A fuzzy taxonomy for e-Health projects[1]

(work in progress)

Pierpaolo D'Urso

Department of Social Sciences, Sapienza - University of Rome, Rome Italy, pierpaolo.durso@uniroma1.it.

Corresponding author.

Livia De Giovanni

Department of Political Science, LUISS Guido Carli University, Rome Italy, ldegiovanni@luiss.it.

Paolo Spagnoletti

CeRSI-LUISS Guido Carli University, Rome Italy, pspagnoletti@luiss.it.

Abstract. Evaluating the impact of Information Technology (IT) projects represents a problematic task for policy and decision makers aiming to define roadmaps based on previous experiences. Especially in the healthcare sector IT can support a wide range of processes and it is difficult to analyze in a comparative way the benefits and results of e-Health practices in order to define strategies and to assign priorities to potential investments. A first step towards the definition of an evaluation framework to compare e-Health initiatives consists in the definition of clusters of homogeneous projects that can be further analyzed through multiple case studies. However imprecision and subjectivity affect the classification of e-Health projects that are focused on multiple aspects of the complex healthcare system scenario. In this paper we apply a method, based on advanced cluster techniques and fuzzy theories, for validating a project taxonomy in the e-Health sector. An empirical test of the method has been performed over a set of European good practices in order to define a taxonomy for classifying e-Health projects.

Keywords: e-health, healthcare, fuzzy clustering, imprecise evaluation scales, soft taxonomy.

1. Introduction

The use of Information Technology (IT) for supporting healthcare organizations in their activities is widespread. Health Information Systems (HIS) such as personal health records (Lafky et al., 2006), asynchronous healthcare communication systems (Wilson, 2003), Internet-based telemedicine and picture

[1] This work has been accepted for publication as D'Urso P., De Giovanni L., Spagnoletti P. (2012) A fuzzy taxonomy for e-Health projects, International Journal of Machine Learning and Cybernetics, 4(6), pp. xx, DOI 10.1007/s13042-012-0118-4

archiving communication systems (Menachemi et al. 2004) have been applied in healthcare to improve the capabilities of physicians and clinical staff and provided increased services to patients caregivers and citizens in general (Mantzana et al. 2008). In the last decade, e-Health has been introduced as an umbrella term, describing the combined use of electronic communication and information technology in the health sector, and also the use of digital data - transmitted, stored and retrieved electronically - for clinical, educational and administrative purposes, both at the local site and at distance (Mitchell 2000). The availability of new technological solutions together with the increasing need for better healthcare services and higher quality of life, raise the interest in this filed on both the demand side and the technology providers. Also governmental institutions are showing an increasing attention towards this field. For instance, e-Health is on the governmental agenda of all EU Members States (EC, 2009). Furthermore, such attention on e-Health investments has created a strong e-Health market with a wide range of applications that span from biomedical technologies to IT platforms supporting healthcare management decisions at all levels of the health system.

Notwithstanding the enthusiastic declarations of e-Health potential, the adoption of IT has been much slower in healthcare than it has been in other sectors such as banking and manufacturing (Bates 2005). Cost is often cited as the primary reason for the slow rate of e-Health adoption, followed by the lack of methods for evaluating the actual benefits provided to the stakeholders (i.e. financial, non-financial, tangible, intangible), and privacy and security concerns (Dixon 2007). In fact, decision makers can benefit from the availability of domain specific evaluation frameworks supporting ex ante and ex post decisions at different levels (i.e. strategic, organizational, group, individual) for different types of systems. For instance at the individual level, where the main issue is the lack of awareness among medical and nursing personnel, Fitterer et al. (2011) have recently proposed a taxonomy for multi-perspective assessment of the value of HIS based on the Unified Theory of Acceptance and Use of Technology (UTAUT). A more comprehensive framework for evaluating HIS has been introduced by Yusof et al. (2008a) building on previous models of IS evaluation, which measure the fit among technological, human and organizational dimensions. In order to validate the proposed framework the authors present a case study on the adoption of a digital imaging software that is used to capture the eyes images of patients with diabetes in UK. In the first case the evaluation framework is focused on a specific perspective (i.e. individual), in the second case it has been validated in a specific type of system.

Among the problems in effectively designing and evaluating the impact of e-health, there is the loose terminology adopted in this field by researchers and practitioners (Barlow et al. 2006). For instance, terms like “telecare”, “telehealthcare”, “telemonitoring” and “telemedicine” are indeed all used interchangeably and have different meanings to different people (Nagendran et al., 2000). With the objective of addressing the inconsistency of terms and definitions used in the HIS literature, Yusof et al. (2008b) provide their classification of different types of HIS: Patient centered information systems, Administrative information systems, Clinical information systems, Radiology information systems, Laboratory information systems, Pharmacy information systems, Telemedicine, Clinical decision support systems, and Hospital information systems. This taxonomy is grounded on concepts and definitions from eleven articles which are focused on one or more HIS classes and adopt different lens or perspectives. Therefore the descriptions of the resulting classes are not homogeneous and refer to either the business processes supported, the organizational units involved, the target users, or the software functionalities.

In this paper we investigate the nature of e-Health systems in terms of their constituent components (i.e. IT capabilities, IT applications, IT platforms) and by analyzing how these components effectively combine with organizational processes and actors to build a successful e-Health system. The outcome of the study is an empirically grounded taxonomy of e-Health projects which addresses the strategic level of e-Health decision making. Our assumption is that a better understanding of the e-Health solution space will provide input to the policy definition and project prioritization processes.

In order to achieve this goal, we analyze with traditional and advanced clustering techniques a dataset related to 94 successful e-health project cases. The dataset is the outcome of an iterative evaluation process in which an expert panel has classified each case with respect to its focus on the different building blocks of the overall healthcare system. To reflect either the intrinsic imprecision of this evaluation process or the inherent subjectivity of the evaluation expressed by the experts, the scale of fuzzy numbers has been used. Furthermore, three clustering analyses based on crisp and fuzzy techniques are comparatively adopted for performing a cross-case synthesis on this dataset.

The paper is structured as follows. In Section 2 we illustrate the theoretical background of the paper. In Section 3 we describe the methodological framework for the statistical analysis. In section 4 we describe the data sources, the data collection and the data analysis process. In Section 5 we present the results of the analysis on the considered dataset. In Section 6 we summarize findings, implications and further research.

2. Theoretical background

It is widely accepted that e-Health can address many of the problems currently faced by the healthcare systems, improving quality of care, increasing efficiency of healthcare work, assuring healthcare services more accessible and better effectiveness of medical interventions and patient care (Fitterer et al. 2011). The benefits of successful e-Health initiatives can be measured in terms of clinical outcomes such as costs reduction (i.e. fewer medication errors and adverse drug effect); improved efficiency in patient care delivery (i.e. number of consultations and length of waiting lists); morbidity (the rate of incidence of a disease) and mortality (death rate). Apart from these quantitative measures, e-Heath systems have also an impact in terms of quality of care, on patient care and communication, such as change in communication style and facilitation of information access (Yusof et al. 2008a). The huge number and the variety of elements in the healthcare arena make difficult to both design new solutions and evaluate their outcomes from multiple perspectives and levels of analysis. In fact, healthcare actors are providers (e.g. medical and nursing professionals and related management personnel), supporters (e.g. suppliers, software providers), healthcare acceptors (e.g. healthy people, patients and their relatives) and controllers (e.g. public institutions, insurance companies) (Mantzana et al. 2007).

With these premises, e-Health solutions can facilitate the transforming of healthcare processes for the benefit of both the patients and the healthcare system by providing a wide variety of solutions which support the whole lifecycle of the health assistance process: health promotion, diagnosis, therapy, rehabilitation or long-term care. e-Health can also underpin support activities like management and administration, logistics and supply of health-related goods, facilities management as well as public health, continued medical education, or medical research and clinical trials.

For the purposes of this paper we define an e-Health system as a set of interrelated IT capabilities implemented to aid in enhancing the efficiency and effectiveness of the healthcare actors in performing their functions and attaining their objectives. An e-Health project is a set of coordinated actions for adding new IT capabilities to an existing healthcare system. In order to better understand the complexity of an e-Health system, it is useful to analyze it along three dimensions: the value of IT, the actors involved and a set of interconnected IT capabilities (figure 1). While previous works have addressed the first two dimensions (Yusof et al. 2008a, Mantzana et al. 2007), in this paper we concentrate on the third dimension by analyzing IT capabilities and their interrelationships to identify classes of e-Health projects.

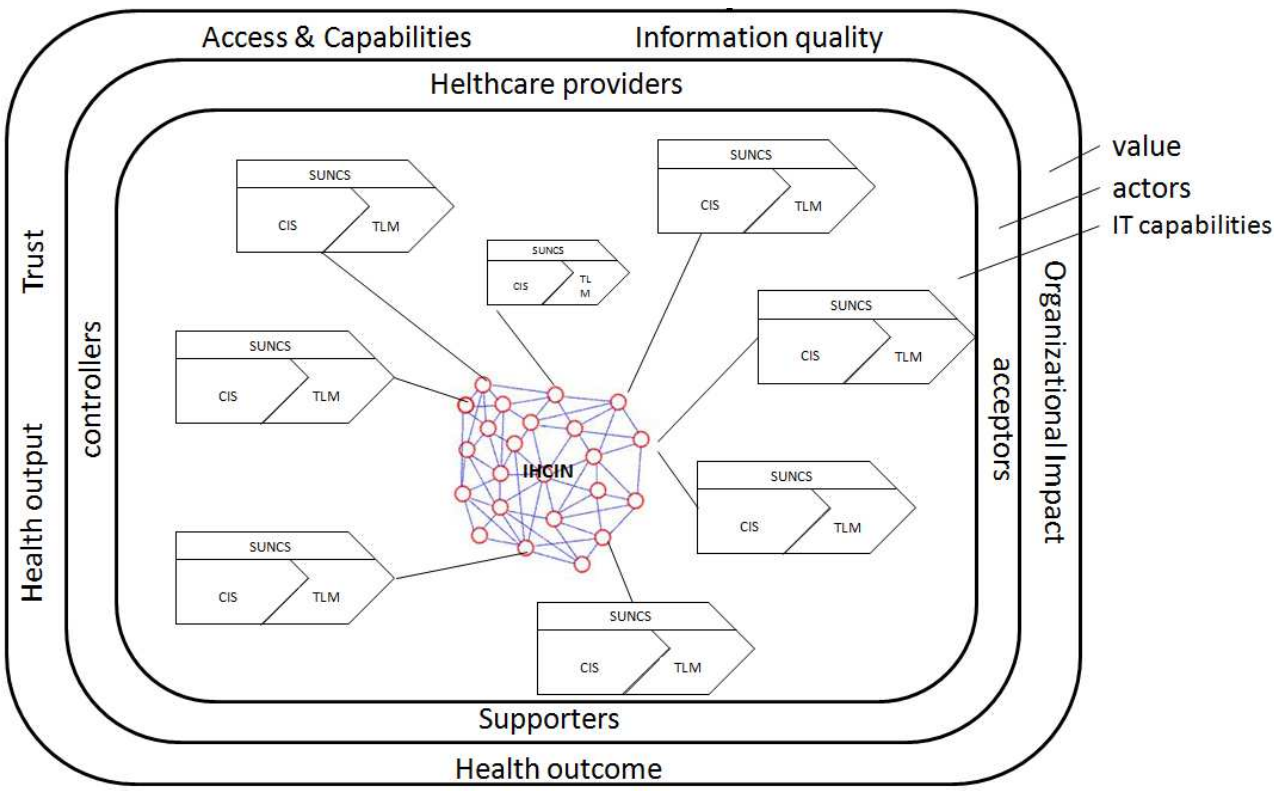

Figure 1. e-Health system

2.1 IT capabilities for e-Health

The above mentioned scenario of e-Health systems emphasizes the key role of IT capabilities which are seen as the elementary components of a complex IT system. An IT capability is defined as the possibility and/or right of the user or a user community to perform a set of actions on a computational object or process (Hanseth and Lyytinen 2010). According with this definition an IT capability is defined and managed locally by single or a small group of designers that typically control its evolution locally. IT capabilities play an important role in IT systems design since they are the basic elements upon which more complex forms of IT artifacts are constructed. From this perspective, IT applications are suites of IT capabilities developed to meet a set of specified user needs within a selected set of communities with a bounded scope. As a further level of complexity, IT platforms are intended as applications with a heterogeneous and growing user base and whose design context is not fixed due to the need of satisfying multiple generic functional specifications based on a mix of IT capabilities.

Since e-Health projects can be focused on many different aspects of the healthcare system, it is important to identify the building blocks and their relationships in order to support strategic decision making in this complex scenario. In fact healthcare information systems are understood to not be standalone entities, but integrated with other information systems and communication technologies, as well as with other technical and non-technical elements (Aanestad and Jensen 2011). The metaphor of “cultivation” and the notion of “installed base” and “modularization” have been introduced for defining design principles of complex IT infrastructures (Ciborra et al., 2000). In this view e-Health projects can be seen as gradual and step-wise transitions in which existing infrastructures cannot be changed instantly but have to be implemented in a gradual fashion and proceed through changing elements or sub-networks (Hanseth and Aanestad, 2003).

The concepts of IT capability, IT application and IT platform are useful for generalizing components of the e-Health system. When applied to the e-Health domain, IT capabilities reflect the distributed structure of the healthcare system in which multiple organizations are tangled at different levels. These capabilities refer to the support of IT solutions to internal core and secondary processes, to remote service delivery and to interorganizational coordination. According with this distinction we define four categories of IT capabilities, whose definitions are provided in Table 1: Clinical Information Systems (CIS), Secondary Usage Non Clinical Systems (SUNCS), Telemedicine (TLM) and Integrated Health Clinical Information Networks (IHCIN).

The contribution of this paper is to define an empirically grounded taxonomy of e-Health projects by testing the following research proposition: e-Health projects can be classified through a taxonomy that takes into account the hierarchical nature of relationships between IT capabilities, IT applications and IT platforms (*Proposition 1*).

Categories	Description
Clinical Information System (CIS)	specialized tools for health professionals within healthcare institutions (e.g. hospitals) tools for primary care and/or for outside care institutions such as general practitioner and pharmacy information systems
Secondary Usage Non-clinical Systems (SUNCS)	systems for health education and health promotion of patients/citizens, such as health portals or online health information services specialised systems for researchers and public health data collection and analysis, such as biostatistical programs for infectious diseases, drug development and outcomes analysis support systems, such as supply chain management, scheduling systems, billing systems, administrative and management systems, which support clinical processes but are not used directly by patients or healthcare professionals
Telemedicine (TLM)	personalized health systems and services, such as disease management services, remote patient monitoring (e.g. at home), teleconsultation, telecare, telemedicine and teleradiology
Integrated Health Clinical Information Network (IHCIN)	distributed electronic health record systems and associated services such as e-prescriptions or e-referrals

Table 1. IT capabilities for e-Health

2.2 Toward an e-Health evaluation framework

The problem of evaluating impacts of IT has represented one of the top issues of concern for both managers and researchers in the Information Systems (IS) domain. Reasons can be found in the complexity of performing an effective evaluation process, in the variety of implications for problem diagnosis and planning, and in the reduced uncertainty (Hawgood and Land, 1988). Several evidences suggest that organizations normally carry out some form of evaluation as part of a feasibility study or investment appraisal, typically using traditional cost-benefit analysis (Smithson and Hirschheim, 1998; Irani and Love, 2008). Among the reasons why organizations appraise their IS investments, there are the need to make comparisons between different projects, to justify investment requests by management, to control expenditures, benefits, risks, development and implementation of projects (Irani and Love 2002). However, managers still struggle with identifying and measuring the strategic implications of IT/IS. The complexity of new technologies asks for comprehensive but understandable methodologies to give a proper solution to project justification and assessment problems (Irani et al 2002). Some authors claim that the evaluation process should take into account both social and technical entities that an organization is confronted with when adopting IT (Smithson and Hirschheim , 1998). Evaluations should be tailored to the needs of individual organizations based on their environment, the context of the evaluation, the object to be evaluated, and the stakeholder's view. In this sense, continuous formative evaluation approaches are finalized to examine the strategic value of systems, and to assess their effectiveness in terms of system use, cost-benefit analysis, comparison with objectives and user satisfaction (Smithson and Hirschheim, 1998). In order to fulfill this goal, the development of conceptual tools and methods for analyzing context related aspects of IT systems implementation, adoption and use are needed (Stockdale et al., 2008).

With respect to the e-Health domain, several studies are available providing information on issues and trends in project implementation. These studies are often commissioned by governmental institutions to consulting companies and research centers in order to support decision makers in their difficult tasks such as policy definition and priority identification in e-Health project investments (i.e. EC 2008). The great variety of possible e-Health initiatives makes difficult to compare cases which are very different in nature in terms of content, context and process (Spagnoletti et al. 2011). Grouping cases through well defined categories in order to compare and analyze their characteristics and impact represents a common approach for performing benchmarks and comparative project evaluation.

The first step for developing an evaluation framework which supports strategic decision making in the e-Health domain, is to provide clear delineation of the uniformities of classes of phenomena to be evaluated through "systematics" (McKelvey 1982). McKelvey refers to "systematics" and to "the science of diversity" for addressing the subject of taxonomies and classification for organizations which are a prerequisite for investigating with a scientific method the fields of biology, zoology, and botany. Given the complexity of the e-Health phenomenon, in terms of IT value, actors, IT systems and their relationships, we apply "systematics" to develop a taxonomy of e-Health systems. A taxonomy is the most basic type of theory where no causal relationships are specified and no predictions are made. It is a conceptual tool for analyzing or summarizing salient attributes of phenomena and the relationships among phenomena. The relationships specified are classificatory, compositional, or associative, not explicitly causal (Gregor 2006).

Instead of building the taxonomy upon previous scientific contributions with a deductive approach, we apply the principles of grounded theory for deriving the taxonomy from empirical data with an inductive approach (Glaser and Strauss 1967). The resulting taxonomy takes into account specific characteristics of e-Health projects by summarizing the similarities found in discrete observations.

Given the incremental nature of complex systems design, it is important to identify classes of homogeneous projects for better understanding the nature of the installed base. This result can be achieved through a careful analysis of successful implementations in which the above mentioned IT capabilities represent the atomic components or building blocks (modules). The identification of these characteristics in successful e-Health projects can benefit from the adoption of a fuzzy approach for encompassing the limitations due to the intrinsic imprecision of IT capabilities definitions and the inherent subjectivity of the evaluation.

This leads to the definition of an additional research proposition to be tested in the empirical part of this paper: the application of fuzzy clustering techniques to fuzzy data allows a better identification of the installed base of e-Health projects (*Proposition 2*).

3. Methodological framework

3.1 Conceptual aspects

As remarked by Coppi et al. (2006) "vagueness may affect the information we use in these processes. In fact, the empirical or theoretical information (respectively, the data and the assumptions) we use in the process of knowledge acquisition is generally affected by uncertainty. This may stem from several sources. In the specific case of statistical reasoning, various features of uncertainty may be considered: (i) the uncertainty related to the link between the observed data and the *universe* of possible data; (ii) the imprecision in measuring the empirical phenomena; (iii) the vagueness connected with the use of linguistic terms in the description of the real world (e.g., when analyzing qualitative data); (iv) the (partial or total) ignorance concerning the values of a phenomenon in a specific observational instance or the validity of a given theoretical assumption (e.g., when adopting a Gaussian model for a stochastic quantity); (v) the imprecision deriving from the granularity of the terms utilized in the description of the physical world (Zadeh, 2005) (e.g., in a sociological investigation we may observe or analyze the variable "age of a person" in terms of *granules* consisting of single years, or intervals of five years, or ordered classes such as "young," "middle age," "old"; an increasing uncertainty is associated with these different *granulations*)."

In this paper, we shall specifically focus on the *vagueness* of the data and on the *uncertainty* in the assignment process in a clustering framework both treated from a fuzzy viewpoint.

In particular, in order to define a taxonomy of e-Health projects by means a clustering approach based on the analysis of empirical information instead of being of the result of a conceptual theory, we consider a cluster analysis (Coppi et al., 2012) formalized in a fuzzy theoretical framework (Bezdek, 1981; Bezdek et al., 1984). In particular, we consider the case in which the empirical information is fuzzy (D'Urso, 2007). Then, we have the situation in which the theoretical information (i.e. the clustering model) is fuzzy and the empirical information (i.e., represented by linguistic terms, qualitative data) is also fuzzy: we have a complete fuzzy information (D'Urso, 2007). Notice that, traditional clustering techniques could be utilized (see., e.g., Graaff, Engelbrecht, 2011; Guo et al., 2011; Liang, Song, 2011). However, we prefer a complete fuzzy clustering approach for the motivations shown in sections 3.1.1 and 3.1.2.

3.1.1 Motivations on the fuzziness of the clustering approach (fuzzy theoretical information)

For our study, we consider a fuzzy clustering approach for classifying e-Health projects. Fuzzy clustering is an overlapping clustering method which allows cases to belong to more than one cluster simultaneously as opposed to traditional clustering which results in mutually exclusive clusters (Bezdek, 1981).

In general, the adopted clustering model suggested by Coppi et al. (2012) inherits the several advantages of the fuzzy approach to cluster analysis. As remarked by Hwang et al. (2007), **"**the fuzzy clustering algorithm is attractive in the context of the proposed method because it is easily compatible with the distribution-free optimization procedure […]. Moreover, due to the difficulty of identifying a clear boundary between clusters in real world problems, the partial classification of fuzzy clustering appears more attractive than the deterministic classification of nonoverlapping clustering methods such as *k*-means (McBratney & Moore, 1985; Wedel & Kamakura, 1998). Furthermore, the fuzzy clustering approach offers other major advantages over traditional clustering methods. Firstly, the fuzzy clustering algorithm is computationally more efficient because dramatic changes in the value of cluster membership are less likely to occur in estimation procedures (McBratney & Moore, 1985). Secondly, fuzzy clustering has been shown to be less afflicted by local optima problems (Heiser & Groenen, 1997). Finally, the memberships for any given set of respondents indicate whether there is a second-best cluster almost as good as the best cluster—a result which traditional clustering methods cannot uncover (Everitt et al. 2001)". Furthermore, as remarked by Hwang et al. (2007), the concept of partial membership underlying the proposed clustering models (Zadeh, 1965) appears more appealing than that of the traditional clustering procedures (also see Wedel & Kamakura, 1998). For more details, see, e.g., Coppi et al., 2012 and D'Urso, 2007.

Another approach to partial membership is Latent Dirichlet Allocation LDA (Blei et al. 2003). It consists in a three level hierarchical Bayesian model which allows probabilistic generation of each item of a collection of discrete data. It has been successfully applied in the context of text modeling and document classification.

A fuzzy clustering approach to partial membership has been adopted to the problem at hand either for the soft modeling nature of the approach or for the fuzziness of the data (Section 3.1.2) besides the fuzziness of the classification. The description of the considered e-Health projects in fact is not directly used for classification but interpreted by experts for the 'fuzzy' assignment of each project to the four categories of IT capabilities described in Section 2.1, as shown in Sections 4.2, 4.3.

As remarked in Section 2, the great variety of possible European e-Health projects makes great difficult to metabolize the complex information connected to features of the e-Health initiatives and to compare cases which are very different in nature in terms of context, content and process. Thus, in this case, it is particularly useful to adopt a fuzzy approach for analyzing this typology of information. In particular, with respect to our specific empirical study, as we will see in Section 4, by considering a fuzzy clustering approach, we define a fuzzy (soft) taxonomy structure for e-Health projects. In this way, we build a flexible (soft) taxonomy characterized by a non rigid clustering structure in which each e-Health project can belong to more than one cluster with different membership degree (between 0 and 1). Each membership degree represents a measure of the level of uncertainty (vagueness) in the assignment process of each e-Health project to each class.

In particular, the two most important motivations justifying the utilization of a fuzzy approach for defining a soft taxonomy of e-Health projects are:

- *Sensitivity* in capturing the details characterizing the European e-Health projects which have been labeled as "good practices" by an expert panel of specialists in different areas of e-Health selected by the European Commission (EC 2008) (see Section 2). In fact, often the e-Health projects present "intermediate" or "quasi- intermediate" or, more in general, different features with respect to well-separated clusters and hence the traditional clustering approaches are likely to miss these underlying structures. On the contrary, the features of the e-Health projects, which are usually vague (fuzzy), can be naturally treated by means of fuzzy clustering. To this purpose, we can notice that all evaluations of the e-Health projects suggests thinking in terms of "degrees" of membership associated with given clusters rather than in terms of total membership versus non-membership. In fact, a traditional definition of clusters contrasts, for example, with the ambiguities presented when e-Health projects with "intermediate" or "quasi- intermediate" or different features may occur.
- *Adaptivity* in defining the "prototype" e-Health projects. This can be better appreciated when the considered e-Health projects do not differ too much from each other. In this case, the fuzzy definition of the clusters allows us to single out underlying prototypes, if these are likely to exist in the given set of e-Health projects.

As mentioned in Section 4, an e-Health project is a set coordinated actions for adding new IT capabilities to an existing healthcare system. The e-Health projects described in the EU database are unique cases since real projects are built upon existing IT infrastructures with their IT capabilities, Applications and Platforms.

This means that two projects belonging to the same class of the taxonomy (the same cluster) may differ in terms of the extent to which one or more IT capabilities are present. That is the installed base of IT capabilities influences e-Health project design and evaluation.

3.1.2. Motivation on the fuzziness of the data (fuzzy empirical information)

In machine learning and knowledge discovery, we usually analyze "precise" (non vague) data, typically exact results of observations and/or of measurements. However, in many real-life situations, the observations may be defined vaguely and measurements may be imprecise. Furthermore, in several fields of knowledge (such as evaluation studies, cognitive sciences, quality rating analysis, decision making, social and political sciences, medical diagnosis, marketing research, neurosciences, ergonomics, and so on), both scientific propositions and empirical data are often formulated in terms of natural language (Coppi et al., 2012). These formulations may be appropriately represented by fuzzy values. For instance, let us consider a set of persons, e.g. a population living in a given area. Each person, from a clinical viewpoint, can be characterized according to her/his "health state". This refers to the "normal" functioning of the various "aspects" of her/his organism. Generally, any "aspect" works correctly to a certain extent. We often use the notion of "insufficiency", related to different relevant functions of parts of the body (e.g. renal or hepatic insufficiency, aortic incompetence, etc.). Insufficiency (as referring to the various relevant above mentioned aspects) is a concept which applies in a certain degree to any individual (depending on several factors such as age, sex, previous illnesses, and so on). This may be expressed by means of a fuzzy value on a continuous standard scale (say, from 0=perfect functioning, to 10=complete insufficiency). Consequently, each individual can be more realistically characterized by a vector of fuzzy variables concerning "insufficiency" of various relevant aspects of her/his organism (Coppi, 2003).

As outlined by Sinova et al. (2012), "the imprecision underlying many available data from surveys, ratings, etc. can be properly formalized in terms of fuzzy values and, in particular, fuzzy numbers. The richness of the scale of fuzzy numbers (including real and interval values as special elements) allows us to cope with a wide set of imprecise data, as those mentioned above. Instead of modeling the type of data by means of either numerical or categorical data [e.g. Likert scales], which would be less accurate or expressive, the fuzzy scale integrates the manageability and diversity/variability of the numerical scale and the interpretability and ability to capture the imprecision of the categorical scale. Furthermore, fuzzy numbers become a flexible and easy-to-use tool which enables us to exploit the subjectivity that is often involved in perceiving and expressing the available information. They have a very intuitive meaning and potential users can friendly understand the required basic notions and ideas to manage fuzzy data."

As we can see in Section 3.2, we formalize the notion of fuzzy data by considering the concept of membership function.

In our empirical study, we define a fuzzy taxonomy for e-Health projects by applying a fuzzy clustering to a Likert-type evaluation scale introduced by an expert panel (i.e., the items are: perfect, good, medium, poor, bad). Then, in our case, we have the type of uncertainty (iii), i.e. we have the vagueness connected with the use of linguistic terms (quality evaluation scales).

To define a taxonomy for e-Health projects based on evaluation scales, we can treat such scales either as categorical (for which statistical methods are rather limited) or coded by and handled as integer numbers (integer coding usually not reflecting the real differences between distinct values, and not capturing the imprecision and subjectivity which is intrinsic to these responses) (Sinova et al., 2012). In our study, we suggest to use instead of Likert-type or integer scales, whenever it is reasonable and feasible, the scale of fuzzy numbers (see below Section 5). This scale enables us to reflect the intrinsic imprecision of the potential evaluations of the e-Health projects, combined with the inherent subjectivity of these evaluations expressed by the experts. In this way, the variability and diversity can be exploited more accurately in the taxonomy process (González-Rodríguez, 2012).

The iterative process for interpreting data and establishing the degree of presence of each IT capability has been carried out by a focus group of experienced researchers whose expertise varies in terms of number of years, level and nature of IT skills (i.e. telecommunication, database, etc.), and level of healthcare skills. With respect to the latter an e-Health expert can have experienced projects in a subset of the possible organizational levels involved. For instance he/she can be familiar with administrative processes of local health authorities, with patient centered applications at a regional level or with more focused applications of telemedicine. These different backgrounds influence the subjectivity of the evaluation which can be addressed by considering fuzzy data in the clustering analysis technique.

For more specific evidences on the usefulness of our clustering approach to define a taxonomy for e-Health projects see Section 5.

3.2 Mathematical formalization

A general class of fuzzy data, called *LR fuzzy data*, can be defined as follows:

$$\tilde{\mathbf{X}} \equiv \left\{ \tilde{x}_{ij} = (c_{1ij}, c_{2ij}, l_{ij}, r_{ij})_{LR} : i = 1,\dots,n;\ j = 1,\dots,p \right\}, \qquad (1)$$

where $\tilde{x}_{ij} = (c_{1ij}, c_{2ij}, l_{ij}, r_{ij})_{LR}$ denotes the *LR* fuzzy variable *j* observed on the *i*-th object , c_{1ij} and c_{2ij} indicate the left and right center and l_{ij} and r_{ij} represent the left and right spread.

For the *LR fuzzy data* (1), we can consider the following membership functions:

$$\mu(\tilde{u}_{ij}) = \begin{cases} L\left(\dfrac{c_{1ij} - \tilde{u}_{ij}}{l_{ij}}\right) & \tilde{u}_{ij} \le c_{1ij} \quad (l_{ij} > 0), \\ 1 & c_{1ij} \le \tilde{u}_{ij} \le c_{2ij}, \\ R\left(\dfrac{\tilde{u}_{ij} - c_{2ij}}{r_{ij}}\right) & \tilde{u}_{ij} \ge c_{2ij} \quad (r_{ij} > 0), \end{cases} \tag{2}$$

where $L(z_{ij})$ (and $R(z_{ij})$) is a decreasing 'shape' function from $\Re^+$ to [0,1] with $L(0)=1$; $L(z_{ij})<1$ for all $z_{ij}>0$, $\forall i,j$; $L(z_{ij})>0$ for all $z_{ij}<1$, $\forall i,j$; $L(1)=0$ (or $L(z_{ij})>0$ for all z_{ij}, $\forall i,j$, and $L(+\infty)=0$)

Notice that, if $x_{ij} \equiv \tilde{x}_{ij} = c_{1ij} = c_{2ij}$ and $l_{ij} = r_{ij} = 0$ then $\tilde{\mathbf{X}} \equiv \mathbf{X}$, i.e. the *fuzzy data* degenerates in a *crisp* or *traditional data*.

A particular case of *LR* fuzzy data is the trapezoidal one, with the following membership function (see figure 2):

$$\mu(\tilde{u}_{ij}) = \begin{cases} 1 - \dfrac{c_{1ij} - \tilde{u}_{ij}}{l_{ij}} & \tilde{u}_{ij} \le c_{1ij} \quad (l_{ij} > 0), \\ 1 & c_{1ij} \le \tilde{u}_{ij} \le c_{2ij}, \\ 1 - \dfrac{\tilde{u}_{ij} - c_{2ij}}{r_{ij}} & \tilde{u}_{ij} \ge c_{2ij} \quad (r_{ij} > 0). \end{cases} \tag{3}$$

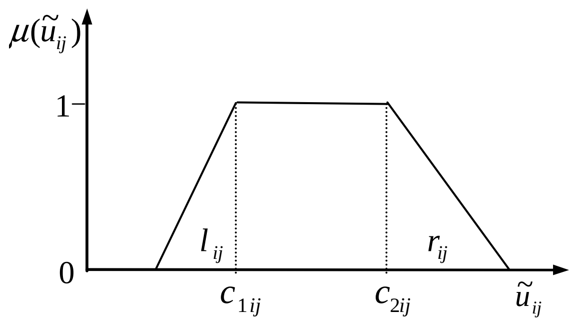

Figure 2. Trapezoidal membership function

On the basis of the family of membership functions (2) and the sub-family (3), we can obtain different particular cases of membership functions, e.g. the triangular membership function (see D'Urso, 2007).

The dissimilarity between each pair of objects is measured by comparing the fuzzy data observed on each object, i.e. by considering, separately, the distances for the centers and the spreads of the fuzzy data and using a suitable weighting system for such distance components. By considering the *i*-th and *i'*-th objects, Coppi et al. (2012) proposed the following squared (Euclidean) distance measure:

$$d_F^2(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_{i'}) = w_C^2[d^2(\mathbf{c}_{1i}, \mathbf{c}_{1i'}) + d^2(\mathbf{c}_{2i}, \mathbf{c}_{2i'})] + w_S^2[d^2(\mathbf{l}_i, \mathbf{l}_{i'}) + d^2(\mathbf{r}_i, \mathbf{r}_{i'})], \tag{4}$$

where $d(\mathbf{c}_{1i},\mathbf{c}_{1i'})=\|\mathbf{c}_{1i}-\mathbf{c}_{1i'}\|$ = Euclidean distance between the left centers $\mathbf{c}_{1i}$ and $\mathbf{c}_{1i'}$; $d(\mathbf{c}_{2i},\mathbf{c}_{2i'})=\|\mathbf{c}_{2i}-\mathbf{c}_{2i'}\|$= Euclidean distance between the right centers $\mathbf{c}_{2i}$ and $\mathbf{c}_{2i'}$; $d(\mathbf{l}_i,\mathbf{l}_{i'})=\|\mathbf{l}_i-\mathbf{l}_{i'}\|$= Euclidean distance between the left spreads $\mathbf{l}_i$ and $\mathbf{l}_{i'}$; $d(\mathbf{r}_i,\mathbf{r}_{i'})=\|\mathbf{r}_i-\mathbf{r}_{i'}\|$ = Euclidean distance between the right spreads $\mathbf{r}_i$ and $\mathbf{r}_{i'}$ $\mathbf{c}_{1i}\equiv(c_{1i1},\ldots,c_{1ij}\ldots,c_{1ip})'$, $\mathbf{c}_{1i'}\equiv(c_{1i'1},\ldots,c_{1i'j}\ldots,c_{1i'p})'$, $\mathbf{c}_{2i}\equiv(c_{2i1},\ldots,c_{2ij}\ldots,c_{2ip})'$, $\mathbf{c}_{2i'}\equiv(c_{2i'1},\ldots,c_{2i'j}\ldots,c_{2i'p})'$, $\mathbf{l}_i\equiv(l_{i1},\ldots,l_{ij}\ldots,l_{ip})'$, $\mathbf{l}_{i'}\equiv(l_{i'1},\ldots,l_{i'j}\ldots,l_{i'p})'$, $\mathbf{r}_i\equiv(r_{i1},\ldots,r_{ij}\ldots,r_{ip})'$, $\mathbf{r}_{i'}\equiv(r_{i'1},\ldots,r_{i'j}\ldots,r_{i'p})'$; $w_C,w_S\geq 0$ are suitable weights for the center component and the spread component of $d_F^2(\tilde{\mathbf{x}}_i,\tilde{\mathbf{x}}_{i'})$, where $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{x}}_{i'}$ denote the fuzzy data vectors, respectively, for the *i*-th and *i'*-th objects, i.e. $\tilde{\mathbf{x}}_i\equiv\{\tilde{x}_{ij}=(c_{1ij},c_{2ij},l_{ij},r_{ij})_{LR}:j=1,\ldots,p\}$ and $\tilde{\mathbf{x}}_{i'}\equiv\{\tilde{x}_{i'j}=(c_{1i'j},c_{2i'j},l_{i'j},r_{i'j})_{LR}:j=1,\ldots p\}$. The weights $w_C,w_S\geq 0$ can be fixed *subjectively* a priori by considering external or subjective conditions or can be computed *objectively* within a suitable clustering procedure. In general it is recommended to estimate the weights in an objective way during the clustering minimization problem (Coppi et al., 2012).

The distance was obtained as a weighted sum of the centers distance and the spreads distance. The weights were constructed in such a way that the centers distance played a more relevant role (at the most an equivalent role) than the spreads distance taking into account that the membership function values within the centers are maximal Then, we have the following conditions: $w_C+w_S=1$ (*normalization condition*) and $w_C\geq w_S\geq 0$ (*coherence condition*) (Coppi et al., 2012).

Notice that, by (4), it assumes that the weights for the left and right center distances and the left and right spreads distances are the same. For more details on (4), see Coppi et al. (2012).

Coppi et al. (2012) proposed a fuzzy clustering model for fuzzy data, in which the weights are obtained objectively. In particular, the clustering model is:

$$\begin{cases}\min:\sum_{i=1}^{n}\sum_{g=1}^{k}u_{ig}^{m}d_F^2(\tilde{\mathbf{x}}_i,\tilde{\mathbf{h}}_g)=\sum_{i=1}^{n}\sum_{g=1}^{k}u_{ig}^{m}[w_C^2[d^2(\mathbf{c}_{1i},\mathbf{h}_g^{C_1})+d^2(\mathbf{c}_{2i},\mathbf{h}_g^{C_2})]+w_S^2[d^2(\mathbf{l}_i,\mathbf{h}_g^{L})+d^2(\mathbf{r}_i,\mathbf{h}_g^{R})]],\\ \text{s.t. }\ u_{ig}\in[0,1];\sum_{g=1}u_{ig}=1,\\ \quad w_C,w_S\geq 0;w_C\geq w_S;w_C+w_S=1,\end{cases}\tag{5}$$

where: *m*>1 is a weighting exponent that controls the fuzziness of the obtained partition; u_{ig} indicates the membership degree of the *i*-th object in the *g*-th cluster; $d_F^2(\tilde{\mathbf{x}}_i,\tilde{\mathbf{h}}_g)$ represents the suggested dissimilarity measure between the *i*-th object and the prototype of the *g*-th cluster; analogously for its components $d^2(\mathbf{c}_{1i},\mathbf{h}_g^{C_1}),d^2(\mathbf{c}_{2i},\mathbf{h}_g^{C_2}),d^2(\mathbf{l}_i,\mathbf{h}_g^{L}),d^2(\mathbf{r}_i,\mathbf{h}_g^{R})$, where the fuzzy vector $\tilde{\mathbf{h}}_g\equiv\{\tilde{h}_{gj}=(h_g^{C_1},\ h_g^{C_2},\ h_g^{L},\ h_g^{R})_{LR}:j=1,\ldots,p\}$ represents the fuzzy prototype of the *g*-th cluster, $\mathbf{h}_g^{C_1}\equiv(h_{g1}^{C_1},\ldots,h_{gj}^{C_1},\ldots h_{gp}^{C_1})'$, $\mathbf{h}_g^{C_2}\equiv(h_{g1}^{C_2},\ldots,h_{gj}^{C_2},\ldots h_{gp}^{C_2})'$,

$\mathbf{h}_g^L \equiv (h_{g1}^L,\dots,h_{gj}^L,\dots h_{gp}^L)'$, $\mathbf{h}_g^R \equiv (h_{g1}^R,\dots,h_{gj}^R,\dots h_{gp}^R)'$ are p-vectors, whose j-th element refers to the j-th variable, that denote, respectively, the (left and right) centers and the (left and right) spreads of the g-th fuzzy prototype.

The *iterative solutions* are (Coppi et al., 2012):

$$u_{ig} = \frac{\left[w_C^2[d^2(\mathbf{c}_{1i},\mathbf{h}_g^{C_1}) + d^2(\mathbf{c}_{2i},\mathbf{h}_g^{C_2})] + w_S^2[d^2(\mathbf{l}_i,\mathbf{h}_g^L) + d^2(\mathbf{r}_i,\mathbf{h}_g^R)]\right]^{-\frac{1}{m-1}}}{\sum_{g'=1}^{k}\left[w_C^2[d^2(\mathbf{c}_{1i},\mathbf{h}_{g'}^{C_1}) + d^2(\mathbf{c}_{2i},\mathbf{h}_{g'}^{C_2})] + w_S^2[d^2(\mathbf{l}_i,\mathbf{h}_{g'}^L) + d^2(\mathbf{r}_i,\mathbf{h}_{g'}^R)]\right]^{-\frac{1}{m-1}}}, \tag{6}$$

$$\mathbf{h}_g^{C_1} = \frac{\sum_{i=1}^n u_{ig}^m \mathbf{c}_{1i}}{\sum_{i=1}^n u_{ig}^m}, \quad \mathbf{h}_g^{C_2} = \frac{\sum_{i=1}^n u_{ig}^m \mathbf{c}_{2i}}{\sum_{i=1}^n u_{ig}^m}, \quad \mathbf{h}_g^{L} = \frac{\sum_{i=1}^n u_{ig}^m \mathbf{l}_{i}}{\sum_{i=1}^n u_{ig}^m}, \quad \mathbf{h}_g^{R} = \frac{\sum_{i=1}^n u_{ig}^m \mathbf{r}_{i}}{\sum_{i=1}^n u_{ig}^m}, \tag{7}$$

$$w_C = \frac{\sum_{i=1}^n\sum_{g=1}^k u_{ig}^m[d^2(\mathbf{l}_i,\mathbf{h}_g^L) + d^2(\mathbf{r}_i,\mathbf{h}_g^R)]}{\sum_{i=1}^n\sum_{g=1}^k u_{ig}^m[d^2(\mathbf{c}_{1i},\mathbf{h}_g^{C_1}) + d^2(\mathbf{c}_{2i},\mathbf{h}_g^{C_2}) + d^2(\mathbf{l}_i,\mathbf{h}_g^L) + d^2(\mathbf{r}_i,\mathbf{h}_g^R)]} \quad \left(w_S = 1 - w_C\right). \tag{8}$$

Notice that, the clustering model (5) represents generalization of the fuzzy clustering model for "precise" (non-vague or non-fuzzy) data proposed by Bezdek (1981).

Furthermore, the model (5) allows us to detect k homogeneous clusters on the basis of n objects described by p fuzzy variables. To characterize every cluster, a fictitious object, i.e. the prototype, has been computed.

A crucial assumption of the clustering model (5) is that the prototypes are of LR fuzzy type, inheriting their typology by the observed data. "Generally speaking, the prototypes are obtained as a weighted mean of the observed objects using the membership degree information as system of weights. In fact, the extent to which an object belongs to a given cluster is expressed by the membership degree (of an object in a cluster). Although every membership degree can range in the unit interval in both the approaches, their meaning remarkably differs. In fact, following the fuzzy approach, the membership degrees can be seen as degrees of sharing of an object among the clusters and their sum for each object over all the clusters must be equal to one." (Coppi et al., 2012).

A flowchart of the steps of the classification via the F*k*M-F method are presented in Figure 3.

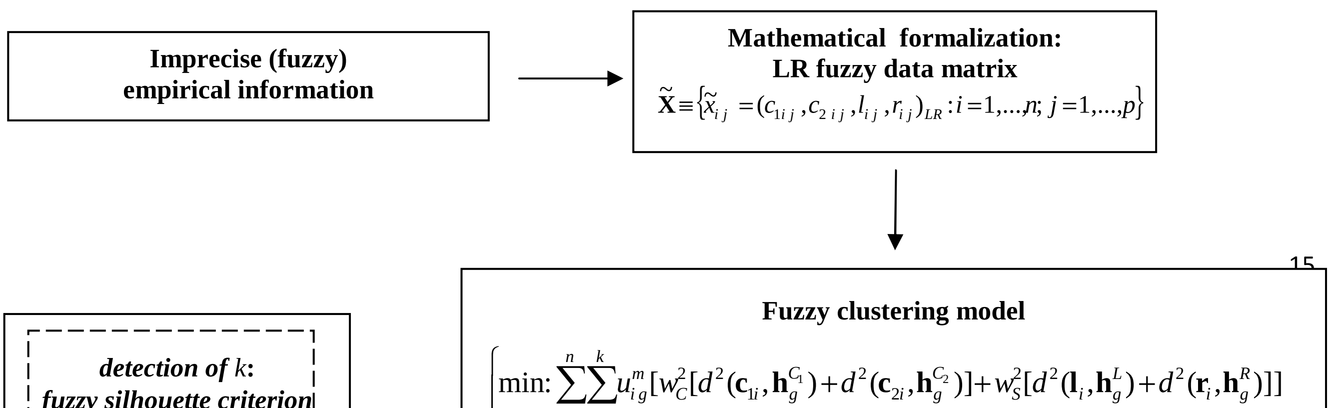

Figure 3. Steps of the classification via the F*k*M-F clustering method

4. A fuzzy taxonomy for the e-Health projects: some empirical evidences

In order to carry on the empirical part of our research, we first perform a qualitative analysis of 94 e-Health projects where a single project corresponds to the unit of analysis. Then we make a cross-case synthesis though advanced cluster techniques to derive a taxonomy of e-Health projects with similar characteristics. With this approach, the taxonomy we obtain is based on the classification of empirical data instead of being the result of a conceptual analysis. More in detail, we define a set of clusters of homogeneous cases based on a given set of case descriptions derived from the analysis of a database of 94 European e-Health projects which have been labeled as "good practices" by an expert panel of specialists in different areas of e-Health selected by the European Commission (EC 2008).

The overall research design (figure 4) corresponds to what is referred in the social science research as a holistic multiple case (Yin 2009, p. 46) with the purpose of setting up the basis for carrying on further embedded multiple case studies, where the cases in each category can be further investigated. In fact, the evidence from multiple cases is often considered more compelling, and the overall study is, therefore, regarded as being more robust (Herriott and Firestone, 1983). Moreover, the results of this preliminary research will set the basis for the development of a rich theoretical framework and for the application of rigorous replication procedures. In this way, the theoretical framework states the conditions under which a particular phenomenon is likely or not likely to be found.

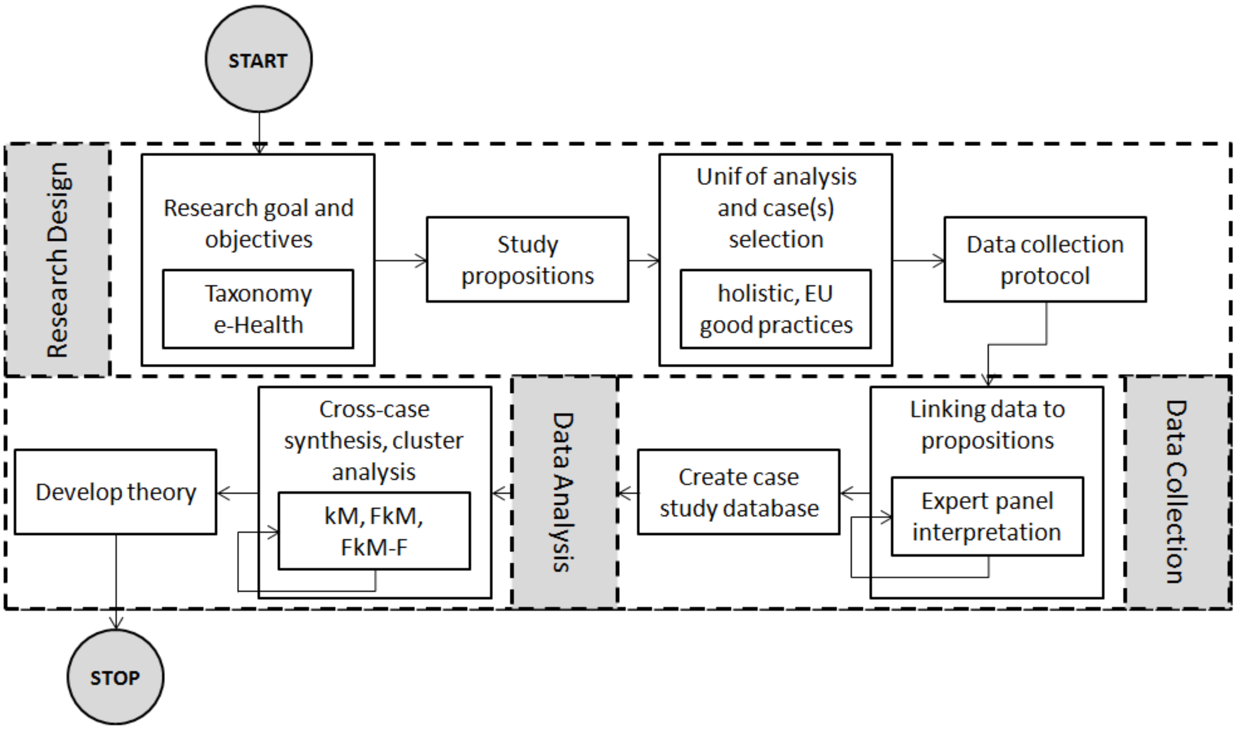

Figure 4. Research process overview

We must specify that the choice of analyzing a dataset with 94 cases does not imply any attempt to pursue some form of statistical generalization. In fact, for case study research, generalization follows the analytic mode, according to which a previously developed theory is used as a template with which to compare the empirical results of the case study (Yin 2009, p. 38).

4.1 Data source

In order to collect information on the characteristics of e-Health projects which have been successfully implemented in the European context, we refer to a public available online database which has been created in the context of an initiative of the European Commission (Good e-Health, EC 2008). The Good eHealth initiative is a three-year study (from 2006 to 2008) which has been financed by the European Commission with the objectives of identifying good practices and their associated benefits, disseminating real life experiences, and fostering accelerated take-up of e-Health.

In order to fulfill these goals, a knowledge base with more than 100 real-life e-Health case studies is made available through an online database. A twofold impact is expected. First, political, clinical, managerial and health professional decision-makers can use this knowledge for implementing more effective e-Health services. Second, patients and citizens can use it to enhance aspects of their own and their families' care.

With the help of an expert panel, Good e-Health has selected a range of e-Health solutions. Among the 132 solutions which were listed in the database at the time of the data collection for this research, 94 cases have been certified as “quality reviewed cases”. The project website (http://kb.good-ehealth.org/search.do) describes in detail the selection process through which cases are analyzed by an expert panel of specialists in different areas of e-Health.

We considered the Good e-Health knowledge base as appropriate with respect to the purposes of this research for three main reasons. First, the wide variety of cases listed in the knowledge base covers a large geographical area with different legal frameworks and socio-economical contexts. Second, the review process through which the submitted cases have been evaluated ensures the quality of available information. In fact, the ratio between proposed cases and selected cases is about 6:1 and projects have been evaluated against 12 criteria such as transformational impacts, current level of deployment, availability, etc. Finally, detailed descriptions are provided for each case based on data gathered from different sources (i.e. a network of country correspondents, secondary source material and telephone interviews with stakeholders). Cases are presented on the website through a common template with an average size for the overall case descriptions of nearly 2000 words.

These cases need not necessarily be the “best” or the most innovative while they are considered as proven real-life good practice examples. Using a qualitative approach the entire field of e-Health solutions can be assessed. The selected solutions illustrate the entire range of the continuum of healthcare and all the European countries. Cases portray the national, community and business levels of health provision. The process pays particular attention to identifying organisational, socio-economic, and stakeholder issues in e-Health.

4.2 Data collection

To generate the dataset on which our statistical analysis applies, a research team composed by five practitioners with experience in the e-health domain and five researchers with experience in IS/IT evaluation has been involved in the data collection process. The objective of this phase has been to achieve a shared understanding on the characteristics of the 94 “quality reviewed cases” from the Good e-Health database. An iterative process with periodic meetings over a six months period has been carried on for this purpose. Each case description has been carefully analyzed and discussed in order to agree on the level of contribution of each project to the four e-health IT capabilities. Therefore the considered cases have been evaluated as *perfect*, *good*, *medium*, *poor* and *bad* with respect to the CIS, SUNCS, TLM, IHCIN dimensions. Therefore the expert panel utilized an ordinal quality scale based on 5 different levels for classifying each case. These quality terms are characterized by the imprecision (vagueness) inherited by human perception.

To reflect either the intrinsic imprecision of the evaluation of e-Health projects or the inherent subjectivity of the evaluation expressed by the experts the scale of fuzzy numbers has been used. In fact, in according with González-Rodríguez et al. (2012) and Sinova et al. (2012), in our case, it is not suitable to utilize the Likert scales) in which the 5 different categorical levels are labeled with numerical values. In fact, using these scales, our statistical analysis for defining the taxonomy of the e-health projects would be limited and the interpretation of the results would be considerably reduced. Conversely, the adoption of fuzzy scale is more expressive and accurate than the utilization of ordinal scales and more accurate. In fact, instead of modeling the quality levels (items) of the qualitative scale utilized by the e-health experts, by means of either numerical or categorical data, which would be less accurate or expressive, the adopted fuzzy scale integrates the manageability and diversity/variability of the numerical scale and the interpretability and ability to capture the imprecision of the considered categorical scale (Sinova et al., 2012). Then, the adopted fuzzy scale enables us to exploit the subjectivity that is involved in perceiving and expressing the available information expressed by the e-Health expert panel. In conclusion, since the fuzzy sets can be suitably utilized for describing the ambiguity and imprecision in natural language, we can represent the quality terms by means of triangular fuzzy numbers, i.e.: $\tilde{Y} = (1,0.25,0)$ (perfect), $\tilde{Y} = (0.75,0.25,0.25)$ (good), $\tilde{Y} = (0.5,0.25,0.25)$ (medium), $\tilde{Y} = (0.25,0.25,0.25)$ (poor), $\tilde{Y} = (0,0,0.25)$ (bad) (Hung and Yang, 2005). The dataset is presented in Table 2.

project	CISr	SUNCSr	TLMr	IHCINr	project	CISr	SUNCSr	TLMr	IHCINr
1	poor	good	bad	bad	**48**	perfect	good	bad	bad
2	bad	good	bad	bad	**49**	poor	good	bad	poor
3	perfect	medium	bad	bad	**50**	bad	perfect	bad	bad
4	medium	medium	bad	perfect	**51**	poor	bad	perfect	bad
5	poor	medium	bad	medium	**52**	medium	bad	bad	medium
6	bad	good	bad	bad	**53**	medium	bad	perfect	poor
7	medium	good	bad	bad	**54**	medium	bad	medium	perfect
8	medium	good	bad	medium	**55**	perfect	good	bad	bad
9	poor	medium	good	medium	**56**	bad	good	bad	bad
10	bad	poor	perfect	bad	**57**	bad	bad	bad	perfect
11	medium	medium	poor	perfect	**58**	bad	good	bad	poor
12	medium	medium	poor	perfect	**59**	medium	poor	bad	good
13	poor	bad	poor	medium	**60**	bad	good	bad	poor
14	medium	poor	bad	perfect	**61**	bad	good	bad	bad
15	medium	poor	bad	poor	**62**	perfect	medium	bad	bad
16	medium	bad	good	poor	**63**	bad	perfect	bad	bad
17	medium	good	bad	good	**64**	bad	good	good	medium
18	medium	poor	bad	perfect	**65**	poor	bad	bad	perfect
19	medium	poor	bad	perfect	**66**	good	good	bad	medium
20	good	medium	bad	bad	**67**	bad	good	bad	bad
21	medium	perfect	bad	bad	**68**	bad	perfect	bad	bad
22	good	perfect	bad	bad	**69**	bad	good	bad	bad
23	good	perfect	bad	bad	**70**	bad	perfect	bad	bad
24	good	perfect	bad	bad	**71**	poor	medium	bad	perfect
25	bad	perfect	bad	medium	**72**	medium	perfect	bad	bad
26	bad	perfect	bad	bad	**73**	poor	perfect	bad	bad
27	perfect	medium	bad	bad	**74**	bad	perfect	poor	bad
28	bad	good	bad	bad	**75**	bad	good	bad	bad
29	good	poor	bad	bad	**76**	medium	bad	bad	perfect
30	good	medium	bad	poor	**77**	bad	bad	perfect	bad
31	bad	poor	medium	bad	**78**	bad	perfect	bad	good
32	medium	bad	perfect	bad	**79**	perfect	perfect	bad	bad
33	bad	bad	perfect	bad	**80**	perfect	good	medium	good
34	bad	good	bad	bad	**81**	medium	perfect	bad	perfect
35	bad	good	bad	bad	**82**	bad	bad	perfect	good
36	medium	perfect	bad	medium	**83**	poor	bad	perfect	bad
37	good	poor	bad	bad	**84**	bad	perfect	medium	bad
38	good	bad	perfect	bad	**85**	bad	medium	medium	bad
39	medium	bad	perfect	bad	**86**	bad	good	bad	bad
40	bad	medium	bad	perfect	**87**	medium	medium	perfect	bad
41	medium	bad	perfect	poor	**88**	perfect	good	bad	bad
42	medium	good	bad	bad	**89**	bad	good	perfect	bad
43	bad	good	bad	medium	**90**	bad	perfect	bad	bad
44	poor	bad	poor	medium	**91**	bad	perfect	bad	bad
45	bad	medium	bad	perfect	**92**	bad	good	bad	good
46	medium	bad	poor	medium	**93**	bad	perfect	bad	bad
47	bad	medium	bad	good	**94**	bad	good	bad	medium

Table 2. The dataset

4.3 Data analysis

As already mentioned the dataset has been created on the basis of the outcome of an iterative interpretation process. Each case corresponds to an EU good practice (project) in the e-Health domain and project characteristics are described on documents publicly available online on the EU database. The interpretation process is finalized to link data to the research propositions that have been identified during the research design phase. More in details the theoretical proposition refers to the relationship between IT capabilities, IT applications and IT platforms within an e-Health project by advocating that a taxonomy of e-health projects should emerge from empirical data with an inductive analytical process (see figure 4). Therefore the cross-case synthesis has three main objectives: i) to identify the number of clusters, ii) to characterize each cluster with its components, iii) to define an empirically grounded taxonomy for e-Health projects.

The subsequent phase of cross-case synthesis has been performed using advanced cluster analysis techniques that takes into account the limitations of the expert panel interpretation. In fact, in the real world, a single case seldom fit only with a single category (i.e. an e-Health project with a single IT capability) and also the level of fit within a category can be biased by the analyst subjective interpretation. This can be explained with the fact that in the real world, an element of information is generally characterized by imprecision (with regards to value) and uncertainty. Imprecision and/or uncertainty define what we may call imperfect information (here the term imperfect indicates that the information presents one or more of the following features: vagueness, roughness, imprecision, ambiguity, and uncertainty) (D'Urso 2007). Hence, since the categories used to perform case classifications are typically derived from *a priori* definitions (i.e. IT capabilities) which are based on abstract conceptualization of systems properties, they may not reflect the characteristics of real cases. This bias can lead to erroneous interpretations on the phenomenon under investigation, raising the risk of misleading and incomplete conclusions with consequences on policy and decision-makers choices.

With these premises, we apply both traditional and more advanced cluster analysis techniques, based on fuzzy theories, to analyze the characteristics of about a hundred successful e-Health projects carried out in European countries in the last ten years. An application of the F*k*M-F clustering model and a comparison with *k*M and F*k*M models on the above mentioned dataset is presented in the next section.

The performances of the F*k*M-F clustering model with respect to three existing clustering techniques for fuzzy data have been compared via a simulation study in D'Urso and Giordani (2006).

5. Results and discussion

In this section we discuss the results of the cross-case synthesis performed on the dataset.

Three methods have been considered: crisp clustering of crisp data (*k*-means, i.e. *k*M) (Mac Queen, 1967), fuzzy clustering of crisp data (F*k*M, i.e. fuzzy *k*-means) (Bezdek, 1981), fuzzy clustering of fuzzy data (F*k*M-F, i.e. fuzzy *k*-means of fuzzy data) (Coppi et al., 2012).

For the F*k*M and F*k*M-F clustering models the value of the fuzzy parameter m should be suitable chosen in advance.

In literature, different empirical heuristic procedures have been suggested, but there seems to exist no theoretically justifiable manner of selecting *m* (Hwang et al., 2007; Maharaj, D'Urso, 2011).
Pal and Bezdek (1995) have given heuristic guidelines regarding the best choice for *m*, suggesting that the value of the level of fuzziness should be between 1.5 and 2.5. Similar recommendations appear in Cannon et al. (1986), Hall et al. (1992), Soreson and Wang (1996) and Fadili et al. (2001).
Based on their analysis, Ozkan and Turksen (2007) suggested that the lower and upper boundary values of m should be, respectively, approximately 1.4 and 2.6.
Different values of *m* between 1.5 and 2.5 have been considered. The choice of *m*=1.5 corresponds to the least fuzziness of the obtained partitions, thus resulting in a clearer reading and interpretation of the results.

For the F*k*M-F clustering method the obtained value of w_C is 0.5.

The present classification of e-health projects considers that four IT capabilities can be combined in different ways in order to achieve the objectives of *k* different classes of e-Health projects.

$$FS = \frac{\sum_{j=1}^{94} (\mu_{rj} - \mu_{qj})^{\alpha} s_j}{\sum_{j=1}^{94} (\mu_{rj} - \mu_{qj})^{\alpha}}$$

$$s_j = \frac{b_{rj} - a_{rj}}{\max\{b_{rj}, a_{rj}\}}$$

For selecting *k*, we adopt the Fuzzy Silhouette *cluster validity* criterion (Campello, Hruschka 2006).
The Fuzzy Silhouette is a generalization to the fuzzy case of the Average Silhouette Width Criterion or Crisp Silhouette. It is a weighted average, with weights that take into account the membership degrees, of the individual silhouettes s_j, where the silhouette of a project is a measure of its closeness to the projects in the highest membership cluster with respect to the distance to projects in other clusters, i.e.:

where μ_{rj} and μ_{qj} are the first and second largest elements of the *j*-th row of the fuzzy partition matrix (the *n*x*k* matrix containing the membership degrees of the *n* projects to the *k* clusters), a_{rj} the average distance of project *j* to all other projects belonging to its highest membership cluster *r* and b_{rj} the minimum (*q*=1,…,*k*) average distance d_{qj} of project *j* to all projects belonging to another cluster *q*, *q*≠*r*, α is an optional user defined weighting coefficient (set to 1).
The higher the value of the Fuzzy Silhouette, the better the assignment of the objects to the clusters.
The values of the Fuzzy Silhouette are presented in figure 5 for 2, 3, 4, 5 clusters.

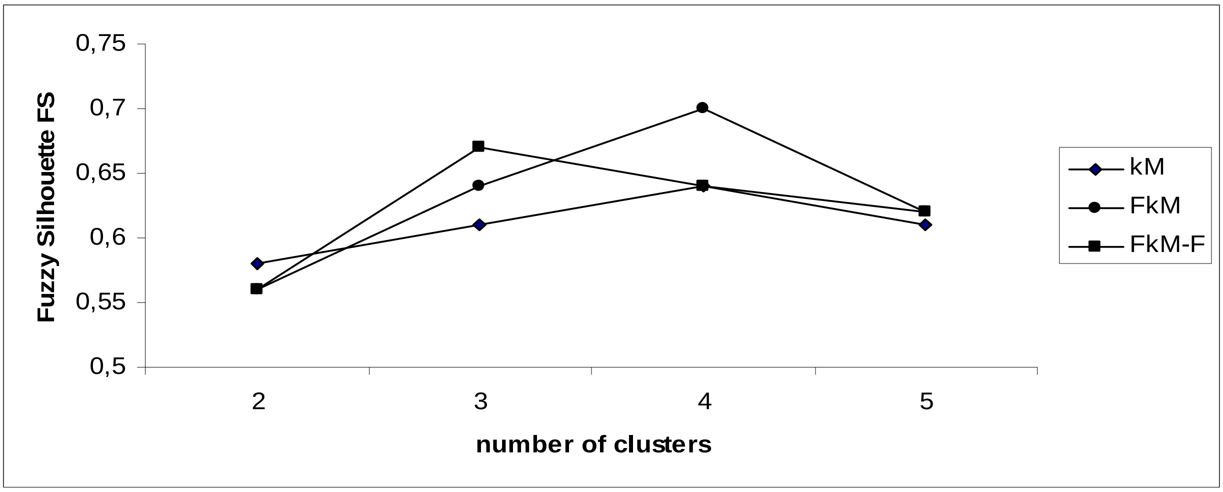

Figure 5. Fuzzy Silhouette – F*k*M-F clustering method

The analysis of the Fuzzy Silhouette values shows that the F*k*M-F cluster method locates the maximum at *k*=3 clusters, the other two methods at *k*=4 clusters. In the F*k*M-F method the information concerning the quantification via the fuzzy variables leads to a parsimonious number of clusters.

Furthermore, as an objective criterion for the evaluation of fuzzy partitions of a data set – provided by a fuzzy clustering algorithm - the Fuzzy Rand index (Anderson et al., 2010) has been considered. It is a fuzzy extension of the original Rand index $\omega=(a+d)/(a+b+c+d)$ based on the comparison of agreements and disagreements (a, d indicate consistent classifications, that is the number of pairs of projects belonging to the same cluster and to different clusters in the two partitions, respectively; b, c indicate inconsistent classifications, that is the number of pairs of projects belonging to the same cluster in the first (second) partition and to different clusters in the second (first) partition in two partitions, the fuzzy partition and the hard partition, or in two fuzzy partitions, which may have different number of clusters. The values of the Fuzzy Rand index are presented in Table 3.

	2 clusters	3 clusters	4 clusters	5 clusters
*k*M *vs* F*k*M	0,66	0,82	0,89	0,88
*k*M *vs* F*k*M-F	0,62	0,66	0,73	0,73
F*k*M *vs* F*k*M-F	0,62	0,66	0,72	0,73

	*k*M	F*k*M	F*k*M-F
3 clusters *vs* 4 clusters	0,82	0,62	0,73
3 clusters *vs* 5 clusters	0,79	0,63	0,73
4 clusters *vs* 5 clusters	0,94	0,68	0,83

Table 3. Fuzzy Rand index between partitions obtained with different clustering methods (left) and different number of clusters (right)

In Table 3 – left – the Fuzzy Rand index has been computed for comparing two partitions obtained with different clustering methods, with the same number of clusters, that is *k*M versus F*k*M, *k*M versus F*k*M-F, F*k*M versus F*k*M-F, considering partitions into 2, 3, 4, 5 clusters. For each number of clusters the values of the Fuzzy Rand index show agreement between the partitions obtained with the clustering methods *k*M and F*k*M. In Table 3 – right – the Fuzzy Rand index has been computed for comparing two partitions obtained with a different number of clusters, with the same clustering method, that is 3 versus 4 clusters, 3 versus 5, 4 versus 5, considering the partitions obtained with *k*M, F*k*M, F*k*M-F clustering methods. For each method the values of the Fuzzy Rand index show more agreement between the partitions into 4 and 5 clusters than between the partitions into 3 and 4 clusters or 3 and 5 clusters.

The prototypes of F*k*M-F clustering method are shown via parallel coordinates plots (figure 6). Notice that the parallel coordinate plot is a graph for representing multivariate data. The dimension in the considered application is the number of variables p=4. To represent a point in a 4-dimensional space, 4 parallel vertical and equally spaced lines are used. A point in a 4-dimensional space is represented as a polyline with vertices on the parallel axes; the position of the vertex on the j-th axis corresponds to the j-th coordinate of the point. For each cluster there are three polylines, one corresponding, for each variable, to the values of the centers, one corresponding to the values of the centers minus the left spreads, one corresponding to the centers plus the right spreads.

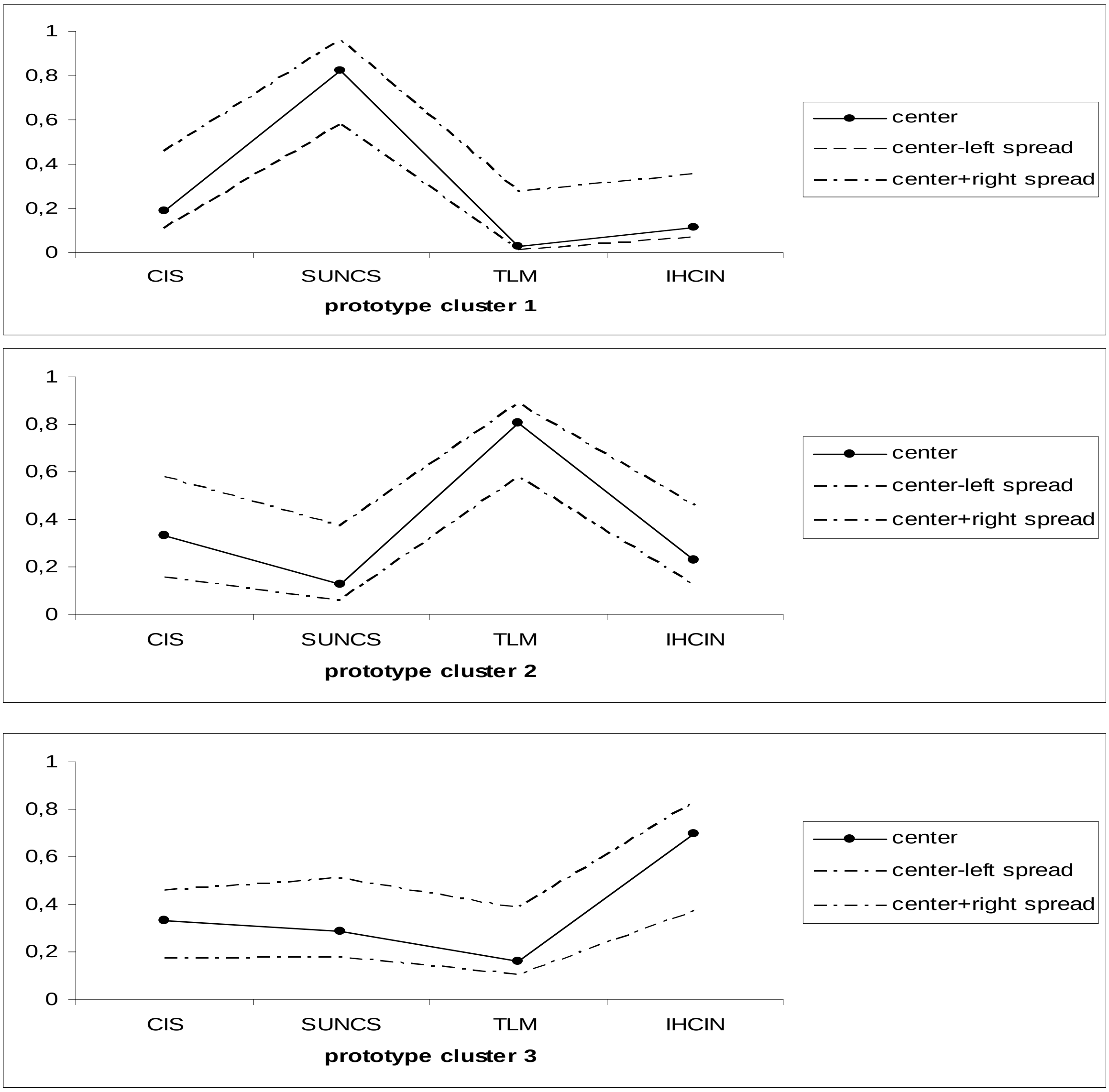

Figure 6. Prototypes for F*k*M-F clustering method

Cluster 1 is characterized by a high value of SUNCS, and non negligible value of CIS. The left spread is small for CIS, TLM and IHCIN; the right spread is small for SUNCS. We associate to this cluster the first category of e-Health projects (eH1) which are aimed to the development of IT applications with both SUNCS and CIS capabilities. These applications are mainly focused on the support of administrative processes within a single organization (i.e. hospital, local health authority, etc.) and with possible interconnections with IT systems supporting clinical processes.

Cluster 2 is characterized by a high value of TLM, and non negligible value of CIS. The left spread is small for SUNCS and IHCIN; the right spread is small for TLM. We associate to this cluster the second category of e-Health projects (eH2) which are aimed to the development of IT applications with both CIS and TLM capabilities. These applications are mainly focused on the support of clinical processes encompassing the physical boundaries of a single healthcare organization (i.e. hospital, laboratory, etc.) through remote data transmission.

Cluster 3 is characterized by a high value of IHCIN , and non negligible values of CIS and SUNCS. The left spread is small for TLM; the right spread is non negligible for all the variables. We associate to this cluster the third category of e-Health projects (eH3) which are aimed to the development of IT platforms with eH1, eH2 applications and IHCIN capabilities. These platforms are mainly focused on providing support to integrated care processes with a patient-centered approach.

The above mentioned classes of e-Health projects (i.e. eH1, eH2, and eH3) validate our research proposition 1. In fact the four IT capabilities identified through the analysis of the knowledge base characterize in a consistent manner the set of e-Health projects analyzed. Furthermore e-Health projects can be classified through a taxonomy whose elements are based on the hierarchical composition of IT capabilities, IT applications, and IT platforms.

As far as the centers are considered, they do not differ from the prototypes of the other two methods.

The analysis of the prototypes shows that exist prototypes with the dominance of only one variable, SUNCS (cluster 1), TLM (cluster 2), IHCIN (cluster 3), with the exception of CIS, that exhibits high dominance in the three prototypes joint with one of the other three variables. By referring to the definitions of the four IT capabilities, this result can be explained by the fact that Clinical Information System are intended either as specialised tools for health professionals within healthcare institutions (e.g. hospitals) or as tools for primary care and/or for outside care institutions such as general practitioner and pharmacy information systems. These systems support the daily operations of health professionals, but they also collect data which are useful for other purposes. For instance, clinical data about patients/citizens can be also used, if properly anonymized, for medical research and public health purposes. Furthermore, within the hospital boundaries, clinical systems are often integrated with systems supporting the administrative and managerial processes.

The partitions via membership degrees for the three clustering methods are illustrated via ternary plots (figure 7).

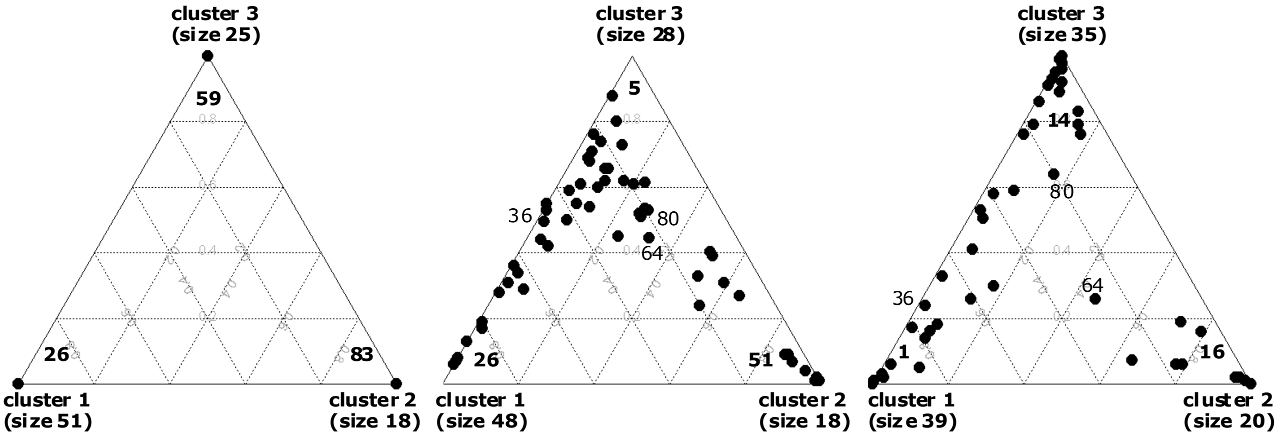

Figure 7. Membership degrees of projects - methods *k*M (left), F*k*M (middle), F*k*M-F (right)

According to the values of the prototypes, the units closest to the prototypes (in bold in the related ternary plot) for the *k*M method are unit 26 for cluster 1, unit 83 for cluster 2 and unit 59 for cluster 3; for the F*k*M method unit 26 for cluster 1, unit 51 for cluster 2 and unit 5 for cluster 3; for the F*k*M-F method are unit 1 for cluster 1, unit 16 for cluster 2 and unit 14 (and 18 and 19) for cluster 3. The values of the variables of the units closest to the prototypes show that the prototypes of the three methods as far as the centers are considered are similar.

The analysis of the membership degrees represented in the ternary plots shows that some units present uncertain classification. In particular these units are units 36, 80 and 64 (figure 7).

The analysis of these units show that unit 80 presents high values of all the four variables, unit 36 high values of three variables (CIS, SUNCS, IHCIN), unit 64 high values of three variables (SUNCS, TLM, IHCIN).

Unit 80 refers to the Ykonos project which is aimed at allowing immediate access to radiological information and medical images of any patient to all health professionals at any healthcare centre in Castilla-La Mancha, Spain. According with the above mentioned objectives, the Ykonos project presents the characteristics of an integrated network for sharing health information, typical of the IHCIN category. Although, this project represents the beginning of an Electronic Health Record (EHR), it provides functionalities for digitalizing radiological images, storing them in an integrated system through a picture archiving and communication system (PACS) and a radiology information system (RIS) which are typical of the CIS and SUNCS categories.

Unit 36 refers to an information system developed to manage the surveillance and control of infectious diseases in Ireland (i.e. CIDR). The system also monitors organisms' ability to resist antibiotic drugs (anti microbial resistance). These properties make the project close to the category SUNCS in that it supports public health data collection and analysis. This case presents characteristics of both CIS and IHCIN categories in that it supports health professionals operations, it allows the control of administrative data and it acts through an integrated network of health information.

Unit 64 refers to a suite of Scottish ambulance communication solutions which allow both transmitting patients' ECG information to specialist cardiac centres that can provide advice to the ambulance crew as remote clinical support (TLM), and link front line ambulance crew to the command and control systems in order to manage the logistic aspects of the incident (SUNCS) and to access the electronic patient reports in an integrated manner (IHCIN).

cluster/method	kM	F*k*M	F*k*M-F
eH1	supply chain optimization (C26)		Wikifood (C1)
eH2	radiology consultations between Sweden and Spain (C83)	Telehome Care for chronically ill patients (C51)	Platform for Chronic Disease Management (C16)
eH3	The Oxford Clinical Intranet regional network (C59)	Flemish vaccination database and Vaccinet (C5)	Internet based EHR system (C14) Southern Ardeche Patient Information Network (C18) Shared and Distributed Patient Record (C19)
uncertain classification	IS for the surveillance of infectious diseases (C36) Scottish Ambulance Communication Solutions (C64) Immediate access to radiological clinical information and medical images (C80)		

Table 4: units closest to the prototypes

The *k*M cluster method fails in identifying prototypes for units that present high values with respect to more than one (or two one of which CIS) variable.

The F*k*M fuzzy clustering method introduces a membership degree of a unit to a cluster, making possible for a unit to exhibit the characteristics of two or more prototypes. Units 36, 80 and 64 are characterized by membership degrees, respectively, (0.48,0.02,0.49), (0.22,0.27,0.51), (0.23,0.32,0.44), showing highest membership to the same cluster.

The F*k*M-F clustering method takes into account the uncertainty concerning quantification via the fuzzy variables, so refining the F*k*M clustering. The F*k*M-F introduces a membership degree of a unit to a cluster, making possible for a unit to exhibit the characteristics of two or more prototypes, and enriches the prototypes with the left and right spreads. Units 36, 80 and 64 are characterized by membership degrees, respectively, (0.70,0.01,0.29), (0.24,0.20,0.56), (0.28,0.46,0.26), showing highest membership to different clusters.

So the information regarding the uncertainty makes it possible to refine the classification. Units 36 and 80 increase the highest membership to a cluster.

The reason why units 36 and 80 are characterized by highest membership to the same cluster in the F*k*M method is that they have high value of more than two variables. Taking into account the spreads adds the information that unit 36 exhibits a small left spread on TLM and right spread on SUNCS, as shown in the prototype of cluster 1; and that unit 80, characterized by a high value of all the four variables, exhibits non negligible values of all the left and right spreads, as shown mostly in the prototype of cluster 3. So unit 36 moves from cluster 3 to cluster 1, whilst unit 80 increases its membership to cluster 3.

Units 64 exhibits the highest membership to cluster 2, and non negligible membership to clusters 1 and 3. The reason is that this unit shows a low level of variable CIS that is high in all the three clusters, and shares the spreads of more than one cluster.

In summary, the refined classification resulting from the application of the F*k*M-F clustering method provides better insights on the nature of e-Health initiatives. Understanding the installed base of an e-Health project is an interpretation process affected by both the intrinsic imprecision of IT capabilities definitions and by the inherent subjectivity of the evaluation process. This result supports proposition 2 and provides a contribution in the direction of developing more powerful tools for supporting strategic decision making in the e-Health domain.

Table 4 summarizes the associations between e-Health project categories and the units closest to their prototypes. The interesting result is that unit 80, which corresponds to a wiki platform for managing food information, belongs to the eH1 project category in which SUNCS and CIS capabilities are provided. Wiki systems are novel platforms which involve users in the production of contents. These platforms belongs to the so called Information Infrastructure (Hanseth and Lyytinen 2010) category of systems which has not been taken into consideration in this research and which adds new capabilities in the basic set. This evidence suggest directions for further investigation.

6. Conclusions

With the aim of building a taxonomy that classifies specific e-Health projects, both traditional and more advanced cluster analysis techniques, based on fuzzy theories, have been applied to analyze the characteristics of about a hundred successful e-Health projects carried out in European countries in the last ten years.

The taxonomy validated through the research process, allows to aggregate similar projects on the basis of their structure and components, representing a first step towards the definition of a set of more context related evaluation frameworks for e-health projects. Strategic decision makers may benefit from the adoption of these frameworks for supporting ex ante and ex post evaluations.

The application of advanced cluster analysis techniques to data collected through the evaluation of good practices in e-Health makes this study innovative from a methodological standpoint. The proposed method combines in fact the qualitative interpretation performed by a team of experts with powerful statistical tools which allows taking into account imperfect information.

The main contribution of this paper consists in an empirically grounded taxonomy for classifying e-Health projects. Evidences have demonstrated that *a priori* classifications fail in providing a description of the general characteristics of e-Health projects when applied to real cases. The application presented in this paper refers to a limited number of IT capabilities which constitute the elementary components of hierarchical IT applications and IT platforms.

Further research may extend the use of this method by considering additional IT capabilities in the data collection protocol, by enlarging the number of project analyzed, and by applying the method to other domains (i.e. e-Government, e-Business, etc.). This will both contribute to the "what is" question related to "e-strategies" and will provide more powerful conceptual tools for policy and decision makers. Another possible future study concerns the investigation of alternative innovative fuzzification approaches of the evaluation scales -suggested in the recent literature (Colubi et al., 2011; González-Rodríguez et al., 2012; Sinova et al., 2012)- for defining a suitable taxonomy of e-Health projects.

Acknowledgements

We wish to thank the referees and the Editor for their useful comments and suggestions which helped to improve the quality and presentation of this manuscript.

References

Aanestad M, Jensen TB (2011) Building nation-wide information infrastructures in healthcare through modular implementation strategies. J Strateg Inf Syst 20(2): 161-176.

Anderson, D.T., Bezdek, J.C., Popescu, M., Keller, J.M. (2010) Comparing Fuzzy, Probabilistic, and Possibilistic Partitions, IEEE Transaction on Fuzzy Systems, 18, 906-918.

Barlow J, Bayer S, Curry R (2006) Implementing complex innovations in fluid multi-stakeholder environments: Experiences of 'telecare', Technovation 26, 396–406.

Bates DW (2005) Physicians and Ambulatory Electronic Health Records. Health Aff: 24/5:1180-1189

Bezdek J.C. (1981) Pattern Recognition with Fuzzy Objective Function Algorithm, Plenum Press.

Blei M., Ng A.Y., Jordan M.I. (2003) Latent Dirichlet allocation. J. Mach. Learn. Res., 3, 993–1022.
Campello, R.J.G.B. (2007) A Fuzzy Extension of the Rand Index and Other Related Indexes for Clustering and Classification Assessment, Pattern Recognition Letters, 28, 833 – 841.
Cannon, R.L., Davè, J.V., Bezdek, J.C. (1986) Efficient implementation of the fuzzy C-means clustering algorithm, IEEE Transactions on Pattern Analysis and Machine Intelligence 8, 248–255.
Ciborra C, Braa K, et al (2000) From Control to Drift: The Dynamics of Corporate Information Infrastructures. Oxford University Press, Oxford
Colubi, A., González-Rodríguez, G., Gil, M.A., Trutschnig, W. (2011) Nonparametric criteria for supervised classification of fuzzy data, International Journal of Approximate Reasoning 52, 1272–1282.
Coppi, R. (2003) The fuzzy approach to multivariate statistical analysis, Technical report, Dipartimento di Statistica, Probabilità e Statistiche Applicate, Sapienza Università di Roma, n. 11.
Coppi, R., D'Urso, P., Giordani, P. (2012) Fuzzy and Possibilistic Clustering Models for Fuzzy Data, Computational Statistics and Data Analysis, 915-927.
Coppi, R., Giordani, P., D'Urso, P. (2006) Component Models for Fuzzy Data, Psychometrika, 71, 733–761.
Dixon BE (2007) A roadmap for the adoption of e-Health. E-Serv J 5(3):3-13
D'Urso, P. (2007) Clustering of fuzzy data, in Advances in Fuzzy Clustering and Its Applications (eds. de Oliveira J.V., Pedrycz W.), J. Wiley and Sons, 155–192.
D'Urso, P., Giordani, P. (2006) A weighted fuzzy c-means clustering model for fuzzy data, Computational Statistics and Data Analysis, 50, 1496–1523.
European Commission (2008) Information Society and Media Directorate-Genaral. Expert Impact Assessment. http://kb.good-ehealth.org/search.do
European Commission (2009) Good eHealth Report-eHealth in Action Good Practice in European Countries, Luxembourg: Office for Official Publications of the European Communities
Everitt, B.S., Landau, S., Leese, M. (2001) Cluster analysis (4th ed.). London: Arnold Press.
Fadili, M.J., Ruan, S., Bloyet, D., Mazoyer, B. (2001) On the number of clusters and the fuzziness index for unsupervised FCA application to BOLD fMRI time series, Medical Image Analysis, 5, 55-67.
Fitterer R, Mettler T, Rohner P, Winter R (2011) Taxonomy for multi-perspective assessment of the value of health information systems. Int J Healthc Technol Manag 12(1): 45–61
Glaser, B. G., and Strauss, A. L. (1967) The Discovery of Grounded Theory: Strategies for Qualitative Research, Aldine Publishing Company, Chicago
González-Rodríguez, G., Colubi, A, Gil, M.A. (2012) Fuzzy data treated as functional data: A one-way ANOVA test approach, Computational Statistics and Data Analysis, in press.
Graaff, A.J., Engelbrecht, A.P. (2012) Clustering data in stationary environments with a local network neighborhood artificial immune system, International Journal of Machine Learning and Cybernetics, DOI: 10.1007/s13042-011-0041-0.
Gregor S,(2006) The nature of theory in information systems, MIS Q 30 (3):611–642
Guo, G., Chen, S., Chen, L. (2012) Soft subspace clustering with an improved feature weight self-adjustment mechanism, International Journal of Machine Learning and Cybernetics, DOI: 10.1007/s13042-011-0038-8.
Hall, L.O., Bensaid, A.M., Clarke, L.P. (1992) A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain, IEEE Transactions on Neural Networks, 3, 672-682.
Hanseth O, Aanestad M (2003) Design as bootstrapping. On the evolution of ICT network in healthcare. Methods Inf Med 42:385–391

Hanseth O, Lyytinen K (2010) Design theory for dynamic complexity in information infrastructures: the case of building internet. J Inf Technol 25:1–19
Hawgood, J and Land, F (1988) A multivalent approach to information systems assessment. In Information Systems Assessment: Issues and Challenges (Bjorn-Andersen N and Davis GB, Eds), pp 103-124, North Holland, Amsterdam
Heiser, W.J., Groenen, P.J.F. (1997) Cluster differences scaling with a within-clusters loss component and a fuzzy successive approximation strategy to avoid local minima, Psychometrika, 62, 63-83.
Herriott, R. E., and W.A. Firestone, 1983, Multisite qualitative policy research: Optimizing description and generalizability. Educational Researcher, 12, 14 -19.
Hung, W.L., Yang M.S. (2005) Fuzzy clustering on LR-type fuzzy numbers with an application in Taiwanese tea evaluation. Fuzzy Sets and Systems, 150, 561-577.
Hwang, H., DeSarbo, W.S., Y. Takane (2007) Fuzzy clusterwise generalized structured component analysis, Psychometrika, 72, 181-198.
Irani Z, Love PED (2002) Developing a frame of reference for ex-ante IT/IS investment evaluation. Eur J Inf Syst 11(1):74-82
Irani Z, Love PED (2008) Evaluating Information Systems: Public and Private Sector. Butterworth-Heinemann, Oxford
Irani Z, Sharif A, Love PED, Kahraman C, (2002) Applying concepts of fuzzy logic cognitive mapping to model: the IT/IS investment evaluation process. Int J ProdEcon 75:199-211
Lafky DB, Tulu B, Horan TA (2006) A User-driven approach to personal health records. Commun Assoc Inf Syst 17(46):1028-1041
Liang, J., Song., W. (2012) Clustering based on Steiner points, International Journal of Machine Learning and Cybernetics, DOI: 10.1007/s13042-011-0047-7.
Mac Queen, J.B. (1967) Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 2, pp. 281–297.
Maharaj, E.A., D'Urso, P. (2011) Fuzzy clustering of time series in the frequency domain, Information Sciences, 2011, 181, 1187-1211.
Mantzana V, Themistocleous M, Irani Z, Morabito V (2007) Identifying healthcare actors involved in the adoption of information systems. Eur J Inf Syst 16(1):91-102
McBratney, A.B., & Moore, A.W. (1985) Application of fuzzy sets to climatic classification. Agricultural and Forest Meteorology, 35, 165–185.
Menachemi N, Burke DE, Ayers D, (2004) Factors affecting the adoption of telemedicine – a multiple adopter perspective. J Med Syst 28(6):617–632
Mitchell J, (2000) Increasing the cost-effectiveness of telemedicine by embracing e-health. J Telemed Telecare 6:S16-S19
Nagendran S, Moores D, Spooner R, Triscott J, (2000) Is telemedicine a subset of medical informatics? J Telemed Telecare 6 (Suppl. 2):50–51
Ozkan, I., Turksen, I.B. (2007) Upper and lower values for the level of fuzziness in FCM, Information Sciences, 177, 5143-5152.
Pal, N.R., Bezdek, J.C. (1995) On cluster validity for the fuzzy c-means model, IEEE Transactions on Fuzzy Systems, 3, 370-379.
Sinova B, Gil MA, Colubi A, Van Aelst S (2012) The median of a random fuzzy number. The 1-norm distance approach, Fuzzy Sets and Systems, in press.
Smithson S, Hirschheim R (1998) Analysing information systems evaluation: Another look at an old problem. Eur J Inf Syst 7(3):158-174

Soreson JA, Wang X (1996) ROC methods for evaluation of fMRI techniques, Magn. Res. Med., 36, 737-744.
Spagnoletti P, Albano V, Caccetta E, Tarquini R, D'Atri A (2011) "Supporting policy definition in the e-health domain: a QCA based method", HEALTHINF – International Conference on Health Informatics, 26-29 January, Roma, Italy
Stockdale R, Standing C, Love PED, Irani Z (2008) Revisiting the content, context and process of IS evaluation. In: Irani Z and Love PED (eds) Evaluating Information Systems, Public and Private Sector. Butterworth-Heinemann, Oxford, pp 35-45
Wedel M and Kamakura WA (1998) Market segmentation: Conceptual and methodological foundations. Boston: Kluwer Academic.
Wilson V (2003) Asynchronous health care communication. Commun ACM 46(6):79-84
Yin RK (2009) Case Study Research: Design and Methods. Fourth Edition. SAGE Publications. California
Yusof MM, Kuljis J, Papazafeiropoulou A, Stergioulas LK (2008a) An evaluation framework for health information systems: human, organization and technology-fit factors (HOT-fit). IntJ Med Inform 77(6): 386–398
Yusof MM, Papazafeiropoulou A, Paul RJ, Stergioulas LK (2008b) Investigating evaluation frameworks for health information systems Int J Med Inform 77(6): 377–385
Zadeh LA (2005) Toward a generalized theory of uncertainty (GTU) - An outline. Information Sciences, 172, 1–40.