields of economic activity (ATECO Codes)	Firms in the province of Milan (% .it registered domains)	Firms in the province of Florence (% .it registered domains)	Firms in the province of Naples (% .it registered domains)
C	21.32%	14.76%	8.26%
I	9.92%	15.06%	6.61%
J	31.86%	25.57%	22.68%
K	12.58%	6.78%	6.87%
M	26.63%	21.17%	18.34%
N	18.33%	15.14%	13.87%

Table 1 shows that among the fields of activity that were analyzed, the companies that use the Internet the most are those that carry out activities of Information and Communication services, ATECO code **J** (this category includes companies that engage in publishing activities, film, video and television program production, companies working in the field of telecommunications and related activities, companies that produce software). This trend occurs in all three provinces, even if the companies of the province of Milan emerge as the most innovative (almost 32% of the companies operating in that field registered a domain name). The second field that uses the Internet more is the field of Professional, Scientific and Technical activities, ATECO code **M**. This area comprises the companies engaged in legal and accounting activities, of architecture and engineering, business consulting services, companies engaged in research and development). The companies in the province of Milan, sector **M**, register 26.63% of domain names, followed by companies in the province of Florence with 21.17% and companies in the province of Naples with 18.34%. An interesting result that we obtained concerns sector **I**, Accommodation and Catering activities. Taking into consideration this field, we expected a higher percentage of registration of domain names. However, this lower percentage could be explained by the wide category that characterizes sector **I**. In fact, this area includes not only hotels, but also restaurants, bars and bakeries. In addition, unlike the other sectors analyzed, the firms in the province of Florence register several domain names (15.06%) compared to the firms in the provinces of Milan (9.92%) and Naples (6.61%). This result could derive from the fact that the province of Florence is considered a more attractive tourist destination compared to the provinces of Milan and Naples, therefore companies working in the field of accommodation and catering are more inclined to appear on the Internet in order to exploit both direct and indirect advantages.

The analysis shows that there is a difference in all areas in the registration of domain names in the three provinces. In general, except for the last mentioned case, the firms of the province of Milan, are more inclined to use the Internet, compared to those of Florence and Naples, with a wider difference in the case of

Naples. This result once again confirms the results obtained in previous studies (Martinelli et al., 2006), in which the South appears to be less inclined to use the Internet than in the North and Centre.

4. CONCLUSION

This paper introduces, for the first time, an experimental methodology for the evaluation of use and the resulting divide between firms, taking into account not only the geographical distribution of users, but also their individual profile. Internet use and the resulting Digital Divide were analyzed using as a proxy the endogenous indicator of domain .it names. These data were extracted from the database of the ccTLD .it Registry managed by the IIT. However, in order to have a greater availability of information, as data are not always extractable automatically from the database managed by IIT, we used the database of Infocamere, which contains all the information of Italian companies. In this analysis, the methodology was applied by taking into account active companies in the provinces of Milan, Florence and Naples.

The association of the information contained in the .it Registry database with those of the Infocamere database enabled us to have, for the first time, and automatically, an exhaustive sample of the firms that register domain names. This means a greater consistency and accuracy in the survey results. For the quantitative assessment of Internet use among firms according to their individual profile, we classified firms that register .it domain names according to legal status, corporate capital, number of employees and field of economic activity. There are six sectors considered, the first concerns manufacturing, while the others regard service activities. Results show that firms engaged in the six surveyed sectors appear to vary quite differently in adopting the Internet. Furthermore, for all sectors considered the firms of the province of Milan are more inclined to use the Internet, compared to those of Florence and Naples, with a wider divide in the case of Naples. This result once again confirms a Digital Divide at the territorial level, as previously reported for individuals. This approach, which aims to measure and analyze the use of the Internet, using data extracted on an automatic basis, could be a useful reference for other similar government agencies, which can use the same methodology for the evaluation of the use of the Internet among businesses, using as a proxy the endogenous indicator of domain names. This methodology aims, therefore, to consider the various aspects of Internet diffusion and the resulting divide, that is to say the phenomena underlying its territorial distribution. These are phenomena determined by the effects that information technology and communication have on the economy and society, and thus represent the relationship between the Internet, the economy and the society, which is ultimately the object of this research. In the future, the research aims to carry out a detailed analysis of the factors that determine the Digital Divide, taking into account not only socio-economic characteristics of the territories, as in previous studies, but also the individual profile of the companies.

REFERENCES

- Diez-Picazo, G.F., 1999. *An Analysis of International Internet Diffusion*. Ph.D. Thesis, MIT
- Hanson, W., 2000. *Principles of Internet Marketing*, South-Western College Publishing, Cincinnati, Ohio
- Greenstein, S. and Prince, J., 2007, Internet diffusion and the geography of the digital divide in the United States, in: Mansell R., Avgerou C., Quah D. and Silverstone, R. (eds), *Oxford Handbook of Information and Communication Technologies*, Oxford University Press, Oxford, U.K., pp. 168-195.
- Martins, Maria Fraga O. and Oliveira, T., 2008. Determinants of Information Technology Diffusion: a Study at the Firm Level for Portugal. *In Electronic Journal Information Systems Evaluation*, Vol. 11, No. 1, pp 27-34.
- Martinelli et al., 2006. Analysis of the Internet diffusion in the non-profit sector: the social digital divide in Italy. *In Scientometrics*, Vol. 66, No. 1, pp 155-170.
- Riggins, F. J. and Dewan, S., 2005. The Digital Divide: Current and Future Research Directions. *In Journal of the Association for Information Systems*, Vol. 6, No. 12.
- Serrecchia M., et al., 2007. ICT adoption by Firms by using Endogenous Metrics, in Ashis K. Pani & Amit Agrahari: (eds), *E- Procurement in Emerging Economies*, Idea Group Publishing: Hershey, PA., pp. 161-196.
- Tan, Z. and Ouyang, W., 2004. Diffusion and Impacts of the Internet and E-Commerce in China. *In Electronic Markets*, Vol.14, No. 1, pp. 25-35.
- Warschauer, M., 2002. Reconceptualizing the Digital Divide. *In first Monday*, Vol. 7, No. 7.